# Fourier Continuation for Exact Derivative Computation in Physics-Informed Neural Operators

Haydn Maust, Zongyi Li, Yixuan Wang, Daniel Leibovici, Oscar Bruno, Thomas Hou, Anima Anandkumar
California Institute of Technology

## Introduction

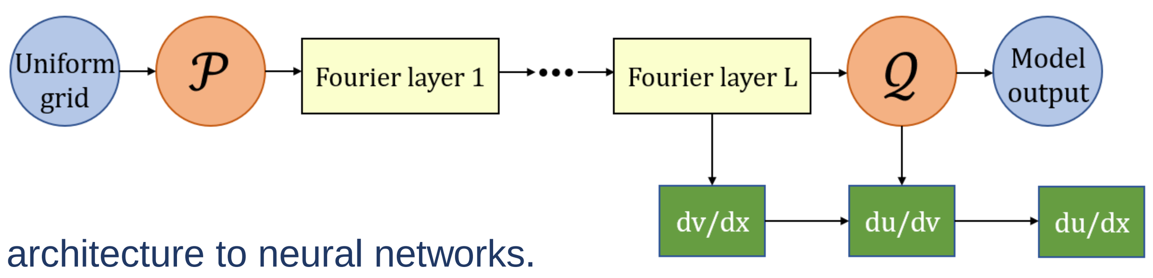
- Physics-informed neural operator (PINO) shows empirical promise for learning PDEs.
- Underlying Fourier neural operator (FNO) architecture improves over PINNs.
- FNO Fourier series representation enables exact gradient computation in frequency space.
- FNO has the expressivity to learn nonperiodic functions, but derivative computations are ill-conditioned, leading to poor optimization behavior.
- We use Fourier continuation (FC) for exact gradient computation on nonperiodic functions.
- Three PINO models using FC are tested on a 1D problem and outperform padded PINO.

## Physics-Informed Learning

$$\mathcal{L}_{\text{pde}}(a, u_\theta) = \int_D |\mathcal{P}(u_\theta(x), a(x))|^2 \mathrm{d}x + \alpha \int_{\partial D} |u_\theta(x) - g(x)|^2 \mathrm{d}x$$

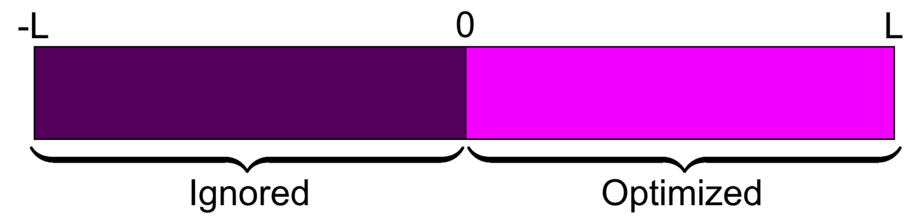
Physics-informed learning uses the differential equation itself to define loss, enabling learning with no solution data.

## Neural Operators




- Alternative architecture to neural networks.
- PINO is neural operator trained on physics-informed loss.
- Each Fourier layer consists of a matrix, an integral kernel operator, and an activation function.
- For FNO, integral kernel operators are linear transformations in frequency space.

## Nonperiodic Functions & Ill-Conditioning




- Want to use Fourier series to take derivatives with high accuracy.
- Fourier series is only well-defined for periodic functions (Gibbs phenomenon in nonperiodic case).
- Want to extend nonperiodic model output to a periodic function.
- Padding (extend domain and ignore the extra part during optimization) is a partially observed optimization problem, which is ill-conditioned because we use discrete Fourier transforms.
- Alternative: deterministic continuation methods.
- Model outputs on the base domain, then we extend afterward using a fixed algorithm.

## Fourier Continuation

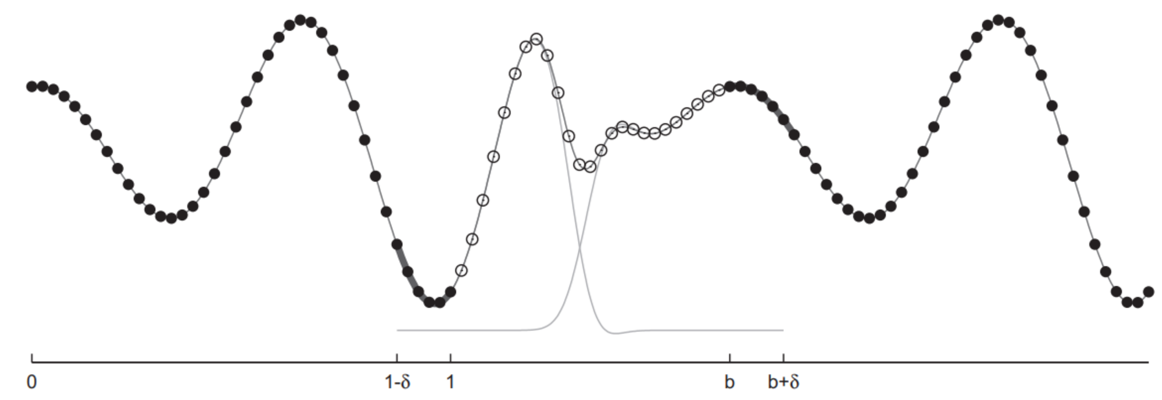
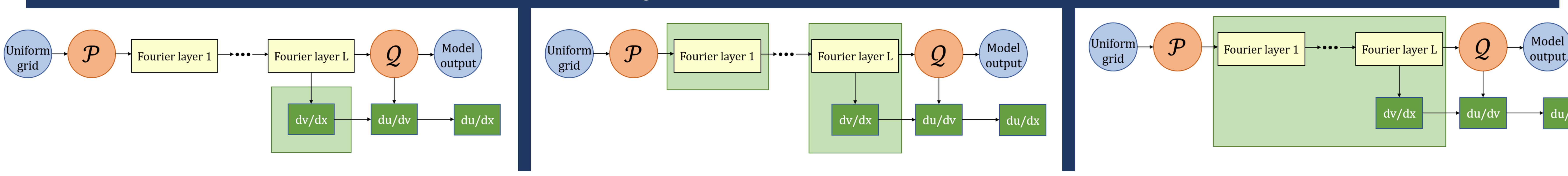



Image adapted from [1].

- We use FC-Gram continuation function.
- FC-Gram interpolates gridpoints near the boundary by trigonometric polynomials that are periodic on an extended domain, which produces a continuation.
- This interpolation has very small error when compared against the original function on the base domain.

## Adding Fourier Continuation to PINO




Three candidate models (from left to right: models 1, 2, 3) that incorporate Fourier continuation into PINO. Light green highlighted regions indicate where Fourier continuation is used. Entering a green highlighted region means applying FC, and exiting a region means truncating to the base domain.

## Baseline Models

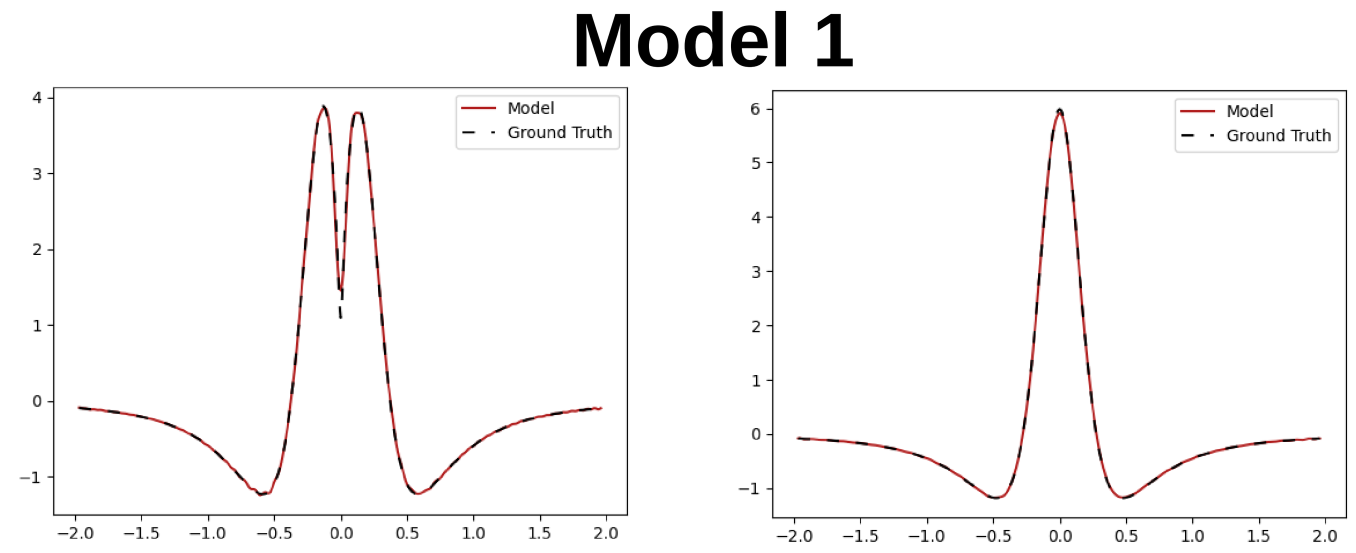
For comparison, we also show results of two baseline models. The first uses a padded domain, and the second uses PINO with the exact gradient method with no domain extension at all.

## Empirical Results: Self-Similar Burgers' Equation

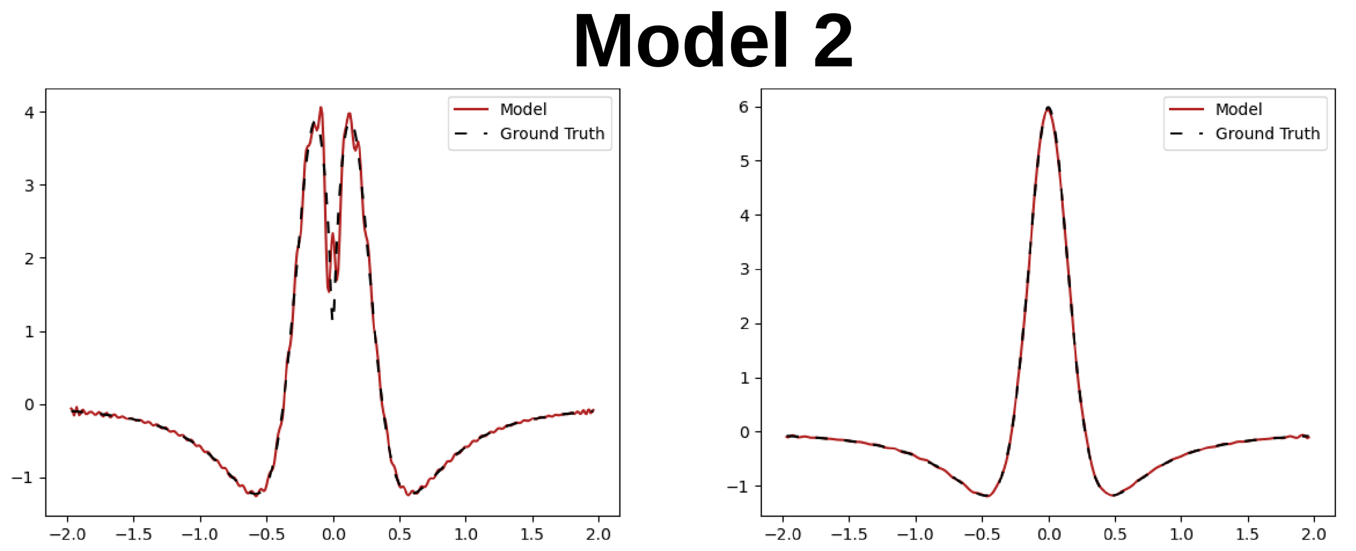
- Train each model on the self-similar Burgers' equation (shown on the right).
- This is a 1D ODE with nonperiodic solution.
- For some values of λ, the solution is nonsmooth: third-order derivative has a cusp at zero.
- Each pair of figures shows the results of a model trained on the self-similar Burgers' equation with λ=0.4 (left) and λ=0.5 (right). λ=0.4 corresponds to a nonsmooth solution.
- To make the performance differences visually apparent, these plots show the third derivative of each model output.

$$-\lambda U + ((1+\lambda)y + U)U_y = 0$$

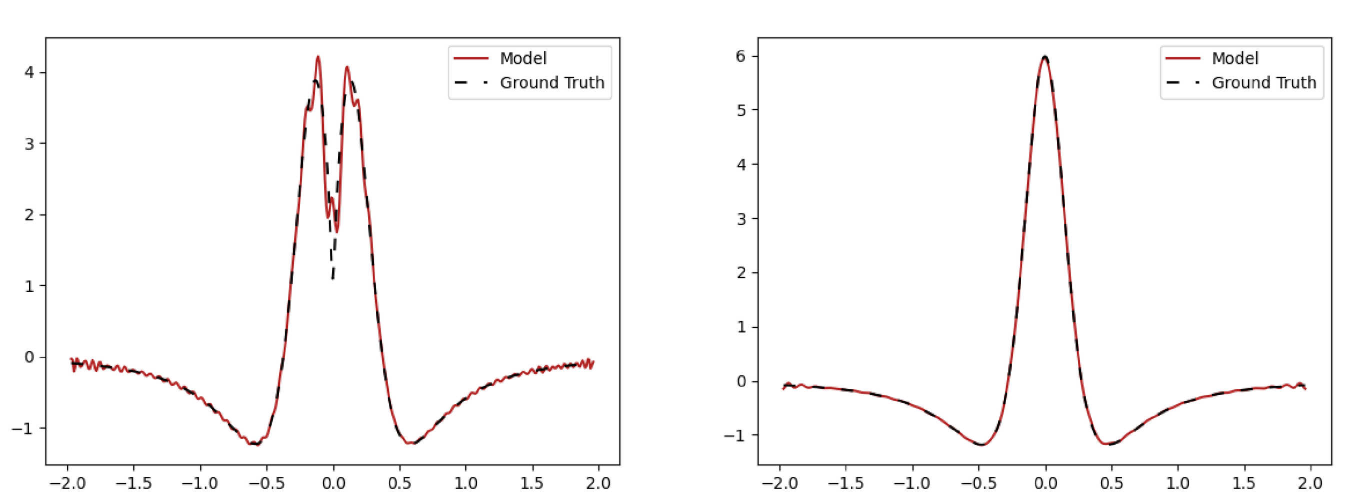
### Model 1



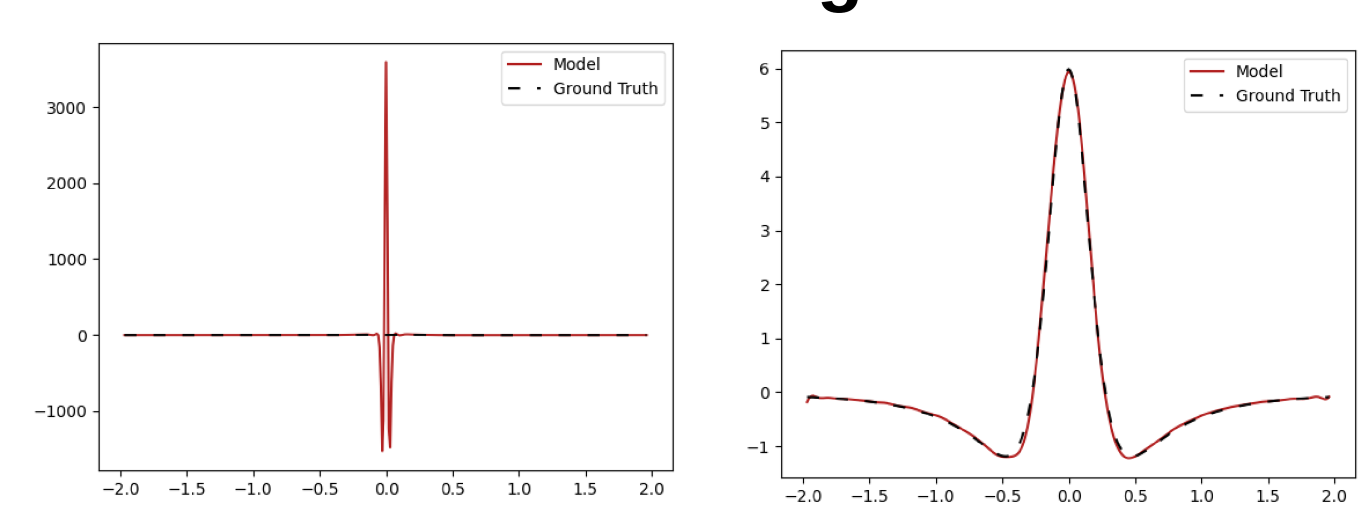

### Model 2



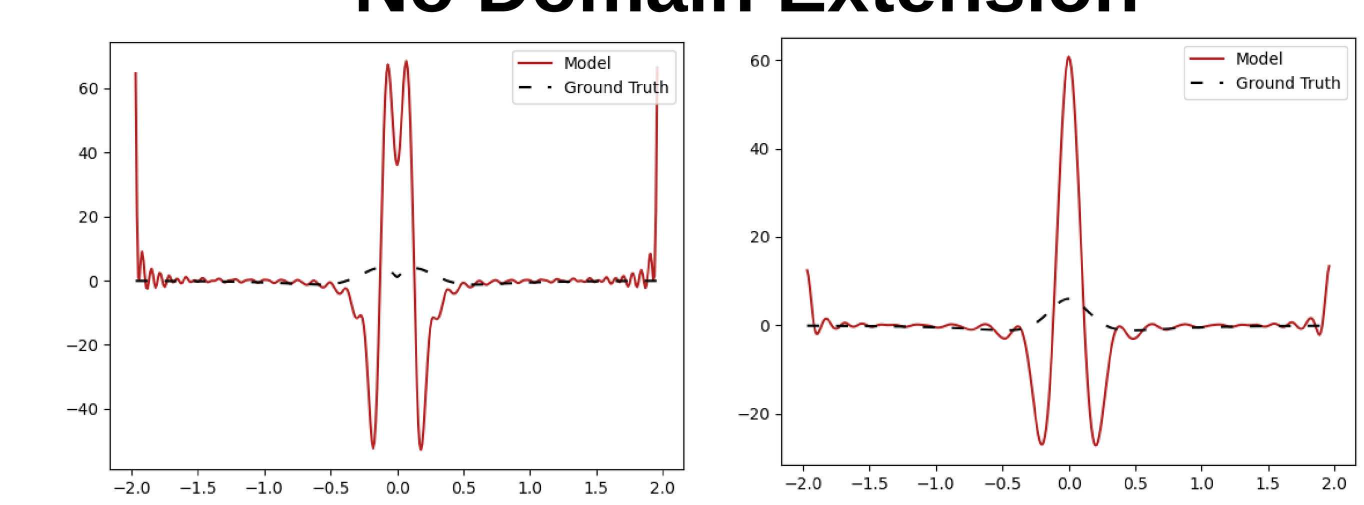

### Model 3




### Padding




### No Domain Extension




From both the plots and the loss values shown in the table, model 1 performs the best on both λ=0.4 and λ=0.5.

Theoretical explanation: In models 2 and 3, the Fourier layers of PINO act on an extended domain, but only the base domain [-2, 2] is observed during optimization. Thus models 2 and 3 produce partially observed optimization problems, and hence are ill-conditioned, so we expect gradient descent to be comparatively ineffective.

| | Equation loss | | $L^2$ loss | |
|---|---|---|---|---|
| | $\lambda = 0.4$ | $\lambda = 0.5$ | $\lambda = 0.4$ | $\lambda = 0.5$ |
| Model 1 (FC only at the end) | $10^{-8.4}$ | $10^{-7.7}$ | $10^{-6.1}$ | $10^{-6.0}$ |
| Model 2 (FC in each Fourier layer) | $10^{-6.5}$ | $10^{-7.3}$ | $10^{-6.1}$ | $10^{-6.0}$ |
| Model 3 (FC after $\mathcal{P}$, truncation before $\mathcal{Q}$) | $10^{-5.5}$ | $10^{-6.6}$ | $10^{-6.1}$ | $10^{-6.0}$ |
| Baseline 1 (PINO on padded domain) | $10^{-2.5}$ | $10^{-3.9}$ | $10^{-5.3}$ | $10^{-5.6}$ |
| Baseline 2 (PINO with no continuation) | $10^{+2.3}$ | $10^{+2.3}$ | $10^{-0.5}$ | $10^{-0.5}$ |

## References

[1] Albin, N. & Bruno, O. P. A spectral FC solver for the compressible Navier–Stokes equations in general domains I: Explicit time-stepping. *Journal of Computational Physics* **230**, 6248–6270 (2011).
[2] Wang, Y., Lai, C.-Y., Gómez-Serrano, J. & Buckmaster, T. Asymptotic self-similar blow up profile for 3-D Euler via physics-informed neural networks. arXiv:2201.06780 [physics] type: article (Mar. 2022).