Two-body problem: in the previous formulation,

$$\min_{\phi} \sum_{t=1}^{T} \| \dot{\boldsymbol{r}}_t \cdot \bigtriangledown \phi \left(\boldsymbol{r}_t \right) \|_2^2 \tag{1}$$

This converged in less than 3 epochs and did not produce any meaningful $\phi()$. As found last week the solution found would simply ignore any inputs and return 0.

As discussed we experimented with adding a term to maximize the gradients of $\phi()$ yielding the minimization:

$$\min_{\phi} \sum_{t=1}^{T} \left(\| \dot{\boldsymbol{r}}_t \cdot \bigtriangledown \phi \left(\boldsymbol{r}_t \right) \|_2^2 + \frac{1}{\| \bigtriangledown \phi \left(\boldsymbol{r}_t \right) \|_2^2} \right) \tag{2}$$

However as $\bigtriangledown \phi()$ was highly variant solving this encountered issues with numeric stability and would underflow/overflow for various training hyper-parameters.

To combat the numerical stability we experimented with log of the norm of the gradient of phi:

$$\min_{\phi} \sum_{t=1}^{T} \left(\| \dot{\boldsymbol{r}}_t \cdot \bigtriangledown \phi \left(\boldsymbol{r}_t \right) \|_2^2 - \log \left(\| \bigtriangledown \phi \left(\boldsymbol{r}_t \right) \|_2^2 \right) \right) \tag{3}$$

This minimization again suffered from numerical stability.

Given our goal of minimizing the dot product $\dot{\boldsymbol{r}}_t \cdot \bigtriangledown \phi \left(\boldsymbol{r}_t \right)$ we minimized the cosine distance between $\dot{\boldsymbol{r}}_t$ and $\bigtriangledown \phi \left(\boldsymbol{r}_t \right)$ and instead of maximizing the gradient we simple enforce that $\| \bigtriangledown \phi(\boldsymbol{r}_t) \| = 1$ along the trajectory:

$$\min_{\phi} \sum_{t=1}^{T} \left(\left| \frac{\dot{\boldsymbol{r}}_t \cdot \bigtriangledown \phi \left(\boldsymbol{r}_t \right)}{\| f_t \|_2^2 * \| \bigtriangledown \phi(\boldsymbol{r}_t) \|_2^2} \right| - \left(1 - \| \bigtriangledown \phi \left(\boldsymbol{r}_t \right) \|_2^2 \right)^2 \right) \tag{4}$$

This objective results in a well defined gradient for minimization. Producing a phi that is constant across $\boldsymbol{r} \in T$. However this phi after convergence decreases consistently for each epoch across multiple planets. This may be a result of over-fitting however as it is only seen after 100k epochs for a small 50 unit fully connected neural network. I will post a histogram of phi as a function of training to validate these claims.

One of the potential concerns with the approach in (5) was unit-gradient magnitude was only enforced along the trajectory. We are currently testing the effect on regularizing the entire field to be unit length:

$$\min_{\phi} \sum_{t=1}^{T} \left| \frac{\dot{\boldsymbol{r}}_t \cdot \bigtriangledown \phi \left(\boldsymbol{r}_t \right)}{\| f_t \|_2^2 * \| \bigtriangledown \phi(\boldsymbol{r}_t) \|_2^2} \right| - \sum_{x \in R^4} \left(1 - \| \bigtriangledown \phi(x) \|_2^2 \right)^2 \tag{5}$$