

A General Accuracy Criterion for Finite Element Approximation

Igor Tsukerman

Electrical Engineering Department, The University of Akron, OH 44325-3904, USA

Abstract — The accuracy of Finite Element approximation depends on the size and shape of elements. The proposed theory clarifies the source of shape related errors. A new general accuracy condition based on the maximum eigenvalue of the element stiffness matrix (or, alternatively, its trace) is presented. Several new and several familiar geometric criteria follow directly from the eigenvalue condition as particular cases.

The new accuracy criterion can be used in engineering practice as an *a priori* measure of the quality of finite element meshes.

Index terms — finite elements, accuracy estimate, approximation, interpolation, stiffness matrix, tetrahedral mesh

I. INTRODUCTION

Since the first application of triangular finite elements in electromagnetics some 30 years ago, it has been well known that the shape of elements affects the accuracy of Finite Element (FE) approximation. In 2D, for example, “very flat” triangular elements are normally avoided. According to Zlámal’s *minimum* angle condition [1], in a family of triangular meshes the minimum angle of elements should be bounded away from zero. Equivalently, the ratio of the maximum element edge to the diameter of the inscribed circle should be bounded [2]. Slightly less restrictive is the Syngé-Babuška-Aziz “*maximum* angle” condition [3, 4], requiring that the maximum angle of triangular elements be bounded away from π .

The situation is less clear in three dimensions, in particular for tetrahedral elements. Although ‘quality’ criteria are abundant (e.g. [5]), no theory is available to explain adequately which parameters are really essential for the approximation accuracy and why.

I argue that shape related approximation errors are caused by large eigenvalues of element stiffness matrices, and an appropriate eigenvalue condition can be used to evaluate the quality of FE meshes in two and three dimensions. Application of the proposed theory to triangular and tetrahedral elements leads to several known and several new accuracy conditions (see also [6, 7]). The results also show that ill-conditioning of the stiffness matrix (due to a large maximum eigenvalue) is a source of not only poor convergence of iterative solvers but of poor approximation as well.

For simplicity, the theory is first presented for tetrahedral node elements (Section II). Then it is extended to arbitrary elements satisfying a few general assumptions (Section III).

In Sections IV, V links between the new criterion and other familiar conditions are discussed. In Section VI, results of a

numerical experiment are presented. This analysis complements and extends the conference paper [7] that focused on 2D triangular elements.

II. THE MAXIMUM EIGENVALUE CONDITION FOR TETRAHEDRAL NODE ELEMENTS

Let u_* be a scalar potential approximated on a tetrahedral FE mesh in a domain Ω . The approximation error is measured in the (semi)norm proportional to the field energy:

$$\|u\|_E^{(e)} = \left[\int_{(e)} |\nabla u|^2 dV \right]^{1/2}; \quad \|u\|_E = \left[\int_{\Omega} |\nabla u|^2 dV \right]^{1/2} \quad (1)$$

Here the subscript ‘E’ stands for ‘Energy’, and the superscript (e) denotes a quantity related to an individual element (e) or the restriction of a function to an element. (For example, $\|u\|_E^{(e)}$ is the energy over a particular element, while $\|u\|_E$ is the energy over the whole domain Ω .)

Let $\{\phi_i\}$ ($i = 1, 2, \dots, n$) be the system of nodal basis functions on a given mesh, and let \tilde{W} be the space spanned by these functions; $\tilde{W}^{(e)}$ is the restriction of \tilde{W} to an element (e). To an Euclidean vector of nodal values $u^{(e)} \in E^4$ there corresponds a function $\tilde{u}^{(e)} = \sum_{k=1}^4 u_k^{(e)} \phi_k^{(e)} \in \tilde{W}^{(e)}$.

A global approximation over the whole FE mesh in Ω can be provided by the usual interpolating function \tilde{u}_I :

$$\tilde{u}_I = \sum_{i=1}^n u_{li} \phi_i \in \tilde{W} \quad (2)$$

where u_{li} are equal to the node values of u_* .

Next, consider a first order Taylor approximation \tilde{u}_T of the exact potential u_* around an arbitrary point $r_0 = (x_0, y_0, z_0)$ in a finite element (e):

$$\tilde{u}_T(r) = u_*(r_0) + \nabla u_*(r_0) \cdot (r - r_0), \quad r_0 \in (e) \quad (3)$$

Informally, the key observations that lead to the proposed criterion can be summarized as follows:

- Approximations of the *node values* of u_* by both the Taylor expansion \tilde{u}_T and the interpolating function \tilde{u}_I are “good” (shape-independent).¹
- However, the *energy norm* approximation of u_* by \tilde{u}_I may be worse than its approximation by \tilde{u}_T : the \tilde{u}_T -approximation is uniform and shape-independent, while the \tilde{u}_I -approximation may depend on shape.

Manuscript received November 3, 1997. The work was supported in part by the National Science Foundation research grant ECS-9702364.

I. Tsukerman, itsukerman@uakron.edu, <http://www.ecgf.uakron.edu/~igor>

¹ The node values of \tilde{u}_I are in fact exact by definition.

(c) Why the apparent inconsistency between (a) and (b)? The energy norm in $\tilde{W}^{(e)}$ is induced by the element stiffness matrix. If this matrix has a large eigenvalue, the small difference $u_I - u_T$ between the node values may translate into a large difference $\tilde{u}_I - \tilde{u}_T$ in the energy norm.

A more formal analysis is as follows. If u_* is sufficiently smooth, the Taylor approximation is *uniform* in the sphere of a radius $\rho < \rho_0$ (and hence independent of the element's shape):

$$|\tilde{u}_T - u_*| \leq c\rho^2, \quad |\nabla \tilde{u}_T - \nabla u_*| \leq c\rho \quad \text{if } |r - r_0| < \rho_0 \quad (4)$$

(c will always denote a generic constant, independent of the element shape, and not necessarily the same in all occurrences). It immediately follows that the \tilde{u}_T -approximation of u_* in the energy (semi)norm is also shape-independent:

$$\|\tilde{u}_T - u_*\|_E^{(e)2} = \int_{(e)} |\nabla \tilde{u}_T - \nabla u_*|^2 dV \leq c h^{(e)2} V^{(e)} \quad (5)$$

where $V^{(e)}$ is the volume of the element, $h^{(e)}$ is its maximum edge.

For the purpose of studying the dependence of the interpolation error on shape, the Taylor approximation \tilde{u}_T may be considered in lieu of u_* . Indeed,

$$\|\tilde{u}_I - u_*\|_E^{(e)} \leq \|\tilde{u}_I - \tilde{u}_T\|_E^{(e)} + \|\tilde{u}_T - u_*\|_E^{(e)} \quad (6)$$

and the second norm in the right hand side is "small" and shape-independent due to (5).

We now focus on the first norm in the right hand side of (6). Since both $\tilde{u}_I^{(e)}$ and $\tilde{u}_T^{(e)}$ are in the FE space $\tilde{W}^{(e)}$, and since the energy norm in that space is induced by the element stiffness matrix $A^{(e)}$, we have

$$\|\tilde{u}_I^{(e)} - \tilde{u}_T^{(e)}\|_E^2 = (A^{(e)}(u_I^{(e)} - u_T^{(e)}), u_I^{(e)} - u_T^{(e)}) \quad (7)$$

The Euclidean quadratic form in the right hand side of (7) can be estimated via the *maximum eigenvalue* of $A^{(e)}$:

$$\|\tilde{u}_I^{(e)} - \tilde{u}_T^{(e)}\|_E^2 \leq \lambda_{\max}(A^{(e)}) \|u_I^{(e)} - u_T^{(e)}\|_2^2 \quad (8)$$

(where $\|\cdot\|_2$ is the standard Euclidean norm). Since u_I is the vector of exact nodal values and u_T is a vector of their Taylor approximations, it follows from (4) that

$$\|u_I^{(e)} - u_T^{(e)}\|_2 \leq c h^{(e)2} \quad (9)$$

Then (8) becomes

$$\|\tilde{u}_I^{(e)} - \tilde{u}_T^{(e)}\|_E^2 \leq c \lambda_{\max}(A^{(e)}) h^{(e)4} \quad (10)$$

and, upon combining with (6) and (5),

$$\|\tilde{u}_I^{(e)} - u_*\|_E^2 \leq c [\lambda_{\max}(A^{(e)}) h^{(e)4} + h^{(e)2} V^{(e)}] \quad (11)$$

Summing up these element-wise estimates over the whole mesh, one obtains a global estimate

$$\|\tilde{u}_I - u_*\|_E^2 \leq c \sum_{(e) \in \text{mesh}} [\lambda_{\max}(A^{(e)}) h^{(e)4} + V^{(e)} h^{(e)2}] \quad (12)$$

One way to simplify this result is by noting that $\lambda_{\max}(A^{(e)}) \leq \text{tr } A^{(e)}$, $\sum_{(e)} \text{tr } A^{(e)} = \text{tr } A$, $\sum_{(e)} V^{(e)} = V$, where A is the global stiffness matrix and V is the volume of the domain:

$$\|\tilde{u}_I - u_*\|_E \leq c (h_{\max}^2 \sqrt{\text{tr } A} + h_{\max} \sqrt{V}) \quad (13)$$

($h_{\max} = \max_{(e)} h^{(e)}$ is the maximum element size²).

Alternatively, after factoring out the element volume $V^{(e)}$ in (12), the global error estimate assumes the form

$$\|\tilde{u}_I - u_*\|_E^2 \leq c \sum_{(e)} (\lambda_{\max}(\hat{A}^{(e)}) h^{(e)4} + h^{(e)2}) V^{(e)} \quad (14)$$

where the 'hat' symbol denotes the *scaled* element stiffness matrix $\hat{A}^{(e)} \equiv A^{(e)} / V^{(e)}$. Estimate (14) simplifies to

$$\|\tilde{u}_I - u_*\|_E \leq c \sqrt{V} \max_{(e) \in \text{mesh}} (h^{(e)2} \sqrt{\lambda_{\max}(\hat{A}^{(e)})} + h^{(e)}) \quad (15)$$

The maximum eigenvalue can again be replaced with the matrix trace which is easier to compute.

For tetrahedral elements, the trace of the scaled nodal stiffness matrix can be interpreted geometrically. The diagonal elements of this matrix are equal to $1/d_i^2$, where d_i 's are the heights of the tetrahedron (Fig. 1), and therefore³

$$\begin{aligned} \lambda_{\max}(\hat{A}^{(e)}) &\leq \text{tr}(\hat{A}^{(e)}) = \sum_{i=1}^4 1/d_i^2 = (1/h^{(e)})^2 \sum_{i=1}^4 (h^{(e)} / d_i)^2 \\ &\leq 2 (1/h^{(e)})^2 \sum_{i \neq j} 1/\sin^2 \alpha_{ij} \end{aligned} \quad (16)$$

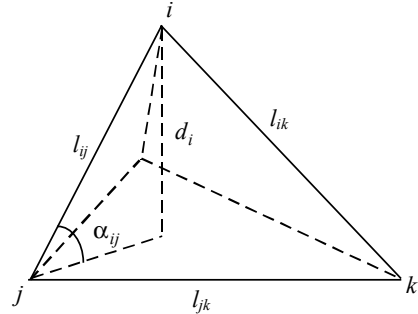


Fig. 1. Geometric parameters of a tetrahedral element.

Formulas (15), (16) lead to the *minimum-maximum angle* condition for the angles α_{ij} between edges and faces (Fig. 1): α_{ij} are to be bounded away from both zero and π . However, a more precise condition can be obtained by applying the eigenvalue criterion to *edge* elements (cf. [6]).

² The first terms in the right hand side of estimates (13), (15) are of the order $O(h_{\max})$, not $O(h_{\max}^2)$ as it may appear, because the maximum eigenvalue and the trace of A are of the order $O(h_{\max}^{-2})$.

³ The extra factor of two in the last term of (16) is due to the fact that, e.g. $(l_{jk}/d_i)^2 \leq ((l_{ji} + l_{ik})/d_i)^2 \leq 2((l_{ji}/d_i)^2 + (l_{ik}/d_i)^2) \leq 2(1/\sin^2 \alpha_{ij} + 1/\sin^2 \alpha_{ik})$

III. GENERALIZATION OF THE MAXIMUM EIGENVALUE CONDITION

The main argument of the previous section does not depend on specific features of tetrahedral elements and holds without any substantial change under the following general assumptions:

(A1) The energy norm is bounded:

$$\|u\|_E^{(e)2} \leq V^{(e)} \sum_{l=0}^p c_l^2 \|u\|_{W_\infty^l}^{(e)2} \quad (17)$$

where $c_l \geq 0$ are some real constants and W_∞^l is the Sobolev space.

(A2) There exists an m -th order ($m \geq p$) Taylor approximation \tilde{u}_T of u_* around a point r_0 within each element, so that for a sufficiently small ρ_0 (the same for all elements) one has

$$\|\tilde{u}_T - u_*\|_{W_\infty^l} \leq c \rho^{m+1-l}, \quad 0 \leq l \leq p, \quad \text{if } \rho \equiv |r - r_0| \leq \rho_0 \quad (18)$$

(A3) The Taylor expansion (A2) within each element (e) belongs to the FE space $\tilde{W}^{(e)}$: $\tilde{u}_T^{(e)} \in \tilde{W}^{(e)}$.

(A4) The Taylor approximation of the FE degrees of freedom is shape independent:

$$\|u_I^{(e)} - u_T^{(e)}\|_2 \leq c h^\kappa, \quad \text{for some } \kappa > 0 \quad (19)$$

Under hypotheses (A1-A4), local (element-wise) estimate (11) and global estimates (13), (15) can be generalized. Estimates (13) and (15) become, respectively,

$$\|\tilde{u}_I - u_*\|_E \leq c (h_{\max}^\kappa \sqrt{tr A} + \sqrt{V} \sum_{l=0}^p c_l h_{\max}^{m+1-l}) \quad (20)$$

and

$$\begin{aligned} \|\tilde{u}_I - u_*\|_E &\leq c \sqrt{V} \max_{(e)} (h^{(e)\kappa} \sqrt{\lambda_{\max}(\hat{A}^{(e)})} \\ &\quad + \sum_{l=0}^p c_l h^{(e), m+1-l}) \end{aligned} \quad (21)$$

For example, tetrahedral node elements satisfy assumptions (A1-A4). Indeed, for the energy norm (1), (A1) holds with $c_0=0$, $c_1=1$. (A2) is satisfied with $m=1$ if the exact potential u_* is sufficiently smooth. (A3) is valid because the Taylor approximation is linear within an element, and $\tilde{W}^{(e)}$ contains all linear functions. Finally, (9) represents (A4), with $\kappa = 2$.

IV. THE MAXIMUM EIGENVALUE CONDITION IN SOME FAMILIAR CASES

A. L_2 -Approximation on Tetrahedral Node Elements

Suppose that the potential u_* itself, rather than the field ∇u_* , is approximated, i.e. the L_2 -norm is of interest:

$$\|u\|_E = \left[\int_{\Omega} u^2 dV \right]^{1/2} \quad (22)$$

This norm in $\tilde{W}^{(e)}$ is induced by the “mass matrix”

$$\begin{aligned} A_{ij}^{(e)} &= \int_{(e)} \phi_i \phi_j dV; \quad \hat{A}_{ij}^{(e)} = \frac{1}{V^{(e)}} \int_{(e)} \phi_i \phi_j dV \\ \hat{A}^{(e)} &= \frac{1}{20} \begin{bmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{bmatrix} \end{aligned} \quad (23)$$

The maximum eigenvalue $\lambda_{\max}(\hat{A}^{(e)})$ of (23) is equal to $1/4$ and does not depend on the element shape. Therefore, according to (21), approximation of the potential is shape-independent. This known result has been obtained here directly from the general estimate (21).

B. First Order Triangular Node Elements

The maximum eigenvalue criterion for first order triangular elements leads directly to Zlámal's minimum angle condition. The derivation is completely analogous to (16).

C. Triangular and Tetrahedral Edge Elements

Two substantially different cases need to be distinguished.

1. Conservative fields are approximated and the energy of the vector field is evaluated in the L_2 -norm (i.e. the field itself, and not its *curl*, is of interest). Then the edge representation of the field is equivalent to the nodal representation of the scalar potential. However, the maximum eigenvalue criterion (21) yields a more precise accuracy estimate when applied to a modified *edge* stiffness matrix rather than to the nodal matrix (see analysis in [6]). For triangular elements, the *maximum* angle condition is obtained along these lines.

2. The energy norm includes the *curl* of the field. The eigenvalue criterion remains formally correct but overestimates the interpolation error for tetrahedral edge elements. The reason is that the error vector $u_I - u_T$ in this case cannot be arbitrary: its component along the maximum eigenvector of the stiffness matrix must be $O(h^2)$. Work is under way to make the theory of Section III effective in such situations.

V. APPROXIMATION AND THE CONDITION NUMBER

The theory of Sections II, III shows that FE interpolation errors are governed by the *maximum* eigenvalue of the stiffness matrix. This establishes a direct link (experimentally observed by Zgainski *et al.* [8]) between the condition number of the stiffness matrix and FE approximation. (The *minimum* eigenvalue, though, is related to the diameter D of the computational domain Ω and is irrelevant to approximation.) Using estimate (12) for a tetrahedral mesh as an example, the approximation error can be related to the condition number as

$$\|\tilde{u}_I - u_*\|_E \propto h_{\max}^2 \left[\sum_{(e)} \lambda_{\max}(A^{(e)}) \right]^{1/2} \propto h_{\max}^2 (\lambda_{\max}(A) h_{\max}^{-3})^{1/2}$$

$$\begin{aligned} \propto h_{\max}^{1/2} (\lambda_{\min}(A))^{1/2} (\text{cond}(A))^{1/2} &\propto h_{\max}^{1/2} (h_{\max}^3 / D^2)^{1/2} (\text{cond}(A))^{1/2} \\ &= D^{-1} h_{\max}^2 (\text{cond}(A))^{1/2} \end{aligned} \quad (24)$$

where the equivalence $X \propto Y$ means that the ratio X/Y is bounded, $0 < c_1 \leq X/Y \leq c_2 < \infty$. The sequence of equivalence relationships in (24) is based on the known properties of the eigenvalues of local and global stiffness matrices (see e.g. [9]) and on the assumption that the ratio of maximum / minimum element size and the number of elements sharing a common node is bounded in a family of FE meshes.

VI. NUMERICAL ILLUSTRATION

Boundary value problems for the Laplace equation were considered in a cubic domain, with four known theoretical solutions (u1-u4, as specified in the caption to Fig. 2) and with the corresponding inhomogeneous Dirichlet boundary conditions. Two families of tetrahedral meshes (labeled T and S) were obtained by successive global refinement of initial meshes⁴ T_1 and S_1 .

Global *a priori* error estimates were calculated as the square root of the sum in (14), the maximum eigenvalue being replaced with the matrix trace. In Fig. 2, these theoretical estimates are plotted vs. the actual errors for test problems. Each test problem is characterized by the mesh family (T or S) and by the exact solution (u1-u4).

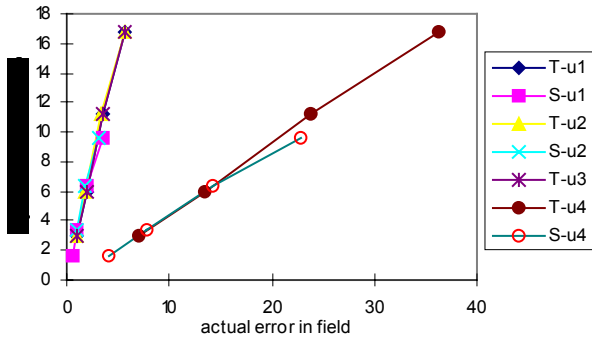


Fig. 2. Global error estimate (14) vs. actual error in field for test problems. T, S – two mesh families. u1-u4 – four test solutions: u1 = $\sin x \sin y \sinh(\sqrt{2} z)$; u2 = $\sin y \sin z \sinh(\sqrt{2} x)$; u3 = $\sin z \sin x \sinh(\sqrt{2} y)$; u4 = $\sin 1.5x \sin 1.5y \sinh(1.5\sqrt{2} z)$.

For each problem, the data points for the estimated vs. exact error fall approximately on a straight line, with the slope of the line depending on the second derivatives of the theoretical solution. (The second derivatives of u4 are substantially higher than those of u1-u3, and hence two distinct line slopes in Fig. 2.) This demonstrates consistency of the eigenvalue criterion as an *a priori* error estimate. However, a more precise criterion for tetrahedra can be obtained by applying the theory of Section III to *edge* elements [6].

⁴ Initial meshes T_1 and S_1 were generated using the CosmosTM package by Structural Research and Analysis Corporation (Los Angeles, CA). T_1 : 27 nodes, 44 elements; T_4 : 4,577 nodes, 22,528 elements; S_1 : 125 nodes, 375 elements; S_4 : 35,181 nodes, 192,000 elements.

VII. CONCLUSIONS

1. A new *a priori* criterion of FE accuracy is based on the maximum eigenvalue or on the trace of the (scaled) FE stiffness matrix. This criterion is valid for arbitrary 2D or 3D elements satisfying a few general assumptions, and can be applied either locally (i.e. element-wise) or globally, to the whole mesh.
2. For first order tetrahedral node elements, the eigenvalue criterion is equivalent to the maximum-minimum angle condition for angles between element edges and faces.
3. Several known results follow from the new criterion as particular cases: the shape independence of the L_2 -approximation of the potential; the minimum and maximum angle conditions for first order triangular elements. For edge element approximation, a new precise criterion is derived from the eigenvalue condition in [6].
4. In practical FE analysis, the maximum eigenvalue / trace condition provides a simple *a priori* criterion of the mesh 'quality'. Computation of the trace requires virtually no additional work.
5. The results may prove to be of particular interest for anisotropic problems and for isoparametric or high order elements, where no other shape conditions are immediately available.

ACKNOWLEDGMENT

I would like to thank Dr. Alain Bossavit for initiating and encouraging this study. Discussions with Dr. Pierre Asselin were very useful. I am also grateful to Alexander Plaks for reading the manuscript and suggesting some improvements, and to the anonymous reviewers for their helpful comments.

REFERENCES

- [1] M. Zlámal, "On the finite element method," *Numer. Math.*, vol. 12, pp.394-409, 1968.
- [2] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland: Amsterdam, 1978.
- [3] J. L. Synge, *The Hypercircle in Mathematical Physics; a Method for the Approximate Solution of Boundary Value Problems*, Cambridge: University Press, 1957, pp.209-213.
- [4] I. Babuška, A.K. Aziz, "On the angle condition in the finite element method," *SIAM J. Numer. Anal.*, vol. 13, No. 2, pp.214-226, 1976.
- [5] V.N. Parthasarathy, C.M. Graichen, and A.F. Hathaway, "A comparison of tetrahedron quality measures," *Finite Elements in Analysis and Design*, vol. 15, 1993, pp.255-261.
- [6] I. Tsukerman, "Approximation of conservative fields and the element 'shape matrix'," this journal.
- [7] I. Tsukerman, A. Bossavit, "Shape of finite elements and approximation in electromagnetics," ISTET'97, Palermo, Italy, June 1997.
- [8] F.-X. Zgainski, Y. Maréchal, J.-L. Coulomb, "An *a priori* indicator of finite element quality based on the condition number of the stiffness matrix," *IEEE Trans. Magn.*, vol.33, No. 2, 1997.
- [9] I. Fried, "The l_2 and l_∞ condition numbers of the finite element stiffness and mass matrices, and the pointwise convergence of the method," in: *The Mathematics of Finite Elements and Applications*, ed. J. R. Whiteman, London, Academic Press, 1973, pp.164-169.