Separating out research-related tidbits from non-research ones.

Contents

1 06/29/19—Collisionless Boltzmann Equation in Galaxies: Landau Damping

Inspired by `https://arxiv.org/pdf/1906.08655.pdf`. The problem is basically formulated as thus: consider a kinetic-theoretic description of a fluid using distribution function $f(t,x,p)$ which obeys collisionless Boltzmann equation $\frac{\mathrm{d}f}{\mathrm{d}t} = 0$ (we use p instead of v to work in Hamiltonian coordinates). Introducing a periodic perturbation to this fluid results in a singular dispersion relation, which can be resolved via the usual Landau prescription (consider a perturbation having grown from zero at $t = -\infty$). The dispersion relation describes *Landau damping* (or growth), in which energy from the fluid is exchanged with the perturber.

1.1 Linearized EOM

The point of the paper is instead to analytically compute the impact of the perturber on the distribution function, to quantify the *scarring* of a galaxy upon encounters with a nearby perturber. The equations of motion coupling the distribution function and gravitational potential are given

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \{*\}f, \mathscr{H} = 0, \tag{1}$$

where $\mathscr{H} = \frac{p^2}{2} + \Phi$ and $\{*\}\dots$ denotes the Poisson bracket $\{*\}f, \mathscr{H} = \vec{\nabla}_x f \cdot \vec{\nabla}_p \mathscr{H} - \vec{\nabla}_p f \cdot \vec{\nabla}_x \mathscr{H}$.

If we linearize for perturbation quantities f_1, Φ_1 where $\Phi_1(x)$ does not depend on the momenta, we obtain

$$\begin{aligned} 0 &= \frac{\partial f_1}{\partial t} + \{*\}f_1, \mathscr{H}_0 - \vec{\nabla}_p f \cdot \vec{\nabla}_x \mathscr{H}_0, \\ &= \frac{\partial f_1}{\partial t} + \vec{\nabla}_x f_1 \cdot \vec{p} - \vec{\nabla}_p f_1 \cdot \vec{\nabla}_x \Phi_0 - \vec{\nabla}_p f_0 \cdot \vec{\nabla}_x \Phi_1. \end{aligned}$$

Oops welp I guess I never solved this.

2 02/16/23—Linear Predictive Coding: Autoregressions and Fourier Transforms

This was a simple enough inquiry initially: given a partial time series that contains sinusoids, how do we extract the frequency? We know one way to do this using the FFT, but there are advantages to other techniques. Courtesy of Jeremy Goodman's pointers.

The trick has to do with autoregressions. Suppose we are looking to extract l frequencies of form $e^{i\omega_m t}$, so that

$$y_n = \sum_m^l C_m e^{i\omega_m n\Delta t}. \tag{2}$$

Thus, if we have at least $2l$ points or so, we should be able to fit for the $2l$ DOF C_m and ω_m. There can be a noise term above if need be, in which case more points will smooth out the noise.

What is the trick? Well, we compute the l-th order autoregression. In other words, for each sequence of l points, we can write down the expression satisfying:

$$y_n = \sum_{m=1}^l a_m y_{n-m}. \tag{3}$$

With l-many such sequences, we have enough equations to solve for the l-many unknowns a_m. These can be written in the matrix form:

$$\begin{bmatrix} y_n \\ y_{n+1} \\ \vdots \\ y_{n+l} \end{bmatrix} = \begin{bmatrix} y_{n-1} & y_{n-2} & \cdots & y_{n-l} \\ y_n & y_{n-1} & \cdots & y_{n-l+1} \\ \vdots & \vdots & \dots & \vdots \\ y_{n+l-1} & y_{n+l-2} & \cdots & y_{n-1} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_l \end{bmatrix}. \tag{4}$$

These $\{a_l\}$ form the AR(l) autoregressive model for y_n.

This is great, but how do we get the frequencies, or also maybe the growth rates? Now, we rewrite the above equation as

$$0 = \begin{bmatrix} y_n & y_{n-1} & \cdots & y_{n-l} \\ y_{n+1} & y_n & \cdots & y_{n-l+1} \\ \vdots & \vdots & \ldots & \vdots \\ y_{n+l} & y_{n+l-1} & \cdots & y_{n-1} \end{bmatrix} \begin{bmatrix} -1 \\ a_1 \\ \vdots \\ a_l \end{bmatrix} \equiv \mathbf{B}\cdot\vec{a}. \tag{5}$$

Now, what if the y_n look like λ^n for some complex λ? Then the λ must satisfy

$$1 - a_1\lambda - a_2\lambda^2 - \cdots - a_l\lambda^l = 0. \tag{6}$$

This is the characteristic equation for this AR(l) model. If we solve for the roots of this equation, we get the possible values of λ that satisfy the model. In other words, if the $y_n = \lambda^n$ indeed, then $\mathbf{B}\cdot\vec{a} = 0$ as requested above. Then, if the data are oscillatory, then $\lambda = e^{i\omega_m}$ as requested above.

2.1 Intuitive Understanding

There's something slightly unintuitive here: we began by seeking the frequencies in the y_n, but we eventually obtained this by solving an equation that has *nothing* to do with the y_n, the characteristic equation for the AR(l) model! Why does this make sense?

Well, it should be noted that a particular AR(l) model does not uniquely specify the y_n; this would be impossible, since there are only l DOF in the model and $2l+1$ in Eq. (4). Indeed, this suggests that the amplitudes of the modes C_m, as well as the initial normalization of the autoregressive chain (y_{n-l}) are free parameters. As such, we can imagine that the AR(l) model permits a family of solutions, any with the correct frequencies. In other words, we could also imagine writing:

$$\mathbf{B}\cdot|a\rangle = \sum_{m=1}^{l} C_m |b_m\rangle\langle b_m||a\rangle. \tag{7}$$

Another way to think about the characteristic equation is as exactly a characteristic equation of a matrix. If we consider the matrix that maps the vector

$$\begin{bmatrix} y_N \\ y_{N-1} \\ \vdots \\ y_{N-l} \end{bmatrix} = \mathbf{M} \begin{bmatrix} y_{N-1} \\ y_{N-2} \\ \vdots \\ y_{N-l-1} \end{bmatrix}, \tag{8}$$

then it's clear that $\mathbf{M}$ has the form

$$\mathbf{M} = \begin{bmatrix} a_1 & a_2 & \ldots & a_{l-1} & a_l \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 1 & 0 \end{bmatrix}. \tag{9}$$

It's then clear $\mathbf{M}$ apparently has exactly the characteristic equation that we prescribe above. This makes sense: the matrix $\mathbf{M}$ tells us whether a vector-valued sequence of y values is growing, decaying, or oscillating.

As such, the final conclusion of this tidbit is this: the autoregression is another way of expressing a Markov chain that allows us to advance the time series. Then, note that any sequence with $y_n = z^n$ where z is complex (allowing periodic or exponential sequences) has z as one of the eigenvalue of its Markov chain matrix, or z is a root of its characteristic polynomial. Turning this on its head, if we compute the autoregression for a sequence and find a root w of its characteristic polynomial, this implies that the sequence has a geometric component with factor w. Applying this to sequence with a periodic component with frequency ω, we see that $e^{i\omega\Delta t}$ must be a root of the characteristic polynomial of its autoregression.

3 02/21/23—Chaotic vs Diffusive Behavior

This is a short section. Dong (and others) talk about "chaotic tides", where the mode amplitude grows stochastically because the forcing occurs with random amplitude. However, this is not chaos, but should be properly termed "diffusive tides."

How can we argue for this? Well, the defining characteristic of chaos is a positive Lyapunov exponent, i.e. an exponential growth rate of the separation between two trajectories with nearly-idential initial conditions.

$$\begin{aligned} \delta y(t) &= y(t; y_0) - y(t; y_0 - \epsilon), && (10)\\ &\sim \mathcal{O}(\epsilon e^{\lambda t}). && (11) \end{aligned}$$

What is the growth rate for a random walk, or diffusive growth?

Let's adopt the simplest model for now, a discrete random walk with step size ± 1. Perhaps, for the sake of consistency, we can imagine that the step is determined based on the current value of x, e.g. whether the rounded value of $10^9 x$ is even or odd. Then, consider two initially adjacent x. It is obvious that

$$|\delta x(t)| \le 2t. \tag{12}$$

So we already see that diffusive behavior is not chaotic.

But now, we have a fun little math problem. Consider two random walks starting at $x = 0$ with step size ± 1. What is the mean and variance of δx? Well, using the usual CLT guidelines, the linearity of expectation gives $\mathrm{E}(X_2 - X_1) = 0$ while linearity of variance gives $\mathrm{Var}(X_2 - X_1) = 2\mathrm{Var}(X_1) = 2t$. Thus, the separation between two walkers grows stochastically and $\sim \sqrt{t}$. This is not chaos, where the separation grows deterministically and $\sim \exp(\lambda t)$.

4 02/23/23—Rayleigh Distribution

4.1 2D

The Rayleigh distribution is commonly used for mutual inclinations. Can we briefly give ourselves some intuition for it?

It's known that the Rayleigh distribution is the magnitude of a 2D vector with two normally-distributed components. Thus, consider if X and Y are drawn from $N(0,\sigma)$. Then the CDF of the magnitude $M = \sqrt{X^2+Y^2}$ of the vector is calculated as

$$F_M(m;\sigma) = \iint\limits_{D(m)} f_U(u;\sigma) f_V(v;\sigma)\,\mathrm{d}u\mathrm{d}v \tag{13}$$

Here, $D(m)$ is the unit disc satisfying $\sqrt{u^2+v^2} \le m$, and $f_U = N(0,\sigma)$ and $f_V = N(0,\sigma)$ are the PDFs of the random variables U,V. Upon changing to polar coordinates and integrating:

$$F_M(m;\sigma) = \int\limits_0^m 2\pi \frac{1}{2\pi\sigma^2} \exp\left[-\frac{u^2+v^2}{2\sigma^2}\right] m'\mathrm{d}m', \tag{14}$$

$$= \frac{1}{\sigma^2} \int\limits_0^m m' \exp\left[-\frac{(m')^2}{2\sigma^2}\right] \mathrm{d}m', \tag{15}$$

$$f_M(m;\sigma) = \frac{m}{\sigma^2} \exp\left[-\frac{m^2}{2\sigma^2}\right]. \tag{16}$$

This is a Rayleigh distribution with width σ.

Now, if we know that two vectors have a magnitude separation that is Rayleigh distributed (like Rayleigh distribution), and we know that they are iid, how can we obtain their individual distributions (under the right assumptions)? Well, let's first assume that the four vector components are all normally distributed with $N(0,\sigma)$. Then their separation vector's components are also normally distributed with:

$$f_{X_1-X_2}(x_1-x_2;\sigma) = N\left(0,\sigma\sqrt{2}\right), \tag{17}$$

$$f_{Y_1-Y_2}(y_1-y_2;\sigma) = N\left(0,\sigma\sqrt{2}\right). \tag{18}$$

Thus, the separation vector's magnitude is Rayleigh distributed with width parameter $\sigma\sqrt{2}$. Thus, to generate a pair of vectors whose separation magnitude is Rayleigh distributed with width δ, we can just generate the vector components from $N\left(0,\delta/\sqrt{2}\right)$.

To briefly comment, this obviously doesn't depend on the number of vectors, as long as the measured Rayleigh distribution is for arbitrary pairs in the system.

4.2 3D

What about in 3D? This is just as much for my own practice with manipulating PDFs and CDFs than anything. Consider a vector $\left(v_x, v_y, v_z\right)$ with all three components drawn from

$N(0,\sigma)$. What is the distribution of the magnitude? Again:

$$F_M(m;\sigma) = \frac{1}{(2\pi\sigma^2)^{3/2}} \iiint_{D(m)} f_{Vx}(v_x;\sigma) f_{Vy}\left(v_y;\sigma\right) f_{Vz}(v_z;\sigma)\, \mathrm{d}^3 v, \tag{19}$$

$$= \frac{4\pi}{(2\pi\sigma^2)^{3/2}} \int_0^m e^{-(m')^2/(2\sigma^2)}\, (m')^2 \mathrm{d}m', \tag{20}$$

$$f_M(m;\sigma) = \frac{4\pi m^2}{\left(2\pi\sigma^2\right)^{3/2}} e^{-m^2/(2\sigma^2)}. \tag{21}$$

This of course can easily generalize: it is clear that in N-D:

$$f_M^{(N)}(m;\sigma) = \frac{S_m^{(N)}}{\left(2\pi\sigma^2\right)^{N/2}} e^{-m^2/(2\sigma^2)}, \tag{22}$$

where $S_m^{(N)}$ is the surface area of the N-sphere with radius m.

And what about the separation vector between some $\vec{v}$ and $\vec{w}$ in 3D? Well, each component of $\vec{v}-\vec{w}$ has distribution $N\left(0,\sigma\sqrt{2}\right)$ again, so if $U = |\vec{v}-\vec{w}|$, then its PDF is

$$f_U(u;\sigma) = \frac{u^2}{\sigma^3\sqrt{4\pi}} e^{-m^2/(4\sigma^2)}. \tag{23}$$

5 03/15/23—Mass Loss and Binary Orbit Change

Without kicks, Hills 1983 seems to have the best prescription. Time to revisit this problem now that I've done it wrong literally every time I've tried to do it.

5.1 Brute Force Circular

We start with a circular orbit and in the rest frame of the binary. Call the pre-ML energy E and post-ML energy E'. These are the sums of kinetic and gravitational potential energies, so

$$E = \frac{m_1 v_1^2}{2} + \frac{m_2 v_2^2}{2} - \frac{G m_1 m_2}{a}, \tag{24}$$

$$E' = E - \frac{m_1 v_1^2}{2}(1-f) + \frac{G m_1 m_2}{a}(1-f). \tag{25}$$

Here, $m_1' = f m_1$ is the post-ML mass. Note then that

$$v_1^2 = \left(\frac{a m_2}{m_{13}}\right)^2 \frac{G m_{12}}{a} = \frac{G m_2^2}{m_{12} a}, \tag{26}$$

$$K_{\mathrm{cm}} = \frac{((1-f) m_1 v_1)^2}{2(f m_1 + m_2)}. \tag{27}$$

Here, K_{cm} is the kinetic energy associated with the motion of the post-ML binary's center of mass. To undergo unbinding, we need f such that $E' = K_{\mathrm{cm}}$, so that in the co-moving frame the binary is unbound. This is a laborious calculation, but we can write ($f' \equiv 1-f$ is the fraction of mass lost from m_1)

$$\begin{aligned}
0 &= E' - K_{\rm cm},\\
&= -\frac{Gm_1m_2}{2a} - \frac{m_1v_1^2}{2}f' + \frac{Gm_1m_2}{a}f' - \frac{\left(f'm_1v_1\right)^2}{2(fm_1+m_2)},\\
&= -\frac{Gm_1m_2}{2a} - \frac{Gm_1m_2^2}{2m_{12}a}f' + \frac{Gm_1m_2}{a}f' - \frac{\left(f'm_1\right)^2}{2(fm_1+m_2)}\frac{Gm_2^2}{m_{12}a},\\
&= -\frac{1}{2} - \frac{m_2}{2m_{12}}f' + f' - \frac{m_1m_2(f')^2}{2(m_{12}-f'm_1)m_{12}},\\
&= -(m_{12}-f'm_1)m_{12} - m_2\left(m_{12}-f'm_1\right)f' + 2(m_{12}-f'm_1)m_{12}f' - m_1m_2(f')^2,\\
&= -m_{12} + f'(m_1 - m_2 + 2m_{12}) - 2m_1(f')^2,\\
&= (f')^2 - \frac{f'}{2}\left(3+\frac{m_2}{m_1}\right) + \frac{1}{2}\left(1+\frac{m_2}{m_1}\right),\\
f' &= \frac{(3+q)/2 \pm \sqrt{(3+q)^2/4 - 2(1+q)}}{2},\\
&= \frac{(3+q)/2 \pm ((q-1)/2)}{2},\\
&= \frac{1+q}{2},\\
&= \frac{m_{12}}{2m_1}.
\end{aligned}$$

Whew. This is the canonical result, that we need to lose $f'm_1 = \frac{m_{12}}{2}$ mass to unbind the system.

5.2 Easier Circular

The primary difficulty above was that we had this stupid center of mass kinetic energy to carry around. This can be simplified if we just recognize that we only need to compute the contribution of the reduced mass to the energy to understand whether the system remains bound. Recall that the the kinetic energy of of the reduced mass component is just

$$K_{\rm red} = \frac{\mu v_{\rm rel}^2}{2}, \tag{28}$$

$$E_{\rm red} = K_{\rm red} - \frac{Gm_{12}\mu}{a} = -\frac{Gm_{12}\mu}{2a}, \tag{29}$$

$$v_{\rm rel}^2 = \frac{Gm_{12}}{a}, \tag{30}$$

$$E'_{\rm red} = \frac{\mu v_{\rm rel}^2}{2} - \frac{Gm'_{12}\mu'}{a} = 0, \tag{31}$$

$$= \frac{Gm_{12}\mu'}{2a} - \frac{Gm'_{12}\mu'}{a}. \tag{32}$$

Thus, we end up with the result that $m'_{12} = m_{12}/2$ results in a reduced-mass energy = 0 and unbinding. It's important to recognize that $v_{\rm rel}$ is the relative velocity of the two particles, given by $\mathbf{v}_{\rm rel} = \mathbf{v}_2 - \mathbf{v}_1$, which does not change with instantaneous mass loss.

5.3 Eccentric Unbinding

The argument in the previous section is much easier to generalize to a general orbit. Consider that the orbit has semimajor axis a and unbinds when the separation is at r. Then

$$\begin{aligned}
E_{\rm red} &= \frac{1}{2}\mu v_{\rm rel}^2 - \frac{Gm_{12}\mu}{r} = -\frac{Gm_{12}\mu}{2a},\\
\frac{v_{\rm rel}^2}{2} &= Gm_{12}\left(-\frac{1}{2a} + \frac{1}{r}\right),\\
E'_{\rm red} &= \frac{1}{2}\mu' v_{\rm rel}^2 - \frac{Gm'_{12}\mu'}{r},\\
&= \mu'\left[-\frac{Gm_{12}}{2a} + \frac{Gm_{12}}{r}\right] - \frac{Gm'_{12}\mu'}{r}.
\end{aligned}$$

Setting this equal to zero, we find

$$0 = \mu'\left[-\frac{Gm_{12}}{2a} + \frac{Gm_{12}}{r}\right] - \frac{Gm'_{12}\mu'}{r}, \tag{33}$$

$$\frac{m'_{12}}{m_{12}} = -\frac{r}{2a} + 1. \tag{34}$$

If we re-express $m'_{12} \equiv m_{12} - \Delta$, then we can rewrite

$$1 - \frac{\Delta}{m_{12}} = 1 - \frac{r}{2a} = 1 - \frac{1-e^2}{2(1+e\cos f)}. \tag{35}$$

This makes sense: since the mass loss effects a torque on the system, we have to give it the maximum torque to unbind the system, which occurs at pericenter. Thus, qualitatively we need a minimum eccentricity of $(1-e)/2 \sim m'_{12}/m_{12}$ to unbind the system of $m'_{12} > m_{12}/2$, i.e. if the mass loss is too little.

5.4 Bound Orbits: Final Eccentricity

Of course, these exercises can be repeated if we would like for bound orbits, and tracking the angular momentum of the system as well to get eccentricity. I may redo this some day, but for now I will just cite the result from Hills 1983, where the final eccentricity is given by

$$e = \left\{1 - (1-e_0^2)\left[\frac{1-(2a_0/r)(\Delta/m_{12})^2}{1-\Delta/m_{12}}\right]\right\}^{1/2}. \tag{36}$$

6 Pendulum Periods

It might be helpful to just do the simple pendulum in a few ways to get its period. Specifically, I mean the oscillation of the nondimensionalized EOM

$$\ddot{\theta} = -\sin\theta. \tag{37}$$

In the small angle approximation $\theta \ll 1$, we have that the frequency of the oscillator is just 1, so the period is 2π.

6.1 Lindstedt-Poincaré

I always get this wrong, so let's try again. The zeroth order solution is $\theta = \epsilon\cos(\omega_0 t)$ where $\omega_0 = 1$. Let's next imagine that the frequency has a small ϵ dependence, so that

$$\begin{aligned}
\theta(t) &= \epsilon\cos((1+\epsilon\omega_1)t),\\
\ddot{\theta} &= -\epsilon(1+\epsilon\omega_1)^2\cos((1+\epsilon\omega_1)t),\\
&\approx -\epsilon\cos((1+\epsilon\omega_1)t) + \frac{1}{6}\epsilon^3\cos^3((1+\epsilon\omega_1)t).
\end{aligned}$$

Using the quick identity

$$\begin{aligned}
\cos^3\theta &= \cos\theta - \cos\theta\sin^2\theta\\
&= \cos\theta + \frac{\cos\theta(\cos 2\theta - 1)}{2}\\
&= \frac{\cos\theta}{2} + \frac{\cos(3\theta) + \sin\theta\sin 2\theta}{2}\\
&= \frac{\cos\theta}{2} + \frac{\cos(3\theta)}{2} + \sin^2\theta\cos\theta\\
&= -\cos^3\theta + \frac{3\cos\theta}{2} + \frac{\cos(3\theta)}{2},\\
\cos^3\theta &= \frac{3\cos\theta}{4} + \frac{\cos(3\theta)}{4},
\end{aligned}$$

we obtain

$$-\epsilon(1+\epsilon\omega_1)^2\cos((1+\epsilon\omega_1)t) \approx -\epsilon\cos((1+\epsilon\omega_1)t) + \frac{1}{6}\epsilon^3\cos^3((1+\epsilon\omega_1)t).$$

Matching coefficients of the first frequency term, we find

$$\begin{aligned}
-\epsilon(1+2\epsilon\omega_1) &\approx -\epsilon + \frac{\epsilon^3}{8},\\
\omega_1 &= -\frac{\epsilon}{16},\\
\omega &= 1 - \frac{\epsilon^2}{16}.
\end{aligned}$$

In our Ph106b class notes, we solve the Duffing oscillator with this technique, for which $\ddot{\theta} = -\theta - \epsilon\theta^3$ and we obtain that $\omega = 1 + (3/8)\epsilon A^2$ for oscillation amplitude A. For the simple pendulum, $\epsilon = -1/6$, and so we recover that $\omega = 1 - A^2/16$. We didn't do this strictly correctly, I guess, since our small parameter was the oscillation amplitude A instead of the perturbing term ϵ, but we obtain the right result: **the oscillation period grows with larger amplitude**. We can already anticipate that $\epsilon \to 4$ will produce problems, and indeed $\epsilon = \pi$ corresponds to the upside-down pendulum.

6.2 Explicit Integral

The period of the pendulum can be solved exactly using the method of quadratures. During oscillation, the total energy of the system is conserved,

$$E = -\cos\theta + \frac{\dot{\theta}^2}{2}. \tag{38}$$

Note that in this notation, $\theta = 0$ is the bottom, so $\theta \in [-\pi, \pi]$. We would formally derive this by writing down the Lagrangian and making the Legendre transform, but in the present case if we just identify $p = \dot{\theta}$ the conjugate momentum to θ then we find immediately that $\dot{p} = -\partial E/\partial\theta = -\sin\theta$ and that $\dot{\theta} = \partial E/\partial p = p = \dot{\theta}$. Thus, we can explicitly write:

$$\begin{aligned}\dot{\theta}(\theta) &= \sqrt{2(E + \cos\theta)},\\ &= \sqrt{2(\cos\theta - \cos\theta_0)}.\end{aligned}$$

The period of the pendulum is formally defined such that

$$P = 4\int_0^{\theta_0} \frac{\mathrm{d}t}{\mathrm{d}\theta}\,\mathrm{d}\theta$$

This is the amount of time it takes for the pendulum to go from an initial condition $\pm\theta_0$ to 0, so a quarter-period. For sufficiently small θ_0, the quadrature expression can be expanded

$$\begin{aligned}\dot{\theta} &\approx \sqrt{-\theta^2 + \theta_0^2}\\ P &= 4\int_0^{\theta_0} \frac{1}{\sqrt{\theta_0^2 - \theta^2}}\,\mathrm{d}\theta.\end{aligned}$$

We know this is arcsin, but is it obvious? Actually, yeah:

$$\begin{aligned}\int^y \frac{1}{\sqrt{1 - x^2}}\,\mathrm{d}x &= \int^{\arcsin(y)} \frac{1}{\sqrt{1 - \sin^2 u}}\,\mathrm{d}(\sin u),\\ &= \int^{\arcsin(y)} \mathrm{d}u,\\ &= \arcsin(y).\end{aligned}$$

So then

$$\begin{aligned}
P &= 4\left[\arcsin\left(\frac{\theta}{\theta_0}\right)\right]_0^{\theta_0},\\
&= 2\pi.
\end{aligned}$$

Great.

What about in the nonlinear limit though? In full generality, we have

$$P = 4\int\limits_0^{\theta_0} \frac{1}{\sqrt{2(\cos\theta - \cos\theta_0)}}\,\mathrm{d}\theta.$$

In the limit where $\theta_0 \to \pm\pi$, we cannot make any approximations since θ will span the full angular interval. However, θ_0 sufficiently close to $\pm\pi$, the unstable points, we recognize that the dominant contribution to P will be near π. Thus, let's instead ask the question: for some fixed $\theta_1 \ll 1$, how long does it take for the trajectory to reach θ_1 as $\theta_0 \to \pi$? We again should be able to make expansions now (let $\phi \equiv \pi - \theta$)

$$P \gtrsim 4\int\limits_{\phi_0}^{\phi_1} \frac{1}{\sqrt{\phi^2 - \phi_0^2}}\,\mathrm{d}\phi.$$

Note here that $\phi > \phi_0$. This one is probably a cosh?

$$\begin{aligned}
\int\limits^{y} \frac{1}{\sqrt{x^2-1}}\,\mathrm{d}x &= \int\limits^{\cosh^{-1}(y)} \frac{1}{\sqrt{\cosh^2(u)-1}}\,\mathrm{d}(\cosh(u)),\\
&= \cosh^{-1}(y),\\
P &\gtrsim 4\left[\cosh^{-1}\left(\frac{\phi}{\phi_0}\right)\right]_{\phi_0}^{\phi_1},\\
&\gtrsim 4\ln\left(\phi_1/\phi_0\right) \sim -4\ln\left(\pi_0\right).
\end{aligned}$$

Here, we've made use of the fact that $\cosh(x) \approx \exp(x)/2$ for large x. Hence, we recover the logarithmic divergence that is expected.

7 Distributions of Functions of Random Variables

I can never remember how to do this, so let me just write it down.

If we have a random variable X and a second random variable Y that satisfies $y = y(x)$, then the PDF of Y is simple:

$$\int f_Y(y)\,\mathrm{d}y = \int f_X(x)\,\mathrm{d}x, \tag{39}$$

$$f_Y(y) = f_X(x)\frac{\mathrm{d}x}{\mathrm{d}y}. \tag{40}$$

What if we have a random variable Z that is a function of two random variables X, Y satisfying $z(x, y)$? This is a little trickier, but we need to write down the CDF

$$\int f_Z(z)\,\mathrm{d}z = \iint f_Y(y) f_X(x)\,\mathrm{d}x\mathrm{d}y. \tag{41}$$

This is no longer solvable in general (but we can often do well in statistical cases with moment-generating functions, CLT, and others). But for sufficiently simple dependencies, we can do this. Let's just consider $z = x + y$, for $x, y \in \mathcal{U}_{[0,1]}$. This is then easy to do:

$$\int_0^z f_Z(w)\,\mathrm{d}w = \int_0^{\min(z,1)} \int_0^{\min(z-y,1)} \mathrm{d}x\,\mathrm{d}y, \tag{42}$$

$$= \int_0^{\min(z,1)} \min(z-y, 1)\,\mathrm{d}y, \tag{43}$$

$$= \begin{cases} \int_0^z z - y\,\mathrm{d}y & z < 1 \\ \int_0^{z-1} \mathrm{d}y + \int_{z-1}^1 (z-y)\,\mathrm{d}y & z > 1, \end{cases} \tag{44}$$

$$= \begin{cases} z^2/2 & z < 1 \\ (z-1) + (z - 1/2) - z(z-1) + (z-1)^2/2 & z > 1, \end{cases} \tag{45}$$

$$f_Z(z) = \begin{cases} z & z < 1 \\ 2 - z & z > 1. \end{cases} \tag{46}$$

We can do the same for $z = xy$:

8 Change in Mutual Inclination in Hierarchical Triples due to SNe

We find that when a hierarchical stellar triple has its inner and outer orbits initially isotropically distributed, the final mutual inclination distribution is not isotropic, but is depleted near 90°. We build a simple quantitative model analogous to this phenomenon and show that it has an exact solution.

The essence of this behavior is that: a symmetric SNe in the inner orbit results in an effective kick to the outer orbit in the plane of the inner orbit, denoted $\vec{v}_{\mathrm{k,eff}}$. Only the component of this kick aligned with the outer orbit normal contributes to realignment of the outer AM. Thus,

$$\Delta I \lesssim \Delta I_{\max} \propto v_{\mathrm{k,eff}} \sin I. \tag{47}$$

The actual change in I is approximately symmetrically distributed over the interval $[-\Delta I_{\max}, \Delta I_{\max}]$. Note that, strictly speaking, neither I nor $\cos I$ are uniformly distributed: imagine that the initial AM is along $\hat{\mathbf{l}}_{\mathrm{b},0}$, then the final AMs are distributed in an azimuthally symmetric way about $\hat{\mathbf{l}}_{\mathrm{b},0}$. This is closer to a symmetric distribution in I than $\cos I$ though.

To see what effect this has on the outer inclination, we imagine a diffusion equation for $I \in [0,\pi]$ with diffusion coefficient $D_0 \sin I$. This can be thought of as the cumulative effect of infinitely many, infinitesimally small effective kicks. This has the form

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial I}\left[D_0 \sin I \frac{\partial f}{\partial I}\right], \tag{48}$$

where $f(t,I)$ is the PDF of I at time t. I really just wanted to solve this PDE lol.

We first consider the steady-state solutions of this PDE. One possible family of solutions requires that

$$D_0 \sin I \frac{\partial f}{\partial I} = C, \tag{49}$$

$$-\sin^2 I \frac{\partial f}{\partial \cos I} = \frac{C}{D_0}, \tag{50}$$

$$\frac{\partial f}{\partial \cos I} = \frac{C}{D_0\left(\cos^2 I - 1\right)}. \tag{51}$$

But since $'(x) = (1-x^2)^{-1}$, we find that this is just

$$f(\cos I) = \frac{C}{D_0}(\cos I). \tag{52}$$

Then C can be set by the normalization of $f(t,I)$. The second homogeneous solution is simply the family of linear solutions $f(t,I) = A \times I$. If the IC is symmetric, then the symmetry is preserved under evolution, and we can require that $f'(t,0) = 0$ at all times. This thus requires that

$$f(\cos I \in [-1,1]) = \frac{C}{D_0}\left|(\cos I) - \cos I\right|. \tag{53}$$

This works since $'(0) = 1$, so the derivative is indeed zero at $\cos I = 0$, and the solution is symmetric. C is again set by the normalization, which I'm too lazy to compute. This is thus the steady-state solution, and we see that it vanishes at the origin and is singular at $\cos I = \pm 1$. This is indeed the behavior we were beginning to see!

Of course, with just a single kick, we don't evolve to this steady state f, but only a little bit. Nevertheless, this gives a quantitative model reflecting the depletion near $\cos I = 0$.