arXiv:2302.13368v1 [cs.CE] 26 Feb 2023

相场 DeepONet：用于快速数值模拟由自由能泛函梯度流所支配的图案形成的物理信息深度算子神经网络

预印本

Wei Li [a, b] Martin Z. Bazant [b, c] Juner Zhu [a, *]

[a]东北大学机械与工业工程系[b]麻省理工学院化学工程系[c]麻省理工学院数学系

2023年2月28日

摘要

科学机器学习的最新进展为图案形成系统的建模带来了新的启示。然而，真实图案的数值模拟仍然需要高昂的计算成本，而利用大型图像数据集有望缓解这一问题。物理信息机器学习和算子学习是针对此类应用新近兴起且颇具前景的两个概念。在此，我们提出“相场 DeepONet”，这是一种物理信息算子神经网络框架，用于预测由自由能泛函的梯度流控制的系统的动态响应。用于验证该方法可行性和准确性的示例包括 Allen-Cahn 方程和 Cahn-Hilliard 方程，它们是化学混合物非平衡热力学反应性相场模型的特殊情形。该方法通过将最小化运动格式纳入框架来实现这一目标；该格式能够优化并控制系统总自由能的演化，而不是直接求解控制方程。训练后的算子神经网络可以充当显式时间步进器，以当前状态作为输入并输出下一状态。这有望促进图案形成动力系统的快速实时预测，例如相分离锂离子电池、乳液、胶体显示器或生物图案。

关键词 物理信息机器学习 · 深度算子神经网络 · 相场方法 · 最小化运动格式 · Allen-Cahn 和 Cahn-Hilliard 方程

1 引言

机器学习（ML）已在许多领域取得巨大成功，尤其是在计算机视觉 [1] 和自然语言处理 [2]方面。然而，对于科学问题，人们往往认为机器学习算法是难以解释的黑箱，也缺乏物理定律提供的严格依据。近年来，科学机器学习（SciML）这一概念应运而生，并迅速受到广泛关注 [3]。它旨在使传统机器学习具备领域感知能力、可解释性和鲁棒性。从某种意义上说，SciML 正在革新计算科学领域，并已应用于多个科学学科，例如：机器学习增强的多物理场与多尺度建模[4, 5, 6], 、机器学习辅助的快速在线预测和引导式数据采集 [7], , 以及复杂系统的最优决策 [8, 7]。在各种科学机器学习架构中，物理信息机器学习（PIML）和算子学习是两个典型代表。PIML 将已知物理知识引入机器学习算法，因此所需的数据集规模更小，有时甚至无需实验数据 [9]。PIML 是一个通用概念，可通过

*通讯作者。电子邮件：w e.li@northeastern.edu（W.L.），bazant@mit.edu（M.Z） .B），j.zhu@northeastern.edu（J.Z.）

多种策略实现 [10]。考虑一般机器学习算法的三个组成要素，物理定律可相应地融入：1）训练数据，例如由基于第一性原理的模拟生成的数据；2）机器学习模型，例如旨在体现某些物理原理（对称性、正定性、层级关系等）的神经网络；以及 3）训练策略，例如将物理定律纳入其中的损失函数。在这些不同方法中，研究最为广泛的一种是物理信息神经网络（PINN） [11, 12, 13], ，它将已知物理知识，即控制方程、初始条件（ICs）和边界条件（BCs），以残差的形式纳入损失函数。研究还表明，PINN 能够处理由纳维–斯托克斯方程控制的流体动力学中的正问题（求解方程）和逆问题（识别参数） [11]。自提出以来，PINN 已广泛应用于各种领域。感兴趣的读者可参阅 [9] ，了解全面综述。除了这些成功应用外，PINN 还经过扩展，以适用于不规则 [14]和多个区域 [15, 16],，利用修改后的神经网络施加硬约束[17], ，引入自适应激活 [18], ，在损失函数中加入梯度增强项 [19]或基于能量的项[20, 21], 等。

尽管取得了巨大成功，物理信息神经网络面临的一个突出挑战是如何高效地确定或优化损失函数中的超参数。现有大多数研究采用试错方案，导致训练耗时。为应对这一挑战，Psaros 等人 [22]最近提出了一种元学习框架，用于离线优化损失函数。同时，还可以通过减少损失项的总数来避免这一问题。一种常见方法是施加硬约束 [17], ，即修改神经网络，使其输出始终满足特定的边界条件和初始条件。另一种方法是利用内在包含偏微分方程和边界条件的能量原理或变分原理。例如，使用物理信息神经网络求解带有诺伊曼型边界条件 $\partial_{\mathbf{n}} u = 0$ 的拉普拉斯方程 $\nabla^2 u = 0$，至少需要两个损失项。或者，该问题等价于求泛函 $\mathcal{J} = \int_\Omega 0.5(\nabla u)^2 \, \mathrm{d}V$（即 $\delta\mathcal{J} = 0$）的最小值，并可将其作为唯一的损失项。通过这种方式，可以减少损失项的数量。从数学上看，该例只是更一般的欧拉-拉格朗日方程（见 A.1）的一个特例；需要注意的是，泛函中的导数阶数低于偏微分方程中的导数阶数，这进一步提高了训练效率。

现有少数研究探索了将能量作为损失函数的策略。E等人 [23] 提出了一种 DeepRiz 神经网络，通过最小化泛函 $\mathcal{J} = \int_\Omega \left[0.5 \cdot (\nabla u)^2 - f(x) \cdot u(x)\right] \mathrm{d}x$，求解具有齐次本质边界条件 ($u(x) = 0, x \in \partial\Omega$) 的泊松方程 ($-\nabla^2 u(x) = f(x)$)；该泛函也是欧拉-拉格朗日方程的一个特例。随后，Wang等人将这一框架扩展至复杂边界几何和多个区域中的非齐次本质边界条件 [15]。另一个探索方向是固体力学中的最小势能原理，该原理指出，在外载荷作用下，固体区域的变形将遵循使总势能最小的路径 [21, 20]。对于准静态弹性响应，总势能 $\mathcal{T}$ 由弹性应变能 $U = \int_\Omega 0.5\sigma : \varepsilon \, \mathrm{d}V$ 和功势 $W = -\int_{\partial\Omega} f \cdot u \, \mathrm{d}S$ 组成。最小化 $\mathcal{T} = U + W (\delta\mathcal{T} = 0)$ 等价于求解相应的欧拉-拉格朗日方程（力平衡）$\nabla \cdot \sigma = 0$，并满足边界条件 $\sigma \cdot \mathbf{n} = 0$。在作者此前的工作中，我们将势能引入损失函数，以实现这一原理，并预测了弹性板的变形 [20]。我们还将这一基于能量的框架与原始残差型物理信息神经网络进行了比较，发现由于损失函数中的导数阶数较低、超参数较少，基于能量的框架在训练时间方面效率更高，而两者的精度相当。值得指出的是，残差型物理信息神经网络是一个更通用的框架，而基于能量的框架仅适用于由能量原理或变分原理支配的系统。据作者所知，以上已有研究仅探索了准静态条件下的简单线性系统。本研究旨在探究高度非线性、耦合储能系统的动力学，并利用变分原理构建物理信息机器学习框架。

算子学习是近年来出现的另一种颇具前景的科学机器学习（SciML）技术；它学习从一个函数到另一个函数的映射，例如序列到序列映射和图像到图像映射。相比之下，许多广泛使用的网络架构，例如全连接神经网络（FNN）和卷积神经网络（CNN），属于有限维算子，可将一个离散信号或图像映射为另一个离散信号或图像。近期，有研究提出了一些用于学习无限维映射的新型架构，例如深度算子网络（DeepONet） [24] 和傅里叶算子网络（FNO） [25]。关于使用神经网络进行算子学习的综合性综述可参见 [25]。本研究将重点关注 Lu 等人开发的 DeepONet 方法[24]。迄今为止，已有研究证明 DeepONet 比 FNN 具有更好的逼近能力和泛化能力，因此已被广泛应用于许多领域。DeepONet 的优势之一在于能够将边界条件（BCs）或初始条件（ICs）作为输入，这使得从理论上讲，可以训练一个适用于所有场景的网络。这意味着网络训练完成后，无需额外训练，便可用于求解具有不同边界条件和/或初始条件的新问题。当边界条件和/或初始条件可能发生变化或存在较大不确定性时，这一特性在相关应用中尤其有用。

科学机器学习（SciML）的这些新进展为储能系统（ESSs）的建模带来了新的启示，尤其是对锂离子电池而言。由于这类系统具有高维特性（例如涉及多种材料、多个尺度和多个物理场），开发完整的基于物理的模型或对大型数据集进行全面解释都十分繁琐。开发一种能够结合部分已知物理知识与小规模数据集的统一物理信息机器学习计算框架，具有很强的吸引力，也十分必要。这里的根本挑战在于数据丰富性与物理定律充分性之间的权衡。在电极或电池单体层面，实验数据相对容易获取，但物理规律大多隐藏在数据之中。因此，可以应用纯数据驱动机器学习算法来预测电池的性能 [7] 和寿命 [26] 。当数据获取成本高昂或数量有限时，例如获取需要数年才能完成的数千次循环中的电池退化数据，一些研究提出了基于物理信息神经网络（PINNs）的框架，用于估计电池单体的状态 [27, 28],、识别电池参数 [29],、预测寿命 [30], 以及识别电极和电池单体层面的退化模式 [31, 32, 33] 。另一方面，尽管在微观尺度上已经发展出许多基础性的电化学-力学物理理论 [34, 35]，但在活性颗粒尺度获取实验数据仍然成本高昂且困难。基于物理的模型通常用于描述单颗粒图案形成，并将结果外推至多孔电极 [36, 37, 38], ，但求解这些模型往往非常耗时。如前文所述，物理信息机器学习（PIML）具有产生高效代理模型或降阶模型的潜在优势，因为机器学习算法在训练完成后的推理速度很快。例如，一些研究使用物理信息神经网络（PINNs）求解相场方法中的两个方程，即 Allen-Cahn 方程和 Cahn-Hilliard 方程 [39, 40, 41], ，相关内容将在第 2 节中详细阐述。

在储能系统（ESS）中应用物理信息机器学习（PIML）的另一个原因是，本构关系和材料常数的确定颇具挑战性。针对电化学非平衡热力学相场模型背景下的全场图像数据大型数据集解读，已有许多先进算法被开发出来 [36]。例如，Zhao等人利用偏微分方程约束优化，根据扫描隧道X射线显微镜的操作过程成像，学习磷酸铁锂纳米颗粒中驱动相分离的物理规律 [42]。Deng等人采用类似方法，根据活性颗粒的微米X射线断层成像和衍射图像，学习本征应变随锂嵌入而变化的本构定律 [43]。这一优化过程通常耗时较长，而物理信息机器学习有望加快识别速度。

如今，整个科学界已认识到，将物理规律与数据整合到统一框架中，作为一种高层次愿景，具有重要价值。储能领域是率先探索这一方向的领域之一。例如，美国能源部（DOE）是首个提出科学机器学习（SciML）概念的联邦机构 [3]。当前仍面临的挑战是找到切实可行的实施方式。要为所研究的系统构建恰当的机器学习架构，首先理解其物理规律始终至关重要。前文提到的 Allen-Cahn 方程和 Cahn-Hilliard 方程是化学系统建模中的关键方程。它们分别能够通过变分定义的化学势描述非守恒和守恒序参量的动力学。这两个方程均可通过变分方法推导得出；这类方法已被广泛确立，可用于求解复杂耦合非线性系统的控制方程 [44]。更具体地说，Allen-Cahn 方程和 Cahn-Hilliard 方程是梯度流的两种特殊情形；梯度流需要确定并构造适当的自由能和内积，以便将动力学纳入变分框架 [44, 45, 46]。梯度流可应用于许多物理过程，包括扩散、相分离、微观结构演化等。因此，为梯度流构建机器学习框架将有益于广泛的应用。正如前文所述，机器学习能够自然地用于求解变分问题：我们可以通过最小化自由能泛函这一损失函数，利用机器学习模型近似求得解。

本研究提出“相场 DeepONet”的概念，将其作为一种适用于由自由能泛函的梯度流所支配的动力系统的通用神经网络框架，并利用基于能量的损失函数、深度算子网络和物理信息学习。本文结构如下：第 2 节介绍相场方法和梯度流理论；第 3 节介绍相场 DeepONet 框架，该框架将最小化运动格式融入物理信息深度算子神经网络；第 4 节研究三种不同的动力系统，包括线性弛豫动力学、Allen-Cahn 动力学和 Cahn-Hilliard 动力学，以验证所提出的框架。

2 相场方法与梯度流

2.1 相场方法

相场方法在材料科学中得到了广泛应用，因为它能够追踪微观结构演化、晶粒生长与粗化、裂纹扩展等。与其他尖锐界面模型不同，相场模型利用相场变量以扩散的方式处理界面，从而能够连续地将整个区域及其所有界面描述为一个整体。相场变量分为两类，即守恒场和非守恒场。两者的演化（或

动力学）均由系统的总自由能 $\mathcal{F}$ 及其关于场变量的变分导数决定，后者可视为扩散化学势。

对于单个场变量 ϕ，范德瓦耳斯 [47] 以及卡恩和希利亚德 [48],提出的非均匀系统标准自由能泛函定义为:

$$\mathcal{F} = \int_{\Omega} [f(\phi) + \frac{1}{2}\kappa_{\phi}(\nabla\phi)^2] \, \mathrm{d}x, \tag{1}$$

其中， $f(\phi)$ 是均匀自由能密度，右侧第二项表示相界面处的梯度能，梯度系数为 κ_{ϕ}。通常，场变量会沿着自由能持续降低的方向演化。对于守恒场变量，其动力学可以表示为梯度驱动通量的守恒定律。由此可得到 Cahn-Hilliard 方程:

$$\frac{\partial\phi}{\partial t} = \nabla \cdot M\nabla\frac{\delta\mathcal{F}}{\delta\phi}, \tag{2}$$

其中， M 是输运系数（迁移率与浓度场变量 [36] 的乘积），而泛函导数（扩散化学势）为:

$$\frac{\delta\mathcal{F}}{\delta\phi} = \frac{\partial f}{\partial\phi} - \kappa_{\phi}\nabla\cdot\nabla\phi. \tag{3}$$

对于非守恒场变量，有 Allen-Cahn 方程:

$$\frac{\partial\phi}{\partial t} = -M\frac{\delta\mathcal{F}}{\delta\phi}, \tag{4}$$

其中，泛函导数也由式 3 给出。艾伦-卡恩方程可视为产生场变量的反应的线性化模型，该反应与亲和力成正比，即与相对于外部储库的扩散化学势差成正比 [49, 50]。

2.2 梯度流的数学

建立系统理论或控制方程的一种常见方法，是从基于实验数据的本构关系出发。随后，需要检验这些理论是否与热力学定律一致。另一方面，变分方法从热力学出发，因此推导出的理论始终与热力学定律一致。此外，变分方法能够更容易地处理极端各向异性、不可微性和非线性。本节将回顾一个通用的数学框架，用于将相场模型推导为自由能泛函的梯度流。

热力学第二定律指出，系统的总自由能始终会降低。因此，应以适当的方式构造描述系统中场变量演化的方程，以保证总自由能单调降低。一种方法是采用梯度流，其可描述为:

$$\frac{\mathrm{d}}{\mathrm{d}t}\, u(x,t) = -\nabla\mathcal{F}(u), \tag{5}$$

其中， u 是随时间演化的场变量（依赖于空间和时间）； $\mathcal{F}$ 是关于 u的光滑凸能量泛函。 $\nabla F(u)$表示 $\mathcal{F}$ 关于 u的泛函梯度。从物理角度看，该方程表明，场变量 u 沿着自由能下降最快的方向演化。泛函梯度是这一动力过程的驱动力。从数学角度看，类似于多变量函数的方向导数，泛函梯度 $\nabla\mathcal{F}$ 与泛函导数 $\delta\mathcal{F}/\delta u$ 通过内积 $<.,\ .>$相关，

$$<\nabla\mathcal{F}, v>= \frac{\mathrm{d}\mathcal{F}(u+v\cdot t)}{\mathrm{d}t}\,|_{t=0} = \int_{\Omega}\frac{\delta\mathcal{F}}{\delta u}v\,\mathrm{d}x, \tag{6}$$

其中， v 是任意函数，可视为流场； t 是一个短时间间隔， $v\cdot t$ 也称为u的变分。

因此，泛函梯度不仅取决于自由能泛函（其泛函导数），还取决于内积（或距离度量）的构造。

对于定义为 $\mathcal{F} = \int_\Omega F(x, u(x), \nabla u(x))\,\mathrm{d}x$ 的能量泛函，其泛函导数可通过以下方式确定（见 A.2）

$$\frac{\delta \mathcal{F}}{\delta u} = \frac{\partial F}{\partial u} - \nabla \cdot \frac{\partial F}{\partial \nabla u}. \tag{7}$$

为了求得泛函梯度，我们还需要构造一个内积。本研究中，我们引入了两种不同的内积。第一种是加权 L^2 内积，

$$< f, g >_{L^2,M} = \int_\Omega \frac{f(x) \cdot g(x)}{M}\,\mathrm{d}x, \tag{8}$$

其中， M 为权重； $f(x)$ 和 $g(x)$ 是两个任意函数。相应的 L^2 范数为

$$|f|_{L^2,M} = \sqrt{\int_\Omega \frac{f^2(x)}{M}\,\mathrm{d}x}. \tag{9}$$

第二种是带权重的 H^{-1} 内积，定义为

$$< f, g >_{H^{-1},M} = \int_\Omega \nabla\phi_f \cdot M \nabla\phi_g\,\mathrm{d}x, \tag{10}$$

其中， ϕ_f 是满足诺伊曼边界条件的泊松方程的解，

$$\begin{cases} \nabla^2 \phi_f = f(x), & \text{in } \Omega, \\ \partial_{\boldsymbol{n}} \phi_f = 0, & \text{on } \partial\Omega. \end{cases} \tag{11}$$

注意，当 $\int_\Omega f\,\mathrm{d}x = \int_\Omega \phi_f\,\mathrm{d}x = 0$ 时， ϕ_f 有唯一解。相应的 H^{-1} 范数可由下式得到

$$|f|_{H^{-1},M} = \sqrt{\int_\Omega \nabla\phi_f \cdot M \nabla\phi_f\,\mathrm{d}x}. \tag{12}$$

接下来，我们可以求得泛函梯度。对于 L^2 内积，由式 6 和式 9 可得

$$< \nabla\mathcal{F}, v >_{L^2,M} = \int_\Omega \frac{\nabla\mathcal{F} v}{M}\,\mathrm{d}x = \int_\Omega \frac{\delta\mathcal{F}}{\delta u} v\,\mathrm{d}x. \tag{13}$$

由于 v 是任意函数，为确保等式始终成立，我们有，

$$\nabla\mathcal{F} = M\frac{\delta\mathcal{F}}{\delta u}. \tag{14}$$

类似地，对于 H^{-1} 内积（式 6 和式 12），我们有，

$$\nabla\mathcal{F} = -\nabla \cdot M\nabla\frac{\delta\mathcal{F}}{\delta u}. \tag{15}$$

将上述两个泛函梯度（方程（14）和方程（15））代入方程（5），即可分别得到 Allen-Cahn 方程和 Cahn-Hilliard 方程。

2.3 最小化运动格式

我们已经展示了如何利用梯度流理论，以数学方式推导相场方程（即 Allen-Cahn 方程和 Cahn-Hilliard 方程）。为了预测由梯度流支配的系统的动态响应，我们可以直接使用有限元（FE）或有限差分（FD）方法求解这些方程。或者，我们也可以利用梯度流的一个重要特征。给定一个固定的小时间步 $\tau > 0$、一个光滑且凸的泛函 $\mathcal{F}(u)$，我们可以从初始条件 u_0^τ出发，通过以下迭代格式找到 u, $[u_1^\tau, u_2^\tau, \ldots, u_n^\tau]$的时间序列；该格式称为最小化运动格式，

$$u_{k+1}^\tau \in \text{argmin}_u \left[\mathcal{F}(u) + \frac{d^2(u, u_k^\tau)}{2\tau} \right], \tag{16}$$

其中， u_k^τ 和 u_{k+1}^τ 表示时间步 k 和 $k+1$时的场分布； d（.,.）表示两个函数之间的距离。通过这种迭代最小化格式得到的时间序列实际上就是相应偏微分方程的解。对于上述最小化问题，我们知道 u_{k+1}^τ 是以下方程的解

$$\frac{\partial}{\partial u} \left[\mathcal{F}(u) + \frac{d(u, u_k^\tau)}{2\tau} \right] = \nabla \mathcal{F}(u) + \frac{d(u, u_k^\tau)}{\tau} = 0, \tag{17}$$

由此可得

$$\frac{d(u, u_k^\tau)}{\tau} = -\nabla \mathcal{F}(u). \tag{18}$$

该方程是方程（5）的离散时间隐式欧拉格式。因此，我们可以通过最小化运动格式获得解，而无需直接求解偏微分方程。

3 相场 DeepONet：用于梯度流的物理信息深度算子神经网络

本节提出相场 DeepONet，这是一种融合上述最小化运动格式的物理信息深度算子神经网络框架，用于求解自由能泛函的梯度流。该框架包含两个重要组成部分：1）算子学习，旨在学习从一个函数到另一个函数的映射。在本研究中，我们利用这一概念学习场变量分布从当前时间步到下一时间步的映射；2）物理信息机器学习，将已知物理规律融入机器学习框架。我们直接利用式 16 作为损失函数，以实现梯度流的物理规律。下面将进一步解释这两个方面。

3.1 深度算子神经网络（DeepONet）

DeepONet 最初由 Lu 等人提出 [13]。如图 1a 所示，它是一种具有两个子网络的高层网络结构，即“分支”网络和“主干”网络。主干网络以坐标 x 作为输入，而分支网络以函数 u 作为输入。最终输出通过将两个网络的输出相乘得到，

$$\mathcal{G}(x, u) = \sum_{k=1}^{p} b_k(u) t_k(x) + b_0, \tag{19}$$

其中， b_k和 $t_k(k = 1, 2, \ldots, p)$分别是分支网络和主干网络的输出； p是两个子网络的输出数量。

关于 DeepONet，有几个重要特征需要注意。首先，由于主干网络以坐标作为输入，因此输出是连续的，这意味着我们可以获得任意位置的预测结果。更重要的是，通过自动微分，可以轻松估计输出相对于输入的梯度。这对于构建物理信息损失函数至关重要。其次，它具有很高的灵活性。这种结构的本质是将向量输入和函数输入分别划分到两个子网络中。它可以根据不同的应用进行定制。一方面，特征

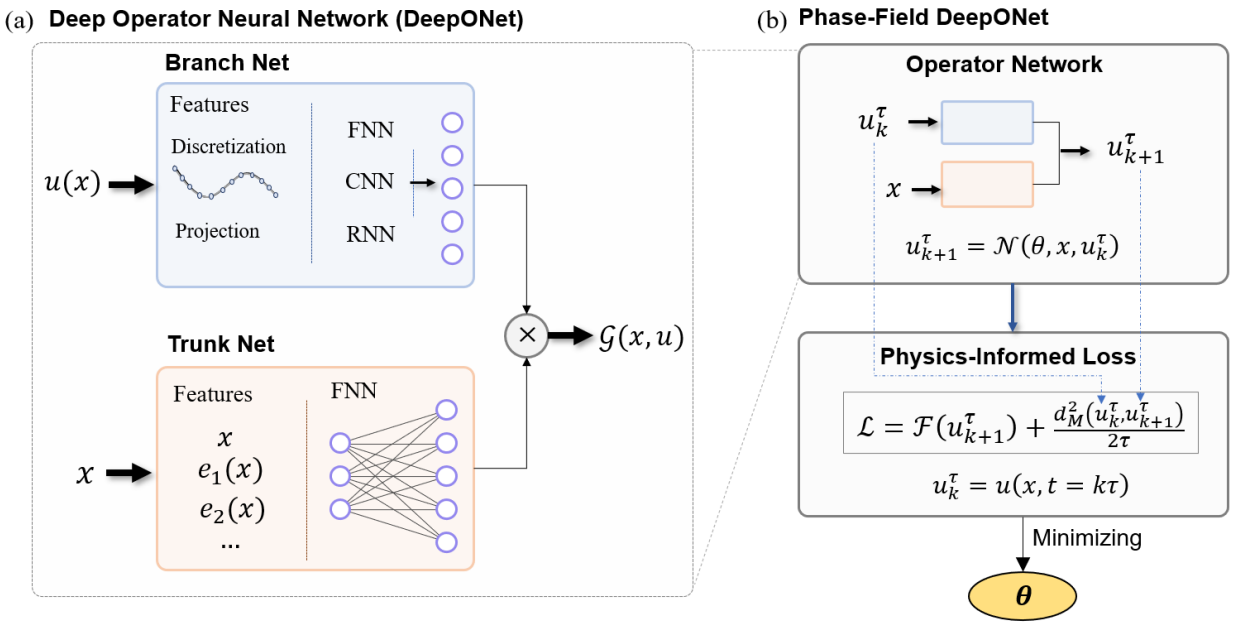

图 1：DeepONet (a) 和采用基于能量的损失函数的物理信息 DeepONet (b) 的示意图。

可以在输入上进行扩展。对于分支网络，我们不必使用函数的离散点向量，而是可以从函数中提取其他任意特征作为输入，例如频域中的幅度和相位。另一方面，子网络可以采用任何类型的神经网络，例如前馈神经网络、卷积神经网络、循环神经网络（RNN）等。

DeepONet 能够将一个函数映射为另一个函数。在实际应用中，输入函数可以是初始条件或边界条件，输出则可以是某个随机时刻的解。在这里，我们旨在预测场变量在任意给定时间步的分布。为此，我们提出将场变量当前的分布作为输入，将其下一时间步的分布作为输出（即映射 $u_k^\tau \to u_{k+1}^\tau$），可表示为：

$$u_{k+1}^\tau = \sum_{k=1}^{p} b_k(u_k^\tau) t_k(x) + b_0, \tag{20}$$

其中，$u_k^\tau = u(x,\ t = k\tau)$表示时间间隔为 τ时 u在第 k^{th}个时间步的分布。所提出的网络结构类似于显式时间步进器。给定第 k^{th} 个时间步 u_k^τ的输入，我们可以得到第$(k+1)^{\text{th}}$ 个时间步 u_{k+1}^τ的输出，随后可将其作为输入以获得 u_{k+2}^τ。按照这一迭代过程，即可获得后续所有时间步的分布序列 $[u_k^\tau, u_{k+1}^\tau, u_{k+2}^\tau, ...]$。

3.2 相场 DeepONet

机器学习算法通常需要大量数据集进行训练，损失函数通常根据数据与预测值之间的误差（例如均方误差、平均绝对误差）构造。然而，物理信息机器学习只需少量数据，甚至无需数据即可运行。这是通过将物理定律（例如偏微分方程）引入损失函数来实现的。物理信息机器学习可以处理两类典型问题，即逆问题和正问题。对于前者，物理信息仅部分已知，且有少量数据可用，目标是学习未知的物理信息；对于后者，我们已知所有物理信息，无需任何数据，目标是求解控制方程。本研究将重点关注正问题，即求解由梯度流控制的系统。

本文提出了一种用于自由能势梯度流的物理信息 DeepONet 通用框架，如图 1b 所示，以下称为“相场 DeepONet”。我们纳入了上述 DeepONet 结构

并根据最小化运动格式（式 16）构建了物理信息损失函数。损失函数可写为：

$$\mathcal{L} = \mathcal{F}(u_{k+1}^{\tau}) + \frac{d^2(u_{k+1}^{\tau}, u_k^{\tau})}{2\tau}. \tag{21}$$

训练过程将最小化该损失函数，这与最小化运动格式等价，因此可以近似真实值。训练该网络时，可以随机采样空间坐标 x 和分布 u_k^{τ} 作为训练数据的输入。训练时不需要输出 u_{k+1}^{τ}。训练好的 DeepONet 可以作为高效的代理模型（显式时间步进器），根据当前场分布预测后续所有时间步的场分布时间序列。详细的训练过程将在以下数值示例中说明。

值得注意的是，物理信息神经网络（PINN）可以视为物理信息 DeepONet 的一种特殊情况，其中省略了分支网络，或者将分支网络的输入固定为初始条件。例如，如果我们在图 1a 中仅保留主干网络，并将式 21 中的 u_k^{τ} 替换为 u_0^{τ}，则得到的框架就是一个用于求解下一步分布 u 的 PINN。

4 数值示例

4.1 弛豫动力学

我们首先给出一个由梯度流支配的一维空间弛豫动力学示例。其自由能为

$$\mathcal{F}(u) = \int_{\Omega} \frac{1}{2} k \cdot u^2(x,t)\,\mathrm{d}x. \tag{22}$$

我们考虑 $[-1,1]$ 内的一个一维区域。利用 L^2 内积，可通过方程（5）、方程（7）和方程（14）（$M=1$）推导出相应的偏微分方程和边界条件，

$$\begin{cases} \dfrac{\partial u}{\partial t} = -ku, \\ u(x=\pm 1) = 0. \end{cases} \tag{23}$$

给定初始条件 $u_0 = u(x, t=0)$，可以求得解析解 $u = -e^{kt}u_0$。在本例中，我们将比较两种不同的框架，即传统物理信息神经网络和所提出的相场 DeepONet，以说明如何将最小化运动格式纳入机器学习框架，并探讨所提出框架的优势。

对于物理信息神经网络框架，需要给定初始条件，设为 $u_0 = \sin(\pi x)$。常数 k 设为 10。按照最小化运动格式，一种直接的方法是将时间域离散为许多具有固定时间间隔 τ 的时间步。如图 2a 所示，对于每个时间步 k，我们构建一个对应的神经网络 $\mathcal{N}_k(x;\theta)$，其中 x 表示输入空间坐标，θ 表示待优化的权重和偏置。子网络的输出是场变量 $u_k^{\tau} = u(x, t=k\tau)$ 的当前分布。方程（16）中的自由能与距离之和被直接视为损失函数，

$$\begin{aligned} \mathcal{L} &= \mathcal{F}(u_k^{\tau}) + \frac{d_{L^2}^2\left(u_k^{\tau}, u_{k-1}^{\tau}\right)}{2\tau} \\ &= \int_{-1}^{1} \left[(u_k^{\tau})^2 + \frac{(u_k^{\tau} - u_{k-1}^{\tau})^2}{2\tau} \right] \mathrm{d}x \\ &\approx \frac{1}{N_F} \sum_{i=1}^{N_F} (u_k^{\tau})^2 + \frac{1}{N_d} \sum_{i=1}^{N_d} \frac{(u_k^{\tau} - u_{k-1}^{\tau})^2}{2\tau}, \end{aligned} \tag{24}$$

其中，${u^{\tau}}_{,k}{}^{i} = u(x = x_i, t = k\tau)$ 表示采样位置 x_i 处的场值；N_F, N_d 分别表示用于计算自由能项和距离项数值积分的样本总数。N_F 和 N_d 可以

不同；为简单起见，本研究将它们设为相等。训练过程遵循最小化运动格式：a) 首先，给定初始条件 u_0，训练第 1 个子网络 $\mathcal{N}_1$；b) 训练后，我们可以得到输出u_1^τ。随后，以 u_1^τ 为输入训练第 2 个子网络 $\mathcal{N}_2$；c) 重复此过程，依次训练所有子网络。

所有子网络均由两个隐藏层构成，每层各有 20 个节点。采用 ReLU 激活函数。我们在空间域 ($N_F = N_d = 1000$) 内随机采样了 1000 个均匀分布的数据点作为训练数据。使用 0.001 的学习率和 Adam 优化器。每个子网络依次训练 500 个训练轮次，整个训练重复 3 轮。在单个图形处理器（NVIDIA T400 4GB）系统上，训练耗时约 25 分钟。将子网络在不同时间步的预测结果与解析结果进行比较，可以观察到二者吻合良好（图 2b）。我们还可以看到系统的自由能在不断降低（图 2c）。这些结果成功验证了将最小化运动格式应用于物理信息机器学习的可行性。然而，这种基于物理信息神经网络的框架效率并不高。首先，所有子网络都必须依次训练；若要预测较大的时间间隔，子网络的数量就会大幅增加。其次，初始条件一旦改变，物理信息神经网络就需要重新训练，这限制了其应用。

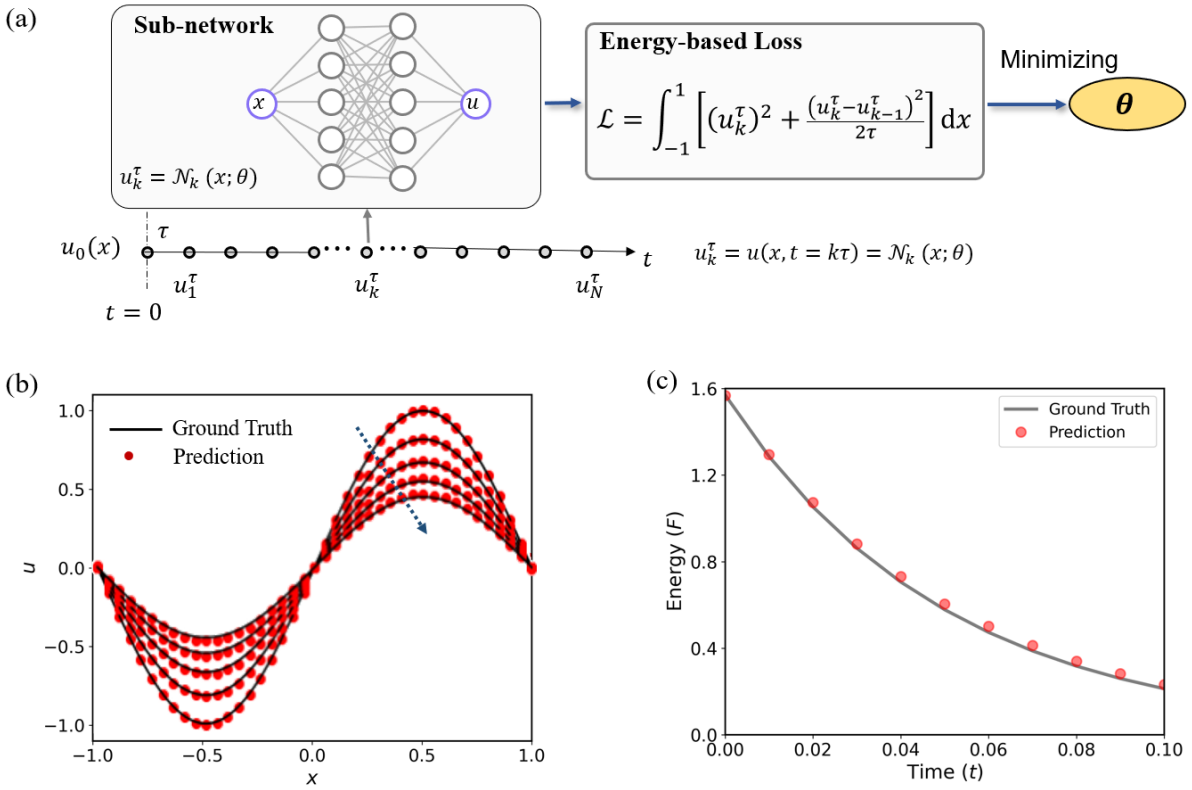

图 2：（a）最小化运动格式的多网络结构；（b）真实值与神经网络预测值在不同时间步下的比较；（c）预测的自由能随时间的变化。

h

为解决这些问题，正如我们在上一节中提出的，我们构建了一个 DeepONet，以当前时间步的分布作为输入，并预测下一时间步的分布，其作用类似于显式时间步进器。如图 3a 所示，分支网络和主干网络均为全连接神经网络。需要注意的是，连续函数无法直接输入分支网络，因此这里需要进行离散化。我们通过在空间域中等间距的位置采样 u 的值，对其进行离散化；随后将所得的数值向量作为输入（图 3b）。损失函数与公式 24 中的相同，不同之处在于，应分别用 $k+1$ 和 k 替换 k 和 $k-1$。

整个网络的输入为 $[x, u_k^\tau]$。物理常数 k 也设为 10；对于 DeepONet，无需给定初始条件。为简化训练过程，我们在空间域 [-1, 1]内均匀采样了 100 个点（$N_F = N_d = 100$）。我们通过高斯随机过程生成了 10,000 种不同的 u 分布，其中一半

表 1：三种不同情形下网络结构的参数

Case	主干网络		分支网络		传感器数量
	深度	宽度	深度	宽度	
弛豫动力学（1D）	3	100	2	100	100
Cahn-Hilliard（1D）	3	40	2	40	40
Allen-Cahn（2D）	滤波器：32 卷积核：(3,3) 步幅：(1, 1)	6 (3,3) (3, 3)	3	120	120

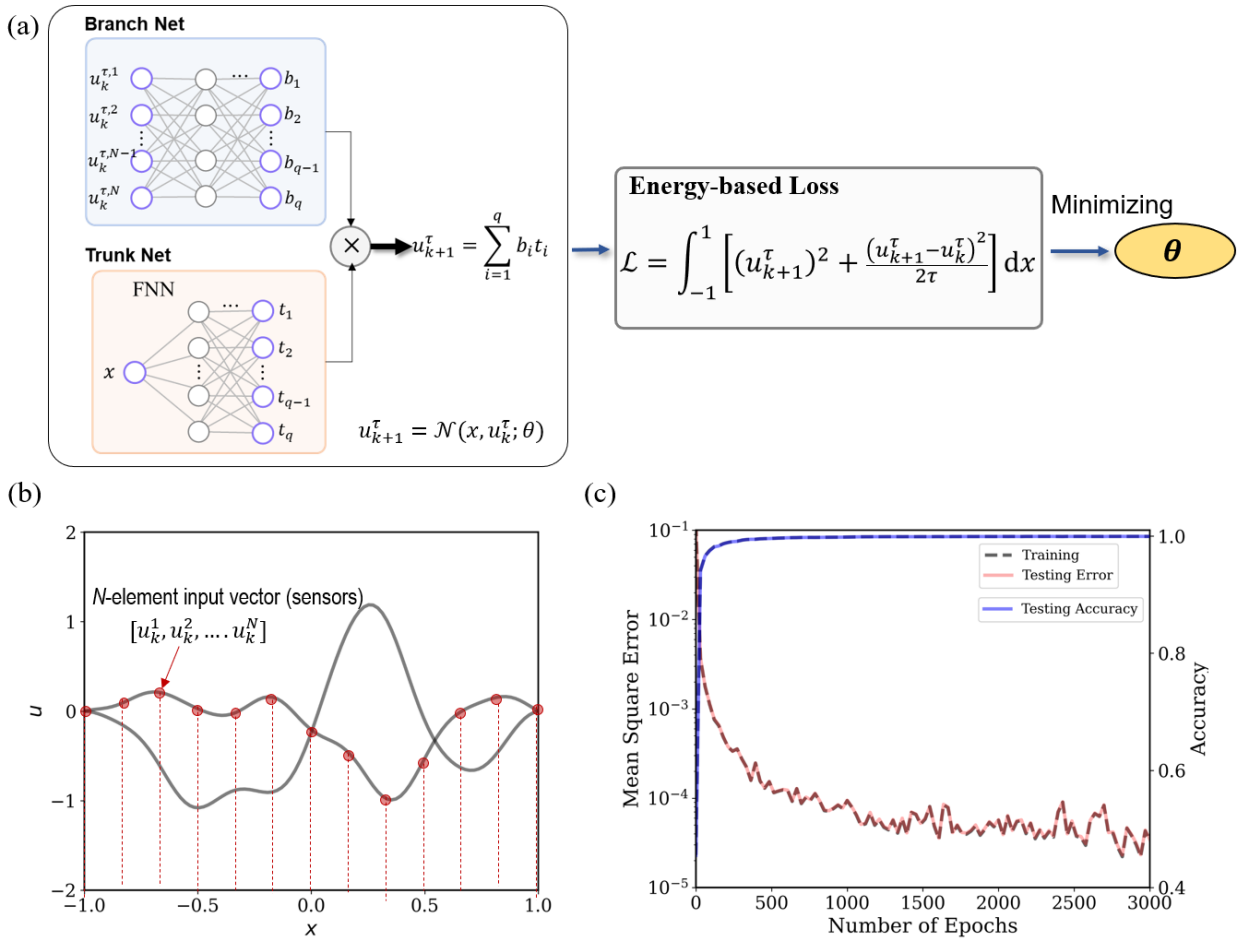

图 3：(a) 将最小化运动方案融入物理信息 DeepONet，用于一维弛豫动力学的示意图；(b) 由高斯随机过程生成的 u 随机分布及离散化输入向量；(c) 训练集和测试集中的均方误差与 r2 值（准确率）。

用于训练，另一半用于测试。DeepONet 的规模和结构见表4.1。该网络经过 3000 个训练轮次，学习率为 0.001。在相同的单图形处理器系统上，训练时间约为 10 分钟。训练和测试误差如图3c所示。测试集上的决定系数（r2 值）达到 0.99。这表明训练后的 DeepONet 可以准确预测弛豫动力学。图4a展示了 u 在不同时间步的预测分布，与真实值吻合良好。请注意， u 的正弦输入不在训练数据集中，这表明训练后的 DeepONet 具有良好的泛化能力。更重要的是，给定任意时间步的随机 u 输入，训练后的网络都能准确预测下一时间步的 u 分布（图4b）。

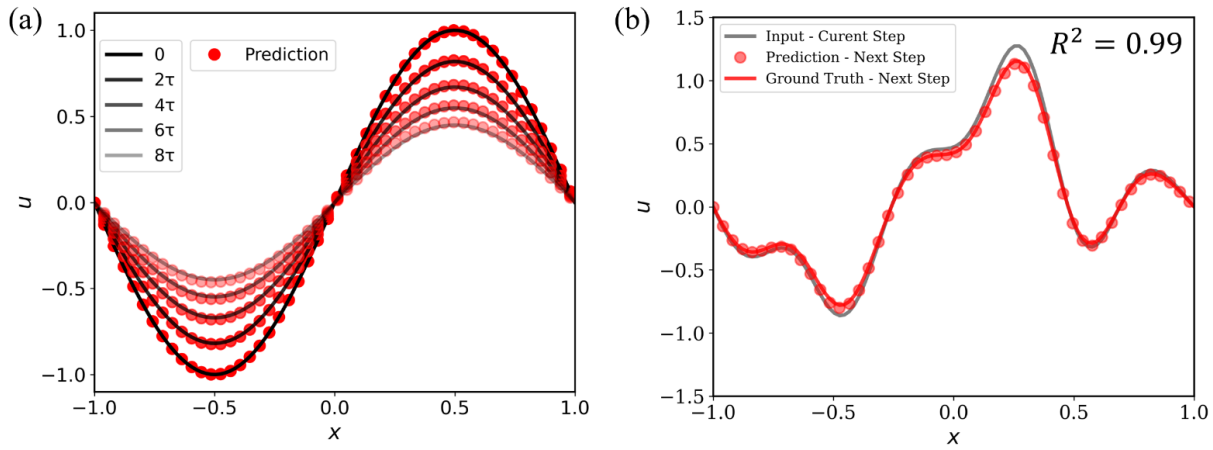

图4：训练后的 DeepONet 的预测结果：(a) 多个时间步的预测；(b) 对随机输入的预测。

4.2 艾伦-卡恩方程

在此示例中，我们研究二维区域中更复杂的艾伦-卡恩方程。该方程所支配的总自由能为：

$$\mathcal{F} = \int_{\Omega} \left[\frac{1}{\varepsilon^2} f(u) + 0.5(\nabla u)^2 \right] \mathrm{d}A, \tag{25}$$

其中，$f(u)$ 是体能量密度，一种常见的取法是 $f(u) = (u^2-1)_4^{\ 2}$；第二项表示界面能。利用 L^2 范数（或内积），即可通过方程（5）、方程（7）和方程（14）导出相应的偏微分方程和边界条件：

$$\begin{cases} \dfrac{\partial u}{\partial t} = \nabla^2 u - \frac{1}{\varepsilon^2} \dfrac{\mathrm{d}f}{\mathrm{d}u}, \\ \nabla u \cdot \mathbf{n} = 0. \end{cases} \tag{26}$$

考虑一个二维区域 $[-1,1] \times [-1,1]$。唯一的物理常数——长度尺度 ϵ——设为 0.25。我们构建了一个 DeepONet 结构，如图5a所示。分支网络将当前时间步的二维分布作为输入，因此采用卷积神经网络（CNN），并将其连接到前馈神经网络（FNN）。主干网络是一个以空间坐标为输入的前馈神经网络。最终输出为下一时间步的分布。此案例的损失函数为

$$\begin{aligned} \mathcal{L} &= \mathcal{F}(u_{k+1}^{\tau}) + \frac{d_{L^2}^2\left(u_{k+1}^{\tau}, u_k^{\tau}\right)}{2\tau} \\ &= \int_0^1 \int_0^1 \left[(\mathcal{F}(u_{k+1}^{\tau}) + \frac{\left(u_{k+1}^{\tau} - u_k^{\tau}\right)^2}{2\tau} \right] \mathrm{d}x\, \mathrm{d}y \\ &\approx \frac{1}{N_F} \sum_{i=1}^{N_F} \mathcal{F}(u_{k+1}^{\tau,i}) + \frac{1}{N_d} \sum_{i=1}^{N_d} \frac{(u_{k+1}^{\tau,i} - u_k^{\tau,i})^2}{2\tau}, \end{aligned} \tag{27}$$

其中，$u^{\tau}{}_k^{\ i} = u(x = x_i,\ y = y_i,\ t = k\tau)$表示采样位置处的场值；$N_F$和 N_d 分别表示用于计算自由能积分和距离项的样本总数。为了训练该网络，在均匀的 28×28 网格（$N_F = 256$）上，通过高斯随机过程生成 u 的随机二维分布，作为

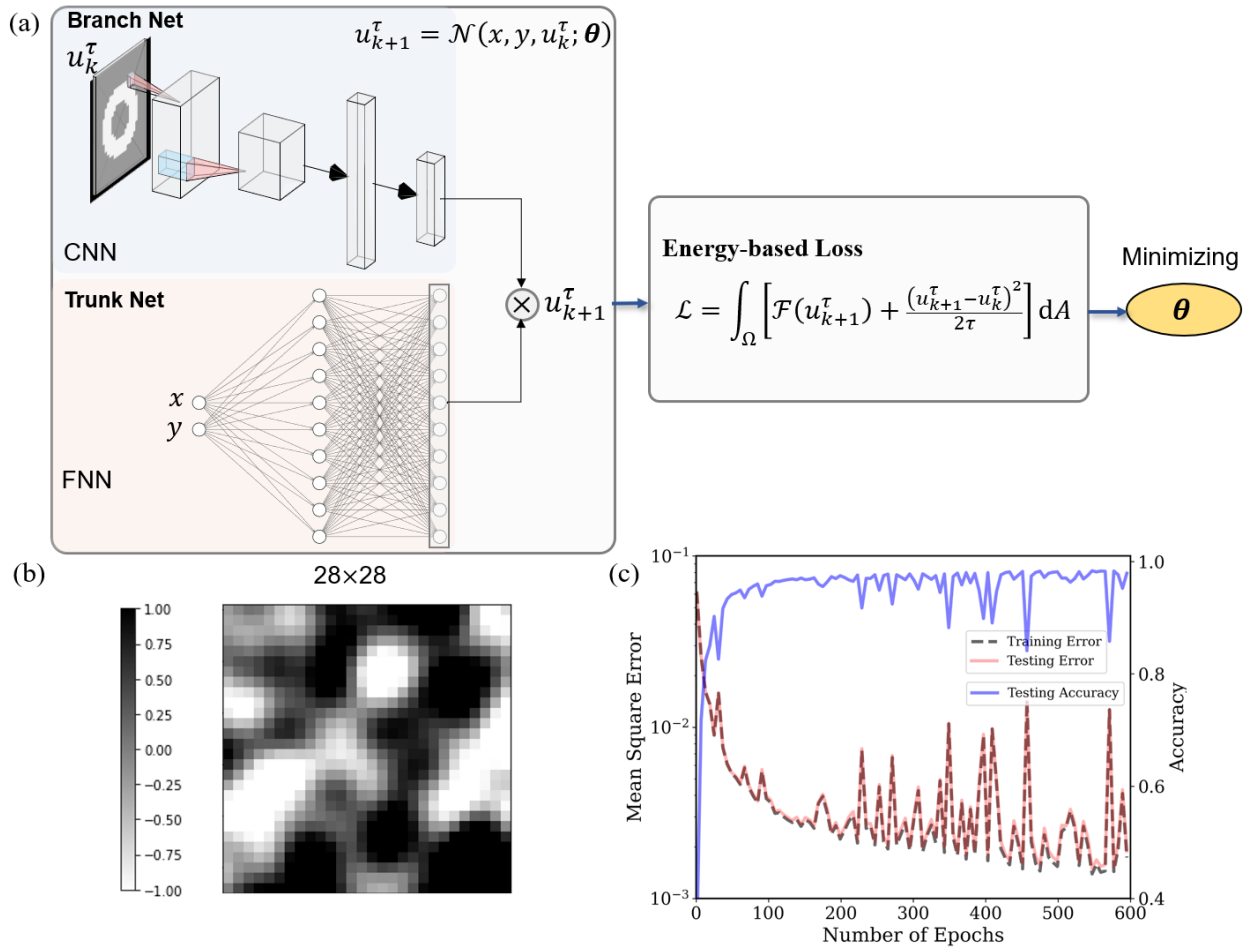

图5：（a）将最小化运动格式融入二维艾伦-卡恩方程的物理信息 DeepONet 示意图；（b）分支网络的图像输入（由高斯随机过程生成的随机分布 u 所构成的 28×28 网格）；（c）训练集和测试集中的均方误差与 r2 值（准确率）。

分支网络的输入（图5b）。总共生成了10,000个分布（图像），其中一半用于训练，另一半用于测试。为简化训练过程，将相同均匀 28×28 网格（$N_d = N_F = 256$）对应的空间坐标作为主干网络的输入。时间步长为 $\tau = 0.005$。网络的大小和结构见表4.1。分支网络中的卷积神经网络由两层组成。第一层包含32个滤波器，卷积核尺寸为3，步幅为1；第二层包含6个滤波器，卷积核尺寸为3，步幅为3。卷积神经网络的输出随后被展平，并全连接到包含120个节点的输出层。主干网络由三个隐藏层组成，每层包含120个节点。由于网络规模相对较大且损失函数中的梯度项较高，在单个图形处理器系统上完成训练约需150分钟。

图5c展示了训练误差和测试误差。在测试数据集中，r2 值最高可达 0.96。我们将训练好的 DeepONet 的预测结果与有限差分法得到的相应解进行比较，并将后者视为真实值（图6）。图6a、b和c比较了测试集中的一个代表性样本，可以看到神经网络的预测与真实值非常吻合。我们还比较了一个差异较大的案例（图6d、e和f），其分布在训练集中未曾出现。两者仍然吻合良好。这进一步验证了所提出框架的可靠性。

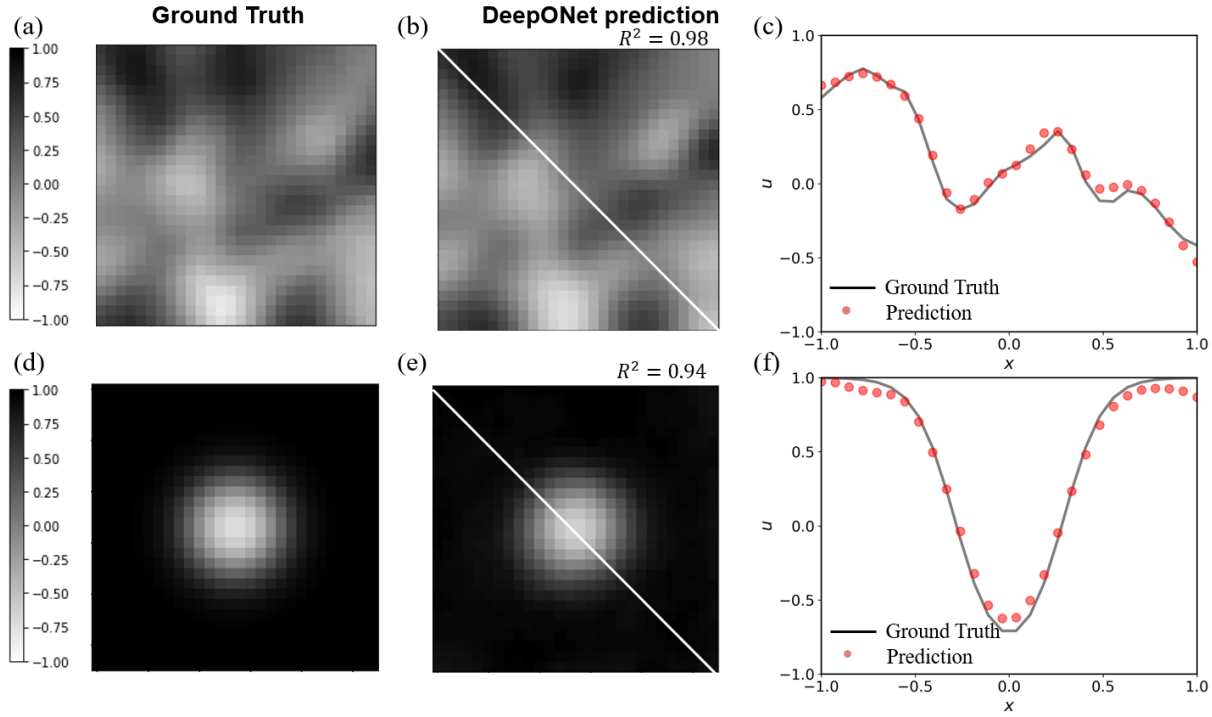

图6：训练后的 DeepONet 预测结果与真实值的对比，展示两个具有代表性的案例：(a)、(b)、(c) 为测试集中的一个案例；(d)、(e)、(f) 为数据集之外的另一个不同案例。(c) 和 (f) 中的数据点取自 (b) 和 (e) 所示的对角线。

4.3 Cahn-Hilliard 方程

一个更具挑战性的案例是 Cahn-Hilliard 方程。虽然自由能与前一个案例相同（公式25），但这里使用了不同的内积，即 H^{-1} 内积，用于描述动力学，由此得到一个带有边界条件的高阶偏微分方程，

$$\begin{cases} \dfrac{\partial u}{\partial t} = \nabla^2 \left(\frac{1}{\varepsilon^2} \dfrac{\mathrm{d}f}{\mathrm{d}u} - \nabla^2 u \right), \\ \nabla \left(\frac{1}{\varepsilon^2} \dfrac{\mathrm{d}f}{\mathrm{d}u} - \nabla^2 u \right) \cdot \mathbf{n} = 0, \\ \nabla u \cdot \mathbf{n} = 0. \end{cases} \tag{28}$$

我们考虑一维区域 $[0, 1]$。损失函数为：

$$\begin{aligned} \mathcal{L} &= \mathcal{F}(u_{k+1}^{\tau}) + \frac{d_{L^2}^2 \left(u_{k+1}^{\tau}, u_k^{\tau} \right)}{2\tau} \\ &= \int_0^1 \left[(\mathcal{F}(u_{k+1}^{\tau}) + \frac{(u_{k+1}^{\tau} - u_k^{\tau})(\phi_{k+1}^{\tau} - \phi_k^{\tau})}{2\tau} \right] \mathrm{d}x \\ &\approx \frac{1}{N_F} \sum_{i=1}^{N_F} \mathcal{F}(u_{k+1}^{\tau,i}) + \frac{1}{N_d} \sum_{i=1}^{N_d} \frac{(u_{k+1}^{\tau,i} - u_k^{\tau,i})(\phi_{k+1}^{\tau,1} - \phi_k^{\tau,1})}{2\tau}, \end{aligned} \tag{29}$$

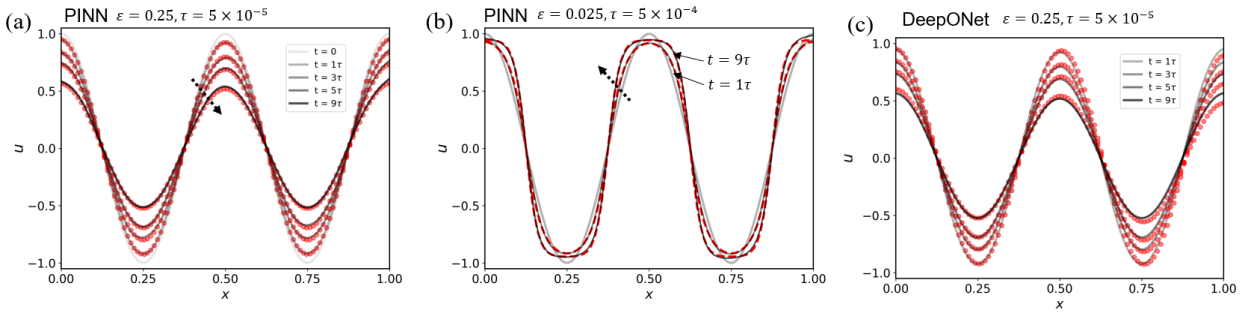

图 7：PINN 对无相分离情形（a）和有相分离情形（b）的预测。（c）DeepONet 对无相分离情形的预测。

其中， ϕ_k^τ 是源项为 u_k^τ 的泊松方程的解。利用有限差分格式，可以得到从离散化的 u_k^τ 到 ϕ_k^τ 的线性映射：

$$\phi_k^\tau = \mathbf{M}^{-1}\mathbf{u}_k^\tau, \tag{30}$$

其中 $\phi_k^\tau = [\phi_k^{\tau,1}, \ldots, \phi_k^{\tau,i}, \ldots, \phi_k^{\tau,N_d}]$、$\mathbf{u}_k^\tau = [u_k^{\tau,1}, \ldots, u_k^{\tau,i}, \ldots, u_k^{\tau,N_d}]$，以及

$$\mathbf{M} = \frac{1}{\Delta x}\begin{bmatrix} -1 & 1 & & & & & & \\ 1 & -2 & 1 & & & & & \\ & \ddots & \ddots & \ddots & & & & \\ & & 1 & -2 & 1 & & & \\ & & & \ddots & \ddots & \ddots & & \\ & & & & 1 & -2 & 1 & \\ & & & & & 1 & -1 & \end{bmatrix}, \Delta x = \frac{1}{N_d - 1}. \tag{31}$$

考虑到该问题的复杂性，我们首先采用了一个相对简单的基于物理信息神经网络的框架，以进一步验证将最小化运动格式引入物理信息机器学习的可行性。给定了初始条件 $u(x,\ t=0) = cos(4\pi x)$ 和边界条件 $u(x=0,\ t) = u(x=1,\ t) = 0$。我们采用了与第一个示例中所述相同的结构。通过改变长度尺度 ϵ，研究了有（$\epsilon = 0.025$）和无（$\epsilon = 0.25$）明显相分离的两种不同情形。两种情形对应的时间步 τ 分别为 5×10^{-4} 和 5×10^{-5}。训练过程与第一个示例相同，训练结果如图 7a、b 所示。可以看出，在这两种情形下，物理信息神经网络的预测结果与真实值均吻合良好。这表明，将最小化运动格式引入机器学习同样适用于这一高 4th 阶方程。

随后，我们训练了一个 DeepONet，其结构见表4.1。类似地，我们随机分布采样了10,000个u，其中一半用于训练，另一半用于测试。图7c展示了无明显相分离情况下的结果。训练后的 DeepONet 预测结果同样能够在多个时间步上很好地捕捉真实值。这些结果进一步验证了所提出框架的可行性。需要注意的是，在当前框架下，准确预测相分离的动态演化仍然是一项重大挑战。这主要是由于相分离的时间尺度相对较大，同时仍需要较小的时间步来捕捉快速的相分离过程。一个时间步的轻微差异使其难以被当前框架捕捉。我们将在下一节进一步讨论这一局限性。

5 讨论

到目前为止，我们已成功开发出一种通用的物理信息深度算子神经网络，并采用基于能量的损失函数，通过三个不同的数值示例对其进行了验证。在此，我们将进一步阐述几个方面，以夯实这一方法的理论基础并拓展其适用范围。

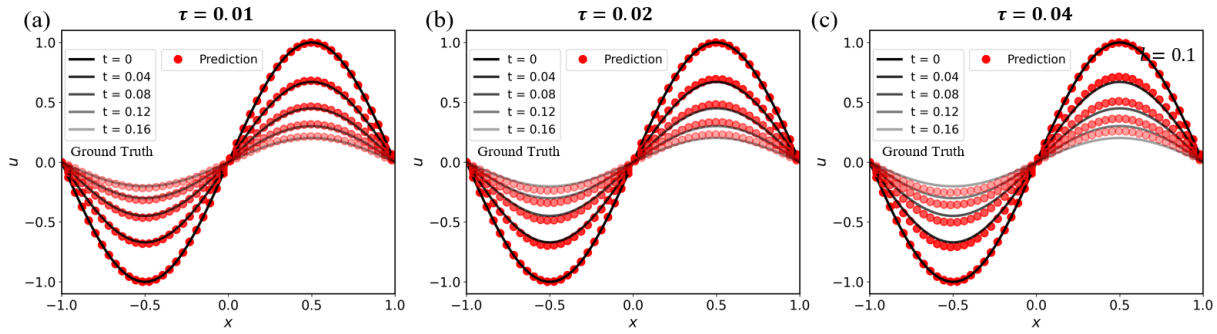

图 8：不同时间步 τ下 DeepONet 的预测结果：(a) $\tau = 0.01$，(b) $\tau = 0.02$，(c) $\tau = 0.04$。

5.1 时间步 τ的估计

时间步 τ需要足够小，才能获得准确的解。然而，据作者所知，目前尚无直接方法可以估计 τ的上限。在本研究中，我们逐步减小时间步，以确保预测结果收敛至真实值。图 8展示了动力学弛豫数值示例中，τ取三个不同值时的训练结果。随着 τ从 0.01（图 8a）增至 0.02（图 8b）和 0.04（图 8c），可以观察到 DeepONet 预测结果与真实值之间的差异逐渐增大。

在实践中，我们对时间步的估计有两点建议。首先，如果有实验数据，应选择一个能够检测到场变量明显变化的时间步。其次，可以基于物理信息神经网络（PINN）通过试错过程优化该时间步。如前文所述，PINN 是仅包含主干网络的 DeepONet 的简化形式。PINN 的训练速度远快于 DeepONet，因此建议使用 PINN 确定 τ。由于时间步的估计应当与初始条件无关，因此，对于 PINN 能够确保解收敛的时间步，也应当适用于 DeepONet。

小时间步长的要求也是所提出框架的一项局限。在较大的时间尺度上进行预测时，需要对训练好的 DeepONet 进行更多次评估，这会降低效率，而且误差预计会不断累积。为避免这一问题，我们可以训练一个不包含最小化运动格式的 DeepONet。因此，所需输出必须来自实验或数值模拟（监督学习）。这样便可以采用较大的时间步长，进一步提高计算效率。

在该框架中实现自适应时间步长的可能性，也是一个值得深入研究的课题。以 Cahn-Hilliard 方程的相分离案例为例（图 7b），相分离开始之前，场变量变化很小，而这一过渡阶段持续时间相对较长。相比之下，相分离期间，过程会在较短时间内迅速展开。因此，若能实现自适应时间步长，即在相分离前阶段采用较大的时间步长，在相分离阶段采用较小的时间步长，则有望提高计算效率。

5.2 训练样本的影响

采样方法是影响准确度的另一个可能因素。场变量的输入函数 $u(x)$ 由零均值高斯随机过程生成，

$$\begin{aligned} u(x) &\sim \mathbf{G}(0, k_l(x_1, x_2)), \\ k_l(x_1, x_2) &= exp(-||x_1 - x_2||^2/2l^2), \end{aligned} \tag{32}$$

其中 $k_l(x_1, x_2)$ 是协方差核，其长度尺度参数为 l。它控制 u 生成分布的平滑度。较大的 l 会使 u 更平滑（图9a）。

为简化起见，本研究将训练集和测试集设为具有相同的“平滑度”。例如，在第一个数值示例中，训练集和测试集均使用 $l = 0.2$。然而，当 u 的输入分布具有不同程度的平滑度时，这可能会影响准确性。图9b和图9c分别展示了同一个 DeepONet 对更平滑的输入（$l = 0.5$）和更尖锐的输入（$l = 0.1$）的预测，其中输入均为受随机扰动的 u。对于更平滑的输入，预测结果仍然高度吻合；而对于更尖锐的输入，则可以观察到尖角处存在偏差。进一步提升性能的一种潜在解决方案是纳入具有不同平滑度的训练数据，

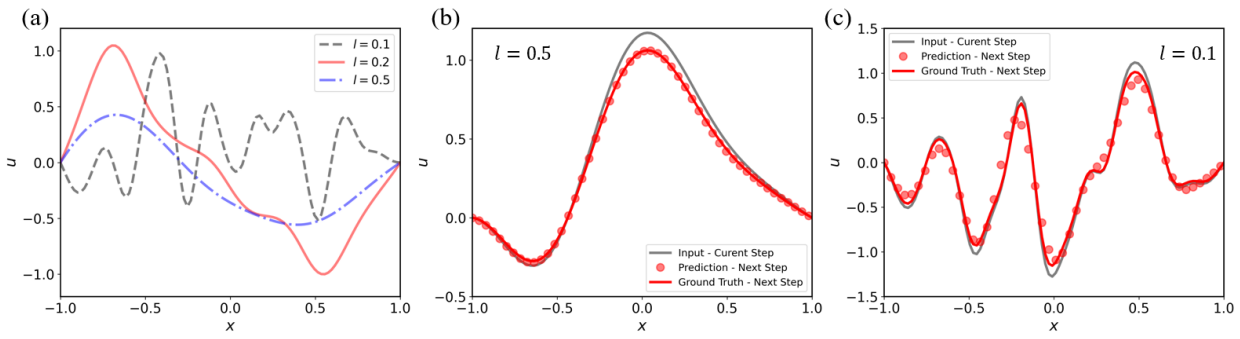

图 9：（a）使用不同长度尺度生成的随机分布 $l = 0.1$，0。2，$and 0.5$；DeepONet 对长度尺度为 u $l = 0.5$ 的随机生成输入的预测结果：（b）和 $l = 0.5$（c）。

即涵盖不同平滑度的数据。值得注意的是，即使输入更尖锐，整体趋势仍能得到较为准确的捕捉。这表明该框架应用于随机输入时具有普适性。

5.3 其他潜在改进

除了优化时间步和采样数据集之外，未来还可以改进该框架的其他几个方面。

a）考虑更广泛的物理常数范围。这将有利于实际应用。目前，为了简便起见，使用了一组固定的物理常数，但实际上，这些数值可能因不同系统而异。为解决这一问题，可以通过添加一个额外的分支网络，将物理常数作为输入，从而训练一个适用于所有可能物理常数的 DeepONet。

b）处理不规则形状区域。这将进一步增强其适用性。DeepONet 能够通过在二维或更高维空间中随机采样位置，并将其作为向量输入网络，来处理任何固定的不规则区域。然而，挑战在于，如何将针对特定区域训练的 DeepONet 应用于具有其他几何形状的区域。

c）将该框架扩展至高维问题。研究在低维度训练的 DeepONet 能否有效应用于更高维度将会很有意义。高维度下的相场模拟因计算成本极高而闻名，利用 DeepONet 扩展数值模拟规模或许能够克服这一问题。

d）基于实验数据的逆向学习。当前框架侧重于正问题，其中所有物理规律和常数均已知，DeepONet 通过无监督学习（无需实验数据）近似求解。若能将该框架扩展至逆问题，将更具意义。在逆问题中，只有部分物理规律和一些实验数据是已知的，目标是识别未知的物理常数或定律。通过修改损失函数以纳入实验数据，是实现这一目标的一种方式。如此一来，物理和实验数据便可纳入同一框架，训练过程既能够近似真实值，也能够学习未知的物理常数和定律。然而，训练过程可能耗时较长，不适用于快速或实时识别。开发一个基于 DeepONet 的快速识别框架仍需付出更多努力。

6 结论

我们提出了一种物理信息驱动的相场 DeepONet 框架，用于求解由自由能泛函梯度流支配的动力系统。该框架融入了最小化运动格式，以求解系统动力学，而不是直接求解控制方程的偏微分方程。三个不同的数值示例验证了所提出的框架，其中包括相场方法的两个主要方程，即 Allen-Cahn 方程和 Cahn-Hilliard 方程。我们可以从这项工作中得出以下主要结论：

1. 变分原理（如梯度流）具有巨大的潜力，可无缝融入物理信息机器学习框架，为数据与物理的融合提供一种新方法，并带来更广泛的实际应用价值。

2. 所提出的相场 DeepONet 框架成功求解了相场方法中的 Allen-Cahn 方程和 Cahn-Hilliard 方程，展现了其在模拟化学系统中图案形成方面的有效性。

3. 本研究训练的相场 DeepONet 可用作高效的显式时间步进器，有望实现对动力系统的快速实时预测。

这项工作展现了对更一般相场模型进行深度算子学习的可能性，包括化学非平衡热力学模型 [51, 50], ，这类模型中通量或反应速率对扩散化学势存在非线性依赖。要超越梯度流近似并考虑非线性动力学，就需要扩展最小化运动格式。如此一来，相场 DeepONet 框架便可利用丰富的图像数据集实现数据驱动学习，并快速模拟图案形成，而不局限于偏微分方程约束优化 [42]。

致谢

W.L. 和 J.Z. 衷心感谢美国国家航空航天局 19-TTT-0103 项目（奖励编号：80NSSC21M0114）对本研究的支持。他们还获得了东北大学及其工程学院提供的启动经费。M.Z.B 和 W.L. 感谢丰田研究院通过 D3BATT可充电电池数据驱动设计中心提供的支持。

[标题]

附录 A 变分法基础

A.1 欧拉-拉格朗日方程

给定一个光滑流形 X 以及一个光滑实值函数 $f = f(x,\ u(x)\ u'(x))$，定义泛函 $\mathcal{F}$ 为

$$\mathcal{F} = \int_{x_a}^{x_b} f(x, u(x), u'(x))\,\mathrm{d}x \tag{33}$$

当欧拉-拉格朗日方程满足时，该泛函具有驻值（最大值、最小值或鞍点），

$$\frac{\partial f}{\partial u_i} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial(u_i)} = 0, \quad i = 1, 2, ..., n. \tag{34}$$

A.2 泛函导数的推导

考虑一种特定类型的能量泛函的泛函导数，该能量泛函仅取决于场变量及其一阶导数，即 $\mathcal{F} = \int_\Omega F(x, u(x)\ \nabla u(x)\ \mathrm{d}x$。我们有

$$\begin{aligned}
\int_\Omega \frac{\delta \mathcal{F}}{\delta u} v\,\mathrm{d}x &= \left[\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega F(x, u + vt, \nabla u + \nabla v t)\,\mathrm{d}x\right]_{t=0} \\
&= \int_\Omega \left(\frac{\partial F}{\partial u} v + \frac{\partial F}{\partial \nabla u}\cdot\nabla v\right)\mathrm{d}x \\
&= \int_\Omega \left(\frac{\partial F}{\partial u} v - (\nabla\cdot\frac{\partial F}{\partial \nabla u})v\right)\mathrm{d}x + \int_{\partial\Omega}\left(\frac{\partial F}{\partial \nabla u}\cdot\mathbf{n}\right) v\,\mathrm{d}x \\
&= \int_\Omega \left(\frac{\partial F}{\partial u} - \nabla\cdot\frac{\partial F}{\partial \nabla u}\right) v\,\mathrm{d}x.
\end{aligned} \tag{35}$$

当 $v = 0$ 或 $\dfrac{\partial F}{\partial \nabla u}\cdot\mathbf{n} = 0$ 在边界上时，第四行成立。由于 v 是任意函数，我们得到泛函导数

$$\frac{\delta \mathcal{F}}{\delta u} = \frac{\partial F}{\partial u} - \nabla\cdot\frac{\partial F}{\partial \nabla u}\,. \tag{36}$$

参考文献

[1] Athanasios Voulodimos、Nikolaos Doulamis、Anastasios Doulamis 和 Eftychios Protopapadakis。《用于计算机视觉的深度学习：简要综述》。计算智能与神经科学，2018年，2018年。[2] Tom Young、Devamanyu Hazarika、Soujanya Poria 和 Erik Cambria。《基于深度学习的自然语言处理最新趋势》。*IEEE*计算智能杂志，13(3):55–75，2018年。[3] Nathan Baker、Frank Alexander、Timo Bremer、Aric Hagberg、Yannis Kevrekidis、Habib Najm、Manish Parashar、Abani Patra、James Sethian、Stefan Wild、Karen Willcox 和 Steven Lee。《科学机器学习基础研究需求研讨会报告：人工智能核心技术》。技术报告，美国能源部科学办公室（SC），华盛顿特区（美国），2019年2月。[4] Shengze Cai、Zhicheng Wang、Lu Lu、Tamer A Zaki 和 George Em Karniadakis。《Deepm&mnet：基于神经网络算子近似推断电对流多物理场》。计算物理学杂志，436:110296，2021年。[5] Chensen Lin、Zhen Li、Lu Lu、Shengze Cai、Martin Maxey 和 George Em Karniadakis。《用于预测多尺度气泡生长动力学的算子学习》。《化学物理学杂志》，154(10):104118，2021年。[6] Minglang Yin、Enrui Zhang、Yue Yu 和 George Em Karniadakis。《将有限元与深度神经算子相结合，实现力学问题的快速多尺度建模》。《计算机在应用力学与工程中的方法》，第115027页，2022年。[7] Benben Jiang、William E Gent、Fabian Mohr、Supratim Das、Marc D Berliner、Michael Forsuelo、Hongbo Zhao、Peter M Attia、Aditya Grover、Patrick K Herring 等。《用于快速预测锂离子电池循环方案的贝叶斯学习》。*Joule*，5(12):3187–3203，2021年。[8] Benben Jiang、Marc D Berliner、Kun Lai、Patrick A Asinger、Hongbo Zhao、Patrick K Herring、Martin Z Bazant 和 Richard D Braatz。《通过贝叶斯优化进行锂离子电池快速充电设计》。应用能源，307:118244，2022年。[9] George Em Karniadakis、Ioannis G. Kevrekidis、Lu Lu、Paris Perdikaris、Sifan Wang 和 Liu Yang。《基于物理的机器学习》。《自然评论：物理学》，3(6):422–440，2021年6月。[10] Donal P Finegan、Juner Zhu、Xuning Feng、Matt Keyser、Marcus Ulmefors、Wei Li、Martin Z Bazant 和 Samuel J Cooper。《数据驱动方法与基于物理的学习在提升电池安全性方面的应用》。*Joule*，5(2):316–329，2021年。[11] M. Raissi、P. Perdikaris 和 G. E. Karniadakis。《物理信息神经网络：用于求解涉及非线性偏微分方程的正问题与逆问题的深度学习框架》。计算物理学杂志，378:686–707，2019年。[12] Maziar Raissi 和 George Em Karniadakis。《隐式物理模型：非线性偏微分方程的机器学习》。计算物理学杂志，357:125–141，2018年3月。[13] Lu Lu、Xuhui Meng、Zhiping Mao 和 George E. Karniadakis。《DeepXDE：用于求解微分方程的深度学习库》。第1–21页，2019年。[14] Han Gao、Luning Sun 和 Jian-Xun Wang。《Phygeonet：用于求解不规则区域上参数化稳态偏微分方程的物理信息几何自适应卷积神经网络》。计算物理学杂志，428:110079，2021年。[15] Yizheng Wang、Jia Sun、Wei Li、Zaiyuan Lu 和 Yinghua Liu。《CENN：基于具有子域的神经网络的守恒能量法，用于求解涉及非均质复杂几何结构的变分问题》。《计算机在应用力学与工程中的方法》，400:115491，2022年。[16] Ameya D Jagtap、Ehsan Kharazmi 和 George Em Karniadakis。《离散区域上的守恒物理信息神经网络：守恒律正问题与逆问题的应用》。《计算机在应用力学与工程中的方法》，365:113028，2020年。[17] Lu Lu、Raphael Pestourie、Wenjie Yao、Zhicheng Wang、Francesc Verdugo 和 Steven G Johnson。《带硬约束的物理信息神经网络用于逆向设计》。《工业与应用数学学会科学计算杂志》，43(6):B1105–B1132，2021年。[18] Ameya D Jagtap、Kenji Kawaguchi 和 George Em Karniadakis。《自适应激活函数加速深度神经网络和物理信息神经网络的收敛》。计算物理学杂志，404:109136，2020年。

[19] Jeremy Yu、Lu Lu、Xuhui Meng 和 George Em Karniadakis。用于偏微分方程正问题与反问题的梯度增强物理信息神经网络。《计算机在应用力学与工程中的方法》，393:114823，2022年。[20] Wei Li、Martin Z Bazant 和 Juner Zhu。用于弹性板的物理引导神经网络框架：基于控制方程的方法与基于能量的方法的比较。《计算机在应用力学与工程中的方法》，383:113933，2021年。[21] E. Samaniego、C. Anitescu、S. Goswami、V. M. Nguyen-Thanh、H. Guo、K. Hamdia、X. Zhuang 和 T. Rabczuk。通过机器学习求解计算力学中的偏微分方程：一种能量方法——概念、实现与应用。《计算机在应用力学与工程中的方法》，362:112790，2020年。[22] Apostolos F Psaros、Kenji Kawaguchi 和 George Em Karniadakis。元学习 PINN 损失函数。计算物理学杂志，458:111121，2022年。[23] 鄂维南和 Bing Yu。深度 Ritz 方法：一种用于求解变分问题的深度学习数值算法。数学与统计通讯，1(6):1–12，2018年。[24] Lu Lu、Pengzhan Jin、Guofei Pang、Zhongqiang Zhang 和 George Em Karniadakis。基于算子通用逼近定理，通过 DeepONet 学习非线性算子。自然·机器智能，3(3):218–229，2021年3月。[25] Nikola Kovachki、Zongyi Li、Burigede Liu、Kamyar Azizzadenesheli、Kaushik Bhattacharya、Andrew Stuart 和 Anima Anandkumar。神经算子：学习函数空间之间的映射。*arXiv:2108.08481 [cs, math]*，2021年12月。[26] Kristen A. Severson、Peter M. Attia、Norman Jin、Nicholas Perkins、Benben Jiang、Zi Yang、Michael H. Chen、Muratahan Aykol、Patrick K. Herring、Dimitrios Fraggedakis、Martin Z. Bazant、Stephen J. Harris、William C. Chueh 和 Richard D. Braatz。在容量衰减之前对电池循环寿命进行数据驱动预测。自然·能源，4(5):383–391，2019年。[27] Weihan Li、Jiawei Zhang、Florian Ringbeck、Dominik Jöst、Lei Zhang、Zhongbao Wei 和 Dirk Uwe Sauer。用于锂离子电池电极级状态估计的物理信息神经网络。电源杂志，506:230034，2021年。[28] Jinpeng Tian、Rui Xiong、Jiahuan Lu、Cheng Chen 和 Weixiang Shen。动态使用情境下基于物理信息深度学习的电池荷电状态估计。储能材料，50:718–729，2022年。[29] Bin Wu、Buyi Zhang、Changyu Deng 和 Wei Lu。在缺乏真实值直接知识的情况下，利用物理编码深度学习识别电池参数。应用能源，321:119390，2022年。[30] Muratahan Aykol、Chirranjeevi Balaji Gopal、Abraham Anapolsky、Patrick K Herring、Bruis van Vlijmen、Marc D Berliner、Martin Z Bazant、Richard D Braatz、William C Chueh 和 Brian D Storey。观点——结合物理与机器学习预测电池寿命。电化学学会志，168(3):030525，2021年。[31] Renato G Nascimento、Matteo Corbetta、Chetan S Kulkarni 和 Felipe AC Viana。用于锂离子电池建模与寿命预测的混合物理信息神经网络。电源杂志，513:230526，2021年。[32] Alexander Bills、Shashank Sripad、William L Fredericks、Matthew Guttenberg、Devin Charles、Evan Frank 和 Venkatasubramanian Viswanathan。适用于电动飞机的通用电池性能与衰减模型。*arXiv* 预印本 *arXiv:2008.01527*，2020年。[33] Bor-Rong Chen、M Ross Kunz、Tanvir R Tanim 和 Eric J Dufek。结合多种基于物理的电化学特征，利用机器学习框架早期检测析锂。细胞报告·物理科学，2(3):100352，2021年。[34] Jongwoo Lim、Yiyang Li、Daan Hein Alsem、Hongyun So、Sang Chul Lee、Peng Bai、Daniel A. Cogswell、Xuzhao Liu、Norman Jin、Young Sang Yu、Norman J. Salmon、David A. Shapiro、Martin Z. Bazant、Tolek Tyliszczak 和 William C. Chueh。电池初级颗粒内部锂组成时空动力学的起源与滞后。科学，353:566–571，2016年。[35] Yiyang Li、Farid El Gabaly、Todd R. Ferguson、Raymond B. Smith、Norman C. Bartelt、Joshua D. Sugar、Kyle R. Fenton、Daniel A. Cogswell、A. L. David Kilcoyne、Tolek Tyliszczak、Martin Z. Bazant 和 William C. Chueh。电流诱导相分离电池电极从逐颗粒嵌入转变为并发嵌入。《自然·材料》，13:1149–1156，2014年。[36] Martin Z Bazant。基于非平衡热力学的化学动力学与电荷转移理论。《化学研究述评》，46(5):1144–1160，2013年。

[37] Todd R. Ferguson 和 Martin Z. Bazant。多孔电极的非平衡热力学。电化学学会志，159:A1967–A1985，2012年1月。NULL。[38] Raymond B. Smith 和 Martin Z. Bazant。多相多孔电极理论。电化学学会志，164:E3291–E3310，2017年6月。[39] Colby L Wight 和 Jia Zhao。利用自适应物理信息神经网络求解 Allen-Cahn 和 Cahn-Hilliard 方程。*arXiv* 预印本 *arXiv:2007.04542*，2020年。[40] Revanth Mattey 和 Susanta Ghosh。一种用于时变非线性高阶偏微分方程的物理信息神经网络。*arXiv* 预印本 *arXiv:2106.07606*，2021年。[41] Revanth Mattey 和 Susanta Ghosh。一种训练物理信息神经网络以求解 Allen-Cahn 和 Cahn-Hilliard 方程的新型序贯方法。《计算机在应用力学与工程中的方法》，390:114474，2022年。[42] Hongbo Zhao、Haitao Deng、Alexander Cohen、Jongwoo Lim、Yiyang Li、Dimitrios Fraggedakis、Benben Jiang、Brian Storey、William Chueh、Richard Braatz 和 Martin Bazant。逐像素地从 X 射线影片中学习非均相反应动力学。2022年。https://doi.org/10.21203/rs.3.rs-2320040/v1。[43] Haitao D Deng、Hongbo Zhao、Norman Jin、Lauren Hughes、Benjamin H Savitzky、Colin Ophus、Dimitrios Fraggedakis、András Borbély、Young-Sang Yu、Eder G Lomeli 等。相变固体中化学力学的相关图像学习。《自然·材料》，21(5):547–554，2022年。[44] J. W. Cahn 和 J. E. Taylor。综述第113篇：表面扩散引起的表面运动。《冶金学与材料学学报》，42(4):1045–1063，1994年4月。[45] Jean E. Taylor 和 John W. Cahn。通过梯度流关联各向异性锐界面与弥散界面运动定律。《统计物理学杂志》，77(1):183–197，1994年10月。[46] Filippo Santambrogio。{ 欧几里得、度量与 Wasserstein }梯度流：概述。*arXiv:1609.03890[math]*,2016年9月。[47] John S Rowlinson。J. D. 范德瓦耳斯《连续密度变化假设下的毛细现象热力学理论》一文的译文。《统计物理学杂志》，20:197–200，1979年。[48] John W. Cahn 和 John E. Hilliard。非均匀系统的自由能。一、界面自由能。《化学物理学杂志》，28(2):258–267，1958年2月。[49] Martin Z. Bazant。基于非平衡热力学的化学动力学与电荷转移理论。《化学研究述评》，46:1144–1160，2013年5月。[50] Martin Z. Bazant。受驱动开放系统的热力学稳定性及通过电化学自催化控制相分离。《法拉第讨论》，199:423–463，2017年。[51] Martin Z Bazant、Brian D Storey 和 Alexei A Kornyshev。离子液体中的双电层：过度屏蔽与拥挤效应。《物理评论快报》，106:046102，2011年1月。