Numerical Solution of Partial Differential Equations

Chenghao Dong

December 2, 2023

CONTENTS

1 Elements of Function Spaces 2
- 1.1 Notations for Multivariable Derivatives . . . 2
- 1.2 Spaces of Continuous Functions . . . 3
- 1.3 Spaces of Integrable Functions . . . 5
- 1.4 Sobolev Spaces . . . 6

2 Elliptic Boundary Value Problems 9
- 2.1 Formulation . . . 9
- 2.2 Weak Solutions . . . 10

3 Finite Difference Schemes for ODEs 11
- 3.1 Finite Differences . . . 11
- 3.2 Existence and Uniqueness . . . 13
- 3.3 Stability and Consistency . . . 15

4 Finite Difference Schemes for Elliptic PDEs 17
- 4.1 Introduction and Terminologies . . . 17
- 4.2 Scheme Formulation . . . 19
- 4.3 Solution Behaviors for Continuous Force Functions . . . 20
- 4.4 Solution Behaviors for Discontinuous Force Functions . . . 23
- 4.5 Discretization for Generalized Elliptic BVPs and Nonaxiparallel Domains with Nonuniform Meshes . . . 29
- 4.6 Maximum Principle . . . 31
- 4.7 Iterative Method for Linear Systems . . . 36

5 Finite Difference Schemes for Parabolic PDEs 40
- 5.1 Introduction . . . 40
- 5.2 θ-methods . . . 43
- 5.3 Stability of $\theta-$methods . . . 46
- 5.4 Boundary Value Problems of Parabolic Problems . . . 53
- 5.5 Maximum Principle and Convergence of $\theta-$methods . . . 54
- 5.6 Parabolic Equations in Two Space Dimensions . . . 57

6 Finite Difference Schemes for Hyperbolic PDEs 60
- 6.1 Introduction . . . 60
- 6.2 Implicit and Explicit Schemes . . . 63
- 6.3 Implicit Scheme : Stability, Consistency and Convergence . . . 65
- 6.4 Explicit Scheme : Stability, Consistency and Convergence . . . 70

7 Appendix 72
- 7.1 Definition and Theorem . . . 72

1 Elements of Function Spaces

1.1 Notations for Multivariable Derivatives

Definition 1.1.1. 偏导数的多重下标 (multi-index)
记 $\boldsymbol{\alpha} = (a_1, \cdots, a_n) \in {\mathbb{N}_0}^n$ 称为偏导数的多重下标；定义该下标的**长度 (*length*)** 为 $|\boldsymbol{\alpha}| = a_1 + \cdots + a_n$，阶乘为 $\boldsymbol{\alpha}! = a_1! \cdots a_2!$；特别地，记 $\mathbf{0} = (0, \cdots, 0)$。于是有以下算符：

$$D^{\boldsymbol{\alpha}} := \partial^{\boldsymbol{\alpha}} := \partial {x_1}^{a_1} \cdots \partial {x_n}^{a_n} := \frac{\partial^{|\boldsymbol{\alpha}|}}{\partial {x_1}^{a_1} \cdots \partial {x_n}^{a_n}}$$

注意有以下几种常见的表达：$\sum_{|\boldsymbol{\alpha}|=k} \boldsymbol{D^{\alpha} u}$ 与 $\sum_{|\boldsymbol{\alpha}|\leq k} \boldsymbol{D^{\alpha} u}$ 分别表示函数 u 的所有 k 阶偏导项，与阶数小于 k 的偏导项；$\sum_{|\boldsymbol{\alpha}|=k} \boldsymbol{D^{\alpha} u}$ 有不超过 $\binom{n+k-1}{k}$ 项，即为 $|\boldsymbol{\alpha}| = k$ 的非负整数解个数；$\sum_{|\boldsymbol{\alpha}|\leq k} \boldsymbol{D^{\alpha} u}$ 有不超过 $\binom{n+k}{k}$ 项。

Theorem 1.1.1. 多项式定理 (multinomial theorem) 对于多元函数自变量 $\boldsymbol{x} = (x_1, \cdots, x_n) \in \mathbb{R}^n$，与多重下标 $\boldsymbol{\alpha} \in \mathbb{N}^n$ 有：

$$(x_1 + \cdots + x_n)^k = \sum_{|\boldsymbol{\alpha}|=k} \frac{k!}{\boldsymbol{\alpha}!} \cdot \boldsymbol{x}^{\boldsymbol{\alpha}} \quad \text{s.t.} \quad \boldsymbol{x}^{\boldsymbol{\alpha}} = x_1^{a_1} \cdots x_n^{a_n}$$

Theorem 1.1.2. 多元函数的泰勒展开 (multivariable taylor's theorem)
记函数 $f : \Omega \longrightarrow \mathbb{R}$ 为开集 $\Omega \subset \mathbb{R}^n$ 上的 $(k+1)$ 阶连续可微函数，即 $f \in C^{k+1}(\Omega)$，则对 $\boldsymbol{x} \in \Omega$ 与充分小的 $\Delta\boldsymbol{x}$ s.t. $\boldsymbol{x} + \Delta\boldsymbol{x} \in \Omega$，有：

$$f(\boldsymbol{x} + \Delta\boldsymbol{x}) = \sum_{|\boldsymbol{\alpha}|\leq k} \frac{D^{\boldsymbol{\alpha}} f(\boldsymbol{x})}{\boldsymbol{\alpha}!} \Delta\boldsymbol{x}^{\boldsymbol{\alpha}} + R_{x,k}(\Delta\boldsymbol{x})$$

其中拉格朗日余项为：

$$R_{x,k}(\Delta\boldsymbol{x}) = \sum_{|\boldsymbol{\alpha}|=k+1} \frac{D^{\boldsymbol{\alpha}} f(x + c)}{\boldsymbol{\alpha}!} \Delta\boldsymbol{x}^{\boldsymbol{\alpha}} \quad \text{for some } c \in (0, 1)$$

特别地，如果满足 $\Delta\boldsymbol{x} = (\Delta x_1, \cdots, \Delta x_n) \quad \text{s.t.} \quad \Delta x_1 = \cdots = \Delta x_n = \Delta x$，则 $R_{x,k}(\Delta\boldsymbol{x}) = O(\Delta x^{k+1})$

Theorem 1.1.3. 多元函数的散度定理 (divergence theorem)
记 Ω 为 $\mathbb{R}^n$ 中的由 $\partial\Omega$ 围成的有界连通区域，$\boldsymbol{n}$ 为曲面 $\partial\Omega$ 上的单位法向量 (诱导定向)。若矢量函数 $\boldsymbol{v}$ 在闭区域 $\bar{\Omega} = \Omega \cup \partial\Omega$ 上连续，在 Ω 内有一阶连续偏导数，则：

$$\int_{\Omega} \nabla \cdot \boldsymbol{v} d\boldsymbol{x} = \oint_{\partial\Omega} \boldsymbol{v} \cdot \boldsymbol{n} d\boldsymbol{S}$$

1.2 Spaces of Continuous Functions

Definition 1.2.1. 闭包（closure） 一个集合 Ω 的闭包是包含其自身的最小闭集合，即包含其自身与该集合所有的聚点（*accumulation point*），记作 $\bar{\Omega}$。

Definition 1.2.2. 完备（complete） 一个集合 Ω，若其在某度量下，其中的序列柯西收敛等价于一般收敛，则称该集合为完备集合。完备集合必然为闭集合；完备集合的子集合也完备当且仅当其为闭集合。
Remark $\left(\mathbb{R}^n, d_p\right)$ 均是完备的，包括对 d_∞。

Definition 1.2.3. 紧致（compact） 在度量空间的语境下，若一个集合 Ω 的任意序列在良定义度量下均存在收敛子列，且子列极限 $l \in \Omega$，则称该集合紧致。紧致集合必然完备；紧致集合的子集合也紧致当且仅当其为闭集合。
Remark 欧式空间语境下 $\mathbb{R}^n$ 的任何有界闭集合均是紧致的。

Definition 1.2.4. 连续函数集合 记 $\Omega \subset \mathbb{R}^n$ 为开集合，则集合 $C^k(\Omega)$ 定义为在 Ω 上的所有 k 阶连续可微函数，其中 $k \in \mathbb{N}_0$；进一步，若 Ω 有界（*bounded*），则 $C^k(\bar{\Omega})$ 表示在闭集合 $\bar{\Omega}$（Ω 的闭包）上的所有 k 阶连续可微函数。

Theorem 1.2.1. 连续函数空间 记 $\Omega \subset \mathbb{R}^n$ 为有界开集合，则 $C^k(\bar{\Omega})$ 为一个线性空间，且其上有范数（*norm*）定义为：

$$||u||_{C^k(\bar{\Omega})} = \sum_{|\boldsymbol{\alpha}| \leq k} \sup_{\boldsymbol{x} \in \Omega} |D^{\boldsymbol{\alpha}} u(\boldsymbol{x})|$$

这表示函数 $u(\boldsymbol{x})$ 的所有 k 阶及以下偏导数绝对值上确界之和。特别地，当 $k = 0$ 时：

$$||u||_{C(\bar{\Omega})} = \max_{\boldsymbol{x} \in \bar{\Omega}} |u(\boldsymbol{x})|$$

注意范数意味着其满足三角不等式（*triangle inequality*），$\forall\, k \in \mathbb{R} : ||ku|| = k||u||$ 与 $||u||_{C^k(\bar{\Omega})} = 0$ 当且仅当 $u(\boldsymbol{x}) \equiv 0$。

Definition 1.2.5. 函数的支集（support） 记连续函数 $u \in C(\Omega)$，其在 Ω 上的支集为：

$$\text{supp } u = \text{closure of } \{\boldsymbol{x} : u(\boldsymbol{x}) \neq 0\}$$

也即 Ω 上包含所有 u 的非零取值点的最小闭集；不难证明它也是 Ω 上最小的补集均为 u 零点的闭集。特别定义：

$$C_0^k(\Omega) = \left\{u \in C^k(\Omega) : (\text{supp } u \subset \Omega) \wedge (\text{supp } u \text{ bounded})\right\}$$

$$C_0^\infty(\Omega) = \bigcap_{k \geq 0} C_0^k(\Omega)$$

后者仅表示满足 $(\text{supp } u \subset \Omega) \wedge (\text{supp } u \text{ bounded})$ 的无穷可微函数。显然 $\forall\, l < k \ :\ C_0^k(\Omega) \subset C_0^l(\Omega)$ 均成立。

此外还定义：$C_c^k(\Omega) = \left\{u \in C^k(\Omega) : (\text{supp } u \subset \Omega) \wedge (\text{supp } u \text{ compact})\right\}$

$$C_c^\infty(\Omega) = \bigcap_{k \geq 0} C_c^k(\Omega)$$

若 $u \in C_c^k(\Omega)$，则称函数 u 在 Ω 上紧支撑（*compactly supported*）。

Theorem 1.2.2. $C_0^k(\Omega)$的常见性质
当 $\Omega \subset \mathbb{R}^n$ 时，有以下常见性质

1) $(\text{supp } u \text{ bounded}) \Longleftrightarrow (\text{supp } u \text{ compact})$，即 $C_0^k(\Omega)$ 与 $C_c^k(\Omega)$ 等价；
2) 若 Ω 为开集，则 $u \in C_0^k(\Omega) \Rightarrow u(\boldsymbol{x}) \equiv 0, \forall \boldsymbol{x} \in \partial\Omega$，反之不成立；
3) 若 Ω 为开集，则 $u \in C_0^k(\Omega) \Rightarrow \partial_{x_i} u \in C_0^{k-1}(\Omega), \forall x_i$。

Proof theorem 1.2.2，2) 假设 $\exists\, \boldsymbol{x}_0 \in \partial\Omega$ s.t. $u(\boldsymbol{x}_0) \neq 0$，由连续性，存在一个 $\boldsymbol{x}_0$ 的邻域 $B(\boldsymbol{x}_0, \epsilon_0)$ s.t. $u(\boldsymbol{x}) \neq 0, \forall \boldsymbol{x} \in B(\boldsymbol{x}_0, \epsilon_0)$。又由于 $\boldsymbol{x}_0$ 为 Ω 的边界点，由边界点的定义：$\forall\, 0 < \epsilon_1 < \epsilon_0$，总存在 $\boldsymbol{x}(\epsilon_1) \in B(\boldsymbol{x}_0, \epsilon_1) \cap \Omega$，且 $u\left(\boldsymbol{x}(\epsilon_1)\right) \neq 0$, 因此 $\boldsymbol{x}(\epsilon_1) \in \text{supp } u$。进一步取 $0 < \epsilon_2 < ||\boldsymbol{x}_0 - \boldsymbol{x}(\epsilon_1)||_n$ 并重复上述操作，可以得到 $\boldsymbol{x}(\epsilon_2) \in \text{supp } u$ 且 $\boldsymbol{x}(\epsilon_2) \neq \boldsymbol{x}(\epsilon_1)$。

重复上述推导，得到 $\{\boldsymbol{x}(\epsilon_n)\}_{n=1}^\infty$ s.t. $\boldsymbol{x}(\epsilon_n) \to \boldsymbol{x}_0$ in norm $||\cdot||_n$。即 $\boldsymbol{x}_0$ 为 $\{\boldsymbol{x} : u(\boldsymbol{x}) \neq 0\}$ 的聚点，从而 $\boldsymbol{x}_0 \in \text{supp } u \text{ (closed)} \subset \Omega$。其中 $\text{supp } u \subset \Omega$ 由 $u \in C_0^k(\Omega)$ 定义自然可得。而 Ω 为开集，必然不包含边界点 $\boldsymbol{x}_0$，即 $\boldsymbol{x}_0 \notin \Omega$，矛盾。

Proof theorem 1.2.2，3) 假设 $u \in C_0^k(\Omega)$ 则将其定义域拓展至全集 $\mathbb{R}^n$，有：$\forall \boldsymbol{x} \in \mathbb{R}^n \setminus (\text{supp } u)\ :\ u(\boldsymbol{x}) = 0$，由于 $\mathbb{R}^n \setminus (\text{supp } u)$ 为开集合，对 $\forall \boldsymbol{x} \in \mathbb{R}^n \setminus (\text{supp } u)$，均存在特定的邻域 $B\left(\boldsymbol{x}, \epsilon(\boldsymbol{x})\right)$ s.t. $u(\boldsymbol{y}) \equiv 0$，$\forall \boldsymbol{y} \in B\left(\boldsymbol{x}, \epsilon(\boldsymbol{x})\right)$，即 u 在该领域上是常数函数。所以：$\partial_{x_i} u(\boldsymbol{y}) \equiv 0, \forall \boldsymbol{y} \in B\left(\boldsymbol{x}, \epsilon(\boldsymbol{x})\right)$ 对任何 x_i 成立。显然 $\boldsymbol{x}$ 不能是集合 $\{\boldsymbol{x}' : \partial_{x_i} u\left(\boldsymbol{x}'\right) \neq 0\}$ 的内点或边界点，因为存在 $B\left(\boldsymbol{x}, \epsilon(\boldsymbol{x})\right)$ 与之无交集。所以 $\boldsymbol{x} \notin \text{supp } \partial_{x_i} u$；$\boldsymbol{x} \in \mathbb{R}^n \setminus \left(\text{supp } \partial_{x_i} u\right)$。

综上，$\mathbb{R}^n \setminus (\text{supp } u) \subset \mathbb{R}^n \setminus \left(\text{supp } \partial_{x_i} u\right) \Longrightarrow \text{supp } \partial_{x_i} u \subset \text{supp } u \subset \Omega$，并且显然有界。

1.3 Spaces of Integrable Functions

Definition 1.3.1. 勒贝格零测集 (Lebesgue null set) 记 m_n 为 $\mathbb{R}^n$ 上的勒贝格测度，则 $N \subset \mathbb{R}^n$ 为（勒贝格）零测集当且仅当 $\forall \epsilon > 0 : \exists$ 可数个开球 $\{B_k\}_{k=1}^{\infty}$ s.t.

$$\left(N \subset \bigcup_{k=1}^{\infty} B_k \right) \wedge \left(\sum_{k=1}^{\infty} m_n \left(B_k \right) < \epsilon \right)$$

两函数（对依勒贝格测度）在 ω 上几乎处处相等 $u_1 \overset{a.e.}{=\!=} u_2$ 当且仅当 $\{\boldsymbol{x} \in \Omega : u_1(\boldsymbol{x}) \neq u_2(\boldsymbol{x})\}$ 为零测集。

Definition 1.3.2. 希尔伯特空间 (Hilbert Space) 记线性空间 $\left(V \ , \ (\cdot,\cdot)_X \right)$ 为一个内积空间，即有良定内积 $(\cdot,\cdot)_X$ 与范数 $||\boldsymbol{v}||_X = \sqrt{(\boldsymbol{v},\boldsymbol{v})_X}$；若其对于度量 $d_X(\boldsymbol{u},\boldsymbol{v}) = ||\boldsymbol{u}-\boldsymbol{v}||_X$ 完备，则称之为一个希尔伯特空间。换句话说，希尔伯特空间就是有完备（范数）度量的内积空间。

Theorem 1.3.1. 可积函数空间 对 $\forall\ p \in \mathbb{R}$ s.t. $p \geq 1$，与开集 $\Omega \subset \mathbb{R}^n$，定义 $\mathcal{L}_p(\Omega) = \{u(\boldsymbol{x}) : \int_\Omega |u|^p d\boldsymbol{x} < \infty\}$ 为其上所有的勒贝格可积函数（*Lebesgue integrable*）。记 $u_1 \overset{a.e.}{=\!=} u_2$ 表示两函数（对勒贝格测度）几乎处处相等，则商集合：$L_p(\Omega) = \left(\mathcal{L}_p(\Omega) \,/\, \overset{a.e.}{=\!=} \right)$ 为一个线性空间，且其上有范数（*norm*）定义为：

$$||u||_{L_p(\Omega)} = \left(\int_\Omega |u|^p d\boldsymbol{x} \right)^{\frac{1}{p}}$$

注意范数意味着其满足三角不等式(*triangle inequality*),$\forall\ k \in \mathbb{R} : ||ku|| = k||u||$ 与 $||u||_{L_p(\Omega)} = 0$ 当且仅当 $u(\boldsymbol{x}) \overset{a.e.}{=\!=} 0$。特别地，当 $p = 2$ 时，有定义在 $L_2(\Omega)$ 上的内积：

$$(u,v)_{L_2(\Omega)} = \int_\Omega u(\boldsymbol{x})\, v(\boldsymbol{x})\, d\boldsymbol{x}$$

如果 $\boldsymbol{u},\boldsymbol{v}$ 是两个同维矢量函数，且每个分量属于 $L_2(\Omega)$：

$$(\boldsymbol{u},\boldsymbol{v})_{L_2(\Omega)} = \int_\Omega \boldsymbol{u}(\boldsymbol{x}) \cdot \boldsymbol{v}(\boldsymbol{x})\, d\boldsymbol{x}$$

那么其自然满足*Cauchy-Schwarz Inequality*：$|(u,v)_{L_2(\Omega)}| \leq ||u||_{L_2(\Omega)} \cdot ||v||_{L_2(\Omega)}$
因此，$L_2(\Omega)$ 为一个希尔伯特空间（*Hilbert Space*）。

Theorem 1.3.2. 可积函数空间的性质 若勒贝格可测集合 $\Omega \subset \mathbb{R}^n$ 满足 $m_n(\Omega) < \infty$，则

1) $\forall\ 1 \leq q \leq p \leq \infty$ ： $L_p(\Omega) \subset L_q(\Omega)$，证明可见 *Hölder inequality*，略；
2) $\forall\ 1 \leq q \leq p \leq \infty$ ： $f_n \xrightarrow{L_p} f \Longrightarrow f_n \xrightarrow{L_q} f$，注意这里 $f_n \xrightarrow{L_p} f$ 指代 $\lim_{n\to\infty} ||f_n - f||_{L_p(\Omega)} = 0$ 。

1.4 Sobolev Spaces

Lemma 1.4.1. 记 $\Omega \subset \mathbb{R}^n$ 为开集合，取 $u \in C^k(\Omega)$ 与 $v \in C_0^\infty(\Omega)$，则对 $\forall\, x_i$：

$$\boldsymbol{\int_\Omega v(x)\,\partial_{x_i} u(x)\,dx = -\int_\Omega u(x)\,\partial_{x_i} v(x)\,dx}$$

Proof lemma 1.4.1. 由 $v \in C_0^\infty(\Omega)$，$\operatorname{supp} v \subset \Omega \Longrightarrow v(\boldsymbol{x}) = 0$，$\forall \boldsymbol{x} \in \partial\Omega$：

$$\begin{aligned}
&\int_\Omega v(\boldsymbol{x})\,\partial_{x_i} u(\boldsymbol{x})\,d\boldsymbol{x} \;+\; \int_\Omega u(\boldsymbol{x})\,\partial_{x_i} v(\boldsymbol{x})\,d\boldsymbol{x} \\
&= \int_\Omega \left(v\partial_{x_i} u + u\partial_{x_i} v\right) d\boldsymbol{x} = \int_\Omega \partial_{x_i}(vu)\,d\boldsymbol{x} \\
&= \int_\Omega \nabla\cdot(0,\cdots,vu,\cdots,0)\,d\boldsymbol{x} \quad (\text{for } i_{th} \text{ entry}) \\
&= \oint_{\partial\Omega} vu\cdot\gamma_i\; d\boldsymbol{S} = \oint_{\partial\Omega} 0\cdot\gamma_i\; d\boldsymbol{S} = 0
\end{aligned}$$

于是有：$\int_\Omega v(\boldsymbol{x})\,\partial_{x_i} u(\boldsymbol{x})\,d\boldsymbol{x} = -\int_\Omega u(\boldsymbol{x})\,\partial_{x_i} v(\boldsymbol{x})\,d\boldsymbol{x}$，证毕。

Theorem 1.4.1. 多元函数的分部积分（integration-by-parts formula） 记 $\Omega \subset \mathbb{R}^n$ 为开集合，取 $u \in C^k(\Omega)$ 与 $v \in C_0^\infty(\Omega)$，则对 $\forall\, \boldsymbol{\alpha} : |\boldsymbol{\alpha}| \le k$：

$$\boldsymbol{\int_\Omega v(x)\left(D^\alpha u(x)\,dx\right) = (-1)^{|\alpha|}\int_\Omega u(x)\left(D^\alpha v(x)\,dx\right)}$$

Proof theorem 1.4.1. 由引理与 theorem 1.2.2. 不难证明有：

$$\int_\Omega v(\boldsymbol{x})\,\partial_{x_i}^{a_i} u(\boldsymbol{x})\,d\boldsymbol{x} = (-1)^{a_i}\int_\Omega u(\boldsymbol{x})\,\partial_{x_i}^{a_i} v(\boldsymbol{x})\,d\boldsymbol{x}$$

故进一步有：

$$\begin{aligned}
&\int_\Omega v(\boldsymbol{x})\left(D^{\boldsymbol{\alpha}} u(\boldsymbol{x})\,d\boldsymbol{x}\right) = \int_\Omega v(\boldsymbol{x})\,\partial_{x_1}^{a_1}\{D^{(a_2,\cdots,a_n)} u(\boldsymbol{x})\}d\boldsymbol{x} \\
&= (-1)^{a_1}\int_\Omega \{D^{(a_2,\cdots,a_n)} u(\boldsymbol{x})\}\cdot\{\partial_{x_1}^{a_1} v(\boldsymbol{x})\}d\boldsymbol{x} \\
&= (-1)^{a_1}\int_\Omega \partial_{x_2}^{a_2}\{D^{(a_3,\cdots,a_n)} u(\boldsymbol{x})\}\cdot\{\partial_{x_1}^{a_1} v(\boldsymbol{x})\}d\boldsymbol{x} \\
&= (-1)^{a_1+a_2}\int_\Omega \{D^{(a_3,\cdots,a_n)} u(\boldsymbol{x})\}\cdot\{\partial_{x_2}^{a_2}\partial_{x_1}^{a_1} v(\boldsymbol{x})\}d\boldsymbol{x} \\
&\cdots = (-1)^{|\boldsymbol{\alpha}|}\int_\Omega u(\boldsymbol{x})\left(D^{\boldsymbol{\alpha}} v(\boldsymbol{x})\,d\boldsymbol{x}\right)
\end{aligned}$$

Theorem 1.4.2. 弱导数（weak derivative） 记 u 在 $\Omega \subset \mathbb{R}^n$ 上局部可积（*locally integrable*），即 $u \in L_1(\Omega), \forall\ \omega$ open and bounded s.t. $\bar{\omega} \subset \Omega$。若存在另一个在 Ω 上局部可积的函数 $w_{\boldsymbol{\alpha}}(\boldsymbol{x})$ 使得 $\forall\ v \in C_0^{\infty}(\Omega)$ 有：

$$\int_{\boldsymbol{\Omega}} \boldsymbol{u}(\boldsymbol{x}) \left(\boldsymbol{D}^{\boldsymbol{\alpha}} \boldsymbol{v}(\boldsymbol{x})\, \boldsymbol{dx}\right) = (\boldsymbol{-1})^{|\boldsymbol{\alpha}|} \int_{\boldsymbol{\Omega}} \boldsymbol{v}(\boldsymbol{x})\, \boldsymbol{w}_{\boldsymbol{\alpha}}(\boldsymbol{x})\, \boldsymbol{dx}$$

称 $w_{\boldsymbol{\alpha}}(\boldsymbol{x})$ 为 u 关于 $\boldsymbol{\alpha}$ 的弱导数，并记为 $w_{\boldsymbol{\alpha}}(\boldsymbol{x}) := D_w^{\boldsymbol{\alpha}} u(\boldsymbol{x})$。$v$ 称为测试函数 *test function*。当 u 足够光滑时，其弱导数等于对应的偏导数。见**Lecture Notes p4 Example** 如何寻找弱导数。

Definition 1.4.1. Sobolev Spaces 记 $k \in \mathbb{N}_0$，$\Omega \subset \mathbb{R}^n$ 为开集合，定义以下 *Sobolev* 函数空间与 *Sobolev* 范数（根号下所有可能导数的 L_2 范数平方和）：

$$H^k(\Omega) = \{u \in L_2(\Omega) : D_w^{\boldsymbol{\alpha}} u \in L_2(\Omega), \forall |\boldsymbol{\alpha}| \leq k\}$$

$$||u||_{H^k(\Omega)} = \left(\sum_{|\boldsymbol{\alpha}| \leq k} ||D_w^{\boldsymbol{\alpha}} u||_{L_2(\Omega)}^2\right)^{1/2} = \left(\sum_{j=0}^{k} |u|_{H^j(\Omega)}^2\right)^{1/2}$$

其中，$|u|_{H^j(\Omega)} = \left(\sum_{|\boldsymbol{\alpha}|=j} ||D_w^{\boldsymbol{\alpha}} u||_{L_2(\Omega)}^2\right)^{1/2}$ 因其不满足 $|u|_{H^j(\Omega)} = 0 \Longrightarrow u = 0$，称为 *Sobolev semi-norm* 表示所有所有 j 阶导数 L_2 范数的平方和开方。*Sobolev Spaces* 是一个希尔伯特空间，其相应的内积定义为：

$$(u, v)_{H^k(\Omega)} = \sum_{|\boldsymbol{\alpha}| \leq k} \left(D_w^{\boldsymbol{\alpha}} u\ ,\ D_w^{\boldsymbol{\alpha}} v\right)_{L_2(\Omega)}$$

特别的，定义子希尔伯特空间 $H_0^1(\Omega) = \{u \in H^1(\Omega) : u = 0 \text{ on } \partial\Omega\}$。

Theorem 1.4.3. Poincare-Friedrichs Inequality $\Omega \subset \mathbb{R}^n$ 为有界开集合，其边界 $\partial\Omega$ 足够平滑，则对 $\forall u \in H_0^1(\Omega)$，存在一个仅与 Ω 有关的常数 $c^*(\Omega)$（该常数与 u 无关），使得：

$$\int_{\Omega} |u(\boldsymbol{x})|^2 d\boldsymbol{x} \leq c^*(\Omega) \sum_{i=1}^{n} \int_{\Omega} |\partial_{x_i} u(\boldsymbol{x})|^2 d\boldsymbol{x}$$

或者可以写作：$||u||_{L_2(\Omega)}^2 \leq c^*(\Omega)\, |u|_{H^1(\Omega)}^2$。下面仅给出 $\Omega = (a, b) \times (c, d) \subset \mathbb{R}^2$ 的证明，一般情形的证明类似。

Proof theorem 1.4.3 显然，对 x 由微积分基本定理：

$$u(x,y)=u(a,y)+u(\xi,y)\,|_a^x=u(a,y)+\int_a^x\partial_\xi u(\xi,y)\,d\xi=\int_a^x\partial_\xi u(\xi,y)\,d\xi$$

最后一个等号是因为 $u\in H_0^1(\Omega)$ 在边界上取值为 0，故 $u(a,y)=0$。

$$\int_\Omega u^2(\boldsymbol{x})\,d\boldsymbol{x}=\int_a^b\left\{\int_c^d u^2(x,y)\,dy\right\}dx=\int_a^b\left\{\int_c^d\left|\int_a^x\partial_\xi u(\xi,y)\,d\xi\right|^2dy\right\}dx$$

对 $\left|\int_a^x\partial_\xi u(\xi,y)\,d\xi\right|^2$ 由 Cauchy-Schwarz 不等式：

$$\left|\int_a^x 1\cdot\partial_\xi u(\xi,y)\,d\xi\right|^2\le\int_a^x 1^2d\xi\cdot\int_a^x|\partial_\xi u(\xi,y)|^2d\xi\le(x-a)\int_a^b|\partial_\xi u(\xi,y)|^2d\xi$$

于是最终有：

$$\begin{aligned}&\int_\Omega u^2(\boldsymbol{x})\,d\boldsymbol{x}=\int_a^b\left\{\int_c^d\left|\int_a^x\partial_\xi u(\xi,y)\,d\xi\right|^2dy\right\}dx\\&\le\int_a^b\left\{\int_c^d(x-a)\int_a^b|\partial_\xi u(\xi,y)|^2d\xi dy\right\}dx\\&=\int_a^b(x-a)\,dx\cdot\left\{\int_c^d\int_a^b|\partial_\xi u(\xi,y)|^2d\xi dy\right\}=\frac{(b-a)^2}{2}\cdot\int_\Omega|\partial_x u(x,y)|^2d\boldsymbol{x}\end{aligned}$$

同理有：

$$\int_\Omega u^2(\boldsymbol{x})\,d\boldsymbol{x}\le\frac{(c-d)^2}{2}\cdot\int_\Omega|\partial_y u(x,y)|^2d\boldsymbol{x}$$

将系数除到左边合并二式可得：

$$\int_\Omega u^2(\boldsymbol{x})\,d\boldsymbol{x}\le c^*(a,b,c,d)\cdot\left\{\int_\Omega|\partial_x u(x,y)|^2d\boldsymbol{x}+\int_\Omega|\partial_y u(x,y)|^2d\boldsymbol{x}\right\}$$

其中 $c^*(a,b,c,d)=\left(\frac{2}{(b-a)^2}+\frac{2}{(c-d)^2}\right)^{-1}$，证毕。

2 Elliptic Boundary Value Problems

2.1 Formulation

在本课程中，我们考虑如下的椭圆型偏微分方程，其定义在有界开集 $\Omega \in \mathbb{R}^n$：

$$-\sum_{i,j=1}^{n} \frac{\partial}{\partial x_j}\left(a_{i,j}(\boldsymbol{x}) \frac{\partial u}{\partial x_i}\right)+\sum_{i=1}^{n} b_i(\boldsymbol{x}) \frac{\partial u}{\partial x_i}+c(\boldsymbol{x}) u=f(\boldsymbol{x}) \qquad (*)$$

或写成 dense form：

$$-\nabla \cdot\left(\nabla^T u A(\boldsymbol{x})\right)+\boldsymbol{b}(\boldsymbol{x}) \cdot \nabla u+c(\boldsymbol{x}) u=f(\boldsymbol{x})$$

其中：$a_{i,j}(\boldsymbol{x}) \in C^1(\bar{\Omega})$；$b_i, c, f \in C(\bar{\Omega})$；且 $\forall \xi \in \mathbb{R}^n,\ \forall \boldsymbol{x} \in \bar{\Omega}$：

$$\xi^T A \xi \geq c^* \xi^T \xi \quad \text{or} \quad \sum_{i,j=1}^{n} a_{i,j}(\boldsymbol{x}) \xi_i \xi_j \geq c^* \sum_{i=1}^{n} \xi_i^2$$

其中常数 $c^*>0$ 与 $\boldsymbol{x}$，ξ 无关。上述条件称为 uniform ellipticity。其中常见的边界条件有（将本笔记的 ∇ 视为**列向量**）：

(a) Dirichlet：$u=g$ on $\partial\Omega$，当 $g=0$ 时称边界条件齐次（homogeneous）；

(b) Neumann：$\frac{\partial u}{\partial \boldsymbol{v}} := \nabla u \cdot \boldsymbol{v}=g$ on $\partial\Omega$，其中 $\boldsymbol{v}$ 表示边界上的单位外法向量;

(c) Robin：$\frac{\partial u}{\partial \boldsymbol{v}}+\sigma u=g$ on $\partial\Omega$，其中 $\forall \boldsymbol{x} \in \Omega : \sigma(\boldsymbol{x}) \geq 0$；

(d) Oblique derivative：$\nabla^T u A \boldsymbol{v}+\sigma(\boldsymbol{x}) u=g$ on $\partial\Omega$。

Definition 2.1.1. 典型解（classical solutions） *考虑齐次 Dirichlet 条件下的椭圆型偏微分方程 $(*)$，若存在解 $u \in C^2(\Omega) \cap C(\bar{\Omega})$，则称其为一个典型解。当 $A, \boldsymbol{b}, c, f$ 和 $\partial\Omega$ 充分平滑时，此类问题的典型解存在且唯一。*

Definition 2.1.2. 二元二阶拟线性偏微分方程 *形如下式的，定义于 $\mathbb{R}^2$ 的偏微分方程称为二元二阶拟线性偏微分方程：*

$$a(x, y) u_{xx}+b(x, y) u_{xy}+c(x, y) u_{yy}=f\left(u_x, u_y, u, x, y\right)$$

其分类如下：

1) $b^2-4ac>0$：双曲型（*hyperbolic*）
2) $b^2-4ac=0$：抛物型（*parabolic*）
3) $b^2-4ac<0$：椭圆型（*elliptic*）

2.2 Weak Solutions

Definition 2.2.1. 弱形式解 (weak solution) 考虑定义在有界开集 $\Omega \subset \mathbb{R}^n$ 下的齐次 *Dirichlet* 条件椭圆型偏微分方程:

$$-\sum_{i,j=1}^{n} \frac{\partial}{\partial x_j}\left(a_{i,j}(\boldsymbol{x})\frac{\partial u}{\partial_{x_i}}\right)+\sum_{i=1}^{n} b_i(\boldsymbol{x})\frac{\partial u}{\partial x_i}+c(\boldsymbol{x})u=f(\boldsymbol{x})$$

$$\text{s.t. } u(\boldsymbol{x})=0,\ \boldsymbol{x}\in\partial\Omega$$

且$a_{i,j}, b_i, c \in C(\bar{\Omega})$, $f \in L_2(\Omega)$。则称满足下列条件的 $u \in H_0^1(\Omega)$ 为该偏微分方程的弱解,$\forall v \in H_0^1(\Omega)$:

$$\int_\Omega \left\{-\sum_{i,j=1}^{n} \frac{\partial}{\partial x_j}\left(a_{i,j}(\boldsymbol{x})\frac{\partial u}{\partial_{x_i}}\right)+\sum_{i=1}^{n} b_i(\boldsymbol{x})\frac{\partial u}{\partial x_i}+c(\boldsymbol{x})u\right\}\cdot v d\boldsymbol{x}=\int_\Omega f(\boldsymbol{x})\cdot v d\boldsymbol{x}$$

上式可简单化为:

$$\sum_{i,j=1}^{n}\int_\Omega a_{i,j}\frac{\partial u}{\partial_{x_i}}\frac{\partial v}{\partial x_j}d\boldsymbol{x}+\sum_{i=1}^{n}\int_\Omega b_i\frac{\partial u}{\partial x_i}\cdot v d\boldsymbol{x}+\int_\Omega cu\cdot v d\boldsymbol{x}=\int_\Omega f\cdot v d\boldsymbol{x}$$

或写成 *dense form* :

$$(\nabla u, A\nabla v)_{L_2(\Omega)}+(\boldsymbol{b}\cdot\nabla u, v)_{L_2(\Omega)}+(cu,v)_{L_2(\Omega)}=(f,v)_{L_2(\Omega)}$$

Remark 上述积分中的二阶偏导项使用了 **Thm 1.4.1.** 中的分部积分转化:

$$\int_\Omega\left\{-\sum_{i,j=1}^{n}\frac{\partial}{\partial x_j}\left(a_{i,j}(\boldsymbol{x})\frac{\partial u}{\partial_{x_i}}\right)\right\}\cdot v d\boldsymbol{x}=-\sum_{i,j=1}^{n}\int_\Omega\frac{\partial}{\partial x_j}\left(a_{i,j}(\boldsymbol{x})\frac{\partial u}{\partial_{x_i}}\right)\cdot v d\boldsymbol{x}$$
$$=-\sum_{i,j=1}^{n}-\int_\Omega\left(a_{i,j}(\boldsymbol{x})\frac{\partial u}{\partial_{x_i}}\right)\frac{\partial v}{\partial x_j}d\boldsymbol{x}=\sum_{i,j=1}^{n}\int_\Omega a_{i,j}(\boldsymbol{x})\frac{\partial u}{\partial_{x_i}}\frac{\partial v}{\partial x_j}d\boldsymbol{x}$$

将方程设置为 齐次 Dirichlet BC 下的 Laplace Equation, 即 $A(\boldsymbol{x})=I$, 可以得到一个简单的分布积分式子:

$$-\int_\Omega v\triangle u d\boldsymbol{x}=\int_\Omega\nabla u\cdot\nabla v d\boldsymbol{x}$$

其中 $u, v \in H_0^1(\Omega)$, Ω open bounded (注: $\triangle u=\sum_{i=1}^{n}\partial_{x_i}^2 u$)。关于弱解的存在性可详见 **Lecture Notes pp. 9-12** 的补充材料。

3 Finite Difference Schemes for ODEs

3.1 Finite Differences

求解微分方程的通常思路是将求解空间离散化后，有有限差分（实际上是差商）近似导数，将差分方程化为一般的线性方程组求解。为了详细讨论这些过程，我们规定以下符号与属术语：首先考虑以下的一般线性问题（$\mathcal{L}$ 与 $\mathcal{B}$ 均为线性算子）：

$$\begin{aligned}\mathcal{L}u(\boldsymbol{x}) &= f(\boldsymbol{x}) \text{ if } \boldsymbol{x} \in \Omega \\ \mathcal{B}u(\boldsymbol{x}) &= g(\boldsymbol{x}) \text{ if } \boldsymbol{x} \in \partial\Omega\end{aligned}$$

我们定义求解区域的离散化网格如下

Definition 3.1.1. 差分网格（finite difference mesh） 方便起见，我们暂时先考虑边界互相平行的区域 Ω。假设对于求解区域 Ω，我们对其的每个轴 Ox_i 取等距的离散化坐标 $x_i^j = x_i^0 + jh_i\ ,\ \forall j = 1\cdots n_i$，得到一个有限点集 $\overline{\Omega}_h = \Omega_h \cup \partial\Omega_h$，称为目标区域的一个网格 *mesh*，其中

1. $h = (h_1, \cdots, h_n)$ 为每根轴上节点的网格间距 *mesh-size*，为了方便后续讨论，当 $h_1 = \cdots = h_n = h$ 时且不引起歧义的情况下，我们直接用 h 表示公共步长；否则，用 $h_{max} = \max_{1\le i\le n} h_i$ 表示最大步长。此外，当我们想强调 h 的矢量/多重下标性质时，我们会使用粗体 $\boldsymbol{h}$，记：$\boldsymbol{h}^* = \prod_i h_i$。

2. $\boldsymbol{x}_{\boldsymbol{\kappa}} \in \Omega_h \subset \Omega \setminus \partial\Omega$ 称为内部网格点 *interior mesh-points/nodes*；$\boldsymbol{x}_{\boldsymbol{\kappa}} \in \partial\Omega_h \subset \partial\Omega$ 称为边界网格点集 *boundary mesh-points/nodes*；其中 $\boldsymbol{\kappa}$ 为节点的 *multi-index*。

在离散网格确定后，我们会对内部网格点 $\boldsymbol{x}_{\boldsymbol{\kappa}} \in \Omega_h$ 套用微分方程，并对边界网格点 $\boldsymbol{x}_{\boldsymbol{\kappa}} \in \partial\Omega_h$ 施加边界条件。这都需要我们对算子 $\mathcal{L}$ 与 $\mathcal{B}$ 中的导数在相应的网格点上进行近似。常见的导数与其有限差分近似为：

Theorem 3.1.1. 一阶导数的有限差分近似（finite difference approximation for first-order derivatives） 我们提供以下常见的差分公式来近似所需的一阶导数，类似地可以推广到任何函数的一阶偏导数上：

1. 一阶导一阶向前差分 $f'(x_i) = D_x^+ f(x_i) + O(h) = \frac{f_{i+1}-f_i}{h} + O(h)$
2. 一阶导一阶向后差分 $f'(x_i) = D_x^- f(x_i) + O(h) = \frac{f_i-f_{i-1}}{h} + O(h)$
3. 一阶导二阶中心差分 $f'(x_i) = D_x^c f(x_i) + O(h^2) = \frac{f_{i+1}-f_{i-1}}{2h} + O(h^2)$

其英文表达分别是：*first/first/second-order forward/backward/central first divided difference*，其中的 *order* 指代近似的精度。

Theorem 3.1.2. 二阶导数的有限差分近似 (finite difference approximation for first-order derivatives) 我们提供以下常见的差分公式来近似所需的二阶导数，类似地可以推广到任何函数的二阶偏导数上：

1. 二阶导二阶中心差分 *second-order symmetric/central second divided difference*，$f''(x_i) = D_{xx}^c f(x_i) + O(h^2)$ 其中

$$D_{xx}^c f(x_i) = D_x^+ D_x^- f(x_i) = \frac{f_{i+1} - 2f_i + f_{i-1}}{h^2}$$

2. 二阶混合导一阶有限差分 $u_{xy}(x_i, y_j) = D_{xy} u(x_i, y_j) + O(h)$ 其中

$$D_{xy} f(x_i, x_j) = D_y^c D_x^c f(x_i, x_j) = \frac{f_{i+1,j+1} - f_{i+1,j-1} - f_{i-1,j+1} + f_{i-1,j-1}}{4h^2}$$

Theorem 3.1.3. 上述差分算子还满足如下的一些关系

1. $D_x^c f_i = \frac{1}{2}\left(D_x^+ f_i + D_x^- f_i\right)$
2. $D_{xx}^c f_i = D_x^+ D_x^- f_i = D_x^- D_x^+ f_i$
3. $D_{xy}^c f_{ij} = D_y^c D_x^c f_{ij} = D_x^c D_y^c f_{ij}$

Theorem 3.1.4. 插值多项式构造的差分公式 取足够光滑的函数 f，其在定义域内有 $(n+1)$ 个等距 (h) 节点 $x_0 < \cdots < x_i < \cdots < x_n$，利用这些节点与其上对应的函数值 f_i 构造 n 阶插值多项式，并用该多项式在 x_i 的导数 $\tilde{f}'_i$ 近似 f'_i。此公式对于任何 $(n+1)$ 个连续节点在相对位置相同的第 i 个节点上均适用，其 k 阶导数的近似的误差为 (记 $(q)_n = q\,(q-1)\cdots(q-i+1)$)

$$\left|\tilde{f}'_i - f'_i\right| = \frac{h^{n+1-k} \cdot \left|f^{(n+1)}(\xi)\right|}{(n+1)!} \left\{\frac{\mathrm{d}^k}{\mathrm{d}q^k}(q)_{n+1}\right\}\Bigg|_{q=i} \sim O(h^{n+1-k})$$

若有奇数个节点，对于中间节点的偶数阶导数，精度可以上升一阶，为$O(h^{n+2-k})$。二阶导数的二阶中心差分 $D_{xx}^c f_i$ 可以用这种方法构造。

一般来说，差分公式的构造还可通过待定系数法与在目标点的泰勒展开实现，详细例子见 MATH 336 笔记。在使用差分公式替换掉方程与边界条件中的导数项后，我们可以将一开始的问题替换为解一个线性方程组：

$$\mathcal{L}_h \tilde{u}(\boldsymbol{x}) = f_h(\boldsymbol{x}) \text{ if } \boldsymbol{x} \in \Omega_h$$
$$\mathcal{B}_h \tilde{u}(\boldsymbol{x}) = g_h(\boldsymbol{x}) \text{ if } \boldsymbol{x} \in \partial\Omega_h$$

我们称 $\{\tilde{u}(\boldsymbol{x_\kappa}) : \boldsymbol{x_\kappa} \in \overline{\Omega}_h\}$ 为 $\{u(\boldsymbol{x_\kappa}) : \boldsymbol{x_\kappa} \in \overline{\Omega}\}$ 的近似解。

3.2 Existence and Uniqueness

本节我们主要讨论形如下的常微分方程边界值问题 $(*)$ 数值解的各种性质：

$$-u'' + c(x)\,u = f(x) \ \text{ if } x \in (0,1)$$
$$u(0) = u(1) = 0$$

其中，$f, c \in C([0,1])$，且 $c(x) \geq 0,\ \forall x \in [0,1]$。我们记 u 的数值解为 $\tilde{u}$。通常，对于定义在一个特定区间 $\overline{\Omega} = [a,b]$ 内的 ODE 问题，其等距差分网格为：$\overline{\Omega}_h = \{x_i \ :\ x_i = a + ih,\ h = (b-a)/N,\ i = 0 \cdots N\}$，其中边界节点为 $\partial\Omega_h = \{x_0\ ,\ x_N\}$。在开始具体分析之前，我们做如下定义：

Definition 3.2.1. 差分网格上的函数 微分方程的数值解均可以定义为在差分网格上的函数。当我们在下文中说 "$\boldsymbol{w}$ 是一个定义在差分网格上的函数" 时，也可以理解为我们在假设其是一个目标方程的潜在数值解。一般使用如下记号表示一个定义为在差分网格上的函数 $\boldsymbol{w}$：

$$\boldsymbol{w}(\boldsymbol{x}_{\boldsymbol{\kappa}}) = w_{\boldsymbol{\kappa}} \quad \forall\ \boldsymbol{x}_{\boldsymbol{\kappa}} \in \overline{\Omega}_h$$

一般来说我们会更关心在解函数在区域内部节点上的取值，也就是定义在 Ω_h 上的函数（数值解）。定义这些内部节点的下标集合为：$K(\Omega_h) = \{\boldsymbol{\kappa} : \boldsymbol{x}_{\boldsymbol{\kappa}} \in \Omega_h\}$
现在，设有两个在内部节点上有定义的函数 $\boldsymbol{w}, \boldsymbol{v}$，我们定义其内积为：

$$(\boldsymbol{w}, \boldsymbol{v})_h = \sum_{\boldsymbol{\kappa} \in K(\Omega_h)} \boldsymbol{h}^* w_{\boldsymbol{\kappa}} v_{\boldsymbol{\kappa}} = \left(\prod_j h_j\right) \cdot \sum_{\boldsymbol{\kappa} \in K(\Omega_h)} w_{\boldsymbol{\kappa}} v_{\boldsymbol{\kappa}}$$

这个定义基本上是函数在 $L_2(\Omega)$ 上内积的离散版本。我们允许有限差分算符应用在这些离散函数的内部节点上（有时边界点也可以），相当于在对应的 $\boldsymbol{x}_{\boldsymbol{\kappa}}$ 上用 $\boldsymbol{v}$ 的值进行 *element-wise* 运算。

回到 *ODE* 的情境，考虑两个个定义在一维网格点 $\overline{\Omega}_h = \{x_i \ :\ x_i = a+ih,\ h = (b-a)/N,\ i = 0 \cdots N\}$ 上的函数 $\boldsymbol{w}, \boldsymbol{v}$，其内积为：

$$(\boldsymbol{w}, \boldsymbol{v})_h = \sum_{i=1}^{N-1} h w_i v_i$$

Lemma 3.2.1. Summation by Parts 假设 $\boldsymbol{v}$ 是定义在一维（等距）网格点 $\overline{\Omega}_h$ 上的函数，且 *test function* $\boldsymbol{\phi}$ 满足齐次边界条件 $\boldsymbol{\phi}_0 = \boldsymbol{\phi}_N = 0$，则：

$$-(D_x^+ \boldsymbol{v}, \boldsymbol{\phi})_h = \sum_{i=1}^{N} h |D_x^- \boldsymbol{\phi}_i|^2 =: (\boldsymbol{v}, D_x^- \boldsymbol{\phi}]_h$$

对于二阶导数，注意有：$D_{xx}^c = D_x^+ D_x^-$。这很像端点值为 0 时候的连续情形的分部积分 $-(u', \phi)_{L_2([a,b])} = (u, \phi')_{L_2([a,b])}$，证明比较简单，见 Lecture Notes p.16。一个简单的推论：当 $\boldsymbol{v}$ 也满足其次边界条件时，$(D_{xx}^c \boldsymbol{v}, \boldsymbol{\phi})_h = (\boldsymbol{v}, D_{xx}^c \boldsymbol{\phi})_h$

Existence and Uniqueness of the Solution

我们第一步将讨论关于常微分方程边界值问题 $(*)$ 数值解的存在与唯一性。整体的思路是将问题化为离散的形式并最终转化为解线性方程组 $A\tilde{u}=f_h$，随后只要证明系数矩阵 A 可逆即可。这等价于证明齐次方程组 $A\tilde{u}=0$ 只有平凡解。考虑常微分方程边界值问题 $(*)$

$$\begin{aligned} -u''+c(x)\,u &= f(x) \ \text{ if } x\in(0,1)\\ u(0)=u(1) &= 0 \end{aligned}$$

其中，$f,c\in C([0,1])$，且 $c(x)\geq 0,\ \forall x\in[0,1]$。我们记 u 的数值解为 $\tilde{u}$。将其离散化为如下的线性方程组：

$$\begin{bmatrix} \frac{2}{h^2}+c_1 & -\frac{1}{h^2} & 0 & \cdots & 0\\ -\frac{1}{h^2} & \frac{2}{h^2}+c_2 & -\frac{1}{h^2} & \ddots & \vdots\\ 0 & \ddots & \ddots & \ddots & 0\\ \vdots & \ddots & -\frac{1}{h^2} & \frac{2}{h^2}+c_{N-2} & -\frac{1}{h^2}\\ 0 & \cdots & 0 & -\frac{1}{h^2} & \frac{2}{h^2}+c_{N-1} \end{bmatrix} \begin{bmatrix} \tilde{u}_1\\ \tilde{u}_2\\ \vdots\\ \tilde{u}_{N-2}\\ \tilde{u}_{N-1} \end{bmatrix} = \begin{bmatrix} f_1\\ f_2\\ \vdots\\ f_{N-2}\\ f_{N-1} \end{bmatrix}$$

将系数矩阵写成 A，方程化为更紧凑的形式 $A\tilde{u}=f_h$，实际上我们不难发现系数矩阵 A 的作用等价于线性差分算子 $L=-D^c_{xx}+c_h$，方程的离散化（在齐次边界条件下）整体等价于将这个算子应用到定义在内部网格点 Ω_h 的离散值函数 $\tilde{u}$ 上

$$A\tilde{u}=f_h \quad\Longleftrightarrow\quad -D^c_{xx}\tilde{u}+c_h\tilde{u}=f_h$$

注意，上述的 c_h,f_h 在差分算符中也视为离散值函数。对任意满足其次边界条件的离散值函数 $\boldsymbol{v}$ 考虑内积

$$\begin{aligned} (A\boldsymbol{v},\boldsymbol{v})_h &= \left(-D^c_{xx}\boldsymbol{v}+c_h\boldsymbol{v},\boldsymbol{v}\right)_h = -\left(D^c_{xx}\boldsymbol{v},\boldsymbol{v}\right)+(c_h\boldsymbol{v},\boldsymbol{v})_h\\ &= \textstyle\sum_{i=1}^N h|D_x^- v_i|^2+(c_h\boldsymbol{v},\boldsymbol{v})_h \geq \ \sum_{i=1}^N h|D_x^- v_i|^2 \ \geq\ 0 \end{aligned}$$

不等号成立是因为我们规定了 $c(x)\geq 0$ 在求解区域内处处成立；第三个等号运用了 lemma 3.2.1。上述的不等关系说明，如果存在 $A\boldsymbol{v}=0$，此时该内积也为 0，则必然有 $\sum_{i=1}^N h|D_x^- v_i|^2=0$ 即 $D_x^- v_i\equiv 0$ 对 $i=1\cdots N$ 恒成立。考虑到有边界条件 $\boldsymbol{v}=0$，则由递推关系能得到 $\boldsymbol{v}\equiv 0$。即 $A\boldsymbol{v}=0\iff\boldsymbol{v}=0$，即 A 为可逆矩阵，数值解存在且唯一。

Theorem 3.2.1. 对于 *ODE* 边界值问题 $(*)$，有限差分法的解存在且唯一。

3.3 Stability and Consistency

Definition 3.3.1. 离散 L_2 范数 (discrete L_2 norm) 设有一个在离散网格节点上有定义的函数 $\boldsymbol{v}$，我们定义其 L_2 范数（只考虑内部节点）为

$$||\boldsymbol{v}||_h = \sqrt{(\boldsymbol{v},\boldsymbol{v})_h} = \sqrt{\sum_{\boldsymbol{\kappa}\in K(\Omega_h)} \boldsymbol{h}^* v_{\boldsymbol{\kappa}}^2}$$

有时候，范数可能或涉及到边界节点。于是对于一维形式，我们定义含右边界内积：

$$(\boldsymbol{w},\boldsymbol{v}]_h = \sum_{i=1}^{N} h w_i v_i$$

含边界范数：

$$||\boldsymbol{v}]||_h = \sqrt{(\boldsymbol{v},\boldsymbol{v}]_h} = \sqrt{\sum_{i=1}^{N} h v_i^2}$$

一维离散 *Sobolev* 范数：

$$||\boldsymbol{v}||_{1,h} = \left(||\boldsymbol{v}||_h^2 + ||D_x^- \boldsymbol{v}]||_h^2\right)^{\frac{1}{2}}$$

Lemma 3.3.1. Discrete Poincare-Friedrichs Inequality (DPFI) 令 $\boldsymbol{v}$ 为定义在一维离散网格点 $\overline{\Omega}_h = \{x_i \ : \ x_i = a + ih,\ h = (b-a)/N,\ i = 0\cdots N\}$ 上的函数，且满足边界条件 $\boldsymbol{v}_0 = \boldsymbol{v}_N = 0$；则存在一个与 $\boldsymbol{v}$，h 无关的常数 $c^* = 1/2$ 使得

$$||\boldsymbol{v}||_h^2 \le \frac{1}{2}||D_x^- \boldsymbol{v}]||_h^2$$

回顾连续版的 *CPFI* 我们有 $||u||_{L_2(\Omega)}^2 \le c^* |u|_{H^1(\Omega)}^2$。证明见 Lecture Notes p.18。

Lemma 3.3.2. 令 $\boldsymbol{v}$ 为定义在一维离散网格点 $\overline{\Omega}_h = \{x_i \ : \ x_i = a + ih,\ h = (b-a)/N,\ i = 0\cdots N\}$ 上的函数，且满足边界条件 $\boldsymbol{v}_0 = \boldsymbol{v}_N = 0$；则，对于边界值问题 $(*)$ 有

$$(A\boldsymbol{v},\boldsymbol{v})_h \ge 2||\boldsymbol{v}||_h^2 \quad \& \quad (A\boldsymbol{v},\boldsymbol{v})_h \ge \frac{2}{3}||\boldsymbol{v}||_{1,h}^2$$

第一个式子由前页存在性讨论中的不等关系 $(**)$ $(A\boldsymbol{v},\boldsymbol{v})_h \ge \sum_{i=1}^{N} h|D_x^- v_i|^2 = ||D_x^- \boldsymbol{v}]||_h^2$ 与 *DPFI* 显然可得。第二个式子直接将新得到的式子与 $(**)$ 相加即可。特别地，对于由差分法得到的解 $\tilde{u}$ 有（证明见 Lecture Notes p.18）

$$||\tilde{u}||_{1,h} \le \frac{3}{2}||f_h||_h$$

Stability and the Error/Consistency of the Solution

在这一部分我们讨论问题 $(*)$ 数值解的稳定性与绝对误差。当我们讨论一个数值解是否稳定时，我们是在讨论初始数据的微小扰动是否也只会引起解（对于某种范数）的微小扰动。这在这里是显然的，因为所有的符合边界条件的数值解 $\boldsymbol{v}$ 的离散 Sobolev 范数都会被 $3/2||f_h||_h$ bounded (lemma 3.2.3)。我们在这里主要讨论方法的全局误差，假设有一个符合边界条件且定义在网格上的离散函数 $\tilde{u}$ 为方法导出的数值解。

定义全局误差为离散函数：$\boldsymbol{e} = \tilde{u} - \boldsymbol{u}$，显然有 $e_0 = e_N = 0$，于是：

$$\begin{aligned} A\boldsymbol{e} &= A(\tilde{u} - \boldsymbol{u}) = f_h - A\boldsymbol{u} = f_h - \left(-D_{xx}^c \boldsymbol{u} + c_h \boldsymbol{u}\right) \\ &= \left(-\boldsymbol{u''} + c_h \boldsymbol{u}\right) - \left(-D_{xx}^c \boldsymbol{u} + c_h \boldsymbol{u}\right) \\ &= D_{xx}^c \boldsymbol{u} - \boldsymbol{u''} := \varphi_h \end{aligned}$$

由 lemma 3.2.3. $||\boldsymbol{e}||_{1,h} \leq \frac{3}{2}||\varphi_h||_h$，取 $||u^{(4)}||_{C([0,1])} = \max\limits_{x\in[0,1]}|u^{(4)}|$，又有

$$\begin{aligned} &\left|D_{xx}^c u_i - u''(x_i)\right| = \left|\frac{1}{h^2}\left(u_{i+1} - 2u_i + u_{i-1}\right) - u''(x_i)\right| \\ =& \left|\frac{1}{h^2}\left(u_i + u_i' h + \frac{1}{2}u_i'' h^2 + \frac{1}{6}u_i^{(3)} h^3 + \frac{1}{24}u^{(4)}(\xi_1) h^4\right)\right. \\ &\left. + \frac{1}{h^2}\left(u_i - u_i' h + \frac{1}{2}u_i'' h^2 - \frac{1}{6}u_i^{(3)} h^3 + \frac{1}{24}u^{(4)}(\xi_2) h^4\right) - \frac{2}{h^2}u_i - u''(x_i)\right| \\ =& \left|\frac{1}{24}u^{(4)}(\xi_1) h^2 + \frac{1}{24}u^{(4)}(\xi_2) h^2\right| \leq \frac{1}{12}h^2 ||u^{(4)}||_{C([0,1])} := Ch^2 \end{aligned}$$

于是

$$\begin{aligned} ||\varphi_h||_h &= \sqrt{\textstyle\sum_{i=1}^{N-1} h\varphi_i^2} \leq \sqrt{\textstyle\sum_{i=1}^{N-1} hC^2h^4} = \sqrt{(N-1)hC^2h^4} \\ &\leq \sqrt{(Nh)C^2h^4} = \sqrt{1 \cdot C^2h^4} = Ch^2 = \frac{1}{12}h^2||u^{(4)}||_{L_\infty([0,1])} \end{aligned}$$

所以综上

$$||\boldsymbol{e}||_{1,h} \leq \frac{3}{2}||\varphi_h||_h \leq \frac{1}{8}h^2||u^{(4)}||_{C([0,1])}$$

Theorem 3.3.1. 对于常微分方程边界值问题 $(*)$，若 $u \in C^4\left([0,1]\right)$，则有限差分法的误差为

$$||\tilde{u} - \boldsymbol{u}||_{1,h} \leq \frac{1}{8}h^2||u^{(4)}||_{C([0,1])}$$

4 Finite Difference Schemes for Elliptic PDEs

4.1 Introduction and Terminologies

前一节的稳定性与一致性分析均可以类似地嵌套到（Elliptic）PDE 的边界值问题上，考虑我们在最开始讨论的一类线性椭圆 BVP $(*)$

$$\begin{aligned}\mathcal{L}u\left(\boldsymbol{x}\right) &= f\left(\boldsymbol{x}\right) \text{ if } \boldsymbol{x} \in \Omega \\ \mathcal{B}u\left(\boldsymbol{x}\right) &= g\left(\boldsymbol{x}\right) \text{ if } \boldsymbol{x} \in \partial\Omega\end{aligned}$$

其离散化后得到的线性方程组表示为 $(**)$

$$\begin{aligned}\mathcal{L}_h\tilde{u}\left(\boldsymbol{x}\right) &= f_h\left(\boldsymbol{x}\right) \text{ if } \boldsymbol{x} \in \Omega_h \\ \mathcal{B}_h\tilde{u}\left(\boldsymbol{x}\right) &= g_h\left(\boldsymbol{x}\right) \text{ if } \boldsymbol{x} \in \partial\Omega_h\end{aligned}$$

假设我们能定义与网格 $\overline{\Omega}_h$ 有关的两个涉及内部节点 Ω_h（或 $\overline{\Omega}_h$）的范数 $||\cdot||'_{\Omega_h}$（for stability）与 $||\cdot||_{\Omega_h}$；以及一个仅涉及边界节点 $\partial\Omega_h$ 的范数 $||\cdot||_{\partial\Omega_h}$，比如在前文中，我们有 $||\cdot||'_{\Omega_h} = ||\cdot||_{1,h}$ 与 $||\cdot||_{\Omega_h} = ||\cdot||_h$。我们定义：

Definition 4.1.1. 差分法的稳定性（stability） 对于任意一个可能的满足边界条件的数值解 $\tilde{u}$，如果存在一个与网格无关的常数 C 使得

$$||\tilde{u}||'_{\Omega_h} \leq C\left(\ ||f_h||_{\Omega_h} + ||g_h||_{\partial\Omega_h}\ \right)$$

则我们称该数值方法产生的数值解 $\tilde{u}$ 是稳定的。比如在前文中，我们由 *lemma 3.3.2* 有

$$||\tilde{u}||_{1,h} \leq \frac{3}{2}||f_h||_h$$

Definition 4.1.2. 对边界条件稳定与适定性 如果对边界条件进行扰动，使其变为 $g_h^1(x)$ 与 $g_h^2(x)$，对这两个边界条件求解原方程 $\mathcal{L}_h\tilde{u}\left(\boldsymbol{x}\right) = f_h\left(\boldsymbol{x}\right)$ 分别得到 $\tilde{u}_1$ 与 $\tilde{u}_2$，如果存在一个与网格无关的常数 C 使得

$$||\tilde{u}_1 - \tilde{u}_2||'_{\Omega_h} \leq C||g_h^1 - g_h^2||_{\partial\Omega_h}$$

则称该方法对边界条件稳定 stability w.r.t the perturbations in the boundary data；这同时也意味着 i.e., the solution's behavior changes continuously w.r.t the initial conditions（*sufficient*）。如果对某个问题满足：解存在且唯一，同时解对初始条件连续（比如此处的对边界条件稳定），则称该问题是适定的 *well-posed*，否则称为 *ill-posed*。

Definition 4.1.3. 差分法的一致性（consistency）考虑一致性误差 *consistency error*，注意下面的 $\boldsymbol{u}$ 表示准确值，与 $\tilde{u}$ 区分开来

$$\begin{aligned}\varphi_{\Omega_h} &= \mathcal{L}_h\left(\boldsymbol{u}-\tilde{u}\right) = \mathcal{L}_h\boldsymbol{u} - f_h\left(\boldsymbol{x}\right) = \mathcal{L}_h\boldsymbol{u} - \left(\mathcal{L}\boldsymbol{u}\right)_h \quad \text{if } \boldsymbol{x}\in\Omega_h \\ \varphi_{\partial\Omega_h} &= \mathcal{B}_h\left(\boldsymbol{u}-\tilde{u}\right) = \mathcal{B}_h\boldsymbol{u} - g_h\left(\boldsymbol{x}\right) = \mathcal{B}_h\boldsymbol{u} - \left(\mathcal{B}\boldsymbol{u}\right)_h \quad \text{if } \boldsymbol{x}\in\partial\Omega_h\end{aligned}$$

如果有这些误差依 $h\to 0$ 收敛，则称该方法一致 *consistent*。

$$||\varphi_{\Omega_h}||_{\Omega_h} + ||\varphi_{\partial\Omega_h}||_{\partial\Omega_h} \longrightarrow 0 \quad \text{as} \quad h\to 0$$

特别地，如果对足够光滑的 u 存在最大的 p，有下面的关系成立，则称该方法有 p-阶一致性 *have order of accuracy/consistency* p。

$$||\varphi_{\Omega_h}||_{\Omega_h} + ||\varphi_{\partial\Omega_h}||_{\partial\Omega_h} \sim O(h^p) \quad \text{as} \quad h\to 0$$

Definition 4.1.4. 差分法的收敛性（convergence）对于有限差分法，如果对其解 $\tilde{u}$ 有

$$||u-\tilde{u}||'_{\Omega_h} \longrightarrow 0 \quad \text{as} \quad h\to 0$$

则称该方法下数值解收敛 *convergent*。特别地，如果对足够光滑的 u 存在最大的 q，有下面的关系成立，则称该方法有 q-阶收敛性 *have order of convergence* q。

$$||u-\tilde{u}||'_{\Omega_h} \sim O(h^q) \quad \text{as} \quad h\to 0$$

在证明收敛性时，我们一般会结合稳定性和一致性，对 $\boldsymbol{e}=\boldsymbol{u}-\tilde{u}$ 同时作用算符 $\mathcal{L}_h$，得到新的离散化方程 $\mathcal{L}_h\boldsymbol{e}=\varphi_{\Omega_h}$，$\mathcal{B}_h\boldsymbol{e}=\varphi_{\partial\Omega_h}$，再对该方程利用其对应的差分策略稳定性和一致性，或者利用类似前文 *Lemma 3.3.2* 和后文 *Lemma 4.3.3* 的不等关系 $||e||'_{\Omega_h} \le (\mathcal{L}_h\boldsymbol{e},\boldsymbol{e})_{\Omega_h}$。

Theorem 4.1.1. 稳定差分法一致性与收敛性的关系 对于线性问题 $(*)$，如果其有限差分策略 $(**)$ 是稳定且一致的，则其一定是收敛的，且收敛阶数 q 不小于一致阶数 p。

Proof theorem 4.1.3 定义全局误差为离散函数：$\boldsymbol{e}=\boldsymbol{u}-\tilde{u}$，显然有

$$\mathcal{L}_h\boldsymbol{e} = \mathcal{L}_h\left(\boldsymbol{u}-\tilde{u}\right) = \mathcal{L}_h\boldsymbol{u} - f_h = \varphi_{\Omega_h}$$

$$\mathcal{B}_h\boldsymbol{e} = \mathcal{B}_h\left(\boldsymbol{u}-\tilde{u}\right) = \mathcal{B}_h\boldsymbol{u} - g_h = \varphi_{\partial\Omega_h}$$

由一致性，$||\varphi_{\Omega_h}||_{\Omega_h} + ||\varphi_{\partial\Omega_h}||_{\partial\Omega_h} \sim O\left(h^p\right)$ as $h\to 0$。再由稳定性

$$||\boldsymbol{u}-\tilde{u}||'_{\Omega_h} = ||\boldsymbol{e}||'_{\Omega_h} \le C\left(\ ||\varphi_{\Omega_h}||_{\Omega_h} + ||\varphi_{\partial\Omega_h}||_{\partial\Omega_h}\ \right) \le C'h^p$$

即至少有 $||\boldsymbol{u}-\tilde{u}||'_{\Omega_h} \sim O\left(h^p\right)$ as $h\to 0$ 依 p 阶收敛。

4.2 Scheme Formulation

在本章中我们具体讨论如下的二元椭圆 BVP，边界条件齐次，求解区域为方形区域 $\Omega = [0,1] \times [0,1]$，$c(x,y)$ 在 $\overline{\Omega}$ 上连续且 $c(x,y) \geq 0$ 恒成立，求解问题定义如下 $(*)$

$$\begin{aligned} -u_{xx} - u_{yy} + c(x,y)\,u &= f(x,y) \quad \text{if } (x,y) \in \Omega \\ u(x,y) &= 0 \qquad\quad \text{if } (x,y) \in \partial\Omega \end{aligned}$$

在随后的两节，我们将就 $f(x,y)$ 的连续性分两种情况讨论差分法的稳定性、一致性和收敛性。第一种情况，$f \in C(\overline{\Omega})$，这种情况下的分析模式与前一章中的类似；后一种情况我们只考虑 $f \in L_2(\Omega)$，此时 BVP 只存在弱解；且在进行一致性分析时，Taylor 展开的使用也会受限。届时我们会引入并介绍另一种分析方法。

在这些分析中，我们默认使用以下的记号与策略，对于离散化网格我们有

$$\overline{\Omega}_h = \{\ (x_i, y_j)\ :\ x_i = ih,\ y_j = jh,\ i,j = 0 \cdots N\ \}$$

$$\Omega_h = \{\ (x_i, y_j)\ :\ x_i = ih,\ y_j = jh,\ i,j = 1 \cdots N-1\ \}$$

考虑如下的差分策略 five-point difference scheme（因为计算一个点的算符 $\mathcal{L}u$ 近似会用到附近五个点的函数值）

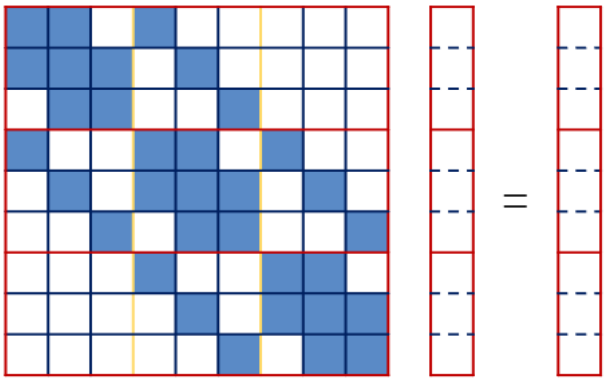
$$\begin{aligned} -D^c_{xx}\tilde{u}_{i,j} - D^c_{yy}\tilde{u}_{i,j} + c_{i,j}\tilde{u}_{i,j} &= f_{i,j} \quad \text{if } (x_i,y_j) \in \Omega_h \\ \tilde{u}_{i,j} &= 0 \qquad \text{if } (x_i,y_j) \in \partial\Omega_h \end{aligned}$$

我们可以通过将各行取出后首尾连接将二维下标 (i,j) 一维化，即规定 (Lexicographic order)

$$\begin{aligned} \tilde{u} &= \begin{bmatrix} \tilde{u}^T_{1.} & \cdots & \tilde{u}^T_{N-1.} \end{bmatrix}^T \\ &= \begin{bmatrix} \tilde{u}_{11} & \cdots & \tilde{u}_{1,N-1} & \tilde{u}_{21} & \cdots & \tilde{u}_{2,N-1} & \cdots & \tilde{u}_{N-1,1} & \cdots & \tilde{u}_{N-1,N-1} \end{bmatrix}^T \end{aligned}$$

$$\begin{aligned} f_h &= \begin{bmatrix} f^T_{1.} & \cdots & f^T_{N-1.} \end{bmatrix}^T \\ &= \begin{bmatrix} f_{11} & \cdots & f_{1,N-1} & f_{21} & \cdots & f_{2,N-1} & \cdots & f_{N-1,1} & \cdots & f_{N-1,N-1} \end{bmatrix}^T \end{aligned}$$

于是可以将 BVP 转换为求解线性方程组 $\mathcal{L}_h\tilde{u} = f_h$，为了保证前后文的连贯性，我们以后默认用 $\mathcal{L}_h$ 表示转化后的 $(N-1)^2 \times (N-1)^2$ 系数矩阵 A。对于五点策略，系数矩阵的大部分行（除了对应了靠近边界的点）应该只有五个非 0 元素（用于计算对应下标点的 $\mathcal{L}u$ 近似），因此其应当是一个稀疏矩阵；一般来说，其还应当是一个带状矩阵。下面是问题 $(*)$ 对应的系数矩阵 $\mathcal{L}_h$，我们假设其形状为 9×9 (i.e, $N = 4$)。



4.3 Solution Behaviors for Continuous Force Functions

类似对一元问题的讨论，当施迫函数 $f \in C(\overline{\Omega})$ 时，我们对节点上的离散函数做如下的范数定义。首先，有在内部节点上 Ω_h 的内积

$$(\boldsymbol{w}, \boldsymbol{v})_h = \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} h^2 w_{ij} v_{ij}$$

分别含行（下）边界与列（右）边界的内积：

$$(\boldsymbol{w}, \boldsymbol{v}]_{x,h} = \sum_{i=1}^{N} \sum_{j=1}^{N-1} h^2 w_{ij} v_{ij} \quad \& \quad (\boldsymbol{w}, \boldsymbol{v}]_{y,h} = \sum_{i=1}^{N-1} \sum_{j=1}^{N} h^2 w_{ij} v_{ij}$$

其次，有范数

$$||\boldsymbol{v}||_h = \sqrt{(\boldsymbol{v}, \boldsymbol{v})_h} = \sqrt{\textstyle\sum_{i=1}^{N-1} \sum_{j=1}^{N-1} h^2 v_{ij} v_{ij}}$$

$$||\boldsymbol{v}]|_{x,h} = \sqrt{(\boldsymbol{v}, \boldsymbol{v}]_{x,h}} \quad \& \quad ||\boldsymbol{v}]|_{y,h} = \sqrt{(\boldsymbol{v}, \boldsymbol{v}]_{y,h}}$$

以及最后，二维离散 Sobolev 范数：

$$||\boldsymbol{v}||_{1,h} = \left(||\boldsymbol{v}||_h^2 + ||D_x^- \boldsymbol{v}]|_{x,h}^2 + ||D_y^- \boldsymbol{v}]|_{y,h}^2 \right)^{\frac{1}{2}}$$

下面和一维情境中类似的，介绍二维离散形式的“分部积分”公式以及 Poincare-Friedrichs Inequality。

Lemma 4.3.1. 一元情形的 *summation by parts* 公式也可以很自然地推广到二元：假设 $\boldsymbol{v}$ 是定义在 *4.2* 节网格点 $\overline{\Omega}_h$ 上的函数，且 *test function* $\boldsymbol{\phi}$ 满足齐次边界条件 $\boldsymbol{\phi} = 0$ on $\partial\Omega_h$，则由 *lemma 3.2.1* 显然可得：

$$-\left(D_x^+\boldsymbol{v}, \boldsymbol{\phi}\right)_h = \left(\boldsymbol{v}, D_x^-\boldsymbol{\phi}\right]_{x,h} \qquad \& \qquad -\left(D_y^+\boldsymbol{v}, \boldsymbol{\phi}\right)_h = \left(\boldsymbol{v}, D_y^-\boldsymbol{\phi}\right]_{x,h}$$

对于二阶导数，注意有：$D_{xx}^c = D_x^+ D_x^-$ 与 $D_{yy}^c = D_y^+ D_y^-$。

Lemma 4.3.2. Discrete Poincare-Friedrichs Inequality (DPFI) 令 $\boldsymbol{v}$ 为定义在 *4.2* 节网格点 $\overline{\Omega}_h$ 上的函数，且满足齐次边界条件 $\boldsymbol{v} = 0$ on $\partial\Omega_h$；则存在一个与 $\boldsymbol{v}$, h 均无关的常数 $c^* = 1/4$ 使得

$$||\boldsymbol{v}||_h^2 \leq \frac{1}{4}\left(\, ||D_x^-\boldsymbol{v}]|_{x,h}^2 + ||D_y^-\boldsymbol{v}]|_{y,h}^2 \, \right)$$

证明见利用了一维形式的 lemma 3.3.1，给出如下。

Proof lemma 4.3.2 固定每一行 (列)，对列使用一维形式的 DPFI (lemma 3.3.1)，分别有

$$\forall\, i = 1 \cdots (N-1) \;:\; \textstyle\sum_{j=1}^{N-1} h|\boldsymbol{v}_{ij}|^2 \leq \frac{1}{2}\sum_{j=1}^{N} h|D_{yy}^-\boldsymbol{v}_{ij}|^2$$

$$\forall\, j = 1 \cdots (N-1) \;:\; \textstyle\sum_{i=1}^{N-1} h|\boldsymbol{v}_{ij}|^2 \leq \frac{1}{2}\sum_{i=1}^{N} h|D_{xx}^-\boldsymbol{v}_{ij}|^2$$

两式分别两端乘上 h，并对另一个坐标从 $1 \sim (N-1)$ 求和之后两式相加后两边同时除以 2 即可得到结果。

$$2 \cdot ||\boldsymbol{v}||_h^2 \leq \frac{1}{2}\left(\, ||D_x^-\boldsymbol{v}]|_{x,h}^2 + ||D_y^-\boldsymbol{v}]|_{y,h}^2 \, \right)$$

Lemma 4.3.3. 令 $\boldsymbol{v}$ 为定义在 *4.2* 节网格点 $\overline{\Omega}_h$ 上的函数，且满足齐次边界条件 $\boldsymbol{v} = 0$ on $\partial\Omega_h$；则，对于边界值问题 $(*)$ 有

$$4||\boldsymbol{v}||_h^2 \leq (\mathcal{L}_h\boldsymbol{v}, \boldsymbol{v})_h \quad \& \quad \frac{4}{5}||\boldsymbol{v}||_{1,h}^2 \leq (\mathcal{L}_h\boldsymbol{v}, \boldsymbol{v})_h$$

第一个式子不难证明 $(\mathcal{L}_h\boldsymbol{v}, \boldsymbol{v})_h \geq \; ||D_x^-\boldsymbol{v}]|_{x,h}^2 + ||D_y^-\boldsymbol{v}]|_{y,h}^2$，结合 *DPFI* 显然可得。第二个式子直接将新得到的不等式与上式相加即可。特别地，对于由差分法得到的解 $\tilde{u}$ 有 (利用 $||\cdot||_h$ 的 *Cauchy-Schwarz*，证明见 Lecture Notes p.26)

$$||\tilde{u}||_{1,h} \leq \frac{5}{4}||f_h||_h$$

现在，我们可以开始讨论对于连续施迫函数 f，差分法得到数值解的各种性质了，见下页讨论。

Existence, Uniqueness, Stability and Consistency of the Solution
讨论与一维情形完全类似，我们只做简单论述

1. **Existence and Uniqueness** 如果存在 $\mathcal{L}_h\boldsymbol{v}=0$，则由 lemma 4.3.3 $||\boldsymbol{v}||_{1,h}^2 \le \frac{5}{4}(\mathcal{L}_h\boldsymbol{v},\boldsymbol{v})_h = 0 \Rightarrow ||\boldsymbol{v}||_{1,h}=0 \Rightarrow \boldsymbol{v}=\boldsymbol{0}$，所以 $\mathcal{L}_h$ 可逆，解存在且唯一。

2. **Stability** 由 lemma 4.3.3 $||\tilde{u}||_{1,h} \le \frac{5}{4}||f_h||_h$ 已经证明了稳定性，我们补充一下这一步证明的细节，利用了范数的柯西不等式：

$$||\tilde{u}||_{1,h}^2 \le \frac{5}{4}(\mathcal{L}_h\tilde{u},\tilde{u})_h \le \frac{5}{4}(f_h,\tilde{u})_h \le \frac{5}{4}||f_h||_h||\tilde{u}||_h \Rightarrow ||\tilde{u}||_{1,h} \le \frac{5}{4}||f_h||_h$$

3. **Consistency** 对于本问题的齐次边界条件，只要证明 $||\mathcal{L}_h\boldsymbol{u}-f_h||_h$ 随 $h\to 0$ 收敛即可。有

$$\begin{aligned}\mathcal{L}_h\boldsymbol{u}-f_h &= \left(-D_{xx}^c\boldsymbol{u}-D_{yy}^c\boldsymbol{u}+c_h\boldsymbol{u}\right)-\left(-\boldsymbol{u}_{xx}-\boldsymbol{u}_{yy}\tilde{u}+c_h\boldsymbol{u}\right)\\ &= \left(\boldsymbol{u}_{xx}-D_{xx}^c\boldsymbol{u}\right)+\left(\boldsymbol{u}_{yy}-D_{yy}^c\boldsymbol{u}\right)=\varphi_h\end{aligned}$$

又有之前证明的

$$\left|\partial_x^2 u_{ij}-D_{xx}^c u_{ij}\right| \le \frac{h^2}{12}||\partial_x^4 u||_{C(\Omega)} = C_x h^2$$

$$\left|\partial_y^2 u_{ij}-D_{yy}^c u_{ij}\right| \le \frac{h^2}{12}||\partial_y^4 u||_{C(\Omega)} = C_y h^2$$

于是 $|\varphi_{ij}| = \left|\partial_x^2 u_{ij}-D_{xx}^c u_{ij}+\partial_y^2 u_{ij}-D_{yy}^c u_{ij}\right| \le \left(C_x+C_y\right)h^2 = C'h^2$，即有

$$\begin{aligned}||\varphi_h||_h = ||\mathcal{L}_h\boldsymbol{u}-f_h||_h &= \sqrt{\textstyle\sum_{i=1}^{N-1}\sum_{j=1}^{N-1}h^2\varphi_{ij}^2} = \sqrt{\textstyle\sum_{i=1}^{N-1}\sum_{j=1}^{N-1}h^2C'^2h^4}\\ &\le \sqrt{(Nh)^2C'^2h^4} = C'h^2 = \left(C_x+C_y\right)h^2 \sim O(h^2)\end{aligned}$$

4. **Convergence** 记 $\boldsymbol{e}=\boldsymbol{u}-\tilde{u}$，不难证明有 $\mathcal{L}_h\boldsymbol{e}=\varphi_h$ 且满足齐次边界条件，由 lemma 4.3.3 有

$$||\boldsymbol{e}||_{1,h} \le \frac{5}{4}||\varphi_h||_h \le \frac{5h^2}{48}\left(C_x+C_y\right) \sim O(h^2)$$

Theorem 4.3.1. *对于边界值问题 $(*)$，若 $u\in C^4(\Omega)$，则有限差分法的误差为*

$$||\tilde{u}-\boldsymbol{u}||_{1,h} \le \frac{5}{48}h^2\left(\,||\partial_x^4 u||_{C(\Omega)}+||\partial_y^4 u||_{C(\Omega)}\,\right)$$

4.4 Solution Behaviors for Discontinuous Force Functions

下面我们讨论当施迫函数 $f \in L_2(\Omega)$ 而不一定为连续函数时的情况。依旧考虑在正方形区域 $\Omega = [0,1] \times [0,1]$ 上的 BVP $-\triangle u + cu = f$ s.t. $u = 0$ if $(x,y) \in \partial\Omega$。对于网格 h 内的每一个节点 $\left(x_i, y_j\right)$，考虑如下的区域

$$K_{ij} = \left[x_i - \frac{h}{2}, x_i + \frac{h}{2}\right] \times \left[y_j - \frac{h}{2}, y_j + \frac{h}{2}\right]$$

为方便讨论，我们记 $x_i \pm \frac{h}{2} = x_{i\pm\frac{1}{2}}$。现在对于原方程在该区域上积分

$$-\int_{K_{ij}} \triangle u \; dxdy \; + \; \int_{K_{ij}} cu \; dxdy \; = \; \int_{K_{ij}} f(x,y)\, dxdy \tag{1}$$

$$-\int_{K_{ij}} \nabla \cdot (\nabla u) \; dxdy \; + \; \int_{K_{ij}} cu \; dxdy \; = \; \int_{K_{ij}} f(x,y)\, dxdy \tag{2}$$

$$-\int_{\partial K_{ij}} \frac{\partial u}{\partial v} \; dl \; + \; \int_{K_{ij}} cu \; dxdy \; = \; \int_{K_{ij}} f(x,y)\, dxdy \tag{3}$$

其中（2）至（3）利用了散度定理；由于边界是二维曲线，所以散度定理中的“面积分”化为了线积分。我们进一步对 K_{ij} 及其边界上的积分做如下的近似，首先在 ∂K_{ij} 的四条边界上（分别记为 $\partial K_{ij}^1 \sim \partial K_{ij}^4$）有（利用 step $= h/2$ 的中心差分）

$$\int_{\partial K_{ij}^1} \frac{\partial u}{\partial v} \; dl = \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \langle u_x, u_y \rangle \cdot \langle 0, 1 \rangle \; dx \approx D_y^c u_{i,j+\frac{1}{2}} \cdot h = u_{i,j+1} - u_{i,j}$$

$$\int_{\partial K_{ij}^2} \frac{\partial u}{\partial v} \; dl = \int_{y_{i-\frac{1}{2}}}^{y_{i+\frac{1}{2}}} \langle u_x, u_y \rangle \cdot \langle 1, 0 \rangle \; dy \approx D_x^c u_{i+\frac{1}{2},j} \cdot h = u_{i+1,j} - u_{i,j}$$

$$\int_{\partial K_{ij}^3} \frac{\partial u}{\partial v} \; dl = \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \langle u_x, u_y \rangle \cdot \langle 0, -1 \rangle \; dx \approx -D_y^c u_{i,j-\frac{1}{2}} \cdot h = u_{i,j-1} - u_{i,j}$$

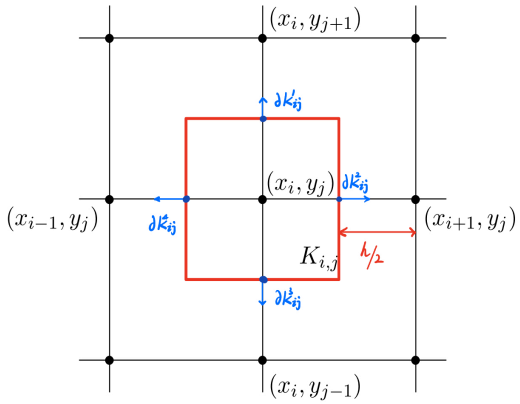
$$\int_{\partial K_{ij}^4} \frac{\partial u}{\partial v} \; dl = \int_{y_{i-\frac{1}{2}}}^{y_{i+\frac{1}{2}}} \langle u_x, u_y \rangle \cdot \langle -1, 0 \rangle \; dy \approx -D_x^c u_{i-\frac{1}{2},j} \cdot h = u_{i-1,j} - u_{i,j}$$

$$\int_{\partial K_{ij}} \frac{\partial u}{\partial v} \; dl \approx u_{i,j+1} - u_{i,j} + u_{i+1,j} - u_{i,j} + u_{i,j-1} - u_{i,j} + u_{i-1,j} - u_{i,j}$$

$$= \left(u_{i+1,j} - 2u_{i,j} + u_{i,j-1}\right) + \left(u_{i,j+1} - 2u_{i,j} + u_{i-1,j}\right) = h^2 D_{xx}^c u_{ij} + h^2 D_{yy}^c u_{ij}$$

接着用 $\int_{K_{ij}} cu \; dxdy \approx h^2 c_{ij} u_{ij}$，整合上式我们得到

$$-D_{xx}^c \tilde{u}_{ij} - D_{yy}^c \tilde{u}_{ij} + c_{ij} \tilde{u}_{ij} = \frac{1}{h^2} \int_{K_{ij}} f(x,y)\, dxdy := T f_{ij}$$



上面这种通过对目标方程积分生成的有限差分策略称为 finite volume method，现在从头整合一下我们的问题和求解策略

$$\begin{aligned} -u_{xx} - u_{yy} + c(x,y)\,u = f(x,y) \quad & \text{if } (x,y) \in \Omega \\ u(x,y) = 0 \quad & \text{if } (x,y) \in \partial\Omega \end{aligned}$$

其中 $f \in L_2(\Omega)$，$c(x,y)$ 在 $\overline{\Omega}$ 上连续且 $c(x,y) \geq 0$ 恒成立；求解区域为正方形区域 $\Omega = [0,1] \times [0,1]$。我们的求解策略 $(*)$ 为

$$\begin{aligned} -D^c_{xx}\tilde{u}_{ij} - D^c_{yy}\tilde{u}_{ij} + c_{ij}\tilde{u}_{ij} = Tf_{ij} \quad & \text{if } (x_i, y_j) \in \Omega_h \\ \tilde{u}_{ij} = 0 \quad & \text{if } (x_i, y_j) \in \partial\Omega_h \end{aligned}$$

其中 $Tf_{ij} = \frac{1}{h^2}\int_{K_{ij}} f(x,y)\,dxdy$，我们可以进一步把求解策略写成更紧凑的形式

$$\mathcal{L}_h\tilde{u} = -D^c_{xx}\tilde{u} - D^c_{yy}\tilde{u} + c_h\tilde{u} = Tf$$

注意，该策略的差分算符和之前的一致，即变换策略没有改变 $\mathcal{L}_h$ 的可逆性，所以该策略的解也存在且唯一。

Theorem 4.4.1. 积分策略的稳定性 对于区域 $s.t.\ \bigcup_{ij} K_{ij} \subseteq \Omega$，有 $||Tf||_h \leq ||f||_{L_2(\Omega)}$，于是策略 $(*)$ 是稳定的，对 $c^* = \frac{5}{4}$，有

$$||\tilde{u}||_{1,h} \leq c^*||Tf||_h \leq c^*||f||_{L_2(\Omega)}$$

Proof theorem 4.4.1 先证第一个不等式

$$||Tf||_h^2 = \sum_{i=1}^{N-1}\sum_{j=1}^{N-1} h^2 \cdot \frac{1}{h^4}(f,1)^2_{L_2(K_{ij})} \le \sum_{i=1}^{N-1}\sum_{j=1}^{N-1} \frac{1}{h^2}||f||^2_{L_2(K_{ij})} \cdot ||1||^2_{L_2(K_{ij})}$$
$$= \sum_{i=1}^{N-1}\sum_{j=1}^{N-1} ||f||^2_{L_2(K_{ij})} \le ||f||^2_{L_2(\Omega)}$$

第二个不等式，因为我们的差分算子较之前并无变化，由 lemma 4.3.3 $||\tilde{u}||_{1,h} \le \frac{5}{4}||Tf||_h \le \frac{5}{4}||f||_{L_2(\Omega)}$。

上述定理说明了积分策略的稳定性；对于其收敛性，我们有如下定理

Theorem 4.4.2. 积分策略的收敛性
上述积分策略的全局误差 $\boldsymbol{e} = \boldsymbol{u} - \tilde{u}$ 符合下述的不等式，记 $c^* = \frac{5}{4}$

$$||\boldsymbol{e}||_{1,h} \le c^* \left(||\varphi_1]|^2_{x,h} + ||\varphi_2]|^2_{y,h} + ||\psi||^2_h \right)^{1/2}$$

其中 for $i = 1 \sim N$, $j = 1 \sim (N-1)$

$$\varphi_1(x_i, y_j) = \frac{1}{h}\int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} \frac{\partial}{\partial x} u(x_{i-\frac{1}{2}}\ ,\ y)\ dy - D_x^- u_{ij}$$

for $i = 1 \sim (N-1)$, $j = 1 \sim N$

$$\varphi_2(x_i, y_j) = \frac{1}{h}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \frac{\partial}{\partial y} u(x\ ,\ y_{j-\frac{1}{2}})\ dx - D_y^- u_{ij}$$

for $i, j = 1 \sim (N-1)$

$$\psi(x_i, y_j) = c_{ij}u_{ij} - \frac{1}{h^2}\int_{K_{ij}} c(x,y)\cdot u(x,y)\ dxdy$$

此时有 $\mathcal{L}_h \boldsymbol{e} = D_x^+ \boldsymbol{\varphi_1} + D_y^+ \boldsymbol{\varphi_2} + \boldsymbol{\psi}$。

Proof theorem 4.4.2 首先考虑将算子 $\mathcal{L}_h$ 作用于 $\boldsymbol{e}$

$$\begin{aligned}\mathcal{L}_h e_{ij} &= \mathcal{L}_h u_{ij} - \mathcal{L}\tilde{u}_{ij} = -D^c_{xx}u_{ij} - D^c_{yy}u_{ij} + c_h u_{ij} - Tf_{ij}\\ &= -D^c_{xx}u_{ij} - D^c_{yy}u_{ij} + c_{ij}u_{ij} + \frac{1}{h^2}\int_{K_{ij}} u_{xx} + u_{yy} - cu\ dxdy\\ &= -D^c_{xx}u_{ij} - D^c_{yy}u_{ij} + \frac{1}{h^2}\int_{K_{ij}} u_{xx}\ dxdy + \frac{1}{h^2}\int_{K_{ij}} u_{yy}\ dxdy + \psi_{ij}\\ &= -D^+_x D^-_x u_{ij} - D^+_y D^-_y u_{ij} + \frac{1}{h}\int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{h}}} D^+_x u_x(x_{i-\frac{1}{2}}, y)\ dy\\ &\quad + \frac{1}{h}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} D^+_y u_y(x, y_{j-\frac{1}{2}})\ dx + \psi_{ij} = (D^+_x\varphi_1)_{ij} + (D^+_y\varphi_2)_{ij} + \psi_{ij}\end{aligned}$$

于是有：$\mathcal{L}_h\boldsymbol{e} = D^+_x\boldsymbol{\varphi_1} + D^+_y\boldsymbol{\varphi_2} + \boldsymbol{\psi}$ 且 $\boldsymbol{e} = 0$ on $\partial\Omega$。运用 lemma 4.3.3 以及 summation by parts 有

$$\begin{aligned}||\boldsymbol{e}||^2_{1,h} &\le \frac{5}{4}(\mathcal{L}_h\boldsymbol{e}, \boldsymbol{e})_h = \frac{5}{4}\left(D^+_x\boldsymbol{\varphi_1} + D^+_y\boldsymbol{\varphi_2} + \boldsymbol{\psi}, \boldsymbol{e}\right)_h\\ &= \frac{5}{4}\left\{\left(D^+_x\boldsymbol{\varphi_1}, \boldsymbol{e}\right)_h + \left(D^+_y\boldsymbol{\varphi_2}, \boldsymbol{e}\right)_h + (\boldsymbol{\psi}, \boldsymbol{e})_h\right\}\\ &= \frac{5}{4}\left\{-\left(\boldsymbol{\varphi_1}, D^-_x\boldsymbol{e}\right]_{x,h} - \left(\boldsymbol{\varphi_2}, D^-_y\boldsymbol{e}\right]_{y,h} + (\boldsymbol{\psi}, \boldsymbol{e})_h\right\}\\ &\le \frac{5}{4}\left\{||\boldsymbol{\varphi_1}]|_{x,h}\cdot||D^-_x\boldsymbol{e}]|_{x,h} + ||\boldsymbol{\varphi_2}]|_{y,h}\cdot||D^-_y\boldsymbol{e}]|_{y,h} + ||\boldsymbol{\psi}||_h||\boldsymbol{e}||_h\right\}\\ &\le \frac{5}{4}\left(||\varphi_1]|^2_{x,h} + ||\varphi_2]|^2_{y,h} + ||\psi||^2_h\right)^{1/2}\cdot\left(||D^-_x\boldsymbol{e}]|^2_{x,h} + ||D^-_y\boldsymbol{e}]|^2_{y,h} + ||\boldsymbol{e}||^2_h\right)^{1/2}\\ &= \frac{5}{4}\left(||\varphi_1]|^2_{x,h} + ||\varphi_2]|^2_{y,h} + ||\psi||^2_h\right)^{1/2}\cdot||\boldsymbol{e}||_{1,h}\end{aligned}$$

其中倒数第二行利用了 Cauchy-Schwarz 不等式。

Theorem 4.4.3. 积分策略的误差估计
当 $f \in L_2(\Omega), c \in C^2(\overline{\Omega})$ 且 $c(x,y) \ge 0$ 在 $\overline{\Omega}$ 上恒成立，且弱解符合 $u \in C^3(\overline{\Omega})$ 时，上述积分策略的全局误差 $\boldsymbol{e} = \boldsymbol{u} - \tilde{u}$ 符合

$$||\boldsymbol{e}||_{1,h} \le \frac{5}{96}h^2 S M_3$$

其中 S 为方形区域 Ω 的面积，M_3 符合

$$\begin{aligned}M_3 = \Big\{&\left(||\partial_{xyy}u||_{C(\overline{\Omega})} + ||\partial^3_x u||_{C(\overline{\Omega})}\right)^2 + \left(||\partial_{xxy}u||_{C(\overline{\Omega})} + ||\partial^3_y u||_{C(\overline{\Omega})}\right)^2\\ &+ \left(||\partial^2_x(cu)||_{C(\overline{\Omega})} + ||\partial^2_y(cu)||_{C(\overline{\Omega})}\right)^2\Big\}^{1/2}\end{aligned}$$

Proof theorem 4.4.3 利用泰勒展开可以证明

$$
\begin{aligned}
|\varphi_1(x_i,y_j)| &\le \frac{h^2}{24}\left(||\partial_{xyy}u||_{C(\overline{\Omega})}+||\partial_x^3 u||_{C(\overline{\Omega})}\right)\\
|\varphi_2(x_i,y_j)| &\le \frac{h^2}{24}\left(||\partial_{xxy}u||_{C(\overline{\Omega})}+||\partial_y^3 u||_{C(\overline{\Omega})}\right)\\
|\psi_1(x_i,y_j)| &\le \frac{h^2}{24}\left(||\partial_x^2(cu)||_{C(\overline{\Omega})}+||\partial_y^2(cu)||_{C(\overline{\Omega})}\right)
\end{aligned}
$$

我们仅演示对第一个不等式的证明，其他的类似

$$
u\left(x_i-\frac{h}{2}\pm\frac{h}{2}\ ,\ y_j\right)=u_{i-\frac12,j}\pm\frac{h}{2}\partial_x u_{i-\frac12,j}+\frac{h^2}{8}\partial_x^2 u_{i-\frac12,j}\pm\frac{h^3}{48}\partial_x^3 u_{\xi_\pm,j}
$$

于是有

$$
\begin{aligned}
D_x^- u(x_i,y_j)&=\frac1h\left[u\left(x_i-\frac{h}{2}+\frac{h}{2}\ ,\ y_j\right)-u\left(x_i-\frac{h}{2}-\frac{h}{2}\ ,\ y_j\right)\right]\\
&=\partial_x u_{i-\frac12,j}+\left(\frac{h^2}{48}\partial_x^3 u_{\xi_+,j}+\frac{h^2}{48}\partial_x^3 u_{\xi_-,j}\right)\le \partial_x u_{i-\frac12,j}+\frac{h^2}{24}||\partial_x^3 u||_{C(\overline{\Omega})}
\end{aligned}
$$

又有对 $\partial_x u(x_i-\frac{h}{2},y)$ 在 y_j 处展开

$$
\partial_x u(x_{i-\frac12}\ ,\ y)=\partial_x u_{i-\frac12,j}+\partial_{xy}u_{i-\frac12,j}(y-y_j)+\frac12\partial_{xyy}u(x_{i-\frac12}\ ,\ \xi(y))(y-y_j)^2
$$

带入 $\varphi_1(x_i,y_j)$ 中的积分项后有

$$
\begin{aligned}
&\frac1h\int_{y_{j-\frac12}}^{y_{j+\frac12}}\frac{\partial}{\partial x}u(x_{i-\frac12}\ ,\ y)\,dy\\
=&\frac1h\int_{y_{j-\frac12}}^{y_{j+\frac12}}\left\{\partial_x u_{i-\frac12,j}+\partial_{xy}u_{i-\frac12,j}(y-y_j)+\frac12\partial_{xyy}u(x_{i-\frac12}\ ,\ \xi(y))(y-y_j)^2\right\}dy\\
=&\partial_x u_{i-\frac12,j}+0+\frac{1}{2h}\int_{y_{j-\frac12}}^{y_{j+\frac12}}\partial_{xyy}u(x_{i-\frac12}\ ,\ \xi(y))(y-y_j)^2dy\\
\le&\partial_x u_{i-\frac12,j}+\frac{1}{2h}||\partial_{xyy}u||_{C(\overline{\Omega})}\int_{y_{j-\frac12}}^{y_{j+\frac12}}(y-y_j)^2dy=\partial_x u_{i-\frac12,j}+\frac{h^2}{24}||\partial_{xyy}u||_{C(\overline{\Omega})}
\end{aligned}
$$

最后由三角不等式

$$
|\varphi_1(x_i,y_j)|\le\frac{h^2}{24}(||\partial_{xyy}u||_{C(\overline{\Omega})}+||\partial_x^3 u||_{C(\overline{\Omega})})
$$

利用上述不等式，我们可以进一步得到

$$\begin{aligned}
||\varphi_1]|_{x,h}^2 &= \sum_{i=1}^{N}\sum_{j=1}^{N-1} h^2|\varphi_1(x_i,y_j)|^2 \\
&\le \sum_{i=1}^{N}\sum_{j=1}^{N} h^2\frac{h^4}{24^2}(||\partial_{xyy}u||_{C(\overline{\Omega})}+||\partial_x^3u||_{C(\overline{\Omega})})^2 \\
&= \frac{h^4}{24^2}(||\partial_{xyy}u||_{C(\overline{\Omega})}+||\partial_x^3u||_{C(\overline{\Omega})})^2\cdot h^2N^2 \\
&= \left\{\frac{h^2}{24}(||\partial_{xyy}u||_{C(\overline{\Omega})}+||\partial_x^3u||_{C(\overline{\Omega})})\cdot S\right\}^2
\end{aligned}$$

同理，有

$$||\varphi_2]|_{y,h}^2 \le \left\{\frac{h^2}{24}\left(||\partial_{xxy}u||_{C(\overline{\Omega})}+||\partial_y^3u||_{C(\overline{\Omega})}\right)\cdot S\right\}^2$$

$$||\psi_h||^2 \le \left\{\frac{h^2}{24}\left(||\partial_x^2(cu)||_{C(\overline{\Omega})}+||\partial_y^2(cu)||_{C(\overline{\Omega})}\right)\cdot S\right\}^2$$

由 theorem 4.4.2

$$\begin{aligned}
||\boldsymbol{e}||_{1,h} &\le c^*\left(||\varphi_1]|_{x,h}^2+||\varphi_2]|_{y,h}^2+||\psi||_h^2\right)^{1/2} \\
&\le \frac{5h^2}{4\times 24}SM_3=\frac{5}{96}h^2SM_3
\end{aligned}$$

实际上，我们可以进一步放宽条件，只要求 $u\in H^3(\Omega)$，有如下的定理

Theorem 4.4.4. 积分策略的最佳误差估计
当 $f\in L_2(\Omega), c\in C^2(\overline{\Omega})$ 且 $c(x,y)\ge 0$ 在 $\overline{\Omega}$ 上恒成立，且弱解符合 $u\in H^3(\Omega)$ 时，上述积分策略的全局误差 $\boldsymbol{e}=\boldsymbol{u}-\tilde{u}$ 符合

$$||\boldsymbol{e}||_{1,h}\le Ch^2||u||_{H^3\Omega}$$

再进一步弱化方程中对 u 的 *regularity hypothesis* 将会导致收敛阶数的下降，因此从对解 u 的连续性假设这一角度来看，这种误差估计是使我们在保持最高阶收敛性时能达到的最优误差估计 *optimal error estimates*。该定理的推导详见**Lecture Notes pp.32-34**。

4.5 Discretization for Generalized Elliptic BVPs and Nonaxiparallel Domains with Nonuniform Meshes

我们简单给出上文 4.3 至 4.4 中对于 BVP $-\triangle u + cu = f \ (*)$ 讨论的两种扩展。一种是保持 $\Omega \in \mathbb{R}^2$ 为方形区域，但对讨论的方程做一个简单的延伸；另一种是延续所求方程 $-\triangle u + cu = f$ 的形式，但考虑 $\Omega \in \mathbb{R}^2$ 为一般区域且离散化节点非均匀分布时的情形。

Discretization for Generalized Elliptic BVPs

我们首先还是考虑 $\Omega \in \mathbb{R}^2$ 为方形区域的情形，但此时目标问题变为解如下的二元椭圆 BVP

$$-\frac{\partial}{\partial x}\left(a^1\frac{\partial u}{\partial x}\right) - \frac{\partial}{\partial y}\left(a^2\frac{\partial u}{\partial y}\right) + b^1\frac{\partial u}{\partial x} + b^2\frac{\partial u}{\partial y} + cu = f \qquad (**)$$

其中：$a^1, a^2 \in C^1(\bar{\Omega})$；$b^1, b^2, c \in C(\bar{\Omega})$，注意上角标仅表示 index。此时，使用如下的二阶差分策略

Theorem 4.5.1. 一般椭圆边界值问题的离散化 当求解区域 $\Omega \in \mathbb{R}^n$ 为方形区域时，对问题 $(**)$ 可采用如下的离散化策略，其有 *2nd order consistency*

$$-D_x^+\left(a^1_{i-\frac{1}{2},j}D_x^-\tilde{u}_{ij}\right) - D_y^+\left(a^2_{i,j-\frac{1}{2}}D_y^-\tilde{u}_{ij}\right) + b^1_{ij}D_x^c\tilde{u}_{ij} + b^2_{ij}D_y^c\tilde{u}_{ij} + c_{ij}\tilde{u}_{ij} = f_{ij}$$

当有 $f \in L_2(\Omega)$ 时，用 $\int_{K_{ij}} f(x,y)\,dxdy$ 代替上式中的 f_{ij}。该策略中对于二阶导数项的近似有 *2* 阶精度，详细证明见 Problem Sheet 1, Q5。

注意上述离散化也可以应用到含有更多自变量的类似形式的椭圆 BVP 中。

Nonaxiparallel Domains and Nonuniform Meshes

重新考虑方程 $-\triangle u + cu = f$，但此时区域 Ω 为非方形区域，因此离散化节点 $\bar{\Omega}_h$ 为非均匀节点。简单定义如下标记

$$h^-_{x,i} = x_i - x_{i-1} \quad , \quad h^+_{x,i} = x_{i+1} - x_i \quad , \quad h_{x,i} = \frac{1}{2}\left(h^-_{x,i} + h^+_{x,i}\right)$$
$$h^-_{y,j} = y_j - y_{j-1} \quad , \quad h^+_{y,j} = y_{j+1} - y_j \quad , \quad h_{y,j} = \frac{1}{2}\left(h^-_{y,j} + h^+_{y,j}\right)$$

注意上述的离散化步长在节点两侧等距时保持不变。进一步给出在此类节点下的有限差分为

Definition 4.5.1. 非均匀节点上的有限差分 若离散化节点 $\overline{\Omega}_h$ 为非均匀节点，使用如下的差分公式近似导数

$$D_x^- \tilde{u}_{ij} = \frac{1}{h_{x,i}^-}\left(\tilde{u}_{ij} - \tilde{u}_{i-1,j}\right) \quad \& \quad D_x^+ \tilde{u}_{ij} = \frac{1}{h_{x,i}}\left(\tilde{u}_{i+1,j} - \tilde{u}_{i,j}\right)$$

于是二阶导数的近似为

$$D_x^+ D_x^- \tilde{u}_{ij} = \frac{1}{h_{x,i}}\left(\frac{\tilde{u}_{i+1,j} - \tilde{u}_{i,j}}{h_{x,i+1}^-} - \frac{\tilde{u}_{ij} - \tilde{u}_{i-1,j}}{h_{x,i}^-}\right)$$

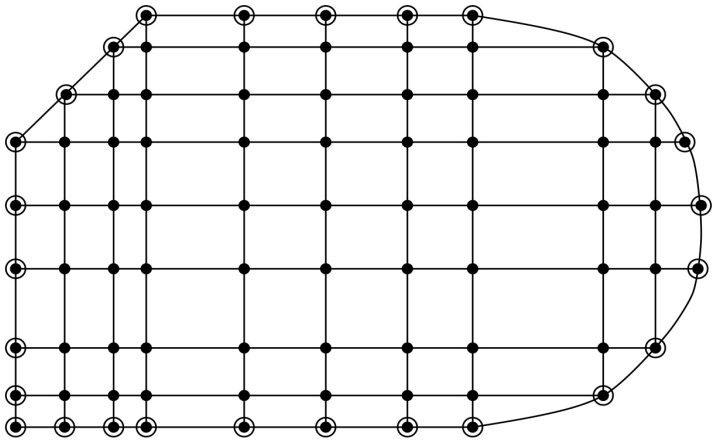
对 y 导数的近似同理，有

$$D_y^- \tilde{u}_{ij} = \frac{1}{h_{y,j}^-}\left(\tilde{u}_{ij} - \tilde{u}_{i,j-1}\right) \quad \& \quad D_y^+ \tilde{u}_{ij} = \frac{1}{h_{y,j}}\left(\tilde{u}_{i,j+1} - \tilde{u}_{i,j}\right)$$

二阶导数的近似为

$$D_y^+ D_y^- \tilde{u}_{ij} = \frac{1}{h_{y,j}}\left(\frac{\tilde{u}_{i,j+1} - \tilde{u}_{i,j}}{h_{y,j+1}^-} - \frac{\tilde{u}_{ij} - \tilde{u}_{i,j-1}}{h_{y,j}^-}\right)$$

注意此时，对于均匀节点的部分结论都不再成立，包括（以对 x 的导数为例）$D_x^- \tilde{u}_{i+1} = D_x^+ \tilde{u}_i$ 以及 $D_x^+ D_x^- \tilde{u}_i = D_x^+ D_x^- \tilde{u}_i$。



4.6 Maximum Principle

本节介绍椭圆 PDE 的一类重要性质 Maximum Principle。其指出，对于特定的一类椭圆 PDE，其典型解 classical solution $u \in C^2(\Omega) \bigcup C(\overline{\Omega})$ 只能在边界 $\partial\Omega$ 上取到最大值，除非其为一个常数函数。我们先给出这一定理在简单情形下的证明；随后，我们给出该定理在有限差分法下的离散化版本。为了方便讨论，先不做证明地给出如下引理

Lemma 4.6.1. 驻点的条件 如果 $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ 为 $C^2(\Omega)$，且对定义域的一个内点 $a \in \text{int}(\Omega)$ 有 $\nabla f(\text{a}) = 0$，记 *Hessian matrix* 在 a 处的值为 $H(f, \text{a})$，则

1) $H(f, \text{a}) \succ 0 \Rightarrow f$ 在 a 取极小值；f 在 a 取极小值 $\Rightarrow H(f, \text{a}) \succeq 0$；
2) $H(f, \text{a}) \prec 0 \Rightarrow f$ 在 a 取极大值；f 在 a 取极大值 $\Rightarrow H(f, \text{a}) \preceq 0$。

下面给出一类椭圆 PDE 的 Maximum Principle，表述如下

Theorem 4.6.1. Maximum Principle 考虑形如下式的不含 *0* 次导数的 *Elliptic PDE*，其定义在有界集合 $\Omega \in \mathbb{R}^n$ 上：

$$-\sum_{i,j=1}^{n} \frac{\partial}{\partial x_j}\left(a_{i,j}(\boldsymbol{x}) \frac{\partial u}{\partial x_i}\right) + \sum_{i=1}^{n} b_i(\boldsymbol{x}) \frac{\partial u}{\partial x_i} = f(\boldsymbol{x})$$

其中：$a_{i,j}(\boldsymbol{x}) \in C^1(\bar{\Omega})$；$b_i, f \in C(\bar{\Omega})$；且 $\forall \xi \in \mathbb{R}^n$，$\forall \boldsymbol{x} \in \bar{\Omega}$：$\xi^T A \xi \geq c^* \xi^T \xi$，常数 $c^* > 0$ 与 $\boldsymbol{x}$，ξ 无关（*uniform ellipticity*）。则典型解 $u \in C^2(\Omega) \bigcup C(\bar{\Omega})$

1) 当 $f(\boldsymbol{x}) \leq 0$ 时，u 只能在边界 $\partial\Omega$ 取到最大值，除非其为一个常数函数；
2) 当 $f(\boldsymbol{x}) \geq 0$ 时，u 只能在边界 $\partial\Omega$ 取到最小值，除非其为一个常数函数。

在边界取到极大值意味着 $\max_{\boldsymbol{x} \in \overline{\Omega}} u(\boldsymbol{x}) = \max_{\boldsymbol{x} \in \partial\Omega} u(\boldsymbol{x})$（*resp.* min）；当有 *Dirichlet* 边界条件 $u = g$ 时，边界最大值（*resp.* 边界最小值）可以化为 $\max_{\boldsymbol{x} \in \partial\Omega} g(\boldsymbol{x})$。在某些特定情况下，可以允许算子包含 *0* 次导数项 $c(\boldsymbol{x}) u(\boldsymbol{x})$，详细的讨论与证明见《*An Introduction to Maximum Principles and Symmetry in Elliptic Problems*》, *Chapter 02, by L. E. Fraenkel.*

为了证明方便，我们仅考虑形如 $-\triangle u = f$ 形式的方程。整体的证明思路于此类似，具体细节可参阅上文提到的参考材料。

Proof theorem 4.6.1 最小值的情况同理，我们只证最大值的情形。当 $f<0$ 时，假设 u 在内点 $\boldsymbol{x}_0 \in \operatorname{int}(\Omega)$ 取到最大值。则由 lemma 4.6.1 有 $\partial_{x_i} u(\boldsymbol{x}_0)=0$ 且 $\partial_{x_i}^2 u(\boldsymbol{x}_0)=e_i^T H e_i \leq 0$，于是 $f=-\triangle u \geq 0$ 与 $f<0$ 矛盾。

当 $f \leq 0$ 时，考虑构造辅助函数 auxiliary function

$$v_\varepsilon(\boldsymbol{x})=u(\boldsymbol{x})+\frac{\varepsilon}{2n}\left(x_1^2+\cdots+x_n^2\right)$$

其中 $\varepsilon>0$ 为任意正实数。不难证明 $-\triangle v_\varepsilon=-\triangle u-\varepsilon=f-\varepsilon<0$，即 $v_\varepsilon(\boldsymbol{x})$ 只能在边界 $\partial\Omega$ 上取到最大值。于是有

$$\begin{aligned}
\max_{\boldsymbol{x}\in\partial\Omega} u(\boldsymbol{x}) &= \max_{\boldsymbol{x}\in\partial\Omega}\left\{v_\varepsilon(\boldsymbol{x})-\frac{\varepsilon}{2n}\left(x_1^2+\cdots+x_n^2\right)\right\} \\
&\geq \max_{\boldsymbol{x}\in\partial\Omega} v_\varepsilon(\boldsymbol{x})-\max_{\boldsymbol{x}\in\partial\Omega}\frac{\varepsilon}{2n}\left(x_1^2+\cdots+x_n^2\right) \\
&= \max_{\boldsymbol{x}\in\overline{\Omega}} v_\varepsilon(\boldsymbol{x})-\frac{\varepsilon}{2n}\max_{\boldsymbol{x}\in\partial\Omega}\left(x_1^2+\cdots+x_n^2\right) \\
&\geq \max_{\boldsymbol{x}\in\overline{\Omega}} u(\boldsymbol{x})-\frac{\varepsilon}{2n}\max_{\boldsymbol{x}\in\partial\Omega}||\boldsymbol{x}||^2
\end{aligned}$$

最后一个等号是因为显然有 $v_\varepsilon \geq u$，取极限 $\varepsilon \to 0$，我们有

$$\max_{\boldsymbol{x}\in\partial\Omega} u(\boldsymbol{x}) \geq \max_{\boldsymbol{x}\in\overline{\Omega}} u(\boldsymbol{x})$$

因为 $\Omega \subseteq \overline{\Omega}$，显然有 $\max_{\boldsymbol{x}\in\partial\Omega} u(\boldsymbol{x}) \leq \max_{\boldsymbol{x}\in\overline{\Omega}} u(\boldsymbol{x})$，综上

$$\max_{\boldsymbol{x}\in\partial\Omega} u(\boldsymbol{x})=\max_{\boldsymbol{x}\in\overline{\Omega}} u(\boldsymbol{x})$$

下面我们给出该定理在离散情形下的对应，定理内容只涉及对 $-\triangle u=f$ 在一般有界区域 Ω 上的离散化（Nonaxiparallel Domains with Nonuniform Meshes），但相似的推导与证明也可以迁移到一般情形。

Theorem 4.6.2. Discrete Maximum Principle 考虑形如下式的离散化 *Elliptic PDE*，其定义在有界 Mesh $\overline{\Omega}_h \in \mathbb{R}^n$（possibly nonaxiparallel and nonuniform）上，记 $\boldsymbol{\kappa}$ 为多重下标：

$$-\sum_{i=1}^{n} D_{x_i}^{+} D_{x_i}^{-} \tilde{u}_{\boldsymbol{\kappa}}=f_{\boldsymbol{\kappa}}$$

1） 当 $f(\boldsymbol{x}) \leq 0$ 时，$\tilde{u}$ 只在边界 $\partial\Omega_h$ 取最大，$\max_{\boldsymbol{x}_{\boldsymbol{\kappa}}\in\overline{\Omega}_h} \tilde{u}_{\boldsymbol{\kappa}}=\max_{\boldsymbol{x}_{\boldsymbol{\kappa}}\in\partial\Omega_h} \tilde{u}_{\boldsymbol{\kappa}}$；

2） 当 $f(\boldsymbol{x}) \geq 0$ 时，$\tilde{u}$ 只在边界 $\partial\Omega_h$ 取最小，$\min_{\boldsymbol{x}_{\boldsymbol{\kappa}}\in\overline{\Omega}_h} \tilde{u}_{\boldsymbol{\kappa}}=\min_{\boldsymbol{x}_{\boldsymbol{\kappa}}\in\partial\Omega_h} \tilde{u}_{\boldsymbol{\kappa}}$。

注意当有 *Dirichlet* 边界条件 $\tilde{u}_{\boldsymbol{\kappa}}=g_{\boldsymbol{\kappa}}$ 时，边界最大值（*resp.* 边界最小值）可以化为 $\max_{\boldsymbol{x}_{\boldsymbol{\kappa}}\in\partial\Omega_h} g_{\boldsymbol{\kappa}}$。

Proof theorem 4.6.2 我们只证最大值的情形。当 $f_{\boldsymbol{\kappa}} < 0$ 对内点 Ω_h 均成立时，假设 $\tilde{u}$ 在内点 $\boldsymbol{x}_{\boldsymbol{\kappa}_0} \in \Omega_h$ 取到最大值。则

$$-\sum_{i=1}^{n} D_{x_i}^{+} D_{x_i}^{-} \tilde{u}_{\boldsymbol{\kappa}} = -\sum_{i=1}^{n} \frac{1}{h_{x_i,\boldsymbol{\kappa}}} \left(\frac{\tilde{u}_{\boldsymbol{\kappa}+e_i} - \tilde{u}_{\boldsymbol{\kappa}}}{h^{-}_{x_i,\boldsymbol{\kappa}+e_i}} - \frac{\tilde{u}_{\boldsymbol{\kappa}} - \tilde{u}_{\boldsymbol{\kappa}-e_i}}{h^{-}_{x_i,\boldsymbol{\kappa}}} \right) = f_{\boldsymbol{\kappa}}$$

将 $\tilde{u}_{\boldsymbol{\kappa}}$ 保留在等式左侧，移项有

$$\tilde{u}_{\boldsymbol{\kappa}} \sum_{i=1}^{n} \frac{1}{h_{x_i,\boldsymbol{\kappa}}} \left(\frac{1}{h^{-}_{x_i,\boldsymbol{\kappa}+e_i}} + \frac{1}{h^{-}_{x_i,\boldsymbol{\kappa}}} \right) = \sum_{i=1}^{n} \frac{1}{h_{x_i,\boldsymbol{\kappa}}} \left(\frac{\tilde{u}_{\boldsymbol{\kappa}+e_i}}{h^{-}_{x_i,\boldsymbol{\kappa}+e_i}} + \frac{\tilde{u}_{\boldsymbol{\kappa}-e_i}}{h^{-}_{x_i,\boldsymbol{\kappa}}} \right) + f_{\boldsymbol{\kappa}}$$

由于 $f_{\boldsymbol{\kappa}} < 0$，$\tilde{u}_{\boldsymbol{\kappa}_0}$ 在内点取最大值，所以

$$\begin{aligned}
\tilde{u}_{\boldsymbol{\kappa}_0} \sum_{i=1}^{n} \frac{1}{h_{x_i,\boldsymbol{\kappa}_0}} \left(\frac{1}{h^{-}_{x_i,\boldsymbol{\kappa}_0+e_i}} + \frac{1}{h^{-}_{x_i,\boldsymbol{\kappa}_0}} \right) &= \sum_{i=1}^{n} \frac{1}{h_{x_i,\boldsymbol{\kappa}_0}} \left(\frac{\tilde{u}_{\boldsymbol{\kappa}_0+e_i}}{h^{-}_{x_i,\boldsymbol{\kappa}_0+e_i}} + \frac{\tilde{u}_{\boldsymbol{\kappa}_0-e_i}}{h^{-}_{x_i,\boldsymbol{\kappa}_0}} \right) + f_{\boldsymbol{\kappa}_0} \\
&< \sum_{i=1}^{n} \frac{1}{h_{x_i,\boldsymbol{\kappa}_0}} \left(\frac{\tilde{u}_{\boldsymbol{\kappa}_0+e_i}}{h^{-}_{x_i,\boldsymbol{\kappa}_0+e_i}} + \frac{\tilde{u}_{\boldsymbol{\kappa}_0-e_i}}{h^{-}_{x_i,\boldsymbol{\kappa}_0}} \right) \\
&\leq \sum_{i=1}^{n} \frac{1}{h_{x_i,\boldsymbol{\kappa}_0}} \left(\frac{\tilde{u}_{\boldsymbol{\kappa}_0}}{h^{-}_{x_i,\boldsymbol{\kappa}_0+e_i}} + \frac{\tilde{u}_{\boldsymbol{\kappa}_0}}{h^{-}_{x_i,\boldsymbol{\kappa}_0}} \right)
\end{aligned}$$

注意不等式两侧实际上应该严格取等，所以矛盾，$\tilde{u}$ 在内点取不到最大值。

进一步考虑当 $f_{\boldsymbol{\kappa}} \leq 0$ 时，考虑构造辅助函数 auxiliary function

$$v_{\varepsilon,\boldsymbol{\kappa}} = \tilde{u}_{\boldsymbol{\kappa}} + \frac{\varepsilon}{2n} ||\boldsymbol{x}_{\boldsymbol{\kappa}}||^2$$

其中 $\varepsilon > 0$ 为任意正实数。不难证明（如下）$\mathcal{L}_h v_\varepsilon = \mathcal{L}_h \tilde{u} - \varepsilon = f - \varepsilon < 0$，即 $v_{\varepsilon,\boldsymbol{\kappa}}$ 只能在边界 $\partial\Omega_h$ 上取到最大值。

$$\begin{aligned}
\mathcal{L}_h \left[\frac{\varepsilon}{2n} ||\boldsymbol{x}_{\boldsymbol{\kappa}}||^2 \right] &= -\frac{\varepsilon}{2n} \sum_{i=1}^{n} D_{x_i}^{+} D_{x_i}^{-} ||\boldsymbol{x}_{\boldsymbol{\kappa}}||^2 = -\frac{\varepsilon}{2n} \sum_{i=1}^{n} D_{x_i}^{+} D_{x_i}^{-} |x_{\boldsymbol{\kappa}_i}|^2 \\
&= -\frac{\varepsilon}{2n} \sum_{i=1}^{n} \frac{1}{h_{x_i,\boldsymbol{\kappa}_i}} \left(\frac{|x_{\boldsymbol{\kappa}_i+1}|^2 - |x_{\boldsymbol{\kappa}_i}|^2}{h^{-}_{x_i,\boldsymbol{\kappa}_i+1}} - \frac{|x_{\boldsymbol{\kappa}_i}|^2 - |x_{\boldsymbol{\kappa}_i-1}|^2}{h^{-}_{x_i,\boldsymbol{\kappa}_i}} \right) \\
&= -\frac{\varepsilon}{2n} \sum_{i=1}^{n} \frac{1}{h_{x_i,\boldsymbol{\kappa}_i}} \left(x_{\boldsymbol{\kappa}_i+1} + x_{\boldsymbol{\kappa}_i} - x_{\boldsymbol{\kappa}_i} - x_{\boldsymbol{\kappa}_i-1} \right) \\
&= -\frac{\varepsilon}{2n} \sum_{i=1}^{n} \frac{2 \cdot h_{x_i,\boldsymbol{\kappa}_i}}{h_{x_i,\boldsymbol{\kappa}_i}} = -\varepsilon
\end{aligned}$$

于是有

$$\begin{aligned}\max_{\boldsymbol{x}_{\boldsymbol{\kappa}}\in\partial\Omega_h}\tilde{u}_{\boldsymbol{\kappa}} &= \max_{\boldsymbol{x}_{\boldsymbol{\kappa}}\in\partial\Omega_h}\left\{v_{\varepsilon,\boldsymbol{\kappa}}-\frac{\varepsilon}{2n}||\boldsymbol{x}_{\boldsymbol{\kappa}}||^2\right\}\\ &\geq \max_{\boldsymbol{x}_{\boldsymbol{\kappa}}\in\partial\Omega_h} v_{\varepsilon,\boldsymbol{\kappa}} - \max_{\boldsymbol{x}_{\boldsymbol{\kappa}}\in\partial\Omega_h}\frac{\varepsilon}{2n}||\boldsymbol{x}_{\boldsymbol{\kappa}}||^2\\ &= \max_{\boldsymbol{x}_{\boldsymbol{\kappa}}\in\overline{\Omega}_h} v_{\varepsilon,\boldsymbol{\kappa}} - \frac{\varepsilon}{2n}\max_{\boldsymbol{x}_{\boldsymbol{\kappa}}\in\partial\Omega_h}||\boldsymbol{x}_{\boldsymbol{\kappa}}||^2\\ &\geq \max_{\boldsymbol{x}_{\boldsymbol{\kappa}}\in\overline{\Omega}_h}\tilde{u}_{\boldsymbol{\kappa}} - \frac{\varepsilon}{2n}\max_{\boldsymbol{x}_{\boldsymbol{\kappa}}\in\partial\Omega_h}||\boldsymbol{x}_{\boldsymbol{\kappa}}||^2\end{aligned}$$

最后一个等号是因为显然有 $v_{\varepsilon,\boldsymbol{\kappa}}\geq\tilde{u}_{\boldsymbol{\kappa}}$，取极限 $\varepsilon\to 0$，我们有

$$\max_{\boldsymbol{x}_{\boldsymbol{\kappa}}\in\partial\Omega_h}\tilde{u}_{\boldsymbol{\kappa}}\geq\max_{\boldsymbol{x}_{\boldsymbol{\kappa}}\in\overline{\Omega}_h}\tilde{u}_{\boldsymbol{\kappa}}$$

因为 $\Omega_h\subseteq\overline{\Omega}_h$，显然有 $\max_{\boldsymbol{x}_{\boldsymbol{\kappa}}\in\partial\Omega_h}\tilde{u}_{\boldsymbol{\kappa}}\leq\max_{\boldsymbol{x}_{\boldsymbol{\kappa}}\in\overline{\Omega}_h}\tilde{u}_{\boldsymbol{\kappa}}$，综上

$$\max_{\boldsymbol{x}_{\boldsymbol{\kappa}}\in\partial\Omega_h}\tilde{u}_{\boldsymbol{\kappa}}=\max_{\boldsymbol{x}_{\boldsymbol{\kappa}}\in\overline{\Omega}_h}\tilde{u}_{\boldsymbol{\kappa}}$$

利用 theorem 4.6.2 我们可以很容易地得出如下结论

Corollary 4.6.1. 形如下方的有限差分策略所得解存在且唯一，且对于边界条件稳定。即下述线性问题 *well-posed*

$$\begin{aligned}-\sum_{i=1}^{n}D_{x_i}^{+}D_{x_i}^{-}\tilde{u}_{\boldsymbol{\kappa}} &= f_{\boldsymbol{\kappa}} \qquad \text{if } \boldsymbol{x}_{\boldsymbol{\kappa}}\in\Omega_h\\ \tilde{u}_{\boldsymbol{\kappa}} &= g_{\boldsymbol{\kappa}} \qquad \text{if } \boldsymbol{x}_{\boldsymbol{\kappa}}\in\partial\Omega_h\end{aligned}$$

其中网格可以为非均匀网格。

Proof theorem 4.6.1** 要证该线性方程组的解存在且唯一，等价于证下述伴随问题只有全 0 解：

$$\begin{aligned}\mathcal{L}_h\tilde{u}=-\sum_{i=1}^{n}D_{x_i}^{+}D_{x_i}^{-}\tilde{u}_{\boldsymbol{\kappa}} &= 0 \qquad \text{if } \boldsymbol{x}_{\boldsymbol{\kappa}}\in\Omega_h\\ \tilde{u}_{\boldsymbol{\kappa}} &= 0 \qquad \text{if } \boldsymbol{x}_{\boldsymbol{\kappa}}\in\partial\Omega_h\end{aligned}$$

这由 thm 4.6.2 显然可得，因为最大最小值只能在边界取，而边界值恒等于 0，所以对于内部的点 $\tilde{u}_{\boldsymbol{\kappa}}$ 有

$$0=\min_{\boldsymbol{x}_{\boldsymbol{\kappa}}\in\partial\Omega_h}\tilde{u}_{\boldsymbol{\kappa}}\leq\tilde{u}_{\boldsymbol{\kappa}}\leq\max_{\boldsymbol{x}_{\boldsymbol{\kappa}}\in\partial\Omega_h}\tilde{u}_{\boldsymbol{\kappa}}=0$$

即 $\tilde{u}_{\boldsymbol{\kappa}}\equiv 0\Rightarrow\mathcal{L}_h$ 可逆，i.e., 差分策略的解存在且唯一。

下面证明该策略的解对于边界条件稳定。对于扰动后的边界条件 g^1 与 g^2, 分别得到解 $\tilde{u}_1$ 与 $\tilde{u}_2$。显然 $\tilde{u} = \tilde{u}_1 - \tilde{u}_2$ 满足下述方程

$$\begin{aligned} \mathcal{L}_h \tilde{u}_{\boldsymbol{\kappa}} &= 0 \quad && \text{if } \boldsymbol{x}_{\boldsymbol{\kappa}} \in \Omega_h \\ \tilde{u}_{\boldsymbol{\kappa}} &= g_{\boldsymbol{\kappa}} && \text{if } \boldsymbol{x}_{\boldsymbol{\kappa}} \in \partial\Omega_h \end{aligned}$$

其中 $g_{\boldsymbol{\kappa}} = g^1_{\boldsymbol{\kappa}} - g^2_{\boldsymbol{\kappa}}$。同理 $-\tilde{u}$ 满足下述方程

$$\begin{aligned} \mathcal{L}_h \left(-\tilde{u}_{\boldsymbol{\kappa}}\right) &= 0 \quad && \text{if } \boldsymbol{x}_{\boldsymbol{\kappa}} \in \Omega_h \\ -\tilde{u}_{\boldsymbol{\kappa}} &= -g_{\boldsymbol{\kappa}} && \text{if } \boldsymbol{x}_{\boldsymbol{\kappa}} \in \partial\Omega_h \end{aligned}$$

于是分别有

$$\tilde{u}_{\boldsymbol{\kappa}} \le \max_{\boldsymbol{x}_{\boldsymbol{\kappa}} \in \overline{\Omega}_h} \tilde{u}_{\boldsymbol{\kappa}} \le \max_{\boldsymbol{x}_{\boldsymbol{\kappa}} \in \partial\Omega_h} g_{\boldsymbol{\kappa}} \le \max_{\boldsymbol{x}_{\boldsymbol{\kappa}} \in \partial\Omega_h} |g_{\boldsymbol{\kappa}}|$$

$$-\tilde{u}_{\boldsymbol{\kappa}} \le \max_{\boldsymbol{x}_{\boldsymbol{\kappa}} \in \overline{\Omega}_h} (-\tilde{u}_{\boldsymbol{\kappa}}) \le \max_{\boldsymbol{x}_{\boldsymbol{\kappa}} \in \partial\Omega_h} (-g_{\boldsymbol{\kappa}}) \le \max_{\boldsymbol{x}_{\boldsymbol{\kappa}} \in \partial\Omega_h} |g_{\boldsymbol{\kappa}}|$$

即有

$$|\tilde{u}_{\boldsymbol{\kappa}}| \le \max_{\boldsymbol{x}_{\boldsymbol{\kappa}} \in \partial\Omega_h} |g_{\boldsymbol{\kappa}}| \;\Rightarrow\; \max_{\boldsymbol{x}_{\boldsymbol{\kappa}} \in \overline{\Omega}_h} |\tilde{u}_{\boldsymbol{\kappa}}| \le \max_{\boldsymbol{x}_{\boldsymbol{\kappa}} \in \partial\Omega_h} |g_{\boldsymbol{\kappa}}|$$

即

$$\max_{\boldsymbol{x}_{\boldsymbol{\kappa}} \in \overline{\Omega}_h} |\tilde{u}_{1,\boldsymbol{\kappa}} - \tilde{u}_{2,\boldsymbol{\kappa}}| \le \max_{\boldsymbol{x}_{\boldsymbol{\kappa}} \in \partial\Omega_h} |g^1_{\boldsymbol{\kappa}} - g^2_{\boldsymbol{\kappa}}|$$

证毕。

4.7 Iterative Method for Linear Systems

之前讨论的所有差分策略，最终都会将所求的 BVP 转化为求解线性方程组 $\mathcal{L}_h\tilde{u} = f_h$。其中一般有 $\mathcal{L}_h$ 为一个正定、对称的稀疏矩阵。在本节中，我们将具体研究如何采用迭代法求解这一线性方程组，即求解

$$\mathcal{L}_h\tilde{u} = f_h \quad \text{s.t.} \quad \mathcal{L}_h \succ 0\ ,\ \text{symmetric}$$

注意，正定实对称矩阵 $\mathcal{L}_h$ 有正实特征根 $0 < \alpha < \lambda_i < \beta$。记矩阵 A 的谱半径 spectral radius 为 $\rho(A) = \max_i |\lambda_i|$，即最大的模特征值。

Theorem 4.7.1. 迭代法的收敛速率 对于对称线性方程组 $A\boldsymbol{x} = \boldsymbol{b}$ s.t. A symmetric，若 $\rho(I - A) < 1$，则下述迭代法收敛于真解 $\boldsymbol{x}$

$$\boldsymbol{x}^{(n+1)} = (I - A)\,\boldsymbol{x}^{(n)} + \boldsymbol{b}$$

准确地，有

$$\begin{aligned} ||\boldsymbol{x} - \boldsymbol{x}^{(n+1)}|| &\leq \{\rho(I - A)\}^n \cdot ||\boldsymbol{x} - \boldsymbol{x}^{(0)}|| \\ ||\boldsymbol{b} - A\boldsymbol{x}^{(n+1)}|| &\leq \{\rho(I - A)\}^n \cdot ||\boldsymbol{b} - A\boldsymbol{x}^{(0)}|| \end{aligned}$$

称 $\boldsymbol{b} - A\boldsymbol{x}^{(n)}$ 为迭代残差 *residual*。

Proof theorem 4.7.1 注意对于实对称矩阵 $I - A$，其矩阵范数 $||I - A|| := ||I - A||_2 = \rho(I - A)$（NLA Problem Sheet 01, Q6 ）

$$\begin{aligned} ||\boldsymbol{x} - \boldsymbol{x}^{(n+1)}|| &= ||\boldsymbol{x} - (I - A)\,\boldsymbol{x}^{(n)} - \boldsymbol{b}|| = ||\boldsymbol{x} - (I - A)\,\boldsymbol{x}^{(n)} - A\boldsymbol{x}|| \\ &= ||\boldsymbol{x} - \boldsymbol{x}^{(n)} + A(\boldsymbol{x}^{(n)} - \boldsymbol{x})|| \\ &= ||\,(I - A)\,(\boldsymbol{x} - \boldsymbol{x}^{(n)})|| \\ &\leq ||I - A|| \cdot ||\boldsymbol{x} - \boldsymbol{x}^{(n)}|| = \rho(I - A) \cdot ||\boldsymbol{x} - \boldsymbol{x}^{(n)}|| \\ &\leq \{\rho(I - A)\}^n \cdot ||\boldsymbol{x} - \boldsymbol{x}^{(0)}|| \longrightarrow 0 \quad \text{if} \quad \rho(I - A) < 1 \end{aligned}$$

对迭代残差 residual $\boldsymbol{b} - A\boldsymbol{x}^{(n)}$，同样有

$$\begin{aligned} ||A\boldsymbol{x} - A\boldsymbol{x}^{(n+1)}|| &= ||\boldsymbol{b} - A\boldsymbol{x}^{(n+1)}|| \\ &= ||A\boldsymbol{x} - A\boldsymbol{x}^{(n)} + A(A\boldsymbol{x}^{(n)} - A\boldsymbol{x})|| \\ &= ||\boldsymbol{b} - A\boldsymbol{x}^{(n)} + A(A\boldsymbol{x}^{(n)} - \boldsymbol{b})|| = ||\,(I - A)\,(\boldsymbol{b} - A\boldsymbol{x}^{(n)})|| \\ &\leq ||I - A|| \cdot ||\boldsymbol{b} - A\boldsymbol{x}^{(n)}|| = \rho(I - A) \cdot ||\boldsymbol{b} - A\boldsymbol{x}^{(n)}|| \\ &\leq \{\rho(I - A)\}^n \cdot ||\boldsymbol{b} - A\boldsymbol{x}^{(0)}|| \longrightarrow 0 \quad \text{if} \quad \rho(I - A) < 1 \end{aligned}$$

Theorem 4.7.2. Richardson Iteration

对于对称线性方程组 $A\boldsymbol{x}=\boldsymbol{b}$ s.t. A symmetric，称下述迭代法为 *Richardson Iteration*

$$\tau>0 \quad : \quad \boldsymbol{x}^{(n+1)}=\boldsymbol{x}^{(n)}-\tau\left(A\boldsymbol{x}^{(n)}-\boldsymbol{b}\right)$$

如果 $A\succ 0$ 还是正定矩阵 s.t. $0<\alpha<\lambda_i<\beta$，则取 $\tau^*=\frac{2}{\beta+\alpha}$ 迭代法总收敛，且收敛速率满足

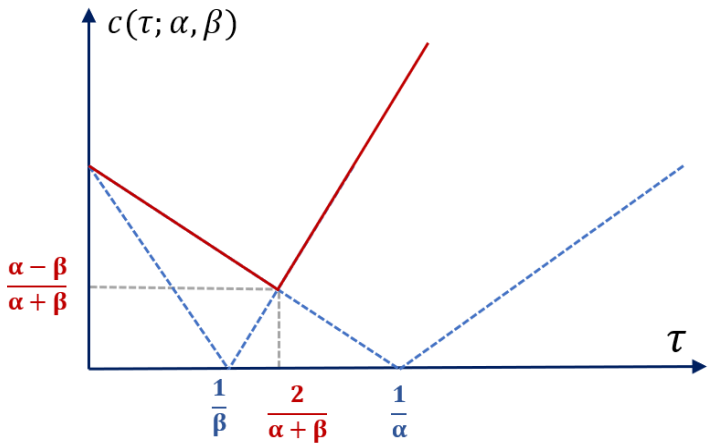
$$\begin{aligned}||\boldsymbol{x}-\boldsymbol{x}^{(n+1)}|| &\leq \left(\tfrac{\beta-\alpha}{\beta+\alpha}\right)^n\cdot||\boldsymbol{x}-\boldsymbol{x}^{(0)}||\\ ||\boldsymbol{b}-A\boldsymbol{x}^{(n+1)}|| &\leq \left(\tfrac{\beta-\alpha}{\beta+\alpha}\right)^n\cdot||\boldsymbol{b}-A\boldsymbol{x}^{(0)}||\end{aligned}$$

Proof theorem 4.7.2 将迭代式子整理成 $\boldsymbol{x}^{(n+1)}=(I-\tau A)\,\boldsymbol{x}^{(n)}+\tau\boldsymbol{b}$，对 $\rho(I-\tau A)$ 有

$$\begin{aligned}\rho(I-\tau A)&=\max_i|\lambda_i(I-\tau A)|=\max_i|1-\tau\lambda_i(A)|\\ &\leq\max\{|1-\tau\alpha|\ ,\ |1-\tau\beta|\}=c(\tau;\alpha,\beta)\end{aligned}$$

我们考虑取 $\tau^*=\arg\min_{\tau>0}c(\tau;\alpha,\beta)$ s.t. $c(\tau^*;\alpha,\beta)<1$。显然可以取 $\tau^*=2/(\alpha+\beta)$（见下图），此时 $c(\tau^*;\alpha,\beta)=(\beta-\alpha)/(\beta+\alpha)<1$。结合 thm 4.7.1 自然有（对迭代残差同理）

$$||\boldsymbol{x}-\boldsymbol{x}^{(n+1)}||\leq\{\rho(I-\tau A)\}^n\cdot||\boldsymbol{x}-\boldsymbol{x}^{(0)}||\leq\left(\tfrac{\beta-\alpha}{\beta+\alpha}\right)^n\cdot||\boldsymbol{x}-\boldsymbol{x}^{(0)}||$$



Convergence Rate of Richardson Iterations

STEP 1 将求解问题离散化为 $\mathcal{L}_h\tilde{u} = f_h \quad \text{s.t.} \quad \mathcal{L}_h \succ 0$, symmetric

STEP 2 代入法 * 求 $\mathcal{L}_h$ 的特征值，找到其上下界 $0 < \alpha < \lambda_i(\mathcal{L}_h) < \beta$

STEP 3 取 $\tau^* = \frac{2}{\beta+\alpha}$，则迭代法 $\tilde{u}^{(n+1)} = \tilde{u}^{(n)} - \tau^*(\mathcal{L}_h\tilde{u}^{(n)} - \boldsymbol{f_h})$ 的收敛速率满足

$$
\begin{aligned}
||\tilde{u} - \tilde{u}^{(n+1)}|| &\leq \left(\frac{\beta-\alpha}{\beta+\alpha}\right)^n \cdot ||\tilde{u} - \tilde{u}^{(0)}|| \\
||\boldsymbol{f_h} - \mathcal{L}_h\tilde{u}^{(n+1)}|| &\leq \left(\frac{\beta-\alpha}{\beta+\alpha}\right)^n \cdot ||f_h - \mathcal{L}_h\tilde{u}^{(0)}||
\end{aligned}
$$

关键是如何找到 $\mathcal{L}_h$ 的特征值，一种思路是先猜测特征向量（或离散特征函数）的形式 $\phi^\alpha = (\phi^\alpha_\kappa)$ 后（其中 α 为区分特征向量的上标，κ 为区分特征向量 ϕ^α 中元素的下标），将其带入 $\mathcal{L}_h\phi^\alpha_\kappa$ 并化成 $\lambda^\alpha\phi^\alpha_\kappa$ 的形式，从而找到特征根 λ^α。

一般我们可以有两种策略得到特征向量的试根：第一种，考虑原 BVP 的特征函数 ϕ^α s.t. $\mathcal{L}\varphi^\alpha = \lambda^\alpha\varphi^\alpha$，则一般离散特征函数，i.e., $\mathcal{L}_h$ 的特征向量为 $\phi^\alpha_\kappa = \varphi^\alpha(\boldsymbol{x}_\kappa)$, $\boldsymbol{x}_\kappa \in \Omega_h$。

第二种方法，记住下面几种情形下的特征向量试根

1） 当 $\mathcal{L}_h = D_n(a, b, a)$ 为 $n \times n$ 三对角矩阵时，可试 $\phi^k_i = sin(k\pi x_i)$；

2） 当为离散化拉普拉斯方程时，二维情况下，若有 $n_x n_y$ 个内部节点，可试

$$\phi^{kl}_{ij} = sin(k\pi x_i)\, sin(l\pi y_j)$$

3） 当求解区域 $\Omega \subseteq \mathbb{R}^n$ 无界时，可试 $\phi_\kappa(\theta) = e^{i||\boldsymbol{x}_\kappa||_1\theta}$，其中 θ 为区分特征向量的参数；即使求解区域有界，有时也可以尝试这种试根。

Example 4.7.1. 寻找差分策略系数矩阵的特征值

$$-\frac{\tilde{u}_{i+1}-2\tilde{u}_i+\tilde{u}_{i-1}}{h^2}+c\tilde{u}_i=f_i \quad \text{s.t.} \quad c\geq 0$$

可设特征向量形如 $\phi_i^k=sin\left(k\pi x_i\right)$ ，$i=1\sim(N-1)$，带入上式左侧有

$$\begin{aligned}
&-\frac{\phi_{i+1}^k-2\phi_i^k+\phi_{i-1}^k}{h^2}+c\cdot\phi_i^k \\
&=-\tfrac{1}{h^2}\left(sin\left(k\pi x_{i+1}\right)-2sin\left(k\pi x_i\right)+sin\left(k\pi x_{i-1}\right)\right)+c\cdot sin\left(k\pi x_i\right) \\
&=-\tfrac{1}{h^2}\left(2sin\left(k\pi x_i\right)cos\left(k\pi h\right)-2sin\left(k\pi x_i\right)\right)+c\cdot sin\left(k\pi x_i\right) \\
&=sin\left(k\pi x_i\right)\left[c-\tfrac{2}{h^2}\left(cos\left(k\pi h\right)-1\right)\right]=\phi_i^k\left[c+\tfrac{4}{h^2}sin^2\left(\tfrac{k\pi h}{2}\right)\right]
\end{aligned}$$

第二个等式为和差化积。于是得到 $\lambda_k=c+\frac{4}{h^2}sin^2\left(\frac{k\pi h}{2}\right)\in\left[c+8\ ,\ c+4/h^2\right]$，上界显然，对于下界，有

$$sin\left(x\right)\geq\tfrac{2\sqrt{2}}{\pi}x\quad,\quad x\in\left[0,\pi/4\right]$$

于是

$$sin^2\left(\tfrac{k\pi h}{2}\right)=sin^2\left(\tfrac{k\pi(b-a)}{2N}\right)\geq sin^2\left(\tfrac{1\cdot\pi(b-a)}{2N}\right)\geq\tfrac{8}{\pi^2}\tfrac{\pi^2h^2}{4}=2h^2$$

这要求 $h=\left(b-a\right)/N<1/2$，同时 $\lambda_k\geq c+8$。综上，我们有 $\alpha=c+8$ ，$\beta=c+4/h^2$。另一种特征向量的构造可以是：$\phi_k^\theta=e^{ix_k\theta}$

$$\begin{aligned}
&-\frac{\phi_{k+1}^\theta-2\phi_k^\theta+\phi_{k-1}^\theta}{h^2}+c\cdot\phi_k^\theta \\
&=-\tfrac{1}{h^2}\left(e^{ix_{k+1}\theta}-2e^{ix_k\theta}+e^{ix_{k-1}\theta}\right)+c\cdot e^{ix_k\theta} \\
&=-\tfrac{1}{h^2}e^{ix_k\theta}\left(e^{ih\theta}-2+e^{-ih\theta}\right)+c\cdot e^{ix_k\theta} \\
&=e^{ix_k\theta}\left[c-\tfrac{2}{h^2}\left(cos\left(h\theta\right)-1\right)\right]=\phi_k^\theta\left[c+\tfrac{4}{h^2}sin^2\left(\tfrac{h\theta}{2}\right)\right]
\end{aligned}$$

在上例中，可以求得 BVP 的特征方程（解 $-\varphi''+c\varphi=\lambda\varphi$ ）为

$$\varphi_k\left(x\right)=sin\left(k\pi x\right)$$

与所设的特征向量（离散特征函数）形式一致。双变量例题见 **Lecture NOtes p.46, Example 1**。

5 Finite Difference Schemes for Parabolic PDEs

5.1 Introduction

在本章中我们主要讨论最简单的热方程，即对 $u(x,t)$，$x \in \mathbb{R}$，$t \geq 0$ 满足

$$u_t = u_{xx} \quad \text{s.t.} \quad u(x,0) = u_0(x) \tag{1}$$

我们将会较多地使用傅里叶变换及其性质，详见 Appendix 或 MATH323 部分笔记。在此处我们只对定义做一些简单的回顾。

Theorem 5.1.1. 傅里叶变换
记函数 $f(x)$ 和 $\hat{f}(\xi)$ 的傅里叶（逆）变换分别为

$$\hat{f}(\xi) = F[f](\xi) = \int_{\infty}^{\infty} f(x) e^{-i\xi x} dx \quad , \quad f(x) \sim F^{-1}[\hat{f}](x) = \frac{1}{2\pi} \int_{\infty}^{\infty} \hat{f}(\xi) e^{i\xi x} d\xi$$

当 f 在 $\mathbb{R}$ 上绝对可积且在任何闭区间上分段可微时

$$F^{-1}[\hat{f}](x) = \frac{1}{2}\left[f(x^-) + f(x^+)\right]$$

在连续点处 $(F^{-1} \circ F)[f] := F^{-1}[F[f]] = f(x)$，未经特殊说明的情况下我们假设该条件一直满足。

Theorem 5.1.2. 热方程的解析解
利用傅里叶变换可以求得热方程 (1) 的解析解为（$x \in \mathbb{R}$，$t \geq 0$）

$$u(x,t) = \left[w(x,t) * u_0(x)\right]_x = \frac{1}{\sqrt{4\pi t}} \int_{\mathbb{R}} e^{\frac{(x-s)^2}{4t}} u_0(s)\, ds$$

$\left[f(x) * g(x)\right] = \int_{\mathbb{R}} f(x-s)\, g(s)\, ds$ 表示对 f, g 卷积 *convolution*；其中 $w(x,t) = \frac{1}{\sqrt{4\pi t}} e^{\frac{x^2}{4t}}$ 称为 *heat kernel*，不难证明有 $\int_{\mathbb{R}} w(x,t)\, dx = 1$。

Theorem 5.1.3. 热方程的解被初值条件 bound
记 $u(x,t)$ 为方程 (1) 的解析解，$u(x,0) = u_0(x)$ 为边界条件。则

1) **Bounded by L_∞** $||u(\cdot,t)||_{L_\infty(\mathbb{R})} \leq ||u_0||_{L_\infty(\mathbb{R})}$ for $\forall t \geq 0$
2) **Bounded by L_2** $||u(\cdot,t)||_{L_2(\mathbb{R})} \leq ||u_0||_{L_2(\mathbb{R})}$ for $\forall t \geq 0$

其中第二条默认了 $u(\cdot,t), u_0(x) \in L_2(\mathbb{R})$。

Corollary 5.1.1. Well-posedness for Solving the Heat Equation 初值问题 (1) 的解存在且唯一；同时解相对于初值条件对 $||\cdot||_{L_\infty(\mathbb{R})}$ 与 $||\cdot||_{L_2(\mathbb{R})}$ 连续：假设 u, u' 分别是初值条件 u_0, u_0' 的解，则对 $\forall t \geq 0$

$$||u(\cdot,t) - u'(\cdot,t)||_{L_\infty(\mathbb{R})} \leq ||u_0 - u_0'||_{L_\infty(\mathbb{R})}$$

如果 $u(\cdot,t), u'(\cdot,t), u_0, u_0' \in L_2(\mathbb{R})$，则还有

$$||u(\cdot,t) - u'(\cdot,t)||_{L_2(\mathbb{R})} \leq ||u_0 - u_0'||_{L_2(\mathbb{R})}$$

为了证明上述适定性的第二条，需要引入以下引理

Lemma 5.1.1. 记定义在 $\mathbb{R}$ 上的复函数 u, v，傅里叶变换均存在，记为 $\hat{u}, \hat{v}$。则

$$\int_{\mathbb{R}} \hat{u}(\xi) v(\xi) d\xi = \int_{\mathbb{R}} u(x) \hat{v}(x) dx$$

Lemma 5.1.2. Parseval's Identity 记定义在 $\mathbb{R}$ 上的复函数 $u \in L_2(\mathbb{R})$，其傅里叶变换为 $\hat{u}$。则 $\hat{u} \in L_2(\mathbb{R})$ 且

$$||u||_{L_2(\mathbb{R})} = \frac{1}{2\pi} ||\hat{u}||_{L_2(\mathbb{R})}$$

下面我们开始逐步证明上述定理，先从引理开始。

Proof lemma 5.1.1/2

$$\begin{aligned}\int_{\mathbb{R}} \hat{u}(\xi) v(\xi) d\xi &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} u(x) e^{-i\xi x} dx \right) v(\xi) d\xi = \int_{\mathbb{R}} \int_{\mathbb{R}} u(x) e^{-i\xi x} v(\xi) dx d\xi \\ &= \int_{\mathbb{R}} u(x) \left(\int_{\mathbb{R}} e^{-i\xi x} v(\xi) d\xi \right) dx = \int_{\mathbb{R}} u(x) \hat{v}(x) dx\end{aligned}$$

取 $v(\xi) = \overline{\hat{u}(\xi)} = 2\pi F^{-1}[\bar{u}](\xi)$（后一个等式由定义易证），带入上式可证 lemma 5.1.2

$$\text{R.H.S} = \int_{\mathbb{R}} \hat{u}(\xi) v(\xi) d\xi = \int_{\mathbb{R}} \hat{u}(\xi) \overline{\hat{u}(\xi)} d\xi = ||\hat{u}||^2_{L_2(\mathbb{R})}$$

$$\text{L.H.S} = \int_{\mathbb{R}} u(x) \hat{v}(x) dx = \int_{\mathbb{R}} u(x) F\left[2\pi F^{-1}[\bar{u}](\xi)\right](x) dx = 2\pi ||u||^2_{L_2(\mathbb{R})}$$

于是 $||\hat{u}||^2_{L_2(\mathbb{R})} = 2\pi ||u||^2_{L_2(\mathbb{R})}$，证毕。

Proof theorem 5.1.2 对方程与边界条件中的 x 做傅里叶变换，转化为对 $\xi \in \mathbb{R}$，$t \geq 0$ 满足

$$\frac{\partial}{\partial t}\hat{u}(\xi,t) = -\xi^2 \hat{u}(\xi,t) \quad \text{s.t.} \quad \hat{u}(\xi,0) = \hat{u}_0(\xi)$$

不难解得变换后的解为 $\hat{u}(\xi,t) = \hat{u}_0(\xi)\, e^{-\xi^2 t}$，对其做（$\xi$ 变量的）逆变换

$$\begin{aligned} u(x,t) &= F^{-1}[\hat{u}](x) = F^{-1}[\hat{u}_0] * F^{-1}[e^{-\xi^2 t}] \\ &= u_0(x) * \frac{1}{2\pi} F\left[e^{-\left(x\sqrt{t}\right)^2}\right](\xi)\,|_{-x} \\ &= u_0(x) * \left\{\frac{1}{2\pi}\frac{\sqrt{\pi}}{\sqrt{t}} exp\left(-\frac{x^2}{4t}\right)\right\} = \left[u_0(x) * w(x,t)\right]_x \end{aligned}$$

Proof theorem 5.1.3 利用 thm 5.1.2 中的结论我们不难有

$$\begin{aligned} |u(x,t)| &= \left|\left[w(x,t) * u_0(x)\right]_x\right| = \left|\int_{\mathbb{R}} w(s,t)\, u_0(x-s)\, ds\right| \\ &\leq ||u_0||_{L_\infty(\mathbb{R})} \int_{\mathbb{R}} w(s,t)\, ds = ||u_0||_{L_\infty(\mathbb{R})} \end{aligned}$$

由上确界的定义 $||u(\cdot,t)||_{L_\infty(\mathbb{R})} = \sup_{x\in\mathbb{R}} |u(x,t)| \leq ||u_0||_{L_\infty(\mathbb{R})}$，证毕。

下证第二个不等式，由 lemma 5.1.2

$$\begin{aligned} ||u(\cdot,t)||_{L_2(\mathbb{R})} &= \frac{1}{\sqrt{2\pi}}||\hat{u}(\cdot,t)||_{L_2(\mathbb{R})} = \frac{1}{\sqrt{2\pi}}||\hat{u}_0(\xi)\, e^{-\xi^2 t}||_{L_2(\mathbb{R})} \\ &\leq \frac{1}{\sqrt{2\pi}}||\hat{u}_0(\xi)||_{L_2(\mathbb{R})} \cdot ||e^{-\xi^2 t}||_{L_2(\mathbb{R})} \\ &\leq \frac{1}{\sqrt{2\pi}}||\hat{u}_0(\xi)||_{L_2(\mathbb{R})} \cdot 1 = ||u_0(\xi)||_{L_2(\mathbb{R})} \end{aligned}$$

Proof corollary 5.1.1 $u - u'$ 为方程对初值条件 $u_0 - u_0'$ 的解，由 theorem 5.1.3 直接可得结论。

5.2 θ-methods

本章中我们目标求解问题为对 $u(x,t)$，$x \in \mathbb{R}$，$0 \le t \le T$ 满足

$$u_t = u_{xx} \quad \text{s.t.} \quad u(x,0) = u_0(x) \tag{1}$$

该方程存在解析解 $u(x,t) = \left[w(x,t) * u_0(x)\right]_x = \frac{1}{\sqrt{4\pi t}} \int_{\mathbb{R}} e^{\frac{(x-s)^2}{4t}} u_0(s)\, ds$。现取离散化网格

$$t_m = m\Delta t \quad \text{for} \quad \Delta t = T/M\ ,\ m = 0 \cdots M$$
$$x_j = j\Delta x \quad \text{for} \quad \Delta x > 0\ ,\ j = 0, \pm 1, \pm 2 \cdots$$

记得到的数值解为 $\tilde{u}_j^m \approx u(x_j, t_m)$。在未经特殊说明的情况下，后文中将默认沿用这些记号规定。

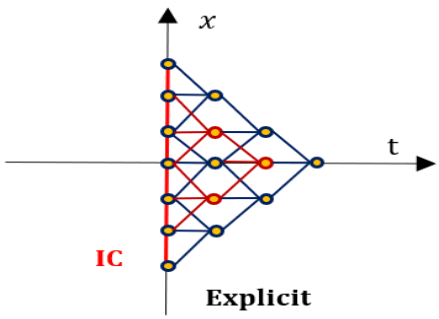
Definition 5.2.1. 欧拉法（Euler Methods） 采用中心差分近似二阶空间导数，一阶向前差分近似时间导数，可以得到显式欧拉法 *Explicit Euler Method*

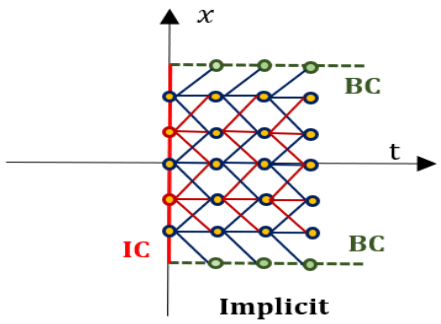
$$D_t^+ \tilde{u}_j^m = D_{xx}^c \tilde{u}_j^m \quad \Longleftrightarrow \quad \tilde{u}_j^{m+1} = \tilde{u}_j^m + \mu\left(\tilde{u}_{j-1}^m - 2\tilde{u}_j^m + \tilde{u}_{j+1}^m\right)$$

其中 $\mu = \frac{\Delta t}{(\Delta x)^2}$；如果使用一阶向后差分近似时间导数，可以得到隐式欧拉法 *Implicit Euler Method*

$$D_t^- \tilde{u}_j^{m+1} = D_{xx}^c \tilde{u}_j^{m+1} \quad \Longleftrightarrow \quad \frac{\tilde{u}_j^{m+1} - \tilde{u}_j^m}{\Delta t} = \frac{\tilde{u}_{j-1}^{m+1} - 2\tilde{u}_j^{m+1} + \tilde{u}_{j+1}^{m+1}}{\Delta x^2}$$

上述 $m = 0 \cdots (M-1)$。实际操作中，显式法由于存在递推关系，只要初值条件覆盖的范围足够大，总是可以得到任意时刻任意位置的估计值。而隐式方法想要求解则必须给出相应时刻的两个边界条件以确保方程数等于未知数个数，如下图所示。





Definition 5.2.2. θ-method

上述 Explicit/Implicit Euler Method 都属于 *θ-method* 的一种，其定义为

$$D_t^+ \tilde{u}_j^m = (1-\theta)\, D_{xx}^c \tilde{u}_j^m + \theta D_{xx}^c \tilde{u}_j^{m+1}$$

其中 $m = 0 \cdots (M-1)$；$\theta \in [0,1]$ 是一个参数。当 $\theta = 0$ 时为 Explicit Euler Method；当 $\theta = 1$ 时为 Implicit Euler Method。特别地，当 $\theta = 1/2$ 时为 *Crank-Nicolson Method*，其代表了 θ-method 所能达到的最高精度 $O(\Delta x^2 + \Delta t^2)$；除此之外其他 θ-method 所能达到的最高精度为 $O(\Delta x^2 + \Delta t)$。此外，对于非齐次方程 $u_t = u_{xx} + f(x,t)$，*θ-method* 有推广为

$$D_t^+ \tilde{u}_j^m = (1-\theta)\left[D_{xx}^c \tilde{u}_j^m + f_j^m\right] + \theta\left[D_{xx}^c \tilde{u}_j^{m+1} + f_j^{m+1}\right]$$

Theorem 5.2.1. θ-method 的精度

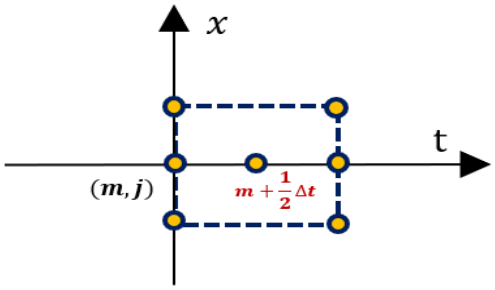
定义 *θ-method* 的一致性误差/截断误差 *consistency/truncation error* 为

$$T_j^m = D_t^+ u_j^m - (1-\theta)\, D_{xx}^c u_j^m - \theta D_{xx}^c u_j^{m+1}$$

其中 u_j^m 表示真值；我们有

$$T_j^m = \begin{cases} O(\Delta x^2 + \Delta t^2) & \text{if } \theta = 1/2 \\ O(\Delta x^2 + \Delta t) & \text{if } \theta \neq 1/2 \end{cases}$$

证明将 T_j^m 在 $(x_j, t_{m+\frac{1}{2}})$ 做展开即可，见下一页。



Definition 5.2.3. Fully-discrete and Semi-discrete Approximations

上述使用的方法统称为 *Fully-discrete Approximation*，因为我们对所有的（时间和空间）变量都进行了离散化。还有一种策略是只对空间坐标进行离散化，因此称为 *Semi-discrete Approximation*。此时，可以将 PDE 化为常微分方程组，并利用常微分方程组的数值法进行求解

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{u}_j(t) = D_{xx}^c \tilde{u}_j(t)$$

此时除了需要初值条件 $u(x,0)=u_0(x)$ 外，还需要将空间坐标限制在有限区间内 $x\in[a,b]$ 并施加边界条件 $u(a,t)=u_a(t)$，$u(b,t)=u_b(t)$。

Proof theorem 5.2.1 利用差分近似的误差公式，我们有

$$D_t^+u_j^m=D_t^cu_j^{m+\frac{1}{2}}=\partial_tu_j^{m+\frac{1}{2}}+O\left(\Delta t^2\right)$$

其中 $D_t^cu_j^{m+\frac{1}{2}}$ 的离散步长为 $\Delta t/2$。对于空间坐标的中心差分，我们有

$$D_{xx}^cu_j^m=\partial_{xx}u_j^m+O\left(\Delta x^2\right)$$

$$D_{xx}^cu_j^{m+1}=\partial_{xx}u_j^{m+1}+O\left(\Delta x^2\right)$$

对其在 $(x_j,t_{m+\frac{1}{2}})$ 做进一步展开，注意 $\partial_{xx}u_j^m=\partial_{xx}u_j^{m+\frac{1}{2}-\frac{1}{2}}$，$\partial_{xx}u_j^{m+1}=\partial_{xx}u_j^{m+\frac{1}{2}+\frac{1}{2}}$，于是

$$\partial_{xx}u_j^{m+\frac{1}{2}\mp\frac{1}{2}}=\partial_{xx}u_j^{m+\frac{1}{2}}\mp\frac{\Delta t}{2}\partial_{xxt}u_j^{m+\frac{1}{2}}+O\left(\Delta t^2\right)$$

带入下式，即有

$$\begin{aligned}
&(1-\theta)D_{xx}^cu_j^m+\theta D_{xx}^cu_j^{m+1}=(1-\theta)\partial_{xx}u_j^m+\theta\partial_{xx}u_j^{m+1}+O\left(\Delta x^2\right)\\
=&(1-\theta)\left[\partial_{xx}u_j^{m+\frac{1}{2}}-\frac{\Delta t}{2}\partial_{xxt}u_j^{m+\frac{1}{2}}\right]+\theta\left[\partial_{xx}u_j^{m+\frac{1}{2}}+\frac{\Delta t}{2}\partial_{xxt}u_j^{m+\frac{1}{2}}\right]\\
&+O\left(\Delta t^2\right)+O\left(\Delta x^2\right)\\
=&\partial_{xx}u_j^{m+\frac{1}{2}}+(2\theta-1)\frac{\Delta t}{2}\partial_{xxt}u_j^{m+\frac{1}{2}}+O\left(\Delta t^2\right)+O\left(\Delta x^2\right)
\end{aligned}$$

于是有

$$\begin{aligned}
T_j^m&=D_t^+u_j^m-(1-\theta)D_{xx}^cu_j^m-\theta D_{xx}^cu_j^{m+1}\\
&=\left[\partial_tu_j^{m+\frac{1}{2}}-\partial_{xx}u_j^{m+\frac{1}{2}}\right]-(2\theta-1)\frac{\Delta t}{2}\partial_{xxt}u_j^{m+\frac{1}{2}}+O\left(\Delta t^2\right)+O\left(\Delta x^2\right)\\
&=(1-2\theta)\frac{\Delta t}{2}\partial_{xxt}u_j^{m+\frac{1}{2}}+O\left(\Delta t^2\right)+O\left(\Delta x^2\right)
\end{aligned}$$

其中 $\partial_tu_j^{m+\frac{1}{2}}-\partial_{xx}u_j^{m+\frac{1}{2}}=0$ 是因为 u 为方程 $u_t=u_{xx}$ 的真解。显然当且仅当 $\theta=1/2$ 时能取到最小误差 $O\left(\Delta t^2\right)+O\left(\Delta x^2\right)$。

5.3 Stability of $\theta-$methods

本节我们将涉及大量的无穷级数求和，在此我们先假设所有的级数都是收敛的（同时所有的积分极限求和号可交换），以避免对复杂的敛散性条件的讨论。在开始讨论前，先对本章中热方程数值解的稳定性做如下定义

Definition 5.3.1. 热方程数值解的稳定性 记 $\tilde{u}_m = (\tilde{u}_j^m)$ 为 5.2 (1) 中热方程的数值解，其中 $m = 0 \cdots M$ 为时间上标，$j = 0, \pm 1 \cdots$ 为空间下标，空间节点的离散化步长为 Δx。定义如下 ℓ_2 范数

$$||\tilde{u}^m||_{\ell_2} = \left(\Delta x \sum_{j=-\infty}^{\infty} |\tilde{u}_j^m|^2 \right)^{\frac{1}{2}}$$

若数值解满足 $\forall m \geq 1 \ : \ ||\tilde{u}^m||_{\ell_2} \leq ||\tilde{u}^0||_{\ell_2}$ 则称数值解是稳定的 (*practically*) *stable*。稳定的一个充分条件是 $\forall m \geq 0 \ : \ ||\tilde{u}^{m+1}||_{\ell_2} \leq ||\tilde{u}^m||_{\ell_2}$。

Definition 5.3.2. 半离散傅里叶变换 SDFT
记 $\boldsymbol{u}$ 为定义在 $\{x_j \ : \ j = 0, \pm 1, \cdots\}$ 上的离散函数，$\boldsymbol{u}_j = \boldsymbol{u}(x_j)$，则定义其半离散傅里叶变换 *Semi-discrete Fourier Transform* 为连续函数

$$\hat{\boldsymbol{u}}(\omega) = F[\boldsymbol{u}](\omega) = \Delta x \sum_{j=-\infty}^{\infty} \boldsymbol{u}_j e^{-i\omega x_j} \quad \text{s.t.} \quad \omega \in \left[-\pi/\Delta x \ , \ \pi/\Delta x \right]$$

反之，定义在 $\left[-\pi/\Delta x \ , \ \pi/\Delta x \right]$ 上的连续函数 $u(\omega)$ 有对应的逆变换，为离散函数

$$F^{-1}[u]_j = \frac{1}{2\pi} \int_{-\pi/\Delta x}^{\pi/\Delta x} u(\omega) e^{i\omega x_j} d\omega$$

SDFT 与其逆变换均为双射，即 $F[\boldsymbol{u}] = F[\boldsymbol{v}] \iff \boldsymbol{u} = \boldsymbol{v}$（逆变换同理）。

Theorem 5.3.1. SDFT 的性质
记 $\boldsymbol{u}$ 为定义在 $\{x_j \ : \ j = 0, \pm 1, \cdots\}$ 上的离散函数，$\hat{\boldsymbol{u}}(\omega)$ 为其 SDFT，则

1) $\int_{-\pi/\Delta x}^{\pi/\Delta x} e^{i\omega(x_j - x_k)} d\omega = \frac{2\pi}{\Delta x} \delta_{jk}$

2) SDFT 有可逆性，即 $\boldsymbol{u} = F^{-1}[\hat{\boldsymbol{u}}] = \frac{1}{2\pi} \int_{-\pi/\Delta x}^{\pi/\Delta x} \hat{\boldsymbol{u}}(\omega) e^{i\omega x_j} d\omega$

3) $\int_{-\pi/\Delta x}^{\pi/\Delta x} \hat{\boldsymbol{u}}(\omega) v(\omega) d\omega = 2\pi \Delta x \sum_j \boldsymbol{u}_j F^{-1}[v(-\omega)]_j$

其中第三条可以用来证明 *Discrete Parseval's Identity*。

Proof theorem 5.3.1 (1) 若 $j=k$，显然成立。只证当 $j\neq k$ 时为 0

$$\begin{aligned}\int_{-\pi/\Delta x}^{\pi/\Delta x} e^{i\omega(x_j-x_k)}d\omega &\propto e^{i\omega(x_j-x_k)}\big|_{-\pi/\Delta x}^{\pi/\Delta x}\\ &\propto e^{i(-\pi/\Delta x)(x_j-x_k)}-e^{i(\pi/\Delta x)(x_j-x_k)}\\ &= e^{i\pi(k-j)}-e^{i\pi(j-k)}\\ &= (-1)^{(k-j)}-(-1)^{(j-k)}=0\end{aligned}$$

(2) 下证可逆性，假设级数均收敛且可以交换积分和极限求和号

$$\begin{aligned}F^{-1}\left[\hat{\boldsymbol{u}}\right]_j &= \frac{1}{2\pi}\int_{-\pi/\Delta x}^{\pi/\Delta x}\hat{\boldsymbol{u}}\left(\omega\right)e^{i\omega x_j}d\omega = \frac{1}{2\pi}\int_{-\pi/\Delta x}^{\pi/\Delta x}\left(\Delta x\sum_{k=-\infty}^{\infty}\boldsymbol{u}_k e^{-i\omega x_k}\right)e^{i\omega x_j}d\omega\\ &= \frac{\Delta x}{2\pi}\sum_{k=-\infty}^{\infty}\boldsymbol{u}_k\left(\int_{-\pi/\Delta x}^{\pi/\Delta x}e^{i\omega\left(x_j-x_k\right)}d\omega\right)\\ &= \frac{\Delta x}{2\pi}\sum_{k=-\infty}^{\infty}\boldsymbol{u}_k\frac{2\pi}{\Delta x}\delta_{jk}=\boldsymbol{u}_j\end{aligned}$$

(3) 同样假设级数均收敛且可以交换积分和极限求和号

$$\begin{aligned}\int_{-\pi/\Delta x}^{\pi/\Delta x}\hat{\boldsymbol{u}}\left(\omega\right)v\left(\omega\right)d\omega &= \int_{-\pi/\Delta x}^{\pi/\Delta x}\left(\Delta x\sum_{j=-\infty}^{\infty}\boldsymbol{u}_j e^{-i\omega x_j}\right)v\left(\omega\right)d\omega\\ &= \Delta x\sum_{j=-\infty}^{\infty}\boldsymbol{u}_j\int_{-\pi/\Delta x}^{\pi/\Delta x}e^{-i\omega x_j}v\left(\omega\right)d\omega\\ &= 2\pi\Delta x\sum_{j=-\infty}^{\infty}\boldsymbol{u}_j\cdot\frac{-1}{2\pi}\int_{-\pi/\Delta x}^{\pi/\Delta x}e^{i(-\omega)x_j}v\left(-\left(-\omega\right)\right)d\left(-\omega\right)\\ &= 2\pi\Delta x\sum_{j=-\infty}^{\infty}\boldsymbol{u}_j\cdot\frac{-1}{2\pi}\int_{\pi/\Delta x}^{-\pi/\Delta x}e^{itx_j}v\left(-t\right)dt\\ &= 2\pi\Delta x\sum_{j=-\infty}^{\infty}\boldsymbol{u}_j\cdot\frac{1}{2\pi}\int_{-\pi/\Delta x}^{\pi/\Delta x}e^{itx_j}v\left(-t\right)dt\\ &= 2\pi\Delta x\sum_j\boldsymbol{u}_j F^{-1}[v(-t)]_j\\ &= 2\pi\Delta x\sum_j\boldsymbol{u}_j F^{-1}[v(-\omega)]_j\end{aligned}$$

Theorem 5.3.2. Discrete Parseval's Identity
记 $\boldsymbol{u}$ 为定义在 $\{x_j \ : \ j=0,\pm1,\cdots\}$ 上的离散函数，$\hat{\boldsymbol{u}}(\omega)$ 为其 SDFT，记$I_{\Delta x}=[-\pi/\Delta x \ , \ \pi/\Delta x]$。若 $||\boldsymbol{u}||_{\ell_2}$ 有限，则 $||\hat{\boldsymbol{u}}||_{L_2(I_{\Delta x})}$ 有限，且

$$||\boldsymbol{u}||_{\ell_2}=\frac{1}{\sqrt{2\pi}}||\hat{\boldsymbol{u}}||_{L_2(I_{\Delta x})}$$

Proof theorem 5.3.2 带入 $v(\omega)=\overline{\hat{\boldsymbol{u}}(\omega)}$ 进 thm 5.3.1 (3) 即可，有

$$\text{R.H.S}=\int_{-\pi/\Delta x}^{\pi/\Delta x}\hat{\boldsymbol{u}}(\omega)\,v(\omega)\,d\omega=\int_{-\pi/\Delta x}^{\pi/\Delta x}\hat{\boldsymbol{u}}(\omega)\,\overline{\hat{\boldsymbol{u}}(\omega)}d\omega=||\hat{\boldsymbol{u}}||^2_{L_2(I_{\Delta x})}$$

$$\text{L.H.S}=2\pi\Delta x\sum_j \boldsymbol{u}_j F^{-1}[v(-\omega)]_j=2\pi\Delta x\sum_j \boldsymbol{u}_j F^{-1}[\overline{\hat{\boldsymbol{u}}}(-\omega)]_j$$

其中有

$$\begin{aligned}F^{-1}[\overline{\hat{\boldsymbol{u}}}(-\omega)]_j&=\frac{1}{2\pi}\int_{I_{\Delta x}}\overline{\hat{\boldsymbol{u}}}(-\omega)\,e^{i\omega x_j}d\omega=\frac{1}{2\pi}\int_{I_{\Delta x}}\left(\Delta x\sum_{k=-\infty}^{\infty}\bar{\boldsymbol{u}}_k e^{-i\omega x_k}\right)e^{i\omega x_j}d\omega\\&=\frac{1}{2\pi}\int_{I_{\Delta x}}\hat{\bar{\boldsymbol{u}}}e^{i\omega x_j}d\omega=F^{-1}[\hat{\bar{\boldsymbol{u}}}(\omega)]_j=\bar{\boldsymbol{u}}_j\end{aligned}$$

于是 $\text{L.H.S}=2\pi||u||^2_{\ell_2}=||\hat{\boldsymbol{u}}||^2_{L_2(I_{\Delta x})}=\text{R.H.S}$，证毕。

Theorem 5.3.3. Stability of Euler Methods
记 $\tilde{u}^m$ 为定义在 $\{x_j \ : \ j=0,\pm1,\cdots\}$ 上的方程 5.2 (1) 在 t_m 时刻的解，回顾稳定的定义为 $\forall m\geq 1 \ : \ ||\tilde{u}^m||_{\ell_2}\leq||\tilde{u}^0||_{\ell_2}$。记 $\mu=\frac{\Delta t}{\Delta x^2}$ 为 *CFL number*，则

1) 若 $\tilde{u}^m$ 是用显式欧拉法得到的解，则其在 $\mu<1/2$ 时条件稳定 *conditionally practically stable*;
2) 若 $\tilde{u}^m$ 是用隐式欧拉法得到的解，则其无条件稳定 *unconditionally practically stable*。

先给出上述结论，证明流程在下页的内容中给出。主要思路是利用傅里叶变换将递推公式化为与空间坐标无关的连续关系 $\hat{\tilde{u}}^m$, $\hat{\tilde{u}}^{m+1}$ 有关的关系之后再对范数放缩。并结合 Discrete Parseval's Identity 将变换后的关系转换为变换前的关系

$$\hat{R}(\hat{\tilde{u}}^m \ , \ \hat{\tilde{u}}^{m+1})\to R(\tilde{u}^m \ , \ \tilde{u}^{m+1})$$

Stability Analysis of the Explicit Euler Scheme

在本页我们讨论显式欧拉法的稳定性。首先考虑利用 SDFT 将数值解表示为如下形式，记 $I_{\Delta x} = [-\pi/\Delta x\ ,\ \pi/\Delta x]$

$$\tilde{u}_j^m = F^{-1} \circ F\left[\tilde{u}^m\right]_j = \frac{1}{2\pi}\int_{I_{\Delta x}} \hat{\tilde{u}}^m(\omega)\, e^{i\omega x_j} d\omega$$

带入欧拉法的关系式 $\frac{\tilde{u}_j^{m+1}-\tilde{u}_j^m}{\Delta t} = \frac{\tilde{u}_{j-1}^m - 2\tilde{u}_j^m + \tilde{u}_{j+1}^m}{\Delta x^2}$，得到

$$\begin{aligned}
\frac{1}{2\pi}\int_{I_{\Delta x}} \left[\hat{\tilde{u}}^{m+1} - \hat{\tilde{u}}^m\right] e^{i\omega x_j} d\omega &= \frac{\mu}{2\pi}\int_{I_{\Delta x}} \hat{\tilde{u}}^m \left[e^{i\omega x_{j-1}} - 2e^{i\omega x_j} + e^{i\omega x_{j+1}}\right] d\omega \\
&= \frac{\mu}{2\pi}\int_{I_{\Delta x}} \hat{\tilde{u}}^m \left[e^{-i\omega\Delta x} - 2 + e^{i\omega\Delta x}\right] e^{i\omega x_j} d\omega \\
&= \frac{\mu}{2\pi}\int_{I_{\Delta x}} 2\hat{\tilde{u}}^m \left[cos\left(\omega\Delta x\right) - 1\right] e^{i\omega x_j} d\omega
\end{aligned}$$

由于逆变换是一一映射，我们得到

$$\hat{\tilde{u}}^{m+1} - \hat{\tilde{u}}^m = 2\mu\hat{\tilde{u}}^m \left[cos\left(\omega\Delta x\right) - 1\right] = -4\mu sin^2\left(\tfrac{\omega\Delta x}{2}\right)\hat{\tilde{u}}^m$$

记 $\lambda(\omega) = 1 - 4\mu sin^2\left(\frac{\omega\Delta x}{2}\right)$ 为 amplification factor，有

$$\hat{\tilde{u}}^{m+1} = \lambda(\omega)\cdot\hat{\tilde{u}}^m$$

于是，结合 Discrete Parseval's Identity

$$\begin{aligned}
||\tilde{u}^{m+1}||_{\ell_2} &= \frac{1}{\sqrt{2\pi}}||\hat{\tilde{u}}^{m+1}||_{L_2(I_{\Delta x})} = \frac{1}{\sqrt{2\pi}}||\lambda(\omega)\cdot\hat{\tilde{u}}^m||_{L_2(I_{\Delta x})} \\
&\le \frac{1}{\sqrt{2\pi}}\max_{\omega\in I_{\Delta x}}|\lambda(\omega)|\cdot||\hat{\tilde{u}}^m||_{L_2(I_{\Delta x})} \\
&= \max_{\omega\in I_{\Delta x}}|\lambda(\omega)|\cdot||\tilde{u}^m||_{\ell_2}
\end{aligned}$$

为满足 $\forall m \ge 1\ :\ ||\tilde{u}^m||_{\ell_2} \le ||\tilde{u}^0||_{\ell_2}$，显然要有 $\max_{\omega\in I_{\Delta x}}|\lambda(\omega)| \le 1$，即等价于 $|\lambda(\omega)| \le 1\ ,\ \forall\ \omega \in I_{\Delta x}$。我们有

$$|\lambda(\omega)| \le 1 \iff \mu sin^2\left(\tfrac{\omega\Delta x}{2}\right) \le \tfrac{1}{2}$$

只要

$$\mu \le \frac{1}{2}\min_{\omega\in I_{\Delta x}}\left\{sin^{-2}\left(\tfrac{\omega\Delta x}{2}\right)\right\} = \frac{1}{2}$$

即可满足要求。

Stability Analysis of the Implicit Euler Scheme

在本页我们讨论隐式欧拉法的稳定性。同样利用 SDFT 将数值解表示为如下形式，记 $I_{\Delta x}=[-\pi/\Delta x\ ,\ \pi/\Delta x]$

$$\tilde{u}_j^m=F^{-1}\circ F\left[\tilde{u}^m\right]_j=\frac{1}{2\pi}\int_{I_{\Delta x}}\hat{\tilde{u}}^m(\omega)\,e^{i\omega x_j}d\omega$$

带入隐式欧拉法的关系式 $\frac{\tilde{u}_j^{m+1}-\tilde{u}_j^m}{\Delta t}=\frac{\tilde{u}_{j-1}^{m+1}-2\tilde{u}_j^{m+1}+\tilde{u}_{j+1}^{m+1}}{\Delta x^2}$，得到

$$\begin{aligned}\frac{1}{2\pi}\int_{I_{\Delta x}}\left[\hat{\tilde{u}}^{m+1}-\hat{\tilde{u}}^m\right]e^{i\omega x_j}d\omega&=\frac{\mu}{2\pi}\int_{I_{\Delta x}}\hat{\tilde{u}}^{m+1}\left[e^{i\omega x_{j-1}}-2e^{i\omega x_j}+e^{i\omega x_{j+1}}\right]d\omega\\&=\frac{\mu}{2\pi}\int_{I_{\Delta x}}2\hat{\tilde{u}}^{m+1}\left[cos\left(\omega\Delta x\right)-1\right]e^{i\omega x_j}d\omega\end{aligned}$$

由于逆变换是一一映射，得到

$$\hat{\tilde{u}}^{m+1}-\hat{\tilde{u}}^m=2\mu\hat{\tilde{u}}^{m+1}\left[cos\left(\omega\Delta x\right)-1\right]=-4\mu sin^2\left(\tfrac{\omega\Delta x}{2}\right)\hat{\tilde{u}}^{m+1}$$

记 $\lambda(\omega)=\left[1+4\mu sin^2\left(\frac{\omega\Delta x}{2}\right)\right]^{-1}$ 为 amplification factor，有

$$\hat{\tilde{u}}^{m+1}=\lambda(\omega)\cdot\hat{\tilde{u}}^m$$

此时显然恒有 $\lambda(\omega)\le 1$，结合 Discrete Parseval's Identity 总有

$$||\tilde{u}^{m+1}||_{\ell_2}\le\max_{\omega\in I_{\Delta x}}|\lambda(\omega)|\cdot||\tilde{u}^m||_{\ell_2}\le||\tilde{u}^m||_{\ell_2}$$

满足稳定性要求。

Theorem 5.3.4. Simple Observation and Application to θ-method

Finite Difference Scheme $P\left(\tilde{u}_{j+q}^{m+p}\right)=0$ 利用 SDFT 转换前后的关系满足

$$P\left(\tilde{u}_{j+q}^{m+p}\right)=0\quad\rightarrow\quad P\left(\hat{\tilde{u}}^{m+p}e^{i\omega q\Delta x}\right)=0$$

可利用此规律快速写出转换后的关系。比如 θ-method 利用 SDFT 转换变成

$$\hat{\tilde{u}}^{m+1}-\hat{\tilde{u}}^m=\mu\cdot(1-\theta)\,\hat{\tilde{u}}^m\left(e^{i\omega\Delta x}-2+e^{-i\omega\Delta x}\right)+\mu\cdot\theta\hat{\tilde{u}}^{m+1}\left(e^{i\omega\Delta x}-2+e^{-i\omega\Delta x}\right)$$

合并同类项 + 三角换元后得到 $\hat{\tilde{u}}^{m+1}=\lambda(\omega)\,\hat{\tilde{u}}^m$ 其中 $\mu=\frac{\Delta t}{\Delta x^2}$ 为 CFL number，$\lambda(\omega)$ 为 amplification factor，且 $\forall\,\omega\in I_{\Delta x}=[-\pi/\Delta x\ ,\ \pi/\Delta x]$

$$\lambda(\omega)=\frac{1-4\mu(1-\theta)\,sin^2\left(\frac{\omega\Delta x}{2}\right)}{1+4\mu\theta sin^2\left(\frac{\omega\Delta x}{2}\right)}$$

Theorem 5.3.5. Stability of θ-method
记 $\tilde{u}^m$ 为定义在 $\{x_j \ : \ j=0,\pm1,\cdots\}$ 上的方程 5.2 (1) 用 θ-method 得到的数值解。记 $\mu=\frac{\Delta t}{\Delta x^2}$ 为 CFL number，则

1) 若 $\theta\in[\frac{1}{2},1]$，数值解无条件稳定 unconditionally practically stable;
2) 若 $\theta\in[0,\frac{1}{2})$，数值解在 $\mu<\frac{1}{2(1-2\theta)}$ 时条件稳定 conditionally practically stable。

Stability Analysis of the θ-scheme
在 thm5.3.4 中我们已经得到 θ-scheme 的 amplification factor 为 $\forall\,\omega\in I_{\Delta x}=[-\pi/\Delta x\ ,\ \pi/\Delta x]$

$$\lambda(\omega)=\frac{1-4\mu(1-\theta)\,sin^2\left(\frac{\omega\Delta x}{2}\right)}{1+4\mu\theta sin^2\left(\frac{\omega\Delta x}{2}\right)}$$

结合 Discrete Parseval's Identity 总有

$$||\tilde{u}^{m+1}||_{\ell_2}\le\max_{\omega\in I_{\Delta x}}|\lambda(\omega)|\cdot||\tilde{u}^m||_{\ell_2}$$

为使解稳定，我们要求

$$|\lambda(\omega)|\le1\iff-1\le\lambda(\omega)\le1$$

不难证明右侧不等号总是成立，于是只要有

$$\lambda(\omega)\ge-1\ \Rightarrow\ 2\mu(1-2\theta)\le[sin^2(\tfrac{\omega\Delta x}{2})]^{-1}$$

即只要有 $2\mu(1-2\theta)\le\min_{\omega\in I_{\Delta x}}[sin^2(\frac{\omega\Delta x}{2})]^{-1}=1$ 即可。显然当 $\theta\in[\frac{1}{2},1]$ 时，$2\mu(1-2\theta)\le0\le1$ 总成立。

Definition 5.3.3. Von Neumann Stability 我们可以进一步放宽对稳定性的要求。记 $\tilde{u}_m=(\tilde{u}_j^m)$ 为 5.2 (1) 中热方程的数值解，然而此时要求时间区间有界 T；有 $m=0\cdots M=\frac{T}{\Delta t}$ 为时间上标，$j=0,\pm1\cdots$ 为空间下标，时间和空间节点的离散化步长分别为 Δt , Δx。若数值解对某个与 T 有关的常数 *stability constant* $C(T)$ 满足

$$\forall m\ge1\ :\ ||\tilde{u}^m||_{\ell_2}\le C(T)\,||\tilde{u}^0||_{\ell_2}$$

则称数值解 *Von Neumann Stable*。一般来说常数 $C(T)\to+\infty$ as $T\to+\infty$，故此稳定性只在有限时间区间上有意义。

Theorem 5.3.6. Von Neumann Stability 的判定准则
若 $\tilde{u}^m$ 为定义在 $\{x_j \ : \ j = 0, \pm 1, \cdots\}$ 上的方程 5.2 (1) 用某差分策略 $P\left(\tilde{u}_{j+q}^{m+p}\right) = 0$ 得到的数值解。若此差分策略利用 SDFT 转换后的关系满足

$$\hat{\tilde{u}}^{m+1} = \lambda(\omega)\, \hat{\tilde{u}}^m$$

且存在非负常数 $C_0 \geq 0$ 使得

$$\forall\ \omega \in I_{\Delta x} \ : \ |\lambda(\omega)| \leq 1 + C_0 \Delta t$$

则该策略对于 stability constant $C(T) = e^{C_0 T}$ Von Neumann stable；当 $C_0 = 0$ 时显然 practically stable。其中 $I_{\Delta x} = [-\pi/\Delta x \ , \ \pi/\Delta x]$，$\lambda(\omega)$ 称为 amplification factor。

Proof theorem 5.3.6 由 Discrete Parseval's Identity 总有

$$||\tilde{u}^{m+1}||_{\ell_2} \leq \max_{\omega \in I_{\Delta x}} |\lambda(\omega)| \cdot ||\tilde{u}^m||_{\ell_2}$$

因此

$$\begin{aligned} ||\tilde{u}^m||_{\ell_2} &\leq \max_{\omega \in I_{\Delta x}} |\lambda(\omega)| \cdot ||\tilde{u}^{m-1}||_{\ell_2} \leq (1 + C_0 \Delta t) \cdot ||\tilde{u}^{m-1}||_{\ell_2} \\ &\leq (1 + C_0 \Delta t)^m \cdot ||\tilde{u}^0||_{\ell_2} \leq \left(e^{C_0 \Delta t}\right)^m \cdot ||\tilde{u}^0||_{\ell_2} \\ &= e^{C_0 m \Delta t} \cdot ||\tilde{u}^0||_{\ell_2} = e^{C_0 T} ||\tilde{u}^0||_{\ell_2} \end{aligned}$$

上述第四个不等号利用了 $1 + x \leq e^x$ 。于是得到 Von Neumann Stability $||\tilde{u}^m||_{\ell_2} \leq e^{C_0 T} ||\tilde{u}^0||_{\ell_2}$。

5.4 Boundary Value Problems of Parabolic Problems

正如前文所述，虽然隐式法稳定性更好，但在实际操作中往往需要对空间坐标施加限制方才能够使用。为此，我们通常考虑如下的 Initial Boundary Value Problem IBVP

Definition 5.4.1. IBVP of Parabolic Heat Equation
考虑求解问题为对 $u(x,t)$，$x \in \mathbb{R}$，$a \leq x \leq b$，$0 \leq t \leq T$ 满足

$$u_t = u_{xx} + f(x,t) \quad \text{s.t.} \quad u(x,0) = u_0(x) \tag{2}$$

并且对 $x = a, b$ 额外附加边界条件

1) **Dirichlet** $u(a,t) = A(t)$ ， $u(b,t) = B(t)$ for $t \in (0,T]$
2) **Neumann** $u_x(a,t) = A(t)$ ， $u_x(b,t) = B(t)$ for $t \in (0,T]$
3) **Dirichlet-Neumann** $u(a,t) = A(t)$ ， $u_x(b,t) = B(t)$ for $t \in (0,T]$

对于 Dirichlet BC，在本章中总是认为其与初值条件是相容的，即 $A(0) = u_0(a)$ ，$B(0) = u_0(b)$。此时，我们取离散化网格

$$\begin{aligned} t_m &= m\Delta t \qquad \text{for} \quad \Delta t = T/M \ , \ m = 0 \cdots M \\ x_j &= a + j\Delta x \quad \text{for} \quad \Delta x = (b-a)/J \ , \ j = 0 \cdots J \end{aligned}$$

并记得到的数值解为 $\tilde{u}_j^m \approx u(x_j, t_m)$。在未经特殊说明的情况下，后文中将默认沿用这些记号规定。

Theorem 5.4.1. θ-method for Dirichlet IBVP of Parabolic Heat Equation 在上述问题 (2) 的设置下，取 Dirichlet BC，θ-method 的每一步解 $\tilde{u}_{1:(J-1)}^m$ 可以通过求解如下线性方程组得到 for $m = 0 \cdots (M-1)$:

$$(I - \theta\mu\mathcal{A})\, \tilde{u}_{1:(J-1)}^{m+1} = \left(I + (1-\theta)\,\mu\mathcal{A}\right) \tilde{u}_{1:(J-1)}^m + \theta \boldsymbol{f}^{m+1} + (1-\theta)\,\boldsymbol{f}^m$$

注意等式的右侧均为已知量，其中 $\mu = \frac{\Delta t}{\Delta x^2}$，$\mathcal{A} = \mathrm{diag}_{J-1}(1, -2, 1)$

$$\boldsymbol{f}^m = \left[f_1^m + \mu A(t_m) \ , \ f_2^m \ , \ \cdots \ , \ f_{J-2}^m \ , \ f_{J-1}^m + \mu B(t_m)\right]^T \in \mathbb{R}^{J-1}$$

。该结论只涉及简单的矩阵变换，证明见 Lecture Notes p.61。

5.5 Maximum Principle and Convergence of $\theta-$methods

未经特别说明，默认本节讨论的对象均为 homogeneous Dirichlet IBVP for the heat equation，i.e，考虑求解问题为对 $u(x,t), x \in \mathbb{R}, a \le x \le b, 0 \le t \le T$ 满足

$$u_t = u_{xx} \quad \text{s.t.} \quad u(x,0) = u_0(x)$$

并且对 $x = a, b$ 额外附加 Dirichlet 边界条件

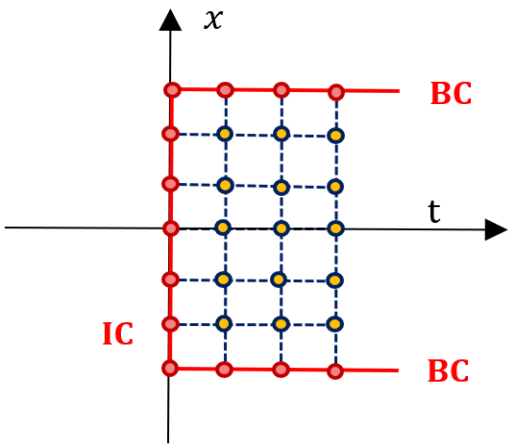
$$u(a,t) = A(t) \quad , \quad u(b,t) = B(t) \quad \text{for} \quad t \in (0,T]$$

规定如下记号，分别表示数值解 $\tilde{u}$ 在三个边界（两个空间边界 $x = a, b$，一个时间边界 $t = 0$）上的最大、最小值

$$\tilde{u}_{\min} = \min\{\ \tilde{u}_j^m \ : \ j \in \{0, J\} \ \text{ or } \ m = 0\ \}$$

$$\tilde{u}_{\max} = \max\{\ \tilde{u}_j^m \ : \ j \in \{0, J\} \ \text{ or } \ m = 0\ \}$$

Theorem 5.5.1. Discrete Maximum Principle for θ-schemes
考虑 Dirichlet IBVP for the heat equation，$\tilde{u}$ 是使用 θ-scheme 得到的数值解，$\mu = \Delta t/\Delta x^2$ 为 CFL number。当 $\mu \le \frac{1}{2(1-\theta)}$ 时，$\tilde{u}_{\min} \le \tilde{u}_j^m \le \tilde{u}_{\max}$ 恒成立。当满足该 *Discrete Maximum Principle* 时，θ-scheme 必然 practically stable，*i.e.*，$\mu \le \frac{1}{2(1-2\theta)}$。



Proof theorem 5.5.1 将 θ-scheme 整理成如下形式

$$(1+2\theta\mu)\,\tilde{u}_j^{m+1} = \theta\mu\left(\tilde{u}_{j+1}^{m+1}+\tilde{u}_{j-1}^{m+1}\right)+(1-\theta)\,\mu\left(\tilde{u}_{j+1}^{m}+\tilde{u}_{j-1}^{m}\right)+\left[1-2\,(1-\theta)\,\mu\right]\tilde{u}_j^{m}$$

假设 $\tilde{u}_j^{m+1}$ 为全局最大点，记 $\tilde{u}^*$ 为 $\tilde{u}_j^{m+1}$ 邻接点中值最大的点，即

$$\tilde{u}^* = \max\{\tilde{u}_{j+1}^{m}, \tilde{u}_j^{m}, \tilde{u}_{j-1}^{m}, \tilde{u}_{j+1}^{m+1}, \tilde{u}_{j-1}^{m+1}\}$$

结合 $1-2\,(1-\theta)\,\mu \geq 0$ 有

$$(1+2\theta\mu)\,\tilde{u}_j^{m+1} \leq 2\theta\mu\tilde{u}^* + 2\,(1-\theta)\,\mu\tilde{u}^* + \left[1-2\,(1-\theta)\,\mu\right]\tilde{u}^* = (1+2\theta\mu)\,\tilde{u}^*$$

故 $\tilde{u}^* \leq \tilde{u}_j^{m+1} \leq \tilde{u}^* \iff \tilde{u}_j^{m+1} = \tilde{u}^*$。所以 $\tilde{u}^*$ 也是全局最大点；对其重复上述步骤直至取到新的 $\tilde{u}^*$ 为边界点即可。同理可证 Minimum 的情形。对于稳定性，当 $\theta \in [1/2, 1]$ 时，方法无条件稳定。当 $\theta \in [0, 1/2)$ 时，$\mu \leq \frac{1}{2(1-\theta)} \leq \frac{1}{2(1-2\theta)}$ 故满足 Maximum Principle 必然条件稳定。

Definition 5.5.1. Maximum Norm
规定形如数值解 $\tilde{u}$ 的离散函数有 *maximum norm*

$$||\tilde{u}||_\infty = \max_{m,j} |\tilde{u}_j^m|$$

记真解与误差分别为 $\boldsymbol{u}$ 与 $\boldsymbol{e} = \boldsymbol{u} - \tilde{u}$，对于 θ-scheme 有如下收敛性定理。

Theorem 5.5.2. $\theta-$method 的收敛性
考虑 homogeneous Dirichlet IBVP for the heat equation，$\tilde{u}$ 是使用 θ-scheme 得到的数值解，$\mu = \Delta t/\Delta x^2$ 为 CFL number；对 $j = 1\cdots(J-1)$，$m = 0\cdots(M-1)$

$$T_j^m = D_t^+ u_j^m - (1-\theta)\,D_{xx}^c u_j^m - \theta D_{xx}^c u_j^{m+1}$$

为 consistency error，记 $\boldsymbol{T} = (T_j^m)$。则当 $\mu \leq \frac{1}{2(1-\theta)}$ 满足 maximum principle 时，

$$||\boldsymbol{e}||_\infty = ||\boldsymbol{u} - \tilde{u}||_\infty \leq T||\boldsymbol{T}||_\infty$$

注意上式中，$T = M\Delta t$ 表示时间尺度上限；$\boldsymbol{e}$ 和 $\boldsymbol{T}$ 的长度不一样，前者上下标范围分别为 $j = 0\cdots J$，$m = 0\cdots M$。于是有

$$||\boldsymbol{e}||_\infty = ||\boldsymbol{u} - \tilde{u}||_\infty = \begin{cases} O(\Delta x^2 + \Delta t^2) & \text{if } \theta = 1/2 \\ O(\Delta x^2 + \Delta t) & \text{if } \theta \neq 1/2 \end{cases}$$

Proof theorem 5.5.2 将 θ-scheme 与其截断误差分别整理成如下形式

$$(1+2\theta\mu)\tilde{u}_j^{m+1} = \theta\mu(\tilde{u}_{j+1}^{m+1}+\tilde{u}_{j-1}^{m+1})+(1-\theta)\mu(\tilde{u}_{j+1}^{m}+\tilde{u}_{j-1}^{m})+\left[1-2(1-\theta)\mu\right]\tilde{u}_j^m$$

$$(1+2\theta\mu)\boldsymbol{u}_j^{m+1} = \theta\mu(\boldsymbol{u}_{j+1}^{m+1}+\boldsymbol{u}_{j-1}^{m+1})+(1-\theta)\mu(\boldsymbol{u}_{j+1}^{m}+\boldsymbol{u}_{j-1}^{m})+\left[1-2(1-\theta)\mu\right]\boldsymbol{u}_j^m+\Delta t T_j^m$$

注意在边界上的误差 $\boldsymbol{e}_j^m$ 总是 0，我们主要考察 $j = 1\cdots(J-1)$，$m = 0\cdots(M-1)$ 时的情况，将上面两个式子相减得到

$$(1+2\theta\mu)\boldsymbol{e}_j^{m+1} = \theta\mu(\boldsymbol{e}_{j+1}^{m+1}+\boldsymbol{e}_{j-1}^{m+1})+(1-\theta)\mu(\boldsymbol{e}_{j+1}^{m}+\boldsymbol{e}_{j-1}^{m})+\left[1-2(1-\theta)\mu\right]\boldsymbol{e}_j^m+\Delta t T_j^m$$

不妨记 $||\boldsymbol{e}^m||_\infty = \max_j|\boldsymbol{e}_j^m|$，$||T^m||_\infty = \max_j|T_j^m|$（注意下标 j 的范围区别），由 $\mu \le \frac{1}{2(1-\theta)} \iff 1-2(1-\theta)\mu \ge 0$ 以及上述误差不等式，我们自然有

$$\begin{aligned}
& (1+2\theta\mu)|\boldsymbol{e}_j^{m+1}| \\
= & \left|\theta\mu(\boldsymbol{e}_{j+1}^{m+1}+\boldsymbol{e}_{j-1}^{m+1})+(1-\theta)\mu(\boldsymbol{e}_{j+1}^{m}+\boldsymbol{e}_{j-1}^{m})+\left[1-2(1-\theta)\mu\right]\boldsymbol{e}_j^m+\Delta t T_j^m\right| \\
\le & 2\theta\mu||\boldsymbol{e}^{m+1}||_\infty+\left[2(1-\theta)\mu+1-2(1-\theta)\mu\right]\cdot||\boldsymbol{e}^m||_\infty+\Delta t||T^m||_\infty \\
\le & 2\theta\mu||\boldsymbol{e}^{m+1}||_\infty+||\boldsymbol{e}^m||_\infty+\Delta t||T^m||_\infty
\end{aligned}$$

对所有 $j = 1\cdots(J-1)$ 成立；由于 $\boldsymbol{e}_0^{m+1} \equiv \boldsymbol{e}_J^{m+1} \equiv 0$ 得到

$$(1+2\theta\mu)||\boldsymbol{e}^{m+1}||_\infty \le 2\theta\mu||\boldsymbol{e}^{m+1}||_\infty+||\boldsymbol{e}^m||_\infty+\Delta t||T^m||_\infty$$

即 $||\boldsymbol{e}^{m+1}||_\infty - ||\boldsymbol{e}^m||_\infty \le \Delta t||T^m||_\infty$ 对 $m = 0\cdots(M-1)$ 成立，叠加后由 $||\boldsymbol{e}^0||_\infty = 0$ 可得

$$||\boldsymbol{e}^m||_\infty \le \Delta t\sum_{n=0}^{m-1}||T^n||_\infty \le \Delta t\cdot||\boldsymbol{T}||_\infty m \le M\Delta t||\boldsymbol{T}||_\infty = T||\boldsymbol{T}||_\infty$$

对 $\forall\, m = 1\cdots M$ 均成立，显然对 $m = 0$ 也成立。不妨记 $m^* = \mathrm{argmax}_m||\boldsymbol{e}^m||_\infty$，有 $||\boldsymbol{e}||_\infty = ||\boldsymbol{e}^{m^*}||_\infty \le T||\boldsymbol{T}||_\infty$ 故结论得证。结合 thm 5.2.1 不难直接证明

$$||\boldsymbol{e}||_\infty = \begin{cases} O(\Delta x^2+\Delta t^2) & \text{if } \theta = 1/2 \\ O(\Delta x^2+\Delta t) & \text{if } \theta \ne 1/2 \end{cases}$$

也成立。

5.6 Parabolic Equations in Two Space Dimensions

未经特别说明，默认本节讨论的对象均为 homogeneous Dirichlet IBVP for the heat equation，i.e，考虑求解问题为对 $u(x,y,t)$，$(x,y)\in\Omega=(a,b)\times(c,d)$，$0\le t\le T$ 满足

$$u_t = u_{xx}+u_{yy} \quad \text{s.t.} \quad u(x,y,0)=u_0(x,y) \tag{3}$$

并且对 $x\in\partial\Omega$ 额外附加 Dirichlet 边界条件

$$u|_{\partial\Omega} = B(x,y,t) \quad \text{for} \quad (x,y)\in\partial\Omega\ ;\ t\in(0,T]$$

考虑对其离散化网格如下

$$\begin{aligned} t_m &= m\Delta t && \text{for} && \Delta t = T/M\ ,\ m=0\cdots M \\ x_i &= a+i\Delta x && \text{for} && \Delta x=(b-a)/J_x\ ,\ i=0\cdots J_x \\ y_j &= c+j\Delta y && \text{for} && \Delta y=(c-d)/J_y\ ,\ j=0\cdots J_y \end{aligned}$$

记得到的数值解 $\tilde{u}_{ij}^m\approx u(x_i,y_j,t_m)=\boldsymbol{u}_{ij}^m$，其中 $\tilde{u}_{ij}^m=B(x_i,y_j,t_m)=\boldsymbol{u}_{ij}^m$ 对边界节点 $(x_i,y_j)\in\partial\Omega_h$ 恒成立；$\tilde{u}_{ij}^0=u_0(x_i,y_j)=\boldsymbol{u}_{ij}^0$ 对边界节点 $t_m=0$ 恒成立。此外，记两个空间坐标的 CFL number 分别为 $\mu_x=\Delta t/\Delta x^2$ 以及 $\mu_y=\Delta t/\Delta y^2$。

Definition 5.6.1. θ**-schemes**
问题 (3) 的 θ-schemes 为对 $i=1\cdots(J_x-1)$，$j=1\cdots(J_y-1)$ 以及 $m=0\cdots(M-1)$ 满足

$$D_t^+\tilde{u}_{ij}^m=(1-\theta)\left[D_{xx}^c\tilde{u}_{ij}^m+D_{yy}^c\tilde{u}_{ij}^m\right]+\theta\left[D_{xx}^c\tilde{u}_{ij}^{m+1}+D_{yy}^c\tilde{u}_{ij}^{m+1}\right]$$

$\theta\in[0,1]$ 是一个参数。当 $\theta=0$ 时为 Explicit Euler Method；当 $\theta=1$ 时为 Implicit Euler Method；当 $\theta=1/2$ 时为 Crank-Nicolson Method

1) **Explicit Euler** $D_t^+\tilde{u}_{ij}^m=D_{xx}^c\tilde{u}_{ij}^m+D_{yy}^c\tilde{u}_{ij}^m$
2) **Implicit Euler** $D_t^-\tilde{u}_{ij}^{m+1}:=D_t^+\tilde{u}_{ij}^m=D_{xx}^c\tilde{u}_{ij}^{m+1}+D_{yy}^c\tilde{u}_{ij}^{m+1}$
3) **Crank-Nicolson** $D_t^+\tilde{u}_{ij}^m=\frac{1}{2}\left[D_{xx}^c\tilde{u}_{ij}^m+D_{yy}^c\tilde{u}_{ij}^m\right]+\frac{1}{2}\left[D_{xx}^c\tilde{u}_{ij}^{m+1}+D_{yy}^c\tilde{u}_{ij}^{m+1}\right]$

下面逐一列举双空间变量下 θ-schemes 对问题 (3) 的稳定性、maximum principle，以及收敛性定理。结论与推导和一维情形类似，可详见 **Lecture Notes pp.67-70**。

Theorem 5.6.1. Stability of θ-method

记 $\tilde{u}^m$ 为问题 (3) 在空间坐标无界时，使用 θ-schemes 得到的数值解，则

1) 若 $\theta \in [\frac{1}{2}, 1]$，数值解无条件稳定 unconditionally practically stable；
2) 若 $\theta \in [0, \frac{1}{2})$，数值解在 $\mu_x + \mu_y < \frac{1}{2(1-2\theta)}$ 时条件稳定 conditionally practically stable。

规定数值解 $\tilde{u}$ 在边界上的最大、最小值为

$$\tilde{u}_{\min} = \min\{\ \tilde{u}_{ij}^m \ :\ (x_i, y_j) \in \partial\Omega_h \ \text{ or } \ m = 0\ \}$$

$$\tilde{u}_{\max} = \max\{\ \tilde{u}_{ij}^m \ :\ (x_i, y_j) \in \partial\Omega_h \ \text{ or } \ m = 0\ \}$$

有 Discrete Maximum Principle

Theorem 5.6.2. Discrete Maximum Principle for θ-schemes

记 $\tilde{u}^m$ 为问题 (3) 使用 θ-schemes 得到的数值解。当 $\mu_x + \mu_y \le \frac{1}{2(1-\theta)}$ 时，$\tilde{u}_{\min} \le \tilde{u}_{ij}^m \le \tilde{u}_{\max}$ 恒成立。当满足该 Discrete Maximum Principle 时，θ-scheme 必然 practically stable，*i.e.*，$\mu_x + \mu_y \le \frac{1}{2(1-2\theta)}$。

Theorem 5.6.3. θ-method 的精度

定义 θ-schemes 的一致性误差/截断误差 consistency/truncation error 为

$$T_{ij}^m = D_t^+ u_{ij}^m - (1-\theta)\left[D_{xx}^c u_{ij}^m + D_{yy}^c u_{ij}^m\right] - \theta\left[D_{xx}^c u_{ij}^{m+1} + D_{yy}^c u_{ij}^{m+1}\right]$$

其中 u_{ij}^m 表示真值；我们有

$$T_{ij}^m = \begin{cases} O(\Delta x^2 + \Delta y^2 + \Delta t^2) & \text{if } \ \theta = 1/2 \\ O(\Delta x^2 + \Delta y^2 + \Delta t) & \text{if } \ \theta \neq 1/2 \end{cases}$$

证明将 T_{ij}^m 在 $(x_i, y_j, t_{m+\frac{1}{2}})$ 做展开即可，省略。

Definition 5.6.2. Maximum Norm
规定形如数值解 $\tilde{u} = \left(\tilde{u}_{ij}^{m}\right)$ 的离散函数有 maximum norm

$$||\tilde{u}||_{\infty} = \max_{m,i,j} |\tilde{u}_{j}^{m}|$$

记真解与误差分别为 $\boldsymbol{u}$ 与 $\boldsymbol{e} = \boldsymbol{u} - \tilde{u}$，对于 θ-scheme 有如下收敛性定理。

Theorem 5.6.4. $\theta-$method 的收敛性
记 $\tilde{u}^{m}$ 为问题 (3) 使用 θ-schemes 得到的数值解；对 $i = 1 \cdots (J_x - 1)$，$j = h1 \cdots \left(J_y - 1\right)$，$m = 0 \cdots (M-1)$，记 $\boldsymbol{T} = (T_{ij}^{m})$，其中 T_{ij}^{m} 为 consistency error。则当 $\mu_x + \mu_y \leq \frac{1}{2(1-\theta)}$ 满足 maximum principle 时，

$$||\boldsymbol{e}||_{\infty} = ||\boldsymbol{u} - \tilde{u}||_{\infty} \leq T||\boldsymbol{T}||_{\infty}$$

上式中，$T = M\Delta t$ 表示时间尺度上限。有

$$||\boldsymbol{e}||_{\infty} = ||\boldsymbol{u} - \tilde{u}||_{\infty} = \begin{cases} O(\Delta x^2 + \Delta y^2 + \Delta t^2) & \text{if } \theta = 1/2 \\ O(\Delta x^2 + \Delta y^2 + \Delta t) & \text{if } \theta \neq 1/2 \end{cases}$$

6 Finite Difference Schemes for Hyperbolic PDEs

6.1 Introduction

本节中我们将主要考虑如下的双曲型 IBVP $(*)$，即对 $u(x,t)$ s.t. $(x,t) \in [a,b] \times [0,T] = \Omega$ 满足

$$
\begin{aligned}
u_{tt} - c^2 u_{xx} &= f(x,t) && \text{for} && (x,t) \in (a,b) \times (0,T] && (1)\\
u(x,0) &= g_0(x) && \text{for} && x \in [a,b] && (2)\\
u_t(x,0) &= g_1(x) && \text{for} && x \in [a,b] && (3)\\
u(a,t) &= u(b,t) = 0 && \text{for} && t \in [0,T] && (4)
\end{aligned}
$$

其中 f，g_0，g_1 均为在定义域上的连续函数，且 g_0，g_1 在端点 a，b 处取值为 0；$c > 0$ 为 wave speed。我们假设该问题的解 u 存在，唯一，且在定义域 Ω 内足够光滑。

在本节中，我们将给出该 IBVP 下 Energy Inequality/Estimate 的推导证明。在后文中，我们将会考虑隐式和显式两种差分策略，它们的稳定性分析将在很大程度上基于离散版本的 Energy Inequality。

Lemma 6.1.1. Gronwall's Lemma
若 $A(t)$，$B(t)$ 均是定义在 $[0,T]$ 上的非负函数，且 B 为 nondecreasing，则

$$
\forall\, t \;:\; A(t) \le B(t) + \int_0^t A(s)\,ds \quad \Rightarrow \quad \forall\, t \;:\; A(t) \le e^t B(t)
$$

Theorem 6.1.1. Energy Inequality
考虑双曲型 IBVP $(*)$，对任意 $t \in [0,T]$，都有以下不等式恒成立

$$
\mathcal{L}^2[u(\cdot,t)] \le e^t \left\{ \mathcal{L}^2[u(\cdot,0)] \;+\; ||f||^2_{L_2((a,b)\times(0,t))} \right\}
$$

$$
\text{s.t.} \quad \mathcal{L}^2[u(\cdot,t)] := \int_a^b u_t^2 + c^2 u_x^2 \; dx \quad ; \quad \mathcal{L}^2[u(\cdot,0)] = ||g_1||^2_{L_2(a,b)} + c^2 |g_0|^2_{H^1(a,b)}
$$

不难证明，给定任意一个 $t \in [0,T]$，$\mathcal{L}[u] = \sqrt{\mathcal{L}^2[u]}$ 为一个定义在 $H^1((a,b)\times(0,T))$ 上的 norm。此外，$u \mapsto \max_{t\in[0,T]} \mathcal{L}[u](t)$ 也是一个 norm，且与 t 的取值无关。

Proof lemma 6.1.1 观察到 $F_A(t)=\int_0^t A(s)\,ds$ 是 $A(t)$ 的原函数，所以对原不等式两边同时乘上 e^{-t}

$$\begin{aligned} e^{-t}B(t) &\geq A(t)e^{-t}-F_A(t)e^{-t} \\ &\geq F_A'(t)e^{-t}+F_A(t)(e^{-t})'=\frac{\mathrm{d}}{\mathrm{d}t}\left[F_A(t)e^{-t}\right] \end{aligned}$$

两边同时从 0 到 t 积分

$$\int_0^t e^{-s}B(s)\,ds \geq \int_0^t \frac{\mathrm{d}}{\mathrm{d}s}\left[F_A(s)e^{-s}\right]\,ds = F_A(t)e^{-t}-F_A(0)e^{-0}$$

显然，$F_A(0)=\int_0^0 A(s)\,ds=0$，又有在 $s\in[0,t]$ 上 $B(t)\geq B(s)$，故

$$F_A(t)e^{-t} \;\leq\; \int_0^t e^{-s}B(s)\,ds \;\leq\; B(t)\int_0^t e^{-s}ds = B(t)\left(1-e^{-t}\right)$$

故 $\int_0^t A(s)\,ds=F_A(t)\leq B(t)(e^t-1)$，由原不等式 $A(t)\leq B(t)+\int_0^t A(s)\,ds$ 易得

$$A(t)\leq B(t)+\int_0^t A(s)\,ds\leq B(t)+B(t)(e^t-1)=B(t)e^t$$

Proof theorem 6.1.1 显然，因为在空间边界处恒有 $u(a,t)=u(b,t)=0$，$u_t(x,t)$ 对变量 $x\in[a,b]$ 在边界上恒有 $u_t(a,t)=u_t(b,t)=0$。对原 PDE 两边同乘 $u_t(x,t)$ 后对 x 在 $[a,b]$ 上积分，得到

$$\int_a^b u_{tt}u_t\,dx - c^2\int_a^b u_{xx}u_t\,dx = \int_a^b u_{tt}u_t\,dx + c^2\int_a^b u_xu_{xt}\,dx = \int_a^b fu_t\,dx$$

其中第二个等号利用了分部积分；又由链式法则，$\frac{1}{2}\partial_t(u_t^2)=u_{tt}u_t$，$\frac{1}{2}\partial_t(u_x^2)=u_{xt}u_x$，于是上式进一步化为

$$\frac{1}{2}\int_a^b \partial_t(u_t^2)\,dx + \frac{c^2}{2}\int_a^b \partial_t(u_x^2)\,dx = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_a^b u_t^2+c^2u_x^2\,dx = \int_a^b fu_t\,dx$$

即 $\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{L}^2[u(\cdot,t)]=\int_a^b fu_t\,dx$，两边同时对 $\tau:=t$ 在 $[0,t]$ 上积分

$$\frac{1}{2}\mathcal{L}^2[u(\cdot,t)]-\frac{1}{2}\mathcal{L}^2[u(\cdot,0)]=\int_0^t\int_a^b f(x,\tau)u_t(x,\tau)\,dxd\tau$$

于是

$$\mathcal{L}^2[u(\cdot,t)]=\mathcal{L}^2[u(\cdot,0)]+2\int_0^t\int_a^b f(x,\tau)u_t(x,\tau)\,dxd\tau$$

由不等式 $2ab \leq a^2 + b^2$ 对最后一项放缩可得

$$\begin{aligned}\mathcal{L}^2[u(\cdot,t)] &= \mathcal{L}^2[u(\cdot,0)] + 2\int_0^t \int_a^b f(x,\tau)u_t(x,\tau)\ dxd\tau \\ &\leq \mathcal{L}^2[u(\cdot,0)] + \int_0^t \int_a^b f^2(x,\tau)\ dxd\tau + \int_0^t \int_a^b u_t^2(x,\tau)\ dxd\tau \\ &\leq \mathcal{L}^2[u(\cdot,0)] + ||f||^2_{L_2((a,b)\times(0,t))} + \int_0^t \mathcal{L}^2[u(\cdot,\tau)]d\tau\end{aligned}$$

由 lemma 6.1.1，易得 $A(t) = \mathcal{L}^2[u(\cdot,t)]$，$B(t) = \mathcal{L}^2[u(\cdot,0)] + ||f||^2_{L_2((a,b)\times(0,t))}$ 故有

$$\mathcal{L}^2[u(\cdot,t)] \leq e^t \left\{ \mathcal{L}^2[u(\cdot,0)] \ + \ ||f||^2_{L_2((a,b)\times(0,t))} \right\}$$

得证。下面我们证明给定任意一个 $t \in [0,T]$，$\mathcal{L}[u] = \sqrt{\mathcal{L}^2[u]}$ 为一个定义在 $H^1((a,b)\times(0,T))$ 上的 norm：显然 $\mathcal{L}[u](t) \geq 0$ 与 $\mathcal{L}[au](t) = |a|\mathcal{L}[u](t)$ 总成立，我们只证明 $\mathcal{L}[u](t) = 0 \iff u \equiv 0$ 与 $\mathcal{L}[u+v](t) \leq \mathcal{L}[u](t) + \mathcal{L}[v](t)$。

当 $\mathcal{L}[u](t) = 0$ 时，必有 $u_t(x,t) = u_x(x,t) = 0$ 对 $x \in [a,b]$ 恒成立。于是对 $u_x(x,t) = 0$ 两边取 x 的积分 $u(x,t) = c(t)$。又 $u_t(x,t) = c'(t) = 0$，有 $u(x,t) = c(t) \equiv \text{const}$。由边界条件 $u(a,t) = u(b,t) = 0$ 恒成立，$\text{const} = 0$，即 $u(x,t) = 0$ 对所取的 t 在 $x \in [a,b]$ 恒成立。

最后证明三角不等式 $\mathcal{L}[u+v](t) \leq \mathcal{L}[u](t) + \mathcal{L}[v](t)$，显然

$$\begin{aligned}\mathcal{L}[u+v] &= \sqrt{||u_t + v_t||^2_{L_2(a,b)} + c^2||u_x + v_x||^2_{L_2(a,b)}} \\ &\leq \sqrt{||u_t||^2_{L_2(a,b)} + ||v_t||^2_{L_2(a,b)} + c^2||u_x||^2_{L_2(a,b)} + c^2||v_x||^2_{L_2(a,b)}} \\ &\leq \sqrt{\mathcal{L}^2[u] + \mathcal{L}^2[v]} \leq \sqrt{\mathcal{L}^2[u]} + \sqrt{\mathcal{L}^2[v]} = \mathcal{L}[u] + \mathcal{L}[v]\end{aligned}$$

最后一个不等式用了 $a,b>0 : \sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$。

Energy Inequality 说明了，该 IBVP 的解 u changes continuously w.r.t the initial data g_0, g_1, f，因此该问题在解存在且唯一时 well-posed。

6.2 Implicit and Explicit Schemes

考虑双曲型 IBVP $(*)$，即对 $u(x,t)$ s.t. $(x,t)\in[a,b]\times[0,T]=\Omega$ 满足

$$
\begin{aligned}
&u_{tt}-c^2u_{xx}=f(x,t) && \text{for} && (x,t)\in(a,b)\times(0,T] && (1)\\
&u(x,0)=g_0(x) && \text{for} && x\in[a,b] && (2)\\
&u_t(x,0)=g_1(x) && \text{for} && x\in[a,b] && (3)\\
&u(a,t)=u(b,t)=0 && \text{for} && t\in[0,T] && (4)
\end{aligned}
$$

其中 f，g_0，g_1 均为在定义域上的连续函数，且 g_0，g_1 在端点 a，b 处取值为 0；$c>0$ 为 wave speed。解 u 存在，唯一，且在定义域 Ω 内足够光滑。我们提供隐式与显式两种差分策略，在介绍这两种策略前，先引入如下的差分网格 $\overline{\Omega}_{\Delta x\times\Delta t}$。

$$
\begin{aligned}
&t_m=m\Delta t && \text{for} && \Delta t=T/M\ ,\ m=0\cdots M\\
&x_j=j\Delta x && \text{for} && \Delta x=(b-a)/J\ ,\ j=0\cdots J
\end{aligned}
$$

并记内部节点为 $\Omega_{\Delta x\times\Delta t}=\Omega_{\Delta x}\times\Omega_{\Delta t}$，其中

$$
\begin{aligned}
&\Omega_{\Delta x}=\{x_j:j=1\cdots J-1\} && ,\ \partial\Omega_{\Delta x}=\{x_j:j=0,J\} && ,\ \overline{\Omega}_{\Delta x}=\Omega_{\Delta x}\cup\partial\Omega_{\Delta x}\\
&\Omega_{\Delta t}=\{t_m:m=1\cdots M-1\} && ,\ \partial\Omega_{\Delta t}=\{t_m:m=0,M\} && ,\ \overline{\Omega}_{\Delta t}=\Omega_{\Delta t}\cup\partial\Omega_{\Delta t}
\end{aligned}
$$

为方便起见，再记 $\Omega^+_{\Delta t}=\{t_{m+1}:m=1\cdots M-1\}=\{t_m:m=2\cdots M\}$（实际上表示所有不能被边界与初值条件所确定的节点离散时间坐标）。最后，记数值解 $\tilde{u}^m_j\approx u^m_j=u(x_j,t_m)$，$f^m_j=f(x_j,t_m)$。

Definition 6.2.1. Implicit Scheme

双曲型 IBVP $(*)$ 的 *Implicit Scheme* 为

$$
\begin{aligned}
&D^c_{tt}\tilde{u}^m_j-c^2D^c_{xx}\tilde{u}^{m+1}_j=f^{m+1}_j && \text{for} && (x_j,t_{m+1})\in\Omega_{\Delta x}\times\Omega^+_{\Delta t} && (5)\\
&\tilde{u}^0_j=g_0(x_j) && \text{for} && x_j\in\Omega_{\Delta x} && (6)\\
&\tilde{u}^1_j=\tilde{u}^0_j+\Delta t g_1(x_j) && \text{for} && x_j\in\Omega_{\Delta x} && (7)\\
&\tilde{u}^m_0=\tilde{u}^m_J=0 && \text{for} && t_m\in\overline{\Omega}_{\Delta t} && (8)
\end{aligned}
$$

其中式 (7) 等价于 $D^+_t\tilde{u}^0_j=g_1(x_j)$。在使用 Implicit Scheme 时，求解 $\tilde{u}^{m+1}$ 需要知道前两排 $\tilde{u}^m,\tilde{u}^{m-1}$ 的值。记 u^m_j 为真值，该策略的一致性误差 *consistency error* 定义为

$$
\begin{aligned}
&T^{m+1}_j=D^c_{tt}u^m_j-c^2D^c_{xx}u^{m+1}_j-f^{m+1}_j && \text{for} && (x_j,t_{m+1})\in\Omega_{\Delta x}\times\Omega^+_{\Delta t}\\
&T^1_j=D^+_tu^0_j-g_1(x_j) && \text{for} && x_j\in\Omega_{\Delta x}
\end{aligned}
$$

Definition 6.2.2. Explicit Scheme - 1

双曲型 IBVP $(*)$ 的 *Explicit Scheme-1* 为

$$D_{tt}^c \tilde{u}_j^m - c^2 D_{xx}^c \tilde{u}_j^m = f_j^m \quad \text{for} \quad (x_j, t_m) \in \Omega_{\Delta x \times \Delta t} \tag{9}$$
$$\tilde{u}_j^0 = g_0(x_j) \quad \text{for} \quad x_j \in \Omega_{\Delta x} \tag{10}$$
$$\tilde{u}_j^1 = \tilde{u}_j^0 + \Delta t g_1(x_j) \quad \text{for} \quad x_j \in \Omega_{\Delta x} \tag{11}$$
$$\tilde{u}_0^m = \tilde{u}_J^m = 0 \quad \text{for} \quad t_m \in \overline{\Omega}_{\Delta t} \tag{12}$$

其中式 (11) 等价于 $D_t^+ \tilde{u}_j^0 = g_1(x_j)$。在使用 Explicit Scheme 时，可以递推求解 $\tilde{u}^{m+1}$，仍需知道前两排 $\tilde{u}^m, \tilde{u}^{m-1}$ 的值。记 u_j^m 为真值，该策略的一致性误差 *consistency error* 定义为

$$T_j^m = D_{tt}^c u_j^m - c^2 D_{xx}^c u_j^m - f_j^m \quad \text{for} \quad (x_j, t_m) \in \Omega_{\Delta x \times \Delta t}$$
$$T_j^0 = D_t^+ u_j^0 - g_1(x_j) \quad \text{for} \quad x_j \in \Omega_{\Delta x}$$

Definition 6.2.3. Explicit Scheme - 2

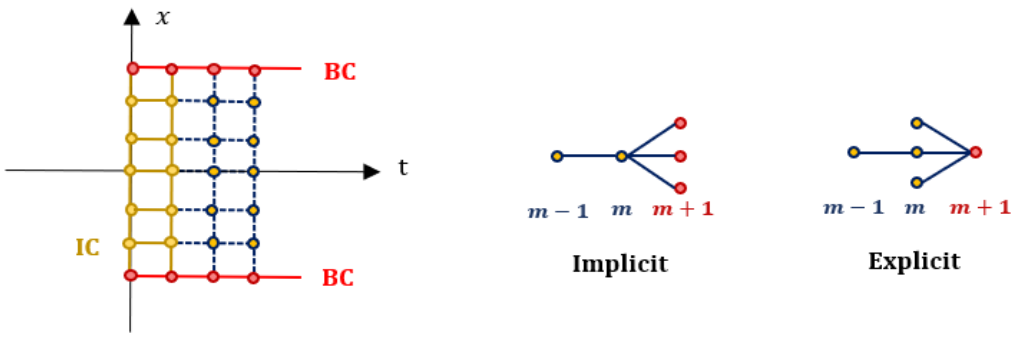
双曲型 IBVP $(*)$ 的另一种更精确的 *Explicit Scheme-2* 将 def 6.2.2 中的 (11) 改为对 $x_j \in \Omega_{\Delta x}$

$$\tilde{u}_j^1 = \tilde{u}_j^0 + \Delta t g_1(x_j) + \frac{1}{2}\Delta t^2 [c^2 D_{xx}^c g_1(x_j) + f_j^0]$$

该策略在边界的一致性误差 *consistency error* 为对 $x_j \in \Omega_{\Delta x}$

$$T_j^0 = D_t^+ u_j^0 - \frac{1}{2}\Delta t^2 [c^2 D_{xx}^c g_1(x_j) + f_j^0] - g_1(x_j) \sim O\left(\Delta t \Delta x^2 + \Delta t^2\right)$$

公式推导见 Lecture Notes p.81。



6.3 Implicit Scheme : Stability, Consistency and Convergence

Definition 6.3.1. 对形如双曲型 IBVP $(*)$（在 t_m 时刻）数值解 $\tilde{u}^m \in \mathbb{R}^{J+1}$（下标从 0 开始）的离散值函数 $\boldsymbol{u}, \boldsymbol{v}$ 定义如下的（半）内积与诱导范数

$$(\boldsymbol{u}, \boldsymbol{v})_{\Delta x} = \sum_{j=1}^{J-1} \Delta x u_j v_j$$

$$(\boldsymbol{u}, \boldsymbol{v}]_{\Delta x} = \sum_{j=1}^{J} \Delta x u_j v_j$$

他们各自的诱导范数定义为 $||\boldsymbol{u}||_{\Delta x} = \sqrt{(\boldsymbol{u}, \boldsymbol{u})_{\Delta x}}$ 和 $||\boldsymbol{u}]|_{\Delta x} = \sqrt{(\boldsymbol{u}, \boldsymbol{u}]_{\Delta x}}$ 我们回忆对于上述内积，有 summation by part，*i.e.*, $\forall \boldsymbol{\phi}$ s.t. $\phi_0 = \phi_J = 0$

$$\left(D_x^+ \boldsymbol{u}, \boldsymbol{\phi}\right)_{\Delta x} = -\left(\boldsymbol{u}, D_x^- \boldsymbol{\phi}\right]_{\Delta x}$$

此外，不难直接验证如下关系 $(**)$ 对上面两个范数也成立

$$(\boldsymbol{u} - \boldsymbol{v}, \boldsymbol{u})_{\Delta x} = \frac{1}{2}\left(||\boldsymbol{u}||_{\Delta x}^2 - ||\boldsymbol{v}||_{\Delta x}^2\right) + \frac{1}{2}||\boldsymbol{u} - \boldsymbol{v}||_{\Delta x}^2$$

$$(\boldsymbol{u} - \boldsymbol{v}, \boldsymbol{u}]_{\Delta x} = \frac{1}{2}\left(||\boldsymbol{u}]|_{\Delta x}^2 - ||\boldsymbol{v}]|_{\Delta x}^2\right) + \frac{1}{2}||\boldsymbol{u} - \boldsymbol{v}]|_{\Delta x}^2$$

Lemma 6.3.1. 若 $\{a_m\}_{m=0}^M$, $\{b_m\}_{m=1}^M$ 均为非负实数列，$M \geq 1$，且 $\exists\, \alpha > 0$ 使得 $\forall\, m \geq 1$: $a_m \leq \alpha a_{m-1} + b_m$，则 $\forall\, m \geq 1$: $a_m \leq \alpha^m a_0 + \sum_{k=1}^m \alpha^{m-k} b_k$

证明由数学归纳法易得。

Theorem 6.3.1. Discrete Energy Inequality for Implicit Scheme (Stability) 考虑双曲型 IBVP $(*)$，若 $\tilde{u}^m$ 表示 t_m 时刻由 Implicit Scheme 得到的数值解，$J, M \geq 2$，则其无条件稳定 *unconditionally stable*，即对 $t_m \in \Omega_{\Delta t}$，都有以下不等式恒成立

$$\mathcal{M}^2[\tilde{u}^m] \leq e^2 \left\{ \mathcal{M}^2[\tilde{u}^0] \;+\; 2T \sum_{k=1}^{m} \Delta t \cdot ||f^{k+1}||_{\Delta x}^2 \right\}$$

$$\text{s.t.} \quad \mathcal{M}^2[\tilde{u}^m] := ||D_t^- \tilde{u}^{m+1}||_{\Delta x}^2 + c^2 ||D_x^- \tilde{u}^{m+1}]|_{\Delta x}^2$$

不难证明，$\mathcal{M}[u] = \sqrt{\mathcal{M}^2[u]}$ 为一个 norm。此外，$\tilde{u} \mapsto \max_{0 \leq m \leq M-1} \mathcal{M}[\tilde{u}^m]$ 也是一个 norm。

Proof theorem 6.3.1 对 $(x_j, t_{m+1}) \in \Omega_{\Delta x} \times \Omega^+_{\Delta t}$ 的差分式两边同时与 $D_t^- \tilde{u}^{m+1}$ 取内积

$$\left(D^c_{tt}\tilde{u}^m, D_t^-\tilde{u}^{m+1}\right)_{\Delta x} - c^2\left(D^c_{xx}\tilde{u}^{m+1}, D_t^-\tilde{u}^{m+1}\right)_{\Delta x} = \left(f^{m+1}, D_t^-\tilde{u}^{m+1}\right)_{\Delta x}$$

显然 $D_t^-\tilde{u}_0^{m+1} = D_t^-\tilde{u}_J^{m+1} \equiv 0$，由 summation by part，上式化为

$$\left(D^c_{tt}\tilde{u}^m, D_t^-\tilde{u}^{m+1}\right)_{\Delta x} + c^2\left(D_x^-\tilde{u}^{m+1}, D_t^-D_x^-\tilde{u}^{m+1}\right]_{\Delta x} = \left(f^{m+1}, D_t^-\tilde{u}^{m+1}\right)_{\Delta x}$$

由 def 6.3.1 $(**)$，有

$$\begin{aligned}
&\left(D^c_{tt}\tilde{u}^m, D_t^-\tilde{u}^{m+1}\right)_{\Delta x} = \left(D_t^+D_t^-\tilde{u}^m, D_t^-\tilde{u}^{m+1}\right)_{\Delta x}\\
=&\frac{1}{\Delta t}\left(D_t^-\tilde{u}^{m+1} - D_t^-\tilde{u}^m, D_t^-\tilde{u}^{m+1}\right)_{\Delta x}\\
=&\frac{1}{2\Delta t}\left(||D_t^-\tilde{u}^{m+1}||^2_{\Delta x} - ||D_t^-\tilde{u}^m||^2_{\Delta x}\right) + \frac{\Delta t}{2}||D_t^-D_t^-\tilde{u}^{m+1}||^2_{\Delta x}
\end{aligned}$$

$$\begin{aligned}
&\left(D_x^-\tilde{u}^{m+1}, D_t^-D_x^-\tilde{u}^{m+1}\right]_{\Delta x} = \left(D_t^-D_x^-\tilde{u}^{m+1}, D_x^-\tilde{u}^{m+1}\right]_{\Delta x}\\
=&\frac{1}{\Delta t}\left(D_x^-\tilde{u}^{m+1} - D_x^-\tilde{u}^m, D_x^-\tilde{u}^{m+1}\right]_{\Delta x}\\
=&\frac{1}{2\Delta t}\left(||D_t^-\tilde{u}^{m+1}]|^2_{\Delta x} - ||D_t^-\tilde{u}^m]|^2_{\Delta x}\right) + \frac{\Delta t}{2}||D_x^-D_t^-\tilde{u}^{m+1}]|^2_{\Delta x}
\end{aligned}$$

带入差分式得到

$$\begin{aligned}
&\frac{1}{2\Delta t}\left(||D_t^-\tilde{u}^{m+1}||^2_{\Delta x} - ||D_t^-\tilde{u}^m||^2_{\Delta x}\right) + \frac{\Delta t}{2}||D_t^-D_t^-\tilde{u}^{m+1}||^2_{\Delta x}\\
+&\frac{c^2}{2\Delta t}\left(||D_x^-\tilde{u}^{m+1}]|^2_{\Delta x} - ||D_x^-\tilde{u}^m]|^2_{\Delta x}\right) + \frac{c^2\Delta t}{2}||D_x^-D_t^-\tilde{u}^{m+1}]|^2_{\Delta x} = \left(f^{m+1}, D_t^-\tilde{u}^{m+1}\right)_{\Delta x}
\end{aligned}$$

整理得

$$\begin{aligned}
\mathcal{M}^2[\tilde{u}^m] - \mathcal{M}^2[\tilde{u}^{m-1}] &= 2\Delta t\left(f^{m+1}, D_t^-\tilde{u}^{m+1}\right)_{\Delta x}\\
&\quad - \Delta t^2\left(||D_t^-D_t^-\tilde{u}^{m+1}||^2_{\Delta x} + c^2||D_x^-D_t^-\tilde{u}^{m+1}]|^2_{\Delta x}\right)\\
&\le 2\Delta t\left(f^{m+1}, D_t^-\tilde{u}^{m+1}\right)_{\Delta x}
\end{aligned}$$

即

$$\mathcal{M}^2[\tilde{u}^m] \le \mathcal{M}^2[\tilde{u}^{m-1}] + 2\Delta t\left(f^{m+1}, D_t^-\tilde{u}^{m+1}\right)_{\Delta x}$$

继续对 $2\Delta t\left(f^{m+1}, D_t^- \tilde{u}^{m+1}\right)_{\Delta x}$ 放缩，结合 $2ab \le a^2 + b^2$，有

$$\begin{aligned}
\mathcal{M}^2[\tilde{u}^m] &\le \mathcal{M}^2[\tilde{u}^{m-1}] + 2\Delta t\left(f^{m+1}, D_t^- \tilde{u}^{m+1}\right)_{\Delta x} \\
&\le \mathcal{M}^2[\tilde{u}^{m-1}] + 2\Delta t||f^{m+1}||_{\Delta x}||D_t^- \tilde{u}^{m+1}||_{\Delta x} \\
&= \mathcal{M}^2[\tilde{u}^{m-1}] + 2\sqrt{T\Delta t}||f^{m+1}||_{\Delta x} \cdot \frac{\sqrt{\Delta t}}{\sqrt{T}}||D_t^- \tilde{u}^{m+1}||_{\Delta x} \\
&\le \mathcal{M}^2[\tilde{u}^{m-1}] + T\Delta t||f^{m+1}||_{\Delta x}^2 + \frac{\Delta t}{T}||D_t^- \tilde{u}^{m+1}||_{\Delta x}^2 \\
&\le \mathcal{M}^2[\tilde{u}^{m-1}] + T\Delta t||f^{m+1}||_{\Delta x}^2 + \frac{\Delta t}{T}\mathcal{M}^2[\tilde{u}^m]
\end{aligned}$$

即 $\left(1 - \frac{\Delta t}{T}\right)\mathcal{M}^2\left[\tilde{u}^m\right] \le \mathcal{M}^2\left[\tilde{u}^{m-1}\right] + T\Delta t||f^{m+1}||_{\Delta x}^2$。由 $\frac{\Delta t}{T} = \frac{1}{M} \in (0, 1/2]$ 与如下不等式

$$\forall\ x \in [\ 0\ ,\ 1/2\ ]\ :\ 1 - x \ge \frac{1}{1+2x}$$

有

$$\begin{aligned}
\mathcal{M}^2[\tilde{u}^m] &\le \left(1 + \tfrac{2\Delta t}{T}\right)\mathcal{M}^2[\tilde{u}^{m-1}] + \left(1 + \tfrac{2\Delta t}{T}\right)T\Delta t||f^{m+1}||_{\Delta x}^2 \\
&\le \left(1 + \tfrac{2\Delta t}{T}\right)\mathcal{M}^2[\tilde{u}^{m-1}] + (1+1)\,T\Delta t||f^{m+1}||_{\Delta x}^2 \\
&= \left(1 + \tfrac{2\Delta t}{T}\right)\mathcal{M}^2[\tilde{u}^{m-1}] + 2T\Delta t||f^{m+1}||_{\Delta x}^2
\end{aligned}$$

由 lemma 6.3.1 可得

$$\begin{aligned}
\mathcal{M}^2[\tilde{u}^m] &\le \left(1 + \tfrac{2\Delta t}{T}\right)^m \mathcal{M}^2[\tilde{u}^0] + 2T\sum_{k=1}^{m}\left(1 + \tfrac{2\Delta t}{T}\right)^{m-k}\Delta t||f^{k+1}||_{\Delta x}^2 \\
&\le e^{\frac{2\Delta t m}{T}}\mathcal{M}^2[\tilde{u}^0] + 2T\sum_{k=1}^{m} e^{\frac{2\Delta t(m-k)}{T}}\Delta t||f^{k+1}||_{\Delta x}^2 \\
&\le e^{\frac{2\Delta t M}{T}}\mathcal{M}^2[\tilde{u}^0] + 2T\sum_{k=1}^{m} e^{\frac{2\Delta t M}{T}}\Delta t||f^{k+1}||_{\Delta x}^2 \\
&= e^2\mathcal{M}^2[\tilde{u}^0] + 2Te^2\sum_{k=1}^{m}\Delta t||f^{k+1}||_{\Delta x}^2
\end{aligned}$$

即 $\mathcal{M}^2[\tilde{u}^m] \le e^2\left\{\ \mathcal{M}^2[\tilde{u}^0]\ +\ 2T\sum_{k=1}^{m}\Delta t \cdot ||f^{k+1}||_{\Delta x}^2\ \right\}$ 得证。

Theorem 6.3.2. Consistency of the Implicit Scheme
考虑双曲型 IBVP $(*)$，若其真解 $u \in C^4(\overline{\Omega})$，则 Implicit Scheme 的一致性误差满足

$$|T_j^1| \le \frac{1}{2}\Delta t M_{2t} \ \sim \ O(\Delta t)$$

$$|T_j^{m+1}| \le \frac{c^2}{12}\Delta x^2 M_{4x} + \frac{5}{3}\Delta t M_{3t} \ \sim \ O(\Delta t + \Delta x^2)$$

其中 $M_{k\cdot} = ||\partial_\cdot^k u||_{C(\overline{\Omega})}$ for $\cdot = t, x$。

Theorem 6.3.3. Convergence of the Implicit Scheme
考虑双曲型 IBVP $(*)$，若其真解 $u \in C^4(\overline{\Omega})$，则 Implicit Scheme 的数值解无条件收敛

$$\max_{0\le m\le M-1} \mathcal{M}[u^m - \tilde{u}^m] = \max_{1\le m\le M-1} \mathcal{M}[u^m - \tilde{u}^m] = O(\Delta x^2 + \Delta t)$$

Proof theorem 6.3.2 先考虑 $m+1 = 2\cdots M$，由泰勒展开（省略），有

$$\begin{aligned} T_j^{m+1} &= D_{tt}^c u_j^m - c^2 D_{xx}^c u_j^{m+1} - f_j^{m+1} \\ &= \left[D_{tt}^c u_j^m - \partial_{tt} u_j^{m+1}\right] - c^2\left[D_{xx}^c u_j^{m+1} - \partial_{xx} u_j^{m+1}\right] \\ &= \frac{1}{3}\Delta t\left\{\partial_t^3 u(x_j, \eta_m) - 4\partial_t^3 u(x_j, \zeta_m)\right\} - \frac{c^2}{12}\Delta x^2 \partial_x^4 u(\xi_j, t_{m+1}) \end{aligned}$$

故有 $|T_j^{m+1}| \le \frac{5}{3}\Delta t M_{3t} + \frac{c^2}{12}\Delta x^2 M_{4x}$ 对 $m+1 = 2\cdots M$ 恒成立。当取 $m+1 = 1$ 时，由泰勒展开的积分余项（见 Appendix thm 7.1.5）

$$T_j^1 = D_t^+ u_j^0 - g_1(x_j) = D_t^+ u_j^0 - \partial_t u_j^0 = \frac{1}{\Delta t}\int_0^{\Delta t} (\Delta t - t)\partial_t^2 u(x_j, t) dt$$

于是 $|T_j^1| \le \frac{M_{2t}}{\Delta t}\int_0^{\Delta t}(\Delta t - t)\, dt = \frac{M_{2t}}{\Delta t}\cdot\frac{1}{2}\Delta t^2 = \frac{1}{2}M_{2t}\Delta t$

Proof theorem 6.3.3 先考虑 $m+1 = 2\cdots M$，记 $\boldsymbol{e}^m = u^m - \tilde{u}^m$，带入 Implicit Scheme 有

$$\begin{aligned} &D_{tt}^c \boldsymbol{e}_j^m - c^2 D_{xx}^c \boldsymbol{e}_j^{m+1} = T_j^{m+1} && \text{for} \quad (x_j, t_{m+1}) \in \Omega_{\Delta x} \times \Omega_{\Delta t}^+ \\ &\boldsymbol{e}_j^0 = 0 && \text{for} \quad x_j \in \Omega_{\Delta x} \\ &D_t^- \boldsymbol{e}^1 = T_j^1 && \text{for} \quad x_j \in \Omega_{\Delta x} \\ &\boldsymbol{e}_0^m = \boldsymbol{e}_J^m = 0 && \text{for} \quad t_m \in \overline{\Omega}_{\Delta t} \end{aligned}$$

由 thm 6.3.1 显然有（注意 $e^2 := exp\,(2)$）

$$\mathcal{M}^2[\boldsymbol{e}^m] \leq e^2 \left\{ \mathcal{M}^2[\boldsymbol{e}^0] \;+\; 2T \sum_{k=1}^{m} \Delta t \cdot ||T^{k+1}||^2_{\Delta x} \right\}$$

显然，由 thm 6.3.2

$$||T^{k+1}||^2_{\Delta x} = \sum_{j=1}^{J-1} \left(T_j^{k+1}\right)^2 \Delta x \leq (b-a) \left[\frac{c^2}{12}\Delta x^2 M_{4x} + \frac{5}{3}\Delta t M_{3t}\right]^2$$

由 $(a+b)^2 = a^2 + 2ab + b^2 \leq a^2 + (a^2+b^2) + b^2 = 2(a^2+b^2) \;\Rightarrow\; O((a+b)^2) = O(a^2+b^2)$ 可得 $||T^{k+1}||^2_{\Delta x} \sim O(\Delta t^2 + \Delta x^4)$。

于是 $2e^2T\sum_{k=1}^{m} \Delta t \cdot ||T^{k+1}||^2_{\Delta x} \sim O(\Delta t^2 \Delta x^2)$。进一步，我们有

$$\mathcal{M}^2[\boldsymbol{e}^0] = ||D_t^- \boldsymbol{e}^1||^2_{\Delta x} + c^2||D_x^- \boldsymbol{e}^1]||^2_{\Delta x} = ||T^1||^2_{\Delta x} + c^2||D_x^- \boldsymbol{e}^1]||^2_{\Delta x}$$

考察 $D_x^- \boldsymbol{e}_j^1$，有

$$\begin{aligned}
D_x^- \boldsymbol{e}_j^1 &= D_x^- \boldsymbol{e}_j^0 + \Delta t D_x^- D_t^- \boldsymbol{e}_j^1 = 0 + \Delta t D_x^- D_t^- T_j^1 \\
&= \Delta t D_x^- \left[\frac{1}{\Delta t}\int_0^{\Delta t} (\Delta t - t)\partial_t^2 u(x_j, t)dt\right] \\
&= \int_0^{\Delta t} (\Delta t - t) D_x^- \partial_t^2 u(x_j, t)dt \\
&= \frac{1}{\Delta x}\int_0^{\Delta t}\int_{x_{j-1}}^{x_j} (\Delta t - t)\partial_x \partial_t^2 u(x,t)dxdt \\
&\leq M_{1x2t} \cdot \frac{1}{\Delta x}\int_0^{\Delta t}\int_{x_{j-1}}^{x_j} (\Delta t - t)dxdt = \frac{1}{2}\Delta t^2 M_{1x2t}
\end{aligned}$$

其中 $M_{1x2t} = ||\partial_x \partial_t^2 u||_{C(\overline{\Omega})}$，于是 $||D_x^- \boldsymbol{e}^1]|^2_{\Delta x} \leq (b-a)\left[\frac{1}{2}\Delta t^2 M_{1x2t}\right]^2 \sim O(\Delta t^4)$；同理，$||T^1||^2_{\Delta x} \sim O(\Delta t^2)$；综上有 $\mathcal{M}^2[\boldsymbol{e}^0] \sim O(\Delta t^2)$。

故

$$\begin{aligned}
\mathcal{M}^2[\boldsymbol{e}^m] = \sqrt{\mathcal{M}^2[\boldsymbol{e}^m]} &\leq \sqrt{C_1\Delta t^2 + C_2\left(\Delta t^2 + \Delta x^4\right)} \\
&\leq \sqrt{C_1}\Delta t + \sqrt{C_2}\left(\Delta t + \Delta x^2\right) \quad \sim \quad O(\Delta x^2 + \Delta t)
\end{aligned}$$

6.4 Explicit Scheme : Stability, Consistency and Convergence

Lemma 6.4.1. 若 $\mathcal{D}$ 是关于对称内积 $(\boldsymbol{u},\boldsymbol{v})=(\boldsymbol{v},\boldsymbol{u})$ 的对称差分算子，即 $(\mathcal{D}\boldsymbol{u},\boldsymbol{v})=(\boldsymbol{v},\mathcal{D}\boldsymbol{u})$，则有

$$\left(\mathcal{D}\left(\boldsymbol{u}-\boldsymbol{v}\right),\boldsymbol{u}+\boldsymbol{v}\right)=\left(\mathcal{D}\left(\boldsymbol{u}+\boldsymbol{v}\right),\boldsymbol{u}-\boldsymbol{v}\right)=(\mathcal{D}\boldsymbol{u},\boldsymbol{u})-(\mathcal{D}\boldsymbol{v},\boldsymbol{v})$$

且两个对称算子的线性组合 $\alpha\mathcal{D}_1+\beta\mathcal{D}_2$ 仍为对称算子。证明较易，省略。

Definition 6.4.1. 定义如下线性空间，其表示所有下标从 0 开始到 J 结束且满足齐次边界条件的实值离散函数集合

$$\mathbb{R}_0^{J+1}:=\{(u_0,\cdots,u_J)\in\mathbb{R}^{J+1}\ :\ u_0=u_J=0\}$$

Lemma 6.4.2. Summation by Parts
若 $\boldsymbol{u},\boldsymbol{v}\in\mathbb{R}_0^{J+1}$，则差分算子 $D_{xx}^c:=D_x^+D_x^-$ 是关于 $\mathbb{R}_0^{J+1}$ 上对称内积 $(\cdot,\cdot)_{\Delta x}$ 的对称算子，即

$$(D_{xx}^c\boldsymbol{u},\boldsymbol{v})_{\Delta x}=(\boldsymbol{u},D_{xx}^c\boldsymbol{v})_{\Delta x}$$

其中 $(\boldsymbol{u},\boldsymbol{v})_{\Delta x}=\sum_{j=1}^{J-1}\Delta x\boldsymbol{u}_j\boldsymbol{v}_j$。证明较易，省略。

Definition 6.4.2. Courant Friedrichs Lewy Condition (CFL)
针对双曲型 IBVP $(*)$，Courant-Friedrichs-Lewy Condition（CFL）指

$$\frac{c\Delta t}{\Delta x}\in(0,1]$$

方便起见，本笔记中 Strict Courant-Friedrichs-Lewy Condition（SCFL）指

$$\exists\ c_0\in(0,1)\ \text{ s.t. }\ 0<\frac{c\Delta t}{\Delta x}\le c_0$$

在 $\mathbb{R}_0^{J+1}$ 上定义运算

$$[\boldsymbol{u},\boldsymbol{v}]=(\mathcal{D}_*\boldsymbol{u},\boldsymbol{v})_{\Delta x}\ \text{ s.t. }\ \mathcal{D}_*=I+\frac{1}{4}c^2\Delta t^2D_{xx}^c$$

总有

$$[\boldsymbol{u},\boldsymbol{u}]\ge\left[1-\left(\frac{c\Delta t}{\Delta x}\right)^2\right]\cdot||\boldsymbol{u}||_{\Delta x}^2$$

当满足 CFL 时，$[\cdot,\cdot]$ 为 $\mathbb{R}_0^{J+1}$ 上的一个 positive semi-definite Hermitian form（即不满足 $[\boldsymbol{u},\boldsymbol{u}]=0\iff\boldsymbol{u}=0$ 的“内积”）。

当满足 SCFL 时，$\mathcal{D}_*$ 可逆，$[\cdot,\cdot]$ 为 $\mathbb{R}_0^{J+1}$ 上的一个良定义内积；记相应的诱导范数为 $|[\boldsymbol{u}]|=[\boldsymbol{u},\boldsymbol{u}]^{\frac{1}{2}}$。

Theorem 6.4.1. Discrete Energy Inequality for Explicit Scheme (Stability) 考虑双曲型 IBVP $(*)$，若 $\tilde{u}^m$ 表示 t_m 时刻由 Explicit Scheme 得到的数值解，$J, M \geq 2$，则其在满足 SCFL 时条件稳定 *conditionally stable*，即对 $t_m \in \Omega_{\Delta t}$，都有以下不等式恒成立

$$\mathcal{N}^2[\tilde{u}^m] \leq e^4 \left\{ \mathcal{N}^2[\tilde{u}^0] \ + \ T \sum_{k=1}^{m} \Delta t \cdot |[\mathcal{D}_*^{-1} f^k]|^2 \right\}$$

$$\text{s.t.} \quad \mathcal{N}^2[\tilde{u}^m] := |[D_t^- \tilde{u}^{m+1}]|^2 + \frac{c^2}{4} ||D_x^-(\tilde{u}^{m+1} + \tilde{u}^m)]|_{\Delta x}^2$$

此时（CFL 满足），$\tilde{u} \mapsto \max_{0 \leq m \leq M-1} \mathcal{N}[\tilde{u}^m]$ 也是一个 norm（Lecture Notes p.84 remark，不要求掌握）。当只满足 CFL 时，若 $f \equiv 0$，则数值解也条件稳定。

7 Appendix

7.1 Definition and Theorem

Theorem 7.1.1. 多元函数的散度定理（divergence theorem）
记 Ω 为 $\mathbb{R}^n$ 中的由 $\partial\Omega$ 围成的有界连通区域，$\boldsymbol{n}$ 为曲面 $\partial\Omega$ 上的单位法向量（诱导定向）。若矢量函数 $\boldsymbol{v}$ 在闭区域 $\bar{\Omega}=\Omega\cup\partial\Omega$ 上连续，在 Ω 内有一阶连续偏导数，则：

$$\int_\Omega \nabla\cdot\boldsymbol{v}d\boldsymbol{x}=\oint_{\partial\Omega}\boldsymbol{v}\cdot\boldsymbol{n}d\boldsymbol{S}$$

Theorem 7.1.2. 涉及散度的高维分部积分（Integration by parts）
记 $\boldsymbol{v}$ 为一矢性函数，u 为一标量函数，$\Omega\subseteq\mathbb{R}^n$ 为一有界闭区域。回顾 product rule for divergence，有

$$\nabla\cdot(u\boldsymbol{v})=\nabla u\cdot\boldsymbol{v}+u\nabla\cdot\boldsymbol{v}$$

对两边在 Ω 上积分并利用 *5.2.1* 散度定理，有

$$\oint_{\partial\Omega}(u\boldsymbol{v})\cdot\boldsymbol{n}d\boldsymbol{S}=\int_\Omega\nabla u\cdot\boldsymbol{v}\ d\boldsymbol{x}+\int_\Omega u\nabla\cdot\boldsymbol{v}d\boldsymbol{x}$$

特别地，$\triangle v=\nabla\cdot(\nabla v)$，于是 $\boldsymbol{v}=\nabla v$ 带入上式，有

$$\int_\Omega u\triangle\boldsymbol{v}d\boldsymbol{x}=-\int_\Omega\nabla u\cdot\nabla v\ d\boldsymbol{x}+\oint_{\partial\Omega}(u\nabla v)\cdot\boldsymbol{n}d\boldsymbol{S}$$

Theorem 7.1.3. 傅里叶变换
记函数 $f(x)$ 和 $\hat{f}(\xi)$ 的傅里叶（逆）变换分别为

$$\hat{f}(\xi)=F[f](\xi)=\int_\infty^\infty f(x)\,e^{-i\xi x}dx\quad,\quad f(x)\sim F^{-1}[\hat{f}](x)=\frac{1}{2\pi}\int_\infty^\infty\hat{f}(\xi)\,e^{i\xi x}d\xi$$

当 f 在 $\mathbb{R}$ 上绝对可积且在任何闭区间上分段可微时

$$F^{-1}[\hat{f}](x)=\frac{1}{2}\left[f(x^-)+f(x^+)\right]$$

在连续点处 $(F^{-1}\circ F)[f]:=F^{-1}[F[f]]=f(x)$，未经特殊说明的情况下我们假设该条件一直满足。

Theorem 7.1.4. 傅里叶变换的常用结论
详细讨论见 MATH 323 笔记第五章，此处只罗列常用的结论。

1) **线性** $F[af+bg]=a\hat{f}+b\hat{g}$; $F^{-1}[af+bg]=aF^{-1}[f]+bF^{-1}[g]$
2) **位移** $F[f(x+a)]=e^{i\xi a}\hat{f}(\xi)$; $F^{-1}[f(\xi+a)]=e^{-ixa}F^{-1}[f](x)$
3) **缩放** $F[f(ax)]=\frac{1}{|a|}\hat{f}(\xi/a)$ for $a\neq 0$
4) **换元** $F[e^{iax}f(x)]=\hat{f}(\xi-a)$; $F^{-1}[e^{ia\xi}f(\xi)]=F^{-1}[f](x+a)$
5) **共轭** $\overline{F[f]}(\xi)=2\pi F^{-1}[\ \bar{f}\ ](\xi)$; $\overline{F^{-1}[f]}(\xi)=\frac{1}{2\pi}F[\ \bar{f}\ ](x)$

6) **微分的变换** $F[f^{(n)}]=(i\xi)^n\hat{f}$; $F^{-1}[f^{(n)}]=(-ix)^nF^{-1}[f]$
7) **偏导数变换** $F_x[\partial_x^n u(x,y)]=(i\xi)^n\hat{u}(\xi,y)$; $F_x[\partial_y^n u(x,y)]=\partial_y^n\hat{u}(\xi,y)$
8) **变换的微分** $\frac{\mathrm{d}}{\mathrm{d}\xi}\hat{f}=-iF[xf]$; $\frac{\mathrm{d}}{\mathrm{d}x}F^{-1}[f]=iF^{-1}[\xi f]$
9) **卷积的变换** $F[f*g]=\hat{f}\hat{g}$; $F^{-1}[f*g]=2\pi F^{-1}[f]F^{-1}[g]$
10) **乘法逆变换** $F^{-1}[\hat{f}\cdot\hat{g}]=F^{-1}[\hat{f}]*F^{-1}[\hat{g}]=f*g$，该公式要求连续函数 $f,g,f*g$ 在 $\mathbb{R}$ 上绝对可积且在任何闭区间上分段可微

11) **指数函数的变换** $F\left[e^{-|a|x}\right]=\frac{2a}{a^2+\xi^2}$
12) **高斯分布的变换** $F[e^{-a^2x^2}]=\frac{\sqrt{\pi}}{|a|}exp(-\frac{\xi^2}{4a^2})$
13) **矩形函数的变换** $F[1_{(-a,a)}]=\frac{2sin(a\xi)}{\xi}$
14) **FT 与逆变换的转换** $\hat{f}(\xi)=2\pi F^{-1}[f](-\xi)$; $F^{-1}[f](x)=\frac{1}{2\pi}\hat{f}(-x)$

利用傅里叶变换求 PDE，先对某个变量对方程与边界条件做变换，化为 ODE，解之后做逆变换。

Theorem 7.1.5. 泰勒展开的积分余项
记函数 $f:\mathbb{R}\mapsto\mathbb{R}$ 满足 $f\in C^{k+1}(U_a)$，其中 U_a 为 $x=a$ 的某个开邻域。则 $\forall x\in U_a$

$$f(x)=f(a)+f'(a)(x-a)+\cdots\frac{f^{(k)}(a)}{k!}(x-a)^k+\int_a^x f^{(k+1)}(t)\frac{(x-t)^k}{k!}dt$$

证明见补充材料 Taylor-integral.pdf。