Publications of the Astronomical Society of Japan, 2025, **00**, 1–11
doi: 10.1093/pasj/xxx000
Advance access publication date:

A new approach for predicting the stability of hierarchical triple systems – I. Coplanar Cases

Ryutaro IWAKURA,[1]* Yoko FUNATO,[2] Junichiro MAKINO,[1]

[1]Department of Planetology, Graduate school of Science, Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe, Hyogo 657-8501, Japan
[2]General Systems, Graduate School of Arts and Science, The University of Tokyo, Komaba, Meguro, Tokyo, 153-8902, Japan
*E-mail: riwakura@stu.kobe-u.ac.jp

ORCID: 0009-0009-8049-2968, 0000-0002-6992-7600, 0000-0002-0411-4297

Abstract

Hierarchical triple systems play a crucial role in various astrophysical contexts, and therefore the understanding of their stability is important. Traditional empirical stability criteria rely on a threshold value of Q, the ratio between the outer orbit's pericenter distance and the inner orbit's semi-major axis. However, determining a single critical value of Q is impossible because there is a range of the value of Q for which both stable and unstable systems exist, referred to as the mixed-region. In this study, we introduce a novel method to assess the stability of triple systems within this mixed-region. We numerically integrate equal-mass, coplanar hierarchical triples within the mixed-region. By performing Fourier analysis of the time evolution of the semi-major axes ratio during the first 1000 inner orbital periods of the systems, we find notable features in stable systems: if the main peaks are periodically spaced in the frequency domain and the continuous components and irregularly spaced peaks are small, the system tends to be stable. This observation indicates that the evolution of stable triples is more periodic than that of unstable ones. We quantified the periodicity of the triples and investigated the correlation between the Fourier power distribution and the system's lifetime. Using this correlation, we show that it is possible to determine if a triple system in the mixed-region is stable or not with very high accuracy. These findings suggest that periodicity in orbital evolution can serve as a robust indicator of stability for hierarchical triples.

Keywords: method:numerical — celestial mechanics — planets and satellites: dynamical evolution and stability

1 Introduction

A hierarchical three-body system consists of a binary and a third body orbiting around the binary. The stability of hierarchical three-body systems is a long-standing problem in the fields of dynamical astronomy, classical mechanics, and mathematics. Understanding its stability is not only of theoretical interest but also a practical necessity for the interpretation of astronomical phenomena.

For example, a highly accurate three-body stability condition is essential for N-body simulations of globular clusters (Aarseth 2003; Heggie & Hut 2003). In N-body simulations, the computational cost of binary and hierarchical three-body systems is very high, because their orbital timescale is many orders of magnitudes shorter than the dynamical or thermal timescale of the parent cluster. Accordingly, in typical N-body simulations, the evolution of binary systems is computed in the following manner. In the case of stable systems, we can integrate them using some variation of orbit averaged perturbation methods such as the slow-down method (e.g., Mikkola & Aarseth 1996; Wang et al. 2020). These methods are designed to reduce the number of time steps needed for orbit calculations while maintaining high accuracy. On the other hand, for unstable systems, we cannot use such approximate methods because such systems should be integrated without approximation. This is because applying approximation to unstable systems can lead to physically incorrect solutions. In principle, when integrating such unstable systems without approximation, the computational cost would not be very large, since such unstable systems disintegrate in relatively short timescales. Thus, prediction of the stability of triple systems with high accuracy is essential for efficient and accurate simulation of globular clusters.

However, there is no such high-accuracy prediction method. A common issue is the misclassification of the dynamical stability of these systems — that is, mistaking a stable system as an unstable one or vice versa. Such misidentification can severely degrade computational efficiency by causing unnecessary direct integrations, effectively stalling the simulation. Moreover, incorrect classification may result in physically inaccurate outcomes.

Here, we provide a brief review of hierarchical triple stability criteria. Traditionally, numerous stability conditions have been formulated as expressions that denote the minimum value of Q:

$$Q = \frac{q_{\rm out}}{a_{\rm in}} = \frac{a_{\rm out}(1-e_{\rm out})}{a_{\rm in}}, \qquad (1)$$

where q is the pericenter distance, a is the semi-major axis, e is the eccentricity, and subscripts "in" and "out" correspond to "inner" and "outer" orbit, respectively. We define the critical value of the stability parameter Q, denoted as $Q_{\rm crit}$, as the minimum value above which the triple system remains dynamically stable. Using semi-analytical or numerical methods, $Q_{\rm crit}$ is fitted as a function of the orbital elements of the system, i.e., $Q_{\rm crit} = Q_{\rm crit}(m_1, m_2, m_3, a_{\rm in}, a_{\rm out}, e_{\rm in}, e_{\rm out}, I, \dots)$, where the variables include the component masses (m_1, m_2 are the masses of inner particles and m_3 is the that of the outer particle), semi-major axes, eccentricities, and orbital inclinations. Dependence on other parameters such as the arguments of periapsis (ω) or orbital phases is usually neglected. This is because including those parameters would make the parameter survey significantly more computation-

Received: , **Accepted:**

ally expensive, and the resulting criterion would become less practical for application. By comparing the actual value of Q for a given three-body system with the fitted $Q_{\rm crit}$, one hopes to judge the system to be unstable if $Q < Q_{\rm crit}$, and stable if $Q > Q_{\rm crit}$.

Harrington (1972) found that the minimum value of Q for stability differs between prograde and retrograde orbits. Harrington also derived a stability criterion for coplanar orbits including three-body mass dependency (Harrington 1975, 1977). Eggleton & Kiseleva (1995, hereafter EK95) introduced a parameter Y, which is defined as the ratio of the apocenter distance of the inner to the pericenter distance of the outer, instead of using Q. They also empirically derived a stability criterion in terms of Y. Both Q and Y represent the distance between the inner and outer binaries, with no substantial difference between them. Mardling & Aarseth (1999, 2001, hereafter MA01) proposed a well-known dynamical stability criterion for hierarchical triples:

$$Q_{\rm crit,MA01} = 2.8\left(1-\frac{0.3I}{\pi}\right)\left[\left(1+\frac{m_3}{m_1+m_2}\right)\frac{1+e_{\rm out}}{\sqrt{1-e_{\rm out}}}\right]^{2/5}, \tag{2}$$

where I is the inclination, m_1 and m_2 are the masses of inner, m_3 is the mass of outer, and e is the eccentricity. Mylläri et al. (2018, hereafter M18) further modified the MA01 criterion by incorporating the dependence on the inner eccentricity, while Vynatheya et al. (2022, hereafter V22) incorporated the effect of ZLK mechanism (von Zeipel 1910; Lidov 1962; Kozai 1962) and refined the criteria proposed by EK95 and MA01. Numerous three-body stability conditions have been reported for use in N-body simulations (e.g., Valtonen et al. 2008; Georgakarakos 2013; Mushkin & Katz 2020; Hayashi et al. 2022). Recently, several studies using machine learning have been reported (e.g., V22; Lalande & Trani 2022, hereafter LT22).

Any formula for Q proposed so far does not really discriminate stable and unstable systems, since for the values of Q close to the "critical" value of Q, $Q_{\rm crit}$, for any of these criteria, some realizations of the three-body systems turned out to be stable while some to be unstable. In this parameter space, both stable and unstable systems coexist, forming what is commonly referred to as a mixed-region (e.g., Dvorak 1986). One reason for the existence of this mixed-region is that in most formulas orbital elements such as the argument of periapsis ω or the initial phase are not taken into account, although these elements can significantly influence the stability of three-body systems (Hayashi et al. 2022, 2023).

Even if one attempts to extend the fitting procedure to include these additional orbital elements, the intrinsic chaotic nature of the three-body problem makes precise classification difficult. This is particularly true in the mixed-region, where the system exhibits strong chaos, and small differences in orbital elements can lead to large variations in the lifetime of the system. Recent studies suggest that stability boundaries in chaotic systems may exhibit fractal-like structures, with stable and unstable regions intricately interwoven across all scales (e.g., Trani et al. 2024). This complexity underscores the limitation of defining deterministic boundaries based solely on initial parameters like Q.

Therefore, to achieve high-accuracy stability classification in the mixed region, it is essential to evaluate not only the initial parameters but also the subsequent orbital evolution itself. LT22 simulates triple systems with Q values 5%–15% below the MA01 critical threshold and employs machine learning on the orbital elements obtained from the simulations, achieving an Area Under Receiver Operating Characteristic Curve (AUC for ROC curve) of about 0.95. However, their approach uses a relatively large time step for sampling, and it remains unclear which aspects of the orbital elements are associated with instability. In addition, the AUC value of 0.95 is rather low to be used in N-body simulations. Hence, in order to evaluate stability by taking into account the orbital evolution, a more detailed analysis of the dynamical behavior of orbital elements is required.

The aim of this study is to develop a new method to determine the stability of hierarchical triples that can overcome the problems discussed above. First, we conducted numerical integrations on coplanar equal-mass hierarchical three-body systems whose Q at the initial condition is in the mixed-region. Here we adopt the critical value Q by MA01 as $Q_{\rm crit}$ of the mixed-region. Then, we investigate the differences between long-lived and short-lived systems from the perspective of orbital element periodicity using Fourier analysis.

Our motivation for employing Fourier analysis stems from the expectation that the fundamental nature of a system's orbital evolution—whether it is regular and quasi-periodic or irregular and chaotic—should be imprinted in its frequency spectrum. We hypothesized that long-term stable hierarchical triple systems, which tend to exhibit quasi-periodic motion, would show a Fourier spectrum dominated by discrete, well-defined peaks corresponding to the principal orbital frequencies and their harmonics. In contrast, systems prone to instability and chaotic behavior were anticipated to display more complex spectra, characterized by broadened peaks, significant continuous components, or an increased power in low-frequency modes, reflecting the irregular and aperiodic nature of their orbits. By examining these spectral characteristics, particularly within a relatively short initial span of the system's evolution (the first $\sim 10^3 P_{\rm in}$), we aimed to identify robust signatures that correlate with long-term stability, thereby providing a more discerning tool for stability assessment in the challenging mixed-region.

The plan of this paper is as follows: In Section 2, we describe our models and method for our numerical integration. We show the result of our simulation and the performance of our new stability criterion in Section 3. Section 4 is for discussion.

2 Methods

In this section, we describe the initial conditions and numerical methods used in our study. We numerically integrated hierarchical triple systems with Q values smaller than the lower limit determined by the MA01 criterion until the hierarchical structure of the systems was disrupted or the time reached to $10^9 P_{\rm in}$. Only Newtonian gravity is considered, neglecting any general relativistic effects.

2.1 Initial Conditions

Our coplanar three-body system consists of a inner binary and a third body, which we refer to as "inner" and "outer". The two bodies that form inner binary has indices 1 and 2, while the outer body has index 3. We consider two types of coplanar cases, prograde orbits with inclination $I = 0$ and retrograde orbits with $I = \pi$.

We fixed some of the parameters and initial orbital elements: mass ($m_1 = m_2 = m_3 = 0.5 M_\odot$), inner semi-major axis ($a_{\rm in} = 1$ au), eccentricity ($e_{\rm in} = 0.5, e_{\rm out} = 0.25$) and inner argument of periapsis ($\omega_{\rm in} = 0$). These parameters are not special values but rather arbitrary ones. We uniformly sample every $2\pi/10$ radian of outer arguments of periapsis ($\omega_{\rm out} \in \mathcal{U}[0, 2\pi)$) and mean anomalies ($M_{\rm in}, M_{\rm out} \in \mathcal{U}[0, 2\pi)$). The list of initial conditions is shown in Table 1.

Table 1. Initial conditions. In the coplanar case, the longitude of the ascending node is not defined. For the convenience of numerical integrations, we set the longitude of the ascending nodes as $\Omega_{\rm in} = \Omega_{\rm out} = 0$. Outer argument of periapsis $\omega_{\rm out}$ and mean anomalies $M_{\rm in}, M_{\rm out}$ take the value $2\pi i/10$, where $i = 0, 1, ...9$.

Parameters	Symbol	Value
mass of particles	m_1, m_2, m_3	0.5 $M_\odot$
inner semi-major axis	$a_{\rm in}$	1 au
inner eccentricity	$e_{\rm in}$	0.5
outer eccentricity	$e_{\rm out}$	0.25
inner argument of periapsis	$\omega_{\rm in}$	0
outer argument of periapsis	$\omega_{\rm out}$	$\mathcal{U}[0, 2\pi)$
inner mean anomaly	$M_{\rm in}$	$\mathcal{U}[0, 2\pi)$
outer mean anomaly	$M_{\rm out}$	$\mathcal{U}[0, 2\pi)$

We set $a_{\rm out}$ so that Q at the initial condition of each run is in the mixed-region. By applying our initial parameters in Table 1 to equation (2) , we can define the stability threshold as $Q \simeq 3.81$ for prograde orbits and $Q \simeq 2.67$ for retrograde orbits. These two Q values represent the stability limits for the coplanar cases determined from MA01 criterion. So we use $Q = 3.80, 3.77, 3.75, 3.73$ for prograde orbits, and $Q = 2.60, 2.58, 2.55, 2.52$ for retrograde orbits. These values of Q are also arbitrary, but they lie in a region where stable and unstable systems coexist, making them appropriate for the purpose of this study. Once Q is determined, the outer semi-major axis ($a_{\rm out}$) can be calculated from equation (1), and all the orbital elements can be obtained. When the Q is determined, then $a_{\rm out}$ can be obtained using equation (1).

Our initial conditions consist of two types of orbital inclination: prograde and retrograde. For each case, there are four different values of Q, along with three parameters that are uniformly distributed. Therefore, there are $4 \times 10 \times 10 \times 10 = 4000$ initial conditions for prograde orbits, and 4000 initial conditions for retrograde orbits as well.

2.2 Numerical Method

We integrate coplanar equal-mass triples for a maximum time of $10^9 P_{\rm in}$, where $P_{\rm in}$ is the initial inner orbital period. We continue the integrations until the system disintegrates or the time reaches to $10^9 P_{\rm in}$. When the binding energy (i.e., the energy needed to disassemble the system) of either the inner or outer binary becomes negative, we regard the system as disintegrate. The snapshots are recorded at intervals of $0.1 P_{\rm in}$ for the first $10^3 P_{\rm in}$. This interval is determined by balancing computational resources with the granularity of the data. By storing snapshots at such fine intervals, we can improve the analysis accuracy in the subsequent Fast Fourier Transform (FFT).

We used Algorithmic regularization (AR) for our integration, which is also called as Time-Transformed Symplectic Integrator (TSI) or LogH method (Preto & Tremaine 1999; Mikkola & Tanikawa 1999). AR is described in the extended phase space. Using AR, time is treated as one of the variables to be integrated in an extended phase space composed of time, positions, and velocities. When applied to time integration of a two-body problem with the leap-frog method, AR gives the exact trajectory, the conservation of energy, and angular momentum of the system. For these reasons, AR is well-suited for long-term integration of few-body systems. We combine AR with a 6th-order symplectic formula by Yoshida (1990) to improve its accuracy. As stated in Wang et al. (2020), the desirable properties of AR are preserved even when combined with the Yoshida 6th-order symplectic method. In this study, we modified the sample code included in the SDAR library [1] (Wang et al. 2020) to create a numerical code for the AR+Yoshida6th method. SDAR is an open-source library for integrating few-body systems, and it is available for anyone to use. In addition to the AR+Yoshida6th method, we performed similar simulations with TSUNAMI (Trani & Spera 2023) and confirmed that our results are independent of the integrator used.

3 Results

3.1 The overall results of all integrations: the distribution of the survival time

Figure 1 shows the distribution of the survival time. In the prograde case, the majority of systems break up in the period between $10^3 P_{\rm in}$ and $10^6 P_{\rm in}$. The number of systems with lifetimes between $10^7 P_{\rm in}$ and $10^9 P_{\rm in}$ is extremely small, and those surviving beyond $10^9 P_{\rm in}$ account for approximately 1% to 4% of the total. In the retrograde case, the distribution is broader than that of the prograde case, with most systems breaking up in the period between $10^2 P_{\rm in}$ and $10^6 P_{\rm in}$. Systems surviving beyond $10^9 P_{\rm in}$ account for approximately 5% of the total, except for the case of $Q = 2.52$.

As shown in figure 1, even systems with the same Q exhibit variations in survival time of several orders of magnitude. Our initial conditions vary only the argument of periapsis (ω) and initial phases for a given Q, so the combination of ω and phases produces the distribution observed in figure 1. The current stability conditions incorporate dependencies on the mass, semi-major axis, eccentricity, and orbital inclination, but the dependence on ω and phases has not yet been thoroughly investigated. Therefore, the distribution of stability and survival time observed in figure 1 cannot be captured at all by the current stability conditions.

3.2 Periodicity of orbital evolution

In this section, we compare the orbital evolution of stable and unstable systems shown in figure 1. First, we explain the mechanism of system destabilization. Figure 2 shows the time evolution of orbital elements and the time evolution of the x-coordinates of each of the three bodies for a prograde orbit with $Q = 3.75 (\omega_{\rm out} = 0, M_{\rm in} = \pi/5, M_{\rm out} = 0)$. Note that in our calculations the orbit is in the x-y plane. In this case, the three-body system breaks down at approximately $10500 P_{\rm in}$, but since this study only collects detailed output up to $10000 P_{\rm in}$, the data is shown up to $10000 P_{\rm in}$, which is sufficient for our explanation. The orbital evolution shows that, initially, the system evolves in a relatively periodic manner up to approximately $7000 P_{\rm in}$. However, around $8000 P_{\rm in}$, fluctuations in the orbital elements occur, which contribute to the destabilization of the system. Eventually, the outer orbit becomes highly eccentric. Systems that reach this state eventually undergo disruption.

Figure 3 shows the trajectories of the system from figure 2 as viewed in the orbital plane (x-y plane). The initial orbital variations occur between $7000 P_{\rm in}$ and $8000 P_{\rm in}$ (upper center and right in figure 3), during which it is evident that the shape of the outer orbit undergoes significant changes before and after this time.

Based on the above considerations, it is anticipated that stable orbits will not undergo significant variations over time. This raises the question: what distinguishes stable orbits from unstable ones? In the following, we conduct a comparative analysis between stable and unstable orbits in order to elucidate their differences.

[1] https://github.com/lwang-astro/SDAR

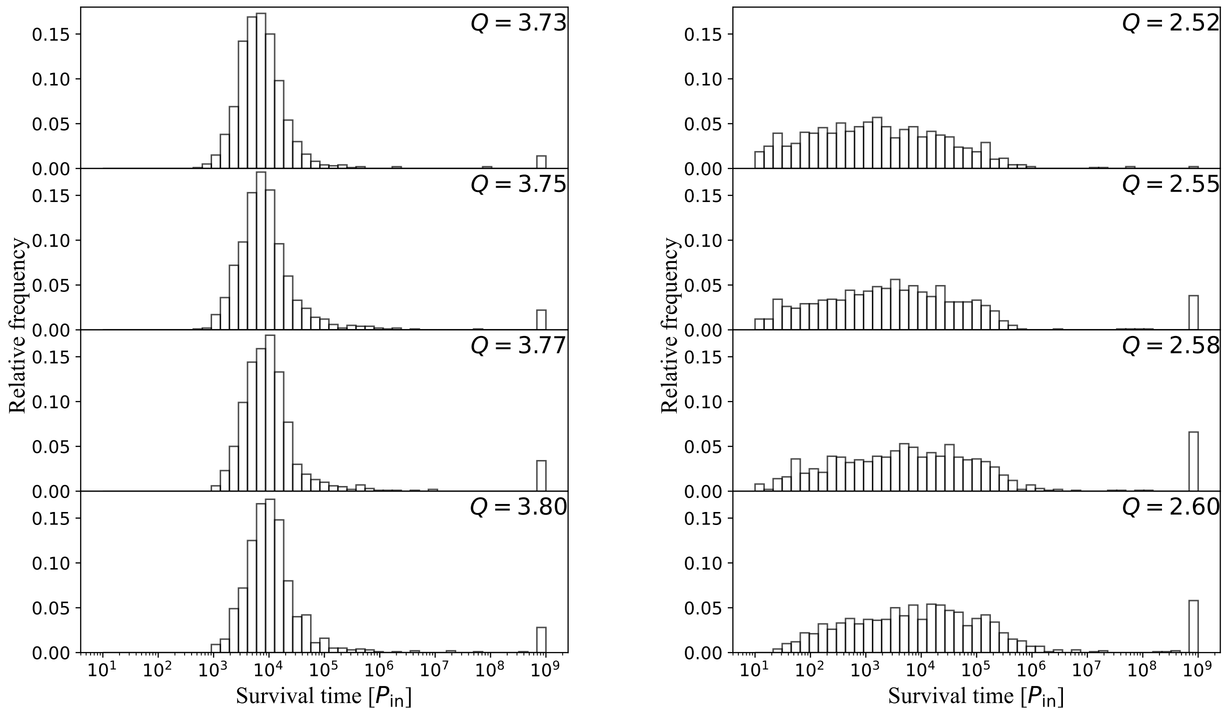

Fig. 1. The distribution of survival time distribution for each Q. The left panel is prograde cases and the right panel is retrograde cases. For each cases, we distribute 50 bins between $10^1 P_{\rm in}$ to $10^9 P_{\rm in}$ on a logarithmic scale, and we plot the relative frequency of each bin.
ALT text: Histogram plots for each of the eight Q values.

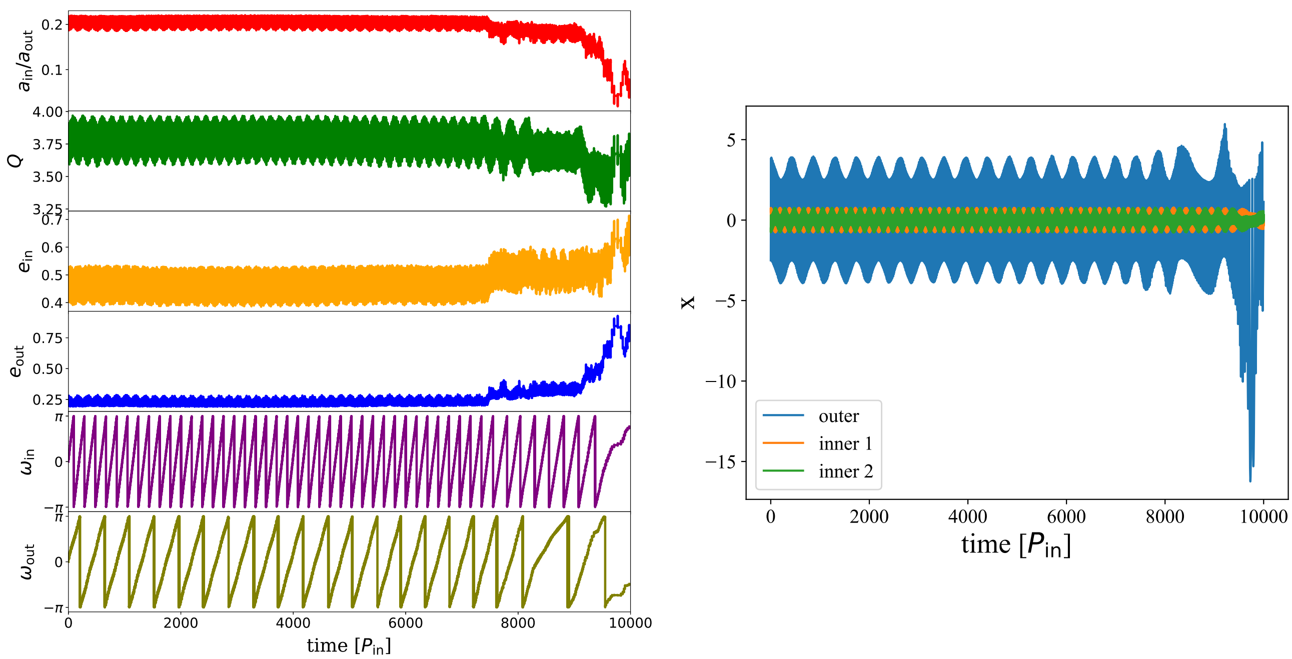

Fig. 2. The orbital evolution of an unstable prograde orbit with $Q = 3.75$ and the time evolution of the x-coordinates of the three bodies.
ALT text: Time evolution of the orbital elements and the x-coordinates of the three bodies.

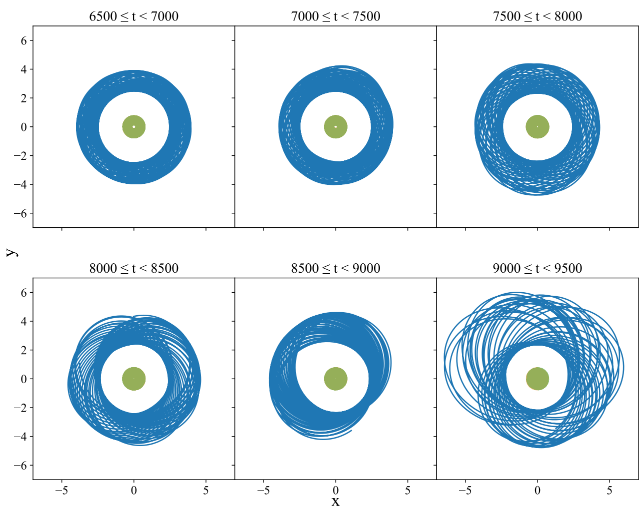

Fig. 3. The trajectories of the three bodies on the orbital plane from $t = 6500P_{\rm in}$ to $t = 9500P_{\rm in}$ for the same system shown in Figure 2.
ALT text: Six snapshots depicting the trajectory of the orbit.

Figure 4 shows the orbital evolution over the first $1000P_{\rm in}$ of the systems shown in figure 2 and that of a stable system having the same value of $Q = 3.75$. The unstable system (the left-hand side panel) exhibits significant irregular variations in all orbital elements while irregularities seem smaller for the stable system(right-hand side panel). This difference might imply that there exists a meaningful relationship between the degree of the periodicity of the orbital elements and the long-term dynamical stability of the system.

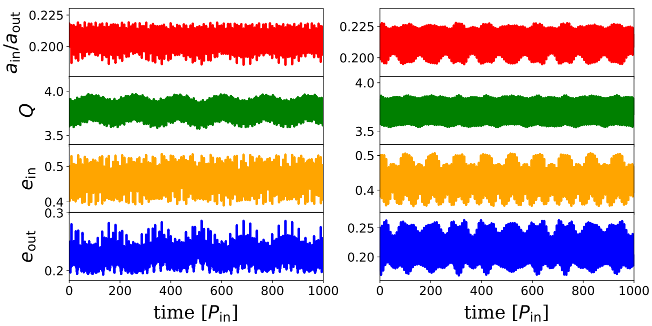

Fig. 4. Examples of orbital elements evolution: the left is for an unstable system, and right is for a stable system. These are for the cases with $Q = 3.75$.
ALT text: A figure comparing four orbital elements between stable and unstable orbits. Four orbital elements are the semi-major axis ratio, Q, inner eccentricity, and outer eccentricity.

In figure 4, we just give one particularly illustrative example. We need some measure for this degree of periodicity. Therefore, we employ the Fast Fourier Transform (FFT) to quantitatively investigate the relationship between periodicity and dynamical stability. In this paper, as the first trial, we focus on the evolution of the ratio of semi-major axes of the inner and outer orbits ($a_{\rm in}/a_{\rm out}$).

Here, we present our procedure to apply FFT to a measured quantity. In this study, we output the orbital elements with a $0.1P_{\rm in}$ interval for the first $10^3 P_{\rm in}$. Since the FFT requires the number of data points to be a power of 2, we extracted the first 8192 data points, corresponding to $819.2P_{\rm in} \sim 10^3 P_{\rm in}$, and performed the FFT. Before performing the FFT, the mean value of the 8192 equally-spaced points in time from $0P_{\rm in}$ to $819.2P_{\rm in}$ is calculated and subtracted from each data point. This step is done to minimize the DC component as much as possible. After this process, a Hanning window function is applied to the data, and then the FFT is performed. The resulting amplitudes are normalized by dividing them by half the total number of data points, i.e., 4096. The sampling frequency of the FFT is 10 $P_{\rm in}^{-1}$, and the Nyquist frequency is 5 $P_{\rm in}^{-1}$. We discarded systems that survive less than $819.2P_{\rm in}$, since we cannot do FFT analysis on such systems for sufficient time. There are 5 prograde systems and 1287 retrograde systems whose survival times did not reach $819.2P_{\rm in}$, so we present the results for 3995 prograde cases and 2713 retrograde cases.

Figure 5 shows 16 examples of our FFT results. The panels are arranged from top to bottom in the order of survival time for four different values of Q (two prograde and two retrograde). If the orbital evolution is perfectly periodic, only the fundamental frequency (f_0) and its integer multiples ($2f_0, 3f_0, \ldots$) would appear in the frequency domain with a small effect of the window function which appears as small broadening of each peak and low-frequency terms. The orbital evolution of the system with evenly spaced peaks and fewer continuous components, which are side-lobes of the peaks, therefore, is more stable. When we compare systems with different lifetimes (for instance, comparing the top and bottom columns), it is evident that the stable systems exhibit more pronounced peaks with smaller side-lobes and non-periodic "noises". This trend indicates that stable systems are more periodic. This statement might sound almost like a tautology, since by definition periodic systems are stable. However, as far as we know this is the first time that the numerical determination of periodicity is applied to the stability of hierarchical triples. Additionally, a consistent feature across almost all cases is the presence of a noticeable peak near 0 $P_{\rm in}^{-1}$. This peak is not the fundamental frequency and its strength seem to be related to the early destabilization before $10^9 P_{\rm in}$.

Figure 6 depicts examples of the FFT results ($Q = 2.60$, the rightmost panel of figure 5), focusing on the frequency range up to 0.45. The four cases in figure 6 all correspond to retrograde cases, but a similar tendency is observed for prograde motion or cases with different Q values.

First, we clarify the physical meaning of the peaks. There are three peaks in figure 6, which correspond to two different types of variations. The middle peak and the right peak originate from the same variation. The peaks around 0.2 $P_{\rm in}^{-1}$ represent the fundamental frequencies (f_0), and the peaks observed around 0.4 $P_{\rm in}^{-1}$ correspond to the second harmonic of the fundamental frequency ($2f_0$). The other peaks observed in figure 5 are also components that are integer multiples of the fundamental frequency. For $Q = 2.6$, the period of the outer orbit is approximately five times that of the inner orbit. Given that the frequency unit in the FFT is $P_{\rm in}^{-1}$, the value $1/5$ $P_{\rm in}^{-1} = 0.2$ $P_{\rm in}^{-1}$ indicates that these peaks are associated with variations caused by the periapsis passage of the outer orbit. In contrast, the leftmost low-frequency peak originates from a different dynamical mechanism. This peak represents the long-term variations caused by the continuous interaction between the inner and outer orbits such as mean-motion or secular resonances. In this paper, the continuous components and peaks in the low-frequency region, as depicted in figure 6, are collectively referred to as *low-frequency components*.

Next, we compare the unstable ones (the top three) with the

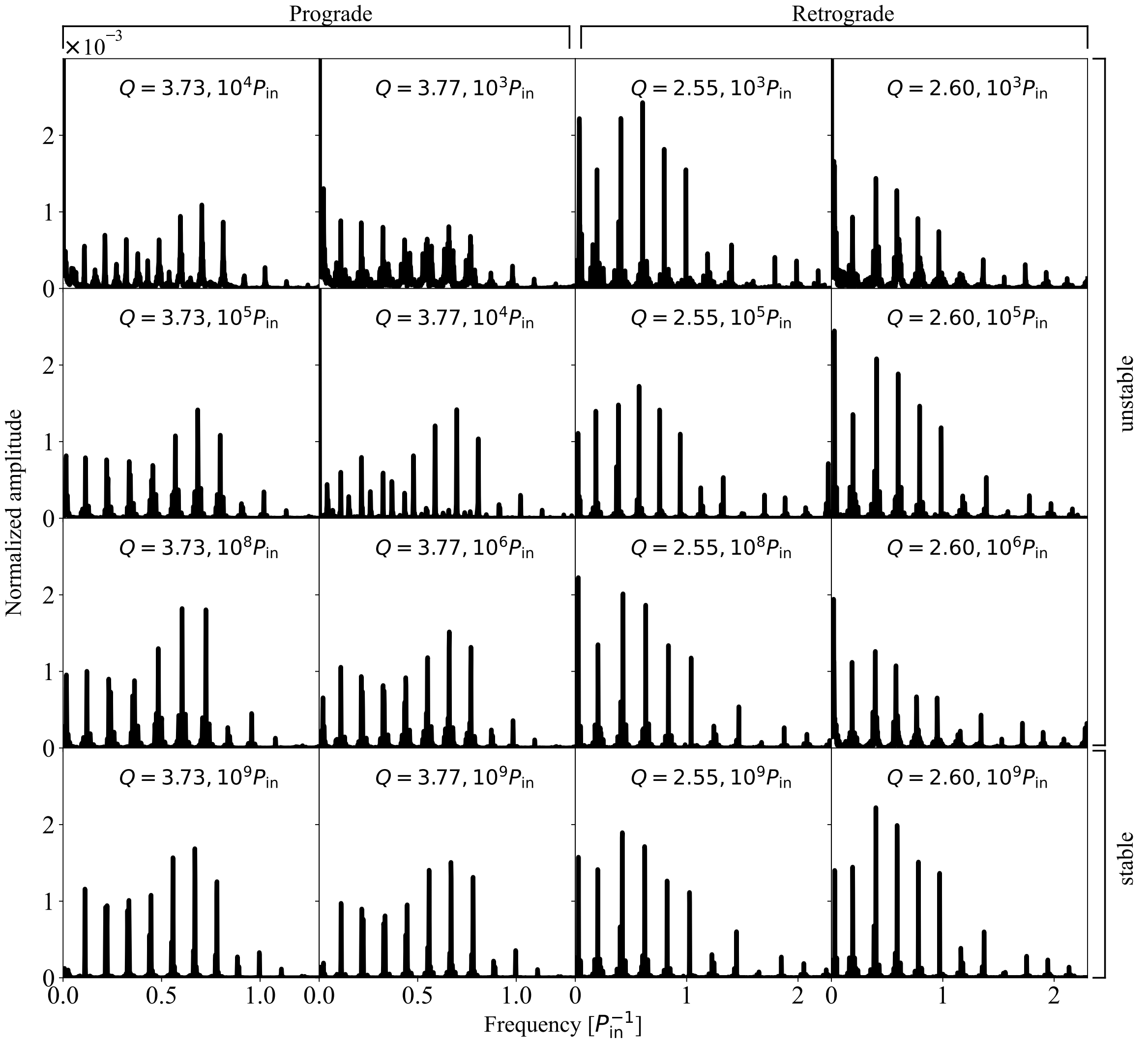

Fig. 5. The frequency distribution of orbital evolution ($a_{\rm in}/a_{\rm out}$). The horizontal and vertical axes correspond to the frequency ($P_{\rm in}^{-1}$) and the normalized amplitude, respectively. From left to right, each column shows 4 cases for prograde ($Q = 3.73$), 4 prograde cases ($Q = 3.77$), 4 retrograde cases ($Q = 2.55$), and 4 retrograde cases ($Q = 2.60$). The survival time of the corresponding system is shown in the upper right of each panel.
ALT text: Sixteen panels showing the results of the FFT of the orbits. Stable orbits exhibit distinct, evenly spaced peaks. In unstable cases, strong continuous components are present, and peaks are not evenly spaced.

stable ones (the bottom). As mentioned above, the system survives longer when peaks at f_0 and $2f_0$ are clearly visible (the bottom of Fig 6). On the other hand, the top three panels in figure 6 show larger continuous components compared to the bottom panel (the most stable one). Such continuous components are characteristic of unstable systems. Additionally, the first and third panels from the top show a DC component at 0 $P_{\rm in}^{-1}$. If the orbital evolution is not periodic, the center of oscillation deviates from the mean value of the data, resulting in the appearance of a DC component and peaks around 0 $P_{\rm in}^{-1}$.

The difference between the second panel from the top and the bottom panel in figure 6 is also rather clear. It can be seen that the second panel from the top has clearly defined peaks, but the low-frequency component is larger than other peaks. Though this feature is common in most cases, there are exceptions such as the second-to-top panel of figure 5 with $Q = 2.55$. Therefore, it is not possible to make a definitive conclusion. Nevertheless, it is plausible that low-frequency components act as the dominant factor, which governs the stability.

To summarize, the FFT analysis of the orbital element evolution of a stable system reveals distinct fundamental frequency components and their integer multiples, with small continuous components. Furthermore, the peak near 0 $P_{\rm in}^{-1}$ is relatively small. It is rather surprising that such trends can be recognized by observing the orbital evolution for only $10^3 P_{\rm in}$.

3.3 Probability analysis of stability prediction using FFT result

3.3.1 Quantification of Periodicity

In this section, we quantitatively evaluate the periodicity of the orbital evolution and show its relationship with the stability. In the previous section, we found that stable triple systems tend to have strong peaks at the fundamental frequency and its integer multi-

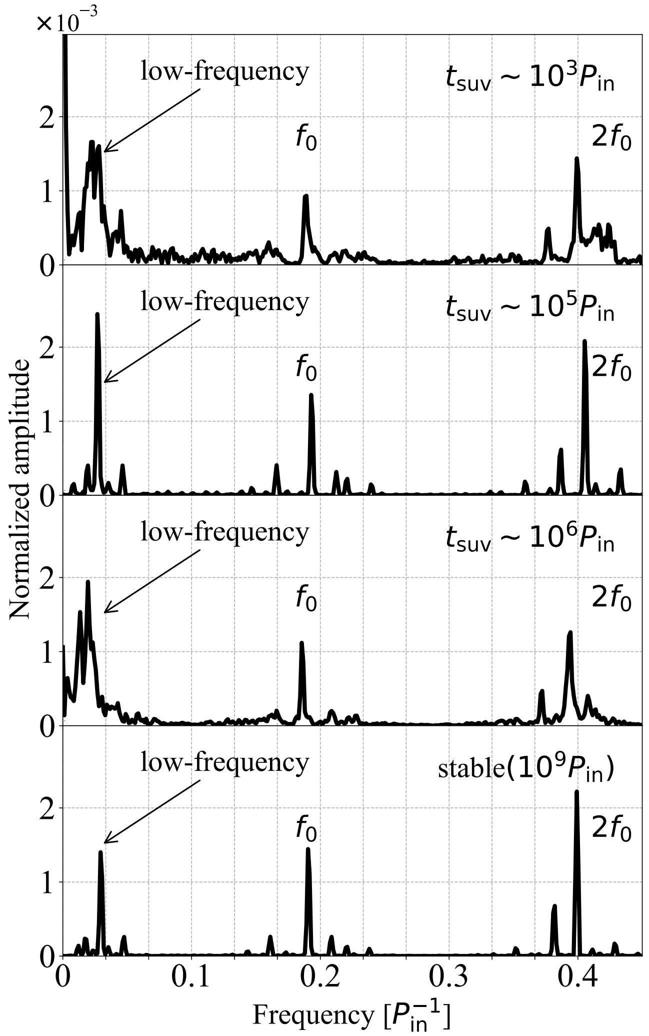

Fig. 6. 4 examples of orbital evolution FFT for retrograde $Q = 2.60$ cases. These figures correspond to the rightmost panel of figure 5. The vertical axes represent the normalized amplitude of FFT, while the horizontal axis represents frequency ($P_{\rm in}^{-1}$). The grid lines on the x-axis are evenly spaced at intervals of 1/30 units.
ALT text: Zoomed-in FFT spectra up to a frequency of 0.45 for four cases with Q equal 2.6.

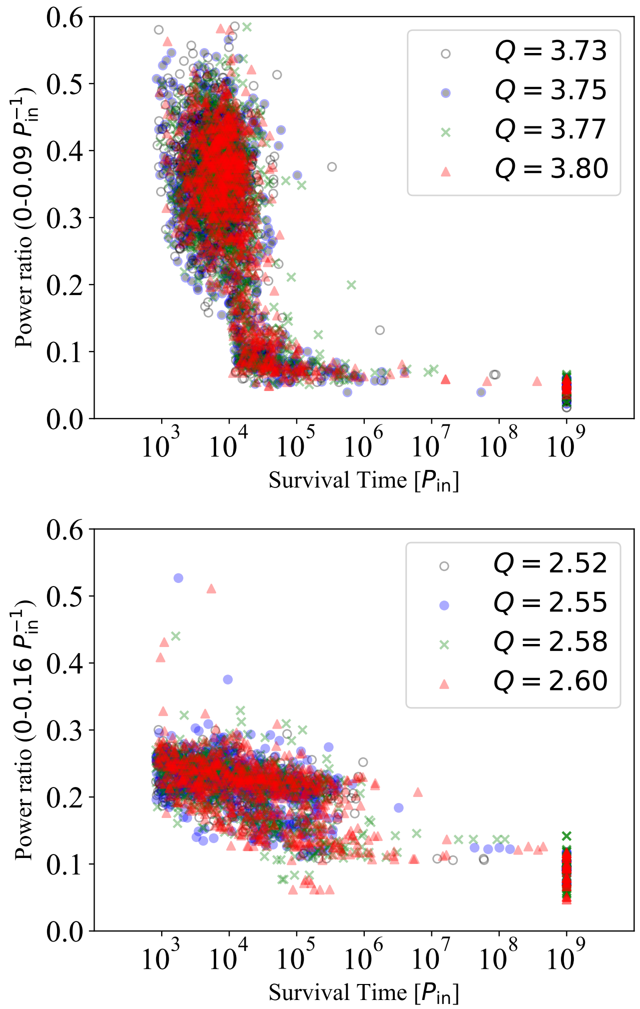

Fig. 7. Relation between survival time and low-frequency components power ratio. The upper panel is for prograde orbits and the lower panel is for retrograde orbits.
ALT text: Two scatter plots showing the survival time versus the fraction of low-frequency components.

ples, with minimal continuous components. Systems that disintegrate tend to have continuous components and/or a large peak between zero and the fundamental frequencies. These components seem to be related to short lifetimes.

We focused on this *low-frequency component* in order to quantify the non-periodicity of the orbital evolutions. Our FFT results show that the fundamental frequency appears around $0.10\ P_{\rm in}^{-1}$ for prograde systems and around $0.19\ P_{\rm in}^{-1}$ for retrograde systems. Taking this into account, we calculated the sum of the power of frequencies lower than this fundamental frequency and expressed it as a ratio to the total power of all peaks. For the prograde case, components with frequencies below $0.09\ P_{\rm in}^{-1}$ were regarded as low-frequency components. For the retrograde case, components with frequencies below $0.16\ P_{\rm in}^{-1}$ were regarded as low-frequency components.

The power ratio of this low-frequency component is plotted against the triple's survival time in figure 7. There is a clear correlation between the power of the low-frequency component and the lifetime of the system. Note that this low-frequency component is not a single factor that determines the lifetime. We can see that there are some systems with small low-frequency power which still exhibit short lifetimes. Even so, it is clear that we can use this measure as one way to predict the lifetime.

3.3.2 Predictability

In the previous subsection, we proposed a method to predict stability using FFT results. Figure 7 illustrates the relationship between the system's survival time, which corresponds to its stability, and the magnitude of its low-frequency component. in the time variation of the orbital element. We classified systems with a specific power ratio. The systems with a power ratio smaller than the threshold are regarded as stable. For example, in the case of retrograde orbits shown in figure 7, drawing a horizontal line at a power ratio of 0.15 allows for the detection of all systems that survive above $10^9 P_{\rm in}$. When the power ratio is divided at a certain value, four patterns can be observed: (i) a stable system is classified as stable (true positive; TP), (ii) an unstable system is classified as stable (false positive; FP), (iii) an unstable system is classified as unstable (true negative; TN), and (iv) a stable system is classified as unstable (false negative; FN). We verify the validity of our results and the accuracy of the predictions using these four classifications.

Before proceeding to the detailed evaluation of accuracy, we introduce the following metrics:

1. Accuracy: the ratio of the number of correct predictions to the total number of systems: (TP + TN)/(TP + TN + FP + FN).
2. Precision: the ratio of the number of actual stable systems to that of systems regarded as stable: TP / (TP + FP).
3. Recall or True Positive Rate (TPR): the ratio of the number of systems correctly identified as stable to that of real stable systems: TP / (TP + FN).
4. False Positive Rate (FPR): the ratio of the number of systems that are incorrectly classified as stable to the total number of real unstable systems: FP / (TN + FP).
5. Specificity: the ratio of the number of systems correctly classified as unstable to that of unstable systems: 1-FPR.
6. F-measure: the harmonic mean of Precision and Recall.

These metrics vary depending on the threshold.

Whole of our dataset is highly imbalanced to the unstable cases, so here we create subsets in which the numbers of stable and unstable systems are equal. These subsets include all systems classified as stable, along with the same number of unstable systems randomly sampled. Note that the definition of "stability" here is based on systems that survive for a certain minimum period of time ($t_{\rm stable}$). For example, in the prograde case with $t_{\rm stable} = 10^9$, there are 98 stable systems and we sampled 98 unstable systems, so the total number of elements in the subset is $98 \times 2 = 196$. For prograde systems, the subset sizes are 436, 238, 212, 198, and 196 for $t_{\rm stable} = 10^5 P_{\rm in}$, $10^6 P_{\rm in}$, $10^7 P_{\rm in}$, $10^8 P_{\rm in}$, and $10^9 P_{\rm in}$, respectively. For retrograde systems, the corresponding sizes are 1232, 414, 366, 342, and 328.

First, we present the evaluation metrics for detecting both stable and unstable systems in table 2 and 3 for the cases of $t_{\rm stable} = 10^7 P_{\rm in}$, $10^8 P_{\rm in}$, and $10^9 P_{\rm in}$. The specific values for metrics reported in these tables are determined at a classification threshold for our power ratio criterion that is chosen to maximize the F-measure for prediction of stable system. This approach ensures we evaluate performance at an optimal balance between precision and recall, which is crucial for practical applications.

For the prediction of stable systems (table 2), the maximum F-measure is approximately 0.8. We can see that the Accuracy is very high, 0.98–0.99. Note that this value is overrated because our dataset is really askew to unstable cases, i.e., 95-97% of the systems are unstable. The precision is around 0.9 and recall rate is around 0.7, which represent the imbalanceness of our datasets. For the prediction of unstable systems (table 3), by definition, the accuracy is identical between Table 2 and Table 3. When instability is treated as the positive class, the maximum F-measure becomes higher—around 0.99—compared to the case where stability is treated as positive. This is also because the majority of the systems we prepared become unstable within $10^6 P_{\rm in}$, and are therefore labeled as unstable.

Since our dataset is highly imbalanced, we also evaluated our model using balanced subsets. Table 4 shows the value of metrics for our subsets. Each metric value represents the average over 10 randomly generated subsets for each cases. Crucially, the metrics reported in Table 4 are also derived from thresholds chosen to maximize the F-measure (for predicting stable systems) within these balanced conditions. In the balanced subset, the maximum F-measure exceeds 95%, indicating highly accurate classification. Furthermore, the standard deviation remains around 1%, indicating that the classification performance is consistent regardless of how unstable systems are sampled.

Next, we evaluate our model's performance using ROC curves. The ROC curve visualizes the TPR values as the threshold is progressively increased, thereby adjusting the FPR from 0 to 1. Before showing ROC curve, we show the TPR and Specificity (1-FPR) as functions of the threshold. Figure 8 shows the TPR and Specificity as functions of the threshold in the case for $t_{\rm stable} = 10^9 P_{\rm in}$ (one of the subsets used). Ideally, the TPR and Specificity should be as close to 1 as possible. In practice, the choice of a threshold must be determined by the user based on what type of problem the stability criterion is intended to address.

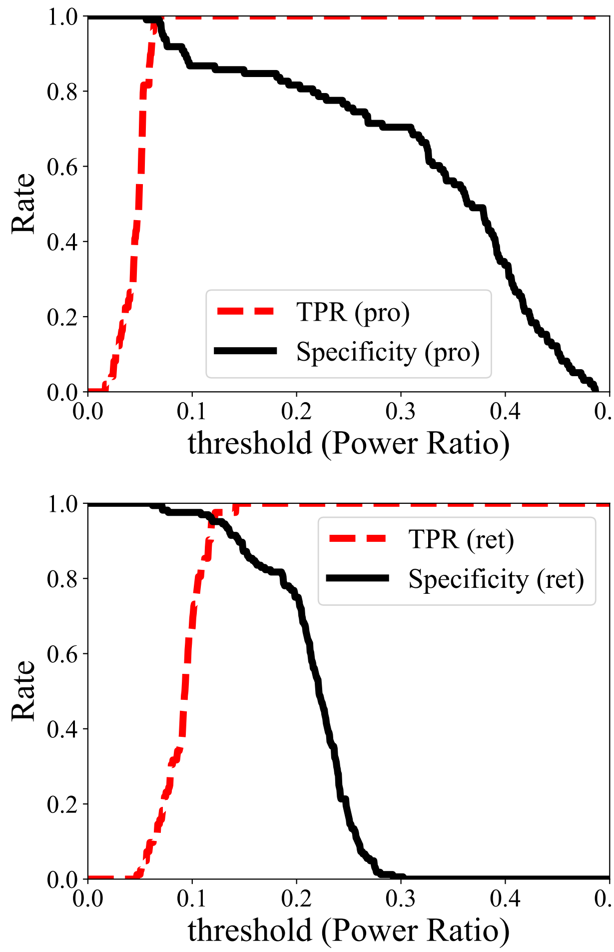

Fig. 8. The changes in TPR and Specificity (1-FPR) with respect to the threshold. The upper panel shows the case for prograde orbits, while the lower panel shows the case for retrograde orbits.
ALT text: Two line plots of the decision threshold versus true positive rate and specificity.

Figure 9 shows ROC curves for our results in the cases for $t_{\rm stable} = 10^5 P_{\rm in}$, $10^6 P_{\rm in}$, $10^7 P_{\rm in}$, $10^8 P_{\rm in}$, and $10^9 P_{\rm in}$. The horizontal axis represents the False Positive Rate (FPR), and the vertical axis represents the True Positive Rate (TPR). The numbers in the legend represent the value of area under the curve (AUC). The AUC is approximately 98-99%, except for the retrograde

Table 2. Metrics for the prediction of stable systems. The definitions of each parameter are provided in the main text.

	$t_{\rm stable}$	threshold	max F-measure	Accuracy	Precision	Recall (TPR)	FPR	Specificity
Prograde	$10^7 P_{\rm in}$	0.054	0.827	0.991	0.900	0.764	0.002	0.998
	$10^8 P_{\rm in}$	0.054	0.847	0.993	0.889	0.808	0.003	0.997
	$10^9 P_{\rm in}$	0.054	0.851	0.993	0.889	0.816	0.0/03	0.997
Retrograde	$10^7 P_{\rm in}$	0.106	0.810	0.977	0.905	0.732	0.006	0.994
	$10^8 P_{\rm in}$	0.105	0.823	0.980	0.914	0.749	0.005	0.995
	$10^9 P_{\rm in}$	0.105	0.842	0.982	0.914	0.780	0.005	0.995

Table 3. Metrics for the prediction of unstable systems. The same thresholds as table 2 are chosen.

	$t_{\rm stable}$	threshold	F-measure	Accuracy	Precision	Recall (TPR)	FPR	Specificity
Prograde	$10^7 P_{\rm in}$	0.054	0.996	0.991	0.994	0.998	0.236	0.764
	$10^8 P_{\rm in}$	0.054	0.996	0.993	0.995	0.997	0.192	0.808
	$10^9 P_{\rm in}$	0.054	0.996	0.993	0.995	0.997	0.184	0.816
Retrograde	$10^7 P_{\rm in}$	0.106	0.988	0.977	0.981	0.994	0.268	0.732
	$10^8 P_{\rm in}$	0.105	0.989	0.980	0.983	0.995	0.251	0.749
	$10^9 P_{\rm in}$	0.105	0.991	0.982	0.986	0.995	0.220	0.780

Table 4. Metrics with standard deviation for the subsets.

	$t_{\rm stable}$	threshold	max F-measure	Accuracy	Precision	Recall (TPR)	FPR	Specificity
P	$10^7 P_{\rm in}$	0.067 ± 0.002	0.987 ± 0.005	0.987 ± 0.005	0.983 ± 0.010	0.992 ± 0.003	0.017 ± 0.010	0.983 ± 0.010
	$10^8 P_{\rm in}$	0.066 ± 0.000	0.993 ± 0.009	0.992 ± 0.009	0.985 ± 0.017	1.000 ± 0.000	0.015 ± 0.018	0.985 ± 0.018
	$10^9 P_{\rm in}$	0.066 ± 0.001	0.991 ± 0.007	0.991 ± 0.007	0.985 ± 0.012	0.998 ± 0.006	0.015 ± 0.012	0.985 ± 0.012
R	$10^7 P_{\rm in}$	0.142 ± 0.000	0.969 ± 0.006	0.968 ± 0.007	0.941 ± 0.012	1.000 ± 0.000	0.063 ± 0.014	0.937 ± 0.014
	$10^8 P_{\rm in}$	0.134 ± 0.008	0.968 ± 0.009	0.968 ± 0.010	0.954 ± 0.016	0.983 ± 0.017	0.047 ± 0.018	0.953 ± 0.018
	$10^9 P_{\rm in}$	0.128 ± 0.009	0.978 ± 0.004	0.978 ± 0.004	0.973 ± 0.008	0.983 ± 0.011	0.027 ± 0.008	0.973 ± 0.008

$t_{\rm stable} = 10^5$ case. As shown in the survival time distribution in figure 1, the majority of systems are destroyed relatively early, while systems with survival times exceeding 10^6-$10^7 P_{\rm in}$ deviate from the majority. These AUC values indicate that the orbital evolution of systems with a lifetime longer than 10^6-$10^7 P_{\rm in}$ looks rather periodic and thus the low-frequency component is small.

While AUC provides a global measure of classification performance (Figure 9), its direct interpretation can be challenging for specific applications. For instance, AUC does not explicitly define an operational threshold or directly translate to the probability of a correct prediction for a given system. Therefore, to offer a more practical evaluation, we present metrics at a specific threshold chosen to maximize the F-measure. The F-measure, as the harmonic mean of precision and recall, represents a well-balanced criterion when these two metrics are in a trade-off relationship, providing a single value to optimize for a robust classification threshold.

In the context of N-body simulations, accurately identifying both stable and unstable hierarchical triple systems is important, though for different reasons. Correctly identifying stable systems as stable (high recall for the stable class) is essential for applying computationally efficient integration methods, thereby reducing overall computational cost. Conversely, misidentifying an unstable system as stable (low specificity for the stable class, or high false positive rate) could lead to the inappropriate application of such approximations and result in physically inaccurate simulation outcomes. For unstable systems, a high recall rate contributes to the efficiency of the N-body simulation by ensuring these typically short-lived systems are integrated directly and their evolution is accurately captured without unnecessary computational effort. High precision for unstable systems ensures that systems flagged for direct integration are indeed unstable, maintaining the reliability of the simulation. Given these considerations, the F-measure serves as a good overall metric for assessing the predictive power for identifying both stable and unstable systems. This is particularly important as our initial dataset is imbalanced; thus, evaluating performance with metrics like F-measure, especially on balanced subsets (as presented in Table 4), provides a more reliable assessment of our method's capabilities.

4 Discussion

4.1 Comparison with previous works and significance of our approach

Our findings demonstrate that using Q value to determine stability in the mixed-region is inadequate. It is consistent with the growing understanding of chaotic dynamics in few-body systems. While traditional Q-based criteria are widely employed, they inherently struggle within this complex parameter space where orbital outcomes are highly sensitive to inital conditions. Our simulations, which show a wide dispersion in survival times for systems with nearly identical inital Qvalues (Figure 1), underscore this limitation. Approximately 5% of systems integrated with initial Q values about 5% smaller than the MA01 critical value remained stable beyond $10^9 P_{\rm in}$, highlighting the significant population of long-lived systems that Q-based criteria alone might misclassify as unstable.

Approach that take into account the early orbital evolution of the system offer a promising avenue for more reliable stability assessment, particularly for identifying long-lived systems within the mixed-region that Q-based criteria may miss. LT22 adopted such an approach, using machine learning to predict lifetime from variations in orbital elements sampled up to $5 \times 10^5 P_{\rm in}$ at intervals of

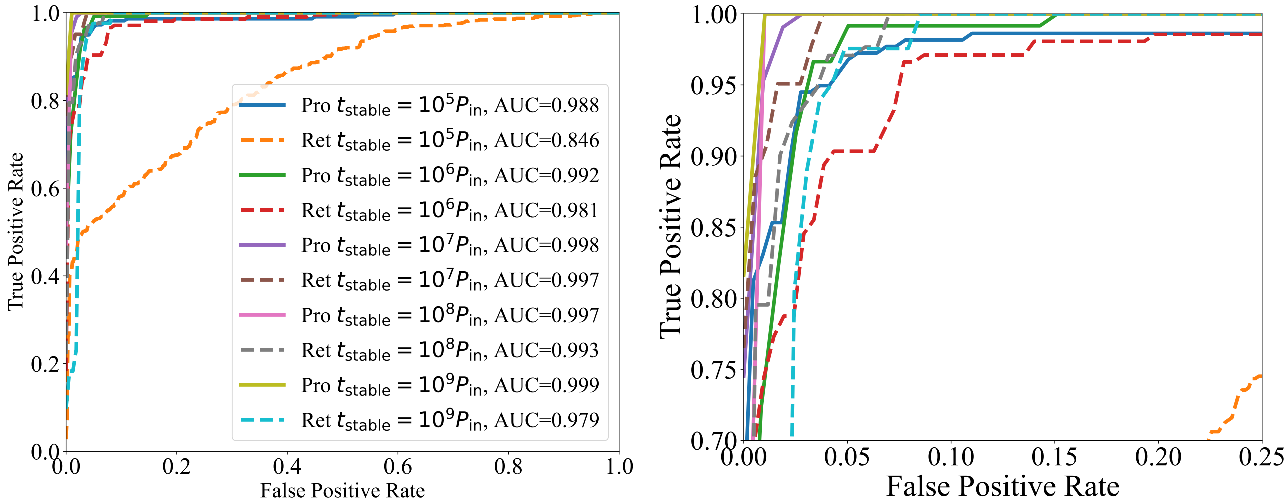

Fig. 9. ROC curves of prograde and retrograde cases. Each lines correspond to different survival time ($t_{\rm stable}$) and inclination (prograde and retrograde). The left panel is a overall look, and the right panel is a closer look at the upper-left corner of the left panel.
ALT text: ROC curve and its zoomed-in view.

$5\pi P_{\rm in}$, achieving an AUC score of around 0.95.

In contrast, our method focuses on a detailed Fourier analysis of a single orbital parameter (the semi-major axes ratio $a_{\rm in}/a_{\rm out}$) sampled at much finer intervals of $0.1 P_{\rm in}$ (approximately 100 times denser than LT22) but over a significantly shorter initial integration time of about $10^3 P_{\rm in}$. This dense, short-timespan analysis results in a significantly reduced computational cost for integration compared to LT22. Despite this efficiency, our Fourier-based approach, when evaluated on a balanced subset to ensure fiar comparison in the context of imbalanced original data, achieved high AUC scores (e.g., > 0.95 as shown in figure 9), demonstrating its strong predictive power.

The significance of our Fourier-based approach lies not only in its predictive accuracy and computational efficiency but also in its physical interpretability. By focusing on the periodicity of orbital motion, as reflected in the frequency spectrum (Figure 5 and 6), we tap into a fundamental dynamical characteristic. For instance, the observation that stabler systems tend to exhibit cleaner spectra with less power in low-frequency components (Figure 7) provides a more direct insight into the system's dynamical state than a "black-box" machine learning model might offer. A notable aspect of this study is the systematic quantification of the link between these early-phase spectral properties and long-term stability in hierarchical triple systems.

In summary, stability criteria based solely on the initial Q parameter, even with improved formulations, exhibit fundamental limitations in the mixed-region. Incorporating information from the system's orbital evolution, as explored by LT22 and the present study, is a more effective strategy for achieving robust stability classification. By analyzing orbital elements in frequency space, our work suggests that it is possible no only to enhance the accuracy of stability predictions but also to potentially gain deeper access to the undelying dynamical mechanisms that govern the stability of hierarchical three-body systems. This approach may contribute to more efficient and physically grounded stability assessments in computationally demanding N-body simulations.

4.2 Summary and future work

In this paper, we have carried out numerical simulations of hierarchical three-body systems in the mixed-region in order to explore a new method for determining the triple stability. Our main findings are as follows:

1. Hierarchical triple systems in the mixed-region exhibit a wide distribution of survival times. The majority of systems in the mixed-region break apart within $10^5 P_{\rm in}$ to $10^6 P_{\rm in}$, but some systems exhibit significantly longer survival times, exceeding $10^9 P_{\rm in}$. Therefore, the current stability criterion based on Q is inadequate for determining stability in the mixed-region.
2. There is a obvious difference in the orbital evolution between stable and unstable triples in the first $10^3 P_{\rm in}$. Stable triples exhibit more distinct peaks in the frequency space at the fundamental frequency and its integer multiples. Additionally, stable ones tend to have smaller low-frequency components. There is a correlation between the triple's instability and the stability of its orbit. It is remarkable that variations in orbital evolution already manifest within a mere $10^3 P_{\rm in}$ during the initial stages.
3. The stability of a system can be predicted using the proportion of low-frequency components relative to the total power. This method facilitates the identification of a limited number of systems that attain stability within the range of Q values situated near the vicinity of instability. Notably, our quantification relies solely on the low-frequency components. Combining this approach with other quantification methods is expected to enable predictions with even higher accuracy.

Here, we present potential future works. First, our findings are limited to coplanar cases, necessitating the verification of whether analogous discussions can be extended to inclination cases. It is also necessary to investigate how the results change when orbital parameters such as mass and eccentricity are varied. Second, it is plausible that incorporating eccentricity and other orbital parameters could yield a more precise prediction of stability. For stable systems, other orbital elements should also exhibit nearly periodic variations. Lastly, as a means to identify components that elude human visual perception, machine learning algorithms could be employed to analyze the FFT results and subsequently predict the stability of triples.

Acknowledgments

We thank the anonymous referee for helpful and valuable comments that helped to improve the quality of this paper. R. I. acknowledges support from JST SPRING, Grant Number JPMJSP2148. Our numerical integration was mainly carried out on the general-purpose PC cluster at the Center for Computational Astrophysics, National Astronomical Observatory of Japan (NAOJ), and the CPU cluster "CPLAB" in Kobe University.

Data availability

The data underlying this article will be shared on reasonable request to the corresponding author.

References

Aarseth, S. J. 2003, in IAU Symposium, Vol. 208, Astrophysical Supercomputing using Particle Simulations, ed. J. Makino & P. Hut, 295,
Dvorak, R. 1986, A&A, 167, 379
Eggleton, P., & Kiseleva, L. 1995, ApJ, 455, 640,
Georgakarakos, N. 2013, New Astronomy, 23, 41,
Harrington, R. S. 1972, Celestial Mechanics, 6, 322,
—. 1975, AJ, 80, 1081,
—. 1977, AJ, 82, 753,
Hayashi, T., Trani, A. A., & Suto, Y. 2022, ApJ, 939, 81,
—. 2023, ApJ, 943, 58,
Heggie, D., & Hut, P. 2003, The Gravitational Million-Body Problem: A Multidisciplinary Approach to Star Cluster Dynamics
Kozai, Y. 1962, AJ, 67, 591,
Lalande, F., & Trani, A. A. 2022, ApJ, 938, 18,
Lidov, M. L. 1962, Planet. Space Sci., 9, 719,
Mardling, R., & Aarseth, S. 1999, in NATO Advanced Study Institute (ASI) Series C, Vol. 522, The Dynamics of Small Bodies in the Solar System, A Major Key to Solar System Studies, ed. B. A. Steves & A. E. Roy, 385
Mardling, R. A., & Aarseth, S. J. 2001, MNRAS, 321, 398,
Mikkola, S., & Aarseth, S. J. 1996, Celestial Mechanics and Dynamical Astronomy, 64, 197,
Mikkola, S., & Tanikawa, K. 1999, MNRAS, 310, 745,
Mushkin, J., & Katz, B. 2020, MNRAS, 498, 665,
Mylläri, A., Valtonen, M., Pasechnik, A., & Mikkola, S. 2018, MNRAS, 476, 830,
Preto, M., & Tremaine, S. 1999, AJ, 118, 2532,
Trani, A. A., Leigh, N. W. C., Boekholt, T. C. N., & Portegies Zwart, S. 2024, A&A, 689, A24,
Trani, A. A., & Spera, M. 2023, in IAU Symposium, Vol. 362, The Predictive Power of Computational Astrophysics as a Discover Tool, ed. D. Bisikalo, D. Wiebe, & C. Boily, 404–409,
Valtonen, M., Mylläri, A., Orlov, V., & Rubinov, A. 2008, in IAU Symposium, Vol. 246, Dynamical Evolution of Dense Stellar Systems, ed. E. Vesperini, M. Giersz, & A. Sills, 209–217,
von Zeipel, H. 1910, Astronomische Nachrichten, 183, 345,
Vynatheya, P., Hamers, A. S., Mardling, R. A., & Bellinger, E. P. 2022, MNRAS, 516, 4146,
Wang, L., Nitadori, K., & Makino, J. 2020, MNRAS, 493, 3398,
Yoshida, H. 1990, Physics Letters A, 150, 262,