



中国科学院大学
University of Chinese Academy of Sciences

博士学位论文

机器学习在太阳无场模型中的应用

作者姓名: 张尧

指导教师: 徐龙 研究员

中国科学院国家空间科学中心

学位类别: 理学博士

学科专业: 天文技术与方法

培养单位: 中国科学院国家空间科学中心

2024 年 8 月

Applications of Machine Learning in Solar Force-Free Modeling

**A dissertation submitted to
University of Chinese Academy of Sciences
in partial fulfillment of the requirement
for the degree of
Doctor of Philosophy
in Astronomical Technology and Methods
By
Yao Zhang
Supervisor: Research Fellow Long Xu**

National Space Science Center, Chinese Academy of Sciences

August, 2024

中国科学院大学 学位论文原创性声明

本人郑重声明：所呈交的学位论文是本人在导师的指导下独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的研究成果。对论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明或致谢。

作者签名：

日 期：

中国科学院大学 学位论文授权使用声明

本人完全了解并同意遵守中国科学院有关保存和使用学位论文的规定，即中国科学院有权保留送交学位论文的副本，允许该论文被查阅，可以按照学术研究公开原则和保护知识产权的原则公布该论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存、汇编本学位论文。

涉密及延迟公开的学位论文在解密或延期后适用本声明。

作者签名：

导师签名：

日 期：

日 期：

摘要

近年来，机器学习迅速发展，并与科学计算密切相关。机器学习和科学计算相互促进，为解决复杂的科学和工程问题提供了新的思路和方法。它们的结合推动了科学的研究的进展，并为数据驱动的科学模拟、预测和优化提供了强大的工具和理论基础。

在无力场模型中，关于解的存在性、唯一性和稳定性尚未得到确切的理论证明。1990 年 Low & Lou 提出了一类针对轴对称非线性无力场的半解析解形式，这些解在数值分析上被广泛用作验证无力场外推模型重建参考磁场的金标准。然而，对于含参的二阶非线性 Low & Lou 方程的初边值问题，人们仍然没有建立关于其解存在性和唯一性的理论。在过去的 30 年中，人们通常使用四阶龙格-库塔方法进行数值计算，并获得固定节点上函数值的近似解。为了更好地处理这个问题，我们提出了通过机器学习方法，在定义域上任意位置得到函数值的方法。

此外，无力场模型的外推问题实质上是通过科学计算推导日冕磁场分布的过程。无力场模型本质上是偏微分方程的边界值问题。然而，一些经典的无力场外推方法引入了时间变量，如优化法和磁流体动力学松弛法，这改变了无力场模型的属性。同时，传统的无力场模型外推方法通常使用有限差分、有限体积、有限元、谱方法等数值计算格式，对计算网格的选择和分布非常敏感。为了解决这个问题，我们提出了无力场神经网络外推法，结合了机器学习和数值方法的优势，用于探索和解决与偏微分方程相关的磁场外推问题。

本论文的主要工作包括以下两个方面：

1. 我们提出了一种基于参数打靶法和机器学习的数值方法，用于求解一类日冕磁场的半解析解。首先，通过参数打靶法求解 Low & Lou 方程中参数 a 的值。然后，利用我们提出的机器学习算法求解 Low & Lou 方程的数值解，可以得到在定义域上任意位置的 F 和 $\frac{dF}{d\mu}$ 数值结果。与传统的经典数值方法不同，传统经典的数值方法通常使用规则的数值网格求解 Low & Lou 方程的数值解，其节点的位置和数量确定，传统经典的数值方法通常只能直接提供计算节点上的值，而无法直接提供其他位置的值。

此外，我们还提出了一种数据驱动的数值算法，通过直接或间接给定 F 的

值来计算 Low & Lou 方程中的参数 n 和 a 的数值。

2. 我们提出了无力场神经网络外推法，利用机器学习方法重建日冕磁场。我们改进了传统的神经网络结构并引入了准输出层的概念，以解决经典机器学习在神经网络外推过程中可能破坏数据符合物理规律的问题。同时，我们采用了二阶优化方法来训练神经网络，相较于经典机器学习中的一阶优化方法更有效率。通过在 Low & Lou 提出的半解析非线性无力场模型中进行验证，结果表明在多个常见的评价指标下，我们的方法能够以高精度重建参考磁场。最后，我们在一个具体的太阳活动区上应用无力场神经网络法进行了磁场外推工作。

这些工作为无力场模型的理论研究和实际应用提供了有益的探索和发展方向。

关键词：机器学习，太阳物理，日冕磁场，非线性无力场，科学计算

Abstract

In recent years, machine learning has experienced rapid development and has become closely intertwined with scientific computing. The combination of machine learning and scientific computing has provided new insights and methodologies for tackling complex scientific and engineering problems. This integration has propelled advancements in scientific research and has offered powerful tools and theoretical foundations for data-driven scientific simulations, predictions, and optimizations.

In the force-free field model, there is currently no definitive theoretical proof regarding the existence, uniqueness, and stability of solutions. Nevertheless, in 1990, Low and Lou proposed a semi-analytical solution form for axisymmetric nonlinear force-free fields, which has been widely used as a benchmark in numerical analysis to validate the reconstruction of reference magnetic fields using force-free extrapolation models. However, for the parameterized second-order nonlinear equations introduced by Low and Lou, a theoretical basis for the existence and uniqueness of solutions is still lacking. Over the past 30 years, conventional numerical methods, such as the fourth-order Runge-Kutta method, have been commonly employed to numerically approximate the values of functions at fixed nodes. To address this issue more effectively, we propose a method that utilizes machine learning techniques to obtain the values of functions at arbitrary positions within the domain.

Additionally, the extrapolation problem in the force-free field model essentially involves deducing the distribution of coronal magnetic fields through scientific computations. The force-free field model is essentially a boundary value problem of partial differential equations. However, some classical extrapolation methods for force-free fields introduce a temporal variable, such as optimization-based methods and magneto-hydrodynamic relaxation methods, which alter the characteristics of the force-free field model. Furthermore, traditional extrapolation methods for force-free fields typically rely on numerical computation formats, such as finite differences, finite volumes, finite elements, and spectral methods, which are highly sensitive to the selection and distri-

bution of computational grids. To overcome these challenges, we propose a force-free neural network extrapolation method that combines the advantages of machine learning and numerical methods. This method aims to explore and solve the extrapolation problems associated with partial differential equations in magnetic field analysis.

The main contributions of this thesis are as follows:

1. We propose a numerical method based on parameter shooting and machine learning to compute numerical approximations of a class of semi-analytical solutions for coronal magnetic fields. Initially, we employ the parameter shooting method to determine the value of the parameter a in the Low & Lou equations. Subsequently, we utilize our proposed machine learning algorithm to obtain numerical results of the functions F and $\frac{dF}{d\mu}$ at arbitrary positions within the domain. In contrast to conventional classical numerical methods that rely on regular numerical grids and can only provide values at specific computation nodes, our approach directly provides values at other positions.

Additionally, we introduce a data-driven numerical algorithm that computes the values of parameters n and a by either directly or indirectly specifying the values of F .

2. We propose a force-free neural network extrapolation method that leverages machine learning techniques to reconstruct coronal magnetic fields. We enhance the classical neural network structure by introducing the concept of a quasi-output layer to address the challenge of preserving physical constraints during the neural network extrapolation process. Furthermore, we employ second-order optimization methods for training the neural network, which are more efficient compared to the first-order optimization methods commonly used in classical machine learning. Through validation experiments based on the semi-analytical nonlinear force-free field model developed by Low and Lou, our method demonstrates high-precision reconstruction of reference magnetic fields. Finally, we apply the force-free neural network extrapolation method to magnetic field extrapolation in a specific solar active region.

These contributions offer valuable exploration and development directions for both theoretical research and practical applications of the force-free field model.

Keywords: Machine Learning, Solar Physics, Coronal Magnetic Field, Nonlinear Force

Free Filed, Scientific Computing

目 录

| | |
|--|----|
| 第1章 绪论 | 1 |
| 1.1 太阳的一生 | 1 |
| 1.2 太阳的内部结构 | 3 |
| 1.3 太阳的大气结构 | 6 |
| 1.4 太阳活动区、黑子及活动周 | 10 |
| 1.5 日冕磁场 | 14 |
| 1.5.1 研究日冕磁场的意义 | 14 |
| 1.5.2 日冕磁场测量的难点 | 15 |
| 1.5.3 等离子方程组 | 16 |
| 1.5.4 经典无力场模型外推法存在的问题 | 18 |
| 1.6 本文各章节之间的逻辑关系 | 19 |
| 第2章 神经网络磁场外推法的数学基础 | 21 |
| 2.1 向量分析的基本知识 | 21 |
| 2.2 逼近定理的基本知识 | 27 |
| 2.3 机器学习的基本知识 | 27 |
| 2.4 本章小节 | 31 |
| 第3章 一类日冕磁场半解析解的数值方法 | 33 |
| 3.1 无力场模型刻画日冕磁场的合理性 | 33 |
| 3.2 日冕活动区中无力场模型的分类 | 35 |
| 3.3 Low & Lou 方程及其数值解 | 38 |
| 3.3.1 固定参数 n 后求解 Low & Lou 方程中参数 a 的数值算法 | 40 |
| 3.3.2 固定参数 n 和 a 后求解 Low & Lou 方程的机器学习算法 | 43 |
| 3.3.3 Low & Lou 方程机器学习数值解对应的无力场 | 48 |
| 3.3.4 Low & Lou 方程中参数的数据驱动求解算法 | 49 |
| 3.4 本章小节 | 52 |
| 第4章 神经网络磁场外推算法 | 53 |
| 4.1 经典无力场外推法中的数学形式 | 54 |
| 4.2 无力场神经网络磁场外推法 | 59 |
| 4.2.1 无力场神经网络外推法的数学模型 | 60 |
| 4.2.2 无力场神经网络外推法模型的数值算法 | 60 |
| 4.2.3 无力场神经网络外推法模型的应用 | 65 |
| 4.3 本章小节 | 78 |

| | |
|--|-----|
| 第 5 章 总结及展望 | 79 |
| 5.1 总结 | 79 |
| 5.2 展望 | 79 |
| 附录 A \mathbb{R}^3 中常见坐标系下的向量分析公式 | 83 |
| 附录 B 一般标准正交基经可逆变换后向量投影变换公式的证 明 | 87 |
| 附录 C 函数相关定理的证明 | 89 |
| 附录 D 机器学习算法在二阶常微分方程不同初边问题上的应 用 | 91 |
| 附录 E 一种三次插值函数的构造方法 | 95 |
| 参考文献 | 97 |
| 致谢 | 109 |
| 作者简历及攻读学位期间发表的学术论文与研究成果 | 111 |

图形列表

| | |
|---|----|
| 1.1 太阳的一生 | 1 |
| 1.2 太阳的结构 | 4 |
| 1.3 太阳标准模型的解..... | 5 |
| 1.4 太阳大气 VAL 模型..... | 7 |
| 1.5 太阳诸相 | 7 |
| 1.6 光球特性 | 8 |
| 1.7 色球特性 | 8 |
| 1.8 日冕特性 | 9 |
| 1.9 太阳活动区 | 11 |
| 1.10 1610 年至 2011 年年平均太阳黑子数..... | 12 |
| 1.11 太阳黑子群极性及极区磁场极性 | 13 |
| 1.12 蒙德蝴蝶图 | 14 |
| 2.1 散度计算示意图 | 23 |
| 2.2 旋度计算示意图 | 25 |
| 2.3 神经网络结构示意图 | 28 |
| 3.1 势场外推模型没有体现 Sigmoid 结构 | 36 |
| 3.2 无力场外推的例子..... | 37 |
| 3.3 $f'_2(-1)$ 的取值对参数 a 的影响 | 41 |
| 3.4 $F(a_i, \mu = 1)$ 零点分布图 | 42 |
| 3.5 MLP 的网络结构示意图..... | 43 |
| 3.6 微分方程求解器的效果..... | 45 |
| 3.7 $F, \frac{dF}{d\mu}$ 分别由 MLP 和 RK 求解器产生 | 47 |
| 3.8 由 na_1 和 na_2 生成磁场的差异 | 51 |
| 3.9 na_1 和 na_2 分别生成磁图 B_x | 51 |
| 3.10 na_1 和 na_2 分别生成磁场的等高线 | 52 |
| 4.1 无力场神经网络外推法采用的网络结构 | 61 |
| 4.2 有无准输出层时，损失函数随迭代次数变化情况的一个例子 | 62 |
| 4.3 可视化例子 1 中的磁力线 | 67 |
| 4.4 可视化例子 2 中的磁力线 | 69 |
| 4.5 关于例子 1，在 $z = 0$ 平面可视化 B_z | 70 |

| | |
|---|----|
| 4.6 关于例子 2，在 $z = 0$ 平面可视化 B_z | 71 |
| 4.7 太阳活动区 AR NOAA 11158 的光球磁场数据可视化 | 73 |
| 4.8 等离子压磁比 β 与高度之间的关系 | 74 |
| 4.9 观测者角度日地关系示意图 | 75 |
| 4.10 磁场分量对应的空间坐标示意图 | 75 |
| 4.11 磁力线分布图 | 76 |
| 4.12 局部磁力线在 xoy 平面投影后与观测数据成像后比对 | 77 |
| D.1 当 $n = 0.5$ 时，不同方法得到 A_n 和 A'_n 的结果比较 | 92 |
| D.2 当 $n = 1$ 时，不同方法得到 A_n 和 A'_n 的结果比较 | 93 |
| D.3 当 $n = 4$ 时，不同方法得到 A_n 和 A'_n 的结果比较 | 94 |

表格列表

| | |
|--|----|
| 1.1 太阳的物理参数 | 2 |
| 1.2 日冕物质能量密度及各能量来源密度的比较 | 10 |
| 3.1 n 给定后 $F(a_i, \mu = 1)$ 的前三个零点 | 41 |
| 3.2 MLP 中的超参数设置 | 45 |
| 3.3 MLP 求解器的性能 | 48 |
| 4.1 无力场神经网络外推法与其他经典无力场外推方法的比较 | 68 |
| D.1 不同方法得到 λ 数值结果的比较 | 91 |

第1章 绪论

太阳是一颗美丽而迷人的恒星，持续不断地向地球万物散发光和热，孕育了地球生命。通常，太阳是平静而祥和的，孕育生命，普照万物。然后，太阳也会时不时的发一下“脾气”，发射出强电磁辐射、抛射出等离子体，对地球环境产生扰动、破坏人类的高科技设施（例如卫星、航天器）、甚至危及人类的健康。

太阳是一颗距离我们最近的、能够对其进行高分辨率成像观测的恒星，在科学上具有重要的研究意义，至今太阳物理研究中仍有许多未解之谜，如日冕加热、太阳活动周形成和太阳爆发机理及触发机制。同时，太阳活动又与人类生产、生活息息相关，剧烈的太阳爆发会对人类生存的地球空间产生灾害性的后果，因此对太阳的研究和实时监测及太阳活动预报具有重大的现实意义。

1.1 太阳的一生

根据宇宙大爆炸理论中的恒星演化理论知：太阳诞生于星际云之中。星际云由漂浮在银河系中的气体和尘埃组成，其中气体中有 92% 的氢、7.8 % 的氦，而尘埃主要由硅酸盐组成^[1]。

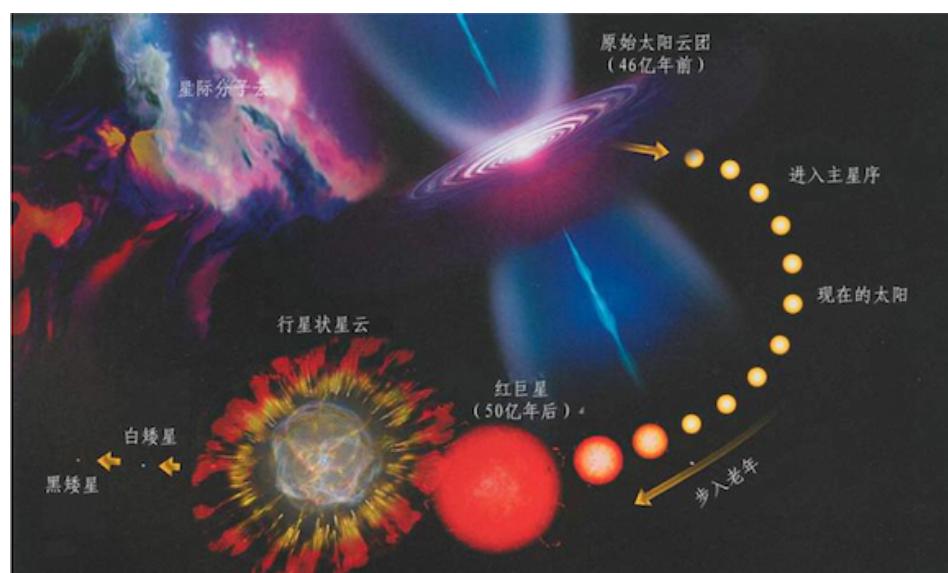


图 1.1 太阳的一生^[2]

大约 46 亿年前，在银河系某个角落发生了一次爆炸，产生了震波。这个震波传播到星际云中，导致星际云开始向密度较高的区域收缩、旋转和聚集。随着

表 1.1 太阳的物理参数

| | | | |
|------|---|--------|---------------------------------------|
| 姓名 | 太阳 | 年龄 | 4.6×10^9 年 |
| 质量 | 1.99×10^{30} 千克 | 半径 | 6.955×10^8 米 |
| 平均密度 | 1.4×10^3 千克·米 ⁻³ | 日地距离 | 1.496×10^{11} 米 |
| 表面重力 | 2.74×10^2 米·秒 ⁻² | 表面逃逸速度 | 6.18×10^2 千米·秒 ⁻¹ |
| 光度 | 3.86×10^{26} 瓦特 | 自转周期 | 26.24 天 |
| 角动量 | 1.7×10^{41} 千克·米 ² ·秒 ⁻¹ | 质量损失率 | 10^9 千克·秒 ⁻¹ |
| 有效温度 | 5.785×10^3 开尔文 | 1 角秒 | 7.26×10^5 米 |

收缩和旋转的稳定下来，星际云逐渐进入了重力和压力梯度平衡状态。持续缓慢的收缩和等离子体加热过程最终导致了太阳的形成。在这个过程中，太阳的核心逐渐形成并开始进行核融合反应。这些反应将氢原子融合成氦，并释放出巨大的能量。太阳的恒星演化使其保持稳定，并持续地向宇宙中散发光和热^[1]。

太阳诞生后立即进入主序星阶段，并经历了漫长的 100 亿年恒星寿命。随着太阳核心的氢燃料耗尽并转变为氦，太阳将演变成为红巨星，并最终发生新星爆炸，形成行星状星云。当太阳耗尽氦之后，其内部充满了氧和碳等元素，此时太阳将成为一颗白矮星。尘归尘，土归土，太阳开始从宇宙中来，最终又回到宇宙中^[1]。

当前，太阳是一颗由等离子体与磁场相互交织在一起的理想球体，是一颗普通的恒星，其光谱类型为 *G2V*。太阳的星等为 4.8，星等是指星体在天空中的相对亮度。值得一提的是，织女星的星等为 0.5，而一般人肉眼可见的星体星等通常小于 6.5。

太阳对地球的气候和空间天气有着深远的影响，其物理参数如表 1.1 所示。相较于地球，太阳的质量是地球的 3.3×10^5 倍，半径是地球的 1.09×10^2 倍，平均密度是地球的 2.7×10 倍。日地距离为一个天文单位，即 1.496×10^{11} 米，光从太阳传播到地球大约需要 8 分钟的时间。地球接收到的太阳辐射仅为每平方米约 10^3 瓦特，关于太阳更多的物理参数信息可参考相关教材^[3,4]。

在过去的几十年里，对太阳的认识在理论上取得了巨大的进步。尤其值得一提的是，在 [圣安德鲁斯大学数学系](#)、[谢菲尔德大学数学系](#)、[利兹大学数学系](#) 等地进行的关于太阳物理中的磁流体动力学研究工作。此外，在地基太阳观测方面，

美国大熊湖太阳天文台、美国国家太阳天文台、中国云南抚仙湖太阳观测站等提供了更多高时空分辨率的太阳观测数据。同时，天基太阳观测方面也取得了显著的成果，例如美国日地关系天文台 (STEREO)、日出号卫星 (Hinode)、美国太阳动态观测台 (SDO)、中国先进天基太阳天文台 (ASO-S) 等项目，它们为我们提供了更多高时空分辨率的太阳空间观测数据。这些重要的进展为我们深入了解太阳的特性和行为提供了有力支持。

1.2 太阳的内部结构

太阳物理学家根据观测和研究，将发出强烈白光且光线无法穿透的球面称为光球层，它是太阳的表面。光球层是太阳内部结构和大气结构之间的分界面^[3]，如图 1.2 所示。

标准太阳模型是基于球对称的准静态恒星模型，用于研究太阳内部的物理参数，如压力 (p)、密度 (ρ)、温度 (T) 和能量 (L)。它由一组数学方程组成，包括理想气体状态方程、流体静力学平衡方程、质量连续方程、能量守恒方程和能量传输方程。标准太阳模型中的理想状态方程是

$$p = \frac{k_B}{m} \rho T, \quad (1.1)$$

其中 m 是离子的质量， k_B 是玻尔兹曼常数。玻尔兹曼常数的数值约为 1.38×10^{-23} 焦耳·开尔文⁻¹，理想气体状态方程描述了太阳内部的压力、密度和温度之间的关系。标准太阳模型中的流体静力学平衡方程是

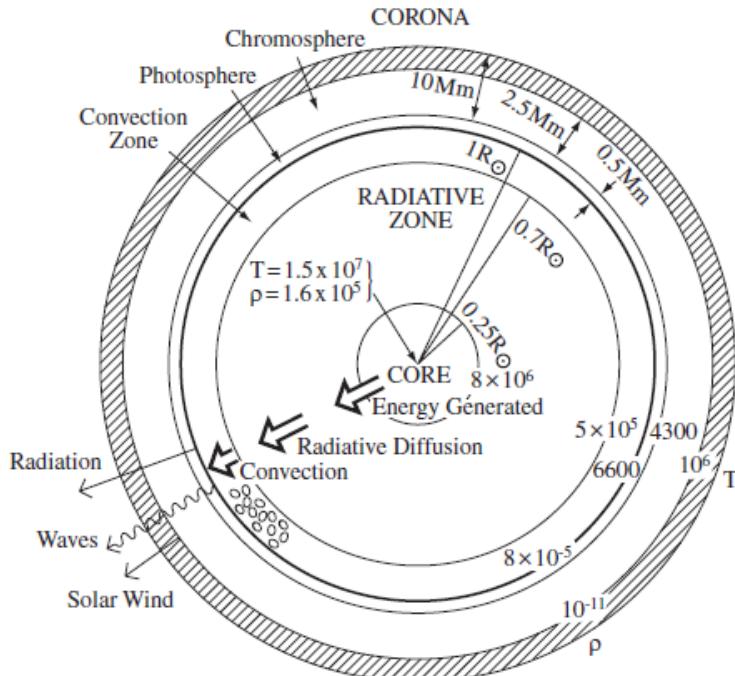
$$\frac{dp}{dr} = -\rho(r) g(r), \quad (1.2)$$

其中 $g(r) = \frac{M(r)G}{r^2}$ 、 $G = 6.67 \times 10^{-11}$ 牛顿·米²·千克⁻²，其中 M 满足质量连续方程 $\frac{dM}{dr} = 4\pi r^2 \rho$ 。流体静力学平衡方程表达了压力和重力在太阳内部的平衡关系，质量连续方程描述了质量在不同半径位置的变化。标准太阳模型中的能量守恒方程是

$$\frac{dL(r)}{dr} = 4\pi r^2 \rho \epsilon, \quad (1.3)$$

其中 ϵ 是单位质量核能发电率。能量守恒方程考虑了太阳内部的能量生成和传输过程。标准太阳模型在对流区以下的能量传输方程为

$$\frac{dT}{dr} = -\frac{3\kappa_R \rho}{16\sigma_{SB} T^3} \frac{L}{4\pi r^2}, \quad (1.4)$$

图 1.2 太阳的结构^[3]

其中 κ_R 是罗斯兰平均不透明度、 σ_{SB} 是斯特藩-玻尔兹曼常数。罗斯兰平均不透明度指的是描述物质对辐射传输的不透明程度的物理量，它用于描述介质中各种能级跃迁引起的吸收和散射对辐射传输的综合影响。斯特藩-玻尔兹曼常数是 5.67×10^{-8} 瓦特 · (米 2 · 开尔文 4) $^{-1}$ 。标准太阳模型在对流区的能量传输方程为

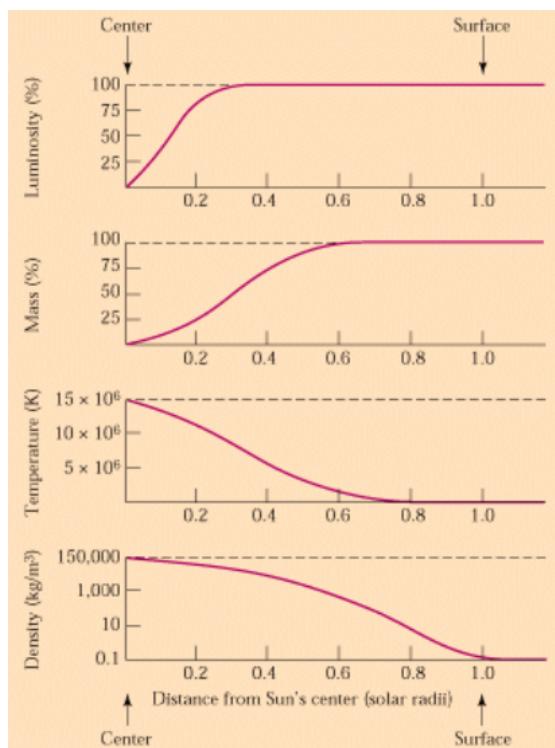
$$\frac{dT}{dr} = -\frac{\gamma - 1}{\gamma} \frac{gm}{k_B}, \quad (1.5)$$

其中 γ 为绝热膨胀系数，对于理想气体其绝热膨胀系数与其比热容比有关。能量传输方程则描述了辐射区和对流区内能量传输的方式。根据给定的边界条件，可以解出上述方程来确定恒星内部的物理特性。标准太阳模型中诸方程的边界条件是

$$M = 0, \frac{dT}{dr} = 0, \text{ 当 } r = 0 \text{ 米时,} \\ M = 1.989 \times 10^{30} \text{ 千克, } L = 3.846 \times 10^{26} \text{ 瓦特, 当 } r = 6.955 \times 10^8 \text{ 米时.} \quad (1.6)$$

根据标准太阳模型(图 1.2)，太阳的内部结构分为以下三个部分：

1. 核心：核心处在 0.25 个太阳半径以内，是核聚变反应发生的地方。核聚变是指四个氢原子核聚变成一个氦原子核，同时释放出两个正电子、两个中微子和能量即 $4^1H \rightarrow ^4He + 2e^+ + 2\nu + 26.7 \text{ MeV}$ 。

图 1.3 太阳标准模型的解^[5]

标准太阳模型计算结果(图 1.3)表明,核心的温度约为 1.5×10^7 开尔文、密度约为 1.5×10^5 千克·米⁻³。尽管核心只占太阳体积的五十分之一,但质量却占了太阳总质量的一半以上,并且产生了 99% 以上的能量。

2. 辐射层: 辐射层处在核心与 0.7 个太阳半径之间。在辐射层,由核心释放出的高能量的 γ 射线与 X 射线会与辐射层的粒子碰撞,并通过辐射的形式传输能量。光子需要花费数百年的时间才能穿透辐射层,并最终以可见光的形式传输到太阳表面。辐射层起到了保护地球上生物免受辐射伤害的作用。

3. 对流层: 对流层位于辐射层的外部,靠近太阳光球表面。根据标准太阳模型计算结果(图 1.3)对流层的厚度约为 1.5×10^5 千米。在对流层中,温度呈迅速下降的趋势。这是因为辐射层外围物质的透明度急剧下降,再加上太阳表面的辐射损失,导致对流层内外存在较大的温差。当满足史瓦西准则时,即温度梯度的绝对值大于一定值时,形成以湍流为主的强烈对流层。史瓦西准则表示为:

$$\left| \frac{dT}{dr} \right| > \frac{\gamma-1}{\gamma} \frac{gm}{k_B}$$
, 其中 γ 是绝热膨胀系数、 k_B 是玻尔兹曼常数、 m 是离子的质量。
当温度梯度超过这个临界值时,对流层中会出现湍流现象。

在对流层中,太阳内部的能量部分会转化为等离子体气流的动能。这些气流的运动和湍流产生了复杂的磁场结构。由于等离子体是带电粒子的集合,它们的

运动会产生电流，而电流又会产生磁场。因此，人们普遍认为对流层是产生太阳磁场的地方。对流层与太阳光球层 (Solar Photosphere)、色球层 (Solar Chromosphere) 和日冕层 (Solar Corona) 的活动及喷流密切相关。这些层之间存在着相互作用，通过对流层中气流的运动，能量和物质从内部向外部传输，并影响太阳表面层的活动。这些运动和湍流的产生也使得对流层中的磁场变得复杂多样。因此，对流层在研究太阳活动和太阳磁场生成机制方面具有重要意义。

1.3 太阳的大气结构

太阳大气的 VAL^[3,6] 模型是一个用于描述太阳大气结构的理论模型，它已成功地拟合了许多来自不同层的谱线观测数据。这个模型最早由 Vernazza 等 在 1981 年提出，并在之后的研究中进行了改进和修正。VAL 模型的主要目的是通过简化描述太阳大气中温度和密度随高度变化的示意图 (图 1.3)，来更好地理解太阳物理过程。该模型将太阳大气分为光球、色球和日冕层。

当我们使用不同波长的滤光片观测太阳时，可以获得不同大气环境下的太阳图像。不同波长的光线在太阳大气层中会受到不同程度的吸收、散射和发射，从而呈现出不同的特征，如图 1.5。通过对这些太阳图像的观察和分析，可以揭示太阳大气的复杂性，并帮助我们更深入地理解太阳物理过程。

太阳的表面就是光球层，其温度约为 5.8×10^3 开尔文，厚度约为 5×10^2 千米。光球表面具有米粒组织和太阳黑子等特征。米粒组织，如图 1.6a 所示，呈现出边缘暗、中间亮的特点，尺度约为 1.5×10^3 千米，而太阳黑子，如图 1.6b 所示，由本影 (暗核部分) 和半影 (周围明亮的边框) 组成，存活时间可长达数日至数月。此外，存在临边昏暗现像，如图 1.6c 所示，太阳圆盘的亮度从中心向边缘逐渐变暗。

在地球所见的太阳光谱中，光球层发射出与其表面温度相当的热辐射光谱，在可见光范围内强度最大，谱型相当于一个温度为 5.8×10^3 开尔文的黑体辐射。而在其他波段如紫外线、X 射线、 γ 射线、远红外和无线电波段的辐射强度则远高于该温度下的黑体辐射强度。通过分析太阳连续光谱上的吸收谱线，科学家们发现太阳大气主要由氢、氦、铁、镁、镍、钠等化学元素组成。在光球层中没有探测到氦的谱线，这与光球层的较低温度有关。

太阳的色球位于光球之上，日冕之下，其厚度约为 2.5×10^3 千米。尽管发光

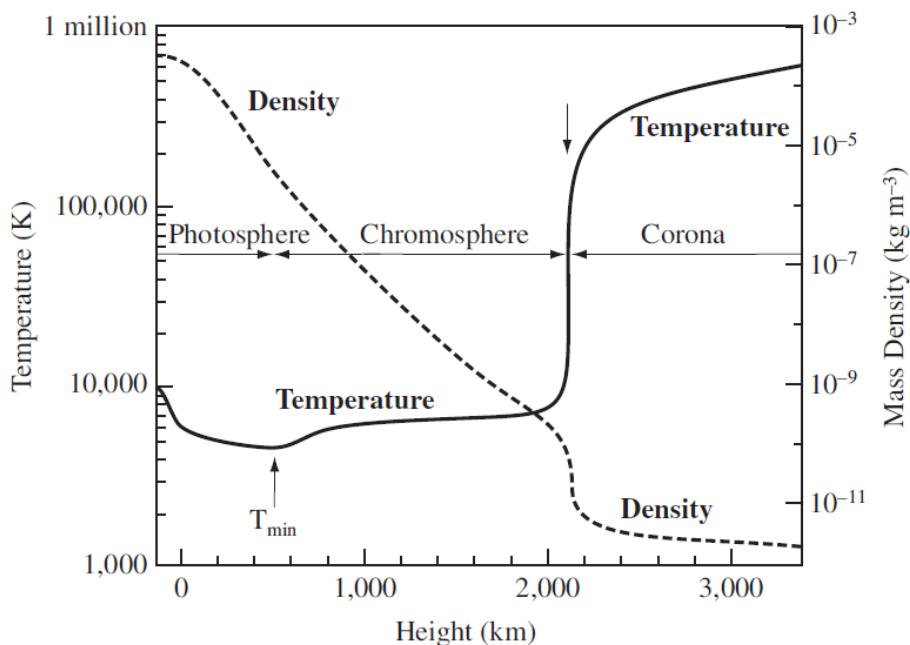
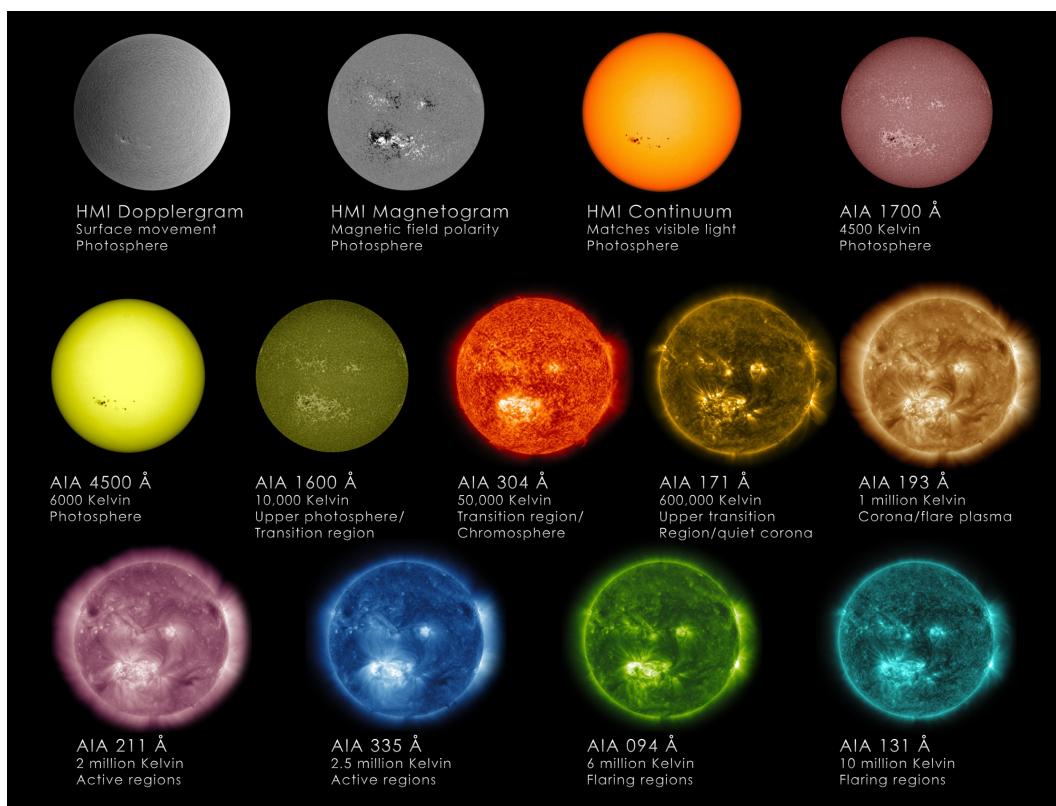
图 1.4 太阳大气 VAL 模型^[3]

图 1.5 太阳诸相

强度只有光球层的万分之一，但色球的温度却比光球高。随着高度的增加，色球的温度也逐渐升高。在色球底层，即厚度约为 400 千米的区域，温度从 4.2×10^3 开尔文升高到 5.5×10^3 开尔文。然后，在厚度约为 1.2×10^3 千米的中层区域，温

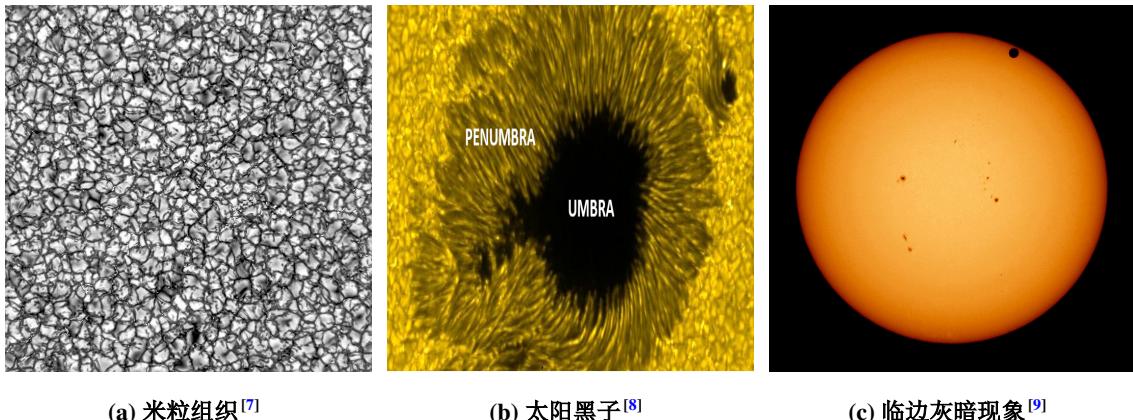


图 1.6 光球特性

度缓慢上升到 8×10^3 开尔文。最后，在高层，温度突然上升到数百万开尔文。因此，在色球层可以观察到氦的谱线。

通过观测发现，针状体 (Solar Spicule) 充满了整个色球层，并且它们的存活时间为 5 到 10 分钟。这些针状体通常与高磁性区域相关联，每时每刻大约有约 3×10^5 个针状体围绕着太阳。

观测结果显示，太阳的光球层会出现突发闪光现象，被称为太阳耀斑 (Solar Flare)，图 1.7b 所示。每次太阳耀斑的爆发持续时间在数分钟到几十分钟之间。一次太阳耀斑的能量释放可达 6×10^{25} 焦耳，相当于同时爆炸数百亿颗氢弹所释放的能量。

在 1990 年代，通过对阳光号卫星 (Yohkoh) 的观测，科学家揭示了太阳耀斑爆发中存在磁重联现象。根据 Tsuneta 等 等人的研究^[12]，太阳耀斑的能量源自于磁重联过程中释放的能量。磁重联是指太阳表面的磁场线重新连接，导致释

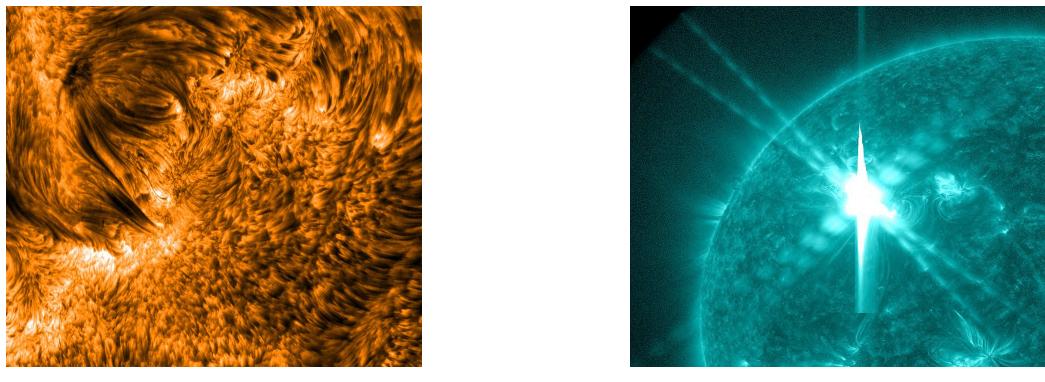


图 1.7 色球特性

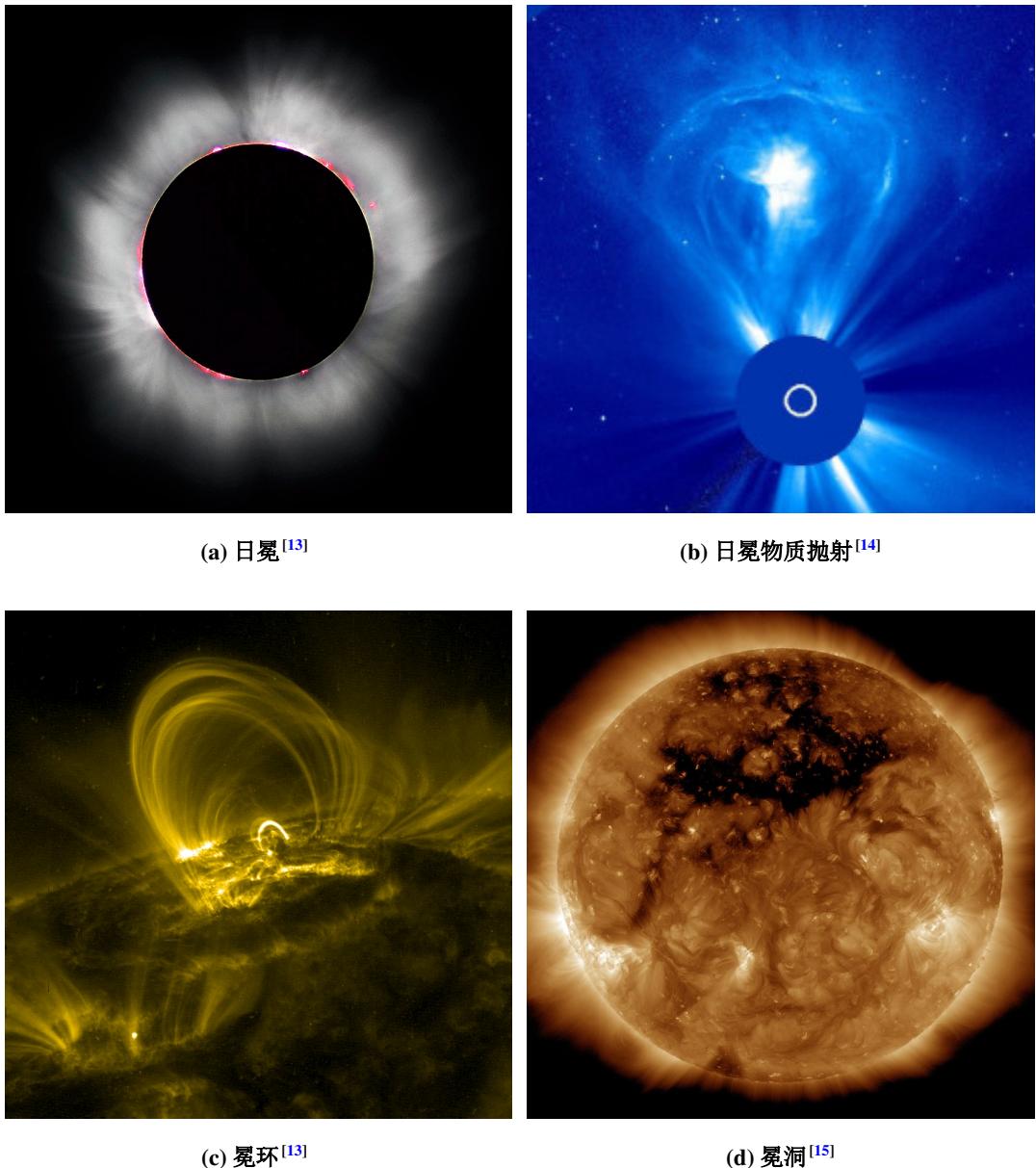


图 1.8 日冕特性

放大量能量的过程。这个过程涉及磁场的重新排列和释放，产生高温、高能粒子和强烈的辐射。因此，磁重联被认为是太阳耀斑能量释放的主要机制之一。

太阳的日冕层是位于色球层之上的外部大气层，由高温等离子体组成。它环绕着太阳形成了一个等离子光环，厚度约为太阳半径的 1.3 倍，如图 1.8a 所示。日冕层的温度非常高，约为 10^6 开尔文，远高于太阳表面和色球层的温度。这种高温主要是由于日冕中存在强烈的磁场活动和能量释放过程。尽管日冕温度很高，但密度相对较低，因为大部分物质以等离子态存在，而不是以固体或液体形式存在。

表 1.2 日冕物质能量密度及各能量来源密度的比较^[17]

| | 计算公式 | 观测均值 | 能量密度(尔格·厘米 ⁻³) |
|--------|-----------------------|---|----------------------------|
| 日冕物质抛射 | | | 100 |
| 日冕的动能 | $\frac{1}{2}m_p nv^2$ | $n = 10^{-9}$ 厘米 ⁻³ , $v = 1$ 千米·秒 ⁻¹ | 10^{-5} |
| 日冕的热能 | nkT | $T = 10^6$ 开尔文 | 0.1 |
| 引力势能 | $m_p n g h$ | $h = 10^5$ 千米 | 0.5 |
| 日冕的磁能 | $\frac{B^2}{8\pi}$ | $B = 100$ 高斯 | 400 |

日冕层经常发生日冕物质抛射 (Coronal Mass Ejection, CME) 现象，即释放等离子体和磁场到太阳系内。图 1.8b 展示了日冕物质抛射事件爆发时的图像。2010 年 5 月 23 日，日地关系天文台 (STEREO) 观测到一次日冕物质抛射事件中的磁场重联和合并过程^[16]。根据 [Forbes](#) 等人的研究^[17]，以中等级别的日冕物质抛射事件中，磁场是主要的能量来源，如表 1.2 所示。从观测的角度来看，在日冕物质抛射事件中，磁场同样起主导作用。

在观测中，人们发现了太阳大气中的一种特殊结构，称为冕环 (Coronal Loop)，如图 1.8c。冕环是由磁力线两端固定在太阳表面上的一种结构。太阳黑子经常出现在冕环的足点处，这可能是由于冕环的磁场活动导致了太阳黑子的形成。

此外，在观测过程中还发现日冕层中的冕洞 (Coronal Hole)，如图 1.8d 的黑暗区域所示。冕洞是由于该区域能量和气体密度低于平均值而形成的。冕洞通常是太阳极小期内空间天气扰动的主要来源之一。冕洞的存在对太阳活动和太阳风的形成具有重要影响。冕洞区域的较低密度使得来自太阳的带电粒子能够更容易地逸出太阳，并形成太阳风。这些太阳风会在宇宙空间中传播，并对地球的磁层和地球附近的空间环境产生影响，例如引发地磁暴和极光等现象。通过对冕环和冕洞的观测和研究，我们能够更好地理解太阳大气的结构和活动以及它们对太阳风和地球空间环境的影响。

1.4 太阳活动区、黑子及活动周

太阳活动区 (Solar Active Region) 是指太阳大气中的局部临时区域，其具有强磁场或复杂的磁场结构，并频繁发生太阳活动，如太阳黑子、耀斑、日冕物质抛射等现象。太阳活动区通常与太阳黑子密切相关，被认为是产生耀斑爆发和日冕物质抛射的主要来源。在给定的时间内，太阳活动区的数量和位置通常与太阳

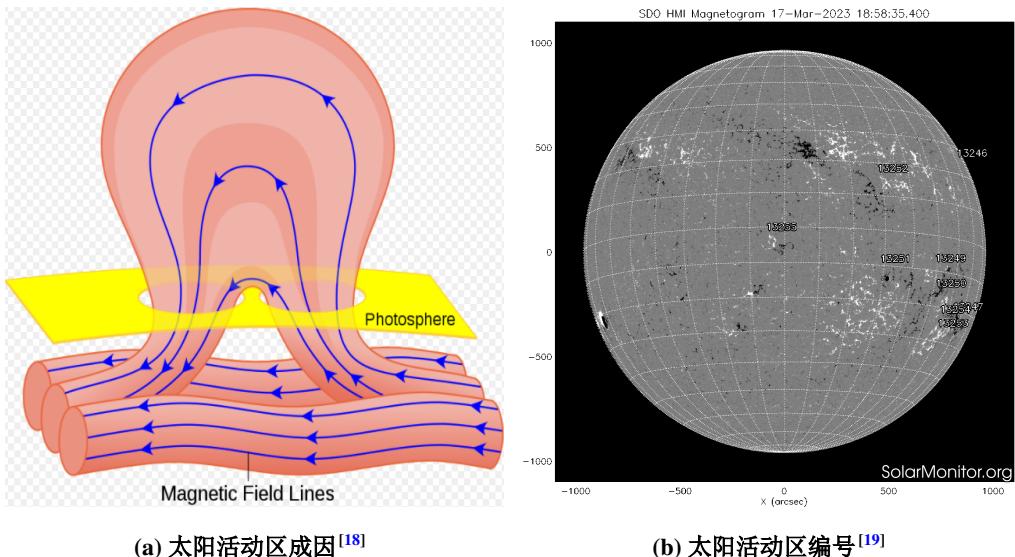


图 1.9 太阳活动区

活动周期（Solar Cycle）相关联^[6,20]。太阳活动区的存活周期在几周到几个月之间不等。

太阳活动区的形成通常伴随着太阳黑子的出现，但并不是所有的活动区都会有太阳黑子出现。当对流层中的磁场浮出光球层时，太阳活动区首先形成，如图 1.9a 所示。美国空间天气预报中心（SWPC）会在第二天为该活动区生成一个活动区编号。例如，2023 年 3 月 14 日，在日面中心生成了编号为 13255 的活动区，如图 1.9b 所示，并且该活动区在 3 月 17 日发生了 C 级耀斑。

太阳活动区的能量聚集通常会伴随着太阳黑子的形成。然而，并非所有的活动区都会产生太阳黑子。随着对流层中的磁场的浮现和演化，能量逐渐积聚在太阳活动区中。如果这个过程演化缓慢或相对稳定，太阳活动区中可能会形成日珥（Solar Prominence）。

当日冕磁场与太阳活动区的磁场发生磁重联 (Magnetic Reconnection) 时，能量会突然释放，通常伴随着耀斑和日冕物质抛射的发生。最终，太阳活动区发生磁对消 (Magnetic Cancellation)，太阳黑子消失，整个太阳活动区逐渐消亡。太阳黑子的中心温度约为 4×10^3 开尔文，比光球表面的温度低，因此它们呈现为较暗的区域，被称为黑子。大多数太阳黑子以群体形式随着太阳的自转运动。每个黑子群通常包含前导黑子和后导黑子，它们的磁极性相反。同时，太阳的南半球和北半球的黑子群的磁极性也是相反的。太阳黑子的存活时间从几天到几个月不等，具体取决于其特定的性质和演化过程。一般来说，黑子的寿命受许多因

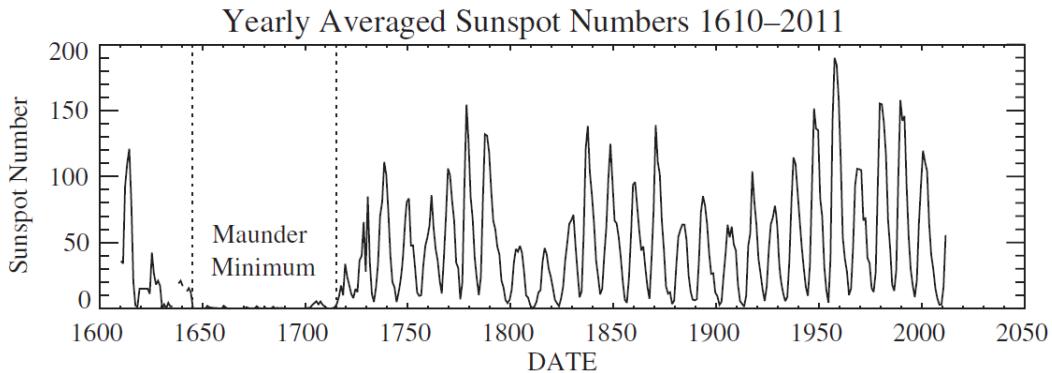


图 1.10 1610 年至 2011 年年平均太阳黑子数

素的影响，包括其大小、形状和环境条件等。有些黑子可能会迅速消失，而其他黑子则可能存在更长的时间。

太阳活动周期或太阳黑子活动周期是指太阳活动特征的变化遵循太阳黑子数目每 11 年波动振荡的规律^[21]。这种周期性可以用沃尔夫数来衡量，沃尔夫数 R 的定义如下：

$$R = K(10g + f), \quad (1.7)$$

其中 K 是观测校正因子，其值会随着观测地点和观测设备的不同而变化。通常情况下， K 取值为 0.6。 g 代表日面上的黑子群数， f 代表日面上的黑子总数。此外，图 1.10 展示了太阳黑子数的规律变化。从 1610 年到 2010 年，太阳黑子数呈现出周期性行为。在蒙德极小期（1645 年到 1715 年）期间，几乎没有太阳黑子出现，与此同时，中国正处于明末自然灾害多发的时期。同样，在道尔顿极小期（1800 年到 1820 年），太阳黑子数也较低，这段时间对应了清朝末年，中国同样经历了多次自然灾害。

太阳黑子数的周期在 7.3 年到 13.3 年之间有所变化，平均约为 10.8 年。因此，人们通常将太阳黑子的周期称为 11 年。

根据长期观测的结果，我们发现在前一个太阳黑子周期中，如果太阳北半球的前导黑子的磁极性为南极，那么后导黑子的磁极性就会是北极，如图 1.11a 所示。在太阳活动极大期，极区冕洞会逐渐收缩并消失，导致太阳黑子向赤道漂移，如图 1.11b 所示。在接下来的一个太阳黑子周期中，太阳北半球的前导黑子的磁极性将变为北极，而后导黑子的磁极性则为南极，如图 1.11c 所示。值得注意的是，太阳南半球的磁极性与北半球正好相反。根据这一观测结果，我们可以

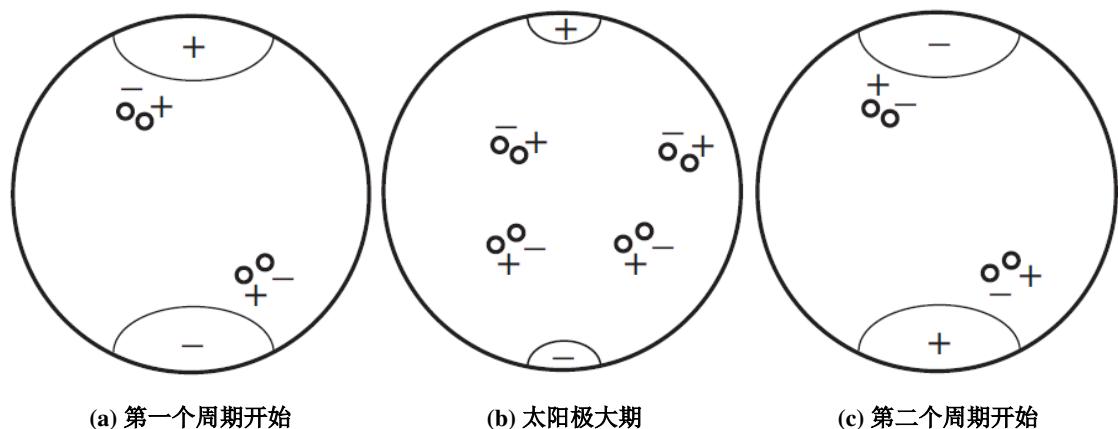
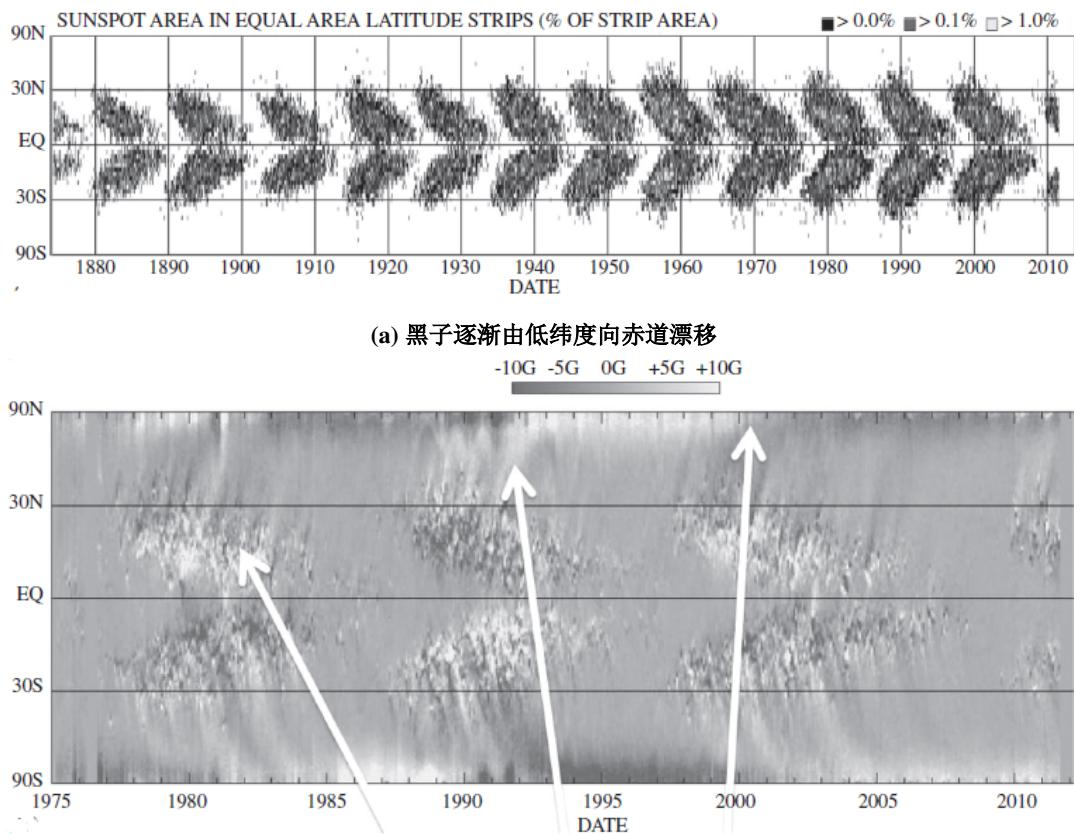


图 1.11 太阳黑子群极性及极区磁场极性

推断出太阳的磁周期约为 22 年。这意味着太阳黑子活动和太阳磁场的变化以大约 22 年的周期重复出现。这种周期性的变化对于理解太阳活动的演化和预测太阳黑子周期具有重要意义。

根据长期观测的结果，如果以年份为横坐标，太阳黑子出现的纬度为纵坐标，可以绘制出太阳黑子的分布图。在太阳黑子活动周期开始时，太阳黑子主要分布在南北纬 35 度处，如图 1.12a 所示。而太阳黑子周期结束后，黑子常常出现在赤道附近，呈现出蝴蝶状的分布图案。这种周期性变化展示了太阳黑子的活动规律。

此外，图 1.12b 说明了太阳磁场的周期大致是太阳黑子周期的两倍。这意味着太阳黑子周期的变化与太阳磁场周期密切相关。通过对太阳黑子和太阳磁场周期的观测和分析，我们可以更好地理解太阳活动的周期性特征以及与之相关的太阳磁场变化。



(b) 在光球纵向磁场(亮的表示正极,暗表示负极)中,第一个箭头太阳黑子向赤道漂移,第二个箭头向极地迁移,最后一个箭头极地磁场磁性反转

图 1.12 蒙德蝴蝶图^[21]

1.5 日冕磁场

日冕磁场是指太阳的日冕层中存在的磁场。日冕磁场在太阳大气的物理过程中起着重要作用,并与太阳活动密切相关。日冕磁场具有复杂的结构和演化。日冕磁场的分布不均匀,形成了各种形状的磁环、磁弧和磁流束等。这些磁场结构反映了太阳磁场的复杂性,涉及磁场的重联、扭曲和释放等过程。

1.5.1 研究日冕磁场的意义

通过研究日冕磁场,可以帮助我们揭示太阳耀斑、日冕物质抛射及其它类型爆发性活动的发生机制,使得我们可以预测这些活动对地球和空间环境的影响,以便人们采取相应措施保护卫星通信、导航系统以及电力网络等关键基础设施的安全运行。

日冕磁场是太阳风形成的重要因素之一,太阳风是由日冕高温等离子体射出的带电粒子流,这影响地球磁场和行星磁场。通过研究日冕磁场,可以帮助我

们理解太阳风的起源和演化规律，进而提高我们对太阳风和空间天气现象(如磁暴和极光等)的预报和监测能力，为空间探测和人类活动提供保障。

日冕磁场是太阳整体磁场结构的重要组成部分，通过研究日冕磁场，可以帮助我们追踪太阳磁场在时空间上的演化过程，揭示太阳活动周期、磁场重联以及磁场反转等现象的发生机制，更深入理解太阳磁场的起源、演化和动力学过程并为恒星磁场和宇宙磁场的研究提供参考。

日冕磁场影响着日冕层中的物质输运、加热和释放机制，研究日冕磁场可以帮助我们揭示日冕中物质和能量的传输机制、形成高温环境、加热和冷却过程、磁重联等的机制，从而进一步了解太阳大气层物理过程和行为。这对于研究太阳和其他恒星的大气层和宇宙等离子体的物理学具有重要意义。

研究日冕磁场对于理解太阳活动、预报太阳爆发、揭示日冕物质抛射机制、探索太阳磁场演化规律和动力学以及深入了解太阳大气层物理过程等方面具有重要的科学意义和应用价值。

1.5.2 日冕磁场测量的难点

在太阳的光球层可以通过观测太阳光谱中的谱线来间接测量太阳光球磁场。当光通过太阳光球区域时受到磁场的影响，光谱线会发生塞曼效应的分裂。通过分析这些分裂的谱线，人们可以推断出太阳光球区域的磁场强度和方向。

具体而言，太阳光谱中的谱线会根据磁场的影响而发生分裂，分裂程度与磁场强度相关。分裂通常呈现为多个成分或多个峰值，其中一部分与无磁场时的谱线重合而其他部分则相对分离，这种分裂即为塞曼效应。根据塞曼效应效应的分裂模式和程度可以确定太阳光球区域的磁场性质，包括磁场强度、方向和结构等...

需要注意的是，观测和分析太阳光谱中的塞曼效应需要使用高分辨率和高灵敏度的仪器，同时还需要针对特定的元素(常用的元素包括氢、钠、铁等)或分子进行观测，因为不同的谱线对应不同的能级跃迁和分裂模式。

然而，测量日冕磁场却是件困难的事，这是由于

1. 日冕的磁场结构非常复杂，日冕磁场具有多种尺度和空间变化。这使得在日冕中进行精确的磁场测量变得非常困难。通过塞曼效应测量磁场通常需要一个较为均匀和稳定的磁场环境，而日冕中的磁场远远没有满足这样的条件。

2. 日冕层是太阳大气层最外层的区域，日冕层温度高达 10^6 到 3×10^6 开尔文，这使得直接观测或将传感器放置在日冕中进行磁场测量非常困难，高温环境可能影响测量设备的性能并导致其损坏使其不能正常工作。

3. 目前的观测技术和仪器还不足以在日冕中进行塞曼效应测量。由于日冕的极高辐射和强烈的光谱连续性，需要具备高分辨率、高灵敏度和耐高温等特征的设备才能进行有效测量。

于是人们通过磁场外推方法来推断日冕磁场，以帮助人们研究日冕磁场的性质和行为。日冕磁场外推指在适当的微分方程模型下，以太阳光球磁场观测值为边界条件，最后由数值计算方法求解日冕层中的磁场，从而揭示日冕磁场的结构。

1.5.3 等离子方程组

等离子方程组 (Plasma Equations) 在研究日冕磁场中起着关键作用。等离子方程组描述了等离子体的运动和行为，包括连续性方程、动量方程和能量方程。

1. 连续性方程 (Mass Continuity):

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{v}) = 0, \quad (1.8)$$

其中 ρ 是物质的密度， \vec{v} 是它的速度。有些时候 (1.8) 也写成

$$\frac{d\rho}{dt} + \rho \nabla \cdot \vec{v} = 0, \quad (1.9)$$

其中 $\frac{d}{dt} = \frac{\partial}{\partial t} + \vec{v} \cdot \nabla$ 。

实际上 (1.8) 等价于

$$\iiint_V \frac{\partial \rho}{\partial t} dV = - \iiint_V \nabla \cdot (\rho \vec{v}) dV = - \iint_V (\rho \vec{v}) \cdot \vec{n} dS, \quad (1.10)$$

故连续性方程的本质是质量守恒定律，即物质的质量变化率和该物质进出表面的通量相同。

2. 动量方程 (Equation of Motion):

$$\rho \frac{d\vec{v}}{dt} = -\nabla p + \vec{j} \times \vec{B} + \vec{F}, \quad (1.11)$$

其中 p 是等离子的压力 (单位面积所受的力)， $\vec{j} \times \vec{B}$ 是受的洛伦兹力， \vec{j} 是电流， \vec{B} 是磁场， \vec{F} 是受的外力。方程 (1.11) 等价于

$$\frac{\partial \vec{v}}{\partial t} + (\nabla \cdot \vec{v}) \vec{v} + \frac{1}{\rho} \nabla p = \vec{f}, \quad (1.12)$$

其中 $\vec{f} = \frac{1}{\rho} (\vec{j} \times \vec{B} + \vec{F})$, 该式是标准的可压缩欧拉方程。若要求解方程, 则需要添加对应的初值和边值条件。

动量方程描述了等离子体中粒子的运动和受力情况。它考虑了压力梯度力、洛伦兹力(由电流和磁场相互作用产生的力)以及外力对等离子体的影响。

在某些情况下, 为了简化分析和数值计算, 动量方程可以进行一些近似和简化。在本文的第3章第1节中, 我们将对磁静力学和无力场模型进行分析, 以评估其在刻画日冕磁场时的合理性。磁静力学和无力场模型是动量方程的不同简化形式, 适用于刻画不同特性的日冕磁场。

近年来, 基于无力场模型 $\nabla \times \vec{B} = \alpha \vec{B}$, 在日冕磁场外推工作中涌现了多种的不同方法。接下来, 我们将重点介绍用于计算非线性无力场的方法: Grad-Rubin (GR) 方法^[22–29]、竖直积分法^[30–32]、优化法^[33–36]、磁流体动力学松弛法^[37–39]和边界积分法^[40–42]。下面我们一一介绍这些方法。

GR方法的是一种交替迭代的方法。首先, 固定 \vec{B} , 通过方程 $\nabla \alpha \cdot \vec{B} = 0$ 求解 α 。接着, 固定 α 通过方程 $(\Delta + \alpha^2) \vec{B} = \vec{0}$ 求解 \vec{B} 。重复上述迭代过程到算法收敛为止。

竖直积分法通过下列微分方程组,

$$\frac{dB_x}{dz} = \frac{\partial B_{z_0}}{\partial x} + J_{y_0}, \quad \frac{dB_y}{dz} = -\frac{\partial B_{y_0}}{\partial y} + J_{x_0}, \quad \frac{dB_z}{dz} = -\frac{\partial B_{x_0}}{\partial x} - \frac{\partial B_{y_0}}{\partial y},$$

来求解磁场 \vec{B} 的三分量 B_x 、 B_y 和 B_z 。

优化法通过求解方程,

$$\frac{\partial \vec{B}}{\partial t} = \mu \left[\nabla \times (\vec{\Omega} \times \vec{B}) - \vec{\Omega} \times (\nabla \times \vec{B}) - \nabla (\vec{\Omega} \cdot \vec{B}) + \vec{\Omega} (\nabla \cdot \vec{B}) + \vec{\Omega}^2 \vec{B} \right],$$

求解磁场 \vec{B} , 其中 $\vec{\Omega} = \vec{B}^{-2} \left[(\nabla \times \vec{B}) \times \vec{B} - (\nabla \cdot \vec{B}) \vec{B} \right]$ 。

磁流体动力学松弛法通过求解微分方程,

$$\frac{\partial \vec{B}}{\partial t} = \mu \nabla \times \frac{[(\nabla \times \vec{B}) \times \vec{B}] \times \vec{B}}{\vec{B}^2},$$

得到磁场 \vec{B} 。

边界积分法利用格林第二等式

$$\int_{\Omega} (Y \Delta B - B \Delta Y) dv = \oint_{\Gamma \cup S_R \cup S_\epsilon} \left(Y \frac{\partial B}{\partial \vec{n}} - B \frac{\partial Y}{\partial \vec{n}} \right) ds,$$

选取适当的格林函数, 代入上述等式后得到磁场的边界积分表达式。

同时，基于磁静力学模型的方法包括了基于 GR 方法的磁静力学方法^[43,44]；基于优化法的磁静力学方法^[45–47]；此外，基于磁流体动力学方法的磁静力学方法^[48–50]。

3. 理想气体状态方程 (Perfect Gas Law) :

$$p = nk_B T,$$

决定，其中 n 是单位体积内的等离子体数目， k_B 是玻尔兹曼常数， T 为热力学温度。

理想气体状态方程描述了等离子体中压力、密度和温度之间的关系。它是一个简化的状态方程，假设等离子体满足理想气体行为。根据该方程，压力与粒子数密度和温度之间成正比关系。

1.5.4 经典无力场模型外推法存在的问题

本文所指的经典无力场模型外推法指的是 GR 方法、竖直积分法、优化法、磁流体动力学松弛法、边界积分方法。

一方面，经典无力场模型外推法在数学建模过程中与无力场模型本身产生了差异，例如优化法和磁流体动力学松弛法引入了伪时间演化过程，而无力场模型本身是偏微分方程组的边界值问题，与时间无关。

另一方面，经典无力场模型外推法的数值解通常使用有限差分、有限体积、有限元、边界元、谱方法等不同的数值方法，然而这些方法在数值求解过程中可能会遇到一些问题。比如：

1. 不同的数值方法对于不同类型的问题具有不同的精度和收敛性。某些方法可能需要更高的网格分辨率或更复杂的算法才能获得准确的结果，而且在某些情况下，数值解可能难以达到预期的收敛性。

2. 很多数值方法对于网格的选择和分布是敏感的，尤其是在存在复杂几何形状或奇点的情况下。不合适的网格选择可能导致数值解的不稳定性或者产生不准确的结果。比如日冕磁场外推法中的竖直积分法对网格的选择和分布及其敏感。

3. 某些数值方法在离散化过程中可能引入数值耗散和数值分散。数值耗散会导致数值解的平滑性降低，而数值分散则可能导致数值解的精度下降。

4. 不同的数值方法在计算效率和存储需求上也存在差异。某些方法可能需

要更多的计算资源和内存空间，而且在处理大规模问题时可能会面临计算时间长、存储需求高等挑战。

我们将在第4章详细讨论偏微分方程各种数值方法的优缺点，并通过机器学习与科学计算结合以部分改进上述问题。

1.6 本文各章节之间的逻辑关系

我们旨在在数学建模过程中，通过利用函数逼近论的知识来减少无力场外推法与无力场模型之间的差异。选择合适的函数空间和逼近方法，尽量减小建模带来的误差。另外，在建模完成后，可以应用机器学习算法进一步减少由数值格式引入的误差。具体内容详见本文第4章，并且第4章需要了解第2章中逼近定理和机器学习的基本知识。

为了评估第4章中无力场神经网络外推法的性能，需要提供无力场方程的解析解或数值解。然而，对于一般形式的无力场方程 $\vec{j} \times \vec{B} = \vec{0}$ ，是否存在解析解或数值解依然没有定论^[3,51–53]。幸运的是，Low & Lou 方程^[54]（将在第3章详细介绍）提供了一类轴对称无力场模型的半解析解。Low & Lou 方程是一类非线性二阶常微分方程。传统经典数值方法通常使用规则的数值网格来求解常微分方程，该网格上的节点位置和数量是固定的。传统方法通常只提供计算节点上的值。然而，有时候需要获得非计算网格节点上的函数值。因此，我们提出了一种通用的二阶微分方程数值算法，该算法能够在定义域上的任意位置提供函数值，以改进传统经典方法无法直接提供这些值的不足。另外，第3章还需要涉及第2章中向量分析、逼近定理和机器学习的基本知识。

在本文的第1章中，我们主要简要介绍了太阳的相关现象和概念，并通过等离子体方程组中的动量方程来铺垫日冕磁场外推。具体地说，在满足一定条件的情况下，通过对方程(1.11)进行简化，我们可以得到静磁学方程 $-\nabla p + \vec{j} \times \vec{B} + \vec{F} = \vec{0}$ 和无力场方程 $\vec{j} \times \vec{B} = \vec{0}$ 。我们将在第3章第1节中讨论这些简化形式所需要的条件。

在前人的基础上，本论文主要进行了以下改进：

1. 提出参数打靶法来确定 Low & Lou 方程(3.24)中的参数 a 。提出一种通用的二阶微分方程机器学习求解数值算法，使得该算法可以在定义域上任意点提供一类无力场半解析解(公式3.23)的函数值，以改进传统方法只能直接提供

计算节点上的函数值。

2. 提出一种数据驱动的数值算法，通过直接或间接给定 F 的值来计算 Low & Low 方程 (3.24) 中参数 n 和 a 的数值。

3. 通过万能逼近定理 (定理 2.5)，建立无力场神经网络外推法的数学模型 (公式 4.11)。在无力场神经网络外推法中，为了遵循物理规律并保持其特性，提出了一种在神经网络内部进行尺度变换的机器学习神经网络结构 (如图 4.1)，以改进在传统神经网络外进行的归一化和逆归一化操作。这样做可以确保输出结果仍然符合数据所遵循的物理规律。在无力场神经网络外推法中，采用二阶优化方法-拟牛顿法 (算法 2) 来训练神经网络。相比传统机器学习中的一阶优化方法，这种方法更高效，可以加快训练过程。

第2章 神经网络磁场外推法的数学基础

本章将介绍论文所需的基本知识，包括向量分析、逼近定理和机器学习的基本概念和原理。这些知识将为后续章节的研究工作提供必要的理论基础。

2.1 向量分析的基本知识

在太阳磁场研究中，磁场被视为三维空间中的向量场，因此向量分析在磁场研究中扮演着重要的角色。然而，在常见的教材或文献中，通常只在三维空间直角坐标系或某种特殊的标准正交基下给出了梯度、散度和旋度的定义。为了提供一个统一的观点，本章在一般的标准正交基下给出梯度、散度和旋度的定义，并详细说明它们之间的联系（定理 2.3）。我们所提出的相关概念适用于所有标准正交基下的梯度、散度和旋度计算，而不局限于直角坐标系、圆柱坐标系和球坐标系。这使得我们能够方便地使用向量分析的方法来研究和分析磁场以及其他向量场的性质。例如，在势场模型的解析解和一类轴对称无力场对应的微分方程推导过程中都涉及到定理 2.3 的应用。

在三维空间 \mathbb{R}^3 中，位置向量 \vec{r} 可以表示为 $\vec{r} = x\vec{i} + y\vec{j} + z\vec{k}$ ，其中 $\vec{i} = (1, 0, 0)$, $\vec{j} = (0, 1, 0)$, $\vec{k} = (0, 0, 1)$ 是直角坐标系的基向量。考虑一般的曲线坐标 u_1, u_2, u_3 ，它与直角坐标 x, y, z 之间的关系为

$$x = x(u_1, u_2, u_3), y = y(u_1, u_2, u_3), z = z(u_1, u_2, u_3).$$

根据全微分形式的不变性，我们有

$$d\vec{r} = \frac{\partial \vec{r}}{\partial u_1} du_1 + \frac{\partial \vec{r}}{\partial u_2} du_2 + \frac{\partial \vec{r}}{\partial u_3} du_3 = h_1 du_1 \vec{e}_1 + h_2 du_2 \vec{e}_2 + h_3 du_3 \vec{e}_3,$$

其中 $h_i = \left| \frac{\partial \vec{r}}{\partial u_i} \right|$, $\vec{e}_i = \frac{1}{h_i} \frac{\partial \vec{r}}{\partial u_i}$, h_i 是单位基向量 \vec{e}_i 的归一化因子, $i = 1, 2, 3$ 。若 $\vec{e}_i \cdot \vec{e}_j = \delta_{ij}$ 则称 $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ 是 \mathbb{R}^3 中的标准正交基。

首先，我们引入方向导数的概念，然后给出梯度的定义。我们考虑函数 f 从 (u_1, u_2, u_3) 到 $(u_1 + \Delta u_1, u_2 + \Delta u_2, u_3 + \Delta u_3)$ 的变化量 Δf 。根据多元微积分的知识， Δf 可表示为

$$\Delta f = \frac{\partial f}{\partial u_1} \Delta u_1 + \frac{\partial f}{\partial u_2} \Delta u_2 + \frac{\partial f}{\partial u_3} \Delta u_3 + \dots. \quad (2.1)$$

接着，考虑位置向量 \vec{r} 的变化量 $\Delta\vec{r}$ ，则我们有

$$\Delta\vec{r} = h_1\Delta u_1\vec{e}_1 + h_2\Delta u_2\vec{e}_2 + h_3\Delta u_3\vec{e}_3 + \dots.$$

然后，我们考虑 $\Delta\vec{r}$ 对应的弧长 Δs ，

$$\Delta s = \sqrt{(h_1\Delta u_1)^2 + (h_2\Delta u_2)^2 + (h_3\Delta u_3)^2} + \dots, \quad (2.2)$$

结合 (2.1) 和 (2.2) 式，我们有 $\frac{\Delta f}{\Delta s} = \left(\frac{1}{h_1} \frac{\partial f}{\partial u_1} \vec{e}_1 + \frac{1}{h_2} \frac{\partial f}{\partial u_2} \vec{e}_2 + \frac{1}{h_3} \frac{\partial f}{\partial u_3} \vec{e}_3 \right) \cdot \frac{\Delta\vec{r}}{\Delta s} + \dots$ 。当 $\Delta s \rightarrow 0$ 时，我们有

$$\lim_{\Delta s \rightarrow \infty} \frac{\Delta f}{\Delta s} = \frac{df}{ds} = \left(\frac{1}{h_1} \frac{\partial f}{\partial u_1} \vec{e}_1 + \frac{1}{h_2} \frac{\partial f}{\partial u_2} \vec{e}_2 + \frac{1}{h_3} \frac{\partial f}{\partial u_3} \vec{e}_3 \right) \cdot \frac{d\vec{r}}{ds}$$

其中 $\frac{d\vec{r}}{ds} = h_1 \frac{du_1}{ds} \vec{e}_1 + h_2 \frac{du_2}{ds} \vec{e}_2 + h_3 \frac{du_3}{ds} \vec{e}_3$ ，是沿着弧长方向的单位向量。于是我们把 $\frac{df}{ds}$ 定义 f 在切向量 $\frac{d\vec{r}}{ds}$ 方向的方向导数。

定义 2.1 (梯度). f 在一般标准正交基下的梯度被定义为

$$\nabla f = \frac{1}{h_1} \frac{\partial f}{\partial u_1} \vec{e}_1 + \frac{1}{h_2} \frac{\partial f}{\partial u_2} \vec{e}_2 + \frac{1}{h_3} \frac{\partial f}{\partial u_3} \vec{e}_3, \quad (2.3)$$

其中 $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ 是 \mathbb{R}^3 中的任意一组标准正交基， h_1, h_2, h_3 对应单位基向量的归一化因子。

定义 2.2. 设 \vec{F} 是 \mathbb{R}^3 中的向量场， V 是 \mathbb{R}^3 中逐片光滑的闭集， \vec{n} 为 ∂V 朝外法向量，记 $\int_{\partial V} \vec{F} \cdot \vec{n} dS$ 为 \vec{F} 在 V 上的通量。

定义 2.3 (散度). 在一般标准正交基下，设 P 是 \vec{F} 定义域 V 中给定的点，在 \vec{P} 点附近任取适当的邻域 D ，如果极限 $\lim_{D \rightarrow \vec{P}} \frac{1}{|D|} \int_D \vec{F} \cdot \vec{n} dS$ 存在，那么则称此极限为 \vec{F} 在 \vec{P} 点的散度，其中 $|D|$ 表示 D 的体积。散度用 $\nabla \cdot \vec{F}(\vec{P})$ 表示，即

$$\nabla \cdot \vec{F}(\vec{P}) = \lim_{D \rightarrow \vec{P}} \frac{1}{|D|} \int_D \vec{F} \cdot \vec{n} dS.$$

根据散度的定义，我们可以直接得到在一般标准正交基下的散度计算公式：

$$\nabla \cdot \vec{F} = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial u_1} (F_1 h_2 h_3) + \frac{\partial}{\partial u_2} (F_2 h_3 h_1) + \frac{\partial}{\partial u_3} (F_3 h_1 h_2) \right], \quad (2.4)$$

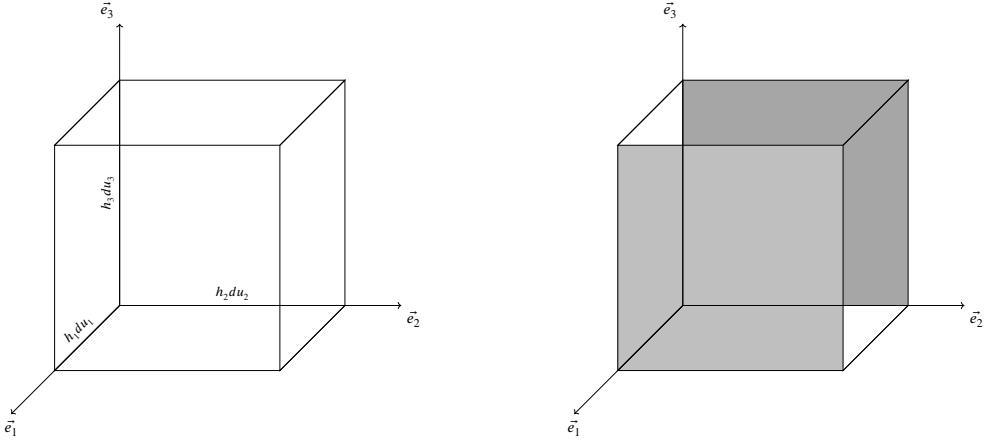


图 2.1 散度计算示意图

其中 h_1, h_2, h_3 对应单位基向量的归一化因子。下面给出由散度的定义推导出这一计算公式的简要过程。考虑以 $\vec{e}_1, \vec{e}_2, \vec{e}_3$ 为三边 (弯曲空间中) 的小长方体中，在该长方体上计算其边界即六个面的通量。首先考虑朝外法向量为 $-\vec{e}_1$ 的面，如图 2.1 所示，这是长方体的背面。在很小的区域内可假设向量 $\vec{F} = F_1 \vec{e}_1 + F_2 \vec{e}_2 + F_3 \vec{e}_3$ 为通过 (u_1, u_2, u_3) 点的向量 $\vec{F}(u_1, u_2, u_3)$ ，则我们有

$$\iint \vec{F} \cdot \vec{n} dS_1 \approx -F_1 [(u_1, u_2, u_3) (h_2 h_3)] du_2 du_3, \vec{F} \approx \vec{F}(u_1, u_2, u_3), \vec{n} = -\vec{e}_1. \quad (2.5)$$

类似地，考虑朝外法向量为 \vec{e}_1 的面，则我们有

$$\iint \vec{F} \cdot \vec{n} dS_1 \approx F_1 (u_1 + du_1, u_2, u_3) (h_2 h_3) du_2 du_3, \vec{F} \approx \vec{F}(u_1 + du_1, u_2, u_3), \vec{n} = \vec{e}_1 \quad (2.6)$$

结合 (2.5) 和 (2.6) 式，我们可以得到 \vec{e}_1 方向的通量

$$\begin{aligned} & F_1 (u_1 + du_1, u_2, u_3) (h_2 h_3) du_2 du_3 - F_1 (u_1, u_2, u_3) (h_2 h_3) du_2 du_3 \\ & \approx \frac{\partial}{\partial u_1} [F_1 h_2 h_3 (u_1, u_2, u_3)] du_1 du_2 du_3. \end{aligned}$$

类似地，我们可以得到 \vec{F} 在 \vec{e}_2 和 \vec{e}_3 方向的通量

$$\vec{e}_2 \text{ 方向: } \frac{\partial}{\partial u_2} (F_2 h_3 h_1) du_1 du_2 du_3, \quad \vec{e}_3 \text{ 方向: } \frac{\partial}{\partial u_3} (F_3 h_1 h_2) du_1 du_2 du_3.$$

因此 \vec{F} 在六个面的通量之和为

$$\iint_S \vec{F} \cdot \vec{n} dS \approx \left(\frac{\partial}{\partial u_1} (F_1 h_2 h_3) + \frac{\partial}{\partial u_2} (F_2 h_3 h_1) + \frac{\partial}{\partial u_3} (F_3 h_1 h_2) \right) du_1 du_2 du_3,$$

令 $du_1 du_2 du_3 \rightarrow 0$ ，则我们可以得到一般标准正交坐标基下的散度公式 (2.4)。

定理 2.1 (高斯散度定理^[55]). 若 \vec{F} 是定义在 V 的一阶连续实值函数, 其边界 ∂V 为逐段光滑的曲面, 则

$$\int_V \nabla \cdot \vec{F} d\vec{x} = \int_{\partial V} \vec{F} \cdot \vec{n} dS, \quad (2.7)$$

其中 \vec{n} 是 ∂V 的单位法向量, dS 是 ∂V 的面积元。

定义 2.4. 设 \vec{F} 为 \mathbb{R}^3 定义在开邻域 Ω 上的向量场, \vec{F} 在 Ω 上一阶连续, $\Gamma \subset \Omega$ 是片段光滑封闭曲线。记 $\int_{\Gamma} \vec{F} \cdot \vec{T} ds$ 为 \vec{F} 在 Γ 上的循环量, 其中 \vec{T} 是 Γ 上单位切向量。

定义 2.5 (旋度). 设 \vec{F} 为 \mathbb{R}^3 定义在开邻域 Ω 上的向量场, \vec{F} 在 Ω 上一阶连续, 若 $\vec{x}_0 \in \Omega$, 给定方向 \vec{n} , 过 \vec{x}_0 以 \vec{n} 为法向作平面 E , 在 \vec{x}_0 点附近取适当的邻域 $S \subset E$, 记 $\lim_{|S| \rightarrow x_0} \frac{1}{|S|} \int_{\partial S} \vec{F} \cdot \vec{T} ds$, 其中 $|S|$ 是 S 的面积。若此极限存在, 则称之为 \vec{F} 在 \vec{x}_0 点以 \vec{n} 为轴的旋度, 并记为

$$\nabla \times \vec{F} \cdot \vec{n} = \lim_{|S| \rightarrow x_0} \frac{1}{|S|} \int_{\partial S} \vec{F} \cdot \vec{T} ds, \quad (2.8)$$

其中 \vec{T} 是 ∂S 上的单位切向量。

根据旋度的定义, 我们可以直接得到在一般标准正交坐标基下的旋度计算公式:

$$\nabla \times \vec{F} = \frac{1}{h_1 h_2 h_3} \begin{vmatrix} h_1 \vec{e}_1 & h_2 \vec{e}_2 & h_3 \vec{e}_3 \\ \frac{\partial}{\partial u_1} & \frac{\partial}{\partial u_2} & \frac{\partial}{\partial u_3} \\ F_1 h_1 & F_2 h_2 & F_3 h_3 \end{vmatrix}, \quad (2.9)$$

其中 $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ 是 \mathbb{R}^3 中的任意一组标准正交基, h_1, h_2, h_3 对应单位基向量的归一化因子。这个公式可以通过散度的定义推导得出。考虑以原点为起点, (弯曲空间中) 长宽高分别为 $h_1 du_1$ 、 $h_2 du_2$ 和 $h_3 du_3$ 的小长方体。按定义 2.5, 我们需要计算向量 $\vec{F} = F_1 \vec{e}_1 + F_2 \vec{e}_2 + F_3 \vec{e}_3$ 在这六个面的循环量。首先考虑以旋转轴 \vec{e}_1 一面边界 $ABCD$ 的线积分, 如图 2.2 所示, 我们有

$$\int \vec{F} \cdot \vec{t} ds \approx F_2(u_1, u_2, u_3) h_2(u_1, u_2, u_3) du_2, \quad \vec{t} = \vec{e}_2, \quad (2.10)$$

$$\int \vec{F} \cdot \vec{t} ds \approx F_2(u_1, u_2, u_3 + du_3) h_2(u_1, u_2, u_3) du_2, \quad \vec{t} = -\vec{e}_2, \quad (2.11)$$

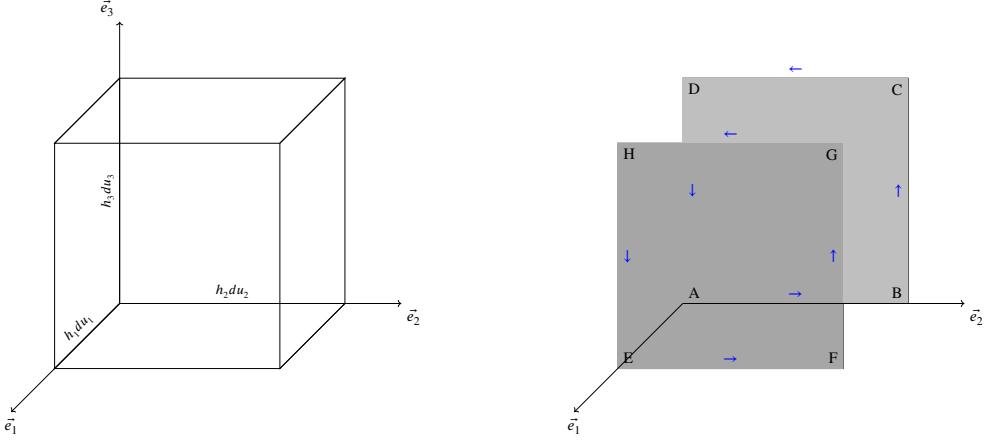


图 2.2 旋度计算示意图

合并上述两式，则由均值定理得 $\left(-\frac{\partial}{\partial u_3} F_2 h_2\right) du_2 du_3$ 。同理，以 \vec{e}_1 为旋转轴对应另一面边界 $EFGHE$ 的线积分为 $\left(-\frac{\partial}{\partial u_2} F_3 h_3\right) du_2 du_3$ ，则

$$\nabla \times \vec{F} \cdot \vec{e}_1 = \frac{1}{h_2 h_3} \left(\frac{\partial}{\partial u_2} (F_3 h_3) - \frac{\partial}{\partial u_3} (F_2 h_2) \right). \quad (2.12)$$

同理可得 \vec{e}_2, \vec{e}_3 方向的分量，

$$\begin{aligned} \nabla \times \vec{F} \cdot \vec{e}_2 &= \frac{1}{h_1 h_3} \left(\frac{\partial}{\partial u_3} (F_1 h_1) - \frac{\partial}{\partial u_1} (F_3 h_3) \right), \\ \nabla \times \vec{F} \cdot \vec{e}_3 &= \frac{1}{h_1 h_2} \left(\frac{\partial}{\partial u_1} (F_2 h_2) - \frac{\partial}{\partial u_2} (F_1 h_1) \right). \end{aligned} \quad (2.13)$$

因此，一般标准正交坐标基下的旋度公式可以通过拉普拉斯降阶法的形式给出，即公式 (2.9)。

定理 2.2 (斯托克斯旋度定理^[55])。设 \vec{F} 为 \mathbb{R}^3 定义在开邻域 Ω 上的一阶连续向量场， $S \subset \Omega$ 为光滑曲面， $\vec{x} \in S$ ， $\vec{n} = \vec{n}(\vec{x})$ 为 S 在 \vec{x} 点朝上单位法向量，则

$$\int_S \nabla \times \vec{F}(\vec{x}_0) \cdot \vec{n} dS = \int_{\partial S} \vec{F} \cdot \vec{T} ds, \quad (2.14)$$

其中 \vec{T} 是 ∂S 上的单位切向量。

注。高斯散度定理和斯托克斯旋度定理实际上是流形上斯托克斯定理的不同表现形式。为了进行严格的数学证明，可以参考微分流形文献^[55,56]。

在一般标准正交基下，梯度、散度和旋度具有以下关系：

定理 2.3 (Rudin^[57])。若 V 是 \mathbb{R}^3 中的开集， u 定义在 E 上有二次连续导数的实值函数， \vec{G} 是 E 上二次连续的向量场，则

1. 若 $\vec{F} = \nabla u$, 则 $\nabla \times \vec{F} = 0$ 。

2. 若 $\vec{F} = \nabla \times \vec{G}$, 则 $\nabla \cdot \vec{F} = 0$ 。

反之, 若 V 是二次连续可微函数的凸定义域, \vec{F} 是 V 上连续的向量场, 则

1'. 若 $\nabla \times \vec{F} = \vec{0}$, 则存在 u 使得 $\vec{F} = \nabla u$ 。

2'. 若 $\nabla \cdot \vec{F} = 0$, 则存在 \vec{G} 使得 $\vec{F} = \nabla \times \vec{G}$ 。

证明. 见文献《Principles of Mathematical Analysis》^[57] 中第 281 页定理 10.43 \square

在研究向量场时, 梯度、散度和旋度都是一阶微分算子。然而, 有时候我们需要使用二阶微分算子来描述向量场的性质。下面我们将通过散度算子作用在梯度算子上定义二阶微分算子-拉普拉斯算子。

定义 2.6 (拉普拉斯算子). 若 $\vec{r} = u_1\vec{e}_1 + u_2\vec{e}_2 + u_3\vec{e}_3$ 其中 $f(\vec{r}) \in \mathbb{R}$ 具有二阶导数。

定义拉普拉斯算子 $\Delta = \nabla \cdot \nabla$ 如下:

$$\Delta = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial u_1} \left(\frac{h_2 h_3}{h_1} \frac{\partial}{\partial u_1} \right) + \frac{\partial}{\partial u_2} \left(\frac{h_3 h_1}{h_2} \frac{\partial}{\partial u_2} \right) + \frac{\partial}{\partial u_3} \left(\frac{h_1 h_2}{h_3} \frac{\partial}{\partial u_3} \right) \right], \quad (2.15)$$

其中 $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ 是 \mathbb{R}^3 中的任意一组标准正交基, h_1, h_2, h_3 对应单位基向量的归一化因子。

在研究向量场的性质时, 有时候我们需要进行坐标变换。下面的定理将给出一般标准正交基经可逆变换后向量的投影变换公式:

定理 2.4. 给定任意位置向量 $\vec{r} = u_1\vec{e}_1 + u_2\vec{e}_2 + u_3\vec{e}_3 \in \mathbb{R}^3$ 。对于向量 $\vec{B}(u_1\vec{e}_1 + u_2\vec{e}_2 + u_3\vec{e}_3) = B_1\vec{e}_1 + B_2\vec{e}_2 + B_3\vec{e}_3$, 其中 B_1, B_2 和 B_3 分别称向量 \vec{B} 在 \vec{e}_1, \vec{e}_2 和 \vec{e}_3 下的投影。若存在可逆映射 $u_i = u_i(u_4, u_5, u_6)$ $i = 1, 2, 3$ 。则存在一组极大线性无关组 $\{\vec{e}_4, \vec{e}_5, \vec{e}_6\}$ 使得

$$u_1\vec{e}_1 + u_2\vec{e}_2 + u_3\vec{e}_3 = u_4\vec{e}_4 + u_5\vec{e}_5 + u_6\vec{e}_6, \quad (2.16)$$

和

$$\vec{B} = B_1\vec{e}_1 + B_2\vec{e}_2 + B_3\vec{e}_3 = B_4\vec{e}_4 + B_5\vec{e}_5 + B_6\vec{e}_6, \quad (2.17)$$

成立, 而且

$$\begin{bmatrix} B_1 & B_2 & B_3 \end{bmatrix} = \begin{bmatrix} B_4 & B_5 & B_6 \end{bmatrix} \begin{bmatrix} \frac{1}{h_4} \frac{\partial u_1}{\partial u_4} & \frac{1}{h_4} \frac{\partial u_2}{\partial u_4} & \frac{1}{h_4} \frac{\partial u_3}{\partial u_4} \\ \frac{1}{h_5} \frac{\partial u_1}{\partial u_5} & \frac{1}{h_5} \frac{\partial u_2}{\partial u_5} & \frac{1}{h_5} \frac{\partial u_3}{\partial u_5} \\ \frac{1}{h_6} \frac{\partial u_1}{\partial u_6} & \frac{1}{h_6} \frac{\partial u_2}{\partial u_6} & \frac{1}{h_6} \frac{\partial u_3}{\partial u_6} \end{bmatrix}. \quad (2.18)$$

证明. 见附录 B。 □

2.2 逼近定理的基本知识

在磁流体静力学模型及其简化的无力场模型中，我们假设磁场 \vec{B} 是位置向量 \vec{r} 的函数。日冕磁场外推的本质是求解偏微分方程的数值解问题。我们希望能够找到磁场 \vec{B} 的解析解，即以显式的形式给出方程的解。然而，在大多数情况下，微分方程很难或者无法直接求得解析解。因此，我们可以转而使用数值方法来近似求解微分方程。

数值解方法的基本思想是将连续的问题离散化，即将定义域划分为离散的网格点，并通过计算在这些网格点上的函数值来逼近微分方程的解。在这个过程中，我们需要选择一组基函数，它们能够在离散的点上近似表示微分方程的解空间。近似的过程即为逼近的过程，因此我们在这一小节讨论函数逼近定理的相关基本知识。

定理 2.5 (万能逼近定理^[58,59]). $K \subseteq \mathbb{R}^d$ 中的紧集， $f : K \rightarrow \mathbb{R}$ 上的连续函数， $\rho : \mathbb{R} \rightarrow \mathbb{R}$ 中的非多项式连续函数。那么，对任意 $\epsilon > 0$ ，存在 $N \in \mathbb{N}, a_k, b_k \in \mathbb{R}, w_k \in \mathbb{R}^d$ 使得

$$\left\| f - \sum_{k=1}^N a_k \rho(\langle w_k, \cdot \rangle - b_k) \right\|_{\infty} < \epsilon.$$

注. 当 $f : K \rightarrow \mathbb{R}^m$ 时，万能逼近定理仍然成立。

注. 当磁场 \vec{B} 是关于位置向量 \vec{r} 的连续函数时，根据万能逼近定理 ??，我们可以使用非多项式函数任意逼近 \vec{B} 。

2.3 机器学习的基本知识

多层感知器 (Multilayer Perceptron, MLP) 是机器学习 (Machine Learning) 中的一种重要神经网络 (Neural Network) 实现形式，它通过建立多层神经元连接来模拟人类神经系统，实现对数据的学习和预测。

多层感知器包括输入 $\vec{x} \in \mathbb{R}^{n_1 \times 1}$ 和最终的输出 $\vec{a} \in \mathbb{R}^{n_L \times 1}$ ，还包含中间隐藏层。第 l 层隐藏层包含权重、偏置和激活函数。其中，权重 $W^{[l]} \in \mathbb{R}^{n_l \times n_{l-1}} (l \geq 2)$ 表示第 l 层与第 $l-1$ 层之间的连接权重矩阵，偏置 $\vec{b} \in \mathbb{R}^{n_l \times 1} (l \geq 2)$ 是第 l 层的

偏置向量, σ 是激活函数, 通常为非线性变换。最后, 输出由 $\vec{a}^{[L]}$ 表示, 其网络结构如图 2.3 所示, 其数学表达式分别为:

1. 输入层:

$$\vec{a}^{[1]} = \vec{x}.$$

2. 隐藏层:

$$\vec{a}^{[l]} = \sigma \left(W^{[l]} \vec{a}^{[l-1]} + \vec{b}^{[l]} \right), \quad l = 2, 3, \dots, L - 1.$$

3. 输出层:

$$\vec{a}^{[L]} = \vec{z}^{[L]} = \left(W^{[L]} \vec{a}^{[L-1]} + \vec{b}^{[L]} \right).$$

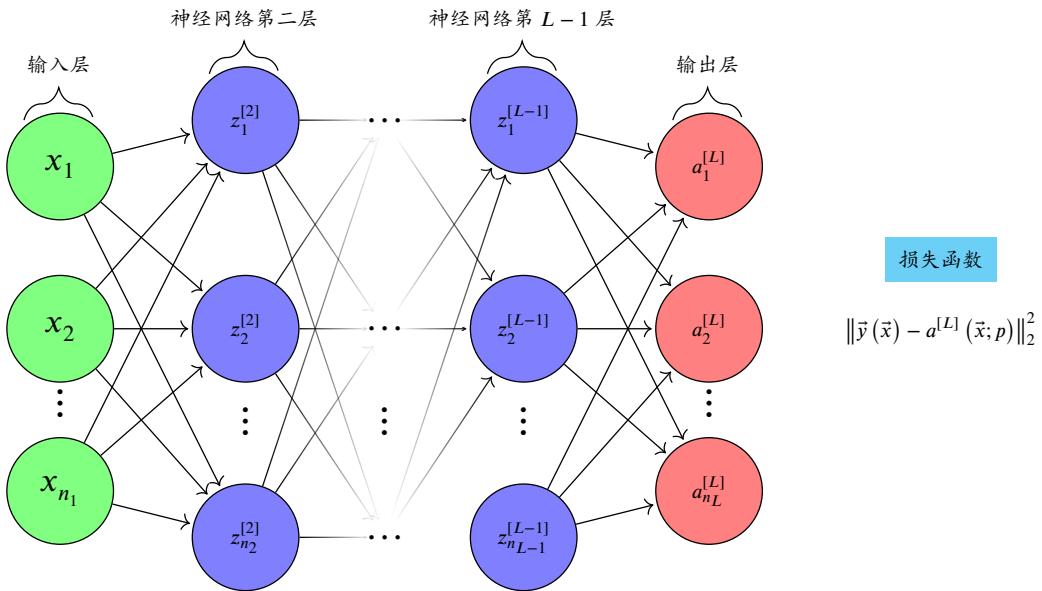


图 2.3 神经网络结构示意图

神经网络的训练需要训练集和目标集。训练集为 $\{\vec{x}^{(i)}\}_{i=1}^N$ (其中 $\vec{x}^{(i)} \in \mathbb{R}^{n_1 \times 1}$), 对应的目标集 $\{\vec{y}(\vec{x}^{(i)})\}_{i=1}^N$ (其中 $\vec{y}^{(i)} \in \mathbb{R}^{n_L \times 1}$)。在神经网络中需要定义损失函数 $L(\vec{p})$, 最简单的损失函数定义如下

$$L(\vec{p}) = \frac{1}{N} \sum_{i=1}^N \frac{1}{2} \|\vec{y}(\vec{x}^{(i)}) - a^{[L]}(\vec{x}^{(i)}; \vec{p})\|_2^2, \quad (2.19)$$

其中 \vec{P} 是 $W^{[2]}, \dots, W^{[L]}, \vec{b}^{[2]}, \dots, \vec{b}^{[L]}$ 组成的列向量。

以 $L = 2$ 为例, 当

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, W^{[2]} = \begin{bmatrix} W_{11}^{[2]} & W_{12}^{[2]} \\ W_{21}^{[2]} & W_{22}^{[2]} \\ W_{31}^{[2]} & W_{32}^{[2]} \end{bmatrix}, \vec{b}^{[2]} = \begin{bmatrix} b_1^{[2]} \\ b_2^{[2]} \\ b_3^{[2]} \end{bmatrix} \quad (2.20)$$

时，则 \vec{p} 为

$$\vec{p} = \left[W_{11}^{[2]}, W_{21}^{[2]}, W_{31}^{[2]}, W_{12}^{[2]}, W_{22}^{[2]}, W_{32}^{[2]}, b_1^{[2]}, b_2^{[2]}, b_3^{[2]} \right]^T. \quad (2.21)$$

神经网络的优化通常使用梯度下降法 (Gradient Decent Method)。梯度下降法通过不断调整参数 \vec{p} 来最小化损失函数 $L(\vec{p})$ 。

$$\vec{p}^{k+1} = \vec{p}^k - \eta \frac{1}{N} \sum_{i=1}^N \nabla L(\vec{p}), \quad (2.22)$$

其中 η 是学习率 (步长), $-\frac{1}{N} \sum_{i=1}^N \nabla L(\vec{p})$ 是损失函数的负梯度方向。

当 $N \gg 1$ 时, 梯度下降法将消耗很多计算资源, 为了解决这一问题, 人们提出了随机梯度下降法 (Stochastic Gradient Descent Method, SGD), 随机选取训练集中的一个样本进行更新

1. 从 $\{1, 2, \dots, N\}$ 中随机选取一个数 i ,
2. 使用 $\vec{p}^{k+1} = \vec{p}^k - \eta \nabla L(\vec{x}^{(i)}; \vec{p}^k)$ 更新 \vec{p} .

或者使用最小批处理的随机梯度下降法 (Minibatch Stochastic Gradient Descent), 将训练集分割成小批次进行更新。

需要注意的是, 随机梯度下降法不能保证每次迭代损失函数 \mathcal{L} 都会下降, 但从统计学的角度来看, 整体上损失函数 \mathcal{L} 是在下降的。为了提高训练的效率和稳定性, 人们提出了一些基于一阶梯度的优化方法。其中, 动量 (Momentum) 梯度下降法通过引入动量项来加速梯度更新, 并且可以帮助跳出局部最优解。动量算法考虑了当前梯度和历史梯度的加权平均, 使得更新具有惯性, 可以更快地收敛到全局最优解。

自适应梯度 (AdaGrad) 下降法针对不同参数的梯度进行自适应调整, 将学习率分别对待每个参数进行缩放, 较大的梯度将有较小的学习率, 而较小的梯度将有较大的学习率。这样可以在训练早期对参数的学习率进行大幅度衰减, 后期则会有更小的学习率, 适应不同参数的变化范围。

均方根传递 (RMSProp) 梯度下降法是对 AdaGrad 方法的改进, 它通过引入衰减系数来平衡历史梯度和当前梯度对学习率的影响, 进一步提高了算法的稳定性。

自适应矩估计 (ADAM) 梯度下降法结合了动量法和 RMSProp 方法的优点, 同时考虑了梯度的一阶矩估计 (均值) 和二阶矩估计 (方差)。ADAM 算法具有较

好的收敛性能和鲁棒性，在实践中得到广泛应用。这些基于一阶梯度的优化方法，利用参数的一阶导数信息进行参数更新。它们在训练神经网络时可以加快收敛速度、提高稳定性，并且适应不同的学习率需求。

在进行梯度下降法时，需要计算梯度。反向传播算法 (Back Propagation) 是一种高效计算神经网络梯度的方法。具体地讲，如果记

$$\vec{z}^{[l]} = W^{[l]} \vec{a}^{[l-1]} + \vec{b}^{[l]}, \vec{a}^{[l]} = \sigma(\vec{z}^{[l]}), \delta_j^{[l]} = \frac{\partial L(\vec{p})}{\partial z_j^{[l]}}, l = 2, 3, \dots, L, \quad (2.23)$$

则下列结论成立。

定理 2.6 (Higham 等^[60]).

1. $\vec{\delta}^{[L]} = \sigma'(\vec{z}^{[L]}) \odot (\vec{a}^{[L]} - \vec{y}).$
2. $\vec{\delta}^{[l]} = \sigma'(\vec{z}^{[l]}) \odot ((W^{[l+1]})^T \vec{\delta}^{[l+1]}), l = 2, 3, \dots, L-1.$
3. $\frac{\partial L(\vec{p})}{\partial b_j^{[l]}} = \delta_j^{[l]}, l = 2, 3, \dots, L.$
4. $\frac{\partial L(\vec{p})}{\partial W_{jk}^{[l]}} = \delta_j^{[l]} a_k^{[l-1]}, l = 2, 3, \dots, L.$

其中 \odot 是哈达玛积，即对应位置按元素相乘。

注. 在计算 $\vec{\delta}^{[l]}$ 时，需要先计算 $\vec{\delta}^{[l+1]}$ 。这种反向计算梯度的方式，即从输出层反向逐层计算误差项，正是反向传播算法的名字来源。

机器学习与微分数值解之间存在密切的关系，以一阶常微分方程的初值问题为例，对于给定的一阶常微分方程初值问题，

$$\begin{aligned} \text{微分方程 : } & f\left(x, y, \frac{dy}{dx}\right) = 0, \\ \text{初值条件 : } & y_0 = y(x_0). \end{aligned} \quad (2.24)$$

其中 $x^i \in X$ 。

我们需要构建一个训练集和相应的目标集来进行机器学习方法的求解。训练集由一组数据点 $\{x^i\}_{i=1}^N$ 组成，而目标集则由函数值和导数值构成。具体而言，目标集中的每个元素包含以下两部分：

1. 当 $x^i \in X \setminus \{x_0\}$ 时，对应的目标值为 $f\left(x^i, y(x^i), \frac{dy}{dx}\Big|_{x^i}\right)$ 。这里的 f 表示微分方程本身的方程式，根据已知的函数值和导数值计算出的目标值。
2. 当 $x^i = x_0$ 时，对应的目标值为 $y_0 - y(x_0)$ 。这是根据给定的初值条件计算出的目标值，用于确保求解器满足初始条件。

通过极小化损失函数来求解微分方程求解器 \mathcal{NN} 的最优参数。损失函数可用常见的 L_2 范数定义：

$$L(\vec{p}) = \frac{1}{N} \sum_{i=1}^N \left[f\left(x^i, \mathcal{NN}(x^i; \vec{p}), \frac{d\mathcal{NN}(x; \vec{p})}{dx} \Big|_{x=x^i}\right) \mathbb{I}_{x^i \in X \setminus \{x_0\}} + (\mathcal{NN}(x^i; \vec{p}) - y_0) \mathbb{I}_{x^i = x_0} \right]^2 \quad (2.25)$$

其中 \vec{p} 是神经网络求解器的参数, \mathbb{I} 是示性函数, 即 $\mathbb{I}_{x^i \in X \setminus \{x_0\}} = \begin{cases} 0, & x^i \notin X \setminus \{x_0\}, \\ 1, & x^i \in X \setminus \{x_0\}. \end{cases}$

和 $\mathbb{I}_{x^i = x_0} = \begin{cases} 0, & x^i \neq x_0, \\ 1, & x^i = x_0. \end{cases}$ 通过最小化损失函数 $L(\vec{p})$, 我们可以利用机器学习方法

找到最优的参数 \vec{p} , 从而使得求解器能够准确地求解给定的一阶常微分方程初值问题。

为了计算 $\frac{\partial L}{\partial \vec{p}}$, 本文使用自动微分 (Automatic Differentiation, AD) 技术^[61], 而不是基于给定网格的数值微分或数值积分方法^[62]。自动微分是一组用于数值求导的技术。它利用符号规则进行微分, 但在特定的数值点上计算导数, 并且不会构造导数的符号表达式。自动微分是一种自动计算导数的强大工具, 特别适用于求解复杂算法和数学函数的导数^[61]。

在 Matlab 中, 可以使用 `dlgradient` 命令进行自动微分操作。自动微分的优势在于它能够高效地计算导数, 特别是对于复杂的算法和数学函数。与传统的数值微分方法相比, 自动微分通过利用符号规则进行微分计算, 避免了数值近似带来的误差和不稳定性。

2.4 本章小节

本章涵盖了论文所需的向量分析、函数逼近定理和机器学习的基本知识。

通过梯度、散度和旋度在一般标准正交基下的计算形式 (公式 2.3、2.4 和 2.9), 并结合定理 2.3, 我们可以得到势场模型解析解的表达式 (公式 3.9)、格拉德-沙弗拉诺夫方程 (公式 3.21), 进一步通过格拉德-沙弗拉诺夫方程推导出一类无力场半解析解满足的 Low & Lou 方程 (公式 3.24)。另外, 根据公式 2.3、2.4 和 2.9 我们可以直接得到直角坐标系、圆柱坐标系和球坐标系下具体的相关计算公式, 详细内容请参见附录 A。

此外, 根据函数逼近定理 (定理 2.5), 我们建立机器学习与求解微分方程的关系。该定理说明机器学习方法可以应用于求解微分方程, 为求解微分方程提供

一种新的数值求解方法。

第3章 一类日冕磁场半解析解的数值方法

由于目前日冕磁场观测技术尚未成熟，但日冕磁场对太阳物理的研究至关重要。为了描述日冕磁场，通常使用无力场模型^[3,63]。因此，我们可以通过经典的无力场模型外推方法来计算日冕磁场。

在评估各种无力场模型的外推方法时，理论上需要提供一组无力场模型的解析解或数值解作为参考磁场。然而，正如文献^[3,51–53]中指出的那样，目前还没有数学理论能够证明一般形式的无力场模型的解具有存在性和唯一性。幸运的是 Low 等 提供了一类轴对称无力场的半解析解（公式 3.17），其中 n 、 a 、 F 和 $\frac{dF}{d\mu}$ 需要满足二阶非线性常微分方程。

我们求解一类日冕磁场半解析解（公式 3.23）的流程如下：首先，固定参数 n ，我们使用参数打靶法来求解 Low & Low 方程（公式 3.24）中参数 a 的值。接着，固定参数 n 和 a ，我们提出使用机器学习算法来求解 Low & Low 方程（公式 3.24）的数值解，该算法可以在定义域上任意位置提供 F 和 $\frac{dF}{d\mu}$ 的值。然而，传统的经典数值方法在求解常微分方程时通常使用规则的数值网格，其中节点的位置和数量确定。传统经典的方法通常只能提供计算节点上的值，而无法直接提供其他位置的值。

此外，在本章中，我们还提出一种数据驱动的数值算法，通过直接或间接给定 F 的值来计算 Low & Low 方程（公式 3.24）中参数 n 和 a 的数值。

3.1 无力场模型刻画日冕磁场的合理性

首先，回顾第 1 章中等离子方程组中的动量方程（1.11）

$$\rho \frac{d\vec{v}}{dt} = \rho \frac{\partial \vec{v}}{\partial t} + \rho (\vec{v} \cdot \nabla) \vec{v} = -\nabla p + \vec{j} \times \vec{B} + \rho \vec{g} \vec{e}_z, \quad (1.11)$$

其中 p 是压力， $\vec{j} \times \vec{B}$ 是洛伦兹力， \vec{j} 是电流且 $\vec{j} = \frac{\nabla \times \vec{B}}{\mu}$ 其中 μ 是磁导率， \vec{B} 是磁场。

为了行文方便，记特征速度为 v_0 、特征时间为 t_0 、特征长度为 l_0 、特征磁场强度为 B_0 、特征压强为 p_0 及特征高度为 H 。

首先对以下物理量进行无量纲化： $\vec{v}^* = \frac{\vec{v}}{v_0}$, $\rho^* = \frac{\rho}{\rho_0}$, $t^* = \frac{t}{t_0}$, $x^* = \frac{x}{l_0}$, $y^* =$

$\frac{y}{l_0}, z^* = \frac{z}{l_0}$, 则方程 (1.11) 可改写为

$$\rho_0 \frac{v_0}{t_0} \left[\rho^* \frac{\partial \vec{v}^*}{\partial t^*} + (\vec{v}^* \cdot \nabla^*) \vec{v}^* \right] = -\nabla p + \frac{B_0^2}{\mu l_0} (\nabla^* \times \vec{B}^* \times \vec{B}^*) + \rho g \vec{e}_z, \quad (3.1)$$

其中 $\nabla^* = \begin{bmatrix} \frac{\partial}{\partial x^*} & \frac{\partial}{\partial y^*} & \frac{\partial}{\partial z^*} \end{bmatrix}^T$, $B_x^* = \frac{B_x}{B_0}$, $B_y^* = \frac{B_y}{B_0}$, $B_z^* = \frac{B_z}{B_0}$ 。

通过对上述方程进行无量纲化, 可以消除物理量的量纲, 并将问题转化为研究无量纲形式下的方程。这有助于简化问题和分析等离子体行为。

当 $\rho_0 \frac{v_0}{t_0} \ll \frac{B_0^2}{\mu l_0}$ 时, 可以忽略 $\rho_0 \frac{v_0}{t_0} \left[\rho^* \frac{\partial \vec{v}^*}{\partial t^*} + (\vec{v}^* \cdot \nabla^*) \vec{v}^* \right]$ 这一项。事实上, $\rho_0 \frac{v_0}{t_0} \ll \frac{B_0^2}{\mu l_0}$ 等价于 $l_0 \frac{v_0}{t_0} \ll \frac{B_0^2}{\mu \rho_0}$, 如果定义阿尔芬 (Alfvén) 速度为 $v_A = \frac{B_0}{\sqrt{\mu \rho_0}}$, 则 $l_0 \frac{v_0}{t_0} \ll \frac{B_0^2}{\mu \rho_0}$ 等价于 $v_0^2 \ll v_A^2$ (由于 $l_0 = v_0 t_0$)。因此, 上述内容可以简单地用数学语言描述为 $\rho_0 \frac{v_0}{t_0} \ll \frac{B_0^2}{\mu l_0} \Leftrightarrow l_0 \frac{v_0}{t_0} \ll \frac{B_0^2}{\mu \rho_0} \Leftrightarrow v_0^2 \ll v_A^2 \Rightarrow v_0^2 \ll v_A^2 \Leftrightarrow \rho_0 \frac{v_0}{t_0} \ll \frac{B_0^2}{\mu l_0}$

因此, 当 $v_0^2 \ll v_A^2$ 时, 方程 (3.1) 可以简化成

$$0 = -\nabla p + \frac{B_0^2}{\mu l_0} (\nabla^* \times \vec{B}^* \times \vec{B}^*) + \rho g \vec{e}_z. \quad (3.2)$$

这个简化后的方程对应于磁流体静力学方程 (Magnetohydrostatics), 即

$$-\nabla p + \left(\nabla \times \frac{\vec{B}}{\mu} \right) \times \vec{B} + \rho g \vec{e}_z = 0. \quad (3.3)$$

在这个简化后的方程中, 压强梯度 $-\nabla p$ 、洛伦兹力项 $\left(\nabla \times \frac{\vec{B}}{\mu} \right) \times \vec{B}$ 和重力项 $\rho g \vec{e}_z$ 之间达到了平衡。这个方程描述了一个在磁场和重力的共同作用下处于静力平衡的磁流体。

根据《Magnetohydrodynamics of the Sun》^[3] 一书中第 109 页的内容知, $-\frac{\partial p}{\partial z} = \rho g$. 通过引入无量纲变量 $p^* = \frac{p}{p_0}$, $z^{**} = \frac{z}{H}$, 方程 (3.2) 可以重写为

$$0 = -\frac{p_0}{l_0} \nabla^* p^* + \frac{B_0^2}{\mu l_0} (\nabla^* \times \vec{B}^* \times \vec{B}^*) - \frac{p_0}{H} \frac{\partial p^*}{\partial z^{**}} \vec{e}_z. \quad (3.4)$$

当 $\frac{p_0}{H} \ll \frac{p_0}{l_0}$ 时, 可以忽略 $-\frac{p_0}{H} \frac{\partial p^*}{\partial z^{**}} \vec{e}_z$ 这一项。由于 $\frac{p_0}{H} \ll \frac{p_0}{l_0} \Leftrightarrow l_0 \ll H$, 所以当 $l_0 \ll H$ 时, 方程 (3.4) 可简化为

$$0 = -\frac{p_0}{l_0} \nabla^* p^* + \frac{B_0^2}{\mu l_0} (\nabla^* \times \vec{B}^* \times \vec{B}^*). \quad (3.5)$$

当 $\frac{p_0}{l_0} \ll \frac{B_0^2}{\mu l_0}$ 时, 可以忽略方程中的 $-\frac{p_0}{l_0} \nabla^* p^*$ 这一项。事实上 $\frac{p_0}{l_0} \ll \frac{B_0^2}{\mu l_0} \Leftrightarrow \frac{\mu p_0}{B_0^2} \ll 1$. 同时, 定义气体压力与磁场的比值为 $\beta = \frac{2\mu}{B_0^2}$, 则 $\frac{p_0}{l_0} \ll \frac{B_0^2}{\mu l_0} \Leftrightarrow \beta \ll 1$ 。因

此当 $\beta \ll 1$ 时, 方程 (3.5) 可简化为

$$0 = \frac{B_0^2}{\mu l_0} \left(\nabla^* \times \vec{B}^* \times \vec{B}^* \right). \quad (3.6)$$

综上所述, 当 $v_0^2 \ll v_A^2$, $l_0 \ll H$ 和 $\beta \ll 1$ 时, 可以简化动量方程 (1.11) 为

$$\vec{j} \times \vec{B} = \vec{0}, \quad (3.7)$$

其中 $\vec{j} = \nabla \times \frac{\vec{B}}{\mu}$ 。上述方程中磁场的洛伦兹力为零, 故称其为无力场模型。

根据 Parnell 在文献 [64] 中指出, 在日冕活动区中, 特征速度通常远小于阿尔芬速度, 即 $v_0 = 30 \text{ km} \cdot \text{s}^{-1} \ll v_A = 2000 \text{ km} \cdot \text{s}^{-1}$; 特征高度 $H = 61 Mm$ 远大于常见冕环的特征长度; 压磁比 $\beta = 0.007 \ll 1$ 。故用无力场模型 (3.7) 可以刻画太阳活动区日冕磁场的分布。

3.2 日冕活动区中无力场模型的分类

在无力场模型中, 假设磁场没有产生电流, 即

$$\nabla \times \vec{B} = \vec{0}, \quad (3.8)$$

则被视作势场 (Potential Field) 模型, 其中磁场 \vec{B} 满足 $\nabla \cdot \vec{B} = 0$ 。

根据定理 2.3, 如果 $\nabla \times \vec{B} = \vec{0}$ 则可以将 \vec{B} 表示为 $\vec{B} = \nabla \Psi$ 。将其代入 $\nabla \cdot \vec{B} = 0$ 中得到拉普拉斯方程 (Laplace equation)

$$\Delta \Psi = 0, \quad (3.9)$$

该方程的边界条件为:

$$\frac{\partial \Psi}{\partial \vec{e}_z} = B_z(x', y', 0), \quad (3.10)$$

其中 $(x', y', 0)$ 是边界 (光球) 上的点, 这一类边值问题在数学上被称作诺伊曼边值问题 (Neumann Problem)。

人们通常使用傅里叶级数展开法和格林函数积分法来解决拉普拉斯方程的上述边值问题。在这里仅给出格林函数法的表达形式 [66],

$$\Psi(x, y, z) = \int \int B_z(x', y', 0) G(x, y, z, x', y') dx' dy', \quad (3.11)$$

其中 $G(x, y, z, x', y') = -\frac{1}{2\pi |\vec{r} - \vec{r}'|}$ 是格林函数, $|\vec{r} - \vec{r}'|$ 是点 (x, y, z) 和点 $(x', y', 0)$ 之间的欧式距离。

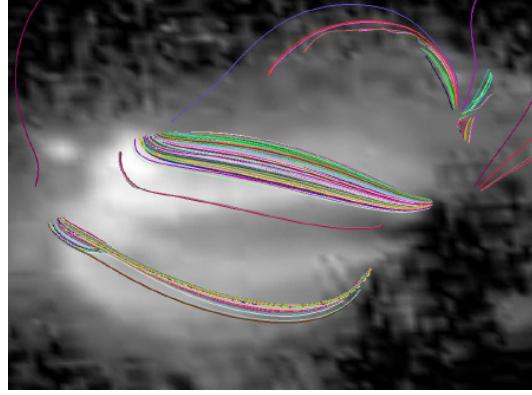


图 3.1 势场外推模型没有体现 Sigmoid 结构^[65]

需要注意的是，格林函数法中要求当 $|\vec{r}| \rightarrow \infty$ 时， $\Psi(\vec{r}) \rightarrow 0$ ^[67]。然而，事实上这只是一个必要条件。严格来说，上述表达式成立的充要条件是 Ψ 在上半空间的积分为有限值。

由 $\vec{B} = \nabla\Psi$ ，则磁场 \vec{B} 可表示成

$$\begin{aligned} B_x &= \frac{1}{2\pi} \iint \frac{(x - x') B_{z0}}{r^3} dx' dy', \\ B_y &= \frac{1}{2\pi} \iint \frac{(y - y') B_{z0}}{r^3} dx' dy', \\ B_z &= \frac{1}{2\pi} \iint \frac{z B_{z0}}{r^3} dx' dy', \end{aligned} \quad (3.12)$$

其中 $r = \sqrt{(x - x')^2 + (y - y')^2 + z^2}$ 。

虽然势场模型的解存在且唯一，计算方法简单，但是实际情况中电流是存在的。在一个外推的例子中^[65]，势场模型没有体现出 Sigmoid 结构，而实际上存在该结构，如图 3.1 所示。

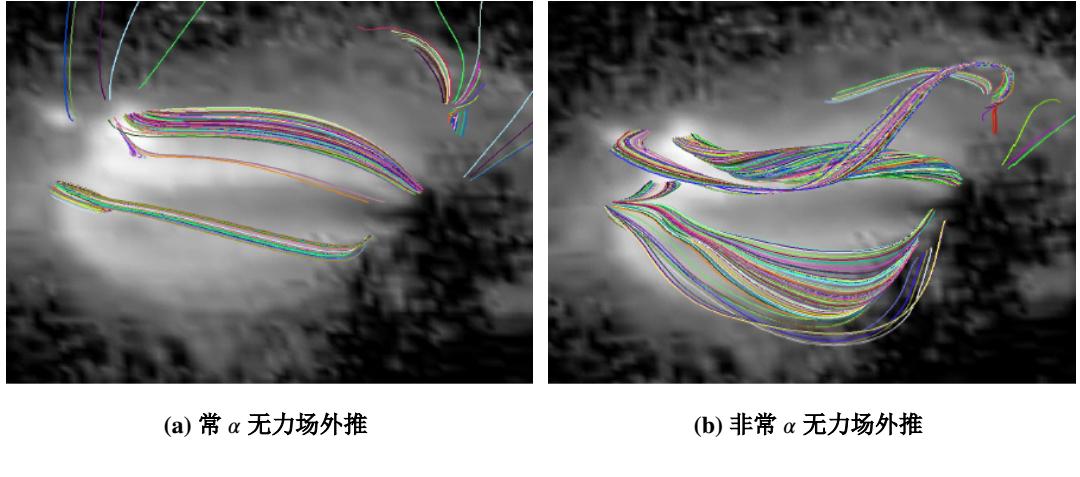
在无力场模型中，假设磁场的变化引起电流产生，即 $\nabla \times \vec{B} = \alpha \vec{B}$ ，如果 α 为常数（即 α 和空间中的位置无关），则被称作常 α 无力场模型。

对 $\nabla \times \vec{B} = \alpha \vec{B}$ 等式两边取旋度并展开，得到 $\nabla (\nabla \cdot \vec{B}) - \Delta \vec{B} = \alpha \nabla \times \vec{B}$ 。由于 $\nabla \cdot \vec{B} = 0$ ，则上式可化简为

$$(\Delta + \alpha^2) \vec{B} = \vec{0}, \quad (3.13)$$

该方程被称为亥姆霍兹方程 (Helmholtz Equation)，它是一个典型的椭圆型偏微分方程。它的边值条件是：

$$\vec{B} \cdot \vec{e}_z = B_z(x', y', 0) \quad \text{在光球表面上.} \quad (3.14)$$

图 3.2 无力场外推的例子^[68]

人们通常使用傅里叶级数法展开法^[30] 和格林函数积分法^[68] 来求解椭圆型偏微分方程的边值问题。格林函数积分法指的是以下形式的表达式

$$\vec{B}(x, y, z) = \frac{1}{2\pi} \int \int \left[\vec{G}_\alpha(x, y, z, x', y') B_z(x', y', 0) + \vec{\bar{G}}_\alpha(x, y, z, x', y') C(x', y') \right] dx' dy' \quad (3.15)$$

其中 $C(x', y')$ 是任意可积分函数， \vec{G}_α 和 $\vec{\bar{G}}_\alpha$ 是标准的格林函数。格林函数的分量表达式为：

$$\left(\frac{x - x'}{\tilde{R}} \frac{\partial \Gamma}{\partial z} + \alpha \Gamma \frac{y - y'}{\tilde{R}}, \frac{y - y'}{\tilde{R}} \frac{\partial \Gamma}{\partial z} - \alpha \Gamma \frac{x - x'}{\tilde{R}}, -\frac{\partial \Gamma}{\partial R} - \frac{\Gamma}{\tilde{R}} \right), \quad (3.16)$$

其中 \vec{G}_α 对应 $\Gamma = \frac{z}{\tilde{R}\tilde{r}} \cos(\alpha\tilde{r}) - \frac{1}{\tilde{R}} \cos(\alpha z)$ ，而 $\vec{\bar{G}}_\alpha$ 对应 $\Gamma = \frac{z}{\tilde{R}\tilde{r}} \sin(\alpha\tilde{r}) - \frac{1}{\tilde{R}} \sin(\alpha z)$ 。这里 \tilde{R} 和 \tilde{r} 分别是点 $(x', y', 0)$ 到点 $(x, y, 0)$ 和 (x, y, z) 的距离，即 $\tilde{R}^2 = (x - x')^2 + (y - y')^2$ 和 $\tilde{r}^2 = \tilde{R}^2 + z^2$ 。

然而，在常 α 无力场模型中解存在不唯一性，这是因为在使用格林函数积分法时， $C(x', y')$ 是任意可积分函数。而在实际的磁场里 α 是关于空间位置的函数。在一个实际的外推例子里，常 α 无力场模型外推的结果没有非常 α 无力场模型外推下的结果好^[65]，如图 3.2 所示。

在无力场模型中，当 α 不是常数时，称为非常 α 无力场模型。遗憾的是，关于非常 α 无力场模型的定解条件、解的存在性和唯一性问题尚未得到解决。然而，人们已经提出了一系列数值方法来求解这个模型。然而，在数学上评估这些

算法的性能、分析其稳定性和准确性是困难的。幸运的是，Low 等^[54] 在球坐标系 (r, θ, ϕ) 提供了一类半解析的非线性无力场。

3.3 Low & Lou 方程及其数值解

在球坐标系下，Low 等^[54] 建模轴对称 ($\frac{\partial}{\partial\phi} = 0$) 的非线性无力场，最终给出附加 $\frac{\partial}{\partial\phi} = 0$ 约束的非线性无力场表达式

$$\vec{B} = \frac{B_0}{R^3} \left[-\frac{dF}{d\mu} \vec{e}_r + \frac{F}{(1-\mu^2)^{\frac{1}{2}}} \vec{e}_\theta + \frac{\eta F^2}{(1-\mu^2)^{\frac{1}{2}}} \vec{e}_\phi \right], \quad (3.17)$$

其中 F 满足二阶非线性含参常微分方程

$$(1-\mu^2) \frac{d^2F}{d\mu^2} + n(n+1)F + a^2 \frac{1+n}{n} F^{1+\frac{2}{n}} = 0, \quad (3.18)$$

其中 $\mu = \cos\theta$ 。我们称方程 (3.18) 为 Low & Lou 方程，然而 Low 等^[54] 省略其数学建模过程。

下面，我们简要补充 Low & Lou 方程的数学建模过程。由于磁场满足 $\nabla \cdot \vec{B} = 0$ ，根据定理 2.3 存在 $\vec{A} = A_r \vec{e}_r + A_\theta \vec{e}_\theta + A_\phi \vec{e}_\phi$ 使得 $\vec{B} = \nabla \times \vec{A}$ 。结合 $\frac{\partial}{\partial\phi} = 0$ ，则 \vec{B} 可表示为

$$\vec{B} = \frac{1}{r \sin\theta} \left(\frac{1}{r} \frac{\partial}{\partial\theta} (r \sin\theta A_\phi) \vec{e}_r - \frac{\partial}{\partial r} (r \sin\theta A_\phi) \vec{e}_\theta + \sin\theta \left(\frac{\partial}{\partial r} (r A_\theta) - \frac{\partial A_r}{\partial\theta} \right) \vec{e}_\phi \right).$$

记 $\tilde{A} = r \sin\theta A_\phi$, $b_\phi = \sin\theta \left(\frac{\partial}{\partial r} (r A_\theta) - \frac{\partial A_r}{\partial\theta} \right)$, 则

$$\vec{B} = \frac{1}{r \sin\theta} \left(\frac{1}{r} \frac{\partial \tilde{A}}{\partial\theta} \vec{e}_r - \frac{\partial \tilde{A}}{\partial r} \vec{e}_\theta + b_\phi \vec{e}_\phi \right). \quad (3.19)$$

然后，将其代入 $\nabla \times \vec{B} = \alpha \vec{B}$ 中，利用 $(\nabla \times \vec{B})_r = (\alpha \vec{B})_r$ 和 $(\nabla \times \vec{B})_\theta = (\alpha \vec{B})_\theta$ 消去 α 后，有

$$\begin{vmatrix} \frac{\partial b_\phi}{\partial r} & \frac{\partial b_\phi}{\partial\theta} \\ \frac{\partial \tilde{A}}{\partial r} & \frac{\partial \tilde{A}}{\partial\theta} \end{vmatrix} = J(b_\phi, \tilde{A}) = 0. \quad (3.20)$$

我们需要使用函数相关定理^[69] 来建立 b_ϕ 和 \tilde{A} 之间的关系。函数相关定理的表述如下：

定理 3.1(函数相关定理). 若 f_1, f_2 是 $(x, y) \rightarrow \mathbb{R}$ 上的一阶连续映射，其雅可比矩阵对应行列式的值处处为零，则 $f_2(x, y) = H(f_1(x, y))$ 或 $f_1(x, y) = H(f_2(x, y))$ ，其中 H 是一阶连续函数。

证明. 见附录 C。 □

根据函数相关定理, 我们可以得到 $b_\phi = b_\phi(\tilde{A})$ 。接着, 对 b_ϕ 关于 r 求导数,

$$\frac{\partial b_\phi}{\partial r} = \frac{\partial b_\phi}{\partial \tilde{A}} \frac{\partial \tilde{A}}{\partial r} = \frac{db_\phi}{d\tilde{A}} \frac{\partial \tilde{A}}{\partial r}.$$

结合该式和 $(\nabla \times \vec{B})_\theta = (\alpha \vec{B})_\theta$ 式可以求得 $\alpha = \frac{db_\phi}{d\tilde{A}}$ 。将其代入 $(\nabla \times \vec{B})_\phi = (\alpha \vec{B})_\phi$ 等式后, 有

$$\frac{\partial^2 \tilde{A}}{\partial r^2} + \frac{\sin \theta}{r^2} \frac{\partial}{\partial \theta} \left(\frac{1}{\sin \theta} \frac{\partial \tilde{A}}{\partial \theta} \right) + \frac{d}{d\tilde{A}} \left(\frac{1}{2} b_\phi^2 \right) = 0, \quad (3.21)$$

该方程被称为格拉德-沙弗拉诺夫 (GS) 方程。在上述方程中, Low 等 作如下假设

$$\tilde{A} = \tilde{A}(r, \theta) = F(\cos \theta) r^{-n}, \quad b_\phi = a \tilde{A}^{1+\frac{1}{n}}, \quad (3.22)$$

其中 n 是给定的奇数, a 是未知的任意常数。则 GS 方程可简化为 Low & Lou 方程 (3.18)。最后, 把 (3.22) 代入到 (3.19) 式, 有

$$\vec{B} = \frac{1}{r^{(n+2)}} \left(-\frac{dF}{d\mu} \vec{e}_r + \frac{nF}{(1-\mu^2)^{\frac{1}{2}}} \vec{e}_\theta + \frac{aF^{(1+\frac{1}{n})}}{(1-\mu^2)^{\frac{1}{2}}} \vec{e}_\phi \right).$$

取 $n = 1$, 上式乘以 $B_0 r_0^3$ 后, 得到 Low & Lou 方程对应的非线性轴对称无力场 (3.17)。

在本章中, 将 Low & Lou 方程对应的流量函数 (flux function) $b_\phi(\tilde{A})$ 定义如下:

$$b_\phi(\tilde{A}) = \begin{cases} a \tilde{A}^{(1+\frac{1}{n})}, & \frac{2}{n} \text{ 为整数时,} \\ a \tilde{A} |\tilde{A}|^{(\frac{1}{n})}, & \text{其它情形,} \end{cases}$$

其中 n 是任意正实数。根据这个定义, 对应的轴对称无力场 \vec{B} 应更新为:

$$\vec{B} = \frac{1}{r^{(n+2)}} \left(-\frac{dF}{d\mu} \vec{e}_r + \frac{nF}{(1-\mu^2)^{\frac{1}{2}}} \vec{e}_\theta + \frac{aC_0}{(1-\mu^2)^{\frac{1}{2}}} \vec{e}_\phi \right), \quad (3.23)$$

其中 C_0 的表达式为:

$$C_0 = \begin{cases} F^{(1+\frac{1}{n})}, & \text{当 } \frac{2}{n} \text{ 是整数时,} \\ F |F|^{\frac{1}{n}}, & \text{其它情形时.} \end{cases}$$

同时，对应的 Low & Lou 方程应更新为：

$$(1 - \mu^2) \frac{d^2 F}{d\mu^2} + n(n+1)F + C = 0, \quad (3.24)$$

其中 C 的表达式为：

$$C = \begin{cases} \left(1 + \frac{1}{n}\right) a^2 F^{1+\frac{2}{n}}, & \text{当 } \frac{2}{n} \text{ 是整数时,} \\ \left(1 + \frac{1}{n}\right) a^2 F(F^2)^{\frac{1}{n}}, & \text{其它情形时.} \end{cases}$$

微分方程 (3.24) 是一类二阶含有参数 n 和 a 的非线性常微分方程，它的初边值条件与 Low & Lou 方程^[54] 初边值条件一致，即 $F(\pm 1) = 0, F'(-1) = 10$ 。

需要注意的是 $F^{1+\frac{2}{n}} = F(F^2)^{\frac{1}{n}}$ 并不恒成立，比如当 $F = -0.5, n = \frac{2}{9}$ 时， $F^{(1+\frac{2}{n})} = 9.7656 \times 10^{-4} > 0$ ，而 $F(F^2)^{\frac{1}{n}} = -9.7656 \times 10^{-4} < 0$ 。在本章中，我们取 $n = 5, 3, 1.5, 1, 0.9, 0.7, 0.5, 0.3, 0.1$ 。

3.3.1 固定参数 n 后求解 Low & Lou 方程中参数 a 的数值算法

求解二阶常微分方程的经典做法是先将其降阶成一阶常微分方程组，于是关于方程 (3.24)，引入中间变量 $f_1 = F(\mu), f_2 = F'(\mu)$ 。则方程 (3.24) 等价于微分方程组

$$f'_1 = f_2, \quad f'_2 = -\frac{n(n+1)f_1 + D}{1 - \mu^2}, \quad (3.25)$$

其中 $D = \begin{cases} \left(1 + \frac{1}{n}\right) a^2 f_1^{1+\frac{2}{n}}, & \frac{2}{n} \text{ 是正整数,} \\ \left(1 + \frac{1}{n}\right) a^2 f_1 (f_1^2)^{\frac{1}{n}}, & \text{其它情形.} \end{cases}$ 及初边值条件为 $f_1(\pm 1) = 0, f_2(-1) = 10$ 。

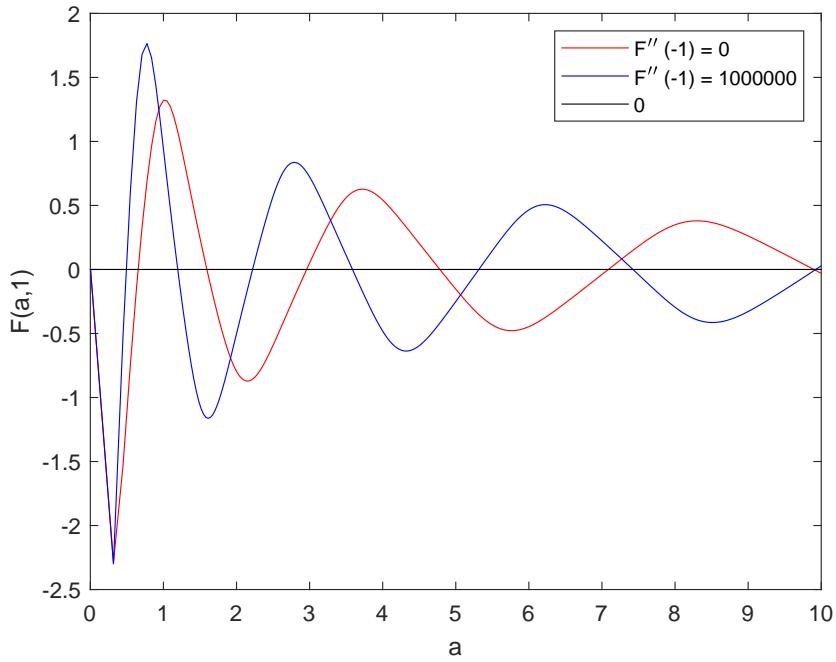
龙格库塔 (RK) 方法^[62,70] 是一种常见高效求解微分方程初值问题的数值方法。针对二阶微分方程等价的一阶微分方程组问题，它的四阶龙格库塔 (RK4) 数值格式如下

$$\vec{F}^{k+1} = \vec{F}^k + \frac{1}{6} \left[\vec{K}_1 + 2\vec{K}_2 + 2\vec{K}_3 + \vec{K}_4 \right],$$

其中 $\vec{K}_1, \vec{K}_2, \vec{K}_3$ 和 \vec{K}_4 的表达式分别为

$$\begin{aligned} \vec{K}_1 &= h \cdot \vec{f} \left(\mu_k, \vec{F}^k \right), & \vec{K}_2 &= h \cdot \vec{f} \left(\mu_k + \frac{1}{2}h, \vec{F}^k + \frac{1}{2}\vec{K}_1 \right), \\ \vec{K}_3 &= h \cdot \vec{f} \left(\mu_k + \frac{1}{2}h, \vec{F}^k + \frac{1}{2}\vec{K}_2 \right), & \vec{K}_4 &= h \cdot \vec{f} \left(\mu_k + h, \vec{F}^k + \vec{K}_3 \right) \end{aligned}$$

其中 $\vec{f} \left(\mu_k, \vec{F}^k \right) = \begin{bmatrix} f_{2,k} & -\frac{(1+\frac{1}{n})f_{1,k}+D_k}{1-\mu_k^2} \end{bmatrix}^T$ 、 $\vec{F}^k = \begin{bmatrix} f_{1,k} & f_{2,k} \end{bmatrix}^T$ 和 $\vec{F}^0 = \begin{bmatrix} 0 & 10 \end{bmatrix}^T$ 。

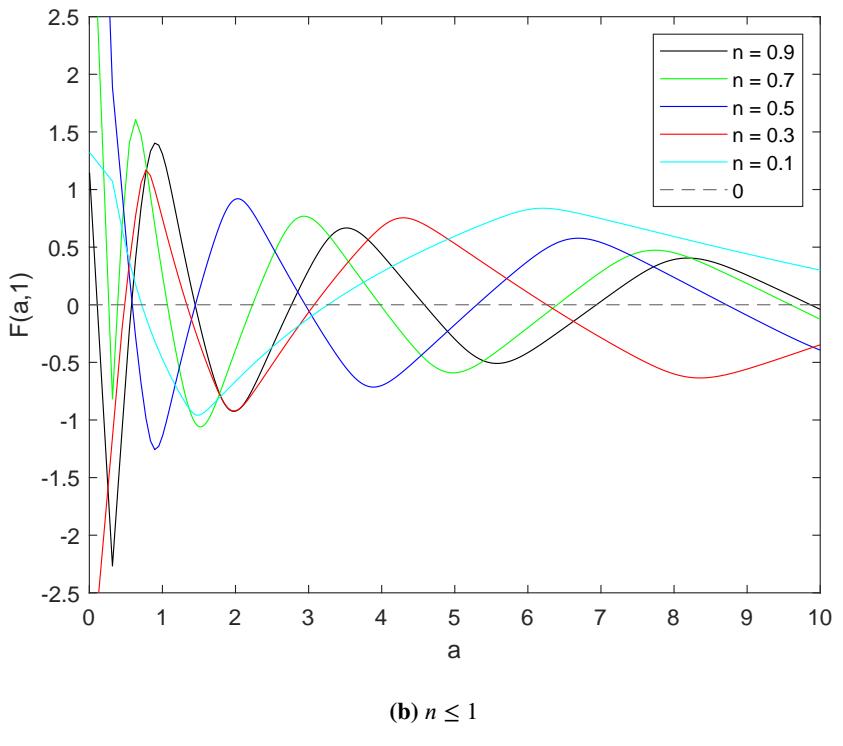
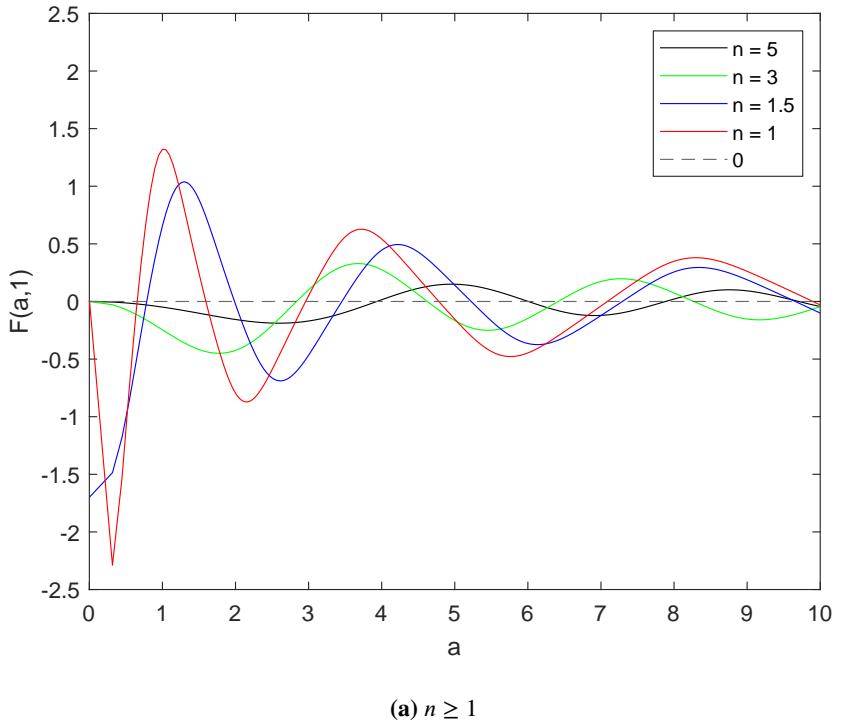
图 3.3 $f'_2(-1)$ 的取值对参数 a 的影响表 3.1 n 给定后 $F(a_i, \mu = 1)$ 的前三个零点

| m/n | 5 | 3 | 1.5 | 1 | 0.9 | 0.7 | 0.5 | 0.3 | 0.1 |
|-------|--------|--------|--------|--------|--------|--------|--------|--------|---------|
| 1 | 3.9341 | 2.8348 | 0.7846 | 0.6538 | 0.0719 | 0.0589 | 0.5865 | 0.4819 | 0.7162 |
| 2 | 6.0052 | 4.6243 | 1.9865 | 1.5947 | 0.5760 | 1.0678 | 1.4512 | 1.3332 | 3.2644 |
| 3 | 7.8851 | 6.4147 | 3.4555 | 2.9597 | 1.4474 | 2.2332 | 2.9533 | 3.0769 | 12.4581 |

需要注意的是 $\mu = -1$ 是 f'_2 的奇异点，故 $f'_2(-1)$ 的取值会影响 a 的数值。比如当 $n = 1$ 时，我们分别取 $f'_2(-1) = 0$ 和 1000000 ，我们把 F 关于 a 的函数呈现在图 3.3 中， F 与横轴的交点即为 a 的值，显然这两种情形下的交点不一致。在这一章中，我们取 $f'_2(-1) = 0$ 。

为了使用打靶法^[62] 确定参数 a 的数值解，我们可以仿照二阶常微分方程边值问题的一阶初值打靶法。首先确定打靶网格区间为 $[a_0, a_{max}]$ 和网格数 N ，这里取 $a_0 = 10^{-5}$, $a_{max} = 10$, $N = 1000$ 。然后通过 RK4 算法求得 $F(a_i, \mu = 1)$ 的值。接下来绘制以 $F(a_i, \mu = 1)$ 为纵轴、 a 为横轴的函数图像，如图 3.4 所示。通过观察这个函数图像以避免遗漏零点，并快速确定零点的大致范围。最后，我们通过二分法求得 $\{a_{n,m}\}$ ($1 \leq m \leq 3$)，参数 a 的数值结果列在表 3.1 中。

当参数 a 确定后，我们需要计算 F 和 F' 的值才能确定一类轴对称无力场磁

图 3.4 $F(a_i, \mu = 1)$ 零点分布图

场 (3.23) 的配置。

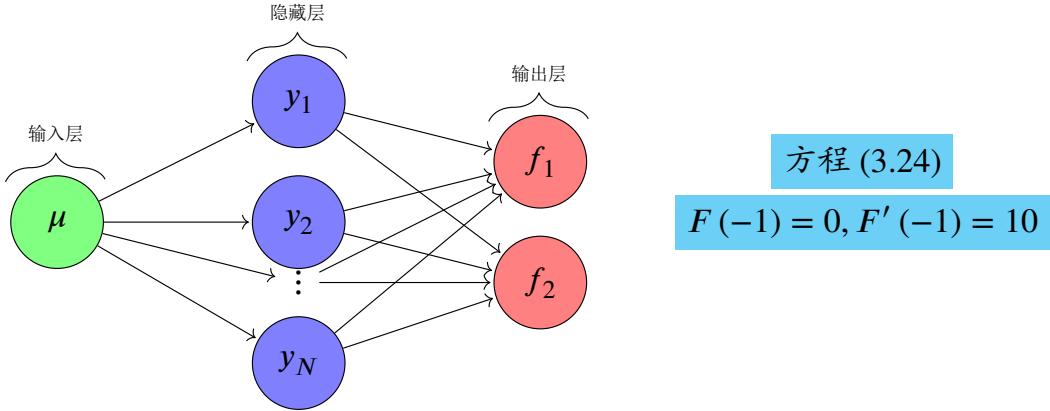


图 3.5 MLP 的网络结构示意图

3.3.2 固定参数 n 和 a 后求解 Low & Lou 方程的机器学习算法

常微分方程初始值问题的经典方法包括欧拉法和龙格库塔法，它们都是基于网格的数值方法。这些方法只能直接计算网格点上的值，而非网格点上的值通常需要通过插值获得。为了解决这个问题，我们可以使用基于神经网络的无网格方法来计算区间 $[-1, 1]$ 上任意点的值。此外为了避免类似经典数值方法出现奇点的问题，即 $\mu_0 = -1$ 时 $-\frac{(1+\frac{1}{n})f_{1,k}+D_k}{1-\mu_k^2}$ 分母出现了零，我们可以使用神经网络避免该奇点问题。

根据定理 2.5，我们使用含有三层全连接层的多层感知机求解微分方程 (3.24) 的初始值问题。采用多层感知器的结构如图 3.5，第一层全连接层对应输入 μ ，第二个全连接层是隐藏层有 N 个神经元，最后输出层分别输出 F 和 $\frac{dF}{d\mu}$ 。它的数学表达式为：

$$\vec{f}(\mu; \theta) = [f_1, f_2]^T = W_2 Y + b_2 = W_2 [\sigma(W_1 \mu + b_1)] + b_2, \quad (3.26)$$

其中 $W_1, b_1 \in \mathbb{R}^{N \times 1}, W_2 \in \mathbb{R}^{2 \times N}, b_2 \in \mathbb{R}^{2 \times 1}$ ， σ 是逐点双曲正切 (Tanh) 函数 $\text{Tanh}(x) = \frac{1-e^{-x}}{1+e^{-x}}$ 。

如果根据微分方程 (3.24) 及其初值条件，我们可以直接定义损失函数为：

$$\mathcal{L}(\mu; \theta) = \lambda_1 \mathcal{L}_1(\mu; \theta) + \lambda_2 \mathcal{L}_2(\mu; \theta), \quad (3.27)$$

其中 \mathcal{L}_1 和 \mathcal{L}_2 分别为

$$\mathcal{L}_1(\mu; \theta) = \frac{1}{M} \sum_{\mu=\mu_j, j=1}^{j=M} \left((1-\mu^2) \frac{d^2 f_1}{d\mu^2} + n(n+1) f_1 + E \right)^2,$$

$$\mathcal{L}_2(\mu; \theta) = (f_1(-1; \theta))^2 + (f_2(-1; \theta) - 10)^2,$$

$$\text{其中 } E = \begin{cases} \left(1 + \frac{1}{n}\right) a^2 f_1^{1+\frac{2}{n}}, & \text{当 } \frac{2}{n} \text{ 是整数时,} \\ \left(1 + \frac{1}{n}\right) a^2 f_1 (f_1^2)^{\frac{1}{n}}, & \text{其它情形.} \end{cases}$$

根据参考文献 Ryck 等 的研究^[71], 我们知道使用具有双曲正切函数作为激活函数的神经网络在一定条件下逼近连续函数及其一定阶导数 (广义导数)。显然, 通过直接定义上述损失函数, 并根据定理 2.5 进行训练, 不能保证 f_1 的导数连续。

我们考虑微分方程 (3.24) 的等价形式:

$$(1 - \mu^2) \frac{df_4}{d\mu} + n(n+1)f_3 + C' = 0, \quad f_4 = \frac{df_3}{d\mu}, \quad f_3 = F, \quad (3.28)$$

$$\text{其中 } C' = \begin{cases} \left(1 + \frac{1}{n}\right) a^2 f_3^{1+\frac{2}{n}}, & \frac{2}{n} \text{ 为正整数,} \\ \left(1 + \frac{1}{n}\right) a^2 f_3 (f_3^2)^{\frac{1}{n}}, & \text{其它情形.} \end{cases}$$

对应于微分方程组 (3.28) 及其初始值条件, 我们可以定义神经网络的损失函数为:

$$\text{Loss} = \mathcal{L}(\mu; \theta) = \lambda_1 \mathcal{L}_1(\mu; \theta) + \lambda_2 \mathcal{L}_2(\mu; \theta) + \lambda_3 \mathcal{L}_3(\mu; \theta), \quad (3.29)$$

其中 \mathcal{L}_1 、 \mathcal{L}_2 和 \mathcal{L}_3 分别为

$$\mathcal{L}_1(\mu; \theta) = \frac{1}{M} \sum_{\mu=\mu_j, j=1}^{j=M} \left((1 - \mu^2) \frac{df_2}{d\mu} + n(n+1)NN_1 + E \right)^2,$$

$$\mathcal{L}_2(\mu; \theta) = \frac{1}{M} \sum_{\mu=\mu_j, j=1}^M \left(\frac{df_1}{d\mu} - f_2 \right)^2,$$

$$\mathcal{L}_3(\mu; \theta) = (f_1(-1; \theta))^2 + (f_2(-1; \theta) - 10)^2,$$

$$\text{其中 } M \text{ 是训练样本数, } E \text{ 的表达式为 } E = \begin{cases} a^2 \left(1 + \frac{1}{n}\right) f_1^{\left(1+\frac{2}{n}\right)}, & \text{当 } \frac{2}{n} \text{ 是整数时,} \\ a^2 \left(1 + \frac{1}{n}\right) f_1 (f_1^2)^{\frac{1}{n}}, & \text{其它情形.} \end{cases}$$

在损失函数的表达式中 $\vec{\theta} \in \mathbb{R}^{(4N+2) \times 1}$ 是网络可学习的参数, \vec{p} 是重排列后 $\{W_1, b_1, W_2, b_2\}$ 的列向量。显然, 当损失函数(3.29) 下降到零时, 神经网络的解满足微分方程 (3.24) 及其初始值条件。此外, 根据万能逼近定理 2.5, 我们知道 f_1 的导数是连续的, 而一阶导数的连续性可以保证二阶导数的存在。

注. 在数值计算的例子中, 我们比较了通过极小化损失函数 (3.27) 和 (3.29) 获得的微分方程 (3.24) 的求解器。其数值结果如图 3.6 所示。通过数值结果的验证,

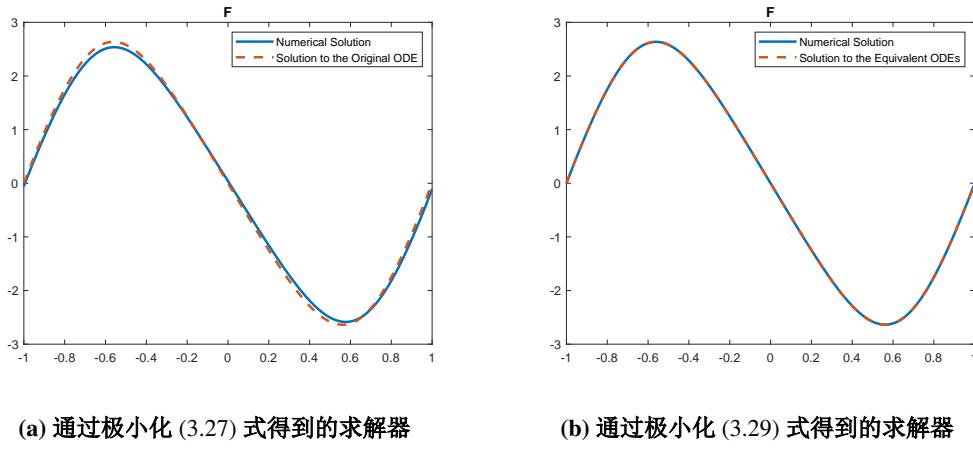


图 3.6 微分方程求解器的效果.

我们可以得出以下结论：直接定义损失函数 (3.27) 来求解微分方程 (3.24) 是不合理的，而通过 (3.29) 式定义的微分方程求解器不仅在理论上，也在实践中都能保证其数值结果的准确性和有效性。

在实际应用中，我们选择小批量的 ADAM 算法^[72] 来极小化损失函数 (3.29)，这个过程即为训练微分方程求解器的过程。我们在区间 $[-1, 1]$ 上随机选取了 10,000 个点作为训练样本，使用多层感知机进行训练。其中， β_1 和 β_2 的取值分别为 0.9 和 0.999， $\lambda_1, \lambda_2, \lambda_3, N_b$ 和 η 的设定如表 3.2 所示。最后，我们得到微分方程求解器的参数 $\vec{\theta}^* = \arg \min_{\vec{\theta}} \text{Loss}$ 。

表 3.2 MLP 中的超参数设置

| n | 5 | 3 | 1.5 | 1 | 0.9 | 0.7 | 0.5 | 0.3 | 0.1 |
|-------------|------|------|-------|------|-------|-------|------|-------|-------|
| λ_1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.005 |
| λ_2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.005 |
| λ_3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| N | 64 | 128 | 50 | 50 | 64 | 64 | 81 | 50 | 81 |
| N_b | 50 | 50 | 50 | 100 | 50 | 50 | 50 | 80 | 50 |
| η | 0.01 | 0.01 | 0.005 | 0.01 | 0.005 | 0.005 | 0.01 | 0.001 | 0.001 |

为了验证我们算法的有效性和可靠性，在 $\mu_i \in [-1, 1]$ 集合上均匀采样 $M = 10,000$ 个点，我们同时利用 RK4 算法计算 $F(\mu_i)$ 和 $\frac{dF}{d\mu}(\mu_i)$ 并记为 $F_{\text{RK}}(\mu_i)$ 和

算法 1 小批量处理的 ADAM 算法

输入: 初始一阶和二阶动量: $m_1 = 0, v_1 = 0$. 移动均值的衰减率: β_1, β_2

1: **while** 不收敛 **do**

2: 从 $\{\tilde{\mu}_1, \tilde{\mu}_2, \dots, \tilde{\mu}_{Nb}\}$ 中随机采样

3: 计算梯度: $\frac{\partial \tilde{L}}{\partial \theta} \Big|_{\theta=\theta_t}$, 其中 $\tilde{L} = \lambda_1 \tilde{L}_1 + \lambda_2 \tilde{L}_2 + \lambda_3 \tilde{L}_3$. \tilde{L}_1, \tilde{L}_2 和 \tilde{L}_3 分别为

$$\tilde{L}_1 = \frac{1}{Nb} \sum_{\tilde{\mu}=\tilde{\mu}_j, j=1}^{Nb} \left[(1 - \tilde{\mu}^2) \frac{dNN_2}{d\mu} + n(n+1) NN_1 + E \right]^2$$

$$\tilde{L}_2 = \frac{1}{Nb} \sum_{\tilde{\mu}=\tilde{\mu}_j, j=1}^{Nb} \left(\frac{dNN_1}{d\mu} - NN_2 \right)^2$$

$$\tilde{L}_3 = L_3$$

4: 更新一阶矩的估计: $m_{t+1} = \beta_1 m_t + (1 - \beta_1) \frac{\partial \tilde{L}}{\partial \theta} \Big|_{\theta=\theta_t}$

5: 更新偏置的二阶矩估计 $v_{t+1} = \beta_2 v_t + (1 - \beta_2) \left(\frac{\partial \tilde{L}}{\partial \theta} \Big|_{\theta=\theta_t} \right)^2$

6: 修正一阶矩估计: $\tilde{m}_{t+1} = \frac{m_{t+1}}{1 - \beta_1^t}$

7: 修正二阶据矩估计: $\tilde{v}_{t+1} = \frac{v_{t+1}}{1 - \beta_2^t}$

8: 更新 $\theta_{t+1} = \theta_t - \eta \frac{\tilde{m}_{t+1}}{\sqrt{\tilde{v}_{t+1} + \epsilon}}$

9: 更新 $t \leftarrow t + 1$

10: **end while**

输出: θ_{t+1}

$\frac{dF_{\text{RK}}}{d\mu}(\mu_i)$, 我们比较这两种算法的解。

当 $n = 5, 3, 1.5, 1, 0.9, 0.7, 0.5, 0.3$ 和 0.1 时, 图 3.7 展示了利用我们提出的机器学习算法和 RK4 算法求解 $F(\mu_i)$ 和 $\frac{dF}{d\mu}(\mu_i)$ 数值结果的接近程度, 可以看到它们非常接近, 几乎重合。

为了量化 MLP 和 RK 数值解之间的差异, 我们定义 $F_{\text{MLP}}(\mu_i)$ 与 $F_{\text{RK}}(\mu_i)$ 及 $\left(\frac{dF}{d\mu}_{\text{MLP}}\right)(\mu_i)$ 和 $\left(\frac{dF}{d\mu}_{\text{RK}}\right)(\mu_i)$ 的均方误差 (MSE):

$$\text{MSE}(F) = \frac{1}{10000} \sum_{i=1}^{10000} [F_{\text{MLP}}(\mu_i) - F_{\text{RK}}(\mu_i)]^2,$$

$$\text{MSE}\left(\frac{dF}{d\mu}\right) = \frac{1}{10000} \sum_{i=1}^{10000} \left[\frac{dF}{d\mu}_{\text{MLP}}(\mu_i) - \frac{dF}{d\mu}_{\text{RK}}(\mu_i) \right]^2,$$

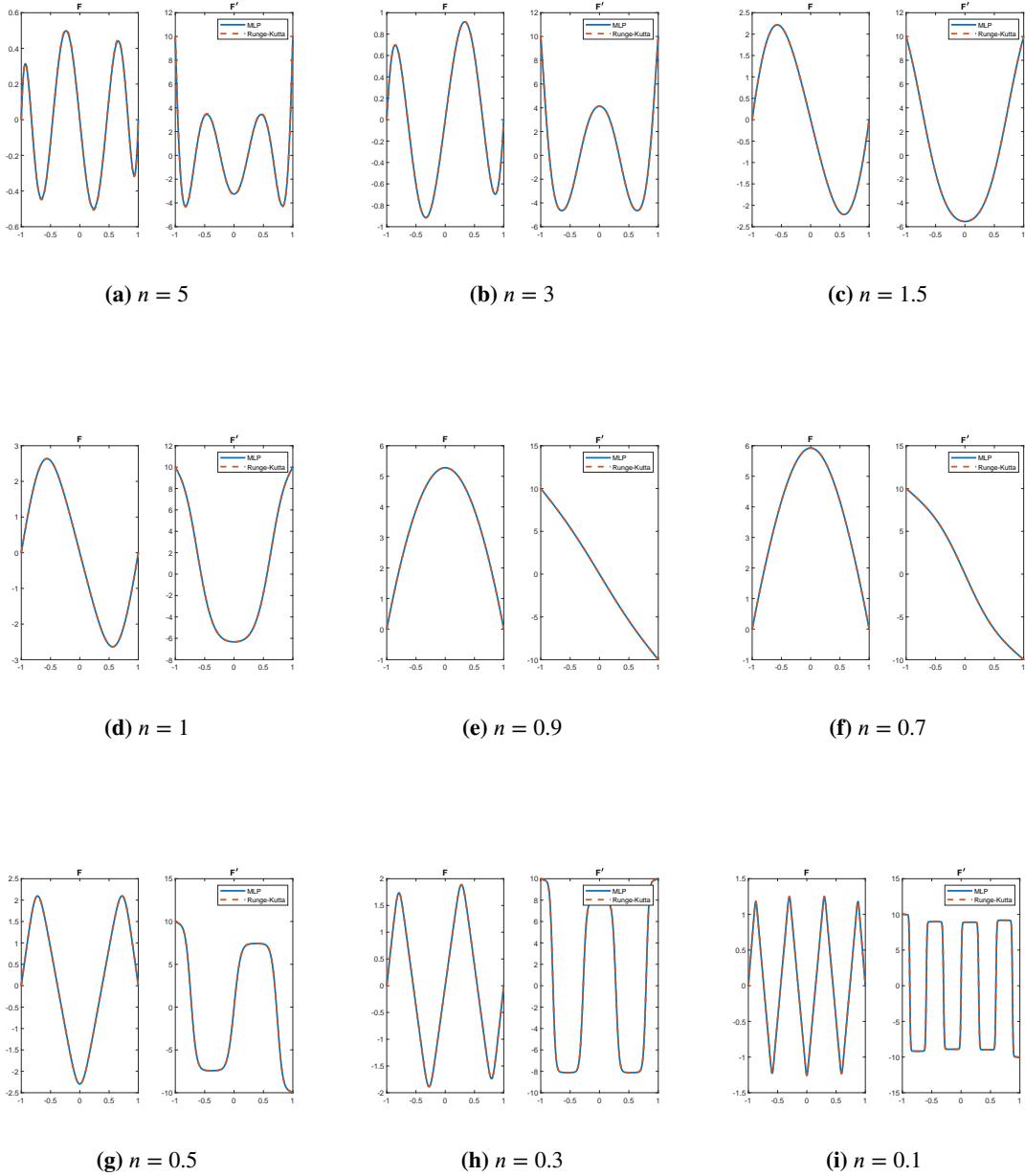
图 3.7 $F, \frac{dF}{d\mu}$ 分别由 MLP 和 RK 求解器产生

表 3.3 中列出了对应的均方误差。从表中的数值可以看出，基于机器学习的方法与经典的四阶龙格库塔方法具有一致的性能。此外，我们的方法还可以计算任意 $\mu \in [-1, 1]$ 对应的 $F(\mu)$ 和 $\frac{dF}{d\mu}$ ，而龙格库塔方法则不能。

本小节的目标是开发一种能够有效求解特定类别的常微分方程的数值算法，特别是针对无力场模型中 $\frac{\partial}{\partial \phi} = 0$ 所诱导的二阶非线性常微分方程。我们算法的一个显著特点是其出色的适应性，通过对损失函数 (3.29) 进行简单修改，可以求解各种初值和边界条件下的微分方程数值解。Lerche 等^[73] 对 Low & Lou 方程 (公式 3.18) 的初始条件进行了修改。在这种情况下，我们的数值算法仍然有效且

表 3.3 MLP 求解器的性能

| n | 5 | 3 | 1.5 | 1 |
|-------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| MSE(F) | 2.99×10^{-5} | 5.14×10^{-5} | 1.22×10^{-6} | 1.25×10^{-6} |
| $\text{MSE}(\frac{dF}{d\mu})$ | 2.70×10^{-3} | 2.58×10^{-4} | 2.61×10^{-5} | 1.39×10^{-5} |

| n | 0.9 | 0.7 | 0.5 | 0.3 | 0.1 |
|-------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| MSE(F) | 5.78×10^{-6} | 8.24×10^{-6} | 1.91×10^{-7} | 2.43×10^{-6} | 6.01×10^{-5} |
| $\text{MSE}(\frac{dF}{d\mu})$ | 5.91×10^{-6} | 2.84×10^{-6} | 2.22×10^{-6} | 4.34×10^{-5} | 2.82×10^{-2} |

适用。为了避免重复，我们把这部分内容放置于附录 D 中。

3.3.3 Low & Lou 方程机器学习数值解对应的无力场

为了得到近似观测到的太阳活动区磁场，Low 等^[54]进行了以下操作：在坐标系 (x, y, z) 中，沿 z 轴向下滑动距离 l 得到新的坐标系 (x', y', z') ，然后将 xoz 平面沿顺时针方向旋转角度 Φ 建立新坐标系 (X, Y, Z) 。为了得到 (x, y, z) 坐标系下的磁场信息，我们需要建立以下关系的数学表达式：

1. 建立坐标系 (X, Y, Z) 与之对应的球坐标系 (r, θ, ϕ) 之间的关系。
2. 建立坐标系 (x, y, z) 与 (r, θ, ϕ) 之间的变换关系。
3. 建立投影坐标 (B_x, B_y, B_z) 与 (B_r, B_θ, B_ϕ) 之间的变换关系。

根据定理 2.4，我们可以得到上述坐标之间的关系式和在其标准正交坐标基下投影的关系式。其中，坐标系的关系式如下：

$$r = \sqrt{x^2 + y^2 + 2x(z+l)\sin\Phi\cos\Phi + (z+l)\cos^2\Phi}, \quad (3.30)$$

$$\theta = \arccos \frac{x\sin\Phi + (z+l)\cos\Phi}{\sqrt{x^2 + y^2 + 2x(z+l)\sin\Phi\cos\Phi + (z+l)\cos^2\Phi}}, \quad (3.31)$$

$$\phi = \begin{cases} \arccos \frac{x\cos\Phi - (z+l)\cos\Phi}{\sqrt{x^2\cos^2\Phi + y^2}}, & y \geq 0, \\ 2\pi - \arccos \frac{x\cos\Phi - (z+l)\cos\Phi}{\sqrt{x^2\cos^2\Phi + y^2}}, & y < 0. \end{cases} \quad (3.32)$$

对应的投影坐标关系式如下：

$$\begin{bmatrix} B_x & B_y & B_z \end{bmatrix} = \begin{bmatrix} B_r & B_\theta & B_\phi \end{bmatrix} \begin{bmatrix} \cos \Phi \sin \theta \cos \phi + \sin \Phi \cos \theta & \sin \theta \sin \phi & -\sin \Phi \sin \theta \cos \phi + \cos \Phi \cos \theta \\ \cos \Phi \cos \theta \cos \phi - \sin \Phi \sin \theta & \cos \theta \sin \phi & -\sin \Phi \cos \theta \cos \phi - \cos \Phi \sin \theta \\ -\cos \Phi \sin \phi & \cos \phi & \sin \Phi \sin \phi \end{bmatrix}. \quad (3.33)$$

我们将上一节得到的 Low & Lou 方程 (3.24) 的数值解 $F(\mu)$ 和 $\frac{dF}{d\mu}$ 代入式 (3.23) 中，得到球坐标系 (r, θ, ϕ) 下在标准正交基下的投影坐标 (B_r, B_θ, B_ϕ) 。在空间直角坐标系 (x, y, z) 下选取计算区域，并根据上述坐标变换建立的关系 (3.30) - (3.33)，确定该计算区域的边界（或底边界）上的坐标及其对应的磁场向量作为磁场外推的边界条件。利用给定的边界条件，我们可以使用磁场外推方法推断计算区域内的磁场分布，并通过与 Low & Lou 方程数值解生成的数据进行比较。最终，我们可以评估磁场外推方法的准确性和可靠性。

然而，经坐标变换后，并非所有计算区域的边界都会落在传统方法（欧拉法或 RK4 法）划分的计算网格节点上。为了解决这个问题，传统上我们采取两种策略。第一种策略是通过插值数值结果来获取未落到计算网格节点的坐标所对应的函数值。具体而言，我们在划分的网络节点上计算数值解，并利用插值方法来获得非节点位置得函数值。通过利用已知节点处得数值结果，我们可以对非节点位置进行插值，从而估计其函数值以处理不在 μ_i 上的点。另一种策略是对坐标点进行多次划分网格节点的操作，以确保它们都能够落在划分的计算节点上。所以，在本文中，我们采用机器学习算法来求解 Low & Lou 方程，这样我们可以直接得到定义域上任意点的函数值，包括非计算网格节点上的函数值。

3.3.4 Low & Lou 方程中参数的数据驱动求解算法

在这一小节中，我们考虑了一个数学上的反问题：在给定光球上测量得到的磁场 \vec{B} 或经过测量算子或经过测量算子 P 映射到 $P\vec{B}$ ，是否可以通过机器学习的方法找到一个神经网络 $f(r_i, \mu_i; a, n)$ 来逼近观测到的光球磁场数据？其中参数 a 和 n 满足 Low & Lou 方程 (3.24)。

在球坐标系 (r, θ, ϕ) 下， \vec{B} 是满足 $\frac{\partial}{\partial \phi} = 0$ 的无力场。假设 $P = \begin{bmatrix} 0 & \frac{r^{n+2}(1-\mu^2)^{\frac{1}{2}}}{n} & 0 \end{bmatrix}$ ，

根据公式 (3.19), 我们有 $P\vec{B} = \begin{bmatrix} 0 & \frac{r^{n+2}(1-\mu^2)^{\frac{1}{2}}}{n} & 0 \end{bmatrix} \begin{bmatrix} B_r & B_\theta & B_\phi \end{bmatrix}^T = F$ 。

关于上述反问题, 神经网络的损失函数可以定义为

$$\mathcal{L} = \lambda_1 \mathcal{L}_1 + \lambda_2 \mathcal{L}_2,$$

其中 λ_1 和 λ_2 是平衡参数, \mathcal{L}_1 和 \mathcal{L}_2 的表达式分是

$$\mathcal{L}_1 = \frac{1}{M} \sum_{i=1}^M \left((1 - \mu_i^2) \frac{d^2 G}{d \mu^2} + \theta_2 (\theta_2 + 1) G + \theta_3 \left(1 + \frac{1}{\theta_2} \right) G (G^2)^{\frac{1}{\theta_2}} \right)^2,$$

和

$$\mathcal{L}_2 = \frac{1}{M} \sum_{i=1}^M (1 - \mu_i^2)^{\frac{1}{2}} \frac{r^{\theta_2+2}}{\theta_2} ((B_\theta)_i - NN(\mu_i; \theta_1))^2,$$

其中 M 训练集的采样点个数, G 的表达式是

$$G(\mu_i) = \frac{r^{n+2}}{n} (1 - \mu_i^2)^{\frac{1}{2}} \mathcal{NN}(\mu_i; \theta_1),$$

另外 $\vec{\theta}_1$ 是对应 B_θ 的可学习向量参数, θ_2 和 θ_3 分别对应 n 和 a 的可学习的参数。

通过一个简单的例子, 我们可以说明基于机器学习的方法可以用来求解上述反问题。假设我们选取满足微分方程 (3.24) 初值问题的数值解, 其中 n 和 a 分别为 5 和 3.9341 (记为 $na_1 = \{n = 5, a = 3.9341\}$), 那么在球坐标系下由 na_1 生成的磁场已知。

我们采用一个具有五层的多层感知机来实现上述神经网络, 其中输入层 r 、 θ 和 B_θ , 中间隐藏层各含有 64 个神经元, 输出层的输出包括对 a 、 n 和 B_θ 的预测值。我们选择 $\lambda_1 = \lambda_2 = 1$ 、 $M = 65,536$ 。在训练完成后, 多层感知机输出 a 和 n 的估计值分别为 3.9830 和 4.9042 (记为 $na_2 = \{n = 4.9042, a = 3.9830\}$)。

图 3.8 展示了归一化误差 $\frac{(B_{n2})_x - (B_{n1})_x}{\max|(B_{n1})_x|}$ 的可视化表示。在图 3.8 中, 最大的差异出现在靠近磁零点的地方, 这是因为当 r 太小时, 在公式 (3.17) 中误差传播迅速增加。根据文献^[74] 提供的相似度计算公式, 由 na_1 和 na_2 生成的磁图之间的相似度为 0.9909。相似度的取值范围在 -1 到 1 之间, 数值越接近 1 表示两个磁图之间的相似度越高。图 3.9 展示了由 na_1 和 na_2 生成的磁图 B_x 。图 3.10 显示了分别由 na_1 和 na_2 生成的 B_x 的等高线图, 它们十分吻合。

Low & Lou 方程在球坐标系下可以很好地刻画满足 $\frac{\partial}{\partial \phi} = 0$ 条件的无力场, 但在其他情况下其表现可能会有所差异。Liu 等^[75] 中的研究指出, 根据他们统

计的数据，大部分光球磁场并不满足无力场条件。因此，在使用我们提出的模型时，需要验证无力性条件和轴对称条件。如果不满足这些条件，则需要修改损失函数的正则项。然而，增加适当的正则项是十分困难的，这个问题将在我们未来的工作中继续研究。

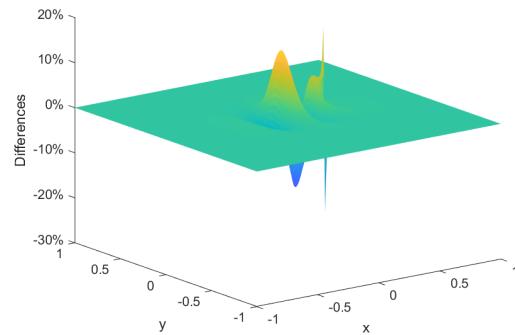


图 3.8 由 na_1 和 na_2 生成磁场的差异

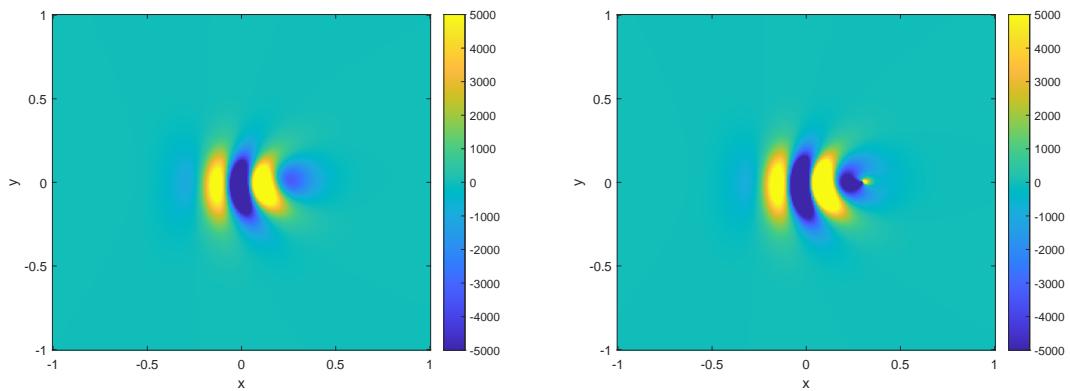
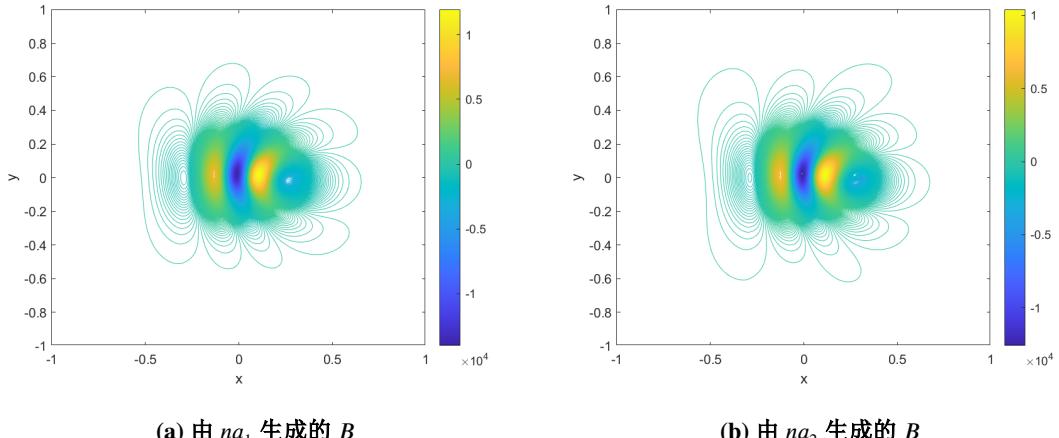


图 3.9 na_1 和 na_2 分别生成磁图 B_x

图 3.10 na_1 和 na_2 分别生成磁场的等高线

3.4 本章小节

本章首先通过无量纲分析给出使用无力场模型刻画日冕磁场的合理性，并简要介绍无力场模型的分类。接着给出 Low & Lou 方程 (公式 3.24) 的解的数值算法。求解 Low & Lou 方程数值解的求解过程包括：给定参数 n 后使用打靶法确定参数 a 的数值；然后利用机器学习算法计算 Low & Lou 方程在定义域 $[-1, 1]$ 上任意点处函数及其导数值。最后介绍了一种数据驱动方法来估计 Low & Lou 方程中的参数 n 和 a 。我们的算法是一种通用的求解二阶非线性常微分方程数值解的方法。

在第 4 章中，我们需要利用给定的边界条件，使用无力场神经网络外推法推断计算区域内的磁场分布。这里计算区域指的是在空间直角坐标系 (x, y, z) 下选定磁场外推的计算区域，而边界条件指的是该计算区域边界上的坐标 (x_i, y_j, z_k) 及其对应的磁场投影坐标 $(B_x(x_i, y_j, z_k), B_y(x_i, y_j, z_k), B_z(x_i, y_j, z_k))$ 。然而 Low & Lou 方程 (公式 3.24) 及其对应的磁场投影坐标形式 (公式 3.23) 建立在球坐标系 (r, θ, ϕ) 下，因此需要使用本章中的公式 (3.31) – (3.33) 进行坐标变换。通过将无力场神经网络外推法的外推结果与 Low & Lou 方程 (公式 3.24) 的数值解生成的磁场数据进行比较，最终评估机器学习磁场外推方法的准确性和可靠性。

第4章 神经网络磁场外推算法

日冕磁场的无力场模型本质上是通过偏微分方程进行描述，具体来说是 $\nabla \times \vec{B} = \alpha \vec{B}$ ，其中磁场 \vec{B} 满足 $\nabla \cdot \vec{B} = 0$ 。当 α 是位置向量的函数时，该模型被称为非线性无力场。尽管其数学形式非常简单，但对于该一般形式解的存在性和唯一性并没有得到证明^[3,51-53]，因此无法提供该偏微分方程的解析表达式。

根据经典非线性无力场外推模型的数学建模方式的不同，可以将其分为以下几种方法：Grad-Rubin (GR) 方法、竖直积分法、优化法、磁流体动力学松弛法和边界积分法。在经过数学建模后，这些经典非线性无力场外推方法的数学表达式对应于 4.1 节中的 (4.1)–(4.5) 式，除了边界积分法之外，它们都可以写成标准的偏微分方程形式 (公式 4.9)。需要注意的是，非线性无力场模型本身与时间无关，属于偏微分方程的边值问题。然而，优化法和磁流体动力学松弛法的建模过程引入时间变量，将磁场外推方法模型转化为求解偏微分方程的初边值问题，从而改变了微分方程的属性。

根据所对应的偏微分方程 (公式 4.1–4.5) 和积分方程 (公式 (4.6) 和 4.7) 的数值计算格式不同，可将其计算格式分类为有限差分、有限元、有限体积法、边界元和谱方法等。我们将在 4.1 节中详细讨论这些问题。

与上述数值格式不同，机器学习中神经网络使用的是自动微分技术，在该方法中空间中的每一个节点相对独立，不直接受周围点的影响。相比之下，传统的数值格式在空间中的点会受到周围点的影响，例如在简单有限差分格式 $\frac{\partial u}{\partial x} \approx \frac{u_{i+1,j,k} - 2u_{i,j,k} + u_{i-1,j,k}}{2\Delta x}$ 、 $\frac{\partial u}{\partial y} \approx \frac{u_{i,j+1,k} - 2u_{i,j,k} + u_{i,j-1,k}}{2\Delta y}$ 和 $\frac{\partial u}{\partial z} \approx \frac{u_{i,j,k+1} - 2u_{i,j,k} + u_{i,j,k-1}}{2\Delta z}$ 中，该数值格式受附近点的影响，在误差传播效应下，最终的解会差很多。

近年来，物理约束神经网络^[76,77]为科学和工程领域提供了一种新的方法，通过结合机器学习和数值方法的优势，探索和解决偏微分方程相关的问题。物理约束神经网络的出现为研究人员提供一个强大而灵活的工具，可以更好地理解和预测物理现象，并推动科学和工程领域的创新和发展，物理约束神经网络的本质仍是机器学习中的神经网络。我们将在 4.2 节讨论这些问题。

在本章中，我们提出的无力场机器学习神经网络外推法仍然是偏微分方程边值问题。然而经典的机器学习算法通常会对数据进行归一化处理，然后将其输

入神经网络进行训练。然而，在应用到无力场模型时，由于其本质是微分方程，对数据进行缩放可能导致不再满足原方程。举个例子来说，考虑一个一般的方程： $a^2 + 2a + 1 = 0$ ，其中 $a = -1$ 是方程的解。如果将 a 缩小两倍，我们得到 $\left(\frac{a}{2}\right)^2 + 2\frac{a}{2} + 1 = \frac{1}{4}$ ，然而，将 $\frac{1}{4}$ 扩大两倍却得到 $\frac{1}{2}$ 。显然， $\frac{1}{2}$ 并不等于 0。然而，文献^[78] 从理论的角度指出，在神经网络输入之前进行归一化操作的必要性，特别是与网络参数优化相关，在这里我们不作详细讨论。

在本章中，为了保持磁场数据的内在物理规律，我们提出无力场神经网络法进行磁场外推。在无力场神经网络外推方法中，我们通过在神经网络内部添加准输出层进行一次尺度变换（如图 4.1 所示），以避免在神经网络外部进行归一化操作。如果不进行归一化，损失函数将很难被有效地最小化。因此，为了克服无法进行归一化带来的优化困难，我们采用二阶优化方法—拟牛顿法来最小化损失函数，而不是使用经典机器学习中的一阶优化方法。在二阶优化方法中，每一步都使用二次函数来逼近原函数，相比一阶优化方法的梯度下降法，二阶优化方法通常更有效，收敛所需的迭代步数也较少^[79,80]。

4.1 经典无力场外推法中的数学形式

在非线性无力场模型中，借助向量磁场的无散条件 $\nabla \cdot \vec{B} = 0$ ，我们有以下关系：

$$\begin{aligned}\nabla \times (\nabla \times \vec{B}) &= \nabla (\nabla \cdot \vec{B}) - \Delta \vec{B} = -\Delta \vec{B}, \quad \nabla \times (\alpha \vec{B}) = \alpha (\nabla \times \vec{B}) = \alpha^2 \vec{B}, \\ \nabla \cdot (\nabla \times \vec{B}) &= 0, \quad \nabla \cdot (\alpha \vec{B}) = \nabla \alpha \cdot \vec{B} + \alpha \nabla \cdot \vec{B} = \nabla \alpha \cdot \vec{B}.\end{aligned}$$

当 α 固定时，我们有 $(\Delta + \alpha^2) \vec{B} = \vec{0}$ ，这是一个椭圆型偏微分方程。当 \vec{B} 固定时， $\nabla \alpha \cdot \vec{B} = 0$ ，则方程属于双曲型偏微分方程。因此，非线性无力场模型对应的方程组是典型的椭圆–双曲复合型偏微分方程。

经典的非线性无旋场外推方法包括 GR 法、竖直积分法、优化法、磁流体动力学松弛法和边界元积分法等。下面我们简要介绍一下这些方法。

GR 方法是一种交替迭代的方法。在固定 \vec{B} 的情况下，我们求解关于 α 的双曲型偏微分方程的边值问题：

$$\begin{aligned}&\text{微分方程: } \nabla \alpha \cdot \vec{B} = 0, \quad \text{在求解区域上,} \\ &\text{边值条件: } \alpha = \alpha_0, \quad \text{在边界上.}\end{aligned}\tag{4.1}$$

其中, 当 $B_z \neq 0$ 时, 根据 $(\nabla \times \vec{B})_z = (\alpha \vec{B})_z$ 可得到边界上的条件 $\alpha_0 = \frac{\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}}{B_z}$ 。此外, 初始磁场 \vec{B} 的值由势场外推模型提供。接下来, 固定 α , 求解关于 \vec{B} 的椭圆型偏微分方程的边值问题:

$$\begin{aligned} \text{微分方程: } & (\Delta + \alpha^2) \vec{B} = \vec{0}, \text{ 在求解区域上,} \\ \text{边值条件: } & B_z = B_{z_0}, \text{ 在边界上.} \end{aligned} \quad (4.2)$$

Amari 等 在文献^[28] 中指出, 他们利用有限元方法实现 GR 磁场外推方法。

竖直积分法是一种简单的积分方法。首先, 计算 $\nabla \times \vec{B} = \alpha \vec{B}$ 在 \vec{e}_x 和 \vec{e}_y 的投影坐标, 得到 $\frac{\partial B_y}{\partial z} = -\frac{\partial B_z}{\partial y} + J_x$, $\frac{\partial B_x}{\partial z} = \frac{\partial B_z}{\partial x} + J_y$, 其中 $J_x = \alpha B_x$, $J_y = \alpha B_y$ 。然后, 根据无散条件 $\nabla \cdot \vec{B} = 0$, 可以得到 $\frac{\partial B_z}{\partial z} = -\frac{\partial B_x}{\partial x} - \frac{\partial B_y}{\partial y}$ 。结合由光球磁场构成的初值条件, 首先得到了第一层 $((x_0, y_0, z)$ 点) 的三组常微分方程的初始值问题:

$$\begin{aligned} \text{微分方程 1: } & \frac{dB_x}{dz} = \frac{\partial B_{z_0}}{\partial x} + J_{y_0}, \quad \text{初始条件 1: } B_{x_0} = (\vec{B}(x_0, y_0, 0))_x. \\ \text{微分方程 2: } & \frac{dB_y}{dz} = -\frac{\partial B_{y_0}}{\partial y} + J_{x_0}, \quad \text{初始条件 2: } B_{y_0} = (\vec{B}(x_0, y_0, 0))_y. \\ \text{微分方程 3: } & \frac{dB_z}{dz} = -\frac{\partial B_{x_0}}{\partial x} - \frac{\partial B_{y_0}}{\partial y}, \quad \text{初始条件 3: } B_{z_0} = (\vec{B}(x_0, y_0, 0))_z. \end{aligned} \quad (4.3)$$

其中 $J_{y_0} = \alpha B_{y_0}$, $J_{x_0} = \alpha B_{x_0}$, $\alpha = \frac{J_{z_0}}{B_{z_0}} = \frac{\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}}{B_{z_0}}$ 。类似地, 沿着 \vec{e}_z 移动一小步, 并更新相应的微分方程和初始条件。整个过程通过对底边界向量磁场沿着 \vec{e}_z 方向进行积分逐层向上, 从而得到该点的磁场值。微分方程初值问题是竖直积分法的核心思想。

优化法是一种基于优化泛函的方法^[34], 其中优化泛函定义为:

$$L = \int \frac{\|J \times \vec{B}\|_F^2}{\|\vec{B}\|_F^2} + \|\nabla \cdot \vec{B}\|_F^2 dV.$$

通过对 L 进行求导, 可以得到 $\frac{dL}{dt} = -2 \int_V \frac{\partial \vec{B}}{\partial t} \cdot \vec{F}_v dV - 2 \int_{\partial V} \frac{\partial \vec{B}}{\partial t} \cdot \vec{F}_s dS$, 其中 \vec{F}_v 、 \vec{F}_s 和 $\vec{\Omega}$ 的表达式分别为

$$\begin{aligned} \vec{F}_v &= \nabla \times (\vec{\Omega} \times \vec{B}) - \vec{\Omega} \times (\nabla \times \vec{B}) - \nabla (\vec{\Omega} \cdot \vec{B}) + \vec{\Omega} (\nabla \cdot \vec{B}) + \vec{\Omega}^2 \vec{B}, \\ \vec{F}_s &= \vec{n} \times (\vec{\Omega} \times \vec{B}) - \vec{n} (\vec{\Omega} \cdot \vec{B}), \\ \vec{\Omega} &= \vec{B}^{-2} \left[(\nabla \times \vec{B}) \times \vec{B} - (\nabla \cdot \vec{B}) \vec{B} \right]. \end{aligned}$$

在计算 $\frac{\partial L}{\partial t}$ 时, 在边界上取 $\frac{\partial \vec{B}}{\partial t} = 0$, 而在求解区域 V 上取 $\frac{\partial \vec{B}}{\partial t} = \mu F_v (\mu > 0)$ 。这样, $\frac{dL}{dt} = - \int_V \mu F_v^2 dV < 0$, 说明 L 是关于 t 的单调递减函数。实际上, 优化法在求解偏微分方程的初边值问题:

$$\begin{aligned} \text{微分方程: } & \frac{\partial \vec{B}}{\partial t} = \mu F_v, \\ \text{初始条件: } & \vec{B}(t=0) = \vec{B}_0, \\ \text{边值条件: } & \vec{B}|_{\partial\Omega} = \vec{B}_{\text{boundary}}, \end{aligned} \tag{4.4}$$

其中, \vec{B}_0 由势场模型提供。如果计算区域是长方体, 底面的数据由光球测量数据提供, 侧面和顶面的数据由势场模型提供。常用的数值方法包括有限差分法、有限元法、有限体积法和谱方法等。[Wiegelmann](#) 在文献^[35] 中指出他们使用五点四阶有限差分格式进行求解该问题。在磁场进行的伪演化过程中, L 在减小, 如果 L 减小到零, 则 \vec{B} 满足非线性无力场的条件 $\nabla \times \vec{B} = \alpha \vec{B}, \nabla \cdot \vec{B} = 0$ 。

磁流体动力学 (MHD) 松弛法的微分方程组由 [Valori 等](#) 等人在文献^[81] 中指出, 该方法包括以下方程:

$$v \vec{V} = (\nabla \times \vec{B}) \times \vec{B}, \quad E + \vec{V} \times \vec{B} = \vec{0}, \quad \frac{\partial \vec{B}}{\partial t} = -\nabla \times E, \quad \nabla \cdot \vec{B} = 0.$$

如果取 $v = \frac{1}{\mu} |\vec{B}|^2$, 结合上述方程组的前三个方程, 可以得到 $\frac{\partial \vec{B}}{\partial t} = \mu \vec{F}_{\text{MHD}}$ 。配合初始条件和边界条件, 构成如下偏微分方程的初边值问题:

$$\begin{aligned} \text{微分方程: } & \frac{\partial \vec{B}}{\partial t} = \mu \nabla \times \frac{[(\nabla \times \vec{B}) \times \vec{B}]}{\vec{B}^2}, \\ \text{初始条件: } & \vec{B}(t=0) = \vec{B}_0, \\ \text{边界条件: } & \vec{B}|_{\partial\Omega} = \vec{B}_{\text{boundary}}. \end{aligned} \tag{4.5}$$

此外, [MacTaggart 等](#) 在文献^[82] 第 118 页中指出, 如果在初始时刻满足无散条件 $\nabla \cdot \vec{B}(x, y, z, t=0) = 0$, 则可以保证在伪演化的过程中仍然满足无散条件即 $\nabla \cdot \vec{B}(x, y, z, t) = 0$ 。这说明磁流体动力学松弛法在演化的过程中可以保持磁场无散性的特性。

边界积分法利用格林第二等式

$$\int_{\Omega} (Y \Delta \vec{B} - B \Delta Y) dv = \oint_{\Gamma \cup S_R \cup S_\epsilon} \left(Y \frac{\partial \vec{B}}{\partial \vec{n}} - B \frac{\partial Y}{\partial \vec{n}} \right) ds,$$

其中 \mathbf{B} 表示 $\vec{\mathbf{B}}$ 的任一磁场分量, $\partial\Omega = \Gamma \cup S_R \cup S_\epsilon$, 其中 Γ 是光球表面、 S_R 是球径为无穷大的球表面和 S_ϵ 以计算点为中心的无穷小邻域的球面。此外 \vec{n} 是边界的朝外法向量。在边界积分法中, 选择不同的格林函数来建模边界元积分外推法。选取适当的格林函数后, 并将其代入格林第二等式, 可以将区域内的体积分转化为边界上的面积分。通过对边界上的积分项进行合理的假设, 可以实现对边界元积分外推法的建模。[Yan 等](#) 在文献^[53] 中引入格林函数 $Y_p = \frac{\cos \lambda_p r}{4\pi r}$, 其中 $p = x, y, z$ 。那么任意位置的磁场满足

$$\text{边界积分方程 1: } c_i B_p = \int_{\Gamma} \left(Y_p \frac{\partial \mathbf{B}_p}{\partial \vec{n}} - \frac{\partial Y_p}{\partial \vec{n}} \mathbf{B}_{0p} \right) ds, \quad (4.6)$$

点落在边界上时 $c_i = \frac{1}{2}$, 而其它情形 $c_i = 1$ 。另外, 在边界积分方程中 Y_p 需要满足 $\int_{\Omega} Y_p \left[\lambda_p^2 \mathbf{B}_p - \alpha^2 \mathbf{B}_P - (\nabla \alpha \times \vec{\mathbf{B}})_p \right] dV = 0$ 。为避免计算法向导数, [Yan 等](#) 在文献^[42] 中引入另一种格林函数 $Y_p = Y_{1p} - Y_{2p} = \frac{\cos(\lambda_p r)}{4\pi r} - \frac{\cos(\lambda_p r')}{4\pi r'}$, 其中 $r = (x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2$, $r' = (x - x_i)^2 + (y - y_i)^2 + (z + z_i)^2$ 。同样任意位置的磁场可表示为边界积分方程的形式

$$\text{边界积分方程 2: } \mathbf{B}_p = \int_{\Gamma} \frac{z_i [\lambda_p r \sin(\lambda_p r) + \cos(\lambda_p r)] \mathbf{B}_{0P}}{2\pi \left[(x - x_i)^2 + (y - y_i)^2 + z_i^2 \right]^{\frac{3}{2}}} ds, \quad (4.7)$$

其中 λ_p 需要满足 $\lambda_p^2 = \frac{\int_{\Omega} Y(x, y, z, x_i, y_i, z_i, \lambda_p) [\alpha^2 \mathbf{B}_p + (\nabla \alpha \times \vec{\mathbf{B}})_p] dV}{\int_{\Omega} Y(x, y, z, x_i, y_i, z_i, \lambda_p) \mathbf{B}_P dV}$ 。实际上, 为避免复杂的计算, [Yan 等](#)^[42] 认为 $\vec{\mathbf{B}}$ 是关于 λ_p 的函数, 他们通过无梯度优化方法下山单纯形方法^[83] 极小化函数 $\frac{|\nabla \times \vec{\mathbf{B}} \times \vec{\mathbf{B}}|}{|\nabla \times \vec{\mathbf{B}}||\vec{\mathbf{B}}|}$ 和 $\frac{|\nabla \cdot \vec{\mathbf{B}}|}{|\vec{\mathbf{B}}|}$ 以求得 λ_p 。这样可以直接得到适用于边界积分方程的格林函数。

在这一节中, 除边界积分法^[40-42,53,84] 外, 其它经典无力场模型外推方法都可以写成下面的形式:

$$\begin{cases} \text{微分方程: } \mathcal{L}(u(\vec{x}, t)) = f(\vec{x}, t), \vec{x} \in \Omega, t \in [t_0, t_1], \\ \text{边界条件: } u(\vec{x}, t) = g(\vec{x}, t), \vec{x} \in \partial\Omega, t \in [t_0, t_1], \\ \text{初始条件: } u(\vec{x}, t_0) = h(\vec{x}), \vec{x} \in \Omega, \end{cases} \quad (4.8)$$

其中 \mathcal{L} 指的是包含微分算子的算子, 其中微分算子指的是 $\frac{d}{dt}, \frac{\partial}{\partial t}, \frac{\partial^2}{\partial t^2}, \frac{\partial}{\partial x}, \frac{\partial^2}{\partial x^2}, \frac{\partial^2}{\partial x \partial y}$ 等, $\mathcal{L}(u)$ 的形式比如 $\frac{du}{dt}, \Delta u, \frac{\partial u}{\partial t} - \Delta u, \frac{\partial^2 u}{\partial t^2} - \Delta u$ 等。

传统上, 人们使用有限差分、有限体积、有限元、边界元和谱方法等数值方法求解 (4.8)。这些方法的特点如下:

有限差分方法主要适用于求解偏微分方程的数值解，它将连续的偏微分方程转化为离散的差分方程并利用差分逼近来近似求解原始问题。有限差分方法在流体力学、电磁学、热传导等领域中得到广泛应用。它的优点是简单易实现、计算效率高，对于规则网格和简单几何形状适用性强，对各向异性的扩散系数较为敏感。它的缺点是对于复杂几何形状的处理能力较弱，数值精度受到网格步长的影响，需要较小的时间和空间步长以获得精确的结果。

有限体积方法适用于求解守恒型方程，如流体动力学方程、输运方程等。有限体积方法将求解区域划分为离散的控制体积，通过对控制体积内的物理量平均值进行计算，在控制体积界面上建立通量的平衡方程从而求解守恒型方程。有限体积方法在流体力学、石油工程、环境模拟等领域中得到广泛应用。它的优点是可以很好地保持守恒关系，对于守恒型问题具有优势。有限体积法适用于复杂几何形状和非结构网格，对守恒量和界面流动的描述较为准确。它的缺点是需要处理更多的控制体积接口，计算量较大，在非守恒型问题上的应用相对较少。需要较小的时间和空间步长以获得精确的结果。

边界元方法适用于求解边界积分方程的数值解，边界元方法将求解域划分为内部和边界两部分，通过在边界上建立积分方程来求解内部问题。边界元方法在电磁学、弹性力学、声学等领域中得到广泛应用。边界元方法的优点是只需在边界上离散，这避免了内部网格划分。边界元方法适用于无穷大区域和特殊边界条件，在二维和三维椭圆型偏微分方程中具有特殊优势。它的缺点是边界上离散需要更加复杂的数学理论、简化假设和边界近似会影响结果的准确性，对于不同类型的偏微分方程的适用性有限。

有限元方法适用于求解连续介质力学问题的数值解，有限元方法将求解域划分为离散的有限元，利用形函数来近似描述物理量的变化，在每个单元上建立局部方程，通过装配和求解整体方程组来得到问题的数值解。有限元方法在结构力学、流体力学、电磁学等领域中得到广泛应用。有限元方法的优点是适用于复杂几何形状和非结构网格，对不规则问题处理能力强，可以灵活处理各向异性问题，可以高阶逼近解，提供较高的数值精度。有限元的缺点是计算量较大特别是对于大规模问题，有限元方法需要进行单元划分和组装过程，实现复杂度较高，对于非线性问题求解相对较为困难。

谱方法适用于求解线性或非线性偏微分方程的数值解，谱方法利用高阶的

基函数展开待求解的函数，通过选取合适的基函数系数来满足边界条件和方程本身从而得到数值解。谱方法在流体力学、量子力学、气象模拟等领域中得到广泛应用。谱方法的优点是高精度和快速收敛。谱方法适用于光滑解和周期性问题，对高阶导数和非线性项具有较好的处理能力。谱方法的缺点是对复杂几何形状的要求较高，网格生成和基函数选取较为困难，计算量较大，在大规模问题上应用相对有限。

4.2 无力场神经网络磁场外推法

自 2019 年以来，Raissi 等^[76] 和 Karniadakis 等^[77] 提出使用物理约束神经网络求解 (4.8)，实际上 Lagaris 等^[85] 早已提出这种概念，只不过当时的计算机性能不够好。

物理约束神经网络主要应用在数值偏微分方程领域中，物理约束神经网络通过将神经网络与物理方程相结合用于求解各种复杂的偏微分方程问题。对于给定的偏微分方程，通常还需要定解条件(边界条件和初始条件)。物理约束神经网络通过将这些条件嵌入到神经网络中，在学习过程中自动满足这些条件，从而得到满足边界条件和初始条件的解。许多实际问题涉及非线性偏微分方程，例如 Navier-Stokes 方程、Klein-Gordon 方程等。物理约束神经网络可以用于求解这些非线性方程，通过将非线性项嵌入到神经网络中并结合适当的边界条件和初始条件来获得数值解或近似解。某些偏微分方程问题可能存在不适定性，即由于数据限制或噪声等原因无法得到唯一的解。物理约束神经网络通过学习过程中的正则化和约束来解决这些不适定问题，从而获得稳定和合理的解。在某些情况下我们需要考虑多个物理过程的相互作用，物理约束神经网络可以通过将多个物理方程结合到同一个神经网络中来处理这些多物理耦合问题，从而实现对复杂耦合系统的建模和求解。当我们缺乏精确的物理方程时可以使用物理约束神经网络进行数据驱动的建模，通过输入已知的数据点、边界条件和初始条件，物理约束神经网络可以学习到近似的物理方程并预测未知区域的解。

综上所述，物理约束神经网络的本质是将偏微分方程作为约束条件引入到神经网络的损失函数中，通过设计适当的神经网络结构和损失函数来求解偏微分方程问题。由于机器学习具有强大的计算能力，物理约束神经网络提供了一种强大的数值偏微分方程求解工具，可以应用于各种复杂的物理问题。

4.2.1 无力场神经网络外推法的数学模型

根据无力场模型的数学表达式，我们可以将其写成以下标准形式的微分方程组：

$$\begin{aligned} \text{微分方程 1: } & \nabla \times \vec{B}(x, y, z) \times \vec{B}(x, y, z) = \vec{0}, \quad (x, y, z) \in \Omega, \\ \text{微分方程 2: } & \nabla \cdot \vec{B}(x, y, z) = 0, \quad (x, y, z) \in \Omega, \\ \text{边界条件: } & \vec{B}\Big|_{\partial\Omega} = \vec{B}(x', y', z'), \quad (x', y', z') \in \partial\Omega. \end{aligned} \quad (4.9)$$

假设存在一个从 \mathbb{R}^3 到 \mathbb{R}^3 的连续函数 $\vec{B}(x, y, z)$ 满足上述微分方程组的边值问题。根据第 2 章的万能逼近定理（定理 2.5），存在一个非多项式连续函数 $\vec{B}^{NN}(x, y, z)$ ，使得

$$\left| \vec{B}^{NN}(x, y, z) - \vec{B}(x, y, z) \right| < \epsilon, \quad (4.10)$$

成立，其中 ϵ 是任意无穷小的正数。因此，无力场外推方法的磁场 $\vec{B}^{NN}(x, y, z)$ 满足以下数学模型：

$$\begin{aligned} & \lambda_0 \left\| \vec{B}^{NN}(x, y, z) - \vec{B}(x, y, z) \right\|_2^2 \cdot \mathbb{I}_{(x,y,z) \in \partial\Omega} + \dots \\ & \dots + \lambda_1 \left\| \nabla \cdot \vec{B}^{NN}(x, y, z) \right\|_2^2 \cdot \mathbb{I}_{(x,y,z) \in \Omega} + \dots \\ & \dots + \lambda_2 \frac{\left\| \nabla \times \vec{B}^{NN}(x, y, z) \times \vec{B}^{NN}(x, y, z) \right\|_2^2}{\left\| \nabla \times \vec{B}^{NN}(x, y, z) \right\|_2^2 \cdot \left\| \vec{B}^{NN}(x, y, z) \right\|_2^2} \cdot \mathbb{I}_{(x,y,z) \in \Omega} \\ & = 0, \end{aligned} \quad (4.11)$$

其中 $(x, y, z) \in \partial\Omega \cup \Omega$, $\lambda_0, \lambda_1, \lambda_2$ 是任意正实数。此外， \mathbb{I} 为示性函数，即 $\mathbb{I}_{(x,y,z) \in A} = \begin{cases} 1, & (x, y, z) \in A, \\ 0, & (x, y, z) \notin A. \end{cases}$ 在这个模型中，当 $(x, y, z) \in \Omega$ 时，无力场神经网络外推法外推重建的磁场满足无力场模型的条件；当 $(x, y, z) \in \partial\Omega$ 时，无力场神经网络外推法外推重建的磁场满足边界条件。

4.2.2 无力场神经网络外推法模型的数值算法

在这一小节中，我们将通过神经网络来实现上一节介绍的无力场神经网络法磁场外推的数学模型。

我们希望通过神经网络使得 $\vec{B}^{NN}(x, y, z; \vec{p}) \approx \vec{B}(x, y, z)$ 。为此，我们定义了

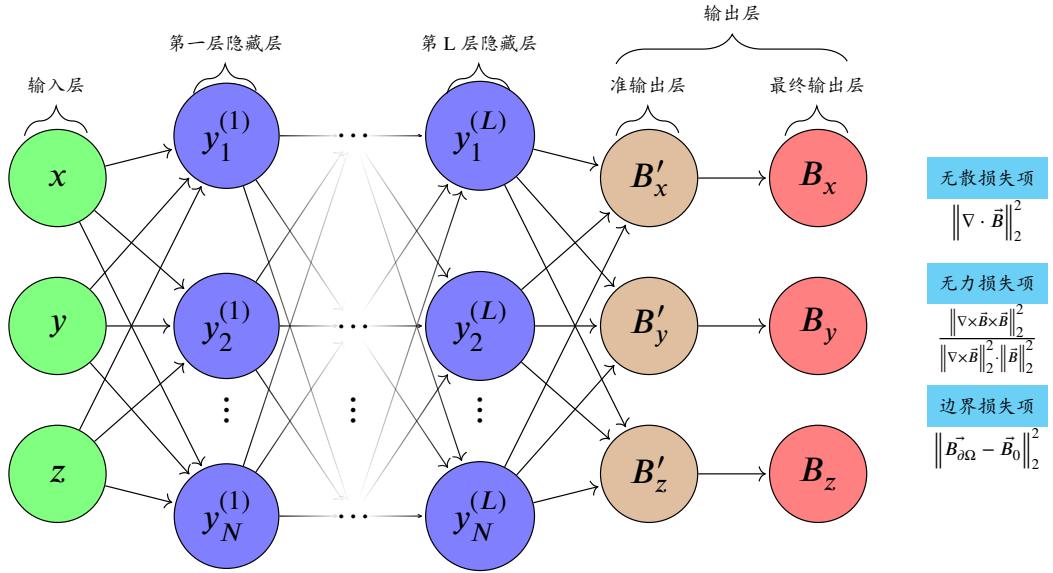


图 4.1 无力场神经网络外推法采用的网络结构

神经网络的损失函数 $f(\vec{p})$, 具体形式如下:

$$\begin{aligned}
 f(\vec{p}) = & \lambda_0 \frac{1}{m} \sum_{i=1}^m \left\| \vec{B}^{NN}(x'_i, y'_i, z'_i; \vec{p}) - \vec{B}(x'_i, y'_i, z'_i) \right\|_2^2 + \\
 & + \lambda_1 \frac{1}{n} \sum_{j=1}^n \left\| \nabla \cdot \vec{B}^{NN}(x_j, y_j, z_j; \vec{p}) \right\|_2^2 + \\
 & + \lambda_2 \frac{1}{n} \sum_{j=1}^n \frac{\left\| \nabla \times \vec{B}^{NN}(x_j, y_j, z_j; \vec{p}) \times \vec{B}^{NN}(x_j, y_j, z_j; \vec{p}) \right\|_2^2}{\left\| \nabla \times \vec{B}^{NN}(x_j, y_j, z_j; \vec{p}) \right\|_2^2 \cdot \left\| \vec{B}^{NN}(x_j, y_j, z_j; \vec{p}) \right\|_2^2},
 \end{aligned} \tag{4.12}$$

其中 $(x'_i, y'_i, z'_i) \in \partial\Omega$, $(x_j, y_j, z_j) \in \Omega$ 是随机产生的训练集。

注. 如果对于任意的训练集, 损失函数 (4.12) 都能减少到零, 那么可以推断出神经网络求解的磁场满足无力场方程 (4.9)。

注. 此外, 需要注意的是可学习参数 \vec{p} 是一个向量, 由神经网络 \vec{B}^{NN} 中的权重和偏置按照一定顺序排列而成。

我们使用的神经网络结构如图 4.1 所示, 它由 $L + 1$ 个全连接层组成。除了输入层和输出层外, 每个隐藏层包含 N 个神经元, 并且每个神经元都带有一个双曲正切激活函数。输入层、准输出层和最终输出层各自有三个神经元。

神经网络的输入层可以用以下数学表达式表示:

$$\left[y_1^{(1)}, y_2^{(1)}, \dots, y_N^{(1)} \right]^T = \tanh(W^{(1)}[x, y, z]^T + b^{(1)}),$$

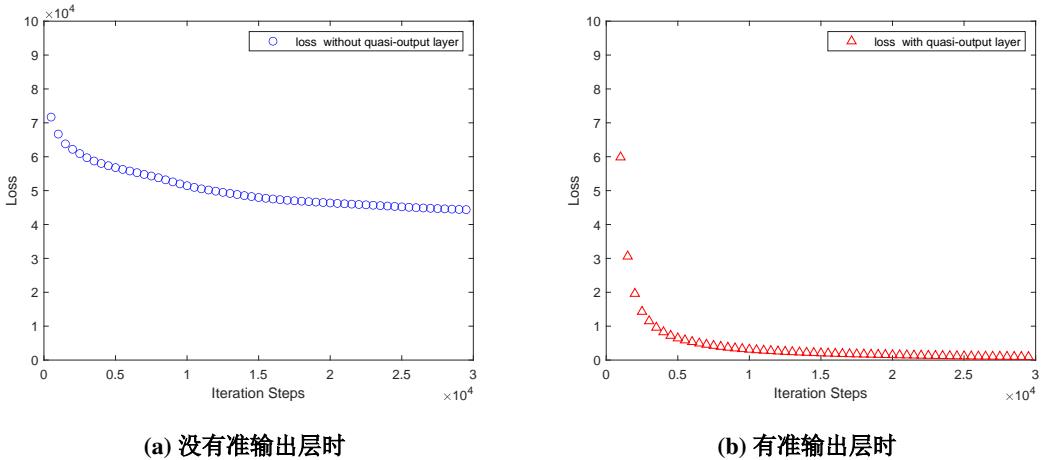


图 4.2 有无准输出层时，损失函数随迭代次数变化情况的一个例子

其中 $W^{(1)} \in \mathbb{R}^{N \times 3}$, $\vec{b}^{(1)} \in \mathbb{R}^{N \times 1}$ 。

隐藏层的数学表达式为：

$$\left[y_1^{(i)}, y_2^{(i)}, \dots y_N^{(i)} \right]^T = \tanh \left(W^{(i)} \left[y_1^{(i-1)}, y_2^{(i-1)}, \dots y_N^{(i-1)} \right]^T + b^{(i)} \right),$$

其中 $W^{(i)} \in \mathbb{R}^{N \times N}$, $\vec{b}^{(i)} \in \mathbb{R}^{N \times 1}$, $2 \leq i \leq L$ 。

准输出层的数学表达式为：

$$\left[B'_x, B'_y, B'_z \right]^T = W^{(L+1)} \left[y_1^{(L)}, y_2^{(L)}, \dots y_N^{(L)} \right]^T + b^{(L+1)},$$

其中 $W^{(L+1)} \in \mathbb{R}^{3 \times N}$, $\vec{b}^{(L+1)} \in \mathbb{R}^{3 \times 1}$ 。

最终输出层的数学表达式为：

$$B_x = a_x B'_x + b_x, \quad B_y = a_y B'_y + b_y, \quad B_z = a_z B'_z + b_z. \quad (4.13)$$

在无力场神经网络外推法的架构中，包含一个准输出层是至关重要的。如果没有这一层，损失函数 $f(\vec{p})$ 在训练过程中可能无法收敛到零。此外，在某些情况下，它甚至可能发散并显著大于 0，如图 4.2 所示。值得注意的是，在图 4.2 (b) 和 (c) 中，我们排除了初始损失值，这是因为初始值与后续值之间存在显著差异。由微积分知，将级数增加、删除、改变有限项并不改变原级数的敛散性。

经计算, 神经网络中可学习参数的维度为 $\vec{p} \in \mathbb{R}^{[(L+1)N^2 + (L+6)N + 9] \times 1}$ 。

在机器学习优化理论中，我们通常采用以下数值格式来通过参数 \vec{p} 最小化损失函数 $f(\vec{p})$ ：

$$\vec{p}_{k+1} = \vec{p}_k + \alpha_k \vec{d}_k, \quad (4.14)$$

其中 α_k 是学习率, \vec{d}_k 是损失函数下降的方向。我们希望通过迭代, 使得 $f(\vec{p}_{k+1}) \leq f(\vec{p}_k)$ 。同时, 当迭代次数足够大时, 我们希望 $f(\vec{p}_k)$ 趋近于 0。

在无场神经网络优化过程中, 我们采用交替变量更新法来交替更新 α_k 和 \vec{d}_k 。为了更新学习率 α_k , 我们使用三次插值函数。经典的机器学习神经网络通常使用固定的学习率或者根据迭代步数来调整学习率。而对于更新损失函数的下降方向 \vec{d}_k , 我们采用二阶优化方法-拟牛顿法, 相比于经典的一阶优化方法更加有效。

具体而言, 第一步是固定 \vec{d}_k , 更新 α_k 。我们针对函数 $\varphi(\alpha) = f(\vec{p}_k + \alpha\vec{d}_k)$, 选择适当的 $a_0, a_1 > 0$, 然后通过对 $\varphi(0), \varphi'(0), \varphi(a_0), \varphi(a_1)$ 进行插值, 得到一个三次函数 $\varphi(\alpha)$ 。最后, 通过极小化 $\varphi(\alpha)$ 来确定 α_k 的最优值, 其中具体的插值函数可以参考附录 E。

第二步是固定 α_k 后, 更新 \vec{d}_k 。在给定 $L = 10, N = 100$ 的情况下, 我们仅考虑梯度 $\nabla f(\vec{p})$ 的维数。由于 $\nabla f(\vec{p}) \in \mathbb{R}^{110,709}$, 计算所需的内存空间较大且计算量也很大。因此, 我们采用有限内存的 Broyden Fletcher Goldfarb Shanno (L-BFGS) 算法^[86] 来更新 \vec{d}_k 。

L-BFGS 源于 BFGS 算法,BFGS 算法通过构造对称矩阵 B_k 使得 $B_{k+1}(\vec{p}_{k+1} - \vec{p}_k) = \nabla f_{k+1} - \nabla f_k$, 将 B_{k+1} 写成迭代形式, 即

$$B_{k+1} = B_k + \frac{(\vec{i}_k - B_k \vec{h}_k)(\vec{i}_k - B_k \vec{h}_k)^T}{(\vec{i}_k - B_k \vec{h}_k)^T \vec{h}_k}, \quad (4.15)$$

是一个低秩矩阵且秩为 2, 其中 \vec{h}_k 和 \vec{i}_k 分别为 $\vec{h}_k = \vec{p}_{k+1} - \vec{p}_k, \vec{i}_k = \nabla f_{k+1} - \nabla f_k$ 。若记 $C_k = B_k^{-1}$, 则

$$C_{k+1} = \left(I - g_k \vec{h}_k \vec{i}_k^T \right) C_k \left(I - g_k \vec{i}_k \vec{h}_k^T \right) + g_k \vec{h}_k \vec{h}_k^T, \quad (4.16)$$

其中 $g_k = \frac{1}{\vec{i}_k^T \vec{h}_k}$ 。BFGS 算法通过逐步近似目标函数的 $(\nabla^2 f_k)$ 且给出更新搜索方向, 从而找到使目标函数最小化的最优解。近似 $(\nabla^2 f_k)^{-1}$ 时会根据每次迭代中的梯度变化情况进行调整以获得更好的搜索方向。

与传统的 BFGS 算法不同, L-BFGS 算法并不需要显式地存储和更新海森矩阵的近似及其近似的逆, L-BFGS 算法通过保留一部分历史搜索方向和梯度差值的信息来进行迭代优化, 该信息被称为记忆。记忆的大小需事先指定。具体而

算法 2 无场神经网络中参数 \vec{p} 的数值算法

输入: $k \leftarrow 0, m \leftarrow m_0, \vec{p}_0 \leftarrow \vec{p}$

- 1: **while** p_k 不收敛 **do**
- 2: 计算 α_k
- 3: 选定 C_0 或 C_k^0
- 4: 计算 $\vec{d}_k \leftarrow -C_k \nabla f_k$
- 5: 计算 $\vec{p}_{k+1} \leftarrow \vec{p}_k + \alpha_k \vec{d}_k$
- 6: 若 $k > m$, 则从内存中丢弃 $\{\vec{h}_{k-m}, \vec{i}_{k-m}\}$
- 7: 计算 \vec{h}_k, \vec{i}_k 并存储
- 8: **end while**

输出: \vec{p}_{k+1}

言, L-BFGS 算法使用有限的记忆来逼近目标函数的海森矩阵乘以搜索方向的乘积。通过利用这个近似值 L-BFGS 算法能够有效地求解大规模非线性优化问题并且减少存储和计算开销。

我们通过 L-BFGS 的优化方式更新 \vec{d}_k , 它的表达式是 $\vec{d}_k = -C_k \nabla f(\vec{p}_k)$, 其中当 $1 \leq k \leq m$ 时,

$$C_k = E_{k-1} C_{k-1} E_{k-1}^T + g_{k-1} \vec{h}_{k-1} \vec{h}_{k-1}^T, \quad (4.17)$$

当 $k > m$ 时,

$$\begin{aligned} C_k = & (E_{k-1}^T \cdots E_{k-m}^T) C_k^0 (E_{k-m} \cdots E_{k-1}) \\ & + g_{k-m} (E_{k-1}^T \cdots E_{k-m+1}^T) \vec{h}_{k-m} \vec{h}_{k-m}^T (E_{k-m+1} \cdots E_{k-1}) \\ & + \cdots + g_{k-1} \vec{h}_{k-1} \vec{h}_{k-1}^T, \end{aligned} \quad (4.18)$$

在 (4.17) 和 (4.18) 中, E_k, g_k, C_k^0 分别为

$$E_k = I - g_k \vec{i}_k \vec{h}_k^T, g_k = \frac{1}{\vec{i}_k^T \vec{h}_k}, C_k^0 = \frac{\vec{h}_{k-1}^T \vec{i}_{k-1}}{\vec{i}_{k-1}^T \vec{i}_{k-1}} I. \quad (4.19)$$

另外, (4.17) 中 C_0 在 \vec{p}_0 的初始化可参考文献^[83]。

求解无场神经网络中参数 \vec{p} 的完整流程见算法 2。

4.2.3 无力场神经网络外推法模型的应用

4.2.3.1 在 Low & Lou 半解析解模型上的应用

首先，根据第3章，我们选取 Low 等^[54] 提供的轴对称无力场来验证我们提出的无力场神经网络外推法的性能。我们分别考虑两个例子。在例子1中，我们选择 $a_{1,1} = 0.6538, l = 0.3, \Phi = \frac{\pi}{4}$ 。而在例子2中，我们选定 $a_{3,1} = 2.8348, l = 0.3, \Phi = \frac{4\pi}{5}$ 。这两个例子都限制在计算区域 $(x, y, z) \in [-1, 1] \times [-1, 1] \times [0, 2]$ 内。

关于例子1，我们给定计算区域边界六个面上的磁场作为边界条件。我们采用一个 $L = 5$ 层和每层 128 个神经元的神经网络，模型的超参数分别设置为 $\lambda_0 = 1, \lambda_1 = 1$ 和 $\lambda_2 = 0.95$ 。此外，将 m_0 设置为 10，并采用 He 等^[87] 提出的初始化方法对 \vec{p} 进行初始化。我们在边界 $\partial\Omega$ 上等间隔选取 23,816 点 ($64 \times 64 \times 6 - (64 - 2) \times 12 - 2 \times 8 = 23,816$)。同时，在空间 Ω 中随机选取 360,000 个点用来训练神经网络。

关于例子2，仅给定计算区域底面上的磁场作为边界条件。Morse 等^[88] 指出这是一个病态问题，因为对于良定义的问题，计算区域的所有边界条件都需要被给定。仅以底面的磁场作为边界条件进行磁场外推会导致严重不符合参考磁场分布（如图4.4所示）。然而，在实际观测中，我们只能测量到太阳光球的磁场。为了补全边界条件的其他五个面的磁场数据，我们使用势场模型的外推结果作为边界上另外五个面的边界条件。我们采用一个包含 $L = 7$ 层神经元的神经网络，每层 156 个神经元，模型的超参数取值为 $\lambda_0 = 1.25, \lambda_1 = 1$ 和 $\lambda_2 = 1, m$ 设置成 50，并采用与例子1相同的方式初始化 \vec{p} 。我们在底边界 $\partial\Omega$ 上均与选取选取 4,096 点 ($64 \times 64 = 4,096$)，并在空间 Ω 中随机选取 360,000 个点用来训练神经网络。另外，我们在其它五个面等间隔采样 20,482 个点的伪磁场数据作为边界条件的一部分。

为了评估算法在外推重建无力场磁场的性能，我们采用 Schrijver 等^[89] 提出的如下评价：基于向量内积的相关性系数 C_{vec} 、基于柯西施瓦兹不等式的 C_{cs} 、向量的误差归一化 E_n 、向量平均误差 E_m 、相对参考磁场的能量 \mathcal{E} 和相对势场

模型的能量的比值 \mathcal{E} 。其中 C_{vec} 的计算公式是

$$C_{\text{vec}} \equiv \frac{\sum_{i=1}^M \vec{B}_i^{\text{Ref}} \cdot \vec{B}_i^{\text{Rec}}}{\left(\sum_{i=1}^M \left| \vec{B}_i^{\text{Ref}} \right|^2 \sum_{i=1}^M \left| \vec{B}_i^{\text{Rec}} \right|^2 \right)^{\frac{1}{2}}},$$

其中 \vec{B}^{Ref} 是参考磁场， \vec{B}^{Rec} 是外推重建的磁场， C_{vec} 描述的是两个向量场在方向上的相似度，类似与统计学中的相关系数。如果参考磁场被完全准确重建，那么 $C_{\text{vec}} = 1$ 。 C_{cs} 的计算公式是

$$C_{\text{cs}} \equiv \frac{1}{M} \sum_{i=1}^M \frac{\vec{B}_i^{\text{Ref}} \cdot \vec{B}_i^{\text{Rec}}}{\left| \vec{B}_i^{\text{Ref}} \right| \left| \vec{B}_i^{\text{Rec}} \right|},$$

其中 M 是在整个空间中选定向量的数量。 C_{cs} 描述了两个向量场方向上的平均相似程度。当外推重建磁场与参考磁场之间的夹角为零时 $C_{\text{cs}} = 1$ 。上述评价指标 C_{vec} 和 C_{cs} 的值越大说明重建参考磁场的能力越好。 E_n 的计算公式是

$$E_n \equiv \frac{\sum_{i=1}^M \left| \vec{B}_i^{\text{Ref}} - \vec{B}_i^{\text{Rec}} \right|}{\sum_{i=1}^M \left| \vec{B}_i^{\text{Ref}} \right|},$$

该指标描述的是两个向量场之间差异归一化的结果。此外 E_m 的计算公式是

$$E_m \equiv \frac{1}{M} \sum_{i=1}^M \frac{\left| \vec{B}_i^{\text{Ref}} - \vec{B}_i^{\text{Rec}} \right|}{\left| \vec{B}_i^{\text{Ref}} \right|}, \quad (4.20)$$

该指标描述的两个向量场中向量差异的平均值。然而， E_n , E_m 和前面两个评价指标不一样，取值越小反而说明重建效果越好。为了保持评价指标的一致性，取 $E'_n = 1 - E_n$ 和 $E'_m = 1 - E_m$ 作为新的评价指标。从能量的角度计算

$$\mathcal{E} \equiv \frac{\sum_{i=1}^M \left| \vec{B}_i^{\text{Rec}} \right|^2}{\sum_{i=1}^M \left| \vec{B}_i^{\text{Ref}} \right|^2},$$

它衡量外推重建磁场的能量占参考磁场的比重。 \mathcal{E} 越靠近 1，重建的效果越好。

从自由能的角度计算

$$\mathcal{E}_P \equiv \frac{\sum_{i=1}^M \left| \vec{B}_i^{\text{Rec}} \right|^2}{\sum_{i=1}^M \left| \vec{B}_i^P \right|^2},$$

其中 \vec{B}_i^P 是势场模型外推的结果。 \mathcal{E}_P 取值越大表明重建磁场蕴含的自由能越多。

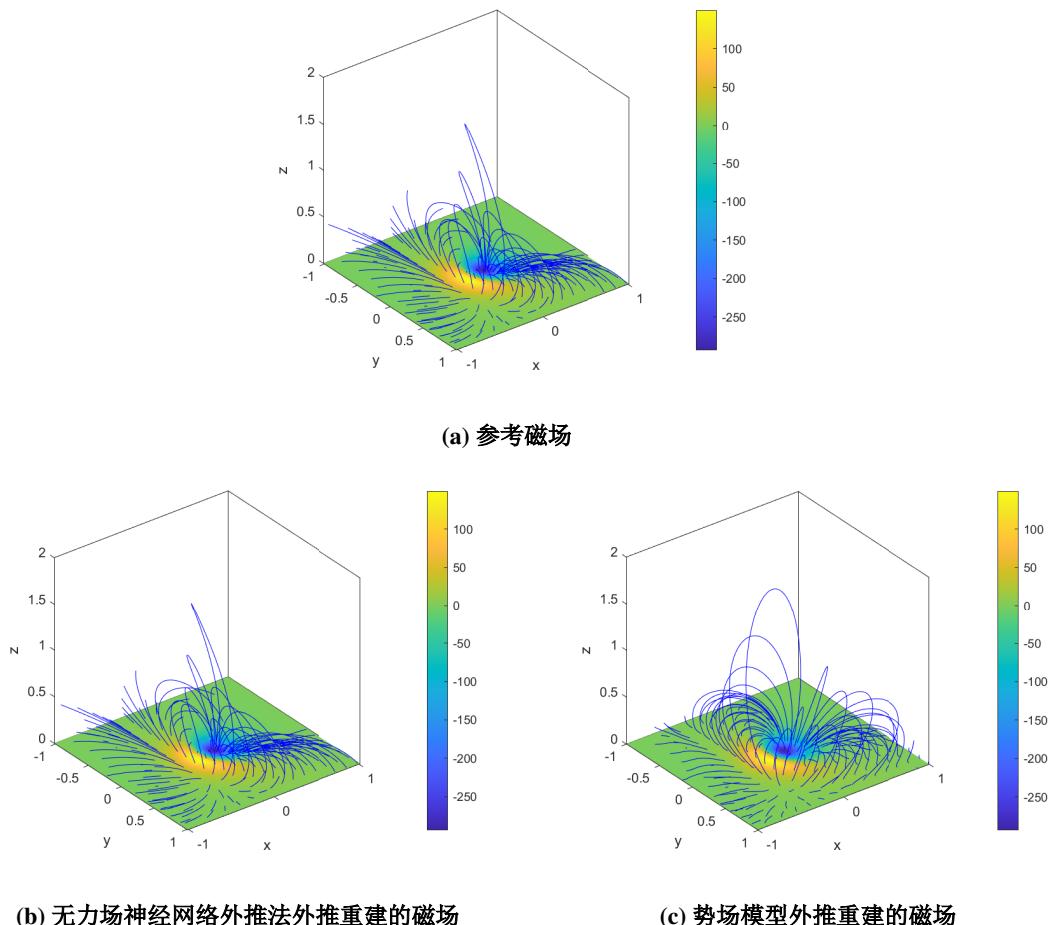


图 4.3 可视化例子 1 中的磁力线

根据上述评价指标, 我们提出的无场神经网络外推算法与以下方法进行了比较: Low 等^[54] 提供的参考磁场、Wiegemann^[35] 实现的含有权重系数的优化法磁场外推结果、McTiernan 实现的原始优化法磁场外推结果、Valori 等^[81] 实现的磁流体动力学松弛法磁场外推结果、Wheatland^[27] 实现的 GR 法磁场外推结果、Regnier 等^[26] 实现的 GR 法磁场外推结果、常 α 线性无场模型法磁场外推结果以及势场模型法磁场外推结果。需要补充说明的是, 在例子 2 中, 所有的外推方法都在评估区域 $[-1, 1] \times [-1, 1] \times [0, 2]$ 上进行。该区域被等间隔划分为 $64 \times 64 \times 64$

表 4.1 无力场神经网络外推法与其他经典无力场外推方法的比较

(a) 在例子 1 上的评价结果

| Model | C_{vec} | C_{cs} | E'_n | E'_m | \mathcal{E} | \mathcal{E}_P |
|------------|------------------|-----------------|--------|--------|---------------|-----------------|
| Low | 1 | 1 | 1 | 1 | 1 | 1.29 |
| Ours | 1.00 | 1.00 | 0.98 | 0.98 | 0.99 | 1.28 |
| Wiegelmann | 1.00 | 1.00 | 0.98 | 0.98 | 1.02 | 1.31 |
| McTiernan | 1.00 | 0.99 | 0.92 | 0.87 | 1.00 | 1.30 |
| Valori | 0.99 | 0.68 | 0.71 | 0.33 | 0.98 | 1.21 |
| Wheatland | 0.98 | 0.83 | 0.64 | 0.42 | 0.90 | 1.17 |
| Regnier | 0.93 | 0.49 | 0.41 | 0.09 | 0.80 | 1.04 |
| LFFF | 0.88 | 0.90 | 0.50 | 0.42 | 0.77 | 1.00 |
| PF | 0.85 | 0.82 | 0.45 | 0.35 | 0.77 | 1 |

(b) 在例子 2 上的评价结果

| Model | C_{vec} | C_{cs} | E'_n | E'_m | \mathcal{E} | \mathcal{E}_P |
|------------|------------------|-----------------|--------|--------|---------------|-----------------|
| Low | 1 | 1 | 1 | 1 | 1 | 1.10 |
| Ours | 1.00 | 0.43 | 0.79 | 0.01 | 0.99 | 1.09 |
| Wiegelmann | 1.00 | 0.57 | 0.86 | -0.25 | 1.04 | 1.14 |
| McTiernan | 1.00 | 0.51 | 0.84 | -0.38 | 1.04 | 1.14 |
| Valori | 0.99 | 0.55 | 0.75 | -0.15 | 1.02 | 1.12 |
| Wheatland | 0.99 | 0.58 | 0.69 | 0.13 | 0.96 | 1.05 |
| Regnier | 0.94 | 0.28 | 0.49 | -1.7 | 0.74 | 0.82 |
| LFFF | 0.93 | 0.08 | -0.80 | -37 | 1.04 | 1.15 |
| PF | 0.92 | 0.35 | 0.47 | -0.63 | 0.91 | 1 |

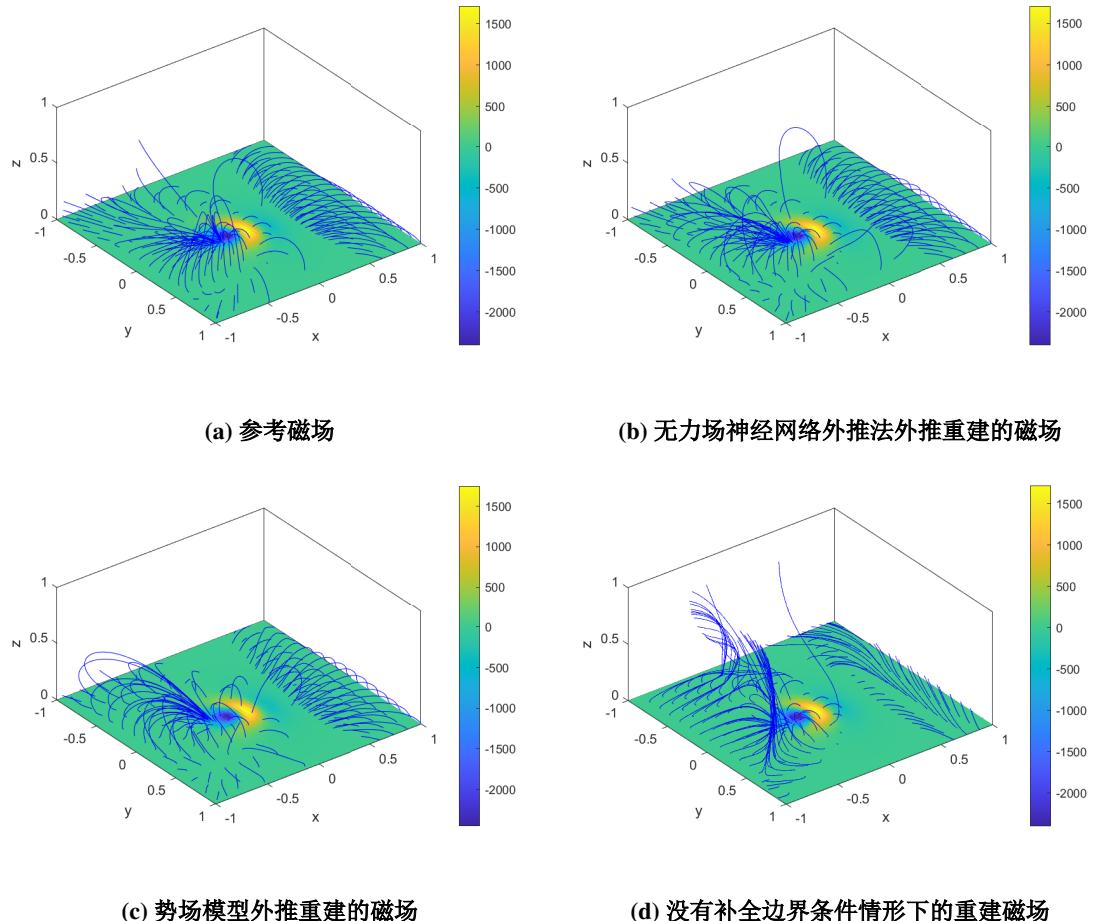
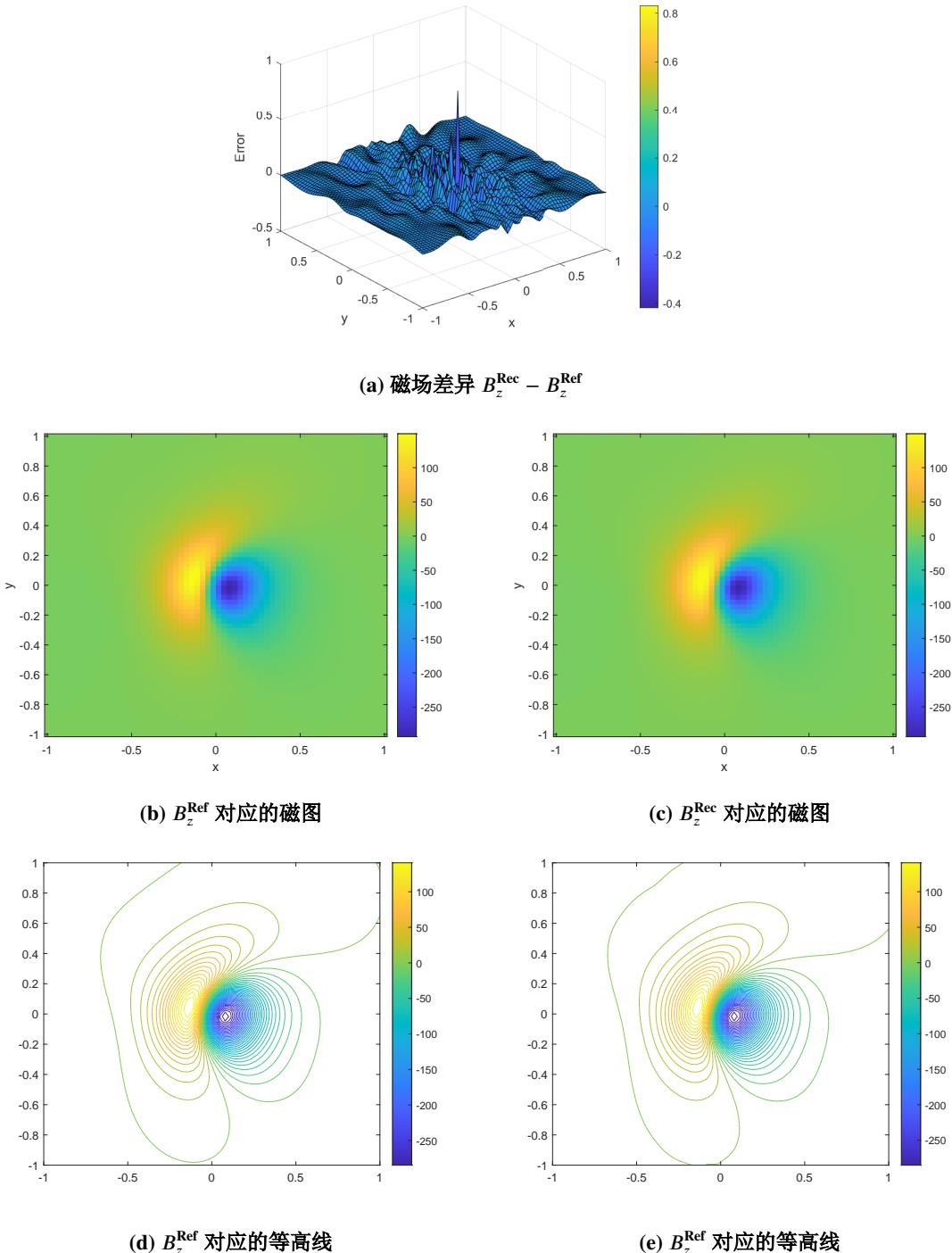


图 4.4 可视化例子 2 中的磁力线

个计算节点，并用于计算上述指标。然而，Wiegelmann^[35]实现的含有权重系数的优化法使用的计算区域是 $[-1.25, 1.25] \times [-1.25, 1.25] \times [0, 2.25]$ ，McTiernan 实现的原始优化法使用的计算区域是 $[-1.625, 1.625] \times [-1.625, 1.625] \times [0, 3.1875]$ ，Valori 等^[81]实现的磁流体动力学松弛法使用的计算区域是 $[-1.3438, 1.3438] \times [-1.3438, 1.3438] \times [0, 2.6875]$ 。这些方法的计算区域都比评估区域大很多，而我们方法的计算区域与评估区域一致，都是 $[-1, 1] \times [-1, 1] \times [0, 2]$ 。

表 4.1 显示了各种无力场模型磁场外推方法在不同评价指标下的性能。从表 4.1 (a) 可以看出，无力场神经网络外推法取得了最好的结果。然而，在表 4.1 (b) 中，并非所有指标上都表现最佳，这可能是因为优化法和磁流体动力学松弛法使用了比我们更大的计算区域，而评估计算的区域却相同造成。

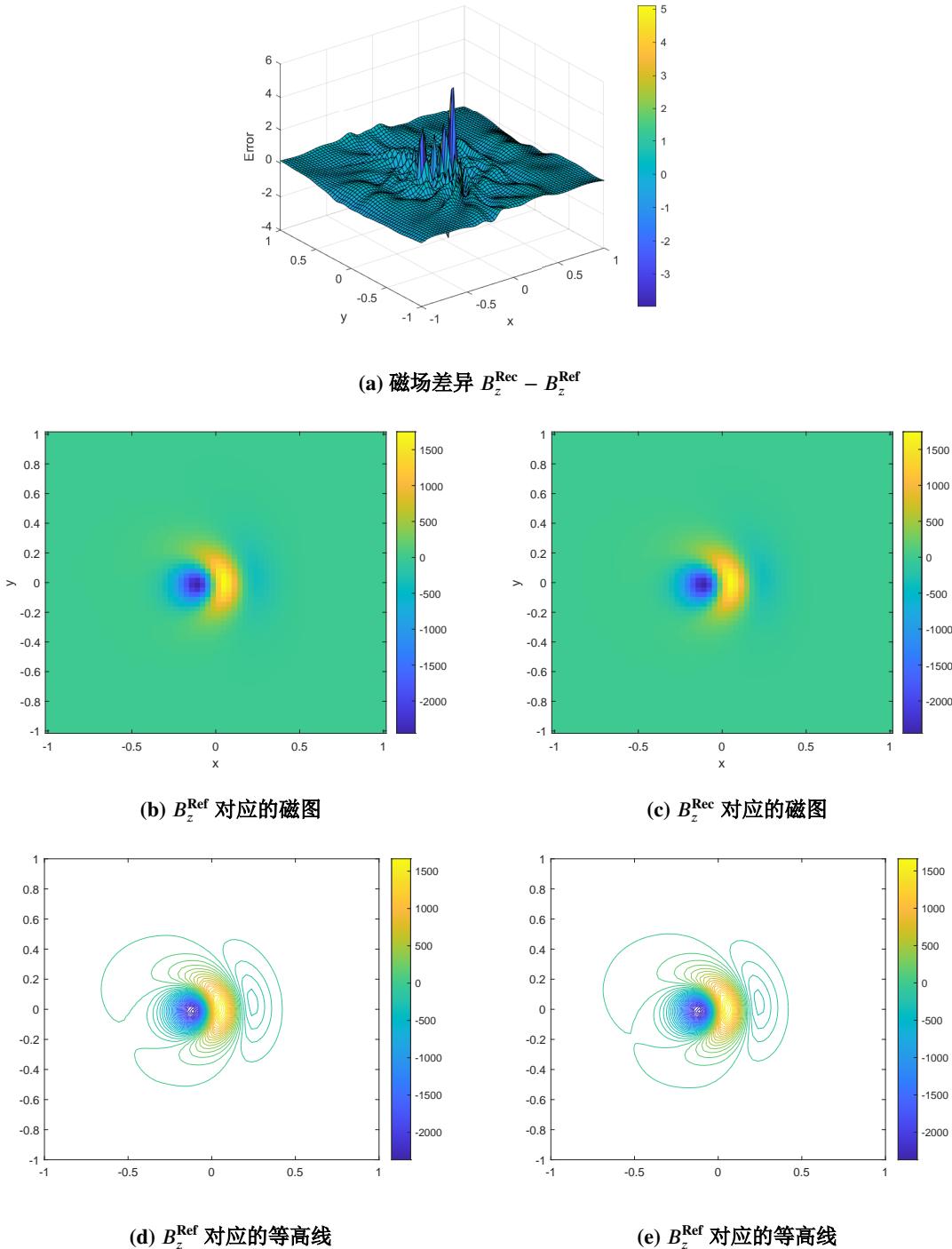
除了根据上述客观指标评估磁场外推方法的性能，我们还需要通过可视化三维磁场来了解其拓扑结构。在这方面，磁力线是最重要的工具之一。磁力线 \vec{r}

图 4.5 关于例子 1，在 $z = 0$ 平面可视化 B_z

在一般标准正交坐标系 $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ 下应满足以下微分方程：

$$\frac{d\vec{r}}{ds} = h_1 \frac{du_1}{ds} \vec{e}_1 + h_2 \frac{du_2}{ds} \vec{e}_2 + h_3 \frac{du_3}{ds} \vec{e}_3 = B_1 \vec{e}_1 + B_2 \vec{e}_2 + B_3 \vec{e}_3. \quad (4.21)$$

通过求解上述微分方程，可以得到磁力线在空间中的分布，从而实现对磁场拓扑结构的可视化。我们通常采用数值方法来求解该微分方程，如欧拉法或龙格-库塔

图 4.6 关于例子 2，在 $z = 0$ 平面可视化 B_z

法等。通过选择适当的初始条件，并沿着磁场方向积分微分方程，可以确定磁力线在三维空间中的分布。为了确定磁力线的具体位置，我们需要给定磁力线的起始位置。在图中，磁力线的起始位置被设定为 $(x_i, y_j, 0)$ 其中 $x_i = -1 + 0.125(i - 1)$, $y_j = -1 + 0.125(j - 1)$ ($1 \leq i, j \leq 17$)。

由于经典无力场模型外推法提供了在计算指标下的数值，但并没有公布实

现磁场外推的代码。因此，在本文中，我们仅展示参考磁场、我们提出的无力场神经网络法对参考磁场进行外推重建的结果和由势场模型对参考磁场外推重建的结果，如图 4.3 和 4.4 所示，结果表明我们提出的无力场神经网络外推法的外推重建结果与参考磁场一致，而势场模型的外推结果与参考磁场的拓扑结果存在很大的差异。

在例子 1 中，无力场神经网络外推法外推重建的磁力线基本和参考磁场高度一致。在例子 2 中，从图中可以看出无力场神经网络外推法外推重建的磁力线基本和参考磁场一致。此外，为了验证补全计算区域边界上的伪磁场数据的必要性，如果仅使用底边界磁场数据作为微分方程的边界条件，重建磁场的磁力线如图 4.4d 所示，与参考磁场 4.4a 存在显著的不一致性。

此外，值得强调的是，无力场神经网络外推法与其他方法有所不同，包括向上积分法、磁流体力学松弛法、优化方法和边界元方法，在综合评述文章^[63] 中进行了讨论。与这些方法的区别在于，通常这些方法在计算过程中使用固定的边界条件，而我们的方法着重于利用部分边界数据学习整个边界上的磁场分布。在训练过程中，边界数据持续调整，直至与边界上的采样数据一致，以确保准确性和连贯性。

于是，我们在底边界取 $(x_i, y_j, 0)$ 其中 $x_i = -1 + \frac{2}{63}(i - 1)$, $y_j = -1 + \frac{2}{63}(j - 1)$ ($1 \leq i, j \leq 64$)。我们绘制了参考磁场和重建磁场在底边界的差异 $B_z^{\text{Rec}} - B_z^{\text{Ref}}$ 的磁场差异、各自磁场等高线和磁图，如图 4.5 和 4.6 所示，它们表明无力场神经网络外推法外重建的磁场在底边界上也能很好地拟合参考磁场。实际上无力场神经网络外推法可以重建边界上任一点的磁场，无力场神经网络外推法的外推区域是边界和外推区域的内部(包括底边界)，而经典无力场外推法只在外推区域内(不包括底边界)重建其磁场。

4.2.3.2 在太阳活动区 AR NOAA 11158 上的应用

在这一部分，我们将利用太阳动力学天文台 (SDO) 上的日震与磁场成像仪 (HMI) 观测到的太阳活动区 AR NOAA 11158 的光球磁场数据作为边界条件来验证我们提出的无力场神经网络外推法的性能。具体来说，我们将使用 hmi.sharp_cea 720s 系列产品在 2011 年 2 月 14 日 20 点的观测数据，如图 4.6 所示。

[卢经纬等^{\[91\]}](#)指出：实际应用中数据存在缺失、裂隙或噪声边界条件等问题，这些问题都极大地影响传统微分方程数值方法的求解精度。具体地讲，日冕非

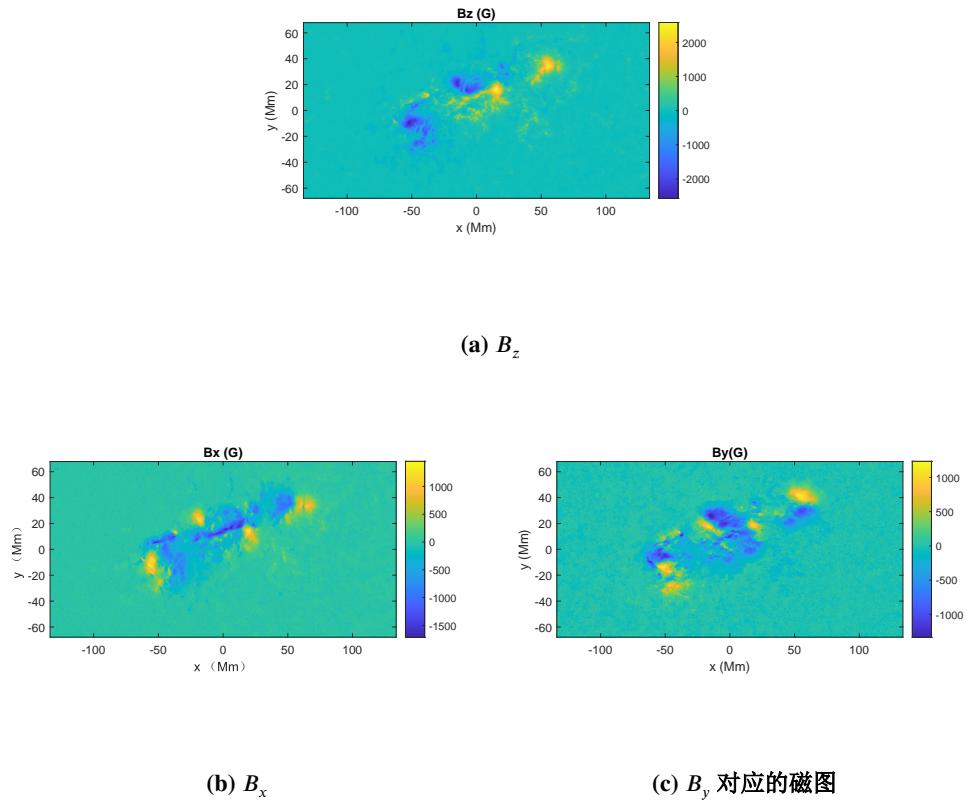
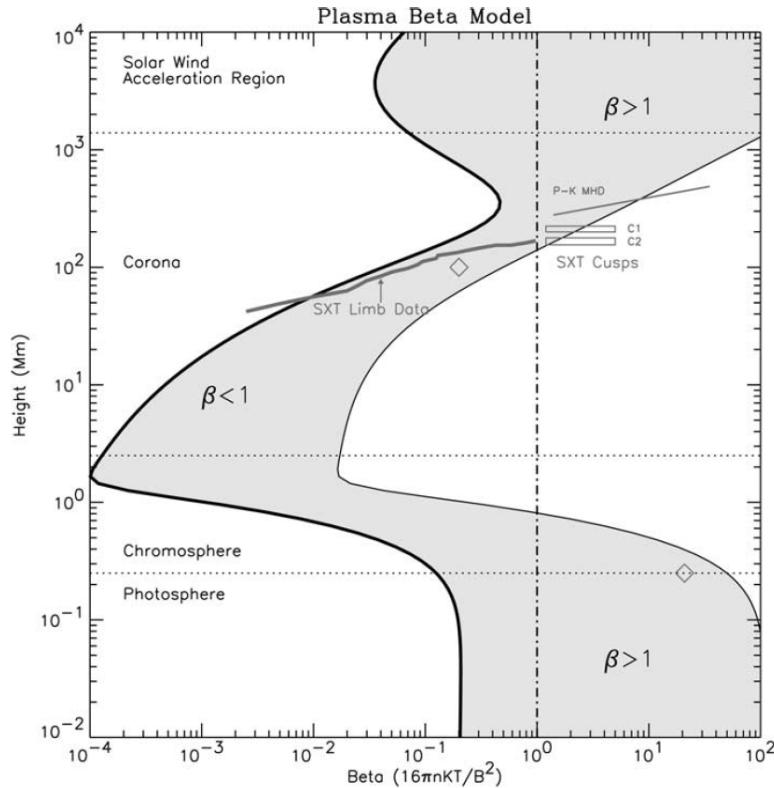


图 4.7 太阳活动区 AR NOAA 11158 的光球磁场数据可视化

线性无力场模型对应的偏微分方程边值问题 (4.9) 中, 如果边界数据 $\vec{B}(x', y', z')$ 不准确、不完整或不符合方程, 可能会导致最终方程的解出现不一致或无法得到明确定义的解。具体的影响取决于问题的性质和具体的边界条件。如果边界数据存在误差或不完整, 解可能会受到这些问题的影响而产生误差或偏离预期结果。这可能导致解的不稳定性、振荡、奇异行为等。此外, 不符合方程的边界数据可能导致方程无法满足, 从而无法得到解。因此, 在求解偏微分方程边值问题时, 确保边界数据的准确性、完整性和与方程的一致性是非常重要的。[Gary^{\[90\]}](#) 指出在光球和色球底部的等离子压磁比并不满足 $\beta \ll 1$ (如图 4.8 所示), 在此情形下等离子方程组中的动量方程 (1.11) 不能简化为无力场模型对应的方程 (3.7)。另一方面, 这也说明光球磁场数据不满足微分方程 (3.7) 的条件。

在求解偏微分方程边值问题 (4.9) 时, 确保边值数据的准确性、完整性和方程的一致性是非常重要的。如果存在不确定性或不准确性, 可以尝试通过数值方法、近似方法或修正边界数据来得到更好的解。当应用非线性无力场模型进行日冕磁场外推时, [Wiegelmann 等^{\[63\]}](#) 指出需要通过无力性边界条件的一致性准则

图 4.8 等离子压磁比 β 与高度之间的关系^[90]

修改光球观测数据。

无力性边界条件的一致性准则包括以下要求：

1. 磁通量守恒要求: $\int_S B_z(x, y, 0) dx dy = 0$, 即在光球上视向方向 (z 轴方向) 的磁场磁通量为零。
2. 无力性要求: $\frac{|F_x|}{F_p} \ll 1, \frac{|F_y|}{F_p} \ll 1, \frac{|F_z|}{F_p} \ll 1$, 其中

$$\begin{aligned} F_x &= -\frac{1}{4\pi} \int_S B_x B_z dx dy, & F_y &= -\frac{1}{4\pi} \int_S B_y B_z dx dy, \\ F_z &= \frac{1}{8\pi} \int_S (B_x^2 + B_y^2 - B_z^2) dx dy, & F_p &= \frac{1}{8\pi} \int_S (B_x^2 + B_y^2 + B_z^2) dx dy. \end{aligned} \quad (4.22)$$

在实际计算中, Moon 等^[92] 提出了衡量边界磁通量守恒的准则。当 $\frac{|F^+ - F^-|}{F^+ + F^-} \leq 0.1$ 时, 可认为在边界磁通量守恒, 表达式中 F^+ 和 F^- 分别代表磁场南北极性的磁通量。另外, Liu 等^[75] 指出, 当 $\frac{|F_z|}{F_p} < 0.1$ 时, 可以认为光球磁场满足无力性的必要条件, 其中 F_z 表示竖直方向的力, F_p 所有的磁压力。

Wiegelmann 等^[63] 指出, 优化法 (4.4) 和磁流体动力学松弛法 (4.5) 的日冕磁场外推模型需要根据上述无力性一致准则进行修改光球观测数据。需要注意的是, 我们没有修改光球磁场观测数据作为无力场神经网络外推法的输入。

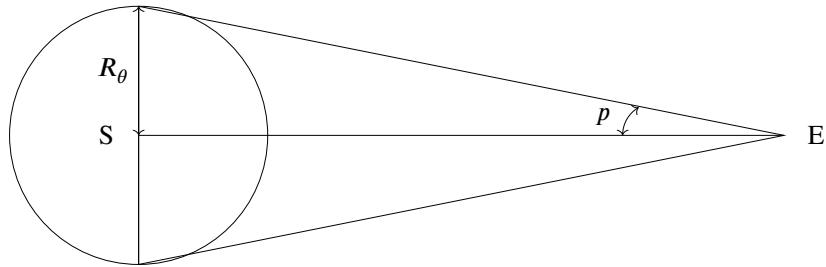
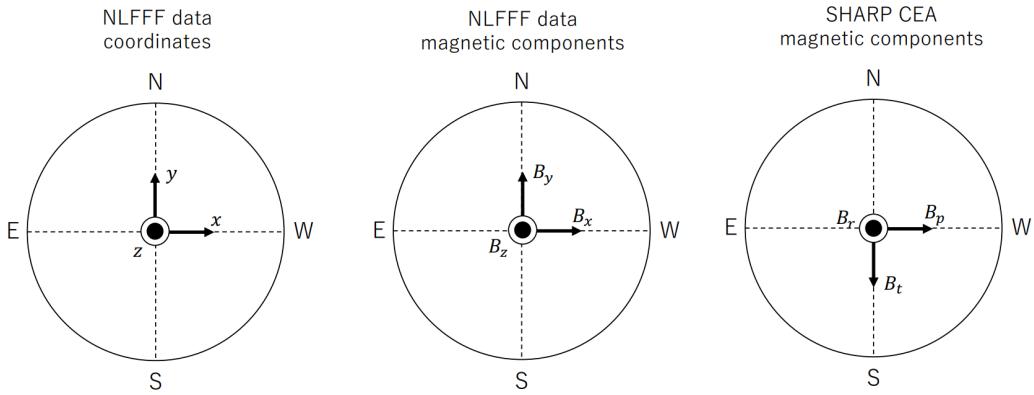


图 4.9 观测者角度日地关系示意图



$$B_x = B_p, B_y = -B_t, B_z = B_r$$

图 4.10 磁场分量对应的空间坐标示意图^[94,95]

HMI 采集磁场数据的像元分辨率为 $0.5''$ ^[93]。根据表格 1.1，我们知道太阳半径和平均日地距离分别为 $R_\theta = 6.955 \times 10^8$ 米和 $SE = 1.496 \times 10^{11}$ 米。因此，由图 4.9 知， $\tan p = \frac{R_\theta}{SE} = \frac{6.955 \times 10^8}{1.496 \times 10^{11}} \approx 4.6491 \times 10^{-3} \ll 1$ 。由此可得， $p \approx 4.6491 \times 10^{-3}$ 弧度。接下来，由弧度制转化成角度制， $p' = p \times \frac{180}{\pi} \approx 0.2265^\circ$ 。然后，将其转换为角秒， $p_0 = p' \times 3600 \approx 960''$ 。最后，我们可以得到 $960''$ 对应于 6.955×10^8 米，那么 $0.5''$ 对应于 $\frac{6.955 \times 10^8 \times 0.5}{960} \approx 3.6224 \times 10^5$ 米，约等于 0.36 兆米。

2011 年 2 月 14 日 20 点观测的数据 hmi.sharp_720s 的大小是 744×377 ，我们删除最后一列数据使得数据的尺寸为 743×377 。根据文献^[94,95] 的记法 (如图 4.10 所示)，我们得到每个像元的坐标 (x_i, y_j) ，其中

$$x_i = 0.36 \times i + \frac{743 + 1}{2} \times 0.36, \quad y_j = 0.36 \times j + \frac{377 + 1}{2} \times 0.36,$$

其中 i 是 x 方向上第 i 个像元， j 是 y 方向的第 j 个像元。空间坐标 $(x_i, y_j, 0)$ 上

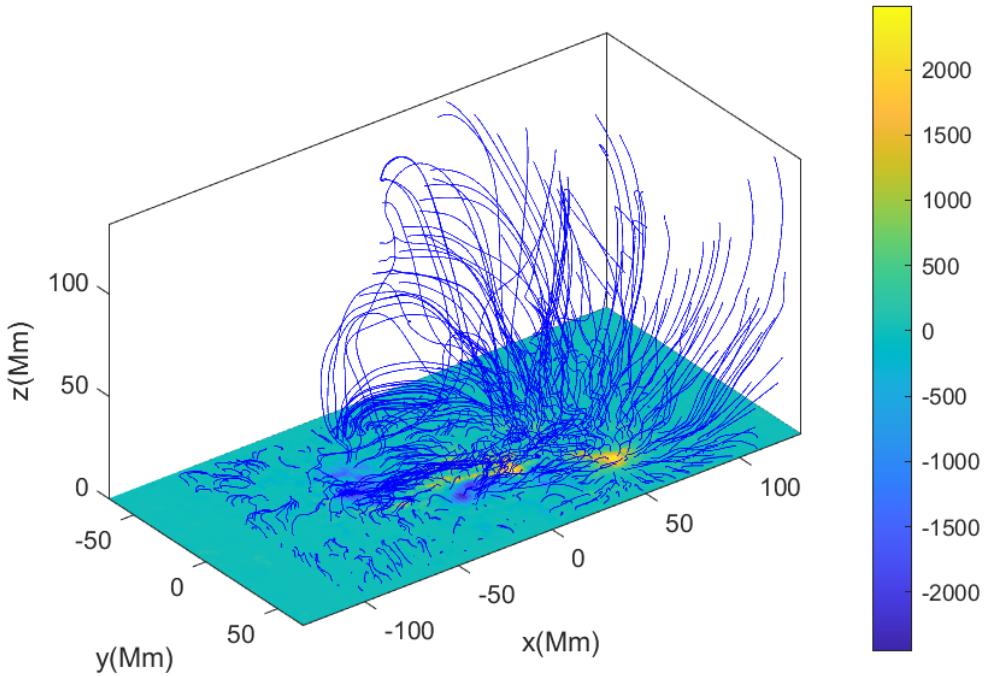


图 4.11 磁力线分布图

的磁场 \vec{B} 在坐标基上的投影分量分别为:

$$\begin{cases} B_x(x_i, y_j, 0) = B_\phi(x_i, y_j, 0) = B_p(x_i, y_j, 0), \\ B_y(x_i, y_j, 0) = -B_\theta(x_i, y_j, 0) = -B_t(x_i, y_j, 0), \\ B_z(x_i, y_j, 0) = B_r(x_i, y_j, 0) = B_r(x_i, y_j, 0). \end{cases}$$

其中 B_p 、 B_t 和 B_r 分别由 hmi.sharp_720s 提供。需要计算的磁场区域 $\bar{\Omega}$ 为:

$$\bar{\Omega} = [\min(x_i), \max(x_i)] \times [\min(y_j), \max(y_j)] \times [0, \max(z_k)],$$

其中 $z_k = 0.36 \times (k - 1)$ 。

为了 $\bar{\Omega} \subseteq [-1, 1] \times [-1, 1] \times [0, 2]$, 我们记 $u_1 = x_0 = \frac{x}{l_0}, u_2 = y_0 = \frac{y}{l_0}, u_3 = z_0 = \frac{z}{l_0}, \nabla_0 = \left[\frac{\partial}{\partial x_0}, \frac{\partial}{\partial y_0}, \frac{\partial}{\partial z_0} \right]^T$, 其中 $l_0 = \max(\max(\max(x_i), \max(y_j)), \max(z_k))$ 。根据散度的计算公式 (2.4)、旋度的计算公式 (2.9) 及在可逆变换下向量投影变换的公式 (2.18), 我们有:

$$\nabla \cdot \vec{B} = \frac{1}{l_0} \nabla_0 \cdot \vec{B}, \quad \frac{\nabla \times \vec{B} \times \vec{B}}{|\nabla \times \vec{B}| |\vec{B}|} = \frac{\nabla_0 \times \vec{B} \times \vec{B}}{|\nabla_0 \times \vec{B}| |\vec{B}|}. \quad (4.23)$$

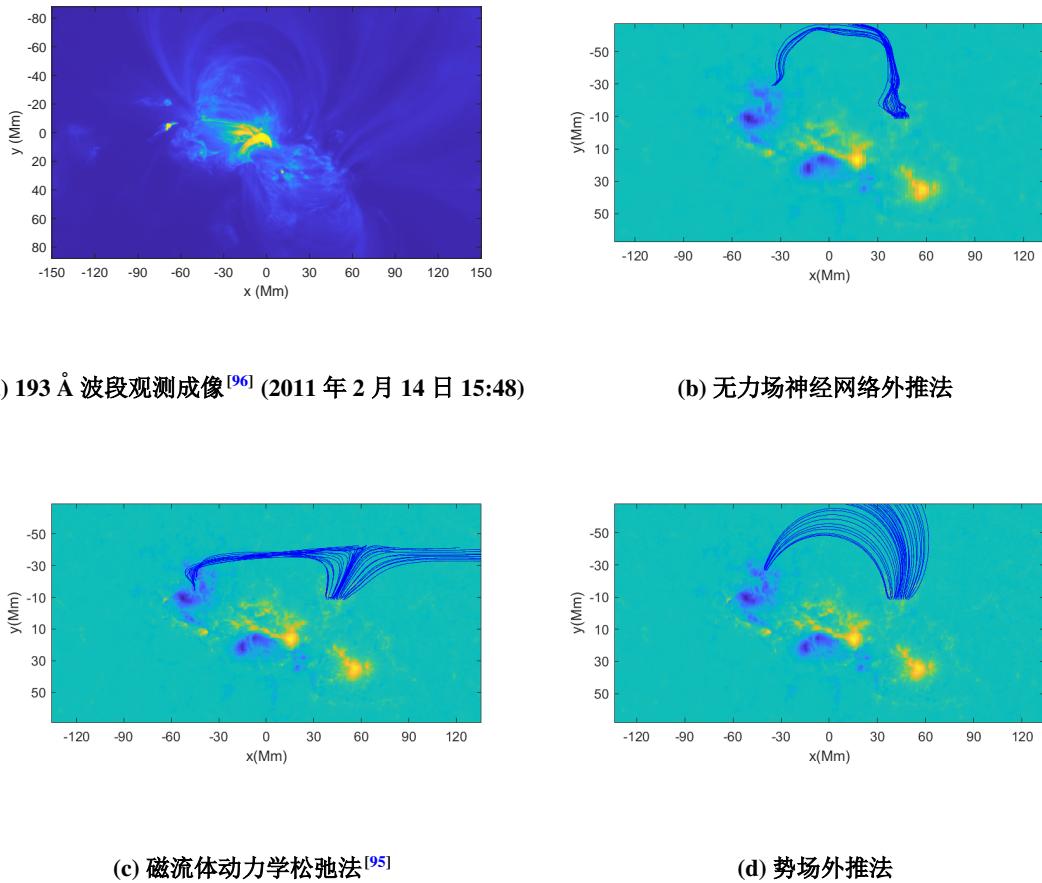


图 4.12 局部磁力线在 xoy 平面投影后与观测数据成像后比对

其中, $\nabla \cdot \vec{B}$ 表示向量场 \vec{B} 的散度, $\nabla \times \vec{B}$ 表示矢量场 \vec{B} 的旋度。根据以上公式, 我们可以使用 ∇_0 来计算散度和旋度。因此, 无力场神经网络的损失函数 $f(\vec{p})$ (见公式(4.12)) 可以等价地表示为以下形式:

$$\begin{aligned}
 f(\vec{p}) = & \lambda_0 \frac{1}{m} \sum_{i=1}^m \left\| \vec{B}^{NN}(x_{0i}', y_{0i}', z_{0i}'; \vec{p},) - \vec{B}(x_{0i}', y_{0i}', z_{0i}'; \vec{p}) \right\|_2^2 + \\
 & + \lambda_1 \frac{1}{n} \sum_{j=1}^n \left\| \frac{1}{l_0} \nabla_0 \cdot \vec{B}^{NN}(x_{0j}, y_{0j}, z_{0j}; \vec{p}) \right\|_2^2 + \\
 & + \lambda_2 \frac{1}{n} \sum_{j=1}^n \frac{\left\| \nabla_0 \times \vec{B}^{NN}(x_{0j}, y_{0j}, z_{0j}; \vec{p}) \times \vec{B}^{NN}(x_{0j}, y_{0j}, z_{0j}; \vec{p}) \right\|_2^2}{\left\| \nabla_0 \times \vec{B}^{NN}(x_{0j}, y_{0j}, z_{0j}; \vec{p}) \right\|_2^2 \cdot \left\| \vec{B}^{NN}(x_{0j}, y_{0j}, z_{0j}; \vec{p}) \right\|_2^2},
 \end{aligned}$$

其中, λ_0 、 λ_1 和 λ_2 分别为神经网络损失函数中的权重系数。

关于 AR NOAA 11158, 光球磁场的测量数据被用作底面的边界条件, 同时

使用势场模型的外推结果作为边界上另外五个面的边界条件。我们采用了一个包含 9 ($L = 7$) 层神经元的神经网络，每层包含 200 个神经元。模型的超参数设定为 $\lambda_0 = 1$, $\lambda_1 = 1$ 和 $\lambda_2 = 3$ ， m 设置为 100。我们在底边界 $\partial\Omega$ 上均匀选择了 15,625 个点，并在空间 Ω 中随机选取了 360,000 个点用于神经网络的训练。此外，在其他五个面上等间隔采样了 20,482 个点的伪磁场数据作为边界条件的一部分。图 4.11 展示了整个活动区经无力场神经网络外推法计算后的磁力线。图 4.11 展示了整个活动区经无力场神经网络外推法计算后的磁力线。图 4.12 (b) 显示了局部磁力线在 xoy 平面上的投影，这一部分与观测到的成像一致，其中图 (a) 是美国航空航天局太阳数据分析中心 (SDAC) [Dissauer 等](#) 发布的大气成像组件 (AIA) 活动区块 (AARPs) 数据集中对应的 AR NOAA 11158 的 193 Å 波段的观测成像，该区域比对应的 SHARPs 区域大一些^[96]。图 (c) 是名古屋大学宇宙地球环境研究所 [Kusano 等](#) 发布的基于磁流体动力学松法的非线性无力场外推数据集^[95] 中对应的投影结果。图 (d) 是基于势场模型外推中在 xoy 平面对应的投影结果。

4.3 本章小节

在本章中，我们提出了一种无力场神经网络外推法来重建日冕磁场。首先，我们利用万能逼近定理 (第 2 章定理 2.5) 给出了无力场神经网络外推法的数学模型 (公式 4.11)。接下来，我们定义了无力场神经网络的损失函数 (公式 4.12)。为了避免在机器学习神经网络外进行数据的归一化和逆归一化操作，我们设计了一种不同于经典机器学习神经网络的结构，其中包含了准输出层 (如图 4.1 所示) 的无力场神经网络。然后，我们采用了二阶优化算法牛顿法来优化无力场神经网络的网络参数，相比经典的一阶算法，这种方法更加高效。最后，我们首先在由经典的一类半解析解生成的轴对称无力场模型 (第 3 章公式 3.23) 上进行评估了我们的外推重建磁场算法的性能，通过与参考磁场以及其他经典无力场外推方法得到的结果进行比较，验证了我们的方法的准确性和可靠性。接着，我们将无力场神经网络外推法应用于在 AR NOAA 11158 上。

第 5 章 总结及展望

5.1 总结

在第 1 章中，我们介绍了太阳物理中的相关概念，并通过等离子体中的动量方程（公式 1.11）对日冕磁场外推法展开了相关讨论。

在第 2 章中，我们介绍了无力场神经网络外推法所需的基本数学知识。这包括在一般正交坐标基下的梯度、散度、旋度以及它们之间的关系（公式 2.3、2.4、2.9 和定理 2.3）。我们还介绍了函数逼近定理（定理 2.5），该定理说明了机器学习方法如何与求解微分方程的数值解联系起来。通过函数逼近定理，我们可以利用机器学习方法来求解微分方程数值解。

在第 3 章中，我们提出了一种用于求解一类日冕磁场半解析解（公式 3.23）的数值算法。根据该半解析解中的参数满足 Low & Lou 方程（公式 3.24），首先，在给定参数 n 的情况下，我们通过参数打靶法求得参数 a 的数值。接下来，在参数 n 和 a 已知的情况下，我们使用机器学习算法计算定义域上任意位置对应的 F 和 $\frac{dF}{d\mu}$ 。这与传统的经典数值方法不同，传统方法只能直接得到计算节点上的函数值。此外，根据给定的 F 值（无论是直接给定还是间接给定），我们可以利用数据驱动的方法求解 Low & Lou 方程（公式 3.24）或者一类日冕磁场半解析解（公式 3.23）中的参数 n 和 a 的数值。

在第 4 章中，我们提出了无力场神经网络外推法用来求解日冕无力场模型所对应的偏微分方程组 $(\nabla \times \vec{B}) \times \vec{B} = \vec{0}, \nabla \cdot \vec{B} = 0$ 。首先，根据函数逼近定理（定理 2.5），我们给出了无力场神经网络的数学模型（公式 4.11）。然后，我们介绍了无力场神经网络外推法中神经网络所对应的损失函数（公式 4.12）。为了保持数据遵循的物理规律，我们在无力场神经网络的结构中引入了准输出层（如图 4.1 所示）的概念。同时，我们使用二阶优化技术-拟牛顿法求解神经网络的表达式，相比于经典机器学习神经网络的一阶优化算法，这种优化方法更加高效。

5.2 展望

机器学习和科学计算是两个密切相关且互相促进的领域。科学计算旨在使用数值方法和计算机技术解决科学和工程问题，例如求解偏微分方程、数值线性

代数问题等，而机器学习则是一种通过从数据中学习模式和规律来进行预测、分类、聚类等任务的方法。

科学计算通常需要基于已知的物理方程建立模型。然而在某些情况下精确的物理方程可能不易获取或难以描述复杂现象，这时机器学习可以用于数据驱动的建模，通过对数据进行训练和学习从中发现隐藏的模式和规律并用于求解科学计算问题。

在科学计算中常常处理高维数据，机器学习提供了强大的特征提取和降维技术，可以帮助人们从复杂数据中提取出最具代表性的特征，从而减少数据维度进而简化科学计算的问题。

机器学习算法通常需要大量的计算资源和处理能力，科学计算对高性能计算的需求非常强烈，科学计算涉及大量的数据和计算量，这为机器学习提供了强大的计算基础。高性能计算平台可以加速机器学习模型的训练和推断过程，从而提高机器学习在科学计算中的效率和应用范围。

科学计算中的数据往往存在噪声和不完整性，机器学习提供了丰富的数据预处理和噪声处理技术，例如去噪、插值等可以帮助科学家更好地处理和利用数据并提高科学计算的准确性和可靠性。

科学计算中通常需要调整模型参数和选择合适的模型来获得最佳结果，机器学习中的优化方法和模型选择技术可以为科学计算提供指导，从而帮助人们找到最优的参数配置和模型选择进而提高科学计算的效果和精度。

在太阳磁场相关问题上，我们将继续努力，并在以下方面进行进一步研究：

1. 推广至有力场神经网络外推法：我们计划将无力场神经网络外推法推广到有力（洛伦兹力）场神经网络外推法。这样可以拓展磁流体静力学模型的应用范围，使其更加适用于描述太阳磁场。

2. 重构日冕三维磁场的时间演化：针对连续时间观测的光球磁场，我们打算利用无力场神经网络外推法来重构日冕三维磁场在时间上的演化过程。同时，结合大气成像组件（AIA）在太阳活动区的数据 AARPs^[96]，我们将验证和提高无力场神经网络外推法的计算精度。这项工作有助于更好地理解太阳活动的变化。

3. 发现隐藏的物理规律并应用于科学计算：基于磁流体动力学方程，我们将通过对数据进行训练和学习，寻找隐藏的物理规律，如磁能释放机制、耀斑爆

发和日冕物质抛射的形成机制等。这些发现可以应用于解决实际的科学计算问题，推动太阳物理学领域的进一步发展。

通过在上述方面的努力，我们期待能够深入研究太阳磁场，并对太阳活动中日冕物质抛射等现象提供合理的诠释。

附录 A \mathbb{R}^3 中常见坐标系下的向量分析公式

直角坐标系 (x, y, z) 下常见的向量分析公式

1. 基底向量

$$\vec{e}_x = (1, 0, 0)^T = \vec{i}, \quad \vec{e}_y = (0, 1, 0)^T = \vec{j}, \quad \vec{e}_z = (0, 0, 1)^T = \vec{k}. \quad (\text{A.1})$$

2. 基向量的归一化因子

$$h_x = 1, \quad h_y = 1, \quad h_z = 1. \quad (\text{A.2})$$

3. 实值函数 f 的梯度

$$\nabla f = \frac{\partial f}{\partial x} \vec{e}_x + \frac{\partial f}{\partial y} \vec{e}_y + \frac{\partial f}{\partial z} \vec{e}_z. \quad (\text{A.3})$$

4. 向量 $\vec{F} = F_x \vec{e}_x + F_y \vec{e}_y + F_z \vec{e}_z$ 的散度

$$\nabla \cdot \vec{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}. \quad (\text{A.4})$$

5. 向量 $\vec{F} = F_x \vec{e}_x + F_y \vec{e}_y + F_z \vec{e}_z$ 的旋度

$$\begin{aligned} \nabla \times \vec{F} &= \left[\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \right] \vec{e}_x + \left[\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \right] \vec{e}_y + \left[\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right] \vec{e}_z, \\ &= \begin{vmatrix} \vec{e}_x & \vec{e}_y & \vec{e}_z \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix}. \end{aligned} \quad (\text{A.5})$$

圆柱坐标系 (r, θ, z) 下常见的向量分析公式

1. 与直角坐标系的关系

$$x = r \cos \theta, \quad y = r \sin \theta, \quad z = z. \quad (\text{A.6})$$

2. 基底向量

$$\vec{e}_r = \cos \theta \vec{i} + \sin \theta \vec{j}, \quad \vec{e}_\theta = -\sin \theta \vec{i} + \cos \theta \vec{j}, \quad \vec{e}_z = \vec{k}. \quad (\text{A.7})$$

3. 基向量的归一化因子

$$h_r = 1, \quad h_\theta = r, \quad h_z = 1. \quad (\text{A.8})$$

4. 实值函数 f 的梯度

$$\nabla f = \frac{\partial f}{\partial r} \vec{e}_r + \frac{1}{r} \frac{\partial f}{\partial \theta} \vec{e}_\theta + \frac{\partial f}{\partial z} \vec{e}_z. \quad (\text{A.9})$$

5. 向量 $\vec{F} = F_r \vec{e}_r + F_\theta \vec{e}_\theta + F_z \vec{e}_z$ 的散度

$$\nabla \cdot \vec{F} = \frac{1}{r} \frac{\partial}{\partial r} (r F_r) + \frac{1}{r} \frac{\partial}{\partial \theta} F_\theta + \frac{\partial}{\partial z} F_z. \quad (\text{A.10})$$

6. 向量 $\vec{F} = F_r \vec{e}_r + F_\theta \vec{e}_\theta + F_z \vec{e}_z$ 的旋度为

$$\begin{aligned} \nabla \times \vec{F} &= \frac{1}{r} \left[\frac{\partial F_z}{\partial \theta} - r \frac{\partial F_\theta}{\partial z} \right] \vec{e}_r + \left[\frac{\partial F_r}{\partial z} - \frac{\partial F_z}{\partial r} \right] \vec{e}_\theta + \frac{1}{r} \left[\frac{\partial}{\partial r} (r F_\theta) - \frac{\partial F_r}{\partial \theta} \right] \vec{e}_z, \\ &= \frac{1}{r} \begin{vmatrix} \vec{e}_r & r \vec{e}_\theta & \vec{e}_z \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial z} \\ F_r & r F_\theta & F_z \end{vmatrix}. \end{aligned} \quad (\text{A.11})$$

球坐标系 (r, θ, ϕ) 下常见的向量分析公式

1. 与直角坐标系的关系

$$x = r \sin \theta \cos \phi, \quad y = r \sin \theta \sin \phi, \quad z = r \cos \theta. \quad (\text{A.12})$$

2. 基底向量

$$\begin{aligned} \vec{e}_r &= \frac{x \vec{i} + y \vec{j} + z \vec{k}}{\sqrt{x^2 + y^2 + z^2}} = \frac{\vec{r}}{|\vec{r}|}, \\ \vec{e}_\theta &= \frac{1}{\sqrt{x^2 + y^2} \sqrt{x^2 + y^2 + z^2}} \left(z x \vec{i} + y z \vec{j} - (x^2 + y^2) \vec{k} \right), \\ \vec{e}_\phi &= \frac{1}{\sqrt{x^2 + y^2}} (-y \vec{i} + x \vec{j}). \end{aligned} \quad (\text{A.13})$$

3. 基向量的归一化因子

$$h_r = 1, \quad h_\theta = r, \quad h_\phi = r \sin \theta. \quad (\text{A.14})$$

4. 实值函数 f 的梯度

$$\nabla f = \frac{\partial f}{\partial r} \vec{e}_r + \frac{1}{r} \frac{\partial f}{\partial \theta} \vec{e}_\theta + \frac{1}{r \sin \theta} \frac{\partial f}{\partial \phi} \vec{e}_\phi. \quad (\text{A.15})$$

5. 向量 $\vec{F} = F_r \vec{e}_r + F_\theta \vec{e}_\theta + F_\phi \vec{e}_\phi$ 的散度为

$$\nabla \cdot \vec{F} = \frac{1}{r^2} \frac{\partial (r^2 F_r)}{\partial r} + \frac{1}{r \sin \theta} \frac{\partial (F_\theta \sin \theta)}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial F_\phi}{\partial \phi}. \quad (\text{A.16})$$

6. 向量 $\vec{F} = F_r \vec{e}_r + F_\theta \vec{e}_\theta + F_\phi \vec{e}_\phi$ 的旋度为

$$\begin{aligned}\nabla \times \vec{F} &= \frac{1}{r^2 \sin \theta} \left[r \frac{\partial}{\partial \theta} (\sin \theta F_\phi) - r \frac{\partial^2 F_\theta}{\partial \phi} \right] \vec{e}_r + \dots \\ &\quad \dots + \frac{1}{r \sin \theta} \left[\frac{\partial F_r}{\partial \phi} - \sin \theta \frac{\partial}{\partial r} (r F_\phi) \right] \vec{e}_\theta + \frac{1}{r} \left[\frac{\partial}{\partial r} (r F_\theta) - \frac{\partial F_r}{\partial \theta} \right] \vec{e}_\phi, \quad (\text{A.17}) \\ &= \frac{1}{r^2 \sin \theta} \begin{vmatrix} \vec{e}_r & r \vec{e}_\theta & r \sin \theta \vec{e}_\phi \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial \phi} \\ F_r & r F_\theta & r \sin \theta F_\phi \end{vmatrix}.\end{aligned}$$

附录 B 一般标准正交基经可逆变换后向量投影变换公式的证明

证明. 取 $\vec{e}_i = \frac{1}{h_i} \frac{\partial(u_1\vec{e}_1 + u_2\vec{e}_2 + u_3\vec{e}_3)}{\partial u_i}$, $h_i = \left| \frac{\partial(u_1\vec{e}_1 + u_2\vec{e}_2 + u_3\vec{e}_3)}{\partial u_i} \right|$, $i = 4, 5, 6$ 。则我们有

$$\begin{bmatrix} \vec{e}_4 \\ \vec{e}_5 \\ \vec{e}_6 \end{bmatrix} = \begin{bmatrix} \frac{1}{h_4} \frac{\partial u_1}{\partial u_4} & \frac{1}{h_4} \frac{\partial u_2}{\partial u_4} & \frac{1}{h_4} \frac{\partial u_3}{\partial u_4} \\ \frac{1}{h_5} \frac{\partial u_1}{\partial u_5} & \frac{1}{h_5} \frac{\partial u_2}{\partial u_5} & \frac{1}{h_5} \frac{\partial u_3}{\partial u_5} \\ \frac{1}{h_6} \frac{\partial u_1}{\partial u_6} & \frac{1}{h_6} \frac{\partial u_2}{\partial u_6} & \frac{1}{h_6} \frac{\partial u_3}{\partial u_6} \end{bmatrix} \begin{bmatrix} \vec{e}_1 \\ \vec{e}_2 \\ \vec{e}_3 \end{bmatrix}. \quad (\text{B.1})$$

向量 \vec{B} 在不同标准正交基下可表示为

$$\begin{aligned} \vec{B} &= \begin{bmatrix} B_1 & B_2 & B_3 \end{bmatrix} \begin{bmatrix} \vec{e}_1 \\ \vec{e}_2 \\ \vec{e}_3 \end{bmatrix} = \begin{bmatrix} B_4 & B_5 & B_6 \end{bmatrix} \begin{bmatrix} \vec{e}_4 \\ \vec{e}_5 \\ \vec{e}_6 \end{bmatrix} \\ &= \begin{bmatrix} B_4 & B_5 & B_6 \end{bmatrix} \begin{bmatrix} \frac{1}{h_4} \frac{\partial u_1}{\partial u_4} & \frac{1}{h_4} \frac{\partial u_2}{\partial u_4} & \frac{1}{h_4} \frac{\partial u_3}{\partial u_4} \\ \frac{1}{h_5} \frac{\partial u_1}{\partial u_5} & \frac{1}{h_5} \frac{\partial u_2}{\partial u_5} & \frac{1}{h_5} \frac{\partial u_3}{\partial u_5} \\ \frac{1}{h_6} \frac{\partial u_1}{\partial u_6} & \frac{1}{h_6} \frac{\partial u_2}{\partial u_6} & \frac{1}{h_6} \frac{\partial u_3}{\partial u_6} \end{bmatrix} \begin{bmatrix} \vec{e}_1 \\ \vec{e}_2 \\ \vec{e}_3 \end{bmatrix}, \end{aligned} \quad (\text{B.2})$$

则在不同标准正交基下投影之间的关系为

$$\begin{bmatrix} B_1 & B_2 & B_3 \end{bmatrix} = \begin{bmatrix} B_4 & B_5 & B_6 \end{bmatrix} \begin{bmatrix} \frac{1}{h_4} \frac{\partial u_1}{\partial u_4} & \frac{1}{h_4} \frac{\partial u_2}{\partial u_4} & \frac{1}{h_4} \frac{\partial u_3}{\partial u_4} \\ \frac{1}{h_5} \frac{\partial u_1}{\partial u_5} & \frac{1}{h_5} \frac{\partial u_2}{\partial u_5} & \frac{1}{h_5} \frac{\partial u_3}{\partial u_5} \\ \frac{1}{h_6} \frac{\partial u_1}{\partial u_6} & \frac{1}{h_6} \frac{\partial u_2}{\partial u_6} & \frac{1}{h_6} \frac{\partial u_3}{\partial u_6} \end{bmatrix}. \quad (\text{B.3})$$

□

附录 C 函数相关定理的证明

假设 $f(x, y)$ 和 $g(x, y)$ 是定义在 $(0, 0)$ 点某个邻域上的实值函数，其中函数 g 满足 $g(0, 0) = 0, \frac{\partial g}{\partial y} \Big|_{x=0, y=0} \neq 0$ 。若条件

$$\left| \frac{\partial(f, g)}{\partial(x, y)} \right| = 0, \quad \forall (x, y) \in U,$$

成立，则存在一个定义原点附近的连续函数 $h : t \rightarrow h(t)$ 使得

$$f(x, y) = h(g(x, y)), \quad \forall (x, y) \in U.$$

证明。引入辅助函数 $F(x, y, t) = g(x, y) - t$ ，则函数 F 满足

$$F(0, 0, 0) = 0, \quad \frac{\partial F(x, y, t)}{\partial y} \Big|_{x=0, y=0, t=0} = \frac{\partial g}{\partial y} \Big|_{x=0, y=0, t=0} \neq 0.$$

根据文献^[57] 中的隐函数定理知：方程 $F(x, y, t) = 0$ 决定唯一的隐函数 $\psi(x, t) = y$ 且 $F(x, \psi(x, t), t) \equiv 0$ 。 $F(x, \psi(x, t), t) \equiv 0$ 意味着

$$F(x, \psi(x, t), t) \equiv 0 \Leftrightarrow g(x, \psi(x, t)) - t \equiv 0,$$

则当对 x 求导数时，我们有

$$\frac{\partial g(x, \psi(x, t))}{\partial x} + \frac{\partial g(x, \psi(x, t))}{\partial \psi} \frac{\partial \psi(x, t)}{\partial x} = 0 \Rightarrow \frac{\partial \psi(x, t)}{\partial x} = -\frac{\frac{\partial g(x, \psi(x, t))}{\partial x}}{\frac{\partial g(x, \psi(x, t))}{\partial \psi}}.$$

记 $\tilde{h}(x, t) = f(x, \psi(x, t))$ ，则

$$\begin{aligned} & \frac{\partial \tilde{h}(x, t)}{\partial x} \\ &= \frac{\partial f(x, \psi(x, t))}{\partial x} + \frac{\partial f(x, \psi(x, t))}{\partial \psi} \frac{\partial \psi(x, t)}{\partial x} \\ &= \frac{\partial f(x, \psi(x, t))}{\partial x} - \frac{\partial f(x, \psi(x, t))}{\partial \psi} \frac{\frac{\partial g(x, \psi(x, t))}{\partial x}}{\frac{\partial g(x, \psi(x, t))}{\partial \psi}} \\ &= \left(\frac{1}{\frac{\partial g(x, \psi(x, t))}{\partial \psi}} \right) \left(\frac{\partial f(x, \psi(x, t))}{\partial x} \frac{\partial g(x, \psi(x, t))}{\partial \psi} - \frac{\partial f(x, \psi(x, t))}{\partial \psi} \frac{\partial g(x, \psi(x, t))}{\partial x} \right) \\ &= 0. \end{aligned}$$

即函数 $f(x, \psi(x, t))$ 与 x 无关。那么 $h(t) = f(x, \psi(x, t))$ 。由上述隐函数定理的结果知 $g(x, y) = t$ 和 $y = \psi(x, t)$, 则最后我们有

$$f(x, y) = h(g(x, y)).$$

□

附录 D 机器学习算法在二阶常微分方程不同初边问题上的应用

通过调整我们模型中相应的损失函数项，可以求解文献^[73] 中的微分方程：
 $(1 - \mu^2) \frac{d^2 A_n}{d\mu^2} + n(n+1) A_n + \lambda A_n (A_n^2)^{\frac{1}{n}} = 0$, 其中 n 和 $\lambda = \frac{n+1}{n} a^2$ 是参变量. 边值条件是 $A_n(-1) = 0$, $A_n(1) = 0$ 。另外，初始条件是

$$\begin{cases} A_n(0) = 1, A'_n(0) = 0; & \text{当 } A_n(\mu) \text{ 是偶函数时,} \\ A_n(0) = 0, A'_n(0) = 1; & \text{当 } A_n(\mu) \text{ 是奇函数时.} \end{cases}$$

该微分方程数值解问题对应的损失函数为：

$$Loss = \mathcal{L}(\mu; \theta) = \lambda_1 \mathcal{L}_1(\mu; \theta) + \lambda_2 \mathcal{L}_2(\mu; \theta) + \lambda_3 \mathcal{L}_3(\mu; \theta), \text{ 其中}$$

$$\mathcal{L}_1(\mu; \theta) = \frac{1}{M} \sum_{\mu=\mu_j, j=1}^{j=M} \left((1 - \mu^2) \frac{df_2}{d\mu} + n(n+1) f_1 + \lambda f_1 (f_1^2)^{\frac{1}{n}} \right)^2,$$

$$\mathcal{L}_2(\mu; \theta) = \frac{1}{M} \sum_{\mu=\mu_j, j=1}^{j=M} \left(\frac{df_1}{d\mu} - f_2 \right)^2,$$

$$\mathcal{L}_3(\mu; \theta) = \begin{cases} (f_1(0; \theta) - 1)^2 + (f_2(0; \theta))^2, & \text{当 } f_1 \text{ 是偶函数时,} \\ (f_1(0; \theta))^2 + (f_2(0; \theta) - 1)^2, & \text{当 } f_1 \text{ 是奇函数时.} \end{cases}$$

首先，我们得到数值求解 λ 的值，结果列在了 D.1 中. 我们的结果和^[73] 中的结果一致，其数值差异可能因计算节点不同或机器精度不同造成。然后，由机器学习算法得到 A_n, A'_n 的值，结果绘制成图 D.1 – D.3，我们的结果和^[73] 十分吻合，我们的算法可以直接得到定义域上任意点的值，然而文献^[73] 中的方法不能。

表 D.1 不同方法得到 λ 数值结果的比较

(a) 当 A_n 是偶函数时

| $m \setminus n = 0.5$ | Lerche & Low | Ours | $m \setminus n = 1$ | Lerche & Low | Ours | $m \setminus n = 4$ | Lerche & Low | Ours |
|-----------------------|--------------|---------|---------------------|--------------|--------------------------|---------------------|--------------|---------|
| 1 | 2.47 | 2.4702 | 1 | 0.0 | 3.7290×10^{-10} | 1 | 12.24 | 12.2523 |
| 2 | 28.69 | 28.6935 | 2 | 17.9 | 17.8673 | 2 | 44.45 | 44.4538 |
| 3 | 79.7 | 79.6805 | 3 | 52.5 | 52.4641 | 3 | -- | 86.8627 |

(b) 当 A_n 是奇函数时

| $m \setminus n = 0.5$ | Lerche & Low | Ours | $m \setminus n = 1$ | Lerche & Low | Ours | $m \setminus n = 4$ | Lerche & Low | Ours |
|-----------------------|--------------|------------|---------------------|--------------|----------|---------------------|--------------|----------|
| 1 | 271 | 270.9715 | 1 | 34.3 | 34.2966 | 1 | 68.5 | 68.5917 |
| 2 | 15700 | 15714.5875 | 2 | 603.0 | 602.9976 | 2 | 186.75 | 186.9304 |

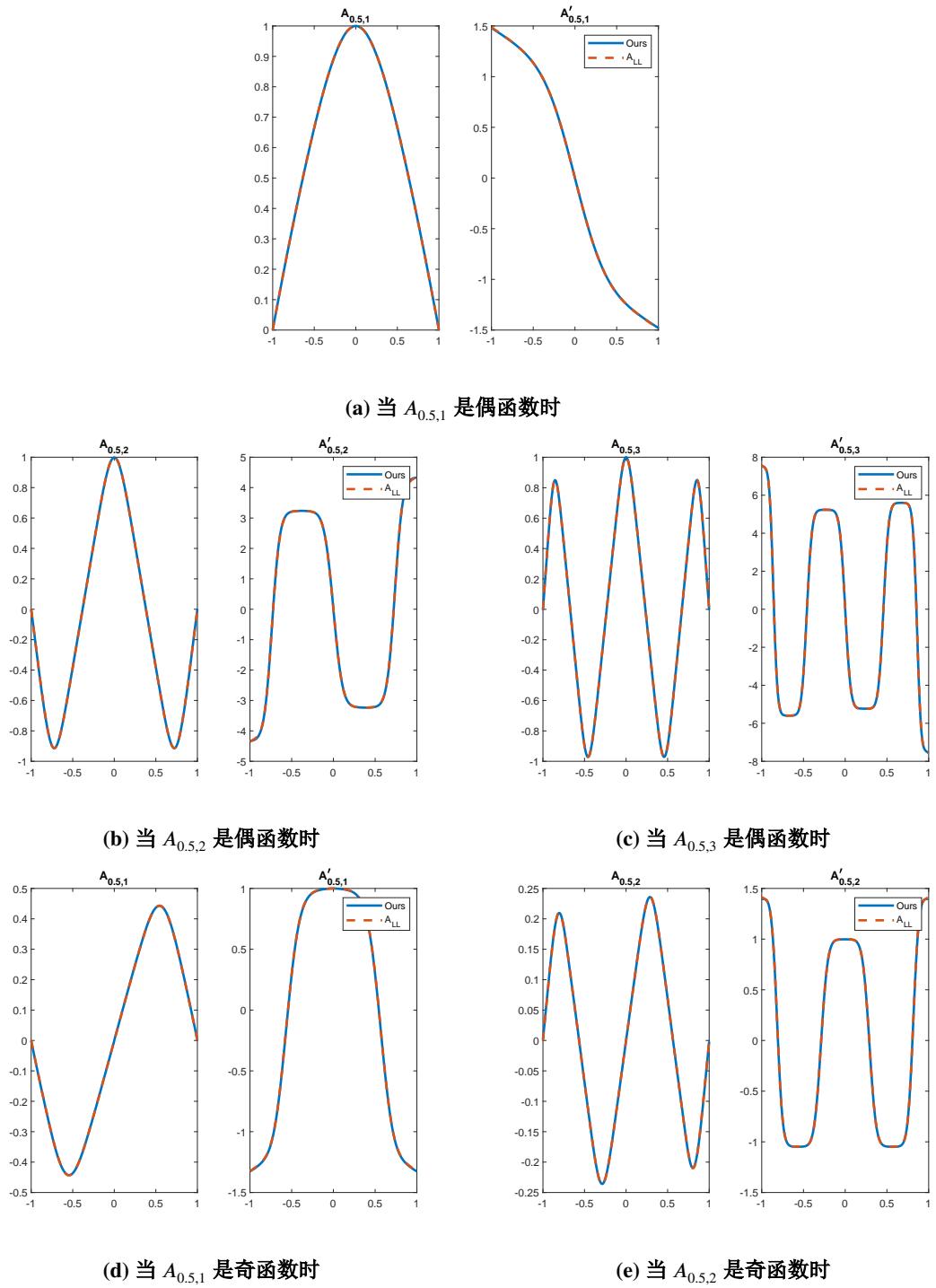


图 D.1 当 $n = 0.5$ 时，不同方法得到 A_n 和 A'_n 的结果比较

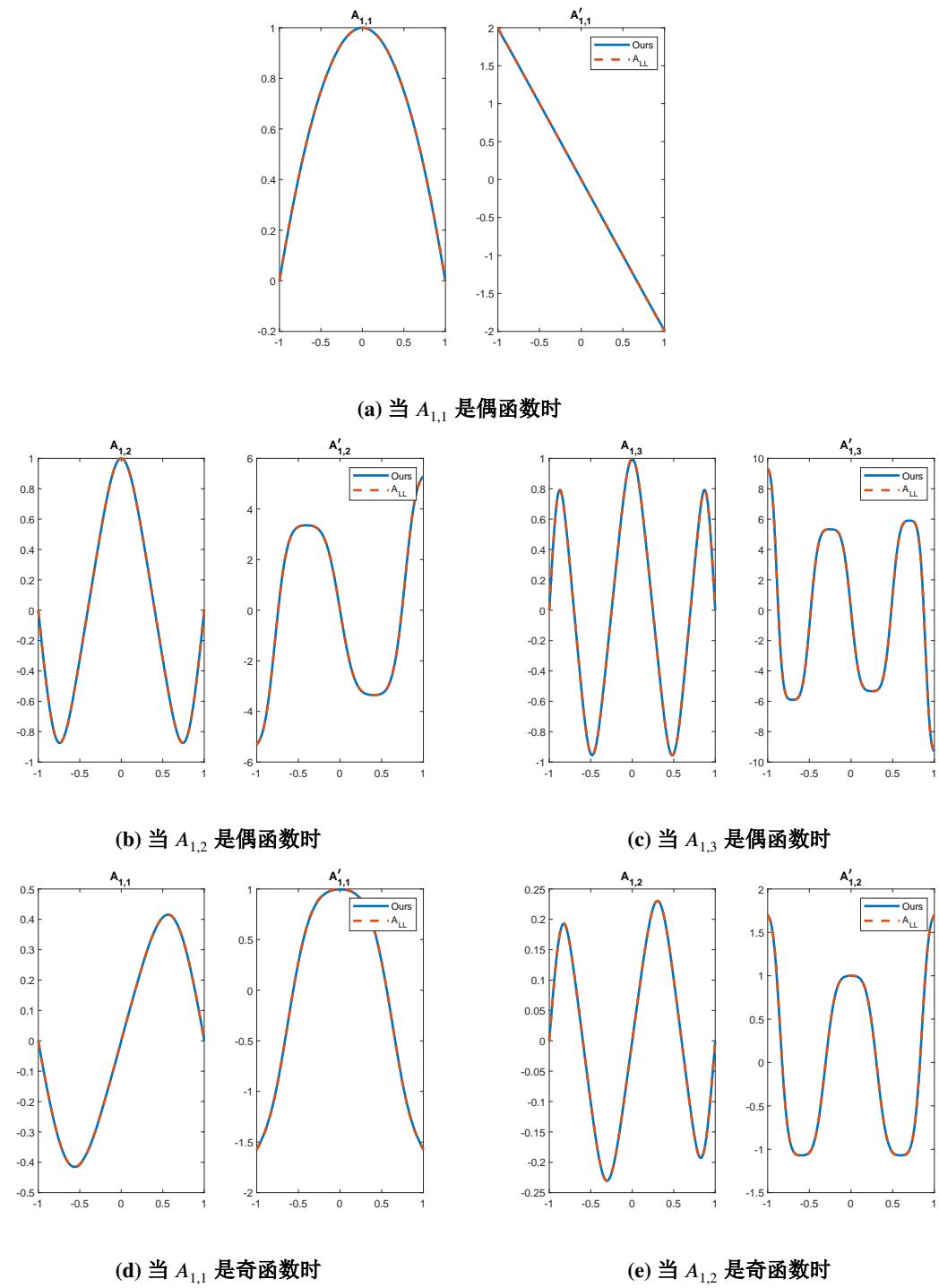
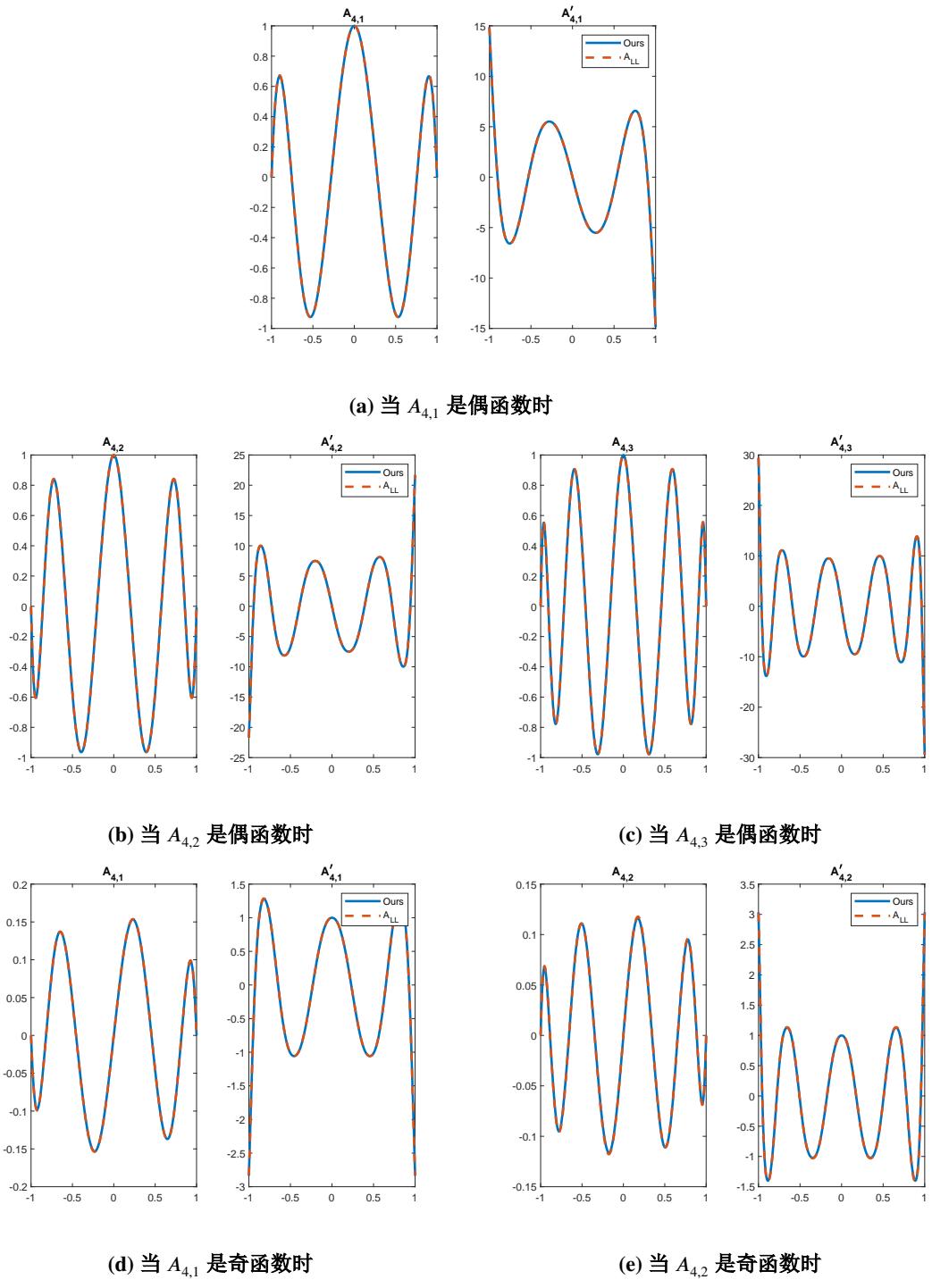


图 D.2 当 $n = 0.5$ 时, 不同方法得到 A_n 和 A'_n 的结果比较

图 D.3 当 $n = 4$ 时，不同方法得到 A_n 和 A'_n 的结果比较

附录 E 一种三次插值函数的构造方法

构造三次函数 $f(x)$, 其中 $f(x)$ 满足插值条件: $f(x_0) = y_0$ 、 $f(x_1) = y_1$ 、 $f(x_2) = y_2$ 和 $f'(x_0) = y'_0 \neq 0$ 。具体的过程如下:

1. 基于拉格朗日插值格式构造如下含待定系数 a 的插值函数,

$$f(x) = l_0(x)y_0 + l_1(x)y_1 + l_2(x)y_2 + ay'_0(x - x_0)(x - x_1)(x - x_2),$$

其中

$$l_0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)}, l_1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)}, l_2(x) = \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)}.$$

2. 确定插值函数中的待定系数, 关于 x 对 $f(x)$ 求导

$$\begin{aligned} f'(x) &= l'_0(x)y_0 + l'_1(x)y_1 + l'_2(x)y_2 + \dots \\ &\quad \dots + ay'_0[(x - x_1)(x - x_2) + (x - x_0)(x - x_2) + (x - x_0)(x - x_1)]. \end{aligned}$$

把 $x = x_0$ 代入上式且使用插值信息 $f'(x_0) = y'_0$, 则我们有

$$\begin{aligned} y'_0 &= \frac{(x_0 - x_1) + (x_0 - x_2)}{(x_0 - x_1)(x_0 - x_2)}y_0 + \frac{(x_0 - x_2)}{(x_1 - x_0)(x_1 - x_2)}y_1 + \frac{(x_0 - x_1)}{(x_2 - x_0)(x_2 - x_1)}y_2 + \dots \\ &\quad \dots + ay'_0(x_0 - x_1)(x_0 - x_2). \end{aligned}$$

最后求得待定系数 a 为

$$a = \frac{y'_0 - \frac{(x_0 - x_1) + (x_0 - x_2)}{(x_0 - x_1)(x_0 - x_2)}y_0 - \frac{(x_0 - x_2)}{(x_1 - x_0)(x_1 - x_2)}y_1 - \frac{(x_0 - x_1)}{(x_2 - x_0)(x_2 - x_1)}y_2}{y'_0(x_0 - x_1)(x_0 - x_2)}.$$

参考文献

- [1] 中央大学应用地质研究所. 太阳系的历史 [EB/OL]. <http://140.115.123.30/earth/age/solar.html#BM4>.
- [2] 陈洁. 太阳的一生 [Z]. 2018.
- [3] Priest E. Magnetohydrodynamics of the sun [M]. Cambridge University Press, 2014.
- [4] Lang K. The sun from space [M]. Springer-Verlag Berlin Heidelberg, 2009.
- [5] 李向东. 恒星的结构 [Z]. 2004.
- [6] Vernazza J, Avrett E, Loeser R. Structure of the solar chromosphere. III - models of the euv brightness components of the quiet-sun [J]. Astrophysical Journal Supplement Series, 1981, 45: 635-725.
- [7] Wikipedia. Solar granule [EB/OL]. https://en.wikipedia.org/wiki/Solar_granule.
- [8] SpaceWeatherLive. What are sunspots? [Z].
- [9] Wikipedia. Limb darkening [EB/OL]. https://en.wikipedia.org/wiki/Limb_darkening.
- [10] Wikipedia. Solar spicule [EB/OL]. https://en.wikipedia.org/wiki/Solar_spicule.
- [11] Wikipedia. Solar flare [EB/OL]. https://en.wikipedia.org/wiki/Solar_flare.
- [12] Tsuneta S, Hara H, Shimizu T, et al. Observation of a solar flare at the limb with the yohkoh soft x-ray telescope [J]. Publications of the Astronomical Society of Japan, 1992, 44: L63-L69.
- [13] Wikipedia. Stellar corona [EB/OL]. https://en.wikipedia.org/wiki/Stellar_corona.
- [14] Wikipedia. Coronal mass ejection [EB/OL]. https://en.wikipedia.org/wiki/Coronal_mass_ejection.
- [15] Wikipedia. Coronal hole [EB/OL]. https://en.wikipedia.org/wiki/Coronal_hole.
- [16] Song H, Chen Y, Li G, et al. Coalescence of macroscopic magnetic islands and electron acceleration from stereo observation [J]. Physical Review X, 2012, 2: 021015.
- [17] Forbes T. A review on the genesis of coronal mass ejections [J]. Journal of Geophysical Research, 2000, 105: 23153-23166.
- [18] Wikipedia. Active region [EB/OL]. https://en.wikipedia.org/wiki/Active_region.
- [19] SolarMonitor. 17 march 2023 [Z].
- [20] Zanna G. The multi-thermal emission in solar active regions [J]. Astronomy & Astrophysics, 2013, 558: A73.
- [21] Hathaway D. The solar cycle [J]. Living Reviews in Solar Physics volume, 2015, 12: 1-87.
- [22] Grad H, Rubin H. Hydromagnetic equilibria and force-free fields [C]//Proceedings of the second United Nations international conference on the peaceful uses of atomic energy V31. Theoretical and experimental aspects of controlled nuclear fusion. 1958: 190-197.

- [23] Sakurai T. Calculation of force-free magnetic field with non-constant α [J]. *Solar Physics*, 1981, 69(2): 343-359.
- [24] Amari T, Aly J, Luciani J, et al. Reconstructing the solar coronal magnetic field as a force-free magnetic field [J]. *Solar Physics*, 1997, 174: 129-149.
- [25] Amari T, Boulmezaoud T, Mikic Z. An iterative method for the reconstructionbreak of the solar coronal magnetic field. i. method for regular solutions [J]. *Astronomy and Astrophysics*, 1999, 350: 1051-1059.
- [26] Regnier S, Amari T, Kersale E. 3d coronal magnetic field from vector magnetograms: non-constant-alpha force-free configuration of the active region noaa 8151 [J]. *Astronomy and Astrophysics*, 2002, 392: 1119-1127.
- [27] Wheatland M. Parallel construction of nonlinear force-free fields [J]. *Solar Physics*, 2004, 222(2): 247-264.
- [28] Amari T, Boulmezaoud T, Aly J. Well posed reconstruction of the solar coronal magnetic field [J]. *Astronomy and Astrophysics*, 2006, 446(2): 691-705.
- [29] Inhester B, Wiegmann T. Nonlinear force-free magnetic field extrapolations: Comparison of the grad rubin and wheatland sturrock roumeliotis algorithm [J]. *Solar Physics*, 2006, 235: 201-221.
- [30] Nakagawa Y. Dynamics of the solar magnetic field. i. method of examination of force-free magnetic fields [J]. *Astrophysical Journal*, 1974, 190: 437-440.
- [31] Demoulin P, Cuperman S, Semel M. Determination of force-free magnetic fields above the photosphere using three-component boundary conditions. ii - analysis and minimization of scale-related growing modes and of computational induced singularities [J]. *Astronomy and Astrophysics*, 1992, 263(1-2): 351-360.
- [32] Amari T, Boulmezaoud T, Maday Y. A regularization method for the extrapolationbreak of the photospheric solar magnetic field. i. linear force-free field [J]. *Astronomy and Astrophysics*, 1998, 339: 252A.
- [33] Sakurai T. A new approach to the force-free field and its application to the mag- netic field of solar active regions [J]. *Publications of the Astronomical Society of Japan*, 1979, 31: 209-230.
- [34] Wheatland M, Sturrock P, Roumeliotis G. An optimization approach to reconstructing force-free fields [J]. *The Astrophysical Journal*, 2000(2): 1150-1155.
- [35] Wiegmann T. Optimization code with weighting function for the reconstruction of coronal magnetic fields [J]. *Solar Physics*, 2004, 219: 87-108.
- [36] Wiegmann T, Inhester B, Sakurai T. Preprocessing of vector magnetograph data for a non-linear force-free magnetic field reconstruction [J]. *Solar Physics*, 2006, 233: 215-232.

- [37] Mikic Z, McClymont A. Deducing coronal magnetic fields from vector magnetograms [J]. Solar active region evolution: comparing models with observations. Astronomical Society of the Pacific Conference Series, 1994, 68: 225-232.
- [38] Roumeliotis G. The "stress-and-relax" method for reconstructing the coronal magnetic field from vector magnetograph data [J]. Astrophysical Journal, 1996, 473: 1095-1103.
- [39] Wiegmann T, Neukirch T. Computing nonlinear force free coronal magnetic fields [J]. Nonlinear Processes in Geophysics, European Geosciences Union (EGU), 2003: 313-322.
- [40] Yan Y. On the application of the boundary element method in coronal magnetic field reconstruction [J]. Space Science Reviews volume, 2003, 107: 119-138.
- [41] Yan Y. On the coronal magnetic field configuration and solar flare/cme process [J]. Space Science Reviews, 2005, 121: 213-221.
- [42] Yan Y, Li Z. Direct boundary integral formulation for solar non-constant- α force-free magnetic fields [J]. Astrophysical Journal, 2006, 636: 1162-1168.
- [43] Gilchrist S, Wheatland M. A magnetostatic grad - rubin code for coronal magnetic field extrapolations [J]. Solar Physic, 2013, 282: 283-302.
- [44] Gilchrist S, Braun D, Barnes G. A fixed-point scheme for the numerical construction of magnetohydrostatic atmospheres in three dimensionss [J]. Solar Physic, 2016, 291: 3583 - 3603.
- [45] Wiegmann T, Neukirch T. An optimization principle for the computation of mhd equilibria in the solar corona [J]. Astronomy and Astrophysics, 2006, 457: 1053-1058.
- [46] Zhu X, Wiegmann T. On the extrapolation of magnetohydrostatic equilibria on the sun [J]. The Astrophysical Journal, 2018, 886: 130.
- [47] Zhu X, Wiegmann T. Testing magnetohydrostatic extrapolation with radiative mhd simulation of a solar flare [J]. Astronomy & Astrophysics, 2019, 631: A162.
- [48] Zhu X, Wang H, Du Z, et al. Forced field extrapolation: Testing a magnetohydrodynamic (mhd) relaxation method with a flux-rope emergence model [J]. The Astrophysical Journal, 2013, 768: 119.
- [49] Zhu X, Wang H, Du Z, et al. Forced field extrapolation of the magnetic structure of the $h\alpha$ fibrils in the solar chromosphere [J]. The Astrophysical Journal, 2016, 826: 51.
- [50] Takahiro M, Kanya K, Satoshi I. A magnetohydrodynamic relaxation method for non-force-free magnetic field in magnetohydrostatic equilibrium [J]. The Astrophysical Journal, 2020, 247: 6.
- [51] Bineau M. On the existence of force-free magnetic fields [J]. Communications on Pure and Applied Mathematics, 1972, 25: 77-84.

- [52] Boulmezaoud T, Amari T. On the existence of non-linear force-free fields in three-dimensional domains [J]. *Zeitschrift für angewandte Mathematik und Physik ZAMP*, 2000, 51: 942-967.
- [53] Yan Y, Sakurai T. New boundary integral equation representation for finite energy force-free magnetic fields in open space above the sun [J]. *Solar Physics*, 2000, 195: 89-109.
- [54] Low B, Lou Y. Modeling solar force-free magnetic fields [J]. *Astrophysical Journal*, 1990, 352: 343-352.
- [55] Parnell C. Differentiability in banach spaces, differential forms and applications [M]. Springer International Publishing, 2021.
- [56] Spivak M. A comprehensive introduction to differential geometry [M]. Publish or Perish, 1999.
- [57] Rudin W. Principles of mathematical analysis [M]. McGraw-Hill, 1976.
- [58] Cybenko G. Approximation by superpositions of a sigmoidal function [J]. *Mathematics of control, signals and systems*, 1989, 2: 303-314.
- [59] Hornik K. Approximation capabilities of multilayer feedforward networks [J]. *Neural networks*, 1991, 4: 251-257.
- [60] Higham C, Higham D. Deep learning: An introduction for applied mathematicians [J]. *Siam review*, 2019, 61: 860-891.
- [61] Baydin A, Pearlmutter B, Radul A, et al. Automatic differentiation in machine learning: a survey [J]. *Journal of Machine Learning Research*, 2018, 18: 1-43.
- [62] Kincaid D, Cheney E. Numerical analysis: Mathematics of scientific computing [M]. American Mathematical Society, 2002.
- [63] Wiegmann T, Sakurai T. Solar force-free magnetic fields [J]. *Living Reviews in Solar Physics*, 2021, 18: 1.
- [64] Parnell C. Solar theory [M]. Lecture Note for MT4510, School of Mathematics and Statistics at the University of St Andrews, 2010.
- [65] Bleybel A, Amari T, Driel-Gesztesy L, et al. Global budget for an eruptive active region i. equilibrium reconstruction approach [J]. *Astronomy & Astrophysics*, 2012, 396: 685-695.
- [66] Schmidt H. On the observable effects of magnetic energy storage and release connected with solar flares [J]. *Proceedings of the AAS-NASA Symposium held 28-30 October, 1963 at the Goddard Space Flight Center*, 1964: 107-114.
- [67] Sakurai T. Green's function methods for potential magnetic fields [J]. *Solar Physics*, 1992, 76: 301-321.
- [68] Chiu Y, Hilton H. Resistive diffusion of force-free magnetic fields in a passive medium. iv. the dynamical theory [J]. *Astrophysical Journal*, 1977, 212: 873-885.

- [69] Tolstykh V. Partial differential equations: An unhurried introduction [M]. De Gruyter, 2020.
- [70] Shen W. An introduction to numerical computation [M]. World Scientific Publishing Company, 2019.
- [71] Ryck T, Lanthaler S, Mishra S. On the approximation of functions by tanh neural networks [J]. Neural Networks, 2021, 143: 732-750.
- [72] Reddi S, Kale S, Kumar S. On the convergence of adam and beyond [C]//Sixth International Conference on Learning Representations(ICLR2018). 2018.
- [73] Lerche I, Low B. A nonlinear eigenvalue problem for self-similar spherical force-free magnetic fields [J]. Physics of Plasmas, 2014, 21: 81.
- [74] Brockmeier A, Mu T, Ananiadou S, et al. Quantifying the informativeness of similarity measurements [J]. Journal of Machine Learning Research, 2017, 18: 1-61.
- [75] Liu S, Su J, Zhang H, et al. A statistical study on force-freeness of solar magnetic fields in the photosphere [J]. Publications of the Astronomical Society of Australia, 2013, 30: e005.
- [76] Raissi M, Perdikaris P, Karniadakis G. 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.
- [77] Karniadakis G, Kevrekidis I, Lu L, et al. Physics-informed machine learning [J]. Nature Reviews Physics, 2021, 3: 422-440.
- [78] Yu J, Spiliopoulos K. Normalization effects on deep neural networks [J]. Foundations of Data Science, 2023, 5: 389-465.
- [79] Luenberger D. Linear and nonlinear programming [M]. Springer, 2003.
- [80] Boyd S, Vandenberghe L. Convex optimization [M]. Cambridge University Press, 2004.
- [81] Valori G, Kliem B, Keppens R. Extrapolation of a nonlinear force-free field containing a highly twisted magnetic loop [J]. Astronomy & Astrophysics, 2005, 433: 335-347.
- [82] MacTaggart D, Hillier A. Topics in magnetohydrodynamic topology, reconnection and stability theory [M]. Springer, 2020.
- [83] Nocedal J, Wright S. Numerical optimization [M]. Springer New York, 2006.
- [84] Yan Y. The 3-d boundary element formulation of linear force-free magnetic fields with finite energy content in semi-infinite space [J]. Solar Physics, 1995, 159: 97-113.
- [85] Lagaris I, Likas A, Fotiadis D. Artificial neural networks for solving ordinary and partial differential equations [J]. IEEE Transactions on Neural Networks and Learning Systems, 1998, 9: 987-1000.
- [86] Liu D, Nocedal J. On the limited memory bfgs method for large scale optimization [J]. Mathematical Programming, 1989, 45: 503-528.

- [87] He K, Zhang X, Ren S, et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification [C]//2015 IEEE International Conference on Computer Vision. 2015: 1026-1034.
- [88] Morse P, Feschbach H. Methods of theoretical physics [M]. McGraw-Hill, 1953.
- [89] Schrijver C, Derosa M, Metcalf T, et al. Reconstructing the solar coronal magnetic field as a force-free magnetic field [J]. Solar Physics, 2006, 235: 161-190.
- [90] Gary G. Plasma beta above a solar active region: Rethinking the paradigm [J]. Solar Physics, 2001, 203: 71-86.
- [91] 卢经纬, 程相, 王飞跃. 求解微分方程的人工智能与深度学习方法: 现状及展望 [J]. 智能科学与技术学报, 2022, 4: 461-476.
- [92] Moon Y, Choe G, Yun H, et al. Force-freeness of solar magnetic fields in the photosphere [J]. The Astrophysical Journal, 2002, 568: 422.
- [93] Hoeksema J, Liu Y, Hayashi K, et al. The helioseismic and magnetic imager (hmi) vector magnetic field pipeline: Overview and performance [J]. Solar Physics, 2014, 289: 3483-3530.
- [94] Sun X. On the coordinate system of space-weather hmi active region patches (sharps): A technical note [J]. arXiv:1309.2392, 2013.
- [95] Kusano K, Iijima H, Kaneko T, et al. Isee database for nonlinear force-free field of solar active regions [M]. Hinode Science Center, Institute for Space-Earth Environmental Research, Nagoya University, 2021.
- [96] Dissauer K, Leka K, Wagner E. Properties of flare-imminent versus flare-quiet active regions from the chromosphere through the corona. i. introduction of the aia active region patches (aarps) [J]. The Astrophysical Journal, 2023, 942(83): 1-15.
- [97] Neukirch T. Introduction to the theory of mhd equilibria [M]. Lecture Note, School of Mathematics and Statistics at the University of St Andrews, 1998.
- [98] Wiegemann T, Thalmann J, Inhester B, et al. How should one optimize nonlinear force-free coronal magnetic field extrapolations from sdo/hmi vector magnetograms? [J]. Solar Physics, 2012, 281: 37-51.
- [99] Sun X. The cgem lorentz force data from hmi vectormagnetograms [J]. eprint arXiv:1405.7353, 2014.
- [100] Gan W, Yan Y, Huang Y. Prospect for space solar physics in 2016–2030 [J]. SCIENTIA SINICA Physica, Mechanica & Astronomica, 2019, 49: 059602.
- [101] 林琦焜. 向量分析 [M]. 沧海书局, 2007.
- [102] 陈金次. 形上观微积分基本定理 [M]. 台湾大学数学系高等微积分讲义, 2013.

- [103] Zhou D. Universality of deep convolutional neural networks [J]. *Applied and Computational Harmonic Analysis*, 2020, 48: 787-794.
- [104] Nocedal J. Updating quasi-newton matrices with limited storage [J]. *Mathematics of Computation*, 1980, 35: 773-782.
- [105] Warren H, Winebarger A, Brooks D. A systematic survey of high-temperature emission in solar active regions [J]. *The Astrophysical Journal*, 2012, 759: 141.
- [106] Charbonneau P. Solar dynamo theory [J]. *Annual Review of Astronomy and Astrophysics*, 2014, 52: 251-290.
- [107] Choudhuri A. *The physics of fluids and plasmas : an introduction for astrophysicists* [M]. Cambridge University Press, 1998.
- [108] Boyd T, Sanderson J. *The physics of plasmas* [M]. Cambridge University Press, 2003.
- [109] Goedbloed J, Poedts S. *The physics of plasmas* [M]. Cambridge University Press, 2004.
- [110] Priest E, Milne A. Force-free magnetic arcades relevant to two-ribbon solar flares [J]. *Solar Physics*, 180, 65: 315-346.
- [111] Zagrodziński J. The solutions of the two-dimensional sine-gordon equation [J]. *Physics Letters A*, 1976, 57(3): 213-214.
- [112] Bateman G. *Mhd instabilities* [M]. MIT Press, 1978.
- [113] Freidberg J. *Ideal magnetohydrodynamics* [M]. Plenum Press, 1987.
- [114] Roberts P. *An introduction to magnetohydrodynamics* [M]. American Elsevier Pub. Co, 1967.
- [115] Cowling T. *Magnetohydrodynamics* [M]. Hilger, 1976.
- [116] Roumeliotis G. A new class of exact, nonlinear solutions to the grad-shafranov equation [J]. *Astrophysical Journal*, 1993, 404(2): 781-787.
- [117] Low B. Evolving force-free magnetic fields. i - the development of the preflare stage [J]. *Astrophysical Journal*, 1977, 212(2): 234-242.
- [118] Birn J, Goldstein H, Schindler K. A theory of the onset of solar eruptive processes [J]. *Solar Physics*, 1978, 57: 81-101.
- [119] Molodensky M. Equilibrium and stability of force-free magnetic field [J]. *Solar Physics*, 1974, 39: 393-404.
- [120] Molodensky M. Equilibrium and stability of force-free magnetic field ii: Stability [J]. *Solar Physics*, 1975, 43: 311-316.
- [121] Molodensky M. Equilibrium and stability of force-free magnetic field iii [J]. *Solar Physics*, 1976, 49: 279-282.
- [122] Bostrom R. Models of force-free magnetic fields in resistive media [J]. *Astrophysics and Space Science*, 1973, 22: 353-380.

- [123] Woltjer L. A theorem on force-free magnetic fields [J]. Proceedings of the National Academy of Sciences of the United States of America, 1958, 44(6): 489-491.
- [124] Taylor J. Relaxation of toroidal plasma and generation of reverse magnetic fields [J]. Physical Review Letters, 1974, 33: 1139-1141.
- [125] Taylor J. Relaxation of toroidal discharges [J]. In: Pulsed high beta plasmas; Proceedings of the Third Topical Conference, Abingdon, Oxon, England, 1976: 59-67.
- [126] Courant R, Hilbert D. Methods of mathematical physics: Partial differential equations, vol2 [M]. Wiley-VCH, 1989: 367.
- [127] Zakharov L. The theory of hydromagnetic stability of a tokamak plasma [J]. Nuclear Fusion, 1978, 18(3): 335-342.
- [128] Atanasiu C, Günter S, Lackner K, et al. Analytical solutions to the grad-shafranov equation [J]. Physics of Plasmas, 2004, 11(7): 3510-3518.
- [129] Lynden-Bell D, Boily C. Self-similar solutions up to flashpoint in highly wound magnetostatics [J]. Monthly Notices of the Royal Astronomical Society, 1994, 267(1): 146-152.
- [130] Gourgouliatos K. Self-similar magnetic arcades [J]. Monthly Notices of the Royal Astronomical Society, 2008, 385(2): 875-882.
- [131] Solovév L. Symmetric magnetohydrodynamic flow and helical waves in a circular plasma cylinder [J]. Reviews of Plasma Physics, 1967: 277-322.
- [132] Edenstrasser J. The only three classes of symmetric mhd equilibria [J]. Journal of Plasma Physics, 1980, 24: 515-518.
- [133] Low B. Magnetic field configurations associated with polarity intrusion in a solar active region. i - the force-free fields [J]. Solar Physics, 1980, 77: 43-61.
- [134] Low B. On the hydromagnetic stability of a class of laminated force-free magnetic fields [J]. Astrophysical Journal, 1988, 330: 992-996.
- [135] Sakurai T. Computational modeling of magnetic fields in solar active regions [J]. Space Science Reviews, 1989, 51: 11-48.
- [136] Chou Y, Low B. A class of force-free magnetic fields for modeling pre-flare coronal magnetic configurations [J]. Solar Physics, 1994, 153: 255-285.
- [137] Low B. Three-dimensional structures of magnetostatic atmospheres. vii. magnetic flux surfaces and boundary conditions [J]. The Astrophysical Journal, 2005, 625(1): 451-462.
- [138] Demoulin P, Henoux J, Mandrini C, et al. Reconstructing the solar coronal magnetic field as a force-free magnetic field [J]. Solar Physics, 1997, 174: 73-89.
- [139] Neukirch T. Magnetic field extrapolation [C]//Proceedings of GLOBECOM 15, San Diego, CA, USA, December 6-10Proceedings of the International Scientific Conference on Chromospheric and Coronal Magnetic Fields (ESA SP-596). 2005.

- [140] Wiegmann T. Nonlinear force-free modeling of the solar coronal magnetic field [J]. *Journal of Geophysical Research: Space Physics*, 2008, 113(A3): 1-19.
- [141] Pevtsov A, Canfield R, Metcalf T. Latitudinal variation of helicity of photospheric magnetic fields [J]. *Astrophysical Journal Letters*, 1995, 440: 109-112.
- [142] Wiegmann T, Neukirch T. Including stereoscopic information in the reconstruction of coronal magnetic fields [J]. *Solar Physics*, 2002, 208: 233-251.
- [143] Carcedo L, Brown D, Hood A, et al. A quantitative method to optimise magnetic field line fitting of observed coronal loops [J]. *Solar Physics*, 2003, 218: 29-40.
- [144] Metcalf T, DeRosa M, Schrijver C, et al. Nonlinear force-free modeling of coronal magnetic fields. ii. modeling a filament arcade and simulated chromospheric and photospheric vector fields [J]. *Solar Physics*, 2008, 247: 269-299.
- [145] Wiegmann T, Inhester B. How to deal with measurement errors and lacking data in nonlinear force-free coronal magnetic field modelling? [J]. *Astronomy and Astrophysics*, 2010, 516: 1-5.
- [146] Wiegmann T, Thalmann J, Schrijver C, et al. Can we improve the preprocessing of photospheric vector magnetograms by the inclusion of chromospheric observations? [J]. *Solar Physics*, 2008, 247: 249-267.
- [147] Chandrasekhar S, Kendall P. On force-free magnetic fields [J]. *Astrophysical Journal*, 1957, 126: 457-460.
- [148] Wilmot-Smith A, Priest E, Hornig G. Magnetic diffusion and the motion of field lines [J]. *Geophysical and Astrophysical Fluid Dynamics*, 2005, 99(2): 177-197.
- [149] Low B. Resistive diffusion of force-free magnetic fields in a passive medium [J]. *Astrophysical Journal*, 1973, 181: 209-226.
- [150] Low B. Resistive diffusion of force-free magnetic fields in a passive medium. iv. the dynamical theory [J]. *Astrophysical Journal*, 1974, 193: 243-252.
- [151] Reid J, Laing E. The resistive evolution of force-free magnetic fields. part 1. slab geometry [J]. *Journal of Plasma Physics*, 1979, 21(3): 501-510.
- [152] Schatzman E. Model of force free fields [J]. *Stellar and Solar Magnetic Fields*, 1965, 22: 337.
- [153] Priest E, Smith E. The structure of coronal arcades and the formation of solar prominences [J]. *Solar Physics*, 1979, 64(2): 267-2867.
- [154] Nakagawa Y, Raadu M. On the practical representation of magnetic field [J]. *Solar Physics*, 1972, 25: 127-135.
- [155] Wheatland M. A better linear force-free field [J]. *Astrophysical Journal*, 1999, 518(2): 948-953.

- [156] Sakurai T. Green's function methods for potential magnetic fields [J]. *Solar Physics*, 1982, 76: 301-321.
- [157] Altschuler M, Jr. G N. Magnetic fields and the structure of the solar corona i: Methods of calculating coronal fields [J]. *Solar Physics*, 1969, 9: 131-14.
- [158] Mikic Z, Linker J, Lionello R, et al. Predicting the structure of the solar corona for the total solar eclipse of march 29, 2006 [J]. *Solar and Stellar Physics Through Eclipses ASP Conference Series*, 2007, 370: 299-307.
- [159] Linker J, Lionello R, Mikic Z, et al. The evolution of open magnetic flux driven by photospheric dynamics [J]. *Astrophysical Journal*, 2011, 731(2): 1-11.
- [160] Riley P, Linker J, Mikic Z. On the application of ensemble modeling techniques to improve ambient solar wind models [J]. *Journal of Geophysical Research: Space Physics*, 2013, 118 (2): 600-607.
- [161] Lundquist S. On the stability of magneto-hydrostatic fields [J]. *Physics Review*, 1951, 83(2): 307-311.
- [162] Lüst R, Schlüter A. Kraftfreie magnetfelder [J]. *Zeitschrift für Astrophysik*, 1954, 34: 263-282.
- [163] Parker E. The origin of solar activity [J]. *Annual Review of Astronomy and Astrophysics*, 1977, 15: 45-68.
- [164] Parker E. The dynamical properties of twisted ropes of magnetic field and the vigor of new active regions on the sun [J]. *Astrophysical Journal*, 1974, 191: 245-254.
- [165] Browning P, Priest E. he structure of untwisted magnetic flux tubes [J]. *Geophysical and Astrophysical Fluid Dynamics*, 1982, 21(3): 237-263.
- [166] Zweibel E, Hundhausen A. Magnetostatic atmospheres: A family of isothermal solutions [J]. *Solar Physics*, 1982, 76: 261-299.
- [167] Hundhausen J, Hundhausen A, Zweibel E. Magnetostatic atmospheres in a spherical geometry and their application to the solar corona [J]. *Journal of Geophysical Research: Space Physics*, 1981, 86(A13): 11117-11126.
- [168] Uchida Y, Low B. Equilibrium configuration of the magnetosphere of a star loaded with accreted magnetized mass [J]. *Journal of Astrophysics and Astronomy*, 1981, 2: 405-419.
- [169] Khater A, Ibrahim R, Shamardan A, et al. Backlund transformations and painleve analysis [J]. *IMA Journal of Applied Mathematics*, 1997, 58(1): 51-69.
- [170] Khater A, Callebaut D, Kamel E. Painlevé analysis and nonlinear periodic solutions for isothermal magnetostatic atmospheres [J]. *Solar Physics*, 1998, 178: 285-315.
- [171] Khater A, Callebaut D, Malfliet W, et al. Bäcklund transformations and exact solutions for

- some nonlinear evolution equations in solar magnetostatic models [J]. Journal of Computational and Applied Mathematics, 2002, 140(1-2): 435-467.
- [172] Low B. Magnetostatic atmospheres with variations in three dimensions [J]. Astrophysical Journal, 1982, 263: 952-969.
- [173] Low B. Three-dimensional structures of magnetostatic atmospheres. vi. examples of coupled electric current systems [J]. Astrophysical Journal, 1993, 408: 693-706.
- [174] Low B. Three-dimensional magnetostatic atmospheres - magnetic field with vertically oriented tension force [J]. Astrophysical Journal, 1984, 227: 415-421.
- [175] Low B. Three-dimensional structures of magnetostatic atmospheres. i. theory [J]. Astrophysical Journal, 1985, 293: 31-43.
- [176] Bogdan T, Low B. The three-dimensional structure of magnetostatic atmospheres. ii. modeling the large-scale corona [J]. Astrophysical Journal, 1986, 306: 271-283.
- [177] Bagenal F, Gibson S. Modeling the large-scale structure of the solar corona [J]. Journal of Geophysical Research: Space Physics, 1991, 96(A10): 17663-17674.
- [178] Zhao X, Hoeksema J. A coronal magnetic field model with horizontal volume and sheet currents [J]. Solar Physics, 1994, 151: 91-105.
- [179] Gibson S, Bagenal F. Large-scale magnetic field and density distribution in the solar minimum corona [J]. Journal of Geophysical Research: Space Physics, 1995, 100(A10): 19865-19880.
- [180] Low B. Three-dimensional structures of magnetostatic atmospheres. iii. a general formulation [J]. Astrophysical Journal, 1991, 370: 427-434.
- [181] Low B. Three-dimensional structures of magnetostatic atmospheres. iv. magnetic structures over a solar active region [J]. Astrophysical Journal, 1992, 399: 300-312.
- [182] Neukirch T, Rastätter L. A new method for calculating a special class of self-consistent three-dimensional magnetohydrostatic equilibria [J]. Astronomy and Astrophysics, 1999, 348: 1000-1004.
- [183] Petrie G, Neukirch T. The green's function method for a special class of linear 3d magnetohydrostatic equilibria [J]. Astronomy and Astrophysics, 2000, 356: 735-746.
- [184] Low B. Three-dimensional structures of magnetostatic atmospheres. v. coupled electric current systems [J]. Astrophysical Journal, 1993, 408: 689-692.
- [185] Gebhardt U, Kiessling M. The structure of ideal magnetohydrodynamics with incompressible steady flow [J]. Physics of Fluids B: Plasma Physics, 1992, 4(7): 1689-1701.
- [186] Petrie G, Neukirch T. Self-consistent three-dimensional steady state solutions of the mhd equations with field-aligned incompressible flow [J]. Geophysical and Astrophysical Fluid Dynamics, 1999, 91(3): 269-302.

- [187] Nickeler D, Wiegmann T. Thin current sheets caused by plasma flow gradients in space and astrophysical plasma [J]. *Annales Geophysicae*, 2010, 28(8): 1523-1532.
- [188] Nickeler D, Wiegmann T. Relation between current sheets and vortex sheets in stationary incompressible mhd [J]. *Annales Geophysicae*, 2012, 30(3): 545-555.

致 谢

即将落笔之时，已夜深人静...

此刻，要衷心感谢我的导师徐龙研究员。徐老师广博的专业知识、严谨的治学态度、因材施教的教育方式、勤奋简朴的生活作风以及无私奉献的精神令我受益匪浅。无论在课题的选择、研究和论文的写作上、还是在生活上，徐老师给予我精心的指导、热情的帮助和深切的关怀，使我能够顺利地完成本论文的工作。

2021年1月初我开始在徐老师的指导下开展太阳日冕磁场外推工作。基于无场模型的太阳日冕磁场外推的本质是求解一类偏微分方程边值问题。在定解条件、解的存在唯一性问题上尚无定论时，太阳物理工作者利用数值方法外推的磁场与观测中的现象吻合，这是件不可思议的事。在后续的学习和研究工作中得到颜毅华老师的许多指导和帮助，在论文完成之际，非常感谢颜老师。

在原国家天文台太阳活动预报团组的大家庭里，我深深体会到坦然接受帮助和热心帮助别人的快乐，也正是在这个集体中，我逐渐体会到了研究太阳物理的乐趣。2022年9月迁徙至国家空间科学中心，感谢国家空间科学中心提供干净舒适的学习工作环境。

感谢国家天文台原预报组老师们：郭娟老师、黄鑫老师、杜占乐老师、闫岩老师、贺晗老师、王华宁老师。感谢国家天文台的原明安图基地的张印老师和谭宝林老师。感谢空间中心的朱小帅老师。感谢北京航空航天大学的姜杰老师。

感谢一起学习的同学们：张凤珍师姐、赵栋师兄、于学鑫师兄、孙文青师姐、窦凤萍师姐、刘思璇师姐、张鑫泽师兄、孙得志师弟、冯凯师弟、汪维师弟、高菲师妹、钟昱丰师弟和赵忠瑞师弟。

感谢新型冠状病毒，让我认识到人类不要企图战胜大自然，我们应怀着一颗敬畏的心，不断地认识大自然并与之和谐相处。在世风日下、人心不古的当下应要向渔父学习：沧浪之水清兮，可以濯吾缨；沧浪之水浊兮，可以濯吾足。

感谢我的父母，年迈的你们，依旧披星戴月奔走于田间与农贸市场，为哥哥和我一直默默无私奉献。

感谢我的爱人，悬壶济世的你对我的无限包容与付出。

Last but not least, 感谢自己，依旧热爱数学，依旧热爱真理。

$$1 \times 1 = 1$$

$$11 \times 11 = 121$$

$$111 \times 111 = 12321$$

$$1111 \times 1111 = 1234321$$

$$11111 \times 11111 = 123454321$$

$$111111 \times 111111 = 12345654321$$

$$1111111 \times 1111111 = 123456787654321$$

$$11111111 \times 11111111 = 12345678987654321$$

作者简历及攻读学位期间发表的学术论文与研究成果

作者简历：

张尧, 江苏沛县人, 中国科学院国家空间科学中心博士研究生

2014 年 6 月毕业于深圳大学数学系信息与计算科学专业

2019 年 6 月毕业于深圳大学数学系应用数学专业

2020 年 9 月起就读中国科学院国家天文台天文方法与技术专业, 2022 年 9 月转入中国科学院国家空间科学中心天文方法与技术专业

已发表（或正式接受）的学术论文：

1. Unsupervised Learning for Salient Object Detection via Minimization of Bilinear Factor Matrix Norm, IEEE Transactions On Neural Networks and Learning Systems, Vol.34 (3), pages 1354 –1366, 2023, 通讯作者.
2. Machine-Learning-Based Numerical Solution for Low and Lou’s Nonlinear Force-Free Field Equilibria, Solar Physics, Vol.299 (108), pages 1 –24, 2024, 第一作者.
3. Physics-informed Neural Network for Force-free Magnetic Field Extrapolation, Research in Astronomy and Astrophysics, 第一作者 (已正式接受).

