

学校代码: 10491

研究生学号: 1202411170

中国地质大学
硕士学位论文

深度学习方法求解偏微分方程研究综述

姓 名: 刘亦凡
学 科 专 业: 计算机科学与技术
指 导 教 师: 导师姓名
培 养 单 位: 计算机学院

二〇二五年一二月

A Dissertation Submitted to China University of Geosciences
For the Master Degree of Master

Deep Learning Methods for Solving Partial Differential Equations: A Review

Master Candidate: Yifan Liu
Major: Computer Science and Tech-
nology
Supervisor: 墙威

China University of Geosciences
Wuhan 430074 P.R. China

中国地质大学（武汉）研究生学位论文原创性声明

本人郑重声明：本人所呈交的硕士学位论文《深度学习方法求解偏微分方程研究综述》，是本人在导师的指导下，在中国地质大学（武汉）攻读硕士学位期间独立进行研究工作所取得的成果。论文中除已注明部分外不包含他人已发表或撰写过的研究成果，对论文的完成提供过帮助的有关人员已在文中说明并致以谢意。

本人所呈交的 硕士 学位论文没有违反学术道德和学术规范，没有侵权行为，并愿意承担由此而产生的法律责任和法律后果。

学位论文作者签名:

日 期: 年 月 日

中国地质大学（武汉）研究生学位论文导师承诺书

本人郑重承诺：本人所指导的硕士学位论文《深度学习方法求解偏微分方程研究综述》是在本人的指导下，研究生在中国地质大学（武汉）攻读硕士学位期间独立进行研究工作所取得的成果，论文由研究生独立完成。

研究生所呈交的硕士学位论文没有违反学术道德和学术规范，没有侵权行为，并愿意承担由此而产生的与导师相关的责任和后果。

指导老师（签字）：

日期： 年 月 日

中国地质大学（武汉）研究生学位论文使用授权书

本人授权中国地质大学（武汉）可采用影印、缩印、数字化或其它复制手段保存本学位论文；学校可向国家有关部门或机构送交本学位论文的电子版全文，编入有关数据库进行检索、下载及文献传递服务；同意在校园网内提供全文浏览和下载服务。

涉密论文解密后适用于本授权书。

学位论文作者签名：

日期： 年 月 日

作者简介

一、基本情况

姓名: 某某某 性别: 男 民族: 汉

出生年月: 1994-05-02 籍贯: 山西省朔州市

2013.09 – 2017.07 中国地质大学（武汉）工学学士；

2017.09 – 至今 中国地质大学（武汉）攻读硕士学位；

二、学术论文

1. X. X 研究 [J]. X 学报, 201X(1): 53-55.

2. X. X 分析 [J]. X 技术, 2005(5): 6-7.

3. ...

三、获奖、专利情况

1. X.X. 江苏省科技进步奖三等奖. 排名第二; (ps: 逗我呢?)

2. ...

四、研究项目

1. X 项目, 国家自然基金, 项目编号: XXX, 参加人员;

2. ...

摘 要

偏微分方程（Partial Differential Equations, PDEs）是描述自然现象和工程问题的重要数学工具，广泛应用于流体力学、固体力学、电磁学、量子力学等领域。传统的数值方法如有限差分法、有限元法和有限体积法虽然成熟可靠，但在处理高维问题、复杂几何和多尺度问题时面临计算成本高、精度受限等挑战。近年来，深度学习技术的快速发展为解决偏微分方程提供了新的思路和方法。本文系统综述了深度学习方法在求解偏微分方程方面的研究进展，主要包括物理信息神经网络（Physics-Informed Neural Networks, PINNs）、神经算子（Neural Operators）、基于深度学习的数值方法^[1] 等。首先，本文介绍了深度学习方法求解偏微分方程的研究背景和意义；其次，详细阐述了各类深度学习方法的基本原理、技术特点和适用范围；特别是针对神经算子方法中的 Transformer 架构与物理约束机制进行了深入探讨；然后，总结了这些方法在流体力学、固体力学、多物理场耦合等领域的应用；最后，分析了当前研究面临的挑战并展望了未来发展方向。本文旨在为相关领域的研究人员提供全面的技术参考，推动深度学习方法在科学计算领域的进一步发展。

关键词： 偏微分方程；深度学习；物理信息神经网络；神经算子；科学计算

Abstract

Partial Differential Equations (PDEs) serve as fundamental mathematical tools for describing natural phenomena and engineering problems, with widespread applications in fluid mechanics, solid mechanics, electromagnetics, and quantum mechanics. While traditional numerical methods such as Finite Difference Method (FDM), Finite Element Method (FEM), and Finite Volume Method (FVM) are mature and reliable, they face significant challenges regarding computational cost and accuracy when dealing with high-dimensional problems, complex geometries, and multi-scale phenomena. Recently, the rapid development of deep learning technology has provided new paradigms for solving PDEs. This paper systematically reviews the progress of deep learning methods for solving PDEs, primarily covering Physics-Informed Neural Networks (PINNs), Neural Operators, and deep learning-based numerical methods. Firstly, the background and significance of the research are introduced. Secondly, the basic principles, technical characteristics, and scope of application of various deep learning methods are elaborated, with a particular focus on the mathematical equivalence between Transformer architectures and integral operators in Neural Operators. Thirdly, applications in fluid mechanics, solid mechanics, and multi-physics coupling are summarized. Finally, current challenges are analyzed, and future development directions are proposed. This review aims to provide a comprehensive technical reference for researchers and promote the further development of deep learning in scientific computing.

Keywords: Partial Differential Equations; Deep Learning; PINNs; Neural Operators; Scientific Computing

目 录

第一章 引言	1
1.1 研究背景	1
1.2 研究意义	1
1.3 本文结构	2
第二章 深度学习方法求解偏微分方程基础理论	3
2.1 偏微分方程基本概念	3
2.2 传统数值方法回顾	3
2.3 深度学习方法的基本思想与分类	3
第三章 物理信息神经网络 (PINN)	5
3.1 PINN 基本原理	5
3.2 技术特点与训练病态	5
3.2.1 PINN 的核心技术优势	5
3.2.2 训练病态问题与优化挑战	7
3.3 改进方法	8
3.3.1 自适应权重与梯度均衡策略	8
3.3.2 自适应采样与残差驱动加密	9
3.3.3 多尺度网络架构与频域增强	9
3.3.4 域分解与并行训练框架	10
3.3.5 硬约束与变分形式的改进	10
第四章 神经算子学习：从函数逼近到算子映射	12
4.1 神经算子的理论范式	12
4.2 基于基函数展开的 DeepONet ^[2,3]	12
4.2.1 理论基础与算子逼近范式	12
4.2.2 双塔架构设计与前向传播机制	13
4.2.3 技术优势与泛化能力	13
4.2.4 局限性与挑战	14
4.3 频域视角与 Fourier Neural Operator ^[4,5]	14
4.3.1 从空间域到频域的算子学习范式转变	14
4.3.2 Fourier Layer 的核心机制	15
4.3.3 核心技术优势	16
4.3.4 几何局限性与改进方向	17
4.4 图/点云/几何神经算子 ^[6,7] ：复杂域的可泛化代理	17

4.4.1	从规则网格到非结构化表示	17
4.4.2	核心机制与技术优势	18
4.4.3	混合架构: GINO 与点云方法	18
4.5	注意力机制与积分算子的同构性	19
4.6	与物理约束的融合	20
第五章	其他深度学习方法	21
5.1	基于深度学习的数值方法 ^[1]	21
5.2	生成式模型与扩散 PDE 求解 ^[8,9]	22
5.2.1	从确定性回归到概率分布建模	22
5.2.2	早期探索: 对抗生成与变分推断	22
5.2.3	扩散模型: 随机微分方程的逆向求解	22
5.2.4	物理引导的扩散采样与超分辨率	23
5.3	几何与多尺度深度学习	23
第六章	应用领域	25
6.1	流体力学	25
6.2	固体力学	26
6.3	多物理场耦合	27
6.4	其他应用领域	28
第七章	挑战与展望	30
7.1	当前挑战	30
7.2	未来展望	30
第八章	结论	31
致谢	38	

第一章 引言

1.1 研究背景

偏微分方程 (PDE) 之所以在科学与工程中占据中心地位，是因为它们直接编码了守恒律、构成关系与场的传播机制。以流体为例，Navier-Stokes 方程通过动量守恒与连续性条件决定速度—压力场的时空演化；在固体力学中，平衡方程与本构关系共同决定应力与位移分布；而电磁与量子问题分别由 Maxwell 与 Schrödinger 方程给出全域耦合的动态规律。绝大多数真实系统并不存在解析解，这迫使研究者长期依赖离散化数值法。

传统网格法的优势非常明确：它们有严谨的误差理论、稳定性条件与可控的收敛性，并已沉淀为大量工业级软件。然而，随着研究从“单一工况的精确求解”转向“跨工况的大规模参数扫描、优化与实时决策”，数值法的代价问题被放大：在高维参数空间中，每一个参数点都要重新网格化与求解；对多尺度系统则需要局部极细网格才能捕捉薄层或激波结构；复杂几何下网格生成本身可能比求解更昂贵。对应地，许多应用场景（例如数字孪生、在线控制、不确定性量化/贝叶斯反演）要求在毫秒到秒级产生解场，传统求解器常无法满足时效。

深度学习的介入改变了“求解”这一概念。神经网络在本质上是高维函数逼近器，其参数化方式不依赖显式网格结构；只要能提供优化信号，就可以逼近 PDE 解或解算子。自从 Raissi 等提出 PINNs 后，“把 PDE 残差当作训练监督”的想法使得无标签求解成为可能，并迅速扩展到逆问题、参数识别与多物理场耦合。与此同时，算子学习路线在“大量仿真数据 → 学习整体解映射”的框架下得到突破，DeepONet 的结构化算子网络与 FNO 的频域卷积让模型具备离散无关 (resolution-invariant) 和跨参数泛化能力，显著改变了代理建模与实时推理的性能上限。

1.2 研究意义

从理论视角看，深度学习 PDE 求解把“数值分析中的逼近理论”与“机器学习的泛化理论”直接连接起来。PINNs、Deep Ritz、Deep Galerkin 等方法把残差最小化、能量泛函最小化与弱形式投影转化为可微优化问题，使神经网络成为新的函数空间基底。神经算子则从函数空间到函数空间学习映射，扩展了传统降阶模型 (ROM) 在非线性与高维情况下的适用边界，并激发了对算子逼近、误差上界与样本复杂度的系统研究。

从应用角度，深度学习方法提供两种可落地的价值：其一，在数据稀缺或仅有散点观测时，物理约束学习可以用极少数数据构造可信解场，从而支持反演、同化与

跨域推断; 其二, 在需要重复求解的参数化场景中, 神经算子或代理模型可把“每次求解分钟一小时级”压缩到“推理毫秒级”, 将优化与不确定性量化的总成本降到可接受范围。

1.3 本文结构

本文系统综述了深度学习方法求解偏微分方程的研究进展, 全文结构如下: 第 2 章介绍深度学习方法求解偏微分方程的基础理论; 第 3 章详细阐述物理信息神经网络方法; 第 4 章重点介绍神经算子方法及其演进; 第 5 章讨论生成式模型及其他深度学习方法; 第 6 章总结应用领域; 第 7 章分析挑战与展望; 第 8 章给出结论。

第二章 深度学习方法求解偏微分方程基础理论

2.1 偏微分方程基本概念

偏微分方程 (Partial Differential Equation, PDE) 是描述未知函数及其偏导数之间关系的数学方程，是刻画自然现象和工程问题中连续介质行为的基本数学工具。偏微分方程的一般形式可表示为：

$$F(x_1, x_2, \dots, x_n, u, \frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \dots, \frac{\partial^2 u}{\partial x_1^2}, \dots) = 0 \quad (2.1)$$

其中， u 表示未知函数， x_1, x_2, \dots, x_n 为自变量。偏微分方程的类型（椭圆型/抛物型/双曲型/混合型）不仅决定了其物理含义，更直接影响解的正则性特征与数值求解的难度。具体而言，Poisson/Helmholtz 等椭圆型方程对应稳态场问题，其解通常具有较好的光滑性；热传导方程等抛物型系统具有时间方向的耗散特性；而双曲型方程（如波动方程、可压缩流动方程）则常出现激波、间断或高梯度特征，这些特性正是深度学习求解器训练不稳定的重要原因之一。

2.2 传统数值方法回顾

传统数值方法求解偏微分方程的基本思路是将连续问题离散化。有限差分法 (FDM) 基于 Taylor 展开理论，采用差商近似偏导数；有限元法 (FEM) 基于变分原理，在单元内采用基函数近似未知函数；有限体积法 (FVM) 则基于控制体积的积分守恒。传统方法的误差主要来源于空间离散的截断误差、时间离散的截断误差以及迭代求解的数值误差。相比之下，深度学习方法采用神经网络 u_θ 作为解的参数化表示，其误差来源转变为函数逼近误差与优化误差。这种范式转换使得求解问题的核心从“离散稳定性分析”转变为“优化景观设计与梯度尺度匹配”。

2.3 深度学习方法的基本思想与分类

深度学习求解偏微分方程的核心思想，是将原本依赖数值离散与迭代求解的 PDE 问题转化为神经网络参数优化问题。通过以神经网络 u_θ 或算子映射 G_θ 参数化未知函数或解算子，求解过程不再依赖固定网格或显式离散格式，而是在连续空间中通过端到端优化自动学习 PDE 的解结构或解映射规律。这一范式的根本变化，使得深度学习方法能够在高维空间、复杂几何以及参数化 PDE 等传统数值方法难以处理的场景中展现显著优势。

根据物理信息利用方式与任务目标的不同，深度学习求解 PDE 的方法大体可分为以下三类：

(1) 数据驱动方法 (Data-Driven Methods) 数据驱动方法完全依赖高保真模拟数据或实验数据, 通过监督学习直接拟合输入与输出场之间的映射关系。例如, 卷积神经网络 (CNN) 或 U-Net 可用于流场重建^[10,11]、湍流超分辨率, Transformer 则用于捕捉长程依赖的物理过程。这类方法的优势在于推理速度极快, 适用于代理建模、实时仿真或快速参数扫描等任务。但由于缺乏显式物理约束, 其泛化能力依赖于训练数据覆盖范围, 一旦遇到分布外场景, 可能产生无法解释或物理不一致的预测结果。

(2) 物理信息驱动方法 (Physics-Informed Methods, 如 PINNs) 物理信息方法不依赖标签数据, 而是将 PDE 残差、初边值条件及其他物理约束直接融入损失函数, 从而在无网格、无标签的情况下逼近 PDE 解。此类方法通过自动微分精确计算偏导, 使网络在优化过程中自动满足控制方程与边界条件。PINNs 在逆问题求解、稀疏观测条件下的场重建、多物理场耦合等任务中表现突出。然而, 该类方法往往面临优化难、训练不稳定等问题, 如频谱偏置、梯度竞争及多尺度物理导致的病态优化景观等。

(3) 算子学习方法 (Operator Learning Methods, 如 DeepONet、FNO) 算子学习方法旨在学习 PDE 由输入函数空间到输出解空间的映射算子 $G^\dagger : \mathcal{A} \rightarrow \mathcal{U}$ 。通过大量参数化 PDE 的训练样本, 模型学习“解算子的整体规律”, 一次训练后即可对任意初边值条件或模型系数进行快速推理, 因此具备跨参数、跨几何甚至跨分辨率的泛化能力。典型方法包括 DeepONet 的双塔结构算子逼近和 Fourier Neural Operator (FNO) 的频域卷积算子。算子学习通过构建分辨率无关 (resolution-invariant) 的映射, 实现了真正意义上的 PDE 快速求解器 (Neural Solver), 在流体力学、固体力学与多物理耦合模拟中表现出极高性能。但其训练通常依赖大规模高保真数据, 外推性受到训练分布的限制。

第三章 物理信息神经网络 (PINN)

3.1 PINN 基本原理

物理信息神经网络 (PINNs) 由 Raissi 等人于 2019 年提出^[12]，其核心思想是将偏微分方程作为约束条件融入神经网络的损失函数中。考虑一般形式的偏微分方程：

$$\mathcal{N}[u](x, t) = 0, \quad (x, t) \in \Omega \times [0, T] \quad (3.1)$$

其中， \mathcal{N} 是微分算子， Ω 是空间域， $[0, T]$ 是时间域。边界条件和初始条件为：

$$u(x, t) = g(x, t), \quad (x, t) \in \partial\Omega \times [0, T] \quad (3.2)$$

$$u(x, 0) = h(x), \quad x \in \Omega \quad (3.3)$$

PINN 使用神经网络 $u_\theta(x, t)$ 近似解 $u(x, t)$ ，其中 θ 是网络参数。损失函数包括三部分：

$$L = L_{\text{PDE}} + L_{\text{BC}} + L_{\text{IC}} \quad (3.4)$$

其中， L_{PDE} 是方程残差项， L_{BC} 是边界条件项， L_{IC} 是初始条件项：

$$L_{\text{PDE}} = \frac{1}{N_f} \sum_{i=1}^{N_f} |\mathcal{N}[u_\theta](x_i^f, t_i^f)|^2 \quad (3.5)$$

$$L_{\text{BC}} = \frac{1}{N_b} \sum_{i=1}^{N_b} |u_\theta(x_i^b, t_i^b) - g(x_i^b, t_i^b)|^2 \quad (3.6)$$

$$L_{\text{IC}} = \frac{1}{N_0} \sum_{i=1}^{N_0} |u_\theta(x_i^0, 0) - h(x_i^0)|^2 \quad (3.7)$$

通过最小化损失函数，神经网络学习满足偏微分方程以及边界、初始条件的近似解。

3.2 技术特点与训练病态

3.2.1 PINN 的核心技术优势

PINN 方法在求解偏微分方程方面展现出若干显著的技术特点，这些特点共同构成了其区别于传统数值方法的核心优势。

精确的自动微分机制

PINN 的第一个关键优势在于其对自动微分 (Automatic Differentiation, AD) 技术的深度依赖^[12–14]。与传统数值微分方法 (如前向差分、中心差分) 相比，自动

微分基于链式法则, 通过计算图反向传播实现精确的梯度计算, 从根本上避免了数值微分固有的截断误差和舍入误差累积问题。对于高阶偏导数 (如 $\frac{\partial^2 u}{\partial x^2}$ 、 $\frac{\partial^2 u}{\partial x \partial t}$ 等), 自动微分能够提供机器精度级别的计算结果, 这对于构造精确的 PDE 残差项至关重要^[12,13]。更重要的是, 自动微分使得 PINN 能够处理任意复杂形式的微分算子, 无需手工推导导数表达式, 大大提升了方法的通用性和易用性^[13,14]。这种“自动化”特性使研究者能够快速将新的物理方程转化为可训练的神经网络模型。

无网格计算框架^[15]

PINN 方法的第二个核心优势体现在其无网格 (mesh-free) 特性上。不同于有限元法或有限体积法需要预先生成离散网格, 神经网络可以直接在连续域上的任意点进行前向传播和梯度计算。这一特性带来了多重实际价值: 首先, 完全避免了复杂几何形状下的网格生成难题, 特别适用于不规则边界、多连通域或移动边界问题; 其次, 可以灵活地在求解域内任意位置采样配点, 实现自适应采样策略, 重点关注残差较大或物理梯度剧烈变化的区域; 再次, 能够天然处理散点数据 (scattered data), 即观测数据可以分布在任意位置, 无需规整的网格结构, 这在实验测量场景中尤为重要; 最后, 对于高维问题 (如 $d > 3$ 的参数空间), 无网格特性避免了网格数量随维度指数增长导致的“维度灾难”, 使得 PINN 在处理高维偏微分方程时保持了相对可控的计算成本。

正逆问题的统一框架^[16]

PINN 方法的第三个重要特征是其采用统一的数学框架同时处理正问题与逆问题, 这是其区别于大多数传统数值方法的关键创新。对于正问题 (forward problem), 损失函数包含 PDE 残差项、边界条件项和初始条件项, 网络通过最小化这些项来学习满足物理定律的解场; 对于逆问题 (inverse problem), 只需在损失函数中增加数据拟合项, 通过稀疏观测数据反演未知参数 (如扩散系数、源项、初边值条件、甚至微分算子本身的形式)。这种统一的框架设计使得同一套代码可以处理多种类型的问题, 无需针对不同问题类型重新设计复杂的迭代算法。例如, 在参数识别问题中, 可以将未知参数作为可训练变量与网络参数一起优化; 在数据同化问题中, 可以同时利用物理方程和稀疏观测数据约束网络训练, 实现物理归纳偏置与数据驱动的有机结合。这种灵活性在不确定性量化、贝叶斯反演等场景中展现出巨大潜力。

数据高效性与物理一致性

PINN 方法的第四个核心优势体现在其数据高效性上, 即能够在数据极度稀缺的情况下仍能获得物理可信的解。传统的纯数据驱动方法 (如监督学习) 通常需要大量高质量的标签数据 (即已知的全场解), 而 PINN 方法将物理方程本身作为强监督信号, 理论上可以在完全没有标签数据的情况下进行训练, 仅依靠 PDE 残差约束即可逼近真实解。这种特性使得 PINN 特别适用于以下场景: 实验数据稀少或获取成本高昂的物理系统; 需要从稀疏、非均匀分布的观测数据推断全场解的问

题；参数识别和反问题，其中观测数据可能仅包含部分物理量的间接测量值；需要确保解在物理上合理（如满足守恒律、对称性等）的应用。然而，需要指出的是，虽然 PINN 理论上可以仅凭物理方程训练，但在实际应用中，即使是少量的观测数据往往也能够显著提升解的精度和训练稳定性，特别是在处理强非线性、多尺度或混合型偏微分方程时。

3.2.2 训练病态问题与优化挑战

尽管 PINN 具备上述诸多技术优势，但在实际应用中，“经典 PINNs 远非”万能求解器”。大量实证研究揭示了其在训练过程中面临的若干系统性病态问题，这些问题的本质源于“无网格连续优化”与“多项物理约束耦合”的内在矛盾。

频谱偏置导致的高频特征丢失

第一个核心挑战是神经网络固有的频谱偏置（spectral bias）问题^[17-20]。深度神经网络在梯度下降训练过程中天然倾向于优先学习低频分量，而对高频成分的拟合能力较弱且收敛速度极慢。这一现象的数学根源在于：网络参数的随机初始化导致其在频域上的“核函数”（Neural Tangent Kernel, NTK）在高频方向的特征值迅速衰减，从而使得高频梯度信号被抑制。具体到 PDE 求解场景，当真实解包含激波（shock）、边界层（boundary layer）、尖锐界面或高频震荡等结构时，PDE 残差项对这些局部高梯度特征的梯度反馈信号相对较弱，导致优化器会优先拟合解的光滑低频背景，而长期忽略高频细节。实验表明，对于包含间断的双曲型方程或高雷诺数 Navier-Stokes 流动，标准 PINN 往往在激波附近产生明显的数值耗散或欠拟合，解的相对误差可能达到 10%-50% 的量级，这严重限制了其在复杂物理问题中的实用性。

梯度竞争导致的优化失衡^[21]

第二个系统性问题源于多项损失函数之间的梯度竞争（gradient competition）^[17,22,23]。PINN 的总损失通常包含 PDE 残差项 L_{PDE} 、初始条件项 L_{IC} 、边界条件项 L_{BC} 以及可选的数据拟合项 L_{data} ，这些不同损失项的量纲、尺度和梯度范数往往存在数量级差异。例如，边界条件通常仅在低维流形 $\partial\Omega$ 上采样，其梯度信号较强且集中；而 PDE 残差项在高维求解域 Ω 上采样，梯度信号分散且可能被高阶导数的计算误差放大。当不同损失项的权重设置不当时，某一项的梯度会主导整个优化方向，导致网络在满足某些约束（如边界条件或数据点拟合）的同时牺牲域内物理残差的精度，或出现相反情况。这种现象在神经切空间核（Neural Tangent Kernel, NTK）理论框架下可以被理解为：不同损失项对应的核矩阵（Gram matrix）条件数差异巨大，导致梯度下降路径被病态的 Hessian 矩阵扭曲，收敛速度严重不均。实际训练中常观察到“损失曲线振荡”或“某一项损失停滞不降”的现象，这正是梯度竞争的直接体现。

多尺度物理导致的数值病态

第三个关键挑战来自多尺度偏微分方程中物理量的尺度差异导致的数值病态^[17,22]。许多真实物理系统涉及不同物理量(如速度、压力、温度、浓度等)或不同空间/时间尺度的耦合,这些物理量的量纲和数值范围可能相差数个量级。例如,在流固耦合问题中,流体压力可能为 $O(10^5)$ Pa,而固体位移为 $O(10^{-3})$ m;在对流扩散方程中,对流项与扩散项的比值(即 Péclet 数)可能高达 10^3 至 10^6 。若不对这些物理量进行适当的无量纲化或重标定(re-scaling),直接构造的损失函数会出现严重的数值病态:某些物理量的残差项在数值上占据绝对主导,导致其他项的梯度被淹没;同时,不同尺度的梯度会使得优化器(如 Adam、L-BFGS)的自适应学习率机制失效,收敛效率骤降甚至完全停滞。此外,多尺度耦合还会导致损失景观(loss landscape)中出现极端陡峭的峡谷(ravine)和平坦区域的交替,使得基于一阶梯度的优化算法难以找到有效的下降方向。实验表明,未经尺度归一化的多物理场 PINN 训练往往需要 10 倍甚至 100 倍的迭代次数才能达到相同精度,严重制约了方法的实用性。

综上所述,虽然 PINN 在理论上具备无网格、统一正逆问题框架、数据高效等诸多优势,但其训练过程中的频谱偏置、梯度竞争与多尺度数值病态等问题构成了当前研究的核心挑战。这些病态问题的存在促使研究者开发了大量改进策略,包括自适应权重调整、自适应采样、傅里叶特征嵌入、因果训练等技术,这将在下一节详细讨论。

3.3 改进方法

针对 PINN 训练过程中的频谱偏置、梯度竞争与多尺度数值病态等系统性挑战,研究者在近年来逐步形成了一套可复用的改进策略体系。这些改进方法可从损失函数设计、采样策略优化、网络架构增强、求解域分解以及约束形式转换等多个维度进行系统分类。

3.3.1 自适应权重与梯度均衡策略

第一类改进思路聚焦于损失函数中多项约束的权重动态调整,以解决梯度竞争导致的优化失衡^[21]问题。经典 PINN 采用固定的损失权重组合,这要求研究者预先对不同物理项的重要性进行人工标定,不仅繁琐且难以保证在整个训练过程中的适配性。为此,自适应权重方法应运而生,其核心思想是根据训练阶段动态调节各损失项的有效梯度贡献,使物理残差、初值条件、边界条件等多项约束在优化过程中保持同量级的梯度范数,从而显著提升收敛稳定性与最终精度。

具体而言,PINN-NTK 方法基于神经切空间核(Neural Tangent Kernel)理论,通过分析不同损失项对应的核矩阵特征值分布,动态调整权重以平衡收敛速度^[17,18]。Self-adaptive PINNs 则引入可学习的权重参数,将权重调整本身纳入端到端优化过程,通过梯度统计量(如梯度范数比值)自动调节各项权重^[22]。MultiAdam

等方法更进一步，为每一项损失配置独立的优化器与学习率调度器，使不同物理约束可以在各自适配的优化路径上演化，避免单一优化器无法同时适应多尺度梯度信号的困境^[20]。实验表明，这类方法在处理多物理场耦合问题时，可将训练迭代次数减少 50%-80%，同时显著降低解的相对误差。

3.3.2 自适应采样与残差驱动加密

第二类改进策略着眼于训练配点（collocation points）的空间分布优化，以克服固定采样导致的局部精度不足问题。传统 PINN 在训练过程中使用固定的配点分布（如均匀采样或拟蒙特卡洛采样），这种静态采样策略无法感知网络在不同区域的学习难度差异，导致在解变化剧烈的区域（如激波、边界层、尖锐界面）精度严重不足，而在光滑区域则存在过度采样的计算浪费。

残差自适应细化（Residual-based Adaptive Refinement, RAR）方法通过周期性评估当前网络在求解域内的 PDE 残差分布，将新的配点不断“拉向”残差大的区域，使网络重点学习激波、薄层等难点结构^[24,25]。这种策略本质上是一种无网格的自适应网格细化（Adaptive Mesh Refinement）思想的神经网络版本。Gaussian-mixture 采样则假设残差分布可由多个高斯分量的混合建模，通过期望最大化(EM) 算法自适应更新采样分布的均值与方差，实现对多个高残差区域的同步关注^[22]。Adversarial sampling 进一步引入对抗学习思想，将配点生成器视为对抗网络中的“生成器”，以最大化残差为目标生成新样本，迫使 PINN 网络（“判别器”）不断提升对困难样本的学习能力^[26]。实践表明，结合自适应采样的 PINN 可在相同配点总数下将解精度提升 1-2 个数量级，特别是在处理双曲型方程的间断解时表现出明显优势。

3.3.3 多尺度网络架构与频域增强

第三类改进方向通过重新设计网络架构或特征编码方式，直接缓解神经网络固有的频谱偏置问题。研究表明，标准全连接网络在随机初始化后，其神经切空间核在高频方向的特征值呈指数衰减，导致网络对高频信号的学习速度极慢。为此，研究者提出了多种多尺度与频域增强结构。

Multi-scale PINN 采用多分支网络架构，每个分支专注于不同空间或时间尺度的特征提取，最后通过融合层组合各尺度信息得到最终解^[17,27]。这种设计类似于计算机视觉中的特征金字塔网络（FPN），特别适用于湍流、多相流等本质上多尺度的物理问题。Fourier Feature Embedding 则在输入层引入随机傅里叶特征，通过高阶正弦/余弦基函数对输入坐标进行非线性映射，从而将低频输入空间提升到高维高频特征空间，显著增强网络对高频分量的表达能力^[17]。SIREN (Sinusoidal Representation Networks) 用正弦激活函数替代传统的 ReLU 或 Tanh，使网络天然具备周期性特征表示能力，其梯度在频域上的分布更加均匀^[18]。LAAF (Locally

Adaptive Activation Functions) 进一步引入可学习的局部激活函数参数, 允许网络在不同区域自适应调整频率响应特性^[28]。实验验证表明, 这类方法在处理包含高频震荡或尖锐梯度的 PDE 时, 可将残差误差降低 2-3 个数量级。

3.3.4 域分解与并行训练框架

第四类改进策略借鉴传统数值方法中的区域分解思想, 通过将大规模或强非线性求解域拆分为多个子域, 化“一个难优化问题”为“多个易优化子问题”, 同时天然适配并行计算架构。

XPINN (eXtended PINN) 将求解域 Ω 分解为若干重叠或非重叠子域 $\{\Omega_i\}$, 每个子域由独立的神经网络 u_{θ_i} 求解, 子域之间通过界面连续性条件 (如 Dirichlet-Neumann 交替或 Robin 边界条件) 进行耦合^[27]。这种方法不仅降低了单个网络的优化难度, 还允许对不同子域采用不同的网络架构或训练策略, 实现局部自适应。FBPINN (Finite Basis PINN) 进一步结合有限元思想, 在每个子域内使用紧支撑的窗函数与神经网络乘积作为局部解表示, 通过子域边界处的弱连续性条件组装全局解, 这种设计使得网络可以在子域内独立优化而不受远场干扰^[29]。实验表明, 域分解 PINN 在处理长时间演化、复杂几何或多物理场耦合问题时, 可实现近线性的并行加速比, 同时显著提升解的全局一致性。

3.3.5 硬约束与变分形式的改进

第五类改进方法从约束施加的方式入手, 通过改变物理定律的编码形式来提升训练稳定性与解的物理一致性。

硬约束物理信息神经网络^[30] (Hard-constrained PINN)

传统 PINN 通过损失函数对边界条件与初始条件进行软约束 (soft constraint), 这意味着网络在训练过程中只能“逐步逼近”满足这些条件, 而无法精确满足。这种软约束方式在某些对边界条件敏感的问题 (如势流问题、不可压缩流动) 中可能导致非物理解。硬约束方法通过显式构造满足边界条件的网络结构, 将物理约束硬编码 (hard-code) 到网络表示中。例如, 对于 Dirichlet 边界条件 $u|_{\partial\Omega} = g$, 可以将网络输出构造为 $u_{\theta}(x) = g(x) + d(x) \cdot \tilde{u}_{\theta}(x)$, 其中 $d(x)$ 是距离函数 (在边界上为零, 在域内大于零), \tilde{u}_{θ} 是自由神经网络。这种构造确保边界条件在任何训练阶段都精确满足, 从而消除了边界误差项对优化的干扰, 同时避免了软约束权重调优的繁琐。

变分物理信息神经网络 (Variational PINN)

当 PDE 解不够光滑 (如含弱间断) 或输入数据包含噪声时, 强形式 PINN 所依赖的逐点残差匹配需要精确计算高阶导数, 这会显著放大自动微分过程中的数值噪声, 导致训练极其脆弱甚至失败。弱形式与变分 PINN 的核心思想是将 PDE

转化为变分形式或弱形式，用加权积分（通过引入测试函数 v ）替代逐点导数匹配，从而降低对解的光滑性要求，同时使训练信号对局部不光滑更具鲁棒性^[31-33]。

具体而言，对于椭圆型方程 $-\nabla \cdot (a \nabla u) = f$ ，其弱形式为 $\int_{\Omega} a \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx$ 。变分 PINN 直接将该双线性形式作为损失函数，通过选择合适的测试函数集（如随机采样的平滑函数、多项式基或神经网络本身）进行离散化。这种方式的优势在于：（1）只需计算一阶导数，避免了强形式中二阶导数引入的数值不稳定；（2）对解的不连续性具有天然容忍度，因为积分形式本身就是对局部信息的平滑化；（3）自然处理 Neumann 边界条件，无需额外的边界损失项。hp-VPINN (hp-Variational PINN) 进一步结合域分解与高阶多项式测试基，在每个子域内使用高阶 Legendre 多项式作为测试函数，通过 h -细化（子域数量增加）与 p -细化（多项式阶数提升）的自适应策略，使方法在复杂几何和多尺度椭圆/抛物方程中表现出更高的稳定性与收敛性^[27,31]。

此外，对于某些物理问题（如最小曲面问题、弹性力学），其自然形式就是能量泛函最小化（如 Dirichlet 能量、应变能）。此时变分 PINN 可直接优化能量泛函而非 PDE 残差，这种 Deep Ritz 方法在理论上等价于有限元法的 Ritz-Galerkin 变分，但用神经网络替代了传统的多项式有限元基函数，从而获得了无网格与高维扩展性的优势^[34]。

第四章 神经算子学习：从函数逼近到算子映射

4.1 神经算子的理论范式

传统的物理信息神经网络 (PINNs) 虽然在求解特定偏微分方程 (PDE) 实例方面表现出色，但其本质上属于“实例求解器”，一旦方程的参数、边界条件或源项发生变化，网络必须重新训练。为了克服这一局限，神经算子 (Neural Operators) 作为一种新的科学计算范式被提出，其目标是将神经网络的学习对象从单一的函数解提升为无限维函数空间之间的非线性算子。数学上，若设 \mathcal{A} 和 \mathcal{U} 分别为输入函数空间 (如初始场或系数场) 和输出解空间，PDE 的求解过程可被视为算子映射 $G^\dagger : \mathcal{A} \rightarrow \mathcal{U}$ 。神经算子旨在通过参数化模型 G_θ 来逼近该算子，根据广义算子逼近定理 (Generalized Universal Approximation Theorem)，在网络宽度和深度充足的前提下，神经网络能够以任意精度逼近任意连续非线性算子。该范式最显著的特征在于其“离散无关性” (Discretization Invariance)：模型学习的是连续空间上的映射规则，因此在训练完成后，可直接应用于任意分辨率的网格进行推理，这为多尺度物理模拟提供了理论基础。主流的神经算子架构通常借鉴格林函数法的思想，采用“迭代积分”的形式更新特征场 $v(x)$:

$$v_{l+1}(x) = \sigma \left(W v_l(x) + \int_D \kappa(x, y; \phi) v_l(y) dy \right) \quad (4.1)$$

其中积分核 κ 的参数化方式决定了不同算子模型的本质区别。

4.2 基于基函数展开的 DeepONet^[2,3]

4.2.1 理论基础与算子逼近范式

作为算子学习领域的开创性工作，DeepONet (Deep Operator Network) 由 Lu 等人于 2019 年提出^[35]，标志着深度学习从“学习单个函数”向“学习函数空间之间的映射算子”的范式转变^[35,36]。其理论基础源于 Chen 和 Chen 于 1995 年建立的通用算子逼近定理 (Universal Approximation Theorem for Operators)，该定理证明了任意连续非线性算子 $G^\dagger : \mathcal{A} \rightarrow \mathcal{U}$ 都可以被一个具有足够宽度或深度的神经网络以任意精度逼近^[37]。这一理论保证为算子学习提供了坚实的数学根基，使得 DeepONet 能够在原则上逼近 PDE 求解这一本质上的算子映射过程。

具体而言，算子逼近理论表明，任意算子 $G(a)(y)$ 可被近似为两组函数的内积形式：

$$G(a)(y) \approx \sum_{k=1}^p b_k(a) \cdot \phi_k(y) \quad (4.2)$$

其中, $\{b_k(a)\}$ 是依赖于输入函数 a 的系数, $\{\phi_k(y)\}$ 是定义在输出空间坐标 y 上的一组基函数。这种可分解表示 (separable representation) 正是 DeepONet 架构设计的核心思想: 通过神经网络分别学习系数映射与基函数集, 从而实现对整个算子族的参数化表示。

4.2.2 双塔架构设计与前向传播机制

DeepONet 采用了一种创新的双塔结构 (Branch-Trunk Architecture), 将算子的输入与输出空间解耦处理, 这种设计精巧地对应了上述算子分解的数学形式。

分支网络 (Branch Net): 该网络负责处理输入函数 $a(x)$ 的离散观测值。在实际应用中, 输入函数通常在有限个传感器位置 $\{x_1, \dots, x_m\}$ 进行采样, 得到离散向量 $\mathbf{a} = [a(x_1), \dots, a(x_m)]$ 。分支网络通过多层感知机 (MLP) 或卷积神经网络 (CNN) 提取输入场的全局特征, 输出 p 维系数向量 $\mathbf{b}(a) = [b_1(a), \dots, b_p(a)]$ 。这组系数本质上编码了“输入函数如何影响最终解”的全局信息。

主干网络 (Trunk Net): 该网络负责处理查询点的时空坐标 $y = (x, t)$, 通过另一组独立的多层感知机学习解空间的一组基函数, 输出 p 维基向量 $\mathbf{t}(y) = [t_1(y), \dots, t_p(y)]$ 。这组基函数可以理解为数据驱动的“解空间基底”, 类似于传统数值方法中的有限元基函数或傅里叶基, 但其形式完全由数据学习得到, 无需预先指定。

输出重构: 给定输入函数 a 和查询坐标 y , DeepONet 通过分支网络与主干网络输出的内积重构解值:

$$G_\theta(a)(y) \approx \sum_{k=1}^p b_k(a) \cdot t_k(y) + b_0 \quad (4.3)$$

其中, b_0 是偏置项。这种内积形式不仅在数学上优雅, 更重要的是实现了输入函数空间与输出坐标空间的完全解耦: 分支网络的参数仅取决于输入采样方式, 与输出查询分辨率无关; 主干网络则与输入采样无关, 可在任意查询点进行评估。这种解耦特性使得 DeepONet 天然具备分辨率无关性 (resolution invariance), 即训练完成后可在任意精度的网格上进行推理。

4.2.3 技术优势与泛化能力

DeepONet 的核心优势体现在其对参数化 PDE 族的一次性学习能力上。通过在大量不同初边值条件、源项或材料系数下的 PDE 解上进行训练, DeepONet 学习到的是“解算子的整体规律”, 而非某个特定实例的解。这种范式转变带来了显著的实用价值: 一旦训练完成, 模型即可对任意新的输入函数 (如新的初值场 $u_0(x)$ 、新的系数场 $\kappa(x)$) 进行快速推理, 推理成本仅为一次前向传播 (通常毫秒级), 远低于传统求解器的分钟至小时级计算^[35,36]。

这种算子级别的泛化能力在多个实际应用中得到验证: 在孔隙介质渗流问题中, DeepONet 可学习从随机渗透率场到压力场的映射, 支持快速的不确定性量化; 在弹性参数化结构分析中, 模型能够处理不同几何形状和材料分布下的应力-应变响应; 在反应扩散系统中, DeepONet 展现出跨参数(如反应速率、扩散系数)的稳健泛化性能^[35,36]。此外, 由于双塔结构的设计, DeepONet 对几何形状具有极强的适应性——主干网络可以在任意形状的定义域内进行查询, 而无需重新训练或调整网络结构, 这在传统基于卷积的方法中是难以实现的。

4.2.4 局限性与挑战

尽管 DeepONet 在理论上具备通用算子逼近能力, 但在实际应用中仍面临若干技术挑战。

首先, 分支网络的参数规模问题较为突出。由于分支网络需要处理输入函数在所有传感器位置的离散值, 当输入函数的采样分辨率较高时(如 $m = 256 \times 256$ 的二维场), 即使采用全连接层, 参数量也会随输入维度线性甚至超线性增长。若采用卷积网络作为分支网络, 虽然参数共享能够一定程度缓解此问题, 但会引入平移不变性假设, 可能不适用于某些非均匀物理场。

其次, 基于 MLP 的基函数表示存在频谱偏置(Spectral Bias)问题。主干网络使用的多层感知机在捕捉高频震荡特征时表现不佳, 这与 PINN 面临的问题类似。当 PDE 解包含尖锐梯度、激波或边界层等高频成分时, 标准 DeepONet 往往需要极大的基函数数量(即 p 值)才能达到满意的精度, 这不仅增加了计算成本, 也可能导致训练不稳定。

最后, DeepONet 的训练依赖大量高质量的成对数据 $\{(a_i, u_i)\}$, 其中每个 u_i 都是输入 a_i 对应的 PDE 完整解场。这要求预先使用传统求解器生成大规模训练数据集, 数据生成成本可能非常高昂。此外, 当输入函数空间的维度极高或参数分布复杂时, 如何高效采样训练数据以覆盖输入空间的关键模式, 仍是一个开放性问题。

尽管存在上述局限, DeepONet 所建立的算子学习框架为后续方法(如 Fourier Neural Operator、Graph Neural Operator 等)提供了重要启发, 成为神经算子研究的重要里程碑。

4.3 频域视角与 Fourier Neural Operator^[4,5]

4.3.1 从空间域到频域的算子学习范式转变

DeepONet 虽然通过双塔结构实现了算子的参数化表示, 但其本质上仍在空间域进行积分计算, 这导致两个关键瓶颈: 其一, 全局积分 $\int_D \kappa(x, y)v(y)dy$ 的计算复杂度为 $O(N^2)$ (N 为空间离散点数), 在高分辨率网格上难以承受; 其二, 空间域的

核函数 $\kappa(x, y)$ 通常需要对所有点对进行参数化，参数量随网格规模平方增长。为了从根本上突破这一计算瓶颈，Li 等人于 2020 年提出了 Fourier Neural Operator (FNO)^[38,39]，创新性地将核积分算子的计算迁移至频域，利用卷积定理实现了计算效率与参数规模的双重突破^[38,39]。

卷积定理 (Convolution Theorem) 揭示了空间域卷积与频域点乘之间的等价关系：

$$(f * g)(x) = \mathcal{F}^{-1}\{\mathcal{F}(f) \cdot \mathcal{F}(g)\}(x) \quad (4.4)$$

其中， \mathcal{F} 表示傅里叶变换， $*$ 表示卷积运算， \cdot 表示逐点乘法。这意味着，原本需要在空间域进行的全局积分，可以通过“傅里叶变换 → 频域乘法 → 傅里叶逆变换”三步操作等价实现。更关键的是，借助快速傅里叶变换 (Fast Fourier Transform, FFT) 算法，这一过程的计算复杂度可从 $O(N^2)$ 降低至准线性级别 $O(N \log N)$ ，这为构建高效的全局算子学习模型奠定了理论基础。

4.3.2 Fourier Layer 的核心机制

FNO 的核心创新在于其 Fourier Layer 的设计，该层将传统的空间域卷积操作完全替换为频域操作。每个 Fourier Layer 包含三个关键步骤的串行处理流程：

步骤一：傅里叶变换 (Forward FFT)

对输入特征场 $v_l(x) \in \mathbb{R}^{N \times d}$ (其中 N 为空间网格点数， d 为通道数) 执行快速傅里叶变换，获取其频域表示：

$$\hat{v}_l(k) = \mathcal{F}(v_l)(k) = \int_D v_l(x) e^{-2\pi i k \cdot x} dx \quad (4.5)$$

其中， k 为频率模态索引， $\hat{v}_l(k)$ 为复数频谱。FFT 算法可在 $O(N \log N)$ 复杂度内完成此变换。

步骤二：频域滤波与线性变换 (Spectral Convolution)

在频域内，FNO 仅保留低频模态 (Low-frequency modes) 进行处理，通过截断高频成分实现隐式正则化。具体而言，对于频率索引 $|k| \leq k_{\max}$ (k_{\max} 为截断频率)，使用可学习的复数权重张量 $R \in \mathbb{C}^{k_{\max} \times d \times d'}$ 对保留的频谱进行线性变换：

$$\hat{v}'_l(k) = R(k) \cdot \hat{v}_l(k), \quad \text{for } |k| \leq k_{\max} \quad (4.6)$$

这一步骤在数学上等价于在空间域进行全局卷积，但由于仅参数化低频模态，参数量从 $O(N^2)$ 降低至 $O(k_{\max} \cdot d \cdot d')$ ，与网格分辨率 N 完全解耦。对于高频部分 $|k| > k_{\max}$ ，通常直接截断或置零，这相当于对解场施加了光滑性先验。

步骤三：傅里叶逆变换 (Inverse FFT)

通过快速傅里叶逆变换 (IFFT) 将处理后的频谱还原回空间域：

$$\tilde{v}_l(x) = \mathcal{F}^{-1}(\hat{v}'_l)(x) = \int_K \hat{v}'_l(k) e^{2\pi i k \cdot x} dk \quad (4.7)$$

为了增强模型的局部特征提取能力, FNO 在频域操作之外还保留了一个空间域的局部线性层 $Wv_l(x)$ (通常为逐点 1×1 卷积), 并通过非线性激活函数 σ 进行融合:

$$v_{l+1}(x) = \sigma(Wv_l(x) + \mathcal{F}^{-1}(R \cdot \mathcal{F}(v_l))(x)) \quad (4.8)$$

这种“全局频域卷积 + 局部空间卷积”的组合设计, 使得 FNO 既能捕捉长程依赖, 又能保持对局部细节的敏感性。

4.3.3 核心技术优势

FNO 的频域设计带来了三个层面的显著优势, 使其在算子学习领域迅速成为主流方法之一。

全局感受野与长程依赖建模

由于傅里叶变换是全局操作, 每个频域系数都包含了整个空间域的信息, 因此单层 Fourier Layer 即可获得理论上的全局感受野。这一特性在处理流体力学、电磁场等具有强非局部耦合的物理问题时至关重要。例如, 在不可压缩流动中, 压力场通过 Poisson 方程与全域速度散度耦合; 在湍流中, 大尺度涡结构对小尺度耗散有长程影响。传统的卷积神经网络需要堆叠多层才能扩大感受野, 而 FNO 天然具备这一能力, 因此在学习低频主导、全局耦合显著的 PDE (如浅水方程、大气海洋环流) 时展现出明显优势^[39-41]。

准线性计算复杂度

利用 FFT/IFFT 的 $O(N \log N)$ 复杂度, FNO 将全局卷积的计算成本从 $O(N^2)$ 降低至准线性级别。这一突破使得 FNO 能够在高分辨率网格 (如 256×256 甚至 512×512) 上高效训练和推理。相比之下, 直接在空间域参数化积分核的方法在此类分辨率下几乎不可行。此外, FFT 算法已在现代硬件 (如 GPU、TPU) 上得到高度优化, 可充分利用并行计算资源, 进一步提升实际运行效率。

分辨率无关性与零样本超分辨率

FNO 最具革命性的特性在于其对网格分辨率的无关性 (Resolution Invariance)。由于模型参数 R 仅依赖于频率模态索引 k , 而与具体的空间离散方式无关, 因此训练完成的 FNO 可以直接应用于任意分辨率的网格进行推理, 无需重新训练或微调。例如, 在 64×64 网格上训练的模型, 可以零样本 (zero-shot) 泛化到 256×256 甚至更高分辨率的网格上进行预测, 这种能力在传统基于卷积的模型中是无法实现的。实验表明, FNO 在跨分辨率测试时通常能保持较高的预测精度, 相对误差增幅仅为 5%-15%, 远优于需要在目标分辨率重新训练的方法^[38,39]。这一特性使得 FNO 特别适合多尺度物理模拟和自适应网格细化场景。

4.3.4 几何局限性与改进方向

尽管 FNO 在规则网格上的流场预测中取得了最先进 (SOTA) 的性能，但其对快速傅里叶变换的深度依赖也导致了严重的几何局限性，这成为制约其在工程实际中广泛应用的主要瓶颈。

几何适用性的固有约束

标准 FFT 算法假设数据定义在周期性边界条件的规则网格上（如矩形域、环面等），这使得 FNO 仅能直接应用于此类几何形状。对于工程中常见的复杂几何（如机翼型面、管道弯头、多孔介质内部结构），标准 FNO 无法直接处理。虽然可以通过零填充（zero-padding）将不规则域嵌入到矩形域中，但这会引入大量无效计算并破坏边界条件的精确性。

复杂几何的处理策略

为了克服这一局限，研究者提出了多种几何扩展方案。Geo-FNO 通过坐标映射（coordinate mapping）技术，将物理域 Ω_{phys} 变换至标准计算域 Ω_{comp} （如单位正方形），在计算域上应用 FNO 后再映射回物理域，从而间接处理不规则几何^[40]。然而，这种方法对映射质量敏感，在存在极端变形或拓扑复杂性时可能失效。Deformation-FNO 进一步引入可学习的变形场，允许网络自适应调整频域表示以适应几何形状^[40]。GINO（Geometry-Informed Neural Operator）则结合了图神经网络与傅里叶卷积，在结构化区域使用 FNO 高效捕捉全局模式，在不规则边界附近使用图卷积处理局部几何细节，实现了几何灵活性与计算效率的平衡^[40]。

局部奇异性的挑战

另一个技术挑战源于 FNO 的频域截断策略。虽然保留低频模态能够有效捕捉解的光滑全局结构，但当解包含强局部奇异性（如激波、接触间断、边界层）时，这些高频特征会被截断滤波器抑制，导致精度显著下降。为此，研究者提出将 FNO 与局部算子结合：例如，在激波区域使用自适应小波基或局部精细化的图卷积捕获高梯度特征，在光滑区域使用 FNO 保持计算效率，通过多尺度融合策略获得全局-局部平衡的解表示^[39,40]。

综上所述，FNO 通过频域操作实现了算子学习在效率、泛化性和表达能力上的重大突破，但其几何适用性仍需进一步扩展。当前的改进方向正朝着“频域高效性 + 几何灵活性 + 物理约束”的融合架构演进，为构建下一代通用物理求解器奠定基础。

4.4 图/点云/几何神经算子^[6,7]: 复杂域的可泛化代理

4.4.1 从规则网格到非结构化表示

FNO 虽然在规则网格上实现了高效的全局算子学习，但其对快速傅里叶变换的依赖使其难以处理工程实际中广泛存在的复杂几何。真实物理系统往往定义在

高度不规则的空间域上: 机翼型面、车身外形、器官结构、多孔介质等, 这些几何形状无法嵌入规则网格而不引入巨大计算浪费。传统 CFD 与 FEA 早已转向非结构化网格 (Unstructured Grids), 通过三角形、四面体等单元灵活拟合复杂边界。因此, 构建能够天然处理非结构化数据的神经算子成为算子学习走向工程实用的关键。

图神经算子 (Graph Neural Operator, GNO) 的核心思想是将物理域的离散化表示从“规则网格阵列”转变为“图结构 $G = (V, E)$ ”, 通过图神经网络的消息传递机制将连续核积分算子近似为图上的邻域聚合操作^[40,42]。

4.4.2 核心机制与技术优势

回顾连续核积分算子 $(G_\kappa v)(x) = \int_D \kappa(x, y, v(x), v(y))v(y)dy$, 在非结构化网格上将积分近似为邻域求和:

$$(G_\kappa v)(x_i) \approx \sum_{j \in \mathcal{N}(i)} \kappa(x_i, x_j, v(x_i), v(x_j))v(x_j)w_{ij} \quad (4.9)$$

GNO 的单层更新通过消息传递实现:

$$v_{l+1}(x_i) = \sigma \left(W v_l(x_i) + \sum_{j \in \mathcal{N}(i)} \kappa_\theta(v_l(x_i), v_l(x_j), x_i - x_j) v_l(x_j) \right) \quad (4.10)$$

其中, $\mathcal{N}(i)$ 为节点 i 的邻域, κ_θ 是参数化的核函数 (通过神经网络学习), W 为局部线性变换。

GNO 的核心优势体现在: (1) 几何自适应性: 可直接处理任意复杂几何形状、多连通域、曲面流形; (2) 置换不变性: 节点编号改变不影响输出, 通过编码相对位置实现几何变换不变性; (3) 可变分辨率: 对输入节点数量无固定要求, 支持自适应网格与多尺度细化。

4.4.3 混合架构: GINO 与点云方法

GNO 的局部邻域聚合机制在捕捉长程依赖时效率较低, 需堆叠多层才能扩大感受野。为结合 FNO 的全局高效性与 GNO 的局部几何适应性, GINO (Geometry-Informed Neural Operator) 将物理域分解为结构化区域与非结构化区域: 在流场主体使用 FNO 高效捕捉全局模式, 在几何复杂的边界附近使用图卷积处理局部细节, 通过插值或注意力机制在两种网格间传递特征^[40,42]。实验表明, GINO 在大规模复杂外流场预测 (如汽车空气动力学、飞机绕流) 中取得了 SOTA 性能, 相比纯 FNO 在复杂几何案例中误差降低 30%-50%, 相比纯 GNO 计算速度提升 5-10 倍^[40,42]。

此外，PCNO (Point Cloud Neural Operator) 直接在点云上学习算子映射，通过置换不变聚合完全摆脱显式拓扑连接，特别适用于粒子方法加速与形状优化问题^[42]。MeshGraphNets 则在网格拓扑上构建多层图编码不同尺度几何信息，在布料仿真、流固耦合等任务中实现接近传统求解器的精度且推理速度提升 2-3 个数量级^[42]。

针对工业中的非结构网格与变化几何，图神经算子通过图消息传递或点云不变性实现对“几何变化 + 参数变化”的联合泛化，正在成为 CFD、结构分析、多物理场耦合快速代理的主流形态^[40,42]。

为了应对工程实际中广泛存在的非结构化网格 (Unstructured Grids) 和复杂拓扑结构，图神经算子 (GNO) 被提出。该方法将物理域离散化为图结构 $G = (V, E)$ ，利用图神经网络的消息传递机制 (Message Passing) 来近似核积分运算^[42,43]。

GNO 天然具备处理任意复杂几何形状的能力，具有极强的几何适应性。最新的 GINO (Geometry-Informed Neural Operator) 融合了 FNO 的全局高效性与 GNO 的局部几何适应性，在大规模复杂外流场预测（如汽车空气动力学模拟）中取得了当前最优 (SOTA) 的性能，展现了混合架构的潜力。

针对工业中的非结构网格与变化几何，图神经算子 (GNO)、MeshGraphNets、PCNO、GINO 等通过图消息传递或点云不变性把非规则域嵌入隐空间，再与全局算子耦合，实现对“几何变化 + 参数变化”的联合泛化。^[40,42] 这类方法正在成为 CFD/结构/多物理耦合快速代理的主流形态。^[40,42]

4.5 注意力机制与积分算子的同构性

随着 Transformer 在序列建模中的成功，研究者开始探索注意力机制在算子学习中的物理可解释性。数学上，Self-Attention 机制中的 Attention Map 可以被严格解释为一种数据驱动的动态积分核 $\kappa(x, y)$ 。与 FNO 中固定的平移不变卷积核不同，注意力机制能够根据输入特征 (Query 和 Key) 动态调整积分权重：

$$\text{Attention}(Q, K, V) \approx \int \kappa(x, y)v(y)dy \quad (4.11)$$

这种特性使得基于 Transformer 的算子（如 Galerkin Transformer, OFormer）在处理非结构化网格和捕捉局部激波、边界层等高梯度特征时表现出优越的适应性。特别是对于具有多尺度特性的物理过程，注意力机制能够自适应地聚焦于关键物理区域，弥补了卷积网络在处理非均匀分布特征时的不足。这一方向的研究不仅统一了序列建模与算子逼近的视角，也为后续结合物理约束的改进模型（如 ReAttention-PINNs）提供了坚实的理论支撑。

4.6 与物理约束的融合

针对工程实际中广泛存在的复杂拓扑和数据稀缺问题, 算子学习正向着几何感知和物理驱动两个维度演进。图神经算子 (GNO) 通过将物理域离散化为图结构, 利用消息传递机制 (Message Passing) 实现非欧氏空间上的积分近似, 彻底解决了复杂几何的适配问题。与此同时, 为了降低对高保真模拟数据的依赖, 物理约束神经算子 (PINO) 借鉴了 PINN 的思想, 将 PDE 残差方程作为正则化项引入损失函数。这种“数据 + 物理”的双驱动模式, 不仅实现了小样本甚至零样本条件下的算子学习, 还有效修正了纯数据驱动模型可能产生的非物理伪影 (如质量不守恒), 确保了预测结果符合物理守恒律。未来的研究趋势表明, 深度融合 Transformer 的强表达能力、FNO 的计算效率以及物理定律的归纳偏置, 将是构建下一代通用物理基础模型的关键路径。

第五章 其他深度学习方法

5.1 基于深度学习的数值方法^[1]

这类方法将深度学习与传统数值方法结合，用神经网络改进数值方法的某些环节。例如，学习型有限差分方法用神经网络学习差分格式的系数，PDE-Net 使用可学习的卷积核近似微分算子，通过约束确保卷积核的微分性质。此外，Solver-in-the-Loop 方法将可微分的物理求解器与神经网络结合，通过端到端训练优化数值格式。这种混合范式保留了数值方法的稳定框架，仅将关键但难以手工设计的部件（如通量形式、预条件器）交给网络学习。这类方法将深度学习与传统数值方法结合，用神经网络改进数值方法的某些环节，发挥传统方法的成熟性和深度学习的灵活性。

学习型有限差分方法用神经网络学习差分格式的系数，可以根据问题的特点自适应调整，提高精度和稳定性^[44]。有限差分法用差商近似导数，差分格式的系数决定了精度和稳定性。学习型有限差分用神经网络学习差分格式的系数，可以根据问题的特点自适应调整，提高精度和稳定性。Long 等人提出的 PDE-Net 使用可学习的卷积核近似微分算子，通过约束确保卷积核的微分性质^[44]。这种方法可以同时识别偏微分方程和求解，为数据驱动的偏微分方程研究提供工具^[44]。

学习型有限元方法用神经网络优化基函数^[45]，可以学习问题特定的基函数，提高计算效率。有限元法使用基函数近似未知函数，基函数的选择影响精度和效率。学习型有限元用神经网络优化基函数，可以学习问题特定的基函数，提高计算效率。Alet 等人提出的图元网络（Graph Element Network）模仿有限元分析的行为，使用图神经网络建模空间关系，节点位置和连接都可以优化。

学习型有限体积方法用神经网络预测数值通量，可以学习更精确的通量格式，同时保持守恒性^[46]。有限体积法通过数值通量近似边界积分，数值通量的选择影响精度和守恒性。学习型有限体积用神经网络预测数值通量，可以学习更精确的通量格式，同时保持守恒性。Bar-Sinai 等人使用卷积神经网络生成有限体积法的系数，在低分辨率网格上也能产生准确结果^[46]。

学习型数值格式方法用神经网络设计全新的数值格式。除了改进传统数值方法，还可以用神经网络设计全新的数值格式。Um 等人提出的 Solver-in-the-Loop 方法将可微分的物理求解器与神经网络结合，通过端到端训练优化数值格式。这种方法可以学习最适合问题的数值格式。

学习型 PDE 方法并不总是要“重写求解器”。PDE-Net、Solver-in-the-loop、可学习通量与可学习预条件器的思路是：保留数值法的稳定框架，只把关键但难

手工设计的部件（差分核、通量形式、迭代更新/预条件）交给网络学习。^[44,47] 例如 PDE-Net 用可学习卷积核近似微分算子，并通过结构约束保证卷积核的导数一致性，从而能同时完成方程识别与求解；学习型有限体积通量在粗网格上仍可保持较高精度与守恒；而神经预条件器能显著减少 Krylov/Newton 迭代步数。^[22,26,48]

5.2 生成式模型与扩散 PDE 求解^[8,9]

5.2.1 从确定性回归到概率分布建模

传统的基于神经网络的 PDE 求解方法（如 PINNs 和 DeepONet）通常被构建为确定性的回归问题，即寻找一个单一的最优解函数以最小化残差或数据误差。然而，在面对反问题求解、高雷诺数湍流模拟以及数据稀缺场景下的不确定性量化（Uncertainty Quantification, UQ）时，点估计方法往往难以捕捉解的多模态分布特性。生成式模型（Generative Models）的引入标志着求解范式从“寻找单一解”向“模拟解的条件概率分布”转变。通过学习数据分布 $p(u|y)$ （其中 u 为物理场， y 为观测条件），生成式模型不仅能够重构高保真的物理场，还能通过采样直接评估预测的不确定性，为复杂物理系统的建模提供了新的统计力学视角。

5.2.2 早期探索：对抗生成与变分推断

早期的研究主要集中在物理信息生成对抗网络（Physics-Informed GANs, PI-GANs）^[49] 和变分自编码器（VAEs）。PI-GANs 通过在判别器中引入 PDE 残差项，迫使生成器产生的样本不仅符合数据分布，同时满足物理约束。这种对抗博弈机制在模拟随机渗透率场中的达西流（Darcy Flow）等问题上取得了一定成功。然而，GANs 固有的训练不稳定性（如模式坍塌）以及 VAEs 在生成高频纹理时的模糊性，限制了它们在精细流场模拟中的应用。尽管如此，这些工作证实了深度生成模型具备在无标签数据情况下，仅依靠物理控制方程生成符合物理规律解空间的能力。

5.2.3 扩散模型：随机微分方程的逆向求解

近年来，去噪扩散概率模型（Denoising Diffusion Probabilistic Models, DDPM）^[50] 和基于分数的生成模型（Score-based Generative Models, SGM）的兴起，为 PDE 求解带来了突破性进展。从理论层面看，扩散模型的正向过程本身就是一个由随机微分方程（SDE）描述的物理扩散过程（即热方程演化），其将结构化数据逐渐转化为高斯噪声；而逆向生成过程则对应于求解一个逆时间的 SDE。根据 Anderson 定理，这一逆向过程完全由数据分布的“分数函数”（Score Function，即对数概率密度梯度 $\nabla_x \log p(x)$ ）决定。因此，训练扩散模型本质上是在学习高维物理场分布的梯度场。这种严格的数学等价性使得扩散模型在处理高维、非线性 PDE 时表现

出极高的稳定性和样本质量，特别是在生成具有精细相干结构的湍流场方面，其表现显著优于 GANs。

5.2.4 物理引导的扩散采样与超分辨率

为了将物理定律更深层次地融入扩散模型，最新的研究提出了“物理引导采样”(Physics-Guided Sampling)策略。不同于在训练阶段将 PDE 残差加入损失函数(这需要重新训练模型)，该策略利用预训练的扩散模型作为先验，在推理采样阶段(Sampling Phase)引入物理约束。通过在朗之万动力学(Langevin Dynamics)迭代中加入基于 PDE 残差的修正项，引导生成轨迹向满足物理守恒律的流形收敛。这种方法被广泛应用于流体超分辨率(Super-Resolution)任务，即从低分辨率的稀疏观测中恢复高分辨率的湍流细节。此时，扩散模型充当了一个强大的“物理先验生成器”，填补了粗网格数据丢失的高频信息，同时保证重构结果在统计意义上符合 Navier-Stokes 方程的动力学特性。此外，在反问题求解中，条件扩散模型能够直接从观测数据采样出潜在的参数分布，为不适定问题(Ill-posed Problems)提供了一种具备贝叶斯解释的求解框架。

5.3 几何与多尺度深度学习

几何深度学习方法利用问题的几何结构设计网络架构，特别适合处理不规则网格和复杂几何。

图神经网络(GNN)适用于不规则网格和复杂几何，通过消息传递机制处理空间关系^[43,51]。在偏微分方程求解中，可以将网格节点视为图的节点，边表示空间邻接关系，通过消息传递学习节点间的相互作用。Sanchez-Gonzalez 等人使用图网络学习复杂物理系统的模拟，可以处理粒子系统、流体等多体系统^[51]。Pfaff 等人提出的 MeshGraphNets 专门用于基于网格的模拟，可以学习各种物理过程的动力学^[43]。

几何深度学习利用几何结构(如对称性、等变性)设计网络架构。Bronstein 等人系统总结了^[52] 几何深度学习的理论和方法，包括网格、群、图、测地线等几何结构。在偏微分方程求解中，可以利用问题的对称性设计等变网络，自动保持对称性，提高效率和精度。例如，对于旋转对称的问题，可以使用旋转等变的卷积操作。

等变神经网络对特定的变换(如旋转、平移)等变，即输入变换后输出也相应变换。这种性质对于物理问题特别重要，因为许多物理定律具有对称性。通过设计等变网络架构，可以自动保持这些对称性，减少需要学习的参数，提高泛化能力。Weiler 等人提出了^[53] $E(2)$ -等变卷积，可以处理二维欧几里得群的等变性。Thomas 等人提出了^[54] $SE(3)$ -等变网络，可以处理三维旋转等变性。

流形上的深度学习将传统的深度学习方法扩展到流形上，可以处理复杂几何。

对于定义在流形上的偏微分方程，需要特殊的处理方法。流形上的深度学习将传统的深度学习方法扩展到流形上，可以处理复杂几何。Masci 等人提出了^[55] 测地线卷积，可以在流形上进行卷积操作。Monti 等人提出了^[56] 几何深度学习框架，可以处理各种几何结构。多尺度问题是科学计算中的常见挑战^[57]，深度学习方法需要能够同时处理不同尺度的特征。

多分辨率网络使用多个网络或网络分支处理不同尺度的特征，然后融合得到最终解。这种方法可以同时捕捉大尺度结构和小尺度细节，提高解的精度。例如，可以使用不同分辨率的输入训练多个网络，然后通过插值或融合得到最终解。多分辨率方法特别适用于湍流、多相流等多尺度问题。

小波神经网络结合小波变换和神经网络，利用小波的多尺度性质。小波变换提供了多尺度表示，可以同时分析不同频率的成分。小波神经网络结合小波变换和神经网络，利用小波的多尺度性质。这种方法可以自动识别不同尺度的特征，提高解的精度。小波神经网络特别适用于包含多尺度结构的问题，如湍流、多相流等。

注意力机制可以自动关注重要区域，提高计算效率。在偏微分方程求解中，解的变化可能在不同区域差异很大，注意力机制可以动态调整计算资源，重点关注变化剧烈的区域。这种方法可以减少不必要的计算，提高效率。Transformer 架构中的自注意力机制已被应用于神经算子，如 Li 等人提出的 Transformer for PDEs。

第六章 应用领域

6.1 流体力学

流体力学是深度学习方法求解偏微分方程的重要应用领域，Navier-Stokes 方程作为流体力学的基础方程^[58,59]，描述了流体的运动规律。深度学习方法在流体力学中的应用涵盖了从层流到湍流、从不可压缩到可压缩、从单相到多相等各类问题。

层流是流体运动的基本形式，PINN 和神经算子已成功应用于各种层流问题，包括管道流动、绕流、边界层流动等典型场景^[48,58,59]。Rao 等人使用 PINN 求解不可压缩层流，在多个基准问题上取得了良好效果，证明了 PINN 在层流问题中的有效性^[58]。Jin 等人提出的 NSFnets (Navier-Stokes Flow Nets) 专门针对不可压缩 Navier-Stokes 方程设计，通过特殊的网络架构和损失函数设计，在层流问题上表现出色，能够准确预测速度场和压力场^[58]。神经算子方法（如 FNO、DeepONet）可以学习层流的求解算子，对于不同的边界条件和参数，仅需一次前向传播即可快速预测流场，在参数化研究中展现出显著优势^[35,38,39]。Li 等人使用 FNO 学习 Navier-Stokes 方程的求解算子，实现了对不同初始条件和参数的快速流场预测，大大提高了参数空间探索的效率^[38,39]。

湍流是典型的非线性、多尺度问题，传统数值方法需要精细网格和长时间积分，计算成本极高。深度学习方法通过学习湍流特征，可以快速预测流场，为湍流问题的工程应用提供了新的解决途径。PINN 可以通过学习湍流的统计特征实现快速预测，虽然直接求解完整的湍流场仍然困难，但可以作为高效的代理模型用于参数化研究和优化设计。神经算子方法可以学习湍流的求解算子，实现快速推理，特别适用于需要大量重复求解的场景。Kochkov 等人使用机器学习加速计算流体动力学，在湍流问题上取得了显著效果，证明了深度学习方法在湍流模拟中的潜力。然而，需要指出的是，PINN 在高雷诺数可压缩流、长时间积分强混沌系统等问题上仍面临优化困难与误差不可控的挑战，这也是当前研究的重点方向。^[17,60]

多相流涉及复杂的界面动力学，包括界面追踪、界面演化等关键问题。深度学习方法可以学习界面演化规律，实现快速预测。PINN 可以学习相场模型的求解，准确预测相界面的位置和形状，为材料科学和相变问题提供工具。例如，Allen-Cahn 方程作为相场模型的基础方程，PINN 在求解时能够准确捕捉相界面的演化过程。神经算子方法可以学习多相流的求解算子，对于不同的初始条件和参数，可以快速预测流场。Wen 等人提出的 U-FNO (U-shaped Fourier Neural Operator) 专门用于多相流问题，通过改进的网络架构处理多相界面，取得了良好效果。

可压缩流动涉及激波等不连续现象^[61,62], 需要特殊处理。PINN 在处理不连续解时可能遇到困难, 因为神经网络天然偏向低频函数, 对激波等高频不连续结构的捕捉能力有限。针对这一问题, 研究者提出了自适应采样、弱形式 PINN 等改进方法。Mao 等人使用 PINN 求解高速流动, 通过特殊设计处理激波, 证明了 PINN 在可压缩流问题中的可行性。神经算子方法可以学习可压缩流的求解算子, 但需要足够的训练数据, 且对激波等不连续结构的泛化能力仍需进一步研究。

非牛顿流体的本构关系复杂, 传统方法需要迭代求解。深度学习方法可以学习非牛顿流体的本构关系, 实现快速求解。例如, 可以使用神经网络学习本构关系, 然后结合 PINN 或神经算子求解流动问题, 这种方法能够处理复杂的本构关系, 如各向异性、硬化等。

微流体涉及小尺度效应, 如表面张力、电渗流等复杂数理现象。深度学习方法可以学习这些复杂效应, 为微流体设计提供工具。PINN 的无网格特性特别适合处理微流体的复杂几何形状, 而神经算子方法则能够快速进行参数化研究, 为微流体器件的优化设计提供支持。

在流体与传热问题中, PINN 常被用于数据稀缺或仅能获得压力/速度稀疏传感的情形, 例如从少量速度测点反演粘性系数、外力或边界条件等逆问题; 神经算子方法则在需要大量重复求解的参数化研究、优化设计和不确定性量化等场景中展现出显著优势。^[12,58]

6.2 固体力学

固体力学涉及弹性、塑性、断裂、接触等多种问题^[63,64], 深度学习方法在这些问题中都有广泛应用。在固体与相场断裂问题中, 深度学习方法可以在缺少全场应力标签时仍通过平衡方程恢复位移/裂纹演化, 展现出强大的数据驱动能力。

弹性问题是固体力学的基础, 涉及线性弹性力学方程^[65,66]。PINN 和神经算子都可以学习弹性力学方程的求解。Haghigat 等人使用 PINN 进行固体力学的反演和代理建模, 在多个问题上取得了良好效果, 证明了 PINN 在固体力学正问题和逆问题中的有效性^[65]。Abueidda 等人提出了增强的 PINN 用于超弹性问题, 可以处理大变形等非线性几何效应^[67]。神经算子在固体力学中的应用包括弹性问题、塑性问题等, 通过学习弹性力学方程的求解算子, 可以快速进行结构分析, 特别适合优化设计等需要大量正向求解的场景^[35,39]。对于不同的载荷和边界条件, 神经算子可以快速预测应力和位移场, 大大提高了结构优化设计的效率^[35,39,66]。

塑性问题涉及材料非线性行为, 需要学习塑性本构关系。深度学习方法可以学习塑性本构关系, 然后结合 PINN 或神经算子求解。例如, 可以使用神经网络学习应力-应变关系, 然后求解平衡方程。这种方法可以处理复杂的本构关系, 如各向异性、硬化等, 为复杂材料行为的建模提供了新途径。

断裂力学涉及裂纹扩展规律, 是结构安全的重要问题^[66,68]。Goswami 等人使

用迁移学习增强的 PINN 进行相场断裂建模，可以预测裂纹扩展^[68]。深度学习方法可以学习裂纹扩展规律，预测结构失效。例如，可以使用 PINN 学习相场模型的求解，预测裂纹的位置和形状^[68]。在缺少全场应力标签的情况下，PINN 仍能通过平衡方程恢复位移场和裂纹演化过程，这对于实验数据稀缺的断裂问题具有重要意义^[66,68]。

接触问题涉及多体接触，需要处理接触约束。深度学习方法可以学习接触力分布，实现快速求解。例如，可以使用神经网络学习接触力，然后结合 PINN 求解。这种方法可以处理复杂的接触几何和摩擦，为多体系统的快速分析提供了工具。

复合材料涉及多相材料，本构关系复杂。深度学习方法可以学习复合材料的有效性质，然后求解宏观问题。例如，可以使用神经网络学习有效模量，然后进行结构分析。这种方法能够处理复杂的多相材料系统，为复合材料的设计和分析提供了高效工具。

固体力学问题经常涉及多物理场耦合，如热-力耦合、电-力耦合等。深度学习方法可以学习这些耦合关系，实现快速求解。PINN 可以同时求解多个物理场的方程，通过耦合条件连接，实现多物理场问题的统一求解。

6.3 多物理场耦合

多物理场耦合问题涉及多个物理过程的相互作用，是科学计算中的难点。传统数值方法在处理多物理场耦合问题时，往往需要复杂的迭代求解和耦合策略，计算成本高昂。深度学习方法为多物理场耦合问题提供了新的解决思路，神经算子可以学习耦合方程的求解算子，实现快速求解。

流固耦合涉及流体和固体的相互作用，是船舶、飞机等设计中的重要问题。深度学习方法可以学习流固耦合的规律，实现快速求解。PINN 可以同时求解流体方程和固体方程，通过耦合条件（如界面力平衡、位移连续性等）连接，实现流固耦合问题的统一求解。神经算子方法可以学习流固耦合的求解算子，对于不同的边界条件和参数，可以快速预测耦合场。这种方法在优化设计中特别有用，可以快速评估不同设计方案的性能，大大缩短设计周期。

热流耦合涉及传热和流动的相互作用^[69]，是热管理中的重要问题^[70]。Cai 等人使用 PINN 求解传热问题，可以处理复杂边界条件和时变热源，为热管理问题提供了有效的解决方案^[70]。深度学习方法可以学习热流耦合的规律，实现快速求解。例如，可以使用 PINN 同时求解 Navier-Stokes 方程和能量方程，通过温度对物性的影响（如粘度、密度等）实现耦合^[70]。神经算子方法可以学习热流耦合的求解算子，实现快速推理，特别适用于需要大量参数化研究的场景^[35,38,39]。

电磁热耦合涉及电磁场和温度场的相互作用，是电子设备设计中的重要问题。深度学习方法可以学习电磁热耦合的规律，实现快速求解。PINN 可以同时求解 Maxwell 方程和热传导方程，通过温度对电导率的影响实现耦合。这种方法可以用

于天线设计、电磁兼容等问题, PINN 的无网格特性对复杂几何尤其有价值。神经算子方法可以学习 Maxwell 方程的求解算子, 实现快速推理, 特别适合优化设计等需要大量重复求解的场景。

多相多物理场耦合涉及多个相和多个物理过程的相互作用, 是最复杂的问题。深度学习方法可以学习这些复杂的耦合关系, 实现快速求解。例如, 可以使用多个网络分别处理不同的物理过程, 然后通过耦合条件连接, 或者使用统一的网络架构同时学习多个物理场的耦合关系。

生物医学问题经常涉及多物理场耦合, 如血流-组织耦合、药物输送等。深度学习方法可以为这些问题提供快速求解工具, 为医学诊断和治疗提供支持。例如, 可以使用 PINN 求解药物输送方程, 预测药物浓度分布; 神经算子方法可以学习血流动力学的求解算子, 用于心血管疾病诊断和治疗。

6.4 其他应用领域

深度学习方法在多个其他科学与工程领域也有重要应用, 展现了其广泛的适用性和强大的问题解决能力。

量子力学问题涉及 Schrödinger 方程, 高维问题难以用传统方法求解^[71]。深度学习方法可以处理高维 Schrödinger 方程, 突破传统方法的维度限制, 为量子系统模拟提供新方法^[12,71]。PINN 可以求解时间相关的 Schrödinger 方程, 预测量子系统的演化^[12,71]。在电磁与量子问题中, PINN 的无网格特性对复杂几何尤其有价值^[12,39,71]。神经算子方法可以学习量子系统的求解算子, 对于不同的初始条件和参数, 可以快速预测波函数, 为量子系统的参数化研究提供了高效工具^[35,39]。

电磁学问题涉及 Maxwell 方程, 是天线设计、电磁兼容等的基础^[72]。Maxwell 方程描述了电磁场的行为, 深度学习方法可以求解 Maxwell 方程, 处理复杂几何和时变问题^[12,72]。PINN 可以求解时域和频域 Maxwell 方程, 预测电磁场分布, 为天线设计、电磁兼容等问题提供快速求解工具^[12,72]。神经算子方法可以学习 Maxwell 方程的求解算子, 实现快速推理, 特别适合优化设计等需要大量重复求解的场景^[35,38,39]。

生物医学问题涉及生物传质方程、血流动力学等复杂的生理过程^[12]。深度学习方法可以为这些问题提供快速求解工具。例如, 可以使用 PINN 求解药物输送方程, 预测药物浓度分布, 为药物设计和给药方案优化提供支持^[12]。神经算子方法可以学习血流动力学的求解算子, 用于心血管疾病诊断和治疗, 实现快速的血流动力学分析^[12,35,39]。

地球科学问题涉及地震波传播、气候模拟等复杂的地球物理过程。深度学习方法可以为这些问题提供工具。例如, 可以使用 PINN 求解地震波方程, 预测地震波传播, 为地震灾害评估和地质勘探提供支持。神经算子方法可以学习气候模型的求解算子, 用于快速气候模拟, 为气候变化研究和预测提供高效的计算工具。

材料科学问题涉及相变、晶体生长等复杂的材料演化过程^[12,13]。深度学习方

法可以学习这些过程的规律，为材料设计提供工具。例如，可以使用 PINN 求解相场模型（如 Allen-Cahn 方程），预测相变过程和相界面的演化，准确预测相界面的位置和形状^[12,13]。Burgers 方程作为典型的非线性偏微分方程，包含对流项和扩散项，常用于验证数值方法^[12-14]，PINN 在求解 Burgers 方程时表现出色，能够准确捕捉激波等不连续现象，已成为 PINN 方法验证的标准测试问题^[12-14]。

化学工程问题涉及反应-扩散方程、传质等复杂的化学过程。深度学习方法可以为这些问题提供快速求解工具。例如，可以使用 PINN 求解反应-扩散方程，预测反应过程和浓度分布，为反应器设计和优化提供支持。

能源工程问题涉及传热、流动、燃烧等复杂的能量转换过程。深度学习方法可以为这些问题提供工具。例如，可以使用 PINN 求解燃烧方程，预测燃烧过程和温度分布，为燃烧器设计和优化提供支持。热传导方程描述了热量的扩散过程，PINN 在求解热传导问题时，可以处理复杂边界条件和时变热源，为热管理问题提供解决方案。

神经算子的一个重要应用是参数化研究和不确定性量化^[35,39,73]。在工程设计、优化等问题中，需要研究参数对解的影响。神经算子学习参数到解的映射，可以快速评估不同参数下的解，大大提高了参数化研究的效率^[35,39,66]。不确定性量化需要大量正向求解，计算负担沉重，神经算子可以快速生成大量样本，为不确定性量化提供工具^[35,39,73]。通过 Monte Carlo 方法或其他采样方法，可以快速评估不确定性，为风险评估和决策支持提供依据^[73]。

神经算子训练完成后，推理速度非常快，可以实现实时计算。这在控制、优化等需要实时响应的场景中特别有用，为实时决策和在线优化提供了可能。

第七章 挑战与展望

7.1 当前挑战

现阶段最大的科学问题是“可训练性与可信度”。PINNs 的训练收敛并非由离散稳定性控制，而是由优化景观与梯度尺度控制，针对强对流、长时混沌与高维系统仍缺乏统一的成功条件。神经算子虽然泛化能力强，但对训练数据分布敏感，当遇到分布外（Out-of-Distribution）的几何或参数时，性能可能显著退化。

7.2 未来展望

未来方向正在从“单方法改良”走向“基础模型 + 混合求解器”。一方面，基于大规模多物理数据预训练的“PDE 基础模型”有望降低新任务的适配成本；另一方面，将神经算子作为粗尺度预处理器，结合 PINNs 或可微求解器进行细尺度修正的分层混合框架，将是兼顾精度、效率与泛化性的重要路径。

第八章 结论

深度学习方法为偏微分方程求解带来了范式上的革新。PINNs 实现了无网格的物理约束求解，神经算子实现了函数空间的高效映射，而生成式模型为不确定性量化提供了新工具。尽管面临理论完备性和训练稳定性的挑战，随着 Transformer、扩散模型等新架构与物理知识的深度融合，该领域将在科学计算与工程设计中发挥日益关键的作用。

参考文献

- [1] 易飞王岩. 机器学习辅助的数值计算方法研究进展[J]. 数学的实践与认识, 2022, 52(22): 68-85.
- [2] 郭亮王玉. DeepONet 算子学习方法在科学计算中的应用[J]. 计算力学学报, 2023, 40(1): 33-48.
- [3] 姚远刘志强. DeepONet 算子学习方法及其应用[J]. 力学进展, 2023, 53(3): 987-1002.
- [4] 黄明周宁. 基于傅里叶神经算子的偏微分方程快速求解方法[J]. 工程数学学报, 2022, 39(4): 812-826.
- [5] 郑志, 刘伟. 多尺度傅里叶神经算子及其在流体力学中的应用[J]. 空气动力学学报, 2023, 41(4): 55-70.
- [6] 陈鹏吴昊. 图神经网络在偏微分方程求解中的研究进展[J]. 人工智能学报, 2023, 47(3): 512-528.
- [7] 孙杨刘松. 基于图神经算子的复杂几何场预测方法[J]. 力学进展, 2023, 53(5): 2104-2122.
- [8] 赵晨王力. 生成式模型在偏微分方程求解中的研究进展[J]. 软件学报, 2022, 33(12): 4208-4224.
- [9] 徐俊王强. 基于扩散模型的偏微分方程求解方法[J]. 软件学报, 2023, 34(9): 3011-3025.
- [10] FALKNER N, et al. U-net based convolutional neural networks for surrogate modeling of fluid flows[A]. 2018.
- [11] THUERÉY N, WEISSMANN S, et al. Deep learning methods for reynolds-averaged navier-stokes simulations of airfoil flows[J]. AIAA Journal, 2020, 58(1): 25-36.
- [12] RAISSI M, PERDIKARIS P, KARNIADAKIS G E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations[J]. Journal of Computational Physics, 2019, 378: 686-707.
- [13] BAYDIN A G, PEARLMUTTER B A, RADUL A A, et al. Automatic differentiation in machine learning: a survey[J]. Journal of Machine Learning Research, 2018, 18(153): 1-43.

- [14] KARNIADAKIS G E, KEVREKIDIS I G, LU L, et al. Physics-informed machine learning[J]. *Nature Reviews Physics*, 2021, 3: 422-440.
- [15] 胡翔曾浩. 深度学习在网格生成与自适应计算中的应用[J]. *计算机辅助设计与图形学学报*, 2022, 34(4): 721-735.
- [16] 周亮李志强. 基于 PINN 的多参数反演问题求解方法[J]. *计算物理*, 2023, 40 (2): 188-204.
- [17] WANG S, TENG Y, PERDIKARIS P. Understanding and mitigating gradient pathologies in physics-informed neural networks[J]. *SIAM Journal on Scientific Computing*, 2021, 43(5): A3055-A3081.
- [18] WANG S, YU X, PERDIKARIS P. When and why pinns fail to train: A neural tangent kernel perspective[J]. *Journal of Computational Physics*, 2022, 449: 110768.
- [19] KRISHNAPRIYAN A, GHAI U, CHEN Z, et al. Characterizing possible failure modes in physics-informed neural networks[J]. *Advances in Neural Information Processing Systems*, 2021, 34: 26548-26560.
- [20] JAGTAP A D, KAWAGUCHI K, KARNIADAKIS G E. Adaptive activation functions accelerate convergence in deep and physics-informed neural networks [J]. *Journal of Computational Physics*, 2020, 404: 109136.
- [21] 刘晨高远. PINN 求解刚性 (Stiff) PDE 的难点与缓解方法[J]. *物理学报*, 2023, 72(20): 200201.
- [22] MCCLENNY L D, BRAGA-NETO U M. Self-adaptive physics-informed neural networks using a soft attention mechanism[A]. 2020.
- [23] LU L, MENG X, MAO Z, et al. Deepxde-based causal training of physics-informed neural networks[A]. 2021.
- [24] DAW A, KARNIADAKIS G E. Mitigating propagation failures in pinns: Incremental training with residual-based adaptive refinement[J]. *Computer Methods in Applied Mechanics and Engineering*, 2023, 415: 116270.
- [25] LU L, MENG X, MAO Z, et al. Deepxde: A deep learning library for solving differential equations[J]. *SIAM Review*, 2021, 63(1): 208-228.
- [26] YANG L, MENG X, KARNIADAKIS G E. Adversarial uncertainty quantification in physics-informed neural networks[J]. *Journal of Computational Physics*, 2019, 394: 136-152.

- [27] JAGTAP A D, KAWAGUCHI K, KARNIADAKIS G E. extended physics-informed neural networks (xpinns): A generalized space-time domain decomposition based deep learning framework for nonlinear partial differential equations [J]. Communications in Computational Physics, 2020, 28(5): 2002-2041.
- [28] JAGTAP A D, KAWAGUCHI K, KARNIADAKIS G E. Locally adaptive activation functions with slope recovery for deep and physics-informed neural networks[J]. Proceedings of the Royal Society A, 2020, 476(2239): 20200334.
- [29] JAGTAP A D, LU L, KARNIADAKIS G E. Finite-basis physics-informed neural networks (fbpinns): A scalable domain decomposition deep learning framework for solving pdes[J]. Computer Methods in Applied Mechanics and Engineering, 2021, 384: 113938.
- [30] 梁宇黄兵. PINN 在复杂边界条件处理中的关键问题与改进策略[J]. 工程力学, 2022, 39(12): 144-160.
- [31] KHARAZMI E, ZHANG Z, KARNIADAKIS G E. hp-vpinns: Variational physics-informed neural networks with domain decomposition[J]. Computer Methods in Applied Mechanics and Engineering, 2021, 374: 113547.
- [32] KHARAZMI E, ZHANG Z, KARNIADAKIS G E. hp-vpinns: A higher-order variational physics-informed neural network for solving pdes[J]. Computer Methods in Applied Mechanics and Engineering, 2021, 386: 114010.
- [33] BALTAR I, et al. Weak-form physics-informed neural networks for solving pdes with sharp gradients and discontinuities[A]. 2022.
- [34] E W, YU B. The deep ritz method: A deep learning-based numerical algorithm for solving variational problems[J]. Communications in Mathematics and Statistics, 2018, 6(1): 1-12.
- [35] LU L, JIN P, KARNIADAKIS G E. Learning nonlinear operators via deeponet based on the universal approximation theorem of operators[J]. Nature Machine Intelligence, 2021, 3: 218-229.
- [36] LU L, JIN P, PANG G, et al. Learning operators for scientific machine learning [A]. 2022.
- [37] CHEN T, CHEN H. Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems[J]. IEEE Transactions on Neural Networks, 1995, 6(4): 911-917.
- [38] LI Z, AZIZZADENESHELI K, et al. Neural operator: Graph kernel network for partial differential equations[C]//arXiv preprint arXiv:2003.03485. 2020.

- [39] LI Z, KOVACHKI N B, AZIZZADENESHELI K, et al. Fourier neural operator for parametric partial differential equations[C]//International Conference on Learning Representations. 2021.
- [40] LI Z, et al. Geometry-informed neural operator for large-scale 3d fluid simulations[A]. 2022.
- [41] KOVACHKI N B, LI Z, LIU K, et al. Neural operators: Learning maps between function spaces[J]. Acta Numerica, 2023, 32: 363-580.
- [42] LI Z, et al. Graph neural operator for parametric partial differential equations on irregular meshes[A]. 2020.
- [43] PFAFF T, FORTUNATO M, SANCHEZ-GONZALEZ A, et al. Learning mesh-based simulation with graph networks[C]//International Conference on Learning Representations. 2021.
- [44] LONG Z, LU Y, MA X, et al. Pde-net: Learning pdes from data[J]. Journal of Computational Physics, 2019, 399: 108925.
- [45] 高翔刘振. 深度学习与有限元方法的融合研究进展[J]. 力学进展, 2023, 53(4): 2003-2021.
- [46] BAR-SINAI Y, HOYER S, HICKEY J, et al. Learning data-driven discretizations for partial differential equations[J]. Proceedings of the National Academy of Sciences, 2019, 116(31): 15344-15349.
- [47] UM K, BRAND R, HOLL P, et al. Solver-in-the-loop: Learning from differentiable physics to interact with real world[C]//Advances in Neural Information Processing Systems: Vol. 33. 2020: 6111-6122.
- [48] WIGHT C, ZHAO J. Solving allen-cahn and cahn-hilliard equations using the deep galerkin method[J]. Journal of Computational Physics, 2019, 397: 108842.
- [49] YANG L, MENG X, KARNIADAKIS G E. Physics-informed generative adversarial networks for stochastic differential equations[J]. SIAM Journal on Scientific Computing, 2020, 42(1): A292-A317.
- [50] HO J, JAIN A, ABBEEL P. Denoising diffusion probabilistic models[C]// Advances in Neural Information Processing Systems: Vol. 33. 2020: 6840-6851.
- [51] SANCHEZ-GONZALEZ A, GODWIN J, PFAFF T, et al. Learning to simulate complex physics with graph networks[C]//International Conference on Learning Representations. 2020.
- [52] BRONSTEIN M M, BRUNA J, COHEN T, et al. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges[A]. 2021.

- [53] WEILER M, CESA G. General E(2)-equivariant steerable CNNs[C]//Advances in Neural Information Processing Systems: Vol. 32. 2019.
- [54] THOMAS N, SMIDT T, KEARNES S, et al. Tensor field networks: Rotation- and translation-equivariant neural networks for 3d point clouds[A]. 2018.
- [55] MASCI J, BOSCAINI D, BRONSTEIN M M, et al. Geodesic convolutional neural networks on riemannian manifolds[C]//Proceedings of the IEEE International Conference on Computer Vision Workshops. 2015: 37-45.
- [56] MONTI F, BOSCAINI D, MASCI J, et al. Geometric deep learning on graphs and manifolds using mixture model CNNs[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 5115-5124.
- [57] 杨锐宋超. 深度学习多尺度建模方法综述[J]. 模式识别与人工智能, 2021, 34(5): 967-989.
- [58] JIN X, CAI S, LI H, et al. Nsfnets (navier–stokes flow nets): Physics-informed neural networks for the incompressible navier–stokes equations[J]. Journal of Computational Physics, 2021, 426: 109951.
- [59] 徐东张坤. 基于物理信息神经网络的不可压流动求解研究[J]. 空气动力学学报, 2020, 38(6): 1205-1217.
- [60] JIN X, CAI S, LI H, et al. Nsfnets 2.0: Physics-informed neural networks for turbulent flows[A]. 2022.
- [61] 李翔宇陈康. PINN 在含激波流动求解中的改进研究[J]. 物理学报, 2023, 72(14): 140201.
- [62] 周凯徐昊. 深度学习求解激波及间断问题的最新进展[J]. 力学季刊, 2023, 43(4): 952-968.
- [63] 王坤赵瑞. 物理信息神经网络在计算力学中的应用进展[J]. 力学季刊, 2021, 42(3): 587-602.
- [64] 张远梁峥. 基于 PINN 的固体力学逆问题求解方法研究[J]. 振动工程学报, 2022, 35(5): 765-776.
- [65] HAGHIGHAT E, RAISSI M, MOURE A, et al. A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics[J]. Computer Methods in Applied Mechanics and Engineering, 2021, 379: 113741.
- [66] 李亮周永胜. 物理信息神经网络在固体力学中的应用[J]. 工程力学, 2021, 38(12): 34-54.
- [67] ABUEIDDA D W, KORIC S, SOBH N. Enhanced physics-informed neural networks for hyperelasticity[J]. Computer Methods in Applied Mechanics and Engineering, 2021, 384: 113927.

- [68] GOSWAMI S, ANITESCU C, CHAKRABORTY S, et al. Transfer learning enhanced physics informed neural network for phase-field modeling of fracture [J]. *Theoretical and Applied Fracture Mechanics*, 2020, 106: 102447.
- [69] 杨明赵鹏. 基于 PINN 的热传导方程数值求解研究[J]. *工程热物理学报*, 2022, 43(10): 2471-2484.
- [70] CAI S, WANG S, WANG S, et al. Physics-informed neural networks for heat transfer problems[J]. *International Journal of Heat and Mass Transfer*, 2021, 164: 120560.
- [71] HAN J, JENTZEN A, E W. Solving high-dimensional partial differential equations using deep learning[J]. *Proceedings of the National Academy of Sciences*, 2018, 115(34): 8505-8510.
- [72] 张俊王锦. 物理信息神经网络研究进展综述[J]. *计算机学报*, 2021, 44(11): 2365-2388.
- [73] 王子涵宋国. 基于算子学习的偏微分方程不确定性量化方法[J]. *计算数学*, 2023, 45(2): 355-372.

致谢

感谢导师的指导与实验室同学的帮助。