

谱方法的算法、分析与应用

蒋凯课题组

数学与计算科学学院

湘潭大学

2019 年 8 月 26 日

目 录

第一章	介绍	1
1.1	加权余量法	1
1.2	谱配置法	4
1.3	Galerkin 型谱方法	6
1.3.1	Galerkin 方法	7
1.3.2	Petrov-Galerkin 方法	9
1.3.3	数值积分 Galerkin 法	10
1.4	分析误差的基本工具	11
1.5	比较数值例子	18
1.5.1	有限差分法与谱配置法相比较	19
1.5.2	谱 Galerkin 法与配置法相比较	23
第二章	周期性问题的傅里叶谱方法	26
2.1	连续和离散傅里叶变换	27
2.1.1	连续傅里叶级数	27
2.1.2	离散傅里叶变换	29
2.1.3	物理空间上的微分	34
2.1.4	频率空间上的微分	36
2.2	傅里叶近似	38
2.2.1	逆不等式	38
2.2.2	正交投影	39
2.2.3	插值	41
2.3	傅里叶谱方法的应用	43
2.3.1	KdV 方程	43

2.3.2	KS 方程	48
2.3.3	Allen-Cahn 方程	50

第一章 介绍

刘双

偏微分方程数值计算方法可以分为局部的与全局的两类。有限差分与有限元方法基于局部参数，而谱方法具有全局性。实践说明，有限元方法特别适用于复杂几何问题，而谱方法在牺牲域的灵活性下可以提供高精度。我们要强调的是存在许多数值方法，如 hp 有限元法与谱元素方法，它们结合了全局及局部方法的优点。然而，在本书中我们将主要讨论全局的谱方法。

在微分方程数值解的背景下，谱方法属于加权余量法 (WRMs) 族，传统上认为它是许多数值计算方法的基础，例如有限元法、谱方法、有限体积法、边界元法 (参见 Finlayson(1972))。WRMs 代表近似技术的一类特殊群体，其中余量 (或误差) 通过某种方式最小化，从而引出特定的方法，包括 Galerkin、Petrov-Galerkin、配置法和 τ 方法。

本介绍章节的主要目的是通过使用余量的概念来制定谱方法。同时 will 介绍一些重要的工具，如离散变换和谱微分。这些都是发展高效谱算法的基本要素。

1.1 加权余量法

在介绍谱方法之前，我们首先给出 WRM 的简单介绍。考虑一般问题：

$$\partial_t u(x,t) - \mathcal{L}u(x,t) = \mathcal{N}(u)(x,t), \quad t > 0, x \in \Omega, \quad (1.1)$$

其中 \mathcal{L} 是主导的空间微分算子， \mathcal{N} 是仅涉及空间微分的低阶线性或非线性算子。这里 Ω 表示 \mathbb{R}^d 的有界域， $d=1,2,3$ 。方程 (1.1) 有补充初始条件和合适的边界条件。

我们仅考虑 WRM 用于空间离散，并且假设时间的微分通过合适的时间步进法而被离散化。在各种时间步进法中（参考附录 D），半隐式法或线性隐式法，其中为减少相关稳定性约束，主要线性算子被隐式处理，为避免在求解非线性方程的每次步进上耗费大量时间，非线性部分被显式处理，这在谱方法中是很常用的。

令 τ 为时间步长， $u^k(\cdot)$ 为 $u(\cdot, k\tau)$ 的近似值。作为一个例子，我们考虑对 (1.1) 使用 Crank-Nicolson 蛙跳步进法：

$$\frac{u^{n+1} - u^{n-1}}{2\tau} - \mathcal{L}\left(\frac{u^{n+1} + u^{n-1}}{2}\right) = \mathcal{N}(u^n), \quad n \geq 1. \quad (1.2)$$

我们可以将 (1.2) 重写为

$$Lu(x) := \alpha u(x) - \mathcal{L}u(x) = f(x), \quad x \in \Omega, \quad (1.3)$$

其中，略微滥用符号，令 $u = \frac{u^{n+1} + u^{n-1}}{2}$ ， $\alpha = \tau^{-1}$ ， $f = \alpha u^{n-1} + \mathcal{N}(u^n)$ 。因此，我们需要在每次步进上解决一个形式 (1.3) 的稳态问题。

在这一点上，需要强调的是为一些形如 (1.3) 的重要方程，例如泊松型方程和对流扩散方程，构造高效的数值求解法，是解决一般非线性偏微分方程的重要步骤。考虑到这点，本书的一个特别重点是设计和分析形如方程 (1.3) 的有效谱算法，其中 \mathcal{L} 是线性椭圆算子。

WRM 一开始是通过有限求和

$$u(x) \approx u_N(x) = \sum_{k=0}^N a_k \phi_k(x), \quad (1.4)$$

来逼近 (1.3) 的解 u ，其中 $\{\phi_k\}$ 是试探函数（或基函数），扩展系数 $\{a_k\}$ 有待确定。在 (1.3) 中用 u_N 代替 u 得到余项：

$$\mathbf{R}_N(x) = Lu_N(x) - f(x) \neq 0, \quad x \in \Omega. \quad (1.5)$$

WRM 的思想是通过要求下式成立而使得余项趋于 0

$$(\mathbf{R}_N, \psi_j)_\omega := \int_{\Omega} \mathbf{R}_N(x) \psi_j(x) \omega(x) dx = 0, \quad 0 \leq j \leq N, \quad (1.6)$$

其中 $\{\psi_j\}$ 是测试函数, ω 是正的权函数; 或

$$\langle \mathbf{R}_N, \psi_j \rangle_{N, \omega} := \sum_{k=0}^N \mathbf{R}_N(x_k) \psi_j(x_k) \omega_k = 0, \quad 0 \leq j \leq N, \quad (1.7)$$

其中 $\{x_k\}_{k=0}^N$ 是预先选好的一组点, $\{\omega_k\}_{k=0}^N$ 是数值求积公式的权重。

试探/测试函数的选择是将谱方法与有限元、有限差分法区分开来的主要特征之一。在后两种方法中, 试探/测试函数在特征上是局部的, 具有有限的规律性。相比之下, 谱方法则采用全局光滑的函数作为试探/测试函数。最常用的试探/测试函数是三角函数或正交多项式 (通常是单个 Sturm-Liouville 问题的特征函数), 其中包括

- $\phi_k(x) = e^{ikx}$ (傅里叶谱方法)
- $\phi_k(x) = T_k(x)$ (切比雪夫谱方法)
- $\phi_k(x) = L_k(x)$ (勒让德谱方法)
- $\phi_k(x) = \mathcal{L}_k(x)$ (拉盖尔谱方法)
- $\phi_k(x) = H_k(x)$ (赫米特谱方法)

这里 T_k, L_k, \mathcal{L}_k 和 H_k 分别是 k 次切比雪夫、勒让德、拉盖尔和赫米特多项式。

以下公式由测试函数的选择来区分:

- Galerkin. 假设边界条件是周期性的或齐次的情况下, 测试函数与试探函数相同 (即在 (1.6) 或 (1.7) 式中 $\phi_k = \psi_k$)。
- Petrov-Galerkin. 测试函数与试探函数不同。
- 配置法. (1.7) 式中的测试函数 $\{\psi_k\}$ 为使得 $\psi_k(x_j) = \delta_{jk}$ 的拉格朗日多项式基, 其中 $\{x_j\}$ 为配置点。因此, 余项在 $\{x_j\}$ 处的值为 0, 即 $\mathbf{R}_N(x_j) = 0$ 。

Remark 1.1. 在文献中, 拟谱方法的术语经常用于一些涉及配置法的运算或产生混淆误差的数值求积 (参考 Gottlieb and Orszag(1977)). 从这个意义上说, 几乎所有实用的谱方法是拟谱方法。在本书中, 我们不会以拟谱或谱来分类方法。相反, 我们将以 Galerkin 型与配置型来分类。

Remark 1.2. 所谓 τ 方法就是特殊的 Petrov-Galerkin 方法。虽然 τ 方法在某些情况下具有一些优势, 但对大多数问题而言, 通常最好使用精心设计的 Galerkin 或 Petrov-Galerkin 方法。因此, 在本书中我们将不会涉及这个主题, 可以参考 El-Daou 和 Ortiz (1998), Canuto 等 (2006) 及对 τ 方法深入讨论的参考文献。

在接下来的部分, 我们将演示如何以一般方式检查基于 Galerkin, Petrov-Galerkin 和配置法的谱方法来构建用于求解微分方程的谱方法。我们将在本书主体中以更严格的方式来重新审视这些说明性的例子。

1.2 谱配置法

为了确定思想, 我们考虑下述线性问题:

$$\begin{aligned} \mathbf{L}u(x) &= -u''(x) + p(x)u'(x) + q(x)u(x) = f(x), \quad x \in (-1, 1), \\ B_{\pm}u(\pm 1) &= g_{\pm}, \end{aligned} \quad (1.8)$$

其中 B_{\pm} 是对应于狄利克雷、诺伊曼或罗宾边界条件的线性算子 (见 4.1 节), p, q, f 和 g_{\pm} 已给出且使得上述问题是适定的。

如前所述, 配置法使得在预先给定点处的余项为 0。更确切地说, 令 $\{x_j\}_{j=0}^N$ (其中 $x_0 = -1, x_N = 1$) 为一组 Gauss-Lobatto 点 (见第三章), P_N 为所有次数小于等于 N 的实系数多项式所构成的集合。(1.8) 的谱配置法就是寻找 $u_N \in P_N$ (a) 使得余项 $\mathbf{R}_N(x_k) = \mathbf{L}u_N(x_k) - f(x_k)$ 在配置点处的值等于 0, 即

$$\mathbf{R}_N(x_k) = \mathbf{L}u_N(x_k) - f(x_k) = 0, \quad 1 \leq k \leq N-1, \quad (1.9)$$

(b) 使得 u_N 满足边界条件, 即

$$B_-u_N(x_0) = g_-, B_+u_N(x_N) = g_+. \quad (1.10)$$

谱配置法通常用于物理空间，通过下式来寻求近似解

$$u_N(x) = \sum_{j=0}^N u_N(x_j) h_j(x) \quad (1.11)$$

其中 $\{h_j\}$ 是拉格朗日因子（也称为节点函数），即 $h_j \in P_N$ 且 $h_j(x_k) = \delta_{kj}$ 。因此，将 (1.11) 式代入到 (1.9) - (1.10) 式得到线性方程组

$$\begin{aligned} \sum_{j=0}^N [\mathbf{L}h_j(x_k)] u_N(x_j) &= f(x_k), \quad 1 \leq k \leq N-1, \\ \sum_{j=0}^N [B_- h_j(x_0)] u_N(x_j) &= g_-, \quad \sum_{j=0}^N [B_+ h_j(x_N)] u_N(x_j) = g_+. \end{aligned} \quad (1.12)$$

上述方程组包含 $N+1$ 个方程和 $N+1$ 个未知量，所以我们可以将其写成矩阵的形式。为了完成这个想法，我们考虑 (1.8) 式的狄利克雷边界条件情形： $u(\pm 1) = g_{\pm}$ 。在这种情况下，令 (1.12) 的第一个方程中 $u_N(x_0) = g_-$, $u_N(x_N) = g_+$ ，我们发现方程组 (1.12) 简化为

$$\sum_{j=1}^{N-1} [\mathbf{L}h_j(x_k)] u_N(x_j) = f(x_k) - \{[\mathbf{L}h_0(x_k)]g_- + [\mathbf{L}h_N(x_k)]g_+\}, \quad (1.13)$$

其中 $1 \leq k \leq N-1$ 。对 (1.11) 式微分 m 次得

$$u_N^{(m)}(x_k) = \sum_{j=0}^N d_{kj}^{(m)} u_N(x_j) \quad (1.14)$$

其中 $d_{kj}^{(m)} = h_j^{(m)}(x_k)$ 。

矩阵 $D^{(m)} = (d_{kj}^{(m)})_{k,j=0,\dots,N}$ 称为关于 $\{x_j\}_{j=0}^N$ 的 m 阶微分矩阵。如果我们用分量为 $u_N^{(m)}$ 在配置点处的值来表示向量 \mathbf{u}^m ，则 (1.14) 可表示为

$$\mathbf{u}^{(m)} = D^{(m)} \mathbf{u}^{(0)}, \quad m \geq 1. \quad (1.15)$$

因此，我们有

$$\mathbf{L}h_j(x_k) = -d_{kj}^{(2)} + p(x_k)d_{kj}^{(1)} + q(x_k)\delta_{kj}. \quad (1.16)$$

用分量为 (1.13) 式右边的部分来表示向量 \mathbf{f} 。令

$$\begin{aligned} \tilde{D}_m &= (d_{kj}^{(m)})_{k,j=1,\dots,N-1}, \quad m = 1, 2, \\ P &= \text{diag}(p(x_1), \dots, p(x_{N-1})), \quad Q = \text{diag}(q(x_1), \dots, q(x_{N-1})), \end{aligned} \quad (1.17)$$

则式 (1.13) 化为

$$(-\tilde{D}_2 + P\tilde{D}_1 + Q)\mathbf{u}^{(0)} = \mathbf{f}. \quad (1.18)$$

注意到一旦微分矩阵已知，配置法将易于实现。此外，由于我们在物理空间中考虑问题，且导数可以通过 (1.14) 直接求得，这是非常方便去解决具有可变系数和/或非线性问题。因此，配置法已广泛的在实践中使用。然而，三个重要的问题应该在配置法的实施与分析上得到考虑：

- 配置法方程组总是充满条件数为 $O(N^{2m})$ 的系数矩阵 (m 是微分方程的阶数)
- 在稳定性、准确性和易于处理边界条件方面，配置点的选取至关重要。通常，选择高斯型求积公式节点 (一般为正交多项式的零点)。
- 上述配置法以强式表达。在误差分析方面，很容易重新以弱式表达 (但并非总是如此)，见 1.3.3 节和第四章。

1.3 Galerkin 型谱方法

上一节叙述了在实体空间实行配置法。这一节，我们将叙述在频率空间的 Galerkin 型谱方法，并介绍 Galerkin 谱方法、Petrov-Galerkin 谱方法和数值积分式 Galerkin 谱方法的基本原理。

1.3.1 Galerkin 方法

不失一般性，我们考虑 (1.8) 中边界条件 $g_{\pm} = 0$ 的情况。对于非齐次边界条件情形，可以通考虑 $v = u - \tilde{u}$ 而简单处理，其中 \tilde{u} 是满足对应非齐次边界条件的简单函数（参考第四章）。

定义有限维近似空间：

$$X_N = \{\phi \in P_N : B_{\pm}\phi(\pm 1) = 0\} \Rightarrow \dim(X_N) = N - 1.$$

令 $\{\phi_k\}_{k=0}^{N-2}$ 为 X_N 的一组基。我们展开近似解为

$$u_N(x) = \sum_{k=0}^{N-2} \hat{u}_k \phi_k(x) \in X_N. \quad (1.19)$$

然后，扩展系数 $\{\hat{u}_k\}_{k=0}^{N-2}$ 可由余项式 (1.6) 决定，其中 $\{\psi_j = \phi_j\}$ ：

$$\int_{-1}^1 (\mathbf{L}u_N(x) - f(x)) \phi_j(x) \omega(x) dx = 0, \quad 0 \leq j \leq N-2, \quad (1.20)$$

等价于

$$\begin{cases} \text{Find } u_N \in X_N \text{ s.t.} \\ (\mathbf{L}u_N, v_N)_{\omega} = (f, v_N)_{\omega}, \quad \forall v_N \in X_N. \end{cases} \quad (1.21)$$

这里 $(\cdot, \cdot)_{\omega}$ 是 $L_{\omega}^2(-1, 1)$ 上的内积（参考附录 B）。

上述方法的线性方程组包含了将 (1.19) 代入到 (1.20)。更为准确的，令

$$\begin{aligned} \mathbf{u} &= (\hat{u}_0, \hat{u}_1, \dots, \hat{u}_{N-2})^T; \quad f_j = (f, \phi_j)_{\omega}, \quad \mathbf{f} = (f_0, f_1, \dots, f_{N-2})^T; \\ s_{jk} &= (\mathbf{L}\phi_k, \phi_j)_{\omega}, \quad S = (s_{jk})_{j,k=0, \dots, N-2}, \end{aligned}$$

方程组 (1.20) 化为

$$S\mathbf{u} = \mathbf{f}. \quad (1.22)$$

因此，选择满足下列条件的基函数是至关重要的：

- 右边项 $(f, \phi_j)_{\omega}$ 可以被有效计算。

- 线性方程组可以被有效解出。

关键思想是使用正交多项式或正交函数的紧组合来构造基函数。为了演示基本原理，我们考虑 Legendre 谱逼近（即 (1.20) - (1.22) 中 $\omega \equiv 1$ ）。令 $L_k(x)$ 为 k 次 Legendre 多项式，并且令

$$\phi_k(x) = L_k(x) + \alpha_k L_{k+1} + \beta_k L_{k+2}, \quad k \geq 0, \quad (1.23)$$

其中常数 α_k 与 β_k 由边界条件 $B_{\pm}\phi_k(\pm 1) = 0$ 唯一决定（参考 4.1 节）。我们将这些基函数称为模态基函数。因此，我们有

$$X_N = \text{span}\{\phi_0, \phi_1, \dots, \phi_{N-2}\}. \quad (1.24)$$

由 Legendre 多项式的性质易验证，如果 $p(x)$ 与 $f(x)$ 是常数，则系数矩阵 S 是稀疏矩阵，所以线性方程组 (1.22) 可以被有效计算。然而，对于更一般的 $p(x)$ 与 $f(x)$ ，系数矩阵 S 是满的则需要求助于迭代法（参考 4.4 节）。

以上，我们仅仅考虑了 Legendre 情形。事实上，对 Chebyshev, Laguerre 和 Hermite 情形构造这样的基函数也是可行的（见第 4-7 章）。使用正交多项式/函数的紧组合这个概念来开发有效的谱求解法将在本书中被反复强调。

现在我们来考虑 $(f, \phi_j)_{\omega}$ 的估值。通常，该项不能被直接计算而是常由 $(I_N f, \phi_j)_{\omega}$ 近似，其中 I_N 是 P_N 上关于 Gauss-Lobatto 点的插值算子。因此，我们可以写为

$$(I_N f)(x) = \sum_{k=0}^N \tilde{f}_k \phi_k(x), \quad (1.25)$$

其中 $\{\phi_k\}$ 是 P_N 的一组正交多项式基（ ω 正交，即 $(\phi_k, \phi_j)_{\omega} = \delta_{jk}$ ）。由于正交性，物理值 $\{f(x_j)\}_{j=0}^N$ 与扩展系数 $\{\tilde{f}_k\}_{k=0}^N$ 之间的离散变化可以被有效计算。特别是 Fourier 与 Chebyshev 离散变化的计算复杂度可以通过快速傅里叶变换被减至 $O(N \log_2 N)$ 。一种为实现离散变化的相关的一般正交多项式在 (3.15) 节给出。

重要的是要指出在解与时间相关的非线性问题时， f 通常包含之前步骤所得数值解 u_N 的微分的非线性项（参见 (1.3)）。因此，频率空间和/或物

理空间的数值微分是必要的。与一般正交多项式相关的微分技术在 3.16 与 3.17 节给出。

1.3.2 Petrov-Galerkin 方法

在 1.1 节已指出，测试函数与试探函数的使用不同使 Petrov-Galerkin 法有别于 Galerkin 法。由于这种灵活性，Petrov-Galerkin 法对一些非自伴问题例如奇阶方程十分有用。

作为一个说明性的例子，我们考虑下列三阶方程：

$$\begin{aligned} \mathbf{L}u(x) &:= u'''(x) + u(x) = f(x), \quad x \in (-1, 1), \\ u(\pm 1) &= u'(1) = 0. \end{aligned} \quad (1.26)$$

与 Galerkin 情形一样，我们在近似解上加上边界条件。所以我们设

$$X_N = \{\phi \in P_N : \phi(\pm 1) = \phi'(1) = 0\} \Rightarrow \dim(X_N) = N - 2.$$

假设 $\{\phi_k\}_{k=0}^{N-3}$ 是 X_N 的一组基，我们将近似解扩展为

$$u_N(x) = \sum_{k=0}^{N-3} \hat{u}_k \phi_k(x) \in X_N.$$

扩展系数 $\{\hat{u}_k\}_{k=0}^{N-3}$ 由余项式 (1.6) 决定（其中 $\omega = 1$ ）：

$$\int_{-1}^1 (\mathbf{L}u_N(x) - f(x)) \psi_j(x) dx = 0, \quad 0 \leq j \leq N - 3. \quad (1.27)$$

由于核心的三阶算子不是自伴的，因此使用 Petrov-Galerkin 法及测试函数空间：

$$X_N^* = \{\psi \in P_N : \psi(\pm 1) = \psi'(-1) = 0\} \rightarrow \dim(X_N^*) = N - 2.$$

假设 $\{\psi_k\}_{k=0}^{N-3}$ 是 X_N^* 的一组基。于是 (1.27) 等价于变分形式：

$$\begin{cases} \text{Find } u_N \in X_N \text{ s.t.} \\ (\mathbf{L}u_N, v_N) = (f, v_N), \quad \forall v_N \in X_N^*, \end{cases} \quad (1.28)$$

其中 (\cdot, \cdot) 是通常 L^2 空间的内积。

上述方法的理论方面将在第六章进行讨论。我们现在考虑实行它。令

$$\begin{aligned} \mathbf{u} &= (\hat{u}_0, \hat{u}_1, \dots, \hat{u}_{N-3})^T; f_j = (f, \psi_j), \mathbf{f} = (f_0, f_1, \dots, f_{N-3})^T; \\ s_{jk} &= (\phi'_k, \psi''_j), S = (s_{jk})_{j,k=0,\dots,N-3}; \\ m_{jk} &= (\phi_k, \psi_j), M = (m_{jk})_{j,k=0,\dots,N-3}, \end{aligned}$$

线性方程组化为

$$(S + M)\mathbf{u} = \mathbf{f}. \quad (1.29)$$

前面的章节阐述过，我们希望构造 X_N 与 X_N^* 的基函数以至于线性方程组 (1.29) 可以有有效的转换。又一次，我们可以通过使用正交多项式的紧组合来实现目标。可以检查对 $0 \leq k \leq N-3$,

$$\begin{aligned} \phi_k &= L_k - \frac{2k+3}{2k+5}L_{k+1} - L_{k+2} + \frac{2k+3}{2k+5}L_{k+3} \in X_N; \\ \psi_k &= L_k + \frac{2k+3}{2k+5}L_{k+1} - L_{k+2} - \frac{2k+3}{2k+5}L_{k+3} \in X_N^*, \end{aligned} \quad (1.30)$$

其中 L_n 是 n 次 Legendre 多项式 (参见 3.3 节)。因此, $\{\phi_k\}_{k=0}^{N-3}(\{\psi_j\}_{j=0}^{N-3})$ 形成 $X_N(X_N^*)$ 的一组基。此外, 运用 Legendre 多项式的性质易验证矩阵 M 是七对角的, 即 $m_{jk} = 0, \forall |j-k| > 3$ 。更重要的是, 矩阵 S 是对角的。

1.3.3 数值积分 Galerkin 法

我们之前在频率空间上考虑的 Galerkin 型方法非常适合常系数 (或多项式系数) 线性问题。然而, 在一般变系数问题上实施它却并不方便。另一方面, 配置法虽然容易实施, 却无法总能重新构成一个稳定的变分方程 (更多的是方便误差分析)。将上述两种方法结合就导出了数值积分 Galerkin 法, 有时也称为弱形式的配置法。

这个方法的主要思想是在 Galerkin 方程中将连续的内积换为离散的。作为一个例子, 我们再一次考虑 (1.8), 其中 $g_{\pm} = 0$ 。数值积分谱方法为:

$$\begin{cases} \text{Find } u_N \in X_N := \{\phi \in P_N : B_{\pm}\phi(\pm 1) = 0\} \text{ s.t.} \\ a_N(u_N, v_N) := \langle L u_N, v_N \rangle_N = \langle f, v_N \rangle_N, \forall v_N \in X_N, \end{cases} \quad (1.31)$$

其中离散内积定义为

$$\langle u, v \rangle_N = \sum_{j=0}^N u(x_j) v(x_j) \omega_j,$$

这里 $\{x_j, \omega_j\}_{j=0}^N$ 为 Legendre-Gauss-Lobatto 节点和权重的集合（参见定理 3.29）。

由于离散内积，上述方法为解变系数问题更易于 spectral-Galerkin 法 (1.21) 而实行。由于弱公式化，它也比谱配置法 (1.12) 更方便用于误差分析。

我们注意到在齐次狄利克雷边界条件的特殊情形下，即 $B_{\pm}u(\pm 1) = u(\pm 1) = 0$ ，通过在 (1.31) 中取 $v_N = h_j, 1 \leq j \leq N-1$ ，并利用 Legendre-Gauss-Lobatto 求积的精确性，即

$$\langle u, v \rangle_N = (u, v), \forall u, v \in P_{2N-1}, \quad (1.32)$$

我们发现 (1.31) 式与配置法 (1.12) 式相等。然而，这对一般的边界条件是不成立的（参见第四章）。

1.4 分析误差的基本工具

在前几节中，我们用加权余量法的概念简单阐述了几类离散谱方法。本节我们提出一些基于弱方程（或变分方程）的数值算法稳定性和收敛性分析的基本工具。

我们考虑线性边界值问题 (1.3)：

$$\mathbf{L}u = f, \text{ in } \Omega; Bu = 0, \text{ on } \partial\Omega, \quad (1.33)$$

其中 \mathbf{L} 和 B 是线性算子， f 是已给的 Ω 上的函数。

如前所示，首先将 (1.33) 重构为一个弱方程：

$$\begin{cases} \text{Find } u \in X \text{ s.t.} \\ a(u, v) = F(v), \forall v \in Y, \end{cases} \quad (1.34)$$

其中 X 是试探函数空间, Y 是测试函数空间, F 是 Y 上的线性函数。 $a(u, v)$ 是定义在 $X \times Y$ 上的双线性形式。常规的假设 X 和 Y 是希尔伯特空间。基础泛函分析集合见附录 B。

现在, 我们考虑应该在 (1.34) 上加哪些条件以确保它在下述意义下的良好性:

- 存在唯一性: 问题有且仅有一个解。
- 稳定性: 解必须是稳定的, 即解是连续依赖数据的。换种说法, 数据发生小的变动相应的解也只产生小的变动。

关于存在唯一性与稳定性的第一个基本结果被称为与抽象问题(1.34)(其中 $X = Y$) 相关的 Lax-Milgram 引理 (见定理 B.1), 即

$$\begin{cases} \text{Find } u \in X \text{ s.t.} \\ a(u, v) = F(v), \forall v \in X. \end{cases} \quad (1.35)$$

更精确的, 如果双线性式 $a(\cdot, \cdot): X \times X \rightarrow \mathbb{R}$ 满足

- 连续性:

$$\exists C > 0 \text{ s.t. } |a(u, v)| \leq C \|u\|_X \|v\|_X, \quad (1.36)$$

- 强制性:

$$\exists \alpha > 0 \text{ s.t. } a(u, u) \geq \alpha \|u\|_X^2, \quad (1.37)$$

然后对于任意 $F \in X'$ (X 的对偶空间, 定义见附录 B), 问题 (1.35) 有一个唯一的解 $u \in X$, 满足

$$\|u\|_X \leq \frac{1}{\alpha} \|F\|_{X'}. \quad (1.38)$$

Remark 1.3. 常数

$$\alpha = \inf_{0 \neq u \in X} \frac{|a(u, u)|}{\|u\|_X^2} \quad (1.39)$$

为 (1.35) 的椭圆常数。

以上结果仅能应用于问题 (1.34) 在 $X = Y$ 的情形。现在我们给出 Lax-Milgram 引理在 $X \neq Y$ 的情况下的推广 (参见例子 *Babuška and Aziz* (1972))。

定理 1.1. 令 X 和 Y 为两个实希尔伯特空间, 分别配有范数 $\|\cdot\|_X$ 和 $\|\cdot\|_Y$ 。假设 $a(\cdot, \cdot): X \times Y \rightarrow \mathbb{R}$ 是一个双线性式, $F(\cdot): Y \rightarrow \mathbb{R}$ 是一个连续线性函数, 即 $F \in Y'$ (Y 的对偶空间) 满足

$$\|F\|_{Y'} = \sup_{0 \neq v \in Y} \frac{|F(v)|}{\|v\|_Y} < \infty. \quad (1.40)$$

进一步, 假设 $a(\cdot, \cdot)$ 满足

• 连续性:

$$\exists C > 0 \text{ s.t. } |a(u, v)| \leq C \|u\|_X \|v\|_Y, \quad (1.41)$$

• *Inf* – *sup* 条件:

$$\exists \beta > 0 \text{ s.t. } \sup_{0 \neq v \in Y} \frac{|a(u, v)|}{\|u\|_X \|v\|_Y} \geq \beta, \forall 0 \neq u \in X, \quad (1.42)$$

• “转置” *Inf* – *sup* 条件:

$$\sup_{0 \neq u \in X} |a(u, v)| > 0, \forall 0 \neq v \in Y. \quad (1.43)$$

然后, 对于任意 $F \in Y'$, 问题 (1.34) 有一个唯一解 $u \in X$ 满足

$$\|u\|_X \leq \frac{1}{\beta} \|F\|_{Y'}. \quad (1.44)$$

Remark 1.4. 条件 (1.42) 也被称为 *Babuška – Brezzi inf – sup* 条件 (参见 *Babuška – Brezzi* (1974)), 其中实数

$$\beta = \inf_{0 \neq u \in X} \sup_{0 \neq v \in Y} \frac{|a(u, v)|}{\|u\|_X \|v\|_Y} \quad (1.45)$$

被称为 $\inf - \sup$ 常数。

Remark 1.5. $X = Y$ 时的定理 1.1 与 Lax-Milgram 引理并不相等。事实上, 很容易验证椭圆常数与 $\inf - \sup$ 常数满足: $\alpha \leq \beta$ 。的确, 通过 (1.37),

$$\alpha \|u\|_X \leq \frac{|a(u, u)|}{\|u\|_X} \leq \sup_{0 \neq v \in X} \frac{|a(u, v)|}{\|v\|_X}, \quad \forall 0 \neq u \in X,$$

可知

$$\alpha \leq \inf_{0 \neq u \in X} \sup_{0 \neq v \in X} \frac{|a(u, v)|}{\|u\|_X \|v\|_X} = \beta.$$

这意味着可以有 $\alpha = 0$ 但 $\beta > 0$ 。换句话说, 双线性式不是强的, 但满足 $\inf - \sup$ 条件。

下面我们来回顾数值逼近 (1.34) 的收敛分析的基本理论。

我们首先考虑 $X = Y$ 的情形。假设 $X_N \subseteq X$ 且

$$\forall v \in X, \quad \inf_{v_N \in X_N} \|v - v_N\|_X \rightarrow 0 \quad (N \rightarrow \infty). \quad (1.46)$$

Galerkin 逼近 (1.35) 为

$$\begin{cases} \text{Find } u_N \in X_N \text{ s.t.} \\ a(u_N, v_N) = F(v_N), \quad \forall v_N \in X_N. \end{cases} \quad (1.47)$$

通过下述引理可以建立该方法的稳定性与收敛性 (参见 Céa(1964)):

定理 1.2.(Céa 引理). 在 Lax-milgram 引理的假设下 (见定理 B.1), 问题 (1.47) 有一个唯一解 $u_N \in X_N$ 使得

$$\|u_N\|_X \leq \frac{1}{\alpha} \|F\|_{X'}. \quad (1.48)$$

此外, 如果 u 是 (1.35) 的解, 我们有

$$\|u - u_N\|_X \leq \frac{C}{\alpha} \inf_{v_N \in X_N} \|u - v_N\|_X. \quad (1.49)$$

这里, 常数 C 与 α 分别在 (1.36) 与 (1.37) 中已给。

证: 由于 X_N 是 X 的一个子空间, 所以对 (1.47) 应用 Lax-Milgram 引理可知 u_N 的存在唯一性与稳定性结果 (1.48)。现在在 (1.35) 中取 $v = v_N$, 然后将所得方程减去 (1.47) 得误差方程

$$a(u - u_N, v_N) = 0, \quad \forall v_N \in X_N, \quad (1.50)$$

结合 (1.36) -(1.37), 由 (1.49) 得

$$\begin{aligned} \alpha \|u - u_N\|_X^2 &\leq a(u - u_N, u - u_N) = a(u - u_N, u - v_N) \\ &\leq C \|u - u_N\|_X \|u - v_N\|_X, \quad \forall v_N \in X_N, \end{aligned}$$

□

Remark 1.6. 如果另外有双线性式是对称的, 即 $a(u, v) = a(v, u)$, 则 Galerkin 法称为 Ritz 法。在这种情况下, (1.49) 中常数的上界可改善为 $\sqrt{C\alpha^{-1}}$ 。

Remark 1.7. 在进行谱方法误差分析时, 通常我们在 (1.49) 中取 $v = \pi_N u$ 作为 u 在 X_N 上的一个合适的正交投影, 由 $\pi_N u$ 得

$$\|u - u_N\|_X \leq \frac{C}{\alpha} \|u - \pi_N u\|_X. \quad (1.51)$$

因此, 误差估计是由 $\|u - \pi_N u\|$ 的近似结果得到的, 它采用典型的形式:

$$\|u - \pi_N u\|_X \leq c N^{-\sigma(m)} \|u\|_{H^m}, \quad (1.52)$$

其中 c 是一个与 N 及任何函数无关的正常数, $\sigma(m) > 0$ 是所谓的正则指数 m 的收敛阶, H^m 是一个合适的 Sobolev 空间, 其范数涉及 u 的高至 m 阶的导数。为每一个正交多项式族/函数族建立这种近似结果将会是本书的另一个重点。

通常, 如果 u 足够光滑, 估计式 (1.52) 对任意 m 成立。然而, 对于一个有限元方法, 收敛阶严格取决于局部基函数的阶。本书还将探讨 (1.52) 类型的值估计对正则指数 m 的显示依赖性。

观察到离散问题 (1.47) 与连续问题 (1.35) 的双线性型和函数 F 相同。然而，通常使用近似双线性型和 (或) 函数是方便的 (作为例子见 (1.31))。因此，考虑下述 (1.35) 的近似是必要的：

$$\begin{cases} \text{Find } u_N \in X_N \text{ s.t.} \\ a_N(u_N, v_N) = F_N(v_N), \forall v_N \in X_N, \end{cases} \quad (1.53)$$

其中 X_N 依旧满足 (1.46)， $a_N(\cdot, \cdot)$ 和 $F_N(\cdot)$ 分别是 $a(\cdot, \cdot)$ 和 $F(\cdot)$ 合适的近似。通常，虽然 X_N 是 X 的子空间，但是离散双线性型的性质不能从连续双线性型的延续出来。因此，它们必须分开推导。

下面的结果称为 First Strang 引理 (见例 Strang and Fix (1973), Ciarlet (1978))，是定理 1.2 的推广。

定理 1.3.(First Strang 引理). 在 Lax-Milgram 引理的假设下，进一步假设离散形式 $F_N(\cdot)$ 和 $a_N(\cdot)$ 在子空间 $X_N \subset X$ 中满足相同的性质，且存在与 N 无关的 $\alpha_* > 0$, s.t.

$$a_N(v, v) \geq \alpha_* \|v\|_X^2, \forall v \in X_N. \quad (1.54)$$

然后，问题 (1.53) 有一个唯一解 $u_N \in X_N$ ，满足

$$\|u_N\|_X \leq \frac{1}{\alpha_*} \sup_{0 \neq v_N \in X_N} \frac{|F_N(v_N)|}{\|v_N\|_X}. \quad (1.55)$$

此外，如果 u 是 (1.35) 的解，我们有

$$\begin{aligned} \|u - u_N\|_X \leq & \inf_{w_N \in X_N} \left\{ \left(1 + \frac{C}{\alpha_*}\right) \|u - w_N\|_X \right. \\ & \left. + \frac{1}{\alpha_*} \sup_{0 \neq v_N \in X_N} \frac{|a(w_N, v_N) - a_N(w_N, v_N)|}{\|v_N\|_X} \right\} \\ & + \frac{1}{\alpha_*} \sup_{0 \neq v_N \in X_N} \frac{|F(v_N) - F_N(v_N)|}{\|v_N\|_X}. \end{aligned} \quad (1.56)$$

这里 C 是 (1.36) 中已给常数。

证：(1.55) 的存在唯一性与稳定性来自 Lax-Milgram 引理。(1.56) 的证明与 (1.49) 略有不同。对任意 $w_N \in X_N$ ，令 $e_N = u_N - w_N$ 。由 (1.54)、(1.35) 和 (1.53) 可得

$$\begin{aligned} \alpha^* \|e_N\|_X^2 &\leq a_N(e_N, e_N) = a(u - w_N, e_N) + a(w_N, e_N) \\ &\quad - a_N(w_N, e_N) + F_N(e_N) - F(e_N). \end{aligned}$$

由于 $e_N = 0$ 的结果是显然的，所以我们从 (1.36) 推导 $e_N \neq 0$,

$$\begin{aligned} \alpha^* \|e_N\|_X &\leq C \|u - w_N\|_X + \frac{|a(w_N, e_N) - a_N(w_N, e_N)|}{\|e_N\|_X} \\ &\quad + \frac{|F(e_N) - F_N(e_N)|}{\|e_N\|_X} \\ &\leq C \|u - w_N\|_X + \sup_{0 \neq v_N \in X_N} \frac{|a(w_N, v_N) - a_N(w_N, v_N)|}{\|v_N\|_X} \\ &\quad + \sup_{0 \neq v_N \in X_N} \frac{|F(v_N) - F_N(v_N)|}{\|v_N\|_X}, \end{aligned}$$

结合三角不等式有

$$\|u - u_N\|_X \leq \|u - w_N\|_X + \|e_N\|_X.$$

最后，在 $w_N \in X_N$ 中取下确界得到期望的结果。 \square

之前的讨论是仅限于近似基于 Galerkin 型方程的抽象问题(1.35)。通过使用定理 1.1 可以同样分析用 Petrov-Galerkin 法近似 (1.34)。令 $X_N \subseteq X, Y_N \subseteq Y$ 。考虑近似 (1.34)：

$$\begin{cases} \text{Find } u_N \in X_N \text{ s.t.} \\ a(u_N, v_N) = F(v_N), \forall v_N \in Y_N. \end{cases} \quad (1.57)$$

与强性质不同， $\inf - \sup$ 性质不能从整个空间延续到其子空间上。实际上，下确界 (1.39) 并不会因为在子空间上取值而变小，但通常上确界 $\inf - \sup$ 常数 (1.45) 在子空间上取值更小。所以，我们必须证明：

- 离散 $\inf - \sup$ 情形：

$$\exists \beta_* > 0 \text{ s.t. } \sup_{0 \neq v_N \in Y_N} \frac{|a(u_N, v_N)|}{\|u_N\|_X \|v_N\|_Y} \geq \beta_*, \forall 0 \neq u_N \in X_N, \quad (1.58)$$

- 离散“转置” $\inf - \sup$ 情形：

$$\sup_{0 \neq u_N \in X_N} |a(u_N, v_N)| > 0, \forall 0 \neq v_N \in Y_N. \quad (1.59)$$

接下来的结果是定理 1.2 的另一个推广，可以在 Babuska and Aziz (1972) 中找到。

定理 1.4. 在定理 1.1 的假设下，进一步假设 (1.58) 和 (1.59) 成立。然后离散问题 (1.57) 有一个唯一解 $u_N \in X_N$ ，满足

$$\|u_N\|_X \leq \frac{1}{\beta_*} \|F\|_{Y'}. \quad (1.60)$$

此外，如果 u 是 (1.34) 的解，我们有

$$\|u - u_N\|_X \leq \left(1 + \frac{C}{\beta_*}\right) \inf_{v_N \in X_N} \|u - v_N\|_X, \quad (1.61)$$

其中 C 是 (1.41) 中已给常数。

Remark 1.8. 如果我们考虑下述 (1.34) 的近似：

$$\begin{cases} \text{Find } u_N \in X_N \text{ s.t.} \\ a_N(u_N, v_N) = F_N(v_N), \forall v_N \in Y_N, \end{cases} \quad (1.62)$$

然后一个与定理 1.3 相似的结果被推导，证明了 (1.58) 和 (1.59) 在子空间 X_N 和 Y_N 上成立。

1.5 比较数值例子

本节的目的是提供一些说明性的数值例子，以便定性的在准确性、计算复杂程度和/或线性方程组条件方面比较：

- 全局近似与局部近似相比
- Spectral-Galerkin 法与 spectral-collocation 法相比

1.5.1 有限差分法与谱配置法相比较

为了说明有限差分法与谱配置法的主要不同之处, 我们利用四阶有限差分法和谱配置法比较周期函数 u 的数值微分。

给定 $h = \frac{2\pi}{N}$, 统一网格 $\{x_0, x_1, \dots, x_N\}, x_j = jh$, 和一组物理值 $\{u_0, u_1, \dots, u_N\}, u_j = u(x_j)$, $u'(x_j)$ 的四阶有限中心差分近似为

$$w_j := \frac{u_{j-2} - 8u_{j-1} + 8u_{j+1} - u_{j+2}}{12h}. \quad (1.63)$$

为了计算 u 的周期, 我们令

$$u_{-2} = u_{N-1}, u_{-1} = u_N, u_0 = u_{N+1}, u_1 = u_{N+2}.$$

然后, 微分过程 (1.63) 可以被表示为

$$\begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ w_{N-1} \\ w_N \end{bmatrix} = \frac{1}{12h} \begin{bmatrix} & & & 1 & -8 \\ & \ddots & & & \\ & \ddots & -1 & & 1 \\ & \ddots & 8 & \ddots & \\ & \ddots & 0 & \ddots & \\ & \ddots & -8 & \ddots & \\ -1 & & & 1 & \ddots \\ 8 & -1 & & & \ddots \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ u_{N-1} \\ u_N \end{bmatrix}. \quad (1.64)$$

注意到系数矩阵是稀疏的, 反映了有限差分法的局部性。

另一方面, 函数 u 的傅里叶配置法近似为

$$\phi(x) = \sum_{k=0}^{N-1} h_k(x) u_k, \quad (1.65)$$

其中 $h_k(x_j) = \delta_{jk}$ 且 (参见引理 2.2)

$$h_k(x) = \frac{1}{N} \frac{\sin(N(x-x_k)/2)}{\sin((x-x_k)/2)} \cos((x-x_k)/2). \quad (1.66)$$

然后, 我们通过下式近似 $u'(x_j)$

$$w_j = \phi'(x_j) = \sum_{k=0}^{N-1} h'_k(x_j) u_k, \quad j = 0, 1, \dots, N-1, \quad (1.67)$$

其中我们有显示表达式 (参见 (2.34)) :

$$h'_k(x_j) = \begin{cases} \frac{(-1)^{k+j}}{2} \cot \left[\frac{(j-k)\pi}{N} \right], & \text{if } j \neq k, \\ 0, & \text{if } j = k. \end{cases} \quad (1.68)$$

因此, 矩阵 (1.67) 变为

$$\begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ w_{N-2} \\ w_{N-1} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \ddots & \vdots & & & & \\ \ddots & -\cot \frac{2h}{2} & & & & \\ & \cot \frac{h}{2} & & & & \\ \ddots & & 0 & \ddots & & \\ & & -\cot \frac{h}{2} & \ddots & & \\ & & \cot \frac{2h}{2} & \ddots & & \\ & & \vdots & & \ddots & \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ u_{N-2} \\ u_{N-1} \end{bmatrix}. \quad (1.69)$$

注意到系数矩阵是满的, 反映了谱配置法的全局性。更多关于谱方法的细节讨论将在第二章进行。

接下来, 我们取 $u(x) = \ln(2 + \sin x)$, 其周期为 2π , 将精确微分 $u'(x) = \cos x / (2 + \sin x)$ 与通过有限差分 (1.64) 及相同网格下的傅里叶配置法 (1.69) 得到的数值微分 $\{w_j\}$ 相比较。在图像 1.1 中, 我们对不同的 N 绘制误差 $\max_{0 \leq j \leq N-1} |u'(x_j) - w_j|$ 的曲线。观察到有限差分 (1.64) 的四阶收敛 $O(h^4)$ (或 $O(N^{-4})$)。同时也观察到傅里叶配置法的收敛速度比有限差分法快。为了更清楚地了解傅里叶配置法 (1.69) 的收敛性, 我们在图 1.2 中绘出半对数标度的误差, 这表明对于某个 $c > 0$ 有指数阶收敛速度 $O(e^{-cN})$ 。

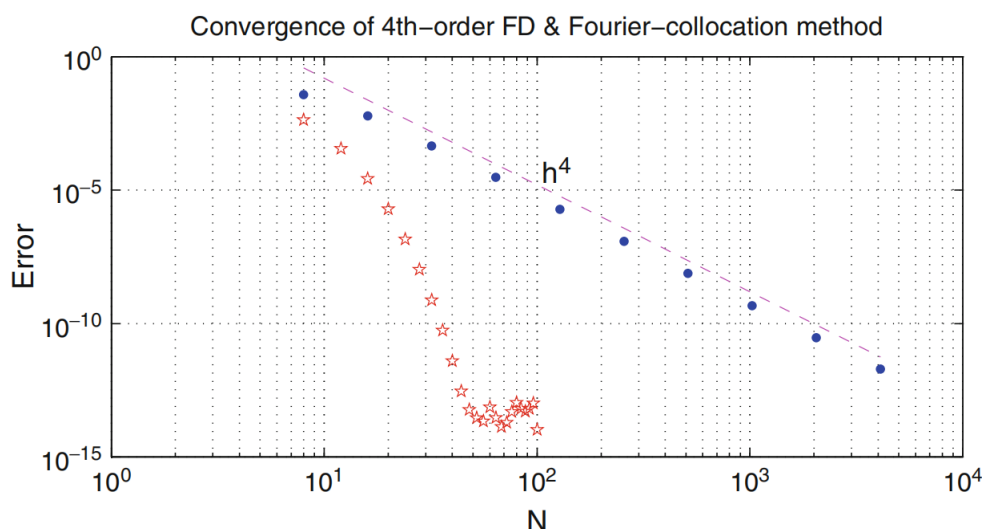


图 1.1: 四阶有限差分 (1.64) 和傅里叶配置法 (1.69) 差分过程的收敛性

Remark 1.9. 谱方法是典型的 $O(N^{-m})$ 阶收敛, 其中 m 是函数的正则性指数。换句话说, 它的收敛速率仅受函数的正则性限制。一个收敛性表现如此的方法通常在文献中被称为具有谱精度。另一方面, 有限元/有限差分法的收敛速度由方法的阶数限制, 而与函数的正则性无关。

谱方法相对于低阶有限元/有限差分法的一个主要优势是: 前者为解已给问题并达到某精确度所需的未知量更少, 这可能大大减少储存空间和

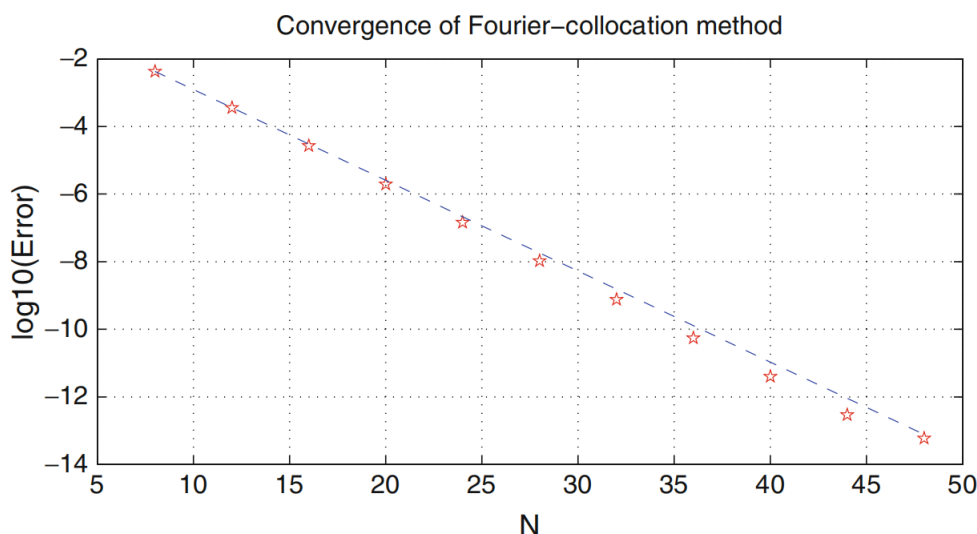


图 1.2: 傅里叶配置法 (1.69) 差分过程的收敛性

CPU 时间。例如，概测法（参见 Gottlieb and Orszag (1977)）为了达到 1% 的工程精度，使用谱方法只需要 π 点每波长，低阶方法却需要大约十点每步长。

谱方法的另一个重要特点是：离散函数的微分可以被准确计算（参见 (1.14)）。因此，谱方法是没有相位误差的，这可能对偏微分方程的长时间积分是很有问题的。

Remark 1.10. 如果一个函数在复平面上一个宽度为 2β 的条带上解析，则此函数的谱近似可以达到 $O(e^{-\beta N})$ 指数阶收敛速率。对在谱投影误差上的此类结果，我们参考了 Davis (1975), Szegő (1975), and Gottlieb 等人 (1992), Gottlieb and Shu (1997), 对谱微分误差我们参考了 Tadmor (1986) (参见 Reddy and Weideman (2005) 有一个更简单的分析，并且还改进了 Tadmor (1986) 中的估计)。由于 $O(e^{-\beta N})$ 阶指数收敛条件非常通用，我们将不会在本书中用指数阶收敛进行分析。

1.5.2 谱 Galerkin 法与配置法相比较

我们在这节从条件作用和舍入误差方面比较两种谱方法：频率空间中的 Galerkin 法和物理空间中的配置法。

作为一个说明性的例子，我们考虑下述问题

$$u - u_{xx} = f, u(\pm 1) = 0$$

其中 $u(x) = \sin(10\pi x)$ 是准确解。比较而言，通过计算 $p=0, q=\alpha, g_{\pm}=0$ 的 (1.17) - (1.18) 得到配置法的解，而通过使用 Legendre 基函数解 (1.22) 得到 Galerkin 法的解（参见 (1.23)）

$$\phi_k(x) = \frac{1}{\sqrt{4k+6}} (L_k(x) - L_{k+2}(x)), 0 \leq k \leq N-2.$$

首先让我们检查这两个线性系统的条件作用。在表格 1.1 中，我们列出了由配置法 (COL) 和 Galerkin 法 (GAL) 得到的矩阵的条件数。

我们发现对于不同的 α ，GAL 系统的条件数相对较小且与 N 无关，但 COL 系统的条件数则像 $O(N^4)$ 一样增加。

表 1.1: COL 和 GAL 的条件数

N	Method	$\alpha = 0$	$\alpha = 10$	$\alpha = 100$	$\alpha = 1,000$
32	COL	1.04E+04	2.05E+03		2.64E+01
32	GAL	1.00	5.07	41.6	396
64	COL	1.71E+05	3.37E+04		4.18E+02
64	GAL	1.00	5.07	41.6	407
128	COL	2.77E+06	5.47E+05		6.78E+03
128	GAL	1.00	5.07	41.7	408
256	COL	4.46E+07	8.81E+06	1.07E+06	1.09E+05
256	GAL	1.00	5.07	41.7	408

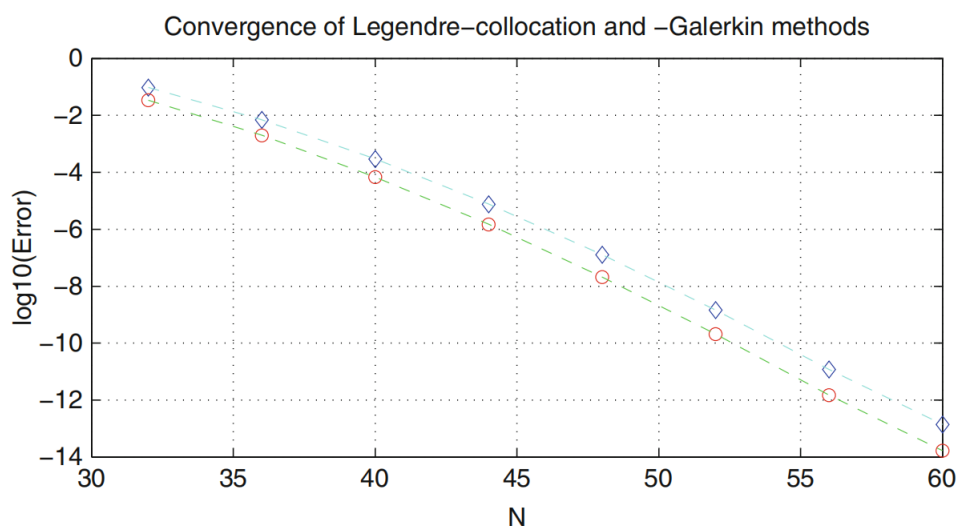


图 1.3: 收敛性: COL(“◇”)vs.GAL(“○”)

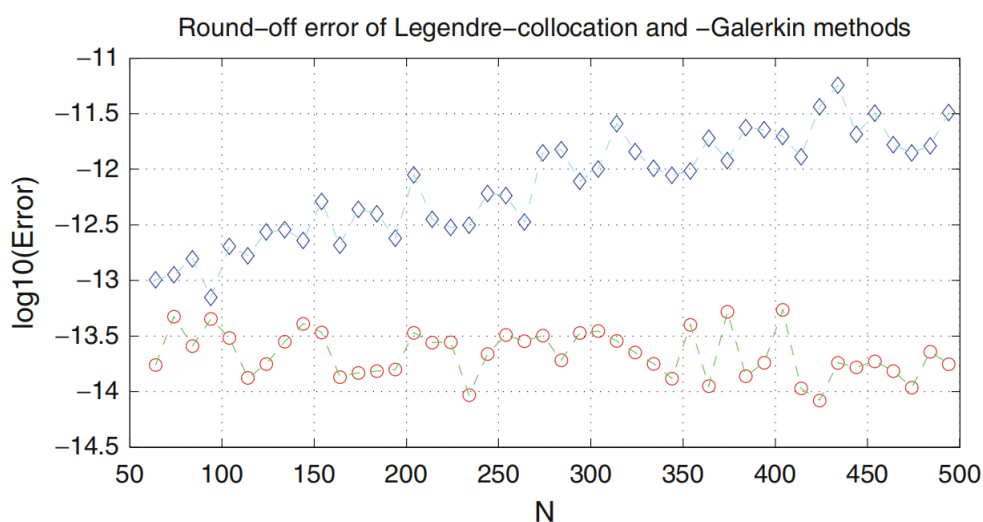


图 1.4: 舍入误差: COL(“◇”)vs.GAL(“○”)

接下来我们比较舍入误差的影响。对于不同的 N 两种方法的最大逐点误差在图 1.3 和 1.4 中示出。从图 1.3 观察到，对于相对小的 N ，两种方法基本上有相同阶的收敛速率。然而，图 1.4 表明在配置法中舍入误差的影响也许会随 N 增大而变严重。

以上比较仅在一维问题上进行。应该指出，对于多维和/或更高阶问题可以同样进行类似的比较。最后，我们想强调在一个配置法中，配置点（求积节

点) 的选取应该与已给微分方程和边界条件一起讨论。例如, Gauss-Lobatto 点并不适合三阶方程 (参见 Huang and Sloan (1992), Merryfield and Shizgal (1993))。然而, 在一个谱 Galerkin 法中, 求积规则的使用仅仅是为了求积, 所以通常的 Gauss-Lobatto 求积法也适用于三阶方程。

第二章 周期性问题的傅里叶谱方法

刘双

谱方法是 Orszag 在大约四十年前使用傅里叶级数模拟不可压缩流体的开创性工作时引入的 (参见 Orszag(1971))。“谱”一词大概是源于傅里叶级数是拉普拉斯算子在周期性边界条件下的特征函数这一事实。该事实和快速傅里叶变换的可用性是使得傅里叶方法广泛应用于具有周期性边界条件问题的两个主要原因。在实践中,许多物理问题表现出周期性。例如,一些问题在几何上和物理上是周期的,比如晶体结构和均匀湍流。另一方面,许多具有科学意义的问题,比如孤立波的相互作用和均匀湍流都能通过具有周期性边界条件的偏微分方程来建模。此外,即使初始问题不是周期性的,也可以通过坐标变换使其是周期性的,比如极坐标变换、球坐标变换及圆柱变换。的确,存在许多仅在一个或两个方向上是周期的而其它方向是非周期的问题。在这种情况下,很自然的在具有周期性的方向上使用傅里叶级数,在非周期性的方向上则使用例如 Legendre 或 Chebyshev 多项式等其它类型的谱展开 (参见第七章)。

本章的目的是针对周期性问题的傅里叶谱方法在计算和理论方面的一些内容。在第一节中,我们介绍了连续和离散的傅里叶级数,并且研究了包括离散傅里叶变换、傅里叶微分矩阵和基于快速傅里叶变换的傅里叶谱微分的基本谱技术。连续和离散傅里叶级数的近似性质在第二节中讨论。在最后一节展示一些线性和非线性问题的傅里叶谱方法应用。更多关于傅里叶近似的细节和其它方面的内容,我们参考 Gottlieb and Orszag (1977),Gottlieb 等 (1984), Boyd(2001) 及其中的参考文献。

2.1 连续和离散傅里叶变换

这节致力于简单回顾傅里叶级数和傅里叶变换的性质。我们的关注点放在离散傅里叶变换和傅里叶微分技术上，它们在傅里叶谱方法中发挥着重要的作用。

2.1.1 连续傅里叶级数

我们用下式来表示复指数

$$E_k(x) := e^{ikx} = \cos kx + i \sin kx = (\cos x + i \sin x)^k, \quad k \in \mathbb{Z}, x \in \mathbb{R},$$

其中 $i = \sqrt{-1}$ 。集合 $\{e^{ikx} : k \in \mathbb{Z}\}$ 构成复希尔伯特空间 $L^2(0, 2\pi)$ 上的一个完全正交组，其中配有内积和范数

$$(u, v) = \frac{1}{2\pi} \int_0^{2\pi} u(x) \bar{v}(x) dx, \quad \|u\| = \sqrt{(u, u)},$$

其中 \bar{v} 是 v 的共轭复数。 $\{E_k : k \in \mathbb{Z}\}$ 的正交性读取

$$(E_k, E_m) = \frac{1}{2\pi} \int_0^{2\pi} e^{i(k-m)x} dx = \delta_{km} \quad (2.1)$$

其中 δ_{km} 表示克罗内克符号。

对任意复值函数 $u \in L^2(0, 2\pi)$ ，它的傅里叶级数定义为

$$u(x) \sim \mathcal{F}(u)(x) := \sum_{k=-\infty}^{\infty} \hat{u}_k e^{ikx}, \quad (2.2)$$

其中傅里叶系数通过下式给定

$$\hat{u}_k = (u, e^{ikx}) = \frac{1}{2\pi} \int_0^{2\pi} u(x) e^{-ikx} dx. \quad (2.3)$$

如果 u 是实值函数，显然它的傅里叶系数满足

$$\hat{u}_{-k} = \bar{\hat{u}}_k, \quad k \in \mathbb{Z}, \quad (2.4)$$

且 \hat{u}_0 显然是实数。

事实上, 傅里叶级数可以定义为 $(0, 2\pi)$ 上通常绝对可积函数, 傅里叶展开的收敛性理论在傅里叶分析中已经有过严谨并彻底的研究 (例如参见 Zygmund (2002), Stein and Shakarchi (2003))。众所周知, 对于任意 $u \in L^2(0, 2\pi)$, 其被截傅里叶级数 $\mathcal{F}_N(u) := \sum_{|k| \leq N} \hat{u}_k e^{ikx}$ 在 L^2 意义下收敛至 u , 且有 Parseval 恒等式:

$$\|u\|^2 = \sum_{k=-\infty}^{\infty} |\hat{u}_k|^2. \quad (2.5)$$

如果 u 在 $[0, 2\pi]$ 上是连续的、周期的和有界变差的, 那么 $\mathcal{F}_N(u)$ 一致收敛于 u 。

注意到截断傅里叶级数也可以用卷积的形式表示, 即

$$\mathcal{F}_N(u)(x) = (\mathcal{D}_N * u)(x) = \frac{1}{2\pi} \int_0^{2\pi} \mathcal{D}_N(x-t)u(t)dt, \quad (2.6)$$

其中 \mathcal{D}_N 是已知的狄利克雷内核

$$\mathcal{D}_N(x) := \sum_{k=-N}^N e^{ikx} = 1 + 2 \sum_{k=1}^N \cos kx = \frac{\sin((N+1/2)x)}{\sin(x/2)}. \quad (2.7)$$

有时用三角函数表示傅里叶级数较方便:

$$u(x) \sim \mathcal{F}(u)(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx), \quad (2.8)$$

其中展开项系数

$$a_k = \frac{1}{\pi} \int_0^{2\pi} u(x) \cos kx dx, \quad b_k = \frac{1}{\pi} \int_0^{2\pi} u(x) \sin kx dx.$$

系数的两种不同表达式 (2.2) 和 (2.8) 有下述关系

$$\hat{u}_0 = \frac{a_0}{2}, \quad \hat{u}_k = \begin{cases} \frac{a_k - ib_k}{2}, & \text{if } k \geq 1, \\ \frac{a_{-k} + ib_{-k}}{2}, & \text{if } k \leq -1. \end{cases} \quad (2.9)$$

特别地, 如果 u 是实值函数, 则

$$a_0 = 2\hat{u}_0, a_k = 2\operatorname{Re}(\hat{u}_k), b_k = -2\operatorname{Im}(\hat{u}_k), k \geq 1. \quad (2.10)$$

2.1.2 离散傅里叶变换

给定一个正整数 N , 令

$$x_j = ih = j\frac{2\pi}{N}, 0 \leq j \leq N-1, \quad (2.11)$$

为 $[0, 2\pi)$ 上的 N 等分间隔网格, 被称为傅里叶配置点。我们通过下式定义离散内积

$$\langle u, v \rangle_N = \frac{1}{N} \sum_{j=0}^{N-1} u(x_j) \bar{v}(x_j). \quad (2.12)$$

下述引理是 (2.1) 的离散情况对应项。

引理 2.1. 令 $E_l(x) = e^{ilx}$ 。对任意整数 $N \leq 1$, 我们有

$$\langle E_k, E_m \rangle_N = \begin{cases} 1, & \text{if } k-m = lN, \forall l \in \mathbb{Z}, \\ 0. & \text{otherwise.} \end{cases} \quad (2.13)$$

证：观察到如果 $k-m$ 不能被 N 整除, 则

$$\begin{aligned} \langle E_k, E_m \rangle_N &= \frac{1}{N} \sum_{j=0}^{N-1} e^{i(k-m)x_j} = \frac{1}{N} \sum_{j=0}^{N-1} \left(e^{2\pi i(k-m)/N} \right)^j \\ &= \frac{1}{N} \frac{e^{2\pi i(k-m)} - 1}{e^{2\pi i(k-m)/N} - 1} = 0. \end{aligned}$$

如果 $k-m$ 能被 N 整除, 我们有 $e^{2\pi i(k-m)/N} = 1$, 所以上面第二行的和等于 1。□

通常, 在 (2.3) 中傅里叶系数 \hat{u}_k 不能被准确估计, 所以我们要用到求积公式。一种简单而精确的针于周期为 2π 函数的求积公式是矩形规则

$$\frac{1}{2\pi} \int_0^{2\pi} v(x) dx \approx \frac{1}{N} \sum_{j=0}^{N-1} v(x_j), \quad \forall v \in C[0, 2\pi), \quad (2.14)$$

其中对 $v \in \text{span}\{e^{ikx} : 0 \leq |k| \leq N-1\}$ 精确成立。

此外，易验证 (2.14) 对与 $v = \sin(\pm Nx)$ 也精确成立，但对 $v = \cos(\pm Nx)$ 不成立。

应用 (2.14) 到 (2.3) 得到近似

$$\hat{u}_k \approx \tilde{u}_k := \frac{1}{N} \sum_{j=0}^{N-1} u(x_j) e^{-ikx_j}, \quad k = 0, \pm 1, \dots \quad (2.15)$$

注意到 $\{\tilde{u}_k\}$ 是 N 周期的，即

$$\tilde{u}_{k \pm N} = \frac{1}{N} \sum_{j=0}^{N-1} u(x_j) e^{-i(k \pm N)x_j} = \frac{1}{N} \sum_{j=0}^{N-1} u(x_j) e^{-ikx_j} e^{\mp 2\pi i j} = \tilde{u}_k,$$

这意味着对于偶数 N ，我们有

$$\tilde{u}_{-N/2} = \tilde{u}_{N/2}. \quad (2.16)$$

因此，对于偶数 N ，网格点 $\{x_j\}_{j=0}^{N-1}$ 不能区分： $k = \pm N/2$ ，因为

$$e^{iNx_j/2} = e^{ij\pi} = (-1)^j = e^{-iNx_j/2}, \quad 0 \leq j \leq N-1. \quad (2.17)$$

换句话说，两种情况 $k = \pm N/2$ 是混叠的。

为了有效实行离散傅里叶变换 (DFT)，最好使用偶数 N ，相应的，考虑对称有限集 $-N/2 \leq k \leq N/2$ 情况下的傅里叶级数（参见如下 (2.20)）。鉴于 (2.16) - (2.17)，我们通过两种情况的 $k = \pm N/2$ 重新定义近似 (2.15)：

$$\hat{u}_k \approx \tilde{u}_k = \frac{1}{N c_k} \sum_{j=0}^{N-1} u(x_j) e^{-ikx_j}, \quad k = -N/2, \dots, N/2, \quad (2.18)$$

其中 $\forall |k| < N/2, c_k = 1; \forall k = \pm N/2, c_k = 2$ 。表达式 (2.18) 被称为与 (2.11) 中网格点相关的 $u(x)$ 的离散傅里叶变换。

由 (2.16) 知只有 N 个独立系数。因此, 我们设

$$\mathcal{T}_N = \{u = \sum_{k=-N/2}^{N/2} \tilde{u}_k e^{ikx} : \tilde{u}_{-N/2} = \tilde{u}_{N/2}\}, \quad (2.19)$$

同时定义映射 $I_N : C[0, 2\pi) \rightarrow \mathcal{T}_N$:

$$(I_N u)(x) = \sum_{k=-N/2}^{N/2} \tilde{u}_k e^{ikx}, \quad (2.20)$$

其中 $\{\tilde{u}_k\}$ 在 (2.18) 中已给出。下述引理说明 I_N 是从 $C[0, 2\pi)$ 映至 \mathcal{T}_N 的插值算子, 且有

$$(I_N u)(x_j) = u(x_j), \quad x_j = \frac{2\pi j}{N}, \quad 0 \leq j \leq N-1. \quad (2.21)$$

引理 2.2. 对任意 $u \in C[0, 2\pi)$,

$$(I_N u)(x) = \sum_{j=0}^{N-1} u(x_j) h_j(x), \quad (2.22)$$

其中

$$h_j(x) = \frac{1}{N} \sin \left[N \frac{x - x_j}{2} \right] \cot \left[\frac{x - x_j}{2} \right] \in \mathcal{T}_N \quad (2.23)$$

满足

$$h_j(x_k) = \delta_{jk}, \quad \forall j, k = 0, 1, \dots, N-1. \quad (2.24)$$

证: 由 (2.18) 与 (2.19),

$$\begin{aligned}
(I_N u)(x) &= \sum_{k=-N/2}^{N/2} \left(\frac{1}{N c_k} \sum_{j=0}^{N-1} u(x_j) e^{-ikx_j} \right) e^{ikx} \\
&= \sum_{j=0}^{N-1} \left(\frac{1}{N} \sum_{k=-N/2}^{N/2} \frac{1}{c_k} e^{ik(x-x_j)} \right) u(x_j) \\
&=: \sum_{j=0}^{N-1} h_j(x) u(x_j).
\end{aligned}$$

由 (2.7) 直接计算得到

$$\begin{aligned}
h_j(x) &= \frac{1}{N} \sum_{k=-N/2}^{N/2} \frac{1}{c_k} e^{ik(x-x_j)} \\
&= \frac{1}{N} \left(D_{N/2-1}(x-x_j) + \cos \left[N \frac{x-x_j}{2} \right] \right) \\
&= \frac{1}{N} \left(\frac{\sin \left[(N-1) \frac{x-x_j}{2} \right]}{\sin \frac{x-x_j}{2}} + \cos \left[N \frac{x-x_j}{2} \right] \right) \\
&= \frac{1}{N} \sin \left[N \frac{x-x_j}{2} \right] \cot \left[\frac{x-x_j}{2} \right].
\end{aligned} \tag{2.25}$$

由 (2.17), 我们有 $h_j(x) \in \mathcal{T}_N$, 显然对于 $i \neq j$ 时有 $h_j(x_j) = 0$. 此外, 在 (2.25) 第一个等式中取 $x = x_j$ 得 $h_j(x_j) = 1$. \square

在 (2.20) 中取 $x = x_j$ 并结合 (2.21) 得傅里叶逆变换 (或向后傅里叶变换):

$$u(x_j) = \sum_{k=-N/2}^{N/2} \tilde{u}_k e^{ikx_j}, \quad j = 0, 1, \dots, N-1. \tag{2.26}$$

显然傅里叶变换 (2.18) 与其傅里叶逆变换可以通过 $O(N^2)$ 阶运算的矩阵向量乘法来实现. 然而, 多亏了 Cooley 和 Tukey(1965) 的快速傅里叶变

换, 这样的过程可以通过 $O(N \log_2 N)$ 阶运算来实现。此外, 如果 u 是实值函数, 那么 $\tilde{u}_{-k} = \tilde{u}_k$, 所以仅 (2.26) 中一半的系数需要计算/储存。

许多软件包都提供 FFT 和 IFFT 的计算程序。这里, 我们将注意力集中在 Matlab 上。给定在 $\{x_k = 2\pi k/N\}_{k=0}^{N-1}$ 上的样本数据 $\{\mathbf{v}(j) = u(x_{j-1})\}_{j=1}^N$, 命令 “ $\tilde{\mathbf{v}} = \text{fft}(\mathbf{v})$ ” 返回通过下式定义的向量 $\{\tilde{\mathbf{v}}(k)\}_{k=1}^N$,

$$\tilde{\mathbf{v}}(k) = \sum_{j=1}^N \mathbf{v}(j) e^{-2\pi i(j-1)(k-1)/N}, \quad 1 \leq k \leq N, \quad (2.27)$$

傅里叶逆变换可以通过命令 “ $\mathbf{v} = \text{ifft}(\tilde{\mathbf{v}})$ ” 计算, 通过下式返回物理值 $\{\mathbf{v}(j)\}_{j=1}^N$

$$\mathbf{v}(j) = \frac{1}{N} \sum_{k=1}^N \tilde{\mathbf{v}}(k) e^{2\pi i(j-1)(k-1)/N}, \quad 1 \leq j \leq N. \quad (2.28)$$

请注意, 必须在模式的顺序上有一些警惕。为了说明这一点, 我们研究了变换 (2.18) 与 (2.26) 之间的一一对应关系。更准确的, 令

$$u(x_j) = \mathbf{v}(j+1), \quad x_j = \frac{2\pi j}{N}, \quad 0 \leq j \leq N-1. \quad (2.29)$$

我们发现

$$\begin{aligned} \tilde{u}_k &= \frac{1}{N} \tilde{\mathbf{v}}(k+1), \quad 0 \leq k \leq \frac{N}{2} - 1, \\ \tilde{u}_k &= \frac{1}{N} \tilde{\mathbf{v}}(k+N+1), \quad -\frac{N}{2} + 1 \leq k \leq -1, \\ \tilde{u}_{-N/2} &= \tilde{u}_{N/2} = \frac{1}{2N} \tilde{\mathbf{v}}(N/2+1). \end{aligned} \quad (2.30)$$

下表给出了上述关系的表格视图。

在表格中, 我们列出 $\{u_j = \mathbf{u}(j)\}_{j=0}^{N-1}$ 和 $\{\tilde{u}_k = \tilde{\mathbf{u}}(k)\}_{k=-N/2}^{N/2}$ 。最后一行给出了基于 FFT 的傅里叶谱微分的频率向量 \mathbf{k} , 见后面 2.1.4 节。注意到在离散傅里叶变换中, 频率 $-N/2$ 与 $N/2$ 相混叠。

表 2.1: Matlab 中 DFT 与 FFT&IFFT 的对应关系

j	1	2	...	N/2-1	N/2	N/2+1	N/2+2	...	N-1	N
$\mathbf{u} = \mathbf{v}$	u_0	u_1	u_{N-2}	u_{N-1}
$\tilde{\mathbf{u}} = \tilde{\mathbf{v}}/N$	\tilde{u}_0	\tilde{u}_1	...	$\tilde{u}_{N/2-2}$	$2\tilde{u}_{N/2-1}$	$\tilde{u}_{N/2}$	$\tilde{u}_{-N/2+1}$...	\tilde{u}_{-2}	\tilde{u}_{-1}
\mathbf{k}	0	1	...	N/2-2	N/2-1	0	-N/2+1	...	-2	-1

2.1.3 物理空间上的微分

在一个傅里叶谱方法中，微分可以在频率空间上进行，也可以在物理空间上进行。

我们从物理空间上的微分开始。令 $\{x_j\}$ 和 $\{h_j\}$ 分别为 (2.11) 和 (2.23) 中定义的。设

$$u(x) = \sum_{j=0}^{N-1} u(x_j) h_j(x), \quad (2.31)$$

同时取 m 次微分得

$$u^{(m)}(x) = \sum_{j=0}^{N-1} u(x_j) h_j^{(m)}(x). \quad (2.32)$$

该过程可以被表示成矩阵向量乘法

$$\mathbf{u}^{(m)} = D^{(m)} \mathbf{u}, \quad m \geq 0, \quad (2.33)$$

其中

$$\begin{aligned} D^{(m)} &= \left(d_{kj}^{(m)} := h_j^{(m)}(x_k) \right)_{k,j=0,\dots,N-1}, \\ \mathbf{u} &= (u(x_0), u(x_1), \dots, u(x_{N-1}))^T, \\ \mathbf{u}^{(m)} &= \left(u^{(m)}(x_0), u^{(m)}(x_1), \dots, u^{(m)}(x_{N-1}) \right)^T. \end{aligned}$$

特别地，有 $D = D^{(1)}$ 。下面给出一阶微分矩阵的紧化形式。

引理 2.3. 一阶傅里叶微分矩阵的元素由下式决定

$$d_{kj}^{(1)} = h'_j(x_k) = \begin{cases} \frac{(-1)^{k+j}}{2} \cot \left[\frac{(k-j)\pi}{N} \right], & \text{if } k \neq j, \\ 0, & \text{if } k = j. \end{cases} \quad (2.34)$$

证：直接微分 (2.23) 中的 Lagrange 基得

$$h'_j(x) = \frac{1}{2} \cos \left[N \frac{x-x_j}{2} \right] \cot \left[\frac{x-x_j}{2} \right] - \frac{1}{2N} \sin \left[N \frac{x-x_j}{2} \right] \csc^2 \left[\frac{x-x_j}{2} \right].$$

显然如果 $x = x_k \neq x_j$, 那么第二项是 0, 第一项正好满足 (2.34) 中的表达。

现在我们考虑 $k = j$ 的情况。为了方便, 令 $\theta = (x - x_j)/2$, 重写上述表达式为

$$h'_j(x) = \frac{1}{2} \frac{\cos(N\theta) \cos \theta \sin \theta - N^{-1} \sin(N\theta)}{\sin^2 \theta}. \quad (2.35)$$

使用泰勒展开, 我们发现

$$\cos(N\theta) \cos \theta \sin \theta = \theta + O(\theta^3), \quad N^{-1} \sin(N\theta) = \theta + O(\theta^3), \quad |\theta| \ll 1.$$

因此, 我们由 (2.35) 推出当 $\theta \rightarrow 0$ 时, $x \rightarrow x_j, h'_j(x) = \lim_{x \rightarrow x_j} h'_j(x) = 0$ 。

□

Remark 2.1. 一阶傅里叶微分矩阵有下述性质：

- D 是实斜对称矩阵, 因为 $\cot(-x) = -\cot(x), d_{kk} = 0$ 。
- D 是循环 Toeplitz 矩阵, 因为 $d_{kj} = d_{k+1, j+1}$ 。
- $\{ik : -N/2 + 1 \leq k \leq N/2 - 1\}$ 是 D 的不同的特征值, 且 0 是二重特征值。

高阶微分的近似有着相同的过程。从 (2.25) 的第一个关系式中，我们发现

$$h_j^{(m)}(x_i) = \frac{1}{N} \sum_{k=-N/2}^{N/2} \frac{(ik)^m}{c_k} e^{2\pi i k(i-j)/N}. \quad (2.36)$$

特别地，二阶微分矩阵 $D^{(2)}$ 的元素由下式给出

$$d_{kj}^{(2)} = h_j''(x_k) = \begin{cases} -\frac{(-1)^{k+j}}{2} \sin^{-2} \left[\frac{(k-j)\pi}{N} \right], & \text{if } k \neq j, \\ -\frac{N^2}{12} - \frac{1}{6}, & \text{if } k = j. \end{cases} \quad (2.37)$$

必须指出的是 $D^{(2)} \neq D^2$ 。确实，我们考虑 $u = \cos(Nx/2)$ 并用 \mathbf{u} 在 $\{x_j\}_{j=0}^{N-1}$ 上的值构成的向量 \mathbf{u} 。由于 $u(x_j) = (-1)^j$ ，立马验证有 $D\mathbf{u} = \mathbf{0}$ 和 $D^2\mathbf{u} = \mathbf{0}$ ，而 $D^{(2)}\mathbf{u} = -N^2\mathbf{u}/6$ 。

显然 (2.33) 中的微分过程需要 $O(N^2)$ 次运算。我们将在后面展示如何在频率空间上用 FFT 使得实行微分只需要 $O(N \log_2 N)$ 次运算。

2.1.4 频率空间上的微分

对于一个通过 (2.31) 给出的函数，我们可以重写为一个有限傅里叶级数

$$u(x) = \sum_{k=-N/2}^{N/2} \tilde{u}_k e^{ikx}, \quad (2.38)$$

其中 $\tilde{u}_{N/2} = \tilde{u}_{-N/2}$ 与之前一样。因此，我们有

$$u'(x_j) = \sum_{k=-N/2}^{N/2} ik \tilde{u}_k e^{ikx_j}, \quad (2.39)$$

其中 $\{x_j = 2\pi j/N\}_{j=0}^{N-1}$ 是 (2.11) 中已给网格点。给出物理值 $\{u(x_j)\}_{j=0}^{N-1}$ ，微分值 $\{w_j = u'(x_j)\}_{j=0}^{N-1}$ 的近似可以通过如下被计算：

- Call $\tilde{\mathbf{v}} = \text{fft}(\mathbf{v})$, 其中输入向量 \mathbf{v} 的分量为 $v(j) = u(x_{j-1}), j = 1, \dots, N$, 返回频率向量:

$$\tilde{\mathbf{v}} = (\tilde{v}_1, \tilde{v}_2, \dots, \tilde{v}_N).$$

- 计算微分的扩展系数:

$$\begin{aligned} \tilde{\mathbf{v}}^{(1)} &= i\mathbf{k} \cdot \tilde{\mathbf{v}} \\ &= i(0, \tilde{v}_2, \dots, (N/2-1)\tilde{v}_{N/2}, 0, (-N/2+1)\tilde{v}_{N/2+2}, \dots, -\tilde{v}_N), \end{aligned}$$

其中乘法向量 \mathbf{k} 在表 2.1 已给出:

$$\mathbf{k} = (0, 1, \dots, N/2-1, 0, -N/2+1, \dots, -1). \quad (2.40)$$

- Call $\mathbf{w} = \text{ifft}(\tilde{\mathbf{v}}^{(1)})$, 产生期望的微分值 $\{w_j\}_{j=0}^{N-1}$.

与上一节所描述的微分过程形成鲜明的对比, 上述过程所需计算量为 $O(N \log_2 N)$ 。此外, 高阶微分可以通过重复使用上述三步来计算。多维情况可以通过使用合适的例程来类似的实行, 例如 Matlab 中的 `fft2.m` 和 `ifft2.m`。

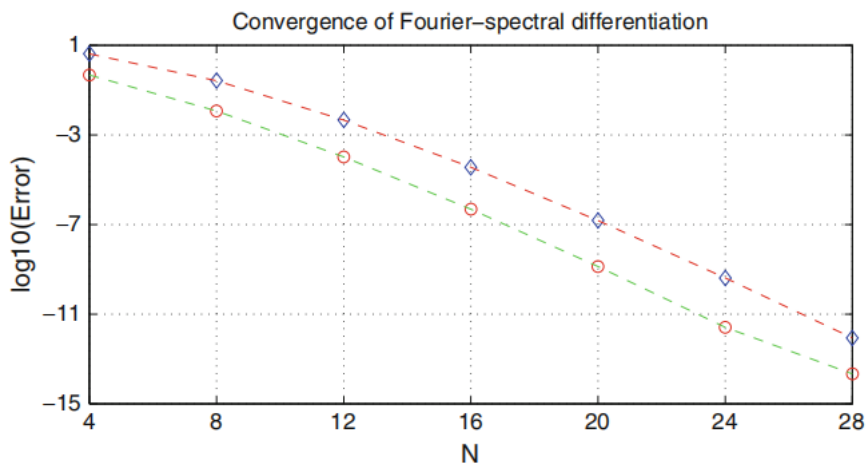


图 2.1: 一阶 (“o”) 与二阶 (“◇”) 傅里叶谱微分的误差

作为数值例子, 我们考虑以 2π 为周期的函数 $u(x) = e^{1+\sin x}$ 的傅里叶谱微分。在图 2.1 中, 我们以半对数标度绘出不同 N 对应的一阶与二阶微分最大逐点误差。图 2.1 清楚表明了傅里叶谱微分的指数收敛过程。

2.2 傅里叶近似

本节，我们将通过连续和离散傅里叶级数来概括周期函数近似的一些基本结论。

2.2.1 逆不等式

由于有限维空间上的所有范数等价，我们可以通过取决于空间维数的约束常数使一个弱范数约束一个强范数。这种类型的不等式被称为逆不等式。本节的主要目的是在逆不等式中寻找一个最佳的常数。

为了表示方便，我们用 $A \lesssim B$ 来表示：存在一个与 N 和任意函数无关的通用正常数 c ，使得 $A \leq cB$ 。我们也用 $\partial_x^m u$ 或 $u^{(m)}$ 来表示常微分 $\frac{d^m u}{dx^m}$ 。令 $I := (0, 2\pi)$ ，定义 $2N+1$ 维复空间

$$X_N := \text{span}\{e^{ikx} : -N \leq k \leq N\}. \quad (2.41)$$

Banach 空间 $L^p(I)$ ，其中 $1 \leq p \leq \infty$ ，范数 $\|\cdot\|_{L^p}$ 定义如附录 B.4。我们首先回忆下述 Nikolski 不等式

引理 2.4. 对任意 $u \in X_N$ 和 $1 \leq p \leq q \leq \infty$,

$$\|u\|_{L^q} \leq \left(\frac{Np_0 + 1}{2\pi} \right)^{\frac{1}{p} - \frac{1}{q}} \|u\|_{L^p}, \quad (2.42)$$

其中 p_0 是最小的 $\geq p$ 的偶数。

另一种类型的逆不等式，即所谓的 Bernstein 不等式，将函数及其导数的 L^p 范数联系起来。

引理 2.5. 对任意 $u \in X_N$ 和 $1 \leq p \leq \infty$,

$$\|\partial_x^m u\|_{L^p} \lesssim N^m \|u\|_{L^p}, \quad m \geq 1. \quad (2.43)$$

特别地，对于 $p = 2$,

$$\|\partial_x^m u\| \lesssim N^m \|u\|. \quad (2.44)$$

上述逆不等式的证明可以在 Butzer and Nessel(1971) 中找到 (也可参考 Guo(1998b))。特别地, 通过使用 $\partial_x^m(e^{ikx}) = (ik)^m e^{ikx}$ 和傅里叶级数的正交性可以直接推导出 (2.44)。

2.2.2 正交投影

令 $P_N: L^2(I) \rightarrow X_N$ 为 L^2 正交投影, 定义为

$$(P_N u - u, v) = 0, \forall v \in X_N. \quad (2.45)$$

显然 $P_N u$ 是截断傅里叶级数, 即

$$(P_N u)(x) = \sum_{k=-N}^N \hat{u}_k e^{ikx},$$

其中 $\{\hat{u}_k\}$ 在 (2.3) 中已给。

接下来在 Sobolev 空间中测量 $P_N u$ 与 u 之间的误差。为了这个目的, 我们用 $H^m(I)$ 的子空间 $H_p^m(I)$ (参见附录 B.4), 它由具有高达 $m-1$ 阶导数的 2π 周期的函数组成。从 $(e^{ikx})' = ike^{ikx}$ 的关系来看, $H_p^m(I)$ 的范数和半范数可以通过下式在频率空间上表示

$$\|u\|_m = \left(\sum_{k=-\infty}^{\infty} (1+k^2)^m |\hat{u}_k|^2 \right)^{1/2}, \quad |u|_m = \left(\sum_{k=-\infty}^{\infty} |k|^{2m} |\hat{u}_k|^2 \right)^{1/2}. \quad (2.46)$$

我们看到带分数 m 的空间 $H_p^m(I)$ 也是很好定义的。

形式上, 对于任意 $u \in H_p^m(I)$, 我们可以对其傅里叶级数微分得到

$$\partial_x^l u(x) = \sum_{k=-\infty}^{\infty} (ik)^l \hat{u}_k e^{ikx}, \quad 0 \leq l \leq m,$$

这暗示了导数算子与投影算子的可交换性:

$$\partial_x^l (P_N u) = P_N (\partial_x^l u), \quad 0 \leq l \leq m. \quad (2.47)$$

主要近似结果如下(参见 Kreiss and Oliger(1979), Canuto and Quarteroni(1982))。

定理 2.1. 对任意 $u \in H_p^m(I)$ 和 $0 \leq \mu \leq m$,

$$\|P_N u - u\|_\mu \lesssim N^{\mu-m} |u|_m. \quad (2.48)$$

证：由 (2.46)

$$\begin{aligned} \|P_N u - u\|_\mu^2 &= \sum_{|k|>N} (1+k^2)^\mu |\hat{u}_k|^2 \\ &\lesssim N^{2\mu-2m} \sum_{|k|>N} |k|^{2m-2\mu} (1+k^2)^\mu |\hat{u}_k|^2 \\ &\lesssim N^{2\mu-2m} \sum_{|k|>N} |k|^{2m} |\hat{u}_k|^2 \\ &\lesssim N^{2\mu-2m} |u|_m^2. \end{aligned}$$

证毕。□

这个定理表明 $P_N u$ 是 u 在所有 Sobolev 空间 $H_p^m(I)$ $m \geq 0$ 上的最佳近似。投影误差的 L^∞ 估计如下。

定理 2.2. 对任意 $u \in H_p^m(I)$, 其中 $m > 1/2$,

$$\max_{x \in [0, 2\pi]} |(P_N u - u)(x)| \leq \sqrt{\frac{1}{2m-1}} N^{1/2-m} |u|_m. \quad (2.49)$$

证：由 Cauchy-Schwarz 不等式有,

$$\begin{aligned} |(P_N u - u)(x)| &\leq \sum_{|k|>N} |\hat{u}_k| \leq \left(\sum_{|k|>N} |k|^{-2m} \right)^{1/2} \left(\sum_{|k|>N} |k|^{2m} |\hat{u}_k|^2 \right)^{1/2} \\ &\leq \sqrt{\frac{1}{2m-1}} N^{1/2-m} |u|_m. \end{aligned}$$

最后一步基于事实：对于 $m > 1/2$,

$$\sum_{|k|>N} |k|^{-2m} \leq \int_N^\infty x^{-2m} dx \leq \frac{N^{1-2m}}{2m-1}.$$

证毕。□

2.2.3 插值

为了保持一致性, 我们考虑 $2N$ 个点 $\{x_j = \pi j/N\}_{j=0}^{2N-1}$ 的傅里叶插值, 但是依然使用 I_N 表示插值算子。即

$$(I_N u)(x) = \sum_{k=-N}^N \tilde{u}_k e^{ikx} \quad (2.50)$$

其中 $\tilde{u}_N = \tilde{u}_{-N}$ 且

$$\tilde{u}_k = \frac{1}{2Nc_k} \sum_{j=0}^{2N-1} u(x_j) e^{-ikx_j}, \quad -N \leq k \leq N. \quad (2.51)$$

插值误差: $I_N u - u$ 由下述定理表示(也可参见 Kreiss and Oliger (1979), Canuto and Quarteroni(1982)):

定理 2.3. 对于任意 $u \in H_p^m(I)$ 和 $m > 1/2$,

$$\|\partial_x^l (I_N u - u)\| \lesssim N^{l-m} |u|_m, \quad 0 \leq l \leq m. \quad (2.52)$$

证: 首先我们说明连续 (参见 (2.3)) 和离散 (参见 (2.51)) 傅里叶级数的展开项系数通过下式相联系

$$c_k \tilde{u}_k = \hat{u}_k + \sum_{|p|>0} \hat{u}_{k+2pN}. \quad (2.53)$$

的确, 将 $u(x_j) = \sum_{|p|=0}^{\infty} \hat{u}_p e^{ipx_j}$ 代入 (2.51) 得

$$c_k \tilde{u}_k = \frac{1}{2N} \sum_{j=0}^{2N-1} \left(\sum_{|p|=0}^{\infty} \hat{u}_p e^{i(p-k)x_j} \right) = \frac{1}{2N} \sum_{|p|=0}^{\infty} \left(\sum_{j=0}^{2N-1} e^{i(p-k)x_j} \right) \hat{u}_p,$$

其中当 $|k| < N$ 时, 常数 $c_k = 1$; 当 $k = \pm N$ 时, 常数 $c_k = 2$ 。

我们从引理 2.1 推出: 当且仅当 $p - k = 2lN, l \in \mathbb{Z}$ 时, $e^{i(p-k)x_j} = 1$, 其余情况都为 0。因此, 我们有

$$c_k \tilde{u}_k = \sum_{|p|=0}^{\infty} \hat{u}_{k+2pN} = \hat{u}_k + \sum_{|p|>0}^{\infty} \hat{u}_{k+2pN},$$

得 (2.53)。因此, 直接计算得

$$\begin{aligned} \|P_N u - I_N u\|^2 &= \sum_{|k| \leq N} |\hat{u}_k - \tilde{u}_k|^2 \\ &= \sum_{|k| < N} |\hat{u}_k - \tilde{u}_k|^2 + \frac{1}{4} \sum_{k=\pm N} |2\hat{u}_k - 2\tilde{u}_k|^2 \\ &\leq \sum_{|k| < N} |\hat{u}_k - \tilde{u}_k|^2 + \frac{1}{2} \sum_{k=\pm N} |\hat{u}_k - 2\tilde{u}_k|^2 + \frac{1}{2} \sum_{k=\pm N} |\hat{u}_k|^2 \\ &\leq \sum_{|k| \leq N} |\hat{u}_k - c_k \tilde{u}_k|^2 + \frac{1}{2} \sum_{k=\pm N} |\hat{u}_k|^2. \end{aligned}$$

最后一项通过下式约束

$$|\hat{u}_N|^2 + |\hat{u}_{-N}|^2 \leq N^{-2m} \sum_{|k|=N}^{\infty} |k|^{2m} |\hat{u}_k|^2 \leq N^{-2m} |u|_m^2,$$

第一项可以通过关系式 (2.53) 和 Cauchy-Schwarz 不等式估得:

$$\begin{aligned} \sum_{|k| \leq N} |\hat{u}_k - c_k \tilde{u}_k|^2 &= \sum_{|k| \leq N} \left| \sum_{|p|>0}^{\infty} \hat{u}_{k+2pN} \right|^2 \\ &\leq \sum_{|k| \leq N} \left\{ \left(\sum_{|p|>0}^{\infty} |k+2pN|^{-2m} \right) \left(\sum_{|p|>0}^{\infty} |k+2pN|^{2m} |\hat{u}_{k+2pN}|^2 \right) \right\} \\ &\leq \max_{|k| \leq N} \left\{ \sum_{|p|>0}^{\infty} |k+2pN|^{-2m} \right\} \left(\sum_{|k| \leq N} \sum_{|p|>0}^{\infty} |k+2pN|^{2m} |\hat{u}_{k+2pN}|^2 \right). \end{aligned}$$

显然有

$$\max_{|k| \leq N} \left\{ \sum_{|p|>0}^{\infty} |k+2pN|^{-2m} \right\} \leq \frac{1}{N^{2m}} \sum_{|p|>0}^{\infty} \frac{1}{|2p-1|^{2m}} \lesssim N^{-2m},$$

和

$$\sum_{|k| \leq N} \sum_{|p| > 0}^{\infty} |k + 2pN|^{2m} |\hat{u}_{k+2pN}|^2 \leq 2|u|_m^2.$$

因此, 结合上述估计得

$$\|P_N u - I_N u\| \lesssim N^{-m} |u|_m.$$

此外, 由逆不等式 (2.44) 有

$$\|\partial_x^l (P_N u - I_N u)\| \lesssim N^l \|P_N u - I_N u\| \lesssim N^{l-m} |u|_m.$$

最后, 由三角不等式和定理 2.1 得

$$\|\partial_x^l (I_N u - u)\| \leq \|\partial_x^l (P_N u - I_N u)\| + \|\partial_x^l (P_N u - u)\| \lesssim N^{l-m} |u|_m.$$

至此证毕。□

上面我们展示了一些在 Sobolev 空间上的基本傅里叶近似。感兴趣的读者可以参考傅里叶分析 (例如 Körner(1988), Folland(1992)), 在不同内容中深入讨论傅里叶近似。

2.3 傅里叶谱方法的应用

本节, 我们将应用傅里叶谱方法到一些有周期性边界条件的非线性 PDEs 上, 包括 KdV 方程 (参见 Korteweg and de Vries(1895)), KS 方程 (参见 Kuramoto and Tsuzuki(1976)) 和 Allen-Cahn 方程 (参见 Allen and Cahn(1979))。重点将放在非线性部分的处理和时间离散化上。

2.3.1 KdV 方程

KdV 方程是浅水表面波的著名数学模型。KdV 方程有一个迷人的性质是它允许孤子型解 (参见 Zabusky and Galvin(1971))。在整个空间上考虑 KdV 方程:

$$\begin{aligned} \partial_t u + u \partial_y u + \partial_y^3 u &= 0, \quad y \in (-\infty, \infty), t > 0, \\ u(y, 0) &= u_0(y), \quad y \in (-\infty, \infty), \end{aligned} \quad (2.54)$$

它有准确解

$$u(y, t) = 12\kappa^2 \operatorname{sech}^2(\kappa(y - y_0) - 4\kappa^3 t), \quad (2.55)$$

其中 y_0 是初始值 $u(y, 0)$ 的中心, κ 是一个与行进速度相关的常数。

由于当 $|y| \rightarrow \infty$ 时 $u(y, t)$ 指数式衰减为 0, 我们可以将无限区间截断为一个有限区间 $(-\pi L, \pi L)$, $L > 0$, 同时通过将周期性边界条件限制在 $(-\pi L, \pi L)$ 上来近似边界条件。期望只要孤立子没有到达边界, 具有周期性边界条件的初边值问题 (2.54) 可以提供原始初值问题一个好的近似。

为了方便, 我们通过下述坐标变换将 $[-\pi L, \pi L]$ 映射成 $[0, 2\pi]$:

$$x = \frac{y}{L} + \pi, \quad y = L(x - \pi), \quad x \in [0, 2\pi], y \in [-\pi L, \pi L],$$

同时取

$$v(x, t) = u(y, t), \quad v_0(x) = u_0(y). \quad (2.56)$$

变换后的 KdV 方程为

$$\begin{aligned} \partial_t v + \frac{1}{L} v \partial_x v + \frac{1}{L^3} \partial_x^3 v &= 0, \quad x \in (0, 2\pi), t > 0, \\ v(\cdot, t) &\text{在 } [0, 2\pi] \text{ 上周期}, \quad t \geq 0; \quad v(x, 0) = v_0(x), \quad x \in [0, 2\pi]. \end{aligned} \quad (2.57)$$

将 v 写为 $v(x, t) = \sum_{|k|=0}^{\infty} \hat{v}_k(t) e^{ikx}$, 作第一个方程与 e^{ikx} 的内积, 并且由 (2.3) 和 $v \partial_x v = \frac{1}{2} \partial_x (v^2)$ 得

$$\frac{d\hat{v}_k}{dt} - \frac{ik^3}{L^3} \hat{v}_k + \frac{ik}{2L} \widehat{(v^2)}_k = 0, \quad k = 0, \pm 1, \dots, \quad (2.58)$$

其中初始条件

$$\hat{v}_k(0) = \frac{1}{2\pi} \int_0^{2\pi} v_0(x) e^{-ikx} dx. \quad (2.59)$$

常微分方程组 (2.58) - (2.59) 可以通过不同的数值方法求解, 例如 Runge-Kutt 法或非线性部分被显式对待而主导的线性部分被隐式对待的半隐式/线性隐式法。

这里, 我们如 Trefethen(2000) 建议的那样, 使用积分因子和 Runge-Kutta 法的组合。更准确的说, 将 (2.58) 乘以积分因子 $e^{-ik^3 t/L^3}$, 重写乘后的方程为

$$\frac{d}{dt} \left[e^{-ik^3 t/L^3} \hat{v}_k \right] = -\frac{ik}{2L} e^{-ik^3 t/L^3} (\widehat{v^2})_k, \quad k = 0, \pm 1, \dots \quad (2.60)$$

这样的操作使得线性项消失且减小了方程组的刚度。方程组 (2.60) 可以由标准 ODE 解法求解。

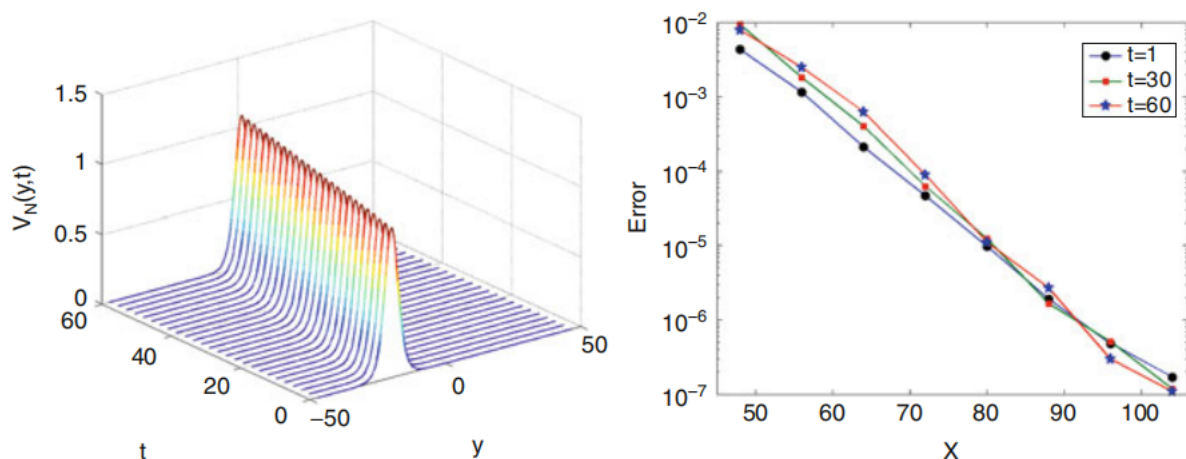
现在我们描述 (2.57) 在 Matlab 中的傅里叶近似。令 \mathbf{k} 为 (2.40) 中的向量, 且有

$$\tilde{\mathbf{v}} = (\tilde{v}_0, \dots, \tilde{v}_{N/2}, \tilde{v}_{-N/2+1}, \dots, \tilde{v}_{-1}), \quad \mathbf{g} = e^{-ik^3 t/L^3}, \quad \tilde{\mathbf{u}} = \mathbf{g} * \tilde{\mathbf{v}},$$

其中对向量的运算是按分量进行的。然后基于 (2.60) 的傅里叶近似法如下:

$$\frac{d\tilde{\mathbf{u}}}{dt} = -\frac{i\mathbf{k}}{2L} * \mathbf{g} * \text{fft} \left(\left[\text{ifft}(\mathbf{g}^{-1} * \tilde{\mathbf{u}}) \right]^2 \right), \quad t > 0. \quad (2.61)$$

因此, Runge-Kutta 法, 例如四阶 Matlab 例程 *rk4.m* 可以直接应用于 (2.61)。我们接下来给出由 Trefethen(2000) 中程序 27 (稍作修改) 得到的一些数值结果。首先取 $\kappa = 0.3, y_0 = -20$ 和 $L = 15$ 。在图 2.2 的左侧, 我们绘出近似解的时间演化 (其中 $N = 256$, 时间步长 $\tau = 0.01, t \in [0, 60]$) ; 在图的右侧, 我们绘出 $t = 1, 30, 60$ 时不同 N 对应的最大误差, 其中 $\tau = 0.001$ 。观察到误差消减速度似 $O(e^{-cN})$, 这是典型的光滑解。该孤子解较高的精度说明在孤立波到达边界之前, 可以用有限区间上的 KdV 方程有效的模拟整个空间的 KdV 方程。

图 2.2: 左: 数值解的时间演化; 右: 不同 N 对应的最大误差

下一个例子, 我们考虑五个孤立波的相互作用。更准确地说, 我们考虑包含五个孤立波的初始条件的 KdV 方程 (2.54),

$$u_0(y) = \sum_{j=1}^5 12\kappa_j^2 \operatorname{sech}^2(\kappa_j(y - y_0)), \quad (2.62)$$

其中

$$\begin{aligned} \kappa_1 &= 0.3, \kappa_2 = 0.25, \kappa_3 = 0.2, \kappa_4 = 0.15, \kappa_5 = 0.1, \\ y_1 &= -120, y_2 = -90, y_3 = -60, y_4 = -30, y_5 = 0. \end{aligned} \quad (2.63)$$

我们修改 $L = 50, N = 512, \tau = 0.01$, 在图 2.3 中, 我们绘出 $t \in [0, 600]$ 时近似解的时间演化, 并在图 2.4 中绘出初始剖面 and $t = 600$ 时的最终剖面。观察到振幅大的孤立波传播速度更快, 在最后时刻五个孤立波的振幅被很好地保留下来。这意味着该方法有优异的保护特性。

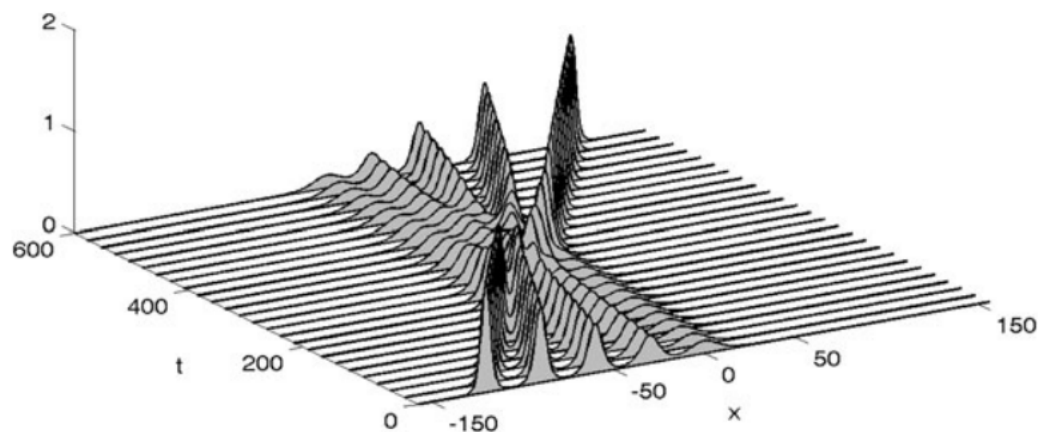
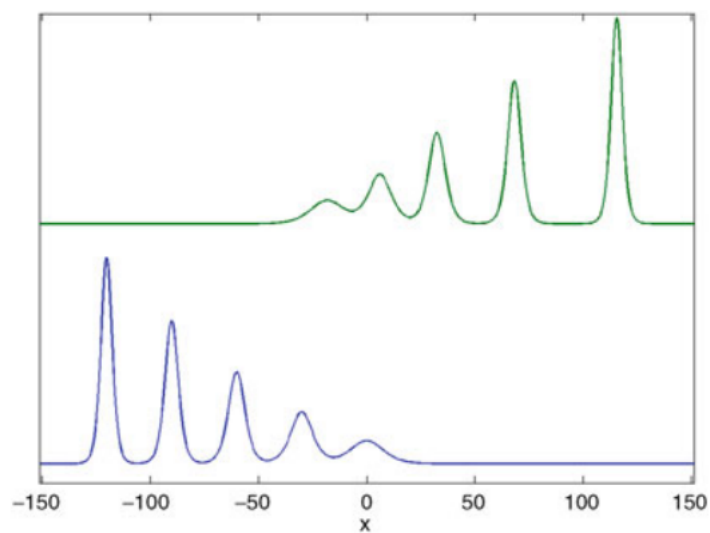


图 2.3: 五个孤立波的相互作用

图 2.4: $t = 0, 600$ 时的剖面

2.3.2 KS 方程

KS 方程已用于各种反应扩散系统的研究(参见 Kuramoto and Tsuzuki(1976)), 同时也是一个可以显示混沌解的有趣的动态 PDE (参见 Hyman and Nicolaenko(1986), Nicolaenko 等人 (1985))。

我们考虑下述形式的 KS 方程

$$\begin{aligned} \partial_t u + \partial_x^4 u + \partial_x^2 u + uu_x &= 0, \quad x \in (-\infty, \infty), t > 0, \\ u(x, t) &= u(x + 2L\pi, t), \quad \partial_x u(x, t) = \partial_x u(x + 2L\pi, t), \quad t \geq 0, \\ u(x, 0) &= u_0(x), \quad x \in (-\infty, \infty), \end{aligned} \quad (2.64)$$

其中已给函数 u_0 是 $2L\pi$ 周期的。

由于周期性, 考虑所涉及区间 $[0, 2L\pi]$ 即可。先通过寻找近似解将 (2.64) 在空间上离散化

$$u_N(x, t) = \sum_{k=-N/2}^{N/2} \tilde{u}_k(t) e^{ikx/L}, \quad t > 0, \quad (2.65)$$

其中 $\tilde{u}_{N/2}(t) = \tilde{u}_{-N/2}(t)$ 。 N 与频率无关由下述方案决定

$$\partial_t u_N + \partial_x^4 u_N + \partial_x^2 u_N = -\frac{1}{2} \partial_x I_N(u_N^2), \quad t > 0, \quad (2.66)$$

其中 I_N 是网格点为 $\{x_j = 2L\pi j/N\}_{j=0}^{N-1}$ 的插值算子。

因此, 对每次频率 k 有

$$\tilde{u}'_k(t) + \left(\frac{k^4}{L^4} - \frac{k^2}{L^2} \right) \tilde{u}_k(t) = -\frac{1}{2L} ik \tilde{w}_k(t), \quad t > 0, \quad (2.67)$$

其中 $\tilde{w}_k(t)$ 是非线性项的离散傅里叶系数, 即

$$I_N(u_N^2) = \sum_{k=-N/2}^{N/2} \tilde{w}_k(t) e^{ikx/L}. \quad (2.68)$$

必须指出的是, 使用傅里叶变换、常系数线性算子总是可以像 (2.67) 一样对角化 (即频率是可分离的)。这导致有充分的时间在频率空间上对所得方程积分。我们参考 Kassam and Trefethen(2005) 来回顾各种时间步进方案。这里, 我们使用 Kassam and Trefethen(2005) 提出的指数时间差分法 (ETD)。

用 $\tilde{\mathbf{u}}(t)$ 表示如表 2.1 所示的扩展系数向量。令 \mathbf{L} 为对角线为 $\mathbf{k}^2/L^2 - \mathbf{k}^4/L^4$ 的对角矩阵, 其中 \mathbf{k} 是根据表 (2.1) 给出的指标向量, $\mathbf{N}(t) := \mathbf{N}(\tilde{\mathbf{u}}, t)$ 为 (2.67) - (2.68) 中非线性项构成的向量。然后, 我们重写 (2.67) - (2.68) 为一个非线性 ODE 方程组

$$\tilde{\mathbf{u}}'(t) = \mathbf{L}\tilde{\mathbf{u}}(t) + \mathbf{N}(t), \quad t > 0. \quad (2.69)$$

令 τ 为时间步长。显然 (2.69) 等价于

$$\tilde{\mathbf{u}}(t_n + \tau) = e^{\mathbf{L}\tau}\tilde{\mathbf{u}}(t_n) + e^{\mathbf{L}\tau} \int_0^\tau e^{-\mathbf{L}s} \mathbf{N}(\tilde{\mathbf{u}}(t_n + s), t_n + s) ds. \quad (2.70)$$

基于如何近似积分, 可以构造不同的 ETD 方法。例如, 令 $\tilde{\mathbf{u}}_n$ 为 $\tilde{\mathbf{u}}(t_n)$ 的近似。以下改进的四阶 ETD Runge-Kutta 法 (ETDRK4) 已被证明是一个对解刚性方程十分稳定且准确的方法 (参见 Kassam and Trefethen(2005)):

$$\begin{aligned} \mathbf{a}_n &= e^{\mathbf{L}\tau/2}\tilde{\mathbf{u}}_n + \mathbf{L}^{-1} (e^{\mathbf{L}\tau/2} - \mathbf{I}) \mathbf{N}(\tilde{\mathbf{u}}_n, t_n), \\ \mathbf{b}_n &= e^{\mathbf{L}\tau/2}\tilde{\mathbf{u}}_n + \mathbf{L}^{-1} (e^{\mathbf{L}\tau/2} - \mathbf{I}) \mathbf{N}(\mathbf{a}_n, t_n + \tau/2), \\ \mathbf{c}_n &= e^{\mathbf{L}\tau/2}\mathbf{a}_n + \mathbf{L}^{-1} (e^{\mathbf{L}\tau/2} - \mathbf{I}) [2\mathbf{N}(\mathbf{b}_n, t_n + \tau/2) - \mathbf{N}(\tilde{\mathbf{u}}_n, t_n)], \\ \tilde{\mathbf{u}}_{n+1} &= e^{\mathbf{L}\tau}\tilde{\mathbf{u}}_n + \{ \alpha \mathbf{N}(\tilde{\mathbf{u}}_n, t_n) + 2\beta [\mathbf{N}(\mathbf{a}_n, t_n + \tau/2) \\ &\quad + \mathbf{N}(\mathbf{b}_n, t_n + \tau/2)] + \gamma \mathbf{N}(\mathbf{c}_n, t_n + \tau) \}, \end{aligned} \quad (2.71)$$

其 t 中系数

$$\begin{aligned} \alpha &= \tau^{-2} \mathbf{L}^{-3} [-4 - \mathbf{L}\tau + e^{\mathbf{L}\tau} (4 - 3\mathbf{L}\tau + (\mathbf{L}\tau)^2)], \\ \beta &= \tau^{-2} \mathbf{L}^{-3} [2 + \mathbf{L}\tau + e^{\mathbf{L}\tau} (-2 + \mathbf{L}\tau)], \\ \gamma &= \tau^{-2} \mathbf{L}^{-3} [-4 - 3\mathbf{L}\tau + e^{\mathbf{L}\tau} (4 - \mathbf{L}\tau) - (\mathbf{L}\tau)^2]. \end{aligned} \quad (2.72)$$

下面的计算, 我们取 $L = 16$, 并且如 Kassam and Trefethen(2005) 一样加上初始条件:

$$u_0(x) = \cos(x/L)(1 + \sin(x/L))$$

在图 2.5 中, 我们通过上述算法取 $\tau = 10^{-4}$, $N = 128$ 绘出 KS 方程 (2.64) 的时间演化。我们在图 2.6 中以瀑布格式绘出了数值解在不同时刻的剖面图。可以观察到如 Kassam and Trefethen(2005) 中所示一样的确定性混沌图案。

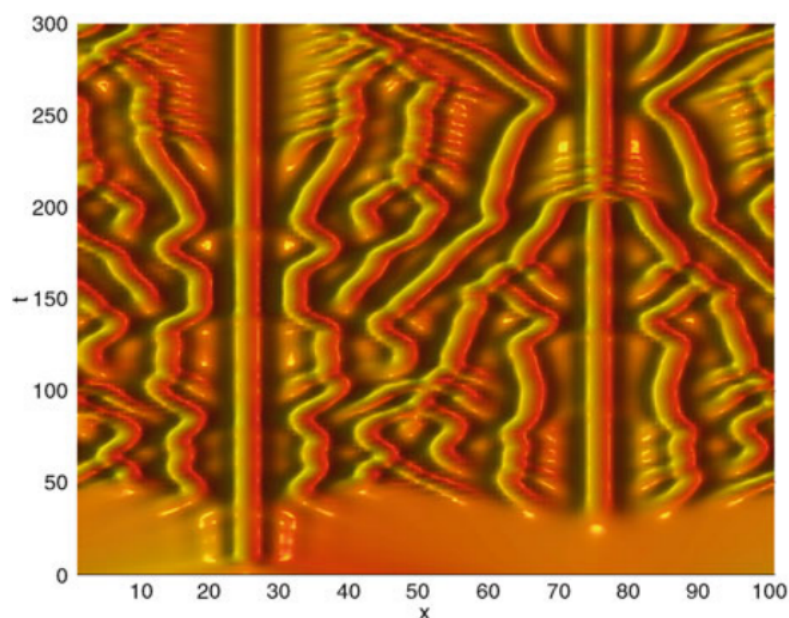


图 2.5: KS 方程的时间演化. 时间由底部 0 跑至顶部 300

2.3.3 Allen-Cahn 方程

Allen-Cahn 方程最初是在 Allen and Cahn(1979) 中用来描述晶体固体中反相界面的运动而引入的。它已广泛应用于材料科学领域。

我们考虑具有周期性边界条件的二维 Allen-Cahn 方程:

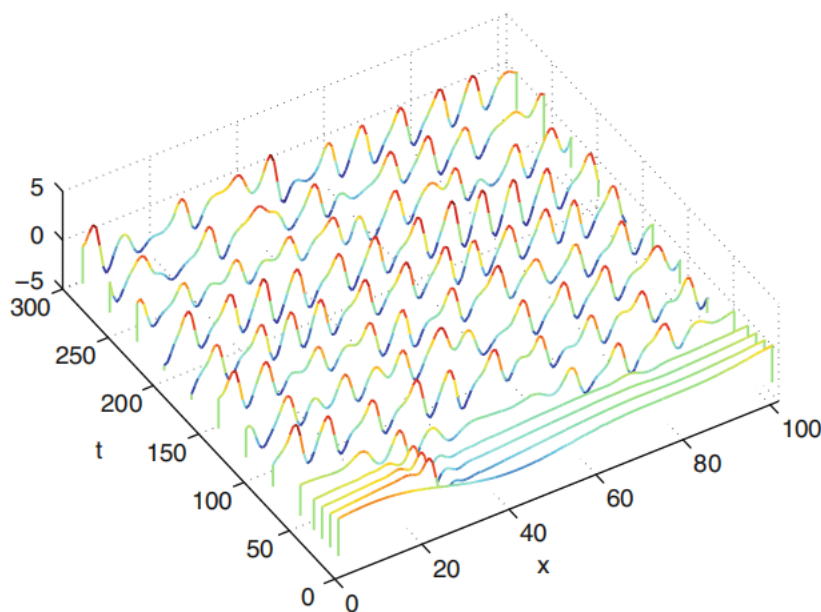


图 2.6: 瀑布式数值解剖面图

$$\begin{aligned}
 &\partial_t u - \varepsilon^2 \Delta u + u^3 - u = 0, \quad (x, y) \in \Omega = (-1, 1)^2, t > 0, \\
 &u(-1, y, t) = u(1, y, t), \quad u(x, -1, t) = u(x, 1, t), \quad t \geq 0, \\
 &u(x, y, 0) = u_0(x, y), \quad (x, y) \in \bar{\Omega},
 \end{aligned} \tag{2.73}$$

其中 ε 是一个描述界面宽度的小参数。对 Allen-Cahn 方程的深入讨论参考 9.3 节。

现在我们写出如下傅里叶近似解

$$u_N(x, y, t) = \sum_{k, l=-N/2}^{N/2-1} \tilde{u}_{kl}(t) e^{i(kx+ly)\pi}, \tag{2.74}$$

并用 u_N^m 表示 u_N 在 $t_m = m\tau$ 时刻的近似值，其中 τ 为时间步长。然后一个时间上的二阶稳定半隐式法为（参见 Shen and Yang(2010)）：

$$\begin{aligned} & \frac{3u_N^{m+1} - 4u_N^m + u_N^{m-1}}{2\tau} - \varepsilon^2 \Delta u_N^{m+1} + (2F_N(u_N^m) - F_N(u_N^{m-1})) \\ & + s(u_N^{m+1} - 2u_N^m + u_N^{m-1}) = 0, \quad m = 1, 2, \dots, \end{aligned} \quad (2.75)$$

其中 $s > 0$ 是一个可调参数, I_N 为计算网格上的二维张量插值算子, $F_N(v) = I_N(v^3) - v$ 。注意到增加额外的耗散项 $s(u_N^{m+1} - 2u_N^m + u_N^{m-1})$ (阶数为 $s\tau^2$) 可以在不改变简单性的同时增加稳定性。在每次时间步进上, 我们只需要解线性问题

$$-2\tau\varepsilon^2 \Delta u_N^{m+1} + (3 + 2s\tau)u_N^{m+1} = \mathbf{N}(u_N^m, u_N^{m-1}), \quad (2.76)$$

其中

$$\begin{aligned} \mathbf{N}(u_N^m, u_N^{m-1}) &= 4(1 + \tau s + \tau)u_N^m - (1 + 2\tau s + 2\tau)u_N^{m-1} \\ &\quad - 2\tau I_N [2(u_N^m)^3 - (u_N^{m-1})^3]. \end{aligned} \quad (2.77)$$

运用 Fourier Galerkin 法得到频率空间上的方程：

$$(2\tau\varepsilon^2(k^2 + l^2)\pi^2 + 3 + 2s\tau)\tilde{u}_{kl}^{m+1} = \tilde{w}_{kl}, \quad (2.78)$$

其中 $\{\tilde{w}_{kl}\}$ 为非线性部分 $\mathbf{N}(u_N^m, u_N^{m-1})$ 的离散傅里叶系数。由于主要成本来自非线性项的处理是可以通过伪谱方法借助 FFT 进行操作, 因此上述半隐式方案可以有效实现。

为了测试数值方法, 我们考虑圆界面的运动并且加上初始条件：

$$u_0(x, y) = \begin{cases} 1, & (x - 0.5)^2 + (y - 0.5)^2 < 1, \\ -1, & \text{otherwise.} \end{cases} \quad (2.79)$$

界面的运动是由圆的平均曲率驱动的, 所以圆会收缩并最终消失, 见图 2.7。注意到圆直径缩小的速率可以通过分析来确定 (参见 Chen and Shen(1998))。

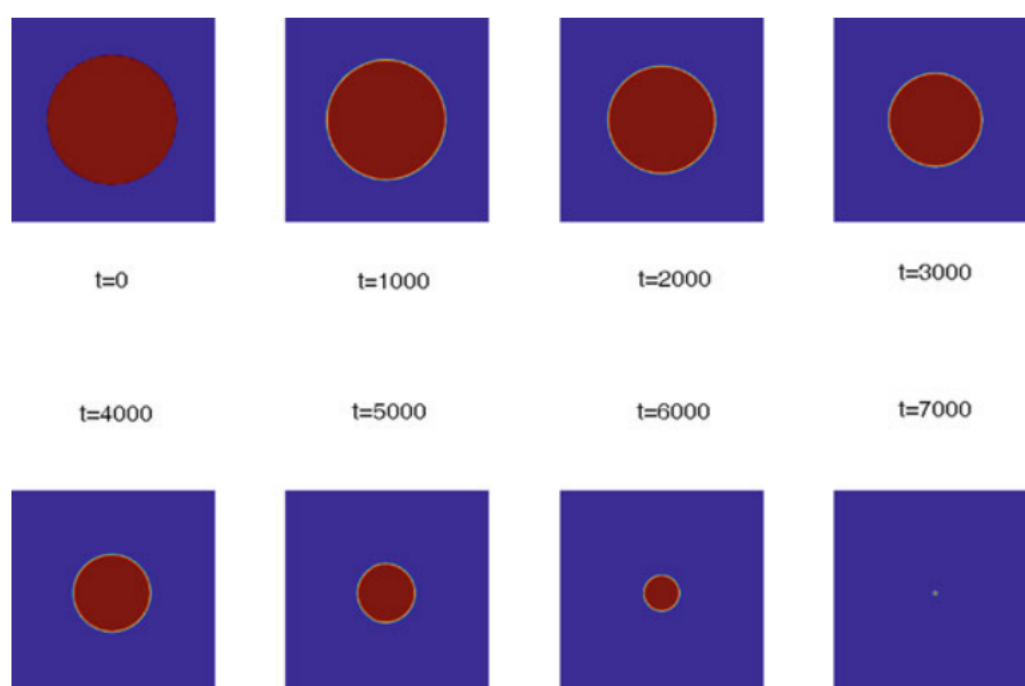


图 2.7: 圆域的时间演化