Octupole-order Lidov-Kozai Population Statistics

Yubo Su

Today

1 10/22/20—Initial Thoughts

1.1 Equations

The equations of motion we want to study come from LML15. Describe the inner binary by $(a, e, I, \Omega, \omega)$ and the outer binary with "out" subscripts, and denote $I_{\rm tot} = I + I_{\rm out}$. Call the inner binary component masses m_1, m_2, and the tertiary mass m_3, define the inner binary total and reduced masses $m_{12} = m_1 + m_2$ and $\mu = m_1 m_2 / m_{12}$, and define the tertiary orbit total and reduced masses $m_{123} = m_{12} + m_3$ and $\mu_{\rm out} = m_{12} m_3 / m_{123}$. The equations of motion are ($j(e) = \sqrt{1-e^2}$)

$$\frac{\mathrm{d}a}{\mathrm{d}t} = -\frac{64}{5}\frac{a}{t_{\rm GW} j^7(e)}\left(1 + \frac{73}{24}e^2 + \frac{37}{96}e^4\right), \tag{1}$$

$$\begin{aligned}\frac{\mathrm{d}e}{\mathrm{d}t} = \frac{j(e)}{64 t_{\rm LK}}\Big\{&120 e \sin^2 I_{\rm tot} \sin(2\omega)\\
&+ \frac{15\epsilon_{\rm oct}}{8}\cos\omega_{\rm out}\left[\left(4 + 3e^2\right)\left(3 + 5\cos(2I_{\rm tot})\right)\right.\\
&\left.\times \sin\omega + 210 e^2 \sin^2 I_{\rm tot} \sin 3\omega\right]\\
&- \frac{15\epsilon_{\rm oct}}{4}\cos I_{\rm tot}\cos\omega\left[15(2 + 5e^2)\cos(2I_{\rm tot})\right.\\
&\left.+ 7\left(30 e^2 \cos(2\omega)\sin^2 I_{\rm tot} - 2 - 9e^2\right)\right]\sin\omega_{\rm out}\Big\}\\
&- \frac{304}{15}\frac{e}{t_{\rm GW} j^5(e)}\left(1 + \frac{121}{304}e^2\right),\end{aligned} \tag{2}$$

$$\begin{aligned}\frac{\mathrm{d}I}{\mathrm{d}t} = -\frac{3e}{32 t_{\rm LK} j(e)}\Big\{&10\sin(2I_{\rm tot})\left[e\sin(2\omega)\right.\\
&\left.+ \frac{5\epsilon_{\rm oct}}{8}\left(2 + 5e^2 + 7e^2\cos(2\omega)\right)\cos\omega_{\rm out}\sin\omega\right]\\
&+ \frac{5\epsilon_{\rm oct}}{8}\cos\omega\left[26 + 37e^2 - 35e^2\cos(2\omega)\right.\\
&\left.- 15\cos(2I_{\rm tot})\left(7e^2\cos(2\omega) - 2 - 5e^2\right)\right]\\
&\times \sin I_{\rm tot}\sin\omega_{\rm out}\Big\}\end{aligned} \tag{3}$$

$$
\begin{aligned}
\frac{\mathrm{d}\Omega}{\mathrm{d}t} = \frac{\mathrm{d}\Omega_{\mathrm{out}}}{\mathrm{d}t} = &-\frac{3\csc I}{32 t_{\mathrm{LK}} j(e)}\Big\{2\big[(2+3e^2-5e^2\cos(2\omega))\\
&+\frac{25\epsilon_{\mathrm{oct}}e}{8}\cos\omega\left(2+5e^2-7e^2\cos(2\omega)\right)\cos\omega_{\mathrm{out}}\big]\\
&\times\sin(2I_{\mathrm{tot}})-\frac{5\epsilon_{\mathrm{oct}}e}{8}\big[35e^2(1+3\cos(2I_{\mathrm{tot}}))\cos 2\omega\\
&-46-17e^2-15\left(6+e_1^2\right)\cos(2I_{\mathrm{tot}})\big]\\
&\times\sin I_{\mathrm{tot}}\sin\omega\sin\omega_{\mathrm{out}}\Big\},
\end{aligned}
\tag{4}
$$

$$
\begin{aligned}
\frac{\mathrm{d}\omega}{\mathrm{d}t} = \frac{3}{8t_{\mathrm{LK}}}&\Big\{\frac{1}{j(e)}\big[4\cos^2 I_{\mathrm{tot}}+(5\cos(2\omega)-1)\\
&\times\left(1-e^2-\cos^2 I_{\mathrm{tot}}\right)\big]+\frac{L\cos I_{\mathrm{tot}}}{L_{\mathrm{out}}j(e_{\mathrm{out}})}\big[2+e^2(3\\
&-5\cos(2\omega))\big]\Big\}+\frac{15\epsilon_{\mathrm{oct}}}{64t_{\mathrm{LK}}}\Big\{\left(\frac{L}{L_{\mathrm{out}}j(e_{\mathrm{out}})}+\frac{\cos I_{\mathrm{tot}}}{j(e)}\right)\\
&\times e\big[\sin\omega\sin\omega_{\mathrm{out}}\left[10(3\cos^2 I_{\mathrm{tot}}-1)(1-e^2)+A\right]\\
&-5B\cos I_{\mathrm{tot}}\cos\Theta\big]-\frac{j(e)}{e}\big[10\sin\omega\sin\omega_{\mathrm{out}}\cos I_{\mathrm{tot}}\\
&\times\sin^2 I_{\mathrm{tot}}\left(1-3e^2\right)+\cos\Theta\left(3A-10\cos^2 I_{\mathrm{tot}}+2\right)\big]\Big\}\\
&+\Omega_{\mathrm{GR}},
\end{aligned}
\tag{5}
$$

$$
\begin{aligned}
\frac{\mathrm{d}e_{\mathrm{out}}}{\mathrm{d}t} = \frac{15eL\,j(e_{\mathrm{out}})\epsilon_{\mathrm{oct}}}{256t_{\mathrm{LK}}e_{\mathrm{out}}L_{\mathrm{out}}}&\Big\{\cos\omega\big[6-13e^2\\
&+5(2+5e^2)\cos(2I_{\mathrm{tot}})+70e^2\cos(2\omega)\sin^2 I_{\mathrm{tot}}\big]\\
&\times\sin\omega_{\mathrm{out}}-\cos I_{\mathrm{tot}}\cos\omega_{\mathrm{out}}\big[5(6+e^2)\cos(2I_{\mathrm{tot}})\\
&+7\left(10e^2\cos(2\omega)\sin^2 I_{\mathrm{tot}}-2+e^2\right)\big]\sin\omega\Big\},
\end{aligned}
\tag{6}
$$

$$
\begin{aligned}
\frac{\mathrm{d}I_{\mathrm{out}}}{\mathrm{d}t} = &-\frac{3eL}{32t_{\mathrm{LK}}j(e_{\mathrm{out}})L_{\mathrm{out}}}\Big\{10\big[2e\sin I_{\mathrm{tot}}\sin(2\omega)\\
&+\frac{5\epsilon_{\mathrm{oct}}}{8}\cos\omega\left(2+5e^2-7e^2\cos(2\omega)\right)\sin(2I_{\mathrm{tot}})\sin\omega_{\mathrm{out}}\big]\\
&+\frac{5\epsilon_{\mathrm{oct}}}{8}\big[26+107e^2+5(6+e^2)\cos(2I_{\mathrm{tot}})\\
&-35e^2(\cos 2(I_{\mathrm{tot}})-5)\cos(2\omega)\big]\cos\omega_{\mathrm{out}}\sin I_{\mathrm{tot}}\sin\omega\Big\},
\end{aligned}
\tag{7}
$$

$$
\begin{aligned}
\frac{\mathrm{d}\omega_{\rm out}}{\mathrm{d}t} = \frac{3}{16t_{\rm LK}}&\left\{\frac{2\cos I_{\rm tot}}{j(e)}\left[2+e^2(3-5\cos(2\omega))\right]\right.\\
&+\frac{L}{L_{\rm out}j(e_{\rm out})}\left[4+6e^2+(5\cos^2 I_{\rm tot}-3)\right.\\
&\left.\left.\times\left[2+e^2(3-5\cos(2\omega))\right]\right]\right\}-\frac{15\epsilon_{\rm oct}e}{64t_{\rm LK}e_{\rm out}}\\
&\times\left\{\sin\omega\sin\omega_{\rm out}\left[\frac{L(4e_{\rm out}^2+1)}{e_{\rm out}L_{\rm out}j(e_{\rm out})}10\cos I_{\rm tot}\sin^2 I_{\rm tot}\right.\right.\\
&\times(1-e^2)-e_{\rm out}\left(\frac{1}{j(e)}+\frac{L\cos I_{\rm tot}}{L_{\rm out}j(e_{\rm out})}\right)\\
&\left.\times\left[A+10\left(3\cos^2 I_{\rm tot}-1\right)\left(1-e^2\right)\right]\right]+\cos\Theta\\
&\times\left[5B\cos I_{\rm tot}e_{\rm out}\left(\frac{1}{j(e)}+\frac{L\cos I_{\rm tot}}{L_{\rm out}j(e_{\rm out})}\right)\right.\\
&\left.\left.+\frac{L\left(4e_{\rm out}^2+1\right)}{e_{\rm out}L_{\rm out}j(e_{\rm out})}A\right]\right\}. \qquad (8)
\end{aligned}
$$

where $n=\sqrt{Gm_{12}/a^3}$ is the mean motion, $L=\mu\sqrt{Gm_{12}a}$ and $L_{\rm out}=\mu_{\rm out}\sqrt{Gm_{123}a_{\rm out}}$ are the circular angular momenta, and

$$t_{\rm LK}^{-1}=n\left(\frac{m_3}{m_{12}}\right)\left(\frac{a}{a_{\rm out}j(e_{\rm out})}\right)^3, \qquad (9)$$

$$t_{\rm GW}^{-1}=\frac{G^3\mu m_{12}^2}{c^5a^4}, \qquad (10)$$

$$\Omega_{\rm GR}=\frac{3Gnm_{12}}{c^2aj^2(e)}, \qquad (11)$$

$$\epsilon_{\rm oct}=\frac{m_2-m_1}{m_{12}}\frac{a}{a_{\rm out}}\frac{e_{\rm out}}{1-e_{\rm out}^2}, \qquad (12)$$

$$A\equiv 4+3e^2-\frac{5}{2}B\sin^2 I_{\rm tot}, \qquad (13)$$

$$B\equiv 2+5e^2-7e^2\cos(2\omega), \qquad (14)$$

$$\cos\Theta\equiv-\cos\omega\cos\omega_{\rm out}-\cos I_{\rm tot}\sin\omega\sin\omega_{\rm out}. \qquad (15)$$

These equations can be nondimensionalized via the following steps (I won't rewrite the equations): (i) multiply through by $t_{\rm LK,0}$ ($a=a_0$ and $e_{\rm out}=0$), and call $\tau\equiv t/t_{\rm LK,0}$ the new variable of differentiation, (ii) re-express all of the timescales as

$$\frac{t_{\rm LK,0}}{t_{\rm LK}} = \left(\frac{a}{a_0}\right)^{3/2} j^{-3}(e_{\rm out}), \tag{16}$$

$$\begin{aligned} \frac{t_{\rm LK,0}}{t_{\rm GW}} &= \frac{G^3 \mu m_{12}^3}{m_3 c^5 a^4} \frac{1}{n_0} \left(\frac{a_{\rm out}}{a_0}\right)^3, \\ &= \epsilon_{\rm GW} \left(\frac{a_0}{a}\right)^4, \end{aligned} \tag{17}$$

$$\epsilon_{\rm GW} \equiv \frac{G^3 \mu m_{12}^3 a_{\rm out}^3}{m_3 c^5 a_0^7 n_0}, \tag{18}$$

$$\begin{aligned} \Omega_{\rm GR} t_{\rm LK,0} &= \frac{3 G n m_{12}^2}{m_3 c^2 a} \frac{1}{n_0} \left(\frac{a_{\rm out}}{a_0}\right)^3, \\ &= \epsilon_{\rm GR} \left(\frac{a_0}{a}\right)^{5/2}, \end{aligned} \tag{19}$$

$$\epsilon_{\rm GR} = \frac{3 G m_{12}^2 a_{\rm out}^3}{m_3 c^2 a_0^4}. \tag{20}$$

(iii) re-express $\mathrm{d}a/\mathrm{d}t$ as

$$\frac{\mathrm{d}(a/a_0)}{\mathrm{d}\tau} = -\frac{64}{5} \frac{\epsilon_{\rm GW}}{j^{7/2}(e)} \left(\frac{a_0}{a}\right)^3 \left(1 + \frac{73}{24} e^2 + \frac{37}{96} e^4\right). \tag{21}$$

As such, the natural unit of length is $a_0 = 1$, the natural unit of time is $t_{\rm LK,0} = 1$, and everything else is dimensionless. When computing these ϵ, I use convention where $1 M_\odot = 1\,\mathrm{AU} = c = 1$, under which $G = 9.87 \times 10^{-9}$.

1.2 Points of Inquiry

The goal is to understand how the merger window varies as a function of $q \equiv m_1/m_2$ ($m_1 < m_2$) when the octupole order LK effects are important.

- First, let's set $\epsilon_{\rm GW} = 0$. It is well known that the octupole order LK is nonintegrable. What does the Fourier spectrum of the eccentricity look like? Will this help us get a delay time distribution between high-e phases?

  When the octupole effect is unimportant, the spectrum falls off exponentially over scales $\tau \simeq P_{\rm LK} j^{-1}(e_{\max})$, where $P_{\rm LK}$ is the quadrupole LK period. One imagines the tail of the spectrum gets heavier when $\epsilon_{\rm oct}$ is increased, and this might help us get the delay time distribution.

  A second way we can postprocess this is to take a histogram of $e(t)$. If there is some regular structure, it's likely this will allow us to compute the average rate of binary coalescence due to GW radiation.

- The goal is to understand the size of the merger window, ΔI, as a function of q. To do this, we numerically sample the merger time function $T_{\rm m}(I_0, q)$. At each I_0, the natural thing to do would be to try for $\sim 5 - 10$ random Ω, ω, and define the merger window to be where $T_{\rm m} \le 10^1 0\,\mathrm{yr}$.

2 11/02/20

We've done a lot more inquiry on this, read the Nov 3 weekly for a recap. In summary though, there should roughly be two ways to have LK-induced mergers:

- Quadrupole LK-induced mergers. The $e_{\max}$ of these systems is well understood, however, and the merger fraction should be easy to calculate: we can get the merger time using the $T_m \propto (1-e_{\max})^{-3}$ physically-justified fitting law from LL18 in the LK-induced regime, and we just have to evaluate where it crosses 10 Gyr.

- Octupole LK-induced mergers. For these, there is a characteristic initial inclination range for which orbit flipping occurs, which is a function of $\epsilon_{\rm oct}$. This can likely be calculated analytically, but I'm not sure yet.

  For these systems, there is a characteristic orbit flipping timescale that is robust up to a factor of a few (since K oscillates on a fixed timescale, and orbit flips occur whenever K crosses $-\eta/2$), call this $t_{\rm ELK}$. Thus, octupole LK-induced mergers occur over characteristic times $t_{\rm ELK}$ within the desired inclination window (as once the system reaches an orbit flipping eccentricity, it executes a one-shot merger, approximately). It is well known that $t_{\rm ELK}$ depends on $\epsilon_{\rm oct}$, see e.g. Antognini 2015.

  There is some small variability, however, in this picture, according to my plots. This can likely be attributed to the exact history of eccentricity maxima prior to the one-shot maximum, since particularly dissipative sequences (i.e. many large eccentricity maxima prior to undergoing "the big one") can shrink the orbit and change the LK cycle pattern.

Armed with this, we should have enough information to compute the indicator function $\mathbb{1}_{\rm merge}(a_{\rm in}, I_0, t_{\rm LK}, \epsilon_{\rm oct})$, whether a system will merge, by simply computing the two ranges of I_0 from above. This should be valid wherever we are in a strongly LK-induced regime, i.e. very large eccentricities are necessary to merge.

In fact, there's a possibility the critical $\epsilon_{\rm oct}$ can be calculated analytically, see Katz et. al. 2011 for the test particle case. Is this generalizable?

2.1 The Quadrupole Conserved Quantity

This is a quick derivation. To quadrupolar order, $L_{\rm out}$ is conserved, as is the total angular momentum, so

$$L_{\rm tot}^2 = L_{\rm out}^2 + L_{\rm in}^2 + 2L_{\rm out}L_{\rm in}\cos I, \tag{22}$$

$$\frac{L_{\rm tot}^2 - L_{\rm out}^2}{L_{\rm out}L_{\rm in,0}} = \eta j_{\rm in}^2 + 2j_{\rm in}\cos I, \tag{23}$$

$$\frac{1}{2}\left[\frac{L_{\rm tot}^2 - L_{\rm out}^2}{L_{\rm out}L_{\rm in,0}} - \eta\right] = j_{\rm in}\cos I - \eta\frac{e_{\rm in}^2}{2}. \tag{24}$$

This is the form of the constant given in LL18. Note that $\eta = \eta_0/j_{\rm out}$, where η_0 is evaluated at $e_{\rm out} = 0$ as well.

3 11/03/20

We found out that the critical inclination window has a clean fitting formula in the test particle limit, MLL16 (which agrees with the leading order expansion of Katz et al 2011). If we take $\epsilon_{\rm oct}$ to be small, then

$$\cos^2 I_0 \le \frac{\epsilon_{\rm oct}}{0.4}, \tag{25}$$

$$\Delta I_0 \le \sqrt{\frac{\epsilon_{\rm oct}}{0.4}}. \tag{26}$$

This marks the inclination window in which the octupole-order effects are expected to be dominant, if the inclination window is similar to the test particle case (it's not exactly, since with finite η the window is offset to larger I_0 than $I_{0,\rm lim}$)

We will make two simplifying assumptions:

- DA is valid.
- $e_{\rm lim}$ is sufficiently large that one-shot mergers occur.

If we assume that DA is valid, then

$$t_{\rm LK}\sqrt{1-e_{\max}^2} \gtrsim \frac{1}{n_{\rm out}}, \tag{27}$$

$$\frac{a^{3/2}}{a_{\rm out}^{3/2}}\frac{m_{12}}{m_3}\left(\frac{a_{\rm out}}{a}\right)^3\sqrt{1-e_{\max}^2}\left(1-e_{\rm out}^2\right)^{3/2} \gtrsim 1, \tag{28}$$

$$\frac{a_{\rm out}}{a} \gtrsim \frac{1}{j_{\min}^{2/3}}. \tag{29}$$

In the last line, we take $e_{\rm out} \sim 0.6$ for which $\left(1-e_{\rm out}^2\right)^{3/2} = 0.5$, and $m_{12}/m_3 = 2$. For some reference $e_{\max} = 10^{-4}$, $j_{\min}^{2/3} = 0.058$ (it's okay for DA to break down at $e_{\rm lim}$, since one shot mergers should occur). TODO this should probably be evaluated for $e_{\max}$ near the edge of the quadrupole merger window.

Recall also that

$$\epsilon_{\rm oct} = \frac{1-q}{1+q}\frac{a}{a_{\rm out}}\frac{e_{\rm out}}{1-e_{\rm out}^2}. \tag{30}$$

Thus, in the DA regime, we can explicitly write down the merger window for ELK-induced mergers

$$f_{\rm oct} \equiv \frac{2\Delta I_0}{\pi} \lesssim \sqrt{\frac{1-q}{1+q}\frac{1}{2}\frac{e_{\rm out}}{1-e_{\rm out}^2}}. \tag{31}$$

This isn't great, since my data suggest that ΔI_0 should be linear in $(1-q)/(1+q)$, where $\epsilon_{\rm oct} \simeq 0.01$ should be small enough to satisfy the assumptions. Bin's data also agree with this, at least up to $e_{\rm out} = 0.6$ (25°, 15°, 10° degree merger windows for $q = 0.2, 0.4, 0.6$ respectively).

So it's clear that the ELK-active window should be revised for the finite-η case.

3.1 Differentiating K

We showed earlier that the quadrupole-conserved quantity is

$$K = j\cos I + j^2\eta = \frac{1}{2}\frac{L_{\rm tot}^2 - L_{\rm out}^2}{L_{\rm out}L_{\rm in,0}}. \tag{32}$$

Differentiating the second expression, we obtain that

$$\frac{\mathrm{d}K}{\mathrm{d}t} = \frac{1}{2L_{\rm in,0}}\left(\frac{-L_{\rm tot}^2}{L_{\rm out}^2}\frac{\mathrm{d}L_{\rm out}}{\mathrm{d}t} - \frac{\mathrm{d}L_{\rm out}}{\mathrm{d}t}\right), \tag{33}$$

$$= -\frac{1 + L_{\rm tot}^2/L_{\rm out}^2}{2L_{\rm in,0}}\hat{\mathbf{L}}_{\rm out}\cdot\frac{\mathrm{d}\mathbf{L}_{\rm out}}{\mathrm{d}t}. \tag{34}$$

Indeed, from LML15, we see that this term vanishes to quadrupolar order, and that the only terms that survive are (I use 1 for in and 2 for out somewhat interchangeably, due to the source materials I'm pulling from)

$$\frac{\hat{\mathbf{L}}_{\rm out}}{L_{\rm in}}\cdot\frac{\mathrm{d}\mathbf{L}_{\rm out}}{\mathrm{d}t} = -\frac{75\epsilon_{\rm oct}}{64}\left[2(\mathbf{e}_1\cdot\hat{\mathbf{n}}_2)\left(\mathbf{j}_1\cdot\hat{\mathbf{n}}_2\right)\hat{\mathbf{L}}_2\cdot\left(\hat{\mathbf{u}}_2\times\mathbf{j}_1\right) + \left[\frac{8}{5}e_1^2 - \frac{1}{5} - 7(\mathbf{e}_1\cdot\hat{\mathbf{n}}_2)^2 + \left(\mathbf{j}_1\cdot\hat{\mathbf{n}}_2\right)^2\right]\hat{\mathbf{L}}_2\cdot(\hat{\mathbf{u}}_2\times\mathbf{e}_1)\right]. \tag{35}$$

Triple products let us simplify a little bit to be in terms of $\hat{\mathbf{v}}_2 = \hat{\mathbf{L}}_2\times\hat{\mathbf{e}}_2$ as in LML15. But this won't be useful: if we don't have the second derivative as well, we can't compute a characteristic oscillation frequency.

We can make one more approximation though: all of the $\mathbf{j}_1\cdot\hat{\mathbf{n}}_2$ terms are generally small, since they are of order j_{1z} which is not conserved exactly but remains small (this is similar to Katz's approximation). This means we're left with

$$\frac{\mathrm{d}K}{\mathrm{d}t} \approx \left(1 + \frac{\eta^2}{4}\right)\frac{75\epsilon_{\rm oct}}{64}\left[\frac{8}{5}e_1^2 - \frac{1}{5} - 7(\mathbf{e}_1\cdot\hat{\mathbf{n}}_2)^2\right]\mathbf{e}_1\cdot\hat{\mathbf{v}}_2. \tag{36}$$

We've assumed $L_{\rm tot}^2/L_{\rm out}^2 \sim 1 + \eta^2/2$, since typically $\mathbf{L}_{\rm in}$ and $\mathbf{L}_{\rm out}$ are misaligned by $I \simeq 90^\circ$. This is again in agreement with the Katz formula except now $\mathbf{n}_2$ and $\mathbf{v}_2$ are permitted to vary (and our prefactor).

Now, their $\Omega_{\rm e}$ is such that $e_z/e = \cos\Omega_{\rm e}$. This is a bit more difficult to generalize for us, but it shouldn't be impossible. We will return later if it proves interesting.

3.2 Timescale Analysis

Let's suppose we only consider systems with $T_{\rm m,0} > 10\,\mathrm{Gyr}$, that cannot merge in a Hubble time. For simplicity, we also require DA hold, and all three masses be comparable ($m_{12} = 2m_3$, say). Then this places constraint

$$\frac{5c^5a_0^4}{256G^3m_{12}^2\mu} > 10\,\mathrm{Gyr}. \tag{37}$$

If we fix $m_3 = 30M_\odot$, so $m_{12}^2\mu = q/(1+q)^2(30M_\odot)^3$, then

$$\frac{a_0}{0.202\,\mathrm{AU}}\left(\frac{2(1+q)^2}{q}\right)^{1/4} > 1. \tag{38}$$

This is a very weak constraint.

Nevertheless, if we are firmly in the LK-induced regime, LL18 can be used to easily compute the quadrupole merger window, $I_{0,\mathrm{merger}}^{+} - I_{0,\mathrm{merger}}^{-}$ by just using their Eq. (42) to compute $e_{\max}$. We know that in the LK-induced regime, ϵ_{GR} is very weak except for at very large $e_{\max}$, so we can likely omit it when computing the $I_{0,\mathrm{merger}}^{\pm}$ since these are very smooth mergers. Then, since $j_{\min}^6 = 10\,\mathrm{Gyr}/T_{\mathrm{m},0}$, we can obtain

$$\cos I_{0,\mathrm{merger}}^{-} - \cos I_{0,\mathrm{merger}}^{+} = \frac{1}{5}\sqrt{\left(5\eta - 4\eta j_{\min}^2\right)^2 - 20\left(\frac{5\eta^2}{4} - j_{\min}^2\left(3 + \frac{9\eta^2}{4}\right) + \eta^2 j_{\min}^4\right)}, \tag{39}$$

$$\approx \frac{j_{\min}\sqrt{60}}{5} + \mathcal{O}\left(\eta j_{\min}\right), \tag{40}$$

$$\approx \frac{\sqrt{60}}{5}\left(\frac{10\,\mathrm{Gyr}}{T_{m,0}}\right)^{1/6} + \mathcal{O}\left(\eta j_{\min}\right), \tag{41}$$

$$\approx \frac{\sqrt{60}}{5}\left(\frac{0.202\,\mathrm{AU}}{a_0}\right)^{2/3}\left(\frac{q}{2(1+q)^2}\right)^{1/6} + \mathcal{O}\left(\eta j_{\min}\right). \tag{42}$$

For $I_{0,\mathrm{merger}} \sim 90^\circ$, we can replace $\cos(x) \approx 90^\circ - x$, and so the difference of these cosines is just the negative difference of their arguments. For $q = 2/3$ and $a_0 = 100\,\mathrm{AU}$, this gives $\Delta I_{0,\mathrm{merger}} = 0.996^\circ$, while LL18 obtain 1.20°. Thus, this seems to be the right scaling, and something more accurate can prboably be obtained by using a more accurate $e_{\max}$.

3.3 Origin of Probabilistic Wings

I had a guess that may the origin of the probabilistic wings would be due to separatrix breaking due to ϵ_{oct}, where the IC would either be circulating or librating depending on ω_{in}. However, running some data, this does not appear to be the case (`emax_omega_sweep`), since the distribution of emaxes between small and large values changes as a function of time. This suggests that there is some characteristic timescale over which systems in this intermediate probabilistic zone (where I_0 gives a probability of merging within 10 Gyr) encounter ELK-induced orbit flips. This is not very surprising: it seems like K basically oscillates sinusoidally (the origin of the criterion I am thinking about above, for whether orbits can flip) and has some small extra bit of wander. Maybe the degree of wander can be quantified, in which case the behavior of these small probabilistic transition zones can be understood.

4 11/04/20

We mentioned that when changing q that the quadrupole strength. This could have entailed holding $t_{\rm LK}$, $\epsilon_{\rm GR}$, $\epsilon_{\rm GW}$, and η constant, where

$$t_{\rm LK} = \frac{1}{n_{\rm in}}\frac{m_{12}}{m_3}\left(\frac{a_{\rm out}\sqrt{1-e_{\rm out}^2}}{a_{\rm in}}\right)^3, \tag{43}$$

$$\eta = \frac{\mu}{\mu_{\rm out}}\sqrt{\frac{m_{12}a_{\rm in}}{m_{123}a_{\rm out}\left(1-e_{\rm out}^2\right)}}, \tag{44}$$

$$\epsilon_{\rm GR} = \frac{3Gm_{12}^2a_{\rm out}^3}{m_3c^2a_0^4}, \tag{45}$$

$$\epsilon_{\rm GW} \equiv \frac{G^3\mu m_{12}^3a_{\rm out}^3}{m_3c^5a_0^7n_{\rm in}}. \tag{46}$$

However, this is a bit contrived. Instead, we will just vary q with everything fixed, and also consider varying $e_{\rm out}$ such that $t_{\rm LK}$ remains fixed. We also should think about

$$\epsilon_{\rm oct} = \frac{1-q}{1+q}\frac{a}{a_{\rm out}}\frac{e_{\rm out}}{1-e_{\rm out}^2}, \tag{47}$$

$$= \frac{1-q}{1+q}\frac{a}{a_{\rm out,eff}}\frac{e_{\rm out}}{\sqrt{1-e_{\rm out}^2}} \tag{48}$$

where $a_{\rm out,eff} \equiv a_{\rm out}\sqrt{1-e_{\rm out}^2}$.

So the goal will be to continue running with all my current $q \in [0.2, 0.3, 0.4, 0.5, 0.7, 1.0]$, but also give some different $e_{\rm out}$ while holding $a_{\rm out,eff}$ constant. Let's try $e_{\rm out} = [0.6, 0.8, 0.9]$. Since used $a_{\rm out} = 4500$ AU for $e_{\rm out} = 0.6$, we will stick to $a_{\rm out,eff} = 3600$ AU, the corresponding $a_{\rm out}$ that we need are $a_{\rm out} = [4500, 6000, 8259]$ AU. This corresponds to increasing $e_{\rm out}$ by roughly $3\times$.

5 11/08/20

While the above simulations are running, we observe three new ideas:

- Notice that, for most parameters, if $e_{\rm in}$ can attain the critical value $e_{\rm lim}$ in the absence of GW, mergers occur when GW is turned on. This can be quantitatively evaluated, the one-shot merger eccentricity $e_{\rm os}$. Consider the scaling

$$\frac{\mathrm{d}\ln a}{\mathrm{d}t} = -\frac{64}{5j^7(e)}\left(1+\frac{73}{24}e^2+\frac{37}{96}e^4\right)\frac{G^3\mu m_{12}^2}{c^5a^4}, \tag{49}$$

$$\sim -j(e_{\rm max})\frac{64}{5j^7(e_{\rm max})}(4)\frac{G^3\mu m_{12}^2}{c^5a^4}. \tag{50}$$

We then seek where $\mathrm{d}\ln a/\mathrm{d}t \sim -1/t_{\rm LK}$, giving

$$j^6(e_{\rm os}) \equiv j_{\rm os} = \frac{256}{5}\frac{G^3\mu m_{12}^2}{c^5 a^4}\frac{1}{n}\frac{m_{12}}{m_3}\left(\frac{a_{\rm out,eff}}{a}\right)^3, \tag{51}$$

$$= \frac{256}{5}\frac{G^3\mu m_{12}^3}{m_3 c^5 a^4 n}\left(\frac{a_{\rm out,eff}}{a}\right)^3. \tag{52}$$

Recall that $n = \sqrt{Gm_{12}/a^3}$.

Then if $j_{\rm lim} \lesssim j_{\rm os}$ (i.e. $e_{\max} > e_{\rm os}$), then one shot mergers occur. We can derive a scaling relation

$$0 = \frac{3}{8}\left(j_{\rm lim}^2 - 1\right)\left(-3 + \frac{\eta^2}{4}\left(\frac{4}{5}j_{\rm lim}^2 - 1\right)\right) + \epsilon_{\rm GR}\left(1 - 1/j_{\rm lim}\right), \tag{53}$$

$$j_{\rm lim} \approx \frac{8\epsilon_{\rm GR}}{9 + 3\eta^2/4}. \tag{54}$$

This holds when $\epsilon_{\rm GR}, j_{\rm lim} \ll 1$. Matching the two expressions, we obtain the criterion for one shot mergers occuring at $e_{\rm lim}$:

$$256\frac{G^{5/2}a_{\rm out,eff}^3 m_{12}^{5/2}\mu}{5a^{11/2}c^5 m_3} \gtrsim \left(\frac{8}{9+3\eta^2/4}\frac{3Gm_{12}^2 a_{\rm out,eff}^3}{c^2 a^4 m_3}\right)^6, \tag{55}$$

$$a^{37/2} \gtrsim 5\cdot 1024\frac{G^{7/2}a_{\rm out,eff}^{15}m_{12}^{19/2}}{c^7 m_3^5 \mu\left(3+\eta^2/4\right)^6}, \tag{56}$$

$$\left(\frac{a}{a_{\rm out,eff}}\right) \gtrsim 0.0118\left(\frac{a_{\rm out,eff}}{3600\,\mathrm{AU}}\right)^{-7/37}\left(\frac{m_{12}}{50M_\odot}\right)^{17/37}\left(\frac{30M_\odot}{m_3}\right)^{10/37}\left(\frac{q/(1+q)^2}{1/4}\right)^{-2/37}. \tag{57}$$

where we have taken $q = 1$. In practice, we need a marginally smaller a than this, since in the one-shot merger regime, GW has to strongly modify the trajectory near $e_{\max}$. Note that both my parameters ($a = 100\,\mathrm{AU}$, $a_{\rm out,eff} = 3600\,\mathrm{AU}$) and Bin's ($a = 10\,\mathrm{AU}$, $a_{\rm out,eff} = 300\,\mathrm{AU}$) roughly satisfy this.

- The above addresses the case of ELK-induced mergers, effectively ($e_{\rm lim}$ induced by ELK). Instead, we can consider the case of ELK-*enhanced* mergers, in analogy with the quadrupole case. Here, quadrupole LK cannot merge the system alone, but with eccentricity cycles, the system is still able to merge. We can quantify this by computing an effective eccentricity over ELK-induced cycles:

$$\left\langle\frac{\mathrm{d}\ln a}{\mathrm{d}t}\right\rangle = -\left\langle\frac{a}{t_{\rm GW}}\right\rangle, \tag{58}$$

$$\approx -\frac{a}{t_{\rm GW,0}}\left\langle\frac{1 + 73e_{\max}^2/24 + 37e_{\max}^4/96}{j^6(e_{\max})}\right\rangle, \tag{59}$$

$$\equiv -\frac{a}{t_{\rm GW,0}}\underbrace{\left(\frac{1 + 73e_{\rm eff}^2/24 + 37e_{\rm eff}^4/96}{j^6(e_{\rm eff})}\right)}_{f(e_{\rm eff})}. \tag{60}$$

where the average is taken over many LK cycles, and $t_{\rm GW,0}$ denotes the $e = 0$ evaluation.

We can then ask what level of $e_{\rm eff}$ is required to induce merger within a Hubble time $t_{\rm Hubb}$. This can also be estimated

$$\frac{t_{\rm GW,0}}{t_{\rm Hubb}} \frac{1}{f(e_{\rm eff})} \lesssim 1, \tag{61}$$

$$\left(\frac{4t_{\rm Hubb}}{t_{\rm GW,0}}\right)^{1/6} \gtrsim j(e_{\rm eff}), \tag{62}$$

$$0.01461\left(\frac{100\,{\rm AU}}{a}\right)^{2/3} \gtrsim j(e_{\rm eff}), \tag{63}$$

$$1 - e_{\rm eff} \lesssim 1.068 \times 10^{-4}. \tag{64}$$

- Finally, we notice in the data that the zone of ELK excitation seems to exhibit a gap near $I_0 \approx 90^\circ$, unlike in the test mass limit, for which ELK is symmetric about $I_0 = 90^\circ$. We can understand this: the ELK-induced eccentricity oscillations in the massless limit occur due to ELK-induced oscillations in the quadrupole-conserved quantity $j_z = j\cos I$.

  When $\eta > 0$, the conserved quantity is instead $j\cos I + \eta j^2$ (LL18, or earlier in notes). Thus, we can understand that if I_0 is such that $j\cos I \lesssim \eta j^2$, that η suppresses ELK oscillations. This can be solved to estimate the range over which the conserved quantity is dominated by the η contribution:

  $$|\cos I| \lesssim \eta. \tag{65}$$

  Within this range, we expect any oscillations in j to be suppressed.

  NB: I don't think $j\cos I + \eta j^2$ is the "generalization" of j_z, since it is the z component of the total angular momentum. Instead, the generalization should be something like $-j\cos I + \eta j^2$, i.e. the z component of $\mathbf{L}_{\rm out} - \mathbf{L}_{\rm in}$, which is only conserved when $\eta \to 0$.

With the above, it seems we should be able to characterize what shapes the $T_{\rm m}$ behavior:

- In the very center, near $I_{\rm lim}$, we have a quadrupole merger window.
- At the very edge of the quadrupole merger window (towards $I > 90^\circ$, as the other direction is η-suppressed), we have some small ELK enhancement of the merger rate due to eccentricity oscillations.
- Farther away from the quadrupole merger window, there are ELK-induced orbit flips, which merger since $e_{\rm lim} < e_{\rm os}$.
- On the boundaries of the ELK orbit flipping region, there are some other ELK-enhanced mergers. This region is generally probabilistic, where the $e_{\max}$ distribution is bimodal. This is expected to be wider for smaller a, as a result.

6 11/15/20

We have a new idea. While the system is nonintegrable, it can only explore a bounded area of phase space, according to our simulations. We try to obtain a bound using the Hamiltonian. Consider the full Hamiltonian, which is $H(e_1, \omega_1, I_1, e_2, \omega_2, I_2)$. We omit the Ω dependencies since we are just evaluating it as a constant of motion, not trying to get EOMs. There are two primary simplifications:

- By the law of sines, $L_1 \sin I_1 = L_2 \sin I_2$, where $L_1 = \mu\sqrt{Gm_{12}a_{\rm in}(1-e_1^2)}$ and for L_2.

- By conservation of total angular momentum:

$$\frac{L_{\rm tot}^2}{L_{\rm out,0}^2} = j_{\rm out}^2 + \eta_0^2 j_{\rm in}^2 + 2 j_{\rm in} j_{\rm out} \eta_0 \cos I_{\rm tot}, \tag{66}$$

  where $\eta_0 = L_{1,0}/L_{2,0}$ is the angular momentum ratio at zero eccentricity.

Putting together these two constraints, these implicitly give (e_2, I_2) as a function of $\left(e_1, I_1, \eta_0, L_{\rm tot,i}\right)$, the initial total angular momentum. Thus, the Hamiltonian is reduced to $H(e_1, \omega_1, I_1, \omega_2)$. Of these, the quadrupole term does not depend on ω_2, so we see that to obtain the Hamiltonian of the entire system to quadrupolar order, we only need to specify e_1, ω_1, I_1. This already hints that oscillations in e_2 are bounded, and that indeed for systems sufficiently far from $I_{\rm lim}$, they cannot be driven to $I_{\rm lim}$ by octupole-induced DA oscillations.

To understand whether a system can be driven to flip orbits, we assume that $H_{\rm quad}$ evolves non-integrably within some bound region, and we wish to understand whether the bound region contains $H_{\rm lim} \equiv H_{\rm quad}(0,0,I_{\rm lim})$. Indeed, if octupole-order effects act slowly, within each orbit, the system should execute an approximately quadrupolar oscillation, so the only way it will reach $e_{\rm lim}$ is if it starts from $I_{\rm lim}$ (this is accurate so long as the $\epsilon_{\rm GR}$-induced flatness in the $e_{\max}$ curve can be neglected; maybe we can incorporate the $\Phi_{\rm GR}$ into the Hamiltonian? unlikely).

A simple criterion exists to compare the two. We assume that $H_{\rm quad}$ evolves randomly within an interval of magnitude $H_{\rm oct}$. By inspection, we can read off the values for which $H_{\rm oct}$ is maximized/minimized:

$$H_{\rm oct} \propto e_1 e_2 \left(A\cos\phi + 10\cos I_{\rm tot} \sin^2 I_{\rm tot} j_1^2 \sin\omega_1 \sin\omega_2\right), \tag{67}$$

$$A = 4 + 3e_1^2 - \frac{5(2 + 5e_1^2 - 7e_1^2\cos 2\omega_1)}{2} \sin I_{\rm tot}, \tag{68}$$

$$\cos\phi = -\cos\omega_1\cos\omega_2 - \cos I_{\rm tot}\sin\omega_1\sin\omega_2. \tag{69}$$

Let's assume $e_2 = e_{2,0}$ for simplicity, accurate to leading order. We can then bound $H_{\rm oct}$ by numerically solving. This proves to be quite good at reproducing the MLL16 fitting formula, I think (to be checked quantitatively)!

7 11/15/20

A clear problem arises though. When we are trying to reach $I_{\rm lim} = 90^\circ$ like in the TP limit, everything is fine. However, when $\eta > 0$, $I_{\rm lim} > 90^\circ$. To leading order, everything makes sense: an ELK-active region expands about $I_{\rm lim}$, and we can probably predict this (TODO try).

However, a conundrum arises: H is symmetric in I. If $|H_{\rm oct}|$ is large enough, everything within a zone centered on 90° should be able to access $I_{\rm lim}$. What is the origin of the gapped region then?

My current hypothesis is that this has to do with the actual amplitude of oscillations in K (returning to an early plot I made). Katz11 shows that $\mathrm{d}K/\mathrm{d}t \propto -K$ roughly, so it's likely that K cannot reach $K(I_{\rm lim})$ when it starts too near zero. This still doesn't explain why the ELK window itself is asymmetric though. It's surprising this becomes insufficient when $\eta > 0$.

For today/tomorrow, our goal will just be to show that this reproduces the correct formula in the TP limit, then we will also make a plot of $\Delta K(I)$ in the nonzero-η case, and use this to support our notion of a "gap." I don't think I can come up with anything quantitative, especially since I've basically only had the weekend to think about this.

Edit (11/17/20): I made a plot of the $\Delta K(I)$ plot and it turned out exactly as expected. It remains to be seen whether an obvious analytical bound can be written down, but it may not be necessary for the purposes of this paper.

8 01/15/21

I've been mostly just writing up notes, I gave up on an analytical expression.

Our notation in the paper is that $P_{\rm merge}$ is the merger probability, for a given I_0, q, and $e_{\rm out}$. Then

$$f_{\rm merge}(q, e_{\rm out}) \propto \int\limits_{-1}^{1} P_{\rm merge}(I_0, q, e_{\rm out})\,\mathrm{d}\cos I_0. \tag{70}$$

We also use the notation

$$f_{\rm merge}(q) \propto \int\limits_{0}^{0.9}\int\limits_{-1}^{1} P_{\rm merge}(I_0, q, e_{\rm out}) P(e_{\rm out})\,\mathrm{d}\cos I_0\,\mathrm{d}e_{\rm out}. \tag{71}$$

We are also motivated to investigate because

$$\frac{\mathrm{d}\ln a}{\mathrm{d}t} \propto \mu = \frac{q}{(1+q)^2}. \tag{72}$$

where

$$q \equiv m_2/m_1. \tag{73}$$

9 02/15/21—Analytic q distribution

I've procrastinated on further work on this paper, since it's almost done, but we add a brief blurb here. Consider if the primordial distribution of masses follows a power law $P(m) \propto m^{-p}$ over some $m_{\rm min}$ and $m_{\rm max}$; what is the distribution of q when two stars are drawn from this distribution? Moe & di Stephano, in some of their papers, point out that $P(q) \propto q^{-p}$, but I didn't see a derivation. We give one below.

WLOG, assume $m_2 \leq m_1$, then

$$
\begin{aligned}
P(m) &\propto m^{-p}, && (74)\\
&= \frac{m^{-p}}{N}, && (75)\\
N &= \frac{1}{p-1}\left(m_{\min}^{1-p} - m_{\max}^{1-p}\right), && (76)\\
P(q) &= \int_{m_{\min}}^{m_{\max}} \int_{m_{\min}}^{m_1} \delta\left(\frac{m_2}{m_1} - q\right) P(m_1) P(m_2)\, \mathrm{d}m_2\, \mathrm{d}m_1, && (77)\\
&= \int_{m_{\min}}^{m_{\max}} m_1 P(m_1) P(q m_1)\, \mathrm{d}m_1, && (78)\\
&= \frac{p-1}{m_{\min}^{1-p} - m_{\max}^{1-p}} \int_{m_{\min}}^{m_{\max}} m_1 \frac{p-1}{m_{\min}^{1-p} - m_1^{1-p}} (m_1)^{-p} (q m_1)^{-p}\, \mathrm{d}m_1, && (79)\\
&= q^{-p} \frac{(p-1)^2}{m_{\min}^{1-p} - m_{\max}^{1-p}} \int_{m_{\min}}^{m_{\max}} \frac{m_1^{-2p+1}}{m_{\min}^{1-p} - m_1^{1-p}}\, \mathrm{d}m_1. && (80)
\end{aligned}
$$

Here, $q_{\min} \equiv m_{\min}/m_{\max}$. The case where $m_2 \geq m_1$ equates to flipping the masses, so $P(q)$ is unaffected, and $P(q) \propto q^{-p}$. In fact, by normalization, we find that

$$
P(q) = \frac{p-1}{q_{\min}^{1-p} - 1} q^{-p}, \tag{81}
$$

where $q_{\min} = m_{\min}/m_{\max}$. These two expressions probably agree? Not sure

What about if, instead of drawing m_1 and m_2 from a distribution randomly, we required that their total mass equal m_{12}? Assume $m_1 \geq m_2$, then

$$
\begin{aligned}
P(q) &= \int_{m_{\min}}^{m_{\max}} \int_{m_{\min}}^{m_1} \delta\left(\frac{m_2}{m_1} - q\right) \delta\left(\frac{m_1 + m_2}{m_{12}} - 1\right) P(m_1) P(m_2)\, \mathrm{d}m_2\, \mathrm{d}m_1, && (82)\\
&= \int_{m_{\min}}^{m_{\max}} m_{12} \delta\left(\frac{m_{12} - m_1}{m_1} - q\right) P(m_1) P(m_{12} - m_1)\, \mathrm{d}m_1, && (83)\\
&= \int_{m_{\min}}^{m_{\max}} m_{12} \delta\left(\frac{m_{12}}{m_1} - (1+q)\right) P(m_1) P(m_{12} - m_1)\, \mathrm{d}m_1, && (84)\\
&= P\left(\frac{m_{12}}{1+q}\right) P\left(m_{12} - \frac{m_{12}}{1+q}\right) \left(\frac{m_{12}}{1+q}\right)^2, && (85)\\
&\propto \frac{q^{-p}}{(1+q)^{-2p+2}}. && (86)
\end{aligned}
$$