Applications of Mathematics

Peter Oswald

Divergence of FEM: Babuška-Aziz triangulations revisited

Applications of Mathematics, Vol. 60 (2015), No. 5, 473–484

Persistent URL: `http://dml.cz/dmlcz/144387`

DIVERGENCE OF FEM: BABUŠKA-AZIZ TRIANGULATIONS REVISITED

Peter Oswald, Dresden

(Received March 28, 2015)

Dedicated to Ivo Babuška

Abstract. By re-examining the arguments and counterexamples in I. Babuška, A. K. Aziz (1976) concerning the well-known maximum angle condition, we study the convergence behavior of the linear finite element method (FEM) on a family of distorted triangulations of the unit square originally introduced by H. Schwarz in 1880. For a Poisson problem with polynomial solution, we demonstrate arbitrarily slow convergence as well as failure of convergence if the distortion of the triangulations grows sufficiently fast. This seems to be the first formal proof of divergence of the FEM for a standard elliptic problem with smooth solution.

Keywords: finite elements; error bounds; divergence; maximum angle condition; triangulation

MSC 2010: 65N30, 65N12, 65N15

1. Introduction

All existing convergence results for the finite element method (FEM) but in the one-dimensional case are based on some regularity assumptions for the underlying partitions. In two dimensions, these regularity conditions were derived by analyzing the energy norm error of the local polynomial interpolation operator. The most widely known sufficient condition, the maximum angle condition, goes back to Babuška and Aziz [2], and Jamet [6]. The fact that the maximum angle condition is not necessary for the convergence of the FEM was demonstrated in [5]. For approaches to proving FEM convergence and error estimates for certain classes of anisotropic partitions based on other local interpolation operators, see [1]. We also note that in [2], Section 3, it was shown that the optimal error estimates for smooth solutions may not hold if the maximum angle condition is violated. Numerical experiments [9] and [4] with the triangulations introduced in [2] suggest that the FEM may not even converge to the solution if the triangles are too distorted.

However, it seems that despite the relatively long history and importance of the finite element method no formal proof of the FEM failing to converge to the solution of standard elliptic boundary value problems appeared in the literature. It is the aim of this note to close this gap by re-examining the examples and arguments from the Babuška-Aziz paper [2]. For this purpose we consider the particular Poisson problem

$$-\Delta u(x,y) = 1, \quad (x,y) \in \Omega = [0,1]^2, \tag{1}$$

on the unit square equipped with homogeneous Dirichlet boundary conditions along its vertical sides $x = 0, 1$, and periodic boundary conditions on its horizontal sides $y = 0, 1$. Its solution is the quadratic polynomial

$$u(x,y) = \frac{1}{2}x(1-x), \quad (x,y) \in \Omega. \tag{2}$$

Denote by $V_{n,m} \subset H^1_D(\Omega)$ the family of linear finite element spaces over triangulations $\mathcal{T}_{n,m}$, $m \geqslant n$, already used in [2] and probably proposed for the first time in 1880 by Schwarz [10] in his construction of triangulated surfaces with arbitrarily large area inscribed into a cylinder (the so-called Schwarz lantern). Figure 1 shows $\mathcal{T}_{4,8}$, the construction of $\mathcal{T}_{n,m}$ for general n, m is obvious from this example. We consider the best approximation error

$$E_{n,m}(u) := \inf_{S \in V_{n,m}} \|u - S\|_{H^1(\Omega)} \tag{3}$$

for the particular u given by (2) and show that

$$E_{n,m}(u) \approx \min(1, m/n^2), \quad m \geqslant n, \ n \to \infty. \tag{4}$$

This implies that the linear FEM with a sequence of FE spaces $V_{n,m}$ does not converge in energy norm to the solution of (1) if the condition $m/n^2 = o(1)$ is violated. The latter fact was conjectured in [9] and [4] based on numerical experiments.

The paper is organized as follows. In the Section 2, we repeat the arguments of [2] and show that they already imply the non-convergence result under the more restrictive assumption $m/n^5 \geqslant c_0$. In Section 3 our main result (4) is proved in full generality, by directly examining the FE solution of (1) in $V_{n,m}$.

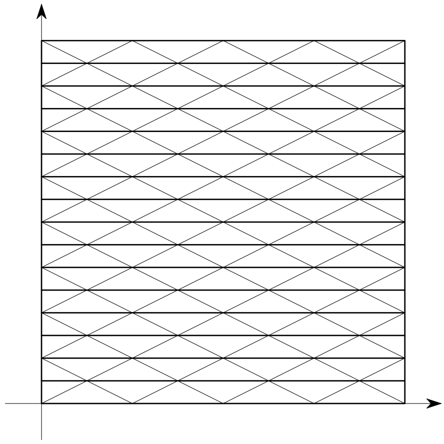

Figure 1. Babuška-Aziz triangulation $\mathcal{T}_{4,8}$.

2. Revisiting the Babuška-Aziz argument

To set the scene, we clarify that the energy space $H^1_D(\Omega)$ is the subspace of all functions $v \in H^1(\Omega)$ 1-periodic in y such that $v(0,y) = v(1,y) = 0$ (in the usual weak sense), and that the linear finite element functions in $V_{n,m}$ are 1-periodic in y and satisfy homogeneous boundary conditions for $x = 0$ and $x = 1$ as well. By C, $c_0, c_1, \ldots$ we denote positive constants that do not depend on the parameters n and m, and the notation $A \approx B$ is used if $c_0 \leqslant A/B \leqslant c_1$ independently of n, m. Further notation will be explained in the proofs.

We will first show a weaker result following line by line the arguments in [2], Section 3.

Theorem 1. *For any sequence of pairs (n,m) such that $m/n^5 \geqslant c_0 > 0$ as $n \to \infty$, there is a positive constant $c_1 > 0$ such that*

$$E_{n,m}(u) \geqslant c_1, \tag{5}$$

where u is given by (2), and $E_{n,m}(u)$ is the error of the best $H^1(\Omega)$-approximation w.r.t. $V_{n,m}$.

P r o o f. Assume on the contrary that $E_{n,m}(u) \to 0$ for some sequence of pairs (n,m) such that $m/n^5 \geqslant c_0$ as $n \to \infty$. Equivalently, assume that there are linear FE functions $S \in V_{n,m}$ such that

$$\|u - S\|_{H^1(\Omega)} = o(1)$$

for this sequence of (n,m). Introduce the notation $h=1/2n$, $k=1/2m$, and

$$x_i=ih,\quad i=0,\dots,2n,\qquad y_j=jk,\quad j=0,\dots,2m.$$

Then our assumptions on (n,m) are equivalent to $k=O(h^5)$ and $h\to 0$. Since $S_y:=\partial S/\partial y$ is constant on any triangle τ in $\mathcal{T}_{n,m}$ and $u_y=0$ (u does not depend on y), we have

$$hk|S_y(x,y)|^2=\int_\tau |S_y|^2\leqslant \|u-S\|^2_{H^1(\Omega)},\quad (x,y)\in\tau,$$

(note that $S_y(x,y)=0$ on the triangles with edges on the vertical Dirichlet boundaries of the square), and thus

$$hk\|S_y(x,y)\|^2_{L_\infty(\Omega)}=o(1). \tag{6}$$

Our next goal is to get bounds on the sequence $Z_i:=\int_0^1(u(x_i,y)-S(x_i,y))\,\mathrm{d}y$ and its second-order differences. On the one hand, since $u(x_i,y)-S(x_i,y)=\int_0^{x_i}(u_x(x,y)-S_x(x,y))\,\mathrm{d}x$ due to the homogeneous Dirichlet boundary conditions at $x=0$, we see that

$$|Z_i|\leqslant\int_0^1\int_0^{x_i}|u_x(x,y)-S_x(x,y)|\,\mathrm{d}x\,\mathrm{d}y\leqslant\|u-S\|_{H^1(\Omega)}.$$

Thus, uniformly in $i=1,\dots,2n-1$ we have

$$|Z_i|=o(1),\quad Z_0=Z_{2n}=0. \tag{7}$$

On the other hand, for estimating the second-order differences $\Delta^2 Z_i=Z_{i-1}-2Z_i+Z_{i+1}$, $i=1,\dots,2n-1$, we use the fact that

$$\Delta_h^2 u(x,y)=u(x-h,y)-2u(x,y)+u(x+h,y)=-h^2,\quad x\in[h,1-h],\ y\in[0,1],$$

is constant and negative. Due to the linearity of $S(x,y)$ along the sides of length $2h$ in x-direction of the triangles in $\mathcal{T}_{n,m}$, the second-order differences $\Delta_h^2 S(x,y)$ vanish at their midpoints, i.e.,

$$\Delta_h^2 S(x_{2i-1},y_{2j})=0,\quad i=1,\dots,n,\ j=0,\dots,m,$$

and

$$\Delta_h^2 S(x_{2i},y_{2j-1})=0,\quad i=1,\dots,n-1,\ j=1,\dots,m.$$

Consequently, for each $y \in [0,1]$ and $i = 1, \ldots, 2n-1$, there is a y_j such that $|y - y_j| \leqslant k$ and $\Delta_h^2 S(x_i, y_j) = 0$. This implies the bound

$$|\Delta_h^2 S(x_i, y)| \leqslant k \|\Delta_h^2 S_y(x_i, y)\|_{L_\infty([0,1])} \leqslant 4k \|S_y(x, y)\|_{L_\infty(\Omega)},$$

which, together with (6), allows us to estimate the remaining integral in

$$\text{(8)}\quad \Delta^2 Z_i = \int_0^1 (\Delta_h^2 u(x_i, y) - \Delta_h^2 S(x_i, y))\,\mathrm{d}y = -h^2 - \int_0^1 \Delta_h^2 S(x_i, y)\,\mathrm{d}y =: -h^2 - D_i$$

by

$$|D_i| \leqslant 4k \|S_y(x, y)\|_{L_\infty(\Omega)} = o(k^{1/2} h^{-1/2}).$$

Using the assumption $k = O(h^5)$, this results in

$$\Delta^2 Z_i = -h^2 + o(k^{1/2} h^{-1/2}) = (-1 + o(1)) h^2 \leqslant -\frac{1}{2} h^2, \tag{9}$$

for large enough n (small enough h), uniformly in $i = 1, \ldots, 2n-1$.

To arrive at a contradiction, we invoke the discrete maximum principle for the new sequence

$$\xi_i := Z_i - \frac{1}{4} x_i (1 - x_i), \quad i = 0, 1, \ldots, 2n.$$

Indeed, $\xi_0 = \xi_{2n} = 0$, and

$$\Delta^2 \xi_i = \Delta^2 Z_i + \frac{1}{2} h^2 \leqslant 0, \quad i = 1, \ldots, 2n-1,$$

for large enough n by (9). Thus, $\xi_i \geqslant 0$ and in particular

$$0 \leqslant \xi_n = Z_n - \frac{1}{16} = -\frac{1}{16} + o(1),$$

which is the desired contradiction due to (7). $\square$

R e m a r k s. 1) One can only wonder why in [2], Section 3, the same condition $m/n^5 \geqslant c_0 > 0$ as in Theorem 1 was assumed to establish a weaker result, namely $E_{n,m}(u) \neq O(n^{-1})$, $n \to \infty$.

2) The above non-convergence result can be extended to the range $m \geqslant c_0 n^{5/2}$ or, equivalently, $k = O(h^{5/2})$ using the following observations: Since $\Delta_h^2 S(x_i, \cdot)$ is a piecewise linear spline on a partition of $[0,1]$ into $2m = k^{-1}$ intervals $\delta_j = [(j-1)k, jk]$, with every second knot value vanishing, we have

$$|D_i| \leqslant \frac{k^2}{2} \sum_j |(\Delta_h^2 S(x_i, y))_y|_{\delta_j}| \leqslant \frac{k^2}{2} (I_{i-1} + 2I_i + I_{i+1}),$$

where $I_i = \sum_j |S_y(x_i, y)\big|_{\delta_j}|$ can be estimated by

$$|I_i| \leqslant \sum_j |(S_y)\big|_{\tau_{ij}}| \leqslant k^{-1/2}\Big(\sum_j (S_y)^2\big|_{\tau_{ij}}\Big)^{1/2}, \quad i = 1, \ldots, 2n-1.$$

Here τ_{ij}, $j = 1, \ldots, 2m$, denote the $2m$ obtuse triangles along the vertical line $x = x_i$. But

$$\sum_j (S_y)^2\big|_{\tau_{ij}} = (hk)^{-1} \sum_j \int_{\tau_{ij}} S_y(x, y)^2 \, \mathrm{d}x \, \mathrm{d}y \leqslant (hk)^{-1} \|u - S\|^2_{H^1(\Omega)},$$

therefore we get by substitution (note that $I_0 = I_{2n} = 0$)

$$|D_i| \leqslant \frac{4k^2}{2(hk)^{1/2}k^{1/2}} \|u - S\|_{H^1(\Omega)} = o(kh^{-1/2}),$$

uniformly in $i = 1, \ldots, 2n-1$. By using this replacement of the weaker $o(k^{1/2}h^{-1/2})$ estimate in the above argument, the claimed improvement follows.

3. Main result

Our main result (4) will be obtained by using $E_{n,m}(u) \approx \|u - S_{n,m}\|_{H^1(\Omega)}$, where $S_{n,m}(x, y)$ denotes the Galerkin solution of the problem (1) w.r.t. the FE space $V_{n,m}$.

Theorem 2. *We have*

$$\|u - S_{n,m}\|_{H^1(\Omega)} \approx \min(1, m/n^2), \quad m \geqslant n. \tag{10}$$

In particular, the FE Galerkin solutions $S_{n,m}$ converge in energy norm to the solution (2) *of* (1) *if and only if the condition $m/n^2 = o(1)$ is fulfilled.*

Proof. To simplify notation, we set $S = S_{n,m}$ for arbitrarily fixed $m \geqslant n$, and

$$\kappa = k/h^2$$

(we silently use the same notation as introduced in the proof of Theorem 1 in the previous section). Thus, proving (10) is equivalent to showing

$$\|u - S\|_{H^1(\Omega)} \approx \min(1, \kappa^{-1}), \quad m \geqslant n. \tag{11}$$

The upper bound in (11) is easy. First of all, the Galerkin projectors are uniformly bounded due to the $H^1_D(\Omega)$-ellipticity of (1), thus $\|u-S\|_{H^1(\Omega)} \leqslant C$, independently of n and m. On the other hand, using Cea's lemma and a Taylor expansion argument on each triangle of $\mathcal{T}_{n,m}$ for the nodal interpolation function $I(x,y) \in V_{n,m}$ of the polynomial u, we get

$$\|u-S\|_{H^1(\Omega)} \leqslant C\|u-I\|_{H^1(\Omega)} \leqslant Ch^2/k = C\kappa^{-1}.$$

We spare the reader the trivial calculations, we just note that the dominating part in the computation of $\|u-I\|^2_{H^1(\Omega)}$ is

$$\|(u-I)_y\|^2_{L_2(\Omega)} = \|I_y\|^2_{L_2(\Omega)} \approx h^4/k^2 = \kappa^{-2}.$$

We now focus on proving a matching lower estimate in (11). By the construction of our problem (1) and the spaces $V_{n,m}$, we can use shift-invariance with stepsize $2k$ in y-direction to see that all nodal values of $S(x,y)$ along the lines $x = x_i$ must be the same, let us denote them by s_i, $i = 1, \ldots, 2n-1$. For convenience, set $s_0 = s_{2n} = 0$ and $s_{-1} = -s_1$, $s_{2n+1} = -s_{2n-1}$. This corresponds to an odd extension of the problem (1) to $\Omega' = [-1,1] \times [0,1]$, followed by a 2-periodic extension in x-direction. Since this extended problem and its discretization w.r.t. the analogously extended FE space are invariant under the coordinate transformation $x \to 1-x$, we also have $s_{2n-i} = s_i$, $i = -1, \ldots, 2n+1$.

We first deal with the case $\kappa \geqslant 1$. A simple calculation shows that

$$\|u-S\|^2_{H^1(\Omega)} \geqslant \|S_y\|^2_{L_2(\Omega)} = \frac{1}{k}\sum_{i=1}^{2n-1}\left(\frac{z_i}{2k}\right)^2 hk = \frac{1}{4h^3\kappa^2}\|z\|_2^2, \tag{12}$$

where the vector z has the entries $z_i = s_{i-1} - 2s_i + s_{i+1}$, $i = 1, \ldots, 2n-1$. Thus, if we establish the inequality

$$\|z\|_2^2 \geqslant c_0 h^3 \tag{13}$$

for some $c_0 > 0$, we arrive at the desired lower estimate for $\|u-S\|_{H^1(\Omega)}$ in the case $\kappa \geqslant 1$.

To do this, we have a closer look at the linear system representing the Galerkin discretization of the problem in terms of the nodal values s_i. Without going into the details, we report the final result of the elementary calculation of the stencils for the stiffness matrix and the right-hand sides:

$$(k^{-2}-h^{-2})\frac{s_{i-2}+s_{i+2}}{2} - 2k^{-2}(s_{i-1}+s_{i+1}) + (3k^{-2}+h^{-2})s_i$$
$$= b_i := \begin{cases} \frac{5}{3}, & i = 1, 2n-1, \\ 2, & 1 < i < 2n-1. \end{cases}$$

Here, a common scaling factor hk has been discarded. Rewriting this system in terms of the variables z_i (note that $z_0 = z_{2n} = 0$), we get

$$(k^{-2} - h^{-2})\frac{z_{i-1} + z_{i+1}}{2} - (k^{-2} + h^{-2})z_i = b_i, \quad i = 1, \ldots, 2n-1,$$

and multiplying by $-2k^2h^{-2}$, we arrive at

$$2(h^{-2} + \kappa^2)z_i - (h^{-2} - \kappa^2)(z_{i-1} + z_{i+1}) = -2\kappa^2h^2b_i, \quad i = 1, \ldots, 2n-1. \tag{14}$$

Now observe that the coefficient matrix A of the Toeplitz system (14) is obviously a positive definite M-matrix and can be decomposed as

$$A = (h^{-2} - \kappa^2)B + 4\kappa^2 \operatorname{Id},$$

where Id is the identity matrix and B the well-known tridiagonal Toeplitz matrix with diagonal entries $b_{ii} = 2$, and $b_{ii'} = -1$ for $|i - i'| = 1$, both of size $2n - 1$. Since the spectral decomposition of B is well-known, we can pick a suitable multiple w of the eigenvector corresponding to the smallest eigenvalue

$$\lambda_{\min}(A) = 4\kappa^2 + 4(h^{-2} - \kappa^2)\sin^2\left(\frac{\pi}{4n}\right) \leqslant 4(\kappa^2 + C)$$

of A, namely, set

$$w_i = -4\kappa^2h^2\sin\left(\frac{\pi i}{2n}\right), \quad i = 1, \ldots, 2n-1.$$

The constant factor $-4\kappa^2h^2$ was chosen such that

$$(Az)_i = -2\kappa^2h^2b_i \leqslant w_i = \lambda_{\min}^{-1}(Aw)_i, \quad i = 1, \ldots, 2n-1.$$

By the M-matrix property, A is inverse isotone, i.e., $Ax \leqslant Ay$ implies $x \leqslant y$ for any two vectors x, y. Therefore, we conclude that

$$z_i \leqslant \lambda_{\min}^{-1}w_i \leqslant -\frac{\kappa^2}{\kappa^2 + C}h^2\sin\left(\frac{\pi i}{2n}\right), \quad i = 1, \ldots, 2n-1.$$

Thus, we obtain the desired estimate (13).

For the range $\kappa \leqslant 1$, we have to argue in a slightly different way, as the above argument fails to give the desired lower bound $\|u - S\|_{H^1(\Omega)} \geqslant c_1$, since in the estimate (12) we neglected the term $\|(u - S)_x\|^2_{L_2(\Omega)}$, which now comes into play. However, instead of examining this term, we return to the original proof idea from [2] outlined in Section 2, and derive a precise estimate for the quantities D_i appearing in the representation (8) of the second-order differences $\Delta^2 Z_i$. It is easy to verify that

$$D_i = \frac{1}{4}(s_{i-2} - 2s_i + s_{i+2}) = \frac{1}{4}(z_{i-1} + 2z_i + z_{i+1}), \quad i = 1, \ldots, 2n-1, \tag{15}$$

just recall that $\Delta_h^2 S(x_i, \cdot)$ is a linear spline with knots at y_j for which every second nodal value vanishes, and compute the remaining non-zero nodal values $\Delta_h^2 S(x_{2i-1}, y_{2j-1})$ and $\Delta_h^2 S(x_{2i}, y_{2j})$, which again do not depend on j. The vector z with the entries z_i satisfies (14), $z_0 = z_{2n} = 0$, and $z_{2n-i} = z_i$, $i = 1, \ldots, 2n-1$, by the above mentioned properties of the s_i. From (14), by introducing a new vector $\hat{z}$ with entries $\hat{z}_i = z_i + h^2$, $i = 1, \ldots, 2n-1$, and recalling that $b_i = 2$ for $i = 2, \ldots, 2n-2$, we obtain the tridiagonal Toeplitz system

$$\begin{aligned} &2\beta\hat{z}_1 - \alpha\hat{z}_2 = 1 - \frac{1}{3}\kappa^2 h^2, \\ &2\beta\hat{z}_i - \alpha(\hat{z}_{i-1} + \hat{z}_{i+1}) = 0, \quad 1 < i < 2n-1, \\ &2\beta\hat{z}_{2n-1} - \alpha\hat{z}_{2n-2} = 1 - \frac{1}{3}\kappa^2 h^2, \end{aligned} \tag{16}$$

where $\alpha = h^{-2} - \kappa^2$, $\beta = h^{-2} + \kappa^2$.

The system (16) can be solved explicitly, by observing that the homogeneous equations with $i = 2, \ldots, 2n-2$ in (16) are simultaneously satisfied by the ansatz

$$\hat{z}_i = a\lambda_1^i + b\lambda_2^i, \quad i = 1, \ldots, 2n-1,$$

where

$$\lambda_1 = \frac{1 + \kappa h}{1 - \kappa h}, \quad \lambda_2 = \frac{1 - \kappa h}{1 + \kappa h}$$

are the roots of the characteristic polynomial $\alpha\lambda^2 - 2\beta\lambda + \alpha$ associated with (16). Since $\hat{z}_{2n-i} = \hat{z}_i$ and $\lambda_2 = \lambda_1^{-1}$, the formula for $\hat{z}_i$ can be simplified to

$$\hat{z}_i = A(\lambda_1^i + \lambda_1^{2n-i}), \quad i = 1, \ldots, 2n-1.$$

The free parameter A is determined from the equation for $i = 1$ in (16). This calculation gives

$$A = \frac{1 - \frac{1}{3}\kappa^2 h^2}{\alpha(1 + \lambda_1^{2n})} = \lambda_1^n \frac{(1 - \frac{1}{3}\kappa^2 h^2)h^2}{(1 - \kappa^2 h^2)(1 + \lambda_1^{2n})}.$$

After substitution we arrive at

$$\hat{z}_i = \frac{\lambda_1^{n-i} + \lambda_1^{i-n}}{\lambda_1^n + \lambda_1^{-n}} \frac{1 - \frac{1}{3}\kappa^2 h^2}{1 - \kappa^2 h^2} h^2 \geqslant c_0 h^2, \quad i = 1, \ldots, 2n-1, \tag{17}$$

for some $c_0 > 0$, since for $-n \leqslant k \leqslant n$ and $\kappa \leqslant 1$ we have $\kappa^2 h^2 \leqslant 1/4$ and

$$0 < \lambda_1^{-n} \leqslant \lambda_1^k \leqslant \lambda_1^n = \Big(\frac{1+\kappa/2n}{1-\kappa/2n}\Big)^n \leqslant C\mathrm{e}^{\kappa}.$$

Now we are ready to conclude. Taking into account $z_i = \hat{z}_i - h^2$, $i = 1, \ldots, 2n-1$, and substituting (15) into (8), we get

$$\Delta^2 Z_i = -h^2 - D_i = -\frac{1}{4}(\hat{z}_{i-1} + 2\hat{z}_i + \hat{z}_{i+1}) \leqslant -c_0 h^2, \quad i = 2, \ldots, 2n-2,$$

and $\Delta^2 Z_1 = \Delta^2 Z_{2n-1} = -h^2 - D_1 = -\frac{1}{4}h^2 - \frac{1}{4}(2\hat{z}_1 + \hat{z}_2) \leqslant -\frac{1}{4}(1+3c_0)h^2$. This gives the bound in (9) with a slightly different constant under the weaker assumption $\kappa = kh^{-2} \leqslant 1$, all other steps in the non-convergence proof in Theorem 1 remain the same. Since the latter statement is equivalent to proving $\|u - S\|_{H^1(\Omega)} \geqslant c_1$ for some $c_1 > 0$ and $\kappa \leqslant 1$, this finishes the proof of Theorem 2. □

R e m a r k s. 3) The result shows that the family $\{\mathcal{T}_{n,m}\}_{m \geqslant n}$ proposed by Babuška and Aziz in [2] is in some sense extremal for the problem of optimal error estimates in the linear FEM on arbitrary triangulations. Indeed, for any triangle τ we have the local estimate

$$|v - I_\tau v|_{H^1(\tau)} \leqslant C R_\tau |v|_{H^2(\tau)} \leqslant C\frac{h_\tau^2}{k_\tau}|v|_{H^2(\tau)}, \quad v \in H^2(\tau), \tag{18}$$

for the standard linear FE interpolation operator I_τ interpolating continuous v at the vertices of τ, where R_τ is the radius of the circumscribed circle, h_τ the length of the longest edge, and k_τ the height associated with the longest edge of τ. Note that $R_\tau \leqslant h_\tau^2/(2k_\tau)$. The last estimate in (18) is well-known (it is usually given in terms of the ratio h_τ^2/ϱ_τ, where ϱ_τ is the radius of the inscribed circle of τ). The sharper first estimate in terms of R_τ can be found in [8], according to [7] it holds with constant $C = 1$. Thus, on any triangulation $\mathcal{T}$ and for any solution $u \in H^2(\Omega)$ of a $H^1(\Omega)$-elliptic boundary value problem, the Galerkin solution $S_{\mathcal{T}}$ in the linear FE space $V_{\mathcal{T}}$ satisfies

$$\|u - S_{\mathcal{T}}\|_{H^1(\Omega)} \leqslant C \min\Big(\|u\|_{H^1(\Omega)}, \max_{\tau \in \mathcal{T}} \frac{h_\tau^2}{k_\tau}|u|_{H^2(\Omega)}\Big). \tag{19}$$

Theorem 2 shows that using this simple mesh characteristics, (19) cannot be improved.

4) The above proof highlights a difference between the uniformly H^1-stable Galerkin projections $S = S_{n,m}$ and the nodal interpolation projection $I = I_{n,m}$ onto $V_{n,m}$: While for $\kappa = k/h^2 \geqslant 1$ we have

$$E_{n,m}(u) \approx \|u - S\|_{H^1(\Omega)} \approx \|u - I\|_{H^1(\Omega)} \approx \|S_y\|_{L_2(\Omega)} \approx \|I_y\|_{L_2(\Omega)} \approx \kappa^{-1},$$

i.e., the interpolation error correctly represents the Galerkin error in the energy norm (as is normally the case, see [3]), for highly distorted meshes $\mathcal{T}_{n,m}$ with $\kappa \to 0$ we have

$$\|u - I\|_{H^1(\Omega)} \approx \|I_y\|_{L_2(\Omega)} \approx \kappa^{-1} \to \infty,$$

and

$$E_{n,m}(u) \approx \|u - S\|_{H^1(\Omega)} \approx 1, \qquad \|S_y\|_{L_2(\Omega)} \approx \kappa \to 0.$$

5) Slight modifications of the partitions $\mathcal{T}_{n,m}$ can change the situation dramatically. E.g., if one adds vertical lines at $x = x_{2i}$, $i = 1, \ldots, n-1$, then convergence in energy norm can be established for any sequence (n, m) with $m \geqslant n$ and $n \to \infty$, despite the fact that many obtuse triangles with arbitrarily large ratio h^2/k are still left, see also [4]. Thus, the problem of finding necessary and sufficient conditions on a sequence of triangulations ensuring FEM convergence remains still open.

6) Extensions of our results to Galerkin discretizations with higher-order Lagrange elements on $\mathcal{T}_{n,m}$ and other types of boundary conditions should be possible, we decided not to pursue them. If the degree of the Lagrange elements is p, non-convergence in the energy norm is expected if the condition $m/n^{p+1} \to 0$ is violated, see [9] for some numerical experiments in this direction.

A c k n o w l e d g e m e n t. We thank H.-G. Roos, who brought this problem to our attention and later recalled that he presented the result of Theorem 1 in one of his classes a couple of years ago (private communication). We also thank L. Ludwig, who made his unpublished manuscript [9] available to us.

References

[1] *T. Apel*: Anisotropic Finite Elements: Local Estimates and Applications. Advances in Numerical Mathematics, Teubner, Leipzig; Technische Univ., Chemnitz, 1999.

[2] *I. Babuška, A. K. Aziz*: On the angle condition in the finite element method. SIAM J. Numer. Anal. *13* (1976), 214–226.

[3] *R. E. Bank, H. Yserentant*: A note on interpolation, best approximation, and the saturation property. Numer. Math. (2014), doi:10.1007/s00211-014-0687-0.

[4] *A. Hannukainen, M. Juntunen, A. Huhtala*: Finite Element Methods I, course notes A.Mat-1.3650, Univ. Helsinki, 2015.

[5] *A. Hannukainen, S. Korotov, M. Křížek*: The maximum angle condition is not necessary for convergence of the finite element method. Numer. Math. *120* (2012), 79–88.

[6] *P. Jamet*: Estimations d'erreur pour des éléments finis droits presque dégénérés. Rev. Franc. Automat. Inform. Rech. Operat. *10*, Analyse numer., R-1 (1976), 43–60.

[7] *K. Kobayashi, T. Tsuchiya*: A Babuška-Aziz type proof of the circumradius condition. Japan J. Ind. Appl. Math. *31* (2014), 193–210.

[8] *M. Křížek*: On semiregular families of triangulations and linear interpolation. Appl. Math., Praha *36* (1991), 223–232.

[9] *L. Ludwig*: A discussion on the maximum angle condition/counterexample for the convergence of the FEM. Manuscript, TU Dresden, 2011.

[10] *H. A. Schwarz*: Sur une définition erroneé de l'aire d'une surface courbe. Gesammelte Mathematische Abhandlungen, vol. 2. Springer, Berlin, 1890, pp. 309–311, 369–370.

Author's address: *Peter Oswald*, Helmholtzstr. 3b, 010 69 Dresden, Germany, e-mail: `agp.oswald@gmail.com`.