

Stability of hierarchical triples – I. Dependence on inner eccentricity and inclination

A. Mylläri,^{1★} M. Valtonen,^{2,3★} A. Pasechnik² and S. Mikkola²

¹*St. George's University, Grenada, West Indies*

²*Tuorla Observatory, Department of Physics and Astronomy, University of Turku, FI-20014 Turku, Finland*

³*Finnish Centre for Astronomy with ESO, University of Turku, FI-20014 Turku, Finland*

Accepted 2018 January 24. Received 2018 January 23; in original form 2016 October 9

ABSTRACT

In simulations it is often important to decide if a given hierarchical triple star system is stable over an extended period of time. We introduce a stability criterion, modified from earlier work, where we use the closest approach ratio Q of the third star to the inner binary centre of mass in their initial osculating orbits. We study by numerical integration the orbits of over 1000 000 triple systems of the fixed masses and outer eccentricities e_{out} , but varying inner eccentricities e_{in} and inclinations i . 12 primary combinations of masses have been tried, representing the range encountered in stellar systems. The definition of the instability is either the escape of one of the bodies, or the exchange of the members between the inner and outer systems. An analytical approximation is derived using the energy change in a single close encounter between the inner and outer systems, assuming that the orbital phases in subsequent encounters occur randomly. The theory provides a fairly good description of the typical Q_{st} , the smallest Q value that allows the system to be stable over $N = 10\,000$ revolutions of the initial outer orbit. The final stability limit formula is $Q_{\text{st}} = 10^{1/3} A [(fg)^2 / (1 - e_{\text{out}})]^{1/6}$, where the coefficient $A \sim 1$ should be used in N -body experiments, and $A = 2.4$ when the absolute long-term stability is required. The functions $f(e_{\text{in}}, \cos i)$ and $g(m_1, m_2, m_3)$ are derived in the paper. At the limit of $e_{\text{in}} = i = m_3 = 0, fg = 1$.

Key words: methods: numerical – celestial mechanics.

1 INTRODUCTION

A hierarchical triple system consists of a binary and a third body in orbit around the centre of mass of the binary. In order to have a clear separation of the inner and outer orbits, the pericentre of the outer orbit should be greater than the major axis of the binary. A stability criterion of hierarchical triples is required in understanding many systems arising in the Universe, for example in understanding the longevity of the Earth–Moon–Sun system (Newton 1687; Clairaut 1752). The stability of our planetary system falls in the same category of problems, initially with the question of the motion of the giant planets Jupiter and Saturn (Euler 1752; Lagrange 1766, 1778; Laplace 1775, 1787). In recent years similar questions have arisen in connection with multiple exoplanets in other stellar systems (see e.g. Funk et al. 2009). With enough information of the initial data, it is always possible to carry out numerical orbit integrations to determine the degree of stability, at least over some period of time (Laskar 2013). However, it is often the case that only some orbital elements are known, and even they have associ-

ated measuring uncertainties. Then the phase space of the unknown elements has so many dimensions that its total coverage may be prohibitive.

Another case where a clear-cut stability criterion would be useful is in the computer simulations of star clusters (Aarseth 1973, 2003; Heggie & Hut 2003). In this case all the elements of a triple subsystem are known, but its integration takes up many resources and will slow the overall cluster simulation. It is more efficient to leave stable triples aside from the main calculation until such time that encounters with other stars or other reasons cause interesting orbital evolution in this subsystem. In this work, we are mainly concerned with the latter situation.

Harrington (1972) realized that the key quantity in assessing the stability of a three-body system is the ratio of the pericentre distance of the outer orbit over the inner orbit semimajor axis. He found that its value, which we call Q , has to be at least 3.5 for direct orbits and 2.75 for retrograde orbits, the exact value depending on the masses of the three bodies. In this paper our main variable to be determined is the minimum value of Q for stability, called Q_{st} . We start by using a sufficiently large value Q and determine that the system is stable. Then the value of Q is

* E-mail: amyllari@gmail.com (AM); mvaltonen2001@yahoo.com (MV)

lowered until an unstable system is found. Going to even smaller Q , the system is likely to be unstable, but even if it is not, we use the Q -value first encountered in the search for the unstable system to define Q_{st} . For example, if we find that the orbit with $Q = 5.0$ is stable and that the next orbit with $Q = 4.9$ is unstable, we define $Q_{\text{st}} = 5.0$, even if the system is stable again, say, in the range $Q = 4.0\text{--}4.8$.

The primary orbital elements to be sampled in this paper are the cosine of the inclination $\cos i$ and the eccentricity of the inner orbit e_{in} , and to a lesser extent the masses of the three bodies, and the outer eccentricity e_{out} . The finite intervals of sampling lead to some scatter in Q_{st} , but the scatter is mainly due to the postulate that the remaining orbital elements arise at random. The scatter in Q_{st} means that our result is not an exact surface but a layer of finite thickness in Q_{st} . Dvorak (1997) and Pilat-Lohinger, Funk & Dvorak (2003) call this layer *mixed region*. The upper surface of the layer is used when we answer the question whether the system is definitely stable under the chosen revolution number N . The lower surface of the layer tells us when the system is definitely unstable, while a surface between them gives a stability criterion that could be useful e.g. in N -body simulations. In later papers we will discuss how the boundary of the layer moves with varying N . In this paper we use $N = 10\,000$ revolutions of the outer orbit.

2 ORBIT INTEGRATIONS

For simplicity, we start with fixed mass values of $m_1 = 1.5$, $m_2 = 0.5$, and $m_3 = 0.5$ units (solar mass, for example), where m_1 and m_2 form the inner binary of unit semimajor axis (e.g. 1 au) and with eccentricity e_{in} , while the body with mass m_3 is in an orbit of eccentricity $e_{\text{out}} = 0.5$ about the centre of mass of the binary. These values have no particular significance. In later papers we will study how the ideas of this paper are extended to other values of masses and e_{out} . In Section 4 we start a brief exploration to this direction.

Most of the orbit integrations are carried out by a symplectic integrator that has the property of conserving total energy, using the Bulirsch–Stoer method (Mikkola 1997). For comparison, we have integrated a smaller number of cases by a three-body regularization code (Aarseth & Zare 1974) and found that the overall result is not integrator dependent.

The Q value, the ratio of the initial pericentre distance of the outer orbit to the inner semimajor axis, is first set so high that the system is stable over the $N = 10\,000$ revolutions. Then for the next orbits the initial parameters are varied, keeping e_{out} constant while the major axis of the outer orbit becomes smaller in steps of equal intervals. Therefore also Q gets smaller values in a stepwise manner, even though the steps are not exactly equal. The typical Q -step is 0.25 units. In later experiments (Figs 5 and 7) it is about 0.1 units. The integrations are continued down to a value of Q where the systems are always unstable. Since this value is not known in advance, in practise the smallest Q is set to about unity. In terms of integration time, the solutions are found very fast once we are below the mixed region.

The process is repeated for 50 values (sometimes 100 values that we call the full resolution) of $\cos i$ equally spaced between -1 and $+1$, and a diagram of Q versus $\cos i$ is generated. In this diagram the Q_{st} versus $\cos i$ line is identified, and is called the stability line. It is always wiggly, not only because of the finite step sizes, but mostly since we sample the mixed region with fixed values of the ‘less-essential’ orbital elements. These are the initial angles of the pericentres, mean anomalies, and the nodes. Unless otherwise stated

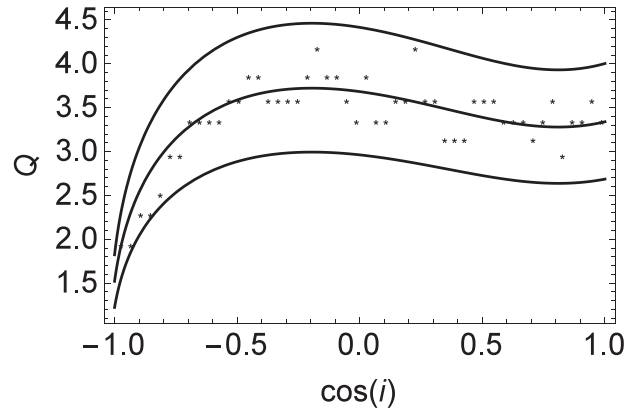


Figure 1. Stability limit Q calculations displayed with Q (y-axis) as a function of $\cos i$ (x-axis). The asterisk (*) shows the first unstable system when the Q -value is reduced in steps from high towards low values. Our standard parameter values are $m_1 = 1.5$, $m_2 = 0.5$, $m_3 = 0.5$, and $e_{\text{out}} = 0.5$. In this case the inner eccentricity $e_{\text{in}} = 0.1$. The curves display the model outlined in the text. The lower curve is the always unstable limit (unstable below the line), the uppermost curve outlines the always stable region (higher up). The middle line is a cut of the middle of the mixed region and corresponds to $A = 1.45$. Upper and lower curves correspond to $A = 1.75$ and 1.15 .

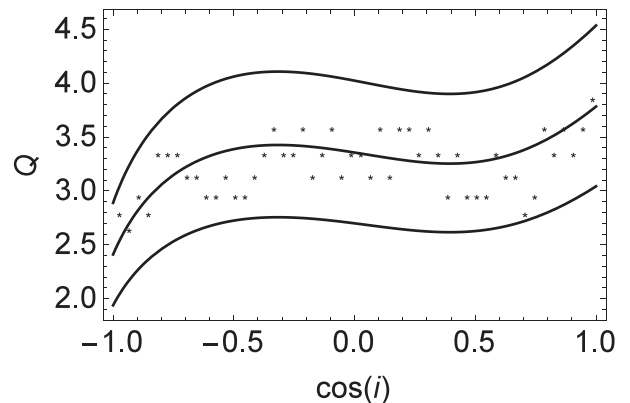


Figure 2. As above, but for inner eccentricity $e_{\text{in}} = 0.5$.

below, their values are set to zero. The asterisk (*) in the plots refer strictly speaking to these fixed values.

Then a new value of e_{in} is selected and a new diagram is generated. In the first part of this paper we report the sample of 13 diagrams that cover the inner eccentricity e_{in} -axis well. The values we use for e_{in} are 0, 0.05, 0.1, 0.15, 0.2, 0.3, 0.4, 0.5, 0.6, 0.75, 0.9, 0.95, and 0.99 (Figs 1–3 give examples of these diagrams).

3 ANALYTIC MODEL

Roy & Haddow (2003), Heggie (2006), and Valtonen & Karttunen (2006) studied the energy change $\delta\varepsilon/\varepsilon$ of the inner binary in a single three-body encounter when the outer orbit is parabolic, hyperbolic, or elliptic, respectively. Here ε is the binary energy and $\delta\varepsilon$ is the change arising from the encounter. As an illustration, let us take the parabolic case. From equation (19) of Roy & Haddow (2003)

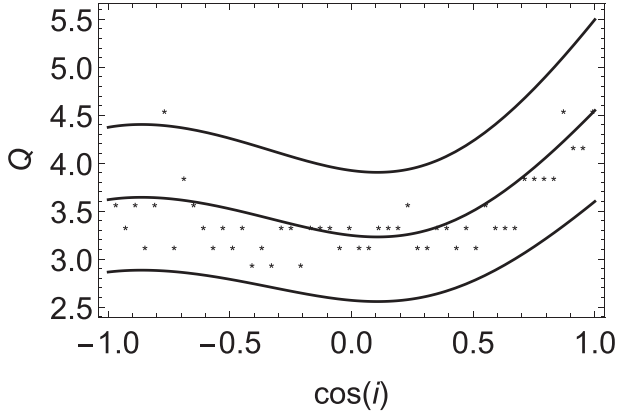


Figure 3. As above, but for inner eccentricity $e_{\text{in}} = 0.9$.

we find

$$\begin{aligned} \frac{\delta \varepsilon}{\varepsilon} = & \frac{m_3}{m_1 + m_2} \frac{\sqrt{\pi}}{4} \left\{ Q^{-3} K^{2.5} e^{-(2/3 K)} \right\} \\ & \times \{ e_1 [-\sin(2\omega + nt_0)(1 - \cos 2i) \\ & - \sin(2\omega + nt_0) \cos 2i \cos 2\Omega \\ & - 3 \sin(2\omega + nt_0) \cos 2\Omega \\ & - 4 \cos(2\omega + nt_0) \cos i \sin 2\Omega] \\ & + e_2 (1 - e_{\text{in}}^2) [\sin(2\omega + nt_0)(1 - \cos 2i) \\ & - \sin(2\omega + nt_0) \cos 2i \cos 2\Omega \\ & - 3 \sin(2\omega + nt_0) \cos 2\Omega \\ & - 4 \cos(2\omega + nt_0) \cos i \sin 2\Omega] \\ & + e_4 \sqrt{1 - e_{\text{in}}^2} [-2 \cos(2\omega + nt_0) \cos 2i \sin 2\Omega \\ & - 6 \cos(2\omega + nt_0) \sin 2\Omega \\ & - 8 \sin(2\omega + nt_0) \cos i \cos 2\Omega] \}. \end{aligned} \quad (1)$$

The functions e_1 , e_2 , and e_4 are defined in terms of the Bessel functions $J_{-1}(e_{\text{in}})$, $J_0(e_{\text{in}})$, $J_2(e_{\text{in}})$, and $J_3(e_{\text{in}})$:

$$e_1 = J_{-1}(e_{\text{in}}) - 2e_{\text{in}}J_0(e_{\text{in}}) + 2e_{\text{in}}J_2(e_{\text{in}}) - J_3(e_{\text{in}}),$$

$$e_2 = J_{-1}(e_{\text{in}}) - J_3(e_{\text{in}}),$$

$$e_4 = J_{-1}(e_{\text{in}}) - e_{\text{in}}J_0(e_{\text{in}}) - e_{\text{in}}J_2(e_{\text{in}}) + J_3(e_{\text{in}}).$$

After substituting the first few terms of power series of the Bessel functions we get

$$-e_1 = \frac{5}{2}e_{\text{in}} - \frac{19}{24}e_{\text{in}}^3 + \frac{41}{768}e_{\text{in}}^5,$$

$$-e_2 = \frac{1}{2}e_{\text{in}} - \frac{1}{24}e_{\text{in}}^3,$$

$$-e_4 = \frac{3}{2}e_{\text{in}} - \frac{5}{24}e_{\text{in}}^3,$$

with the accuracy of better than 1 per cent for all eccentricities up to $e_{\text{in}} = 1$.

We use the definition

$$K = Q^{1.5} \sqrt{2(m_1 + m_2)/(m_1 + m_2 + m_3)},$$

and i , ω , and Ω are the inclination, the argument of the pericentre, and the longitude of the ascending node of the third body orbit relative to the binary centre. The mean motion of the binary is n and t_0 is the time measured from the third body pericentre passage.

At the limit of small inner eccentricity (i.e. putting $e_{\text{in}}^2 = 0$) we get

$$\begin{aligned} \frac{\delta \varepsilon}{\varepsilon} = & \frac{m_3}{m_1 + m_2} \frac{\sqrt{\pi}}{4} \left\{ Q^{-3} K^{2.5} e^{-(2/3 K)} \right\} e_{\text{in}} \\ & \times \left\{ 6 \sin(2\omega + nt_0)(1 + \cos i)^2 \right. \\ & \left. + 4 \sin(2\omega + nt_0)(1 - \cos^2 i) \right\}. \end{aligned} \quad (2)$$

Here we are interested in the absolute value of the energy change when the sinusoidal factors are essentially random. Their absolute values average to $2/\pi$, and apart from this factor, the inclination dependence is of the form

$$6(1 + \cos i)^2 + 4(1 - \cos^2 i) \approx 3.6 \left(\frac{5}{3} + \cos i \right)^2. \quad (3)$$

The latter expression agrees with the former at better than 1 per cent level relative to its maximum value at $\cos i = 1$. Besides being a slightly simpler function than the first one, the latter also makes more physical sense since using it $\frac{\delta \varepsilon}{\varepsilon}$ goes to a non-zero value at $\cos i = -1$.

The quantity in the first curly brackets is well approximated by a factor proportional to

$$(Q/Q_1)^{-7}, \quad (4)$$

where

$$Q_1 = 2.5 (1 + m_3/(m_1 + m_2))^{1/3} \quad (5)$$

(Valtonen & Karttunen 2006). This results from a numerical evaluation of the function in the first curly brackets of equation (1), and then finding the best power-law fit over a limited range of Q . Figs 1–3 show what the range of interest of Q is for this particular combination of masses. Thus the power -7.0 is not exact. For example, in our primary study of systems with $m_3 = 0.5$, in the range $Q = 3\text{--}3.5$ the power-law index is -6.7 , while in the range $Q = 3\text{--}4$ it is -7.4 . The slope is exactly -7.0 at $Q = 3.35$. For other mass values, the slope of -7.0 occurs at somewhat different Q that scales with Q_1 .

We may then calculate the coefficient $\frac{\sqrt{\pi}}{4} \left\{ Q^{-3} K^{2.5} e^{-(2/3 K)} \right\} e_{\text{in}}$ at the representative value $Q = 3.35$ and put $e_{\text{in}} = 0.05$ to represent a small initial eccentricity. Multiplied by the factor 3.6 from $3.6(\frac{5}{3} + \cos i)^2$ we get 2×10^{-3} . The same numerical value is given by $0.0092(Q/Q_1)^{-7}$ for the masses and Q in question. Considering various uncertainties during this derivation, we replace 0.0092 by 0.01 and get

$$\frac{\delta \varepsilon}{\varepsilon} \approx 0.01 \frac{m_3}{m_1 + m_2} (Q/Q_1)^{-7} \left(\frac{5}{3} + \cos i \right)^2. \quad (6)$$

This result was based on the study of parabolic encounters, but in fact very similar results arise if the outer orbit elliptic. In particular, we may confirm the numerical factor 0.01 and the splitting of the functional dependence into several factors that each depends only on a smaller number of parameters (Valtonen & Karttunen 2006). Thus in the following we apply this result to multiple elliptic outer-body encounters.

The sinusoidal factor in curly brackets is phase dependent. Since we may assume that the repeated encounters do not keep the phase, this factor causes a drift in the energy space. We may model the drift by a random walk, with the step size of the order of the amplitude of the phase factor. There is also a drift in other orbital elements; in this work we are not able to calculate this effect, but refer below to some of the possible consequences of the drift in the $e_{\text{in}}\text{--}\cos i$ plane.

By energy conservation, the relative energy change of the binary $\delta\varepsilon/\varepsilon$ translates into the corresponding change in the outer orbit $\delta E/E$:

$$\begin{aligned} \frac{\delta\varepsilon}{\varepsilon} &= -\frac{\delta E}{\varepsilon} = -\frac{E}{\varepsilon} \frac{\delta E}{E} \\ &= -\frac{m_3(m_1+m_2)}{m_1 m_2} \frac{a_{\text{in}}}{a_{\text{out}}} \frac{\delta E}{E} \\ &= -\frac{m_3}{m_1+m_2} \frac{(m_1+m_2)^2}{m_1 m_2 Q_1} \frac{1-e_{\text{out}}}{Q/Q_1} \frac{\delta E}{E} \\ &= -\frac{m_3}{m_1+m_2} M \frac{1-e_{\text{out}}}{Q/Q_1} \frac{\delta E}{E}, \end{aligned} \quad (7)$$

where a_{in} and a_{out} are the inner and out semimajor axes, respectively, and

$$M = \frac{0.4(m_1+m_2)^{7/3}}{(m_1+m_2+m_3)^{1/3} m_1 m_2}. \quad (8)$$

The functional dependence on the masses is split in two factors in order to help at the next step. After N encounters we expect the total energy to change by $\sqrt{N}\delta E$. When this amounts to E , we definitely have an unstable situation. We may also use a more strict definition of the instability, saying that the instability arises when $E = \lambda\sqrt{N}\delta E$. This defines a strictness factor λ such that it equals to unity when the instability means the total destruction of the hierarchy of the system (Huang & Innanen 1983), while a greater value of the strictness factor, e.g. $\lambda = 10$, is closer to the definition of Mardling & Aarseth (1999). We use $\lambda = 1$ throughout in this paper.

We equate the amplitude of equation (6) to the absolute value of the right-hand side of equation (7) and solve for Q :

$$\begin{aligned} Q &\approx (\lambda\sqrt{N})^{1/6} M^{-1/6} \left(1 + \frac{m_3}{m_1+m_2}\right)^{1/3} \\ &\quad \times (1-e_{\text{out}})^{-1/6} \left(\frac{5}{3} + \cos i\right)^{1/3}. \end{aligned} \quad (9)$$

This is called the stability limit Q_{st} . The quantity $M^{-1/6}$ is set to unity in the following. It is 0.925 if the binary masses are equal and the third body mass is zero. For the range of masses considered in this paper, it is within 24 percent of this value.

Even in other respects this cannot be considered an exact formula, as is obvious after several steps of simplifications. Rather it gives a motivation to search for a particular type of result. In the end we introduce a scaling coefficient A that is determined purely from experiments.

Let us now consider arbitrary inner eccentricities e_{in} . Then the form in the second curly brackets in equation (2) becomes, taking terms up to the order e_{in}^5 in the derivation above (Roy & Haddow 2003),

$$\begin{aligned} &12 \left\{ \left[(1 - 0.444e_{\text{in}}^2 + 0.032e_{\text{in}}^4) \cos i + 0.5\sqrt{1-e_{\text{in}}^2} \right. \right. \\ &\quad \times (1 - 0.139e_{\text{in}}^2) (1 + \cos i^2) \left. \right] \cos(2\omega + nt_0) \sin 2\Omega \\ &\quad + \left[\sqrt{1-e_{\text{in}}^2} (1 - 0.139e_{\text{in}}^2) \cos i \right. \\ &\quad \left. + 0.5 (1 - 0.444e_{\text{in}}^2 + 0.032e_{\text{in}}^4) (1 + \cos i^2) \right] \\ &\quad \times \sin(2\omega + nt_0) \cos 2\Omega + 1/3(1 - 0.12e_{\text{in}}^2) \\ &\quad \left. \times (1 - \cos i^2) \sin(2\omega + nt_0) \right\}. \end{aligned} \quad (10)$$

Now it is apparent that the parameters e_{in} and $\cos i$ are not separable, but are contained inside a single factor. In equation (2) the

two parameters separated because e_{in} was small and was a common multiplier in all the terms. In general this is not the case.

There is one more consideration that we must worry about. During the evolution of the triple system it may wonder widely in the $e_{\text{in}}\text{-}\cos i$ plane due to the eccentric Kozai resonance (Ford, Kozinsky & Rasio 2000; Katz, Dong & Malhotra 2011; Lithwick & Naoz 2011) whereby the stability limit tends to be determined by the region where the limit is highest during this wondering. The process tends to equalize the stability limit within $-0.75 \leq \cos i \leq 0.75$, the effective classical (i.e. quadrupole) Kozai cycle range (Kozai 1962; Innanen et al. 1997; Valtonen et al. 2008). Fitting the calculated data points to different functional forms quickly shows that we need to introduce $\cos^3 i$ terms into our formula that do not appear in the expression (10). It helps to describe the rather high Q_{st} values in retrograde orbits in the range of $-0.5 \leq \cos i \leq 0$. We do not find need for order higher than three in $\cos i$ in order to describe the results of our data set. In particular, it is clear that the second-order expression above is not useful as such. Also the universal $(\frac{5}{3} + \cos i)^2$ term for all e_{in} is obviously not valid.

Therefore we look for a form that is of third order in $\cos i$ and has coefficients that are polynomials of e_{in} . The simplest form that could be used is of the third order in both $\cos i$ and e_{in} . The coefficients of the polynomials have been determined by fitting our experiments (the 13 diagrams mentioned above, $m_1 = 1.5$) to the polynomials of this order, and determining the best values for the coefficients. We used a least-squares fit to all our data points in these diagrams. Subsequently the coefficients of some high-order terms were found to be so small as to be negligible. Then these terms were dropped, and new fits were performed without them. This procedure was continued until all remaining terms were significant. The coefficients were then rounded to nearest simple fractions, making sure that this does not worsen the fit. In this way we find that $\frac{5}{3} + \cos i$ in the stability limit formula should be replaced by

$$\begin{aligned} f(e_{\text{in}}, \cos i) &= \left\{ 1 - \frac{2}{3}e_{\text{in}} \left[1 - \frac{1}{2}e_{\text{in}}^2 \right] - 0.3 \cos i \left[1 \right. \right. \\ &\quad \left. \left. - \frac{1}{2}e_{\text{in}} + 2 \cos i \left(1 - \frac{5}{2}e_{\text{in}}^{3/2} - \cos i \right) \right] \right\}. \end{aligned} \quad (11)$$

In the last bracket the terms containing e_{in} and e_{in}^2 have been combined to a term $e_{\text{in}}^{3/2}$ for simplicity, and without loss of accuracy.

Remember that in our original formula of equation (1) the e_{in} and $\cos i$ factors were not separable; their separation in different factors happened only at the low eccentricity limit in equation (2). Here we have returned to a factor containing both variables after making extensive use of numerical experiments.

Let us denote the mass factor

$$g(m_1, m_2, m_3) = \left(1 + \frac{m_3}{m_1+m_2} \right). \quad (12)$$

In addition to the other factors of equation (9) there may be a coefficient of the order of unity on the right-hand side of this equation; we call it A . This coefficient has to be obtained by a fit to our data set because our complete formula was not derived from first principles; we used a number of simplifying assumptions. We first find the function f from theory, and then look for suitable values of A . Our complete formula is now

$$Q_{\text{st}} = A \left(\lambda\sqrt{N}/(1-e_{\text{out}}) \right)^{1/6} (f g)^{1/3}. \quad (13)$$

Note that the relative simplicity of this formula is based on equation (1). Even though it describes energy changes only in parabolic encounters, the studies for elliptic encounters give qualitatively sim-

ilar results. One of the main advantages of this approach is the factorization where each factor depends on rather few parameters. There is a factor for masses, for the normalized pericentre distance, and a mixed inner eccentricity/inclination factor. The latter also has other angular dependences in the form of sinusoidal functions that we have averaged out. It is this feature of equation (1) that carries through the different stages of our derivation, and finally leads to a rather simple end result.

The value of A is determined during the fit of the analytical functions to our data set. In our simulations we start with high values of Q and reduce values in finite steps until we hit the first unstable system. These points are plotted in our figures. Consequently, the last stable system is one Q -step higher. So, the stability border lies in between these two points. The value of A that is determined from the first unstable points is too low when we are looking to represent this instability line. From the fit we find $A = 1.40$. However, instability is one half-a-step higher or at $A = 1.45$. The step size comes from our initial choice for sampling along the Q -axis. The scatter about the midsurface is roughly Gaussian with standard deviation approximately 0.1 units of Q . Most of the data points should be within the layer of thickness of three standard deviations of the central surface, which is within a layer defined by $A = 1.45 \pm 0.3$. This is true for all except two points.

In N -body simulations one may want to use the smaller limit of $A = 1.15$. Then no computer time is wasted in the calculation of stable subsystems. The drawback of this choice is that many unstable systems from the mixed region will be treated as stable, and the integration of these subsystems is halted. A compromise may be to use the central plane of the mixed region, $A = 1.45$. One would need to experiment with actual N -body simulations to decide the optimum value of A . In other types of problems we may want to be absolutely sure of the stability of the triple system. Then we will want to use the upper value, $A = 1.75$. However, our experiments so far may be too sparse to determine this absolute upper limit.

In order to make a more careful study of the upper and lower limits of the mixed region, we increased the number of simulations. Instead of using just one value of the longitude of the node Ω ($\Omega = 0$ in practice), we started covering the whole range of its values. We note from expression (10) that $\sin 2\Omega$ and $\cos 2\Omega$ are independent factors that should be varied to get the whole range of possible outcomes for a given $\cos i$. This contrasts with the longitude of the pericentre ω that appears only in combination with m_0 and is therefore, at least in theory, a random variable.

In order to study the lower limit of the mixed region, the limit for the absolutely unstable systems, we made simulations at a given $\cos i$ and starting from a low Q -value. By orbit integration we checked the stability for each Ω -value. Initially experiments are unstable for all Ω -values. Then the Q -value is increased by one step. This is continued until at least one Ω -value results in a stable system. At that point we say that the absolute instability boundary has been found and that it lies between the last two Q -values.

Similarly, when we study the upper limit of the mixed region, the limit for absolute stability, we start from a sufficiently high Q -value so that systems with all Ω are stable. Then we lower the Q -value by one step and repeat the process. When we discover first unstable system, for any value of Ω , we have crossed the stability boundary that lies between the last two Q -values.

In this way we generate two boundary lines, well separated from each other. Fig. 4 shows an example. Here 50 equally spaced values of $\cos i$ and 10 values of randomly generated Ω were used for $m_2 = 0.2$, $m_3 = 10$, $e_{\text{in}} = 0.25$, and $e_{\text{out}} = 0.5$. We see that the line for the upper limit is effectively raised. Typically the A -value is

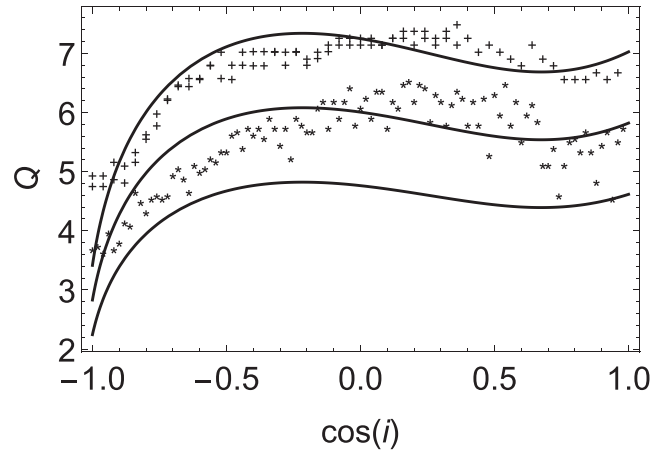


Figure 4. Stability limit points for the upper limit (stability, crosses) and lower limit (instability, stars) for $m_1 = 1.8$, $m_2 = 0.2$, $m_3 = 10$, $e_{\text{in}} = 0.25$, and $e_{\text{out}} = 0.5$.

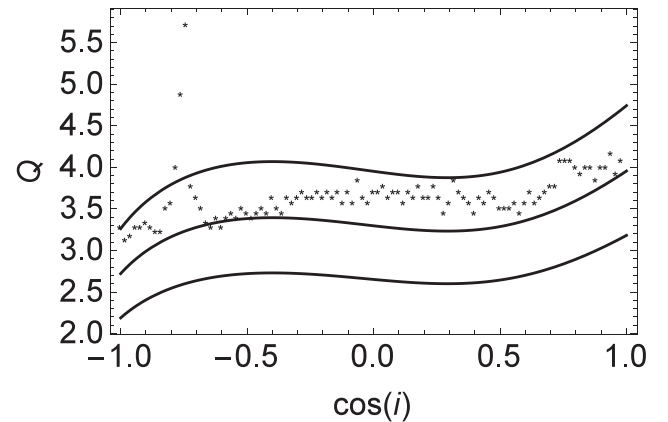


Figure 5. The upper (stability) border for the inner eccentricity $e_{\text{in}} = 0.6$ and masses $m_1 = 1.5$, $m_2 = 0.5$, $m_3 = 0.5$, and $e_{\text{out}} = 0.5$. The three lines are drawn for $A = 1.15$, 1.45 , and 1.75 .

increased by 0.15 units higher by using this method than by using points for single values of Ω .

10 similar diagrams were generated by varying m_2 (either 0.2, 0.5, or 1.0) and e_{in} (either 0.1, 0.25, 0.5, 0.75, 0.9, or 0.99) for fixed values of $m_3 = e_{\text{out}} = 0.5$. In all cases the upper boundary points follow the $A = 1.45$ line rather well, with the highest point at $A = 2.3$. The inner boundary line is of the same general shape as the upper boundary but typically 0.6 units lower in A .

In this paper we are primarily concerned with the upper envelope of the mixed region. The study of the lower boundary is left for future work since it could be very sensitive to islands of stability (Dvorak 1997).

In Fig. 5 we give an example of the upper boundary using 100 equally spaced $\cos i$ values and 20 randomly picked Ω values. Except for a resonance at $i = 140^\circ$, the points are well confined between $A = 1.45$ and 1.75 , i.e. the average A value for the upper boundary is $A \approx 1.6$.

The limit $A = 1.75$ is good everywhere except at the resonance around $i = 140^\circ$. No similar feature is seen at $i = 40^\circ$. A rather safe upper boundary of the instability layer (mixed region) in experiments reported so far is therefore

$$Q_{\text{st}} = 2.4 \left(\lambda \sqrt{N} / (1 - e_{\text{out}}) \right)^{1/6} (fg)^{1/3}.$$

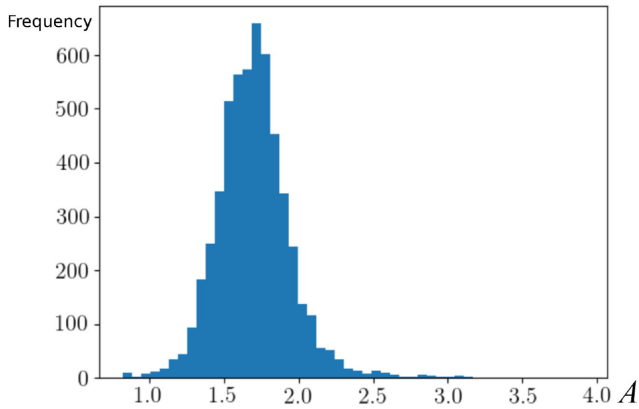


Figure 6. Distribution of A -values (stability limit points) for all 5445 points in the experiments with varying masses and eccentricities.

The coefficient 2.4 (rather than 2.3) covers also the more extensive set of experiments reported in the next section, and illustrated in Appendix A.

4 UNIVERSALITY

In order to test the validity of our function f , we have carried out a set of simulations to cover the range of masses at several inner eccentricities and longitudes of the ascending node with the constant value of the outer eccentricity of 0.5. We used a grid composed of five equally spaced values of e_{in} (uniform from 0 to 0.99), 11 equally spaced values of both Ω and $\cos i$ (the latter uniform from -1 to $+1$), three values of m_2 (0.1, 0.5, and 1.0), and three values of m_3 (0.1, 1.0, and 10). Since $m_1 = 2 - m_2$, this creates a grid of nine specific combinations of mass. We generate 45 diagrams similar to Fig. 1, and for each point in these diagrams (5445 in all) we calculate the A -value using equation (13). The distribution of A -values is shown in Fig. 6. The distribution is well described by a Gaussian with the mean of $A = 1.6$ and the standard deviation $\sigma = 0.26$. Therefore all points are below $A = 2.4$ at the 3σ level. We may use this as the limit of absolute stability, with the exception of resonances that occur in a few diagrams, similar to the $i = 140^\circ$ resonance in our previous Fig. 5. The exceptionally high values of A arise only at a relatively narrow range of Ω usually close to 0° and 180° ; a further study of these regions will appear in Paper II.

At this point we went back to equation (11) to find out if the third-order dependence on e_{in} and $\cos i$ is visible in the 45 new diagrams (Appendix A). In particular, we checked if another choice of coefficients $2/3, 1/2, \dots$, etc. would improve the fit to this larger data set, by reducing the standard deviation. The lowest value found was practically equal to 0.26; thus in this way we confirm the validity of the coefficients in equation (11) at the level of ± 4 per cent.

In addition to this grid, we have studied some individual cases, for instance $m_1 = m_2 = 1, m_3 = 0.1$, and $e_{\text{in}} = e_{\text{out}} = 0$, that are quite different from our standard case. The result of the study with the full resolution, using 3×10^5 three-body solutions, is displayed in Fig. 7. Another case has the same initial values, except $e_{\text{out}} = 0.3$ and the inclination resolution is lower (Fig. 8). These cases show that the general formula is applicable even though the outer orbit starts by being circular or at a small value. In an earlier paper (Valtonen et al. 2008) we have covered a wider range of e_{out} . We also checked the dependence on the mass of the external body over a wider range in our initial system, i.e. when $m_1 = 1.5, m_2 = 0.5$,

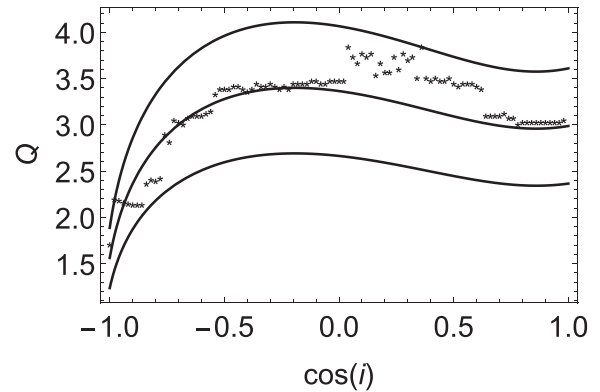


Figure 7. As in Fig. 5, but the 20 Ω simulations for $m_1 = m_2 = 1, m_3 = 0.1$, and $e_{\text{in}} = e_{\text{out}} = 0$.

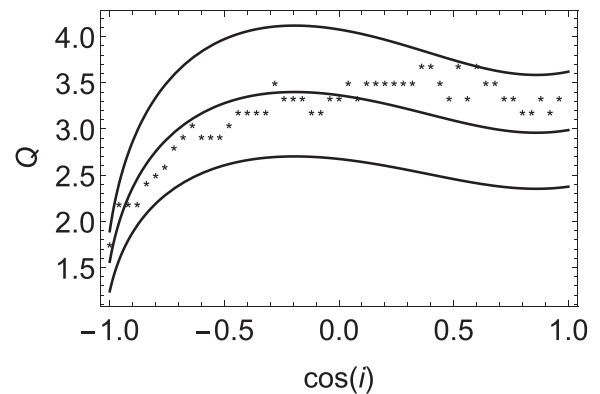


Figure 8. As above, but for the eccentricity $e_{\text{out}} = 0.3$.

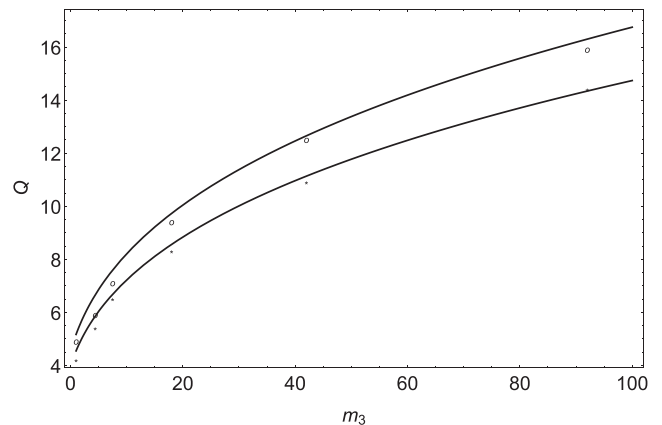


Figure 9. Upper (stability) limit Q (y-axis) as a function of m_3 (mass of the external body) when $m_1 = 1.5, m_2 = 0.5, e_{\text{in}} = 0.6$, and $e_{\text{out}} = 0.5$. Curves for $\cos i = 0.8$ (higher) and -0.8 (lower) are shown. The curves are displayed for $A = 2$.

$e_{\text{in}} = 0.6$, and $e_{\text{out}} = 0.5$. The results at two definite values of the inclination $\cos i = \pm 0.8$ are shown in Fig. 9.

5 DISCUSSION

The inner eccentricity–inclination term has not been published previously. Considering only the e_{in} dependence, and taking Q_{st} to be

a constant as a function of $\cos i$, Eggleton & Kiseleva (1995) and Mardling & Aarseth (1999) give

$$Q_{\text{st}} \propto (1 + e_{\text{in}}),$$

while Baily (private communication) suggests

$$Q_{\text{st}} \propto \left(1 + \frac{1}{2}e_{\text{in}}^2\right).$$

Our data for $\cos i = 1$ agree with an approximate formula (simpler than equation 11 with $\cos i = 1$)

$$Q_{\text{st}} \propto \left(1 - \frac{2}{3}e_{\text{in}} + 1.2e_{\text{in}}^2\right).$$

Therefore the formulae given by earlier studies are not very satisfactory. Moreover, we note that the $\cos i$ dependence is clearly a function of e_{in} in the experimental data set (Figs 1–3). The $\cos i$ dependence given in Valtonen & Karttunen (2006),

$$f(\cos i) = \left\{ \frac{1}{3} + [(1 + \cos i)(1.97 - \cos i)]^{0.8} \right\},$$

is a fair representation of the $f(\cos i)$ function at low eccentricity but not applicable to mid-range or high eccentricities. The same is true for the simpler form by Valtonen et al. (2008):

$$f(\cos i) = \left\{ \frac{7}{4} + \frac{1}{2} \cos i - \cos i^2 \right\}.$$

Mardling (2008) and Mardling (private communication; a computer code distributed privately) suggested a practically flat function at high eccentricity and a function that varies from flat at positive $\cos i$ to gently sloping (towards $\cos i = -1$) at negative $\cos i$ for mid-range and low eccentricity. It is an improvement over the flat function of Mardling & Aarseth (1999), but not a good description of the experimental data. Mardling & Aarseth (2001) keep Q_{st} constant with e_{in} , but assume $f(\cos i) = \{1 + 0.3i/\pi\}$. This is also a rough approximation to the numerically calculated data.

These expressions and our current formula and computer simulations show that retrograde orbits are more stable than direct orbits at the same Q -distance. The amount of difference in stability depends on the eccentricity of the inner orbit. The main exception to this rule occurs at the $i = 140^\circ$ resonance (Fig. 5). In a future paper we will also discuss the apparent resonance feature at $i = 140^\circ$. At this resonance the general formula fails, even though in a very small volume of the phase space.

The end result of this investigation is an analytical formula for determining the stability of a triple system, equation (13). The stability limit depends on the orbital parameters of the inner and outer binary and on the number of outer revolutions N required for stability. Since the latter depends on the problem at hand, its uncertainty may be absorbed in the coefficient A . Thus putting $A \sim 1$ in N -body simulations, and $A = 2.4$ in tests for absolute stability, together with $\Lambda = 1$, may constitute a practical stability formula:

$$Q_{\text{st}} = 10^{1/3} A [(f g)^2 / (1 - e_{\text{out}})]^{1/6},$$

where the functions f and g are given by equations (11) and (12), respectively. At the limit of $e_{\text{in}} = i = m_3 = 0$, $f g = 1$. The factor $10^{1/3}$ corresponds to $N = 10^4$, the standard value used in this study.

ACKNOWLEDGEMENTS

The authors thank the anonymous referee for thorough and helpful comments that helped to improve the paper considerably.

REFERENCES

- Aarseth S. J., 1973, *Vistas Astron.*, 15, 13
Aarseth S. J., 2003, *Gravitational N-body Simulations: Tools and Algorithms*. Cambridge Univ. Press, Cambridge
Aarseth S. J., Zare K., 1974, *Celest. Mech.*, 10, 185
Clairaut A., 1752, *Theorie de la Lune*. Imperial Academy of Science, St. Petersburg
Dvorak R., 1997, *Celest. Mech. Dyn. Astron.*, 68, 63
Eggleton P., Kiseleva L., 1995, *ApJ*, 455, 640
Euler L., 1752, *Recherches sur les irrégularités du mouvement de Jupiter et Saturne*
Ford E. B., Kozinsky B., Rasio F. A., 2000, *ApJ*, 535, 385
Funk B., Schwarz R., Pilat-Lohinger E., Suli A., Dvorak R., 2009, *Planet. Space Sci.*, 57, 434
Harrington R. S., 1972, *Celest. Mech.*, 6, 322
Heggie D. C., 2006, in Flynn C., ed., *Few Body Problem*. Ann. Univ. Turkuensis, Ser. A, Vol. 358, p. 20
Heggie D. C., Hut P., 2003, *The Gravitational Million-Body Problem: A Multidisciplinary Approach to Star Cluster Dynamics*. Cambridge Univ. Press, Cambridge
Huang T. Y., Innanen K. A., 1983, *AJ*, 88, 1064
Innanen K. A., Zheng J. Q., Mikkola S., Valtonen M. J., 1997, *AJ*, 113, 1915
Katz B., Dong S., Malhotra R., 2011, *Phys. Rev. Lett.*, 107, 181101
Kozai Y., 1962, *AJ*, 67, 591
Lagrange J. L., 1766, *Miscellanea Taurinensia*, 3, 1762
Lagrange J. L., 1778, *Hist. de l'acad. des sciences, 1774 = Lagrange 1867–92*, 6, 635
Laplace P. S., 1775, *Mem. Acad. Sci. Paris*, 1772
Laplace P. S., 1787, *Mem. Acad. Sci. Paris*, 1784, *Oeuvres* 11 (Paris, 1895), 49
Laskar J., 2013, *Progress Math. Phys.*, 66, 239
Lithwick Y., Naoz S., 2011, *ApJ*, 742, 94
Mardling R. A., 2008, in Vesperini E., Sills A., Giersz M., eds, *Proc. IAU Symp. Vol. 246, Dynamical Evolution of Dense Stellar Systems*. Cambridge Univ. Press, Cambridge, p. 199
Mardling R., Aarseth S., 1999, in Steves B. A., Roy A. E., eds, *The Dynamics of Small Bodies in the Solar System*. Kluwer, Dordrecht, p. 385
Mardling R. A., Aarseth S. J., 2001, *MNRAS*, 321, 398
Mikkola S., 1997, *Celest. Mech. Dyn. Astron.*, 67, 145
Newton I., 1687, *Philosophiae Naturalis Principia Mathematica*. Royal Society, London
Pilat-Lohinger E., Funk B., Dvorak R., 2003, *A&A*, 400, 1085
Roy A., Haddow M., 2003, *Celest. Mech. Dyn. Astron.*, 87, 411
Valtonen M., Karttunen H., 2006, *The Three-Body Problem*. Cambridge Univ. Press, Cambridge
Valtonen M., Mylläri A., Orlov V., Rubinov A., 2008, in Vesperini E., Sills A., Giersz M., eds, *Proc. IAU Symp. Vol. 246, Dynamical Evolution of Dense Stellar Systems*. Cambridge Univ. Press, Cambridge, p. 209

APPENDIX A: SAMPLE ILLUSTRATIONS FOR THE RANGE OF MASSES

Here we present results of simulations described in the beginning of Section 4 for different masses and uniformly sampled Ω . Everywhere in what follows $e_{\text{out}} = 0.50$. In addition to the three curves corresponding to $A = 1.15, 1.45$, and 1.75 , one more curve for $A = 2.0$ was added.

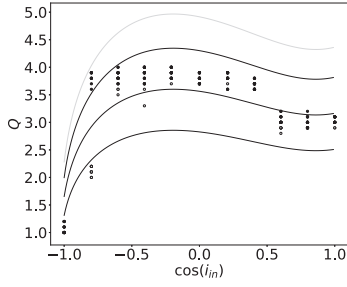


Figure A1. $m_1 = 1.00, m_2 = 1.00, m_3 = 0.10,$ and $e_{in} = 0.00.$

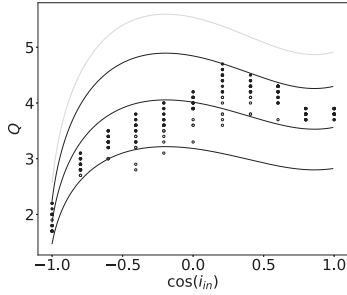


Figure A2. $m_1 = 1.00, m_2 = 1.00, m_3 = 1.00,$ and $e_{in} = 0.00.$

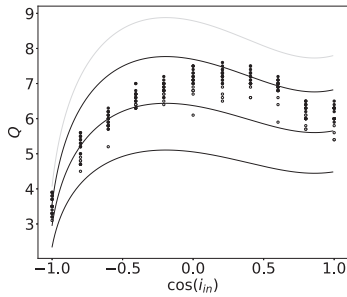


Figure A3. $m_1 = 1.00, m_2 = 1.00, m_3 = 10.00,$ and $e_{in} = 0.00.$

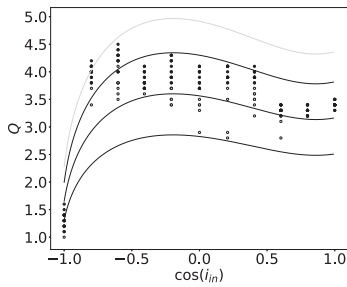


Figure A4. $m_1 = 1.50, m_2 = 0.50, m_3 = 0.10,$ and $e_{in} = 0.00.$

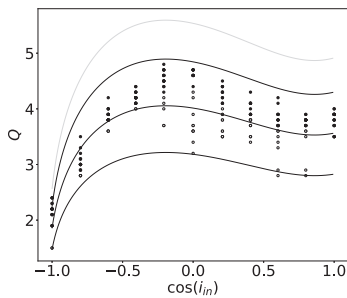


Figure A5. $m_1 = 1.50, m_2 = 0.50, m_3 = 1.00,$ and $e_{in} = 0.00.$

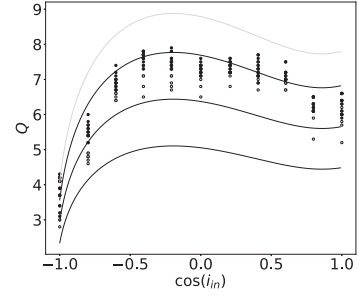


Figure A6. $m_1 = 1.50, m_2 = 0.50, m_3 = 10.00,$ and $e_{in} = 0.00.$

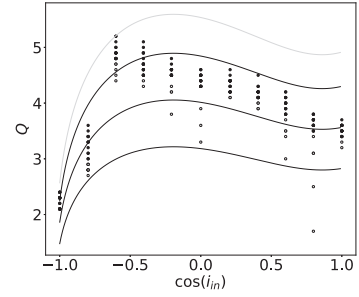


Figure A7. $m_1 = 1.90, m_2 = 0.10, m_3 = 1.00,$ and $e_{in} = 0.00.$

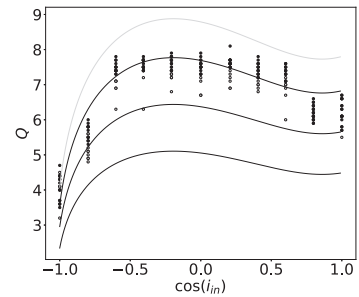


Figure A8. $m_1 = 1.90, m_2 = 0.10, m_3 = 10.00,$ and $e_{in} = 0.00.$

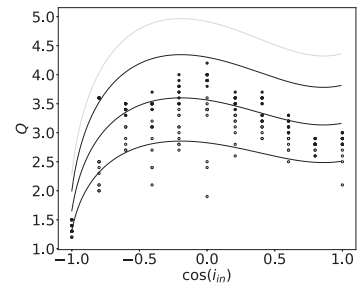


Figure A9. $m_1 = 1.90, m_2 = 0.10, m_3 = 0.10,$ and $e_{in} = 0.00.$

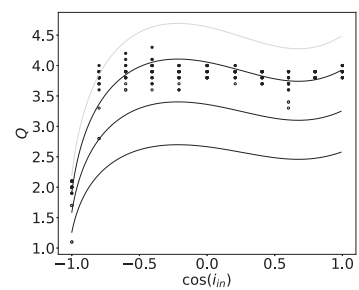


Figure A10. $m_1 = 1.00, m_2 = 1.00, m_3 = 0.10,$ and $e_{in} = 0.25.$

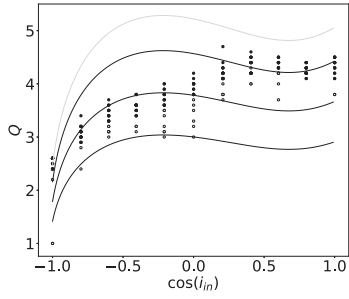


Figure A11. $m_1 = 1.00, m_2 = 1.00, m_3 = 1.00,$ and $e_{in} = 0.25.$

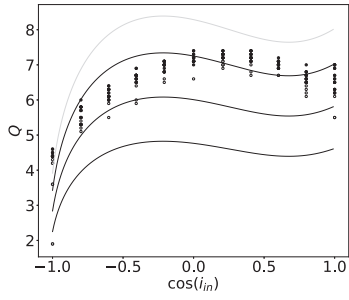


Figure A12. $m_1 = 1.00, m_2 = 1.00, m_3 = 10.00,$ and $e_{in} = 0.25.$

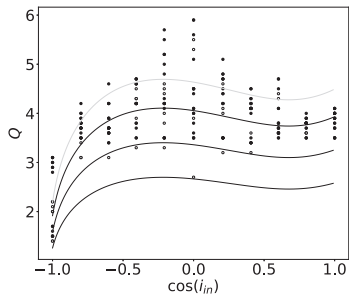


Figure A13. $m_1 = 1.50, m_2 = 0.50, m_3 = 0.10,$ and $e_{in} = 0.25.$

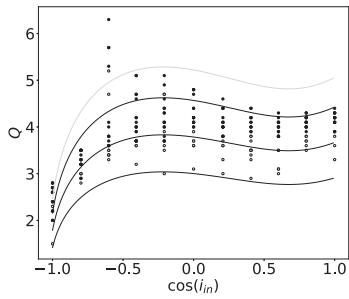


Figure A14. $m_1 = 1.50, m_2 = 0.50, m_3 = 1.00,$ and $e_{in} = 0.25.$

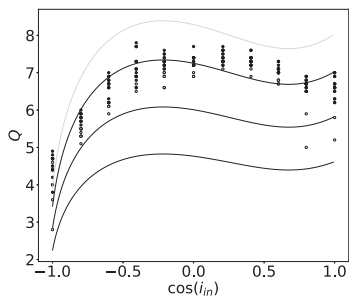


Figure A15. $m_1 = 1.50, m_2 = 0.50, m_3 = 10.00,$ and $e_{in} = 0.25.$

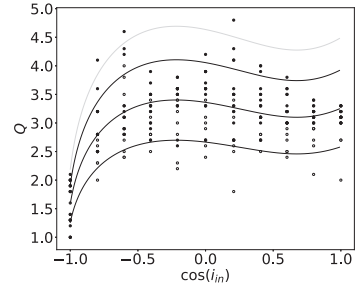


Figure A16. $m_1 = 1.90, m_2 = 0.10, m_3 = 0.10,$ and $e_{in} = 0.25.$

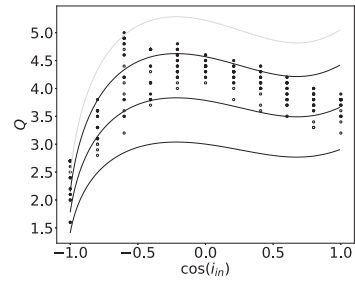


Figure A17. $m_1 = 1.90, m_2 = 0.10, m_3 = 1.00,$ and $e_{in} = 0.25.$

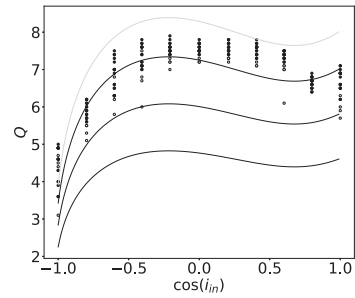


Figure A18. $m_1 = 1.90, m_2 = 0.10, m_3 = 10.00,$ and $e_{in} = 0.25.$

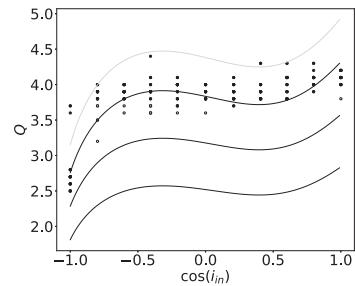


Figure A19. $m_1 = 1.00, m_2 = 1.00, m_3 = 0.10,$ and $e_{in} = 0.50.$

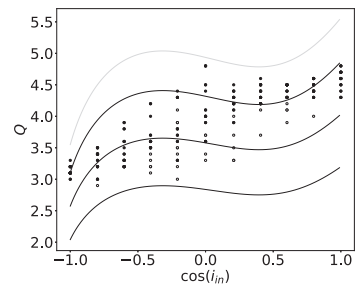


Figure A20. $m_1 = 1.00, m_2 = 1.00, m_3 = 1.00,$ and $e_{in} = 0.50.$

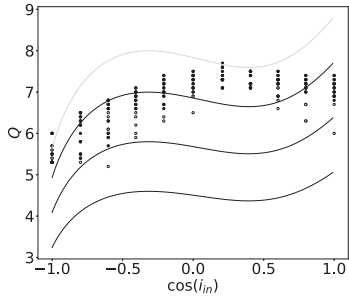


Figure A21. $m_1 = 1.00, m_2 = 1.00, m_3 = 10.00,$ and $e_{in} = 0.50.$

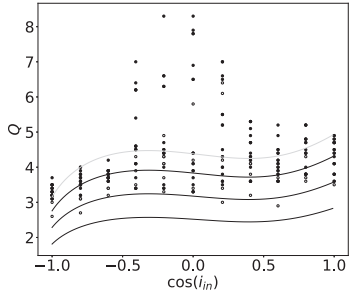


Figure A22. $m_1 = 1.50, m_2 = 0.50, m_3 = 0.10,$ and $e_{in} = 0.50.$

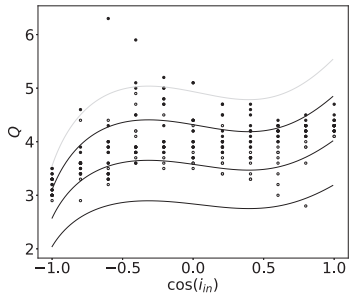


Figure A23. $m_1 = 1.50, m_2 = 0.50, m_3 = 1.00,$ and $e_{in} = 0.50.$

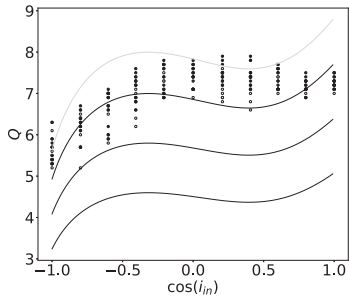


Figure A24. $m_1 = 1.50, m_2 = 0.50, m_3 = 10.00,$ and $e_{in} = 0.50.$

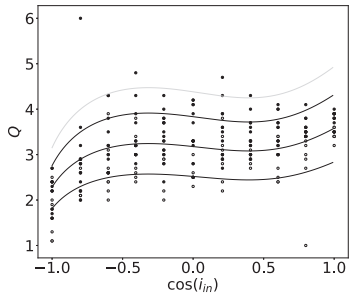


Figure A25. $m_1 = 1.90, m_2 = 0.10, m_3 = 0.10,$ and $e_{in} = 0.50.$

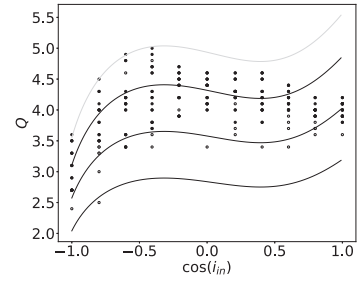


Figure A26. $m_1 = 1.90, m_2 = 0.10, m_3 = 1.00,$ and $e_{in} = 0.50.$

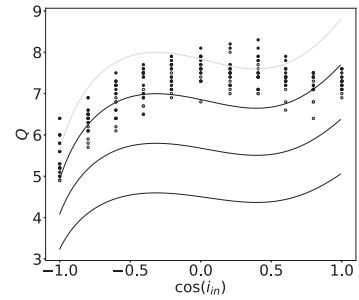


Figure A27. $m_1 = 1.90, m_2 = 0.10, m_3 = 10.00,$ and $e_{in} = 0.50.$

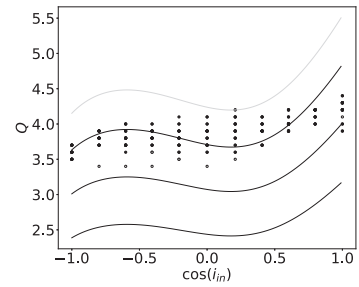


Figure A28. $m_1 = 1.00, m_2 = 1.00, m_3 = 0.10,$ and $e_{in} = 0.75.$

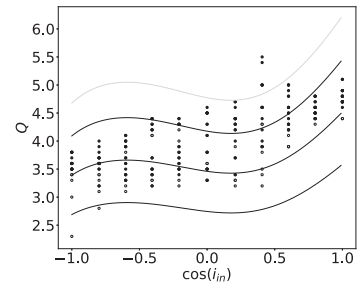


Figure A29. $m_1 = 1.00, m_2 = 1.00, m_3 = 1.00,$ and $e_{in} = 0.75.$

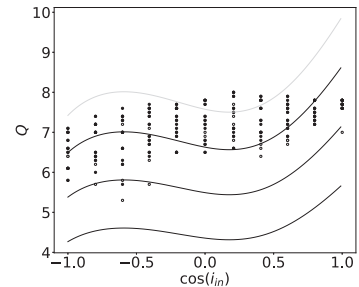


Figure A30. $m_1 = 1.00, m_2 = 1.00, m_3 = 10.00,$ and $e_{in} = 0.75.$

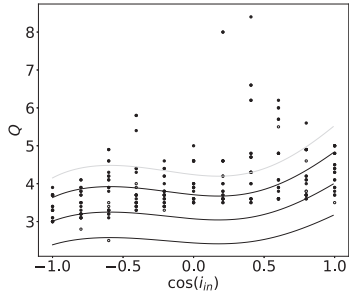


Figure A31. $m_1 = 1.50, m_2 = 0.50, m_3 = 0.10,$ and $e_{in} = 0.75.$

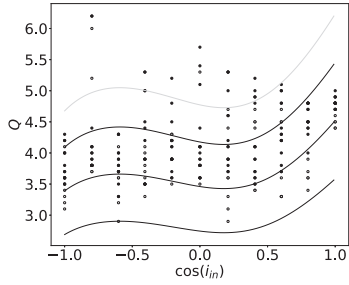


Figure A32. $m_1 = 1.50, m_2 = 0.50, m_3 = 1.00,$ and $e_{in} = 0.75.$

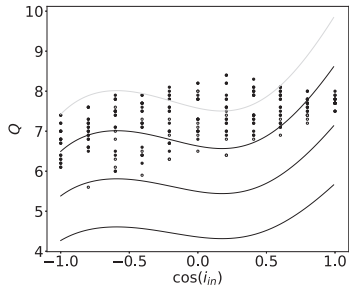


Figure A33. $m_1 = 1.50, m_2 = 0.50, m_3 = 10.00,$ and $e_{in} = 0.75.$

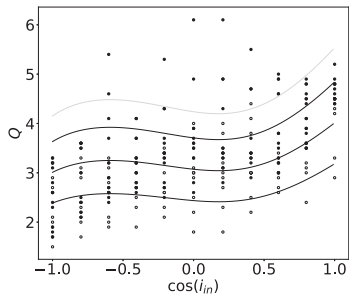


Figure A34. $m_1 = 1.90, m_2 = 0.10, m_3 = 0.10,$ and $e_{in} = 0.75.$

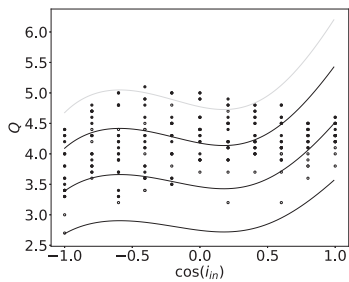


Figure A35. $m_1 = 1.90, m_2 = 0.10, m_3 = 1.00,$ and $e_{in} = 0.75.$

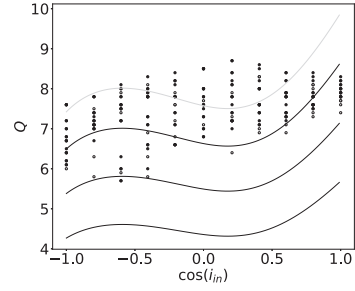


Figure A36. $m_1 = 1.90, m_2 = 0.10, m_3 = 10.00,$ and $e_{in} = 0.75.$

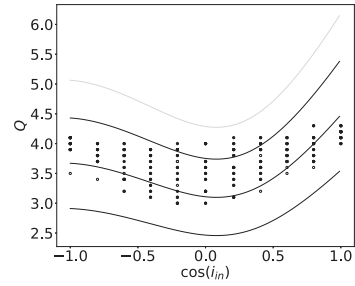


Figure A37. $m_1 = 1.00, m_2 = 1.00, m_3 = 0.10,$ and $e_{in} = 0.99.$

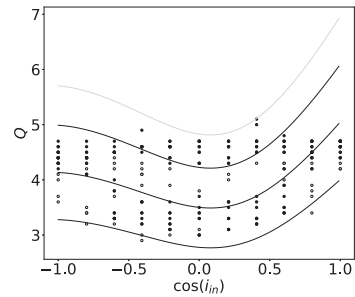


Figure A38. $m_1 = 1.00, m_2 = 1.00, m_3 = 1.00,$ and $e_{in} = 0.99.$

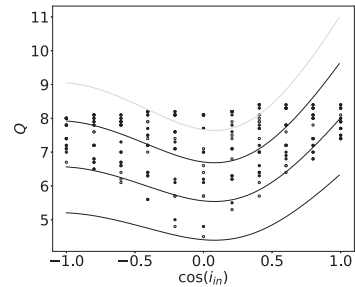


Figure A39. $m_1 = 1.00, m_2 = 1.00, m_3 = 10.00,$ and $e_{in} = 0.00.$

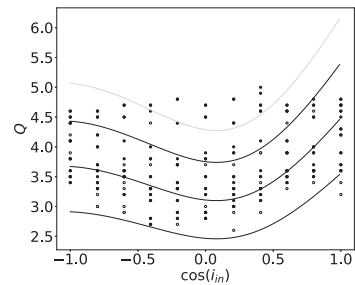


Figure A40. $m_1 = 1.50, m_2 = 0.50, m_3 = 0.10,$ and $e_{in} = 0.99.$

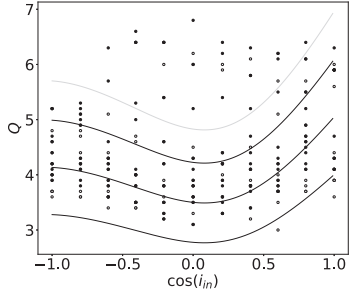


Figure A41. $m_1 = 1.50$, $m_2 = 0.50$, $m_3 = 1.00$, and $e_{in} = 0.99$.

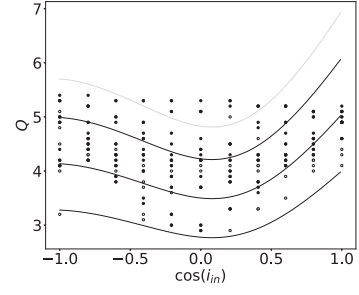


Figure A44. $m_1 = 1.90$, $m_2 = 0.10$, $m_3 = 1.00$, and $e_{in} = 0.99$.

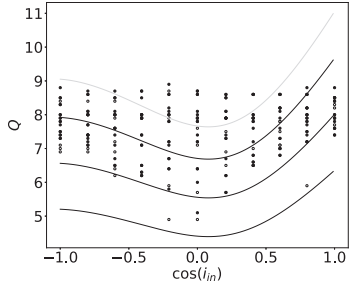


Figure A42. $m_1 = 1.50$, $m_2 = 0.50$, $m_3 = 10.00$, and $e_{in} = 0.99$.

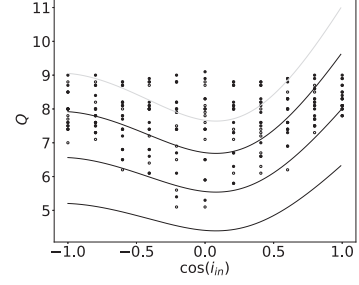


Figure A45. $m_1 = 1.90$, $m_2 = 0.10$, $m_3 = 10.00$, and $e_{in} = 0.99$.

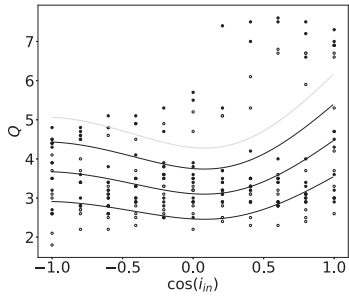


Figure A43. $m_1 = 1.90$, $m_2 = 0.10$, $m_3 = 0.10$, and $e_{in} = 0.99$.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.