Václav Kučera

A note on necessary and sufficient conditions for convergence of the finite element method

In: Jan Brandts and Sergej Korotov and Michal Křížek and Karel Segeth and Jakub Šístek and Tomáš Vejchodský (eds.): Application of Mathematics 2015, In honor of the birthday anniversaries of Ivo Babuška (90), Milan Práger (85), and Emil Vitásek (85), Proceedings. Prague, November 18-21, 2015. Institute of Mathematics CAS, Prague, 2015. pp. 132–139.

Persistent URL: http://dml.cz/dmlcz/702971

Conference Applications of Mathematics 2015, in honor of the birthday anniversaries of Ivo Babuška (90), Milan Práger (85), and Emil Vitásek (85)
J. Brandts, S. Korotov, M. Křížek, K. Segeth, J. Šístek, T. Vejchodský (Eds.)
Institute of Mathematics CAS, Prague 2015

A NOTE ON NECESSARY AND SUFFICIENT CONDITIONS FOR CONVERGENCE OF THE FINITE ELEMENT METHOD

Václav Kučera

Charles University in Prague, Faculty of Mathematics and Physics
Sokolovská 83, 186 75 Praha, Czech Republic
kucera@karlin.mff.cuni.cz

Abstract: In this short note, we present several ideas and observations concerning finite element convergence and the role of the maximum angle condition. Based on previous work, we formulate a hypothesis concerning a necessary condition for $O(h)$ convergence and show a simple relation to classical problems in measure theory and differential geometry which could lead to new insights in the area.

Keywords: finite element method, a priori error estimates, maximum angle condition

MSC: 65N30, 65N15, 53A05

1. Introduction

The finite element method (FEM) is among the most popular, if not the single most popular numerical method for the solution of partial differential equations. The theory and practice of the FEM has a long and rich history. One of the main questions is, of course, "when does it work". Specifically, for piecewise linear FEM, we ask when does the FEM have optimal $O(h)$ convergence in the $H^1(\Omega)$-norm. It was believed that the so-called *maximum angle condition* is a sufficient as well as necessary condition for $O(h)$ convergence. While the first is true, cf. [1, 2], the maximum angle condition is not *necessary* for $O(h)$ convergence, as was recently shown in [2] by a simple argument.

While the author believes that we are still far away from formulating a necessary and also sufficient condition for $O(h)$ convergence, in this short note we present some ideas and observations related to this question.

In Section 2.1, we investigate the refinement procedure from [2], where maximum angle condition satisfying triangulations are arbitrarily subdivided to obtain maximum angle violating triangulations with $O(h)$ convergence. We show that by such refinement, one cannot obtain triangulations containing *only* degenerate elements.

In Section 2.2, we review the only known counterexample of Babuška and Aziz [1], which together with the results of Section 2.1 leads to formulating a hypothesis on a necessary condition for $O(h)$ convergence. Namely, we hypothesize that elements satisfying the maximum angle condition must be "dense" in Ω.

We are unable to prove the presented hypothesis, in fact as far as the author knows, no nontrivial necessary condition is known in the literature. In Section 3, we present a simple connection between the question of FEM convergence and differential geometry, which could possibly give alternative insight into these and related questions of FEM convergence.

2. A hypothesis on a necessary condition for $O(h)$ convergence

We treat the following problem: Find $u : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ such that

$$-\Delta u = f, \quad u|_{\partial\Omega} = 0, \tag{1}$$

where Ω is a bounded polygonal domain with a Lipschitz boundary and $f \in L^2(\Omega)$. Defining $V = H_0^1(\Omega)$ and the associated bilinear form $a(u,v) = \int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x$, the corresponding weak form of (1) reads: Find $u \in V$ such that

$$a(u,v) = (f,v) \quad \forall v \in V.$$

The finite element method constructs a sequence of spaces $\{V_h\}_{h\in(0,h_0)}$ on conforming triangulations $\{\mathcal{T}_h\}_{h\in(0,h_0)}$ of Ω, where $V_h \subset V$ consists of globally continuous piecewise linear functions on $\mathcal{T}_h$. The FEM formulation then reads: Find $u_h \in V_h$ such that

$$a(u_h, v_h) = (f, v_h) \quad \forall v_h \in V_h.$$

Denoting h as the maximal diameter of all elements K in $\mathcal{T}_h$, the natural measure of convergence of u_h to u is estimation by powers of h. Specifically, in the energy norm of (1), we obtain at most $O(h)$ convergence if $u \in H^2(\Omega)$, i.e.

$$|u - u_h|_{H^1(\Omega)} \le C(u)h \quad \forall u \in H^2(\Omega) \cap V, \tag{2}$$

where the constant $C(u)$ is typically written in the form $C|u|_{H^2(\Omega)}$. The question is, when can such a result be proved. Currently, the most general sufficient condition known for (2) is the maximum angle condition:

Definition 1. *A system of triangulations* $\{\mathcal{T}_h\}_{h\in(0,h_0)}, h_0 > 0$ *satisfies the* maximum angle condition*, if there exists* $\alpha < \pi$ *such that all angles in all triangles* $K \in \mathcal{T}_h$ *are less than* α *for all* $h \in (0, h_0)$.

Recently it was shown that the maximum angle condition is not necessary for $O(h)$ convergence. In [2], triangulations $\mathcal{T}_h$ satisfying Definition 1 are refined by subdividing each triangle $K \in \mathcal{T}_h$ arbitrarily, thus obtaining new triangulations $\widetilde{\mathcal{T}}_h$.

Since $\mathcal{T}_h$ satisfies Definition 1 and $\widetilde{\mathcal{T}}_h$ is a refinement of $\mathcal{T}_h$, then since $\mathcal{T}_h$ allows $O(h)$ convergence, so must $\widetilde{\mathcal{T}}_h$ by Céa's lemma.

The natural question arises, how far can one take these refinements. For example, taking $\mathcal{T}_h$ satisfying the maximum angle condition, can one construct $\widetilde{\mathcal{T}}_h$ such that the maximal angles of all triangles are arbitrarily close to π? We answer this question negatively in the following section.

2.1. Mesh subdivisions

In the following, we will need to distinguish between triangles $K \in \mathcal{T}_h$ "satisfying" and "violating" the maximum angle condition. Of course, this depends on the choice of α in Definition 1. For this purpose, we fix some α arbitrarily close to π. We will call K with maximum angle larger than α *degenerate* and *non-degenerate* otherwise. This terminology is clear: if $\{\mathcal{T}_h\}_{h\in(0,h_0)}$ violates Definition 1, than there exist triangles K in some $\mathcal{T}_h$ such that their maximum angle is arbitrarily close to π. Hence Definition 1 is violated for any choice of α. Therefore, in the end, the "maximum angle violating" property is independent of the specific choice of α.

Definition 2. *Let $\alpha \in (0, \pi)$ be close to π. A triangle $K \in \mathcal{T}_h$ is called* degenerate*, if the maximum angle in K is $> \alpha$. Otherwise, K is called* non-degenerate*.*

Now we show that a non-degenerate triangle K cannot be cut into degenerate triangles *only*. Hence the construction from [2] cannot give triangulations containing *only* degenerate triangles.

Lemma 1. *Let $\alpha \in (\frac{2}{3}\pi, \pi)$. Let K be a triangle with all angles less than α. Then there does not exist a finite conforming partition of K into triangles which all contain an angle greater than or equal to α.*

Proof. Assume on the contrary that such a partition $\mathcal{P}$ exists. Let $t =$ number of triangles in $\mathcal{P}$, $v_I =$ number of vertices of $\mathcal{P}$ contained in the interior of K and $v_B =$ number of vertices of $\mathcal{P}$ lying on ∂K.

On one hand, the sum of all angles in $\mathcal{P}$ is πt. On the other hand, the same sum can be calculated by summing all angles surrounding all interior, boundary and corner vertices in $\mathcal{P}$. Thus we get

$$\pi t = 2\pi v_I + \pi v_B + \pi,$$

which simplifies to the Euler-type identity

$$t = 2v_I + v_B + 1. \tag{3}$$

Now, we calculate the number of angles in $\mathcal{P}$ that are greater than or equal to α. On one hand, we obtain t, since each triangle in $\mathcal{P}$ contains exactly one such angle. On the other hand, since $\alpha > \frac{2}{3}\pi$, each interior vertex of $\mathcal{P}$ can be the vertex of (at most) two such angles. Similarly, each boundary vertex can be the vertex of (at most) one such angle and the corner vertices of T are all $< \alpha$. Thus

$$t \leq 2v_I + v_B. \tag{4}$$

Substituting (3) into (4), we get $1 \leq 0$, which is a contradiction. □

Lemma 1 can be interpreted informally in the following way: A non-degenerate triangle cannot be cut into degenerate triangles only, one always has at least one non-degenerate triangle in the resulting partition.

In [2], triangulations violating the maximum angle condition but possessing the $O(h)$ convergence property are constructed by taking a system of triangulations satisfying the maximum angle condition and subdividing each of its triangles arbitrarily. However, Lemma 1 states that each of these subdivided triangles K contains a triangle $\tilde{K}$ satisfying the same maximum angle condition as K. Therefore, using this procedure, one cannot produce large regions of degenerate triangles in the following sense.

Definition 3. *We say that the set of non-degenerate triangles is* dense *in* $\{\mathcal{T}_h\}_{h\in(0,h_0)}$, *if for all* $x \in \Omega$ *and all neighbourhoods* $\mathcal{U} \in \mathcal{O}(x)$ *there exists* $\tilde{h} \in (0, h_0)$ *such that for all* $h \in (0, \tilde{h})$ *there exists a non-degenerate* $K \in \mathcal{T}_h$ *such that* $K \subset \mathcal{U}$.

Now we will prove the main result, that using the procedure of [2], one can obtain only triangulations, where the set of non-degenerate triangles is dense in the sense of Definition 3. In particular, one cannot obtain triangulations with only degenerate triangles by subdividing triangulations satisfying the maximum angle condition.

Theorem 2. *Let* $\{\mathcal{T}_h\}_{h\in(0,h_0)}$ *satisfy the maximum angle condition. Let* $\{\widetilde{\mathcal{T}}_h\}_{h\in(0,h_0)}$ *be a set of conforming triangulations of* Ω *obtained from* $\{\mathcal{T}_h\}_{h\in(0,h_0)}$ *by subdividing each triangle in each* $\mathcal{T}_h$ *into a finite number of triangles. Then non-degenerate triangles are dense in* $\{\widetilde{\mathcal{T}}_h\}_{h\in(0,h_0)}$.

Proof. Choose $x \in \Omega$ and $\mathcal{U} \in \mathcal{O}(x)$. Then for sufficiently small $\tilde{h}$, for each $\mathcal{T}_h$, $h \in (0, \tilde{h})$ there exists $K \in \mathcal{T}_h$ such that $K \subset \mathcal{U}$. Since $\widetilde{\mathcal{T}}_h$ is obtained from $\mathcal{T}_h$ by subdividing each triangle, by Theorem 1 the partition of K must contain a non-degenerate triangle $\tilde{K}$. Since $K \subset \mathcal{U}$, then also $\tilde{K} \subset \mathcal{U}$, hence the set of non-degenerate triangles is dense in $\{\widetilde{\mathcal{T}}_h\}_{h\in(0,h_0)}$. □

2.2. The Babuška-Aziz counterexample

As far as the author is aware of, there exists only one counterexample to $O(h)$ convergence of the finite element method. This is the counterexample of Babuška and Aziz [1], further refined in [4]. The counterexample consists of a series of triangulations $\mathcal{T}_{m,n}$ of the unit square, where m and $2n$ denote the number of intervals into which the horizontal and vertical sides of the unit square are divided, cf. Figure 1. On these triangulations, the piecewise linear FEM is used to discretize Poisson's problem with the exact solution $\frac{1}{2}x(1-x)$. While the original paper [1] uses this

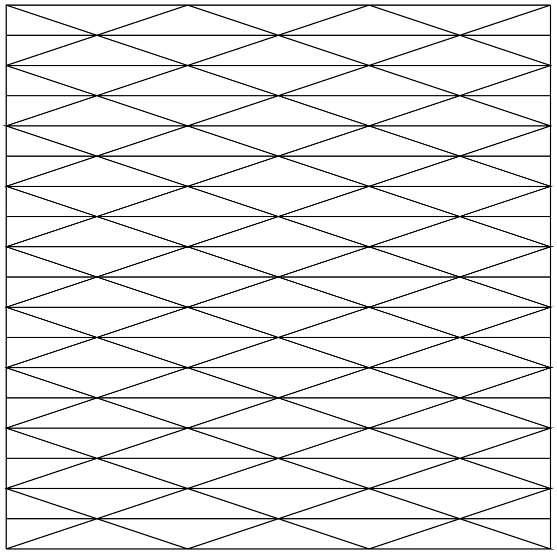

Figure 1: Babuška-Aziz triangulation $\mathcal{T}_{3,9}$ of the unit square

problem only to provide a counterexample to $O(h)$ convergence, in [4] a more detailed analysis is carried out and the error of the discrete solution $u_{m,n}$ is shown to satisfy

$$\|u - u_{m,n}\|_{H^1(\Omega)} \approx \min\{1, m/n^2\}. \tag{5}$$

In $\mathcal{T}_{m,n}$, the maximal edge length satisfies $h = 1/m$. Estimate (5) means that if the maximal condition is violated, i.e. $n \to \infty$ faster than m, then $O(h)$ convergence does not hold.

In the Babuška-Aziz counterexample, if all triangles (except the few triangles near the vertical boundaries) violate the maximal angle condition, then we lose $O(h)$ convergence. Therefore, in this counterexample, large open subsets of Ω containing only *degenerate* triangles destroy $O(h)$ convergence. In general, if another system of triangulations $\mathcal{T}_h$ coincided with $\mathcal{T}_{m,n}$ on a fixed open subset of Ω, then $\mathcal{T}_h$ would also not admit $O(h)$ convergence. Of course, $\mathcal{T}_{m,n}$ are highly structured, even periodic, and therefore represent only one possibility of triangulations containing only degenerate triangles. Theorem 2 states that such triangulations cannot be obtained by subdivision. Therefore, based on these considerations, we state the following hypothesis, which says that the result of Theorem 2 is a necessary condition for $O(h)$ convergence.

Hypothesis. *Let $\{\mathcal{T}_h\}_{h\in(0,h_0)}$ be a system of triangulations and $u_h \in V_h$ the corresponding discrete solutions. If $|u - u_h|_{H^1(\Omega)} \leq C(u)h$ for all $u \in H^2(\Omega)$, or some dense subset thereof, then triangles satisfying the maximum angle condition (non-degenerate triangles) are dense in $\{\mathcal{T}_h\}_{h\in(0,h_0)}$ in the sense of Definition* 3.

The strategy how to prove this hypothesis is as follows: show that triangulations similar to $\mathcal{T}_{m,n}$ violate $O(h)$ convergence. Here "similar" means that $\mathcal{T}_h$ should contain only degenerate triangles (up to perhaps a few adjoining the boundary). However, for such general triangulation, we lack the simple structure of $\mathcal{T}_{m,n}$, which is an essential ingredient in the proofs presented in [1, 4].

3. Relation to differential geometry

The Babuška-Aziz counterexample is the only known counterexample to finite element convergence to date. Moreover, it is interesting that this counterexample coincides with a classical counterexample from the early stages of development of measure theory, the so-called *Schwarz lantern* [5]. The purpose of Schwarz's counterexample is to show that even for smooth surfaces, surface area cannot be defined as the limit of areas of approximating polyhedral surfaces. In fact, in Schwarz's counterexample, the surface areas of the approximating polyhedral surfaces tend to infinity, although the limit surface (in the Hausdorff metric) has finite surface area. This surprised the contemporary mathematical community, since this limit definition was standardly used, based on a flawed analogy with curve length.

Since two different areas of mathematics share the same counterexample, it is natural to ask whether there is some connection. Unsurprisingly, this is the case. The purpose of this section is to point out this connection and that this opens the door to obtain a different insight into the convergence of FEM via differential geometry and measure theory.

Definition 4. *Let $\Omega \subset \mathbb{R}^2$ and $v : \Omega \to \mathbb{R}$. Then the* graph *of v is defined as*

$$\mathrm{graph}(v) = \{(x, y, z) \in \mathbb{R}^3 : \ z = v(x, y) \text{ for } (x, y) \in \Omega\}.$$

In the following theorem, we show that FEM convergence implies convergence of surface areas of the corresponding graphs. By $A(v)$, we denote the surface area of $\mathrm{graph}(v)$ of a function $v : \Omega \to \mathbb{R}$, if it is well defined.

Theorem 3. *Let $u \in H^2(\Omega)$ and $u_h \in V_h$ for $h \in (0, h_0)$. If $u_h \to u$ in $H^1(\Omega)$ as $h \to 0$, then for a subsequence,* $\mathrm{graph}(u_{h_n}) \to \mathrm{graph}(u)$ *in the Hausdorff metric and $A(u_{h_n}) \to A(u)$.*

Proof. Since $H^2(\Omega), V_h \subset C(\bar{\Omega})$, then $u, u_h \in C(\bar{\Omega})$. Since $u_h \to u$ in $H^1(\Omega)$, also $u_h \to u$ in $L^2(\Omega)$. Therefore there exists a subsequence u_{h_n} converging to u pointwise almost everywhere in Ω. Since $u, u_{h_n} \subset C(\bar{\Omega})$, we have $u_{h_n} \to u$ uniformly in Ω. By the definition of the Hausdorff metric and uniform convergence, we immediately have $\mathrm{graph}(u_{h_n}) \to \mathrm{graph}(u)$ in the Hausdorff metric.

It remains to prove the convergence of surface areas. For a function $v \in H^1(\Omega)$, the area of $\mathrm{graph}(v)$ is given by

$$A(v) = \int_\Omega \sqrt{1 + \left(\tfrac{\partial v}{\partial x}\right)^2 + \left(\tfrac{\partial v}{\partial y}\right)^2} \mathrm{d}(x, y).$$

Therefore,

$$\big|A(u) - A(v)\big| \le \int_\Omega \left|\sqrt{1 + \left(\tfrac{\partial u}{\partial x}\right)^2 + \left(\tfrac{\partial u}{\partial y}\right)^2} - \sqrt{1 + \left(\tfrac{\partial v}{\partial x}\right)^2 + \left(\tfrac{\partial v}{\partial y}\right)^2}\right| \mathrm{d}(x, y). \qquad (6)$$

Using the easily verifiable inequality $|\sqrt{1 + a^2 + b^2} - \sqrt{1 + c^2 + d^2}| \le |a - c| + |b - d|$ for all $a, b, c, d \in \mathbb{R}$, we obtain

$$\big|A(u) - A(v)\big| \le \int_\Omega \Big(\big|\tfrac{\partial u}{\partial x} - \tfrac{\partial v}{\partial x}\big| + \big|\tfrac{\partial u}{\partial y} - \tfrac{\partial v}{\partial y}\big|\Big)\mathrm{d}(x, y) \le \sqrt{2}|\Omega|^{1/2}|u - v|_{H^1(\Omega)} \qquad (7)$$

by Hölder's inequality. Therefore, $u_{h_n} \to u$ in $H^1(\Omega)$ implies $A(u_{h_n}) \to A(u)$. □

Theorem 3 thus establishes a simple connection to the theory of approximation of surface area. Similar questions have been dealt with in the past decade in the field of discrete differential geometry. The question is, how do classical differential-geometric objects defined on polygonal (polyhedral) surfaces converge to those of the limit surface. In the case of Theorem 3, we are interested in results of the following type:

Theorem 4 ([3])**.** *If a sequence of polyhedral surfaces $\{M_n\}$ converges to a smooth surface M in the Hausdorff metric, then the following conditions are equivalent:*
convergence of area,
convergence of normal fields,
convergence of metric tensors,
convergence of Laplace-Beltrami operators.
Here convergence is always meant in the L^∞-sense for the corresponding term.

Theorems 3 and 4 yield a potential strategy for proving necessary conditions for FEM convergence. In the context of discrete differential geometry, the problem can be attacked from other points of view using different techniques than usual in the FEM community. We note that in [3], the theory used to prove Theorem 4 is used also to investigate convergence of other objects such as geodesics and mean curvature vectors. Unfortunately, it seems that there are no suitable more general results directly applicable to the convergence of FEM in existing discrete differential literature, although the analogy of the minimum angle condition is known in the community.

Acknowledgements

The research is supported by the Grant No. P201/13/00522S of the Czech Science Foundation. V. Kučera is a junior researcher at the University Center for Mathematical Modelling, Applied Analysis and Computational Mathematics (Math MAC).

References

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

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

[3] Hildebrandt, K., Polthier, K., and Wardetzky, M.: On the convergence of metric and geometric properties of polyhedral surfaces. Geom. Dedicata **123** (2006), 89–112.

[4] Oswald, P.: Divergence of the FEM: Babuška-Aziz revisited. Appl. Math. **60 (5)** (2015).

[5] Schwarz, H. A.: Sur une définition erronée de l'aire d'une surface courbe. Ges. Math. Abhandl. II (1890), Springer-Verlag, 309–311.