Lidov-Kozai 90° Attractor

Yubo Su

Date

1 Equations

1.1 Bin's Papers

Our major references will be Bin's paper with Diego + Dong in 2015 (LML15) and Bin's later paper with Dong on spin-orbit misalignment (LL18). The target of study is §4.3 of LL18, where a 90° attractor in spin-orbit misalignment seems to appear when the octupole effect is negligible.

The easiest formulation is just to express everything in terms of **L** and **e**, following LL18. We drop octupole terms and hold the third perturber constant. These equations come out to be (Eqs. 4–5 w/ substitutions)

$$\frac{\mathrm{d}\mathbf{L}}{\mathrm{d}t} = \frac{3}{4t_{LK}}\mu\sqrt{Gm_{12}a}\left[(\mathbf{j}\cdot\hat{n}_2)(\mathbf{j}\times\hat{n}_2) - 5(\mathbf{e}\cdot\hat{n}_2)(\mathbf{e}\times\hat{n}_2)\right], \tag{1}$$

$$\frac{\mathrm{d}\mathbf{e}}{\mathrm{d}t} = \frac{3}{4t_{LK}}\left[(\mathbf{j}\cdot\hat{n}_2)(\mathbf{e}\times\hat{n}_2) + 2\mathbf{j}\times\mathbf{e} - 5(\mathbf{e}\cdot\hat{n}_2)(\mathbf{j}\times\hat{n}_2)\right]. \tag{2}$$

Note that $\mathbf{j} \equiv \sqrt{1-e^2}\hat{\mathbf{L}} = \frac{\mathbf{L}}{\mu\sqrt{Gm_{12}a}}$. $m_{12} = m_1 + m_2$ and $\mu = m_1 m_2/m_{12}$. We've defined

$$t_{LK} \equiv \frac{L_1}{\mu_1\Phi_0} = \frac{1}{n_1}\left(\frac{m_{12}}{m_3}\right)\left(\frac{a_2}{a}\right)^3\left(1-e_2^2\right)^{3/2}. \tag{3}$$

Here, $n_1 \equiv \sqrt{Gm_{12}/a^3}$. Thus, $1/t_{LK} \propto a^{3/2}$.

GW radiation (Peters 1964) cause decays of **L** and **e** as

$$\left.\frac{\mathrm{d}\mathbf{L}}{\mathrm{d}t}\right|_{GW} = -\frac{32}{5}\frac{G^{7/2}}{c^5}\frac{\mu^2 m_{12}^{5/2}}{a^{7/2}}\frac{1+7e^2/8}{\left(1-e^2\right)^2}\hat{L}, \tag{4}$$

$$\left.\frac{\mathrm{d}\mathbf{e}}{\mathrm{d}t}\right|_{GW} = -\frac{304}{15}\frac{G^3}{c^5}\frac{\mu m_{12}^2}{a^4\left(1-e^2\right)^{5/2}}\left(1+\frac{121}{304}e^2\right)\mathbf{e}. \tag{5}$$

Here, $m_{12} \equiv m_1 + m_2$, and a is implicitly defined by **L** and e. The last GR effect is precession of $\vec{e}$, which acts as

$$\left.\frac{\mathrm{d}\mathbf{e}}{\mathrm{d}t}\right|_{GR} = \frac{1}{t_{GR}}\hat{\mathbf{L}}\times\mathbf{e}, \tag{6}$$

$$\frac{1}{t_{GR}} \equiv \frac{3Gnm_{12}}{c^2 a\left(1-e^2\right)}. \tag{7}$$

Note that $t_{GR}^{-1} \propto a^{-5/2}$.

Given this system (from LML15 + LL18), we can then add the spin-orbit coupling term (from de Sitter precession), which is given in LL18 to be

$$\frac{\mathrm{d}\hat{\mathbf{S}}}{\mathrm{d}t} = \frac{1}{t_{SL}}\hat{\mathbf{L}} \times \hat{\mathbf{S}}, \tag{8}$$

$$\frac{1}{t_{SL}} \equiv \frac{3Gn\left(m_2 + \mu/3\right)}{2c^2 a\left(1-e^2\right)}. \tag{9}$$

Note that μ is the reduced mass of the inner binary. We can drop the back-reaction term since $S \ll L$. Thus, $t_{SL}^{-1} \propto a^{-5/2}$ as well.

Finally, an adiabaticity parameter can be defined:

$$\mathscr{A} \equiv \left|\frac{\Omega_{SL}}{\Omega_L}\right|. \tag{10}$$

Here, $\Omega_L \simeq \frac{3(1+4e^2)}{8t_{LK}\sqrt{1-e^2}}\left|\sin 2I\right|$ is an approximate rate of change of L during an LK cycle

It's natural to redimensionalize to the initial LK time such that

$$\frac{1}{t_{LK,0}} \equiv \left(\frac{a}{a_0}\right)^{3/2}\frac{1}{t_{LK}}, \tag{11}$$

since nothing else in t_{LK} is changing. The next natural timescale for gravitational waves is

$$\frac{1}{t_{GW}} \equiv \frac{G^3\mu m_{12}^2}{c^5 a^4} \equiv \frac{1}{t_{GW,0}}\left(\frac{a_0}{a}\right)^4 \equiv \epsilon_{GW}\frac{1}{t_{LK,0}}\left(\frac{a_0}{a}\right)^4. \tag{12}$$

We can repeat the procedure for the GR precession term and the spin-orbit coupling terms:

$$\frac{1}{t_{GR}} = \epsilon_{GR}\frac{1}{t_{LK,0}}\left(\frac{a_0}{a}\right)^{5/2}, \tag{13}$$

$$\frac{1}{t_{SL}} = \epsilon_{SL}\frac{1}{t_{LK,0}}\left(\frac{a_0}{a}\right)^{5/2}. \tag{14}$$

Thus, finally, if we let $\tau = t/t_{LK,0}$, then we obtain full equations of motion (note that $a_0 = 1$ below)

$$\begin{aligned}
\frac{\mathrm{d}\mathbf{L}}{\mathrm{d}\tau} &= \left(\frac{a}{a_0}\right)^{3/2}\frac{3}{4}\sqrt{a}\left[(\mathbf{j}\cdot\hat{n}_2)(\mathbf{j}\times\hat{n}_2) - 5(\mathbf{e}\cdot\hat{n}_2)(\mathbf{e}\times\hat{n}_2)\right] \\
&\quad - \epsilon_{GW}\left(\frac{a_0}{a}\right)^4\frac{32}{5}\frac{1+7e^2/8}{\left(1-e^2\right)^{5/2}}\mathbf{L},
\end{aligned} \tag{15}$$

$$\begin{aligned}
\frac{\mathrm{d}\mathbf{e}}{\mathrm{d}\tau} &= \left(\frac{a}{a_0}\right)^{3/2}\frac{3}{4}\left[(\mathbf{j}\cdot\hat{n}_2)(\mathbf{e}\times\hat{n}_2) + 2\mathbf{j}\times\mathbf{e} - 5(\mathbf{e}\cdot\hat{n}_2)(\mathbf{j}\times\hat{n}_2)\right] \\
&\quad - \epsilon_{GW}\left(\frac{a_0}{a}\right)^4\frac{304}{15}\frac{1}{\left(1-e^2\right)^{5/2}}\left(1+\frac{121}{304}e^2\right)\mathbf{e} \\
&\quad + \epsilon_{GR}\left(\frac{a_0}{a}\right)^{5/2}\frac{1}{1-e^2}\hat{\mathbf{L}}\times\mathbf{e},
\end{aligned} \tag{16}$$

$$\frac{\mathrm{d}\hat{\mathbf{S}}}{\mathrm{d}t} = \epsilon_{SL}\left(\frac{a_0}{a}\right)^{5/2}\frac{1}{1-e^2}\hat{\mathbf{L}}\times\hat{\mathbf{S}}. \tag{17}$$

For reference, we note that $a = |\mathbf{L}|^2/\left(\mu^2 G m_{12}\left(1-e^2\right)\right)$, while $\mathbf{j} = \mathbf{L}/\left(\mu\sqrt{G m_{12} a}\right)$. To invert $a(\mathbf{L})$ and $\mathbf{J}(\mathbf{L})$ in this coordinate system where $a_0 = 1$, it is easiest to choose the angular momentum dimensions such that $\mu\sqrt{G m_{12}} = 1$, such that now

$$|\mathbf{L}(t=0)| \equiv \mu\sqrt{G m_{12} a_0\left(1-e_0^2\right)} = \sqrt{\left(1-e_0^2\right)}, \tag{18}$$

$$a = \frac{|\mathbf{L}|^2}{1-e^2}, \tag{19}$$

$$\mathbf{j} = \frac{\mathbf{L}}{\sqrt{a}} = \hat{\mathbf{L}}\sqrt{1-e^2}. \tag{20}$$

Finally, the timescales are

$$t_{LK,0} = \frac{1}{n}\frac{m_{12}}{m_3}\left(\frac{a_2}{a(t=0)}\right)^3\left(1-e_2^2\right)^{3/2}, \tag{21}$$

$$\epsilon_{GW} \equiv \frac{t_{LK,0}}{t_{GW,0}} = \frac{1}{n}\frac{m_{12}}{m_3}\frac{a_2^3}{(a(t=0))^7}\left(1-e_2^2\right)^{3/2}\frac{G^3\mu m_{12}^2}{c^5}, \tag{22}$$

$$\epsilon_{GR} \equiv \frac{t_{LK,0}}{t_{GR,0}} = \frac{m_{12}}{m_3}\frac{a_2^3}{(a(t=0))^4}\left(1-e_2^2\right)^{3/2}\frac{3Gm_{12}}{c^2}, \tag{23}$$

$$\epsilon_{SL} \equiv \frac{t_{SL,0}}{t_{GR,0}} = \frac{m_{12}}{m_3}\frac{a_2^3}{(a(t=0))^4}\left(1-e_2^2\right)^{3/2}\frac{3G\left(m_2+\mu/3\right)}{2c^2}. \tag{24}$$

The adiabacitity parameter

$$\mathscr{A} \equiv \left|\frac{\Omega_{SL}}{\Omega_L}\right| = \frac{\epsilon_{SL}}{t_{LK,0}}\left(\frac{a_0}{a}\right)^{5/2}\frac{1}{1-e^2}\left[\frac{3(1+4e^2)}{8t_{LK,0}\sqrt{1-e^2}}\left(\frac{a}{a_0}\right)^{3/2}|\sin 2I|\right]^{-1}, \tag{25}$$

(note that Ω_L is a somewhat averaged sense, see LL18) can be evaluated in these units as

$$\mathscr{A} = \epsilon_{SL}\left(\frac{a_0}{a}\right)^4\frac{1}{\sqrt{1-e^2}}\frac{8}{3\left(1+4e^2\right)|\sin 2I|}. \tag{26}$$

Note also that the Hamiltonian is just

$$H = \Omega_{SL}\hat{\mathbf{L}}\cdot\hat{\mathbf{S}}, = \epsilon_{SL}\left(\frac{a_0}{a}\right)^{5/2}\frac{1}{1-e^2}\hat{\mathbf{L}}\cdot\hat{\mathbf{S}}. \tag{27}$$

1.2 Maximum Eccentricity and Merger Time

Note that, since we are only evolving $\mathbf{L}$ and $\mathbf{e}$, and not $\mathbf{L}_2$ and $\mathbf{e}_2$, we are in the test mass approximation, under which we set $\eta = 0$ in Bin's equations. As such, the maximum eccentricity satisfies (Eq 42 of LL18 with $\eta \to 0$)

$$\frac{3}{8}\frac{j_{\min}^2-1}{j_{\min}^2}\left[5\cos I_0^2 - 3j_{\min}^2\right] + \epsilon_{GR}\left(1-\frac{1}{j_{\min}}\right) = 0. \tag{28}$$

Note that ϵ_{GR} is exactly as we defined above, incidentally, and that when GR is negligible, this reduces to the classic $j_{\min} \equiv \sqrt{1-e_{\max}^2} = \sqrt{\frac{5}{3}\cos^2 I_0}$. Since ϵ_{GR} is generally very small for most of the evolution, this generally reduces to the well known

$$e_{\max} = \sqrt{1-\frac{5}{3}\cos^2 I_0}. \tag{29}$$

This only fails to saturate for extremely high eccentricities, so $I_0 \to 90°$.

1.3 Attractor Behavior

Proposal: The reason the 90° attractor appears is that the initial θ_{sb} is roughly stationary for $\mathscr{A} \ll 1$ (only small kicks during each LK cycle, as long as the maximum eccentricity isn't too large), then as we enter the transadiabatic regime, the L-K cycles die down and we simply have conservation of adiabatic invariant.

The latter half of this follows the LL18 claim, where the requirement that $\epsilon_{GR} \lesssim 9/4$ (GR precession of pericenter is slow enough that L-K survives) equates to $\mathscr{A} \lesssim 3$. The former half of this is somewhat tricky, but we can understand what is happening if we consider what is happening in the frame corotating with $\Omega_{SL,e=0}$ about $\hat{z}$: every time that a LK cycle appears, Ω_{SL} becomes much larger, and the axis of precession changes from $\hat{z}$ to the location of $\hat{L}$ very briefly. We can imagine this as a kick in this corotating frame (which is the right frame to consider for $\mathscr{A} \ll 1$). In the limit that I does not change very much between L-K cycles, and the azimuthal angle of $\hat{L}$ is roughly symmetric, the impulses roughly cancel out in the θ_{sb} direction. In other words, after two LK cycles, θ_{sb} does not change much in the corotating frame. This is indeed the picture that we obtain when we observe the plot.

As such, the hypothesis is that if $\mathscr{A} \gtrsim 1$ is satisfied while the kicks are still *small*, then deviations about 90° cannot be very large, and adiabatic invariance tilts us right over. On the other hand, if the kicks have become *large*, then θ_{sb} after any particular LK cycle is far from 90°, and this is frozen in during the adiabatic invariance phase. This explains the key observation that the initial θ_{sb} eventually becomes the final θ_{sl}, regardless of whether it is 90°. Furthermore, it explains why the kicks to θ_{eff} become larger over time, but peak smaller for larger I_0.

The natural way to think about this is to consider the evolution of the trajectory in (a,e) space. There are two curves that can be drawn on here, $\mathscr{A} \sim 1$ and $|\Delta\theta_{sb}| \sim 1$ (the kick size), and then we can see which one gets crossed first. The hypothesis is that the first always gets crossed first, but if $e_{\max}$ is too large, then the second gets crossed in the same LK cycle, and we get kicked far away from the starting θ, and have this frozen into θ_{sl}. We need to find out how to draw these boundaries in (a,e) space. Drawing $\mathscr{A}$ is very easy, since we have the explicit formulla for it.

To get the kick size, we have to integrate one of the LK peaks. This is easiest done by considering the evolution of the $\delta e \equiv 1-e$ variable by dotting $\vec{e}$ into $\frac{\mathrm{d}\vec{e}}{\mathrm{d}t}$, such that

$$2e\frac{\mathrm{d}e}{\mathrm{d}t} = \frac{\mathrm{d}(\vec{e}\cdot\vec{e})}{\mathrm{d}t} = 2\vec{e}\cdot\frac{\mathrm{d}\vec{e}}{\mathrm{d}t}, \tag{30}$$

$$= -\frac{15}{2t_{LK}}(\mathbf{e}\cdot\hat{n}_2)(\hat{n}_2\cdot(\mathbf{e}\times\mathbf{j})), \tag{31}$$

$$\lesssim \pm\frac{15}{2t_{LK}}e^2\sqrt{1-e^2}, \tag{32}$$

$$\frac{\mathrm{d}e}{\mathrm{d}t} \sim -\frac{15}{4t_{LK}}e\sqrt{1-e^2}, \tag{33}$$

$$\frac{\mathrm{d}(\delta e)}{\mathrm{d}t} \sim \frac{15}{4t_{LK}}\sqrt{2\delta e}, \tag{34}$$

$$\delta e(t) \sim \left(\frac{15t}{4\sqrt{2}t_{LK}}\right)^2. \tag{35}$$

The finding of a power law/quadratic seems in accordance w/ my simulations, though I have to plot

$\delta e - \delta e_{\min}$. Then, we can simply integrate

$$\Delta\theta_{sb} \sim \oint_{LK} \Omega_{SL}\,\mathrm{d}t, \tag{36}$$
$$\sim \frac{\epsilon_{SL}}{2}\left(\frac{a_0}{a}\right)^{5/2} \oint_{LK} \frac{1}{\delta e}\,\mathrm{d}t, \tag{37}$$
$$\sim \frac{\epsilon_{SL}}{2}\left(\frac{a_0}{a}\right)^{5/2} \oint_{LK} \frac{1}{\delta e_{\min} + \left(\frac{15}{4\sqrt{2}t_{LK}}\right)^2 t^2}\,\mathrm{d}t, \tag{38}$$
$$\sim \frac{\epsilon_{SL}}{2\delta e_{\min}}\left(\frac{a_0}{a}\right)^{5/2} \pi \frac{4t_{LK}\sqrt{2\delta e_{\min}}}{15}, \tag{39}$$
$$\sim \frac{\epsilon_{SL}}{\sqrt{2\delta e_{\min}}}\frac{a_0}{a}\pi\frac{4}{15}. \tag{40}$$

In the last few steps, we've just taken the bounds of integration to be $t \in [-\infty, \infty]$ for simplicity (they contribute negligibly), and used $t_{LK} = (a/a_0)^{3/2}$ since $t_{LK,0} = 0$.

If we now explicitly write down the criteria where $\mathscr{A} \sim 1$ and $\Delta\theta_{sb} \sim 1$ in the (a, e) plane, then we obtain

$$a_{c,\theta} \sim \frac{\epsilon_{SL}}{\sqrt{2\delta e_{\min}}}\frac{4\pi}{15}, \tag{41}$$
$$a_{c,\mathscr{A}} \sim \left[\epsilon_{SL}\frac{8/3}{\sqrt{1 - e_{\min}^2}\left(1 + 4e_{\min}^2\right)\left|\sin 2I\right|}\right]^{1/4}. \tag{42}$$

The key difference between the two is that kicks occur at $e_{\max}$ or $\delta e_{\min}$, while the adiabaticity parameter is moreso evaluated at $e_{\min}$.

Update: This cannot be the correct mechanism since it would generate symmetric scatter of θ_{sl} about $\theta_{sb,0}$, which is not the case (see Fig 19 of LL18). Instead, it must really be how quickly the axis of precession of $\frac{\mathrm{d}\hat{\mathbf{S}}}{\mathrm{d}t}$ moves compared to the precession frequency, or indeed $\left|\Omega_{eff}\right|$ as compared to $\frac{\mathrm{d}\hat{\mathbf{\Omega}}_{eff}}{\mathrm{d}t}$.

1.4 More analysis on LL18's proposal

Note that, since θ_{sb} during the LK oscillations will receive a sequence of kicks, it randomizes the ordering a bit, so exact conservation of $\theta_{sb,i}$ to $\theta_{sl,f}$ is not maintained (i.e. the ordering can change somewhat).

But ultimately, it must boil down indeed to comparison of the change in precession axis vs the precession frequency. One of the key difficulties in this conclusion in LL18 is neglect of nutation in Equations 64 and 65. However, in the transadiabatic regime, the LK cycles are of small amplitude ($1 - e$ typically at least $\lesssim 0.1$, often $\lesssim 0.01$ throughout the cycles) and are fast, and as a result I is to good approximation constant and nutation can likely be neglected; at worst an average value of $\mathbf{L}$ can be used. The final spread in $\theta_{sl,f}$ probably comes from the spread in $\theta_{sb,i}$ upon exiting the nonadiabatic regime, due to the kicks during the LK cycles. **NB:** Another way to argue that the fast nutation can be ignored is if $\Delta I \ll \theta_{\text{eff,S}}$, since then the spin vector just precesses around a fuzzy vector, which isn't a huge deal. If the precession frequencies are equal, it's possible to hit an SHO-like resonance, which should probably be dealt with TODO.

Let's suppose this is the case for the time being, where e, I, and a are all approximately slowly varying going into the transadiabatic regime. Then let's go to the co-rotating frame with $\mathbf{L}$ (fix this in the $\hat{x}, \hat{z}$ plane) and look at the evolution of the components of $\mathbf{\Omega}_{\rm eff}$:

$$\mathbf{\Omega}_{\rm eff} = \Omega_{SL}(\sin I \hat{x} + \cos I \hat{z}) + \Omega_{pl}\hat{z}, \tag{43}$$

$$\hat{\mathbf{\Omega}}_{\rm eff} \cdot \hat{z} = \frac{\Omega_{SL}\cos I + \Omega_{pl}}{\sqrt{\Omega_{SL}^2 \sin^2 I + \left(\Omega_{SL}\cos I + \Omega_{pl}\right)^2}}. \tag{44}$$

Then, we just have to compare $\frac{\mathrm{d}\arccos \hat{\Omega}_{\rm eff,z}}{\mathrm{d}t}$ to $\Omega_{\rm eff}$ the magnitude, and this tells us whether $\hat{S}$ can track $\mathbf{\Omega}_{\rm eff}$ as it moves. This tracks the polar angle, and the z component doesn't have a singularity during the evolution and is preferable (compared to the x component). If all quantities are slowly varying (at roughly constant speeds), the characteristic speed at which the polar angle varies occurs when it is $\sim 90^\circ$, or when $\Omega_{\rm eff,z} \approx 0$, so we can simply the expression a bit

$$\frac{\mathrm{d}\arccos \hat{\Omega}_{\rm eff,z}}{\mathrm{d}t} \lesssim \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\Omega_{SL}\cos I + \Omega_{pl}}{\Omega_{SL}\sin I}\right) \sim \frac{1}{\sin I}\frac{\mathrm{d}(\mathscr{A}^{-1})}{\mathrm{d}t} \tag{45}$$

In summary, the picture is as follows:

- Starting from some initial $\theta_{sb,0}$, there are some random kicks (which cancel slightly better than a random walk, i.e. the variance does not seem to grow), we exit the nonadiabatic regime with some random value $\in \theta_{sb,0} \pm \Delta\theta_{sb}$.
- Under the influences of ϵ_{GW} on $e_{\max}$ and ϵ_{GR} on $e_{\min}$, the trajectory flows towards a single point in (a, e) space. Note that I should be fixed by approximate conservation of the Kozai constant, since the GW effect is much weaker than the GR effect, and the GR effect preserves the Kozai constant.
- If the system hasn't exited the Kozai regime or merged at this point, it will evolve with small LK oscillations about GR decay of e and a, coupled to convergence in I. As GR acts, this fuzz gets smaller and smaller amplitude until GR breaks the LK resonance.

  During this fuzzy phase, so long as $\Omega_{\rm eff} \gtrsim \frac{\mathrm{d}(\mathscr{A})^{-1}}{\mathrm{d}t}/\sin I$, then θ_{sb} gets sent to $\theta_{sl,f}$. The fuzz timescale is so short that it can be averaged over (the spin can't see it), so we should just have to consider the GR decay timescale when making this comparison.
- Regardless of whether the transadiabatic phase conserves the adiabatic invariant $\theta_{\rm eff}$, the final value is conserved once LK entirely disappears and it's just slow GR decay (which will evolve $\Omega_{\rm eff}$ slightly, but more obviously slowly).

1.5 Timescales for My Picture

NB: we are in the circulating regime of the L-K mechanism!

We now make some comments on the dynamics in each of these regimes:

- During the initial pure-LK phase, there are small perturbations to θ_{sb} as we derived above.

  But furthermore, we can estimate the characteristic number of LK cycles by observing that the decay in the range of ω oscillation is what drives $e_{\min}$ to increase over time. We can integrate one Kozai cycle

$$\Delta\omega = \oint_{LK} \frac{3Gnm_{12}e}{c^2a\left(1-e^2\right)}\,\mathrm{d}t, \tag{46}$$

$$\approx \frac{Gnm_{12}e}{2c^2a}\int\limits_{-\infty}^{\infty} \frac{1}{\delta e_{\min} + \left(\frac{15t}{4\sqrt{2}t_{LK}}^2\right)}\,\mathrm{d}t, \tag{47}$$

$$\approx \frac{3Gnm_{12}4\pi t_{LK}}{c^2a\sqrt{2\delta e_{\min}}}. \tag{48}$$

It's likely we can replace $\sqrt{2\delta e_{\min}} \to \sqrt{1-e_{\max}^2}$.

We should be able to determine the number of Kozai cycles before coalescence by computing $\frac{\partial H}{\partial \omega}$ at $e_{\max}$, which I haven't done.

If we ignore GW effects, the final state for this phase is where $e \approx e_{\max}, I \approx I_{\min}$. There are small corrections due to (i) GW decay near the high-e phases; we can estimate the former, since we also know the number of high-e cycles, but it may not be very important.

- During the fuzz phase, let's assert that ω, I make small amplitude oscillations about mean values that evolve slowly under GW emission (which also affects a): note that I is affected because GW emission is approximately adiabatic compared to the LK timescale. What sets the frequency and amplitude of these oscillations?

  Gave up: some online references seem to suggest that oscillations get to order $\sim t_{LK}/6$ as we have defined it[1], while in our simulations, each LK cycle actually is much longer than this initially, so that gives us a decent idea of the timescale of the "fuzz." **Edit:** It's probably even faster, since this is the librating timescale, so let's just assume the fuzz is very short scale. **Edit 2:** It is bound from below by Ω_{GR}, since that's one component of $\dot{\omega}$, and it must be at least as fast.

  Looking at the phase portrait, it's more clear that the GR precession will eventually just send entire trajectory to be roughly constant at $e_{\max} \to 1$. It's not clear that the amplitude of these oscillations ever saturates, but it's obvious that they are small and continue to decrease. One way to see that this is the case is to consider the $H(\omega, x)$ surface, where we drop constant of proportionality

$$H \propto \left(2+3e^2\right)\left(3\cos^2 I - 1\right) + 15e^2\sin^2 I\cos 2\omega, \tag{49}$$

  and I is implicitly defined by conservation of the Kozai constant $K = \sqrt{1-e^2}\cos I$. We can see that along the separatrix, $H = -2$, if we give a kick at the location of maximum eccentricity ($\sin^2 I = 2/5, e = 1, \omega = \pi/2$), the change in H is quadratic like

$$\delta H = \frac{1}{2}\frac{\partial^2 H}{\partial \omega^2}(\delta\omega)^2, \tag{50}$$

  where $\delta\omega \sim \left(1-e_{\max}^2\right)^{-1/2}$ was an earlier result we showed. The sign of this term is negative, so H is being driven towards oscillating at large e with small amplitude.

  In any case, the fuzz decreases in amplitude over time and oscillates faster and faster, probably $\ll t_{LK}$ (indeed so, according to my plots).

[1] https://arxiv.org/pdf/1504.05957.pdf

- As we evolve through the fuzz, we want to understand whether $\theta_{\mathrm{eff,S}}$ evolves adiabatically. We need to evaluate the precession frequency and the rate of change of the precession axis, but for this we need expressions for $\dot{e}, \dot{I}, \dot{a}$ through the fuzz. Based on the final observation that $\dot{\omega}_{GR}$ doesn't affect the mean eccentricity, we can assert that $\dot{e} = \dot{e}_{GW}$, $\dot{a} = \dot{a}_{GW}$, while I is constrained implicitly by conservation of the Kozai constant (so long as Kozai still is active). Thus, to order of magnitude, $\dot{\hat{\Omega}}_{\mathrm{eff}} \sim \Omega_{GW}$ while $\Omega_{\mathrm{eff}} \sim \Omega_{GR}$, and since $\Omega_{GW} \propto 1/\left(a^4 x^{7/2}\right)$ while $\Omega_{GR} \propto 1/\left(a^{5/2} x\right)$, it's clear that for sufficiently large eccentricities exiting the fuzz regime that $\hat{S}$ will not keep up with $\hat{\Omega}_{\mathrm{eff,S}}$.

  If so, what is the predicted $\theta_{\mathrm{sl,f}}$? Well, suppose that $\hat{L}$ ends up on the ring with uniform I (probably. . .), and take the limit where $\hat{S}$ is not able to respond at all, then $\theta_{\mathrm{sl,f}} \in \left[I - \theta_{\mathrm{sb,i}}, 2\pi - I - \theta_{\mathrm{sb,i}}\right]$ and is roughly centered on $\pi - \theta_{\mathrm{sb,i}}$.

  NB: Above, we said $\Omega_{\mathrm{eff}} \sim \Omega_{GR}$ based on saying that $\hat{L}$ precesses around $\hat{L}_{\mathrm{out}}$ with $\dot{\omega}_{GR}$, but of course, if evolution is sufficiently abrupt, we should really use $\Omega_{\mathrm{eff}} \sim \Omega_{\mathrm{pl}}$, and if evolution is abrupt this is $\ll \Omega_{GR}$, further contributing to making the nonadiabatic criterion easy to satisfy.

- Note that there is one more way that this picture can break down, as we saw examples of in Bin's paper: we can get trapped in the LK resonance such that ω librates instead of circulating. This is not in general easy to do, since we start with $e \neq 0, \omega = 0$. Furthermore, to linear order, $\oint \frac{\partial H}{\partial \omega} \frac{\mathrm{d}\omega}{\mathrm{d}t}\,\mathrm{d}t = 0$ along the separatrix (it cancels during the increasing e and decreasing e phases). But we can imagine that if ω_{GR} is so strong, then $\dot{\omega}$ will drive the Kozai cycle inside the resonance during just the increasing e phase alone, and takes a very different route back to low e such that it is captured. That this is a nonlinear effect in $\delta\omega_{GR}$ might be important, since otherwise the resonance capture dynamics would only depend on the initial condition: the separatrix would open a gap like in the CS problem and for arbitrarily weak ω_{GR} we could still experience separatrix capture, which is obviously not the case?

  The advantage of invoking this mechanism is twofold: (i) if we look at Fig. 19 of LL18, it's clear that the distribution of $\theta_{sl,f}$ is roughly symmetric for a stronger companion (faster LK cycles), but becomes markedly asymmetric for a weaker companion (LK is weak). The violation of adiabiticity proposed above is generally expected to generate a $\theta_{sl,f}$ distribution symmetric about its mean. But capturing $\hat{L}$ into the $\omega = \pi/2$ resonance means $\hat{S}$ precesses *towards* it as it becomes dominant, meaning that $\theta_{sl,f} \lesssim 90^\circ$ is enforced. (ii), the above mechanism does not depend on the properties of the perturber or of the Kozai timescale, so there should be no change in distribution of $\theta_{sl,f}$ as a function of a_{out}. This resonance capture mechanism provides a way for the outcome to be sensitive to the perturber properties.

A j Equations

We define vectors

$$\mathbf{j} = \sqrt{1-e^2}\,\hat{n}, \tag{51}$$

$$\mathbf{e} = e\hat{u}. \tag{52}$$

Here, $\mathbf{j}$ is the dimensionless angular momentum vector and $\mathbf{e}$ is the eccentricity vector; see LML15 for precise definitions. Note that $\mathbf{j} \cdot \mathbf{e} = 0$, $j^2 + e^2 = 1$. Then, the EOM for the inner and outer vectors satisfy to quadrupolar order

$$\frac{\mathrm{d}\mathbf{j}}{\mathrm{d}t} = \frac{3}{4t_{LK}}\left[(\mathbf{j}\cdot\hat{n}_2)(\mathbf{j}\times\hat{n}_2) - 5(\mathbf{e}\cdot\hat{n}_2)(\mathbf{e}\times\hat{n}_2)\right], \tag{53}$$

$$\frac{\mathrm{d}\mathbf{e}}{\mathrm{d}t} = \frac{3}{4t_{LK}}\left[(\mathbf{j}\cdot\hat{n}_2)(\mathbf{e}\times\hat{n}_2) + 2\mathbf{j}\times\mathbf{e} - 5(\mathbf{e}\cdot\hat{n}_2)(\mathbf{j}\times\hat{n}_2)\right]. \tag{54}$$

Let's assume for the time being that $L_1 \ll L_2$, so the system is sufficiently hierarchical that $\mathbf{j}_2$, $\mathbf{e}_2$ are constants. Note for reference that

$$t_{LK} \equiv \frac{L_1}{\mu_1\Phi_0} = \frac{1}{n_1}\left(\frac{m_1+m_2}{m_3}\right)\left(\frac{a_2}{a}\right)^3\left(1-e_2^2\right)^{3/2}. \tag{55}$$

Here, $n_1 \equiv \sqrt{G(m_1+m_2)/a^3}$. Thus, $1/t_{LK} \propto a^{3/2}$.

GW radiation (Peters 1964) cause decays of $\mathbf{L}$ and $\mathbf{e}$ as

$$\left.\frac{\mathrm{d}\mathbf{L}}{\mathrm{d}t}\right|_{GW} = -\frac{32}{5}\frac{G^{7/2}}{c^5}\frac{\mu^2 m_{12}^{5/2}}{a^{7/2}}\frac{1+7e^2/8}{\left(1-e^2\right)^2}\hat{L}, \tag{56}$$

$$\left.\frac{\mathrm{d}\mathbf{e}}{\mathrm{d}t}\right|_{GW} = -\frac{304}{15}\frac{G^3}{c^5}\frac{\mu m_{12}^2}{a^4\left(1-e^2\right)^{5/2}}\left(1+\frac{121}{304}e^2\right)\mathbf{e}, \tag{57}$$

$$\left.\frac{\dot{a}}{a}\right|_{GW} = -\frac{64}{5}\frac{G^3\mu m_{12}^2}{c^5a^4}\frac{1}{\left(1-e^2\right)^{7/2}}\left(1+\frac{73}{24}e^2+\frac{37}{96}e^4\right). \tag{58}$$

Here, $m_{12} \equiv m_1 + m_2$. The last GR effect is precession of $\vec{e}$, which acts as

$$\left.\frac{\mathrm{d}\mathbf{e}}{\mathrm{d}t}\right|_{GR} = \frac{3Gnm_{12}}{c^2a\left(1-e^2\right)}. \tag{59}$$

Given this system (from LML15 + LL18), we can then add the spin-orbit coupling term (from de Sitter precession), which is given in LL18 to be

$$\frac{\mathrm{d}\hat{S}}{\mathrm{d}t} = \Omega_{SL}\hat{L}\times\hat{S}, \tag{60}$$

$$\Omega_{SL} \equiv \frac{3Gn\left(m_2+\mu/3\right)}{2c^2a\left(1-e^2\right)}. \tag{61}$$

Note that μ is the reduced mass of the inner binary. We can drop the back-reaction term since $S \ll L$. Thus, $\Omega_{SL} \propto a^{-5/2}$.

What is observed is that, as this system is evolved forward in time and GR coalesces the inner binary, $\theta_{sl} \equiv \arccos\left(\hat{S}\cdot\hat{L}\right)$ goes to 90° consistently. The relevant figure is Fig. 19 of LL18, which shows that for a close-in, low-eccentricity perturber ($\bar{a}_{\mathrm{out,eff}} \propto a_{out}$), the focusing is significantly stronger. Note that initially, $I \equiv \arccos\left(\hat{L}\cdot\hat{L}_2\right) \approx 90^\circ$ while $\theta_{sl} \approx 0$.

Next, when accounting for GR, we should let a evolve as above. Note that since $\mathbf{j}$ and $\vec{e}$ are our dynamical variables, we should use $\mathbf{j} \equiv \sqrt{1-e^2}\hat{L} = \sqrt{1-e^2}\frac{\mathbf{L}}{\mu\sqrt{Gm_{12}a(1-e^2)}}$ and rewrite

$$\left.\frac{\mathrm{d}\mathbf{j}}{\mathrm{d}t}\right|_{GW} = \frac{1}{\mu\sqrt{GMa}}\left.\frac{\mathrm{d}\mathbf{L}}{\mathrm{d}t}\right|_{GW} - \frac{\mathbf{j}}{2a}\left.\frac{\mathrm{d}a}{\mathrm{d}t}\right|_{GW}. \tag{62}$$

To double check, we should verify that $\left.\frac{\mathrm{d}(j^2+e^2)}{\mathrm{d}t}\right|_{GW} = 0$, which can be verified as (Let's set $G = M = \mu = a = c = 1$ for convenience)

$$\frac{1}{2}\frac{\mathrm{d}(j^2+e^2)}{\mathrm{d}t} = \mathbf{j}\cdot\frac{\mathrm{d}\mathbf{j}}{\mathrm{d}t} + \mathbf{e}\cdot\frac{\mathrm{d}\mathbf{e}}{\mathrm{d}t}, \tag{63}$$

$$= \mathbf{j}\cdot\left[\left(-\frac{32}{5}\frac{1+7e^2/8}{\left(1-e^2\right)^2}\right)\hat{L} - \frac{\mathbf{j}}{2}\left(-\frac{64}{5}\frac{1+73e^2/24+37e^4/96}{\left(1-e^2\right)^{7/2}}\right)\right] + \mathbf{e}\cdot\left(-\frac{304}{15}\frac{1+121e^2/304}{\left(1-e^2\right)^{5/2}}\right)\mathbf{e}, \tag{64}$$

$$= \left(-\frac{32}{5}\frac{1+7e^2/8}{\left(1-e^2\right)^{3/2}}\right) + \left(\frac{32}{5}\frac{1+73e^2/24+37e^4/96}{\left(1-e^2\right)^{5/2}}\right) + e^2\left(-\frac{304}{15}\frac{1+121e^2/304}{\left(1-e^2\right)^{5/2}}\right), \tag{65}$$

$$= \frac{1}{15\left(1-e^2\right)^{5/2}}\left[-96\left(1-e^2\right)\left(1+\frac{7e^2}{8}\right) + 96\left(1+\frac{73e^2}{24}+\frac{37e^4}{96}\right) - 304e^2\left(1+\frac{121e^2}{304}\right)\right]. \tag{66}$$

This can be verified to vanish upon term-by-term examination indeed.