

# Finite Volume Method in 2D

Anh Ha Le, Huong Lan Tran

University of Sciences

April 8, 2014

## Introduction

### Rectangular mesh

Rectangular mesh

Scheme

Existence and uniqueness of discrete solution

Error estimation

### Admissible Mesh

Introduction

Admissible Mesh

Scheme

Properties of Scheme

Definition of differential operators and its product scalars

Discrete variational formulation

Existence and uniqueness of discrete solution

Discrete maximum principle

Error Estimation

Energy norm

$L^2$  norm

Let  $\Omega \subset \mathbb{R}^2$  and  $f \in L^2(\Omega)$ . We will use the finite volume method to discretize the following Poisson equation

$$-\Delta u = f(x, y) \quad \text{in } \Omega \quad (2.1)$$

subject to a Dirichlet boundary condition:

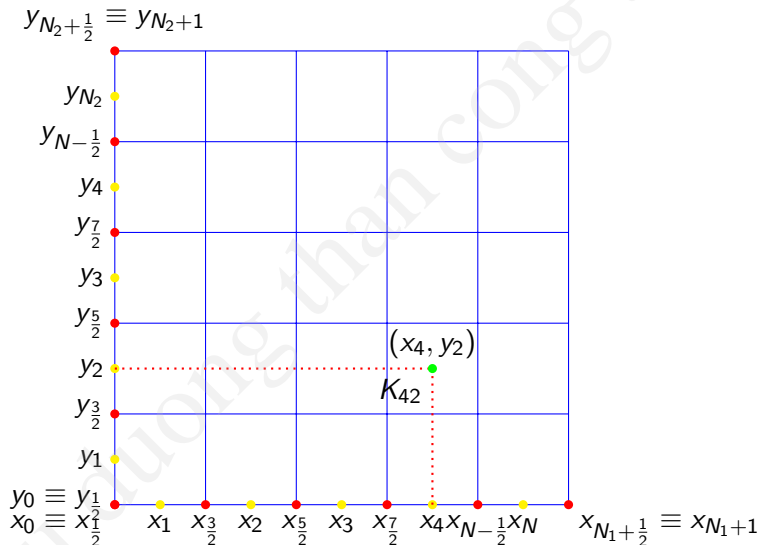
$$u(x, y) = 0 \quad \text{on } \partial\Omega. \quad (2.2)$$

The Laplacian  $\Delta$  is defined in Cartesian coordinates by

$$\Delta = \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

When  $f(x, y) = 0$ , (2.1) is called Laplace equation.

The existence and uniqueness of the solution to equation (2.1) is proved. Our purpose is to find the discrete solution.



Consider  $\Omega = (0, 1) \times (0, 1)$ . On interval  $[0, 1]$ , we make two partition  $(x_{i+\frac{1}{2}})_{i \in \overline{0, N_1}}$ ,  $(y_{j+\frac{1}{2}})_{j \in \overline{0, N_2}}$  such that

$$0 = x_{\frac{1}{2}} < x_{\frac{3}{2}} < \cdots < x_{N_1-\frac{1}{2}} < x_{N_1+\frac{1}{2}} = 1,$$

$$0 = y_{\frac{1}{2}} < y_{\frac{3}{2}} < \cdots < y_{N_2-\frac{1}{2}} < y_{N_2+\frac{1}{2}} = 1.$$

Let  $\mathcal{T} = (T_{ij})_{i \in \overline{1, N_1}, j \in \overline{1, N_2}}$  be an admissible mesh of  $(0, 1) \times (0, 1)$  such that

$$T_{ij} = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}]$$

$T_{ij}$  is called a **control volume** of  $\mathcal{T}$ . The points  $x_{i+\frac{1}{2}}$ ,  $y_{j+\frac{1}{2}}$  are called **mesh points**.

Choosing the sequences  $(x_i)_{i \in \overline{0, N_1+1}}$  and  $(y_j)_{j \in \overline{0, N_2+1}}$  such that

$$x_0 \equiv x_{\frac{1}{2}}, \quad x_i = \frac{1}{2} \left( x_{i-\frac{1}{2}} + x_{i+\frac{1}{2}} \right), \quad x_{N_1+1} \equiv x_{N_1+\frac{1}{2}},$$

$$y_0 \equiv y_{\frac{1}{2}}, \quad y_j = \frac{1}{2} \left( y_{j-\frac{1}{2}} + y_{j+\frac{1}{2}} \right), \quad x_{N_2+1} \equiv y_{N_2+\frac{1}{2}}.$$

The point  $(x_i, y_j)$  is the **control point** of control volume  $T_{ij}$ .

Let

$$h_i = |x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}|, \quad k_j = |y_{j+\frac{1}{2}} - y_{j-\frac{1}{2}}| \quad \text{for all } i \in \overline{1, N_1} \text{ and } j \in \overline{1, N_2}$$

and

$$h_{i+\frac{1}{2}} = |x_{i+1} - x_i|, \quad k_{j+\frac{1}{2}} = |y_{j+1} - y_j| \quad \text{for all } i \in \overline{0, N_1} \text{ and } j \in \overline{0, N_2}.$$

Then, the **area** of control volume  $|T_{ij}| = h_i k_j$ .

$h = \max\{h_i, k_j\}$  is the **mesh size**.

The finite volume scheme is found by integrating the first equation of (2.1) over each control volume  $T_{ij}$ , which gives

$$\frac{1}{|T_{ij}|} \int_{T_{ij}} -\Delta u(x, y) dx dy = \frac{1}{|T_{ij}|} \int_{T_{ij}} f(x, y) dx dy \quad (3.1)$$

Following the definition of  $-\Delta$  operator, we have

$$-\frac{1}{|T_{ij}|} \int_{T_{ij}} u_{xx}(x, y) dx dy - \frac{1}{|T_{ij}|} \int_{T_{ij}} u_{yy}(x, y) dx dy = \frac{1}{|T_{ij}|} \int_{T_{ij}} f(x, y) dx dy \quad (3.2)$$

We can rewrite clearer that

$$\begin{aligned} \frac{-1}{|T_{ij}|} \int_{y_{j-1/2}}^{y_{j+1/2}} \int_{x_{i-1/2}}^{x_{i+1/2}} u_{xx}(x, y) dx dy - \frac{1}{|T_{ij}|} \int_{x_{i-1/2}}^{x_{i+1/2}} \int_{y_{j-1/2}}^{y_{j+1/2}} u_{yy}(x, y) dy dx \\ \frac{1}{|T_{ij}|} \int_{T_{ij}} f(x, y) dx dy \end{aligned} \quad (3.3)$$

Applying the integral formulation, we obtain:

$$\int_{y_{j-1/2}}^{y_{j+1/2}} \int_{x_{i-1/2}}^{x_{i+1/2}} u_{xx}(x, y) dx dy = \int_{y_{j-1/2}}^{y_{j+1/2}} u_x(x_{i+\frac{1}{2}}, y) dy - \int_{y_{j-1/2}}^{y_{j+1/2}} u_x(x_{i-\frac{1}{2}}, y) dy \quad (3.4)$$

and

$$\int_{x_{i-1/2}}^{x_{i+1/2}} \int_{y_{j-1/2}}^{y_{j+1/2}} u_{yy}(x, y) dy dx = \int_{x_{i-1/2}}^{x_{i+1/2}} u_y(x, y_{j+\frac{1}{2}}) dx - \int_{x_{i-1/2}}^{x_{i+1/2}} u_y(x, y_{j-\frac{1}{2}}) dx \quad (3.5)$$



Then, we get

$$\begin{aligned}
 & -\frac{1}{|T_{ij}|} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} u_x(x_{i+\frac{1}{2}}, y) dy + \frac{1}{|T_{ij}|} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} u_x(x_{i-\frac{1}{2}}, y) dy \\
 & - \frac{1}{|T_{ij}|} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} u_y(x, y_{j+\frac{1}{2}}) dx + \frac{1}{|T_{ij}|} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} u_y(x, y_{j-\frac{1}{2}}) dy \\
 & = \frac{1}{|T_{ij}|} \int_{T_{ij}} f(x, y) dx dy
 \end{aligned} \tag{3.6}$$

The following approximations are made:

$$\frac{1}{|T_{ij}|} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} u_x(x_{i+\frac{1}{2}}, y) dy = \frac{k_j u_x(x_{i+\frac{1}{2}}, y_j)}{|T_{ij}|} = \frac{u_{i+1,j} - u_{i,j}}{h_i h_{i+\frac{1}{2}}},$$

$$\frac{1}{|T_{ij}|} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} u_x(x_{i-\frac{1}{2}}, y) dy = \frac{k_j u_x(x_{i-\frac{1}{2}}, y_j)}{|T_{ij}|} = \frac{u_{i,j} - u_{i-1,j}}{h_i h_{i-\frac{1}{2}}},$$

$$\frac{1}{|T_{ij}|} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} u_y(x, y_{j+\frac{1}{2}}) dy = \frac{h_i u_y(x_i, y_{j+\frac{1}{2}})}{|T_{ij}|} = \frac{u_{i,j+1} - u_{i,j}}{k_j k_{j+\frac{1}{2}}},$$

$$\frac{1}{|T_{ij}|} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} u_y(x, y_{j-\frac{1}{2}}) dy = \frac{h_i u_y(x_i, y_{j-\frac{1}{2}})}{|T_{ij}|} = \frac{u_{i,j} - u_{i,j-1}}{k_j k_{j-\frac{1}{2}}}.$$

The approximate equation (3.6) becomes

$$-\frac{u_{i+1,j} - u_{i,j}}{h_i h_{i+\frac{1}{2}}} + \frac{u_{i,j} - u_{i-1,j}}{h_i h_{i-\frac{1}{2}}} - \frac{u_{i,j+1} - u_{i,j}}{k_j k_{j+\frac{1}{2}}} + \frac{u_{i,j} - u_{i,j-1}}{k_j k_{j-\frac{1}{2}}} = f_{ij} \quad (3.7)$$

where  $f_{ij} = \frac{1}{|T_{ij}|} \int_{T_{ij}} f(x, y) dx$  is mean-value of  $f$  on  $T_{ij}$ .

Rearranging equation (3.7) gives

$$\begin{aligned} & -\frac{1}{h_i h_{i-\frac{1}{2}}} u_{i-1,j} - \frac{1}{h_i h_{i+\frac{1}{2}}} u_{i+1,j} - \frac{1}{k_j k_{j-\frac{1}{2}}} u_{i,j-1} - \frac{1}{k_j k_{j+\frac{1}{2}}} u_{i,j+1} \\ & + \left( \frac{1}{h_i h_{i-\frac{1}{2}}} + \frac{1}{h_i h_{i+\frac{1}{2}}} + \frac{1}{k_j k_{j-\frac{1}{2}}} + \frac{1}{k_j k_{j+\frac{1}{2}}} \right) u_{i,j} = f_{ij} \end{aligned} \quad (3.8)$$

By setting

$$a_i = -\frac{1}{h_i h_{i-\frac{1}{2}}}, \quad b_i = -\frac{1}{h_i h_{i+\frac{1}{2}}}$$

$$c_j = -\frac{1}{k_j k_{j-\frac{1}{2}}}, \quad d_j = -\frac{1}{k_j k_{j+\frac{1}{2}}}$$

$$s_{i,j} = a_i + b_i + c_j + d_j$$

At cell  $(i, j)$  for  $i \in [1, N_1]$  and  $j \in [1, N_2]$  the discrete equation is written as

$$-a_i u_{i-1,j} - b_i u_{i+1,j} - c_j u_{i,j-1} - d_j u_{i,j+1} + s_{i,j} u_{i,j} = f_{ij} \quad (3.9)$$

The system is closed with boundary conditions

$$u_{0,j} = u_{N_1+1,j} = 0, \quad j \in \overline{1, N_2}$$

and

$$u_{i,0} = u_{i,N_2+1} = 0, \quad i \in \overline{1, N_1} \quad (3.10)$$

## Matrix form of the discrete equation

We arrange the discrete unknowns

$(u_{i,j}), i = 1, \dots, N_1, j = 1, \dots, N_2$  in the following form

$$u = (u_{1,1}, u_{1,2}, \dots, u_{1,N_2}; u_{2,1}, u_{2,2}, \dots, u_{2,N_2}; \dots; u_{N_1,1}, u_{N_1,2}, \dots, u_{N_1,N_2})^T$$

and

$$f = (f_{1,1}, f_{1,2}, \dots, f_{1,N_2}; f_{2,1}, f_{2,2}, \dots, f_{2,N_2}; \dots; f_{N_1,1}, f_{N_1,2}, \dots, f_{N_1,N_2})^T$$

Then the discrete equation is written in the matrix form  $Au = f$  where

$$A = \begin{pmatrix} A_1 & D_2 & 0 & \dots & 0 & 0 \\ C_1 & A_2 & D & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & & \vdots & \vdots \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & A_{N_1-1} & D_{N_1-1} \\ 0 & 0 & 0 & \dots & C_{N_1} & A_{N_1} \end{pmatrix}$$

where

$$A_i = \begin{pmatrix} s_{i,1} & -b_1 & 0 & \dots & 0 & 0 \\ -a_2 & s_{i,2} & -b_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & & \vdots & \vdots \\ \vdots & \vdots & & \ddots & \vdots & \\ 0 & 0 & 0 & \dots & s_{i,N_2-1} & -b_{N_2-1} \\ 0 & 0 & 0 & \dots & -a_{N_2} & s_{i,N_2} \end{pmatrix}$$

$$C_i = \begin{pmatrix} -c_i & 0 & \dots & 0 \\ 0 & -c_i & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -c_i \end{pmatrix}, \quad D_i = \begin{pmatrix} -d_i & 0 & \dots & 0 \\ 0 & -d_i & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -d_i \end{pmatrix}$$

The scheme is set as a system of  $N_1 \times N_2$  equations with  $N_1 \times N_2$  unknowns  $(u_{i,j})_{i \in [1, N_1], j \in [1, N_2]}$ .

In  $\mathbb{R}^{N_1 \times N_2}$ , existence and uniqueness are equivalent for square systems.

**Uniqueness of the solution:** We prove that if  $f_{ij} = 0$  for  $i \in [1, N_1]$  and  $j \in [1, N_2]$  then  $u_{i,j} = 0$  for  $i \in [1, N_1]$  and  $j \in [1, N_2]$ .

Multiplying two sides of (3.7) by  $u_{i,j}$  and taking sum over  $i \in [1, N_1]$  and  $j \in [1, N_2]$

$$\begin{aligned} & \sum_{i,j=1}^{N_1, N_2} \left[ -\frac{(u_{i+1,j} - u_{i,j})k_j}{h_{i+\frac{1}{2}}} + \frac{(u_{i,j} - u_{i-1,j})k_j}{h_{i-\frac{1}{2}}} \right] \\ & + \sum_{i,j=1}^{N_1, N_2} \left[ -\frac{(u_{i,j+1} - u_{i,j})h_i}{k_{j+\frac{1}{2}}} + \frac{(u_{i,j} - u_{i,j-1})h_i}{k_{j-\frac{1}{2}}} \right] u_{i,j} = 0 \end{aligned}$$

Or

$$\sum_{i,j=1}^{N_1,N_2} \left[ -\frac{(u_{i+1,j} - u_{i,j})u_{i,j}k_j}{h_{i+\frac{1}{2}}} + \frac{(u_{i,j} - u_{i-1,j})u_{i,j}k_j}{h_{i-\frac{1}{2}}} \right] \\ \sum_{i,j=1}^{N_1,N_2} \left[ -\frac{(u_{i,j+1} - u_{i,j})u_{i,j}h_i}{k_{j+\frac{1}{2}}} + \frac{(u_{i,j} - u_{i,j-1})u_{i,j}h_i}{k_{j-\frac{1}{2}}} \right] = 0$$

Changing index and using boundary condition, we get

$$\sum_{i=0,j=1}^{N_1,N_2} \frac{(u_{i+1,j} - u_{i,j})^2 k_j}{h_{i+\frac{1}{2}}} + \sum_{i=1,j=0}^{N_1,N_2} \frac{(u_{i,j+1} - u_{i,j})^2 h_i}{k_{j+\frac{1}{2}}} = 0$$

From this equality and combining with boundary condition, we get  $u_{i,j} = 0$  for all  $i \in [1, N_1]$  and  $j \in [1, N_2]$



## Error estimate

Let  $\Omega = (0, 1) \times (0, 1)$  and  $f \in L^2(\Omega)$ . Let  $u$  be the unique solution of (2.1). We assume that there exist  $\zeta > 0$  such that  $h_i \geq \zeta h$  for  $i \in [1, N_1]$  and  $k_j \geq \zeta k$  for  $j \in [1, N_2]$ . Let  $(u_{i,j})_{i \in [1, N_1], j \in [1, N_2]}$  be the unique discrete solution of (3.7). There exists  $C > 0$  only depending on  $u$ ,  $\Omega$  and  $\zeta$  such that

$$\sum_{i=0, j=1}^{N_1, N_2} \frac{(e_{i+1,j} - e_{i,j})^2}{h_{i+1/2}} k_j + \sum_{i=1, j=0}^{N_1, N_2} \frac{(e_{i,j+1} - e_{i,j})^2}{k_{j+1/2}} h_i \leq Ch^2, \quad (3.11)$$

and

$$\sum_{i,j=1}^{N_1, N_2} (e_{i,j})^2 h_i k_j \leq Ch^2, \quad (3.12)$$

where  $e_{i,j} = u(x_i, y_j) - u_{i,j}$  for all  $i \in [0, N_1]$  and  $j \in [0, N_2]$

Relation (3.11) can be seen as an estimate of a discrete  $H_0^1$  norm of the error, while relation (3.12) gives an estimate of the  $L^2$  norm of the error.

Consider first the case  $u \in \mathbb{C}^2(\overline{\Omega})$ . We can prove as in 1D case.

### Consistency

If  $u \in \mathbb{C}^2([0, 1] \times [0, 1], \mathbb{R})$ , there exists  $C \in \mathbb{R}_+$  only depending on  $u$  such that

$$|R_{i+1/2,j}| = \left| \frac{1}{k_j} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} u_x(x_{i+1/2}, y) dy - \frac{u(x_{i+1}, y_j) - u(x_i, y_j)}{h_{i+\frac{1}{2}}} \right| \leq Ch$$

## Proof

Using Taylor series expansion, there exist  $\eta_{i+1/2} \in (x_{i+1/2}, x_{i+1})$  and  $\theta_{i+1/2} \in (x_i, x_{i+1/2})$  such that

$$\begin{aligned} & \frac{u(x_{i+1}, y_j) - u(x_{i+1/2}, y_j)}{h_{i+\frac{1}{2}}} - \frac{x_{i+1} - x_{i+1/2}}{h_{i+\frac{1}{2}}} u_x(x_{i+1/2}, y_j) \\ &= \frac{1}{2} \frac{(x_{i+1} - x_{i+1/2})^2}{h_{i+\frac{1}{2}}} u_{xx}(\eta_{i+1/2}, y_j) \end{aligned}$$

$$\begin{aligned} & \frac{u(x_{i+1/2}, y_j) - u(x_i, y_j)}{h_{i+\frac{1}{2}}} - \frac{x_{i+1/2} - x_i}{h_{i+\frac{1}{2}}} u_x(x_{i+1/2}, y_j) \\ &= -\frac{1}{2} \frac{(x_{i+1/2} - x_i)^2}{h_{i+\frac{1}{2}}} u_{xx}(\theta_{i+1/2}, y_j) \end{aligned}$$

Taking sum of the two expressions gives

$$\begin{aligned} R_{i+\frac{1}{2},j}^* &= \frac{u(x_{i+1}, y_j) - u(x_i, y_j)}{h_{i+\frac{1}{2}}} - u_x(x_{i+\frac{1}{2}}, y_j) \\ &= -\frac{(x_{i+1} - x_{i+\frac{1}{2}})^2}{2(h_{i+\frac{1}{2}})} u_{xx}(\eta_{i+\frac{1}{2}}, y_j) + \frac{(x_{i+\frac{1}{2}} - x_i)^2}{2(h_{i+\frac{1}{2}})} u_{xx}(\theta_{i+\frac{1}{2}}, y_j) \end{aligned}$$

The following inequality holds

$$\begin{aligned} |R_{i+\frac{1}{2},j}^*| &\leq \frac{(x_{i+1} - x_{i+\frac{1}{2}})^2}{2(h_{i+\frac{1}{2}})} |u_{xx}(\eta_{i+\frac{1}{2}}, y_j)| + \frac{(x_{i+\frac{1}{2}} - x_i)^2}{2(h_{i+\frac{1}{2}})} |u_{xx}(\theta_{i+\frac{1}{2}}, y_j)| \\ &\leq C \left( \frac{(x_{i+1} - x_{i+\frac{1}{2}})^2}{h_{i+\frac{1}{2}}} + \frac{(x_{i+\frac{1}{2}} - x_i)^2}{h_{i+\frac{1}{2}}} \right) \\ &\leq C \frac{(h_{i+\frac{1}{2}})^2}{h_{i+\frac{1}{2}}} \end{aligned}$$

According to mean-value theorem, the following equality holds

$$u_x(x_{i+\frac{1}{2}}, y) - u_x(x_{i+\frac{1}{2}}, y_j) = (y - y_j)u_{xy}(x_{i+\frac{1}{2}}, \zeta), \quad \zeta \in (y, y_j)$$

Then

$$\begin{aligned} |u_x(x_{i+\frac{1}{2}}, y) - u_x(x_{i+\frac{1}{2}}, y_j)| &= |y - y_j| |u_{xy}(x_{i+\frac{1}{2}}, \zeta)| \\ &\leq k_j C \end{aligned}$$

Incorporating with the equality on  $|R_{i+\frac{1}{2}j}^*|$  it implies that

$$\begin{aligned} |R_{i+\frac{1}{2}j}| &\leq \left| \frac{1}{k_j} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} u_x(x_{i+\frac{1}{2}}, y) dy - u_x(x_{i+\frac{1}{2}}, y_j) \right| + |R_{i+\frac{1}{2}j}^*| \\ &\leq \frac{1}{k_j} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} |u_x(x_{i+\frac{1}{2}}, y) - u_x(x_{i+\frac{1}{2}}, y_j)| dy + |R_{i+\frac{1}{2}j}^*| \\ &\leq k_j C + Ch \\ &\leq 2Ch \end{aligned}$$

**Prove the error estimate**

1.

$$\sum_{i=0, j=1}^{N_1, N_2} \frac{(e_{i+1,j} - e_{i,j})^2 k_j}{h_{i+1/2}} + \sum_{i=1, j=0}^{N_1, N_2} \frac{(e_{i,j+1} - e_{i,j})^2 h_i}{k_{j+1/2}} \leq Ch^2$$

Integrating equation  $-\Delta u = f$  over  $K_{i,j}$  yields

$$\begin{aligned} & -\frac{1}{|K_{i,j}|} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} u_x(x_{i+\frac{1}{2}}, y) dy + \frac{1}{|K_{i,j}|} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} u_x(x_{i-\frac{1}{2}}, y) dy \\ & -\frac{1}{|K_{i,j}|} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} u_y(x, y_{j+\frac{1}{2}}) dx + \frac{1}{|K_{i,j}|} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} u_y(x, y_{j-\frac{1}{2}}) dy \\ & = f_{i,j} \end{aligned} \quad (3.13)$$

The approximate solution  $U$  satisfies

$$-\frac{u_{i+1,j} - u_{i,j}}{h_i h_{i+\frac{1}{2}}} + \frac{u_{i,j} - u_{i-1,j}}{h_i h_{i-\frac{1}{2}}} - \frac{u_{i,j+1} - u_{i,j}}{k_j k_{j+\frac{1}{2}}} + \frac{u_{i,j} - u_{i,j-1}}{k_j k_{j-\frac{1}{2}}} = f_{ij}$$

Therefore,

$$\begin{aligned}
 & -\frac{1}{|K_{i,j}|} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} u_x(x_{i+\frac{1}{2}}, y) dy + \frac{u_{i+1,j} - u_{i,j}}{h_i h_{i+\frac{1}{2}}} \\
 & + \frac{1}{|K_{i,j}|} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} u_x(x_{i-\frac{1}{2}}, y) dy - \frac{u_{i,j} - u_{i-1,j}}{h_i h_{i-\frac{1}{2}}} \\
 & - \frac{1}{|K_{i,j}|} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} u_y(x, y_{j+\frac{1}{2}}) dx + \frac{u_{i,j+1} - u_{i,j}}{k_j k_{j+\frac{1}{2}}} \\
 & + \frac{1}{|K_{i,j}|} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} u_y(x, y_{j-\frac{1}{2}}) dx - \frac{u_{i,j} - u_{i,j-1}}{k_j k_{j-\frac{1}{2}}} = 0
 \end{aligned}$$

Setting

$$-\int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} u_x(x_{i+\frac{1}{2}}, y) dy + \frac{(u_{i+1,j} - u_{i,j})k_j}{h_{i+\frac{1}{2}}} = -k_j R_{i+1/2,j} - \frac{(e_{i+1,j} - e_{i,j})k_j}{h_{i+\frac{1}{2}}}$$

$$\int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} u_x(x_{i-\frac{1}{2}}, y) dy - \frac{(u_{i,j} - u_{i-1,j})k_j}{h_{i-\frac{1}{2}}} = k_j R_{i-1/2,j} + \frac{(e_{i,j} - e_{i-1,j})k_j}{h_{i-\frac{1}{2}}}$$

$$-\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} u_y(x_i, y_{j+\frac{1}{2}}) dy + \frac{(u_{i,j+1} - u_{i,j})h_i}{k_{j+\frac{1}{2}}} = -h_i R_{i,j+1/2} - \frac{(e_{i,j+1} - e_{i,j})h_i}{k_{j+\frac{1}{2}}}$$

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} u_y(x_i, y_{j-\frac{1}{2}}) dy - \frac{(u_{i,j} - u_{i,j-1})h_i}{k_{j-\frac{1}{2}}} = h_i R_{i,j-1/2} + \frac{(e_{i,j} - e_{i,j-1})h_i}{k_{j-\frac{1}{2}}}$$

Then

$$\begin{aligned} & -\frac{(e_{i+1,j} - e_{i,j})k_j}{h_{i+\frac{1}{2}}} + \frac{(e_{i,j} - e_{i-1,j})k_j}{h_{i-\frac{1}{2}}} - \frac{(e_{i,j+1} - e_{i,j})h_i}{k_{j+\frac{1}{2}}} + \frac{(e_{i,j} - e_{i,j-1})h_i}{k_{j-\frac{1}{2}}} \\ & = k_j R_{i+1/2,j} - k_j R_{i-1/2,j} + h_i R_{i,j+1/2} - h_i R_{i,j-1/2} \end{aligned}$$



Multiplying by  $e_{i,j}$  and summing over  $i = 1, \dots, N_1$ ;  $j = 1, \dots, N_2$  gives

$$\begin{aligned}
 & - \sum_{i=1, j=1}^{N_1, N_2} \frac{(e_{i+1,j} - e_{i,j})e_{i,j}k_j}{h_{i+\frac{1}{2}}} + \sum_{i=1, j=1}^{N_1, N_2} \frac{(e_{i,j} - e_{i-1,j})e_{i,j}k_j}{h_{i-\frac{1}{2}}} \\
 & - \sum_{i=1, j=1}^{N_1, N_2} \frac{(e_{i,j+1} - e_{i,j})e_{i,j}h_i}{k_{j+\frac{1}{2}}} + \sum_{i=1, j=1}^{N_1, N_2} \frac{(e_{i,j} - e_{i,j-1})e_{i,j}h_i}{k_{j-\frac{1}{2}}} \\
 & = \sum_{i=1, j=1}^{N_1, N_2} R_{i+1/2,j} e_{i,j} k_j - \sum_{i=1, j=1}^{N_1, N_2} R_{i-1/2,j} e_{i,j} k_j + \sum_{i=1, j=1}^{N_1, N_2} R_{i,j+1/2} e_{i,j} h_i \\
 & - \sum_{i=1, j=1}^{N_1, N_2} R_{i,j-1/2} e_{i,j} h_i
 \end{aligned}$$

Changing the index gives

$$\begin{aligned}
 & - \sum_{i=1, j=1}^{N_1, N_2} \frac{(e_{i+1, j} - e_{i, j}) e_{i, j} k_j}{h_{i+\frac{1}{2}}} + \sum_{i=0, j=1}^{N_1-1, N_2} \frac{(e_{i+1, j} - e_{i, j}) e_{i+1, j} k_j}{h_{i+\frac{1}{2}}} \\
 & - \sum_{i=1, j=1}^{N_1, N_2} \frac{(e_{i, j+1} - e_{i, j}) e_{i, j} h_i}{k_{j+\frac{1}{2}}} + \sum_{i=1, j=0}^{N_1, N_2-1} \frac{(e_{i, j+1} - e_{i, j}) e_{i, j+1} h_i}{k_{j+\frac{1}{2}}} \\
 & = \sum_{i=1, j=1}^{N_1, N_2} R_{i+1/2, j} e_{i, j} k_j - \sum_{i=0, j=1}^{N_1-1, N_2} R_{i+1/2, j} e_{i+1, j} k_j + \sum_{i=1, j=1}^{N_1, N_2} R_{i, j+1/2} e_{i, j} h_i \\
 & - \sum_{i=1, j=0}^{N_1, N_2-1} R_{i, j+1/2} e_{i, j+1} h_i
 \end{aligned}$$

Reordering and using  $e_{0,j} = e_{N_1,j} = e_{i,0} = e_{i,N_2} = 0$  yields

$$\begin{aligned} & \sum_{i=0,j=1}^{N_1,N_2} \frac{(e_{i+1,j} - e_{i,j})^2 k_j}{h_{i+\frac{1}{2}}} + \sum_{i=1,j=0}^{N_1,N_2} \frac{(e_{i,j+1} - e_{i,j})^2 h_i}{k_{j+\frac{1}{2}}} \\ &= \sum_{i=0,j=1}^{N_1,N_2} R_{i+1/2,j} (e_{i,j} - e_{i+1,j}) k_j + \sum_{i=1,j=0}^{N_1,N_2} R_{i,j+1/2} (e_{i,j} - e_{i,j+1}) h_i \end{aligned}$$

Using the consistency property, it implies

$$\begin{aligned} & \sum_{i=0,j=1}^{N_1,N_2} \frac{(e_{i+1,j} - e_{i,j})^2 k_j}{h_{i+\frac{1}{2}}} + \sum_{i=1,j=0}^{N_1,N_2} \frac{(e_{i,j+1} - e_{i,j})^2 h_i}{k_{j+\frac{1}{2}}} \\ & \leq Ch \sum_{i=0,j=1}^{N_1,N_2} |(e_{i,j} - e_{i+1,j}) k_j| + Ch \sum_{i=1,j=0}^{N_1,N_2} |(e_{i,j} - e_{i,j+1}) h_i| \end{aligned} \tag{3.15}$$

## Applying Cauchy-Schwarz inequality

$$\sum_{j=1}^{N_2} k_j \sum_{i=0}^{N_1} |(e_{i,j} - e_{i+1,j})| \leq \sum_{j=1}^{N_2} k_j \left( \sum_{i=0}^{N_1} \frac{(e_{i+1,j} - e_{i,j})^2}{h_{i+\frac{1}{2}}} \right)^{1/2} \left( \sum_{i=0}^{N_1} h_{i+\frac{1}{2}} \right)^{1/2}$$

e.g.

$$\sum_{i=0, j=1}^{N_1, N_2} |(e_{i,j} - e_{i+1,j}) k_j| \leq \sum_{j=1}^{N_2} k_j \left( \sum_{i=0}^{N_1} \frac{(e_{i+1,j} - e_{i,j})^2}{h_{i+\frac{1}{2}}} \right)^{1/2}$$

$$\sum_{i=1}^{N_1} h_i \sum_{j=0}^{N_2} |(e_{i,j} - e_{i,j+1})| \leq \sum_{i=1}^{N_1} h_i \left( \sum_{j=0}^{N_2} \frac{(e_{i,j+1} - e_{i,j})^2}{k_{j+\frac{1}{2}}} \right)^{1/2} \left( \sum_{j=0}^{N_2} k_{j+\frac{1}{2}} \right)^{1/2}$$

e.g.

$$\sum_{i=1, j=0}^{N_1, N_2} |(e_{i,j} - e_{i,j+1}) h_i| \leq \sum_{i=1}^{N_1} h_i \left( \sum_{j=0}^{N_2} \frac{(e_{i,j+1} - e_{i,j})^2}{k_{j+\frac{1}{2}}} \right)^{1/2}$$

By setting

$$P_j = \sum_{i=0}^{N_1} \frac{(e_{i+1,j} - e_{i,j})^2}{h_{i+\frac{1}{2}}}, \quad Q_i = \sum_{j=0}^{N_2} \frac{(e_{i,j+1} - e_{i,j})^2}{k_{j+\frac{1}{2}}}$$

we have

$$\sum_{j=1}^{N_2} k_j P_j + \sum_{i=1}^{N_1} h_i Q_i \leq Ch \left( \sum_{j=1}^{N_2} k_j \sqrt{P_j} + \sum_{i=1}^{N_1} h_i \sqrt{Q_i} \right) \quad (3.16)$$

Applying Cauchy-Schwarz inequality for the RHS of (3.16) it yields

$$\sum_{j=1}^{N_2} k_j P_j + \sum_{i=1}^{N_1} h_i Q_i \leq Ch \left( \sum_{j=1}^{N_2} k_j + \sum_{i=1}^{N_1} h_i \right)^{1/2} \left( \sum_{j=1}^{N_2} k_j P_j + \sum_{i=1}^{N_1} h_i Q_i \right)^{1/2}$$

It implies that

$$\sum_{j=1}^{N_2} k_j P_j + \sum_{i=1}^{N_1} h_i Q_i \leq 2C^2 h^2$$

$$2. \sum_{i,j=1}^{N_1, N_2} (e_{i,j})^2 h_i k_j \leq Ch^2$$

We have

$$e_{i,j} = \sum_{i_1=0}^{i-1,j} (e_{i_1+1,j} - e_{i_1-1,j})$$

Then

$$|e_{i,j}| \leq \sum_{i_1=0}^{i-1} |e_{i_1+1,j} - e_{i_1-1,j}|$$

Applying Cauchy-Schwarz inequality, we have

$$\begin{aligned} |e_{i,j}| &\leq \left( \sum_{i_1=0}^{i-1} \frac{|e_{i_1+1,j} - e_{i_1,j}|^2}{h_{i_1+1/2}} \right)^{1/2} \left( \sum_{i_1=0}^{i-1} h_{i_1+1/2} \right)^{1/2} \\ &\leq \left( \sum_{i_1=0}^{N_1} \frac{|e_{i_1+1,j} - e_{i_1,j}|^2}{h_{i_1+1/2}} \right)^{1/2} \end{aligned}$$

It is similar, we have

$$|e_{i,j}| \leq \left( \sum_{j_1=0}^{N_2} \frac{|e_{i,j_1+1} - e_{i,j_1}|^2}{k_{j_1+1/2}} h_i \right)^{1/2}$$

Multiplying two previous inequality, there holds

$$|e_{i,j}|^2 \leq \left( \sum_{i_1=0}^{N_1} \frac{|e_{i_1+1,j} - e_{i_1,j}|^2}{h_{i_1+1/2}} \right)^{1/2} \left( \sum_{j_1=0}^{N_2} \frac{|e_{i,j_1+1} - e_{i,j_1}|^2}{h_{j_1+1/2}} \right)^{1/2}$$

Applying Cauchy-Schwarz inequality, we have

$$|e_{i,j}|^2 \leq \frac{1}{2} \sum_{i_1=0}^{N_1} \frac{|e_{i_1+1,j} - e_{i_1,j}|^2}{h_{i_1+1/2}} + \frac{1}{2} \sum_{j_1=0}^{N_2} \frac{|e_{i,j_1+1} - e_{i,j_1}|^2}{h_{j_1+1/2}}$$

Or

$$|e_{i,j}|^2 h_i k_j \leq \frac{h_i}{2} \sum_{i_1=0}^{N_1} \frac{|e_{i_1+1,j} - e_{i_1,j}|^2}{h_{i_1+1/2}} k_j + \frac{k_j}{2} \sum_{j_1=0}^{N_2} \frac{|e_{i,j_1+1} - e_{i,j_1}|^2}{h_{j_1+1/2}} h_i$$

Summing over  $i$  and  $j$ , we have

$$\begin{aligned} \sum_{i,j=1}^{N_1,N_2} |e_{i,j}|^2 h_i k_j &\leq \frac{1}{2} \sum_{i=1}^{N_1} h_i \sum_{j=1}^{N_2} \sum_{i_1=0}^{N_1} \frac{|e_{i_1+1,j} - e_{i_1,j}|^2}{h_{i_1+1/2}} k_j \\ &\quad + \frac{1}{2} \sum_{j=1}^{N_2} k_j \sum_{i=1}^{N_1} \sum_{j_1=0}^{N_2} \frac{|e_{i,j_1+1} - e_{i,j_1}|^2}{h_{j_1+1/2}} h_i \end{aligned}$$

Putting

$$P = \sum_{j=1}^{N_2} \sum_{i_1=0}^{N_1} \frac{|e_{i_1+1,j} - e_{i_1,j}|^2}{h_{i_1+1/2}} k_j \quad Q = \sum_{i=1}^{N_1} \sum_{j_1=0}^{N_2} \frac{|e_{i,j_1+1} - e_{i,j_1}|^2}{h_{j_1+1/2}} h_i$$



Then

$$\sum_{i,j=1}^{N_1, N_2} |e_{i,j}|^2 h_i k_j \leq \frac{1}{2} \sum_{i=1}^{N_1} h_i P + \frac{1}{2} \sum_{j=1}^{N_2} k_j Q$$

Since

$$\sum_{i=1}^{N_1} h_i = 1 \qquad \sum_{j=1}^{N_2} k_j = 1$$

Thus

$$\sum_{i,j=1}^{N_1, N_2} |e_{i,j}|^2 h_i k_j \leq \frac{P+Q}{2} \leq C^2 h^2$$

Let  $\Omega \subset \mathbb{R}^2$  and  $f \in L^2(\Omega)$ . We will use the finite volume method to discretize the following Poisson equation

$$-\Delta u = f(x, y) \quad \text{in } \Omega \quad (4.1)$$

subject to a Dirichlet boundary condition:

$$u(x, y) = u_d(x, y) \quad \text{on } \Gamma = \partial\Omega. \quad (4.2)$$

CuuDuongThanCong.com

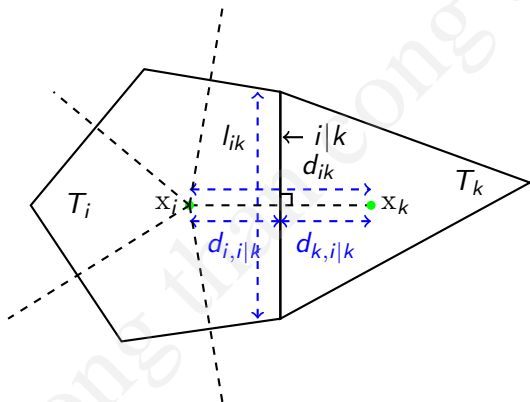
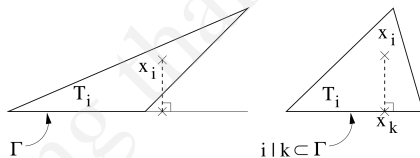


Figure: Two neighboring cells of admissible mesh.

With each  $T_i$ , we associate a point  $x_i \in \overset{\circ}{T}_i$ . We will denote by  $i|k$  the common edge of  $T_i$  and  $T_k$  when these two elements are

The mesh is said to be admissible mesh if  $[x_i, x_k]$  is orthogonal to  $i|k$  for any couple  $(T_i, T_k)$  of neighboring elements, and if, for any element  $T_i$  which has an edge on  $\Gamma$ , the orthogonal projection of the associated point  $x_i$  on the straight line going over the considered edge belongs to this edge.



**Figure:** Non admissibility (left) and admissibility (right) at the boundary

In that case, the orthogonal projections on the edge is still denote by  $x_k$  with  $k \in [I + 1, I + I^\Gamma]$ , where  $I^\Gamma$  denotes the number of boundary edges of the mesh, and edge is still denoted by  $i|k$ .

For  $i \in [1, I]$ , we denote by  $V(i) \subset [1, I + I^\Gamma]$  the set of the neighboring indexes of the element  $T_i$ , by  $|T_i|$  the area of  $T_i$ . Let  $h_i$  denote the diameter of  $T_i$  and  $\rho_i$  denote the diameter of the largest ball inscribed in  $T_i$ .

We make the following shape regularity assumption on the mesh.

**Assumption** (shape-regularity of the meshes). There exist the a constant positive  $\theta$  independent of mesh such that  $\min_{i \in [1, I]} h_i / \rho_i \leq \theta$ .

We next denote by  $E$  the set of all edges in the mesh, by  $E^{int}$  the set of interior, by  $E^{ext}$  the set of exterior. Let  $l_{ik}$  be the length of  $i|k$  (note that  $l_{ik} = l_{ki}$ ), and let  $n_{ik}$  be the unit vector orthogonal to  $i|k$  pointing from  $T_i$  to  $T_k$  (note that  $n_{ik} = -n_{ki}$ ). We will denote by  $d_{i,i|k} = d(x_i, i|k)$  (resp  $d_{k,i|k}$ ) is the distance between point  $x_i$  (resp  $x_k$ ) and edge  $i|k$ , and  $d_{ik} = d_{ki} = \|\overrightarrow{x_i x_k}\|$ .

We associate with any finite volume  $T_i$  of the mesh an unknown denoted by  $u_i$ , which will approximate the value  $u(x_i)$ , and we integrate the first equation in (4.1) over  $T_i$ , there holds

$$\frac{-1}{|T_i|} \int_{T_i} \Delta u(x) dx = \frac{1}{|T_i|} \int_{T_i} f(x) dx. \quad (4.3)$$

The left-hand side of (4.3) may be evaluated thanks to Green formula

$$\frac{-1}{|T_i|} \int_{T_i} \Delta u(x) dx = \frac{-1}{|T_i|} \sum_{k \in V(i)} \int_{i|k} \nabla u(\sigma) \cdot n_{ik} d\sigma \quad (4.4)$$

When the mesh is admissible, a reasonable approximation of  $\nabla u \cdot n_{ik}$  on the edge  $i|k$  is given by:

$$\nabla u \cdot n_{ik} \approx \frac{u(x_k) - u(x_i)}{d_{ik}}. \quad (4.5)$$

Thus as soon as  $k \in [1, I]$  (i.e if  $i|k$  is interior edge), we can approach

$$\int_{i|k} \nabla u(\sigma) \cdot n_{ik} d\sigma = \frac{l_{ik}}{d_{ik}} (u_k - u_i) \quad \forall k \in [1, I]. \quad (4.6)$$

If  $k \in [I + 1, I + I^\Gamma]$  (i.e  $i|k$  is exterior edge), the formula (4.5) is still a good approximation of the gradient in the normal direction. However, the value  $u(\mathbf{x}_k)$  is known since  $k \in \Gamma$ . We may approach

$$\int_{i|k} \nabla u(\sigma) \cdot n_{ik} d\sigma = \frac{l_{ik}}{d_{ik}} (u_d(\mathbf{x}_k) - u_i) \quad \forall k \in \Gamma. \quad (4.7)$$



Finally, the  $i$  equation of the scheme thus writes

$$\frac{-1}{|T_i|} \sum_{i|k \notin \Gamma} \frac{l_{ik}}{d_{ik}} (u_k - u_i) - \frac{1}{|T_i|} \sum_{i|k \in \Gamma} \frac{l_{ik}}{d_{ik}} (u_d(x_k) - u_i) = f_i, \quad (4.8)$$

where  $f_i$  is the mean-value of the function  $f$  over  $T_i$  it mean that

$$f_i = \frac{1}{|T_i|} \int_{T_i} f(x) dx.$$

## Divergence and Gradient operators

Let  $E$  be the all edges in the mesh and  $N_E$  be the number in the set  $E$  ( $N_b$  be number of boundary edges). We recall that  $I$  is the number of elements in the mesh. We define the following discrete divergence operator

$$\begin{aligned} d : \mathbb{R}^{N_E} &\rightarrow \mathbb{R}^I \\ (v_{ik})_{ik \in E} &\mapsto (dv)_i := \frac{1}{|T_i|} \sum_{k \in V(i)} l_{ik} v_{ik} \end{aligned} \quad (4.9)$$

We also introduce the discrete gradient operator

$$\begin{aligned} g : \mathbb{R}^{I+I^\Gamma} &\rightarrow \mathbb{R}^{N_E} \\ (u_i)_{i \in I} &\mapsto (gu)_{ik} := \frac{u_k - u_i}{d_{ik}} \end{aligned} \quad (4.10)$$

From two operator and our scheme (4.8), we get

$$-d(gu)_i = f_i \quad \forall i \in [1, I] \text{ and } u_k = u_d(x_k) \quad \forall x_k \in \Gamma \quad (4.11)$$

## Product scalars

We also define the discrete product scalar  $(\cdot, \cdot)_T$  by

$$(u_i)_{i \in [1, I]}, (w_i)_{i \in [1, I]} \mapsto (u, w)_T = \sum_{i=1}^I T_i u_i w_i \quad (4.12)$$

and  $\|u\|_{0,T}^2 = (u, u)_T$

With each edge, we define discrete product scalar  $(\cdot, \cdot)_D$

$$(a_{ik})_{ik \in E}, (b_{ik})_{ik \in E} \mapsto (a, b)_D = \sum_{ik \in E} \frac{l_{ik} d_{ik}}{2} a_{ik} b_{ik} \quad (4.13)$$

and  $\|v\|_{0,D}^2 = (v, v)_D$ ,  $\|u\|_{1,D}^2 = \|g(u)\|_{0,D}^2 = (g(u), g(u))$

Finally, we define boundary discrete product scalar on  $\Gamma$

$$(a_{ik})_{ik \in \Gamma}, (b_{ik})_{ik \in \Gamma} \mapsto (a, b)_\Gamma = \sum_{ik \in \Gamma} l_{ik} a_{ik} b_{ik} \quad (4.14)$$

## Proposition

Let  $(u_i)_{i \in [I+I^\Gamma]}$  and  $(v_{ik})_{ik \in E}$  be given. There holds

$$(dv, u)_T = -2(v, gu)_D + (v, \gamma u)_\Gamma \quad (4.15)$$

where the discrete trace operator  $\gamma$  is defined following way:

$$\begin{aligned} \gamma : \mathbb{R}^{I+I^\Gamma} &\rightarrow \mathbb{R}^{I^\Gamma} \\ (u_i)_{i \in I+I^\Gamma} &\mapsto (\gamma u)_{ik} = u_k \text{ when } x_k \in \Gamma \end{aligned} \quad (4.16)$$

The discrete Green formula is discrete equivalent of

$$(\nabla \cdot v, u)_{L^2(\Omega)} = -(v, \nabla u)_{L^2(\Omega)} + (v \cdot n, u)_{L^2(\Gamma)}$$

## Discrete variational formulation

Consider any  $(w_i)_{i \in [1, I+I^\Gamma]}$  with  $w_k = 0$  for all  $k \in [I+1, I+I^\Gamma]$ .  
There holds

$$2(g(u), g(w))_D = (f, w)_T \quad (4.17)$$

**Proof:** Let us start from (4.11), multiple by  $|T_i|w_i$  and sum over  $i \in [1, I]$ , we obtain

$$-(dgu, w)_T = (f, w)_T \quad (4.18)$$

Combining this with the discrete Green formula (4.15), we have

$$2(gu, gw)_D - (gu, \gamma w)_\Gamma = (f, w)_T \quad (4.19)$$

Now, on  $\Gamma$ ,  $\gamma w$  is vanish. Therefore (4.19) implies (4.17).

## Existence and uniqueness of discrete solution

The finite volume scheme may be written as system of  $I + I^\Gamma$  and  $I + I^\Gamma$  unknowns given by (4.11)/ Therefore, the existence and uniqueness are equivalent. If  $f_i = 0$  for all  $i \in [1, I]$  and  $u_k = 0$  for all  $k \in [I + 1, I + I^\Gamma]$ , then we can choose  $w = u$ , there holds, thanks to (4.17),

$$(g(u), g(u))_D = 0 = \sum_{ik \in E} \frac{l_{ik} d_{ik}}{2} (gu)_{ik}^2$$

This implies  $(gu)_{ik} = 0$  for all  $ik \in E$ . By the definition of  $(gu)_{ik}$ ,  $u_i = u_k$  for all  $i|k$  (including the boundary edges). Combining with boundary condition,  $u_i = 0$  for all  $i \in [1, I + I^\Gamma]$

## Discrete maximum principle

### Proposition

*We suppose that  $f$  is positive on  $\Omega$  and that  $u_k = 0$  for all  $k \in [I + 1, I + I^\Gamma]$ . There holds*

$$u_i \geq 0 \quad \forall i \in [1, I] \quad (4.20)$$

We shall give error estimate for both in discrete  $H_0^1(\Omega)$  norm and discrete  $L^2(\Omega)$  norm. We shall define mesh size  $h = \max_{i \in [1, I]} \text{diam}(T_i)$ . We suppose that  $u$ , exact solution of (4.1) with  $u_d = 0$ , belong to  $C^2(\bar{\Omega})$ . We condition the pointwise projection of  $u$  as follows

$$(\Pi u)_k = u(x_k) \quad \text{for all } k \in [1, I + I^\Gamma]$$

and this implies that

$$(\Pi u)_k = 0 \quad \text{for all } k \in [I + 1, I + I^\Gamma] \quad (4.21)$$

We shall consider the difference between the projection  $((\Pi u)_i)_{i \in [1, I + I^\Gamma]}$  the value  $(u_i)_{i \in [1, I + I^\Gamma]}$  obtained by from the finite volume scheme in (4.11) with boundary condition

$$u_k = 0 \quad \text{for all } k \in [I + 1, I + I^\Gamma] \quad (4.22)$$



## Energy norm

We would like to estimate the norm of the discrete gradient of error

$$|u - \Pi u|_{1,D}^2 = (g(u) - g(\Pi u), g(u) - g(\Pi u))_D \quad (4.23)$$

For first, we will define averaged normal gradient on each edge  $i|k$

$$(\delta u)_{ik} = \frac{1}{l_{ik}} \int_{i|k} \nabla u \cdot n_{ik}(\sigma) d\sigma \quad (4.24)$$

Note that this implies that  $(\delta u)_{ik} = -(\delta u)_{ki}$

## Lemma

For any  $(w_i)_{i \in [1, l+l^\Gamma]}$  with  $w_k = 0$  for all  $k \in [l+1, l+l^\Gamma]$ , there holds

$$(f, w)_T = 2(\delta u, g(w))_D \quad (4.25)$$

Thanks to (4.17) and (4.25), we have

$$(\delta u, g(w))_D = (g(u), g(w))_D \quad (4.26)$$

for all vanishing on the boundary. Setting  $w = u - \Pi u$ . Thanks to (4.26), we have

$$\begin{aligned} |w|_{1,D}^2 &= |u - \Pi u|_{1,D}^2 = (g(u - \Pi u), g(w))_D \\ &= (g(u), g(w))_D - (g(\Pi u), g(w))_D \\ &= (\delta u, g(w))_D - (g(\Pi u), g(w))_D \\ &= (\delta u - g(\Pi u), g(w))_D \leq \|\delta u - g(\Pi u)\|_{0,D} |w|_{1,D} \end{aligned}$$

Thus, we have

$$|u - \Pi u|_{1,D} \leq |\delta u - g(\Pi u)|_{0,D} \quad (4.27)$$

By definition of  $|\cdot|_{0,D}$ , we have

$$\|\delta u - g(\Pi u)\|_{0,D}^2 = \sum_{ik \in E} \frac{l_{ik} d_{ik}}{2} ((\delta u)_{ik} - (g(\Pi u))_{ik})^2$$

Setting

$$\begin{aligned} e_{ik} &= (\delta u)_{ik} - (g(\Pi u))_{ik} = \frac{1}{l_{ik}} \int_{l_{ik}} \nabla u \cdot n_{ik}(\sigma) d\sigma - \frac{u_k - u_i}{d_{ik}} \\ &= \frac{1}{l_{ik}} \int_{l_{ik}} \nabla u \cdot n_{ik}(\sigma) d\sigma - \frac{1}{d_{ik}} \int_{[x_k, x_i]} \nabla u \cdot n_{ik}(\sigma) d\sigma \end{aligned}$$

Let  $x_{ik}$  be intersection between  $l_{ik}$  and  $[x_k, x_i]$ , then

$$e_{ik} = \frac{1}{l_{ik}} \int_{l_{ik}} (\nabla u(\sigma) - \nabla u(x_{ik})) \cdot n_{ik}(\sigma) d\sigma \\ - \frac{1}{d_{ik}} \int_{[x_k, x_i]} (\nabla u(\sigma) - \nabla u(x_{ik})) \cdot n_{ik}(\sigma) d\sigma$$

There exists constant  $C > 0$  (depending on  $u$ ) such that

$$\left| \frac{1}{l_{ik}} \int_{l_{ik}} (\nabla u(\sigma) - \nabla u(x_{ik})) \cdot n_{ik}(\sigma) d\sigma \right| \leq Cl_{ik} \leq Ch \\ \left| \frac{1}{d_{ik}} \int_{[x_k, x_i]} (\nabla u(\sigma) - \nabla u(x_{ik})) \cdot n_{ik}(\sigma) d\sigma \right| \leq Cd_{ik} \leq 2Ch$$

Thus, this implies that

$$|e_{ik}| \leq 3Ch \quad (4.28)$$

Using the estimate of  $e_{ik}$ , there holds

$$\|\delta u - g(\Pi u)\|_{0,D}^2 = \sum_{ik \in E} \frac{l_{ik} d_{ik}}{2} (3Ch)^2 \quad (4.29)$$

$$\leq (3Ch)^2 \sum_{ik \in E} \frac{l_{ik} d_{ik}}{2} = (3Ch)^2 |\Omega| \quad (4.30)$$

Then,

$$\|\delta u - g(\Pi u)\|_{0,D} \leq 3C |\Omega|^{1/2} h \quad (4.31)$$

This implies

$$|u - \Pi u|_{1,D} \leq 3C |\Omega|^{1/2} h$$

## $L^2$ norm

We will prove the discrete Poincare inequality: Let  $w_{i \in [1, I+1]^{\Gamma}}$  such that  $w_k = 0$  for all  $k \in \Gamma$ . Then there exists  $C = |\Omega|^{1/2}$  such that

$$\|w\|_{0,T} \leq C|w|_{1,D} \quad (4.32)$$

Let us define the following function

$$\begin{aligned} \omega : \Omega &\rightarrow \mathbb{R} \\ x &\mapsto \omega(x) = w_i \text{ if } x \in T_i \end{aligned}$$

Let  $x \in \Omega$  be given, we define by  $D_x^1$  and  $D_x^2$  the two straight line going through with direction  $(1, 0)$  and  $(0, 1)$ , respectively.

## $L^2$ norm

For a given  $i|k$  in the set of edges  $E$  and a given  $x \in \Omega$ , we also define the function  $\chi_{ik}^j$  with  $j = 1$  and  $j = 2$  by

$$\chi_{ik}^j : \Omega \rightarrow \mathbb{R}$$

$$x \mapsto \chi_{ik}^j(x) = \begin{cases} 1 & \text{if } i|k \cap D_x^j \neq \emptyset \\ 0 & \text{if } i|k \cap D_x^j = \emptyset \end{cases}$$

Then, for a given  $T_i$  and for all  $x \in T_i$ , there holds

$$\omega(x) = w_i = (w_i - w_{k_1}) + (w_{k_1} - w_{k_2}) + \cdots + (w_{k_{q-1}} - w_{k_q}) + (w_{k_q} - w_k)$$

where the index  $k$  is such that  $x_k \in \Gamma$  and belongs to an edge of the mesh which intersects  $D_x^1$ , so that  $w_k = 0$ . Since

$$w_{k_l} - w_{k_{l+1}} = d_{k_l k_{l+1}}(gw)_{k_l k_{l+1}}$$

## $L^2$ norm

there holds

$$\omega(x) \leq \sum_{k_1 | k_2 \in E} d_{k_1 k_2}(gw)_{k_1 k_2} \chi_{k_1 k_2}^1(x)$$

Performing the same calculation with  $j = 2$  instead of  $j = 1$ , multiplying the two inequalities, there holds

$$\omega^2(x) = \left( \sum_{k_1 | k_2 \in E} d_{k_1 k_2}(gw)_{k_1 k_2} \chi_{k_1 k_2}^1(x) \right) \left( \sum_{k_1 | k_2 \in E} d_{k_1 k_2}(gw)_{k_1 k_2} \chi_{k_1 k_2}^2(x) \right)$$



## $L^2$ norm

Integrating over  $\Omega$  and taking into account that  $\omega$  is a constant over each  $T_i$ , there holds

$$\sum_{i \in [1, I]} |T_i| w_i^2 \leq \int_{\Omega} \left( \sum_{k_1 | k_2 \in E} d_{k_1 k_2}(g w)_{k_1 k_2} \chi_{k_1 k_2}^1(x) \right) \left( \sum_{k_1 | k_2 \in E} d_{k_1 k_2}(g w)_{k_1 k_2} \chi_{k_1 k_2}^2(x) \right) dx_1 dx_2$$

Now, it is easily seen that  $\chi_{k_1 k_2}^1(x)$  only depends on  $x_2$  and  $\chi_{k_1 k_2}^2(x)$  only depends on  $x_1$ , so that, setting

$$\begin{aligned} a &:= \min\{x_1 \text{ s.t. } (x_1, x_2) \in \Omega\}, & b &:= \max\{x_1 \text{ s.t. } (x_1, x_2) \in \Omega\} \\ \alpha &:= \min\{x_2 \text{ s.t. } (x_1, x_2) \in \Omega\}, & \beta &:= \max\{x_2 \text{ s.t. } (x_1, x_2) \in \Omega\} \end{aligned}$$

## $L^2$ norm

we get

$$\sum_{i \in [1, I]} |T_i| w_i^2 \leq \int_{\alpha}^{\beta} \left( \sum_{k_1 | k_2 \in E} d_{k_1 k_2} (gw)_{k_1 k_2} \chi_{k_1 k_2}^1(x_2) \right) dx_2$$

$$\int_a^b \left( \sum_{k_1 | k_2 \in E} d_{k_1 k_2} (gw)_{k_1 k_2} \chi_{k_1 k_2}^2(x_1) \right) dx_1$$

It is easily seen that

$$\int_{\alpha}^{\beta} \chi_{k_1 k_2}^1(x_2) dx_2 \leq l_{k_1 k_2} \text{ and } \int_a^b \chi_{k_1 k_2}^2(x_1) dx_1 \leq l_{k_1 k_2}$$

## $L^2$ norm

So that we finally get

$$\sum_{i \in [1, I]} |T_i| w_i^2 \leq \left( \sum_{k_1 k_2 \in E} d_{k_1 k_2} l_{k_1 k_2} (gw)_{k_1 k_2} \right)^2$$

Applying the discrete Cauchy-Schwarz inequality, there holds

$$\sum_{i \in [1, I]} |T_i| w_i^2 \leq 4 \left( \sum_{k_1 k_2 \in E} \frac{d_{k_1 k_2} l_{k_1 k_2}}{2} \right) \left( \sum_{k_1 k_2 \in E} \frac{d_{k_1 k_2} l_{k_1 k_2}}{2} (gw)_{k_1 k_2}^2 \right)$$

which is exact the discrete Poincare inequality (4.32).