Foundation Models for Phase-Field Dynamics

The Allen-Cahn equation is a fundamental model in materials science that describes phase separation dynamics in binary systems:

$$\frac{\partial u}{\partial t} = \Delta_x u - \frac{1}{\epsilon^2}(u^3 - u), \quad t \in [0,1], \quad x \in [-1,1] \tag{6}$$

with periodic boundary conditions. The parameter ϵ controls the width of transition layers between phases, with smaller values leading to sharper interfaces. This equation exhibits rich dynamics, from smooth diffusion-dominated behavior at large ϵ to rapid phase separation with sharp interfaces at small ϵ.

In this task, you will develop a neural foundation model capable of solving the Allen-Cahn equation across different parameter regimes and initial conditions. Template code is provided for generating training data with various initial conditions drawn from three distributions: random Fourier series, Gaussian mixtures, and piecewise linear functions. The code also includes the numerical solver using `scipy.integrate.solve_ivp`.

Your task: Explore how neural architectures can learn and generalize across different dynamical regimes. Starting from the code templates from this **folder** (which you are **NOT** forced to use, it's just there to give you a starting point):

- Complete the data generation code to create training datasets with different ϵ values (e.g. 0.1, 0.05, 0.02) and initial condition types. The data should capture the full range of behaviors from smooth evolution to sharp interface dynamics (you can start from `allen-cahn-template.py`).
- Develop a time-dependent neural solver that can handle the entire trajectory of the solution. Consider how to effectively embed both the time dependency and the ϵ parameter in your architecture.
- Investigate the model's ability to generalize across:
  - Different ϵ values, including interpolation and extrapolation
  - Various initial condition types
  - Higher frequency components and sharper transitions than seen in training

The provided template includes functions for generating different types of initial conditions and solving the PDE. You will need to implement the core components, including the Allen-Cahn right-hand side and the initial condition generators.

Data generation specifics: Each training trajectory consists of 5 temporal snapshots that capture as much as possible of the dynamics before reaching a steady state, on a spatial grid of 128 points. Therefore both the time-scale t and ϵ are to be considered as hyper-parameters you need to tune to ensure capturing the full-range of the dynamics.

Note: Think carefully about your choices of ϵ and the expected behavior of solutions as $\epsilon \to 0$ – how do the dynamics change, in terms of the competing diffusive versus nonlinear effects?

Your final submission should include:

- The completed data generation code
- Your neural PDE solver implementation
- A report (maximum 2 pages) that includes:
  - Visualization of the generated data across different regimes
  - Analysis of your model's performance, including error metrics and convergence behavior
  - Discussion of generalization properties, supported by plots comparing predictions with true solutions
  - Investigation of how the model handles challenging cases, such as very small ϵ values or high-frequency initial conditions

Deliverables: For evaluation, focus on the relative L2 error between predicted and true solutions, but also consider qualitative physical aspects such as interface width.

Your report (max 2 pages long) should demonstrate understanding of both the mathematical properties of the Allen-Cahn equation and the capabilities and limitations of your neural approach.

Bonus Task: Stability Analysis

Consider the following stability theorem for the Allen-Cahn equation:

Theorem (Stability)**.** *Let $u \in H^1([0,T]; H^{-1}(\Omega)) \cap L^\infty([0,T]; H^1(\Omega))$ be a weak solution of the Allen-Cahn equation with $|u| \leq 1$ almost everywhere in $[0,T] \times \Omega$. Let $\tilde{u} \in H^1([0,T]; H^{-1}(\Omega)) \cap L^2([0,T]; H^1(\Omega))$ satisfy $|\tilde{u}| \leq 1$ almost everywhere in $[0,T] \times \Omega$, and $\tilde{u}(0) = \tilde{u}_0$, $f \in H^1([0,T]; W^{1\infty}_{loc}(\Omega))$ and solve*

$$(\partial_t \tilde{u}, v) + (\nabla \tilde{u}, \nabla v) = -\varepsilon^{-2}(f(\tilde{u}), v)$$

for almost every $t \in [0,T]$, all $v \in H^1(\Omega)$. Then we have

$$\sup_{t\in[0,T]} \|u - \tilde{u}\|^2_{L^2(\Omega)} + \int_0^T \|\nabla(u - \tilde{u})\|^2 dt \leq 2\|u_0 - \tilde{u}_0\|^2 \exp((1 + 2c_f \varepsilon^{-2})T).$$

Provide a complete proof of this stability result. Your proof should:

1. Explain the choice of test function for deriving the energy estimate
2. Show how to handle the nonlinear term $f(u)$
3. Use appropriate Gronwall-type arguments
4. Carefully track the dependence on ε

Hint: Consider the difference $w = u - \tilde{u}$ as a test function and analyze how the nonlinear term contributes to the energy estimate through appropriate bounds on f'.