

PINNuclear-Neutrons 项目笔记

刘洋¹

1. 四川师范大学数学科学学院, 四川成都 610066

E-mail: mathliuyang@163.com

通信作者简介: 刘洋 (1997-), 男, 研究生, 主要研究方向: 偏微分方程与数学物理

摘要 PINNuclear-Neutrons 项目旨在学习和复现核反应堆中子相关问题的解决方法, 使用物理反应神经网络 (PINN) 和深度机器学习技术。我们的主要目标是重新实现刘东老师团队发表的“基于 PINN 深度机器学习技术求解多维中子学扩散方程”, 以更准确地模拟中子传输。鼓励对核能技术和深度学习感兴趣的研究人员和学生参与到这个项目中。在各个子目录中, 您将找到有关每个主题的详细说明和示例。

关键词 基于物理信息指引的神经网络模型; 深度学习; 核反应堆; 中子学扩散方程

引言

随着人工智能技术的迅猛发展, 深度学习神经网络 (DNNs) 在解决复杂偏微分方程 (PDE) 的领域引起了广泛的兴趣。在核反应堆设计和工程领域, 中子学扩散模型是核反应堆工程设计的重要组成部分。传统的中子扩散模型采用有限差分法 (FDM) 和有限元法 (FEM) 等数值方法进行求解, 这些方法在工程实践中取得了显著的成功。然而, 随着深度学习技术的不断发展, 一种新的方法已经崭露头角, 即基于物理信息指导的神经网络模型 (PINN), 它在解决微分方程方面展现出巨大潜力。DNNs 首次用于解微分方程是在 1998 年由 Lagaris 提出的 [1], 后来由 Raissi 和 Karniadakis 进一步发展成为基于物理信息的神经网络 (PINNs) [2], 将方程和边界条件等物理信息作为损失函数引入到神经网络训练当中, 通过这种方法将问题的物理信息考虑进去, 不再仅仅依靠数据驱动的方法来使得网络拟合偏微分方程的解函数; 同时, 此研究也指出如果将方程中的系数当作未知参数也可以完成发现方程的任务。此后, 便有许多研究围绕基于物理信息的神经网络展开, 而这些研究通常可以总结为调整神经网络结构, 更改损失函数的形式, 挑选不同的激活函数, 采用不同的优化策略。PINN 模型充分利用了深度学习神经网络的通用逼近能力, 同时能够处理多维高阶复杂微分方程。与传统数值方法相比, PINNs 具有多方面的优势, 包括无需事先收集大量训练数据、适用于正向和反向问题、无空间几何限制等。因此, PINNs 已经引起了广泛的研究兴趣, 并在热传递、结构动力学、流体力学、固体力学以及核反应堆动力学等多个领域取得了重要突破。

本 PINNuclear-Neutrons 项目笔记旨在深入理解 PINN 模型的原理和应用，通过推导和复现刘东老师的研究工作 [3]，进一步探讨 PINN 模型在核反应堆设计和工程中的潜在应用。我们将重点关注 PINN 模型的工作原理、关键参数的影响以及可能的优化方向，以便更好地理解和应用这一创新方法。通过本学习笔记，我们希望能够为核反应堆设计和工程领域的研究人员提供有关 PINNs 的深入见解，以应对复杂问题和挑战。

1 相关理论和技术

本章的主要内容是介绍相关理论和技术，首先介绍基于物理信息的神经网络，通过这种方法可以将物理信息引入到神经网络当中，然后介绍了本文方法中所用到的二阶优化函数 L-BFGS，为了网络更快更好的优化。最后介绍了核反应堆物理的相关知识，为后文的方法介绍做好理论上的准备。

1.1 基于物理信息的神经网络

使用神经网络逼近函数从万能逼近定理被提出之后就一直是科学计算的重要方向，然而随着数据时代的到来，大数据在人工智能领域取得了巨大的成功，于是人们想通过数据驱动的方式来拟合函数，但对于某些难以求解的函数，函数的解析解和数值解很难获得，无法大量产生训练神经网络所需要的数据。并且数据驱动的方法还会降低模型的泛用性，只有在有监督的点的位置能达到对应的精度，而在做推断位置的点所求出的值精度就会与训练时相差很多，直至采点够密集才能使求解精度达到对应要求。

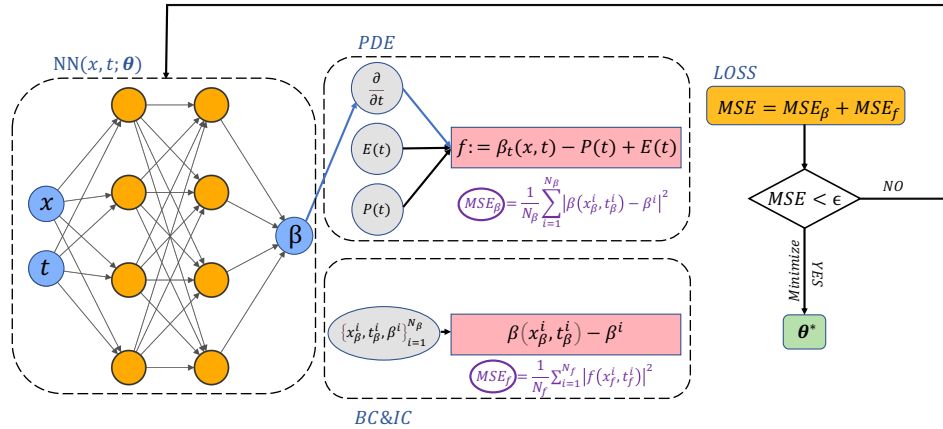


图 1 PINN 网络结构图
Figure 1 PINN Network Structure

为此 Raissi 等人在 2019 年提出了基于物理信息的神经网络 (PINN) [2]，通过将物理方程条件作为损失函数，将物理方程信息参与到神经网络的训练过程之中。当网络训练时，神经网络迭代优化所反向传播的不仅仅有数值上的损失，同时还有物理方程的损失，这样就可以对神经网络所表示的空间进行限制，逐渐去逼近真正的解空间。而且有了物理信息的引入，网络对数据的要求也逐渐降低，不再完全依赖数据驱动，在有很少的数据，甚至没有数据的情况下依然能的到很好的回归效

果,大大增加了网络的泛用性和稳定性。具体神经网络结构如下图 1 所示。如图中所示定义损失函数为:

$$MSE = MSE_u + MSE_{BC,IC} + MSE_R \quad (1.1)$$

其中 MSE_R 是解的数值约束,只对能取到解的值的点进行约束,通过数值约束来使得神经网络拟合对应的函数; $MSE_{BC,IC}$ 是初边值条件的约束,当一个点来自于初边值中任意一个点时,需要满足其对应的初边值条件,通过边值条件来限制解空间; MSE_u 是方程本身的约束,对于内部区域的点,都需要满足微分方程,通过微分方程来限制解空间。当损失函数整体趋于零时,可以认为神经网络很好的拟合了满足初边值条件和方程约束的解,所以此时解方程问题变成了满足对应损失函数的优化问题,可以使用神经网络的优化方法来求得对应方程的解。

1.2 L-BFGS 优化算法

常用的人工智能优化算法如 Adam,SGD 等都是在一阶梯度下降法 (Gradient Descent) 上改进的,而 L-BFGS 是基于二阶优化算法牛顿法改进的。牛顿法作为二阶的优化算法收敛速度要远快于一阶算法的,但是牛顿法需要计算海森矩阵的逆矩阵,运算成本很高,所以针对这个问题,人们提出了 BFGS 算法和 L-BFGS 算法降低内存的需求,在介绍此算法前,先介绍牛顿法的过程。

设 $f(x)$ 的零点为 \tilde{x} , 随机选取初始点 x_0 , 做 $f(x)$ 的切线得到 $y = f(x_0) + f'(x_0)(x - x_0)$, 令 $y = 0$, 得到 $x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$, 由此我们可以得到递推公式 $x_{n+1} =$ 用目标函数 $f(x)$ 的二阶泰勒展开求驻点, 即 $f(x) \approx f(x_n) + f'(x_n)(x - x_n) + \frac{1}{2}f''(x_n)(x - x_n)^2$, 对其进行求导可得 $f'(x_n) + f''(x_n)(x - x_n) = 0$, 由此更新 x 来得到是目标函数最小的点。

当 $f(x)$ 是多元函数时, x 与 $f'(x)$ 为向量, 记 $f'(x)$ 为 $g, f''(x)$ 为海森矩阵记为 H , 则牛顿法在多元情况下的迭代公式变为 $x_{k+1} = x_k - g_k H_k^{-1}, k = 0, 1, \dots$, 其中 H_k^{-1} 是海森矩阵的逆, 而直接求海森矩阵的逆十分困难, 所以 BFGS 算法提出通过迭代来近似 H_k^{-1} , 即:

$$B_{k+1}^{-1} = \left(I - \frac{\delta_k y_k^T}{y_k^T \delta_k} \right) B_k^{-1} \left(I - \frac{y_k \delta_k^T}{y_k^T \delta_k} \right) + \frac{\delta_k \delta_k^T}{y_k^T \delta_k} \quad (1.2)$$

其中 B_{k+1}^{-1} 是对 H_{k+1}^{-1} 的逼近, $B_0^{-1} = I, \delta_k = x_{k+1} - x_k, y_k = g_{k+1} - g_k$ 。利用 B_k^{-1} 的迭代公式, 每次都需要储存 B , 当数据维度很大时, 对计算机的内存需求也变得很高, 于是 L-BFGS 方法提出只储存 δ_k 与 y_k , 每个 B_k^{-1} 都可由 $\delta_1, \dots, \delta_k$ 与 y_1, \dots, y_k 迭代计算得出, 但当 k 过于大时, 计算机内存仍然会面临内存不足的风险, 于是在算法中可以设定一个超参数 N , 当 k 超过 N 时, 我们释放前 $k - N$ 项, 仅用最近的 N 项来逼近 B_k^{-1} , 虽然会损失部分精度, 但可以在有限的内存条件下完成函数的优化。

L-BFGS 仍然面临一个问题, 就是当初始点与函数最优解的点相距过远时, L-BFGS 算法可能会收敛到鞍点, 所以在本文实际应用时会先用 Adam 优化算法寻找一个最优解可能存在的区域, 再使用 L-BFGS 优化方法加快收敛的速度, 得到更好的优化结果。

1.3 核反应堆物理

核反应堆是一种能以可控方式实现自续链式核反应的装置。根据原子核产生能量的方式, 可以分为裂变反应堆和聚变反应堆两种 [4]。当今世界上已建成和广泛使用的反应堆都是裂变反应堆, 聚

变反应堆目前尚处于研究设计阶段。裂变反应堆通过把一个重核裂变为两个中等质量核而释放能量。它是由核燃料、冷却剂、慢化剂、结构材料和吸收剂等材料组成的一个复杂系统。按用途不同, 裂变反应堆可分为生产堆、实验堆和动力堆。按冷却剂或慢化剂的种类不同可分为轻水堆、重水堆、气冷堆和液态金属冷却快中子增殖堆。按引起裂变反应的中子能量不同, 又可分为热中子反应堆和快中子反应堆。

1.3.1 链式裂变反应

核反应堆中的核裂变反应是一种自持续的链式反应。当中子与裂变物质作用而发生核裂变反应时, 裂变物质的原子核通常分裂为两个中等质量数的核 (称为裂变碎片)。与此同时, 还将平均地产生两个以上新的裂变中子, 并释放出蕴藏在原子核内部的核能。在适当的条件下, 这些裂变中子又会引起周围其它裂变同位素的裂变, 如此不断地继续下去。这种反应过程称为链式裂变反应, 其反应过程如下图 2 所示。

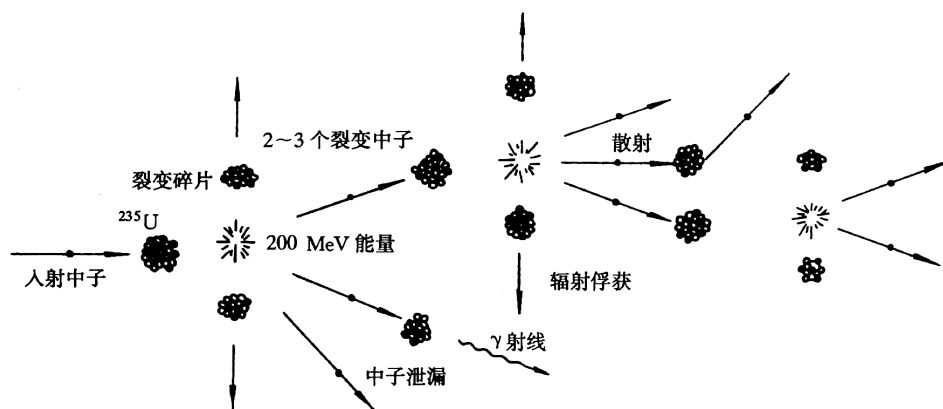


图 2 PINN 网络结构图
Figure 2 PINN Network Structure

如果每次裂变反应产生的中子数目大于引起核裂变所消耗的中子数目, 那么一旦在少数的原子核中引起了裂变反应之后, 就有可能不再依靠外界的作用而使裂变反应不断地进行下去。这样的裂变反应称作自续链式裂变反应。裂变核反应堆就是一种能以可控方式产生自续链式裂变反应的装置。它能够以一定的速率将蕴藏在原子核内部的核能释放出来。

从上面的讨论可以看出, 实现自续链式裂变反应的条件是: 当一个裂变核俘获一个中子产生裂变以后, 在新产生的中子中, 平均至少应该再有一个中子去引起另外一个核的裂变。由于裂变物质每次裂变时平均地放出两个以上裂变中子, 因而实现自续的链式裂变反应是有可能的。但是, 因为核反应堆是由核燃料、慢化剂、冷却剂以及结构材料等所组成的装置, 所以在反应堆内, 不可避免地有一部分中子要被非裂变材料吸收。同时, 还有一部分中子要从反应堆中泄漏出去。因此, 在实际的反应堆中, 并不是全部的裂变中子都能引起新的核裂变反应。一个反应堆能否实现自续的链式裂变反应, 就取决于上述裂变、非裂变吸收和泄漏等过程中中子的产生率与消失率之间的平衡关系。如果在上述的反应过程中, 产生的中子数等于或多于消耗掉的中子数, 则链式裂变反应将会自续地进行下去。

反应堆内自续链式裂变反应的条件可以很方便地用有效增殖系数 k_{eff} 来表示。它的定义是：对给定系统，新生一代的中子数和产生它的直属上一代中子数之比，即

$$k_{\text{eff}} = \frac{\text{新生一代中子数}}{\text{直属上一代中子数}} \quad (1.3)$$

上式的定义是直观地从中子的“寿命一循环”观点出发的。然而，该式在实用上是不太方便的，因为在实际问题中很难确定中子每“代”的起始和终了时间。例如，在芯部中有的中子从裂变产生后立即就引起新的裂变，有的中子则需要经过慢化过程成为热中子之后才引起裂变，有的中子在慢化过程中便泄漏出系统或者被辐射俘获。所以，实际上从中子的平衡关系来定义系统的有效增殖系数更为方便，即

$$k_{\text{eff}} = \frac{\text{系统内中子的产生率}}{\text{系统内中子的总消失 (吸收 + 泄漏) 率}} \quad (1.4)$$

若芯部的有效增殖系数 k_{eff} 恰好等于 1，则系统内中子的产生率便恰好等于中子的消失率。这样，在系统内已经进行的链式裂变反应，将以恒定的速率不断地进行下去，也就是说，链式裂变反应过程处于稳态状况，这种系统称为临界系统。若有效增殖系数 k_{eff} 小于 1，这时系统内的中子数目将随时间而不断地衰减，链式裂变反应是非自续的，这种系统便称为次临界系统。若有效增殖系数 k_{eff} 大于 1，则系统内的中子数目将随时间而不断地增加，我们称这种系统为超临界系统。

显然，有效增殖系数 k_{eff} 与系统的材料成分和结构（例如易裂变同位素的富集度，燃料-慢化剂的比例等）有关。同时，它还与中子的泄漏程度，或反应堆的大小有关。当反应堆的尺寸为无限大时，中子的泄漏损失便等于零，这时增殖系数将只与系统的材料成分和结构有关。通常，我们把无限大介质的增殖系数称为无限介质增殖系数，以 k_{∞} 表示。

对于实际的有限大小的反应堆，中子的泄漏损失总是不可避免的。假定中子的不泄漏概率为 Λ ，它的定义是

$$\Lambda = \frac{\text{系统内中子的吸收率}}{\text{系统内中子的泄漏率}} \quad (1.5)$$

不泄漏概率 Λ 主要取决于反应堆芯部的大小和几何形状，当然它也和芯部成分有关。一般说来，芯部愈大，不泄漏概率也愈大。于是，由式 (1.4) 和式 (1.5) 可知，有限尺寸芯部的有效增殖系数为

$$k_{\text{eff}} = k_{\infty} \Lambda \quad (1.6)$$

当系统为无限大时， $\Lambda = 1$ ，这时有效增殖系数 $k_{\text{eff}} = k_{\infty}$ 。根据以上的讨论，立即可以得出反应堆维持自续链式裂变反应的条件是

$$k_{\text{eff}} = k_{\infty} \Lambda = 1 \quad (1.7)$$

式 (1.7) 称为反应堆的临界条件。可以看出，要使有限大小反应堆维持临界状态，首先必须要求 $k_{\infty} > 1$ 。如果对于由特定的材料组成和布置的系统，它的无限介质增殖系数 $k_{\infty} > 1$ ，那么，对于这种系统必定可以通过改变反应堆芯部的大小，找到一个合适的芯部尺寸，恰好使 $k_{\infty} \Lambda = 1$ ，亦就是使反应堆处于临界状态，这时反应堆芯部的大小称为临界大小。在临界情况下，反应堆内所装载的燃料质量叫作临界质量。

反应堆的临界大小取决于反应堆的材料组成与几何形状。例如，对于采用富集铀的反应堆，它的 k_{∞} 比较大，所以即使其不泄漏概率小一点，仍然可能满足 $k_{\infty} \Lambda = 1$ 的条件。这样，用富集铀做燃料

的反应堆, 其临界大小必定小于用天然铀做燃料的反应堆。决定临界大小的另一个因素是反应堆的几何形状。由于中子总是通过反应堆的表面泄漏出去, 而中子的产生则发生在反应堆的整个体积中, 因而, 要减少中子的泄漏损失, 增加不泄漏概率, 就需要减少反应堆的表面积与体积之比。在体积相同的所有几何形状中, 球形的表面积最小, 亦即球形反应堆的中子泄漏损失最小。然而, 实际中出于工程上的考虑, 动力反应堆是做成圆柱形的。

1.3.2 单能中子扩散方程

在物理学中, 我们已熟悉分子的扩散现象, 即分子间的无规则碰撞运动。分子从浓度大的地方向浓度小的地方扩散, 并且分子扩散的速率与分子密度的梯度成正比, 也就是服从分子扩散现象中的“菲克扩散定律”。同样地, 若把一个中子源 (例如 Ra – Be 源) 放到某一介质内, 我们可以通过测量仪器观察到, 中子不断地从源点扩散开来, 经过一段时间后, 介质内到处都有中子了。由于中子密度 (在热中子反应堆内约为 10^{16} m^{-3} 数量级) 比起介质的原子核密度 (一般约为 10^{28} m^{-3}) 要小得多, 因而它与分子扩散现象不同, 其主要差别在于分子扩散是由于分子间的相互碰撞引起的, 而中子的扩散主要是中子与介质原子核间的散射碰撞的结果, 中子之间的相互碰撞可以略去不计。在中子密度大的地方, 中子与原子核碰撞的次数就多, 而每次碰撞以后, 中子通常要改变运动方向离开碰撞中心, 因此与分子的扩散相似, 中子总是从中子密度高的地方向密度低的地方扩散。下面即将证明, 与分子扩散相类似, 中子的扩散也服从与分子扩散相类似的定律, 它是中子扩散近似模型的基础。

显然中子扩散问题属于统计问题, 因而可以像气体动力论一样, 把它发展成为一种处理大量中子运动的宏观理论。例如, 在气体分子或热的扩散中扩散物质有一种倾向, 就是从分子密度大的区域向分子密度小的区域移动 (扩散)。中子的行为也一样, 从中子密度大的区域向中子密度小的区域移动 (扩散)。它的规律服从气体动力学中的扩散规律, 因此可以以某种扩散方程为基础来处理大量中子行为的 (扩散) 宏观理论。

在扩散宏观理论中, 中子密度 $n(\mathbf{r}, E)$ 是一个重要的参量, 它描述了中子群在介质内的分布情况。另外, 还有以下两个参量也是我们经常要用到的重要参数。

(1) $\phi(\mathbf{r}, E)$ 。如前所述它与中子在介质中的核反应率密切相关。

(2) $\mathbf{J}(\mathbf{r}, E)$ 。它定义为中子的净流密度

$$\mathbf{J}(\mathbf{r}, E) = \int_{4\pi} \Omega \Phi(\mathbf{r}, \Omega) d\Omega \quad (1.8)$$

从物理上看, 它是由很多具有不同方向的微分中子束矢量合成的量, 表示该处中子的净流动情况, 因而它是一个矢量。设 \mathbf{n} 表示该点的法线方向, 则

$$J_n(\mathbf{r}, E) = \mathbf{J}(\mathbf{r}, E) \cdot \mathbf{n} \quad (1.9)$$

就表示每秒穿过该点的净中子数目 (净中子流)。它表示中子在介质中的流动情况。

链式核反应堆理论用到的一个基本原理, 就是所谓的“中子数守恒”或者“中子数平衡”, 即在一定的体积 V 内, 中子总数对时间的变化率应等于该体积内中子的产生率减去该体积内中子的吸收率和泄漏率。如果 $n(\mathbf{r}, t)$ 是 t 时刻在 \mathbf{r} 处的中子密度, 则在 V 内中子数守恒的方程可写成

$$\frac{d}{dt} \int_V n(\mathbf{r}, t) dV = \text{产生率}(S) - \text{泄漏率}(L) - \text{吸收率}(A) \quad (1.10)$$

中子扩散方程就是根据这一平衡原则建立的。首先计算中子的泄漏率。让我们观察空间 \mathbf{r} 处的微体积元 dV 。我们知道 $\mathbf{J}(\mathbf{r}, t) \cdot \mathbf{n}$ 等于在 \mathbf{r} 处中子通过垂直于外法线 \mathbf{n} 的单位面积的净流动率。如果取 \mathbf{n} 为体积 V 的表面积 dS 上的外法线单位向量, 那么 $\mathbf{J}(\mathbf{r}, t) \cdot \mathbf{n} dS$ 就等于中子通过表面积元 dS 向外的净流率, 于是中子从体积 V 的整个表面泄漏出去的总速率为

$$\text{泄漏率} = \int_S \mathbf{J}(\mathbf{r}, t) \cdot \mathbf{n} dS \quad (1.11)$$

应用高斯散度公式可以把面积分变换成体积分, 于是

$$\text{泄漏率} = \int_S \mathbf{J}(\mathbf{r}, t) \cdot \mathbf{n} dS = \int_V \nabla \cdot \mathbf{J}(\mathbf{r}, t) dV = \int_V \text{div } \mathbf{J}(\mathbf{r}, t) dV \quad (1.12)$$

设源分布函数以 $S(\mathbf{r}, t)$ 表示, 它等于 t 时刻 \mathbf{r} 处每秒、每单位体积由源放出的中子数。因此, 在 V 内中子的产生率为

$$\text{产生率} = \int_V S(\mathbf{r}, t) dV \quad (1.13)$$

在 V 内中子的吸收率可以表示为

$$\text{吸收率} = \int_V \sum_a \phi(\mathbf{r}, t) dV \quad (1.14)$$

于是式 (1.10) 可以写成

$$\frac{d}{dt} \int_V n(\mathbf{r}, t) dV = \int_V S(\mathbf{r}, t) dV - \int_V \sum_a \phi(\mathbf{r}, t) dV - \int_V \text{div } \mathbf{J}(\mathbf{r}, t) dV \quad (1.15)$$

由于所有积分都是在相同的积分体积内进行的, 所以方程 (1.15) 两边被积函数必然相等, 即

$$\frac{\partial n(\mathbf{r}, t)}{\partial t} = S(\mathbf{r}, t) - \sum_a \phi(\mathbf{r}, t) - \text{div } \mathbf{J}(\mathbf{r}, t) \quad (1.16)$$

方程 (1.16) 叫做连续方程, 它实际上表征单位体积 (或微元) 内的中子数的平衡关系, 在反应堆理论中具有极为重要的意义。

连续方程 (1.16) 是具有普遍意义的, 但其中具有两个未知函数 $\phi(\mathbf{r}, t)$ 及 $\mathbf{J}(\mathbf{r}, t)$, 因此无法求解这个方程。如果在求解之前我们能找出 $\phi(\mathbf{r}, t)$ 和 $\mathbf{J}(\mathbf{r}, t)$ 之间的关系, 将其代入式 (1.16), 这样将其变成只含一个, 例如 $\phi(\mathbf{r}, t)$ 的方程, 问题就解决了。

当考虑稳态情况, 也就是中子通量密度不随时间变化时, 同时我们假设以下条件成立:

- (1) 介质是无限的、均匀的;
- (2) 在实验室坐标系中, 散射是各向同性的;
- (3) 介质的吸收截面很小, 即 $\Sigma_a \ll \Sigma_s$;
- (4) 中子通量密度是随空间位置缓慢变化的函数。

在这种情况下, 可以得到以下定理:

定理 1 (菲克定律) 中子流密度 \mathbf{J} 正比于负的中子通量密度梯度, 即

$$\mathbf{J} = -D \text{grad } \phi \quad (1.17)$$

其中比例常数 $D = \frac{\lambda_s}{3}$, 并称作扩散系数。

应该指出, 在上面的推导过程中, 曾经假设在实验室坐标系中中子散射是各向同性的, 实际上, 这一假设只有对于重核才近似成立, 在一般情况下, 这种假设是不正确的。由中子输运理论可以证明, 为了对散射的各向异性作适当的修正, 扩散系数中, 必须用输运平均自由程 λ_{tr} 来代替式中的散射平均自由程 λ_s 。菲克定律中的扩散系数 D 为

$$D = \frac{\lambda_{tr}}{3} \quad (1.18)$$

λ_{tr} 为输运平均自由程, 它等于

$$\lambda_{tr} = \frac{\lambda_s}{1 - \bar{\mu}_0} \quad (1.19)$$

其中: $\bar{\mu}_0$ 是实验室系统内的平均散射角余弦 [4], $\bar{\mu}_0 = \frac{2}{3A}$ 。对于重核, $\bar{\mu}_0 \ll 1$, 则 λ_{tr} 便近似地等于 λ_s 。式 (1.17) 表明: 任一处净中子流动的方向与中子通量密度分布的梯度的方向相反。因为 $\text{grad } \phi$ 的方向指向 ϕ 的增加方向, 所以 J 的方向指向 ϕ 减小最快的方向。

在求出中子流密度 $J(r, E)$ 和中子通量 ϕ 的关系后, 就可以从中子连续方程直接导出只含有中子通量 ϕ 的中子扩散方程。如前所述, 根据连续方程

$$\frac{\partial n(\mathbf{r}, t)}{\partial t} = S(\mathbf{r}, t) - \Sigma_a \phi(\mathbf{r}, t) - \text{div } \mathbf{J}(\mathbf{r}, t) \quad (1.20)$$

它实际上表征单位体积 (或微元) 内的中子数的平衡关系。其中方程的右端第三项表示泄漏项。根据菲克定律它可以写成

$$\text{div } \mathbf{J}(\mathbf{r}, t) = -\text{div } D \text{ grad } \phi = D \nabla^2 \phi \quad (1.21)$$

同时由于假定所有中子都具有相同的能量, 所以中子通量密度

$$\phi = nv \quad (1.22)$$

方程变为

$$D \nabla^2 \phi - \Sigma_a \phi + S = \frac{1}{v} \frac{\partial \phi}{\partial t} \quad (1.23)$$

式中: ∇^2 是拉普拉斯算符。在反应堆计算常用的几种坐标系中, ∇^2 的表达式如下:

$$\text{直角坐标系} \quad \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad (1.24)$$

$$\text{柱坐标系} \quad \nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} + \frac{\partial^2}{\partial z^2} \quad (1.25)$$

$$\text{球坐标系} \quad \nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} + \frac{1}{r^2} \cot \theta \frac{\partial}{\partial \theta} + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \quad (1.26)$$

这个方程称为单能中子扩散方程, 它是借助中子扩散的菲克定律从中子数守恒的连续方程推导而得的, 是反应堆理论中的一个基本方程, 在反应堆理论中广泛应用并占有重要的位置。

扩散方程边界条件

扩散方程式 (1.23) 是一个微分方程, 它适用于普遍情况, 在它的普遍解中将包含有任意的积分常数。为了确定这些积分常数的数值, 就要根据具体问题在普遍解上加一些限制条件, 也就是问题本身的物理或几何特性所规定的边界条件。边界条件的数目应恰好足以确定方程有唯一解。

下面讨论求解扩散方程时经常用到的几种边界条件。

- 在扩散方程适用的范围内, 中子通量密度的数值必须是正的、有限的实数。
- 在两种不同扩散性质的介质交界面上, 垂直于分界面的中子流密度相等, 中子通量密度相等。

设有两种不同介质的分界面 (见图 3), 在分界面上所有沿正 x (或负 x) 方向穿过 A 介质的中子数必定等于同一方向穿过 B 介质的中子数 (因为在分界面上不能有中子的积累或消耗), 即

$$J_x^+|_A = J_x^+|_B \quad (1.27)$$

和

$$J_x^-|_A = J_x^-|_B \quad (1.28)$$

将 J_x^+ 及 J_x^- 的表示式代入式 (1.27) 及式 (1.28), 然后两式相减便可得到

$$D_A \frac{d\phi}{dx} \Big|_A = D_B \frac{d\phi}{dx} \Big|_B \quad (1.29)$$

若两式相加则有

$$\phi_A = \phi_B \quad (1.30)$$

式 (1.29) 及式 (1.30) 便是扩散方程在分界面上的边界条件。

- 介质与真空交界的外表面上, 根据物理上的要求, 自真空返回介质的中子流等于零 (见图 4), 即

$$J_x^-|_{x=0} = 0 \quad (1.31)$$

反应堆的外表面就属于这种情况。空气虽然不是真空, 但是, 由于单位体积内的分子数比非气体介质要小得多, 所以, 中子在空气中的平均自由程比在非气体介质中的平均自由程要大得多, 因而可以把空气近似地当作真空来处理。由于中子不可能自真空中散射回到介质中来, 所以在 $x = 0$ 处沿负 x 方向上的中子流等于零。有

$$J_x^-|_{x=0} = \frac{\phi_0}{4} + \frac{\lambda_{tr}}{6} \frac{d\phi}{dx} \Big|_{x=0} = 0 \quad (1.32)$$

或者写成

$$\frac{d\phi}{dx} \Big|_{x=0} = -\frac{3\phi_0}{2\lambda_{tr}} \quad (1.33)$$

这个条件可以写成更方便的形式。如果我们假想从交界面处将中子通量密度的分布曲线按它在交界面处的斜率向真空作直线外推 (见图 4 中的虚线), 则在离开交界面某个距离 d 处的位置上中子通量密度将等于零, 于是有

$$\frac{d\phi}{dx} \Big|_{x=0} = -\frac{\phi_0}{d} \quad (1.34)$$

而根据式 (1.33) 有

$$d = \frac{2}{3}\lambda_{tr} \quad (1.35)$$

这里, d 称为直线外推距离。应当指出, 上述的直线外推距离 d 是从扩散理论求出的, 但是扩散理论在真空交界处附近不适用, 因而由此求出的 d 也是不精确的。按照更精确的中子输运理论, 对于平面边界的情况, 可以求得 $d = 0.7104\lambda_{tr}$, 以后在计算中, 我们将采用这一数值。严格地讲, 外推距离 d 与表面曲率半径有关, 但对于曲率较小的表面, 近似地应用 $0.7104\lambda_{tr}$ 这一数值不会产生很大的误差。

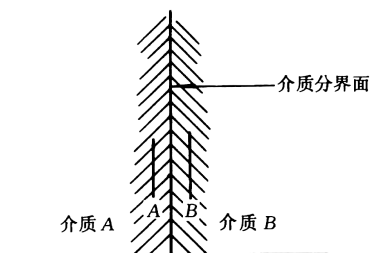


图 3 在两种介质分界面上的中子扩散

Figure 3 Neutron diffusion at the interface of two media

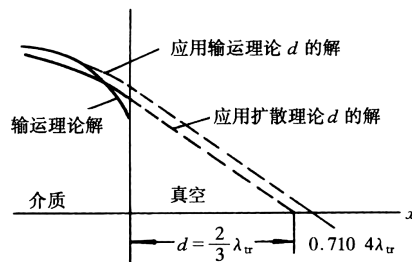


图 4 外推距离求得的扩散方程的解

Figure 4 The solution of the diffusion equation

这样, 自由外表面 (真空边界) 的边界条件可以用更简单的形式表示: 在自由表面外推距离 d 处, 中子通量密度等于零。最后必须强调指出, 这个边界条件仅仅是为了简化扩散方程的求解而采取的一种数学处理方法, 并不是说在自由表面外, 中子通量密度的变化是真的按照直线外推的那样, 在外推边界上等于零。实际上, 在自由表面以外中子通量密度的分布是没有意义的。这种外推边界仅仅是一种假想的边界概念, 只是表示利用上述外推距离作为扩散方程的边界条件可以在许多情况下得出比较满意的近似结果, 如图 4 所示。从图中可以看到, 应用输运理论的外推距离求得扩散方程的解, 除靠近边界附近有偏差外, 其它部分与输运理论的严格解吻合良好。

1.4 本章小结

本章主要对本文所需要用到的理论基础与技术进行了简单地介绍。了解了本文方法将物理信息引入神经网络的理论基础, 并对本文使用的网络结构的理论基础进行介绍, 最后介绍了一种不常用的二阶优化算法 L-BFGS, 为后文的方法介绍做好了理论方面的准备。

2 结束语

本文主要介绍了 PINNuclear-Neutrons 项目的学习笔记, 主要是对 PINNuclear-Neutrons 项目的学习过程进行了总结, 复习了论文中的算例代码, 并对项目中的一些关键代码进行了解读, 对项目中的一些关键参数进行了调整, 以便更好地理解论文中的算法。

参考文献

- 1 LAGARIS, I. E., LIKAS, A., AND FOTIADIS, D. I. Artificial Neural Networks for Solving Ordinary and Partial Differential Equations. *IEEE Trans. Neural Netw.* 9, 5 (Sept./1998), 987–1000.
- 2 RAISSI, M., PERDIKARIS, P., AND KARNIADAKIS, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics* 378 (Feb. 2019), 686–707.
- 3 刘东, 罗琦, 唐雷, 安萍, AND 杨帆. 基于 PINN 深度机器学习技术求解多维中子学扩散方程. *核动力工程* 43, 2 (2022), 1–8.
- 4 谢仲生. 核反应堆物理分析, 第五版 ed. 西安交通大学出版社, 2020.

PINNuclear-Neutrons Project Notes

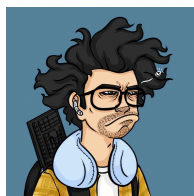
Liu Yang¹

1. *School of Mathematical Sciences, Sichuan Normal University, Chengdu 610068, Sichuan*

E-mail: mathliuyang@163.com

Abstract PINNuclear-Neutrons project aims to learn and reproduce the solution of neutron related problems in nuclear reactor, using physical reaction neural network (PINN) and deep machine learning technology. Our main goal is to re-implement the "Solving Multidimensional Neutron Diffusion Equation Based on PINN Deep Machine Learning Technology" published by Liu Dong's team, so as to simulate neutron transport more accurately.

Keywords PINN; Deep Learning; Nuclear Reactor; Neutron Diffusion Equation



Liu Yang was born in Sichuan in 1997. He is currently a master's student in the School of Mathematical Sciences, Sichuan Normal University. His research interests include partial differential equations and mathematical physics.