In this problem, you will practice using automatic differentiation to construct the Fisher Information Matrix of a non-linear system and using said matrix to make informed decisions about sampling.

Consider the model $f(\vec{x}) = \begin{pmatrix} e^{-\theta_1 x_1} \tan \theta_2 x_2 \\ \cosh \theta_1 x_1 x_2 \end{pmatrix}$

(a) Use automatic differentiation to create the Jacobian of this function. Note that your result will be a programmatic object, not something that you can write out by hand

(b) The experiment that you want to test this model on has three possible sample points: 0, 0.5, and 1, but due to cost, you can only choose to sample at two points. Explore the behavior of your model by plotting the condition number of the Fisher Information Matrix as a function of $\vec{\theta}$. Use the following procedure:

- Create three plots, one for each pair of possible sampling points: (0,0.5), (0.5, 1), and (0, 1)
- Sample θ_1 from (-3, 3)
- Sample θ_2 from (-1.5, 1.5) to avoid singularities
- Use a sparse grid to sample θ_i to keep computational time low ($\sim$ 20 samples from each θ_i)
- Report the base-10 logarithm of the condition number

(c) What does a low condition number tell you about your model? Which of your plots has the greatest area with the lowest condition number? With that in mind, which set of sampling points is best for this model and why?