

# PINNuclear-Neutrons 项目笔记

刘洋<sup>1</sup>

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. 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

## 目录

<b>1 预备知识: PINN 相关理论和技术</b>	<b>3</b>
1.1 基于物理信息的神经网络	3
1.2 L-BFGS 优化算法	4
<b>2 预备知识: 核反应堆物理分析</b>	<b>4</b>
2.1 裂变能量的释放	5
2.2 链式裂变反应	5
2.3 热中子反应堆内的中子循环	7
2.4 单能中子扩散方程	9
2.4.1 扩散方程边界条件	12
2.4.2 菲克定律和扩散理论的适用范围	13
2.5 均匀裸堆的单群理论	14
2.5.1 均匀裸堆的单群扩散方程的解	15
2.5.2 热中子反应堆的临界条件	17
2.5.3 反应堆曲率和临界计算任务	20
2.5.4 单群理论的修正	21
2.5.5 稳态单群扩散方程	21

<b>3 学习笔记：基于 PINN 深度机器学习技术求解多维中子学扩散方程</b>	<b>22</b>
3.1 理论原理	22
3.1.1 中子学扩散方程模型	22
3.1.2 利用 PINN 求解扩散方程的关键技术	24
3.1.3 扩散方程特征向量加速收敛方法	24
3.1.4 $k_{\text{eff}}$ 并行搜索方法	24
3.1.5 几何空间网格点密度不均匀分布策略	25
3.2 数值验证程序复现	26
3.2.1 深度机器学习的超参数设置	26
3.2.2 临界条件下稳态扩散方程的验证	26
3.2.3 扩散方程特征向量加速收敛方法验证	34
<b>4 结束语</b>	<b>34</b>

## 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.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.通过本学习笔记本,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 预备知识：PINN 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.

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得到很好的回归效果, 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 1.1 所示.

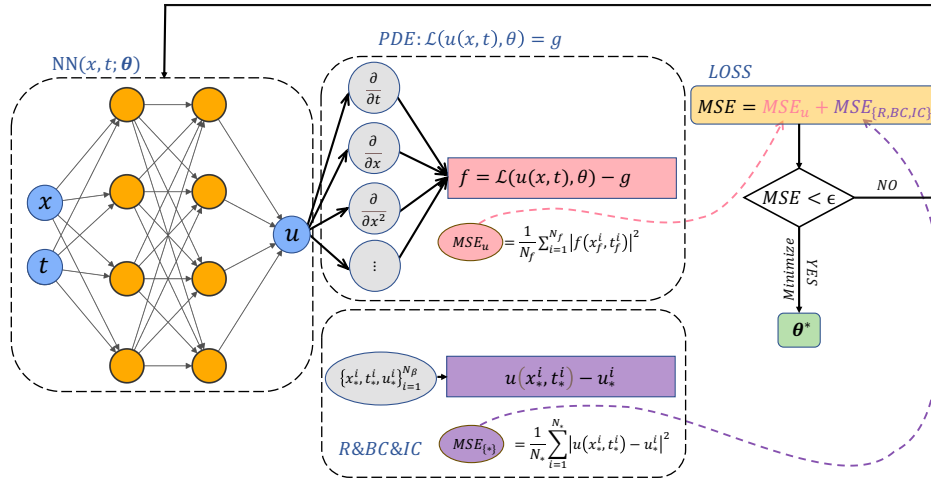


图 1.1 PINN 网络结构图  
Figure 1.1 PINN Network Structure

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  $\delta_k$  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, 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.

## 2 预备知识: 核反应堆物理分析

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.

2.1 裂变能量的释放

根据结合能公式可以算出, 实验上也已测出,  $^{235}\text{U}$  核一次裂变大约释放 200MeV 的能量, 其中裂变碎片的动能约占总释放能量的 80%, 裂变能量的大致分配如表 1 所示. 裂变能量的绝大部分都在堆内转变为热能. 由于中微子不带电, 其质量几乎为零, 因而它几乎不与堆内任何物质作用. 因此, 中微子所具有的 12MeV 的能量, 是不能利用的. 另一方面, 裂变中子将被堆内各种材料的核吸收而发生  $(n, \gamma)$  反应, 这要释放出 3 ~ 12MeV 的能量. 虽然这部分能量并不是核裂变时直接放出来的, 但它也是裂变带来的后果, 并且这部分能量的绝大部分也在堆内转变为热能, 故通常把这一部分能量也归入裂变能量. 可利用的裂变能量中大约 97% 分配在燃料内, 不到 1% (为  $\gamma$  射线能量) 在堆屏蔽层内, 其余能量则分配在冷却剂和结构材料内. 确切地讲, 每次裂变可利用的能量因堆型的不同而不同, 一般计算时, 可以近似地认为  $^{235}\text{U}$  核每次裂变可利用的能量约为 200MeV .

表 1 核裂变释放的能量

能量形式	能量 /MeV	能量形式	能量 /MeV
裂变碎片的动能	168	裂变产物 $\gamma$ 衰变-缓发 $\gamma$ 能量	7
裂变中子的动能	5	裂变产物 $\beta$ 衰变-缓发 $\beta$ 能量	8
瞬发 $\gamma$ 能量	7	中微子能量	12
共计		207	

应该指出, 裂变产物的衰变  $\beta$  和  $\gamma$  射线的能量约占总裂变能量的 4% ~ 5%, 它们是裂变碎片在衰变过程中发射出来的, 即这部分能量释放是有一段时间延迟的. 因而停堆后, 仍然会有衰变热量产生, 仍需进行冷却和屏蔽. 这种停堆后衰变余热的导出问题是反应堆安全研究中重要的问题之一.

2.2 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 2.1 所示.

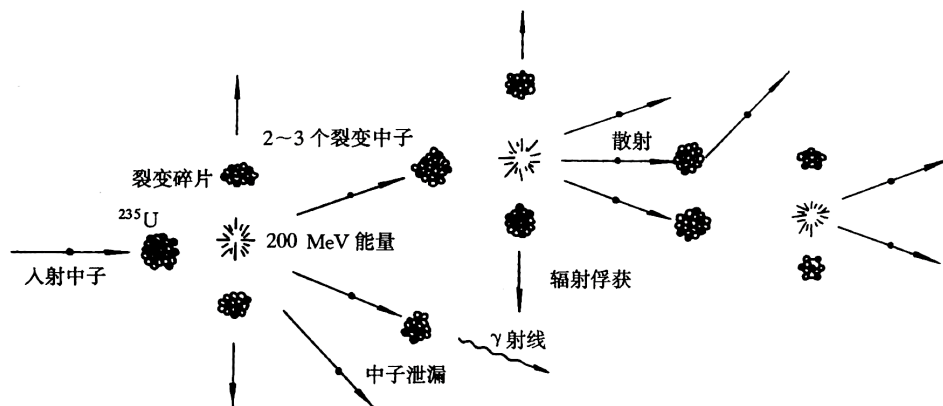


图 2.1 链式裂变反应示意图

Figure 2.1 Chain fission reaction diagram

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  $k_{\text{eff}}$  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

$$k_{\text{eff}} = \frac{\text{新生一代中子数}}{\text{直属上一代中子数}} \quad (2.1)$$

上式的定义是直观地从中的“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

$$k_{\text{eff}} = \frac{\text{系统内中子的产生率}}{\text{系统内中子的总消失 (吸收 + 泄漏) 率}} \quad (2.2)$$

若芯部的有效增殖系数  $k_{\text{eff}}$  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  $k_{\text{eff}}$  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  $k_{\text{eff}}$  大于 1, 则系统内的中子数目将随时间而不断地增加, 称这种系统为超临界系统.

显然, 有效增殖系数  $k_{\text{eff}}$  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, 把无限大介质的增殖系数称为无限介质增殖系数, 以  $k_{\infty}$  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  $\Lambda$ , 它的定义是

$$\Lambda = \frac{\text{系统内中子的吸收率}}{\text{系统内中子的泄漏率}} \quad (2.3)$$

不泄漏概率  $\Lambda$  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 (2.2) 和式 (2.3) 可知, 有限尺寸芯部的有效增殖系数为

$$k_{\text{eff}} = k_{\infty} \Lambda \quad (2.4)$$

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

$$k_{\text{eff}} = k_{\infty} \Lambda = 1 \quad (2.5)$$

式 (2.5) 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  $k_{\infty}$  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.

### 2.3 热中子反应堆内的中子循环

为了讨论反应堆内中子产生和消亡之间的平衡关系, 人为地将中子分成一代一代来处理是有益的. 热中子反应堆内每代中子循环, 可以近似地用图 2.2 表示. 下面将对这一循环过程作进一步讨论. 从以上的讨论可以知道, 反应堆内中子数目的增减与平衡, 主要取决于下列几个过程.

- (1)  $^{238}\text{U}$  的快中子增殖.
- (2) 慢化过程中的共振吸收.
- (3) 慢化剂以及结构材料等物质的辐射俘获.

(4) 燃料吸收热中子引起的裂变.

(5) 中子的泄漏. 包括: 慢化过程中的泄漏; 热中子扩散过程中的泄漏.

反应堆内中子数目的变化取决于上述 5 种过程竞争的结果. 其中快中子增殖和燃料裂变将使反应堆内的中子数目增加, 其它过程使中子数目减少. 为了定量计算, 就以上 5 个过程, 定义以下 5 个参量.

(1) 快中子增殖系数  $\epsilon$ . 它的定义: 由一个初始裂变中子所得到的、慢化到  $^{238}\text{U}$  裂变阈能以下的平均中子数. 由于初始裂变中子中, 大约有 60% 的中子的能量在  $^{238}\text{U}$  裂变阈能 (1.1MeV) 以上, 这些中子与  $^{238}\text{U}$  核作用时, 有一部分能引起  $^{238}\text{U}$  裂变而产生快中子, 这一过程称为  $^{238}\text{U}$  的快中子增殖效应.

(2) 逃脱共振俘获概率  $p$ . 裂变产生的快中子的平均能量为 2MeV, 在它们慢化的过程中, 要经过共振能区 ( $1 \sim 10^4\text{eV}$ ), 而  $^{238}\text{U}$  核在该能区有许多共振峰 (见图 2.2). 因而在慢化中, 裂变产生的快中子中必然有一部分被  $^{238}\text{U}$  核共振吸收而损失掉, 只有一部分快中子慢化至热中子. 在慢化过程中逃脱共振吸收的中子份额就称为逃脱共振俘获概率, 用  $p$  表示.

(3) 热中子利用系数  $f$ . 它表示被燃料吸收的热中子数占被芯部中所有物质 (包括燃料在内) 吸收的热中子总数的份额.  $f$  定义为

$$f = \frac{\text{燃料吸收的热中子数}}{\text{燃料吸收的热中子数} + \text{慢化剂吸收的热中子数} + \text{冷却剂吸收的热中子数} + \text{结构材料吸收的热中子数}} \quad (2.6)$$

这里分母中包括被燃料、慢化剂、冷却剂和结构材料等所有物质吸收的热中子总数. 对于均匀堆, 由于各种材料内的热中子通量密度相等, 因而

$$f = \frac{N_f \sigma_{a,f}}{N_f \sigma_{a,f} + N_m \sigma_{a,m} + N_c \sigma_{a,c} + N_s \sigma_{a,s}} \quad (2.7)$$

其中,  $N_f$   $N_m$   $N_c$  和  $N_s$  分别为在均匀混合物单位体积中燃料、慢化剂、冷却剂和结构材料的核子数.

(4) 有效裂变中子数  $\eta$ . 它的定义: 核燃料每吸收一个热中子所产生的平均裂变中子数. 设  $\Sigma_a$  和  $\Sigma_f$  分别为燃料的热中子宏观吸收截面和宏观裂变截面. 由于燃料每吸收一个热中子引起裂变的概率为  $\Sigma_f/\Sigma_a$ , 若设每次裂变所产生的平均裂变中子数为  $\nu$ , 则显然有

$$\eta = \frac{\Sigma_f}{\Sigma_a} \nu \quad (2.8)$$

(5) 不泄漏概率  $\Lambda$ . 它是中子在慢化过程和热中子在扩散过程中不泄漏概率的乘积, 为

$$\Lambda = \Lambda_s \Lambda_d \quad (2.9)$$

其中,  $\Lambda_s$  为慢化过程中的不泄漏概率;  $\Lambda_d$  为热中子扩散过程中的不泄漏概率.

定义了以上这些量以后, 便可以进一步定量地来研究热中子反应堆内中子的平衡过程. 假设在某一代开始时有  $n$  个裂变中子, 由于  $^{238}\text{U}$  快中子增殖的结果, 中子数目将增加到  $n\epsilon$  个. 这些中子继续

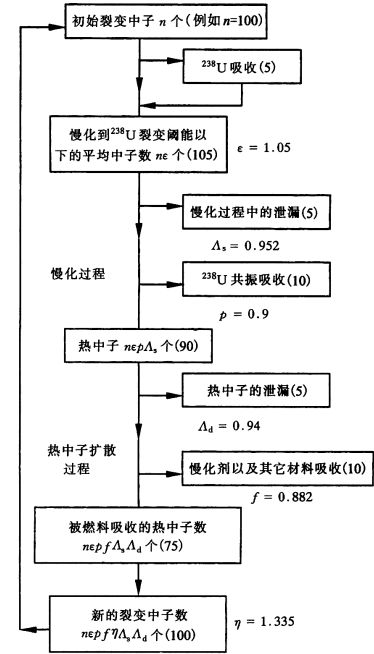


图 2.2 热中子反应堆内的中子平衡

Figure 2.2 Neutron balance in thermal neutron reactor



慢化,但是由于共振吸收将损失一部分中子,所以只有  $n\epsilon p$  个中子能够逃脱共振吸收而慢化成热中子. 如果考虑到中子的泄漏损失,那么,实际上被吸收的热中子数目将只有  $n\epsilon p\Lambda_s\Lambda_d$  个. 显然,其中被燃料所吸收的热中子数目等于  $n\epsilon pf\Lambda_s\Lambda_d$  个,其余部分的热中子被其它材料所吸收. 被燃料吸收的热中子将使燃料发生核裂变反应,而又重新放出新的裂变中子. 由于燃料每吸收一个热中子将产生  $\eta$  个裂变中子,因而新的裂变中子数目等于  $n\epsilon pf\eta\Lambda_s\Lambda_d$ . 作为一个例子,在图 2.2 中给出了各个过程的具体数值,以便读者有定量的概念. 根据有效增殖系数的定义,便可得出

$$k_{\text{eff}} = \frac{n\epsilon pf\eta\Lambda_s\Lambda_d}{n} = k_{\infty}\Lambda \quad (2.10)$$

其中,  $\Lambda = \Lambda_s\Lambda_d$ . 若根据图 2.2 中给出的数值,则  $\Lambda = 0.8987$ . 由式 (2.4) 可以看出

$$k_{\infty} = \epsilon pf\eta \quad (2.11)$$

式 (2.11) 称为四因子公式. 以上对热中子反应堆内中子平衡的分析方法称为“四因子模型”. 在早期反应堆物理分析和计算中,它曾被广泛地应用. 同时,它对热中子反应堆内中子的循环过程可以给出清晰的物理概念和形象的图形. 但是它仅仅是对充分热化的热中子反应堆内的物理过程的一种近似的、简单的描述,它并不能严格地描述一些更为复杂的反应堆内的物理过程. 目前,伴随着对核反应堆物理过程的深入了解,同时也随着计算机的发展,它已被更为精确的建立在 neutron 输运理论基础上的多群模型所代替,而直接用数值方法求解分群扩散或中子输运方程. 尽管如此,在反应堆物理中无限介质增殖系数  $k_{\infty}$  仍然是一个经常被用到的重要概念参数.

## 2.4 单能中子扩散方程

在物理学中,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,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  $10^{16} \text{ m}^{-3}$  数量级)比起介质的原子核密度(一般约为  $10^{28} \text{ m}^{-3}$ )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经常要用到的重要参数.

(1)  $\phi(r, E)$  . 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

$$\mathbf{J}(\mathbf{r}, E) = \int_{4\pi} \Omega \Phi(\mathbf{r}, \Omega) d\Omega \quad (2.12)$$

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

$$J_n(\mathbf{r}, E) = \mathbf{J}(\mathbf{r}, E) \cdot \mathbf{n} \quad (2.13)$$

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

$$\frac{d}{dt} \int_V n(\mathbf{r}, t) dV = \text{产生率}(S) - \text{泄漏率}(L) - \text{吸收率}(A) \quad (2.14)$$

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. 让观察空间  $\mathbf{r}$  处的微体积分元  $dV$  . 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

$$\text{泄漏率} = \int_S \mathbf{J}(\mathbf{r}, t) \cdot \mathbf{n} dS \quad (2.15)$$

应用高斯散度公式可以把面积分变换成体积分, 于是

$$\text{泄漏率} = \int_S \mathbf{J}(\mathbf{r}, t) \cdot \mathbf{n} dS = \int_V \nabla \cdot \mathbf{J}(\mathbf{r}, t) dV = \int_V \text{div } \mathbf{J}(\mathbf{r}, t) dV \quad (2.16)$$

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

$$\text{产生率} = \int_V S(\mathbf{r}, t) dV \quad (2.17)$$

在  $V$  内中子的吸收率可以表示为

$$\text{吸收率} = \int_V \sum_a \phi(\mathbf{r}, t) dV \quad (2.18)$$

于是式 (2.14) 可以写成

$$\frac{d}{dt} \int_V n(\mathbf{r}, t) dV = \int_V S(\mathbf{r}, t) dV - \int_V \sum_a \phi(\mathbf{r}, t) dV - \int_V \text{div } \mathbf{J}(\mathbf{r}, t) dV \quad (2.19)$$

由于所有积分都是在相同的积分体积内进行的, 所以方程 (2.19) 两边被积函数必然相等, 即

$$\frac{\partial n(\mathbf{r}, t)}{\partial t} = S(\mathbf{r}, t) - \sum_a \phi(\mathbf{r}, t) - \text{div } \mathbf{J}(\mathbf{r}, t) \quad (2.20)$$

方程 (2.20) 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 (2.20) 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. 如果在求解之前能找出  $\phi(\mathbf{r}, t)$  和  $\mathbf{J}(\mathbf{r}, t)$  之间的关系, 将其代入式 (2.20), 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. 为此, 考虑稳态情况, 也就是中子通量密度不随时间变化时, 同时假设以下条件成立:

- (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

$$\mathbf{J} = -D \text{grad } \phi \quad (2.21)$$

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

应该指出, 在上述定理的推导过程中, 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  $\lambda_{tr}$  来代替式中的散射平均自由程  $\lambda_s$ . 菲克定律中的扩散系数  $D$  为

$$D = \frac{\lambda_{tr}}{3} \quad (2.22)$$

$\lambda_{tr}$  为输运平均自由程, 它等于

$$\lambda_{tr} = \frac{\lambda_s}{1 - \bar{\mu}_0} \quad (2.23)$$

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  $\lambda_{tr}$  便近似地等于  $\lambda_s$ . 式 (2.21) 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  $\phi$  的增加方向, 所以  $\mathbf{J}$  的方向指向  $\phi$  减小最快的方向.

在求出中子流密度  $\mathbf{J}(\mathbf{r}, E)$  和中子通量  $\phi$  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  $\phi$  的中子扩散方程. 如前所述, 根据连续方程

$$\frac{\partial n(\mathbf{r}, t)}{\partial t} = S(\mathbf{r}, t) - \Sigma_a \phi(\mathbf{r}, t) - \text{div } \mathbf{J}(\mathbf{r}, t) \quad (2.24)$$

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

$$\text{div } \mathbf{J}(\mathbf{r}, t) = -\text{div } D \text{grad } \phi = D \nabla^2 \phi \quad (2.25)$$

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

$$\phi = nv \quad (2.26)$$

方程变为

$$D\nabla^2\phi - \Sigma_a\phi + S = \frac{1}{v}\frac{\partial\phi}{\partial t} \quad (2.27)$$

其中,  $\nabla^2$  是拉普拉斯算符. 在反应堆计算常用的几种坐标系中,  $\nabla^2$  的表达式如下:

$$\text{直角坐标系} \quad \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad (2.28)$$

$$\text{柱坐标系} \quad \nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} + \frac{\partial^2}{\partial z^2} \quad (2.29)$$

$$\text{球坐标系} \quad \nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} + \frac{1}{r^2}\cot\theta\frac{\partial}{\partial \theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial \varphi^2} \quad (2.30)$$

方程 (2.27) 称为单能中子扩散方程, 它是借助中子扩散的非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很重要的位置.

**注释 1** 在扩散理论中有两个重要的物理参数: 扩散系数  $D$  和扩散长度  $L$ . 它们是确定中子在介质内扩散过程的重要参数. 根据定义, 扩散长度为

$$L^2 = \frac{D}{\Sigma_a} = \frac{\lambda_a\lambda_{tr}}{3} \quad (2.31)$$

或

$$L^2 = \frac{\lambda_a\lambda_s}{3(1-\bar{\mu}_0)} = \frac{1}{3\Sigma_a\Sigma_s(1-\bar{\mu}_0)} \quad (2.32)$$

这便是扩散长度的计算公式. 对于混合物, 则式 (2.32) 中的  $\Sigma_a$   $\Sigma_s$  和  $\bar{\mu}_0$  等都是指混合物的平均值. 对于热中子, 式中的  $\Sigma_a$  和  $\Sigma_s$  应该是热中子能谱的平均值.

### 2.4.1 扩散方程边界条件

扩散方程式 (2.27) 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 2.3), 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

$$J_x^+|_A = J_x^+|_B \quad (2.33)$$

和

$$J_x^-|_A = J_x^-|_B \quad (2.34)$$

将  $J_x^+$  及  $J_x^-$  的表示式代入式 (2.33) 及式 (2.34), 然后两式相减便可得到

$$D_A \frac{d\phi}{dx}\bigg|_A = D_B \frac{d\phi}{dx}\bigg|_B \quad (2.35)$$

若两式相加则有

$$\phi_A = \phi_B \quad (2.36)$$

式 (2.35) 及式 (2.36) 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 2.4), 即

$$J_x^-|_{x=0} = 0 \quad (2.37)$$

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

$$J_x^-|_{x=0} = \frac{\phi_0}{4} + \frac{\lambda_{tr}}{6} \frac{d\phi}{dx} \Big|_{x=0} = 0 \quad (2.38)$$

或者写成

$$\frac{d\phi}{dx} \Big|_{x=0} = -\frac{3\phi_0}{2\lambda_{tr}} \quad (2.39)$$

这个条件可以写成更方便的形式. 如果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 2.4 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

$$\frac{d\phi}{dx} \Big|_{x=0} = -\frac{\phi_0}{d} \quad (2.40)$$

而根据式 (2.39) 有

$$d = \frac{2}{3}\lambda_{tr} \quad (2.41)$$

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, 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. 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 2.4 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.

#### 2.4.2 菲克定律和扩散理论的适用范围

在推导菲克定律的过程中, 曾经作了一些假设, 因而它和以它为基础的扩散理论只能在一定范围内适用.

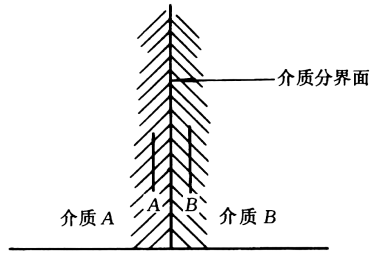


图 2.3 在两种介质分界面上的中子扩散

Figure 2.3 Neutron diffusion at the interface of two media

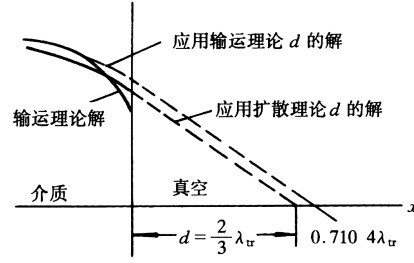


图 2.4 外推距离求得的扩散方程的解

Figure 2.4 The solution of the diffusion equation

首先, 假定了扩散介质是无限的, 这样, 推导过程的积分延伸到无穷远处. 但是, 在积分中因子  $e^{-\Sigma_s |l|}$  是随距离  $|l|$  按指数函数规律迅速减小的, 例如, 在距离为三个平均自由程 ( $\Sigma_s l \approx 3$ ) 时,  $e^{-\Sigma_s |l|}$  的数值便已小于 0.05 了, 因此, 对推导过程中积分的贡献主要是来自距所讨论点  $P$  几个平均自由程以内的中子. 所以, 在有限介质内, 在其距离表面几个自由程以外的内部区域菲克定律便是成立的, 而在距真空边界两三个自由程内的区域, 它是不适用的.

其次, 在推导中把中子通量密度展成泰勒级数并只取到一阶项 (实际上, 保留二阶项, 其结果也是一样的, 因为二阶项在积分中或是等于零, 或者在  $J^+$  和  $J^-$  中以相等形式出现, 因而在求净中子流密度  $J$  时便互相抵消了), 因此, 要求各点的中子通量密度能够用泰勒级数展开前三项表示便足够准确. 这就要求在所讨论点的几个平均自由程内, 中子通量密度必须是缓慢地变化或者它的梯度变化不大. 实际上, 在强吸收体 (如控制棒) 附近, 或者在两种扩散性质显著不同的交界面附近的几个自由程内, 中子通量密度分布将发生急剧的变化, 因此, 菲克定律在上述区域内不能适用, 此外, 正如假设条件所述, 它只适用于  $\Sigma_a \ll \Sigma_s$  的弱吸收介质.

最后, 在推导过程的积分中假定, 对中子流密度的贡献仅仅来自中子与介质核的散射碰撞, 并没有考虑中子源的贡献. 在强的中子源附近, 中子通量密度的变化率往往是很大的, 这就破坏了假设的条件. 但是, 由于在中子流密度的积分中含有  $e^{-\Sigma_s |l|}$  衰变因子, 因此在离开中子源几个平均自由程以外的区域, 源中子的贡献和影响便可忽略了. 由此得出: 在距强的中子源两三个平均自由程的区域内, 菲克定律不适用.

菲克定律是扩散理论的基础, 因而在反应堆物理分析中应用这一近似时, 应牢记上述这些限制条件. 在上述菲克定律不能适用的地方, 必须应用其它更精确的方法 (例如输运理论), 或者对扩散理论所得结果作一些必要的修正.

## 2.5 均匀裸堆的单群理论

对于由燃料和慢化剂组成的均匀增殖介质的反应堆系统, 根据裂变反应率的物理含义, 反应堆内单位时间、单位体积内的裂变中子源强可写为

$$S_F(\mathbf{r}, t) = \nu \Sigma_f \phi(\mathbf{r}, t) \quad (2.42)$$

或者根据无限介质增殖系数的定义, 可表示为介质  $k_\infty$  和中子吸收率的乘积, 即

$$S_F(\mathbf{r}, t) = k_\infty \Sigma_a \phi(\mathbf{r}, t) \quad (2.43)$$

以上两个表示式虽然形式不同,但其实质却是完全相同的. 因此,在简单单群近似下有

$$k_{\infty} = \frac{\nu \Sigma_f \phi(\mathbf{r})}{\Sigma_a \phi(\mathbf{r})} = \frac{\nu \Sigma_f}{\Sigma_a} \quad (2.44)$$

将以上给出的裂变中子源项代入式 (2.27), 为不失一般性, 还考虑系统内可能存在独立的外中子源  $S_0(\mathbf{r}, t)$ , 这样与时间相关的反应堆单群中子扩散方程就可以写成

$$\frac{1}{v} \frac{\partial \phi(\mathbf{r}, t)}{\partial t} = D \nabla^2 \phi(\mathbf{r}, t) - \Sigma_a \phi(\mathbf{r}, t) + k_{\infty} \Sigma_a \phi(\mathbf{r}, t) + S_0(\mathbf{r}, t) \quad (2.45)$$

通过式(2.31), 将吸收截面  $\Sigma_a$  用扩散长度  $L$  和扩散系数  $D$  表示出来, 有  $\Sigma_a = \frac{D}{L^2}$ , 则上式可写成

$$\frac{1}{v} \frac{\partial \phi(\mathbf{r}, t)}{\partial t} = D \nabla^2 \phi(\mathbf{r}, t) - \frac{D}{L^2} \phi(\mathbf{r}, t) + k_{\infty} \frac{D}{L^2} \phi(\mathbf{r}, t) + S_0(\mathbf{r}, t) \quad (2.46)$$

$$\frac{1}{v} \frac{\partial \phi(\mathbf{r}, t)}{\partial t} = D \nabla^2 \phi(\mathbf{r}, t) + (k_{\infty} - 1) \frac{D}{L^2} \phi(\mathbf{r}, t) + S_0(\mathbf{r}, t) \quad (2.47)$$

$$\frac{1}{Dv} \frac{\partial \phi(\mathbf{r}, t)}{\partial t} = \nabla^2 \phi(\mathbf{r}, t) + \frac{k_{\infty} - 1}{L^2} \phi(\mathbf{r}, t) + \frac{1}{D} S_0(\mathbf{r}, t) \quad (2.48)$$

注意式中出现的  $D$  和  $\Sigma_a$  都是对中子能谱平均后的数值. 值得指出的是, 在工程实际中, 除了反应堆启动的初期阶段, 由于反应堆内中子通量密度水平很低, 必须考虑外源中子的影响之外, 在大多数情况下都可以忽略外中子源的影响, 认为核裂变是反应堆内中子的唯一来源.

### 2.5.1 均匀裸堆的单群扩散方程的解

本节以一个长、宽为无限大, 厚度 (包括外推距离在内) 为  $a$  的平板裸堆为例 (见图 2.5), 来讨论单群扩散方程所描述的核反应堆特性. 选取这种假想的“平板反应堆”的主要原因在于一维平板几何条件十分简单, 有利于详细求解单群扩散方程.

在无外源的情况下, 描述平板反应堆内中子通量密度变化规律的方程为

$$\frac{1}{v} \frac{\partial \phi(x, t)}{\partial t} = D \frac{\partial^2 \phi(x, t)}{\partial x^2} - \Sigma_a \phi(x, t) + k_{\infty} \Sigma_a \phi(x, t) \quad (2.49)$$

其初始条件为

$$\phi(x, 0) = \phi_0(x) \quad (2.50)$$

边界条件为在反应堆外推边界处, 中子通量密度等于零, 即

$$\phi\left(\frac{a}{2}, t\right) = \phi\left(-\frac{a}{2}, t\right) = 0 \quad (2.51)$$

注意这里已假设初始中子通量密度是对称的. 用  $D$  除式 (2.49) 各项, 并注意到  $L^2 = D/\Sigma_a$ , 最后可得到

$$\frac{1}{Dv} \frac{\partial \phi(x, t)}{\partial t} = \nabla^2 \phi(x, t) + \frac{k_{\infty} - 1}{L^2} \phi(x, t) \quad (2.52)$$

这是一个二阶的偏微分方程, 对这类方程常可以采用分离变量法求解, 即求如下形式的解

$$\phi(x, t) = \varphi(x)T(t) \quad (2.53)$$

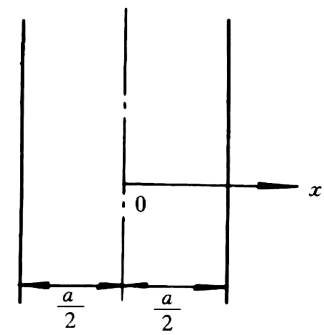


图 2.5 无限平板形反应堆  
Figure 2.5 Infinite plate reactor

将式 (2.53) 代入式 (2.52) 中, 并用  $\varphi(x)T(t)$  除方程中的各项, 便得到

$$\frac{\nabla^2 \varphi(x)}{\varphi(x)} = \frac{1}{DvT(t)} \frac{dT(t)}{dt} - \frac{k_\infty - 1}{L^2} \quad (2.54)$$

上式左端是仅含  $x$  的函数, 而右端是仅含  $t$  的函数, 因此要使等式成立, 等式两端必须都等于某一常数 (例如,  $-B^2$ ). 这样有

$$\frac{\nabla^2 \varphi(x)}{\varphi(x)} = -B^2 \quad (2.55)$$

或

$$\nabla^2 \varphi(x) + B^2 \varphi(x) = 0 \quad (2.56)$$

这里,  $B^2$  为一待定常数. 式 (2.56) 为典型的波动方程,  $B^2$  称为方程的“特征值”. 容易得出其通解为

$$\varphi(x) = A \cos Bx + C \sin Bx \quad (2.57)$$

其中,  $A$ 、 $C$  为待定常数. 由于初始通量密度分布  $\phi_0(x)$  关于  $x = 0$  平面对称, 因此, 针对的问题只能选用满足对称性条件的解, 即

$$\varphi(x) = A \cos Bx \quad (2.58)$$

由边界条件式 (2.51), 可导出  $\varphi(x)$  应满足如下的边界条件:  $\varphi(\pm \frac{a}{2}) = 0$

$$A \cos \frac{Ba}{2} = 0 \quad (2.59)$$

因此要求

$$B_n = \frac{n\pi}{a} \quad n = 1, 3, 5, \dots \quad (2.60)$$

或

$$B_n = \frac{(2n-1)\pi}{a} \quad n = 1, 2, 3, \dots \quad (2.61)$$

对应于其中的任一  $B_n$  值, 都可以给出满足微分方程 (2.56) 及边界条件 (2.51) 的解

$$\varphi_n(x) = A_n \cos B_n x = A_n \cos \frac{(2n-1)\pi}{a} x \quad (2.62)$$

从上面的求解过程可以得出结论, 齐次方程 (2.56) 只对某些特定的特征值  $B_n$  才有解. 相应的解  $\varphi_n(x)$  称为此问题的特征函数.

下面接着来讨论方程 (2.54) 的求解, 由于特征函数的正交性<sup>1)</sup>, 对于每一个  $n$  值的项都是线性独立的, 因而对应每一个  $B_n^2$  值和  $\varphi_n(x)$ , 都有一个  $T_n(t)$  与之对应, 由方程 (2.54) 得

$$\frac{1}{DvT_n(t)} \frac{dT_n(t)}{dt} = \frac{k_\infty - 1}{L^2} - B_n^2 \quad (2.63)$$

用  $L^2/(1 + L^2 B_n^2)$  乘上式的每一项, 可以化简得到下列等式

$$\frac{1}{T_n(t)} \frac{dT_n(t)}{dt} = \frac{k_n - 1}{l_n} \quad (2.64)$$

1) 单群特征函数的正交性可表示为  $\int_V \varphi_n(\mathbf{r}) \varphi_m(\mathbf{r}) dV = 0$ , 当  $m \neq n$  时.



其中

$$l_n = \frac{L^2}{Dv(1 + L^2 B_n^2)} = \frac{l_\infty}{1 + L^2 B_n^2} \quad (2.65)$$

$$k_n = \frac{k_\infty}{1 + L^2 B_n^2} \quad (2.66)$$

$l_\infty = \frac{\lambda_a}{v}$  为无限介质的热中子寿命,  $\lambda_a$  是热中子的平均吸收自由程. 方程 (2.64) 的解为

$$T_n = C_n e^{(k_n - 1)t/l_n} \quad (2.67)$$

其中,  $C_n$  为待定常数. 这样, 对于一维平板反应堆, 其中子通量密度的完全解就是对  $n = 1$  到  $n = \infty$  的所有项的总和, 即

$$\phi(x, t) = \sum_{n=1}^{\infty} A'_n \left[ \cos \frac{(2n-1)\pi}{a} x \right] e^{(k_n-1)t/l_n} \quad (2.68)$$

根据问题的初值条件式 (2.50), 可确定出上式中的待定系数  $A'_n$ . 为此, 令式中的  $t = 0$ , 可得

$$\phi_0(x) = \sum_{n=1}^{\infty} A'_n \left[ \cos \frac{(2n-1)\pi}{a} x \right] \quad (2.69)$$

利用正交关系, 可求得

$$A'_n = \frac{2}{a} \int_{-\frac{a}{2}}^{\frac{a}{2}} \phi_0(x) \cos \frac{(2n-1)\pi}{a} x \, dx \quad (2.70)$$

将其代入式 (2.68), 可得到图 2.5 所示无限平板反应堆内的中子通量密度分布为

$$\phi(x, t) = \sum_{n=1}^{\infty} \left[ \frac{2}{a} \int_{-\frac{a}{2}}^{\frac{a}{2}} \phi_0(x') \cos \frac{(2n-1)\pi}{a} x' \, dx' \right] \cos \frac{(2n-1)\pi}{a} x e^{(k_n-1)t/l_n} \quad (2.71)$$

### 2.5.2 热中子反应堆的临界条件

通过前一小节的讨论可以知道, 特征值  $B_n^2$  随  $n$  的增加而单调增大, 最小特征值是  $n = 1$  时的  $B_1^2$  值, 而同时又从式 (2.66) 可以看出, 当  $n$  增加时,  $k_n$  单调递减, 也就是说对应于最小特征值  $B_1^2$  的  $k_1$  是  $k_1, \dots, k_n$  中的最大值. 另外, 从上一小节对平板反应堆的讨论可看出  $B_n^2$  与系统尺寸有关, 当系统尺寸加大时,  $B_n^2$  便减小, 因而改变系统的尺寸就可以改变  $B_n^2$  值, 从而也就改变了  $k_n$  值. 分以下几种情况对式 (2.68) 进行讨论.

第一种情况: 对于一定几何形状和体积的反应堆芯部, 若  $B_1^2$  对应的  $k_1$  小于 1, 那么, 其余的  $k_2, \dots, k_n$  都将小于 1, 这时所有的  $(k_n - 1)$  都是负值, 从式 (2.68) 可以看出,  $\phi(x, t)$  将随时间  $t$  按指数规律衰减, 因而系统处于次临界状态.

第二种情况: 若  $k_1 > 1$ , 则  $(k_1 - 1) > 0$ , 这时中子通量密度将随时间不断地增长, 反应堆将处于超临界状态.

第三种情况: 若通过调整反应堆堆芯尺寸或改变反应堆内的材料成分, 使  $k_1$  恰好等于 1, 则其余  $k_2, \dots, k_n (n > 1)$  的值都将小于 1. 这时式 (2.68) 中  $n = 1$  的项将与时间无关, 而  $n > 1$  的各项将随时间而衰减. 因而当时间足够长时,  $n > 1$  各项都已衰减到零, 系统达到稳态, 这时中子通量密度按基波形式 ( $B = B_1$ ) 分布, 反应堆处于临界状态. 从上面的讨论, 得到以下两个重要结果.

(1) 裸堆单群近似的“临界条件”为

$$k_1 = \frac{k_\infty}{1 + L^2 B^2} = 1 \quad (2.72)$$

其中,  $B^2$  是波动方程 (2.56) 的最小特征值  $B_1^2$ , 通常把它记为  $B_g^2$ , 并称为几何曲率, 式 (2.72) 称为单群理论的临界方程. 显然,  $k_1$  便是前面定义的有效增殖系数.

(2) 当反应堆处于临界状态时, 中子通量密度按最小特征值  $B_g^2$  所对应的基波特征函数分布, 也就是说稳态反应堆的中子通量密度空间分布满足波动方程

$$\nabla^2 \phi(\mathbf{r}) + B_g^2 \phi(\mathbf{r}) = 0 \quad (2.73)$$

这两点结论是非常重要的, 它们回答了反应堆临界时, 其材料组成、几何形状及大小之间必须怎样匹配? 还告诉一个临界反应堆内中子通量密度的分布.

下面仍以一维平板反应堆为例来加以具体讨论, 由式 (2.61) 及式 (2.72) 可得出无限平板反应堆的临界条件为

$$k_1 = \frac{k_\infty}{1 + L^2 \left(\frac{\pi}{a}\right)^2} = 1 \quad (2.74)$$

显然, 若系统的材料组成给定 (即  $k_\infty$ 、 $L^2$  给定), 则只有一个唯一的尺寸  $a_0$  能使  $k_1 = 1$ ,  $a_0$  即为反应堆的临界大小. 当反应堆尺寸  $a > a_0$  时, 则  $k_1 > 1$ , 反应堆处于超临界状态; 反之, 若  $a < a_0$ , 则处于次临界状态.

另一方面, 若反应堆的尺寸  $a$  给定, 则从式 (2.74) 可看出, 必然可以找到一种燃料富集度 (材料组成), 使得由其所确定的  $k_\infty$  及  $L^2$  值能使式 (2.74) 成立, 即  $k_1 = 1$ , 从而反应堆达到临界状态. 临界时, 反应堆内的中子通量密度分布为

$$\phi(x) = A \cos \frac{\pi}{a} x \quad (2.75)$$

现在来讨论一下式 (2.72) 的物理意义. 将证明式 (2.72) 中  $1/(1 + L^2 B_g^2)$  项的物理意义就是单群近似下反应堆内中子的不泄漏概率  $\Lambda$ . 在反应堆中单位时间、单位体积内的中子泄漏率等于  $-D\nabla^2 \phi$ , 根据式 (2.73) 有  $-D\nabla^2 \phi = DB_g^2 \phi$ , 而单位时间、单位体积内中子的吸收率为  $\Sigma_a \phi$ . 显然, 以单群理论观点来看, 反应堆内的中子不是从堆内泄漏出去, 就是在堆内被吸收, 因而有

$$\begin{aligned} \Lambda &= \frac{\text{中子吸收率}}{\text{中子吸收率} + \text{中子泄漏率}} \\ &= \frac{\Sigma_a \int_V \phi dV}{\Sigma_a \int_V \phi dV + DB_g^2 \int_V \phi dV} \\ &= \frac{1}{1 + L^2 B_g^2} \end{aligned} \quad (2.76)$$

这样, 式 (2.72) 可以写成

$$k_1 = k_\infty \Lambda = 1 \quad (2.77)$$

它与临界条件式 (2.5) 完全一样. 由此可见, 这里的  $k_1$  也就是前面所定义的反应堆有效增殖系数  $k$ . 同时, 由式 (2.65) 可知,  $l_1$  为考虑热中子泄漏影响后的中子寿命.

另外, 从式 (2.76) 可以看出, 反应堆的中子泄漏不仅与扩散长度有关, 而且与几何曲率有关. 从前面平板形反应堆的例子中可以看到, 当反应堆体积增大时,  $B_g^2$  就减小, 因而正如所预期的那样, 不泄漏概率也就增大. 此外, 扩散长度  $L$  愈大, 意味着中子自产生到被吸收所穿行的距离也愈大, 因而从反应堆中泄漏出去的概率也就愈大, 不泄漏概率  $P_L$  就要减小, 这和式 (2.76) 的结果是一致的. 下面以一个简单例子说明上述结果的应用.

**例 1** 设有如图 2.5 所示的一维石墨慢化反应堆,  $k_\infty = 1.06$ ,  $L^2 = 300 \text{ cm}^2$ ,  $\lambda_{tr} = 2.8 \text{ cm}$ . 试求:

(1) 达到临界时反应堆的厚度  $H$  和中子通量密度的分布;

(2) 设取  $H = 250 \text{ cm}$ , 试求反应堆的有效增殖系数  $k_{eff}$ .

解 (1) 根据式 (2.72) 的临界条件, 求得临界时反应堆的几何曲率  $B_g^2$  应为

$$B_g^2 = \frac{k_\infty - 1}{L^2} = \frac{1.06 - 1}{300} = 2.0 \times 10^{-4} \text{ cm}^{-2} \quad (2.78)$$

因而  $B_g = 0.01414 \text{ cm}^{-1}$ , 另一方面根据式 (2.61) 有  $B_g = \pi/a$ , 因而有

$$a = \frac{\pi}{B_g} = \frac{\pi}{0.01414} = 222.2 \text{ cm} \quad (2.79)$$

由于外推距离  $d = 0.7104\lambda_{tr} = 0.7104 \times 2.8 \approx 2 \text{ cm}$ , 因而求得临界时反应堆的厚度

$$H = a - 2d = 222.2 - 4 = 218.2 \text{ cm} \quad (2.80)$$

临界时中子通量密度分布为

$$\phi(x) = A \cos \frac{\pi}{a} x \quad (2.81)$$

(2) 若  $H = 250 \text{ cm}$ , 则反应堆的几何曲率

$$B_g^2 = \left( \frac{\pi}{H + 2d} \right)^2 = 1.530 \times 10^{-4} \text{ cm}^{-2} \quad (2.82)$$

反应堆的有效增殖系数为

$$k_{eff} = \frac{k_\infty}{1 + L^2 B_g^2} = \frac{1.06}{1 + 300 \times 1.53 \times 10^{-4}} = 1.0135 \quad (2.83)$$

从前面讨论知道, 平板形均匀裸堆内临界时的中子通量密度分布由方程 (2.75) 确定, 其分布形状只取决于反应堆的几何形状, 而与反应堆的功率大小无关. 这些式中都包含有待定的常数  $C$ , 也就是并未确定出反应堆内确切的中子通量密度大小. 系数  $C$  则需由反应堆的功率大小来确定. 这是为什么呢? 这可以从波动方程 (2.73) 找到原因, 因为该方程是齐次方程, 所以, 若  $\phi(\mathbf{r})$  是方程的一个解, 则  $\phi(\mathbf{r})$  的任意倍数都是方程的解, 这样纯粹通过求解方程, 自然只能给出临界反应堆内中子通量密度的形状, 却不能给出通量密度的大小. 实际上, 这蕴含着反应堆理论一个十分重要的结论, 那就是, 临界反应堆内中子通量密度的基波特征函数分布可以在任意功率水平下得到稳定. 若不考虑各种工程因素的限制, 从物理原理上讲, 反应堆功率可以是任意大小, 一个临界反应堆所能发出的功率并无任何限制.

反过来, 如已知反应堆的功率, 那么就可以确定出反应堆内中子通量密度的确切数值. 设已知反应堆的功率为  $P$ , 芯部的体积为  $V$ , 每次裂变产生的能量为  $E_f$ , 则整个堆芯的总功率为

$$P = E_f \int_V \Sigma_f \phi(\mathbf{r}) dV \quad (2.84)$$

把有关反应堆内中子通量密度的表达式 (2.75) 代入上式, 便可确定出常数  $C$  .

### 2.5.3 反应堆曲率和临界计算任务

从前面的讨论知道, 稳态反应堆内中子通量密度的空间分布满足波动方程, 即

$$\nabla^2 \phi(\mathbf{r}) + B_g^2 \phi(\mathbf{r}) = 0 \quad (2.85)$$

这里,  $B_g^2$  为方程的最小特征值, 即几何曲率. 对于裸堆来讲, 几何曲率只与反应堆的几何形状和尺寸大小有关, 例如对于平板形裸堆,  $B_g^2 = \left(\frac{\pi}{a}\right)^2$ , 它与反应堆的材料成分和性质没有关系. 对于处在临界状态的反应堆, 它的几何曲率应能满足临界方程 (2.72).

另一方面, 由于  $k_\infty$ 、 $L^2$  等参数都仅仅取决于反应堆芯部材料特性, 因而, 对于一定材料成分 (即给定  $k_\infty$ 、 $L^2$  等值) 的反应堆, 便有一个确定的  $B^2$  值能满足临界方程 (2.72). 为了计算和讨论方便, 把它叫作材料曲率, 它是满足临界方程 (2.72) 的  $B^2$  值, 并记作  $B_m^2$ . 因而, 对于单群扩散理论, 材料曲率  $B_m^2$  为

$$B_m^2 = \frac{k_\infty - 1}{L^2} \quad (2.86)$$

显然, 材料曲率  $B_m^2$  反映增殖介质材料的特性, 它的数值只取决于反应堆的材料成分和特性 (如  $L^2$ 、 $\tau$ 、 $k_\infty$  等), 而与反应堆的几何形状及大小无关. 显然, 对于任意给定材料成分和几何尺寸的反应堆, 几何曲率不一定等于材料曲率.

因而引进了材料曲率的概念之后, 反应堆的临界条件可以描写如下: 反应堆达到临界的条件是材料曲率等于几何曲率, 即

$$B_m^2 = B_g^2 \quad (2.87)$$

在一般情况下, 对于给定尺寸的反应堆, 其几何曲率  $B_g^2$  不一定等于材料曲率  $B_m^2$ . 若  $B_g^2 < B_m^2$ , 这时  $k$  将大于 1, 反应堆处于超临界状态; 反之, 若  $B_g^2 > B_m^2$ , 则反应堆处于次临界状态. 最后, 由上面的讨论可清楚地看到, 反应堆临界问题的计算可以归纳为下面两类问题.

第一类问题: 给定反应堆材料成分, 确定它的临界尺寸. 当反应堆的材料成分给定以后,  $k_\infty$ 、 $L^2$  等物理参数便可以很容易地计算出来. 因而, 材料曲率就成为已知的量, 问题便成了求临界尺寸问题. 临界时,  $B_m^2 = B_g^2$ , 可以根据  $B_g^2$  与尺寸的关系式计算出临界尺寸来.

第二类问题: 给定了反应堆的形状和尺寸, 确定临界时反应堆的材料成分. 这时, 事先给定反应堆的几何形状和尺寸 (例如, 由热工计算给出), 而要求临界时所需的材料成分 (一般是燃料的富集度). 事实上这是第一种情况的逆问题, 因为这时几何曲率  $B_g^2$  是已知的, 所需要的是改变反应堆的燃料—慢化剂成分和比例, 以使得临界方程 (2.72) 得到满足.

在具体计算中, 有时还会遇到这样的情况, 即不仅反应堆的材料成分 (如燃料—慢化剂) 已经给定, 而且反应堆的几何尺寸亦已由工程或热工的需要所确定. 这时要求的是反应堆的有效增殖系数

$$k_{\text{eff}} = \frac{k_\infty}{1 + L^2 B_g^2} \quad (2.88)$$

或

$$\rho = \frac{k - 1}{k} \quad (2.89)$$

其中,  $\rho$  通常称为反应性. 对于临界反应堆,  $\rho = 0$ ; 若  $\rho > 0$ , 则反应堆处于超临界状态; 若  $\rho < 0$ , 则反应堆处于次临界状态. 因此  $|\rho|$  的大小表示反应堆偏离临界状态的程度.

#### 2.5.4 单群理论的修正

前面提到, 单群理论是一种近似的方法. 计算表明, 对于热中子反应堆, 直接应用式 (2.72) 或式 (2.87) 进行计算, 将带来比较大的误差. 但是, 计算若用  $M^2 = L^2 + \tau$  来替换上述式中的  $L^2$ , 则可以改善计算结果, 使其更令人满意. 这样, 临界条件和材料曲率便改写成下列形式:

$$\begin{aligned} k_{\text{eff}} &= \frac{k_{\infty}}{1 + M^2 B_g^2} = 1 \\ B_m^2 &= \frac{k_{\infty} - 1}{M^2} \end{aligned} \quad (2.90)$$

这就是所谓的热中子反应堆的修正单群理论. 修正单群理论之所以能改善计算结果, 其主要原因可以从物理上解释如下: 在单群理论中, 把所有中子都看成是热中子, 因而该理论没有考虑慢化过程对中子泄漏的影响. 知道,  $L^2$  与中子从变成热中子地点开始到它被吸收为止所移动过的距离有关,  $1/(1 + L^2 B_g^2)$  为热中子扩散过程中的不泄漏概率. 但是, 在反应堆内, 实际上中子由裂变中子慢化成为热中子之前, 已经移动过了一个距离  $\tau$ , 而徙动面积  $M^2 = L^2 + \tau$  跟中子由核裂变产生直到它被吸收所穿行距离的均方值有关, 故若用徙动面积  $M^2 = L^2 + \tau$  来代替  $L^2$ , 便可初步地考虑慢化过程对泄漏的影响, 因而使计算精度得到改善. 实际计算结果也证明了这一结论. 因此, 以后提到单群理论一般都是指应用修正的单群理论公式, 它既简单又具有较好的精度.

#### 2.5.5 稳态单群扩散方程

当反应堆在稳态 (临界) 时, 由式 (2.45) 得的中子扩散方程为

$$D\nabla^2\phi(\mathbf{r}) - \Sigma_a\phi(\mathbf{r}) + k_{\infty}\Sigma_a\phi(\mathbf{r}) = 0 \quad (2.91)$$

显然, 它只对于临界系统才成立. 由反应堆临界理论, 对于给定的参数  $D$ ,  $\Sigma_a$  和  $\Sigma_f$  (或  $k_{\infty}$ ), 只有在一定几何形状和尺寸的情况下, 系统才能达到稳态, 方程 (2.91) 才能成立, 并且有解. 一般来说, 对于同时任意给定的几何形状大小和材料特性的系统, 它不一定处于稳态, 方程 (2.91) 不一定都成立. 但是, 从物理上, 可以通过引入一个特征参数  $k$  来进行调整使其达到临界, 即人为地把  $k_{\infty}$  或每次裂变的中子产额  $\nu$  除以参数  $k$ . 根据具体情况, 改变  $k$  值的大小便可以使系统达到临界. 因为若原来的问题是次临界的, 则将  $k_{\infty}$  除以一个小于 1 的正数  $k$ , 亦即人为地提高裂变产生的中子数. 改变  $k$  值大小必然可以使得系统恰好到达临界. 反之, 若系统是超临界的, 则应除以大于 1 的  $k$  值, 也可以使其达到临界. 当系统临界时,  $k = 1$ . 这样, 对于任意系统, 都可以写出它的稳态单群扩散方程如下

$$D\nabla^2\phi(\mathbf{r}) - \Sigma_a\phi(\mathbf{r}) + \frac{k_{\infty}}{k}\Sigma_a\phi(\mathbf{r}) = 0 \quad (2.92)$$

或者写成

$$\nabla^2\phi(\mathbf{r}) + B^2\phi(\mathbf{r}) = 0 \quad (2.93)$$

其中<sup>2)</sup>

$$B^2 = \frac{k_\infty/k - 1}{L^2} \quad (2.94)$$

其中,  $L^2$  是扩散长度. 从物理上可以看到, 这里引进的  $k$  就是系统的有效增殖系数  $k_{\text{eff}}$ . 这可以证明如下: 把式 (2.92) 对整个体积积分, 注意到

$$-\int_V D \nabla^2 \phi(\mathbf{r}) dV = -\int_S D \text{grad } \phi(\mathbf{r}) \cdot d\mathbf{S} = \int_S \mathbf{J} \cdot \mathbf{n} ds \quad (2.95)$$

它就是单位时间内从整个表面泄漏出去的中子数, 这样由式 (2.92) 有

$$\begin{aligned} k_{\text{eff}} &= \frac{\int_V \nu \Sigma_f \phi(\mathbf{r}) dV}{-\int_V D \nabla^2 \phi(\mathbf{r}) dV + \int_V \Sigma_a \phi(\mathbf{r}) dV} \\ &= \frac{\text{新一代中子产生率}}{\text{中子总消失率 (泄漏率 + 被吸收率)}} \end{aligned} \quad (2.96)$$

这正是前面式 (2.2) 所定义的有效增殖系数.

### 3 学习笔记: 基于 PINN 深度机器学习技术求解多维中子学扩散方程

在前面的学习中, 已经掌握了有关物理反应神经网络 (PINN) 和核反应堆物理的知识. 现在, 将深入学习刘东老师的论文, 该论文使用 PINN 深度学习技术来解决多维中子学扩散方程的问题. 通过对论文内容的分析, 将理解深度学习在核反应堆物理中的具体应用, 包括模型构建、关键技术、数值验证等方面.

#### 3.1 理论原理

##### 3.1.1 中子学扩散方程模型

核反应堆工程堆芯中子学设计计算中采用中子扩散模型的方程形式如式 (3.1) 所示:

$$\begin{aligned} \frac{1}{v} \frac{\partial \phi(r, E, t)}{\partial t} &= \nabla \cdot D \nabla \phi(r, E, t) - \Sigma_t(r, E) \phi(r, E, t) \\ &+ x(E) \int_0^\infty \vartheta(E') \Sigma_f(r, E') \phi(r, E', t) dE' \\ &+ S(r, E, t) + \int_0^\infty \Sigma_s(r, E' \rightarrow E) \phi(r, E', t) dE' \end{aligned} \quad (3.1)$$

方程 (3.1) 是核反应堆中描述中子行为的重要方程, 其各项的物理意义 (表 2) 对于理解中子守恒和核反应堆的性能至关重要.

---

2) 对于修正单群理论, 应为  $B^2 = \frac{k_\infty/k - 1}{M^2}$ .

方程项	物理含义
$\frac{1}{v} \frac{\partial \phi(r, E, t)}{\partial t}$	通量密度变化率
$\nabla \cdot D \nabla \phi(r, E, t)$	泄漏项
$-\Sigma_t(r, E) \phi(r, E, t)$	吸收项
$x(E) \int_0^\infty \vartheta(E') \Sigma_f(r, E') \phi(r, E', t) dE'$	裂变项
$S(r, E, t)$	源项
$\int_0^\infty \Sigma_s(r, E' \rightarrow E) \phi(r, E', t) dE'$	散射项

表 2 多群中子学扩散方程模型各项解释

式中,  $\phi(r, E, t)$  是  $t$  时刻在  $r$  位置处能量为  $E$  的中子注量率<sup>3)</sup>;  $v$  为中子速率;  $D$  为中子扩散系数;  $\vartheta$  为每次裂变放出的平均中子数;  $x(E)$  是裂变中子的能谱分布;  $\Sigma_t$   $\Sigma_s$  和  $\Sigma_f$  分别是总的反应截面、散射截面和裂变截面;  $S(r, E, t)$  为外源。

其单能群存在独立的外中子源  $S_0(\mathbf{r}, t)$  的扩散方程<sup>4)</sup>就可以写成

$$\frac{1}{Dv} \frac{\partial \phi(\mathbf{r}, t)}{\partial t} = \nabla^2 \phi(\mathbf{r}, t) + \frac{k_\infty - 1}{L^2} \phi(\mathbf{r}, t) + \frac{1}{D} S_0(\mathbf{r}, t) \quad (3.2)$$

其单能群无外源扩散方程<sup>5)</sup>形式可写为:

$$\frac{1}{Dv} \frac{\partial \phi(r, t)}{\partial t} = \nabla^2 \phi(r, t) + \frac{k_\infty - 1}{L^2} \phi(r, t) \quad (3.3)$$

式中,  $L$  为扩散长度;  $k_\infty$  为无限介质增殖系数。

当系统处于稳态<sup>6)</sup>时, 式 (3.3) 形式为:

$$\nabla^2 \phi(r) + \frac{k_\infty/k_{\text{eff}} - 1}{L^2} \phi(r) = 0 \quad (3.4)$$

当系统处于临界状态 ( $k_{\text{eff}} = 1$ ) 时, 式 (3) 的形式为:

$$\nabla^2 \phi(r) + B_g^2 \phi(r) = 0 \quad (3.5)$$

式中,  $B_g^2$  为系统临界时的几何曲率, 与系统的几何特性相关, 临界时等于材料曲率。

式 (3.1) ~ 式 (3.4) 可描述为统一的两阶偏微分方程形式:

$$F[\vec{x}, \phi(\vec{x}), \nabla \phi(\vec{x}), \nabla^2 \phi(\vec{x})] = 0 \quad (3.6)$$

式中,  $\vec{x}$  为扩散方程几何描述空间向量。

3) 在研究辐射和与能量沉积有关的问题时, 相应通量密度的总照射量常用注量 (fluence) 来度量。根据国际辐射单位和测量委员会定义, 在空间  $r$  处单位时间内进入以该点为中心的单位横截面的小球体内的中子数称为该点的中子注量率 (neutron fluence rate), 一般以  $\phi$  表示, 中子 / ( $\text{m}^2 \cdot \text{s}$ )。因而  $\Delta t$  时间内的注量  $F(\mathbf{r})$  为

$$F(\mathbf{r}) = \int_{t_1}^{t_1 + \Delta t} \phi(\mathbf{r}, t) dt$$

根据中子通量密度定义, 显然中子注量率就等于中子通量密度。在我国国家标准 (GB/T 3102.10-1993 和 GB/T 4960.2-1996) 中, 这两个名词都是允许使用的。

4) 推导细节见第2.5节式(2.48)

5) 推导细节见第7节式(2.52)

6) 推导细节见第2.5.5节式(2.93)

### 3.1.2 利用 PINN 求解扩散方程的关键技术

利用 PINN 基本模型 (PINN 网络结构见图 1.1) 求解微分方程具有一定的普适性, 但将其直接应用于中子学扩散方程, 精度与收敛性上会出现一定的困难, 需要针对扩散方程的特点进行改进与完善。

### 3.1.3 扩散方程特征向量加速收敛方法

当扩散方程是稳态临界状态时, 扩散方程为式 (3.4)、式 (3.5) 形式, 这在数理方法上属于特征值方程。对同一特征值来说, 特征向量是无穷多个<sup>7)</sup>, 对应于其中的任一  $B_n$  值, 都可以给出满足微分方程 (3.3) 及边界条件 (2.51) 的解

$$\varphi_n(x) = A_n \cos B_n x = A_n \cos \frac{(2n-1)\pi}{a} x \quad (3.7)$$

从上面的求解过程可以得出结论, 齐次方程 (3.3) 只对某些特定的特征值  $B_n$  才有解。相应的解  $\varphi_n(x)$  称为此问题的特征函数。

实验证明, 这种特征值方程解的不确定性会对神经网络收敛过程造成大幅度振荡, 收敛时间会大幅度增加。同时, 神经网络初始值的不同会造成收敛特征向量的不同, 而且, 在同样的初始值下, 因为学习率、批处理值的差异等因素, 收敛的特征向量也可能不同, 相同的初始值难以重现计算结果。

一种改进的方法是将扩散方程的特殊空间点、多个区域平均值/积分值与预设值的均方差作为的额外加权因素, 采用与初始条件、边界条件相似的处理方法进行加权处理, 将大幅度提高机器学习收敛效率, 即便采用不同的神经网络初始值参数也会收敛到相近的结果。

特定空间点/区域理论上可选择任意注量率不为零的点/区域, 实践中推荐选取求解几何域的中心处, 或者具有相同注量率对称分布的不同几何点/区域, 预设值可设定为任意固定的归一化注量率值。

此外, 采用已经收敛的神经网络作为类似问题的神经网络初始值, 将很大程度加速收敛过程, 这是由于深度神经网络具备一定的迁移学习特性所导致的。

### 3.1.4 $k_{\text{eff}}$ 并行搜索方法

稳态方程式 (3.4) 需要引入  $k_{\text{eff}}$ , 并进行搜索达到临界。由于式 (3.4) 是特征值方程, 常用的是源迭代方法。基于 PINN 模型的深度学习方程求解方法, 可以直接作为其内迭代环节进行  $k_{\text{eff}}$  搜索计算。但由于内迭代次数较多, 若每一次内迭代均需要完整的机器学习, 迭代计算量比较大,  $k_{\text{eff}}$  的搜索效率很低。

本文采用式 (3.3) 进行  $k_{\text{eff}}$  的搜索, 将式 (3.4) 写成式 (3.3) 的形式:

$$\frac{1}{Dv} \frac{\partial \phi(r, t)}{\partial t} = \nabla^2 \phi(r, t) + \frac{k_{\infty}/k - 1}{L^2} \phi(r, t) \quad (3.8)$$

给出  $k$  与  $\phi(r, t_0)$  的任意初始值, 考察经过一定时间  $t_t$ , 即  $t > t_{\tau}$  时,  $\phi(r, t)$  对  $t$  的偏导数是否接近 0, 从而进行临界判断。可设定合适的  $\varepsilon$ , 若  $\partial \phi(r, t)/\partial t > \varepsilon$ , 判断为超临界状态, 需要增大  $k$  值;

<sup>7)</sup> 细节说明见第 3 节式 (3.7)



若  $\partial\phi(r,t)/\partial t < -\varepsilon$ , 为次临界状态, 需要减小  $k$  值。当  $-\varepsilon < \partial\phi(r,t)/\partial t < \varepsilon$ , 则认为系统已达临界, 此时  $k = k_{\text{eff}}$ ,  $\phi(r, t > t_r)$  即为稳态时系统的  $\phi(r)$  分布 (式 (3.4))。

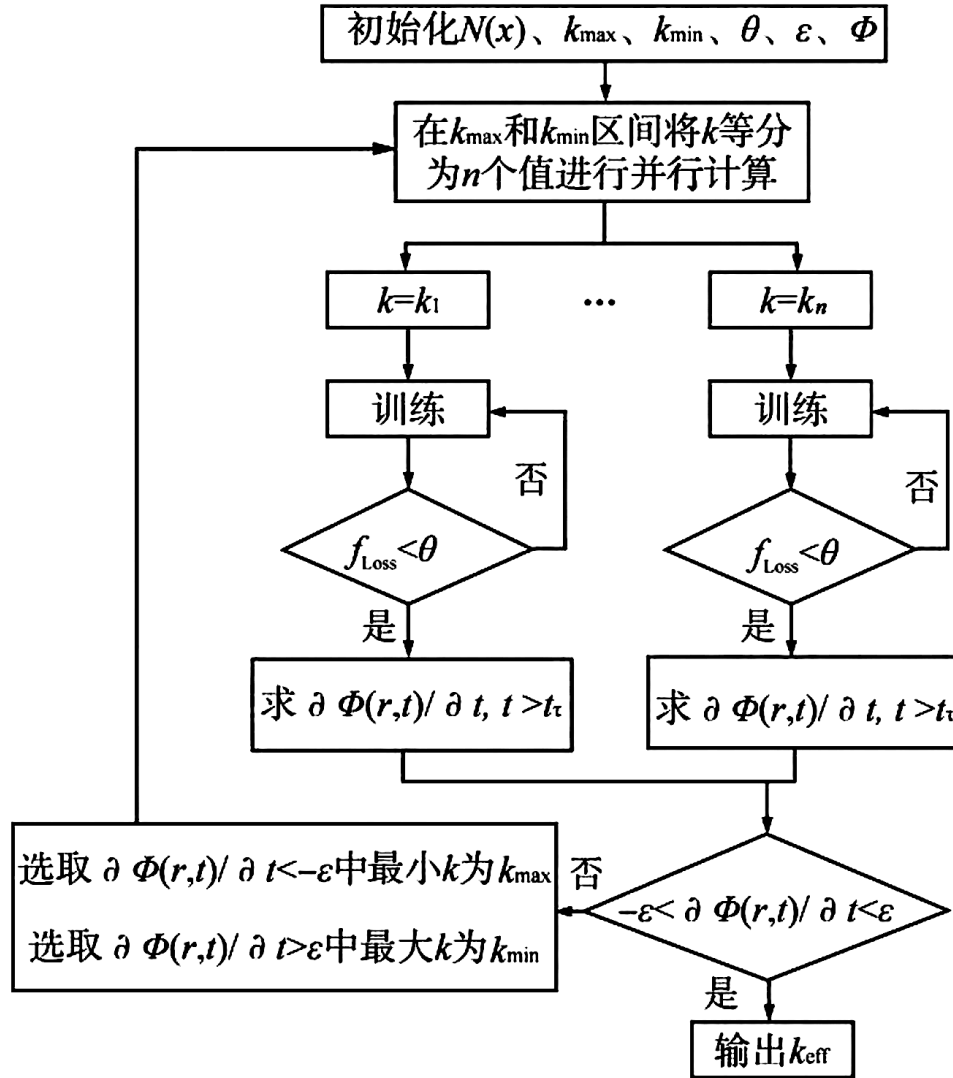


图 3.1  $k_{\text{eff}}$  并行搜索流程图

Figure 3.1 Flow chart of  $k_{\text{eff}}$  parallel search

$t_r$  的取值应大于注量率从任意初始状态到达“稳定”变化状态的时间, 根据扩散方程中反应堆类型 [表现为式 (3.8) 中的  $Dv$ ] 与初始边界条件形式会有所不同;  $k$  值的搜索方法可采用二分法或者将  $k$  等分成一定数量后, 大规模并行搜索, 可有效提高效率。 $k_{\text{eff}}$  的并行搜索流程如图 3.1 所示。

### 3.1.5 几何空间网格点密度不均匀分布策略

式 (3.1) ~ 式 (3.4) 中空间网格几何点  $i$  的分布与数量, 对模型精度与机器学习效率的影响很大。通常来说, 网格点的数量越多, 计算精度越高, 但计算量越大, 收敛效率降低。同时, 初始条件、边

界条件及  $\phi(\vec{x})$  变化很大的地方网格点密度对模型的精度与收敛效率影响很大。

为了平衡机器学习训练精度与效率, 推荐的策略是在相同网格点数量情况下,  $\phi(\vec{x})$  变化大的区域网格点相对紧密, 变化平滑的区域相对稀疏, 以  $\phi(\vec{x})$  在几何空间区域网格点偏导数的变化值作为网格点密度设定的判别依据。

据此, 在方程的边界条件和初始条件附近, 或者系统  $\phi(\vec{x})$  “畸性”比较大处 (即区域偏导较大处), 应加密网格点提高精度, 而在变化平缓处 (即几何区域点偏导较小处), 应减小网格密度以提高收敛效率。

### 3.2 数值验证程序复现

为了便于计算和讨论, 本文选择了结构相对简单的平板几何形式的扩散方程进行数值验证。这样的选择使我们能够更清晰地理解和分析神经网络在特定几何条件下的性能, 为进一步的讨论提供基础。

#### 3.2.1 深度机器学习的超参数设置

为了验证计算结果, 我们选择具有解析解的特定几何形式的扩散方程进行数值验证。这些验证结果不仅对当前方程形式有所验证, 同时也为其他形式的方程和几何形状提供了参考。以下是用到的神经网络的具体超参数设置, 如表 3 所示:

超参数		设置
网络结构	$\mathcal{N}(\vec{x})$	全连接神经网络
深度	$l$	16
中间层神经元数量	$s$	20
激活函数	$\sigma$	双曲正切函数 $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$
损失函数	$Loss$	$Loss : MSE_u + MSE_{\{R, BC, IC\}}$
边界权重	$P_b$	100
学习率	$lr$	0.001
优化器		Adam
网络初始化方法		高斯分布随机采样
训练停止条件		达到设定的训练次数

表 3 神经网络超参数及相关设置

#### 3.2.2 临界条件下稳态扩散方程的验证

选取平板几何结构的扩散方程进行验证, 该几何结构的解析解为  $C \cdot \cos(x \cdot \pi/a)$ , 其中  $C = 0.5$  为常数,  $a = 1$  为平板的长度。我们验证计算神经网络的超参数设置见表 3。这部分将逐步讲解上述

描述的泊松方程求解器的实现。作为示例，此处将使用 Python 语言和深度学习框架 DeepXDE<sup>8)</sup> 进行求解。具体代码如下所示：

首先，导入 DeepXDE 和 numpy 等必要模块

```
1 import deepxde as dde
2 import numpy as np
3 import os
4 import matplotlib.pyplot as plt
```

初始化相关参数

```
1 # 初始化参数
2 k_eff = 1 # 有效增殖系数
3 a = 1 # 平板的宽度
4 B2 = (np.pi / a) ** 2 # 系统临界时的几何曲率
5 l = 16 # 神经网络的深度
6 s = 20 # 神经网络的中间层隐藏神经元数量
7 Pb = 100 # 边界权重
8 C = 0.5 # 解析解参数
```

定义解析解  $C \cdot \cos(x \cdot \pi/a)$

```
1 # 定义解析解
2 def phi_analytical(x):
3     return C * np.cos(x * np.pi / a)
```

使用内置的 Interval 类定义一个几何求解区域  $[-a/2, a/2]$ ，如下所示：

```
1 # 定义几何网格
2 geom = dde.geometry.Interval(-a / 2, a / 2)
```

接下来，表达扩散方程(3.5)的 PDE 残差函数  $\nabla^2 \phi(r) + B_g^2 \phi(r)$ ，如下所示：

```
1 # 定义微分方程
2 def pde(x, phi):
3     dphi_xx = dde.grad.hessian(phi, x, i=0, j=0)
4     return dphi_xx + B2 * phi
```

pde 的第一个参数是网络输入，即横坐标。第二个参数是网络输出，即解，但这里我们使用  $y$  作为变量的名称。

接下来，考虑平板反应堆的外推边界条件 (2.51)，即  $\phi(\pm \frac{a}{2}) = 0$ 。我们使用一个简单的 Python 函数，返回一个布尔值，来定义平板反应堆的外推边界条件的子域 ( $\{-a/2, a/2\}$ )。该函数应该对子域内的点返回 **True**，对于外部的点返回 **False**。在我们的情况下，平板反应堆的外推边界条件的点  $x$  是  $-a/2$  和  $a/2$ 。

```
1 # 定义边界条件
2 bc = dde.icbc.DirichletBC(geom, lambda x: 0, lambda _, on_boundary: on_boundary)
```

8) DeepXDE 是一个用于求解微分方程的 Python 包，其主要特点使用深度学习技术求解微分方程，其官方网站为 <https://deepxde.readthedocs.io>。

现在,我们已经指定了几何形状、扩散方程的 PDE 残差和外推边界条件。然后,我们定义 PDE 问题:

```
1 # 定义数据
2 data = dde.data.PDE(geom, pde, bc, num_domain=898, num_boundary=2, solution=phi_analytical
    , num_test=100)
```

数字 `num_domain=898` 是在域内采样的训练残差点数量,数字 `num_boundary=2` 是在边界上采样的训练点的数量。参数 `solution=phi_analytical` 是用于计算我们解的误差的参考解,如果我们没有参考解,可以忽略它。我们使用 `num_test=100` 个残差点来测试 PDE 残差。

接下来,我们选择网络。在这里,我们使用深度为  $l = 16$  (即 16 个隐藏层)宽度为  $s = 20$  的全连接神经网络,其网络初始值权重采用高斯分布随机采样,并选用 `tanh` 函数作为激活函数:

```
1 # 定义神经网络
2 layer_size = [1] + [s] * l + [1]
3 activation = "tanh"
4 # 网络初始值权重采用高斯分布随机采样
5 initializer = "Glorot uniform"
6 net = dde.nn.PFNN(layer_size, activation, initializer)
```

现在,有了 PDE 问题和网络。下面构建一个 Model,并选择 Adam 作为优化器和初始学习率 `lr=0.001`:

```
1 # 定义模型
2 model = dde.Model(data, net)
3 # 定义求解器
4 model.compile("adam", lr=0.001, metrics=["l2 relative error"], loss_weights=[1, Pb])
```

这里通过 `loss_weights=[1, Pb]` 将边界处的损失权重调整为 100,同时还选用 `metrics=["l2 relative error"]` 计算在测试点上训练好的神经网络输出与对应解析解的  $L_2$  相对误差作为训练过程中的一个指标。然后,对模型进行 3500 次迭代训练,并使用 `dde.saveplot` 回调函数在训练期间保存模型和图片,最后输出在  $x = 0$  处的值 (即 C),以此来显示特征值方程解 (特征向量) 的不确定性。

```
1 # 训练模型
2 losshistory, train_state = model.train(epochs=3500)
3 # 保存和可视化训练结果
4 dde.saveplot(losshistory, train_state, issave=True, isplot=True)
5
6 # 输出在 x=0 处的值 (即 C)
7 print("Predicted value at x=0:", model.predict(np.array([0])))
```

为了深入了解模型的训练过程、监控关键指标的变化,并直观地呈现训练效果,我们采用了以下可视化策略:

```
1 # 可视化
2 # 创建保存图像的文件夹
3 output_folder = "figure/算例1"
4 os.makedirs(output_folder, exist_ok=True)
5
6 # 设置中文字体
7
```

```

8     # 替换为您的字体文件路径
9     font_path = '/System/Library/Fonts/STHeiti Light.ttc'
10
11     # 添加字体路径
12     font_properties = FontProperties(fname=font_path)
13     plt.rcParams['font.sans-serif'] = [font_properties.get_name()]
14     plt.rcParams['axes.unicode_minus'] = False
15
16     # 绘制损失函数变化图
17     loss_train = np.sum(losshistory.loss_train, axis=1)
18     loss_test = np.sum(losshistory.loss_test, axis=1)
19
20     plt.figure(figsize=(6, 5)) # 设置图像大小
21     plt.semilogy(losshistory.steps, loss_train, label="训练损失")
22     plt.semilogy(losshistory.steps, loss_test, label="测试损失")
23     for i in range(len(losshistory.metrics_test[0])):
24         plt.semilogy(
25             losshistory.steps,
26             np.array(losshistory.metrics_test)[: , i],
27             label="测试指标",
28         )
29     plt.xlabel("步骤")
30     plt.legend()
31     # 保存为png
32     plt.tight_layout()
33     plt.savefig(os.path.join(output_folder, "losshistory.png"), format="png", dpi=200)
34
35
36     # 绘制训练状态变化图
37     def _pack_data(train_state):
38         def merge_values(values):
39             if values is None:
40                 return None
41             return np.hstack(values) if isinstance(values, (list, tuple)) else values
42
43         y_train = merge_values(train_state.y_train)
44         y_test = merge_values(train_state.y_test)
45         best_y = merge_values(train_state.best_y)
46         best_ystd = merge_values(train_state.best_ystd)
47         return y_train, y_test, best_y, best_ystd
48
49
50     if isinstance(train_state.X_train, (list, tuple)):
51         print(
52             "错误：网络有多个输入，尚未实现绘制此类结果。"
53         )
54
55     y_train, y_test, best_y, best_ystd = _pack_data(train_state)
56     y_dim = best_y.shape[1]
57
58     # 回归分析图
59     if train_state.X_test.shape[1] == 1:

```

```

60     idx = np.argsort(train_state.X_test[:, 0])
61     X = train_state.X_test[idx, 0]
62     plt.figure(figsize=(6, 5)) # 设置图像大小
63     for i in range(y_dim):
64         if y_train is not None:
65             plt.plot(train_state.X_train[:, 0], y_train[:, i], "ok", label="训练点分布")
66         if y_test is not None:
67             plt.plot(X, y_test[idx, i], "-k", label="真实")
68             plt.plot(X, best_y[idx, i], "--r", label="预测")
69             if best_ystd is not None:
70                 plt.plot(
71                     X, best_y[idx, i] + 2 * best_ystd[idx, i], "-b", label="95% 置信区间"
72                 )
73                 plt.plot(X, best_y[idx, i] - 2 * best_ystd[idx, i], "-b")
74             plt.xlabel("x")
75             plt.ylabel("y")
76             plt.legend()
77             plt.savefig(os.path.join(output_folder, "solution_state.png"), format="png", dpi=200)
78 # 神经网络输出的残差
79 if y_test is not None:
80     plt.figure(figsize=(6, 5)) # 设置图像大小
81     residual = y_test[:, 0] - best_y[:, 0]
82     plt.plot(best_y[:, 0], residual, "o", zorder=1)
83     plt.hlines(0, plt.xlim()[0], plt.xlim()[1], linestyle="dashed", zorder=2)
84     plt.xlabel("预测值")
85     plt.ylabel("残差 = 观测 - 预测")
86     plt.tight_layout()
87     plt.savefig(os.path.join(output_folder, "y_test.png"), format="png", dpi=200)
88 # 神经网络输出的标准差
89 if best_ystd is not None:
90     plt.figure(figsize=(6, 5)) # 设置图像大小
91     for i in range(y_dim):
92         plt.plot(train_state.X_test[:, 0], best_ystd[:, i], "-b")
93         plt.plot(
94             train_state.X_train[:, 0],
95             np.interp(
96                 train_state.X_train[:, 0], train_state.X_test[:, 0], best_ystd[:, i]
97             ),
98             "ok",
99         )
100     plt.xlabel("x")
101     plt.ylabel("std(y)")
102     plt.tight_layout()
103     plt.savefig(os.path.join(output_folder, "best_ystd.png"), format="png", dpi=200)
104
105 # 绘制数值解与解析解的图像并比较误差
106 x = np.linspace(-a / 2, a / 2, 100).reshape((-1, 1))
107 y_pred = model.predict(x)
108 y_exact = phi_analytical(x)
109
110 plt.figure(figsize=(12, 5))
111 plt.subplot(1, 2, 1)

```

```

112     plt.plot(x, y_pred, label="神经网络解", color="tab:blue")
113     plt.plot(x, y_exact, label="解析解", color="tab:orange")
114     plt.title("数值解与解析解的比较")
115     plt.xlabel("x")
116     plt.ylabel("$\phi(x)$")
117     plt.legend()
118
119     plt.subplot(1, 2, 2)
120     plt.plot(x, np.abs(y_pred - y_exact), label="绝对误差", color="tab:red")
121     plt.title("数值解与解析解的绝对误差")
122     plt.xlabel("x")
123     plt.ylabel("绝对误差")
124     plt.legend()
125
126     plt.tight_layout()
127     plt.savefig(os.path.join(output_folder, "solution_comparison.png"), format="png", dpi=200)
128
129     plt.show()

```

通过深度学习网络训练的结果可视化展示了在训练过程和最终效果。首先，我们关注损失函数变化图，如图 3.2 所示。图中横轴表示训练步骤，纵轴表示误差数量级，分别展示了训练损失和测试损失的变化情况。从图中可以清晰地观察到随着训练的进行，训练损失逐渐降低，而测试损失也呈现下降趋势，这表明网络在学习过程中取得了良好的效果。此外，图中还展示了测试阶段的指标（数值解与解析解比较）变化，图中显示，随着训练的进行，测试指标几乎保持稳定，且其值相对较大，这表明数值解与解析解始终保持较大误差。这为我们提供了对网络性能的更全面认识。

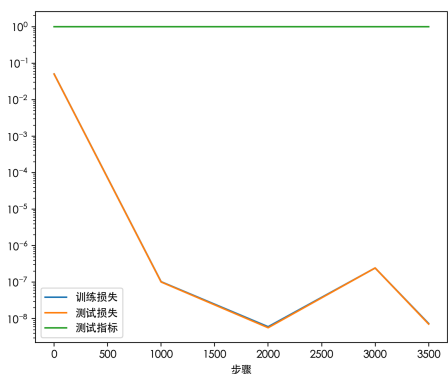


图 3.2 损失函数变化图

Figure 3.2 Loss function change

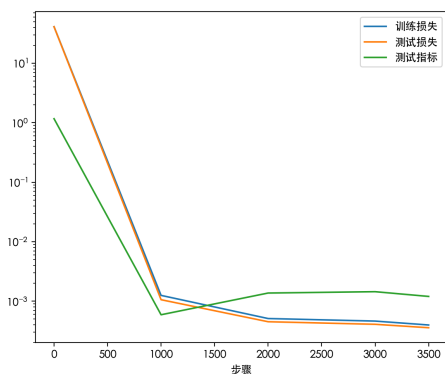


图 3.3 优化后损失函数变化图

Figure 3.3 Loss function change after optimization

接下来，我们关注回归分析图，图 3.4 展示了神经网络训练点的采样分布情况，真实数据的曲线以及网络的预测曲线。从图中可以看出，网络预测曲线与真实曲线趋势相差很大。

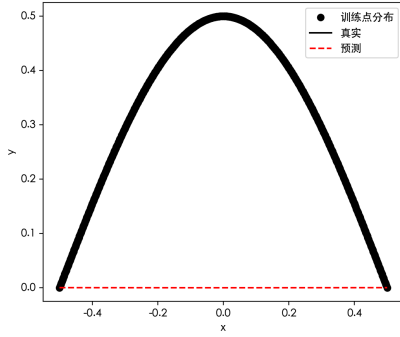


图 3.4 回归分析图  
Figure 3.4 Regression analysis

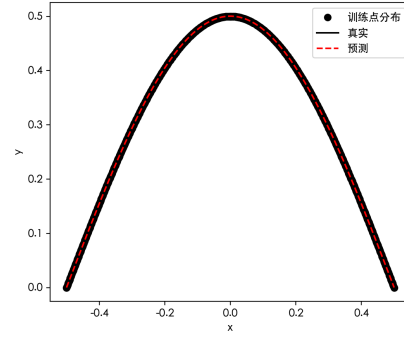


图 3.5 优化后回归分析图  
Figure 3.5 Regression analysis after optimization

另外，我们还通过图 3.6 展示了针对测试数据的残差分析。从图中可以看出，残差的分布情况较为随机，且残差的绝对值较大，这表明网络的预测效果并不理想。这些现象归因于由于特征值方程的特征向量不确定性，收敛得到的特征向量是不同的（表现为收敛时  $C$  值不同），且不易收敛。同时可以发现，通过神经网络的输出值几乎全为 0，这表明神经网络从众多解中选择一个近似为 0 的解，这复合神经网络优先偏好学习低频解的特性。

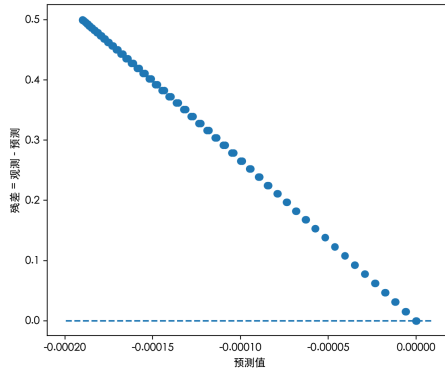


图 3.6 残差分析图  
Figure 3.6 Regression analysis

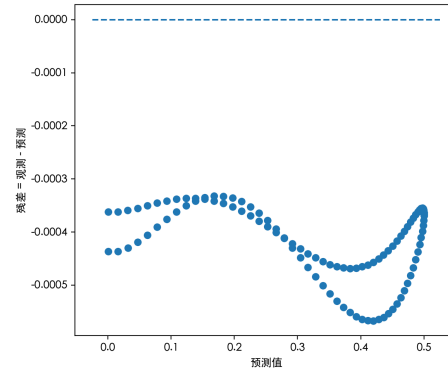


图 3.7 优化后残差分析图  
Figure 3.7 Regression analysis after optimization

为了解决此问题，我们采用 3.1.3 节中提到的改进方法：将扩散方程的特殊空间点、多个区域平均值/积分值与预设值的均方差作为的额外加权因素，采用与初始条件、边界条件相似的处理方法进行加权处理。代码段如下所示：

```
1 # 扩散方程特征向量加速收敛方法验证
2 observe_x = np.array([0])
3 observe_phi0 = dde.icbc.PointSetBC(observe_x, phi_analytical(observe_x))
4 # 定义数据
5 data = dde.data.PDE(geom, pde, [bc, observe_phi0], num_domain=898, num_boundary=2, anchors=
    observe_x, solution=phi_analytical, num_test=100)
```



首先，我们关注添加了额外加权因素损失函数变化图，如图 3.3 所示，从图中可以清晰地观察到随着训练的进行，三种损失均呈现下降趋势，这表明网络在学习过程中取得了良好的效果。接下来，我们关注回归分析图 3.5 以及测试数据的残差分析图 3.7，从图中可以看出，预测曲线与真实曲线趋势相符。

此外，我们注意到如图 3.7 所示预测曲线与真实曲线在边界处的残差很大。这直接使用经典 PINN 对边界条件的约束方法并不理想，于是考虑通过硬边界约束的方法来改进。具体做法是定义神经网络输出的转换

$$\phi(\vec{x}) = \vec{x} \cdot (\pi - \vec{x}) \cdot y + \vec{x} \quad (3.9)$$

并将其应用于网络。当  $\vec{x} = \pm \frac{a}{2}$  时，边界条件  $\phi(\pm \frac{a}{2}) = 0$  得到满足。这表明边界约束的两端都是硬条件。代码段如下所示：

```
1 # 定义硬边界条件
2 def output_transform(x, y):
3     return (x + a / 2) * (x - a / 2) * y
4 net.apply_output_transform(output_transform)
```

在这段代码中，我们定义了一个名为 `output_transform` 的函数，它对输出进行了转换。具体而言，对于给定的输入  $x$  和输出  $y$ ，转换后的输出是  $\vec{x} \cdot (\pi - \vec{x}) \cdot y + \vec{x}$ 。

然后，通过 `net.apply_output_transform(output_transform)` 将这个输出转换函数应用于网络。这表示我们通过这个转换函数考虑了边界条件的硬约束，确保网络的输出满足问题的物理边界条件。

最后，我们再次关注回归分析图 3.8 以及测试数据的残差分析图 3.9，从图中可以看出，预测曲线与真实曲线趋势相符，残差的绝对值较小，且在边界处误差为 0，这表明使用硬边界条件的网络的预测效果最好，且实验发现在本文的其他算例中存在类似的现象。鉴于此，我们将在后边的算例中使用硬边界条件的网络进行计算。

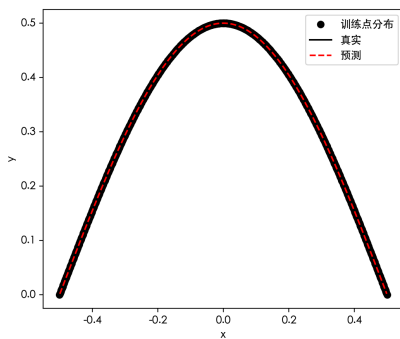


图 3.8 硬边界条件回归分析图  
Figure 3.8 Regression analysis with hard BC

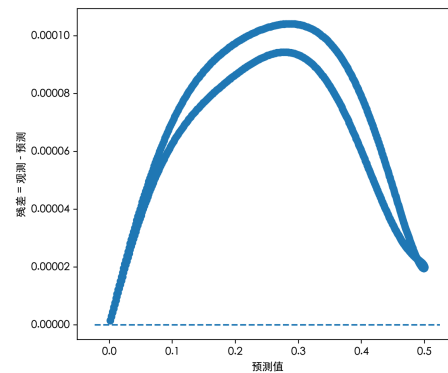


图 3.9 硬边界条件残差分析图  
Figure 3.9 Residual analysis with hard BC

### 3.2.3 扩散方程特征向量加速收敛方法验证

为了验证扩散方程特征向量加速收敛方法的有效性,我们选取了平板几何结构的扩散方程进行验证,该几何结构的解析解为  $C \cdot \cos(x \cdot \pi/a)$ , 其中  $C = 0.5$  为常数,  $a$  为平板的长度。我们验证计算神经网络的超参数设置见表 3。

验证计算的目标是: 统计并分析不同初始权重值  $\{\vec{w}, \vec{b}\}$  的神经网络  $\mathcal{N}(\vec{x})$ , 在训练到相似精度时, 所需要的收敛时间。

算例 1、算例 2 网络初始值权重  $\{\vec{w}, \vec{b}\}$  随机选择; 算例 3 也选择随机初始值, 但将  $x = 0$  时,  $\phi(0)$  值 [ (??) 解析解中的  $C$  值] 设定为 0.5。算例 4、算例 5、算例 6 初始状态为具有不同  $C_0$  值、已经事先训练好的精度小于  $10^{-7}$  网络, 训练方式与算例 3 类似, 将  $\phi(0) = 0.5$  作为加权损失函数组成部分进行训练。各算例训练精度小于  $10^{-7}$  即停止, 记录所需要的训练次数及精度等相关参数, 结果如表 4 所示。

算例	网络初始值 / $C_0$ 值	目标 $C$ 值设定	$\sigma_{\text{MSE},1}/10^{-8}$	学习次数 ( $n$ )	学习完成时 $C$ 值
1	随机/随机	无	9.9998	2405	0.015
2	随机/随机	无	9.1095	1606	0.0022
3	随机/随机	0.5	9.6124	379	0.4995
4	已训练 / $C_0 = 0.01$	0.5	9.1010	211	0.4996
5	已训练 / $C_0 = 0.05$	0.5	9.3309	55	0.4996
6	已训练 / $C_0 = 0.1$	0.5	9.5475	39	0.4995

表 4 神经网络初始值敏感性分析结果

可见, 算例 1、算例 2 分别收敛到不同的  $C$  值, 训练次数较多, 这说明了若不给定目标  $C$  值作为训练的加权  $Loss$ , 由于特征值方程的特征向量不确定性, 收敛得到的特征向量是不同的 (表现为收敛时  $C$  值不同), 且不易收敛。算例 3 训练次数大幅度下降, 而算例 4、算例 5、算例 6 由于有预训练初始值, 最终训练次数进一步降低, 且初始  $C$  值越接近目标值  $C = 0.5$ , 收敛越快, 这表明了采用已训练好的神经网络作为相似问题神经网络初始值, 具有良好的迁移学习特性, 可大幅度加快收敛进程。

## 4 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 Yang<sup>1</sup>

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.