Lectures on the Error Analysis of Interpolation on Simplicial Triangulations without the Shape Regularity Assumption and Its Applications to Finite Element Methods

Part 1: Lagrange Interpolation on Triangles

Kenta Kobayashi 小林健太¹ Takuya Tsuchiya 土屋卓也²

 ¹ Graduate School of Business Administration, Hitotsubashi University, Kunitachi, 186-8601, JAPAN, 一橋大学経営管理研究科
 ² Graduate School of Science and Engineering, Ehime University, Matsuyama, 790-8577, JAPAN, 愛媛大学理工学研究科

Abstract

In the error analysis of finite element methods, the shape regularity assumption on triangulations is typically imposed to obtain *a priori* error estimations. In practical computations, however, very "thin" or "degenerated" elements that violate the shape regularity assumption may appear when we use adaptive mesh refinement. In this survey, we attempt to establish an error analysis approach without the shape regularity assumption on triangulations.

We have presented several papers on the error analysis of finite element methods on non-shape regular triangulations. The main points in these papers are that, in the error estimates of finite element methods, the circumradius of the triangles is one of the most important factors.

The purpose of this survey is to provide a simple and plain explanation of the results to researchers and, in particular, graduate students who are interested in the subject. Therefore, this survey is not intended to be a research paper. We hope that, in the near future, it will be merged into a textbook on the mathematical theory of the finite element methods.

Key words: Lagrange interpolation, shape-regurality, maximum angle, finite elements

2021年11月16日受付 2022年1月28日受理

§1. Introduction: Lagrange interpolation on triangles

Lagrange interpolation on triangles and the associated error estimates are important subjects in numerical analysis. In particular, they are crucial in the error analysis of finite element methods. Throughout this survey, $K \subset \mathbb{R}^2$ denotes a triangle with vertices \mathbf{x}_i , i = 1, 2, 3. In this survey, we always assume that triangles are closed sets. Let λ_i be the barycentric coordinates of K with respect to \mathbf{x}_i . By definition, $0 \leq \lambda_i \leq 1$, $\sum_{i=1}^3 \lambda_i = 1$. Let \mathbb{N}_0 be the set of nonnegative integers, and $\gamma = (a_1, a_2, a_3) \in \mathbb{N}_0^3$ be a multi-index. Let k be a positive integer. If $|\gamma| := \sum_{i=1}^{d+1} a_i = k$, then $\gamma/k := (a_1/k, a_2/k, a_3/k)$ can be regarded as a barycentric coordinate in K. The set $\Sigma^k(K)$ of points on K is defined as ¹

(1) $\Sigma^{k}(K) := \left\{ \frac{\gamma}{k} \in K \mid |\gamma| = k, \ \gamma \in \mathbb{N}_{0}^{3} \right\}.$



Figure 1. Set $\Sigma^{k}(K)$, k = 1, k = 2, k = 3.

Let $\mathcal{P}_k(K)$ be a set of polynomials defined on K whose degree is at most k. For a continuous function $v \in C^0(K)$, the kth-order Lagrange interpolation $\mathcal{I}_K^k v \in \mathcal{P}_k(K)$ is defined as

$$v(\mathbf{x}) = (\mathcal{I}_K^k v)(\mathbf{x}), \quad \forall \mathbf{x} \in \Sigma^k(K).$$

To enable the error analysis of Lagrange interpolation, we typically introduce the following condition [8, 6, 10]. Let $h_K := \operatorname{diam} K$ and ρ_K be the diameter of its inscribed circle. Suppose that X is a set of (possibly infinitely many) triangles.

Assumption 1 (Shape regularity). The set X is called **shape regular** if there exists a constant $\sigma > 0$ such that

$$\frac{h_K}{\rho_K} \le \sigma, \qquad \forall K \in X$$

The maximum of the ratio h_K/ρ_K in X is called its **chunkiness parameter** [6]. The shape regularity condition is sometimes also called the **inscribed ball condition**. For more information on the conditions equivalent to shape regularity, see [9].

Let \widehat{K} be a **reference element**. The triangle with vertices $(0,0)^{\top}$, $(1,0)^{\top}$, and $(0,1)^{\top}$ is typically taken as the reference triangle \widehat{K} . Let $\varphi(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ be an affine transformation that maps \widehat{K} to K, where A is a 2 × 2 regular matrix and $\mathbf{b} \in \mathbb{R}^2$.

¹The set $\Sigma^k(K)$ is sometimes called a *stencil*.

Error analysis is first performed on the reference element \widehat{K} . Then, the "pull back" with $v \circ \varphi$ is used to transfer the result obtained on \widehat{K} to the "physical element" K.

Let ||A|| denote the matrix norm of A associated with the Euclidean norm of \mathbb{R}^2 , and let $1 \leq p \leq \infty$. The function $v \in W^{k+1,p}(K)$ is pulled back by φ as $\hat{v} := v \circ \varphi$. Let k and m be integers such that $k \geq 1$ and $0 \leq m \leq k$. The following theorem is standard.

Theorem 2 ([8], Theorem 3.1.4). Let $\sigma > 0$ be a constant. If $h_K/\rho_K \leq \sigma$, then there exists a constant $C = C(\widehat{K}, p, k, m)$ independent of K such that, for $v \in W^{k+1,p}(K)$,

(2)
$$\begin{aligned} |v - \mathcal{I}_{K}^{k}v|_{m,p,K} &\leq C \|A\|^{k+1} \|A^{-1}\|^{m} |v|_{k+1,p,K} \\ &\leq C \frac{h_{K}^{k+1}}{\rho_{K}^{m}} |v|_{k+1,p,K} \leq (C\sigma^{m}) h_{K}^{k+1-m} |v|_{k+1,p,K}. \end{aligned}$$

To derive the second inequality in (2), we use the following lemma.

Lemma 3 ([8], Theorem 3.1.3). We have
$$||A|| \le h_K \rho_{\widehat{K}}^{-1}$$
, $||A^{-1}|| \le h_{\widehat{K}} \rho_K^{-1}$.

Let K be an arbitrary triangle, and $h_K \ge \alpha \ge \beta > 0$ be the lengths of its three edges. Note that $h_K/2 < \alpha \le h_K$. Using translation, rotation, and mirror imaging, K is transformed into a triangle with vertices $\mathbf{x}_1 = (0, 0)^{\top}$, $\mathbf{x}_2 = (\alpha, 0)^{\top}$, and $\mathbf{x}_3 = (\beta s, \beta t)^{\top}$, where $s = \cos \theta$, $t = \sin \theta$, and $0 < \theta < \pi$ is the inner angle of K at \mathbf{x}_1 . This triangle is called the **standard position** of K. By the law of cosines,

$$h_K^2 = \alpha^2 + \beta^2 - 2\alpha\beta\cos\theta$$
 and $\cos\theta = \frac{\beta}{2\alpha} + \frac{\alpha^2 - h_K^2}{2\alpha\beta} \le \frac{\beta}{2\alpha} \le \frac{1}{2}$

Hence, $\pi/3 \le \theta < \pi$.



Figure 2. General triangle K in the standard position. The vertices are $\mathbf{x}_1 = (0, 0)^{\top}$, $\mathbf{x}_2 = (\alpha, 0)^{\top}$, and $\mathbf{x}_3 = (\beta s, \beta t)^{\top}$, where $s^2 + t^2 = 1$, t > 0. We assume that $0 < \beta \le \alpha \le h_K$.

These assumptions imply that the affine transformation φ can be written as $\varphi(\mathbf{x}) = A\mathbf{x}$ with the matrix

(3)
$$A = \begin{pmatrix} \alpha \ \beta s \\ 0 \ \beta t \end{pmatrix}.$$

We set $t = \sin \theta = 1$, for example (i.e., K is a right triangle). Then, s = 0, $||A|| = \alpha$, $||A^{-1}|| = 1/\beta$, and the inequalities in (16) can be rearranged as

(4)
$$|v - \mathcal{I}_K^k v|_{m,p,K} \le C \frac{\alpha^{k+1}}{\beta^m} |v|_{k+1,p,K} \le C \left(\frac{\alpha}{\beta}\right)^m h_K^{k+1-m} |v|_{k+1,p,K}.$$

Thus, we might consider that the ratio α/β should not be too large, or K should not be too "flat." This consideration is expressed as the *minimum angle condition* (Zlámal [28], Ženíšek [27]), which is equivalent to the shape regularity condition for triangles.

Theorem 4 (Minimum angle condition). Let θ_0 , $(0 < \theta_0 \le \pi/3)$ be a constant. If any angle θ of K satisfies $\theta \ge \theta_0$ and $h_K \le 1$, then there exists a constant $C = C(\theta_0)$ independent of h_K such that

$$|v - \mathcal{I}_K^1 v|_{1,2,K} \le Ch_K |v|_{2,2,K}, \qquad \forall v \in H^2(K).$$

However, the minimum angle condition and shape regularity are not necessarily needed to obtain an error estimate. The following condition is well known (Babuška–Aziz [4]).

Theorem 5 (Maximum angle condition). Let θ_1 , $(\pi/3 \leq \theta_1 < \pi)$ be a constant. If any angle θ of K satisfies $\theta \leq \theta_1$ and $h_K \leq 1$, then there exists a constant $C = C(\theta_1)$ that is independent of h_K such that

(5)
$$|v - \mathcal{I}_K^1 v|_{1,2,K} \le Ch_K |v|_{2,2,K}, \quad \forall v \in H^2(K).$$

Křížek [19] introduced the semiregularity condition, which is equivalent to the maximum angle condition (see *Remark* below). Let R_K be the circumradius of K.

Theorem 6 (Semiregularity condition). Let p > 1 and $\sigma > 0$ be a constant. If $R_K/h_K \leq \sigma$ and $h_K \leq 1$, then there exists a constant $C = C(\sigma)$ that is independent of h_K such that

$$|v - \mathcal{I}_K^1 v|_{1,p,K} \le Ch_K |v|_{2,p,K}, \qquad \forall v \in W^{2,p}(K).$$

We mention a few more known results. Jamet [13] presented the following results.

Theorem 7. Let $1 \le p \le \infty$. Let $m \ge 0$, $k \ge 1$ be integers such that k+1-m > 2/p $(1 or <math>k - m \ge 1$ (p = 1). Then, the following estimate holds: $|v - \mathcal{I}_K^k v|_{m,p,K} \le C \frac{h_K^{k+1-m}}{\cos^m \theta_K/2} |v|_{k+1,p,K}, \quad \forall v \in W^{k+1,p}(K),$ where θ_K is the maximum angle of K, and C depends only on k and p. Remark: (1) In Theorem 7, the restriction on p comes from the Sobolev imbedding theorem. Note that in [13, Théorème 3.1] the case p = 1 is not mentioned explicitly but clearly holds for triangles (see Section 2.5). For the case of the maximum angle condition, we set k = m = 1 and find that Jamet's result (Theorem 7) does not imply the estimation (5) because the case p = 2 is excluded.

(2) Let an arbitrary triangle K be in its standard position (Figure 2). Then θ is the maximum internal angle of K, and

(6)
$$\frac{R_K}{h_K} = \frac{1}{2\sin\theta}, \qquad \frac{\pi}{3} \le \theta < \pi$$

by the law of sines. Thus, the dimensionless quantity R_K/h_K represents the maximum internal angle of K, and the boundedness of R_K/h_K , which is the semiregularity of K, is equivalent to the maximum angle condition $\theta \leq \theta_1 < \pi$ with a fixed constant θ_1 . \Box

For further results of the error estimations on "skinny elements", see the monograph by Apel [2].

Recently, Kobayashi, one of the authors, obtained the following epoch-making result [14]. Let A, B, and C be the lengths of the three edges of K and S be the area of K.

Theorem 8 (Kobayashi's formula). We define the constant
$$C(K)$$
 as

$$C(K) := \sqrt{\frac{A^2 B^2 C^2}{16S^2} - \frac{A^2 + B^2 + C^2}{30} - \frac{S^2}{5} \left(\frac{1}{A^2} + \frac{1}{B^2} + \frac{1}{C^2}\right)}.$$
Then the following holds:
 $|v - \mathcal{I}_K^1 v|_{1,2,K} \le C(K) |v|_{2,2,K}, \quad \forall v \in H^2(K).$

Recall that R_K is the circumradius of K and is written as ²

(7)
$$R_K = \frac{ABC}{4S}$$

Then, we immediately realize that $C(K) < R_K$ and obtain a corollary of Kobayashi's formula.

Corollary 9.	For any triangle $K \subset \mathbb{R}^2$, the following estimate holds:
(8)	$ v - \mathcal{I}_K^1 v _{1,2,K} \le R_K v _{2,2,K}, \forall v \in H^2(K).$

This corollary demonstrates that even if the minimum angle is very small or the maximum angle is very close to π , the error $|v - \mathcal{I}_K^1 v|_{1,K}$ converges to 0 if R_K converges

 $^{^2\}mathrm{This}$ formula is proved using the law of sines.

to 0. We consider the isosceles triangle K shown in Figure 3 (left). Using (7), we realize that $R_K = h^{\alpha}/2 + h^{2-\alpha}/8 = \mathcal{O}(h^{2-\alpha})$ ($\alpha \ge 1, h \le 1$). Thus, if $\alpha < 2, R_K \to 0$ as $h \to 0$.

As another example, let $\alpha, \beta \in \mathbb{R}$ satisfy $1 < \alpha < \beta < 1 + \alpha$. We consider the triangle K whose vertices are $(0,0)^{\top}$, $(h,0)^{\top}$, and $(h^{\alpha}, h^{\beta})^{\top}$ (Figure 3 (right)). With (7), it is straightforward to see

$$R_{K} = \frac{h \left(h^{2\alpha} + h^{2\beta}\right)^{1/2} \left((h^{\alpha} - h)^{2} + h^{2\beta}\right)^{1/2}}{2h^{1+\beta}}$$

= $\frac{h^{1+\alpha}}{2h^{\beta}} \left(1 + h^{2\beta-2\alpha}\right)^{1/2} \left(1 + h^{2\alpha-2} - 2h^{\alpha-1} + h^{2\beta-2}\right)^{1/2} = \mathcal{O}(h^{1+\alpha-\beta})$
 $\rho_{K} = \frac{h^{1+\beta}}{h + (h^{2\alpha} + h^{2\beta})^{1/2} + ((h^{\alpha} - h)^{2} + h^{2\beta})^{1/2}}, \quad \frac{h^{\beta}}{3} < \rho_{K} < h^{\beta}.$

Hence, if $h \to 0$, the convergence rates that (2) and (8) yield are $\mathcal{O}(h^{2-\beta})$ and $\mathcal{O}(h^{1+\alpha-\beta})$, respectively. Therefore, (8) obtains a better convergence rate than (2). Moreover, if $\beta \geq 2$, (2) does not yield convergence whereas (8) does. Note that, when $h \to 0$, the maximum angles of K approach to π in both cases.



Figure 3. Examples of triangles that violate the maximum angle condition but satisfy $R_K \to 0$ as $h \to 0$.

Although Kobayashi's formula is remarkable, its proof is long and needs validated numerical computation. We began this research to provide a "paper-and-pencil" proof of (8), and recently reported an error estimation in terms of the circumradius of a triangle [15, 17, 18].

Theorem 10 (Circumradius estimates). Let K be an arbitrary triangle. Then, for the kth-order Lagrange interpolation \mathcal{I}_{K}^{k} on K, the estimation

$$|v - \mathcal{I}_K^k v|_{m,p,K} \le C \left(\frac{R_K}{h_K}\right)^m h_K^{k+1-m} |v|_{k+1,p,K} = C R_K^m h_K^{k+1-2m} |v|_{k+1,p,K}$$

holds for any $v \in W^{k+1,p}(K)$, where the constant C = C(k,m,p) is independent of the geometry of K.

We recall that a general triangle K may be written using the settings in Figure 2.

The essence of the proof of Theorem 10 is that the matrix A in (3) is decomposed as

$$A = \widetilde{A}D_{\alpha\beta}, \qquad \widetilde{A} := \begin{pmatrix} 1 \ s \\ 0 \ t \end{pmatrix}, \quad D_{\alpha\beta} := \begin{pmatrix} \alpha \ 0 \\ 0 \ \beta \end{pmatrix}$$

With this decomposition, the estimate (2) is rearranged as

$$|v - \mathcal{I}_{K}^{k}v|_{m,p,K} \leq C \|\widetilde{A}\|^{k+1} \|\widetilde{A}^{-1}\|^{m} \|D_{\alpha\beta}\|^{k+1} \|D_{\alpha\beta}^{-1}\|^{m} |v|_{k+1,p,K}.$$

As indicated by us [18] and Babuška–Aziz [4], the linear transformation by $D_{\alpha\beta}$ does not reduce the approximation property of Lagrange interpolation, and only \tilde{A} could make it "bad." This means that the term

$$\|D_{\alpha\beta}\|^{k+1}\|D_{\alpha\beta}^{-1}\|^{m} = \frac{(\max\{\alpha,\beta\})^{k+1}}{(\min\{\alpha,\beta\})^{m}} \text{ may be replaced with } C_{1}h_{K}^{k+1-m}$$

Furthermore, $\|\widetilde{A}\|$ and $\|\widetilde{A}^{-1}\|$ (the maximum singular values of \widetilde{A} and \widetilde{A}^{-1}) are bounded using the circumradius R_K and h_K as

$$\|\widetilde{A}\|^{k+1}\|\widetilde{A}^{-1}\|^m \le C_2 \left(\frac{R_K}{h_K}\right)^m, \quad \frac{R_K}{h_K} = \frac{1}{2\sin\theta},$$

where θ is the maximum internal angle of K (see Figure 2 and (6)). We emphasize that the constants C_i (i = 1, 2) only depend on k, m, and p. Note that, by setting t = 1 and $\beta = \alpha^2$ in (2) (and (4)), we realize that, regardless of how much we try to analyze $||A||^{k+1} ||A^{-1}||^m$, we cannot prove Theorem 10. In the sequel of this survey, we will explain the proof of Theorem 10 in detail.

§2. Preliminaries

§2.1. Notation

Let $n \geq 1$ be a positive integer and \mathbb{R}^n be *n*-dimensional Euclidean space. We denote the Euclidean norm of $\mathbf{x} \in \mathbb{R}^n$ by $|\mathbf{x}|$. Let $\mathbb{R}^{n*} := \{l : \mathbb{R}^n \to \mathbb{R} : l \text{ is linear}\}$ be the dual space of \mathbb{R}^n . We always regard $\mathbf{x} \in \mathbb{R}^n$ as a column vector and $\mathbf{a} \in \mathbb{R}^{n*}$ as a row vector. For a matrix A and $\mathbf{x} \in \mathbb{R}^n$, A^{\top} and \mathbf{x}^{\top} denote their transpositions. For matrices $A = (a_{ij})_{i,j=1,\dots,n}$ and $B = (b_{ij})_{i,j=1,\dots,n}$, their Kronecker product $A \otimes B$ is an $n^2 \times n^2$ matrix defined as

$$A \otimes B := \begin{pmatrix} a_{11}B \cdots a_{1n}B \\ \vdots & \vdots \\ a_{n1}B \cdots a_{nn}B \end{pmatrix}.$$

For matrices A_i , $i = 1, \dots, k$, the Kronecker product $A_1 \otimes \dots \otimes A_k$ is defined recursively.

For a differentiable function f with n variables, its gradient $\nabla f = \text{grad} f \in \mathbb{R}^{n*}$ is the row vector defined as

$$\nabla f = \nabla_{\mathbf{x}} f := \left(\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_n}\right), \qquad \mathbf{x} := (x_1, \cdots, x_n)^\top.$$

Let \mathbb{N}_0 be the set of nonnegative integers. For $\delta = (\delta_1, ..., \delta_n) \in (\mathbb{N}_0)^n$, the multiindex ∂^{δ} of partial differentiation (in the sense of distribution) is defined by

$$\partial^{\delta} = \partial^{\delta}_{\mathbf{x}} := \frac{\partial^{|\delta|}}{\partial x_1^{\delta_1} \cdots \partial x_n^{\delta_n}}, \qquad |\delta| := \delta_1 + \cdots + \delta_n$$

For two multi-indices $\eta = (\eta_1, \dots, \eta_n)$, $\delta = (\delta_1, \dots, \delta_n)$, $\eta \leq \delta$ means that $\eta_i \leq \delta_i$ $(i = 1, \dots, n)$. Additionally, $\delta \cdot \eta$ and δ ! are defined as $\delta \cdot \eta := \eta_1 \delta_1 + \dots + \eta_n \delta_n$ and δ ! $:= \delta_1! \cdots \delta_n!$, respectively.

Let $\Omega \subset \mathbb{R}^n$ be a (bounded) domain. The usual Lebesgue space is denoted by $L^p(\Omega)$ for $1 \leq p \leq \infty$. For a positive integer k, the Sobolev space $W^{k,p}(\Omega)$ is defined by $W^{k,p}(\Omega) := \{v \in L^p(\Omega) \mid \partial^{\delta} v \in L^p(\Omega), |\delta| \leq k\}$. For $1 \leq p < \infty$, the norm and semi-norm of $W^{k,p}(\Omega)$ are defined as

$$|v|_{k,p,\Omega} := \left(\sum_{|\delta|=k} |\partial^{\delta} v|_{0,p,\Omega}^{p}\right)^{1/p}, \quad \|v\|_{k,p,\Omega} := \left(\sum_{0 \le m \le k} |v|_{m,p,\Omega}^{p}\right)^{1/p},$$

and $|v|_{k,\infty,\Omega} := \max_{|\delta|=k} \left\{ \operatorname{ess\,sup}_{\mathbf{x}\in\Omega} |\partial^{\delta} v(\mathbf{x})| \right\}, \, \|v\|_{k,\infty,\Omega} := \max_{0 \le m \le k} \{|v|_{m,\infty,\Omega}\}.$

§2.2. Preliminaries from matrix analysis

We introduce some facts from the theory of matrix analysis. For their proofs, refer to textbooks on matrix analysis such as [12] and [26].

Let $n \ge 2$ be an integer and A be an $n \times n$ regular matrix. Note that $A^{\top}A$ is symmetric positive-definite and has n positive eigenvalues $0 < \mu_1 \le \cdots \le \mu_n$. The square roots of μ_i are called the *singular values* of A. Let $\mu_m := \mu_1$ and $\mu_M := \mu_n$ be the minimum and maximum eigenvalues. Then,

$$\mu_m |\mathbf{x}|^2 \le |A\mathbf{x}|^2 \le \mu_M |\mathbf{x}|^2, \quad \mu_M^{-1} |\mathbf{x}|^2 \le |A^{-1}\mathbf{x}|^2 \le \mu_m^{-1} |\mathbf{x}|^2, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

For A, the matrix norm ||A|| with respect to the Euclidean norm is defined by

$$||A|| := \sup_{\mathbf{x} \in \mathbb{R}^n} \frac{|A\mathbf{x}|}{|\mathbf{x}|}.$$

From these definitions, we realize that $||A|| = \mu_M^{1/2}$ and $||A^{-1}|| = \mu_m^{-1/2}$.

For the Kronecker product of matrices, we have the following lemma whose proof is straightforward (see the textbooks mentioned above).

Lemma 11. Let A, B, C, and D be $n \times n$ matrices. Then, the following equations hold:

$$(A \otimes B)(C \otimes D) = (AC \otimes BD), \qquad (A \otimes B)^{\top} = A^{\top} \otimes B^{\top}.$$

Furthermore, if A and B have eigenvalues λ_i and μ_j , $i, j = 1, \dots, n$, respectively, then $\lambda_i \mu_j$ are eigenvalues of $A \otimes B$.

Exercise: Prove Lemma 11.

From Lemma 11, we realize that the minimum and maximum eigenvalues of $(A^{\top}A) \otimes (A^{\top}A) = (A \otimes A)^{\top} (A \otimes A)$ are $0 < \mu_m^2 \leq \mu_M^2$. Hence, for any $\mathbf{w} \in \mathbb{R}^{n^2}$,

$$\mu_m^2 |\mathbf{w}|^2 \le |(A \otimes A)\mathbf{w}|^2 \le \mu_M^2 |\mathbf{w}|^2, \quad \mu_M^{-2} |\mathbf{w}|^2 \le |(A^{-1} \otimes A^{-1})\mathbf{w}|^2 \le \mu_m^{-2} |\mathbf{w}|^2.$$

The above facts can be extended straightforwardly to the case of the higher-order Kronecker product $A \otimes ... \otimes A$. For $A \otimes ... \otimes A$, $A^{-1} \otimes ... \otimes A^{-1}$ (the *k*th Kronecker products), and we have, for $\mathbf{w} \in \mathbb{R}^{n^k}$,

$$\begin{split} \mu_m^k |\mathbf{w}|^2 &\leq |(A \otimes \dots \otimes A)\mathbf{w}|^2 \leq \mu_M^k |\mathbf{w}|^2, \\ \mu_M^{-k} |\mathbf{w}|^2 &\leq |(A^{-1} \otimes \dots \otimes A^{-1})\mathbf{w}|^2 \leq \mu_m^{-k} |\mathbf{w}|^2. \end{split}$$

These inequalities imply that

$$||A \otimes ... \otimes A|| = ||A||^k, \qquad ||A^{-1} \otimes ... \otimes A^{-1}|| = ||A^{-1}||^k.$$

§ 2.3. Useful inequalities

For N positive real numbers $U_1, ..., U_N$, the following inequalities hold:

(9)
$$\sum_{k=1}^{N} U_k^p \le N^{\tau(p)} \left(\sum_{k=1}^{N} U_k^2 \right)^{p/2}, \quad \tau(p) := \begin{cases} 1-p/2, & 1 \le p \le 2\\ 0, & 2 \le p < \infty \end{cases},$$

(10)
$$\left(\sum_{k=1}^{N} U_{k}^{2}\right)^{p/2} \leq N^{\gamma(p)} \sum_{k=1}^{N} U_{k}^{p}, \quad \gamma(p) := \begin{cases} 0, & 1 \leq p \leq 2\\ p/2 - 1, & 2 \leq p < \infty \end{cases}$$

Exercise: Prove the inequalities (9) and (10).

\S 2.4. The affine transformation defined by a regular matrix

Let A be an $n \times n$ matrix with detA > 0. We consider the affine transformation $\varphi(\mathbf{x})$ defined by $\mathbf{y} = \varphi(\mathbf{x}) := A\mathbf{x} + \mathbf{b}$ for $\mathbf{x} = (x_1, \dots, x_n)^\top$, $\mathbf{y} = (y_1, \dots, y_n)^\top$ with $\mathbf{b} \in \mathbb{R}^n$. Suppose that a reference region $\widehat{\Omega} \subset \mathbb{R}^n$ is transformed to a domain Ω by φ ; $\Omega := \varphi(\widehat{\Omega})$. Then, a function $v(\mathbf{y})$ defined on Ω is pulled-back to the function $\hat{v}(\mathbf{x})$ on $\widehat{\Omega}$ as $\hat{v}(\mathbf{x}) := v(\varphi(\mathbf{x})) = v(\mathbf{y})$. Then, we have $\nabla_{\mathbf{x}} \hat{v} = (\nabla_{\mathbf{y}} v)A$, $\nabla_{\mathbf{y}} v = (\nabla_{\mathbf{x}} \hat{v})A^{-1}$, and $|\nabla_{\mathbf{y}} v|^2 = |(\nabla_{\mathbf{x}} \hat{v})A^{-1}|^2 = (\nabla_{\mathbf{x}} \hat{v})A^{-1}A^{-\top}(\nabla_{\mathbf{x}} \hat{v})^\top$.

The Kronecker product $\nabla \otimes \nabla$ of the gradient ∇ is defined by

$$\nabla \otimes \nabla := \left(\frac{\partial}{\partial x_1} \nabla, ..., \frac{\partial}{\partial x_n} \nabla\right) = \left(\frac{\partial^2}{\partial x_1^2}, \frac{\partial^2}{\partial x_1 \partial x_2}, ..., \frac{\partial^2}{\partial x_{n-1} \partial x_n}, \frac{\partial^2}{\partial x_n^2}\right).$$

We regard $\nabla \otimes \nabla$ to be a row vector. From this definition, it follows that

$$\sum_{|\delta|=2} (\partial^{\delta} v)^2 = \sum_{i,j=1}^n \left(\frac{\partial^2 v}{\partial x_i \partial x_j} \right)^2 = |(\nabla \otimes \nabla) v|^2$$

and $(\nabla_{\mathbf{x}} \otimes \nabla_{\mathbf{x}})\hat{v} = ((\nabla_{\mathbf{y}} \otimes \nabla_{\mathbf{y}})v) (A \otimes A), (\nabla_{\mathbf{y}} \otimes \nabla_{\mathbf{y}})v = ((\nabla_{\mathbf{x}} \otimes \nabla_{\mathbf{x}})\hat{v}) (A^{-1} \otimes A^{-1}).$ Thus, we have $||A||^{-2} |\nabla_{\mathbf{x}}\hat{v}|^2 \le |\nabla_{\mathbf{y}}v|^2 \le ||A^{-1}||^2 |\nabla_{\mathbf{x}}\hat{v}|^2$ and

$$\begin{split} \sum_{|\delta|=2} (\partial_{\mathbf{y}} v)^2 &= |(\nabla_{\mathbf{y}} \otimes \nabla_{\mathbf{y}}) v|^2 \\ &= ((\nabla_{\mathbf{x}} \otimes \nabla_{\mathbf{x}}) \hat{v}) \left(A^{-1} \otimes A^{-1}\right) (A^{-1} \otimes A^{-1})^\top \left((\nabla_{\mathbf{x}} \otimes \nabla_{\mathbf{x}}) \hat{v}\right)^\top \\ &= ((\nabla_{\mathbf{x}} \otimes \nabla_{\mathbf{x}}) \hat{v}) \left(A^{-1} A^{-\top} \otimes A^{-1} A^{-\top}\right) \left((\nabla_{\mathbf{x}} \otimes \nabla_{\mathbf{x}}) \hat{v}\right)^\top, \\ \|A\|^{-2} \sum_{|\delta|=2} (\partial_{\mathbf{x}}^{\delta} \hat{v})^2 &\leq \sum_{|\delta|=2} (\partial_{\mathbf{y}}^{\delta} v)^2 \leq \|A^{-1}\|^2 \sum_{|\delta|=2} (\partial_{\mathbf{x}}^{\delta} \hat{v})^2. \end{split}$$

The above inequalities can be easily extended to higher-order derivatives, and we obtain the following inequalities: for $k \ge 1$,

(11)
$$\begin{aligned} \|A\|^{-2k} \sum_{|\delta|=k} (\partial_{\mathbf{x}}^{\delta} \hat{v})^2 &\leq \sum_{|\delta|=k} (\partial_{\mathbf{y}}^{\delta} v)^2 \leq \|A^{-1}\|^{2k} \sum_{|\delta|=k} (\partial_{\mathbf{x}}^{\delta} \hat{v})^2, \\ &\det A|^{1/2} \|A\|^{-k} |\hat{v}|_{k,2,\widehat{\Omega}} \leq |v|_{k,2,\Omega} \leq |\det A|^{1/2} \|A^{-1}\|^k |\hat{v}|_{k,2,\widehat{\Omega}}. \end{aligned}$$

Using the inequalities (9) and (10), we can extend (11) for the case of arbitrary p, $1 \le p < \infty$:

$$\begin{split} |v|_{k,p,\Omega}^{p} &= \int_{\Omega} \sum_{|\delta|=k} |\partial_{\mathbf{y}}^{\delta} v(\mathbf{y})|^{p} \mathrm{d}\mathbf{y} \leq n^{k\tau(p)} \int_{\Omega} \left(\sum_{|\delta|=k} |\partial_{\mathbf{y}}^{\delta} v(\mathbf{y})|^{2} \right)^{p/2} \mathrm{d}\mathbf{y} \\ &\leq n^{k\tau(p)} \|A^{-1}\|^{kp} \int_{\Omega} \left(\sum_{|\delta|=k} |\partial_{\mathbf{x}}^{\delta} \hat{v}(\mathbf{x})|^{2} \right)^{p/2} \mathrm{d}\mathbf{y} \\ &= n^{k\tau(p)} |\det A| \|A^{-1}\|^{kp} \int_{\widehat{\Omega}} \left(\sum_{|\delta|=k} |\partial_{\mathbf{x}}^{\delta} \hat{v}(\mathbf{x})|^{2} \right)^{p/2} \mathrm{d}\mathbf{x} \\ &\leq n^{k(\tau(p)+\gamma(p))} |\det A| \|A^{-1}\|^{kp} \int_{\widehat{\Omega}} \sum_{|\delta|=k} |\partial_{\mathbf{x}}^{\delta} \hat{v}(\mathbf{x})|^{p} \mathrm{d}\mathbf{x} \\ &= n^{k(\tau(p)+\gamma(p))} |\det A| \|A^{-1}\|^{kp} |\hat{v}|_{k,p,\widehat{\Omega}}^{p} \end{split}$$

and

$$\begin{aligned} |v|_{k,p,\Omega}^{p} &= \int_{\Omega} \sum_{|\delta|=k} |\partial_{\mathbf{y}}^{\delta} v(\mathbf{y})|^{p} \mathrm{d}\mathbf{y} \geq n^{-k\gamma(p)} \int_{\Omega} \left(\sum_{|\delta|=k} |\partial_{\mathbf{y}}^{\delta} v(\mathbf{y})|^{2} \right)^{p/2} \mathrm{d}\mathbf{y} \\ &\geq n^{-k\gamma(p)} |\det A| ||A||^{-kp} \int_{\Omega} \left(\sum_{|\delta|=k} |\partial_{\mathbf{x}}^{\delta} \hat{v}(\mathbf{x})|^{2} \right)^{p/2} \mathrm{d}\mathbf{y} \end{aligned}$$

$$= n^{-k\gamma(p)} |\det A| ||A||^{-kp} \int_{\widehat{\Omega}} \left(\sum_{|\delta|=k} |\partial_{\mathbf{x}}^{\delta} \hat{v}(\mathbf{x})|^2 \right)^{p/2} d\mathbf{x}$$

$$\geq n^{-k(\tau(p)+\gamma(p))} |\det A| ||A||^{-kp} \int_{\widehat{\Omega}} \sum_{|\delta|=k} |\partial_{\mathbf{x}}^{\delta} \hat{v}(\mathbf{x})|^p d\mathbf{x}$$

$$= n^{-k(\tau(p)+\gamma(p))} |\det A| ||A||^{-kp} |\hat{v}|_{k,p,\widehat{\Omega}}^p,$$

where we use the fact that $|v|_{k,2,\Omega}$ contains n^k terms. Therefore, we obtain the following lemma:

Lemma 12. In the above setting of the linear transformation, we have $(12)n^{-k\mu(p)} |\det A|^{1/p} ||A||^{-k} |\hat{v}|_{k,p,\widehat{\Omega}} \leq |v|_{k,p,\Omega} \leq n^{k\mu(p)} |\det A|^{1/p} ||A^{-1}||^k |\hat{v}|_{k,p,\widehat{\Omega}}.$ where $\mu(p) := \frac{\tau(p) + \gamma(p)}{p} = \begin{cases} 1/p - 1/2, & 1 \leq p \leq 2\\ 1/2 - 1/p, & 2 \leq p \leq \infty \end{cases}.$

Proof: We only need to prove the case of $p = \infty$, and it is done just by letting $p \to \infty$ in (12). \Box

Let us apply (12) to the case $A \in O(n)$, where O(n) is the set of orthogonal matrices. That is, $A^{\top}A = AA^{\top} = I_n$. In this case, $|\det A| = ||A|| = ||A^{-1}|| = 1$. Thus, we have

(13)
$$n^{-k\mu(p)} |\hat{v}|_{k,p,\widehat{\Omega}} \le |v|_{k,2,\Omega} \le n^{k\mu(p)} |\hat{v}|_{k,p,\widehat{\Omega}}.$$

Those inequalities mean that, if p = 2, the Sobolev norms $|v|_{k,2,\Omega}$ are not affected by rotations. If $p \neq 2$, however, they are affected by rotations up to the constants $n^{-k\mu(p)}$ and $n^{k\mu(p)}$.

$\S 2.5.$ The Sobolev imbedding theorem

If 1 , Sobolev's imbedding theorem and Morrey's inequality imply that

$$\begin{split} W^{2,p}(K) &\subset C^{1,1-2/p}(K), \quad p > 2, \\ H^2(K) &\subset W^{1,q}(K) \subset C^{0,1-2/q}(K), \quad \forall q > 2, \\ W^{2,p}(K) &\subset W^{1,2p/(2-p)}(K) \subset C^{0,2(p-1)/p}(K), \quad 1$$

For proofs of the Sobolev imbedding theorems, see [1] and [7]. For the case p = 1, we still have the continuous imbedding $W^{2,1}(K) \subset C^0(K)$. For proof of the critical imbedding, see [1, Theorem 4.12] and [6, Lemma 4.3.4].

§2.6. Gagliardo–Nirenberg's inequality

Theorem 13 (Gagliardo–Nirenberg's inequality). Let $1 \le p \le \infty$. Let k, m be integers such that $k \ge 2$ Then, for $\alpha := m/k, \ 0 < \alpha < 1$, the following inequality holds: $|v|_{m,p,\mathbb{R}^n} \le C |v|_{0,p,\mathbb{R}^n}^{1-\alpha} |v|_{k,p,\mathbb{R}^n}^{\alpha}, \quad \forall v \in W^{k,p}(\mathbb{R}^n),$

where the constant C depends only on k, m, p, and n.

For the proof and the general cases of Galliardo–Nirenberg's inequality, see [7] and the references therein.

§2.7. A standard error analysis of Lagrange interpolation

In this subsection, we explain a standard error analysis of Lagrange interpolation. First, we prepare a theorem from Ciarlet[8]. Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with the Lipschitz boundary $\partial \Omega$. Let k be a positive integer and p be a real with $1 \leq p \leq \infty$. We consider the quotient space $W^{k+1,p}(\Omega)/\mathcal{P}_k(\Omega)$. As usual, we introduce the following norm to the space:

$$\begin{aligned} \|\dot{v}\|_{k+1,p,\Omega} &:= \inf_{q \in \mathcal{P}_k(\Omega)} \|v + q\|_{k+1,p,\Omega}, \quad \forall \dot{v} \in W^{k+1,p}(\Omega) / \mathcal{P}_k(\Omega), \\ \dot{v} &:= \left\{ w \in W^{k+1,p}(\Omega) \mid w - v \in \mathcal{P}_k(\Omega) \right\}. \end{aligned}$$

We also define the seminorm of the space by $|\dot{v}|_{k+1,p,\Omega} := |v|_{k+1,p,\Omega}$. Take an arbitrary $q \in \mathcal{P}_k(\Omega)$. If $1 \leq p < \infty$, we have

$$||v+q||_{k+1,p,\Omega}^p = |v|_{k+1,p,\Omega}^p + ||v+q||_{k,p,\Omega}^p \ge |v|_{k+1,p,\Omega}^p,$$

and if $p = \infty$, we have

$$||v+q||_{k+1,\infty,\Omega} = \max\{|v|_{k+1,\infty,\Omega}, ||v+q||_{k,\infty,\Omega}\} \ge |v|_{k+1,\infty,\Omega}.$$

Thus the following inequality follows:

$$|\dot{v}|_{k+1,p,\Omega} \le ||\dot{v}||_{k+1,p,\Omega}, \quad \forall \dot{v} \in W^{k+1,p}(\Omega)/\mathcal{P}_k(\Omega).$$

The next theorem claims the seminorm is actually a norm of $W^{k+1,p}(\Omega)/\mathcal{P}_k(\Omega)$.

Theorem 14 (Ciarlet[8], Theorem 3.1.1). There exists a positive constant $C(\Omega)$ depending only on $k, p \in [1, \infty]$, and Ω , such that the following estimations hold:

$$\|\dot{v}\|_{k+1,p,\Omega} \leq C(\Omega)|\dot{v}|_{k+1,p,\Omega}, \qquad \forall \dot{v} \in W^{k+1,p}(\Omega)/\mathcal{P}_k(\Omega),$$

(14)
$$\inf_{q \in \mathcal{P}_k(\Omega)} \|v+q\|_{k+1,p,\Omega} \leq C(\Omega)|v|_{k+1,p,\Omega}, \qquad \forall v \in W^{k+1,p}(\Omega).$$

Proof: Let N be the dimension of $\mathcal{P}_k(\Omega)$ as a vector space, and $\{q_i\}_{i=1}^N$ be its basis and $\{f_i\}_{i=1}^N$ be the dual basis of $\{q_i\}$. That is, $f_i \in \mathcal{L}(\mathcal{P}_k(\Omega), \mathbb{R})$ and they satisfy $f_i(q_j) = \delta_{ij}$,

 $i, j = 1, \dots, N$ (δ_{ij} are Kronecker's deltas). By Hahn-Banach's theorem, f_i is extended to $f_i \in \mathcal{L}(W^{k+1,p}(\Omega), \mathbb{R})$. For $q \in \mathcal{P}_k(\Omega)$, we have

$$q = 0 \iff f_i(q) = 0, \quad 1 \le i \le N.$$

Now, we claim that there exists a constant $C(\Omega)$ such that

(15)
$$||v||_{k+1,p,\Omega} \le C(\Omega) \left(|v|_{k+1,p,\Omega} + \sum_{i=1}^{N} |f_i(v)| \right), \quad \forall v \in W^{k+1,p}(\Omega).$$

Suppose that (15) holds. For given $v \in W^{k+1,p}(\Omega)$, let $q \in \mathcal{P}_k(\Omega)$ be defined with the extended $f_i \in \mathcal{L}(W^{k+1,p}(\Omega), \mathbb{R})$ by

$$q = \sum_{i=1}^{N} \lambda_i q_i, \quad \lambda_i := -f_i(v), \quad i = 1, \cdots, N.$$

Then, we have $f_i(v+q) = 0$, $i = 1, \dots, N$. Therefore, The inequality (14) follows from (15).

We now show the inequality (15) by contradiction. Assume that (15) does not hold. Then, there exists a sequence $\{v_l\}_{l=1}^{\infty} \subset W^{k+1,p}(\Omega)$ such that

$$||v_l||_{k+1,p,\Omega} = 1, \ \forall l, \qquad \lim_{l \to \infty} \left(|v_l|_{k+1,p,\Omega} + \sum_{i=1}^N |f_i(v_l)| \right) = 0.$$

By the compactness of the inclusion $W^{k+1,p}(\Omega) \subset W^{k,p}(\Omega)$, there exists a subsequence $\{v_{l_m}\}$ and $v \in W^{k,p}(\Omega)$ such that

$$\lim_{l_m \to \infty} \|v_{l_m} - v\|_{k,p,\Omega} = 0, \qquad \lim_{l_m \to \infty} |v_{l_m}|_{k+1,p,\Omega} = 0.$$

Here, $\{v_{l_m}\}$ is a Cauchy sequence in $W^{k,p}(\Omega)$. We show that it is also a Cauchy sequence in $W^{k+1,p}(\Omega)$ as well. If, for example, $1 \leq p < \infty$, we have

$$\lim_{l_m, l_n \to \infty} \|v_{l_m} - v_{l_n}\|_{k+1, p, \Omega} = \lim_{l_m, l_n \to \infty} \left(\|v_{l_m} - v_{l_n}\|_{k, p, \Omega}^p + |v_{l_m} - v_{l_n}|_{k+1, p, \Omega}^p \right)^{1/p} \\ \leq \lim_{l_m, l_n \to \infty} \left(\|v_{l_m} - v_{l_n}\|_{k, p, \Omega}^p + 2^{p-1} (|v_{l_m}|_{k+1, p, \Omega}^p + |v_{l_n}|_{k+1, p, \Omega}^p) \right)^{1/p} \\ = 0.$$

The case for $p = \infty$ is similarly shown. Hence, v belong s to $W^{k+1,p}(\Omega)$, and $\{v_{l_m}\}$ satisfies

 $\lim_{l_m \to \infty} \|v_{l_m} - v\|_{k+1, p, \Omega} = 0.$

This $v \in W^{k+1,p}(\Omega)$ satisfies

$$|\partial^{\beta} v|_{0,p,\Omega} = \lim_{l_k \to 0} |\partial^{\beta} v_{l_k}|_{0,p,\Omega} = 0, \quad \forall \beta, \quad |\beta| = k+1,$$

and thus $v \in \mathcal{P}_k(\Omega)$. Therefore, because

$$\sum_{i=1}^{N} |f_i(v)| = \lim_{l_m \to \infty} \sum_{i=1}^{N} |f_i(v_{l_m})| = 0,$$

we conclude v = 0. However, this contradicts to $||v||_{k+1,p,\Omega} = \lim_{l_m \to \infty} ||v_{l_m}||_{k+1,p,\Omega} = 1$.

We are now ready to prove the first inequality in Theorem 2. Recall that \widehat{K} is the reference triangle and K is mapped as $K = \varphi(\widehat{K})$ with $\varphi(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$.

Theorem 15. Suppose that $||A^{-1}|| \ge 1$. Then, there exists a constant $C = C(\widehat{K}, p, k, m)$ independent of K such that (16) $||v - \mathcal{I}_K^k v||_{m,p,K} \le C ||A||^{k+1} ||A^{-1}||^m |v|_{k+1,p,K}, \quad \forall v \in W^{k+1,p}(K).$

Proof: Note that, for arbitrary $\hat{v} \in W^{k+1,p}(\widehat{K})$ and $\hat{p} \in \mathcal{P}_k(\widehat{K})$), we have

$$\hat{v} - \mathcal{I}_{\widehat{K}}^k \hat{v} = (I - \mathcal{I}_{\widehat{K}}^k \hat{v})(\hat{v} + \hat{p}),$$

where $I: W^{k+1,p}(\widehat{K}) \to W^{m,p}(\widehat{K})$ is the identity mapping, which is obviously continuous. Therefore, it follows from (14) that

$$\begin{split} \left\| \hat{v} - \mathcal{I}_{\widehat{K}}^{k} \hat{v} \right\|_{m,p,\widehat{K}} &\leq \left\| I - \mathcal{I}_{\widehat{K}}^{k} \right\|_{\mathcal{L}(W^{k+1,p}(\widehat{K}),W^{m,p}(\widehat{K}))} \inf_{\hat{p} \in \mathcal{P}_{k}(\widehat{K})} \left\| \hat{v} + \hat{p} \right\|_{k+1,p,\widehat{K}} \\ &\leq C_{1} \left| \hat{v} \right|_{k+1,p,\widehat{K}}, \end{split}$$

where the constant C_1 depends on \widehat{K} , m, k, p, (and $\mathcal{I}_{\widehat{K}}^k$).

Note that the mapping between $W^{n,p}(K)$ and $W^{n,p}(\widehat{K})$ (n = m or n = k + 1)defined by the pull-back $\hat{v} = v \circ \varphi$ is an isomorphism. By (12), we have

$$\begin{split} \left\| v - \mathcal{I}_{K}^{k} v \right\|_{m,p,K} &= \left(\sum_{l=0}^{m} \left| v - \mathcal{I}_{K}^{k} v \right|_{l,p,K}^{p} \right)^{1/p} \\ &\leq \left(\sum_{l=0}^{m} n^{l\mu(p)} |\det A| \|A^{-1}\|^{lp} \left| \hat{v} - \mathcal{I}_{\widehat{K}}^{k} \hat{v} \right|_{l,p,\widehat{K}}^{p} \right)^{1/p}, \\ &\leq n^{m\mu(p)} |\det A|^{1/p} \|A^{-1}\|^{m} \left\| \hat{v} - \mathcal{I}_{\widehat{K}}^{k} \hat{v} \right\|_{m,p,\widehat{K}}, \\ &|\hat{v}|_{k+1,p,K} \leq n^{(k+1)\mu(p)} |\det A|^{-1/p} \|A\|^{k+1} \left| v \right|_{k+1,p,\widehat{K}}, \end{split}$$

because of the assumption $||A^{-1}|| \ge 1$. Combining these inequalities, the proof is completed with $C := n^{(k+1+m)\mu(p)}C_1$. \Box

Combining these propositions with Lemma 3, we see that, for arbitrary $v \in W^{k+1,p}(K)$,

$$\begin{aligned} \|v - \mathcal{I}_{K}^{k}v\|_{m,p,K} &\leq C \|A\|^{k+1} \|A^{-1}\|^{m} |v|_{k+1,p,K} \leq C \left(\frac{h_{K}}{\rho_{\widehat{K}}}\right)^{k+1} \left(\frac{h_{\widehat{K}}}{\rho_{K}}\right)^{m} |v|_{k+1,p,K} \\ &\leq C \frac{h_{\widehat{K}}^{m}}{\rho_{\widehat{K}}^{k+1}} \frac{h_{K}^{k+1}}{\rho_{K}^{m}} |v|_{k+1,p,K}. \end{aligned}$$

If there exists a constant σ such that $h_K/\rho_K \leq \sigma$, then $\rho_K^{-1} \leq \sigma h_K^{-1}$, and we obtain the following standard error estimation.

Theorem 16. Let $K \subset \mathbb{R}^2$ be a triangle with $h_K \leq 1$. Suppose that $h_K/\rho_K \leq \sigma$, where σ is a positive constant. Then, there exists a constant $C = C(\hat{K}, p, k, m, \sigma)$ independent of K such that (17) $\|v - \mathcal{I}_K^k v\|_{m,p,K} \leq Ch_K^{k+1-m} |v|_{k+1,p,K}, \quad \forall v \in W^{k+1,p}(K).$

§3. Babuška–Aziz's technique

In the previous section, we have proved the standard error estimation (16), (17). To improve them, we introduce the technique given by Babuška–Aziz [4].

Let \widehat{K} be the reference triangle with the vertices $(0,0)^{\top}$, $(1,0)^{\top}$, and $(0,1)^{\top}$. For \widehat{K} , the sets $\Xi_p^i \subset W^{1,p}(\widehat{K})$, $i = 1, 2, p \in [1,\infty]$ are defined by

$$\begin{split} &\Xi_p^{(1,0),1} := \left\{ v \in W^{1,p}(\widehat{K}) \ \Big| \ \int_0^1 v(s,0) \mathrm{d}s = 0 \right\}, \\ &\Xi_p^{(0,1),1} := \left\{ v \in W^{1,p}(\widehat{K}) \ \Big| \ \int_0^1 v(0,s) \mathrm{d}s = 0 \right\}. \end{split}$$

The constant A_p is then defined by

$$A_p := \sup_{v \in \Xi_p^{(1,0),1}} \frac{|v|_{0,p,\widehat{K}}}{|v|_{1,p,\widehat{K}}} = \sup_{v \in \Xi_p^{(0,1),1}} \frac{|v|_{0,p,\widehat{K}}}{|v|_{1,p,\widehat{K}}}, \qquad 1 \le p \le \infty$$

The second equation in the above definition follows from the symmetry of \widehat{K} . The constant A_p (and its reciprocal $1/A_p$) is called the **Babuška–Aziz constant** for $p \in [1, \infty]$. According to Liu–Kikuchi [22], A_2 is the maximum positive solution of the equation $1/x + \tan(1/x) = 0$, and $A_2 \approx 0.49291$.

In the following, we show that $A_p < \infty$ (Babuška–Aziz [4, Lemma 2.1] and Kobayashi–Tsuchiya [15, Lemma 1]).

Lemma 17. We have $A_p < \infty$, $p \in [1, \infty]$.

Proof: The proof is by contradiction. Assume that $A_p = \infty$. Then, there exists a sequence $\{u_k\}_{i=1}^{\infty} \subset \Xi_p^{(1,0),1}$ such that

$$|u_k|_{0,p,\widehat{K}} = 1, \qquad \lim_{k \to \infty} |u_k|_{1,p,\widehat{K}} = 0.$$

From the inequality (14), for an arbitrary $\varepsilon > 0$, there exists a sequence $\{q_k\} \subset \mathcal{P}_0(\widehat{K})$ such that

$$\inf_{q \in \mathcal{P}_{0}(\widehat{K})} \|u_{k} + q\|_{1,p,\widehat{K}} \leq \|u_{k} + q_{k}\|_{1,p,\widehat{K}} \leq \inf_{q \in \mathcal{P}_{0}(\widehat{K})} \|u_{k} + q\|_{1,p,\widehat{K}} + \frac{\varepsilon}{k} \leq C |u_{k}|_{1,p,\widehat{K}} + \frac{\varepsilon}{k},$$
$$\lim_{k \to \infty} \|u_{k} + q_{k}\|_{1,p,\widehat{K}} = 0.$$

Since the sequence $\{u_k\} \subset W^{1,p}(\widehat{K})$ is bounded, $\{q_k\} \subset \mathcal{P}_0(\widehat{K}) = \mathbb{R}$ is also bounded. Therefore, there exists a subsequence $\{q_{k_i}\}$ such that q_{k_i} converges to $\overline{q} \in \mathcal{P}_0(\widehat{K})$. Thus, in particular, we have

$$\lim_{k_i \to \infty} \|u_{k_i} + \bar{q}\|_{1,p,\hat{K}} = 0.$$

Let Γ be the edge of \widehat{K} connecting $(1,0)^{\top}$ and $(0,0)^{\top}$ and $\gamma: W^{1,p}(\widehat{K}) \to W^{1-1/p,p}(\Gamma)$ be the trace operator. The continuity of γ and the inclusion $W^{1-1/p,p}(\Gamma) \subset L^1(\Gamma)$ yield

$$0 = \lim_{k_i \to \infty} \int_{\Gamma} \gamma(u_{k_i} + \bar{q}) \mathrm{d}s = \int_{\Gamma} \bar{q} \mathrm{d}s$$

because $u_{k_i} \in \Xi_p^1$. Thus, we find that $\bar{q} = 0$ and $\lim_{k_i \to \infty} \|u_{k_i}\|_{1,p,\hat{K}} = 0$. This contradicts $\lim_{k_i \to \infty} \|u_{k_i}\|_{1,p,\hat{K}} \ge \lim_{k_i \to \infty} |u_{k_i}|_{0,p,\hat{K}} = 1$. \Box

We define the bijective linear transformation $F_{\alpha\beta}: \mathbb{R}^2 \to \mathbb{R}^2$ by

$$(x^*, y^*)^\top = (\alpha x, \beta y)^\top, \qquad (x, y)^\top \in \mathbb{R}^2, \quad \alpha, \beta > 0.$$

The map $F_{\alpha\beta}$ is called the squeezing transformation.

Now, we consider the "squeezed" triangle $K_{\alpha\beta} := F_{\alpha\beta}(\widehat{K})$. Take an arbitrary $v \in W^{2,p}(K_{\alpha\beta})$, and pull-back v to $u := v \circ F_{\alpha\beta} \in W^{2,p}(\widehat{K})$. For, $p, 1 \leq p < \infty$, we have

(18)
$$\frac{|v|_{0,p,K_{\alpha\beta}}^{p}}{|v|_{1,p,K_{\alpha\beta}}^{p}} = \frac{|u|_{0,p,\widehat{K}}^{p}}{\frac{1}{\alpha^{p}}|u_{x}|_{0,p,\widehat{K}}^{p} + \frac{1}{\beta^{p}}|u_{y}|_{0,p,\widehat{K}}^{p}}$$

(19)
$$\frac{|v|_{0,p,K_{\alpha\beta}}^{p}}{|v|_{2,p,K_{\alpha\beta}}^{p}} = \frac{|u|_{0,p,\widehat{K}}^{p}}{\frac{1}{\alpha^{2p}}|u_{xx}|_{0,p,\widehat{K}}^{p} + \frac{2}{\alpha^{p}\beta^{p}}|u_{xy}|_{0,p,\widehat{K}}^{p} + \frac{1}{\beta^{2p}}|u_{yy}|_{0,p,\widehat{K}}^{p}},$$

(20)
$$|v|_{2,p,K_{\alpha\beta}} = \frac{1}{\alpha^{2p}} |u_{xx}|_{0,p,\widehat{K}}^{2} + \frac{1}{\alpha^{p}\beta^{p}} |u_{xy}|_{0,p,\widehat{K}}^{2} + \frac{1}{\beta^{2p}} |u_{yy}|_{0,p,\widehat{K}}^{2} + \frac{1}{\beta^{p}} |u_{y}|_{0,p,\widehat{K}}^{p} + \frac{1}{\beta^{p}}$$

(20)
$$\frac{1}{|v|_{2,p,K_{\alpha\beta}}^{p}} = \frac{1}{\frac{1}{\alpha^{2p}}|u_{xx}|_{0,p,\widehat{K}}^{p} + \frac{2}{\alpha^{p}\beta^{p}}|u_{xy}|_{0,p,\widehat{K}}^{p} + \frac{1}{\beta^{2p}}|u_{yy}|_{0,p,\widehat{K}}^{p}}}.$$

In the following we explain how these equations are derived.

Note that, for $(x, y)^{\top} \in \widehat{K}$ and $(x^*, y^*)^{\top} = (\alpha x, \beta y)^{\top} \in K_{\alpha\beta}$, we have $u_x = \alpha v_{x^*}, \qquad u_y = \beta v_{y^*},$

and

$$|v_{x^*}|_{0,p,K_{\alpha\beta}}^p = \int_{K_{\alpha\beta}} |v_{x^*}|^p \mathrm{d}\mathbf{x}^* = \frac{1}{\alpha^p} \int_{K_{\alpha\beta}} |u_x|^p \mathrm{d}\mathbf{x}^* = \frac{\beta}{\alpha^{p-1}} \int_{\widehat{K}} |u_x|^p \mathrm{d}\mathbf{x} = \frac{\beta}{\alpha^{p-1}} |u_x|_{0,p,\widehat{K}}^p.$$

Here, $d\mathbf{x} := dxdy$, $d\mathbf{x}^* := dx^*dy^*$, and used the fact $det(DF_{\alpha\beta}) = \alpha\beta$, where $DF_{\alpha\beta}$ is the Jacobian matrix of $F_{\alpha\beta}$. Similarly, we obtain

$$|v_{y^*}|_{0,p,K_{\alpha\beta}}^p = \frac{\alpha}{\beta^{p-1}} |u_y|_{0,p,\widehat{K}}^p, \qquad |v|_{0,p,K_{\alpha\beta}}^p = \alpha\beta |u|_{0,p,\widehat{K}}^p.$$

Therefore, these equations yield (18):

$$\frac{|v|_{0,p,K_{\alpha\beta}}^p}{|v|_{1,p,K_{\alpha\beta}}^p} = \frac{\alpha\beta|u|_{0,p,\widehat{K}}^p}{\frac{\beta}{\alpha^{p-1}}|u_x|_{0,p,\widehat{K}}^p + \frac{\alpha}{\beta^{p-1}}|u_y|_{0,p,\widehat{K}}^p} = \frac{|u|_{0,p,\widehat{K}}^p}{\frac{1}{\alpha^p}|u_x|_{0,p,\widehat{K}}^p + \frac{1}{\beta^p}|u_y|_{0,p,\widehat{K}}^p}$$

Similarly, the equations

$$\begin{aligned} |v_{x^*x^*}|^p_{0,p,K_{\alpha\beta}} &= \frac{\beta}{\alpha^{2p-1}} |u_{xx}|^p_{0,p,\widehat{K}}, \\ |v_{x^*y^*}|^p_{0,p,K_{\alpha\beta}} &= \frac{1}{\alpha^{p-1}\beta^{p-1}} |u_{xy}|^p_{0,p,\widehat{K}}, \\ |v_{y^*y^*}|^p_{0,p,K_{\alpha\beta}} &= \frac{\alpha}{\beta^{2p-1}} |u_{yy}|^p_{0,p,\widehat{K}} \end{aligned}$$

are obtained and yield (19) and (20) as

$$\begin{aligned} \frac{|v|_{0,p,K_{\alpha\beta}}^{p}}{|v|_{2,p,K_{\alpha\beta}}^{p}} &= \frac{\alpha\beta|u|_{0,p,\hat{K}}^{p}}{\frac{\beta}{\alpha^{2p-1}}|u_{xx}|_{0,p,\hat{K}}^{p} + \frac{2}{\alpha^{p-1}\beta^{p-1}}|u_{xy}|_{0,p,\hat{K}}^{p} + \frac{\alpha}{\beta^{2p-1}}|u_{yy}|_{0,p,\hat{K}}^{p}}} \\ &= \frac{|u|_{0,p,\hat{K}}^{p}}{\frac{1}{\alpha^{2p}}|u_{xx}|_{0,p,\hat{K}}^{p} + \frac{2}{\alpha^{p}\beta^{p}}|u_{xy}|_{0,p,\hat{K}}^{p} + \frac{1}{\beta^{2p}}|u_{yy}|_{0,p,\hat{K}}^{p}}, \\ \frac{|v|_{1,p,K_{\alpha\beta}}^{p}}{|v|_{2,p,K_{\alpha\beta}}^{p}} &= \frac{\frac{\beta}{\alpha^{p-1}}|u_{x}|_{0,p,\hat{K}}^{p} + \frac{\alpha}{\beta^{2p-1}}|u_{y}|_{0,p,\hat{K}}^{p}}{\frac{\beta}{\alpha^{2p-1}}|u_{xx}|_{0,p,\hat{K}}^{p} + \frac{\alpha}{\beta^{2p-1}}|u_{yy}|_{0,p,\hat{K}}^{p} + \frac{2}{\alpha^{p-1}\beta^{p-1}}|u_{xy}|_{0,p,\hat{K}}^{p}} \\ &= \frac{\frac{1}{\alpha^{p}}|u_{x}|_{0,p,\hat{K}}^{p} + \frac{\alpha}{\beta^{2p-1}}|u_{yy}|_{0,p,\hat{K}}^{p} + \frac{2}{\alpha^{p-1}\beta^{p-1}}|u_{yy}|_{0,p,\hat{K}}^{p}}{\frac{1}{\alpha^{2p}}|u_{xx}|_{0,p,\hat{K}}^{p} + \frac{2}{\alpha^{p}\beta^{p}}|u_{xy}|_{0,p,\hat{K}}^{p} + \frac{1}{\beta^{2p}}|u_{yy}|_{0,p,\hat{K}}^{p}}. \end{aligned}$$

Next, let $p = \infty$. Then, we have

$$|v|_{0,\infty,K_{\alpha\beta}} = |u|_{0,\infty,\widehat{K}}, \qquad |v|_{1,\infty,K_{\alpha\beta}} = \max\left\{|u_x|_{1,\infty,\widehat{K}}/\alpha, |u_y|_{1,\infty,\widehat{K}}/\beta\right\},\\ |v|_{2,\infty,K_{\alpha\beta}} = \max\left\{|u_{xx}|_{2,\infty,\widehat{K}}/\alpha^2, |u_{xy}|_{2,\infty,\widehat{K}}/(\alpha\beta), |u_{yy}|_{2,\infty,\widehat{K}}/\beta^2\right\},$$

and obtain

$$\frac{|v|_{0,\infty,K_{\alpha\beta}}}{|v|_{1,\infty,K_{\alpha\beta}}} = \frac{|u|_{0,\infty,\widehat{K}}}{\max\left\{\frac{1}{\alpha}|u_{x}|_{0,\infty,\widehat{K}},\frac{1}{\beta}|u_{y}|_{0,\infty,\widehat{K}}\right\}},$$

$$\frac{|v|_{0,\infty,K_{\alpha\beta}}}{|v|_{2,\infty,K_{\alpha\beta}}} = \frac{|u|_{0,\infty,\widehat{K}}}{\max\left\{\frac{1}{\alpha^{2}}|u_{xx}|_{0,\infty,\widehat{K}},\frac{1}{\alpha\beta}|u_{xy}|_{0,\infty,\widehat{K}},\frac{1}{\beta^{2}}|u_{yy}|_{0,\infty,\widehat{K}}\right\}},$$

$$(21) \qquad \frac{|v|_{1,\infty,K_{\alpha\beta}}}{|v|_{2,\infty,K_{\alpha\beta}}} = \frac{\max\left\{\frac{1}{\alpha}|u_{x}|_{0,\infty,\widehat{K}},\frac{1}{\beta}|u_{y}|_{0,\infty,\widehat{K}},\frac{1}{\beta^{2}}|u_{yy}|_{0,\infty,\widehat{K}}\right\}}{\max\left\{\frac{1}{\alpha^{2}}|u_{xx}|_{0,\infty,\widehat{K}},\frac{1}{\alpha\beta}|u_{xy}|_{0,\infty,\widehat{K}},\frac{1}{\beta^{2}}|u_{yy}|_{0,\infty,\widehat{K}}\right\}}.$$

For a triangle K and $1 \le p \le \infty$, we define $\mathcal{T}_p^1(K) \subset W^{2,p}(K)$ by

$$\mathcal{T}_p^1(K) := \left\{ v \in W^{2,p}(K) \mid v(\mathbf{x}_i) = 0, \ i = 1, 2, 3 \right\}.$$

Note that if $v \in \mathcal{T}_p^1(K_{\alpha\beta})$, then $u := v \circ F_{\alpha\beta} \in \mathcal{T}_p^1(\widehat{K})$. The following lemma is from Babuška–Aziz [4, Lemma 2.2] and Kobayashi–Tsuchiya [15, Lemma 3].

Lemma 18. The constant
$$B_p^{1,1}(K_{\alpha\beta})$$
 is defined by

$$B_p^{1,1}(K_{\alpha\beta}) := \sup_{v \in \mathcal{T}_p^1(K_{\alpha\beta})} \frac{|v|_{1,p,K_{\alpha\beta}}}{|v|_{2,p,K_{\alpha\beta}}}, \qquad 1 \le p \le \infty.$$
Then, we have $B_p^{1,1}(K_{\alpha\beta}) \le \max\{\alpha,\beta\}A_p$.

Proof: Suppose first that $1 \leq p < \infty$. Take an arbitrary $v \in \mathcal{T}_p^1(K_{\alpha\beta})$ and define $u \in \mathcal{T}_p^1(\widehat{K})$ by $u(x,y) := v(x^*, y^*), (x^*, y^*)^\top = (\alpha x, \beta y)^\top$. By (20), we find

$$\begin{aligned} \frac{|v|_{1,p,K_{\alpha\beta}}^{p}}{|v|_{2,p,K_{\alpha\beta}}^{2}} &= \frac{\frac{1}{\alpha^{p}}|u_{x}|_{0,p,\widehat{K}}^{p} + \frac{1}{\beta^{p}}|u_{y}|_{0,p,\widehat{K}}^{p}}{\frac{1}{\alpha^{2p}}|u_{xx}|_{0,p,\widehat{K}}^{p} + \frac{1}{\alpha^{p}\beta^{p}}|u_{xy}|_{0,p,\widehat{K}}^{p} + \frac{1}{\beta^{p}}|u_{y}|_{0,p,\widehat{K}}^{p} + \frac{1}{\beta^{2p}}|u_{yy}|_{0,p,\widehat{K}}^{p}} \\ &\leq \frac{\max\{\alpha^{p},\beta^{p}\}\left(\frac{1}{\alpha^{p}}|u_{x}|_{0,p,\widehat{K}}^{p} + \frac{1}{\beta^{p}}|u_{y}|_{0,p,\widehat{K}}^{p}\right)}{\frac{1}{\alpha^{p}}\left(|u_{xx}|_{0,p,\widehat{K}}^{p} + |u_{xy}|_{0,p,\widehat{K}}^{p}\right) + \frac{1}{\beta^{p}}\left(|u_{xy}|_{0,p,\widehat{K}}^{p} + |u_{yy}|_{0,p,\widehat{K}}^{p}\right)} \\ &= \max\{\alpha^{p},\beta^{p}\}\frac{\frac{1}{\alpha^{p}}|u_{x}|_{0,p,\widehat{K}}^{p} + \frac{1}{\beta^{p}}|u_{y}|_{0,p,\widehat{K}}^{p}}{\frac{1}{\alpha^{p}}|u_{x}|_{1,p,\widehat{K}}^{p} + \frac{1}{\beta^{p}}|u_{y}|_{1,p,\widehat{K}}^{p}}. \end{aligned}$$

Here, we used the fact that, for X, Y > 0,

$$\frac{1}{\frac{1}{\alpha^p}X + \frac{1}{\beta^p}Y} \le \frac{\max\{\alpha^p, \beta^p\}}{X+Y}.$$

Note that u(0,0) = u(1,0) = 0 by the definition of $\mathcal{T}_p^1(\widehat{K})$ and $u_x \in \Xi_p^{(1,0),1}$. Thus, by Lemma 17, we realize that

$$|u_x|_{0,p,\widehat{K}}^p \le A_p^p |u_x|_{1,p,\widehat{K}}.$$

By the same reason, we realize that $u_y \in \Xi_p^{(0,1),1}$ and

$$|u_y|_{0,p,K}^p \le A_p^p |u_y|_{1,p,K}^p.$$

Inserting those inequalities into the above estimation, we obtain

$$\frac{|v|_{1,p,K_{\alpha\beta}}^{p}}{|v|_{2,p,K_{\alpha\beta}}^{p}} \le \max\{\alpha^{p},\beta^{p}\}\frac{\frac{A_{p}^{p}}{\alpha^{p}}|u_{x}|_{1,p,\widehat{K}}^{p} + \frac{A_{p}^{p}}{\beta^{p}}|u_{y}|_{1,p,\widehat{K}}^{p}}{\frac{1}{\alpha^{p}}|u_{x}|_{1,p,\widehat{K}}^{p} + \frac{1}{\beta^{p}}|u_{y}|_{1,p,\widehat{K}}^{p}} = (\max\{\alpha,\beta\})^{p}A_{p}^{p},$$

and conclude

$$B_p^{1,1}(K_{\alpha\beta}) = \sup_{v \in \mathcal{T}_p^1(K_{\alpha\beta})} \frac{|v|_{1,p,K_{\alpha\beta}}}{|v|_{2,p,K_{\alpha\beta}}} \le \max\{\alpha,\beta\}A_p.$$

Next, let $p = \infty$. By (21), we immediately obtain

$$\begin{aligned} \frac{|v|_{1,\infty,K_{\alpha\beta}}}{|v|_{2,\infty,K_{\alpha\beta}}} &= \frac{\max\left\{\frac{1}{\alpha}|u_x|_{0,\infty,\widehat{K}},\frac{1}{\beta}|u_y|_{0,\infty,\widehat{K}}\right\}}{\max\left\{\max\left\{\frac{|u_{xx}|_{0,\infty,\widehat{K}}}{\alpha^2},\frac{|u_{xy}|_{0,\infty,\widehat{K}}}{\alpha\beta}\right\},\max\left\{\frac{|u_{xy}|_{0,\infty,\widehat{K}}}{\alpha\beta},\frac{|u_{yy}|_{0,\infty,\widehat{K}}}{\beta^2},\right\}\right\}} \\ &\leq \frac{\max\{\alpha,\beta\}\max\left\{\frac{1}{\alpha}\max\left\{|u_{xx}|_{0,\infty,\widehat{K}},|u_{xy}|_{0,\infty,\widehat{K}}\right\},\frac{1}{\beta}\max\left\{|u_{xy}|_{0,\infty,\widehat{K}},|u_{yy}|_{0,\infty,\widehat{K}},\right\}\right\}}{\max\left\{\frac{1}{\alpha}|u_x|_{0,\infty,\widehat{K}},\frac{1}{\beta}|u_y|_{0,\infty,\widehat{K}}\right\}} \\ &= \max\{\alpha,\beta\}\frac{\max\left\{\frac{1}{\alpha}|u_x|_{0,\infty,\widehat{K}},\frac{1}{\beta}|u_y|_{0,\infty,\widehat{K}}\right\}}{\max\left\{\frac{1}{\alpha}|u_x|_{1,\infty,\widehat{K}},\frac{1}{\beta}|u_y|_{1,\infty,\widehat{K}}\right\}} \\ &\leq \max\{\alpha,\beta\}\frac{A_{\infty}\max\left\{\frac{1}{\alpha}|u_x|_{1,\infty,\widehat{K}},\frac{1}{\beta}|u_y|_{1,\infty,\widehat{K}}\right\}}{\max\left\{\frac{1}{\alpha}|u_x|_{1,\infty,\widehat{K}},\frac{1}{\beta}|u_y|_{1,\infty,\widehat{K}}\right\}} = \max\{\alpha,\beta\}A_{\infty}.$$

The following lemma is from Babuška–Aziz [4, Lemma 2.3,2.4] and Kobayashi–Tsuchiya [15, Lemma 4,5].

Lemma 19. The constants
$$B_p^{0,1}(K_{\alpha\beta})$$
, \widetilde{A}_p are defined by
 $B_p^{0,1}(K_{\alpha\beta}) := \sup_{v \in \mathcal{T}_p^1(K_{\alpha\beta})} \frac{|v|_{0,p,K_{\alpha\beta}}}{|v|_{2,p,K_{\alpha\beta}}}, \quad \widetilde{A}_p := B_p^{0,1}(\widehat{K}) := \sup_{v \in \mathcal{T}_p^1(\widehat{K})} \frac{|v|_{0,p,\widehat{K}}}{|v|_{2,p,\widehat{K}}}, \quad 1 \le p \le \infty.$
Then, we have the estimation $B_p^{0,1}(K_{\alpha\beta}) \le \max\{\alpha^2, \beta^2\}\widetilde{A}_p < +\infty.$

Proof: The proof of $\widetilde{A}_p < +\infty$ is very similar to that of Lemma 17 and is by contradiction. Suppose that $\widetilde{A}_p = \infty$. Then, there exists $\{u_m\}_{m=1}^{\infty} \subset \mathcal{T}_p^1(\widehat{K})$ such that

$$|u_m|_{0,p,\widehat{K}} = 1, \qquad \lim_{m \to \infty} |u_m|_{2,p,\widehat{K}} = 0.$$

Then, by (14), there exists $\{q_m\} \subset \mathcal{P}_1(\widehat{K})$ such that

$$\lim_{m \to \infty} \left\| u_m + q_m \right\|_{2,p,\widehat{K}} = 0$$

Since $|u_m|_{0,p,\widehat{K}}$ and $|u_m|_{2,p,\widehat{K}}$ are bounded, $|u_m|_{1,p,\widehat{K}}$ and $||u_m||_{2,p,\widehat{K}}$ are bounded as well by Gagliardo–Nirenberg's inequality (Theorem 13). Hence, $\{q_m\} \subset \mathcal{P}_1(\widehat{K})$ is also bounded. Thus, there exists a subsequence $\{q_{m_i}\}$ which converges to $\overline{q} \in \mathcal{P}_1(\widehat{K})$. In particular, we have

$$\lim_{m_i \to \infty} \left\| u_{m_i} + \bar{q} \right\|_{2,p,\widehat{K}} = 0.$$

Since $\{u_m\} \subset \mathcal{T}_p^1(\widehat{K})$, we conclude that $\overline{q} \in \mathcal{T}_p^1(\widehat{K}) \cap \mathcal{P}_1(\widehat{K})$ and $\overline{q} = 0$. Therefore, we reach $\lim_{m_i \to \infty} \|u_{m_i}\|_{2,p,\widehat{K}} = 0$ which contradicts to $\lim_{m_i \to \infty} \|u_{m_i}\|_{2,p,\widehat{K}} \ge \lim_{m_i \to \infty} |u_{m_i}|_{0,p,\widehat{K}} = 1$.

We now consider the estimation for the case $1 \le p < \infty$. From (19) we have

$$\frac{|v|_{0,p,K_{\alpha\beta}}^{p}}{|v|_{2,p,K_{\alpha\beta}}^{p}} = \frac{|u|_{0,p,\widehat{K}}^{p}}{\frac{1}{\alpha^{2p}}|u_{xx}|_{0,p,\widehat{K}}^{p} + \frac{2}{\alpha^{p}\beta^{p}}|u_{xy}|_{0,p,\widehat{K}}^{p} + \frac{1}{\beta^{2p}}|u_{yy}|_{0,p,\widehat{K}}^{p}} \\
\leq \frac{\max\{\alpha^{2p},\beta^{2p}\}|u|_{0,p,\widehat{K}}^{p}}{|u_{xx}|_{0,p,\widehat{K}}^{p} + 2|u_{xy}|_{0,p,\widehat{K}}^{p} + |u_{yy}|_{0,p,\widehat{K}}^{p}} \leq \left(\max\{\alpha^{2},\beta^{2}\}\right)^{p}\widetilde{A}_{p}^{p}.$$

and Lemma is shown for this case. The proof for the case $p = \infty$ is very similar. \Box

Exercise: In Lemma 19, prove the case $p = \infty$.

We may apply Lemmas 18 and 19 to $v - \mathcal{I}^1_{K_{\alpha\beta}} v \in \mathcal{T}^1_p(K_{\alpha\beta})$ for $v \in W^{2,p}(K_{\alpha\beta})$, and obtain the following corollary.

Corollary 20. For arbitrary $v \in W^{2,p}(K_{\alpha\beta})$ $(1 \le p \le \infty)$, the following estimations hold: $|v - \mathcal{I}^{1}_{K_{\alpha\beta}}v|_{1,p,K_{\alpha\beta}} \le \max\{\alpha,\beta\}A_{p}|v|_{2,p,K_{\alpha\beta}},$ $|v - \mathcal{I}^{1}_{K_{\alpha\beta}}v|_{0,p,K_{\alpha\beta}} \le (\max\{\alpha,\beta\})^{2} \widetilde{A}_{p}|v|_{2,p,K_{\alpha\beta}}.$

\S 4. Extending Babuška-Aziz's technique to the higher order Lagrange interpolation

In this section, we prove the following theorem using Babuška-Aziz's technique. Let k be a positive integer and p be such that $1 \le p \le \infty$. The set $\mathcal{T}_p^k(K)$ is defined by

$$\mathcal{T}_p^k(K) := \left\{ v \in W^{k+1,p}(K) \mid v(\mathbf{x}) = 0, \forall \mathbf{x} \in \Sigma^k(K) \right\},\$$

where $\Sigma^k(K)$ is defined by (1). Note that if $v \in \mathcal{T}_p^k(K_{\alpha\beta})$, then $u = v \circ F_{\alpha\beta} \in \mathcal{T}_p^k(\widehat{K})$.

Theorem 21. Take arbitrary $\alpha > 0$ and $\beta > 0$. Then, there exists a constant $C_{k,m,p}$ such that, for $m = 0, 1, \dots, k$,

$$B_p^{m,k}(K_{\alpha\beta}) := \sup_{v \in \mathcal{T}_p^k(K_{\alpha\beta})} \frac{|v|_{m,p,K_{\alpha\beta}}}{|v|_{k+1,p,K_{\alpha\beta}}} \le \left(\max\{\alpha,\beta\}\right)^{k+1-m} C_{k,m,p}.$$

Here, $C_{k,m,p}$ depends only on k, m, and p, and is independent of α and β .



Figure 4. Squeezing the reference triangle \hat{K} perpendicularly does not deteriorate the approximation property of Lagrange interpolation.

Applying Theorem 21 to $v - \mathcal{I}_{K_{\alpha\beta}}^k v \in \mathcal{T}_p^k(K_{\alpha\beta})$ for $v \in W^{k+1,p}(K_{\alpha\beta})$, and obtain the following corollary.

Corollary 22. For arbitrary $v \in W^{k+1,p}(K_{\alpha\beta})$ $(1 \le p \le \infty)$, the following estimations hold: $|v - \mathcal{I}_{K_{\alpha\beta}}^{k}v|_{m,p,K_{\alpha\beta}} \le C_{k,m,p} \left(\max\{\alpha,\beta\}\right)^{k+1-m} |v|_{k+1,p,K_{\alpha\beta}}.$

The manner of the proof of Theorem 21 is exactly similar as in the previous section. The ratio $|v|_{m,p,K_{\alpha\beta}}^p/|v|_{k+1,p,K_{\alpha\beta}}^p$ is written using the seminorms of u on \hat{K} , and is bounded by a constant that does not depend on v.

First, let $1 \leq p < \infty$. For a multi-index $\gamma = (a, b) \in \mathbb{N}_0^2$ and a real $t \neq 0$, set $(\alpha, \beta)^{\gamma t} := \alpha^{at} \beta^{bt}$. Then, we have

$$\frac{|v|_{m,p,K_{\alpha\beta}}^{p}}{|v|_{k+1,p,K_{\alpha\beta}}^{p}} = \frac{\sum_{|\gamma|=m} \frac{m!}{\gamma!} (\alpha,\beta)^{-\gamma p} \left| \partial^{\gamma} u \right|_{0,p,\widehat{K}}^{p}}{\sum_{|\delta|=k+1} \frac{(k+1)!}{\delta!} (\alpha,\beta)^{-\delta p} \left| \partial^{\delta} u \right|_{0,p,\widehat{K}}^{p}} \\
= \frac{\sum_{|\gamma|=m} \frac{m!}{\gamma!} (\alpha,\beta)^{-\gamma p} \left| \partial^{\gamma} u \right|_{0,p,\widehat{K}}^{p}}{\sum_{|\gamma|=m} \frac{m!}{\gamma!} (\alpha,\beta)^{-\gamma p} \left(\sum_{|\eta|=k+1-m} \frac{(k+1-m)!}{\eta! (\alpha,\beta)^{\eta p}} \left| \partial^{\eta} (\partial^{\gamma} u) \right|_{0,p,\widehat{K}}^{p} \right)} \\
\leq \frac{(\max\{\alpha,\beta\})^{(k+1-m)p} \sum_{|\gamma|=m} \frac{m!}{\gamma!} (\alpha,\beta)^{-\gamma p} \left(\sum_{|\eta|=k+1-m} \frac{(k+1-m)!}{\eta!} \left| \partial^{\eta} (\partial^{\gamma} u) \right|_{0,p,\widehat{K}}^{p} \right)}{\sum_{|\gamma|=m} \frac{m!}{\gamma!} (\alpha,\beta)^{-\gamma p} \left(\sum_{|\eta|=k+1-m} \frac{(k+1-m)!}{\eta!} \left| \partial^{\eta} (\partial^{\gamma} u) \right|_{0,p,\widehat{K}}^{p} \right)}$$

Kenta Kobayashi · Takuya Tsuchiya

(22)
$$= \left(\max\{\alpha,\beta\}\right)^{(k+1-m)p} \frac{\sum_{|\gamma|=m} \frac{m!}{\gamma!} (\alpha,\beta)^{-\gamma p} \left|\partial^{\gamma} u\right|_{0,p,\widehat{K}}^{p}}{\sum_{|\gamma|=m} \frac{m!}{\gamma!} (\alpha,\beta)^{-\gamma p} \left|\partial^{\gamma} u\right|_{k+1-m,p,\widehat{K}}^{p}}.$$

Here, we used the fact that, for a multi-index η , $(\alpha, \beta)^{\eta p} \leq (\max\{\alpha, \beta\})^{|\eta|p}$ and, for a multi-index δ with $|\delta| = k + 1$,

$$\frac{|\delta|!}{\delta!} = \sum_{\substack{\gamma+\eta=\delta\\ |\gamma|=m, |\eta|=k+1-m}} \frac{|\gamma|!}{\gamma!} \frac{|\eta|!}{\eta!}.$$

For example, if k = 2, then we see

$$\begin{split} \frac{|v|_{1,p,K_{\alpha}}^{p}}{|v|_{3,p,K_{\alpha}}^{p}} &= \frac{\frac{1}{\alpha^{p}} |u_{x}|_{0}^{p} + \frac{1}{\beta^{p}} |u_{y}|_{0}^{p}}{\frac{1}{\alpha^{3p}} |u_{xxx}|_{0}^{p} + \frac{3}{\alpha^{2p}\beta^{p}} |u_{xxy}|_{0}^{p} + \frac{3}{\alpha^{p}\beta^{2p}} |u_{xyy}|_{0}^{p} + \frac{1}{\beta^{3p}} |u_{yyy}|_{0}^{p}}{\frac{1}{\alpha^{p}} \left(\frac{|u_{xxx}|_{p}^{p}}{\alpha^{2p}} + 2\frac{|u_{xxy}|_{p}^{p}}{\alpha^{p}\beta^{p}} + \frac{|u_{xyy}|_{p}^{p}}{\beta^{2p}}\right) + \frac{1}{\beta^{p}} \left(\frac{|u_{xxy}|_{p}^{p}}{\alpha^{2p}} + 2\frac{|u_{xyy}|_{p}^{p}}{\beta^{2p}} + \frac{|u_{yyy}|_{p}^{p}}{\beta^{2p}}\right) \\ &\leq \frac{\max\{\alpha^{2p},\beta^{2p}\}\left(\frac{1}{\alpha^{p}} |u_{x}|_{0}^{p} + \frac{1}{\beta^{p}} |u_{y}|_{0}^{p}\right)}{\frac{1}{\alpha^{p}} \left(|u_{xxx}|_{0}^{p} + 2|u_{xxy}|_{0}^{p} + |u_{xyy}|_{0}^{p}\right) + \frac{1}{\beta^{p}} \left(|u_{xxy}|_{0}^{p} + 2|u_{xyy}|_{0}^{p} + u_{yyy}|_{0}^{p}\right)}{\frac{1}{\alpha^{p}} |u_{x}|_{0,p,\hat{K}}^{p} + \frac{1}{\beta^{p}} |u_{y}|_{0,p,\hat{K}}^{p}}, \end{split}$$

and

$$\begin{split} \frac{|v|_{2,p,K_{\alpha\beta}}^{p}}{|v|_{3,p,K_{\alpha\beta}}^{p}} &= \frac{\frac{1}{\alpha^{2p}} |u_{xx}|_{0}^{p} + \frac{2}{\alpha^{p}\beta^{p}} |u_{xy}|_{0}^{p} + \frac{1}{\beta^{2p}} |u_{yy}|_{0}^{p}}{\frac{1}{\alpha^{3p}} |u_{xxx}|_{0}^{p} + \frac{3}{\alpha^{2p}\beta^{p}} |u_{xxy}|_{0}^{p} + \frac{3}{\alpha^{p}\beta^{2p}} |u_{xyy}|_{0}^{p} + \frac{1}{\beta^{3p}} |u_{yyy}|_{0}^{p}} \\ &= \frac{\frac{1}{\alpha^{2p}} |u_{xx}|_{p}^{p} + \frac{2}{\alpha^{p}\beta^{p}} |u_{xy}|_{p}^{p} + \frac{1}{\beta^{2p}} |u_{yy}|_{p}^{p}}{\frac{1}{\alpha^{2p}} \left(\frac{|u_{xxx}|_{p}^{p}}{\alpha^{p}} + \frac{|u_{xxy}|_{p}^{p}}{\beta^{p}} \right) + \frac{2}{\alpha^{p}\beta^{p}} \left(\frac{|u_{xxy}|_{p}^{p}}{\alpha^{p}} + \frac{|u_{xyy}|_{p}^{p}}{\beta^{p}} \right) + \frac{1}{\beta^{2p}} \left(\frac{|u_{xyy}|_{p}^{p}}{\alpha^{p}} + \frac{|u_{yyy}|_{p}^{p}}{\beta^{p}} \right) \\ &\leq \frac{\max\{\alpha^{p}, \beta^{p}\} \left(\frac{1}{\alpha^{2p}} |u_{xx}|_{0}^{p} + \frac{2}{\alpha^{p}\beta^{p}} (|u_{xxy}|_{0}^{p} + \frac{2}{\alpha^{p}\beta^{p}} |u_{xy}|_{0}^{p} + \frac{1}{\beta^{2p}} |u_{yy}|_{0}^{p} \right)}{\frac{1}{\alpha^{2p}} \left(|u_{xxx}|_{0}^{p} + |u_{xxy}|_{0}^{p} \right) + \frac{2}{\alpha^{p}\beta^{p}} (|u_{xxy}|_{0}^{p} + \frac{2}{\alpha^{p}\beta^{p}} |u_{xy}|_{0}^{p} + \frac{1}{\beta^{2p}} |u_{yy}|_{0}^{p} \right)} \\ &= \max\{\alpha^{p}, \beta^{p}\} \frac{\frac{1}{\alpha^{2p}} |u_{xx}|_{0}^{p} + \frac{2}{\alpha^{p}\beta^{p}} (|u_{xxy}|_{0}^{p} + |u_{xyy}|_{0}^{p} + \frac{1}{\beta^{2p}} |u_{yy}|_{0}^{p}, \hat{K} + \frac{1}{\beta^{2p}} |u_{yy}|_{0}^{p}, \hat{K} \right)}{\frac{1}{\alpha^{2p}} |u_{xx}|_{0,p,\hat{K}}^{p} + \frac{2}{\alpha^{p}\beta^{p}} |u_{xy}|_{0,p,\hat{K}}^{p} + \frac{1}{\beta^{2p}} |u_{yy}|_{1,p,\hat{K}}^{p} \right)}. \end{split}$$

In the above, we use the notation $|\cdot|_0$ instead of $|\cdot|_{0,p,\widehat{K}}$ for simplicity. **Exercise:** Confirm the details of the above inequalities, in particular, (22).

Now suppose that, for $\mathcal{T}_p^k(\widehat{K})$ and a multi-index γ , the set $\Xi_p^{\gamma,k}$ is defined so that

(23)
$$u \in \mathcal{T}_p^k(\widehat{K}) \Longrightarrow \partial^{\gamma} u \in \Xi_p^{\gamma,k}$$

and

(24)
$$A_{p}^{\gamma,k} := \sup_{v \in \Xi_{p}^{\gamma,k}} \frac{|v|_{0,p,\widehat{K}}}{|v|_{k+1-|\gamma|,p,\widehat{K}}} < \infty$$

hold. Then, from (22), we would conclude that

$$\frac{|v|_{m,p,K_{\alpha\beta}}^{p}}{|v|_{k+1,p,K_{\alpha\beta}}^{p}} \leq \left(\max\{\alpha,\beta\}\right)^{(k+1-m)p} \frac{\sum_{|\gamma|=m} \frac{m!}{\gamma!} (\alpha,\beta)^{-\gamma p} \left|\partial^{\gamma} u\right|_{0,p,\widehat{K}}^{p}}{\sum_{|\gamma|=m} \frac{m!}{\gamma!} (\alpha,\beta)^{-\gamma p} \left|\partial^{\gamma} u\right|_{k+1-m,p,\widehat{K}}^{p}} \\ \leq \left(\max\{\alpha,\beta\}\right)^{(k+1-m)p} \frac{\sum_{|\gamma|=m} \left(A_{p}^{\gamma,k}\right)^{p} \frac{m!}{\gamma!} (\alpha,\beta)^{-\gamma p} \left|\partial^{\gamma} u\right|_{k+1-m,p,\widehat{K}}^{p}}{\sum_{|\gamma|=m} \frac{m!}{\gamma!} (\alpha,\beta)^{-\gamma p} \left|\partial^{\gamma} u\right|_{k+1-m,p,\widehat{K}}^{p}} \\ (25) \leq \left(\max\{\alpha,\beta\}\right)^{(k+1-m)p} C_{k,m,p}^{p}, \qquad C_{k,m,p} := \max_{|\gamma|=m} A_{p}^{\gamma,k}.$$

Our task now is to define $\Xi_p^{\gamma,k}$ that satisfies (23) and (24). We will explain the details in the following sections.

§5. Difference quotients

In this section, we define the difference quotients for two-variable functions. Our treatment is based on the theory of difference quotients of one-variable functions given in standard textbooks such as [3] and [25]. All statements in this section can be readily proved.

§ 5.1. Difference quotients of one-variable functions

For a function f(x) and nodal points $x_0, x_1, \dots, x_n \in \mathbb{R}$, the difference quotients of f are defined recursively by

$$f[x_0, x_1] := \frac{f(x_0) - f(x_1)}{x_0 - x_1}, \qquad f[x_0, x_1, x_2] := \frac{f[x_0, x_1] - f[x_1, x_2]}{x_0 - x_2},$$

$$f[x_0, x_1, \cdots, x_m] := \frac{f[x_0, \cdots, x_{m-1}] - f[x_1, \cdots, x_m]}{x_0 - x_m}.$$

A simplest case is $x_i := x_0 + hi$, $i = 1, \dots, m$, with h > 0. In this case, the difference quotients are

$$f[x_0, x_1] := \frac{f(x_1) - f(x_0)}{h}, \qquad f[x_0, x_1, x_2] := \frac{f(x_0) - 2f(x_1) + f(x_2)}{h^2},$$

and so on. The difference quotients are expressed by integration:

$$f[x_0, x_1] = \frac{1}{x_1 - x_0} \int_{x_0}^{x_1} f'(t) dt = \int_0^1 f'(x_0 + t_1(x_1 - x_0)) dt_1,$$

$$f[x_0, x_1, x_2] = \frac{f[x_2, x_1] - f[x_1, x_0]}{x_2 - x_0} = \frac{f[x_2, x_0] - f[x_0, x_1]}{x_2 - x_1}$$

$$= \frac{1}{x_2 - x_1} \int_0^1 (f'(x_0 + t_1(x_2 - x_0)) - f'(x_0 + t_1(x_1 - x_0))) dt_1$$

$$= \int_0^1 \int_0^{t_1} f''(x_0 + t_1(x_1 - x_0) + t_2(x_2 - x_1)) dt_2 dt_1.$$

For $n \ge 1$, the following formula holds:

$$(26)f[x_0, x_1, \cdots, x_n] = \int_0^1 \int_0^{t_1} \cdots \int_0^{t_{n-1}} f^{(n)} \left(x_0 + \sum_{i=1}^n t_i (x_i - x_{i-1}) \right) dt_n \cdots dt_2 dt_1.$$

Exercise: Prove (26) by induction.

§ 5.2. Difference quotients of two variable functions

We now extend the difference quotient to functions with two variables. For a positive integer k, the set $\widehat{\Sigma}^k \subset \widehat{K}$ is defined by

$$\widehat{\Sigma}^k := \Sigma^k(\widehat{K}) := \left\{ \mathbf{x}_{\gamma} := \frac{\gamma}{k} \in \widehat{K} \mid \gamma \in \mathbb{N}_0^2, \ 0 \le |\gamma| \le k \right\},\$$

where $\gamma/k = (a_1/k, a_2/k)$ is understood as the coordinate of a point in $\widehat{\Sigma}^k$.

For $\mathbf{x}_{\gamma} \in \widehat{\Sigma}^k$ and a multi-index $\delta \in \mathbb{N}_0^2$ with $|\gamma| \leq k - |\delta|$, we define the correspondence Δ^{δ} between nodes by

$$\Delta^{\delta} \mathbf{x}_{\gamma} := \mathbf{x}_{\gamma+\delta} \in \widehat{\Sigma}^k.$$

For example, $\Delta^{(1,1)}\mathbf{x}_{(0,0)} = \mathbf{x}_{(1,1)}$ and $\Delta^{(2,1)}\mathbf{x}_{(0,1)} = \mathbf{x}_{(2,2)}$. Using Δ^{δ} , we define the **difference quotients** on $\widehat{\Sigma}^k$ for $f \in C^0(\widehat{K})$ by

$$f^{|\delta|}[\mathbf{x}_{\gamma}, \Delta^{\delta} \mathbf{x}_{\gamma}] := k^{|\delta|} \sum_{\eta \le \delta} \frac{(-1)^{|\delta| - |\eta|}}{\eta! (\delta - \eta)!} f(\Delta^{\eta} \mathbf{x}_{\gamma}).$$

For simplicity, we denote $f^{|\delta|}[\mathbf{x}_{(0,0)}, \Delta^{\delta}\mathbf{x}_{(0,0)}]$ by $f^{|\delta|}[\Delta^{\delta}\mathbf{x}_{(0,0)}]$. The following are examples of $f^{|\delta|}[\Delta^{\delta}\mathbf{x}_{(0,0)}]$:

$$\begin{split} f^2[\Delta^{(2,0)}\mathbf{x}_{(0,0)}] &= \frac{k^2}{2}(f(\mathbf{x}_{(2,0)}) - 2f(\mathbf{x}_{(1,0)}) + f(\mathbf{x}_{(0,0)})), \\ f^2[\Delta^{(1,1)}\mathbf{x}_{(0,0)}] &= k^2(f(\mathbf{x}_{(1,1)}) - f(\mathbf{x}_{(1,0)}) - f(\mathbf{x}_{(0,1)}) + f(\mathbf{x}_{(0,0)})), \\ f^3[\Delta^{(2,1)}\mathbf{x}_{(0,0)}] &= \frac{k^3}{2}(f(\mathbf{x}_{(2,1)}) - 2f(\mathbf{x}_{(1,1)}) + f(\mathbf{x}_{(0,1)}) - f(\mathbf{x}_{(2,0)}) \\ &\quad + 2f(\mathbf{x}_{(1,0)}) - f(\mathbf{x}_{(0,0)})). \end{split}$$

Let $\eta \in \mathbb{N}_0^2$ be such that $|\eta| = 1$ and $\eta \leq \delta$. The difference quotients clearly satisfy the following recursive relations:

$$f^{|\delta|}[\mathbf{x}_{\gamma}, \Delta^{\delta}\mathbf{x}_{\gamma}] = \frac{k}{\delta \cdot \eta} \left(f^{|\delta|-1}[\mathbf{x}_{\gamma+\eta}, \Delta^{\delta-\eta}\mathbf{x}_{\gamma+\eta}] - f^{|\delta|-1}[\mathbf{x}_{\gamma}, \Delta^{\delta-\eta}\mathbf{x}_{\gamma}] \right).$$

If $f \in C^k(\widehat{K})$, the difference quotient $f^{|\delta|}[\mathbf{x}_{\gamma}, \Delta^{\delta}\mathbf{x}_{\gamma}]$ is written as an integral of f. Setting d = 2 and $\delta = (0, s)$, for example, we have

$$\begin{split} f^{1}[\mathbf{x}_{(l,q)}, \Delta^{(0,1)}\mathbf{x}_{(l,q)}] &= k(f(\mathbf{x}_{(l,q+1)}) - f(\mathbf{x}_{(l,q)})) = \int_{0}^{1} \partial^{(0,1)} f\left(\frac{l}{k}, \frac{q}{k} + \frac{w_{1}}{k}\right) \mathrm{d}w_{1}, \\ f^{1}[\mathbf{x}_{(l,q)}, \Delta^{(0,2)}\mathbf{x}_{(l,q)}] &= \frac{k^{2}}{2}(f(\mathbf{x}_{(l,q+2)}) - 2f(\mathbf{x}_{(l,q+1)}) + f(\mathbf{x}_{(l,q)})) \\ &= \int_{0}^{1} \int_{0}^{w_{1}} \partial^{(0,2)} f\left(\frac{l}{k}, \frac{q}{k} + \frac{1}{k}(w_{1} + w_{2})\right) \mathrm{d}w_{2} \mathrm{d}w_{1} \\ &= k \int_{0}^{1} \left[\partial^{(0,1)} f\left(\frac{l}{k}, \frac{q}{k} + \frac{2}{k}w_{1}\right) - \partial^{(0,1)} f\left(\frac{l}{k}, \frac{q}{k} + \frac{1}{k}w_{1}\right)\right] \mathrm{d}w_{1}, \\ f^{s}[\mathbf{x}_{(l,p)}, \Delta^{(0,s)}\mathbf{x}_{(l,q)}] \end{split}$$

$$=\int_0^1\int_0^{w_1}\cdots\int_0^{w_{s-1}}\partial^{(0,s)}f\left(\frac{l}{k},\frac{q}{k}+\frac{1}{k}(w_1+\cdots+w_s)\right)\mathrm{d}w_s\cdots\mathrm{d}w_2\mathrm{d}w_1.$$

To provide a concise expression for the above integral, we introduce the s-simplex

$$\mathbb{S}_s := \{ (t_1, t_2, \cdots, t_s) \in \mathbb{R}^s \mid t_i \ge 0, \ 0 \le t_1 + \cdots + t_s \le 1 \},\$$

and the integral of $g \in L^1(\mathbb{S}_s)$ on \mathbb{S}_s is defined by

$$\int_{\mathbb{S}_s} g(w_1, \cdots, w_k) \mathrm{d}\mathbf{W}_{\mathbf{s}} := \int_0^1 \int_0^{w_1} \cdots \int_0^{w_{s-1}} g(w_1, \cdots, w_s) \mathrm{d}w_s \cdots \mathrm{d}w_2 \mathrm{d}w_1,$$

where $d\mathbf{W}_{\mathbf{s}} = dw_s \cdots dw_2 dw_1$. Then, $f^s[\mathbf{x}_{(l,q)}, \Delta^{(0,s)}\mathbf{x}_{(l,q)}]$ becomes

$$f^{s}[\mathbf{x}_{(l,q)}, \Delta^{(0,s)}\mathbf{x}_{(l,q)}] = \int_{\mathbb{S}_{s}} \partial^{(0,s)} f\left(\frac{l}{k}, \mathbf{W}_{s}\right) \mathrm{d}\mathbf{W}_{s}, \quad \mathbf{W}_{s} := \frac{q}{k} + \frac{1}{k}(w_{1} + \dots + w_{s}).$$

For a general multi-index (t, s), we have

$$f^{t+s}[\mathbf{x}_{(l,q)}, \Delta^{(t,s)}\mathbf{x}_{(l,q)}] = \int_{\mathbb{S}_s} \int_{\mathbb{S}_t} \partial^{(t,s)} f(\mathbf{Z}_t, \mathbf{W}_s) \, \mathrm{d}\mathbf{Z}_t \, \mathrm{d}\mathbf{W}_s.$$
$$\mathbf{Z}_t := \frac{l}{k} + \frac{1}{k}(z_1 + \dots + z_t), \quad \mathrm{d}\mathbf{Z}_t := \mathrm{d}z_t \cdots \mathrm{d}z_2 \mathrm{d}z_1.$$

Let \Box_{γ}^{δ} be the rectangle defined by \mathbf{x}_{γ} and $\Delta^{\delta} \mathbf{x}_{\gamma}$ as the diagonal points. If $\delta = (t, 0)$ or (0, s), \Box_{γ}^{δ} degenerates to a segment. For $v \in W^{1,1}(\widehat{K})$ and \Box_{γ}^{δ} with $\gamma = (l, q)$, we denote the integral as

$$\int_{\square_{\gamma}^{(t,s)}} v := \int_{\mathbb{S}_s} \int_{\mathbb{S}_t} v \left(\mathbf{Z}_t, \mathbf{W}_s \right) \mathrm{d} \mathbf{Z}_t \mathrm{d} \mathbf{W}_s.$$

If \Box_{γ}^{δ} degenerates to a segment, the integral is understood as an integral on the segment. By this notation, the difference quotient $f^{t+s}[\mathbf{x}_{\gamma}, \Delta^{(t,s)}\mathbf{x}_{\gamma}]$ is written as

$$f^{(t+s)}[\mathbf{x}_{\gamma}, \Delta^{(t,s)}\mathbf{x}_{\gamma}] = \int_{\Box_{\gamma}^{(t,s)}} \partial^{(t,s)} f.$$

Therefore, if $u \in \mathcal{T}_p^k(\widehat{K})$, then we have

(27)
$$0 = u^{t+s}[\mathbf{x}_{\gamma}, \Delta^{(t,s)}\mathbf{x}_{\gamma}] = \int_{\square_{\gamma}^{(t,s)}} \partial^{(t,s)}u, \quad \forall \square_{\gamma}^{(t,s)} \subset \widehat{K}.$$

Exercise: Confirm that all the equations in this section certainly hold.

§ 6. The proof of Theorem 21

By introducing the notation in the previous section, we now be able to define $\Xi_p^{\gamma,k} \subset W^{k+1-|\gamma|,p}(\widehat{K})$ and $A_p^{\gamma,k}$ for $p \in [1,\infty]$, which satisfy (23) and (24). For multiindex γ , define

$$\Xi_p^{\gamma,k} := \left\{ v \in W^{k+1-|\gamma|,p}(\widehat{K}) \ \Big| \ \int_{\square_{lp}^{\gamma}} v = 0, \quad \forall \square_{lp}^{\gamma} \subset \widehat{K} \right\}.$$

From the definition and (27), it is clear that (23) holds. Define

$$A_p^{\gamma,k} := \sup_{v \in \Xi_p^{\gamma,k}} \frac{|v|_{0,p,\widehat{K}}}{|v|_{k+1-|\gamma|,p,\widehat{K}}}, \qquad 1 \le p \le \infty.$$

Then, the following lemma holds.

Lemma 23. We have
$$\Xi_p^{\gamma,k} \cap \mathcal{P}_{k-|\gamma|} = \{0\}$$
. That is, if $q \in \mathcal{P}_{k-|\gamma|}$ belongs to $\Xi_p^{\gamma,k}$, then $q = 0$.

Proof: We notice that $\dim \mathcal{P}_{k-|\delta|} = \#\{\Box_{lp}^{\delta} \subset \widehat{K}\}$. For example, if k = 4 and $|\delta| = 2$, then $\dim \mathcal{P}_2 = 6$. This corresponds to the fact that, in \widehat{K} , there are six squares with size 1/4 for $\delta = (1,1)$ and there are six horizontal segments of length 1/2 for $\delta = (2,0)$. All their vertices (corners and end-points) belong to $\Sigma^4(\widehat{K})$ (see Figure 5). Now, suppose that $v \in \mathcal{P}_{k-|\delta|}$ satisfies $\int_{\Box_{lp}^{\delta}} q = 0$ for all $\Box_{lp}^{\delta} \subset \widehat{K}$. This condition is linearly independent and determines q = 0 uniquely. \Box

To understand the above proof clearly, we consider the cases k = 2 and 3. Let k = 2 and $\gamma = (1, 0)$. Then, $k - |\gamma| = 1$. Set q(x, y) = a + bx + cy. If the three integrals

$$\int_{\Box_{00}^{(1,0)}} q(x,y) = a + \frac{b}{4}, \qquad \int_{\Box_{10}^{(1,0)}} q(x,y) = a + \frac{3b}{4}, \qquad \int_{\Box_{01}^{(1,0)}} q(x,y) = a + \frac{b}{4} + \frac{c}{2}$$



Figure 5. The six squares of size 1/4 for $\delta = (1,1)$ and the (union of) six segments of length 1/2 for $\delta = (2,0)$ in \widehat{K} .

are equal to 0, then we have a = b = c = 0, that is, q(x, y) = 0. The case $\gamma = (0, 1)$ is similar.

Let k = 3 and $\gamma = (1,0)$. Then, $k - |\gamma| = 2$. Set $q(x,y) = a + bx + cy + dx^2 + ey^2 + fxy$. If the integrals

$$\begin{split} \int_{\Box_{00}^{(1,0)}} q(x,y) &= a + \frac{b}{6} + \frac{d}{27}, \qquad \int_{\Box_{10}^{(1,0)}} q(x,y) &= a + \frac{b}{2} + \frac{7}{27}d, \\ \int_{\Box_{20}^{(1,0)}} q(x,y) &= a + \frac{5}{6}b + \frac{19}{27}d \end{split}$$

are all equal to 0, we have a = b = d = 0. Moreover, if the integrals

$$\begin{split} \int_{\Box_{01}^{(1,0)}} q(x,y) &= \frac{c}{3} + \frac{e}{9} + \frac{f}{18}, \qquad \int_{\Box_{11}^{(1,0)}} q(x,y) = \frac{c}{3} + \frac{e}{9} + \frac{f}{6}, \\ \int_{\Box_{02}^{(1,0)}} q(x,y) &= \frac{2}{3}c + \frac{4}{9}e + \frac{f}{9} \end{split}$$

are equal to 0 as well, we have c = e = f = 0. Hence, we conclude that q(x, y) = 0. The case $\gamma = (0, 1)$ is similar.

Lemma 24. We have
$$A_p^{\gamma,k} < \infty$$
, $p \in [1,\infty]$. That is, (24) holds.

Proof: The proof is by contradiction. Suppose that $A_p^{\gamma,k} = \infty$. Then, there exists a sequence $\{u_n\}_{n=1}^{\infty} \subset \Xi_p^{\gamma,k}$ such that

$$|u_n|_{0,p,\widehat{K}} = 1, \qquad \lim_{n \to \infty} |u_n|_{k+1-|\gamma|,p,\widehat{K}} = 0.$$

By the inequality (14), for an arbitrary $\varepsilon > 0$, there exists a sequence $\{q_n\} \subset \mathcal{P}_{k-|\gamma|}$

such that

$$\begin{split} \inf_{q\in\mathcal{P}_{k-|\gamma|}} \|u_n+q\|_{k+1-|\gamma|,p,\widehat{K}} &\leq \|u_n+q_n\|_{k+1-|\gamma|,p,\widehat{K}} \\ &\leq \inf_{q\in\mathcal{P}_{k-|\gamma|}} \|u_n+q\|_{k+1-|\gamma|,p,\widehat{K}} + \frac{\varepsilon}{n} \\ &\leq C|u_n|_{k+1-|\gamma|,p,\widehat{K}} + \frac{\varepsilon}{n}, \quad \lim_{n\to\infty} \|u_n+q_n\|_{k+1-|\gamma|,p,\widehat{K}} = 0. \end{split}$$

Since $|u_n|_{k+1-|\gamma|,p,\widehat{K}}$ and $|u_n|_{0,p,\widehat{K}}$ are bounded, $|u_n|_{m,p,\widehat{K}}$ $(1 \le m \le k - |\gamma|)$ is bounded as well by Gagliardo–Nirenberg's inequality (Theorem 13). That is, $||u_n||_{k+1-|\gamma|,p,\widehat{K}}$ and $\{q_n\} \subset \mathcal{P}_{k-|\gamma|}$ are bounded. Thus, there exists a subsequence $\{q_{n_i}\}$ such that q_{n_i} converges to $\bar{q} \in \mathcal{P}_{k-|\gamma|}$. In particular, we see

$$\lim_{n_i \to \infty} \|u_{n_i} + \bar{q}\|_{k+1 - |\gamma|, p, \widehat{K}} = 0.$$

Therefore, for any \Box_{lp}^{γ} , we notice that

$$0 = \lim_{n_i \to \infty} \int_{\Box_{lp}^{\gamma}} (u_{n_i} + \bar{q}) = \int_{\Box_{lp}^{\gamma}} \bar{q}$$

and $\bar{q} = 0$ by Lemma 23. This yields

$$\lim_{n_i \to \infty} \left\| u_{n_i} \right\|_{k+1-|\gamma|, p, \widehat{K}} = 0$$

which contradicts $\lim_{n_i \to \infty} \|u_{n_i}\|_{k+1-|\gamma|,p,\widehat{K}} \ge \lim_{n_i \to \infty} |u_{n_i}|_{0,p,\widehat{K}} = 1.$

Now, we have defined the set $\Xi_p^{\gamma,k}$ that satisfies (23) and the estimate (24) has been shown. Therefore, Theorem 21 has been proved by (25).

Exercise: We have shown the Theorem 21 for the case $1 \le p < \infty$. Prove Theorem 21 for the case $p = \infty$.

\S 7. The error estimation on general triangles in terms of circumradius

Using the previous results, we can obtain the error estimations on general triangles. Recall the reference triangle and the definition of the standard position of an aribtrary triangle K (Figure 2). Let $K_{\alpha\beta}$ be the triangle with the vertices $(0,0)^{\top}$, $(\alpha,0)^{\top}$, and $(0,\beta)^{\top}$. Let \hat{K} be the reference triangle with the vertices $(0,0)^{\top}$, $(1,0)^{\top}$, and $(0,1)^{\top}$.

We consider 2×2 matrices

$$A := \begin{pmatrix} \alpha \ \beta s \\ 0 \ \beta t \end{pmatrix} = \begin{pmatrix} 1 \ s \\ 0 \ t \end{pmatrix} \begin{pmatrix} \alpha \ 0 \\ 0 \ \beta \end{pmatrix}, \quad \widetilde{A} := \begin{pmatrix} 1 \ s \\ 0 \ t \end{pmatrix}, \quad D_{\alpha\beta} := \begin{pmatrix} \alpha \ 0 \\ 0 \ \beta \end{pmatrix},$$
$$A^{-1} = \begin{pmatrix} \alpha^{-1} - \alpha^{-1} s t^{-1} \\ 0 \ \beta^{-1} t^{-1} \end{pmatrix} = \begin{pmatrix} \alpha^{-1} \ 0 \\ 0 \ \beta^{-1} \end{pmatrix} \begin{pmatrix} 1 - s t^{-1} \\ 0 \ t^{-1} \end{pmatrix},$$



Figure 6. The standard position of a general triangle (reprint). The vertices are $\mathbf{x}_1 = (0,0)^{\top}$, $\mathbf{x}_2 = (\alpha,0)^{\top}$, and $\mathbf{x}_3 = (\beta s, \beta t)^{\top}$, where $s^2 + t^2 = 1$, t > 0. We assume that $0 < \beta \leq \alpha \leq h_K$. Then, $\pi/3 \leq \theta < \pi$.

and the linear transformation $\mathbf{y} = A\mathbf{x}$. The reference triangle \widehat{K} is transformed to $K_{\alpha\beta}$ by $\mathbf{y} = D_{\alpha\beta}\mathbf{x}$, and $K_{\alpha\beta}$ is transformed to K by $\mathbf{y} = \widetilde{A}\mathbf{x}$. Accordingly, $\mathcal{T}_p^k(K)$ is pulled-back to $\mathcal{T}_p^k(K_{\alpha\beta})$ by the mapping $\mathcal{T}_p^k(K) \ni v \mapsto \hat{v} := v \circ \widetilde{A}$, and $\mathcal{T}_p^k(K_{\alpha\beta})$ is pulled-back to $\mathcal{T}_p^k(\widehat{K})$ by the mapping $\mathcal{T}_p^k(K) \ni v \mapsto \hat{v} := v \circ D_{\alpha\beta}$.

By Theorem 21, for arbitrary $\alpha \geq \beta > 0$ and arbitrary $p, 1 \leq p \leq \infty$, there exists a constant $C_{k,m,p}$ depending only on k, m, p such that

(28)
$$B_p^{m,k}(K_{\alpha\beta}) := \sup_{v \in \mathcal{T}_p^k(K_{\alpha\beta})} \frac{|v|_{m,p,K_{\alpha\beta}}}{|v|_{k+1,p,K_{\alpha\beta}}} \le \alpha^{k+1-m} C_{k,m,p}$$

A simple computation confirms that $\widetilde{A}^{\top}\widetilde{A}$ has the eigenvalues $1 \pm |s|$, and $\widetilde{A}^{-1}\widetilde{A}^{-\top}$ has the eigenvalues $(1 \pm |s|)^{-1}$. That is, $\|\widetilde{A}\| = (1 + |s|)^{1/2}$, $\|\widetilde{A}^{-1}\| = (1 - |s|)^{-1/2}$, and det $\widetilde{A} = t$. Therefore, defining $\hat{v}(\mathbf{x}) = v(\widetilde{A}\mathbf{x})$ for $v \in \mathcal{T}_p^k(K)$, it follows from (12) that

$$|v|_{m,p,K} \le 2^{m\mu(p)} t^{1/p} \| \widetilde{A}^{-1} \|^m | \hat{v} |_{m,p,K_{\alpha\beta}},$$

$$2^{-(k+1)\mu(p)} t^{1/p} \| \widetilde{A} \|^{-(k+1)} | \hat{v} |_{k+1,p,K_{\alpha\beta}} \le |v|_{k+1,p,K}.$$

Combining the above inequalities and (28), we obtain

$$\frac{|v|_{m,p,K}}{|v|_{k+1,p,K}} \le c_{k,m,p} \|\widetilde{A}\|^{k+1} \|\widetilde{A}^{-1}\|^m \frac{|\hat{v}|_{m,p,K_{\alpha\beta}}}{|\hat{v}|_{k+1,p,K_{\alpha\beta}}} \le c_{k,m,p} C_{k,m,p} \|\widetilde{A}\|^{k+1} \|\widetilde{A}^{-1}\|^m \alpha^{k+1-m}$$

where $c_{k,m,p} := 2^{(k+1+m)\mu(p)}$. Hence, we obtain the following lemma.

Lemma 25. For an arbitrary triangle K in the standard position, we have

$$B_p^{m,k}(K) \leq c_{k,m,p} \|\widetilde{A}\|^{k+1} \|\widetilde{A}^{-1}\|^m B_p^{m,k}(K_{\alpha\beta})$$

$$\leq c_{k,m,p} C_{k,m,p} \|\widetilde{A}\|^{k+1} \|\widetilde{A}^{-1}\|^m \alpha^{k+1-m},$$
where $\|\widetilde{A}\| = (1+|s|)^{1/2}$ and $\|\widetilde{A}^{-1}\| = (1-|s|)^{-1/2}.$

Applying Lemma 25 to $v - \mathcal{I}_K^k v \in \mathcal{T}_p^k(K)$, we have the following corollary.

Corollary 26. For an arbitrary triangle K in the standard position, we have $|v - \mathcal{I}_{K}^{k}v|_{m,p,K} \leq c_{k,m,p}C_{k,m,p} \|\widetilde{A}\|^{k+1} \|\widetilde{A}^{-1}\|^{m} \alpha^{k+1-m} |v|_{k+1,p,K}, \quad \forall v \in W^{k+1,p}(K).$

We would like to obtain upper bounds of $\|\widetilde{A}\|$ and $\|\widetilde{A}^{-1}\|$. From Lemma 25, we obviously have $\|\widetilde{A}\| \leq \sqrt{2}$. For $\|\widetilde{A}^{-1}\|$, we observe that

$$\begin{split} \|\widetilde{A}^{-1}\| &= \frac{1}{(1-|s|)^{1/2}} = \frac{(1+|s|)^{1/2}}{t} \qquad (\because \ s^2 + t^2 = 1) \\ &\leq \frac{2^{1/2}\alpha\beta h_K}{\alpha\beta h_K t} = \frac{\alpha\beta h_K}{2^{1/2}h_K|K|} \qquad \left(\because |K| = \frac{1}{2}\alpha\beta t\right) \\ &= \frac{2^{3/2}R_K}{h_K} \qquad \left(\because R_K = \frac{\alpha\beta h_K}{4|K|}\right). \end{split}$$

Thus, redefining the constant $C_{k,m,p}$, we obtain the following theorem.

Theorem 27. Suppose that a triangle K is in the standard position. Let k, m be integers with $k \ge 1$, $m = 0, \dots, k$ and $1 \le p \le \infty$. Then, the following estimate holds:

$$B_p^{m,k}(K) := \sup_{v \in \mathcal{T}_p^k(K)} \frac{|v|_{m,p,K}}{|v|_{k+1,p,K}} \le C_{k,m,p} \left(\frac{R_K}{h_K}\right)^m \alpha^{k+1-m},$$

where R_K is the circumradius of K, and $C_{k,m,p}$ is a constant depending only on k, m, and p.

Now, let K be an arbitrary triangle. Note that $\alpha \leq h_K$ and the Sobolev norms are affected by rotations if $p \neq 2$ up to an constant (see (13)). Then, with rewriting the constant, we obtain the following corollary from Theorem 27, that is the main theorem of this survey (reprint of Theorem 10).

Corollary 28. Let K be an arbitrary triangle with circumradius R_K . Let k and m be intergers with $k \ge 1$ and $m = 0, \dots, k$. Let p, $1 \le p \le \infty$. For the Lagrange interpolation $\mathcal{I}_K^k v$ of degree k on K, the following estimate holds: for any $v \in W^{2,p}(K)$,

$$B_{p}^{m,k}(K) := \sup_{u \in \mathcal{T}_{p}^{k}(K)} \frac{|u|_{m,p,K}}{|u|_{k+1,p,K}} \leq C_{k,m,p} \left(\frac{R_{K}}{h_{k}}\right)^{m} h_{K}^{k+1-m}$$
$$|v - \mathcal{I}_{K}^{k}v|_{m,p,K} \leq C_{k,m,p} \left(\frac{R_{K}}{h_{K}}\right)^{m} h_{K}^{k+1-m} |v|_{k+1,p,K},$$

where $C_{k,m,p}$ depends only on k, m, and p.

Remarks: (1) Let $\Omega \subset \mathbb{R}^2$ be a bounded polygonal domain. We compute a numerical solution of the Poisson equation

$$-\Delta u = f \text{ in } \Omega, \quad u = 0 \text{ on } \partial \Omega$$

by the conforming piecewise kth-order finite element method on simplicial elements. To this end, we construct a triangulation \mathcal{T}_h of Ω and consider the piecewise \mathcal{P}_k continuous function space $S_h \subset H_0^1(\Omega)$. The weak form of the Poisson equation is

$$\int_{\Omega} \nabla u \cdot \nabla v \mathrm{d}\mathbf{x} = \int_{\Omega} f v \mathrm{d}\mathbf{x}, \quad \forall v \in H_0^1(\Omega),$$

and the finite element solution is defined as the unique solution $u_h \in S_h$ of

$$\int_{\Omega} \nabla u_h \cdot \nabla v_h \mathrm{d}\mathbf{x} = \int_{\Omega} f v_h \mathrm{d}\mathbf{x}, \quad \forall v_h \in S_h.$$

Céa's Lemma implies that the error $|u - u_h|_{1,2,\Omega}$ is estimated as

(29)
$$|u - u_h|_{1,2,\Omega} \le \left(\sum_{K \in \mathcal{T}_h} |u - \mathcal{I}_K^k u|_{1,2,K}^2\right)^{1/2}.$$

Combining (29) and Corollary 28 with $p = 2, k \ge 2, m = 1$, we have

$$|u - u_h|_{1,2,\Omega} \le C \left(\sum_{K \in \mathcal{T}_h} |u - \mathcal{I}_K^k u|_{k+1,2,K}^2 \right)^{1/2}$$

$$\le C \left(\sum_{K \in \mathcal{T}_h} (R_K h_K^{k-1})^2 |u|_{k+1,2,K}^2 \right)^{1/2}$$

$$\le C \max_{K \in \mathcal{T}_h} (R_K h_K^{k-1}) |u|_{k+1,2,\Omega}.$$

Therefore, if $\max_{K \in \mathcal{T}_h} (R_K h_K^{k-1}) \to 0$ as $h \to 0$ and $u \in H^{k+1}(\Omega)$, the finite element solution u_h converges to the exact solution u even if there exist many skinny elements violating the shape regularity condition or the maximum angle condition in \mathcal{T}_h .

Recall the triangle depicted in Figure 3 (right) with vertices $(0,0)^{\top}$, $(h,0)^{\top}$, and $(h^{\alpha}, h^{\beta})^{\top}$ with $R_K = \mathcal{O}(h^{1+\alpha-\beta})$. Suppose now that $\alpha + 1 \leq \beta < 2 + \alpha$. If a sequence of triangulations contains those triangles, and k = 1, then $\max_{K \in \mathcal{T}_h} R_K = \mathcal{O}(1)$ and the piecewise linear Lagrange FEM might not converge. However, if k = 2, then $\max_{K \in \mathcal{T}_h} (R_K h_K) = \mathcal{O}(h^{2+\alpha-\beta})$, and the finite element solution certainly converges to the exact solution, although the convergence rate is worse than expected. This means that "bad" triangulations with many very skinny triangles can be remedied by using higher-order Lagrange elements.

§8. Numerical experiments

To confirm the results obtained, we perform numerical experiments similar to those in [11]. Let $\Omega := (-1, 1) \times (-1, 1)$, $f(x, y) := a^2/(a^2 - x^2)^{3/2}$, and $g(x, y) := (a^2 - x^2)^{1/2}$ with a := 1.1. Then we consider the following Poisson equation: Find $u \in H^1(\Omega)$ such that

(30)
$$-\Delta u = f \quad \text{in } \Omega, \qquad u = g \quad \text{on } \partial \Omega.$$

The exact solution of (30) is u(x, y) = g(x, y) and its graph is a part of the cylinder. For a given positive integer N and $\alpha > 1$, we consider the isosceles triangle with base length h := 2/N and height $2/\lfloor 2/h^{\alpha} \rfloor \approx h^{\alpha}$, as shown in Figure 7. Let R be the circumradius of the triangle. For comparison, we also consider the isosceles triangle with base length h and height h/2 for $\alpha = 1$. We triangulate Ω with this triangle, as shown in Figure 7. Let τ_h be the triangulation. As usual, the set S_h of piecewise linear functions on τ_h and its subsets are defined by

$$S_h := \left\{ v_h \in C(\overline{\Omega}) \mid v|_K \in \mathcal{P}_1(K), \ \forall K \in \tau_h \right\},$$

$$S_{hg} := \left\{ v_h \in S_h \mid v_h = g \text{ at boundary nodes} \right\},$$

$$S_{h0} := \left\{ v_h \in S_h \mid v_h = 0 \text{ on } \partial\Omega \right\}.$$

Then, the piecewise linear finite element method for (30) is defined as follows: Find $u_h \in S_{hg}$ such that

$$(\nabla u_h, \nabla v_h)_{\Omega} = (f, v_h)_{\Omega}, \quad \forall v_h \in S_{h0},$$

where $(\cdot, \cdot)_{\Omega}$ is the inner product of $L^2(\Omega)$. By Céa's lemma and the result obtained, we obtain the estimation

$$|u - u_h|_{1,2,\Omega} \le \inf_{v_h \in S_{hg}} |u - v_h|_{1,2,\Omega} \le \left(\sum_{K \in \tau_h} |u - \mathcal{I}_K^1 u|_{1,2,K}^2\right)^{1/2} \le CR|u|_{2,2,\Omega}.$$

The behavior of the error is given in Figure 7. The horizontal axis represents the mesh size measured by the maximum diameter of triangles in the meshes and the vertical axis represents the error associated with FEM solutions in the H^1 semi-norm. The graph clearly shows that the convergence rates worsen as α approaches 2.0. For $\alpha = 2.1$, the FEM solutions even diverge. This is a counterexample to the vaguely believed dogma that "FEM solutions always converge to the exact solution if $h \to 0$ ". See also [23].

We replot the same data in Figure 8, in which the horizontal axis represents the maximum of the circumradius of triangles in the meshes. Figure 8 shows convergence rates are almost the same in all cases if we measure these with the circumradius. These experiments strongly support that our theoretical results are correct and optimal.



Figure 7. Triangulation of Ω with N = 12 and $\alpha = 1.6$, and the errors for FEM solutions in the H^1 -norm. The horizontal axis represents the maximum diameter of the triangles and the vertical axis represents the H^1 -norm of the errors of the FEM solutions. The number next to the symbol indicates the value of α .



Figure 8. Replotted data: the errors in the H^1 -norm of FEM solutions measured using the circumradius. The horizontal axis represents the maximum circumradius of the triangles.

Acknowledgments

We thank Dr. Théophile Chaumont-Frelet and the referees for their valuable comments.

References

- [1] R.A. Adams, J.J.F. Fournier: Sobolev Spaces, 2nd edition, Pure and Applied Mathematics 140, Elsevier/Academic Press, New York, 2003.
- [2] T. Apel: Anisotropic Finite Element: Local estimates and applications, Advances in Numerical Mathematics. B.G. Teubner, Stuttgart, 1999.

- [3] K.E. Atkinson: An Introduction to Numerical Analysis, 2nd edition, John Wiley & Sons, New York, 1989.
- [4] I. Babuška, A.K. Aziz: On the angle condition in the finite element method, SIAM J. Numer. Anal. 13 (1976), 214–226.
- [5] R.E. Barnhill, J.A. Gregory: Sard kernel theorems on triangular domains with application to finite element error bounds. Numer. Math., **25** (1976), 215-229.
- [6] S.C. Brenner, L.R. Scott: The Mathematical Theory of Finite Element Methods. 3rd edition. Texts in Applied Mathematics 15, Springer, New York, 2008.
- [7] H. Brezis: Functional Analysis, Sobolev Spaces and Partial Differential Equations. Universitext, Springer, New York, 2011.
- [8] P.G. Ciarlet: The Finite Element Methods for Elliptic Problems. Classics in Applied Mathematics 40, SIAM, Philadelphia, 2002, Reprint of the 1978 original (North Holland, Amsterdam).
- [9] J. Brandts, S. Korotov, M. Křížek: On the equivalence of regularity criteria for triangular and tetrahedral finite element partitions, Comput. Math. Appl. 55 (2008), 2227–2233.
- [10] A. Ern, J-L. Guermond: Theory and Practice of Finite Elements. Applied Mathematical Sciences 159, Springer, New York, 2004.
- [11] A. Hannukainen, S. Korotov, M. Křížek: The maximum angle condition is not necessary for convergence of the finite element method, Numer. Math., 120, 79–88 (2012)
- [12] R.A. Horn, C.R. Johnson: Topics in Matrix Analysis, Cambridge Univ. Press, 1991.
- [13] P. Jamet: Estimations d'erreur pour des elements finis droits presque degeneres. R.A.I.R.O. Anal. Numer., 10 (1976), 43–61.
- [14] K. Kobayashi: On the interpolation constants over triangular elements (in Japanese), RIMS Kokyuroku, 1733 (2011), 58-77.
- [15] K. Kobayashi, T. Tsuchiya: A Babuška-Aziz type proof of the circumradius condition, Japan J. Indust. Appl. Math., **31** (2014), 193-210.
- [16] K. Kobayashi, T. Tsuchiya: On the circumradius condition for piecewise linear triangular elements, Japan J. Indust. Appl. Math., 32 (2015), 65–76.
- [17] K. Kobayashi, T. Tsuchiya: A priori error estimates for Lagrange interpolation on triangles. Appl. Math., Praha 60 (2015), 485–499.
- [18] K.Kobayashi, T.Tsuchiya: Extending Babuška-Aziz's theorem to higher order Lagrange interpolation, Appl. Math., Praha 61 (2016), 121–133.
- [19] M. Křížek: On semiregular families of triangulations and linear interpolation. Appl. Math., Praha 36 (1991), 223–232.
- [20] M. Křížek: On the maximum angle condition for linear tetrahedral elements. SIAM J. Numer. Anal., 29 (1992), 513–520.
- [21] A. Kufner, O. John, S. Fučík: Function Spaces. Noordhoff International Publishing, Leyden, 1977.
- [22] X. Liu, F. Kikuchi: Analysis and estimation of error constants for P_0 and P_1 interpolation over triangular finite element, J. Math. Sci. Univ. Tokyo, **17** (2010), 27–78.
- [23] *P. Oswald*: Divergence of fem: Babuška-Aziz triangulations revisited, Appl. Math., Praha **60** (2015), 473–484.
- [24] N.A. Shenk: Uniform error estimates for certain narrow Lagrange finite elements. Math. Comp., 63 (1994), 105–119.
- [25] T. Yamamoto: Introduction to Numerical Analysis, 2nd edition, (In Japanese), Saiensusha, 2003.
- [26] T. Yamamoto: Elements of Matrix Analysis (in Japanese), Saiensu-sha, 2010.
- [27] A. Zeníšek: The convergence of the finite element method for boundary value problems of a system of elliptic equations. (in Czech) Appl. Math., 14 (1969), 355–377.
- [28] M. Zlámal: On the finite element method, Numer. Math. 12 (1968), 394–409.