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 $\mathbf{j} = \mathbf{L}/\left(\mu\sqrt{Gm_{12}a}\right)$. 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 $\mu\sqrt{Gm_{12}} = 1$, such that now

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

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

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

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

$$
\begin{aligned}
2e\frac{\mathrm{d}e}{\mathrm{d}t} &= \frac{\mathrm{d}(\vec{e}\cdot\vec{e})}{\mathrm{d}t} = 2\vec{e}\cdot\frac{\mathrm{d}\vec{e}}{\mathrm{d}t}, && (30)\\
&= -\frac{15}{2t_{LK}}(\mathbf{e}\cdot\hat{n}_2)(\hat{n}_2\cdot(\mathbf{e}\times\mathbf{j})), && (31)\\
&\lesssim \pm\frac{15}{2t_{LK}}e^2\sqrt{1-e^2}, && (32)\\
\frac{\mathrm{d}e}{\mathrm{d}t} &\sim -\frac{15}{4t_{LK}}e\sqrt{1-e^2}, && (33)\\
\frac{\mathrm{d}(\delta e)}{\mathrm{d}t} &\sim \frac{15}{4t_{LK}}\sqrt{2\delta e}, && (34)\\
\delta e(t) &\sim \left(\frac{15t}{4\sqrt{2}t_{LK}}\right)^2. && (35)
\end{aligned}
$$

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

$$a_{c,\theta} \sim \frac{\epsilon_{SL}}{\sqrt{2\delta e_{\min}}}\frac{4\pi}{15}, \tag{41}$$

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

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 $\Delta I \ll \theta_{\mathrm{eff,S}}$, 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 $\mathbf{\Omega}_{\rm eff}$:

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

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

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

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

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

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

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 $\theta_{\rm eff,S}$ 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, $\dot{\hat{\Omega}}_{\rm eff} \sim \Omega_{GW}$ while $\Omega_{\rm eff} \sim \Omega_{GR}$, 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 $\hat{\Omega}_{\rm eff,S}$.

  If so, what is the predicted $\theta_{\rm sl,f}$? 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 $\theta_{\rm sl,f} \in \left[I - \theta_{\rm sb,i}, 2\pi - I - \theta_{\rm sb,i}\right]$ and is roughly centered on $\pi - \theta_{\rm sb,i}$.

  NB: Above, we said $\Omega_{\rm eff} \sim \Omega_{GR}$ based on saying that $\hat{L}$ precesses around $\hat{L}_{\rm out}$ with $\dot{\omega}_{GR}$, but of course, if evolution is sufficiently abrupt, we should really use $\Omega_{\rm eff} \sim \Omega_{\rm pl}$, 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 $a_{\rm out}$. This resonance capture mechanism provides a way for the outcome to be sensitive to the perturber properties.

Edit: Looks like I got Ω, ω confused, and most of the above is either wrong or not new.

2 Fresh Start

NB: I think these orbital elements Kozai actually give much slower inspirals. It could be because my atol/rtol params were too loose when I was doing the vector simulations, so we should prefer the `4sims` line of results, which qualitatively seem to agree with Bin's.

2.1 Useful Kozai Results

I have a bunch of formulas that I need to write down before I forget them, so I'll do that here. We have begun analyzing the EOM (at quadrupolar order) in Keplerian orbital elements, so I'll reproduce them here

$$\frac{\mathrm{d}a}{\mathrm{d}t} = -\frac{64}{5}\frac{a}{t_{GW,0}}\frac{1}{\left(1-e^2\right)^{7/2}}\left(1+\frac{73}{24}e^2+\frac{37}{96}e^4\right), \tag{51}$$

$$\frac{\mathrm{d}e}{\mathrm{d}t} = \frac{15}{8t_{LK}}e\sqrt{1-e^2}\sin 2\omega\sin^2 I - \frac{304}{15}\frac{e}{t_{GW,0}}\frac{1}{\left(1-e^2\right)^{5/2}}\left(1+\frac{121}{304}e^2\right), \tag{52}$$

$$\frac{\mathrm{d}\Omega}{\mathrm{d}t} = \frac{3}{4t_{LK}}\frac{\cos I\left(5e^2\cos^2\omega - 4e^2 - 1\right)}{\sqrt{1-e^2}}, \tag{53}$$

$$\frac{\mathrm{d}I}{\mathrm{d}t} = -\frac{15}{16t_{LK}}\frac{e^2\sin 2\omega\sin 2I}{\sqrt{1-e^2}}, \tag{54}$$

$$\frac{\mathrm{d}\omega}{\mathrm{d}t} = \frac{3}{4t_{LK}}\frac{2\left(1-e^2\right)+5\sin^2\omega(e^2-\sin^2 I)}{\sqrt{1-e^2}} + \frac{\Omega_{GR,0}}{1-e^2}, \tag{55}$$

$$\frac{\mathrm{d}\hat{\mathbf{S}}}{\mathrm{d}t} = \frac{\Omega_{SL,0}}{1-e^2}\hat{\mathbf{L}}\times\hat{\mathbf{S}}. \tag{56}$$

Here, we have defined

$$t_{LK}^{-1} = n\left(\frac{m_3}{m_{12}}\right)\left(\frac{a}{\bar{a}_3}\right)^3, \tag{57}$$

$$t_{GW,0}^{-1} = \frac{G^3\mu m_{12}^2}{c^5a^4}, \tag{58}$$

$$\Omega_{GR,0} = \frac{3Gnm_{12}}{c^2a}, \tag{59}$$

$$\Omega_{SL,0} = \frac{3Gn\left(m_2+\mu/3\right)}{2c^2a}. \tag{60}$$

and $n = \sqrt{Gm_{12}/a^3}$ is the mean motion of the inner binary. We define/recall the following:

- $K = \sqrt{1-e^2}\cos I$ is conserved, and we will sometimes write $x = 1-e^2$.
- Kozai eccentricity excursions occur at $\omega = \pi/2, 3\pi/2$.
- If we ever need this, Natalia's paper gives "closed" forms for the eccentricity evolution

$$x = x_0 + (x_1 - x_0)\operatorname{cn}^2(\theta, k^2), \tag{61}$$

$$\theta = \frac{K}{\pi}\left(n_e t + \pi\right), \tag{62}$$

$$n_e = \frac{6\pi\sqrt{6}}{8Kt_{LK}}\sqrt{x_2 - x_1}, \tag{63}$$

$$k^2 = \frac{x_0 - x_1}{x_2 - x_1}. \tag{64}$$

Here, x_0 and x_1 are the maximum/minimum x respectively (corresponding to min/max eccentricity), and x_2 is the other root to the quadratic (x_1 is one of them). K is not the Kozai constant but is approximately $\pi/(2\mathrm{agm}(1,\sqrt{1-k^2}))$, the *arithmetic-geometric mean*.

$$x^2 - \frac{1}{3}(5+5K-2x_0)x + \frac{5K}{3} = 0. \tag{65}$$

Note that this implies $x_1 + x_2 = \frac{5+5K-2x_0}{3}$.

- In terms of this x parameter, the LK components of the EOM take on particularly simple form

$$\dot{\Omega} = \frac{3\sqrt{h}}{4t_{LK}}\left(1 - 2\frac{x_0 - h}{x - h}\right), \tag{66}$$
$$\dot{I} = \frac{\dot{x}\cos I}{2x\sin I}. \tag{67}$$

It is not so hard to solve for the Kozai resonance location in the absence of GW radiation; we know this occurs at $\omega = \pi/2$, which forces $\dot{e} = \dot{I} = 0$, then we set $\dot{\omega} = 0$ and find

$$\frac{\mathrm{d}\omega}{\mathrm{d}t} = \frac{3}{4t_{LK}}\frac{2(1-e^2) + 5(e^2 - \sin^2 I)}{\sqrt{1-e^2}} + \Omega_{GR}, \tag{68}$$
$$0 = \frac{\Omega_{GR}\sqrt{1-e^2}4t_{LK}}{3} + 2\left(1-e^2\right) - 5\left(1 - \cos^2 I - e^2\right), \tag{69}$$
$$5\cos^2 I = 3\left(1-e^2\right) + \mathcal{O}\left(\Omega_{GR}\right). \tag{70}$$

Then, given some K, which is conserved even with GR precession, we know (I, e) the Kozai resonance. To obtain the $\mathcal{O}(\Omega_{GR})$ correction, we have to solve a quadratic, which yields

$$\sqrt{1-e^2} = \frac{5\cos^2 I}{6}\left(2 + \sqrt{1 + \frac{16\Omega_{GR}t_{LK}}{25\cos^4 I}}\right). \tag{71}$$

If Ω_{GR} is strong, the equilibrium condition drives $\cos^2 I \to 0$, and simplifying in this limit we get the familiar condition $\Omega_{GR}t_{LK} \le 9/4$ for the Kozai resonance itself to exist (of course, the separatrix about which trajectories librate will begin shrinking much earlier).

We can use this to understand what $\mathscr{A}$ looks like when Kozai disappears. Assuming $m_1 = m_2$, we can find $\Omega_{SL} = \Omega_{GR}\frac{7}{24}$, and so when Kozai dies we obtain constraints

$$\Omega_{SL}t_{LK} = \frac{21}{32}, \tag{72}$$
$$\mathscr{A} \simeq \left|\frac{\Omega_{SL}}{\dot{\Omega}}\right| = \frac{7}{8}\frac{\sqrt{1-e^2}}{\cos I\left(4e^2+1\right)}. \tag{73}$$

If we plug in the values near the Kozai equilibrium when it disappears, we find rough scaling

$$\mathscr{A} \simeq \frac{7}{8\left(1+4e^2\right)\cos I}. \tag{74}$$

Thus, indeed, typically $\mathscr{A} \sim 1$ when Kozai dies.

Finally, it bears noting that

$$\mathscr{A} = \mathscr{A}_0\frac{1}{\left(1+4e^2\right)\sqrt{1-e^2}\left|\sin 2I\right|}, \tag{75}$$
$$= \mathscr{A}_0\frac{1}{2\left(1+4e^2\right)K\sin I}. \tag{76}$$

Thus, over the course of a Kozai cycle, where $\sin I \in [\sqrt{2/5}, 1]$, and $e \in [0, 1]$, the adiabaticity does not actually change very much, unless $\mathscr{A}_0 \propto a^{-4}$ changes significantly due to $\dot{a}_{GW}$.

2.2 Hamiltonian Approach 1: Natalia Style

We go to the frame where $\hat{\mathbf{L}}$ is stationary. The rotation vector is the same as in SL15, and we obtain Hamiltonian ($\hat{\mathbf{L}}_o$ is the outer angular momentum, is constant in nonrotating frame)

$$H = \Omega_{SL}\cos\theta - \mathbf{R}\cdot\hat{\mathbf{S}}, \tag{77}$$

$$\mathbf{R} \equiv \left(\dot{\Omega}\hat{\mathbf{L}}_o + \dot{I}\left(\frac{\hat{\mathbf{L}}_o\times\hat{\mathbf{L}}}{\sin I}\right)\right). \tag{78}$$

If we break down all the vectors into component form, such that $\hat{\mathbf{L}} = \hat{\mathbf{z}}$, $\hat{\mathbf{L}}_o = -\sin I\hat{\mathbf{x}} + \cos I\hat{\mathbf{z}}$, then we obtain

$$H = \Omega_{SL}\cos\theta - \dot{\Omega}\left(-\sin I\sin\theta\cos\phi + \cos I\cos\theta\right) + \dot{I}\sin\theta\sin\phi. \tag{79}$$

Note ICs $\theta = 0$, $\phi = 0$, $I = 90°$. The EOM are

$$\dot{\phi} = \frac{\partial H}{\partial\cos\theta} = \Omega_{SL} - \dot{\Omega}\cos I - \cot\theta\left(\dot{\Omega}\sin I\cos\phi - \dot{I}\sin\phi\right), \tag{80}$$

$$\frac{\mathrm{d}(\cos\theta)}{\mathrm{d}t} = -\frac{\partial H}{\partial\phi} = -\dot{\Omega}\sin I\sin\theta\sin\phi + \dot{I}\sin\theta\cos\phi, \tag{81}$$

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = -\frac{1}{\sin\theta}\frac{\mathrm{d}(\cos\theta)}{\mathrm{d}t} = \dot{\Omega}\sin I\sin\phi - \dot{I}\cos\phi. \tag{82}$$

If we assume $\Omega_{SL} \ll \dot{\Omega}, \dot{I}$ initially, even during LK peaks (which is true by our experience), then we can imagine breaking down the trajectory of $\hat{\mathbf{S}}$ into a zeroth order precession about $\mathbf{R}$ (which is very complicated, since $\mathbf{R}$ is both moving and changing in magnitude) and a leading order perturbation due to Ω_{SL}. The perturbation Hamiltonian is then given

$$H^{(1)} = \Omega_{SL}(t)[\cos\theta](t). \tag{83}$$

If we're brave like Natalia, we would expand $\cos\theta(t)$ in Fourier components, and $\Omega_{SL}(t)$ in Fourier components, but there is clearly no chance for a resonance here since there is no ϕ dependence, so the level curves of this Hamiltonian are azimuthally symmetric and there can be no resonance.

Conversely, if we're in the other regime $\Omega_{SL} \gg \dot{\Omega}, \dot{I}$, we must be in the regime where Kozai cycles have died out, which implies $\dot{I} = 0$. Here, the Hamiltonian is much more similar to Natalia's problem. Let's consider that $\theta = \theta_0$ and $\phi = \Omega_{SL}t$, then the perturbing Hamiltonian is

$$H^{(1)} = \dot{\Omega}\left(-\sin I\sin\theta_0\cos\phi + \cos I\cos\theta_0\right). \tag{84}$$

Again, there is no resonance condition since there is only one ϕ dependent term, and we really need two so we can get a form $\cos\left(\phi - Mt\right)$ like Natalia's problem. Thus, there are no resonances to investigate here.

We can identify the key reasons that we don't have a similar problem:

- LK is not a perturbation for us (compared to $\hat{\mathbf{S}}\cdot\hat{\mathbf{L}}$ dynamics), it is significantly dominant. This corresponds to the $\mathscr{A} \ll 1$ regime of SL15. They obtain a neat bifurcation due to separatrix crossing, which is not observed in our LK simulations, so this cannot in spirit be a similar mechanism.

- SL15 focuses on adiabatically changing $\mathscr{A}$ and seeing how it encounters resonances. In our problem, nothing nontrivial can occur if $\mathscr{A}$ changes slowly.

- Our Hamiltonian takes on form $H = (\mathbf{\Omega}_{SL} - \mathbf{R})\cdot\hat{\mathbf{S}}$. This will never have any resonances since it's perfectly linear; anything that looks nonlinear is a pure consequence of coordinates (e.g. multiplication of θ and ϕ terms).

2.3 Hamiltonian 2: Rotating Style

Why is the previous Hamiltonian hard to use? Well, since there are no resonances, and H is linear in $\hat{\mathbf{S}}$, it makes much more sense to just analyze the EOM. As such, it's better to just find the right set of rotations such that we have a convenient coordinate. LL18 proposed this $\theta_{\rm S,eff}$, and I think this is the right idea, but it can be expounded on.

Let's consider the following: we clearly want to rotate by at least $\dot{\Omega}\hat{\mathbf{L}}_o$, so that $\mathbf{L}$ does not precess any more, but it still nutates ($\dot{I}$) about fixed $\hat{\mathbf{L}}_o = \hat{\mathbf{z}}$. Let's first write down the Hamiltonian and the EOM for this case:

$$H = \Omega_{SL}\hat{\mathbf{S}}\cdot\hat{\mathbf{L}} - \dot{\Omega}\hat{\mathbf{L}}_o\cdot\hat{\mathbf{S}}, \tag{85}$$
$$= \Omega_{SL}\left(\sin I\sin\theta\cos\phi + \cos I\cos\theta\right) - \dot{\Omega}\cos\theta, \tag{86}$$
$$\frac{\mathrm{d}\cos\theta}{\mathrm{d}t} = -\frac{\partial H}{\partial\phi} = -\Omega_{SL}\sin I\sin\theta\sin\phi, \tag{87}$$
$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \Omega_{SL}\sin I\sin\phi, \tag{88}$$
$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = \frac{\partial H}{\partial\cos\theta} = -\Omega_{SL}\sin I\cot\theta\cos\phi + \Omega_{SL}\cos I - \dot{\Omega}. \tag{89}$$

This is an even stupider example than the previous section, since the desired final angle $\theta_{sl,f}$ is almost impossible to measure, and writing down $\dot{\theta}_{sl,f}$ would give huge excursions early in the evolution due to $\dot{I}$ (we've seen this plot before, in LL18). But similarly, the EOM from the previous section is also very difficult to use, since it's very unclear how $\theta = \theta_{sl}$ evolves through the Kozai phase; with the benefit of hindsight, we know that this $\dot{\theta}$ equation is just along a great circle normal to $\mathbf{R}$, but it's hard to say anything quantitative other than "this angle gets frozen by conservation of adiabatic invariant."

Instead, let's consider applying an arbitrary rotation in the $\hat{\mathbf{y}}$ direction for the time being, let's call it $\mathbf{R} = \dot{I}_o\hat{\mathbf{y}}$; taking this to equal either 0 or I equates to taking $\hat{\mathbf{L}}_o$ or $\hat{\mathbf{L}}$ as $\hat{\mathbf{z}}$ respectively. We can write down this Hamiltonian, calling $I_L = I - I_o$ (these have interpretation of $\cos I_L = \hat{\mathbf{L}}\cdot\hat{\mathbf{z}}$ and $\cos I_o = \hat{\mathbf{L}}_o\cdot\hat{\mathbf{z}}$ respectively)

$$H = \Omega_{SL}\hat{\mathbf{S}}\cdot\hat{\mathbf{L}} - \hat{\mathbf{R}}\cdot\hat{\mathbf{S}}, \tag{90}$$
$$= \Omega_{SL}\left(\sin I_L\sin\theta\cos\phi + \cos I_L\cos\theta\right) - \dot{\Omega}\left(-\sin I_o\sin\theta\cos\phi + \cos I_o\cos\theta\right) + \dot{I}_o\sin\theta\sin\phi, \tag{91}$$
$$\frac{\mathrm{d}\cos\theta}{\mathrm{d}t} = -\left(\Omega_{SL}\sin I_L + \dot{\Omega}\sin I_o\right)\left(-\sin\theta\sin\phi\right) - \dot{I}_o\sin\theta\cos\phi, \tag{92}$$
$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = -\left(\Omega_{SL}\sin I_L + \dot{\Omega}\sin I_o\right)\sin\phi + \dot{I}_o\cos\phi, \tag{93}$$
$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = -\cot\theta\left(\Omega_{SL}\sin I_L\cos\phi + \dot{\Omega}\sin I_o\cos\phi - \dot{I}_o\sin\phi\right) + \left(\Omega_{SL}\cos I_L - \dot{\Omega}\cos I_o\right). \tag{94}$$

Immediately, we can see both of the terms that made the above equations hard to work with: both $\Omega_{SL}\sin I_L$ and $\dot{\Omega}\sin I_o$ become large at some point or another, making it hard to consider the effect on the final θ. But if we choose some $\dot{I}_o$ such that $\theta(t=0) = \theta_{sb}$ while $\theta(t=\infty) = \theta_{sl}$, then the EOM is easy to analyze in both regimes.

Instead, the obvious thing to do is as follows: choose $\dot{I}_o$ such that $\Omega_{SL}\sin I_L + \dot{\Omega}\sin I_o = 0$, also satisfying $I_L + I_o = I$. This ensures that the two terms Ω_{SL} and $\dot{\Omega}$ are almost always small, while $\dot{I}_o$ is generally symmetric per cycle. On the other hand, $\dot{\phi} \approx \max\left(\Omega_{SL}, \dot{\Omega}\right)$. Note that if the initial inclination $> 90^\circ$, then $\dot{\Omega} > 0$, and we want $I_o < 0$. This ensures that the Ω-dependent terms in $\frac{\mathrm{d}\phi}{\mathrm{d}t}$

have the same sign, and the EOM become

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \dot{I}_o \cos\phi, \tag{95}$$
$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = \dot{I}_o \cot\theta \sin\phi + \left(\Omega_{SL}\cos I_L - \dot{\Omega}\cos I_o\right). \tag{96}$$

This should be almost analytically solvable, if we take a simple parameterized form for $\dot{I}_o$. In particular, let's realize that $\dot{I}_o$ will always go through negative-positive signs every Kozai cycle, so we can factor this out by considering $\dot{I}_0 = F(t)\sin(2\pi t/t_{\mathrm{LK}})$.

Let's be a bit more quantitative and write down this $\dot{I}_o$ rotation. It must satisfy (recall $I_L = I - I_o$)

$$\Omega_{SL}\sin I_L = \dot{\Omega}\sin I_0, \tag{97}$$
$$\dot{\Omega}_{SL}\sin I_L + \Omega_{SL}\cos I_L\left(\dot{I} - \dot{I}_o\right) = \ddot{\Omega}\sin I_o + \dot{\Omega}\cos I_0 \dot{I}_o, \tag{98}$$
$$\dot{\Omega}_{SL}\sin I_L - \ddot{\Omega}\sin I_o + \Omega_{SL}\cos I_L \dot{I} = \dot{I}_o\left(\dot{\Omega}\cos I_0 + \Omega_{SL}\cos I_L\right), \tag{99}$$
$$\dot{I}_o = \frac{\dot{\Omega}_{SL}\sin I_L - \ddot{\Omega}\sin I_o + \Omega_{SL}\cos I_L \dot{I}}{\dot{\Omega}\cos I_0 + \Omega_{SL}\cos I_L} \tag{100}$$

This is an absolute mess, but does it hold up?

- Well, if we are in the $\dot{\Omega} \sim \dot{I} \gg \Omega_{SL}$ limit, then $I_o \approx 0$, and things simplify to

  $$\frac{\dot{I}_o}{\dot{I}} \approx \frac{\dot{\Omega}_{SL}\sin I/\dot{I} + \Omega_{SL}\cos I}{\dot{\Omega}} \ll 1, \tag{101}$$

  since all time derivatives are the same, $\sim 1/t_{LK}$. This is correct, since the rotation should basically not be acting in this limit.

- And in the other limit, $\dot{\Omega} \sim \dot{I} \ll \Omega_{SL}$ limit, then $I_l \approx 0$ and we have

  $$\frac{\dot{I}_o}{\dot{I}} \approx \frac{-\ddot{\Omega}\sin I_0/\dot{I} + \dot{I}\Omega_{SL}\cos I}{\Omega_{SL}\cos I} \approx 1. \tag{102}$$

  This is also correct, since in this limit we should have to rotate by $\dot{I}$.

Thus, it seems like we have the correct expression, as ugly as it is. In fact, numerically, this turns out to be exactly $\boldsymbol{\Omega}_{\mathrm{eff}}$, which shouldn't be a huge surprise, since that's exactly our definition. In this case, it seems easier to just use Dong's suggestion along with Natalia's Hamiltonian and EOM.

2.4 Finding a Resonance

Recall EOM from when we rotated $\hat{L} \propto \hat{z}$ (here, $\theta = \theta_{sl}$):

$$\dot{\phi} = \frac{\partial H}{\partial \cos\theta} = \Omega_{SL} - \dot{\Omega}\cos I - \cot\theta\left(\dot{\Omega}\sin I\cos\phi - \dot{I}\sin\phi\right), \tag{103}$$
$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = -\frac{1}{\sin\theta}\frac{\mathrm{d}(\cos\theta)}{\mathrm{d}t} = \dot{\Omega}\sin I\sin\phi - \dot{I}\cos\phi. \tag{104}$$

If we directly substitute our known forms for $\dot{\Omega}$ and $\dot{I}$, we obtain

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{3\sin 2I}{8t_{LK}\sqrt{x}}\left[\left(5e^2\cos^2\omega - 4e^2 - 1\right)\sin\phi - \left(\frac{5e^2\sin 2\omega}{2}\right)\cos\phi\right], \tag{105}$$
$$= \frac{3\sin 2I}{8t_{LK}\sqrt{x}}\left(-\sin\phi\left(\frac{3e^2}{2} - 1\right) + \frac{5e^2}{2}\sin\left(\phi - 2\omega\right)\right). \tag{106}$$

If ϕ is slowly varying compared to ω, the second term just becomes an $\dot{I}$, and we indeed find the total change in θ is indeed just I. Thus, we put together $\theta_{sl,i} + I = \theta_{sl,f}$, in the peaceful limit.

But there does seem to be a resonance here, if $\dot{\phi} = 2\dot{\omega}$, or if just $\dot{\phi} = 0$, then some kicks will add rather than cancel out. It's not obvious what the final value will be, but this is a breakdown condition to the above equality.

However, this doesn't seem to be the entire picture, since θ_{sb} seems to be conserved to different extents in my $I = 90.45^\circ$ and $I = 90.5^\circ$ simulations. This again highlights the importance of choosing a good coordinate system, since θ_{sb} is very difficult to analyze in this coordinate system; we can't find the resonance that causes this.

We can see the origin of the θ_{sb} resonance as well: recall EOM

$$\frac{\mathrm{d}\theta_{sb}}{\mathrm{d}t} = \Omega_{SL}\sin I \sin\phi. \tag{107}$$

Since $\Omega_{SL}\sin I$ is periodic in t_{LK}, we can write

$$\frac{\mathrm{d}\theta_{sb}}{\mathrm{d}t} = \sin\phi \sum_{N=-\infty}^{\infty} \tilde{\Omega}_{SL}\cos\left(\frac{Nt}{t_{LK}}\right), \tag{108}$$

$$= \sum_{N=-\infty}^{\infty} \frac{\tilde{\Omega}_{SL}}{2}\left[\sin\left(\phi - \frac{Nt}{t_{LK}}\right) + \sin\left(\phi + \frac{Nt}{t_{LK}}\right)\right], \tag{109}$$

$$\frac{\mathrm{d}\phi_{sb}}{\mathrm{d}t} = \frac{\partial H}{\partial \cos\theta} = -\Omega_{SL}\sin I\cot\theta\cos\phi + \Omega_{SL}\cos I - \dot{\Omega}. \tag{110}$$

Thus, there can be a resonance if $\dot{\phi}$ matches one of the harmonics of t_{LK}. However, these resonances are indeed weaker, since they go with Ω_{SL}.

Edit: I don't think these are the resonances that I am seeing in the simulations. Rather, it is much simpler: when $\theta_{sb} \approx 90^\circ$, then $\frac{\mathrm{d}\phi_{\mathrm{sb}}}{\mathrm{d}t} \approx 0$ in between Kozai cycles. Then, when ϕ_{sb} attains substantial values, the $\frac{\mathrm{d}\theta_{sb}}{\mathrm{d}t}$ term activates even off LK peaks and θ_{sb} drifts from its initial value. When $\theta_{sb} \neq 90^\circ$, then there is a slow and steady Ω_{SL} term in $\dot{\phi}_{\mathrm{sb}}$ that prevents substantial drift of θ_{sb}.

Edit 2: It seems like once I increased atol and rtol, this resonance behavior also died out. This isn't super surprising, since the cause was that ϕ wasn't precessing enough during the LK peaks, and $\dot{\phi}$ should never be zero (we match signs of $\Omega_{SL}\hat{L}$ to $\dot{\Omega}\hat{L}_o$).

3 Take 2: Resonance Crossing

3.1 Finding Resonances in the EOM: Deriving a Slow-Merger Criterion

The goal of this calculation will be to identify a criterion by which we may classify the nonadiabatic $\to$ adiabatic transition as sufficiently slow that θ_{eff} is conserved. We want to consider the problem where some constant characteristic precession frequency Ω crosses through some other resonance frequencies $m\Omega'$, then define "sufficiently slowly" such that the transition is adiabatic. In practice this is tricky since there are two strongly varying precession rates, Ω_{SL} and all the LK cycles.

Let's go back to the Hamiltonian in the inertial frame, $H_i = \Omega_{SL}\hat{\mathbf{S}}\cdot\hat{\mathbf{L}}$. We then consider rotating to the frame where $\hat{\mathbf{L}}$ only nutates, so this gives us

$$H = \Omega_{SL}\hat{\mathbf{L}}\cdot\hat{\mathbf{S}} - \dot{\Omega}\hat{\mathbf{z}}\cdot\hat{\mathbf{S}}. \tag{111}$$

The following is in the spirit of Natalia's work. Note that both $\Omega_{SL}\hat{\mathbf{L}}$ and $\dot{\Omega}$ vary greatly over the course of an LK cycle, but are non-zero mean. Define angle brackets and prime quantities to be the mean and mean-subtracted values, then perform Fourier decomposition over T_{LK} Kozai period. WLOG, assume the phases are such that we can just use cosines, then

$$H = \left(\langle \Omega_{SL} \hat{\mathbf{L}} \rangle - \langle \dot{\Omega} \rangle \hat{\mathbf{z}}\right) \cdot \hat{\mathbf{S}} + \left\{ \sum_{m=1}^{\infty} \left[\Omega_{SL} \hat{\mathbf{L}}\right]_m' \cos\left(\frac{2\pi m t}{T_{LK}}\right) - \sum_{N=1}^{\infty} \left[\dot{\Omega}\right]_N' \cos\left(\frac{2\pi N t}{T_{LK}}\right) \hat{\mathbf{z}} \right\} \cdot \hat{\mathbf{S}}. \tag{112}$$

The first term corresponds to the intuitively averaged terms, where $\hat{\mathbf{S}}$ precesses about some effective axis, and the second terms correspond to terms that should average out to zero unless there are resonant terms. Thus, there are clearly possible resonances when $\hat{\mathbf{S}}$ has average precession rate equal to a half-integer multiple of the Kozai period.

3.1.1 Aside: Toy Problem

What actually happens at these commensurabilities though? The Hamiltonian is linear, so there can be no separatrix. The toy model to consider is

$$H = \left[\omega_0 \hat{\mathbf{z}} + \epsilon(\cos(\omega t)\hat{\mathbf{x}} + \sin(\omega t)\hat{\mathbf{y}})\right] \cdot \hat{\mathbf{s}}, \tag{113}$$

$$\frac{\mathrm{d}\hat{\mathbf{s}}}{\mathrm{d}t} = \left[\omega_0 \hat{\mathbf{z}} + \epsilon(\cos(\omega t)\hat{\mathbf{x}} + \sin(\omega t)\hat{\mathbf{y}})\right] \times \hat{\mathbf{s}}, \tag{114}$$

$$\approx \omega_0 \hat{\boldsymbol{\phi}} + \epsilon \hat{\mathbf{z}} \left[\cos(\omega t) s_y - \sin(\omega t) s_x\right]. \tag{115}$$

The zeroth order solution is just $s_x = \sin\theta\cos\phi$ and $s_y = \sin\theta\sin\phi$ and $\phi \approx \omega_0 t$, so the $\hat{\mathbf{z}}$ component simplifies to

$$\frac{\mathrm{d}s_z}{\mathrm{d}t} \approx \epsilon \sin\theta \sin(\omega t - \omega_0 t). \tag{116}$$

This obviously breaks down for $\omega = \omega_0$. What breaks? Well, consider Hamiltonian in rotating frame $H = (\omega_0 \hat{\mathbf{z}} + \epsilon(\dots) - \omega\hat{\mathbf{z}}) \cdot \hat{\mathbf{s}}$ such that the second vector is stationary in space, maybe $\epsilon\hat{\mathbf{x}}$. Then it is very clear that when $\omega = \omega_0$ that rotation is just about $\hat{\mathbf{x}}$ with very reduced frequency ϵ. This explains the coordinate singularity, since some trajectories pass through the poles; if we hadn't dropped the s_ϕ contributions at ϵ order we probably could have seen the trajectory traces out a great circle.

What about when the second axis is just nutating, e.g. if the total rotation is $\omega_0\hat{\mathbf{z}}+\epsilon(A + B\cos(\omega t))\hat{\mathbf{x}}$? Well, we can write down the equation of motion again

$$\frac{\mathrm{d}\hat{\mathbf{s}}}{\mathrm{d}t} \approx \omega_0 \hat{\boldsymbol{\phi}} + \epsilon \hat{\mathbf{z}}(A + B\cos\omega t) s_y. \tag{117}$$

Again, if we assume $s_y \approx \sin\theta \sin\omega_0 t$, then we can use the trick from before to factorize and keep only slowly varying terms

$$\frac{\mathrm{d}s_z}{\mathrm{d}t} \approx \epsilon \sin\theta (A + B\cos\omega t)\sin\omega_0 t, \tag{118}$$

$$\approx \epsilon \sin\theta \frac{B}{2}\left[\sin(\omega_0 t + \omega t) + \sin(\omega_0 t - \omega t)\right]. \tag{119}$$

We have had to introduce an extra term, but again there is a divergence if $\omega_0 = \omega$, which implies that terms that don't start at the pole can reach the pole.

If we want to get the quantitative behavior of this toy model (specifically, we want to know how far the precession axis gets from $\hat{\mathbf{z}}$ when we're near resonance), we just have to write down the EOM carefully. The Hamiltonian in spherical coordinates is

$$H = \omega_0 \cos\theta + \epsilon(A + B\cos\omega t)\sin\theta\cos\phi. \tag{120}$$

This yields EOM

$$\frac{\mathrm{d}\cos\theta}{\mathrm{d}t} = \epsilon(A + B\cos\omega t)\sin\theta\sin\phi, \tag{121}$$

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = \omega_0 + \epsilon(A + B\cos\omega t)\cot\theta\cos\phi. \tag{122}$$

Compare to EOM for the case we could solve exactly above (with the $-\omega\hat{\mathbf{z}}$ rotation)

$$\frac{\mathrm{d}\cos\theta}{\mathrm{d}t} = \epsilon\sin\theta\cos(\omega t)\sin\phi - \epsilon\sin\theta\sin(\omega t)\cos\phi, \tag{123}$$

$$= \epsilon\sin\theta\sin\left(\omega t - \phi\right), \tag{124}$$

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = \omega_0 + \epsilon(\cos\omega t + \sin\omega t)\cot\theta. \tag{125}$$

Thus, it becomes clear: the rotation axis will still tilt all the way into the xy plane, but it has some variations on the time scale of ω_0 (due to the A term) and $\omega_0 + \omega$ (due to the additive term).

What about if ω doesn't exactly equal ω_0, how far are we from tilting all the way? Let's consider the first toy problem for simplicity, then it's very obvious the total rotation axis is just $\epsilon\hat{\mathbf{x}} + (\omega_0 - \omega)\hat{\mathbf{z}}$. Therefore, the "resonance width" is just the strength of the time-varying perturbation.

3.1.2 Solving Toy Problem Exactly

Let's consider again the first toy problem. We can write down the EOM explicitly:

$$\frac{\mathrm{d}(\cos\theta)}{\mathrm{d}t} = \epsilon\sin\theta\sin\phi, \tag{126}$$

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = -\epsilon\sin\phi, \tag{127}$$

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = \omega_0 - \omega. \tag{128}$$

Consider now $\omega = \omega_0 + \dot{\omega}t$. Then, $\frac{\mathrm{d}\phi}{\mathrm{d}t} = \dot{\omega}t$ and $\phi = \dot{\omega}t^2/2 + \phi_0$. Thus, we can integrate explicitly

$$\Delta\theta = \int_{-\infty}^{\infty} \epsilon\sin\left(\frac{\dot{\omega}t^2}{2} + \phi_0\right)\mathrm{d}t, \tag{129}$$

$$= \epsilon\,\mathrm{Im}\left[e^{i\phi_0}\int_{-\infty}^{\infty} e^{i\dot{\omega}t^2/2}\,\mathrm{d}t\right], \tag{130}$$

$$\approx \epsilon\,\mathrm{Im}\left[e^{i\phi_0}\sqrt{2\pi i/\dot{\omega}}\right], \tag{131}$$

$$= \epsilon\sqrt{2\pi/\dot{\omega}}\sin\left(\phi_0 + \pi/4\right). \tag{132}$$

Indeed, this matches our qualitative assertions from before: if $\epsilon \to 0$, for instance, when the Fourier coefficients are all tiny, then there is no time to rotate during the resonance, while if $\dot{\omega} \to \infty$ then the resonance is crossed too quickly to generate misalignment.

3.2 Untangling the Parametric Resonance (LL17)

We proceed by solving for the second toy model. In the second toy model, let's set $A = 0$, since we will consider resonances with the $N = 1$ Fourier component which is zero mean. Also shifting into the rotating frame (by replacing $\phi + \omega t \to \phi$), we obtain

$$\frac{\mathrm{d}\cos\theta}{\mathrm{d}t} = \epsilon B \cos\omega t \sin\theta \sin\left(\phi - \omega t\right), \tag{133a}$$

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = -\epsilon B \left[\sin\phi - \sin\omega t \cos\left(\phi - \omega t\right)\right], \tag{133b}$$

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = \omega_0 - \omega + \epsilon B \cos\omega t \cot\theta \cos\left(\phi - \omega t\right). \tag{133c}$$

This resembles the EOM for the first toy model above, for which a $-\omega\hat{\mathbf{z}}$ rotation gave us a very clean EOM (total rotation axis $\epsilon\hat{\mathbf{x}} + (\omega_0 - \omega)\hat{\mathbf{z}}$), but there are a few fast-varying terms. It's intuitively clear that the two possibilities for the rotation axis in this second toy model are either the same or the $\hat{\mathbf{x}}$ component is something like halved.

To verify this numerically, we compute the range of $\cos\theta$ excited for a given ϵ and fixed $\omega - \omega_0$. For simplicity, we start both $\hat{\mathbf{\Omega}}_1$ and $\hat{\mathbf{S}}$ at $\phi = 0$, and call their polar angles ψ, θ respectively. Under our formalism, if the polar angle of the tilted axis is ψ' (which will also be in the $\phi = 0$) plane, we see that θ ranges between θ and $2\psi' - \theta$, and ψ' is given by

$$\psi' = \arctan\frac{\epsilon}{\omega_0 - \omega}. \tag{134}$$

Upon plotting, we find generally very good agreement for small ϵ, particularly near $\omega \approx \omega_0$. But for $\epsilon = 1$, the total change in $\cos\theta$ is much larger than predicted, saturating at 2 over a wide range of parameters. Furthermore, for $\epsilon = 0.1$, we find a distinct peak at $\epsilon = 0.5\omega_0$. Further investigation shows that reducing ϵ significantly lengthens the timescale of this change in θ, strongly suggesting a parametric-type instability. Also further investigation shows the appearance of smaller peaks at all integer fractions of ϵ/ω_0. Most interestingly, this instability *vanishes* when analyzing the first toy model. Thus, it is behavior unique to the second toy model. Let's recap the behavior we are observing so far:

- Occurs at $\epsilon/\omega_0 = 1/N$.
- Timescale and amplitude of oscillations are ϵ-dependent.
- Only occur for toy model 2.
- Is much stronger for $\psi = 0$ than $\psi = 90$.

The combination of the first and second properties strongly suggest a parametric instability, which explains the timescale growth, while Lindstedt-Poincaré perturbation theory likely gives a bound on the amplitude. The first property may be a result of the identity

$$\int\limits_0^{2\pi} \sin(Nx)\sin(x)\,\mathrm{d}x = \frac{\sin(2\pi N)}{N^2 - 1}. \tag{135}$$

This is consistent with the third bullet as well since the parametric instability relies on a time-modulating amplitude. The final bullet suggests that it is the modulation of the $\hat{\mathbf{z}}$ precession frequency that is generating the parametric instability, completing the picture.

It is tempting to use the analytic LK solution to get ϵ (the $N = 1$ Fourier coefficient) but we should do it numerically, as we know that GR is significant in our regime of interest (e.g. the discrepancy of $\langle\dot{\Omega}\rangle_{LK}$ versus the analytical form).

3.3 Few-Shot Merger Exact (LL18)

We solve for the second source of non-adiabicity, the short-merger limit.

To do this problem exactly, let's start with the LK-averaged $\frac{\mathrm{d}\hat{\mathbf{S}}}{\mathrm{d}t}$ but then transform to the frame where $\hat{\mathbf{\Omega}}_{\mathrm{eff}} \propto \hat{\mathbf{z}}$. Again, this is effected by some rotation about the $\hat{\mathbf{y}}$ axis; let's call it $\dot{I}_{\mathrm{out}}$ for simplicity. Here, I_{out} is the angle between the old and new z axis and is defined such that

$$-\dot{\Omega}\sin I_{\mathrm{out}} + \Omega_{SL}\sin\left(I + I_{\mathrm{out}}\right) = 0. \tag{136}$$

Here, positive I_{out} means a *positive* component along the x axis.

Recall that $\mathrm{sgn}(\dot{\Omega}) = -\mathrm{sgn}(\cos I)$. Let's first specialize to $I < \pi/2$, then $\dot{\Omega} < 0$. Then, when $\dot{\Omega}$ dominates, $I_{\mathrm{out}} < 0, \to 0$, while when Ω_{SL} dominates, $I_{\mathrm{out}} > -I \to -I$. Similarly, when $I > \pi/2$: when $\dot{\Omega}$ dominates, $I_{\mathrm{out}} > 0, \to 0$, and when Ω_{SL} dominates, $I_{\mathrm{out}} < \pi - I \to \pi - I$.

In terms of this, the EOM are then

$$\frac{\mathrm{d}\hat{\mathbf{S}}}{\mathrm{d}t} = \left[-\dot{\Omega}\cos I_{\mathrm{out}}\hat{\mathbf{z}} + \Omega_{SL}\cos\left(I + I_{\mathrm{out}}\right)\hat{\mathbf{z}} - \dot{I}_{\mathrm{out}}\hat{\mathbf{y}}\right] \times \hat{\mathbf{S}}. \tag{137}$$

This gives the anticipated behavior: for $I < \pi/2$, the system stays precessing about $\hat{\mathbf{z}}$, and otherwise around $-\hat{\mathbf{z}}$ (which, in the final state, corresponds to precession around $\pm\hat{\mathbf{L}}$ respectively).

Now, note that Eq. (137) has no resonance structure, but it has decomposed very cleanly: θ dependence depends only on the $\hat{\mathbf{y}}$ component, and the rest only on the $\hat{\mathbf{z}}$ component. As such, we obtain explicit closed form

$$\phi(t) = \int_0^t \left(-\dot{\Omega}\cos I_{\mathrm{out}} + \Omega_{SL}\left(\cos\left(I + I_{\mathrm{out}}\right)\right)\right)\,\mathrm{d}\tau, \tag{138}$$

$$\Delta\theta_{\mathrm{eff}} = -\int\limits_0^\infty \dot{I}_{\mathrm{out}}\cos\phi(t) + (\ldots)\cot\theta\sin\phi\,\mathrm{d}t. \tag{139}$$

Again, it appears we will need a closed form for $\dot{I}_{\mathrm{out}}$. Implicit differentiation gives

$$\dot{I}_{\mathrm{out}}\left(-\dot{\Omega}\cos I_{\mathrm{out}} + \Omega_{SL}\cos\left(I + I_{\mathrm{out}}\right)\right) = \ddot{\Omega}\sin I_{\mathrm{out}} - \dot{\Omega}_{SL}\sin\left(I + I_{\mathrm{out}}\right), \tag{140}$$

$$= \ddot{\Omega}\frac{\Omega_{SL}\sin\left(I + I_{\mathrm{out}}\right)}{\dot{\Omega}} - \dot{\Omega}_{SL}\sin\left(I + I_{\mathrm{out}}\right), \tag{141}$$

$$\dot{I}_{\mathrm{out}}\left(-\cot I_{\mathrm{out}} + \cot\left(I + I_{\mathrm{out}}\right)\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\ln\left(-\dot{\Omega}\right) - \ln\Omega_{SL}\right). \tag{142}$$

It is clear that the LHS coefficients are large except where $I_{\mathrm{out}} \approx (\pi - I)/2$, where $\dot{I}_{\mathrm{out}}$ formally diverges (except for the $\dot{I}$ neglected term), as the RHS is very peaceful ($\sim 1/T_{GW}$, no large variations). This occurs where $\left|\dot{\Omega}\right| \sim \Omega_{SL}$. To regularize the divergence, we smear it out over the characteristic width over which $\dot{I}_{\mathrm{out}}$ is substantial, which hopefully is given by the slope of the coefficient near the singularity.

Consider modeling $\dot{I}_{\mathrm{out}}(t)$ as a Gaussian with some width σ and peaking at some time $t_0 > 0$. σ is then probably roughly $\sigma \sim I/\dot{I}_{\mathrm{out}}$ (the total change in I_{out} is I). The height of the Gaussian for $\dot{I}_{\mathrm{out}}$ is just its peak value, which conveniently is just $\sim \dot{\mathcal{A}}/\sqrt{2}$. Requiring $\dot{I}_{\mathrm{out}}$ have total integral I gives us

$$\dot{I}_{\mathrm{out}} \approx \frac{\dot{\mathcal{A}}}{\sqrt{2}}\exp\left[-\frac{t^2}{2\sigma^2}\right], \tag{143}$$

$$\sigma^2 \approx \frac{I}{\dot{\mathcal{A}}\sqrt{\pi}}. \tag{144}$$

There are two regimes in which this integral is evaluated: (i) σ is small (specifically $\dot\phi\sigma \ll 1$), and we perform the Gaussian integral with a slowly varying phase, or (ii) σ is large and we perform stationary phase integration for fixed $\dot I_{\rm out}$ near its peak value. Don't think the first regime is valid (and I did it wrong below), so we just calculate the second regime.

In the other regime, where σ is large, let's just take $\dot I_{\rm out} \approx \dot{\mathscr{A}}/2$ and integrate the stationary phase, where $\dot\phi$ has a minimum for $t_0 > 0$, so $\tau = t - t_0$

$$\Delta\theta_{\rm eff} = \mathrm{Im}\,\frac{\dot{\mathscr{A}}}{\sqrt{2}}\int\limits_0^\infty e^{i\phi}\,\mathrm{d}t, \tag{145}$$

$$\approx \mathrm{Im}\,\frac{\dot{\mathscr{A}}}{\sqrt{2}}\int\limits_{-\infty}^\infty \exp\left[i\phi_0 + i\dot\phi\tau + \frac{i\ddot\phi}{2}\tau^2\right]\,\mathrm{d}\tau, \tag{146}$$

$$\approx \mathrm{Im}\,\frac{\dot{\mathscr{A}}}{\sqrt{2}}e^{i\phi_0}\int\limits_{-\infty}^\infty \exp\left[\frac{i\ddot\phi}{2}\left(\tau^2 + \frac{2\dot\phi}{\ddot\phi}\tau\right)\right]\,\mathrm{d}\tau, \tag{147}$$

$$\approx \mathrm{Im}\,\frac{\dot{\mathscr{A}}}{\sqrt{2}}e^{i\tilde\phi_0}\sqrt{2\pi i/\ddot\phi}. \tag{148}$$

We want to rewrite $\ddot\phi$, for which we observe

$$\dot\phi = -\dot\Omega\cos I_{\rm out} + \Omega_{SL}\cos\left(I + I_{\rm out}\right), \tag{149}$$

$$= -\mathscr{A}^{-3/8}\dot\Omega_0\cos I_{\rm out} + \mathscr{A}^{5/8}\Omega_{SL,0}\cos\left(I + I_{\rm out}\right), \tag{150}$$

$$\dot\phi(\tau = 0) \approx -\mathscr{A}^{-3/8}\dot\Omega_0\cos\left(\frac{I}{2}\right) + \mathscr{A}^{5/8}\Omega_{SL,0}\cos\left(\frac{I}{2}\right), \tag{151}$$

$$\ddot\phi(\tau = 0) \approx -\cos\left(\frac{I}{2}\right)\dot\phi(\tau = 0)\left(-\frac{3\dot{\mathscr{A}}}{8\mathscr{A}} + \frac{5\dot{\mathscr{A}}}{8\mathscr{A}}\right) = \dot\phi(\tau = 0)\frac{\dot{\mathscr{A}}}{4}, \tag{152}$$

$$\Delta\theta_{\rm eff} \approx \sqrt{\frac{4\pi\dot{\mathscr{A}}}{\dot\phi(\tau = 0)}}\sin\left(\phi_0 + \frac{\pi}{4}\right). \tag{153}$$

Indeed, this now has the desired scaling: $\sqrt{\dot{\mathscr{A}}} \propto T_{\rm m}^{-1/2} \propto \left(1 - e_{\max}^2\right)^{-3/2}$, so $\propto \cos^{-3}(I_0)$. We denote $\dot\phi = -\dot\Omega_0 = \dot\Omega_{\rm SL,0}$ the precession rates sans projection effect.

Finally, let's remember in our simulations that, for identical initial phases ϕ_{sb}, the $\Delta\theta_{\rm eff}(I_0)$ curves are identical (past the one-shot case). This is again in agreement with our prediction above.

Finally, I noticed one more thing: in the above calculation, we somewhat assume I is constant, in these averaged equations. How valid is this? Well, numerically, one of the vector-averaged Is is pretty constant. In particular, $\langle\dot\Omega\hat{\mathbf{L}}\rangle \sim \oint\cos I\,\mathrm{d}\Omega$, which is one of the phase space areas. Since the Kozai constant only changes slowly, we should expect this area to be conserved indeed. Unfortunately, I got confused, and we really want $\langle\Omega_{SL}\mathbf{L}\rangle$. This means that the averaged I is slightly more dominated during the high-e phases, so the averaged I being used above will start slightly above its final value ($\sim 125^\circ$) and the LK max of 141°.

We can probably predict the final value of I by the above argument though, and know its rough range (around 120°), which tells us that the I in our $\Delta\theta_{\rm eff}$ calculation indeed does not change too much.

3.4 Analytical Forms for Average Precession Frequencies

This is just an attempt. We seek to write down, given e_0 and I_0, the analytical form for $\langle \dot{\Omega}\hat{\mathbf{z}} + \Omega_{SL}\hat{\mathbf{L}}\rangle$. To do so, the key expressions are ($a_0 = 1$)

$$x \equiv 1 - e^2 = x_0 + (x_1 - x_0)\mathrm{cn}^2(u, k^2), \tag{154}$$

$$\mathrm{cn}\left(u, k^2\right) \equiv \mathrm{cn}\left(\int_0^\phi \frac{\mathrm{d}\theta}{\sqrt{1 - k^2\sin^2\theta}}\right) = \cos\phi, \tag{155}$$

$$\frac{\mathrm{d}u}{\mathrm{d}\phi} = \frac{1}{\sqrt{1 - k^2\sin^2\phi}}, \tag{156}$$

$$u = \frac{K}{\pi}\left(n_e t + \pi\right), \tag{157}$$

$$n_e = \frac{6\pi\sqrt{6}}{8Kt_{LK}}\sqrt{x_2 - x_1}, \tag{158}$$

$$\hat{\boldsymbol{\Omega}}_{\mathrm{eff}} \equiv \dot{\Omega}\hat{\mathbf{z}} + \Omega_{SL}\hat{\mathbf{L}} = \frac{3\sqrt{h}}{4t_{LK}}\left[1 - \frac{2(x_0 - h)}{x - h}\right]\hat{\mathbf{z}} + \frac{\epsilon_{SL}}{a^4 t_{LK}}\frac{1}{x}\left(\cos I\hat{\mathbf{z}} + \sin I\hat{\mathbf{x}}\right), \tag{159}$$

$$k^2 = \frac{x_0 - x_1}{x_2 - x_1}, \tag{160}$$

$$x_1^2 < \frac{5}{3}\cos^2 I_0, \tag{161}$$

$$x_2 = \frac{5}{3}\frac{h}{x_1} > 1, \tag{162}$$

$$K = \frac{\pi/2}{\mathrm{agm}\left(1, \sqrt{1 - k^2}\right)}, \tag{163}$$

$$h = x\cos^2 I. \tag{164}$$

We've followed the definitions of SL15, such that $2\pi/n_e$ is the eccentricity cycle period. The goal is to integrate $\hat{\boldsymbol{\Omega}}_{\mathrm{eff}}$ over one eccentricity cycle. Let's first try the $\dot{\Omega}$ piece, where $\epsilon = x_1/x_0$:

$$\int_0^{T_{LK}} \dot{\Omega}\,\mathrm{d}t = \frac{3\sqrt{h}}{4t_{LK}}\int_0^{T_{LK}} 1 - \frac{2x_0\left(1 - \cos^2 I_0\right)}{x_0\left(1 - \cos^2 I_0\right) + (x_1 - x_0)\,\mathrm{cn}^2(u, k^2)}\,\mathrm{d}t, \tag{165}$$

$$= \frac{3\sqrt{h}}{4t_{LK}}\left[T_{LK} - \int_0^\pi \frac{2\left(1 - \cos^2 I_0\right)}{1 - \cos^2 I_0 + (\epsilon - 1)\cos^2\phi}\,\mathrm{d}\phi\frac{\mathrm{d}u}{\mathrm{d}\phi}\frac{\mathrm{d}t}{\mathrm{d}u}\right], \tag{166}$$

$$= \frac{3\sqrt{h}}{4t_{LK}}\left[\frac{2\pi}{n_e} - \frac{2\pi}{Kn_e}\int_0^\pi \frac{\sin^2 I_0/\sqrt{1 - k^2\sin^2\phi}}{\sin^2 I_0 + (\epsilon - 1)\cos^2\phi}\,\mathrm{d}\phi\right]. \tag{167}$$

This should let us tease out a scaling later.

The two components of the Ω_{SL} are perhaps somewhat trickier to do but can be written down

$$\int_0^{T_{LK}} \Omega_{SL}\cos I\,\mathrm{d}t = \frac{\epsilon_{SL}}{a^4 t_{LK}}\int_0^{T_{LK}} \frac{h^{1/2}}{x^{3/2}}\,\mathrm{d}t, \tag{168}$$

$$= \frac{\epsilon_{SL}\sqrt{h}}{a^4 t_{LK}}\frac{\pi}{Kn_e}\int_0^{\pi}\left(x_0 + (x_1 - x_0)\cos^2\phi\right)^{-3/2}\left(1 - k^2\sin^2\phi\right)^{-1/2}\,\mathrm{d}\phi, \tag{169}$$

$$\int_0^{T_{LK}} \Omega_{SL}\sin I\,\mathrm{d}t = \frac{\epsilon_{SL}}{a^4 t_{LK}}\int_0^{T_{LK}} \frac{\sqrt{1 - h/x}}{x}\,\mathrm{d}t, \tag{170}$$

$$= \frac{\epsilon_{SL}\sqrt{h}}{a^4 t_{LK}}\frac{\pi}{Kn_e}\int_0^{\pi}\frac{\sqrt{1 - h/\left(x_0 + (x_1 - x_0)\cos^2\phi\right)}}{\left(x_0 + (x_1 - x_0)\cos^2\phi\right)\sqrt{1 - k^2\sin^2\phi}}\,\mathrm{d}\phi. \tag{171}$$

The general expression takes on similar form. The above integrals have been numerically validated to agree with explicit integrations, though seem to breakdown for substantial ϵ_{GR}, and give useful guidance as to the scalings. This expression as a whole will not be terribly useful, since the two ways adiabaticity breaks down is for (i) fast, wide mergers (in which case we need to evaluate $\dot{\phi}(\mathscr{A} \sim 1)$), or for (ii) slow mergers that cross resonances, which must be compact if they are to merge within a Hubble time.

3.5 Using Resonances and Numerical LK Solutions

I spent a few weeks away from the writeup doing mostly numerical work; these are captured in my weeklies. The present focus is on understanding the deviation from exact conservation of $\theta_{\rm eff}$. We did a Fourier decomposition of the $\mathrm{d}\hat{\mathbf{S}}/\mathrm{d}t$ EOM. The analysis when $N \geq 1$ can be ignored is given above, and we find very swift decay of $\Delta\theta_{\rm eff}$.

The next question to ask is how resonances can affect this conservation, when we reintroduce terms $N \geq 1$. This analysis can only be relevant in the slow merger regime, where an instantaneous Fourier decomposition of the LK inspiral (where we neglect GW but include GR) is a good approximation to the trajectory. We considered this analysis from two perspectives, the toy model approach (analyzed above) and the numerical LK approach, which was performed largely offline. In both setups, the key quantity we consider is the amplitude of oscillation of $\theta_{\rm eff}$, denoted A (confusingly, sometimes also $\theta_{\rm eff}$, but I'll fix this for the paper) if the orbit is fixed in time indefinitely. The two results are generally in good agreement, and we summarize the results below:

- Most importantly, A can be computed using a particular a_0, e_0, I_0 numerically (e_0 and I_0 are defined when eccentricity is smallest). These results seem to agree to quite good accuracy when using a_0, e_0, and I_0 to compute $\mathbf{\Omega}_{\rm eff,0}$, $\mathbf{\Omega}_{\rm eff,1}$ and $P_{\rm LK}$ for the toy model and measuring A that way. I need to make a plot for this.

  They are also in agreement with the original $\Omega_{\rm eff,0}P_{\rm LK}$ plots that we made a bit back, further supporting the conclusion that we're on the right track.

- The locations of the resonances in the Paper I physical regime are consistent with intermediate inclinations deviating significantly from the adiabatic prediction.

- There is a gross overprediction of $\Delta\theta_{\rm eff}$ by examining $A\left(a_0(t), e_0(t), I_0(t)\right)$ in the Paper II regime, since numerically, an initial inclination of 90.5° for the inspiral simulation sees $\Delta\theta_{\rm eff} < 0.01°$, but A gets to be as large as 8°. Thus, a mechanism is needed to ensure the $\theta_{\rm eff}$ oscillations go back down in amplitude and maintain the same value as initially.

It is furthermore worth noting that the two resonances in the toy model when $\Omega_{\rm eff,0} P_{\rm LK} = \pi$ have much longer oscillation timescales (almost 100 orbits) than the 2π ones (which are well understood, and have periods of just a few orbits as $\epsilon \lesssim 1$ for these highly eccentric orbits).

The key puzzle at this point is however why both the $\pi, 2\pi$ resonances are able to generate significant $\Delta\theta_{\rm eff}$ values while the substantian A in the Paper II regime (which is likely caused by ψ, the misalignment angle between the $N = 1$ and $N = 0$ components, growing as $\Omega_{\rm SL}/\dot{\Omega}$ oscillates about 1 and $\mathbf{\Omega}_{\rm eff}$ nutates a lot in space) fails to generate any $\Delta\theta_{\rm eff}$ at all. I hypothesize that the key mechanism is that as $a_0(t)$ decreases, $P_{\rm LK}$ decreases tremendously (really just the period in ω), so it is actually very easy to *adiabatically* move the effective effective precession axis. We must next analyze the relevant timescales and see whether this is still consistent with a failed prediction in the Paper I regime.

We find currently:

- The 90.5° Paper II simulation goes from $A \simeq 9^\circ$ to $A = 0^\circ$ over the span of a few cycles. This implies the precession axis has to change by about 4° over a few LK periods. In this limit, $\Omega_{\rm eff} P_{\rm LK}/(2\pi) \sim 0.5$, so we have just a few cycles to change the precession axis by a few degrees. This seems like it could be adiabatic.

- The 80° Paper I simulation should have passed a resonance quite early in its evolution (we indeed predict this using the model above, but there is a "missing" resonance in `6_Iscan`, when comparing to the N1 plot). This implies the resonance is exceedingly narrow, and explains why deviation from the adiabatic prediction is quite small when I_0 for the inspiral gets closer to 90°: GW becomes more efficient, so the time spent inside the resonance is shorter. This is a $\Omega_{\rm eff} P_{\rm LK}/2\pi = 0.5$ resonance, so its growth rate is somewhat slower, so resonance crossing is always in the nonadiabatic regime, and faster passings yield smaller kicks.

  It is surprising this resonance should be narrow, since in the high eccentricity regime, the $N = 1$ component should have a substantial perpendicular angle (relatively speaking, ψ is larger), and the $N = 1$ component should be larger (larger temporal variations since the oscillation is less mild than at $\sim 125^\circ$). Furthermore, we see from the N1 plots that, graphically, the slope of $\Omega_{\rm eff} P_{\rm LK}$ is quite flat at this resonance crossing, implying there should be a larger range of inclinations that satisfy the resonant condition. TODO investigate.

The next orders of business are: (i) ensure GW-driven passage through the fringes of the resonance (Paper II regime) is always adiabatic, (ii) find the missing resonance, (iii) check agreement of numerical LK coefficients + toy model vs. real A plots, (iv) understand the nature of the toy model resonances (at least identify the locations of resonances and instability timescales).

It seems plausible that (i) the Paper II regime system is adiabatic at the very end, but it's not obvious why precession cuts off.

4 Floquet Theory

4.1 Formal Results

Consider that we have EOM

$$\frac{\mathrm{d}\mathbf{S}}{\mathrm{d}t} = \mathbf{\Omega}_{\rm e} \times \mathbf{S}. \tag{172}$$

Note that we solve for $\mathbf{S} \in \mathbb{R}^3$, not $\hat{\mathbf{S}} \in \mathscr{S}^2$, so the EOM is guaranteed to be linear (translations/rotations in $\mathscr{S}^2$ do not commute, but translations commute in $\mathbb{R}^3$).

For a linear, periodic EOM of form

$$\dot{x} = A(t)x, \tag{173}$$

where x is a N-vector and A is an $N \times N$ matrix, we can apply Floquet theory. In our problem, $N = 3$ and $A_{ij} = \epsilon_{ijk}[\Omega_{\mathrm{e}}]_k$ is the antisymmetric 3×3 matrix.

Floquet theory also requires N linearly independent solutions, to form the *fundamental matrix solution* $\phi(t)$, the matrix whose columns are linearly independent solutions. More convenient to use is $\Phi(t)$, the principal fundamental matrix solution, which satisfies $\Phi(0) = \mathbb{I}_N$. In our problem, this is very easy to construct: since the cross product is guaranteed to be non-degenerate, we simply evolve the three unit vectors forwards in time. Linearity guarantees they remain orthogonal.

The fundamental matrix solution is used to write down formal solution

$$x(t) = \Phi(t)\Phi^{-1}(0)x(0) = \Phi(t)x(0). \tag{174}$$

In the Floquet problem, Φ has the further property

$$\Phi(t+T) = \Phi(t)\Phi^{-1}(0)\Phi(T) = \Phi(t)\Phi(T). \tag{175}$$

Generally, $M \equiv \phi^{-1}(0)\phi(T)$ is known as the *monodromy matrix*, particularly simple for the principal fundamental matrix solution. Thus, we can notice that

$$x(nT) = \Phi(nT)x(0) = [\Phi(T)]^n\, x(0). \tag{176}$$

Recall that $\Phi(0) = [\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}]$, while $\Phi(T)$ propagates these forward in time by T. But note that the columns of $\Phi(T)$ must still be unit vectors (our particular case; $|\mathbf{S}|$ is conserved by EOM), therefore M must be an orthogonal 3×3 matrix, or a rotation matrix. Furthermore, since M is built of many infinitesimal *proper* rotations, we argue M must be a proper rotation matrix as well. Thus, its eigenvalues are $1, (a \pm bi)$ satisfying $a^2 + b^2 = 1$, and M is a rotation matrix about its eigenvector with eigenvalue 1. Numerically, Eq. (176) seems to hold to numerical precision, and so $[x \cdot \hat{\mathbf{v}}_1](nT)$ are constant, where $\hat{\mathbf{v}}_1$ is the rotation axis.

Now, let's try to write down the coordinate transform into a time-independent ODE. Define complex matrix $\mathbf{B}$ such that $e^{T\mathbf{B}} = \mathbf{M}$. In the eigenbasis of $\mathbf{M}$, it is clear $\mathbf{B}$ will take on form $[[0, i\theta], [-i\theta, 0]] \otimes 0$, a rotation generator and a constant. We then know that

$$\Phi(t) = P(t)e^{tB}, \tag{177}$$

$$\equiv Q(t)e^{tR}. \tag{178}$$

Here, $P(t)$ is complex with period T while Q is real with period $2T$, and R, B are constant. With this construction, we can use

$$y = Q^{-1}(t)x, \tag{179}$$

$$\dot{y} = Ry. \tag{180}$$

This is obvious since $y = y_0 e^{tR}$ and $x = Q(t)y = Q(t)e^{tR}(Q^{-1}(0)y_0)$ or $x(t) = \Phi(t)x_0$.

The primary advantages of this Floquet formulation are two-fold:

- Proves that evolution is periodic, and cannot exhibit chaos.

- Proves that a characteristic period (I think $2T_{\mathrm{LK}}$, but might be T_{LK}) can be used to generate an iterative map, over which evolution looks just like a rotation in $\mathbb{R}^3$. This bypasses all the complex harmonic terms, and boils all the dynamics down to the monodromy matrix (a rotation matrix).

The rotation axis $\mathbf{R}$ is of key importance, as it plays the role of our current $\bar{\mathbf{\Omega}}_{\mathrm{e}}$, and satisfies

$$\mathbf{S}(t=0) = \mathbf{R} \tag{181}$$

$$0 = \int\limits_0^{T_{\mathrm{LK}}} \mathbf{\Omega}_{\mathrm{e}} \times \mathbf{S}\,\mathrm{d}t. \tag{182}$$

We generally would expect this to $\approx \bar{\mathbf{\Omega}}_{\mathrm{e}}$, but it can differ. Plots (`4sims`) show the location of this rotation axis compared to $\bar{\mathbf{\Omega}}_{\mathrm{e}}$.