CSCI 145 Problem 22: Stretch the RoPE Clock

August 27, 2026

Submission Instructions

Please upload your work to Gradescope by **11:59pm on DATE**.

- Turn in a single PDF of *handwritten* solutions. Typed (including LaTeX) solutions will not be accepted.
- Problems are grouped into three-week **cycles**. Once per cycle, you will meet with the grader and present the solution to *one* problem from that cycle, selected at random when you arrive; bring your work and be ready to explain any of them.

Problem 22: Stretch the RoPE Clock 2 points

A model with even head dimension $d \geq 4$ was trained with RoPE frequencies

$$\theta_j = B^{-2j/d}, \qquad j = 0, \ldots, d/2 - 1,$$

and context length L. You must deploy it at length $4L$ without retraining. One engineer proposes increasing the base B just enough to quadruple the slowest clock's period. Another proposes leaving B fixed and replacing every position m by $m/4$ inside the rotations.

Derive both modified frequency ladders and compare them as context-extension rules. Your comparison must account for the fastest and slowest periods, local positional resolution as measured by the phase change between adjacent positions, and the change each rule makes to every relative-position score that the trained model saw for offsets $|m - n| \leq L$. Determine the new base required by the first proposal as a function of B and d.

Both rules preserve some structural properties of RoPE and disrupt others. Give a mathematically supported deployment recommendation, including one failure mode that cannot be resolved from the frequency formulas alone and therefore requires empirical validation.

Results you may use. You may quote $\mathbf{R}_m^\top \mathbf{R}_n = \mathbf{R}_{n-m}$, norm preservation of rotations, and the fact that a two-dimensional clock with angular frequency θ has period $2\pi/\theta$.