Axiverse Machine

Masha Baryakhtar, David Cyncynates, Ella Henry

August 4, 2025

1 Anthropics

The bounds for the possible ratio of the dark matter abundance to baryon abundance, ζ, are given in [1] to be $2.5 < \zeta < 100$. The lower bound is related to the value for which perturbations near the size of our galaxy's stop growing, and the upper bound is related to the value that ultimately prevents star formation. All values within this range are allowed in the sense that they would lead to our existence today, but their likelihoods depend on the theory for dark matter, as we will see below. In section 5 of [2], the authors consider the case of n axions and the fraction of the dark matter that they constitute. If the axions make up all of the dark matter, then we can write ζ as:

$$\zeta = \sum_{a=1}^{n} c(m_a) F(\theta_a) \tag{1}$$

Here, $c(m_a)F(\theta_a)$ is the fractional abundance in the a-th axion, where $F(\theta_a)$ encodes the dependence of the fractional abundance on the axion's intial misalignment angle. For small enough angles, $F(\theta_a) \sim \theta_a^2$. The $c(m_a)$ factor takes into account the fractional abundance's dependence on the mass and on the scale of symmetry breaking, e.g. for axions that start to oscillate during radiation domination, $c(m_a) \sim \left(\frac{m_a}{H_{eq}}\right)^{1/2} \left(\frac{f_a}{M_{pl}}\right)^2$.

1.1 The anthropic measure

The authors of [2] then calculate the probability P of finding ourselves in a causal patch where the dark matter abundance is no larger than the observed value of $\zeta \sim 5$, without being so small that we could not exist, in other words, the probability that $2.5 < \zeta < 5$. The calculation uses the causal diamond measure (details in [1]) to circumvent the problem of computing probabilities in an infinite-dimensional space, which leads to difficulties in normalizing these probabilities. The measure allows us to restrict ourselves to one causal patch and consider the number of observations made in that patch. To determine this number, the author of [1] counts the number of baryons in a causal patch and then the number of observations per baryon. The quantitative expression

the author finds for the number of baryons in a causal patch is $(1 + \zeta)^{-1}$ with the assumption of fixed baryon-to-photon ratio. The number of observers per baryon is approximated to be constant within the range of allowed ζ values. The overall contribution to the probability density from the causal diamond measure is therefore $C(1 + \zeta)^{-1}$, for some constant C.

We must also assume some distribution for the initial conditions θ_a, which will depend on the scale of inflation relative to the axion's potential: for high scale inflation, the distribution is flat between $[-\pi, \pi]$, while for low scale inflation, $p(\theta_a) \propto \exp\left(-\frac{8\pi^2 V(\theta_a)}{3H^4}\right)$, where $V(\theta_a) = \Lambda^4(1 - \cos\theta_a)$. We can approximate this distribution from low scale inflation as flat between $[-\sigma, \sigma]$, where $\sigma^2 = 3H^4/(8\pi^2 m^2 f^2)$ is the spread in this distribution. So in general, we can consider $p(\theta_a)$ to be flat between two values, which we'll call $\pm f(H)$, where:

$$f(H) = \begin{cases} \pi & \text{for } H > \Lambda \\ \sigma & \text{for } H < \Lambda \end{cases} \tag{2}$$

Now, using the causal diamond measure described above and the flat distribution between $\pm f(H)$ for the initial conditions, we can write down an expression for the probability P:

$$P = N^{-1} \frac{C}{(2f(H))^n} \int_{2.5 < \zeta(\theta_a) < 5} \frac{\prod_{a=1}^{n} d\theta_a}{1 + \zeta(\theta_a)} \tag{3}$$

where again n is the number of axions and C is the constant factor from the number of observations per baryon. The normalization N is such that the probability is normalized to 1 and is given by:

$$N = \frac{C}{(2f(H))^n} \int_{2.5 < \zeta(\theta_a) < 100} \frac{\prod_{a=1}^{n} d\theta_a}{1 + \zeta(\theta_a)} \tag{4}$$

We can compute this probability for any value of $f(H)$ by rescaling the integration variables as $\theta_a' = \theta_a/f(H)$ and writing $\zeta(\theta_a) = c_a\theta_a^2$ (summing over the a index):

$$P = \mathcal{N}^{-1} \int_{-1}^{1} \frac{\prod_{a=1}^{n} d\theta_a'}{1 + c_a'\theta_a'^2} \Theta(c_a'\theta_a'^2 - 2.5)\Theta(5 - c_a'\theta_a'^2) \tag{5}$$

where the θ_a-independent factors out front in Eq. (3) have canceled out with those in the normalization N, and so $\mathcal{N}$ is the same as N, without those factors. Θ here is the Heaviside function and $c_a' \equiv c_a f(H)^2$.

1.2 1 axion

The probability for one axion is given by Eq. (5) for $n = 1$ and can be computed numerically as a function of c_1. It is plotted below:

Fig. 1 shows that once $c_1' > \zeta_{\text{upper}} = 100$, the probability becomes independent of the axion parameters (recall that $c' \propto m^{1/2}f^2$), and acquires the
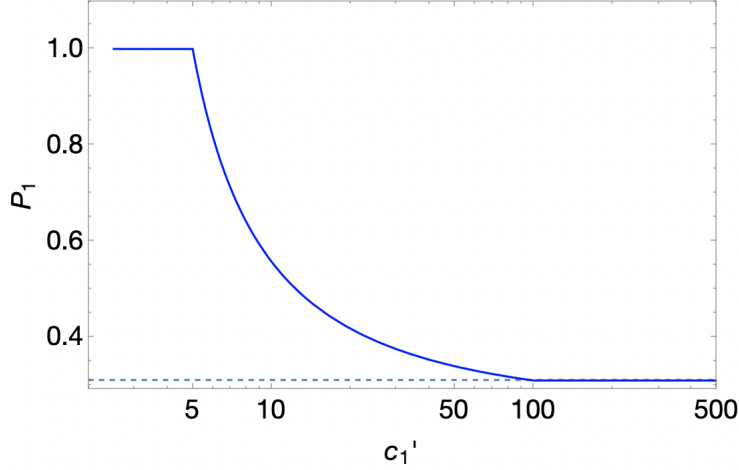
2

Figure 1: Probability of measuring $2.5 < \zeta < 5$ in the case of one axion, as a function of c_1'. The dashed horizontal line corresponds to $P = 0.3$, the probability found for 1 axion in [2]

.

constant value of $P = 0.3$. This corresponds to the $n = 1$ case of the result presented in Eq. 79 of [2], reproduced below:

$$P_n = \frac{\int_{2.5}^{5} \frac{d\zeta \zeta^{(n-2)/2}}{1+\zeta}}{\int_{2.5}^{100} \frac{d\zeta \zeta^{(n-2)/2}}{1+\zeta}} = 0.3,\ 0.16,\ 0.06,\ 0.02,\ 0.006,\ \dots \tag{6}$$

In this result, the authors focus on n axions at the GUT scale ($f \sim 10^{16}$ GeV), with masses heavy enough ($m_a \geq 10^{-19}$ eV) so that the axions would find themselves in this constant P regime. Indeed, one can check that for $f \sim 10^{16}$ GeV and $m_a \geq 10^{-19}$ eV, the criterion $c_a > 100$ is satisfied.

1.3   2 axions

We can numerically compute the probability in Eq. (5) now for $n = 2$, and plot the results as a function of c_2', for a few chosen values of c_1'. This is shown below in Fig. 2

The behavior of P_2 shows that there is a little more nuance, but the result that an axion contributes to the anthropic measure once $c' > \zeta_{\text{upper}}$ holds true. We see that if $c_1' \ll \zeta_{\text{upper}}$ and $c_2' > \zeta_{\text{upper}}$ (see red line for larger c_2' in Fig. 2), the the first axion does not contribute while the second one does, and thus we recover the $P = 0.3$ probability from the 1-axion scenario. The same is true if $c_1' > \zeta_{\text{upper}}$ and $c_2' \ll \zeta_{\text{upper}}$ (Fig. 2, blue line for small c_2'). If both $c_1', c_2' > \zeta_{\text{upper}}$ (Fig. 2, blue line for large c_2'), then the two axions contribute to the measure and so the probability is $P = 0.16$, in agreement with Eq. (6). There is a weird
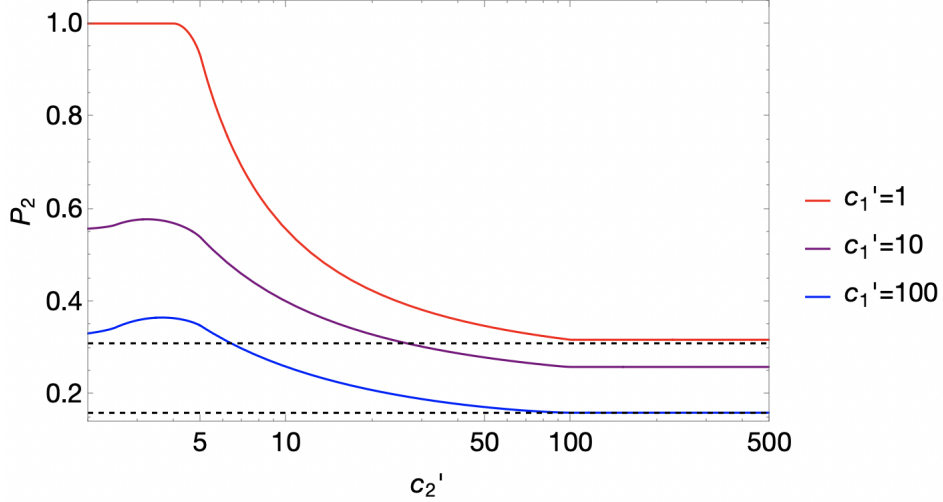
Figure 2: Probability of measuring $2.5 < \zeta < 5$ in the case of two axion as a function of c_2', for different values of c_1'. The dashed horizontal lines correspond to the probabilities for 1 and 2 axions as given by Eq. (6): 0.3 and 0.16 respectively.

.

middle regime (purple line in Fig. 2, as well as region in all three lines where there is a bump), where I am not really sure what's going on yet.

We have thus seen, in the case of 1 and 2 axions, that the condition that an axion must satisfy to contribute to the anthropic measure is $c' > \zeta_{\text{upper}}$. Recalling that $c' = c f(H)^2$ (with $f(H)$ given in Eq. (2)), and that $c = (m/H_{eq})^{1/2}(f/M_{pl})^2$ we can write this condition in terms of the physical parameters as:

$$\left(\frac{m}{H_{eq}}\right)^{1/2}\left(\frac{f}{M_{pl}}\right)^2 \min\left\{\pi^2, \frac{3H^4}{8\pi^2 m^2 f^2}\right\} > \zeta_{\text{upper}} \tag{7}$$

The takeaway is that if the scale of inflation is low, heavier axions will evade this criterion and therefore will not contribute to the anthropic measure: including them in the theory does not require more fine-tuning. On the other hand, if the scale of inflation is high, these axions will quickly satisfy this condition and require fine-tuning, while lighter axions won't (the words "lighter" and "heavier" depend on the values of H and f in each case respectively; this will need to be further investigated).

Note that Eq. (7) is approximate in that: (1) it assumes that $F(\theta_a) \sim \theta^2$, which is not true if $\theta_a \sim O(1)$, in which case there is a small correction; (2) it approximates the distribution of θ_a in a low scale inflation scenario to be flat between $\pm\sigma$, instead of treating it fully as $\propto \exp\left(-V(\theta_a)/H^4\right)$; and (3) it only accounts for regimes where $H < \Lambda$ or $H > \Lambda$, and glosses over the $H \sim \Lambda$ regime.

1.4 Exact anthropic probability with low scale inflation

In the previous section, we used an approximate form for the distribution of initial conditions from low scale inflation. We can compute the anthropic probability *for one axion* more precisely in the case of low scale inflation by numerically calculating the following:

$$P = N^{-1} \int_{2.5 < \zeta < 5} \frac{d\theta}{1 + \zeta} \exp\left(-\frac{8\pi^2 V(\theta)}{3H^4}\right) \tag{8}$$

where the normalization is the same integral but with integration bound $2.5 < \zeta < 100$. This probability is plotted against $\sqrt{mf}$ for certain values of H in Fig. 3 (and f is set to 1).
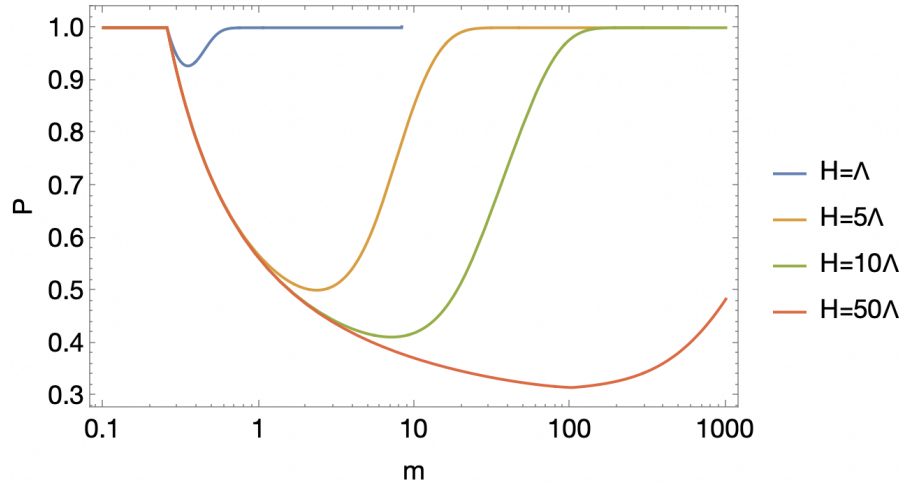

Figure 3: Probability of $2.5 < \zeta < 5$ for a stochastic axion scenario (low scale, long-lasting inflation), with different vales of H relative to $\Lambda = (mf)^{1/2}$. In this plot, $f = 1$.

As expected, for each curve, we see that when the scale of inflation is low relative to the size of the potential (i.e. $H < \Lambda$), the probability that $2.5 < \zeta < 5$ grows, since you are sampling θ from a distribution peaked closer and closer to zero (effectively acts like a small c). When $H > \Lambda$, you recover the expected behavior from flat initial conditions: the probability drops as Λ increases (until $\Lambda \sim H$, and then you enter the first regime).

Things to do:

1. could include correction to the energy density from cosine potential in the probability calc

2. what happens below $c = \zeta_{\text{upper}}$?

5

3. and related to the previous point, what is the bump? what does it looks like with more axions?

4. visualize/better understand the parts of parameter space that are viable via anthropics using Eq. (7)

5. keep in mind dependence of results on choice of measure (causal diamond vs others?)

2 GSification

Consider a theory of two interacting axions $\varphi_1/f_1 \equiv \theta_1$ and $\varphi_2/f_2 \equiv \theta_2$ with a Lagrangian of the following form:

$$\mathcal{L} = \frac{1}{2}\partial_\mu\theta_1\partial^\mu\theta_1 + \frac{1}{2}\partial_\mu\theta_2\partial^\mu\theta_2 + \Lambda_b \cos\left(a\theta_1 + b\theta_2\right) + \Lambda_s \cos\left(c\theta_1 + d\theta_2\right)) \quad (9)$$

where we can assume $\Lambda_b > \Lambda_s$ and have set $f_1 = f_2 = 1$ for now. The mixing matrix between the two fields can be written as follows:

$$A \equiv \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (10)$$

Following the Gram-Schmidtificaiton (GS) procedure, we want to perform a change of basis in order to rewrite this matrix as a lower triangular one, call it T. In other words, we are looking to redefine the fields in the above Lagrangian, so that the first cosine term contains only one redefined field, say ϕ_1, while the second contains a linear combination of ϕ_1 and the second redefined field, ϕ_2:

$$\begin{cases} a\theta_1 + b\theta_2 = t_{11}\phi_1 \\ c\theta_1 + d\theta_2 = t_{21}\phi_1 + t_{22}\phi_2 \end{cases} \quad\Leftrightarrow\quad A\begin{pmatrix}\theta_1 \\ \theta_2\end{pmatrix} = \begin{pmatrix} t_{11} & 0 \\ t_{21} & t_{22} \end{pmatrix}\begin{pmatrix}\phi_1 \\ \phi_2\end{pmatrix} \quad (11)$$

We want to choose the d_{ij} so that the kinetic terms remain in canonical form after redefinition, so we will impose $\frac{1}{2}\partial_\mu\theta_i\partial^\mu\theta_i = \frac{1}{2}\partial_\mu\phi_i\partial^\mu\phi_i$ and no cross terms. Dropping factors of $1/2$ and derivatives for less cumbersome computation, we impose $\theta_i^2 = \phi_i^2$. Using Eq. (11), we write θ_1 and θ_2 in terms of the redefined fields:

$$\begin{pmatrix}\theta_1 \\ \theta_2\end{pmatrix} = A^{-1}T\begin{pmatrix}\phi_1 \\ \phi_2\end{pmatrix} \quad (12)$$

Using the matrices A and T as defined above, and then perform an inner product of this vector with itself to find:

$$\theta_1^2 + \theta_2^2 = \frac{1}{(\det A)^2}\Big[\phi_1^2\left(t_{21}^2\left(a^2 + b^2\right) - 2t_{11}t_{21}(ac + bd) + t_{11}^2\left(c^2 + d^2\right)\right)$$
$$+2t_{22}\phi_2\phi_1\left(t_{21}\left(a^2 + b^2\right) - t_{11}(ac + bd)\right) + t_{22}^2\phi_2^2\left(a^2 + b^2\right)\Big]$$

The condition from the canonical kinetic terms means that we must impose:

$$\begin{cases} \text{coefficient of } \phi_1^2 = (\det A)^{-2}\big((d - bx)^2 + (ax - c)^2\big) = 1 \\ \text{coefficient of } \phi_1\phi_2 = (\det A)^{-2}\big(2(d - bx)by - 2(ax - c)ay\big) = 0 \\ \text{coefficient of } \phi_2^2 = (\det A)^{-2}y^2(b^2 + a^2) = 1 \end{cases} \tag{13}$$

Solving for t_{ij}, we find:

$$\begin{cases} t_{11} = \sqrt{a^2 + b^2} \\ t_{21} = \frac{ca + db}{(\det A)^2} \\ t_{22} = \pm\frac{\det A}{\sqrt{a^2 + b^2}} \end{cases} \tag{14}$$

Thus, the Lagrangian is:

$$\begin{aligned} \mathcal{L} =& \frac{1}{2}\partial_\mu\phi_1\partial^\mu\phi_1 + \frac{1}{2}\partial_\mu\phi_2\partial^\mu\phi_2 + \Lambda_b\cos\left(\sqrt{a^2 + b^2}\phi_1\right) + \Lambda_s\cos\left(\frac{(ca + db)\phi_1 + \det A\phi_2}{\sqrt{a^2 + b^2}}\right) \\ \sim& \frac{1}{2}\partial_\mu\phi_1\partial^\mu\phi_1 + \frac{1}{2}\partial_\mu\phi_2\partial^\mu\phi_2 + \Lambda_b\cos\left(\sqrt{a^2 + b^2}\phi_1\right) + \Lambda_s\cos\left(\frac{\det A}{\sqrt{a^2 + b^2}}\phi_2\right) \end{aligned}$$

Because $\Lambda_b > \Lambda_s$, ϕ_1 will receive the significant part of its dynamics from the first cosine term and so we can view this as the Lagrangian for two decoupled axions, ϕ_1 and ϕ_2, each in their own cosine potential. The fields have a mass hierarchy set by $\Lambda_b > \Lambda_s$ and have decay constants rescaled from those of the original theory as follows:

$$\begin{cases} f_1 \to f_1/\sqrt{a^2 + b^2)} \\ f_2 \to f_2(a^2 + b^2)/\det A \end{cases} \tag{15}$$

Now consider the slope, γ, in the $\log\left(\rho_2/\rho_1\right)$ vs. $\log\left(m_2/m_1\right)$ plane, where $\rho_{1(2)} \sim m_{1(2)}^{1/2}(f_{1(2)}\theta_{i,1(2)})^2$ is the energy density in the decoupled fields $\phi_{1(2)}$, and $m_{1(2)}$ is its mass. For uniformly distributed misalignment angles, $\phi_{i,1(2)}$, the expectation for this power law is given by:

$$\langle\gamma\rangle = \frac{\log\left(m_2^{1/2}f_2^2\right) - \log\left(m_1^{1/2}f_1^2\right)}{\log\left(m_2/m_1\right)} \tag{16}$$

After the GS procedure, we send the decay constants to their rescaled values as given in Eq. (15) and we find that we can write:

$$\langle\gamma\rangle = \frac{\log\left(F(m_2/m_1)^{1/2}\right)}{\log\left(m_2/m_1\right)} \tag{17}$$

where $F \equiv ((a^2 + b^2)/\det A)^2$ and we have set the decay constants $f_1 = f_2 = 1$. In our problem, the mixing coefficients $\{a, b, c, d\}$ are each chosen from $\{-1, 0, 1\}$ with equal probability. We would like to compute an average value of the power

law, given that our potential is chosen at random in this manner. We find that there are only four possible values for F: 0, undefined,1,1/4. The cases where $F = 0$ or F is undefined are not physically interesting as they correspond to cases where either the linear combinations in the cosine terms are linearly dependent, or where too many of these coefficients are zero, and so the fields may not appear at all in the potential, or may not be mixed. If these values of F are drawn, then we redraw mixing coefficients (if the potential had more cosine terms, this would be equivalent to moving to the cosine with the next largest Λ). We can do this as many times as it takes to get $F = 1$ or $F = 1/4$, which corresponds to a Markov chain process, and we can calculate the probability of ending up in these configurations by constructing a transition matrix, where the rows correspond to the incoming state, and the columns correspond to the outgoing state. As an example, I will work out the problem for the configuration/potential-choosing method outlined above.

There are 81 possible choices for the pairs (a, b) and (c, d). Out of those 81, 24 correspond to cases where $F = 0$, 9 to $F =$ undefined, 32 to $F = 1$ and 16 to $F = 1/4$. The transition probability matrix is given by:

$$\mathbb{P} = \begin{pmatrix} 24/81 & 9/81 & 32/81 & 18/81 \\ 24/81 & 9/81 & 32/81 & 18/81 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{18}$$

. Here, the choice of basis is such that the probabilities in a row or column correspond to the following order: $(F = 0, F =$ undefined$, F = 1, F = 1/4)$. Notice that the rows add to 1 to conserve probability. The last two rows correspond to the fact that if we chose $a, b, c,$ and d such that $F = 1$ or $1/4$, then we do not re-draw. Probability theory tells us that:

$$P(X_t = j \mid X_0 = i) = (\mathbb{P}^t)_{ij} \tag{19}$$

and so we can calculate the probability of starting with either $F = 0$ or undefined, and ending in either $F = 1$ or $1/4$ by looking at the appropriate entries of some high power of $\mathbb{P}$, which will end up converging to something close to:

$$\begin{pmatrix} 0 & 0 & 2/3 & 1/3 \\ 0 & 0 & 2/3 & 1/3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{20}$$

Therefore the overall probability of having either $F = 1$ or $F = 1/4$ is (this is bad notation, inconsistent with notation in Eq. (19) but anyway)

$$P(F = 1) = \frac{32}{81} + \frac{2}{3} \frac{24 + 9}{81} = \frac{54}{81} \tag{21}$$

$$P(F = \frac{1}{4}) = \frac{16}{81} + \frac{1}{3} \frac{24 + 9}{81} = \frac{27}{81} \tag{22}$$

where the second terms in the intermediate step of each line correspond to the contribution form choices of $F = 0$ or undefined. We can therefore determine the weighted average of $\langle \gamma \rangle$ using Eq. (17) as:

$$
\begin{aligned}
\langle \gamma \rangle_{GS} &= \frac{54}{81} \frac{1}{2} + \frac{27}{81} \frac{\log \left[(1/4)(m_2/m_1)^{1/2} \right]}{\log (m_2/m_1)} \\
&= \frac{1}{3} + \frac{27}{81} \left(\frac{-\log 4 + \frac{1}{2} \log (m_2/m_1)}{\log (m_2/m_1)} \right) \\
&= \frac{1}{2} - \frac{1}{3} \frac{\log 4}{\log (m_2/m_1)}
\end{aligned}
$$

A   Introduction

Consider the general task of determining how the initial condition of a collection of scalar fields $\boldsymbol{\theta}(0)$ maps onto the final partition of energy density among the various mass eigenstates, where the dynamics are controlled by the following Lagrangian:

$$
\mathcal{L} = \frac{1}{2} \dot{\boldsymbol{\theta}}^T \boldsymbol{K} \dot{\boldsymbol{\theta}} - V(\boldsymbol{\theta}) \, . \tag{23}
$$

The kinetic term matrix $\boldsymbol{K}$ may be diagonalized through the following field redefinition

$$
\boldsymbol{\psi} \equiv \boldsymbol{F} \boldsymbol{R}^T \boldsymbol{\theta} \, , \qquad \boldsymbol{F}^2 \equiv \boldsymbol{R}^T \boldsymbol{K} \boldsymbol{R} \, , \tag{24}
$$

where $\boldsymbol{R}^T \boldsymbol{R} = 1$, and $\boldsymbol{F}^2$ is the diagonal matrix of $\boldsymbol{K}$ eigenvalues. In this basis, the Lagrangian becomes

$$
\mathcal{L} = \frac{1}{2} \dot{\boldsymbol{\psi}}^T \dot{\boldsymbol{\psi}} - V(\boldsymbol{R} \boldsymbol{F}^{-1} \boldsymbol{\psi}) \, . \tag{25}
$$

Although one may as well have started with Equation (25), starting with Equation (23) is essential if the form of V is specified, since mixing induced by the off-diagonal kinetic terms will play an important role in the dynamics.

The potential $V(\boldsymbol{\psi}) \equiv V(\boldsymbol{R} \boldsymbol{F}^{-1} \boldsymbol{\psi})$ is in general an arbitrarily complicated function of the fields, and it may have multiple, physically inequivalent minima, from which $\boldsymbol{\psi}$ may be initialized arbitrarily far away. Therefore, at early times there is no sense in which one may define any particular linear combination of the fields to carry some definite fraction of the energy density of the system. Only at asymptotically late times is one guaranteed that the fields settle into a local minimum of the potential, though I emphasize that one does not know the location or curvature of this part of the potential *a priori*. Therefore, our algorithm to separate out the various late-time mass eigenstates of the fields must avoid making explicit reference to the shape of the potential about any particular point.

A.1 Separable energy density

To simplify notation, we will use V'' and V' to refer to the second and first derivatives $\partial^2 V(\psi)/\partial\psi\partial\psi$ and $\partial V(\psi)/\partial\psi$ respectively. The matrix V'' defines the square of the local mass eigenvalues, and so we must redefine the fields to work in the mass eigenstates:

$$\phi \equiv S^T \psi, \qquad M^2 \equiv S^T V''(\psi)S, \tag{26}$$

where $S^T S = 1$ and M^2 is the diagonal matrix of $V''(\psi)$ eigenvalues. The energy density contained in the field oscillations is then

$$\rho_{\text{tot}} \equiv \frac{1}{2}\dot{\phi}^T\dot{\phi} + \frac{1}{2}(\phi - \phi_0)M^2(\phi - \phi_0), \tag{27}$$

where ϕ_0 is the location of the minimum of the potential, which approximately satisfies

$$0 = V''(\psi)(\psi - \psi_0) + V'(\psi) = SM^2(\phi - \phi_0) + SV'(\phi) \tag{28}$$

Thus, the energy density is

$$\rho_{\text{tot}} \equiv \frac{1}{2}\dot{\phi}^T\dot{\phi} + \frac{1}{2}V'(\phi)^T M^{-2}V'(\phi). \tag{29}$$

Labeling each of the mass eigenstates by their mass eigenvalue m, we may decompose the energy density:

$$\rho_m \equiv \frac{1}{2}\dot{\phi}_m^2 + \frac{1}{2}[M^{-2}(V'(\phi))^2]_m. \tag{30}$$

I assume a convention where the modes are sorted from lightest to heaviest, with $m = m_0$ being the lightest eigenvalue and m_{-1} the heaviest.

A.2 Pruning trivial modes

The procedure of Section A.1 is sufficient to determine the final partition of energy densities at asymptotically late times. In practice, integrating the equations of motion to very late times where one may assume that all states have settled to the bottom of some local minimum is impossible if the masses span more than a few orders of magnitude, since the number of time steps that must be taken is proportional to the ratio of the largest to smallest masses. On the other hand, heavy particles will settle to the minimum of their potential long before the lightest mass eigenstates have begun to oscillate, and so we may remove them from the integration by carefully checking that these modes are dynamically decoupled from all others.

To eliminate the heavy modes, it is enough to replace the damping term $3H\dot{\phi}_{m_i}$ with the critically damped term $2m_i\dot{\phi}_{m_i}$. Note, though, that at every step, the heavy modes will still involve the heavy mass scale. Therefore, one should re-scale the heavy-mode EOM by the mass (squared) of the lightest mode that has been integrated out. This will keep the heavy modes dynamically forced to the bottom of their potentials, while removing any fast timescales.
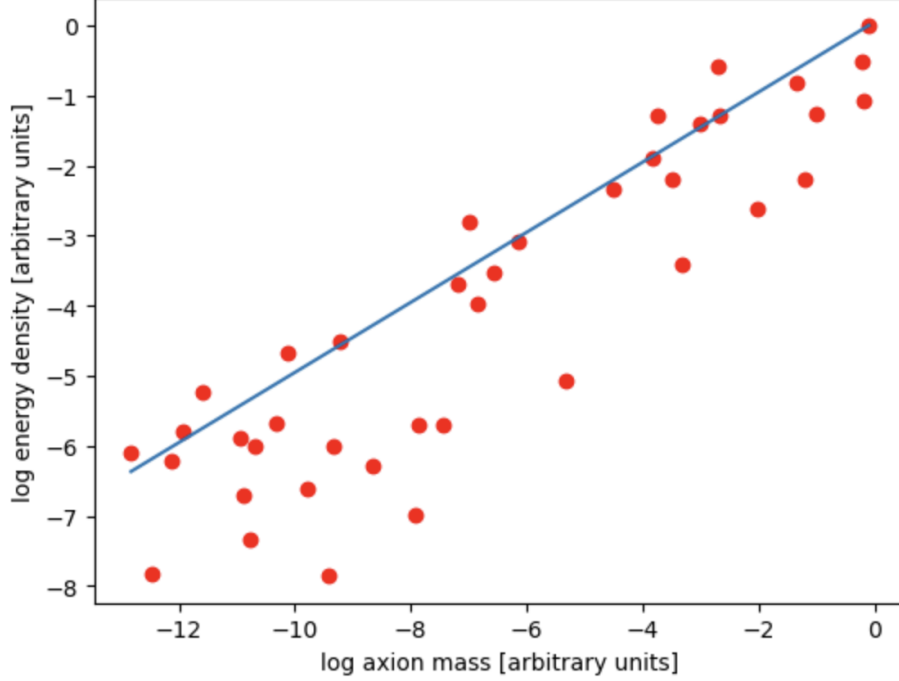
Figure 4: This is a simulation of the trivial case where each axion has its own instanton and the kinetic term is diagonal. The blue line represents the expectation that $\rho \propto m^{1/2}$. The horizontal axis is the log of the masses normalized against some arbitrary mass scale, while the vertical axis is the asymptotic relic abundance of each mass eigenstate normalized against some arbitrary energy density.

B Coupling to photons

In the interaction basis, there is a single mode that couples to photons, which we take to be θ_1 without loss of generality:

$$\mathcal{L}_{\text{int}} = \frac{\alpha_{\text{EM}}}{8\pi}\theta_1 \tilde{F}F = \frac{\alpha_{\text{EM}}}{8\pi}[\boldsymbol{R}\boldsymbol{F}^{-1}\boldsymbol{S}\boldsymbol{\phi}]_1 \tilde{F}F. \tag{31}$$

For each mass eigenstate, the coupling to the photon is then given by the corresponding coefficient in the first row of this matrix.

C Directions

1. Plot: $g_{a\gamma\gamma}\sqrt{\rho/\rho_{\text{DM}}}$ vs mass (like what we have in the friendship paper) for all the axions.
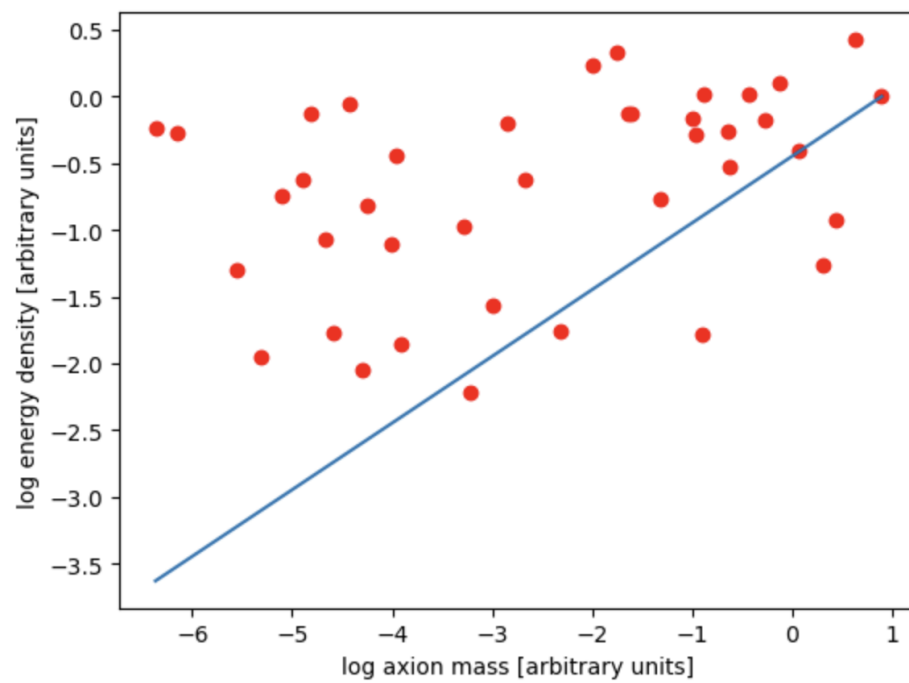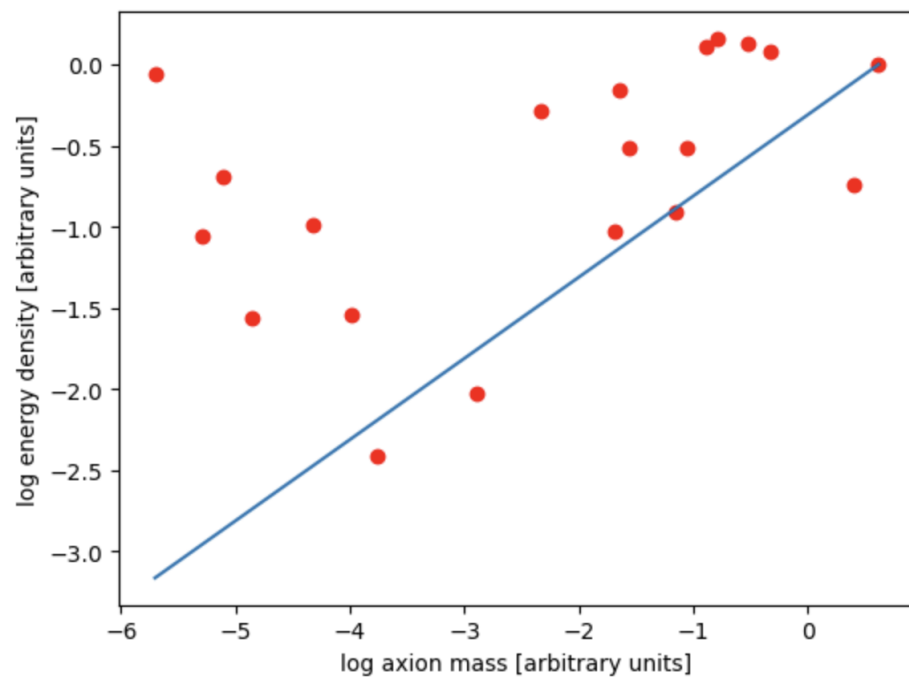
Figure 5: 40 axions!

Figure 6: 20 axions!

2. Theoretical expectation for thermalization rate of nonlinearly coupled chain of oscillators – extend to time-varying parameters with adiabatic approximation

3. Theoretical expectations for the distribution of energy density among the mass eigenstates

4. Effect of QCD Axion?

5. Statistics of the CC (what is the distribution of vacuum energies?)

D   Being more clever

Let us consider the algorithm more carefully, and consider how it may be improved.

1. Heavy modes are still kept. Can we drop them entirely using an analytical method without introducing numerical error? Perhaps we simply don't include them in the rotation matrix?

2. Light modes are included in the rotation matrix even though they don't need to be included. Perhaps only including the oscillating modes in the rotation matrix would improve numerical stability.

3. In order to avoid using arbitrary precision during numerical integration, consider why arbitrary precision may be considered necessary at all. It is only the case that arbitrary precision would be necessary if terms with a huge dynamical range entered into the light field equations of motion in the first place. These terms necessarily cancel – how can we eliminate them without losing precision? IDEA: Suppose there are N dynamical degrees of freedom – using arbitrary precision, precompute an N dimensional subvolume of the potential

References

[1] B. Freivogel, "Anthropic Explanation of the Dark Matter Abundance," *JCAP*, vol. 03, p. 021, 2010.

[2] A. Arvanitaki and S. Dubovsky, "Exploring the string axiverse with precision black hole physics," *Phys. Rev. D*, vol. 83, p. 044026, Feb 2011.