The Poisson equation

May 18, 2020

1 The discrete Poisson problem

We consider the 2D elliptic Boundary Value Problem (BVP) given by *the Poisson equation with Dirichlet boundary condition*

$$\begin{cases} \Delta u\left(x,y\right) = f\left(x,y\right), \ (x,y) \in \Omega, \\ u\left(x,y\right) = g\left(x,y\right), \ (x,y) \in \Gamma, \end{cases} \tag{1}$$

where $\Omega = \left(0,1\right)^2$ is the open unit square in $\mathbb{R}^2$ and Γ is the boundary of Ω (see Figure 1).

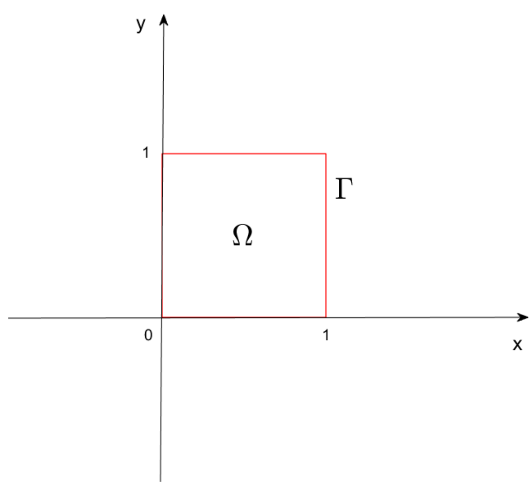


Figure 1: $\Omega = (0,1)^2$ and its boundary Ω.

In the following, we show how to numerically solve the Poisson problem by using finite differences.

Let N be a positive integer and set the *stepsize* $h = \frac{1}{N}$.
The plane $\mathbb{R}^2$ is discretized by the mesh (see Figure 2)

$$\mathbb{R}_h^2 := \{(mh, nh) : m, n \in \mathbb{Z}\}$$

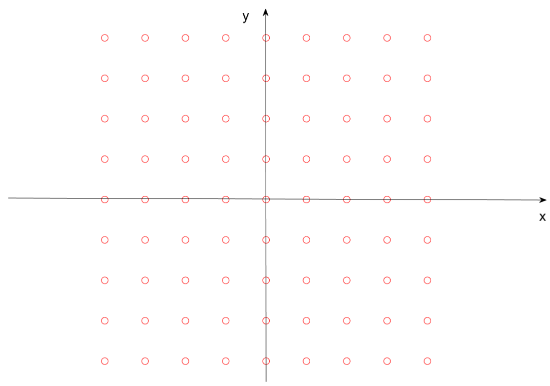


Figure 2: The mesh $\mathbb{R}_h^2$.

Each mesh point $(mh, nh) \in \mathbb{R}_h^2$ has four nearest neighbors in the mesh (see Figure 3):

- $((m-1)h, nh)$ to the left,
- $((m+1)h, nh)$ to the right,
- $(mh, (n-1)h)$ below,
- $(mh, (n+1)h)$ above.

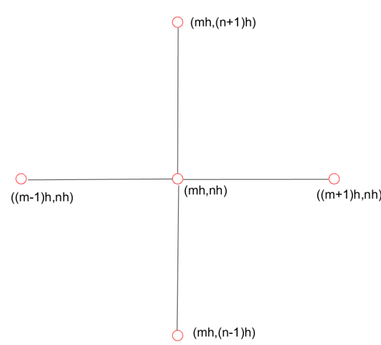


Figure 3: The four nearest neighbors of (mh, nh).

Now, we introduce discretizazions of Ω, Γ and $\overline{\Omega}$ (see Figure 4).

- Let $\Omega_h = \Omega \cap \mathbb{R}_h^2$. We consider Ω_h as a discretization of Ω.
- Let Γ_h be the set of the mesh points not in Ω_h but with a nearest neighbor in Ω_h. We consider Γ_h as a discretization of Γ. Note that $\Gamma_h \subseteq \Gamma$ but $\Gamma_h \neq \Gamma \cap \mathbb{R}_h^2$: the four corners of Γ are in $\mathbb{R}_h^2$ but not in Γ_h.
- Let $\overline{\Omega}_h = \Omega_h \cup \Gamma_h$. We consider $\overline{\Omega}_h$ as a discretization of $\overline{\Omega}$.

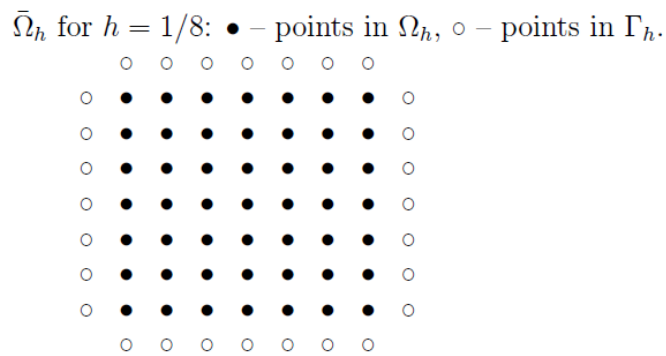


Figure 4: The discretizations Ω_h, Γ_h and $\overline{\Omega}_h$.

In the discretization of the BVP (1), we look for a function $u_h : \overline{\Omega}_h \to \mathbb{R}$ such that

$$\begin{cases} \Delta_h u_h (x, y) = f (x, y) \, , \; (x, y) \in \Omega_h, \\ u_h (x, y) = g (x, y) \, , \; (x, y) \in \Gamma_h, \end{cases}$$

Here Δ_h is a *discrete Laplacian*, i.e. an operator such that:

- it associates to a mesh function $\overline{\Omega}_h \to \mathbb{R}$ an interior mesh function $\Omega_h \to \mathbb{R}$;
- it approximates the Laplacian Δ in the sense that if $v : \overline{\Omega} \to \mathbb{R}$ is a sufficiently smooth function, then

  $$\Delta_h v|_{\overline{\Omega}_h} (x, y) \approx \Delta v (x, y) \, , \; (x, y) \in \Omega_h,$$

  where $v|_{\overline{\Omega}_h}$ is the restriction of v to $\overline{\Omega}_h \subseteq \overline{\Omega}$.

Now, we define a discrete Laplacian Δ_h by finite differences.

1.1 The five-point discretization of the Laplacian

Given a sufficiently smooth function $v(t)$ of one real variable t, we want to construct a finite difference approximating the second derivative $v''(t)$.

Let $h > 0$. We have

$$\begin{aligned} v''(t) &\approx \frac{v'(t+h) - v'(t)}{h} \approx \frac{\frac{v(t)-v(t+h)}{-h} - \frac{v(t-h)-v(t)}{-h}}{h} \\ &= \frac{\frac{v(t+h)-v(t)}{h} - \frac{v(t)-v(t-h)}{h}}{h} = \frac{v(t-h) - 2v(t) + v(t+h)}{h^2}. \end{aligned}$$

So, an approximation of $v''(t)$ is given by

$$v''(t) \approx \frac{v(t-h) - 2v(t) + v(t+h)}{h^2}. \tag{2}$$

The scheme (2) for approximating $v''(t)$ is called the *second central difference*.

What about the error of this approximation? If v is of class C^4, we have

$$\begin{aligned} v(t-h) &= v(t) - hv'(t) + \frac{h^2}{2}v''(t) - \frac{h^3}{6}v'''(t) + \frac{h^4}{24}v^{(4)}(\alpha_h) \\ v(t+h) &= v(t) + hv'(t) + \frac{h^2}{2}v''(t) + \frac{h^3}{6}v'''(t) + \frac{h^4}{24}v^{(4)}(\beta_h), \end{aligned}$$

where $\alpha_h \in (t-h, t)$ and $\beta_h \in (t, t+h)$. Thus, we obtain

$$v(t-h) + v(t+h) = 2v(t) + h^2 v''(t) + \frac{h^4}{24}\left(v^{(4)}(\alpha_h) + v^{(4)}(\beta_h)\right).$$

and then

$$\frac{v(t-h) - 2v(t) + v(t+h)}{h^2} - v''(t) = \frac{h^2}{24}\left(v^{(4)}(\alpha_h) + v^{(4)}(\beta_h)\right).$$

Now, we define a discrete Laplacian. Since for function $v : \overline{\Omega} \to \mathbb{R}$ of class C^2 in Ω

$$\Delta v(x,y) = \frac{\partial^2 v}{\partial x^2}(x,y) + \frac{\partial^2 v}{\partial y^2}(x,y), \ (x,y) \in \Omega,$$

we can set, for a mesh function $v_h : \overline{\Omega}_h \to \mathbb{R}$

$$\begin{aligned} \Delta_h v_h(x,y) &= \frac{v_h(x-h,y) - 2v_h(x,y) + v_h(x+h,y)}{h^2} \\ &\quad + \frac{v_h(x,y-h) - 2v_h(x,y) + v_h(x,y+h)}{h^2}, \\ &\quad (x,y) \in \Omega_h. \end{aligned}$$

Note that, for $(x,y) \in \Omega_h$, we have $(x \pm h, y), (x, y \pm h) \in \overline{\Omega}_h$. This the reason for which $\Delta_h v_h : \Omega_h \to \mathbb{R}$ is defined for $v_h : \overline{\Omega}_h \to \mathbb{R}$.

This discrete Laplacian Δ_h is known as the *five-point discretization of the Laplacian.*

Now, we study how well Δ_h approximates Δ.

Let $v : \overline{\Omega} \to \mathbb{R}$ be of class C^4. We have, for $(x,y) \in \Omega_h$,

$$\begin{aligned}
&\left|\Delta_h v|_{\overline{\Omega}_h}(x,y) - \Delta v(x,y)\right| \\
&= \left|\frac{v(x-h,y) - 2v(x,y) + v(x+h,y)}{h^2} + \frac{v(x,y-h) - 2v(x,y) + v(x,y+h)}{h^2}\right. \\
&\qquad \left. - \left(\frac{\partial^2 v}{\partial x^2}(x,y) + \frac{\partial^2 v}{\partial y^2}(x,y)\right)\right| \\
&\le \left|\frac{v(x-h,y) - 2v(x,y) + v(x+h,y)}{h^2} - \frac{\partial^2 v}{\partial x^2}(x,y)\right| \\
&\quad + \left|\frac{v(x,y-h) - 2v(x,y) + v(x,y+h)}{h^2} - \frac{\partial^2 v}{\partial y^2}(x,y)\right| \\
&= \frac{h^2}{24}\left|\frac{\partial^4 v}{\partial x^4}(\alpha_h,y) + \frac{\partial^4 v}{\partial x^4}(\beta_h,y)\right| \\
&\quad + \frac{h^2}{24}\left|\frac{\partial^4 v}{\partial y^4}(x,\gamma_h) + \frac{\partial^4 v}{\partial y^4}(x,\delta_h)\right|
\end{aligned}$$

with $\alpha_h \in (x-h,x)$, $\beta_h \in (x,x+h)$, $\gamma_h \in (y-h,y)$ and $\delta_h \in (y,y+h)$.

Thus: for $v : \overline{\Omega} \to \mathbb{R}$ of class C^4, we have

$$\begin{aligned}
&\max_{(x,y)\in\Omega_h} \left|\Delta_h v|_{\overline{\Omega}_h}(x,y) - \Delta v(x,y)\right| \\
&\le \frac{h^2}{6}\max\left\{\max_{(x,y)\in\overline{\Omega}}\left|\frac{\partial^4 v}{\partial x^4}(x,y)\right|, \max_{(x,y)\in\overline{\Omega}}\left|\frac{\partial^4 v}{\partial y^4}(x,y)\right|\right\}.
\end{aligned}$$

Exercise. Given $v : \overline{\Omega} \to \mathbb{R}$ of class C^6, find a function $C(x,y)$, $(x,y) \in \overline{\Omega}$, and a nonnegative constant M such that

$$\max_{(x,y)\in\Omega_h} \left|\Delta_h v|_{\overline{\Omega}_h}(x,y) - \Delta v(x,y) - C(x,y)h^2\right| \le Mh^4.$$

Just above we have shown the expansion

$$\Delta_h v|_{\overline{\Omega}_h}(x,y) = \Delta v(x,y) + E_2(x,y),\ (x,y) \in \Omega,$$

where

$$\max_{(x,y)\in\Omega_h} |E_2(x,y)| \le \frac{h^2}{6}\max\left\{\max_{(x,y)\in\overline{\Omega}}\left|\frac{\partial^4 v}{\partial x^4}(x,y)\right|, \max_{(x,y)\in\overline{\Omega}}\left|\frac{\partial^4 v}{\partial y^4}(x,y)\right|\right\}.$$

By this exercise , we can expand one more term:

$$\Delta_h v|_{\overline{\Omega}_h}(x,y) = \Delta v(x,y) + C(x,y)h^2 + E_4(x,y), (x,y) \in \Omega,$$

where

$$\max_{(x,y)\in\Omega_h} |E_4(x,y)| \le Mh^4.$$

1.2 The discrete problem as a linear system

The discrete problem

$$\begin{cases} \Delta_h u_h(x,y) = f(x,y)\,, \ (x,y) \in \Omega_h, \\ u_h(x,y) = g(x,y)\,, \ (x,y) \in \Gamma_h, \end{cases}$$

is a system of $(N-1)^2$ scalar linear equations into $(N-1)^2$ scalar unknowns.

In fact, the discrete Poisson equation is

$$\begin{aligned} &\frac{u_h(x-h,y) - 2u_h(x,y) + u_h(x+h,y)}{h^2} + \frac{u_h(x,y-h) - 2u_h(x,y) + u_h(x,y+h)}{h^2} \\ &= \frac{u_h(x,y-h) + u_h(x-h,y) - 4u_h(x,y) + u_h(x+h,y) + u_h(x,y+h)}{h^2} \\ &= f(x,y), \ (x,y) \in \Omega_h. \end{aligned}$$

By setting

$$\begin{aligned} &u_{ij} := u_h(ih,jh)\,, \ (ih,jh) \in \overline{\Omega}_h, \ \text{i.e.} \ (i,j) \in \{0,1,\ldots,N-1,N\}^2 \setminus \{(0,0),(N,0),(N,N),(0.N)\}, \\ &f_{ij} := f(ih,jh)\,, \ (ih,jh) \in \Omega_h, \ \text{i.e.} \ (i,j) \in \{1,\ldots,N-1\}^2\,, \\ &g_{ij} := g(ih,jh)\,, \ (ih,jh) \in \Gamma_h, \ \text{i.e.} \quad (i,j) \in \{0,N\} \times \{1,\ldots,N-1\} \cup \{1,\ldots,N-1\} \times \{0,N\}\,, \end{aligned}$$

the discrete Poisson problem becomes

$$\begin{cases} u_{ij-1} + u_{i-1j} - 4u_{ij} + u_{i+1j} + u_{ij+1} = h^2 f_{ij}, \ (i,j) \in \{1,\ldots,N-1\}^2 \\ u_{ij} = g_{ij}, \ (i,j) \in \{0,N\} \times \{1,\ldots,N-1\} \cup \{1,\ldots,N-1\} \times \{0,N\}\,. \end{cases}$$

This is the system of the $(N-1)^2$ scalar linear equations

$$u_{ij-1} + u_{i-1j} - 4u_{ij} + u_{i+1j} + u_{ij+1} = h^2 f_{ij}, \ (i,j) \in \{1,\ldots,N-1\}^2\,,$$

into the $(N-1)^2$ scalar unknowns

$$u_{ij}, \ (i,j) \in \{1,\ldots,N-1\}^2\,.$$

We want to write the linear system in the form

$$\mathcal{A}U = b,$$

where

- $\mathcal{A}$ is the square matrix of the system of order $(N-1)^2$.
- U is the column vector of the unknowns of dimension $(N-1)^2$.
- b is the column vector of the known terms of dimension $(N-1)^2$.

In order to do this, the $(N-1)^2$ unknowns

$$\begin{array}{l} u_{1,N-1}, u_{2,N-1}, \ldots, u_{N-1,N-1} \\ \qquad\qquad \vdots \\ u_{1,2}, u_{2,2}, \ldots, u_{N-1,2} \\ u_{1,1}, u_{2,1}, \ldots, u_{N-1,1}, \end{array}$$

which are arranged in a square because they are values at the mesh points in

$$\Omega_h : \quad \begin{array}{cccc} \circ & \circ & \cdots & \circ \\ & \vdots & & \\ \circ & \circ & \cdots & \circ \\ \circ & \circ & \cdots & \circ \end{array} \quad ,$$

need to be ordered in a column vector.

For example, in the case $N = 4$, by ordering the $(N-1)^2 = 9$ unknowns as

$$U = (u_{11}, u_{21}, u_{31}, \quad u_{12}, u_{22}, u_{32}, \quad u_{13}, u_{23}, u_{33})$$

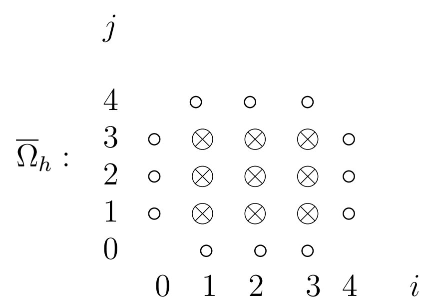
we obtain the following linear system:

$$\begin{bmatrix} -4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -4 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -4 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & -4 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & -4 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & -4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 \end{bmatrix} \begin{bmatrix} u_{11} \\ u_{21} \\ u_{31} \\ u_{12} \\ u_{22} \\ u_{32} \\ u_{13} \\ u_{23} \\ u_{33} \end{bmatrix} = \begin{bmatrix} h^2 f_{11} - g_{10} - g_{01} \\ h^2 f_{21} - g_{20} \\ h^2 f_{31} - g_{30} - g_{41} \\ h^2 f_{12} - g_{02} \\ h^2 f_{22} \\ h^2 f_{32} - g_{42} \\ h^2 f_{13} - g_{03} - g_{14} \\ h^2 f_{23} - g_{24} \\ h^2 f_{33} - g_{43} - g_{34} \end{bmatrix}.$$

In fact, we have the nine equations

$$u_{ij-1} + u_{i-1j} - 4u_{ij} + u_{i+1j} + u_{ij+1} = h^2 f_{ij}, \ (i,j) \in \{1,2,3\}^2,$$

namely



,

$$\begin{aligned}
&u_{10} + u_{01} - 4u_{11} + u_{21} + u_{12} = h^2 f_{11}, \ u_{10} = g_{10} \text{ and } u_{01} = g_{01} \\
&u_{20} + u_{11} - 4u_{21} + u_{31} + u_{22} = h^2 f_{21}, \ u_{20} = g_{20} \\
&u_{30} + u_{21} - 4u_{31} + u_{41} + u_{32} = h^2 f_{31}, \ u_{30} = g_{30} \text{ and } u_{41} = g_{41} \\
&u_{11} + u_{02} - 4u_{12} + u_{22} + u_{13} = h^2 f_{12}, \ u_{02} = g_{02} \\
&u_{21} + u_{12} - 4u_{22} + u_{32} + u_{23} = h^2 f_{22} \\
&u_{31} + u_{22} - 4u_{32} + u_{42} + u_{33} = h^2 f_{32}, \ u_{42} = g_{42} \\
&u_{12} + u_{03} - 4u_{13} + u_{23} + u_{14} = h^2 f_{13}, \ u_{03} = g_{03} \text{ and } u_{14} = g_{14} \\
&u_{22} + u_{13} - 4u_{23} + u_{33} + u_{24} = h^2 f_{23}, \ u_{24} = g_{24} \\
&u_{32} + u_{23} - 4u_{33} + u_{43} + u_{34} = h^2 f_{33}, \ u_{43} = g_{43} \text{ and } u_{34} = g_{34}.
\end{aligned}$$

The matrix $\mathcal{A}$ of the system can be rewritten as a 3×3 block matrix:

$$\mathcal{A} = \begin{bmatrix} A & I & O \\ I & A & I \\ O & I & A \end{bmatrix}$$

where the blocks are 3×3 and they are

$$A = \begin{bmatrix} -4 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & -4 \end{bmatrix}, \quad \text{the identity matrix } I \text{ and the zero matrix } O.$$

For a general N, by ordering the $(N-1)^2$ unknowns as

$$U = (u_{11}, u_{21}, \ldots, u_{N-1\ 1}, \quad u_{12}, u_{22}, \ldots, u_{N-1\ 2}, \quad \ldots \quad, u_{1\ N-1}, u_{2\ N-1}, \ldots, u_{N-1\ N-1})$$

the matrix $\mathcal{A}$ of the system is a $(N-1) \times (N-1)$ tridiagonal block matrix:

$$\mathcal{A} = \begin{bmatrix} A & I & O & . & . & O \\ I & A & I & . & & . \\ O & I & . & . & . & . \\ . & . & . & . & . & O \\ . & & . & . & . & I \\ O & . & . & O & I & A \end{bmatrix},$$

where the blocks are $(N-1) \times (N-1)$ and they are

$$A = \begin{bmatrix} -4 & 1 & 0 & . & . & 0 \\ 1 & . & . & . & & . \\ 0 & . & . & . & . & . \\ . & . & . & . & . & 0 \\ . & & . & . & -4 & 1 \\ 0 & . & . & 0 & 1 & -4 \end{bmatrix}, \quad \text{the identity matrix } I \text{ and the zero matrix } O.$$

The matrix $\mathcal{A}$ is *sparse*, i.e the number of nonzero elements of $\mathcal{A}$ is $O(n)$, $n = (N-1)^2$ being the order of $\mathcal{A}$. In fact, in each row there are at most 5 nonzero elements and so the total number of nonzero elements is not larger than $5n$.

Observe that the matrix $\mathcal{A}$ has only 5 nonzero diagonals: the principal diagonal, the first lower diagonal, the first upper diagonal, the $(N-1)$-th lower diagonal and the $(N-1)$-th upper diagonal. All the diagonals below the $(N-1)$-th lower diagonal and above the $(N-1)$-th upper diagonal are zero. This means that $\mathcal{A}$ is a *band matrix* with bandwidth $w = N - 1 = \sqrt{n}$.

By solving the linear system $\mathcal{A}U = b$ by gaussian elimination, the *fill-in phenomenon* will appear: in the LU factorization of a band matrix, the band structure is preserved in the matrix L and U with the same bandwidth. So, in the $\mathcal{A} = \mathcal{L}\mathcal{U}$ factorization of the matrix $\mathcal{A}$, the lower triangular matrix $\mathcal{L}$ is a band matrix with bandwidth w and the upper triangular matrix $\mathcal{U}$ is a band matrix with bandwidth w. But, what happens for the matrix $\mathcal{A}$ is that all the lower diagonals of $\mathcal{L}$ up to the order w and all the upper diagonals of $\mathcal{U}$ up to the order w are filled with nonzero elements. This means that gaussian elimination cannot take advantage of the sparsity of the matrix and the computational cost for solving $\mathcal{A}U = b$ is the cost when a general linear system with a band matrix of bandwidth w is solved, namely

$$\text{number of arithmetic operations} = O(w^2 n) = O(n^2) = O((N-1)^4).$$

Since N is a large number, the gaussian elimination is too expensive. Iterative methods for solving linear systems should be used.

The matrix $\mathcal{A}$ is symmetric because the blocks A, I and O are symmetric. Moreover, we have the following fact.

Proposition 1 *$\mathcal{A}$ is negative definite, i.e. for any $v \in \mathbb{R}^{(N-1)^2}$ we have $\langle v, \mathcal{A}v\rangle \le 0$ and $\langle v, \mathcal{A}v\rangle < 0$ if $v \neq 0$.*

Proof. We begin to prove that, for $v \in \mathbb{R}^{N-1}$, we have

$$\langle v, Av\rangle \le -2\,\|v\|^2$$

for the block

$$A = \begin{bmatrix} -4 & 1 & 0 & . & . & 0 \\ 1 & . & . & . & & . \\ 0 & . & . & . & . & . \\ . & . & . & . & . & 0 \\ . & & . & . & -4 & 1 \\ 0 & . & . & 0 & 1 & -4 \end{bmatrix}.$$

We have

$$\begin{aligned}
&\langle v, Av\rangle = v^T A v = \sum_{i,j=1}^{N-1} v_i a_{ij} v_j \\
&= -4v_1^2 + v_1v_2 + v_2v_1 - 4v_2^2 + v_2v_3 + v_3v_2 - 4v_3^2 \\
&\quad + \cdots \\
&\quad -4v_{N-2}^2 + v_{N-2}v_{N-1} + v_{N-1}v_{N-2} - 4v_{N-1}^2 \\
&= -4v_1^2 - 4v_2^2 - 4v_3^2 + \cdots - 4v_{N-2}^2 - 4v_{N-1}^2 \\
&\quad +2v_1v_2 + 2v_2v_3 + \cdots + 2v_{N-2}v_{N-1} \\
&= -2v_1^2 - 2v_2^2 - 2v_3^2 + \cdots - 2v_{N-2}^2 - 2v_{N-1}^2 \\
&\quad -v_1^2 - (v_1^2 - 2v_1v_2 + v_2^2) - (v_2^2 - 2v_2v_3 + v_3^2) + \cdots - (v_{N-2}^2 - 2v_{N-2}v_{N-1} + v_{N-1}^2) - v_{N-1}^2 \\
&= -2v_1^2 - 2v_2^2 - 2v_3^2 + \cdots - 2v_{N-2}^2 - 2v_{N-1}^2 \\
&\quad -v_1^2 - (v_1 - v_2)^2 - (v_2 - v_3)^2 + \cdots - (v_{N-2} - v_{N-1})^2 - v_{N-1}^2 \\
&\le -2\left(v_1^2 + v_2^2 + \cdots + v_{N-2}^2 + v_{N-1}^2\right) = -2\|v\|^2.
\end{aligned}$$

Now, we prove that the matrix $\mathcal{A}$ is negative semi-definite. For

$$v = (v_1, v_2, \ldots, v_{N-1}) \in \mathbb{R}^{(N-1)^2}$$

with $v_i \in \mathbb{R}^{N-1}$, $i \in \{1, \ldots, N-1\}$, we have

$$\begin{aligned}
&\langle v, \mathcal{A}v \rangle = v^T \mathcal{A} v \\
&= \begin{bmatrix} v_1^T & v_2^T & . & . & . & v_{N-1}^T \end{bmatrix} \begin{bmatrix} A & I & O & . & . & O \\ I & . & . & . & & . \\ O & . & . & . & . & . \\ . & . & . & . & . & O \\ . & & . & . & A & I \\ O & . & . & O & I & A \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ . \\ . \\ . \\ v_{N-1} \end{bmatrix} \\
&= v_1^T A v_1 + v_1^T v_2 + v_2^T v_1 + v_2^T A v_2 + v_2^T v_3 + v_3^T v_2 + v_3^T A v_3 \\
&\quad + \cdots \\
&\quad + v_{N-2}^T A v_{N-2} + v_{N-2}^T v_{N-1} + v_{N-1}^T v_{N-2} + v_{N-1}^T A v_{N-1} \\
&= \langle v_1, Av_1 \rangle + 2 \langle v_1, v_2 \rangle + \langle v_2, Av_2 \rangle + 2 \langle v_2, v_3 \rangle + \langle v_3, Av_3 \rangle \\
&\quad + \cdots \\
&\quad + \langle v_{N-2}, Av_{N-2} \rangle + 2 \langle v_{N-2}, v_{N-1} \rangle + \langle v_{N-1}, Av_{N-1} \rangle \\
&\leq -2 \|v_1\|^2 + 2 \langle v_1, v_2 \rangle - 2 \|v_2\|^2 + 2 \langle v_2, v_3 \rangle - 2 \|v_3\|^2 \\
&\quad + \cdots \\
&\quad -2 \|v_{N-2}\|^2 + 2 \langle v_{N-2}, v_{N-1} \rangle - 2 \|v_{N-1}\|^2 \\
&= - \|v_1\|^2 - \left(\|v_1\|^2 - 2 \langle v_1, v_2 \rangle + \|v_2\|^2 \right) - \left(\|v_2\|^2 - 2 \langle v_2, v_3 \rangle + \|v_3\|^2 \right) \\
&\quad + \cdots \\
&\quad - \left(\|v_{N-2}\|^2 - 2 \langle v_{N-2}, v_{N-1} \rangle + \|v_{N-1}\|^2 \right) - \|v_{N-1}\|^2 \\
&= - \|v_1\|^2 - \|v_1 - v_2\|^2 - \|v_2 - v_3\|^2 + \cdots - \|v_{N-2} - v_{N-1}\|^2 - \|v_{N-1}\|^2 \leq 0.
\end{aligned}$$

So, the matrix $\mathcal{A}$ is negative semidefinite.

Finally, we show that $\mathcal{A}$ is negative definite. Our previous derivation shows that for $v = (v_1, v_2, \ldots, v_{N-1}) \in \mathbb{R}^{(N-1)^2}$ such that

$$\langle v, \mathcal{A}v \rangle = 0$$

we have

$$- \|v_1\|^2 - \|v_1 - v_2\|^2 + \ldots - \|v_{N-2} - v_{N-1}\|^2 - \|v_{N-1}\|^2 = 0.$$

Then

$$v_1 = 0,\ v_1 - v_2 = 0,\ \ldots,\ v_{N-2} - v_{N-1} = 0,\ v_{N-1} = 0,$$

and so $v_1 = v_2 = \ldots = v_{N-1} = 0$.

So, the matrix is negative definite.

■

Since the matrix $\mathcal{A}$ is negative definite, it is nonsingular. As a consequence, we have that the discrete problem

$$\begin{cases} \Delta_h u_h(x,y) = f(x,y), & (x,y) \in \Omega_h, \\ u_h(x,y) = g(x,y), & (x,y) \in \Gamma_h, \end{cases}$$

has a unique solution u_h.

Exercise. Write the 1D BVP given by the Poisson equation with Dirichlet boundary condition on $\Omega = (0,1)$. Solve analytically this problem. Propose a corresponding discrete problem and write the associated linear system. Compute the computational cost for solving the linear system by gaussian elimination. Prove that the discrete problem has a unique solution.

Exercise. Consider the 3D BVP given by the Poisson equation with Dirichlet boundary condition on $\Omega = (0,1)^3$. Propose a discrete Laplacian and a consequent discrete problem. Describe, in the matrix of the linear system associated to the discrete problem, the row corresponding to a mesh point in Ω_h with nearest neighbors in Ω_h. Compute the computational cost for solving the linear system by gaussian elimination.

2 Convergence Analysis

Let us assume that the BVP (1) given by the Poisson equation with Dirichlet boundary condition has a unique solution u.

Now, by using discrete maximum and minimum principles introduced below,

- we prove that the discrete problem has a unique solution u_h; indeed, this is already known since we have proved that the matrix of the linear system corresponding to the discrete problem is negative definite and then nonsingular;

- we give an upper bound for the error

$$\max_{(x,y)\in\overline{\Omega}_h} |u_h(x,y) - u(x,y)|.$$

2.1 The Discrete Maximum principle

We remind *the Maximum Principle for the Laplacian*: let $v : \overline{\Omega} \to \mathbb{R}$ be a continuous function of class C^2 in Ω such that

$$\Delta v(x,y) \geq 0, \ (x,y) \in \Omega,$$

i.e. v is a *subharmonic function.* Then

$$\sup_{(x,y)\in\Omega} v(x,y) \leq \max_{(x,y)\in\Gamma} v(x,y)$$

with equality if and only if v is constant.

The Discrete Maximum Principle is the discrete analogue of the previous fact.

Theorem 2 (Discrete Maximum Principle)*. Let $v_h : \overline{\Omega}_h \to \mathbb{R}$ be a mesh function such that*

$$\Delta_h v_h(x,y) \geq 0, \ (x,y) \in \Omega_h.$$

Then

$$\max_{(x,y)\in\Omega_h} v_h(x,y) \leq \max_{(x,y)\in\Gamma_h} v_h(x,y) \tag{3}$$

with equality if and only if v_h is constant.

Proof. If v_h is constant, then (3) holds with equality. Now, we prove that

$$v_h \text{ is not constant } \Rightarrow \max_{(x,y)\in\Omega_h} v_h(x,y) < \max_{(x,y)\in\Gamma_h} v_h(x,y)$$

i.e.

$$\max_{(x,y)\in\Omega_h} v_h(x,y) \geq \max_{(x,y)\in\Gamma_h} v_h(x,y) \Rightarrow v_h \text{ is constant }.$$

So, we assume

$$\max_{(x,y)\in\Omega_h} v_h(x,y) \geq \max_{(x,y)\in\Gamma_h} v_h(x,y),$$

from which we have

$$\max_{(x,y)\in\Omega_h} v_h(x,y) = \max_{(x,y)\in\overline{\Omega}_h} v_h(x,y),$$

and we show that v_h is constant.

Let $(x_0, y_0) \in \Omega_h$ the point where the maximum of v_h is obtained.
Since

$$\begin{aligned}&0 \leq h^2 \cdot \Delta_h v_h(x_0,y_0)\\&= -4v_h(x_0,y_0) + v_h(x_0-h,y_0) + v_h(x_0+h,y_0) + v_h(x_0,y_0-h) + v_h(x_0,y_0+h),\end{aligned}$$

we obtain

$$\begin{aligned}&4v_h(x_0,y_0)\\&\leq v_h(x_0-h,y_0) + v_h(x_0+h,y_0) + v_h(x_0,y_0-h) + v_h(x_0,y_0+h)\\&\leq 4v_h(x_0,y_0).\end{aligned}$$

This means that

$$\begin{aligned}&4v_h(x_0,y_0)\\&= v_h(x_0-h,y_0) + v_h(x_0+h,y_0) + v_h(x_0,y_0-h) + v_h(x_0,y_0+h).\end{aligned}$$

So

$$\begin{aligned}&v_h(x_0,y_0)\\&= v_h(x_0-h,y_0) = v_h(x_0+h,y_0) = v_h(x_0,y_0-h) = v_h(x_0,y_0+h).\end{aligned}$$

Thus, v_h achieves its maximum also at all the nearest neighbors

$$(x_0 - h, y_0), \ (x_0 + h, y_0), \ (x_0, y_0 - h), \ (x_0, y_0 + h) \tag{4}$$

of (x_0, y_0).

Applying the same argument to points in (4) belonging to Ω_h and then to the nearest neighbors in Ω_h of these points in (4) belonging to Ω_h , etc., we conclude that the maximum is obtained at all points in $\overline{\Omega}_h$, i.e. v_h is constant. ■

Exercise. Prove (in a very short proof) the *Discrete Minimum Principle*: if $v_h : \overline{\Omega}_h \to \mathbb{R}$ is a mesh function such that

$$\Delta_h v_h(x,y) \leq 0, \ (x,y) \in \Omega_h,$$

then

$$\min_{(x,y)\in\Omega_h} v_h(x,y) \geq \min_{(x,y)\in\Gamma_h} v_h(x,y)$$

with equality if and only if v_h is constant.

As an immediate consequence of the Discrete Maximum and Minimum Principles, we obtain the already proved existence and uniqueness of a solution for the discrete problem.

Theorem 3 *The discrete BVP*

$$\begin{cases} \Delta_h u_h(x,y) = f(x,y), \ (x,y) \in \Omega_h, \\ u_h(x,y) = g(x,y), \ (x,y) \in \Gamma_h, \end{cases}$$

has a unique solution.

Proof. Since this problem corresponds to a square linear system

$$\mathcal{A}U = b,$$

it is sufficient to prove that the matrix $\mathcal{A}$ of the system is nonsingular, i.e.

$$\mathcal{A}U = 0 \ \Rightarrow \ U = 0.$$

In other words, we have to prove that the discrete homogeneous problem

$$\begin{cases} \Delta_h u_h(x,y) = 0, \ (x,y) \in \Omega_h, \\ u_h(x,y) = 0, \ (x,y) \in \Gamma_h, \end{cases}$$

has the unique solution $u_h = 0$.

Let u_h be a solution of such a problem.

By the discrete maximum and minimum principles,

$$\Delta_h u_h(x,y) = 0, \ (x,y) \in \Omega_h,$$

implies

$$\begin{aligned} 0 = \min_{(x,y)\in\Gamma_h} u_h(x,y) &\leq \min_{(x,y)\in\Omega_h} u_h(x,y) \\ \leq \max_{(x,y)\in\Omega_h} u_h(x,y) &\leq \max_{(x,y)\in\Gamma_h} u_h(x,y) = 0. \end{aligned}$$

We conclude that $u_h = 0$. ■

2.2 The stability theorem

We measure the size of functions by using the ∞ *norm*, also called the L^{∞} *norm*.

Given a function $f : A \to \mathbb{R}$, we set

$$\|f\|_{L^{\infty}(A)} := \sup_{x \in A} |f(x)| .$$

Of course, if A is finite or $A \subseteq \mathbb{R}^d$ is closed bounded (i.e. compact) and f is continuous, we have

$$\|f\|_{L^{\infty}(A)} = \max_{x \in A} |f(x)| .$$

The set $L^{\infty}(A)$ is the set of the functions $f : A \to \mathbb{R}$ such that $\|f\|_{L^{\infty}(A)} < +\infty$. So, when A is finite, $L^{\infty}(A)$ is simply the set of functions $A \to \mathbb{R}$.

Note that

$$\left\|\Delta_h(v|_{\overline{\Omega}_h}) - (\Delta v)|_{\Omega_h}\right\|_{L^{\infty}(\Omega_h)} \leq \frac{h^2}{6} \max\left\{\left\|\frac{\partial^4 v}{\partial x^4}\right\|_{L^{\infty}(\overline{\Omega})}, \left\|\frac{\partial^4 v}{\partial y^4}\right\|_{L^{\infty}(\overline{\Omega})}\right\}$$

for $v : \overline{\Omega} \to \mathbb{R}$ of class C^4.

Next result states that the mapping

$$(f, g) \mapsto u_h,\ f \in L^{\infty}(\Omega_h) \text{ and } g \in L^{\infty}(\Gamma_h),$$

where u_h is the solution of the discrete BVP

$$\begin{cases} \Delta_h u_h(x, y) = f(x, y),\ (x, y) \in \Omega_h, \\ u_h(x, y) = g(x, y),\ (x, y) \in \Gamma_h, \end{cases} \tag{5}$$

is uniformly bounded with respect to h, i.e. there exists a constant C, independent of h, such that for any $f \in L^{\infty}(\Omega_h)$ and for any $g \in L^{\infty}(\Gamma_h)$ we have

$$\|u_h\|_{L^{\infty}(\overline{\Omega}_h)} \leq C \max\left\{\|f\|_{L^{\infty}(\Omega_h)}, \|g\|_{L^{\infty}(\Gamma_h)}\right\}.$$

Theorem 4 (The Stability Theorem) *Let $f \in L^{\infty}(\Omega_h)$, let $g \in L^{\infty}(\Gamma_h)$ and let u_h be the solution of the discrete BVP (5). Then*

$$\|u_h\|_{L^{\infty}(\overline{\Omega}_h)} \leq \frac{1}{8}\|f\|_{L^{\infty}(\Omega_h)} + \|g\|_{L^{\infty}(\Gamma_h)} .$$

Proof. Let us introduce the mesh function

$$\phi(x, y) = \frac{\left(x - \frac{1}{2}\right)^2 + \left(y - \frac{1}{2}\right)^2}{4},\ (x, y) \in \overline{\Omega}_h.$$

Exercise. Prove that

$$\Delta_h \phi(x, y) = 1,\ (x, y) \in \Omega_h,$$

and

$$0 \le \phi(x,y) < \frac{1}{8}, \ (x,y) \in \overline{\Omega}_h.$$

Exercise. Prove that Δ_h is a linear operator, i.e.

$$\Delta_h(v_h + w_h) = \Delta_h v_h + \Delta_h w_h$$

and

$$\Delta_h(\alpha v_h) = \alpha \Delta_h v_h$$

for any mesh functions $v_h, w_h : \overline{\Omega}_h \to \mathbb{R}$ and $\alpha \in \mathbb{R}$.

Set

$$A := \|f\|_{L^\infty(\Omega_h)}, \ B := \|g\|_{L^\infty(\Gamma_h)}.$$

We have, for $(x,y) \in \Omega_h$,

$$\begin{aligned} \Delta_h (u_h + A\phi)(x,y) &= \Delta_h u_h(x,y) + A\Delta_h \phi(x,y) \\ &= f(x,y) + A \ge -|f(x,y)| + A \\ &\ge 0. \end{aligned}$$

By the discrete maximum principle, we have

$$\begin{aligned} \max_{(x,y)\in\Omega_h} u_h(x,y) &\le \max_{(x,y)\in\Omega_h} (u_h + A\phi)(x,y) \le \max_{(x,y)\in\Gamma_h} (u_h + A\phi)(x,y) \\ &= \max_{(x,y)\in\Gamma_h} (g(x,y) + A\phi(x,y)) \le B + \frac{1}{8}A. \end{aligned}$$

Moreover, for $(x,y) \in \Omega_h$,

$$\begin{aligned} \Delta_h (u_h - A\phi)(x,y) &= \Delta_h u_h(x,y) - A\Delta_h \phi(x,y) \\ &= f(x,y) - A \le |f(x,y)| - A \\ &\le 0. \end{aligned}$$

By the discrete minimum principle, we have

$$\begin{aligned} \min_{(x,y)\in\Omega_h} u_h(x,y) &\ge \min_{(x,y)\in\Omega_h} (u_h - A\phi)(x,y) \ge \min_{(x,y)\in\Gamma_h} (u_h - A\phi)(x,y) \\ &= \min_{(x,y)\in\Gamma_h} (g(x,y) - A\phi(x,y)) \ge -B - \frac{1}{8}A. \end{aligned}$$

Thus

$$-B - \frac{1}{8}A \le u_h(x,y) \le B + \frac{1}{8}A, \ (x,y) \in \overline{\Omega}_h.$$

■

2.3 The Convergence Theorem

Now, we are in position to give a bound for the error

$$\left\|u_h - u|_{\overline{\Omega}_h}\right\|_{L^\infty\left(\overline{\Omega}_h\right)} = \max_{(x,y)\in\overline{\Omega}_h} \left|u_h(x,y) - u(x,y)\right|.$$

Theorem 5 (The Convergence Theorem) *Let u be the solution of the continuous BVP*

$$\begin{cases} \Delta u\left(x,y\right) = f\left(x,y\right), \; \left(x,y\right) \in \Omega, \\ u\left(x,y\right) = g\left(x,y\right), \; \left(x,y\right) \in \Gamma, \end{cases}$$

and let u_h be the solution of the discrete BVP

$$\begin{cases} \Delta_h u_h\left(x,y\right) = f\left(x,y\right), \; \left(x,y\right) \in \Omega_h, \\ u_h\left(x,y\right) = g\left(x,y\right), \; \left(x,y\right) \in \Gamma_h. \end{cases}$$

Then

$$\left\|u_h - u|_{\overline{\Omega}_h}\right\|_{L^\infty\left(\overline{\Omega}_h\right)} \leq \frac{1}{8}\left\|\Delta_h(u|_{\overline{\Omega}_h}) - (\Delta u)|_{\Omega_h}\right\|_{L^\infty(\Omega_h)}.$$

Proof. Let

$$e_h := u_h - u|_{\overline{\Omega}_h}.$$

Note that

$$\begin{aligned} \Delta_h e_h &= \Delta_h\left(u_h - u|_{\overline{\Omega}_h}\right) = \Delta_h u_h - \Delta_h\left(u|_{\overline{\Omega}_h}\right) \\ &= \Delta_h u_h - (\Delta u)|_{\Omega_h} + (\Delta u)|_{\Omega_h} - \Delta_h(u|_{\overline{\Omega}_h}) \\ &= (\Delta u)|_{\Omega_h} - \Delta_h(u|_{\overline{\Omega}_h}), \end{aligned}$$

where the last equality follows by

$$\Delta_h u_h(x,y) - \Delta u(x,y) = f(x,y) - f(x,y) = 0, \; (x,y) \in \Omega_h.$$

Moreover, we have

$$e_h\left(x,y\right) = u_h\left(x,y\right) - u\left(x,y\right) = g\left(x,y\right) - g\left(x,y\right) = 0, \; \left(x,y\right) \in \Gamma_h.$$

Then e_h solves the discrete boundary value problem

$$\begin{cases} \Delta_h e_h\left(x,y\right) = \left((\Delta u)|_{\Omega_h} - \Delta_h(u|_{\overline{\Omega}_h})\right)\left(x,y\right), \; \left(x,y\right) \in \Omega_h, \\ e_h\left(x,y\right) = 0, \; \left(x,y\right) \in \Gamma_h. \end{cases}$$

The thesis now follows by the Stability Theorem:

$$\begin{aligned} \left\|e_h\right\|_{L^\infty(\overline{\Omega}_h)} &\leq \frac{1}{8}\left\|(\Delta u)|_{\Omega_h} - \Delta_h(u|_{\overline{\Omega}_h})\right\|_{L^\infty(\Omega_h)} + \left\|0\right\|_{L^\infty(\Gamma_h)} \\ &= \frac{1}{8}\left\|\Delta_h(u|_{\overline{\Omega}_h}) - (\Delta u)|_{\Omega_h}\right\|_{L^\infty(\Omega_h)}. \end{aligned}$$

■

When u is of class C^4, since

$$\left\|\Delta_h(u|_{\overline{\Omega}_h}) - (\Delta u)|_{\Omega_h}\right\|_{L^\infty(\Omega_h)} \le \frac{h^2}{6} \max\left\{\left\|\frac{\partial^4 u}{\partial x^4}\right\|_{L^\infty(\overline{\Omega})}, \left\|\frac{\partial^4 u}{\partial y^4}\right\|_{L^\infty(\overline{\Omega})}\right\}$$

holds, we obtain

$$\left\|u_h - u|_{\overline{\Omega}_h}\right\|_{L^\infty(\overline{\Omega}_h)} \le \frac{h^2}{48} \max\left\{\left\|\frac{\partial^4 u}{\partial x^4}\right\|_{L^\infty(\overline{\Omega})}, \left\|\frac{\partial^4 u}{\partial y^4}\right\|_{L^\infty(\overline{\Omega})}\right\}.$$

2.4 Consistency, Stability and Convergence

We conclude this section with some remarks about the procedure used for obtaining the upper bound of $\left\|u_h - u|_{\overline{\Omega}_h}\right\|_{L^\infty(\overline{\Omega}_h)}$, where u and u_h are the solutions of the continuous and discrete, respectively, BVP.

Here, we assume that a general discrete Laplacian Δ_h is used, not necessarily the five-point discretization.

The mesh function

$$e_h = u_h - u|_{\overline{\Omega}_h} \in L^\infty\left(\overline{\Omega}_h\right)$$

is called the *convergence error* of the discretization. The mesh function

$$\varepsilon_h = \Delta_h\left(u|_{\overline{\Omega}_h}\right) - (\Delta u)\,|_{\Omega_h} \in L^\infty\left(\Omega_h\right)$$

is called the *consistency error* of the discretization.

The convergence error is the important error, but it is difficult to estimate its size. On the other hand, the size of the consistency error is not difficult to estimate.

Let p be a positive integer. The discretization is called *convergent of order* p if

$$\text{size of the convergence error} = \|e_h\|_{L^\infty(\overline{\Omega}_h)} = O(h^p),\ N \to \infty,$$

for a sufficiently smooth solution u. The discretization is called *consistent of order* p if

$$\text{size of the consistency error} = \|\varepsilon_h\|_{L^\infty(\Omega_h)} = O(h^p),\ N \to \infty,$$

for a sufficiently smooth solution u.

In case of the five-point discretization,

$$\text{size of the convergence error} = \|e_h\|_{L^\infty(\overline{\Omega}_h)} = O\left(h^2\right),\ N \to \infty,$$

and

$$\text{size of the consistency error} = \|\varepsilon_h\|_{L^\infty(\Omega_h)} = O\left(h^2\right),\ N \to \infty,$$

for u of class C^4. This means that the five-point discretization is convergent and consistent of order 2.

An upper bound

$$\|v_h\|_{L^\infty(\overline{\Omega}_h)} \le C_h \|f\|_{L^\infty(\Omega_h)},$$

where v_h is the solution of

$$\begin{cases} \Delta_h v_h(x,y) = f(x,y), \ (x,y) \in \Omega_h, \\ v_h(x,y) = 0, \ (x,y) \in \Gamma_h, \end{cases}$$

and C_h is independent of f, is called a *stability estimate*.

The discretization is called *stable* if it has a stabilty estimate with C_h independent of h.

The five-point discretization is stable, since we have a stability estimate with $C_h = \frac{1}{8}$.

We have

$$\begin{gathered} \text{consistency of order } p \text{ and stability} \\ \Downarrow \\ \text{convergence of order } p. \end{gathered}$$

In fact, as in the proof of The Convergence Theorem for the five-point discretization, the relation between the convergence error e_h and the consistency error ε_h is

$$\begin{cases} \Delta_h e_h(x,y) = -\varepsilon_h(x,y), \ (x,y) \in \Omega_h, \\ e_h(x,y) = 0, \ (x,y) \in \Gamma_h. \end{cases}$$

So, if

$$\|\varepsilon_h\|_{L^\infty(\Omega_h)} = O(h^p), \ N \to \infty, \quad \text{(consistency of order } p\text{)}$$

and

$$\|e_h\|_{L^\infty(\overline{\Omega}_h)} \le C_h \|\varepsilon_h\|_{L^\infty(\Omega_h)} \text{ with } C_h \text{ independent of } h \quad \text{(stability)},$$

then

$$\|e_h\|_{L^\infty(\Omega_h)} = O(h^p), \ N \to \infty, \quad \text{(convergence of order } p\text{)}.$$

3 Fourier Analysis

In the following, let $L(\Omega_h)$ be the set of the discrete functions $\Omega_h \to \mathbb{R}$. This set is a finite dimensional space isomorphic to $\mathbb{R}^n$, $n = (N-1)^2$.

We assume that these functions are extended to $\overline{\Omega}_h$ by setting their values to zero on Γ_h. Then, we can see the discrete Laplacian Δ_h as a linear operator $L(\Omega_h) \to L(\Omega_h)$.

The operator Δ_h is invertible. In fact, for any $f \in L(\Omega_h)$, the equation

$$\Delta_h v_h = f$$

has a unique solution $v_h \in L(\Omega_h)$ since it is equivalent to the discrete BVP

$$\begin{cases} \Delta_h v_h(x,y) = f(x,y), & (x,y) \in \Omega_h, \\ v_h(x,y) = 0, & (x,y) \in \Gamma_h, \end{cases} \tag{6}$$

which has a unique solution, as we have previously seen.

The relation between the convergence error $e_h \in L(\Omega_h)$ and the consistency error $\varepsilon_h \in L(\Omega_h)$ of the discretization of the BVP is

$$\begin{cases} \Delta_h e_h(x,y) = -\varepsilon_h(x,y), & (x,y) \in \Omega_h, \\ e_h(x,y) = 0, & (x,y) \in \Gamma_h \end{cases}$$

namely

$$\Delta_h e_h = -\varepsilon_h.$$

So, we have

$$e_h = -\Delta_h^{-1}\varepsilon_h$$

and then

$$\|e_h\|_{L^\infty(\Omega_h)} \leq \left\|\Delta_h^{-1}\right\| \|\varepsilon_h\|_{L^\infty(\Omega_h)}$$

where $\left\|\Delta_h^{-1}\right\|$ is the operator norm relevant to the L^∞ norm on $L(\Omega_h)$. By looking at (6), we have

$$\left\|\Delta_h^{-1}\right\| = \sup_{f \in L(\Omega_h),\ f \neq 0} \frac{\|v_h\|_{L^\infty(\Omega_h)}}{\|f\|_{L^\infty(\Omega_h)}}.$$

Our previous stability estimate

$$\|v_h\|_{L^\infty(\Omega_h)} = \|v_h\|_{L^\infty(\overline{\Omega}_h)} \leq \frac{1}{8}\|f\|_{L^\infty(\Omega_h)}$$

for the solution v_h shows that

$$\left\|\Delta_h^{-1}\right\| \leq \frac{1}{8}.$$

The stability of the discretization amounts to be able to bound $\left\|\Delta_h^{-1}\right\|$ with a constant independent of h.

The L^∞ norm on $L(\Omega_h)$ is the discrete analogous of the L^∞ norm on $L^\infty(\Omega)$.

Now, we introduce on $L(\Omega_h)$ another norm, which is the discrete analogous of the L^2 norm $\|\cdot\|$ on $L^2(\Omega)$ given by

$$\|v\| = \sqrt{\int_\Omega v(x,y)^2 \, d(x,y)},\ v \in L^2(\Omega),$$

$L^2(\Omega)$ being the set of the functions $\Omega \to \mathbb{R}$ such that $\|v\| < +\infty$.

We recall that the L^2 norm on $L^2(\Omega)$ is derived by the scalar product:

$$\langle v, w \rangle = \int_{\Omega} v(x,y)\, w(x,y) d(x,y), \ v, w \in L^2(\Omega).$$

Now, we define the L^2 norm on $L(\Omega_h)$. Then, by using a *Fourier Analysis*, namely an analysis based on the eigenvalues and the orthogonal eigenvectors of Δ_h, we will give a bound for the operator norm $\|\Delta_h^{-1}\|$ relevant to the L^2 norm on $L(\Omega_h)$.

But, first, we consider the 1D version of our problem.

3.1 The 1D problem

In the 1D version

$$\Omega = I := (0,1)$$

and

$$\Omega_h = I_h := \{h, 2h, \ldots, (N-1)h\}.$$

On $L(\Omega_h) = L(I_h)$, we define the scalar product

$$\langle v_h, w_h \rangle_h = h \sum_{k=1}^{N-1} v_h(kh)\, w_h(kh), \ v_h, w_h \in L(I_h),$$

with corresponding norm $\|\cdot\|_h$ given by

$$\|v_h\|_h = \sqrt{h \sum_{k=1}^{N-1} v_h(kh)^2}, \ v_h \in L(I_h),$$

called the L^2 norm on $L(I_h)$.

Exercise. Explain why the previous scalar product and norm are discretizations of the scalar product and norm on $L^2(I)$.

3.1.1 The continuous case

On the space $L^2(I)$, the functions ϕ_m, $m \in \{1, 2, 3, \ldots\}$, given by

$$\phi_m(x) = \sin(m\pi x), \ x \in I,$$

constitute an orthogonal basis of $L^2(I)$, i.e. we have

$$\langle \phi_m, \phi_n \rangle = \int_0^1 \phi_m(x)\, \phi_n(x)\, dx = 0, \ m, n \in \{1, 2, 3, \ldots\} \text{ with } m \neq n,$$

and, for any $v \in L^2(I)$, we have the *Fourier series*

$$v = \sum_{m=1}^{\infty} c_m \phi_m, \ c_m = \frac{\langle v, \phi_m \rangle}{\|\phi_m\|^2},$$

of v, where the convergence of the series is in the norm of $L^2(I)$, i.e.

$$\lim_{M \to \infty} \left\| v - \sum_{m=1}^{M} c_m \phi_m \right\| = 0.$$

Moreover, we have the *Parseval's identity*

$$\|v\|^2 = \sum_{m=1}^{\infty} c_m^2 \|\phi_m\|^2.$$

The function ϕ_m, $m \in \{1, 2, 3, \ldots\}$, is an eigenfunction of the operator

$$\Delta = D^2 = \text{second derivative}$$

relevant to the eigenvalue $\lambda_m := -\pi^2 m^2$.

In fact

$$\frac{d^2}{dx^2} \sin(\pi m x) = \pi m \frac{d}{dx} \cos(\pi m x) = -\pi^2 m^2 \sin(\pi m x), \ x \in I.$$

Now we establish the discrete analogous of these results.

3.1.2 The discrete case

In the space $L(I_h)$, we consider the functions $\phi_{m,h}$, $m \in \{1, 2, \ldots, N-1\}$, given by

$$\phi_{m,h}(x) = \phi_m(x) = \sin(m\pi x), \ x \in I_h.$$

Recall that the functions of $L(I_h)$ are extended to the boundary points 0 and 1 with value 0. So, $\phi_{m,h}(x) = \sin(m\pi x)$ also for $x = 0$ or $x = 1$.

The functions $\phi_{m,h}$, $m \in \{1, 2, \ldots, N-1\}$, are eigenvectors of the one-dimensional discrete Laplacian $\Delta_h : L(I_h) \to L(I_h)$ given by

$$\Delta_h v_h(x) = \frac{v_h(x-h) - 2v_h(x) + v_h(x+h)}{h^2}, \ x \in I_h \ \text{ and } \ v_h \in L(I_h).$$

In fact, for $x \in I_h$,

$$\begin{aligned}
&\Delta_h \phi_{m,h}(x) \\
&= \frac{\sin(\pi m (x-h)) - 2\sin(\pi m x) + \sin(\pi m (x+h))}{h^2} \\
&= \frac{2\sin(\pi m x)\cos(\pi m h) - 2\sin(\pi m x)}{h^2} \qquad \text{since } \sin\alpha + \sin\beta = 2\sin\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2} \\
&= -2\frac{1-\cos(\pi m h)}{h^2}\sin(\pi m x) \\
&= -2\frac{1-\cos(\pi m h)}{h^2}\phi_{m,h}(x).
\end{aligned}$$

The eigenvalue relevant to the eigenvector $\phi_{m,h}$ is

$$\lambda_{m,h} := -2\frac{1-\cos(\pi m h)}{h^2} = -4\frac{\sin^2\left(\frac{\pi m h}{2}\right)}{h^2} \qquad \text{since } 1-\cos\alpha = 2\sin^2\frac{\alpha}{2}.$$

Note that

$$0 > \lambda_{1,h} > \cdots > \lambda_{N-1,h}.$$

Exercise. Prove that

$$\begin{aligned} &\lim_{h\to 0}\lambda_{m,h} = \lambda_m \\ &\lambda_{N-1,h} > -\frac{4}{h^2} \quad \text{and} \quad \lambda_{N-1,h} = -\frac{4}{h^2}\left(1+O(h^2)\right) \\ &-8 \geq \lambda_{1,h}; \end{aligned}$$

for the last one, observe that $|\lambda_{1,h}|$ is a decreasing function of h.

The matrix form of the one-dimensional discrete Laplacian is

$$\Delta_h = \frac{1}{h^2}\begin{bmatrix} -2 & 1 & 0 & . & . & 0 \\ 1 & -2 & 1 & . & & . \\ 0 & 1 & . & . & . & . \\ . & . & . & . & . & 0 \\ . & & . & . & . & 1 \\ 0 & . & . & 0 & 1 & -2 \end{bmatrix} \in \mathbb{R}^{(N-1)\times(N-1)}.$$

Since the $N-1$ eigenvalues $\lambda_{m,h}$ of the matrix Δ_h are distinct, the $N-1$ eigenvectors $\phi_{m,h}$ are a basis of $\mathbb{R}^{N-1} = L(I_h)$. So, for any $v_h \in L(I_h)$, we have

$$v_h = \sum_{m=1}^{N-1} c_{m,h}\phi_{m,h}$$

for some unique coefficients $c_{m,h}$. This is the *discrete Fourier series* of v_h.

Moreover, since the matrix Δ_h is symmetric, the eigenvectors $\phi_{m,h}$ are orthogonal in the usual scalar product

$$(v_h, w_h) \mapsto \sum_{k=1}^{N-1} v_h(kh)\, w_h(kh),\ v_h, w_h \in L(I_h),$$

and then also in the scalar product

$$\langle v_h, w_h\rangle_h = h\sum_{k=1}^{N-1} v_h(kh)\, w_h(kh),\ v_h, w_h \in L(I_h).$$

Exercise. For $v_h \in L(I_h)$, show that

$$c_{m,h} = \frac{\langle v_h, \phi_{m,h}\rangle_h}{\|\phi_{m,h}\|_h^2},\ m \in \{1, 2, \ldots, N-1\},$$

and the *discrete Parseval's identity*

$$\|v_h\|_h^2 = \sum_{m=1}^{N-1} c_{m,h}^2 \|\phi_{m,h}\|_h^2 .$$

holds.

Now, we can obtain immediately a stability result for the one-dimensional discrete Laplacian.

Let $v_h, f_h \in L(I_h)$ such that

$$\Delta_h v_h = f_h$$

and let

$$v_h = \sum_{m=1}^{N-1} c_{m,h}\phi_{m,h}$$

be the discrete Fourier series of v_h.

Then

$$\begin{aligned} f_h &= \Delta_h v_h = \Delta_h \left(\sum_{m=1}^{N-1} c_{m,h}\phi_{m,h} \right) \\ &= \sum_{m=1}^{N-1} c_{m,h}\Delta_h \phi_{m,h} = \sum_{m=1}^{N-1} c_{m,h}\lambda_{m,h}\phi_{m,h} \end{aligned}$$

is the discrete Fourier series of f_h.

Thus, by the discrete Parseval's identity and $-8 \geq \lambda_{1,h} > \ldots > \lambda_{N-1,h}$, we have

$$\begin{aligned} \|f_h\|_h^2 &= \sum_{m=1}^{N-1} (c_{m,h}\lambda_{m,h})^2 \|\phi_{m,h}\|_h^2 = \sum_{m=1}^{N-1} \lambda_{m,h}^2 c_{m,h}^2 \|\phi_{m,h}\|_h^2 \\ &\geq \sum_{m=1}^{N-1} 8^2 c_{m,h}^2 \|\phi_{m,h}\|_h^2 = 8^2 \sum_{m=1}^{N-1} c_{m,h}^2 \|\phi_{m,h}\|_h^2 = 8^2 \|v_h\|_h^2 . \end{aligned}$$

We conclude that

$$\|v_h\|_h \leq \frac{1}{8} \|f_h\|_h .$$

This means

$$\left\|\Delta_h^{-1}\right\| \leq \frac{1}{8},$$

where the operator norm

$$\left\|\Delta_h^{-1}\right\| = \sup_{f_h \in L(I_h),\ f_h \neq 0} \frac{\|v_h\|_h}{\|f_h\|_h}$$

is relevant to the L^2 norm on $L(I_h)$.

3.2 The 2D problem

The extension to the 2D case, where

$$\Omega = I^2 = (0,1)^2$$

and

$$\Omega_h = I_h^2 = \{h, 2h, \ldots, (N-1)\,h\}^2,$$

is now straightforward.

On $L(\Omega_h)$, we define the scalar product

$$\langle v_h, w_h \rangle_h = h^2 \sum_{k=1}^{N-1} \sum_{l=1}^{N-1} v_h(kh, lh)\, w_h(kh, lh)\,,\ \ v_h, w_h \in L(\Omega_h)\,,$$

with corresponding norm

$$\|v_h\|_h = \sqrt{h^2 \sum_{k=1}^{N-1} \sum_{l=1}^{N-1} v_h(kh, lh)^2},\ \ v_h \in L(\Omega_h)\,.$$

Consider the functions $\phi_{m,n,h}$, $(m,n) \in \{1, 2, \ldots, N-1\}^2$, in $L(\Omega_h)$ given by

$$\phi_{m,n,h}(x,y) = \phi_{m,h}(x)\, \phi_{n,h}(y)\,,\ \ (x,y) \in \Omega_h,$$

where $\phi_{m,h}$ and $\phi_{n,h}$ have been defined in the one-dimensional case. Recall that the functions of $L(\Omega_h)$ are extended to the boundary points in Γ_h with value 0. So, $\phi_{m,n,h}(x,y) = \phi_{m,h}(x)\, \phi_{n,h}(y)$ also for $(x,y) \in \Gamma_h$.

These $(N-1)^2$ functions are orthogonal: for $(m,n), (p,q) \in \{1, 2, \ldots, N-1\}$, we have

$$\begin{aligned}
&\langle \phi_{m,n,h}, \phi_{p,q,h} \rangle_h \\
&= h^2 \sum_{k=1}^{N-1} \sum_{l=1}^{N-1} \phi_{m,h}(kh)\, \phi_{n,h}(lh)\, \phi_{p,h}(kh)\, \phi_{q,h}(lh) \\
&= h^2 \left(\sum_{k=1}^{N-1} \phi_{m,h}(kh)\, \phi_{p,h}(kh) \right) \left(\sum_{l=1}^{N-1} \phi_{n,h}(lh)\, \phi_{q,h}(lh) \right) \\
&= \langle \phi_{m,h}, \phi_{p,h} \rangle_h \langle \phi_{n,h}, \phi_{q,h} \rangle_h = 0
\end{aligned}$$

if $(m,n) \neq (p,q)$.

Hence, these functions constitute an orthogonal bases of $L(\Omega_h)$ and we have, for any $v_h \in L(\Omega_h)$, the discrete Fourier series

$$v_h = \sum_{m=1}^{N-1} \sum_{n=1}^{N-1} c_{m,n,h} \phi_{m,n,h},\ \ c_{m,n,h} = \frac{\langle v_h, \phi_{m,n,h} \rangle_h}{\|\phi_{m,n,h}\|_h^2},$$

and the discrete Parseval's identity

$$\|v_h\|_h^2 = \sum_{m=1}^{N-1} \sum_{n=1}^{N-1} c_{m,n,h}^2 \|\phi_{m,n,h}\|_h^2 .$$

The function $\phi_{m,n,h}$, $m, n \in \{1, 2, 3, \ldots, N-1\}$, is an eigenvector of the two-dimensional discrete Laplacian $\Delta_h : L(\Omega_h) \to L(\Omega_h)$:

$$\begin{aligned}\Delta_h v_h(x,y) = &\frac{v_h(x-h,y) - 2v_h(x,y) + v_h(x+h,y)}{h^2} \\ &+ \frac{v_h(x,y-h) - 2v_h(x,y) + v_h(x,y+h)}{h^2} \\ (x,y) \in \Omega_h, v_h \in L(\Omega_h),&\end{aligned}$$

whose relevant eigenvalue is $\lambda_{m,h} + \lambda_{n,h}$.

In fact, for $(x,y) \in \Omega_h$,

$$\begin{aligned}&\Delta_h \phi_{m,n,h}(x,y) \\ &= \frac{\phi_{m,n,h}(x-h,y) - 2\phi_{m,n,h}(x,y) + \phi_{m,n,h}(x+h,y)}{h^2} \\ &\quad + \frac{\phi_{m,n,h}(x,y-h) - 2\phi_{m,n,h}(x,y) + \phi_{m,n,h}(x,y+h)}{h^2} \\ &= \frac{\phi_{m,h}(x-h) - 2\phi_{m,h}(x) + \phi_{m,h}(x+h)}{h^2}\phi_{n,h}(y) \\ &\quad + \frac{\phi_{n,h}(y-h) - 2\phi_{n,h}(y) + \phi_{n,h}(y+h)}{h^2}\phi_{m,h}(x) \\ &= \lambda_{m,h}\phi_{m,h}(x)\phi_{n,h}(y) + \lambda_{n,h}\phi_{n,h}(y)\phi_{m,h}(x) \\ &= (\lambda_{m,h} + \lambda_{n,h})\phi_{m,n,h}(x,y).\end{aligned}$$

By proceeding as in the one-dimensional case, for $v_h, f_h \in L(\Omega_h)$ such that

$$\Delta_h v_h = f_h,$$

starting with the discrete Fourier series

$$v_h = \sum_{m=1}^{N-1} \sum_{n=1}^{N-1} c_{m,n,h}\phi_{m,n,h}$$

of v_h, we obtain the discrete Fourier series

$$f_h = \Delta_h v_h = \sum_{m=1}^{N-1} \sum_{n=1}^{N-1} c_{m,n,h}(\lambda_{m,h} + \lambda_{n,h})\phi_{m,n,h}$$

of f_h and then, by using two times the discrete Parseval's identity, we obtain

$$\begin{aligned}\|f_h\|_h^2 &= \sum_{m=1}^{N-1} \sum_{n=1}^{N-1} c_{m,n,h}^2 (\lambda_{m,h} + \lambda_{n,h})^2 \|\phi_{m,n,h}\|_h^2 \\ &\geq 16^2 \|v_h\|_h^2 .\end{aligned}$$

Thus

$$\left\|\Delta_h^{-1}\right\| \leq \frac{1}{16}.$$

follows, where the operator norm is relevant to the L^2 norm on $L(\Omega_h)$.

For the convergence error, we obtain

$$\|e_h\|_h \leq \|\Delta_h^{-1}\|\|\varepsilon_h\|_h \leq \frac{1}{16}\|\varepsilon_h\|_h$$

by using the L^2 norm. Observe that we have obtained

$$\|e_h\|_{L^\infty(\Omega_h)} \leq \|\Delta_h^{-1}\|\|\varepsilon_h\|_{L^\infty(\Omega_h)} \leq \frac{1}{8}\|\varepsilon_h\|_{L^\infty(\Omega_h)}.$$

for the L^∞ norm.

Exercise. Prove that, for $v_h \in L(\Omega_h)$, we have

$$\|v_h\|_h \leq \|v_h\|_{L^\infty(\Omega_h)}$$

and then conclude with an estimate

$$\|e_h\|_h = O(h^2), \ h \to 0,$$

for the convergence error in the discrete L^2 norm.

4 General domains

Up to now, we have studied, as a model problem, the discretization of the Poisson problem on the square by the five-point discretization of the Laplacian.

Now, we consider a variant of this discretization, which can be used for the Poisson problem on a general domain Ω of $\mathbb{R}^2$.

Let Ω be an open bounded connected subset of $\mathbb{R}^2$ with boundary Γ. We also assume that Γ is smooth or Γ is piecewise-smooth and Ω is convex. This assumption on the domain Ω guarantees existence and uniqueness for the solution of the Poisson problem.

Given $h > 0$, we set

$$\Omega_h := \Omega \cap \mathbb{R}_h^2,$$

where

$$\mathbb{R}_h^2 = \{(mh, nh) : m.n \in \mathbb{Z}\}$$

is the discrete plane.

Now, we define the neighbors of a point $(x, y) \in \Omega_h$.

If $(x + th, y) \in \Omega$ for any $t \in [0, 1]$, then $(x + h, y) \in \Omega_h$ is called the *right neighbor* of (x, y).

When it is not true that $(x + th, y) \in \Omega$ for any $t \in [0, 1]$, then we define as *right neighbor* of (x, y) the point $(x + t^*h, y)$, where $t^* \in (0, 1]$ satisfies:

- $(x + t^*h, y) \in \Gamma$;
- $(x + th, y) \in \Omega$ for any $t \in [0, t^*)$.

So $(x + t^*h, y) \notin \Omega_h$ and $(x + t^*h, y) \in \Gamma$.

In the same manner, we define the *left neighbor* $(x - h, y) \in \Omega_h$ or $(x - t^*h, y) \in \Gamma$, as well as the *upper and lower neighbors* $(x, y \pm h) \in \Omega_h$ or $(x, y \pm t^*h) \in \Gamma$.

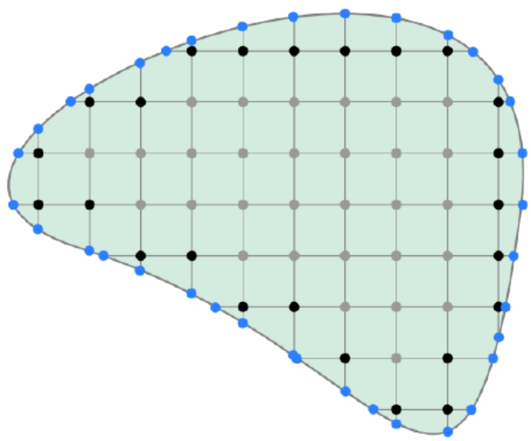
We define Γ_h the set of all neighbors of type $(x \pm t^*h, y)$ or $(x, y \pm t^*h)$. Note that $\Gamma_h \subseteq \Gamma$.

Moreover, we set

$$\overline{\Omega}_h := \Omega_h \cup \Gamma_h.$$

Finally, let $\mathring{\Omega}_h$ be the set of points $(x, y) \in \Omega_h$ whose four neighbors are in Ω_h: these neighbors are $(x \pm h, y)$ and $(x, y \pm h)$.

So, $\Omega_h \setminus \mathring{\Omega}_h$ is the set of points $(x, y) \in \Omega_h$ with some neighbor in Γ_h, i.e. of type $(x \pm t^*h, y)$ or $(x, y \pm t^*h)$.



In this figure:

- The dots are the points of $\overline{\Omega}_h$;
- The gray or black dots are the points of Ω_h.
- The blue dots are the points of Γ_h.
- The gray dots are the points of $\mathring{\Omega}_h$.
- The black dots are the points of $\Omega_h \setminus \mathring{\Omega}_h$.

Observe that, in case of the square $\Omega = (0, 1)^2$ with $h = \frac{1}{N}$ and N positive integer, the sets Ω_h, Γ_h and $\overline{\Omega}_h$ are the same as those previously introduced. Note that the points in Γ_h are of type $(x \pm t^*h, y)$ or $(x, y \pm t^*h)$ with $t^* = 1$. Moreover

$$\mathring{\Omega}_h = \{(mh, nh) : m, n \in \{2, \ldots, N - 2\}\}.$$

4.1 The Shortley-Weller formula

Now, we define a discrete Laplacian Δ_h in this general situation.

Let $v_h : \overline{\Omega}_h \to \mathbb{R}$ be a mesh function. Of course, for $(x,y) \in \mathring{\Omega}_h$, we can use the classic five-point discretization

$$\begin{aligned}\Delta_h v_h(x,y) &= \frac{v_h(x-h,y) - 2v_h(x,y) + v_h(x+h,y)}{h^2} \\ &\quad + \frac{v_h(x,y-h) - 2v_h(x,y) + v_h(x,y+h)}{h^2}.\end{aligned}$$

For $(x,y) \in \Omega_h \setminus \mathring{\Omega}_h$, we have to define $\Delta_h v_h(x,y)$ in terms of

$$v_h(x,y),\ v_h(x-h_1,y),\ v_h(x+h_2,y),\ v_h(x,y-h_3),\ v_h(x,y+h_4),$$

where $h_1, h_2, h_3, h_4 \in (0,h]$.

So, given a smooth function v of one real variable t and fixed stepsizes h_- and h_+, we look for an approximation of $v''(t)$ of the form

$$\alpha_- v(t-h_-) + \alpha_0 v(t) + \alpha_+ v(t+h_+).$$

If v is of class C^3, we have

$$\begin{aligned}\alpha_- v(t-h_-) &= \alpha_- v(t) - \alpha_- h_- v'(t) + \alpha_- \frac{h_-^2}{2} v''(t) - \alpha_- \frac{h_-^3}{6} v'''(\xi_h) \\ \alpha_+ v(t+h_+) &= \alpha_+ v(t) + \alpha_+ h_+ v'(t) + \alpha_+ \frac{h_+^2}{2} v''(t) + \alpha_+ \frac{h_+^3}{6} v'''(\eta_h),\end{aligned}$$

where $\xi_h \in (t-h_-, t)$ and $\eta_h \in (t, t+h_+)$. Then

$$\begin{aligned}&\alpha_- v(t-h_-) + \alpha_0 v(t) + \alpha_+ v(t+h_+) \\ &= (\alpha_- + \alpha_0 + \alpha_+) v(t) + (-\alpha_- h_- + \alpha_+ h_+) v'(t) \\ &\quad + \frac{1}{2}\left(\alpha_- h_-^2 + \alpha_+ h_+^2\right) v''(t) - \alpha_- \frac{h_-^3}{6} v'''(\xi_h) + \alpha_+ \frac{h_+^3}{6} v'''(\eta_h).\end{aligned}$$

Now, we require

$$\begin{cases} \alpha_- + \alpha_0 + \alpha_+ = 0 \\ -\alpha_- h_- + \alpha_+ h_+ = 0 \\ \frac{1}{2}\left(\alpha_- h_-^2 + \alpha_+ h_+^2\right) = 1. \end{cases}$$

By solving this system one finds

$$\begin{aligned}\alpha_- &= \frac{2}{h_-(h_- + h_+)} \\ \alpha_+ &= \frac{2}{h_+(h_- + h_+)} \\ \alpha_0 &= -\frac{2}{h_-(h_- + h_+)} - \frac{2}{h_+(h_- + h_+)} \\ &= -\frac{2h_+ + 2h_-}{h_- h_+ (h_- + h_+)} = -\frac{2}{h_- h_+}.\end{aligned}$$

Therefore, we have the discretization scheme for $v''(t)$ given by

$$\frac{2}{h_-(h_-+h_+)}v(t-h_-)-\frac{2}{h_-h_+}v(t)+\frac{2}{h_+(h_-+h_+)}v(t+h_+).$$

Observe that, when $h_- = h_+ = h$, we have the second central difference since we have

$$\frac{2}{h_-(h_-+h_+)}=\frac{1}{h^2},\quad -\frac{2}{h_-h_+}=-\frac{2}{h^2}\quad\text{and}\quad\frac{2}{h_+(h_-+h_+)}=\frac{1}{h^2}.$$

For the error we have:

$$\begin{aligned}
&|\alpha_-v(t-h_-)+\alpha_0v(t)+\alpha_+v(t+h_+)-v''(t)|\\
&\le \alpha_-\frac{h_-^3}{6}|v'''(\xi_h)|+\alpha_+\frac{h_+^3}{6}|v'''(\eta_h)|\\
&\le \frac{1}{6}\left(\alpha_-h_-h_-^2+\alpha_+h_+h_+^2\right)\max_{\xi\in[t-h_-,t+h_+]}|v'''(\xi)|\\
&= \frac{1}{6}\cdot\frac{2}{h_-+h_+}\left(h_-^2+h_+^2\right)\max_{\xi\in[t-h_-,t+h_+]}|v'''(\xi)|\\
&\qquad\text{since } \alpha_-h_-=\alpha_+h_+=\frac{2}{h_-+h_+}\\
&= \frac{1}{6}\cdot\frac{2}{h_{\min}+h_{\max}}\left(h_{\min}^2+h_{\max}^2\right)\max_{\xi\in[t-h_-,t+h_+]}|v'''(\xi)|\\
&\qquad\text{by introducing } h_{\min}=\min\{h_-,h_+\}\text{ and } h_{\max}=\max\{h_-,h_+\}\\
&= \frac{1}{3}h_{\max}\cdot\frac{1+\left(\frac{h_{\min}}{h_{\max}}\right)^2}{1+\frac{h_{\min}}{h_{\max}}}\max_{\xi\in[t-h_-,t+h_+]}|v'''(\xi)|\\
&\le \frac{1}{3}h_{\max}\max_{\xi\in[t-h_-,t+h_+]}|v'''(\xi)|.
\end{aligned}$$

So, we have the *Shortley-Weller formula*: for $v_h : \overline{\Omega}_h \to \mathbb{R}$ and $(x,y) \in \Omega_h \setminus \overset{\circ}{\Omega}_h$,

$$\begin{aligned}
&\Delta_h v_h(x,y)\\
&= \frac{2}{h_1(h_1+h_2)}v_h(x-h_1,y)-\frac{2}{h_1h_2}v_h(x,y)+\frac{2}{h_2(h_1+h_2)}v_h(x+h_2,y)\\
&+\frac{2}{h_3(h_3+h_4)}v_h(x,y-h_3)-\frac{2}{h_3h_4}v_h(x,y)+\frac{2}{h_4(h_3+h_4)}v_h(x,y+h_4).
\end{aligned}$$

For $h_1 = h_2 = h_3 = h_4$, we have the classic five-point discretization (used in $\overset{\circ}{\Omega}_h$).

For $v : \overline{\Omega} \to \mathbb{R}$ of class C^4, we have, for points in $\overset{\circ}{\Omega}_h$,

$$\left\|\Delta_h\left(v|_{\overline{\Omega}_h}\right)-(\Delta v)|_{\Omega_h}\right\|_{L^\infty\left(\overset{\circ}{\Omega}_h\right)}\le\frac{h^2}{6}\max\left\{\left\|\frac{\partial^4 v}{\partial x^4}\right\|_{L^\infty(\overline{\Omega})},\left\|\frac{\partial^4 v}{\partial y^4}\right\|_{L^\infty(\overline{\Omega})}\right\}.$$

On the other hand, for points in $\Omega_h \setminus \overset{\circ}{\Omega}_h$ the order is $O(h)$ only, in general. This is clarified by the following two exercises.

Exercise. Given $v : \overline{\Omega} \to \mathbb{R}$ of class C^4, find a function $C(x,y)$, $(x,y) \in \Omega_h \setminus \overset{\circ}{\Omega}_h$, and a nonnegative constant M such that

$$\max_{(x,y)\in\Omega_h\setminus\overset{\circ}{\Omega}_h} \left|\Delta_h v|_{\overline{\Omega}_h}(x,y) - \Delta v(x,y) - C(x,y)\right| \leq Mh^2.$$

So, we have

$$\Delta_h v|_{\overline{\Omega}_h}(x,y) - \Delta v(x,y) = C(x,y) + E(x,y),\ (x,y) \in \Omega_h \setminus \overset{\circ}{\Omega}_h,$$

with

$$\max_{(x,y)\in\Omega_h\setminus\overset{\circ}{\Omega}_h} |E(x,y)| \leq Mh^2.$$

Exercise. Prove that, for $v : \overline{\Omega} \to \mathbb{R}$ of class C^3, we obtain

$$\left\|\Delta_h\left(v|_{\overline{\Omega}_h}\right) - (\Delta v)|_{\Omega_h}\right\|_{L^\infty(\Omega_h)} \leq \frac{2}{3}h\max\left\{\left\|\frac{\partial^3 v}{\partial x^3}\right\|_{L^\infty(\overline{\Omega})}, \left\|\frac{\partial^3 v}{\partial y^3}\right\|_{L^\infty(\overline{\Omega})}\right\}.$$

4.2 The discrete problem

The discrete problem is

$$\begin{cases} \Delta_h u_h(x,y) = f(x,y),\ (x,y) \in \Omega_h \\ u_h(x,y) = g(x,y),\ (x,y) \in \Gamma_h. \end{cases}$$

We obtain a square system of linear equations: we have as unknowns the values of u_h at the points of Ω_h and there is an equation for each point in Ω_h.

In general, differently from the case of the square domain $\Omega = (0,1)^2$, the matrix is not symmetric. However, as in the case of the square domain, it is sparse, with at most five nonzero elements per row, it has negative diagonal elements and positive off-diagonal elements and it is diagonal dominant, i.e. the modulus of each diagonal element is not smaller than the sum of the moduli of the elements in the same row.

Exercise. Prove that the matrix is diagonally dominant and explain why in general it is not symmetric.

Discrete maximum and minimum principles can be obtained also for a general domain.

Exercise. Prove the discrete maximum principle for a general domain by following the steps of the proof given for the square domain.

Then existence and uniqueness for the discrete problem follows.

4.3 Convergence Analysis

Since discrete maximum and minimum principles hold, the same stability estimate obtained for the square domain is also valid for general domains. The proof of this estimate is the same as for the square domain.

In this way, we show that

$$\|e_h\|_{L^\infty(\overline{\Omega}_h)} = O\left(h\right), \ h \to 0,$$

holds for the convergence error $e_h = u_h - u|_{\overline{\Omega}_h}$ whenever u is of class C^3, where u and u_h are the solutions of the continuous and discrete Poisson problems, respectively. This is a consequence of the fact that, for the consistency error $\varepsilon_h = \Delta_h \left(u|_{\overline{\Omega}_h}\right) - \left(\Delta u\right)|_{\Omega_h}$, we have

$$\|\varepsilon_h\|_{L^\infty(\overline{\Omega}_h)} = O\left(h\right), \ h \to 0.$$

With respect to the case of the square domain, we have only order one with respect to h. But this result can be improved by a more refined argument: we have order two with respect to h, as in the case of the square domain.

Theorem 6 *If u is of class C^4, then*

$$\|e_h\|_{L^\infty(\overline{\Omega}_h)} = O\left(h^2\right), \ h \to 0.$$

Proof. We have

$$\begin{cases} \Delta_h e_h\left(x,y\right) = -\varepsilon_h(x,y), \ (x,y) \in \Omega_h \\ e_h\left(x,y\right) = 0, \ (x,y) \in \Gamma_h \end{cases}$$

where the consistency error ε_h satisfy

$$\|\varepsilon_h\|_{L^\infty\left(\mathring{\Omega}_h\right)} = O\left(h^2\right) \text{ and } \|\varepsilon_h\|_{L^\infty\left(\Omega_h \setminus \mathring{\Omega}_h\right)} = O(h), \ h \to 0.$$

Now, for the discrete problem

$$\begin{cases} \Delta_h u_h\left(x,y\right) = f\left(x,y\right), \ (x,y) \in \Omega_h, \\ u_h\left(x,y\right) = g\left(x,y\right), \ (x,y) \in \Gamma_h, \end{cases}$$

we prove the stability estimate

$$\|u_h\|_{L^\infty(\overline{\Omega}_h)} \le c\,|||f||| + \|g\|_{L^\infty(\Gamma_h)},$$

where c is a constant independent of h and

$$|||f||| := \max\left\{\|f\|_{L^\infty\left(\mathring{\Omega}_h\right)}, h^2\|f\|_{L^\infty\left(\Omega_h \setminus \mathring{\Omega}_h\right)}\right\}.$$

By this stability estimate, the thesis of the theorem immediately follows: we have

$$\|\varepsilon_h\|_{L^\infty\left(\overset{\circ}{\Omega}_h\right)} = O\left(h^2\right) \text{ and } h^2 \|\varepsilon_h\|_{L^\infty\left(\Omega_h \setminus \overset{\circ}{\Omega}_h\right)} = O(h^3), \ h \to 0.$$

Let (p, q) be the center of a circle of radius r including Ω. Let ϕ be the mesh function

$$\phi(x, y) = \begin{cases} \frac{(x-p)^2+(y-q)^2}{4} \text{ if } (x, y) \in \Omega_h \\ \frac{(x-p)^2+(y-q)^2}{4} + 1 \text{ if } (x, y) \in \Gamma_h. \end{cases}$$

Exercise. Prove that

$$\Delta_h \phi(x, y) = 1, \ (x, y) \in \overset{\circ}{\Omega}_h,$$

and

$$\Delta_h \phi(x, y) \geq h^{-2}, \ (x, y) \in \Omega_h \setminus \overset{\circ}{\Omega}_h$$

and

$$0 \leq \phi(x, y) \leq \frac{r^2}{4} + 1 =: c$$

Let

$$A = |||f||| \text{ and } B = \|g\|_{L^\infty(\Gamma_h)}.$$

We have

$$\Delta_h (u_h + A\phi)(x, y) \geq 0 \text{ and } \Delta_h (u_h - A\phi)(x, y) \leq 0, \ (x, y) \in \Omega_h.$$

In fact, for $(x, y) \in \overset{\circ}{\Omega}_h$, we have

$$\begin{aligned} \Delta_h (u_h + A\phi)(x, y) &= \Delta_h u_h(x, y) + A\Delta_h \phi(x, y) \\ &= f(x, y) + A \\ &\geq -|f(x, y)| + \|f\|_{L^\infty(\Omega_h)} \\ &\geq 0 \end{aligned}$$

and

$$\begin{aligned} \Delta_h (u_h - A\phi)(x, y) &= \Delta_h u_h(x, y) - A\Delta_h \phi(x, y) \\ &= f(x, y) - A \\ &\leq |f(x, y)| - \|f\|_{L^\infty(\Omega_h)} \\ &\leq 0. \end{aligned}$$

Moreover, for $(x, y) \in \Omega_h \setminus \overset{\circ}{\Omega}_h$, we have

$$\begin{aligned} \Delta_h (u_h + A\phi)(x, y) &= \Delta_h u_h(x, y) + A\Delta_h \phi(x, y) \\ &\geq f(x, y) + Ah^{-2} \\ &\geq -|f(x, y)| + \|f\|_{L^\infty\left(\Omega_h \setminus \overset{\circ}{\Omega}_h\right)} \\ &\geq 0 \end{aligned}$$

and

$$\begin{aligned}\Delta_h (u_h - A\phi)(x,y) &= \Delta_h u_h (x,y) - A\Delta_h \phi (x,y) \\ &\leq f(x,y) - Ah^{-2} \\ &\leq |f(x,y)| - \|f\|_{L^\infty\left(\Omega_h \setminus \mathring{\Omega}_h\right)} \\ &\leq 0.\end{aligned}$$

By the discrete maximum principle, we have

$$\begin{aligned}\max_{(x,y)\in\Omega_h} u_h(x,y) &\leq \max_{(x,y)\in\Omega_h} (u_h + A\phi)(x,y) \leq \max_{(x,y)\in\Gamma_h} (u_h + A\phi)(x,y) \\ &= \max_{(x,y)\in\Gamma_h} (g(x,y) + A\phi(x,y)) \leq B + cA.\end{aligned}$$

Moreover, by the discrete minimum principle, we have

$$\begin{aligned}\min_{(x,y)\in\Omega_h} u_h(x,y) &\geq \min_{(x,y)\in\Omega_h} (u_h - A\phi)(x,y) \geq \min_{(x,y)\in\Gamma_h} (u_h - A\phi)(x,y) \\ &= \min_{(x,y)\in\Gamma_h} (g(x,y) - A\phi(x,y)) \geq -B - cA.\end{aligned}$$

Thus

$$-B - cA \leq u_h(x,y) \leq B + cA, \ (x,y) \in \overline{\Omega}_h.$$

■

By looking to the proof of the previous theorem, we see that the points in $\Omega_h \setminus \mathring{\Omega}_h$ contribute to the convergence error only with a term $O\left(h^3\right)$, despite the fact that the consistency error is $O(h)$ in $\Omega_h \setminus \mathring{\Omega}_h$.