

# Hill stability in the AMD framework

Antoine C. Petit, Jacques Laskar, and Gwenaél Boué

IMCCE, CNRS-UMR8028, Observatoire de Paris, PSL University, Sorbonne Université, 77 Avenue Denfert-Rochereau,  
75014 Paris, France  
e-mail: [antoine.petit@obspm.fr](mailto:antoine.petit@obspm.fr)

Received 23 March 2018 / Accepted 19 June 2018

## ABSTRACT

In a two-planet system, a topological boundary that is created by [Sundman \(1912, Acta Math., 36, 105\)](#) inequality can forbid close encounters between the two planets for an infinite time. A system is said to be Hill stable if it verifies this topological condition. Hill stability is widely used in the study of extrasolar planet dynamics. However, the coplanar and circular orbit approximation is often used. In this paper, we explain how the Hill stability can be understood in the framework of angular momentum deficit (AMD). In the secular approximation, AMD allows us to discriminate between a priori stable systems and systems for which a more in-depth dynamical analysis is required. We show that the general Hill stability criterion can be expressed as a function of only semimajor axes, masses, and total AMD of the system. The proposed criterion is only expanded in the planets-to-star mass ratio  $\varepsilon$  and not in the semimajor axis ratio, eccentricities, nor the mutual inclination. Moreover, the expansion in  $\varepsilon$  remains excellent up to values of about  $10^{-3}$  even for two planets with very different mass values. We performed numerical simulations in order to highlight the sharp change of behavior between Hill stable and Hill unstable systems. We show that Hill stable systems tend to be very regular, whereas Hill unstable systems often lead to rapid planet collisions. We also note that Hill stability does not provide protection from the ejection of the outer planet.

**Key words.** planets and satellites: general – planets and satellites: dynamical evolution and stability – celestial mechanics

## 1. Introduction

The chaotic nature of planetary dynamics has made the development of stability criteria necessary because the complete study of individual systems for the lifetime of the central star would require too much computational power. Moreover, in the context of exoplanet dynamics, the orbital parameters are often not known with great precision, making it impossible to conduct a precise dynamical study. For multiplanetary systems, the best solutions yet are empirical stability criteria based on minimum spacing between planets obtained from numerical simulations ([Chambers et al. 1996](#); [Pu & Wu 2015](#)). For tightly packed systems, [Tamayo et al. \(2016\)](#) have suggested a solution based on short integrations and a machine-learning algorithm.

Another approach consists in the study of the angular momentum deficit (AMD; see [Laskar 1997, 2000](#)). The AMD is a weighted sum of the eccentricities and mutual inclinations of planets and can be interpreted as a dynamical temperature of the planetary system. The AMD is conserved at all orders of averaging over the mean motions. In [Laskar & Petit \(2017\)](#), it was shown that if the AMD is small enough, collisions are impossible. We can define a sufficient stability condition from the secular conservation of the AMD, i.e., the AMD-stability. Systems that are AMD-stable are long-lived, whereas a more in-depth dynamical study is necessary for AMD-unstable systems.

The initial AMD-stability definition is based on the secular approximation. In [Petit et al. \(2017\)](#), the criterion was slightly modified to exclude systems experiencing short-term chaos due to first-order mean motion resonances (MMR) overlap. The MMR overlap is the main source of chaos and instability in planetary dynamics. Based on [Chirikov \(1979\)](#) resonance overlap criterion, [Wisdom \(1980\)](#) has derived an analytical stability criterion for the two-planet systems that has since been widely

used. Wisdom's criterion is obtained for circular and coplanar planets. It was improved to take into account moderate eccentricities ([Mustill & Wyatt 2012](#); [Deck et al. 2013](#)). We have shown in [Petit et al. \(2017\)](#) that this MMR overlap criterion can be expressed solely as a function of the AMD, semimajor axes, and masses for almost coplanar and low eccentricity systems.

The MMR overlap criterion gives a clear limit between regular and chaotic orbits for small eccentricities. However, this criterion is based on first-order expansions in the planets-to-star mass ratio, the spacing between planets, eccentricities, and inclinations. As a result, for wider orbit separations, the secular collision criterion from [Laskar & Petit \(2017\)](#) remains a better limit ([Petit et al. 2017](#)). Moreover, the MMR overlap criterion only takes into account the interaction between a couple of planets, making it really accurate only for two-planet systems.

In the case of two-planet systems, the topology of the phase space gives a far simpler criterion of stability, that is the Hill stability. Based on a work by [Sundman \(1912\)](#) on the moment of inertia in the three-body problem, [Marchal & Saari \(1975\)](#) noted the existence of forbidden zones in the configuration space. In [Marchal & Bozis \(1982\)](#), they extended the notion of Hill stability to the general three-body problem and showed that some systems can forbid close encounters between the outer body and any of the inner bodies. In particular, the Hill stability ensures that collisions between the outer body and the close binary (in a two-planet system, between the outer planet and the inner planet or the star) are impossible for infinite time.

The results of [Marchal & Saari \(1975\)](#) and [Marchal & Bozis \(1982\)](#) have many applications outside of the Hill stability. It is possible to cite a sufficient condition for the ejection of a body from the system ([Marchal et al. 1984a,b](#)) or the determination of

the limit of a triple close approach for bounded orbits (Laskar & Marchal 1984).

Marchal & Bozis (1982) presented the planetary problem (one body with a much larger mass) as a particular case, but the result was mainly popularized by Gladman (1993) who introduced a minimal spacing for initially circular and coplanar systems. Gladman also proposed some criteria for eccentric orbits in some particular configurations of masses. The result by Gladman was refined to cover other situations as, for example, the case of inclined orbits (Veras & Armitage 2004). Georgakarakos (2008) provided a review of stability criteria for hierarchical three-body problems.

The Hill stability is consistent with numerical integrations, where a sharp transition between Hill stable and Hill unstable systems is often observed (Gladman 1993; Barnes & Greenberg 2006; Deck et al. 2013). Moreover, Deck et al. (2013) analyzed the differences between the MMR overlap and Hill stability criteria. They remarked that there exists an area where orbits are chaotic because of the overlap of MMR but long-lived owing to Hill stability.

In this paper, we show that Marchal & Bozis (1982) Hill stability criterion fits extremely well in the AMD-stability framework. In Sect. 2, we derive a criterion for Hill stability solely expressed as a function of the total AMD, semimajor axes, and masses. Our criterion does not need any expansion in the spacing, eccentricities, or inclinations of the orbits and admits all previous Hill criteria as particular approximations. In order to do so, we follow the reasoning proposed by Marchal & Bozis (1982).

We then compare the Hill stability criterion with the AMD-stability criteria proposed in Laskar & Petit (2017) and Petit et al. (2017). In the last section, we carry numerical integrations of two-planet systems over a large part of the phase space. We show that the only parameters of importance are the initial AMD and semimajor axis ratio.

## 2. Hill stability in the three-body problem

### 2.1. Generalized Hill curves

Let us use the formalism proposed in Marchal & Saari (1975) and Marchal & Bozis (1982) for the definition of the Hill regions for the general three-body problem. We mainly consider the planetary case and therefore slightly adapt their notations to this particular problem. Let us consider two planets of masses  $m_1, m_2$  orbiting a star of mass  $m_0$ . Let  $\mathcal{G}$  be the constant of gravitation,  $\mu = \mathcal{G}m_0$ ,

$$\varepsilon = \frac{m_1 + m_2}{m_0}, \quad (1)$$

the planet mass to star mass ratio and

$$\gamma = \frac{m_1}{m_2}. \quad (2)$$

As in Laskar & Petit (2017), we use the heliocentric canonical coordinates  $(\tilde{\mathbf{r}}_j, \mathbf{r}_j)_{j=1,2}$ . In those coordinates, the Hamiltonian is

$$\mathcal{H} = \sum_{j=1}^2 \left( \frac{1}{2} \frac{\|\tilde{\mathbf{r}}_j\|^2}{m_j} - \frac{\mu m_j}{r_j} \right) + \frac{1}{2} \frac{\|\tilde{\mathbf{r}}_1 - \tilde{\mathbf{r}}_2\|^2}{m_0} - \frac{\mathcal{G}m_1 m_2}{r_{12}}, \quad (3)$$

where  $r_{12} = \|\mathbf{r}_1 - \mathbf{r}_2\|$ . We denote  $\mathbf{G}$  the total angular momentum of the system:

$$\mathbf{G} = \sum_{j=1}^2 \mathbf{r}_j \times \tilde{\mathbf{r}}_j, \quad (4)$$

and  $G$  its norm. From now, we assume  $\mathbf{G}$  to be aligned along the  $z$ -axis. We also define

$$\Lambda_j = m_j \sqrt{\mu a_j}, \quad (5)$$

where  $a_j$  is the semimajor axis of the  $j$ th planet. We also note  $e_j$ , the eccentricity, and  $i_j$ , the inclination of the orbital plane with the horizontal plane. The AMD,  $C$  (Laskar 1997, 2000) is expressed

$$C = \sum_{j=1}^2 \Lambda_j \left( 1 - \sqrt{1 - e_j^2} \cos(i_j) \right). \quad (6)$$

Following Marchal & Bozis (1982), we define a generalized semimajor axis:

$$a = -\frac{\mathcal{G}M^*}{2\mathcal{H}}, \quad (7)$$

and a generalized semilatus rectum:

$$p = \frac{m_0(1 + \varepsilon)}{\mathcal{G}M^{*2}} G^2, \quad (8)$$

where

$$M^* = m_0 m_1 + m_0 m_2 + m_1 m_2 = m_0^2 \varepsilon \left( 1 + \varepsilon \frac{\gamma}{(1 + \gamma)^2} \right). \quad (9)$$

The values  $a$  and  $p$  are the two length units that can be built from the first integrals  $\mathcal{H}$  and  $G$ .

We finally define two variable lengths. First,  $\rho$  the mean quadratic distance

$$M^* \rho^2 = m_0 m_1 r_1^2 + m_0 m_2 r_2^2 + m_1 m_2 r_{12}^2, \quad (10)$$

that is proportional to the moment of inertia  $I$  computed in the center of mass frame (Marchal & Bozis 1982):

$$I = \frac{1}{2} \frac{M^* \rho^2}{m_0 + m_1 + m_2}. \quad (11)$$

We also define  $\nu$ , the mean harmonic distance:

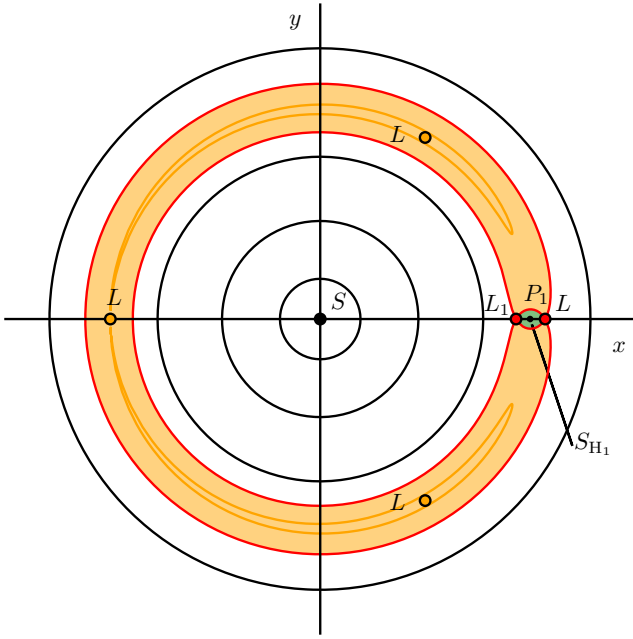
$$\frac{M^*}{\nu} = \frac{m_0 m_1}{r_1} + \frac{m_0 m_2}{r_2} + \frac{m_1 m_2}{r_{12}}, \quad (12)$$

which is proportional to the potential energy:

$$U = -\frac{\mathcal{G}M^*}{\nu}. \quad (13)$$

For a system with given  $\mathcal{H}$  and  $G$ , some configurations of the planets are forbidden. Indeed, the value of the ratio  $\rho/\nu$  is constrained by the inequality (Marchal & Saari 1975):

$$\frac{\rho}{\nu} \geq \frac{\rho}{2a} + \frac{p}{2\rho}, \quad (14)$$



**Fig. 1.** Some levels of the function  $R$  defined in (16) for  $\varepsilon = 10^{-3}$  and  $\gamma = 1$  in the  $\mathcal{P}$  plane. The two red points correspond to the Lagrange points  $L_1$  and  $L_2$  and the three orange points to  $L_3$ ,  $L_4$ , and  $L_5$ . The orange-filled area corresponds to the region where  $R(x, y) < R(L_1)$ . The star  $S$  is at the origin, the planet  $P_1$  at  $(1, 0)$ , and  $(x, y)$  are the coordinates of the second planet. The green region  $S_{H_1}$  represents the Hill sphere of the first planet.

derived from Sundman's inequality (Sundman 1912). Moreover, if the system has a negative energy, the right-hand side of Eq. (14) has a minimum value obtained for  $\rho = \sqrt{a\bar{p}}$ . We therefore have the inequality (Marchal & Bozis 1982):

$$\frac{\rho^2}{v^2} \geq \frac{p}{a} = -\frac{2m_0(1+\varepsilon)}{\mathcal{G}^2 M^*{}^3} \mathcal{H} G^2. \quad (15)$$

If  $p/a$  is high enough, the inequality (15) makes some regions of the phase space inaccessible. In this case, we can ensure that certain initial conditions forbid collisions between the two planets for all times.

Let us study the values of the function  $(\rho/v)^2$ . Since this ratio only depends on the ratios of mutual distances, we can always place ourselves in the plane generated by the three bodies. We can also choose to place the first planet on the  $x$ -axis and normalize the lengths by  $r_1$ . Let us call this plane  $\mathcal{P}$  and note  $(x, y)$  the coordinates of the second planet. In the plane  $\mathcal{P}$  (see Fig. 1), the star  $S$  is at the origin, the first planet  $P_1$  is situated at the point  $(1, 0)$ . For the planar restricted three-body problem, this reduction is equivalent to study the dynamics in the corotating frame. We note

$$R(x, y) = \left(\frac{\rho}{v}\right)^2. \quad (16)$$

The shape of the function  $R$  in  $\mathcal{P}$  only depends on the mass distribution, for example, the two ratios  $\varepsilon$  and  $\gamma$ . In Fig. 1, we can see level curves of the function  $R$  plotted in  $\mathcal{P}$  for  $\varepsilon = 10^{-3}$  and  $\gamma = 1$ . The value  $R$  is minimal and equal to 1 at the two Lagrange points  $L_4$  and  $L_5$  and goes to  $+\infty$  for  $\|\mathbf{r}_2\| \rightarrow 0$ ,  $\|\mathbf{r}_2 - \mathbf{r}_1\| \rightarrow 0$  or  $\|\mathbf{r}_2\| \rightarrow +\infty$ . The function has three saddle nodes at the Lagrange points  $L_1$ ,  $L_2$ , and  $L_3$ . We give the method to compute the position of the Lagrange points in Appendix A. Let

us denote  $x_1$ ,  $x_2$ , and  $x_3$  their abscissa, respectively. At the lowest order in  $\varepsilon$ , we have

$$\begin{aligned} x_1 &= 1 - \left(\frac{\varepsilon}{3}\right)^{1/3} + O(\varepsilon^{2/3}), \\ x_2 &= 1 + \left(\frac{\varepsilon}{3}\right)^{1/3} + O(\varepsilon^{2/3}), \\ x_3 &= -1 + \frac{7}{12} \frac{\gamma - 1}{\gamma + 1} \varepsilon + O(\varepsilon^2). \end{aligned} \quad (17)$$

At the lowest order in  $\varepsilon$ , the value of  $R$  at the Lagrange points are as follows:

$$\begin{aligned} R(L_1) &= 1 + 3^{4/3} \varepsilon^{2/3} \gamma / (\gamma + 1)^2 + O(\varepsilon), \\ R(L_2) &= 1 + 3^{4/3} \varepsilon^{2/3} \gamma / (\gamma + 1)^2 + O(\varepsilon), \\ R(L_3) &= 1 + 2\varepsilon\gamma / (\gamma + 1)^2 + O(\varepsilon^2). \end{aligned} \quad (18)$$

The values  $R(L_1)$  and  $R(L_2)$  have the same first-order term but differ in the expansion of higher order. Indeed, if  $m_0 \geq m_1 \geq m_2$ , we have  $R(L_1) \geq R(L_2)$  (Marchal & Bozis 1982). From now on, let us assume  $R(L_1) \geq R(L_2)$  (if not, we can just substitute  $R(L_2)$  to  $R(L_1)$  in further equations).

For  $p/a \geq R(L_1)$ , the accessible domain is split into three parts: the Hill sphere of the star  $S$ , which is around the origin; the first planet Hill sphere<sup>1</sup>  $S_{H_1}$  (in green in Fig. 1); and the outer region. In this case, if the second planet is not initially inside  $S_{H_1}$ , it will never be able to enter this region. However, if  $P_2$  is in the outer region, the Hill stability cannot constrain the possibility of ejection. Similarly, if  $P_2$  is closer to the star (in the inner region), a collision with the star is still possible.

The study of the function  $R$  and the inequality (15) gives a noncollision criterion for an infinite time. Marchal and Bozis called it the Hill stability.

**Proposition 1 (Marchal & Bozis 1982).** *Let us consider a negative energy three-body problem with a body  $S$  of mass  $m_0$  and two others  $P_1$  and  $P_2$  of mass  $m_1$  and  $m_2$  such that  $m_0 \geq m_1 \geq m_2$ . We place ourselves in the  $\mathcal{P}$  plane defined by  $S$ ,  $P_1$ , and  $P_2$  (Fig. 1). If  $P_2$  is not initially inside the Hill sphere  $S_{H_1}$  of  $P_1$ , the system is Hill stable if*

$$\frac{p}{a} > R(L_1), \quad (19)$$

where  $a$  is defined in (7),  $p$  in (8), and  $R$  in (16).

From this inequality, Gladman (1993) obtained criteria for initially circular orbits and for two particular cases of eccentric orbits: the case of equal masses and small eccentricities and the case of equal masses and large, but equal eccentricities.

While Gladman's Hill stability criterion for initially circular orbits is useful, the eccentric criteria are too particular to be used in the context of a generic system. It is, however, possible to obtain a very general Hill stability criterion using the AMD to take into account the eccentricities and inclinations of the orbits.

## 2.2. AMD condition for Hill stability

The total energy of the system can be written

$$\mathcal{H} = -\frac{m_2^3 \mu^2}{2\Lambda_2^2} \left( \frac{\gamma}{\alpha} + 1 + h_1 \right), \quad (20)$$

<sup>1</sup> The Hill region is usually called the Hill sphere, although it is not technically a sphere.

where  $\alpha = a_1/a_2$  and

$$h_1 = -\frac{2\Lambda_2^2}{m_2^3\mu^2} \left( \frac{1}{2} \frac{\|\tilde{\mathbf{r}}_1 + \tilde{\mathbf{r}}_2\|^2}{m_0} - \frac{\mathcal{G}m_1m_2}{r_{12}} \right). \quad (21)$$

From now on, we assume that initially  $\alpha \leq 1$  (if not we can just renumber the two planets). Similarly, the angular momentum can be rewritten:

$$G = \Lambda_2 (\gamma \sqrt{\alpha} + 1 - \mathcal{C}), \quad (22)$$

where

$$\mathcal{C} = \frac{C}{\Lambda_2} = \gamma \sqrt{\alpha} \left( 1 - \sqrt{1 - e_1^2 \cos i_1} \right) + 1 - \sqrt{1 - e_2^2 \cos i_2} \quad (23)$$

is the relative AMD (Laskar & Petit 2017). Combining Eqs. (15), (20), and (22), we obtain

$$\frac{p}{a} = \frac{(1 + \varepsilon)}{(\gamma + 1)^3 (1 + \varepsilon\gamma/(\gamma + 1)^2)^3} \left( \frac{\gamma}{\alpha} + 1 + h_1 \right) (\gamma \sqrt{\alpha} + 1 - \mathcal{C})^2. \quad (24)$$

The Hill stability criterion (19) can be rewritten without any approximation as a condition on  $\mathcal{C}$  and we have the following formulation of the Hill stability.

**Proposition 2** (Hill stability). *With the hypotheses of the proposition 1, assuming the elliptical elements can be defined (i.e., both Keplerian energies are negative), a system is Hill stable if*

$$\mathcal{C} < C_c^{\text{Ex}} = \gamma \sqrt{\alpha} + 1 - (\gamma + 1)^{3/2} \sqrt{\frac{R(L_1)(1 + \varepsilon\gamma/(\gamma + 1)^2)^3}{(1 + \varepsilon)(\gamma/\alpha + 1 + h_1)}}, \quad (25)$$

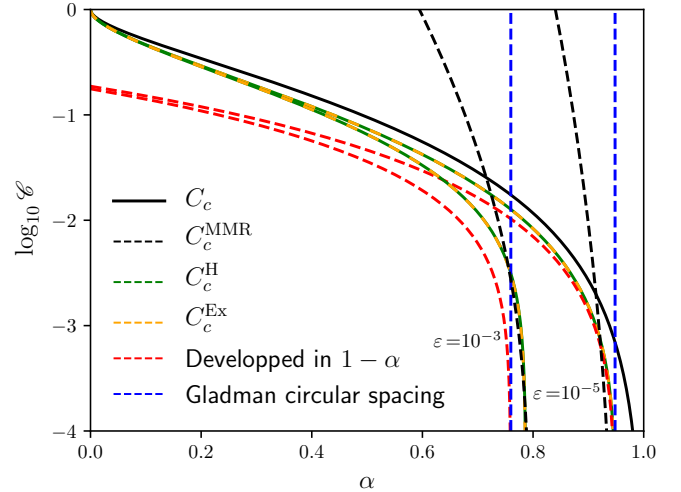
where  $\mathcal{C}$  is the relative AMD (23) and  $h_1$  is the normalized perturbation part (21).

The inequality (25) is equivalent to Proposition 1, but we isolated the contribution from the AMD on the left-hand side. Up to the perturbation term  $h_1$ , the right-hand side of Eq. (25) only depends on the masses and the semimajor axis ratio  $\alpha$ . If we only keep the terms of leading order in  $\varepsilon$  in the square root of the right-hand side of Eq. (25), we obtain an expression that depends only on  $\alpha$ ,  $\varepsilon$ , and  $\gamma$ .

**Proposition 3** (Hill stability, planetary case). *For small enough  $\varepsilon$ , a two-planet system is Hill stable if the relative AMD  $\mathcal{C}$  verifies the inequality:*

$$\mathcal{C} < \gamma \sqrt{\alpha} + 1 - (1 + \gamma)^{3/2} \sqrt{\frac{\alpha}{\gamma + \alpha} \left( 1 + \frac{3^{4/3} \varepsilon^{2/3} \gamma}{(1 + \gamma)^2} \right)} + \text{O}(\varepsilon). \quad (26)$$

As explained in Appendix B,  $h_1$  is of smaller order in  $\varepsilon$  and can be neglected if the criterion is verified. We want to stress that the expression (26) is obtained with only an expansion in  $\varepsilon$  and only depends on  $\alpha$ ,  $\mathcal{C}$ , and the masses of the bodies. The term of order  $\varepsilon^{2/3}$  in Eq. (26) also depends on  $\gamma/(\gamma + 1)^2$ , but we show in Appendix B that (26) is still valid for  $\gamma \ll 1$  or  $\gamma \gg 1$ .



**Fig. 2.** Comparison of the right-hand sides of the inequalities – Eq. (25) (orange), Eq. (26) (green), and Eq. (27) (red) – as a function of  $\alpha$  for  $\gamma = 1$  and  $\varepsilon = 10^{-5}$  (upper three curves) or  $\varepsilon = 10^{-3}$  (lower three curves). The black curve is the critical AMD for the collision condition (Laskar & Petit 2017). The green and orange curves are on top of each other (see Fig. 3). The critical AMD from MMR overlap  $C_c^{\text{MMR}}$  and circular criterion of Gladman (1993) are also plotted for comparison.

### 2.3. Close planets approximation

Assuming  $1 - \alpha \ll 1$ , and a small AMD value, further approximations can be made. At leading order in  $\mathcal{C}$ ,  $1 - \alpha$ , and  $\varepsilon$ , the inequality (26) becomes

$$\mathcal{C} < \frac{3\gamma}{8(\gamma + 1)} (1 - \alpha)^2 - \frac{3^{4/3} \gamma}{2(\gamma + 1)} \varepsilon^{2/3}. \quad (27)$$

We can isolate  $1 - \alpha$  in this expression to obtain an approximate minimum spacing for Hill stable systems.

**Proposition 4** (Hill stability, close planets case). *For a close planets system, the minimum spacing criterion for Hill stability is*

$$\frac{a_2 - a_1}{a_2} = 1 - \alpha > \sqrt{4 \times 3^{1/3} \varepsilon^{2/3} + \frac{8}{3} \frac{\gamma + 1}{\gamma} \mathcal{C}}. \quad (28)$$

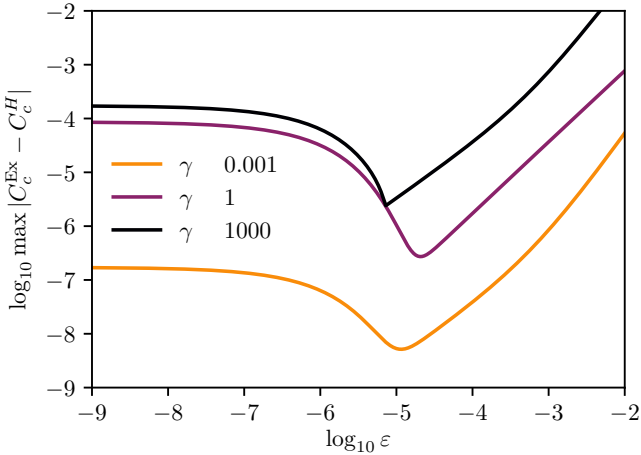
Gladman's eccentric criteria can be recovered from Eq. (28) if the AMD is developed under the assumptions (same planet masses, small or large, and equal eccentricities) made in Gladman (1993). However, Eq. (28) is more general as it takes into account mutual inclinations or uneven mass distribution.

In the case of circular orbits, we also get the well-known formula (Gladman 1993):

$$1 - \alpha > 2 \times 3^{1/6} \varepsilon^{1/3} = 2.40 \varepsilon^{1/3}. \quad (29)$$

### 2.4. Comparison of the Hill criteria

We can compare the right-hand side of Eqs. (25)–(27), and Gladman's circular approximation (29) to test how relevant are the approximations made here. In Fig. 2, we plot the exact expression  $C_c^{\text{Ex}}$  (Eq. (25); in green), the expansion in  $\varepsilon$  (Eq. (26); in orange), the approximation for close planets (Eq. (27); in red), and the minimum spacing for circular orbits (Eq. (29); in blue). We see that the expansion in  $\varepsilon$  (Eq. (26)) cannot be distinguished from the exact curve (Eq. (25)). In order to better quantify this,



**Fig. 3.** Maximum difference between  $C_c^{\text{Ex}}$  (Eq. (25)) and  $C_c^{\text{H}}$  (Eq. (30)) on the right-hand sides of Eqs. (25) and (26), respectively, as a function of  $\varepsilon$  for various values of  $\gamma = m_1/m_2$ . In Eq. (25), the term  $h_1$  is evaluated using the approximation of the kinetic term given in Appendix B.

we plot in Fig. 3 the maximum difference between the two curves as a function of  $\varepsilon$  for various values of the mass ratio  $\gamma$ .

We see in Fig. 3 that for the range of  $\varepsilon$  used in planetary dynamics (typically from  $10^{-6}$ – $10^{-2}$ ), the expression (26) developed in  $\varepsilon$  is accurate even for very uneven planet mass distribution. From now, we use Eq. (26) to define the Hill stability.

### 3. Comparison with the AMD-stability

In Laskar & Petit (2017), the AMD-stability of a system is defined by comparing its AMD to a critical value  $C_c$  above which the stability of the system cannot be guaranteed. The AMD-stability is a sufficient condition for long-lived stability of a system. The result from Laskar & Petit (2017) was obtained by finding the minimal AMD needed for a system to allow for collisions in the secular approximation, i.e., assuming the semimajor axes are fixed. We completed, in Petit et al. (2017), the AMD-stability definition by excluding the configurations such that the overlap of first-order MMR occurs. The modified critical AMD is hereafter noted  $C_c^{\text{MMR}}$ .

We can use Eq. (26) to define a critical AMD  $C_c^{\text{H}}$  for the Hill stability:

$$C_c^{\text{H}} = \gamma \sqrt{\alpha} + 1 - (1 + \gamma)^{3/2} \sqrt{\frac{\alpha}{\gamma + \alpha} \left( 1 + \frac{3^{4/3} \varepsilon^{2/3} \gamma}{(1 + \gamma)^2} \right)}. \quad (30)$$

A system is Hill stable if its initial relative AMD is smaller than the initial critical AMD  $C_c^{\text{H}}$ .

We see that for two planets, the Hill stability definition fits extremely well in the AMD-stability framework. We can also compare  $C_c^{\text{H}}$  to the previously proposed critical AMD. The collision critical AMD  $C_c$  is plotted in Fig. 2 with two values of  $C_c^{\text{H}}$  for  $\varepsilon = 10^{-5}$  (resp.  $10^{-3}$ ). We can see that the Hill stability criterion is stricter ( $C_c^{\text{H}} < C_c$ ) than the collision condition for secular dynamics. It can be easily understood since the Hill stability forbids the planets to approach each other.

Indeed, let us consider a Hill stable system, i.e., such that  $\mathcal{C} < C_c^{\text{H}}$ . As a result, the two planets cannot approach each other by less than their mutual Hill radius for any variation of semimajor axes and thus also in the secular system. In particular, a configuration such that the two orbits intersect is impossible.

Therefore, the system is AMD-stable and  $\mathcal{C} < C_c$ . Since we are not making any additional hypothesis regarding  $\mathcal{C}$ , we have  $C_c^{\text{H}} < C_c$ . The strict inequality comes from the positive minimal distance between the two planets.

As a comparison, we also plot in Fig. 2 the MMR critical AMD  $C_c^{\text{MMR}}$ . For small relative AMD,  $\mathcal{C}$ ,  $C_c^{\text{MMR}}$ , and  $C_c^{\text{H}}$  are almost identical.

### 4. Numerical simulations

The Hill criterion proposed by Marchal & Bozis (1982) has already been tested numerically in particular cases (Gladman 1993; Veras & Armitage 2004; Barnes & Greenberg 2006). In their comparison between the Hill and the overlap of MMR criteria, Deck et al. (2013) noted a sharp transition in the proportion of chaotic orbits at the Hill limit  $(p/a) = (p/a)|_c$ . It also appears that for small  $\varepsilon$  and  $\alpha$  close to 1, the overlap of MMR criterion provides a better limit for the chaotic region.

We want to test whether the Hill criterion gives a good limit to the chaotic region for wider separations. Moreover, we want our initial conditions to sample homogeneously the phase space. Indeed, the Hill stability criterion studied in this paper only depends on few quantities, the relative AMD  $\mathcal{C}$ , and the ratio of semimajor axis and not the angles or the actual distribution of the AMD between the degrees of freedom of eccentricities or inclinations. Choosing an homogeneous sampling of the initial conditions also avoids giving too much importance to regions protected by MMR owing to particular combinations of angles while another choice of angles would have given an unstable orbit.

#### 4.1. Numerical setup

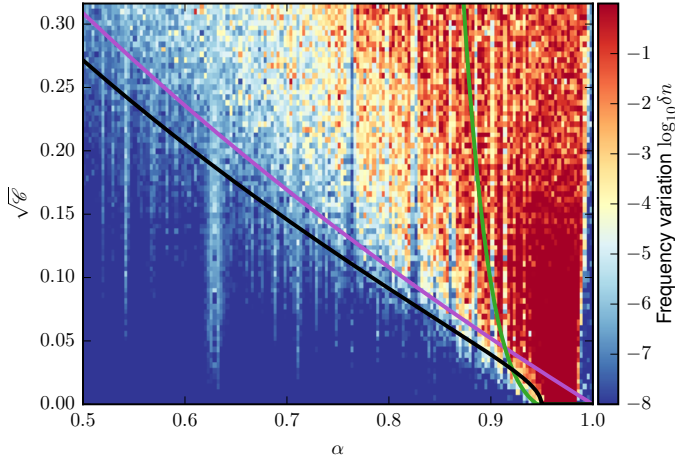
We ran numerical simulations using the symplectic scheme ABAH1064 from Farrés et al. (2013). We chose our initial conditions such that

- the outer planet semimajor axis  $a_2$  is fixed at 1 au;
- the AMD and inner planet semimajor axis  $a_1$  are chosen such that, we have a regular grid in the plane  $(\alpha, \sqrt{\mathcal{C}})$ . Such a scaling in  $\mathcal{C}$  is chosen to have an approximately uniform distribution in terms of eccentricities and inclination;
- the AMD is on average equipartitioned between the eccentricity and inclination degrees of freedom;
- the inclinations are chosen such that the angular momentum is on the  $z$ -axis;
- the angles are chosen randomly;
- the star mass is taken as  $1 M_\odot$  and the planets masses do not vary for each grid of initial conditions.

We then integrate each initial condition for 500 kyr using a time-step of  $10^{-3}$  yr. The numerical integration is stopped if the planets approach each other by less than a quarter of their mutual Hill radius, if a planet reaches  $10^{-2}$  or 20 AU or if the relative variation of energy is higher than  $10^{-8}$ .

In order to measure the stability of a system, we use frequency map analysis (Laskar 1990, 1993). Our criterion is based on the relative variation of the main frequencies in the quasiperiodic best fit. More precisely, let  $n_k^{(i)}$  (respectively  $n_k^{(f)}$ ) be the frequency obtained by frequency analysis for the planet  $k$  for the first (resp. last) 100 kyr of integration. We consider an orbit to be chaotic if

$$\delta n = \max_k \left| \frac{n_k^{(f)} - n_k^{(i)}}{n_k^{(i)}} \right| \quad (31)$$



**Fig. 4.** Frequency variation  $\delta n$  (31) for 1 00 000 initial conditions binned in a  $160 \times 75$  grid with  $\varepsilon = 10^{-5}$  and  $\gamma = 1$ . The black curve is the Hill critical AMD  $C_c^H$  (30), the purple curve is the collisional critical AMD  $C_c$  (Laskar & Petit 2017), and the green curve is the critical AMD obtained from the overlap of first-order MMR  $C_c^{\text{MMR}}$  (Petit et al. 2017). Each bin is average over about eight initial conditions.

is greater than  $10^{-4}$ . The chosen threshold is such that the variation of semimajor axis changes by about 1% in a few Gyr if we assume a constant diffusion process as it would happen for a random walk. Indeed,  $\delta n$  measures the variation of frequency over 500 kyr. If the diffusion rate remains constant, a variation on the order of 1% on average needs a time 10 000 times larger, i.e., 5 Gyr.

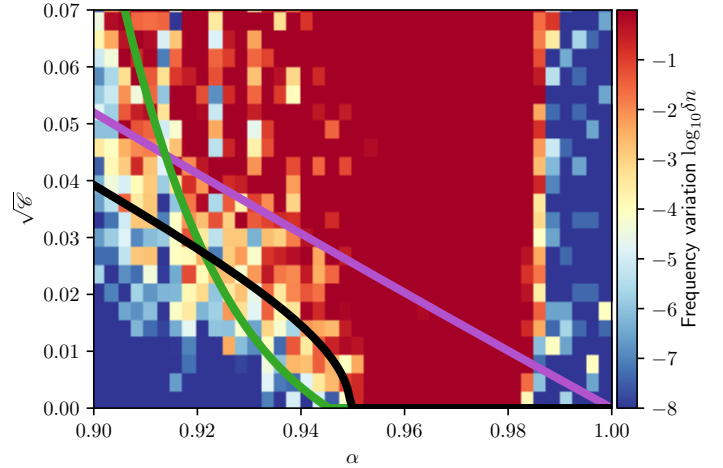
If the integration time is shorter due to a collision or ejection, we set  $\delta n$  to 1. Since we randomly draw most of the initial parameters, we bin the results into a two-dimensional grid in  $(\alpha, \sqrt{\mathcal{C}})$  and average the frequency variation in each bin.

#### 4.2. Results

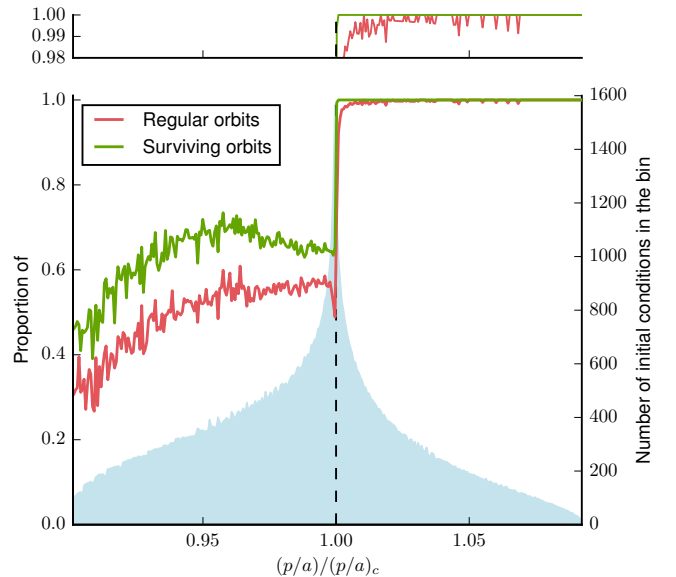
We first integrate 1 00 000 initial conditions on a uniform grid with  $\alpha$  taking values from 0.5 to 1 and  $\mathcal{C}$  from 0 to 0.1. The masses of the two planets are equal to  $0.5 \times 10^{-5} M_\odot$ , such that  $\varepsilon = 10^{-5}$  and  $\gamma = 1$ . In this simulation, 78.4% of the orbits survive up to 500 kyr, 21.1% end up in a collision between the two planets, and 0.5% of the integrations are stopped because of the nonconservation of energy due to an unresolved close encounter.

The results of the frequency analysis are shown in Fig. 4. We see that the chaotic region is well constrained by the Hill curve  $C_c^H$ . Indeed, very few orbits with  $\mathcal{C} < C_c^H$  appear to be chaotic. The region where Hill stable orbits ( $\mathcal{C} < C_c^H$ ) are chaotic seems restricted to the region, where  $C_c^{\text{MMR}} < \mathcal{C} < C_c^H$  around  $\alpha \approx 0.94$  and low  $\mathcal{C}$ , i.e., for orbits experiencing MMR overlap (a zoomed view of Fig. 4 is given in Fig. 5). The behavior of planets initially in this region was already discussed in (Deck et al. 2013). Orbits that are Hill unstable ( $\mathcal{C} > C_c^H$ ) appear to be largely chaotic up to some resonant islands situated at  $\alpha \approx 1$  (co-orbital resonance) and near the 3:2 and 4:3 resonances ( $\alpha \approx 0.76$  and  $0.82$ ). We also see that for larger separations ( $\alpha \lesssim 0.6$ ), orbits are less chaotic. However, it is probable that for longer integration times, these orbits would end up unstable.

In order to highlight how the Hill criterion separates the chaotic orbits from the stable orbits, we can plot the fraction of regular orbits as a function of  $(p/a)/(p/a)_c$  as suggested in



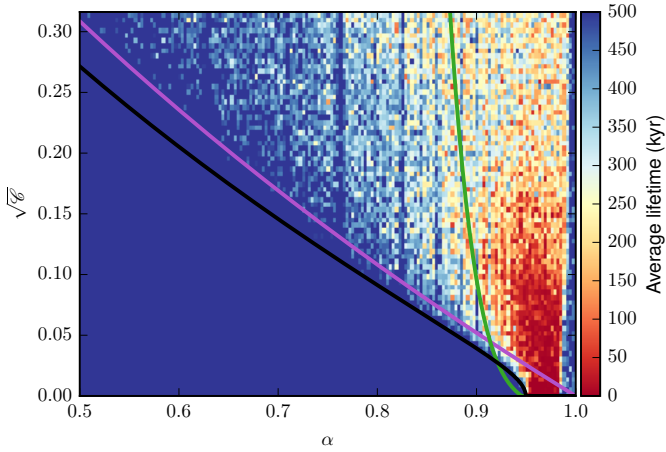
**Fig. 5.** Zoomed view of Fig. 4 for  $0.9 < \alpha < 1$  and  $\sqrt{\mathcal{C}} < 0.07$ . In the region where  $C_c^{\text{MMR}} < \mathcal{C} < C_c^H$ , we see that orbits are chaotic but Hill stable.



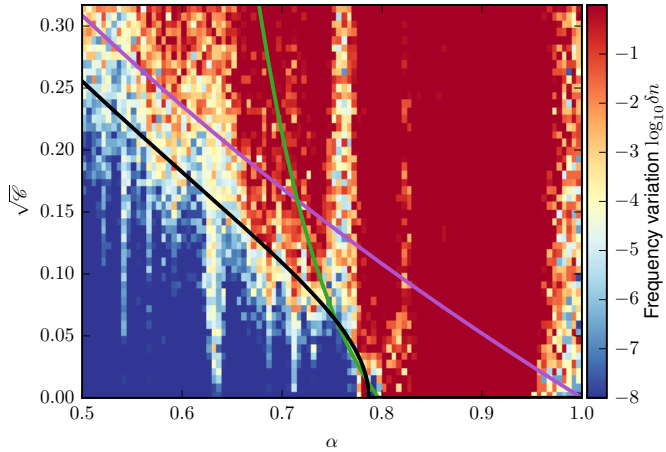
**Fig. 6.** Proportion of orbits that survived for 500 kyr (green curve) and nonchaotic orbits ( $\delta n < 10^{-4}$ ; red curve) as a function of  $(p/a)/(p/a)_c$ .  $(p/a)/(p/a)_c > 1$  means that the initial condition verifies the Hill criterion. Above the main figure, we added a zoomed view of the upper part of the plot. The light blue histogram represents the number of initial conditions in each of the 300 bins used to compute the fractions plotted.

Barnes & Greenberg (2006). We see in Fig. 6 that there is a sharp limit at  $p/a = (p/a)_c$ . All Hill stable integrations go to 500 kyr and very few are chaotic. For Hill unstable orbits ( $\mathcal{C} > C_c^H$ ), we see a slight increase of regular and surviving orbits with  $p/a$  but the decrease is not as significant as the change at the Hill limit. On average 64.9% of the Hill unstable orbits survive and 52.3% are regular.

We plot in Fig. 7 the average time of the integrations as a function of initial  $\alpha$  and  $\mathcal{C}$ . The initial conditions are binned in a  $160 \times 75$  grid. We see in this figure that the average lifetime of a system is almost always smaller than 500 kyr in the Hill unstable region. We also remark that the average lifetime is less than 50 kyr for a system with  $\alpha \gtrsim 0.9$  and not too large  $\mathcal{C}$  but increases for wider separations or greater AMD. Indeed, for



**Fig. 7.** Average lifetime of the system as a function of  $\alpha$  and  $\mathcal{C}$ . The parameters are similar to Fig. 4. Each bin represents an average over eight initial conditions. Dark blue implies that all integration ended at 500 kyr.



**Fig. 8.** Frequency variation  $\delta n$  (31) for 40 000 initial conditions binned in a  $100 \times 50$  grid with  $\varepsilon = 10^{-3}$  and  $\gamma = 1$ . The three curves are similar to Fig. 4.

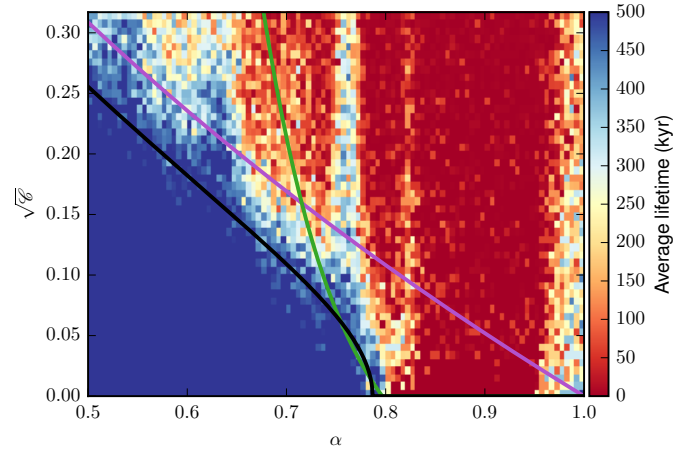
higher  $\mathcal{C}$ , the planets are on average initially further away from crossing orbits because the initial choice of eccentricities and inclinations is random.

#### 4.3. Influence of the masses

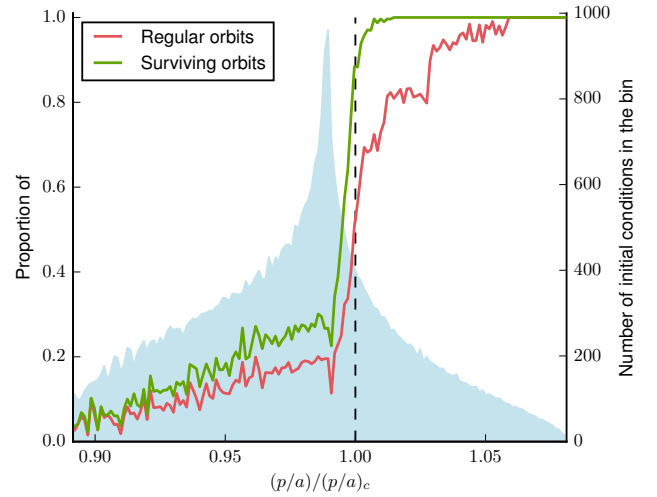
To highlight the role of the mass of the planets, we ran another simulation with the same method but with  $\varepsilon = 10^{-3}$  and  $\gamma = 1$ . We integrate 40 000 initial conditions and carry out a similar analysis. After 500 kyr, 54.6% of the systems leads to a collision between the two planets, 2% to an ejection, and 0.5% is stopped due to nonconservation of energy. If we consider only the Hill unstable systems ( $\mathcal{C} > C_c^H$ ), only 24.0% have survived and 15.7% are regular. As expected, larger masses lead to much more unstable systems.

The results of the frequency analysis are plotted in Fig. 8. We observe that the 3:2 MMR is still very stable in comparison to the surrounding regions. The 2:1 MMR appears to create an area of moderate chaos in the Hill stable region but is not as marked in the Hill unstable part. The co-orbital resonance also appears more unstable.

We see in Fig. 9 that some of the Hill stable systems do not survive for 500 kyr. These systems have led to an ejection of the



**Fig. 9.** Average lifetime of the system as a function of  $\alpha$  and  $\mathcal{C}$ . The parameters are similar to Fig. 8. We see that some Hill stable initial conditions stopped before 500 kyr owing to ejections.



**Fig. 10.** Proportion of orbits that survived for 500 kyr (green curve) and nonchaotic orbits ( $\delta n < 10^{-4}$ ; red curve) as a function of  $(p/a)/(p/a)_c$ .  $(p/a)/(p/a)_c > 1$  means that the initial condition verifies the Hill criterion. The light blue histogram represents the number of initial conditions in each of the 150 bins used to compute the fractions plotted.

outer planet. As a result, the Hill stability line appears to be more porous in comparison to the case  $\varepsilon = 10^{-5}$ .

We also see that a larger percentage of Hill stable orbits appear to be chaotic. This can be quantified, thanks to Fig. 10. We see that the proportion of regular orbits drops before it reaches the Hill stability limit. However, the more important change of behavior is still at the Hill stability limit.

## 5. Conclusions

In a two-planet system, Hill stability is a topological limit that forbids close encounters between the outer planet and the inner planet or the star. If verified, the system remains stable if the outer planet does not escape or if the inner planet does not collide with the star. Moreover, since a minimal distance between planets is imposed, the planet perturbation remains moderate and the system is most likely regular.

We have generalized Gladman's Hill stability criterion and have shown that it is natural to express the Hill criterion in the AMD framework. Indeed, we obtained a simple expression for this criterion (Proposition 3) with only an expansion in  $\varepsilon$  (Eq. (1)). Moreover, it is easy to recover all former published

criteria as particular cases of this expression. Because of its formulation as a function of the AMD, the expression (26) is valid in the general spatial case for any value of eccentricities and inclinations. Moreover, our Hill criterion is accurate even for very different planet masses. We also highlight that the AMD and the semimajor axis ratio  $\alpha$  are the main parameters to consider in a stability study.

We show that the Hill stability allows us to give an accurate stability limit up to large orbital separations. The sharp change of behavior at the Hill stability limit has already been studied in Barnes & Greenberg (2006) or in Deck et al. (2013). Nevertheless, our numerical integrations confirm it for a much larger range of  $\alpha$  and AMD with randomized initial conditions for the parameters not taken into account.

Our simulations for large planet masses also show that the expansion in  $\varepsilon$  is valid even for larger planets and that the Hill stability accurately segregates between regular and short-lived initial conditions. However, it appears that for large mass values, ejections cannot be neglected and a model should be developed to understand this further behavior.

As shown in several works on tightly packed systems (Chambers et al. 1996; Pu & Wu 2015), Gladman's Hill criterion is no longer adapted in these cases. Such a sharp limit between almost eternal and short-lived systems no longer exists. Instead, it appears that there exists a scaling between the initial orbital separation and the time of instability, wider separated orbits becoming unstable after a longer time. In the cited works, the empirical stability criteria give stability spacing as a function of the mutual radius

$$\frac{a_{n+1} - a_n}{a_{n+1}} > KR_H, \quad (32)$$

where  $R_H = (\varepsilon/3)^{1/3}$  is the Hill radius. In the context of multi-planetary systems, an analytical work on long-term stability is still necessary. This will be the subject of future work.

*Acknowledgements.* This work was granted access to the HPC resources of MesoPSL financed by the Region Ile de France and the project Equip@Meso (reference ANR-10-EQPX-29-01) of the programme Investissements d'Avenir supervised by the Agence Nationale pour la Recherche. We acknowledge the support of the PNP and the CS of the Observatoire de Paris.

## References

- Barnes, R., & Greenberg, R. 2006, *ApJL*, 647, L163  
 Chambers, J., Wetherill, G., & Boss, A. 1996, *Icarus*, 119, 261  
 Chirikov, B. V. 1979, *Phys. Rep.*, 52, 263  
 Deck, K. M., Payne, M., & Holman, M. J. 2013, *ApJ*, 774, 129  
 Farrés, A., Laskar, J., Blanes, S., et al. 2013, *Celest. Mech. Dyn. Astron.*, 116, 141  
 Georgakarakos, N. 2008, *Celest. Mech. Dyn. Astron.*, 100, 151  
 Gladman, B. 1993, *Icarus*, 106, 247  
 Laskar, J. 1990, *Icarus*, 88, 266  
 Laskar, J. 1993, *Physica D Nonlinear Phenom.*, 67, 257  
 Laskar, J. 1997, *A&A*, 317, L75  
 Laskar, J. 2000, *Phys. Rev. Lett.*, 84, 3240  
 Laskar, J., & Marchal, C. 1984, *Celest. Mech. Dyn. Astron.*, 32, 15  
 Laskar, J., & Petit, A. C. 2017, *A&A*, 605, A72  
 Marchal, C., & Bozis, G. 1982, *Celest. Mech. Dyn. Astron.*, 26, 311  
 Marchal, C., & Saari, D. G. 1975, *Celest. Mech. Dyn. Astron.*, 12, 115  
 Marchal, C., Yoshida, J., & Yi-Sui, S. 1984a, *Celest. Mech. Dyn. Astron.*, 33, 193  
 Marchal, C., Yoshida, J., & Yi-Sui, S. 1984b, *Celest. Mech. Dyn. Astron.*, 34, 65  
 Mustill, A. J., & Wyatt, M. C. 2012, *MNRAS*, 419, 3074  
 Petit, A. C., Laskar, J., & Boué, G. 2017, *A&A*, 607, A35  
 Pu, B., & Wu, Y. 2015, *ApJ*, 807, 44  
 Sundman, K. F. 1912, *Acta Math.*, 36, 105  
 Tamayo, D., Silburt, A., Valencia, D., et al. 2016, *ApJ*, 832, L22  
 Veras, D., & Armitage, P. J. 2004, *Icarus*, 172, 349  
 Wisdom, J. 1980, *AJ*, 85, 1122

## Appendix A: Computation of $R$ at the Lagrange points

The function  $R(x, y) = (\rho/\nu)^2$  defined in Eq. (16) admits three saddle points situated on the  $x$ -axis, which are the Lagrange points  $L_1, L_2$ , and  $L_3$ . If  $x_j$  is the abscissa of the point  $L_j$ , we have  $0 < x_1 < 1$ ,  $x_2 > 1$ , and  $x_3 < 0$ . The  $x_j$  quantities only depends on the mass ratio  $\varepsilon$  and on

$$\zeta = \frac{\gamma}{\gamma + 1} = \frac{m_1}{m_1 + m_2}. \quad (\text{A.1})$$

They are the roots of the three different polynomial equations:

$$\begin{aligned} (1 + \zeta\varepsilon)x_1^5 - (2 + 3\zeta\varepsilon)x_1^4 \\ + (1 + 2\varepsilon + \zeta\varepsilon)x_1^3 - (1 + 3\varepsilon(1 - \zeta))x_1^2 \\ + (2 + 3\varepsilon(1 - \zeta))x_1 - (1 + \varepsilon(1 - \zeta)) = 0, \end{aligned} \quad (\text{A.2})$$

$$\begin{aligned} (1 + \zeta\varepsilon)x_2^5 - (2 + 3\zeta\varepsilon)x_2^4 \\ + (1 + 3\zeta\varepsilon)x_2^3 - (1 + 3\varepsilon - \zeta\varepsilon)x_2^2 \\ + (2 + 3\varepsilon(1 - \zeta))x_2 - (1 + \varepsilon(1 - \zeta)) = 0, \end{aligned} \quad (\text{A.3})$$

$$\begin{aligned} (1 + \zeta\varepsilon)x_3^5 - (2 + 3\zeta\varepsilon)x_3^4 \\ + (1 + 3\zeta\varepsilon)x_3^3 + (1 + 3\varepsilon(1 - \zeta))x_3^2 \\ + (2 + 3\varepsilon(1 - \zeta))x_3 + (1 + \varepsilon(1 - \zeta)) = 0, \end{aligned} \quad (\text{A.4})$$

At up to terms of order  $\varepsilon^{4/3}$ , we have

$$\begin{aligned} x_1 &= 1 - \left(\frac{\varepsilon}{3}\right)^{1/3} + \frac{\gamma + 2}{3(\gamma + 1)} \left(\frac{\varepsilon}{3}\right)^{2/3} + \frac{(\gamma - 2)\varepsilon}{27(\gamma + 1)} + \text{O}(\varepsilon^{4/3}), \\ x_2 &= 1 - \left(\frac{\varepsilon}{3}\right)^{1/3} + \frac{\gamma + 2}{3(\gamma + 1)} \left(\frac{\varepsilon}{3}\right)^{2/3} - \frac{(\gamma - 2)\varepsilon}{27(\gamma + 1)} + \text{O}(\varepsilon^{4/3}), \\ x_3 &= -1 + \frac{7}{12} \frac{\gamma - 1}{\gamma + 1} \varepsilon + \text{O}(\varepsilon^2). \end{aligned} \quad (\text{A.5})$$

We can then evaluate  $R$  (Eq. (16)) at those points and we have at order,  $\varepsilon^{4/3}$ :

$$\begin{aligned} R(L_1) &= 1 + \frac{3^{4/3}\varepsilon^{2/3}\gamma}{(\gamma + 1)^2} - \frac{(11 + 7\gamma)\gamma\varepsilon}{3(\gamma + 1)^3} + \text{O}(\varepsilon^{4/3}), \\ R(L_2) &= 1 + \frac{3^{4/3}\varepsilon^{2/3}\gamma}{(\gamma + 1)^2} - \frac{(11\gamma + 7)\gamma\varepsilon}{3(\gamma + 1)^3} + \text{O}(\varepsilon^{4/3}), \\ R(L_3) &= 1 + \frac{2\varepsilon\gamma}{(\gamma + 1)^2} + \text{O}(\varepsilon^2). \end{aligned} \quad (\text{A.6})$$

## Appendix B: Expansion of $C_c^{\text{Ex}}$ and $h_1$

In Sect. 2, the Hill stability criterion (26) is obtained by the expansion at the leading order in  $\varepsilon$  of

$$F = (\gamma + 1)^{3/2} \sqrt{\frac{R(L_1)(1 + \varepsilon\gamma/(\gamma + 1)^2)^3}{(1 + \varepsilon)(\gamma/\alpha + 1 + h_1)}}. \quad (\text{B.1})$$

In  $F$ , the main term depending on  $\varepsilon$  comes from the expansion of

$$R(L_1) = 1 + \frac{3^{4/3}\gamma\varepsilon^{2/3}}{(\gamma + 1)^2} - \frac{(11 + 7\gamma)}{3(\gamma + 1)} \frac{\gamma\varepsilon}{(\gamma + 1)^2} + \text{O}(\varepsilon^{4/3}). \quad (\text{B.2})$$

However, we also need to make sure that  $h_1$  remains small in comparison to this term. Thus, the expansion of  $F$  requires an estimate of  $h_1$  with respect to  $\varepsilon$  and  $\gamma$ . As explained in Sect. 2,  $h_1$  is the renormalized perturbation part of the Hamiltonian (3):

$$h_1 = -\frac{2\Lambda_2^2}{m_2^3\mu^2} \left( \frac{1}{2} \frac{\|\tilde{\mathbf{r}}_1 + \tilde{\mathbf{r}}_2\|^2}{m_0} - \frac{\mathcal{G}m_1m_2}{r_{12}} \right). \quad (\text{B.3})$$

If we note  $\dot{\mathbf{r}}_j = \tilde{\mathbf{r}}_j/m_j$  and simplify the expression (B.3), we obtain  $h_1 = h_1^T + h_1^P$ , where

$$h_1^T = -\frac{\varepsilon a_2}{\mu} \frac{\|\gamma\dot{\mathbf{r}}_1 + \dot{\mathbf{r}}_2\|^2}{\gamma + 1}, \quad \text{and} \quad h_1^P = \frac{2a_2}{r_{12}} \frac{\varepsilon\gamma}{\gamma + 1}. \quad (\text{B.4})$$

As one can see in Fig. 2, the value of  $\varepsilon$  is only relevant for small values of the critical AMD  $C_c^{\text{Ex}}$ , i.e., for close planets. Using the expansion in  $1 - \alpha$  (27), we can estimate that we need to compute  $h_1$  for systems such that  $\mathcal{C}$  is of order  $\varepsilon^{2/3}$ . This corresponds to systems with eccentricities of order  $\varepsilon^{1/3}$ . We can therefore use the circular approximation in our estimation of  $h_1$ . In particular, we have  $r_j = a_j(1 + \text{O}(\varepsilon^{1/3}))$ .

We first consider the term coming from the gravitational interaction between the two planets,  $h_1^P$ . If the system is Hill stable, the distance  $r_{12}$  is greater than the radius of the Hill sphere  $S_{H_1}$ :

$$\begin{aligned} r_{12} > r_1 \max_{j=1,2} |1 - x_j| &= r_1(\varepsilon/3)^{1/3} + \text{O}(\varepsilon^{2/3}) \\ &= a_1(\varepsilon/3)^{1/3} + \text{O}(\varepsilon^{2/3}), \end{aligned} \quad (\text{B.5})$$

where  $x_j$  is defined in (A.5). For all times, we therefore have

$$h_1^P = \frac{2a_2}{r_{12}} \frac{\varepsilon\gamma}{\gamma + 1} \leq \frac{2 \times 3^{1/3}}{\alpha} \varepsilon^{2/3} \frac{\gamma}{\gamma + 1} + \text{O}(\varepsilon). \quad (\text{B.6})$$

The gravitational potential term  $h_1^P$  is at most of the same order as the leading term in  $\varepsilon$  of  $R(L_1)$ . However, we can always choose to estimate the energy and actions values when the two planets are far from each other. In this case,  $a_2/r_{12} = \text{O}(1)$  and  $h_1^P$  is linear in  $\varepsilon$ . From now on, we assume that we have

$$h_1^P = \frac{\varepsilon\gamma}{\gamma + 1} p_{12}, \quad (\text{B.7})$$

with  $p_{12} = \text{O}(1)$ .

Moreover,  $F$  is a decreasing function of  $h_1$ , so  $C_c^{\text{Ex}}$  is an increasing function of  $h_1$ . Since  $h_1^P$  is positive, neglecting it is equivalent to have a more conservative criterion.

Let us now consider the kinetic term  $h_1^T$ . We develop (B.4) for almost circular orbits. In this limit, we have  $\|\dot{\mathbf{r}}_j\| = \sqrt{\mu/a_j} + \text{O}(\varepsilon^{1/3})$ . We have

$$h_1^T = -\frac{\varepsilon}{\gamma + 1} \left( \frac{\gamma^2}{\alpha} + 1 + \frac{2\gamma}{\sqrt{\alpha}} \cos(\lambda_1 - \lambda_2) \right) + \text{O}(\varepsilon^{4/3}). \quad (\text{B.8})$$

Therefore,  $h_1^T$  is always linear in  $\varepsilon$ . Combining the estimations (B.7) and (B.8), we see that  $h_1 = \text{O}(\varepsilon)$ .

We can now use the expansions of  $h_1$ , (B.8) and (B.7), to obtain the expansion of  $F$  in function of  $\varepsilon$  and  $\gamma$ . Because of the term  $3^{4/3}\varepsilon^{2/3}\gamma/(\gamma + 1)^2$  from  $R(L_1)$ , we do not keep terms of order  $\text{O}(\varepsilon\gamma/(\gamma + 1)^2)$ . We keep all other terms depending on  $\varepsilon$  up to the order  $\text{O}(\varepsilon^{4/3})$ . We obtain

$$F = \sqrt{\frac{\alpha(\gamma + 1)^3}{\gamma + \alpha} \left( 1 + \frac{3^{4/3}\varepsilon^{2/3}\gamma}{(\gamma + 1)^2} \right)} D + \text{O}\left(\frac{\varepsilon\gamma}{(\gamma + 1)^2}, \varepsilon^{4/3}\right), \quad (\text{B.9})$$

where  $D = (1 - \varepsilon) \left( 1 - \frac{\alpha h_1}{\gamma + \alpha} \right)$

and

$$O\left(\frac{\varepsilon\gamma}{(\gamma+1)^2}, \varepsilon^{4/3}\right) = O\left(\frac{\varepsilon\gamma}{(\gamma+1)^2}\right) + O\left(\varepsilon^{4/3}\right).$$

Let us develop  $D$  on the same order than  $F$ . We have

$$\begin{aligned} D &= 1 - \varepsilon - \frac{\alpha h_1^\Gamma}{\gamma + \alpha} + O\left(\frac{\varepsilon\gamma}{(\gamma+1)^2}, \varepsilon^{4/3}\right) \\ &= 1 + \varepsilon \left( \frac{\gamma^2 + \alpha}{(\gamma+1)(\gamma+\alpha)} - 1 \right) + O\left(\frac{\varepsilon\gamma}{(\gamma+1)^2}, \varepsilon^{4/3}\right) \\ &= 1 - \frac{\varepsilon\gamma}{(\gamma+1)^2} \frac{(\gamma+1)(\alpha+1)}{\gamma+\alpha} + O\left(\frac{\varepsilon\gamma}{(\gamma+1)^2}, \varepsilon^{4/3}\right) \\ &= 1 + O\left(\frac{\varepsilon\gamma}{(\gamma+1)^2}, \varepsilon^{4/3}\right). \end{aligned} \tag{B.10}$$

We see that  $D$  has no contribution to  $F$  at the considered orders. Therefore, we can write

$$F = \sqrt{\frac{\alpha(\gamma+1)^3}{\gamma+\alpha} \left( 1 + 3^{4/3} \frac{\varepsilon^{2/3}\gamma}{(\gamma+1)^2} \right)} + O\left(\frac{\varepsilon\gamma}{(\gamma+1)^2}, \varepsilon^{4/3}\right). \tag{B.11}$$

As an immediate consequence of (B.11), the expression proposed in Eq. (26) remains valid in the case of a very uneven mass distribution between the two planets (e.g., for  $O(\varepsilon^{1/3}) < \gamma < O(\varepsilon^{-1/3})$ ).