Lecture Notes: African Institute of Mathematics–Senegal, January 2016

Topic Title: A short introduction to numerical methods for elliptic PDEs

Authors and Lecturers: Gerard Awanou (University of Illinois-Chicago) and Johnny Guzmán (Brown University)

1. Description of Course

In this course we will mostly focus on second-order elliptic Partial Differential Equations (PDEs). We start by developing finite difference methods (a five point stencil method) for Poisson problems in one and two dimensions. Then, we develop finite element methods for general second order problems. Finally, we develop finite element methods for Stokes problems.

Students are expected to write computer code to implement the numerical methods. Also, theoretical aspects such as stability and error analysis will be covered and students are expected to be able to follow all the main arguments.

CHAPTER 1

Two Point Boundary Value Problem: finite difference and finite element methods

Before we jump into higher dimensional problems it is instructive to consider a two point boundary value problem

1. Two Point Boundary Value problem

$$-u''(x) = f(x) \quad \text{for all } 0 < x < 1, \tag{1.1a}$$

$$u(0) = a, \tag{1.1b}$$

$$u(1) = b, \tag{1.1c}$$

Here the function $f(x)$ is given. This is commonly known as the *source term*. The data a, b is also normally given an this is known as the Dirichlet boundary data. The unknown function is u. For now we are assuming that u belongs to $C^2([0,1]) = \{u : u, u', \text{ and } u'' \text{ are continuous on } [0,1]\}$.

We would like to prove a crucial property of these type of equations *weak maximum principle* for this problem. Before we do that let us recall basic properties of calculus.

PROPOSITION 1. *Suppose that $w \in C^2((0,1))$ an suppose that w has a local maximum at $0 < z < 1$ then*

$$w'(z) = 0 \text{ and } w''(z) \leq 0 \tag{1.2}$$

Similarly if w has a local minimum at z then

$$w'(z) = 0 \text{ and } w''(z) \geq 0. \tag{1.3}$$

Theorem 1. *Suppose u solves (1.1) and $f(x) \leq 0$ for all $x \in (0,1)$ then*

$$\max_{x\in[0,1]} u(x) \leq \max\{a, b\}. \tag{1.4}$$

On the other hand, if $f(x) \geq 0$ for all $x \in (0,1)$ then

$$\min_{x\in[0,1]} u(x) \geq \min\{a, b\}. \tag{1.5}$$

PROOF. We only prove (1.4) and leave the proof of (1.5) to the reader. We start by defining the function $\phi(x) = \frac{(x-1/2)^2}{2}$ and we let $w_\epsilon(x) = u(x) + \epsilon\phi(x)$ where $\epsilon > 0$. Noting that $\phi''(x) = 1$ for all x. Then we see that w_ϵ satisfies

$$-w_\epsilon''(x) = f(x) - \epsilon \quad \text{for all } 0 < x < 1, \tag{1.6a}$$

$$w_\epsilon(0) = a + \frac{\epsilon}{8}, \tag{1.6b}$$

$$w_\epsilon(1) = b + \frac{\epsilon}{8}, \tag{1.6c}$$

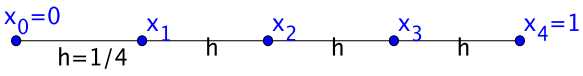

FIGURE 1. Example: $N = 3$, $h = 1/4$

Since $f(x) \leq 0$ for all $0 < x < 1$ we have that $w_\epsilon''(x) = -(f(x) - \epsilon) > 0$. We then see that w_ϵ cannot have a maximum at any point $0 < x < 1$. Indeed, if w_ϵ had a maximum at x then by (1.2) we have $w_\epsilon''(x) \leq 0$ contradicting that $w_\epsilon''(x) > 0$. Hence, the maximum must occur at $x = 0$ or $x = 1$. That is,

$$u(x) \leq u(x) + \epsilon\phi(x) = w_\epsilon(x) \leq \max\{w_\epsilon(0), w_\epsilon(1)\} = \max\{a + \epsilon/8, b + \epsilon/8\} \text{ for all } 0 \leq x \leq 1.$$

Therefore,

$$u(x) \leq \max\{a, b\} + \epsilon/8 \text{ for all } 0 \leq x \leq 1.$$

Now taking the limit at $\epsilon \to 0$ we have

$$u(x) \leq \max\{a, b\} \text{ for all } 0 \leq x \leq 1.$$

Taking the maximum we get (1.4). □

1.1. Finite Difference method. Now we develop a computational method to approximate (1.1). To do that, we let N be an integer and define $h = \frac{1}{N+1}$ and define $x_n = n\,h$ for $n = 0, 1, 2, \ldots N + 1$ (see Figure 1). We will approximate $u(x_n)$ for $n = 0, 1, \ldots, N + 1$. Indeed, the finite difference method will find numbers $v_0, v_1, \ldots, v_{N+1}$ such that $v_n \approx u(x_n)$ for $n = 0, 1, \ldots, N + 1$.

In order to derive the method, we use Taylor expansions on the solution u

$$\begin{aligned} u''(x_i) &\approx (u'(x_{i+1}) - u'(x_i))/h \\ &\approx (\frac{u(x_{i+1}) - u(x_i))}{h} - \frac{u(x_i) - u(x_{i-1})}{h})/h \\ &= (u(x_{i+1}) - 2u(x_i) + u(x_{i-1}))/h^2. \end{aligned}$$

In fact, we can show the following lemma (we leave the details to the reader).

LEMMA 1. *It holds*

$$u''(x_i) = (u(x_{i+1}) - 2u(x_i) + u(x_{i-1}))/h^2 - \frac{h^2}{24}(u^{(4)}(\theta_1) + u^{(4)}(\theta_2)) \tag{1.7}$$

where $x_{i-1} \le \theta_1 \le x_i$, $x_i \le \theta_2 \le x_{i+1}$.

Therefore, we define

Now, the finite difference method finds $v_0, v_1, \ldots, v_{N+1}$ such that:

$$(-v_{n+1} + 2v_n - v_{n-1})/h^2 = f(x_n) \quad \text{for all } n = 1, 2, \ldots, N, \tag{1.8a}$$

$$v_0 = a, \tag{1.8b}$$

$$v_{N+1} = b. \tag{1.8c}$$

This will give rise to a system of N equations with N unknowns.

$$\begin{aligned} \frac{1}{h^2}(2v_1 - v_2) &= f(x_1) + a\frac{1}{h^2} \\ \frac{1}{h^2}(-v_1 + 2v_2 - v_3) &= f(x_2) \\ &\vdots \\ \frac{1}{h^2}(-v_{N-1} + 2v_N) &= f(x_N) + b\frac{1}{h^2} \end{aligned}$$

In matrix form we can write

$$\frac{1}{h^2}\begin{bmatrix} 2 & -1 & 0 & \ldots & 0 \\ -1 & 2 & -1 & & \\ 0 & \ddots & \ddots & \ddots & \\ \vdots & & & & -1 \\ 0 & \ldots & & -1 & 2 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_N \end{bmatrix} = \begin{bmatrix} f(x_1) \\ f(x_2) \\ f(x_3) \\ \vdots \\ f(x_N) \end{bmatrix} + \frac{1}{h^2}\begin{bmatrix} a \\ 0 \\ \vdots \\ 0 \\ b \end{bmatrix}$$

You can easily write a computer code to implement this method.

Next, we will analyze the finite difference method described above. To do this we use some notation. We let $S = \{x_0, x_1, \ldots, x_{N+1}\}$ and we define all *discrete functions* as

$$P_h = \{v : v \text{ is a real valued function with domain } S\}.$$

That is, a function in P_h only has to be defined on the discrete points S. We use the notation $v_n = v(x_n)$.

We also denote the i-th derivative of a function as $D^i u(x) = u^{(i)}(x)$. In particular, we denote the second derivative $D^2 u(x) = u''(x)$. We then define the *discrete* second derivative as

$$D_h^2 v(x_n) \equiv \frac{1}{h^2}(v(x_{n+1}) - 2v(x_n) + v(x_{n-1})) \text{ for all } n = 1, \ldots, N.$$

We also use the notation $D_h^2 v_n \equiv D_h^2 v(x_n)$. We can then rewrite (1.8) with the new notation. Find $v \in P_h$ satisfying

$$-D_h^2 v_n = f(x_n) \quad \text{for all } n = 1, 2, \ldots, N, \tag{1.9a}$$

$$v_0 = a, \tag{1.9b}$$

$$v_{N+1} = b. \tag{1.9c}$$

1.1.1. *Discrete Maximum Principle.* We now state and prove the discrete analogue to Theorem 1.

Theorem 2. *Suppose v solves (1.9) and $f(x_n) \leq 0$ for all $n = 1, 2, \ldots, N$ then*

$$\max_{n=0,1,\ldots,N+1} v_n \leq \max\{a, b\}. \tag{1.10}$$

On the other hand, if $f(x_n) \geq 0$ for all $n = 1, 2, \ldots, N$ then

$$\min_{n=0,1,\ldots,N+1} v_n \geq \min\{a, b\}. \tag{1.11}$$

Proof. We will prove (1.10) and leave the proof (1.11) to the reader. Let $M = \max_{n=0,1,\ldots,N+1} v_n$. First suppose that $\max_{n=1,\ldots,N} v_n < M$ then (1.10) holds.

On the other hand, suppose there that there exists an $1 \leq n \leq N$ such that $v_n = M$. Then from (1.9a) we have

$$M = v_n = \frac{1}{2}(v_{n+1} + v_{n-1}) + h^2 f(x_n) \leq \frac{1}{2}(v_{n+1} + v_{n-1}) \leq M$$

since$f(x_n) \leq 0$ and by our hypothesis $\frac{1}{2}(v_{n+1} + v_{n-1}) \leq M$. Hence, we must have

$$M = \frac{1}{2}(v_{n+1} + v_{n-1}),$$

or that

$$\frac{1}{2}(M - v_{n-1}) + \frac{1}{2}(M - v_{n+1}) = 0.$$

Since by our hypothesis $(M - v_{n-1}) \geq 0$ and $(M - v_{n+1}) \geq 0$ it must be $v_{n-1} = v_n = v_{n+1} = M$. We can continue this process to show that $v_i = M = \max_{n=0,1,\ldots,N+1} v_n$ for all $i = 0, 1, \ldots, N+1$. In other words, $v \equiv M$ is a constant discrete function. Therefore, (1.10) trivially holds. □

We can use the discrete maximum principle to prove a stability result.

Theorem 3. *Let $v \in P_h$ solve (1.9) then we have*

$$\max_{n=0,1,\ldots,N+1} |v_n| \leq \max\{|a|, |b|\} + \frac{1}{8} \max_{n=1,\ldots,N} |f(x_n)|. \tag{1.12}$$

Proof. Let $Q = \max_{n=1,\ldots,N} |f(x_n)|$ and define $\phi \in P_h$ as follows

$$\phi_n = \phi(x_n) \equiv \frac{Q}{2}(x_n - 1/2)^2 \quad \text{for all } n = 0, 1, 2, \ldots, N+1.$$

Not difficult to see that $D_h^2 \phi_n = Q$ for $n = 1, \ldots, N$. Define $w \in P_h$ as $w = v + \phi$. Then we see that w satisfies

$$-D_h^2 w_n = f(x_n) - Q \quad \text{for all } n = 1, 2, \ldots, N, \tag{1.13a}$$

$$v_0 = a + \frac{Q}{8}, \tag{1.13b}$$

$$v_{N+1} = b + \frac{Q}{8}. \tag{1.13c}$$

Now, note that $f(x_n) - Q \leq 0$ for all n and hence by (1.10) we have

$$\max_{n=0,1,\ldots,N+1} w_n \leq \max\{a + \frac{Q}{8}, b + \frac{Q}{8}\}.$$

Since $\phi_n \geq 0$ we have,

$$\max_{n=0,1,\ldots,N+1} v_n \leq \max_{n=0,1,\ldots,N+1} w_n$$

Moreover, we have

$$\max\{a + \frac{Q}{8}, b + \frac{Q}{8}\} \leq \max\{|a|, |b|\} + \frac{Q}{8}.$$

Hence, combining the last two inequalities we get □

$$\max_{n=0,1,\ldots,N+1} v_n \leq \max\{|a|, |b|\} + \frac{Q}{8}. \tag{1.14}$$

Noting that

$$\begin{aligned} -D_h^2(-v_n) &= -f(x_n) \quad \text{for all } n = 1, 2, \ldots, N, \\ -v_0 &= -a, \\ -v_{N+1} &= -b. \end{aligned}$$

we can apply the previous argument verbatim to get

$$\max_{n=0,1,\ldots,N+1} (-v_n) \leq \max\{|a|, |b|\} + \frac{Q}{8}. \tag{1.15}$$

Combining inequality (1.14) and (1.15) we get (1.12).

1.1.2. *Error Estimate for Finite Difference Method.* Now that we have established the stability result (1.12), we can prove an error estimate for the finite difference method. Let us define the max norm for functions in $w \in C([0,1])$ as

$$\|w\|_{C([0,1])} = \max_{0 \leq x \leq 1} |w(x)|.$$

Theorem 4. *Suppose that u solves* (1.1) *and $u \in C^4([0,1])$. If v solves* (1.9) *we have*

$$\max_{n=0,1,\ldots,N+1} |u(x_n) - v_n| \leq \frac{h^2}{104} \|D^4 u\|_{C([0,1])}.$$

Proof. Define $w \in P_h$ as $w_n = u(x_n) - v_n$ for $n = 0, 1, \ldots, N+1$. Then, we see that

$$\begin{aligned} -D_h^2(w_n) &= \tau_n \quad \text{for all } n = 1, 2, \ldots, N, \\ w_0 &= 0, \\ w_{N+1} &= 0. \end{aligned}$$

where $\tau_n = -D_h^2 u(x_n) - f(x_n) = D^2 u(x_n) - D_h^2 u(x_n)$. By (1.7) we have

$$|\tau_n| \leq \frac{h^2}{12} \|D^4 u\|_{C([0,1])} \quad \text{for all } n = 1, 2, \ldots, N. \tag{1.16}$$

Now applying (1.12) we have

$$\max_{n=0,1,\ldots,N+1} |w_n| \leq \frac{1}{8} \max_{n=1,\ldots,N} |\tau_n|.$$

Combining this with (1.16) proves the result. □

1.2. Finite element method. In this section we derive the finite element method for (1.1). The methodology is quite different from the finite difference method but at the end they give similar methods.

We first need to define the weak formulation of (1.1). Let us recall an integration formula

$$\int_0^1 v'(x)w(x)dx = -\int_0^1 v(x)w'(x)dx + v(1)w(1) - v(0)w(0). \tag{1.17}$$

Now let $v \in C^1([0,1])$ with $v(0) = 0$ and $v(1) = 0$ then if we multiply (1.1a) by v and integrate we get

$$-\int_0^1 u''(x)v(x)dx = \int_0^1 f(x)v(x)dx$$

If we apply (1.17) we get

$$-\int_0^1 u''(x)v(x)dx = \int_0^1 u'(x)v'(x)dx,$$

where we used that $v(0) = 0 = v(1)$. Hence, we have that if u solves (1.1) then it solves

$$\int_0^1 u'(x)v'(x)dx = \int_0^1 f(x)v(x)dx \quad \text{for all } v \in C^1([0,1]), v(0) = 0 = v(1). \tag{1.18}$$

The finite element method is based on (1.18).

The next step is to build a finite dimensional space of functions.

We let

$$V_h = \{v \in C([0,1]) : \text{ for all } i = 0, 1, \ldots, N, v|_{(x_i, x_{i+1})} \text{ is a linear function}\}.$$

We can define a basis easily for this space in the following way. For each $i = 0, 1, 2, \ldots$ we define ψ_i (see Figures 1 and 4)

$$\psi_i(x_j) = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{if } j \neq i \end{cases}$$

Then we see that if $v \in V_h$ then we have

$$v(x) = \sum_{i=0}^{N+1} v(x_i)\psi_i(x) \text{ for } 0 \leq x \leq 1.$$

We see that the dimension of V_h is exactly $N + 1$. We also need to define

$$V_h^0 = \{v \in V_h : v(0) = 0 = v(1)\}.$$

The finite element method is as follows:

Find $u_h \in V_h$ such that

$$\int_0^1 u_h'(x)v_h'(x)dx = \int_0^1 f(x)v_h(x)dx \quad \text{for all } v_h \in V_h^0, \tag{1.19}$$

where $u_h(0) = u_h(x_0) = a$ and $u_h(1) = u_h(x_{N+1}) = b$.

Of course, since the space V_h^0 is finite dimensional then (1.21) is equivalent to:

$$\int_0^1 u_h'(x)\psi_j'(x)dx = \int_0^1 f(x)\psi_j(x)dx \quad \text{for all } j = 1, 2, \ldots, N, \tag{1.20}$$

and $u_h(0) = u_h(x_0) = a$ and $u_h(1) = u_h(x_{N+1}) = b$. Writing

$$u_h'(x) = \sum_{i=0}^{N+1} u_h(x_i)\psi_i'(x) \text{ for } 0 \leq x \leq 1.$$

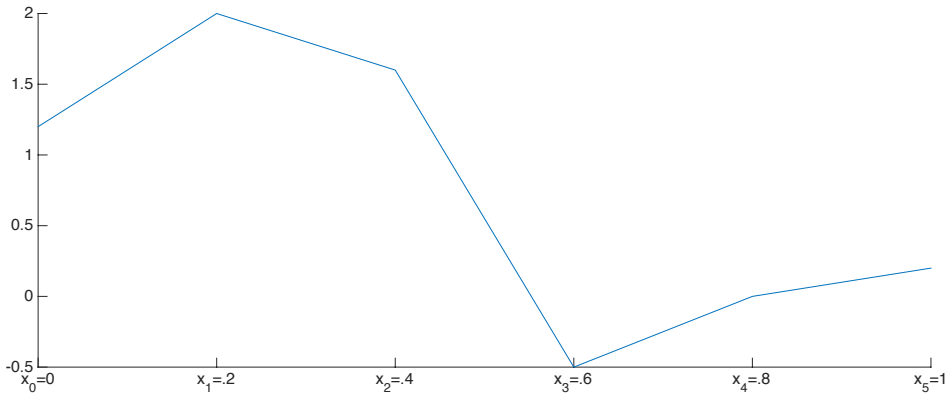

FIGURE 2. Example of a function in V_h, h=.2

we get

$$\int_0^1 u_h'(x)\psi_j'(x)dx = \int_0^1 \sum_{i=0}^{N+1} u_h(x_i)\psi_i'(x)\psi_j'(x)dx = \sum_{i=0}^{N+1} u_h(x_i)\int_0^1 \psi_i'(x)\psi_j'(x)dx. \tag{1.21}$$

Therefore, we have

$$\sum_{i=1}^{N} u_h(x_i)\int_0^1 \psi_i'(x)\psi_j'(x)dx = \int_0^1 f(x)\psi_j(x)dx$$
$$- a\int_0^1 \psi_0'(x)\psi_j'(x) - b\int_0^1 \psi_{N+1}'(x)\psi_j'(x) \quad \text{for all } j = 1, 2, \ldots, N,$$

To write this method in matrix form we define

$$A = \begin{bmatrix} \int_0^1 \psi_1'(x)\psi_1'(x)dx & \int_0^1 \psi_1'(x)\psi_2'(x)dx & \ldots & \int_0^1 \psi_1'(x)\psi_N'(x)dx \\ \int_0^1 \psi_2'(x)\psi_1'(x)dx & \int_0^1 \psi_2'(x)\psi_2'(x)dx & \ldots & \\ \vdots & & \ddots & \\ \int_0^1 \psi_N'(x)\psi_1'(x)dx & \ldots & & \int_0^1 \psi_N'(x)\psi_N'(x)dx \end{bmatrix}$$

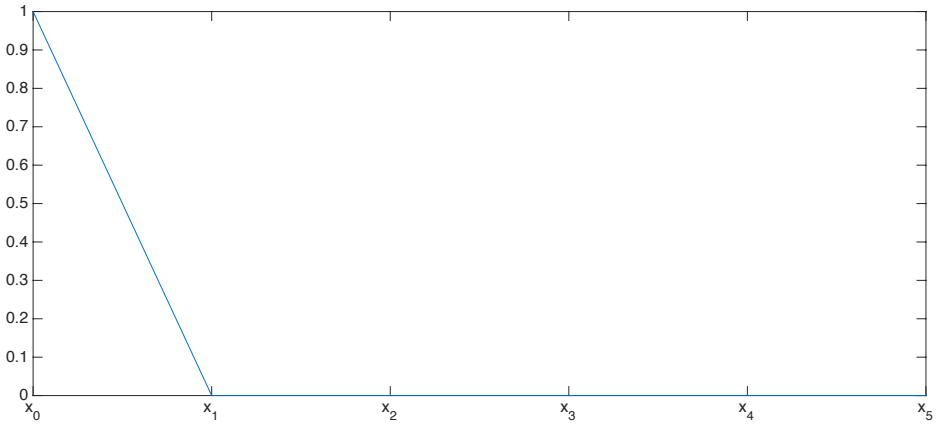

FIGURE 3. Graph of ψ_0

$$U = \begin{bmatrix} u_h(x_1) \\ u_h(x_2) \\ \vdots \\ u_h(x_N) \end{bmatrix}$$

$$F = \begin{bmatrix} \int_0^1 f(x)\psi_1(x)dx \\ \int_0^1 f(x)\psi_2(x)dx \\ \vdots \\ \int_0^1 f(x)\psi_N(x)dx \end{bmatrix}$$

$$G = -a \begin{bmatrix} \int_0^1 \psi_0'(x)\psi_1'(x)dx \\ \int_0^1 \psi_0'(x)\psi_2'(x)dx \\ \vdots \\ \int_0^1 \psi_0'(x)\psi_N'(x)dx \end{bmatrix} - b \begin{bmatrix} \int_0^1 \psi_{N+1}'(x)\psi_1'(x)dx \\ \int_0^1 \psi_{N+1}'(x)\psi_2'(x)dx \\ \vdots \\ \int_0^1 \psi_{N+1}'(x)\psi_N'(x)dx \end{bmatrix}$$

Then, U solves

$$AU = F + G.$$

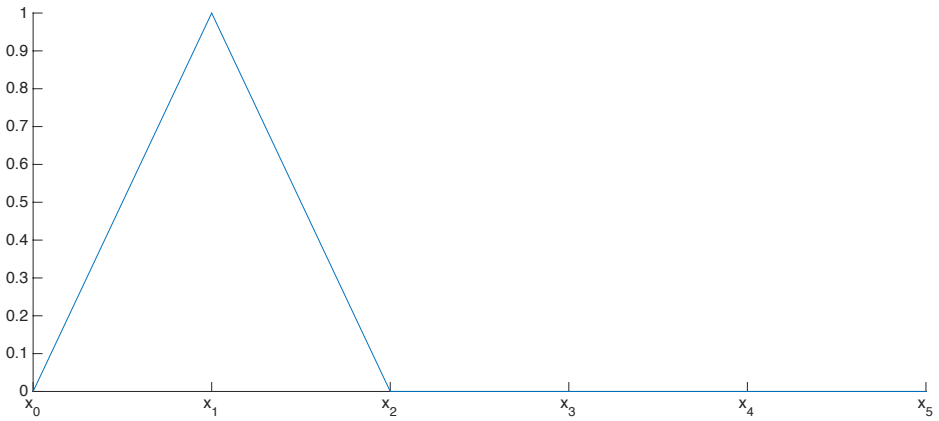

FIGURE 4. Graph of ψ_1

In order to be able to implement the method we need to find need to be able to evaluate the entries of the matrix A and the vectors F and G. Lets start with A. An easy calculation shows that (see Figure 5

$$\psi_i'(x) = \begin{cases} 0 & x \in [0, x_{i-1}) \\ 1/h & x \in [x_{i-1}, x_i) \\ -1/h & x \in [x_i, x_{i+1}) \\ 0 & x \in (x_{i+1}, 1]. \end{cases}$$

Since

$$A_{ij} = \int_0^1 \psi_i'(x)\psi_j'(x)dx.$$

First note that since the support of ψ_i' is in (x_{i-1}, x_{i+1}) we see that

$$A_{ij} = 0 \text{ when } j = 1, \ldots, i-2 \text{ and } j = i+2, \ldots, N.$$

We thus only have to consider the case $j = i-1, i, i+1$. Let us start the $j = i$, In this case

$$A_{ii} = \int_{x_{i-1}}^{x_i} (1/h)^2 dx + \int_{x_i}^{x_{i+1}} (-1/h)^2 dx = 2/h.$$

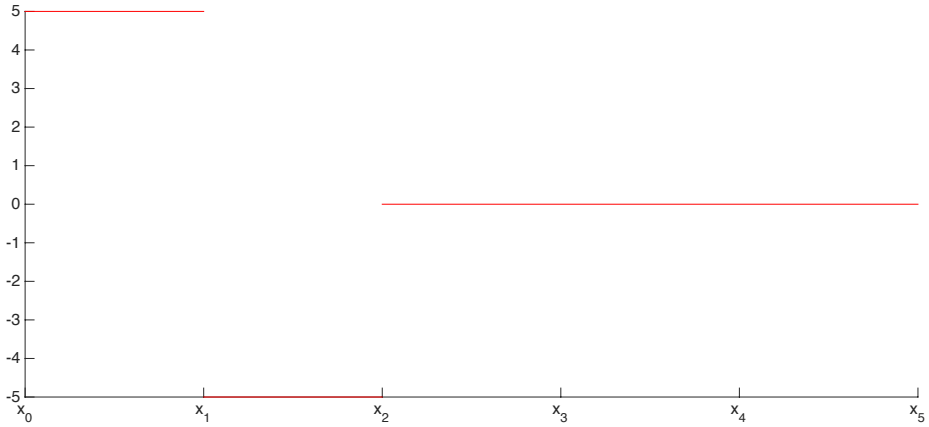

FIGURE 5. Graph of ψ_1', h=.2

For $j = i + 1$ we have

$$A_{i,i+1} = \int_{x_i}^{x_{i+1}} (-1/h)(1/h)dx = -1/h.$$

Finally, for $j = i - 1$ we have

$$A_{i,i-1} = \int_{x_{i-1}}^{x_i} (1/h)(-1/h)dx = -1/h.$$

In other words, we get

$$A = \frac{1}{h}\begin{bmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & & \\ 0 & \ddots & \ddots & \ddots & \\ \vdots & & & & -1 \\ 0 & \dots & & -1 & 2 \end{bmatrix}$$

We can easily calculate G to get

$$G = \frac{1}{h}\begin{bmatrix} a \\ 0 \\ \vdots \\ 0 \\ b \end{bmatrix}$$

Finally, in general we cannot compute exactly F. Instead we approximate F_i using the midpoint rule:

$$F_i = \int_0^1 f(x)\psi_i(x)dx = \int_{x_{i-1}}^{x_i} f(x)\psi_i(x)dx + \int_{x_i}^{x_{i+1}} f(x)\psi_i(x)dx \approx h\frac{f(\overline{x_i}) + f(\overline{x_{i+1}})}{2}.$$

where $\overline{x_i} = \frac{x_{i-1}+x_i}{2}$. Hence, computationally we can solve

$$\frac{1}{h^2}\begin{bmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & & \\ 0 & \ddots & \ddots & \ddots & \\ \vdots & & & & -1 \\ 0 & \dots & & -1 & 2 \end{bmatrix}\begin{bmatrix} u_h(x_1) \\ u_h(x_2) \\ u_h(x_3) \\ \vdots \\ u_h(x_N) \end{bmatrix} = \begin{bmatrix} \frac{f(\overline{x_1})+f(\overline{x_2})}{2} \\ \frac{f(\overline{x_2})+f(\overline{x_3})}{2} \\ \frac{f(\overline{x_3})+f(\overline{x_4})}{2} \\ \vdots \\ \frac{f(\overline{x_N})+f(\overline{x_{N+1}})}{2} \end{bmatrix} + \frac{1}{h^2}\begin{bmatrix} a \\ 0 \\ \vdots \\ 0 \\ b \end{bmatrix}$$

It is worth noting that the finite element method (where we approximate F with the midpoint rule) is very similar to the finite difference method. The only difference is the right hand side.

CHAPTER 2

Finite difference method for poisson problem in two dimensions

In this chapter we consider a standard finite difference method for Poisson's problem. For simplicity, we will consider the domain $\Omega = (0,1)^2$. We let $\partial\Omega$ denote the boundary of Ω. Throughout we let $x = (x_1, x_2) \in \mathbb{R}^2$. We define the Laplacian Δ applied to a smooth function w:

$$\Delta w(x) = \partial_{x_1}^2 w(x) + \partial_{x_2}^2 w(x).$$

The **Poisson** problem is given by

$$-\Delta u(x) = f(x) \quad \text{for } x \in \Omega \tag{0.1a}$$

$$u(x) = g(x) \quad \text{for } x \in \partial\Omega. \tag{0.1b}$$

Here the functions f and$g$ are given. The function f is known as the source term and g as the Dirichlet data. We will assume that f and g are smooth functions. The unknown is the function u for this section assume it belongs to $C^2(\Omega) \cap C(\overline{\Omega})$.

As we did in the one-dimensional case we can easily prove the weak maximum principle

Theorem 5. *Suppose u solves (0.1) and $f(x) \le 0$ for $x \in \Omega$ then*

$$\max_{x\in\overline{\Omega}} u(x) \le \max_{x\in\partial\Omega} g(x). \tag{0.2}$$

On the other hand, if $f(x) \ge 0$ for all $x \in \Omega$ then

$$\min_{x\in\overline{\Omega}} u(x) \ge \min_{x\in\partial\Omega} g(x). \tag{0.3}$$

Proof. We only prove (0.2) and leave (0.3) to the reader. Define $\phi(x) = \frac{(x_1-1/2)^2+(x_2-1/2)^2}{4}$ and let $w_\epsilon(x) = u(x) + \epsilon\phi(x)$ where $\epsilon > 0$. We see $\Delta\phi(x) = 1$ for all x and therefore we get

$$\begin{aligned} -\Delta w_\epsilon(x) &= f(x) - \epsilon && \text{for } x \in \Omega, \\ w_\epsilon(x) &= g(x) + \epsilon\phi(x) && \text{for } x \in \partial\Omega. \end{aligned}$$

Assume that w_ϵ attains a local maximum at $x \in \Omega$. Then by (1.2) we must have $\partial_{x_1}^2 w_\epsilon(x) \le 0$ and $\partial_{x_2}^2 w_\epsilon(x) \le 0$ which would imply that $\Delta w_\epsilon(x) \le 0$. However, $\Delta w_\epsilon(x) = -(f(x) - \epsilon) > 0$ by our hypothesis which reaches a contradiction. Therefore, w_ϵ cannot attain a local maximum in Ω let alone its global maximum in Ω. Hence, w_ϵ must attain its global maximum on $\partial\Omega$. In other words,

$$u(x) \le u(x) + \epsilon\phi(x) = w_\epsilon(x) \le \max_{y\in\partial\Omega} w_\epsilon(y) \quad \text{for all } x \in \overline{\Omega}.$$

On the other hand,

$$\max_{y\in\partial\Omega} w_\epsilon(y) \le \max_{y\in\partial\Omega} g(y) + \epsilon \max_{y\in\partial\Omega} \phi(y) \le \max_{x\in\partial\Omega} g(y) + \frac{\epsilon}{8}.$$

Where we used that $\max_{y\in\partial\Omega} \phi(y) \le \frac{1}{8}$. Combining the last two inequalities and taking the limit as $\epsilon \to 0$ gives

$$u(x) \le \max_{y\in\partial\Omega} g(y) \quad \text{for all } x \in \overline{\Omega}.$$

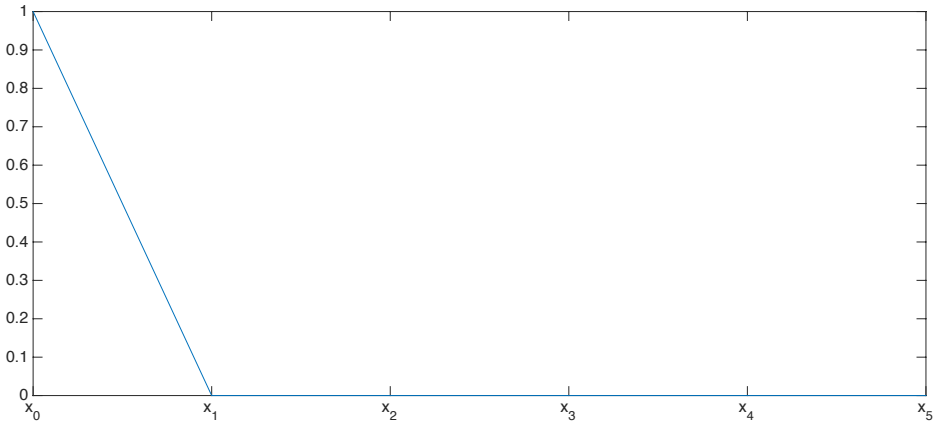

FIGURE 1. Graph of ψ_0

This proves (0.2). □

1. Finite difference method

In order to define a finite difference we need a grid defined on Ω. Let $N+1$ be given and define $h = 1/(N+1)$. Then, we define the grid as the collection of points

$$S = \{(nh, mh) : n, m = 0, 1, \ldots, N+1\}.$$

We denote the the interior points as S_I

$$S_I = \{(nh, mh) : n, m = 1, \ldots, N\}.$$

and the boundary points as $S_b = S \backslash S_I$. As we did in the one-dimensional case we define the space of discrete functions:

$$P_h = \{v : v \text{ is a real valued function with domain } S\}.$$

We denote the principle directions of $\mathbb{R}^2$ by $e_1 = (1, 0)$ and $e_2 = (0, 1)$.

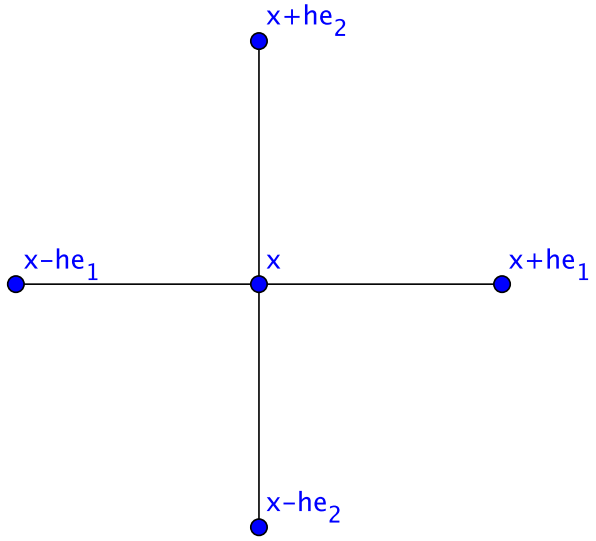

FIGURE 2. Five point stencil

We define the discrete Laplacian for every $x \in S_I$ as follows

$$\begin{aligned}\Delta_h v(x) \equiv & \frac{1}{h^2}(v(x+he_1) - 2v(x) + v(x-he_1)) + \frac{1}{h^2}(v(x+he_2) - 2v(x) + v(x-he_2)) \\ = & \frac{1}{h^2}(-4v(x) + v(x-he_1) + v(x-he_2) + v(x+he_1) + v(x+he_2)).\end{aligned}$$

In fact, we can prove the following result using Taylor's theorem

LEMMA 2. *Let* $u \in C^4(\overline{\Omega})$

$$\max_{x \in S_I} |\Delta u(x) - \Delta_h u(x)| \leq C\, h^2 \max\{\|\partial_{x_1}^4 u\|_{C(\Omega)}, \|\partial_{x_2}^4 u\|_{C(\Omega)}\} \tag{1.1}$$

where the constant C is independent of h and u. Here we define

$$\|w\|_{C(\Omega)} = \max_{x \in \overline{\Omega}} |w(x)|.$$

The standard finite difference method finds $v \in P_h$ such that

$$-\Delta_h v(x) = f(x) \quad \text{for all } x \in S_I \tag{1.2a}$$

$$v(x) = g(x) \quad \text{for all } x \in S_b. \tag{1.2b}$$

In order to be able to write a computer code we need to write the equivalent linear system that the finite difference gives rise to. To do that we should enumerate the vertices of S_I (note that there are N^2 vertices belonging to S_I) . We define $z_1, z_2, \ldots, z_{N^2}$ as (see Figure 3)

$$z_{n+(m-1)N} = (nh, mh) \quad \text{for all } n, m = 1, \ldots, N.$$

We then define $v(z_k) = v_k$ for $k = 1, \ldots, N^2$.

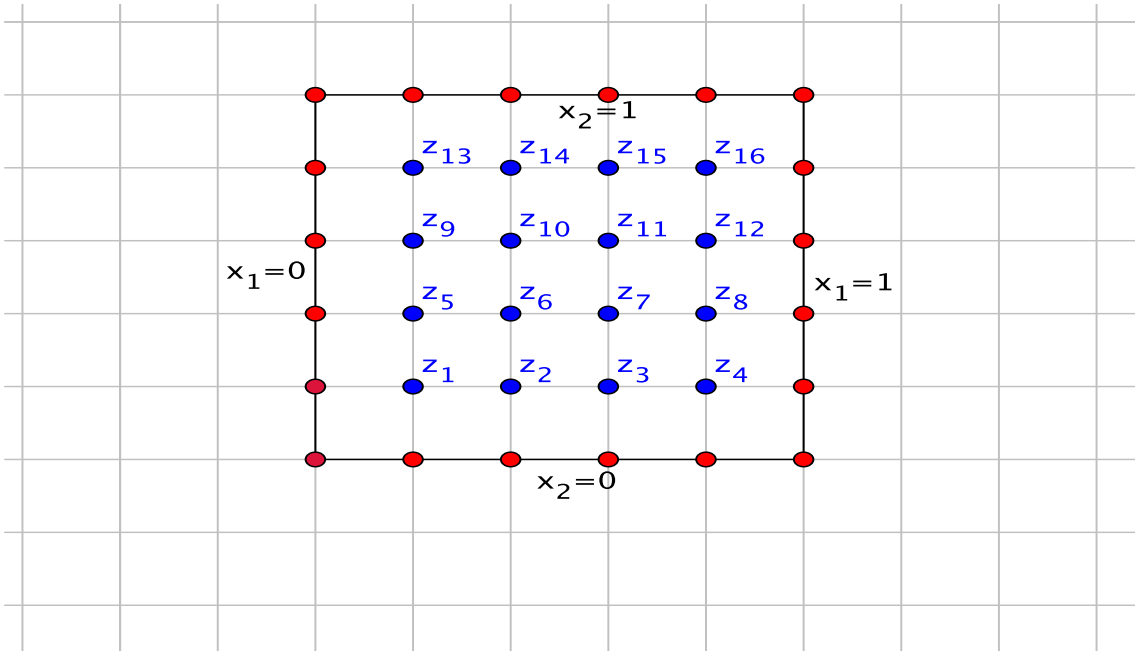

FIGURE 3. Grid, $N = 4$, grid points S_b are in red and S_I in blue

To write the linear system we need to distinguish between vertices that share an edge with a boundary vertex and those that don't. For vertices that don't share an edge with a boundary vertex we have the following linear equation

$$\frac{1}{h^2}(4v_k - v_{k+1} - v_{k+N} - v_{k-1} - v_{k-N}) = f(z_k)$$

for $k = n + (m-1)N$ with $2 \leq n \leq N-1$ and $2 \leq m \leq N-1$.

Then for those interior vertices that share an edge with vertices belonging to line $x = 0$ we have the following linear system

$$\begin{aligned}
\frac{1}{h^2}(4v_1 - v_2 - v_{N+1}) &= f(z_1) + \frac{1}{h^2}(g(z_1 - he_1) + g(z_1 - he_2)) \\
\frac{1}{h^2}(4v_2 - v_3 - v_{N+2} - v_1) &= f(z_2) + \frac{1}{h^2}(g(z_1 - he_2)) \\
&\vdots \\
\frac{1}{h^2}(4v_{N-1} - v_N - v_{2N-1} - v_{N-2}) &= f(z_{N-1}) + \frac{1}{h^2}(g(z_{N-1} - he_2)) \\
\frac{1}{h^2}(4v_N - v_{2N} - v_{N-1}) &= f(z_N) + \frac{1}{h^2}(g(z_N - he_2) + g(z_N + he_1))
\end{aligned}$$

Similarly, we can write the linear equations for the rest of vertices sharing an edge with a boundary vertex.

2. Error Analysis of Finite difference method

In order to obtain the error estimates of the finite difference method we need a stability result. And, before that we need a discrete maximum principle

2.1. Discrete Maximum Principle.

Theorem 6. *Suppose v solves* (1.2) *and $f(x) \leq 0$ for $x \in S_I$ then*

$$\max_{x\in S} v(x) \leq \max_{x\in S_b} g(x). \tag{2.1}$$

On the other hand, if $f(x) \geq 0$ for all $x \in S_I$ then

$$\min_{x\in S} v(x) \geq \min_{x\in S_b} g(x). \tag{2.2}$$

Proof. We will prove (2.1) and leave the proof (2.2) to the reader. Let $M = \max_{x\in S} v(x)$. First suppose that then $\max_{x\in S_I} v(x) < M$ then (2.1) holds.

On the other hand, suppose there that there exists $x \in S_I$ such that $v(x) = M$. Then from (1.2a) we have

$$\begin{aligned} M =& v(x) \\ =& \frac{1}{4}(v(x+he_1) + v(x+he_2) + v(x-he_1) + v(x-he_2)) + h^2 f(x) \\ \leq& \frac{1}{4}(v(x+he_1) + v(x+he_2) + v(x-he_1) + v(x-he_2)) \leq M \end{aligned}$$

since$f(x) \leq 0$ and by our hypothesis $\frac{1}{4}(v(x+he_1) + v(x+he_2) + v(x-he_1) + v(x-he_2)) \leq M$. Hence, we must have

$$M = \frac{1}{4}(v(x+he_1) + v(x+he_2) + v(x-he_1) + v(x-he_2)),$$

or that

$$\frac{1}{4}(M - v(x+he_1)) + \frac{1}{4}(M - v(x+he_2)) + \frac{1}{4}(M - v(x-he_1)) + \frac{1}{4}(M - v(x-he_2)) = 0.$$

By our hypothesis each term on the left hand side is non-negative as so it must be $v(x) = v(x+he_1) = v(x+he_2) = v(x-he_1) = v(x-he_2) = M$. We can repeat the argument for the $v(x+he_1), v(x+he_2), v(x-he_1), v(x-he_2)$ and have all its neighbors also equal to M. If we repeat this process we will arrive at $v \equiv M$ is a constant discrete function. Therefore, (2.1) trivially holds. □

2.2. Discrete Stability Result.

Theorem 7. *Suppose v solves* (1.2) *then*

$$\max_{x\in S} |v(x)| \leq \max_{x\in S_b} |g(x)| + \frac{1}{8} \max_{x\in S_I} |f(x)|. \tag{2.3}$$

Proof. Let $\phi(x) = \frac{Q}{4}(x_1 - 1/2)^2 + (x_2 - 1/2)^2$ where $Q = \max_{x\in S_I} |f(x)|$. The it is not difficult to show that $\Delta_h \phi(x) = Q$ for all $x \in S_I$. Then we define $w \in P_h$ as follows $w(x) = v(x) + \phi(x)$. We see that w solves □

$$-\Delta_h w(x) = f(x) - Q \quad \text{for all } x \in S_I \tag{2.4a}$$

$$w(x) = g(x) + \phi(x) \quad \text{for all } x \in S_b. \tag{2.4b}$$

Since $f(x) - Q \leq 0$ *for all* $x \in S_I$ *we can apply* (2.1) *to get*

$$v(x) \leq v(x) + \phi(x) \leq \max_{y \in S_b}(g(y) + \phi(y)) \quad \text{for all } x \in S.$$

However,

$$\max_{y \in S_b}(g(y) + \phi(y)) \leq \max_{y \in S_b}|g(y)| + \max_{y \in S_b}|\phi(y)| \leq \max_{y \in S_b}|g(y)| + Q/8.$$

Hence, we have

$$v(x) \leq \max_{y \in S_b}|g(y)| + Q/8. \quad \text{for all } x \in S.$$

Similarly, we can show that

$$-v(x) \leq \max_{y \in S_b}|g(y)| + Q/8. \quad \text{for all } x \in S.$$

Hence, we have

$$|v(x)| \leq \max_{y \in S_b}|g(y)| + Q/8. \quad \text{for all } x \in S,$$

which proves the result.

2.3. Error Estimate. We can prove the desired error estimate.

Theorem 8. *Suppose* v *solves* (1.2) *and suppose that* $u \in C^4(\overline{\Omega})$ *solves* (0.1) *then we have*

$$\max_{x \in S}|u(x) - v(x)| \leq C\, h^2 \max\{\|\partial^4_{x_1} u\|_{C(\Omega)}, \|\partial^4_{x_2} u\|_{C(\Omega)}\}$$

where C *is independent of* h *and* u

PROOF. Let $w \in P_h$ and let $w(x) = u(x) - v(x)$ for all $x \in S$.
Then, w solves

$$\begin{aligned} -\Delta_h w(x) &= -\Delta_h u(x) + f(x) && \text{for all } x \in S_I \\ w(x) &= 0 && \text{for all } x \in S_b. \end{aligned}$$

Therefore, we get by (2.3) we get

$$\max_{x \in S}|w(x)| \leq \frac{1}{8}\max_{x \in S_I}|-\Delta_h u(x) + f(x)| = \frac{1}{8}\max_{x \in S_I}|\Delta u(x) - \Delta_h u(x)|.$$

The result now follows from (1.1). □

CHAPTER 3

Finite Element Methods for second order elliptic Problems

In this chapter we study finite element methods for second order elliptic problems.

1. Function Spaces

We first need to study some function spaces. We let $\Omega \subset \mathbb{R}^2$ be a bounded domain with Lipschitz boundary. We recall the space of m-th order continuous functions:

$C^m(\Omega) = \{v : \text{ all the partial derivatives of order less than or equal to } m \text{ of } v \text{ are continuous in } \Omega\}.$

$$C_c^m(\Omega) = \{v \in C^m(\Omega) : v \text{ vanishes outside of } \Omega_0 \subset\subset \Omega\}.$$

Here Ω_0 is a compact domain and $\Omega_0 \subset\subset \Omega$ means $\partial\Omega_0 \cap \partial\Omega = \emptyset$.

We will need to make sense of derivatives of functions that are not differentiable everywhere. For example, the function $u(x) = |x|$. To do that we need to define the *weakderivative* of a function if it exists. This is done via integration by parts.

Let us recall integration by parts formula for smooth functions. Let $\mathbf{n} = (\mathbf{n_1}, \mathbf{n_2})$ be the unit outward pointing normal to Ω. Suppose that $u, v \in C^\infty(\mathbb{R}^2)$

$$\int_\Omega \partial_{x_i} u(x)v(x)dx = -\int_\Omega u(x)\partial_{x_i} v(x)dx - \int_{\partial\Omega} u(x)v(x)n_i ds \quad \text{for } i = 1, 2.$$

Let u be an integrable function on Ω then we say that $\partial_{x_i} u$ exists if there is an integrable function ϕ such that

$$\int_\Omega \phi(x)v(x)dx = -\int_\Omega u(x)\partial_{x_i} v(x)dx \text{ for all } v \in C_c^\infty(\Omega).$$

in which case we define $\partial_{x_i} u(x) = \phi(x)$. In other words, if $\partial_{x_i} u$ then

$$\int_\Omega \partial_{x_i} u(x)v(x)dx = -\int_\Omega u(x)\partial_{x_i} v(x)dx \text{ for all } v \in C_c^\infty(\Omega). \tag{1.1}$$

Or course $\partial_{x_i} u(x)$ are uniquely defined up to a set of measure zero.

We will also need to define the following function spaces

$$L^2(\Omega) = \{v : v \text{ is integrable on } \Omega \text{ and } \int_\Omega v^2(x)dx < \infty\}.$$

We define the Sobolev space H^m as follows

$$H^m(\Omega) = \{v \in L^2(\Omega) : \partial_{x_1}^{\alpha_1}\partial_{x_2}^{\alpha_2} u \in L^2(\Omega), \text{ for all } \alpha_1 + \alpha_2 \le m\}.$$

The associated norm is given as follows

$$\|u\|_{H^m(\Omega)}^2 = \sum_{j=0}^m |u|_{H^j(\Omega)}^2 \text{ where } \quad |u|_{H^j(\Omega)}^2 = \sum_{\alpha_1+\alpha_2=j} \|\partial_{x_1}^{\alpha_1}\partial_{x_2}^{\alpha_2} u\|_{L^2(\Omega)}^2.$$

2. Linear second-order elliptic problems

Recall that div $F(x) = \partial_{x_1} F_1(x) + \partial_{x_2} F_2(x)$ where $F(x) = (F_1(x), F_2(x))$. We will study second order elliptic problems of the form

$$-\text{div } (A(x)\nabla u(x)) = f(x) \quad \text{for } x \in \Omega \tag{2.1a}$$

$$u(x) = g(x) \quad \text{for } x \in \partial\Omega. \tag{2.1b}$$

Here we assume that $A(x) \in \mathbb{R}^{2\times 2}$ is a symmetric matrix for every $x \in \overline{\Omega}$ and is smooth on $\overline{\Omega}$. Notice that if $A(x)$ is the identity matrix for all x then div $(A(x)\nabla u(x)) = -\Delta u(x)$. We can write out div $(A(x)\nabla u(x))$

$$\text{div } (A(x)\nabla u(x)) = \partial_{x_1}(A_{11}(x)\partial_{x_1}u(x) + A_{12}(x)\partial_{x_2}u(x)) + \partial_{x_1}(A_{21}(x)\partial_{x_1}u(x) + A_{22}(x)\partial_{x_2}u(x)).$$

In particular, if A is a constant matrix independent of x we have

$$\text{div } (A(x)\nabla u(x)) = A_{11}\partial^2_{x_1}u(x) + 2A_{12}\partial_{x_1}\partial_{x_2}u(x) + A_{22}\partial^2_{x_2}u(x).$$

We assume that A is uniformly elliptic. That, is there exists a constant $\theta > 0$ such that

$$y^t A(x) y \geq \theta y^t y \quad \text{for all } y \in \mathbb{R}^2 \text{ and } x \in \Omega. \tag{2.2}$$

One can easily show that this is equivalent to $A(x)$ having uniformly positive eigenvalues on Ω. Since A is smooth we also have that there exists a constant $\gamma > 0$ such that

$$y^t A(x) z \leq \gamma |y||z| \quad \text{for all } y, z \in \mathbb{R}^2 \text{ and } x \in \Omega. \tag{2.3}$$

2.1. Weak Formulation of second-order problems. We define $H^1_0(\Omega) = \{v \in H^1(\Omega) : v \text{ vanishes on } \partial\Omega\}$. If we use the integration by parts formula (1.1) we can derive Gauss' integral formula

$$\int_\Omega \text{div } F(x)v(x)dx = -\int_\Omega F(x)\cdot\nabla v(x)dx + \int_{\partial\Omega} F(x)\cdot \mathbf{nv(x)ds}.$$

If we apply this to (2.1a) we have

$$\int_\Omega A(x)\nabla u(x)\cdot\nabla v(x)dx = \int_\Omega f(x)v(x)dx \quad \text{for all } v \in H^1_0(\Omega).$$

Hence, we can state the weak formulation of (2.1) as follows Find $u \in H^1(\Omega)$ with $u = g$ on $\partial\Omega$ such that

$$\int_\Omega A(x)\nabla u(x)\cdot\nabla v(x)dx = \int_\Omega f(x)v(x)dx \quad \text{for all } v \in H^1_0(\Omega). \tag{2.4}$$

We use the notation

$$a(u,v) = \int_\Omega A(x)\nabla u(x)\cdot\nabla v(x)dx. \tag{2.5}$$

and

$$(f,u) = \int_\Omega f(x)v(x)dx.$$

Hence, we can write the weak form as follows Find $u \in H^1(\Omega)$ with $u = g$ on $\partial\Omega$ such that

$$a(u,v) = (f,v) \quad \text{for all } v \in H^1_0(\Omega). \tag{2.6}$$

The nice property of the weak formulation is that solutions could exists even though they do not have second derivatives. They only require that $u \in H^1(\Omega)$. Therefore, the weak formulation is a generalization of (2.1).

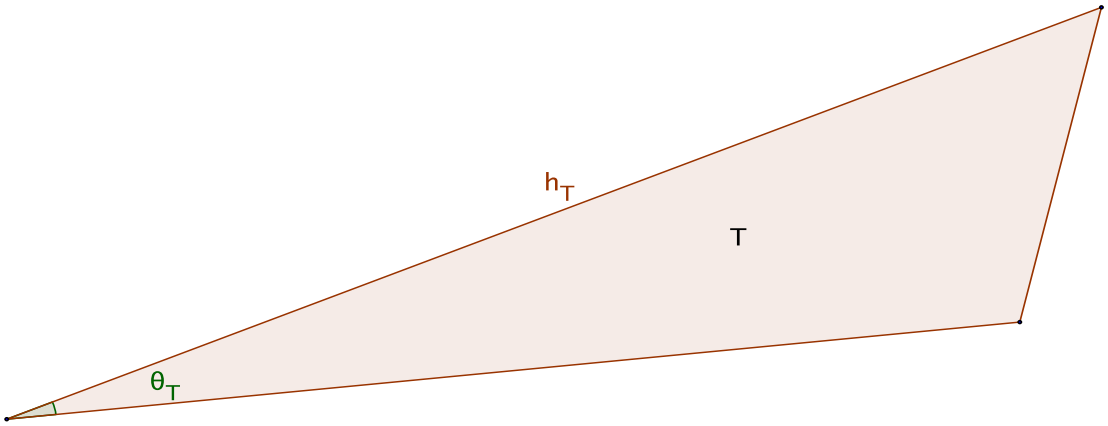

FIGURE 1. Illustration of θ_T, h_T

3. Finite Element Formulation

In this section we develop finite element method for (2.6). To do that we need to define some concepts. For simplicity we will assume that Ω is a polygonal domain. We will assume that we have a family of triangulations of Ω of $\{\mathcal{T}_h\}$. For every h we assume that

$$\overline{\Omega} = \cup_{T\in\mathcal{T}_h}\overline{T}.$$

If $T, K \in \mathcal{T}_h$ and $T \neq K$ then $\overline{T} \cap \overline{K}$ is either empty, a vertex or an edge of the triangulation.

We assume that the mesh is *shape regular*: There exists a constant $\kappa > 0$ such that

$$\theta_T \geq \kappa \quad \text{for all } T \in \{\mathcal{T}_h\}, \tag{3.1}$$

where θ_T denotes the smallest angle of the triangle T.

We also also define

$$h_T = \operatorname{diam}(T),$$

and

$$h = \sup_{T\in\mathcal{T}_h} h_T.$$

We now define finite element spaces:

$$V_h = \{v \in C(\Omega) : v|_T \in P^1(T) \text{ for all } T \in \mathcal{T}_h\},$$

where $P^1(T)$ is the space of affine function defined on T. We also define

$$V_h^0 = \{v \in V_h, v \equiv 0 \text{ on } \partial\Omega\}.$$

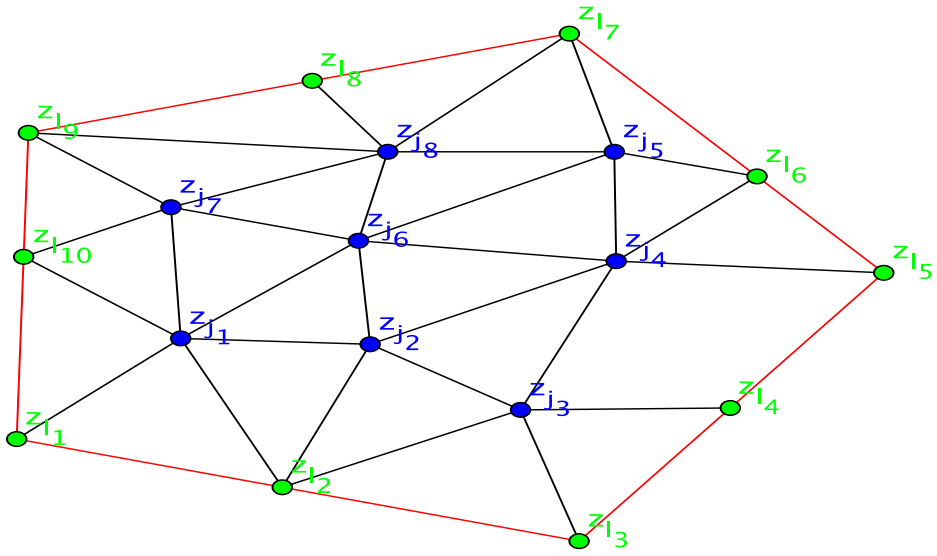

FIGURE 2. Example of a triangulation of a polygonal domain the vertices S_b are in green, and S_I in blue

We can also define a basis for the spaces V_h. Let us denote by

$$\begin{aligned} S =& \{ \text{ set of all vertices of } \mathcal{T}_h\} \\ S_I =& \{x \in S : x \notin \partial\Omega\} \\ S_b =& S \backslash S_I. \end{aligned}$$

Let us enumerate the vertices in S, S_I, S_b:

$$\begin{aligned} S =& \{z_1, z_2, \dots, z_M\} \\ S_I =& \{z_{j_1}, z_{j_2}, \dots, z_{j_Q}\} \\ S_b =& \{z_{\ell_1}, z_{\ell_2}, \dots, z_{\ell_R}\} \end{aligned}$$

here $1 \le j_1 \le \cdots \le j_Q \le M$, $1 \le \ell_1 \le \cdots \le \ell_R \le M$, and $Q + R = M$.

For each $i = 1, \dots, M$ we define $\psi_i \in V_h$

$$\psi_i(z_j) = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{if } j \neq i \end{cases}$$

Then we see that if $v \in V_h$ we have

$$v(x) = \sum_{1 \le i \le 1} v(z_i)\psi_i(x) \quad \text{for all } x \in \overline{\Omega}.$$

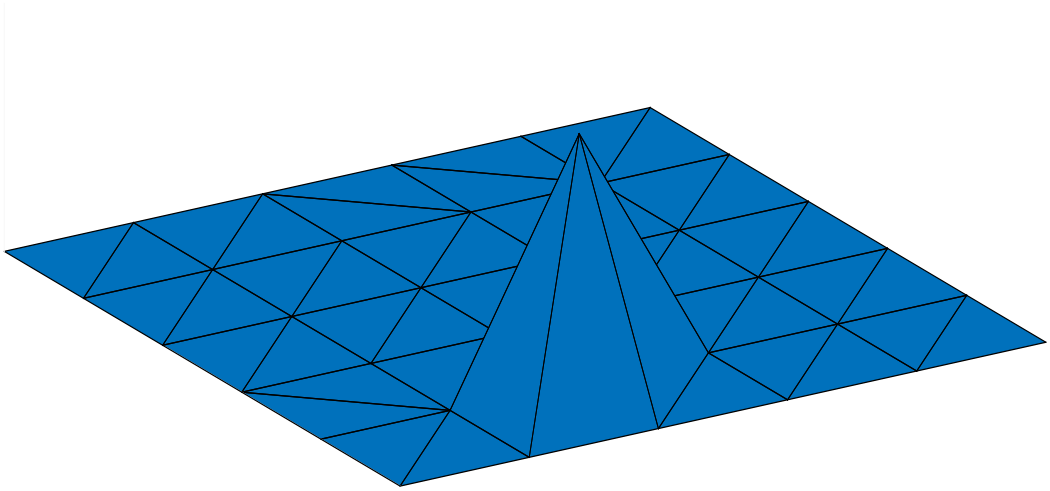

FIGURE 3. An example of an ψ_i basis function

The finite element method solves: Find $u_h \in V_h$ with $u(z) = g(z)$ for all $z \in S_b$ such that

$$\int_\Omega A(x)\nabla u_h(x) \cdot \nabla v_h(x)dx = \int_\Omega f(x)v_h(x)dx \quad \text{for all } v_h \in V_h^0. \tag{3.2}$$

or using the $a(\cdot,\cdot)$ notation Find $u_h \in V_h$ with $u(z) = g(z)$ for all $z \in S_b$ such that

$$a(u_h, v_h) = (f, v) \quad \text{for all } v_h \in V_h^0. \tag{3.3}$$

We can easily show, using that $a(u_h,\cdot)$, $(f,\cdot)$ are linear forms that

$$a(u_h, \psi_{j_k}) = (f, \psi_{j_k}) \quad 1 \le k \le Q. \tag{3.4}$$

If we write

$$u_h(x) = \sum_{1\le i\le M} u_h(z_i)\psi_i(x) \quad \text{for all } x \in \overline{\Omega}.$$

we have

$$\sum_{1\le i\le M} u_h(z_i)a(\psi_i, \psi_{j_k}) = (f, \psi_{j_k}) \quad 1 \le k \le Q. \tag{3.5}$$

or using that $u_h(z) = g(z)$ for $z \in S_b$ we get:

$$\sum_{1\le s\le Q} u_h(z_{j_s})a(\psi_{j_s}, \psi_{j_k}) = (f, \psi_{j_k}) - \sum_{1\le s\le R} g(z_{\ell_s})a(\psi_{\ell_s}, \psi_{j_k}) \quad 1 \le k \le Q. \tag{3.6}$$

We then have the matrix system

$$\begin{bmatrix} a(\psi_{j_1},\psi_{j_1}) & a(\psi_{j_1},\psi_{j_2}) & \dots & a(\psi_{j_1},\psi_{j_Q}) \\ a(\psi_{j_2},\psi_{j_1}) & a(\psi_{j_2},\psi_{j_2}) & \dots & a(\psi_{j_2},\psi_{j_Q}) \\ \vdots & & \ddots & \vdots \\ a(\psi_{j_Q},\psi_{j_1}) & \dots & & a(\psi_{j_Q},\psi_{j_Q}) \end{bmatrix} \begin{bmatrix} u_h(z_{j_1}) \\ u_h(z_{j_2}) \\ \vdots \\ u_h(z_{j_Q}) \end{bmatrix} = \begin{bmatrix} (f,\psi_{j_1}) \\ (f,\psi_{j_2}) \\ \vdots \\ (f,\psi_{j_Q}) \end{bmatrix} + G,$$

where

$$G = \begin{bmatrix} a(\psi_{j_1},\psi_{\ell_1}) & a(\psi_{j_1},\psi_{\ell_2}) & \dots & a(\psi_{j_1},\psi_{\ell_R}) \\ a(\psi_{j_2},\psi_{\ell_1}) & a(\psi_{j_2},\psi_{\ell_2}) & \dots & a(\psi_{j_2},\psi_{\ell_R}) \\ \vdots & & \ddots & \vdots \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ a(\psi_{j_Q},\psi_{\ell_1}) & \dots & & a(\psi_{j_Q},\psi_{\ell_R}) \end{bmatrix} \begin{bmatrix} g(z_{\ell_1}) \\ g(z_{\ell_2}) \\ \vdots \\ g(z_{\ell_R}) \end{bmatrix}.$$

3.1. Implemenation of FEM.

3.2. Error Analysis of FEM. In this section we perform an error analysis of the finite element method. We will see that error estimates follows easily from, coercivity and boundedness of bilinear form and from Galerkin Orthogonality of the FEM.

LEMMA 3. **Coercivity***: It holds*

$$\theta\|\nabla v\|^2_{L^2(\Omega)} \le a(v,v) \quad \textit{for all } v \in H^1(\Omega). \tag{3.7}$$

PROOF. Recall that

$$a(v,v) = \int_\Omega A(x)\nabla v(x)\cdot\nabla v(x)dx = \int_\Omega (\nabla v(x))^t A(x)\nabla v(x)dx$$

then the result follows from (2.2). □

LEMMA 4. **Boundedness***: It holds,*

$$a(v,w) \le \gamma\|\nabla v\|_{L^2(\Omega)}\|\nabla w\|_{L^2(\Omega)} \le \quad \textit{for all } v,w \in H^1(\Omega). \tag{3.8}$$

PROOF. Similar to the previous proof the result follows from (2.3). □

LEMMA 5. **Galerkin Orthogonality** *Let u solve* (2.5) *and let u_h solve* (3.3) *then we have*

$$a(u-u_h,v_h) = 0 \quad \textit{for all } v_h \in V_h^0. \tag{3.9}$$

PROOF. Since $V_h^0 \subset H_0^1(\Omega)$ we have by (2.5)

$$a(u,v_h) = (f,v_h) \quad \text{for all } v_h \in V_h^0.$$

However, (3.3) states that

$$a(u_h,v_h) = (f,v_h) \quad \text{for all } v_h \in V_h^0.$$

Therefore, we have

$$a(u,v_h) - a(u_h,v_h) = 0 \quad \text{for all } v_h \in V_h^0.$$

Using that $a(\cdot,v_h)$ is linear we get

$$a(u,v_h) - a(u_h,v_h) = a(u-u_h,v_h) \quad \text{for all } v_h \in V_h^0.$$

The result now follows. □

Now we can easily prove the following result.

Theorem 9. *Let u solve (2.5) and let u_h solve (3.3) then we have*

$$\|\nabla(u-u_h)\|_{L^2(\Omega)} \leq \frac{\gamma}{\theta} \min_{v\in V_h^g} \|\nabla(u-v_h)\|_{L^2(\Omega)}. \tag{3.10}$$

where $V_h^g = \{v_h \in V_h : v_h(z) = g(z) \quad \text{for all } z \in S_b\}$.

Proof. By Coercivity (3.7) we obtain

$$\theta\|\nabla(u-u_h)\|^2_{L^2(\Omega)} \leq a(u-u_h, u-u_h).$$

Now by linearity of $a(u-u_h,\cdot)$ we have

$$a(u-u_h, u-u_h) = a(u-u_h, u-v_h) - a(u-u_h, u_h-v_h) \quad \text{for all } v_h \in V_h^g.$$

Since $u_h, v_h \in V_h^g$ we have $u_h - v_h \in V_h^0$ and therefore, we get

By Galerkin orthogonality (3.9)

$$a(u-u_h, u_h-v_h) = 0 \quad \text{for all } v_h \in V_h^g.$$

Combining the above results we get

$$\theta\|\nabla(u-u_h)\|^2_{L^2(\Omega)} \leq a(u-u_h, u-v_h) \text{ for all } v_h \in V_h^g.$$

Using (3.8) we get

$$\theta\|\nabla(u-u_h)\|^2_{L^2(\Omega)} \leq \gamma\|\nabla(u-u_h)\|_{L^2(\Omega)}\|\nabla(u-v_h)\|_{L^2(\Omega)} \text{ for all } v_h \in V_h^g.$$

Dividing by $\|\nabla(u-u_h)\|_{L^2(\Omega)}$ we obtain

$$\theta\|\nabla(u-u_h)\|_{L^2(\Omega)} \leq \frac{\gamma}{\theta}\|\nabla(u-v_h)\|_{L^2(\Omega)} \text{ for all } v_h \in V_h^g.$$

The result follows by taking the minimum over $v_h \in V_h^g$. □

The result (3.12) says that the FEM approximation u_h is almost the best approximation to u in the space V_h when measured in the $\|\nabla\cdot\|_{L^2(\Omega)}$ norm. Next, we give an approximation result in terms of the mesh size h. The proof of this result which is quite technical will proved in the next section.

Proposition 2. *If $u \in H^2(\Omega)$ then*

$$\min_{v_h\in V_h^g} \|\nabla(u-v_h)\|_{L^2(\Omega)} \leq Ch|u|_{H^2(\Omega)}. \tag{3.11}$$

Then the following corollary follows from (3.11) and (3.12).

Corollary 1. *Let u solve (2.5) and let u_h solve (3.3) and assume that $u \in H^2(\Omega)$, then*

$$\|\nabla(u-u_h)\|_{L^2(\Omega)} \leq C\, h|u|_{H^2(\Omega)}. \tag{3.12}$$

where the constant C is independent of u and h.

We now wan to study the error estimates $\|u-u_h\|_{L^2(\Omega)}$ instead. For simplicity we assume that $g \equiv 0$. To do this we need a result from PDE theory. The following is a standard H^2 regularity resullt.

Proposition 3. *Let u solve (2.5) with $g \equiv 0$ and suppose that Ω is a a convex polygon then*

$$\|u\|_{H^2(\Omega)} \leq C\,\|f\|_{L^2(\Omega)}. \tag{3.13}$$

The result seems very reasonable since formally $\Delta u = -f$. So, we that some combination of second derivatives are controlled by f. The result says that all the second derivatives of u are controlled by f.

We can now state the desired estimate.

Theorem 10. *Let u solve (2.5) with $g \equiv 0$ and let u_h solve (3.3) and assume that Ω is a a convex polygon, then*

$$\|u - u_h\|_{L^2(\Omega)} \leq Ch^2 \|f\|_{L^2(\Omega)}.$$

Please note that this result says that $\|u - u_h\|_{L^2(\Omega)}$ converges to zero with one order higher than $\|\nabla(u - u_h)\|_{L^2(\Omega)}$.

PROOF. Let $\phi \in H_0^1(\Omega)$ solve the following problem

$$a(\phi, v) = (u - u_h, v) \quad \text{for all } v \in H_0^1(\Omega).$$

Then, setting $v = u - u_h$ we get

$$\|u - u_h\|^2_{L^2(\Omega)} = a(\phi, u - u_h) = a(u - u_h, \phi).$$

If we use Galerkin Orthogonality (3.9) we get

$$\|u - u_h\|^2_{L^2(\Omega)} = a(u - u_h, \phi - v_h) \quad \text{for all } v_h \in V_h^0.$$

By (3.8) we get

$$\|u - u_h\|^2_{L^2(\Omega)} \leq \gamma \|\nabla(u - u_h)\|_{L^2(\Omega)} \|\nabla(\phi - v_h)\|_{L^2(\Omega)} \quad \text{for all } v_h \in V_h^0.$$

Using (3.11) and (3.13) we obtain

$$\|u - u_h\|^2_{L^2(\Omega)} \leq C\, h \|\nabla(u - u_h)\|_{L^2(\Omega)} \|u - u_h\|_{L^2(\Omega)}.$$

After dividing by $\|u - u_h\|_{L^2(\Omega)}$, using (3.12) and (3.13) we obtain the result. □

3.3. Approximation of Piecewise linear functions. We are left to prove (3.11). To be precise we will use standard results that are used in the analysis of elliptic PDEs. These results are useful in many other context and not only important for the approximation theory we will give here.

We will fix a reference triangle: $\hat{T}$ which is the triangle with vertices (0,0), (1, 0) , (0,1). We then state a few results.

The first is Poincare's inequality.

PROPOSITION 4. **Poincare's inequality***: Suppose that $w \in H^1(\hat{T})$ and let $\overline{w} = \frac{1}{|\hat{T}|} \int_{\hat{T}} w(x)dx$ then,*

$$\|w - \overline{w}\|_{L^2(\hat{T})} \leq C\, |w|_{H^1(\hat{T})}, \tag{3.14}$$

where C is independent of w.

The next result is one type Sobolev inequality. There are many Sobolev embedding and inequalities.

PROPOSITION 5. **A Sobolev inequality***: If $v \in H^2(\hat{T})$ then $v \in C(\overline{\hat{T}})$ and the following inequality holds*

$$\|v\|_{L^\infty(\hat{T})} \leq C \|v\|_{H^2(\hat{T})}. \tag{3.15}$$

We will are going to transfer estimates from $T \in \mathcal{T}_h$ to $\hat{T}$. The important point here is that $\hat{T}$ is fixed. To do this, we need to define an affine transformation from $F_T : \hat{T} \to T$ define as

$$F_T(\hat{x}) = B_T \hat{x} + b_T.$$

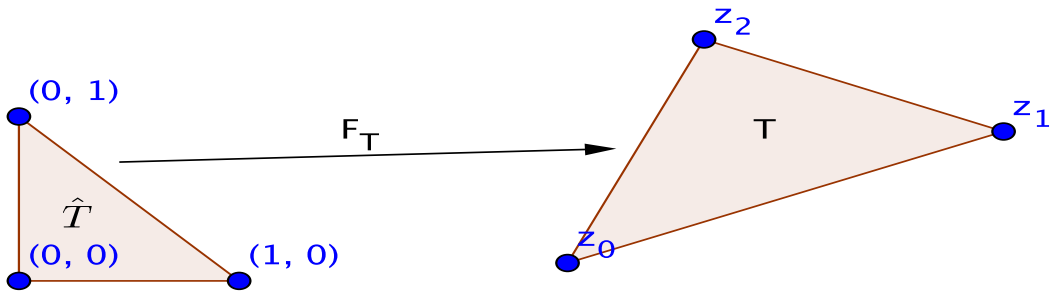

FIGURE 4. An example of an ψ_i basis function

Not that if T has vertices z_0, z_1, z_2 then the first column of B_T is $z_1 - z_0$ while the second column is $z_2 - z_0$ and finally $b_T = z_0$. That is,

$$B_T = \begin{bmatrix} z_1 - z_0 & z_2 - z_0 \end{bmatrix}, \quad b_T = z_0.$$

We will state a proof of the size of entries of B_T and B_T^{-1}. Here we use strongly the shape regularity, (3.1)

LEMMA 6. *There exists a constant C independent of h and $T \in \mathcal{T}_h$ such that*

$$|(B_T)_{ij}| \leq C\, h_T \quad |(B_T^{-1})_{ij}| \leq C h_T^{-1} \quad \text{for all } T \in \mathcal{T}_h, i, j = 1, 2. \tag{3.16}$$

We also use the following notation. Let w be defined as a function on T then define the function $\hat{w}$ as function on $\hat{T}$ given by

$$\hat{w}(\hat{x}) = w(F_T(\hat{x})).$$

Using a change of variable formula we have

$$\int_T w(x)dx = \int_{\hat{T}} w(F_T(\hat{x}))\det(B_T)d\hat{x}. \tag{3.17}$$

or

$$\int_T w(x)dx = \int_{\hat{T}} \hat{w}(\hat{x})\det(B_T)d\hat{x}. \tag{3.18}$$

Now notice that using the chain rule we have

$$\nabla_x w(x) = \nabla_x \hat{w}(F_T^{-1}(x)) = B_T^{-t}\nabla_{\hat{x}}\hat{w}(F_T^{-1}(x)) = B_T^{-t}\nabla_{\hat{x}}\hat{w}(\hat{x}), \tag{3.19}$$

where $x = F_T(\hat{x})$. Note that we write ∇_x to denote the gradient with respect to the variable x. Whereas, $\nabla_{\hat{x}}$ is gradient with respect to $\hat{x}$.

Hence using (3.17) and (3.19) we obtain

$$\int_T \nabla_x w(x) \cdot \nabla_x v(x) dx = \int_{\hat{T}} (B_T^{-t} \nabla_{\hat{x}} \hat{w}(\hat{x})) \cdot (B_T^{-t} \nabla_{\hat{x}} \hat{v}(\hat{x})) \det(B_T) d\hat{x}. \tag{3.20}$$

We can then prove the following results

LEMMA 7. *It holds,*

$$\frac{c_1}{h_T} \|w\|_{L^2(T)} \leq \|\hat{w}\|_{L^2(\hat{T})} \leq \frac{c_2}{h_T} \|w\|_{L^2(T)}, \tag{3.21}$$

$$c_1 |w|_{H^1(T)} \leq |\hat{w}|_{H^1(\hat{T})} \leq c_2 |w|_{H^1(T)}, \tag{3.22}$$

$$c_1 h_T |w|_{H^2(T)} \leq |\hat{w}|_{H^2(\hat{T})} \leq c_2 h_T \|w\|_{L^2(T)}. \tag{3.23}$$

PROOF. We only prove (3.22). The proof of the other two proofs follow a similar line of argument. By (3.20) we have

$$|w|^2_{H^1(T)} = \|\nabla w\|^2_{L^2(T)} = \det B_T \|B_T^{-t} \nabla \hat{w}\|^2_{L^2(\hat{T})}.$$

By (3.16) we have

$$\det B_T \leq C\, h_T^2,$$

and

$$\|B_T^{-t} \nabla \hat{w}\|^2_{L^2(\hat{T})} \leq C h_T^{-2} \|\nabla \hat{w}\|^2_{L^2(\hat{T})}.$$

The result now follows. □

We can now state a local approximation result known as the Bramble-Hilbert lemma. The result will follow

LEMMA 8. **Bramble–Hilbert***: Let $w \in H^2(T)$ then there exists a $v \in P^1(T)$ such that*

$$\|w - v\|_{L^2(T)} + h_T \|\nabla(w - v)\|_{L^2(T)} \leq C h_T^2 |w|_{H^2(T)}.$$

PROOF. Let $v(x) = c_0 + c_1 x_1 + c_2 x_2$ where $x = (x_1, x_2)$. We define first c_1, c_2

$$c_i = \frac{1}{|T|} \int_T \partial_{x_i} w(x) dx \quad \text{for all } i = 1, 2.$$

Note that since $c_i = \partial_{x_i} v(x)$ we have that

$$\frac{1}{|T|} \int_T \partial_{x_i} v(x) dx = \frac{1}{|T|} \int_T \partial_{x_i} w(x) dx \quad \text{for all } i = 1, 2. \tag{3.24}$$

Now that we have define c_1, c_2 we define c_0 so that

$$\frac{1}{|T|} \int_T v(x) dx = \frac{1}{|T|} \int_T w(x) dx. \tag{3.25}$$

That is,

$$c_0 = \frac{1}{|T|} \int_T (w(x) - (c_1 x_1 + c_2 x_2)) dx.$$

If we define $q(x) = w(x) - v(x)$ and consider $\hat{q}(\hat{x})$. Note that by (3.24)

$$\int_T \partial_{x_i} q(x) = 0 \quad \text{for } i = 1, 2$$

Again, using (3.17) and (3.19) we get

$$\det B_T((B_T^{-t})_{i1}\int_{\hat{T}}\partial_{\hat{x}_1}\hat{q}(\hat{x})d\hat{x}+(B_T^{-t})_{i2}\int_{\hat{T}}\partial_{\hat{x}_2}\hat{q}(\hat{x})d\hat{x})=\int_T\partial_{x_i}q(x)dx,$$

for $i=1,2$. Hence, we get

$$(\det B_T)B_T^{-1}\begin{bmatrix}\int_T\partial_{\hat{x}_1}\hat{q}(\hat{x})d\hat{x}\\ \int_T\partial_{\hat{x}_2}\hat{q}(\hat{x})d\hat{x}\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

Hence, we get

$$\int_{\hat{T}}\partial_{\hat{x}_i}\hat{q}(\hat{x})d\hat{x}=0\quad\text{ for } i=1,2$$

Similarly, we have

$$\int_T\hat{q}(\hat{x})d\hat{x}=0.$$

Thus, $\overline{\partial_{\hat{x}_i}\hat{q}}\equiv 0$ and $\overline{\hat{q}}\equiv 0$ and by Poincare's inequality we get

$$|\hat{q}|^2_{H^1(\hat{T})}=\|\partial_{\hat{x}_1}\hat{q}\|^2_{L^2(\hat{T})}+\|\partial_{\hat{x}_2}\hat{q}\|^2_{L^2(\hat{T})}\le C\sum_{\alpha_1+\alpha_2=2}\|\partial^{\alpha_1}_{\hat{x}_1}\partial^{\alpha_2}_{\hat{x}_2}\hat{q}\|^2_{L^2(\hat{T})}=C|\hat{q}|^2_{H^2(\hat{T})}.$$

Using this result with (3.22) gives

$$|q|_{H^1(T)}\le C\,|\hat{q}|_{H^1(\hat{T})}\le C\,|\hat{q}|^2_{H^2(\hat{T})}\le h_T|q|_{H^2(T)}.$$

Also, using (3.21), Poincare's inequality and the previous inequality we get

$$\|q\|_{L^2(T)}\le Ch_T\|\hat{q}\|_{L^2(\hat{T})}\le Ch_T|\hat{q}|_{H^1(\hat{T})}\le Ch_T\|q\|_{L^2(T)}\le Ch_T^2|q|_{H^2(T)}.$$

□

We will also need an inverse estimate. We can prove this result with more elementary results. We give a proof that is generalizable to other finite element spaces.

LEMMA 9. **Inverse inequality** *: Let $T\in\mathcal{T}_h$ then for every $v\in P^1(T)$*

$$|v|_{H^1(T)}\le\frac{C}{h_T}\|v\|_{L^2(T)}$$

where the constant C is independent of T and v.

PROOF. By (3.22)

$$|v|_{H^1(T)}\le C|\hat{v}|_{H^1(\hat{T})}\tag{3.26}$$

Note that $\hat{v}\in P^1(\hat{T})$. Since $P^1(\hat{T})$ is a finite dimensional space we know that norms are equivalent. Hence,

$$|\hat{v}|_{H^1(\hat{T})}\le\hat{C}\|\hat{v}\|_{L^2(\hat{T})}.$$

where the constant $\hat{C}$ is independent $\hat{v}$. Note that we are using that $\hat{T}$ is fixed. Using (3.21) we get

$$\|\hat{v}\|_{L^2(\hat{T})}\le\frac{C}{h_T}\|v\|_{L^2(T)}.$$

Combining the above inequalities gives the result. □

We can state a simple corollary.

COROLLARY 2. *Let $T\in\mathcal{T}_h$ with vertices z_i $(i=0,1,2)$ then for every $v\in P^1(T)$*

$$|v|_{H^1(T)}\le C\max\{v(z_0),v(z_1),v(z_2)\}.\tag{3.27}$$

PROOF. We easily see that

$$\|v\|^2_{L^2(T)} = \int_T v^2(x)dx \leq |T|\|v\|^2_{L^\infty(T)}.$$

Hence,

$$\|v\|_{L^2(T)} \leq \sqrt{|T|}\|v\|_{L^\infty(T)}.$$

Since the maximum of $|v|$ occurs at the vertices and so we have

$$\|v\|_{L^2(T)} \leq \sqrt{|T|}\max\{v(z_0), v(z_1), v(z_2)\}.$$

Using that the mesh is shape regular (3.1) we easily have that $\sqrt{|T|} \leq Ch_T$. Hence, combining the above inequality with (3.26). □

In order to prove our main result (3.11) we need to define the interpolant operator.

DEFINITION 3.1. *For $w \in C(\overline{\Omega})$ we define $I_h w = V_h$ by*

$$I_h w(z) = w(z) \quad \text{for all } z \in S.$$

3.4. Proof of (3.11). First we see that $I_h u \in V_h^g$ and so

$$\inf_{v \in V_h^g} \|\nabla(u - v)\|_{L^2(\Omega)} \leq \|\nabla(u - I_h u)\|_{L^2(\Omega)}.$$

Let us write

$$\|\nabla(u - I_h u)\|^2_{L^2(\Omega)} = \sum_{T \in \mathcal{T}_h} \|\nabla(u - I_h u)\|^2_{L^2(T)}. \tag{3.28}$$

From the Bramble-Hilbert Lemma we have there exists $q \in P^1(T)$ so that

$$\|u - q\|_{L^2(T)} + h_T\|\nabla(u - q)\|_{L^2(T)} \leq C\, h_T^2 |u|_{H^2(T)}. \tag{3.29}$$

Using the triangle inequality we get

$$\|\nabla(u - I_h u)\|_{L^2(T)} \leq C(\|\nabla(u - q)\|_{L^2(T)} + \|\nabla(q - I_h u)\|_{L^2(T)}). \tag{3.30}$$

We let $w = (q - I_h u)|_T$ and use (3.27) to get

$$\|\nabla(q - I_h u)\|_{L^2(T)} = \|\nabla w\|_{L^2(T)} \leq C\, \max\{w(z_0), w(z_1), w(z_2)\},$$

where z_i $(i = 0, 1, 2)$ are the vertices of T. However, $w(z_i) = q(z_i) - I_h u(z_i) = q(z_i) - u(z_i)$ by the definition of I_h. Hence,

$$\max\{w(z_0), w(z_1), w(z_2)\} \leq \|u - q\|_{L^\infty(T)}.$$

Therefore, we have

$$\|\nabla(q - I_h u)\|_{L^2(T)} \leq C\, \|u - q\|_{L^\infty(T)}. \tag{3.31}$$

Clearly we have

$$\|u - q\|_{L^\infty(T)} \leq \|\hat{u} - \hat{q}\|_{L^\infty(\hat{T})}.$$

By (3.15) and definition of H^2-norm

$$\|\hat{u} - \hat{q}\|_{L^\infty(\hat{T})} \leq C\|\hat{u} - \hat{q}\|_{H^2(\hat{T})} \leq C(|\hat{u} - \hat{q}|_{H^2(\hat{T})} + |\hat{u} - \hat{q}|_{H^1(\hat{T})} + \|\hat{u} - \hat{q}\|_{L^2(\hat{T})}).$$

Then, combining the above inequalities and using (3.21), (3.22), (3.23) we have

$$\|u - q\|_{L^\infty(T)} \leq C(h_T|u - q|_{H^2(T)} + |u - q|_{H^1(T)} + \frac{1}{h_T}\|u - q\|_{L^2(T)}).$$

If we combine this inequality with this inequality with (3.31) and (3.30) we get

$$\|\nabla(u - I_h u)\|_{L^2(T)} \leq C(h_T |u - q|_{H^2(T)} + |u - q|_{H^1(T)} + \frac{1}{h_T}\|u - q\|_{L^2(T)}).$$

Using (3.29) we get

$$\|\nabla(u - I_h u)\|_{L^2(T)} \leq C\, h_T |u|_{H^2(T)},$$

where we also used that the second derivatives of $q \in P^1(T)$ are zero since. Hence, using (3.28) we get

$$\|\nabla(u - I_h u)\|^2_{L^2(\Omega)} \leq C\, h^2 \sum_{T \in \mathcal{T}_h} |u|^2_{H^2(T)} = Ch^2 |u|^2_{H^2(\Omega)}.$$

The result now follows after taking the square root of both sides.

Bibliography

[1] C. Johnson, *Numerical solution of partial differential equations by the finite element method*, Cambridge University Press, Cambridge, 1987.

[2] D. Gilbarg and N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, 2001, Springer.