Finite Element Approximation Theory Using Families of Reference Elements

Douglas N. Arnold
Richard S. Falk

March 28, 2003

Standard approach to error estimates for **affine families** of finite elements:

Let $\hat{K}$ be a reference element, $K = \mathbf{F}(\hat{K})$, J_F the Jacobian of $\mathbf{F}$.

$$\|v - \pi_K v\|_{L^2(K)} = |J_F|^{1/2} \|\hat{v} - \hat{\pi}\hat{v}\|_{L^2(\hat{K})}$$

$$\leq |J_F|^{1/2} C(\hat{K}) |\hat{v}|_{r+1,\hat{K}}$$

$$\leq |J_F|^{1/2} C(\hat{K}) h^{r+1} |J_F^{-1}|^{1/2} |v|_{r+1,K}$$

$$\leq C(\hat{K}) h^{r+1} |v|_{r+1,K}.$$

Quadrilateral Finite Elements:

Mapping $\mathbf{F}$ now bilinear isomorphism from unit square to convex quadrilateral K.

Affine maps from reference square only produce parallelograms.

Define $\hat{V}$ set of shape functions on reference element $\hat{K}$. For $x = \mathbf{F}(\hat{x})$, $\hat{x} \in \hat{K}$:

Define $V_F(K) = \{v : v(x) = \hat{v}(\hat{x}) : \hat{v} \in \hat{V}\}$.

Approximation theory estimates using previous approach:

Now have scaling problem from reference to physical element:

Example: $r = 1$.

$$|\widehat{v}|^2_{2,\widehat{K}} = \int_{\widehat{K}} \left(\frac{\partial^2 \widehat{v}}{\partial \widehat{x}^2} + \frac{\partial^2 \widehat{v}}{\partial \widehat{x} \partial \widehat{y}} + \frac{\partial^2 \widehat{v}}{\partial \widehat{y}^2} \right)$$

Now

$$\widehat{v}(\widehat{x}, \widehat{y}) = v(x, y) = v(F_1(\widehat{x}, \widehat{y}), F_2(\widehat{x}, \widehat{y})).$$

Writing $F_i = a_i + b_i \widehat{x} + c_i \widehat{y} + d_i \widehat{x} \widehat{y}$,

$$\frac{\partial F_i}{\partial \widehat{x}} = b_i + d_i \widehat{x} = O(h), \qquad \frac{\partial F_i}{\partial \widehat{y}} = c_i + d_i \widehat{y} = O(h)$$

$$\frac{\partial^2 F_i}{\partial \widehat{x}^2} = 0, \qquad \frac{\partial^2 F_i}{\partial \widehat{y}^2} = 0, \qquad \frac{\partial^2 F_i}{\partial \widehat{x} \partial \widehat{y}} = d_i = O(h).$$

Applying the chain rule, we find:

$$\frac{\partial^2 \widehat{v}}{\partial \widehat{x}^2} = \frac{\partial v}{\partial x}\frac{\partial^2 F_1}{\partial \widehat{x}^2} + \frac{\partial v}{\partial y}\frac{\partial^2 F_2}{\partial \widehat{x}^2}$$

$$+\frac{\partial^2 v}{\partial x^2}\left(\frac{\partial F_1}{\partial \widehat{x}}\right)^2 + 2\frac{\partial^2 v}{\partial x \partial y}\frac{\partial F_1}{\partial \widehat{x}}\frac{\partial F_2}{\partial \widehat{x}} + \frac{\partial^2 v}{\partial y^2}\left(\frac{\partial F_2}{\partial \widehat{x}}\right)^2$$

$$\frac{\partial^2 \widehat{v}}{\partial \widehat{y}^2} = \frac{\partial v}{\partial x}\frac{\partial^2 F_1}{\partial \widehat{y}^2} + \frac{\partial v}{\partial y}\frac{\partial^2 F_2}{\partial \widehat{y}^2}$$

$$+\frac{\partial^2 v}{\partial x^2}\left(\frac{\partial F_1}{\partial \widehat{y}}\right)^2 + 2\frac{\partial^2 v}{\partial x \partial y}\frac{\partial F_1}{\partial \widehat{y}}\frac{\partial F_2}{\partial \widehat{y}} + \frac{\partial^2 v}{\partial y^2}\left(\frac{\partial F_2}{\partial \widehat{y}}\right)^2$$

$$\frac{\partial^2 \widehat{v}}{\partial \widehat{x} \partial \widehat{y}} = \frac{\partial v}{\partial x}\frac{\partial^2 F_1}{\partial \widehat{x} \partial \widehat{y}} + \frac{\partial v}{\partial y}\frac{\partial^2 F_2}{\partial \widehat{x} \partial \widehat{y}}$$

$$+\frac{\partial^2 v}{\partial x^2}\frac{\partial F_1}{\partial \widehat{x}}\frac{\partial F_1}{\partial \widehat{y}} + \frac{\partial^2 v}{\partial y^2}\frac{\partial F_2}{\partial \widehat{x}}\frac{\partial F_2}{\partial \widehat{y}}$$

$$+\frac{\partial^2 v}{\partial x \partial y}\left(\frac{\partial F_1}{\partial \widehat{x}}\frac{\partial F_2}{\partial \widehat{y}} + \frac{\partial F_1}{\partial \widehat{y}}\frac{\partial F_2}{\partial \widehat{x}}\right)$$

Problem: first order terms scale like h, not h^2.

Solution well known: use special semi-norm.

If error functional exact for Q_1, instead of only P_1, no mixed partial derivatives occur in application of Bramble-Hilbert lemma and scaling is correct.

Consider another approach:
Avoid use of a special semi-norm

Approximation Theory Using Families of Reference Elements

Consider shape-regular sequence of quadrilateral meshes.

Theorem: Suppose

$$\pi_K v(x) = \widehat{\pi}\widehat{v}(\widehat{x}),$$

$$\|\widehat{\pi}\widehat{v}\|_{L^2(\widehat{K})} \leq C(\widehat{K})\|\widehat{v}\|_{r+1,\widehat{K}},$$

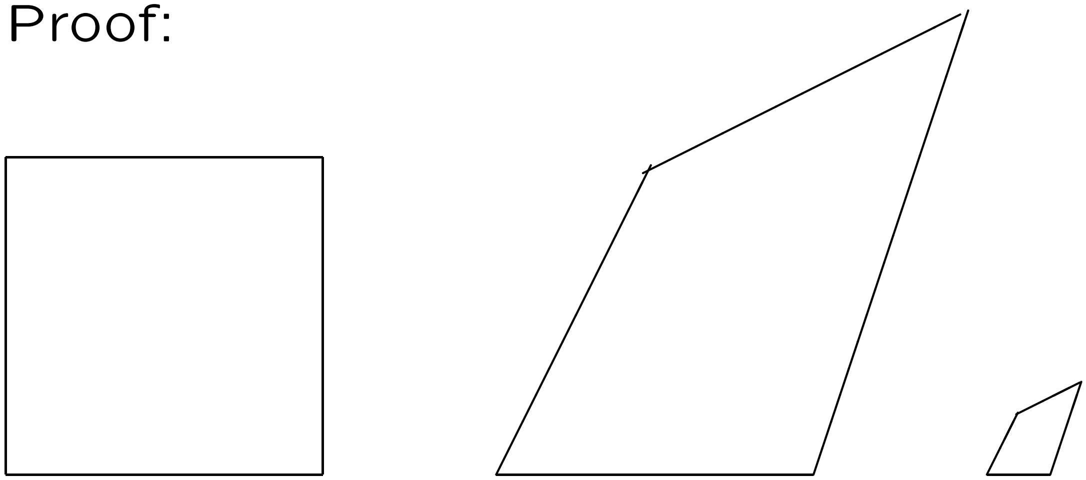
$$\widehat{v} - \widehat{\pi}\widehat{v} = 0 \quad \forall \widehat{v} \in \mathcal{Q}_r.$$

Then for all $v \in H^{r+1}(K)$,

$$||v - \pi_K v||_{L^2(K)} \leq C h_K^{r+1} |v|_{r+1,K},$$

C depends only on shape-regularity constant.

Proof:



$$\hat{K} \to \mathbf{G}_{\tilde{K}} \to \tilde{K} \to \mathbf{M}_K \to K$$

$$\hat{V} \qquad V_{\tilde{G}} \qquad V_F$$

Let $\tilde{K}$ denote quadrilateral related to K by $K = \mathbf{M}_K(\tilde{K})$, where $\mathbf{M}_K(\tilde{x}) = x_K + h_K\tilde{x}$, $h_K =$ diameter of K, $x_K =$ vertex of K.

Define $\mathbf{G}_{\tilde{K}}$ so that $\mathbf{F}_K = \mathbf{M}_K \circ \mathbf{G}_{\tilde{K}}$,
i.e., $\tilde{x} = \mathbf{G}_{\tilde{K}}(\hat{x}) = \mathbf{M}_{K}^{-1}\,\mathbf{F}_K(\hat{x})$.

Define

$$\tilde{v}(\tilde{x}) := v(\,\mathbf{M}_K(\tilde{x})) = v(x) = \hat{v}(\hat{x}).$$

and

$$V_G(\tilde{K}) = \{\,\tilde{v} : \tilde{K} \to \mathbb{R} \,|\, v \in V_F(K)\}.$$

Since $\mathcal{T}_h$ shape-regular, consider elements $\tilde{K}$ as set of reference elements of diameter $O(1)$ and obtain the desired error estimate by Bramble-Hilbert lemma (with $\tilde{K}$ as reference element), provided (for C independent of $\tilde{K}$).

$$(i) \qquad \mathcal{P}_r(\tilde{K}) \subseteq V_G(\tilde{K}),$$

$$(ii) \qquad \tilde{v} - \tilde{\pi}_{\tilde{K}}\tilde{v} = 0 \text{ if } \tilde{v} \in \mathcal{P}_r(\tilde{K}),$$

$$(iii) \qquad \|\tilde{\pi}_{\tilde{K}}\tilde{v}\|_{L^2(\tilde{K})} \leq C\|\tilde{v}\|_{r+1,\tilde{K}},$$

To establish (i), enough to show that if $\tilde{v} \in \mathcal{P}_r(\tilde{K})$, then $\hat{v} \in \mathcal{Q}_r(\hat{K}) = \hat{V}$, since then $v \in V_F(K)$ and so $\tilde{v} \in V_G(\tilde{K})$.

Components of $\mathbf{F}(\hat{x}, \hat{y})$ are bilinear functions of $\hat{x}$ and $\hat{y}$, so if $\tilde{v}(\tilde{x}, \tilde{y})$ (and hence $v(x, y)) \in \mathcal{P}_r$, then $v(\mathbf{F}(\hat{x}, \hat{y}))$ is of degree at most r in $\hat{x}$ and $\hat{y}$ separately, i.e., $\hat{v}(\hat{x}, \hat{y}) = v \circ \mathbf{F} \in \mathcal{Q}_r(\hat{K}) = \hat{V}$.

Using this result, (ii) follows from exactness of interpolant on $\hat{K}$ for $\hat{v} \in \mathcal{Q}_r(\hat{K})$ and identity

$$\tilde{v}(\tilde{x}) - \tilde{\pi}\tilde{v}(\tilde{x}) = \hat{v}(\hat{x}) - \hat{\pi}\hat{v}(\hat{x}).$$

Inequality (iii) follows from analogous estimate for $\widehat{\pi}$ on $\widehat{K}$, assumed in theorem, and a change of variables.

Check: all norms on mapping $\mathbf{G}_{\tilde{K}}$ and its Jacobian, which enter due to change of variable, are bounded by a constant independent of $\tilde{K}$.

For family of shape-regular meshes, constant only depends on shape-regularity constant σ.

Example 2: Quadrilateral Raviart-Thomas (RT) elements

Same problem in proof: Use of full semi-norm in Bramble-Hilbert lemma produces mixed derivatives which do not scale correctly under change of variable.

Using cleverly chosen semi-norm, Thomas showed for quad RT elements of order r:

$\| \mathbf{u} - \pi \mathbf{u}\|_{L^2(\Omega)} \leq Ch^{r+1}(\| \mathbf{u}\|_{r+1} + h\| \operatorname{div} \mathbf{u}|_{r+1})$.

Optimal order of h, but requires more than optimal regularity.

Define $\hat{\mathbf{RT}}_r$ = RT elements of order r on $\hat{K}$, i.e., $\hat{\mathbf{RT}}_r = \mathcal{P}_{r+1,r} \times \mathcal{P}_{r,r+1}$

For functions in $H(\operatorname{div}, \Omega)$, the natural way to transform functions from $\hat{K}$ to K is via *Piola transform*.

Given $\hat{\mathbf{u}} : \hat{K} \to \mathbb{R}^2$, define $\mathbf{u} : K \to \mathbb{R}^2$ by

$$\mathbf{u}(x) = J(\hat{x})^{-1} D\,\mathbf{F}(\hat{x})\,\hat{\mathbf{u}}(\hat{x}) = \mathbf{P}_F\,\hat{\mathbf{u}}.$$

Theorem: Suppose

$$(\pi_K \mathbf{v})(x) = \mathbf{P}_{F_K} \widehat{\pi} \widehat{\mathbf{v}}(\widehat{x}).$$

If $\widehat{\mathbf{V}} = \widehat{\mathbf{RT}}_r$ and $\widehat{\pi}$ satisfies

$$\| \widehat{\pi} \widehat{\mathbf{v}} \|_{L^2(\widehat{K})} \leq C(\widehat{K}) \| \widehat{\mathbf{v}} \|_{r+1,\widehat{K}}$$

and $\widehat{\mathbf{v}} - \widehat{\pi} \widehat{\mathbf{v}} = 0$ for all $\widehat{\mathbf{v}} \in \widehat{\mathbf{V}}$, then for all $\mathbf{v} \in H^{r+1}(K, \mathbb{R}^2)$,

$$|| \mathbf{v} - \pi_K \mathbf{v} ||_{L^2(K)} \leq C h^{r+1} | \mathbf{v} |_{r+1,K},$$

where C depends only on shape-regularity constant.

Proof: Following proof of previous theorem, enough to establish the analogues of (i)-(iii), i.e.,

$$(i) \qquad \mathcal{P}_r(\tilde{K}, \mathbb{R}^2) \subseteq \mathbf{V}_G(\tilde{K}),$$

$$(ii) \qquad \tilde{\mathbf{v}}(\tilde{x}) - \tilde{\pi}_{\tilde{K}} \tilde{\mathbf{v}}(\tilde{x}) = 0 \quad \text{if } \tilde{\mathbf{v}} \in \mathcal{P}_r(\tilde{K}, \mathbb{R}^2),$$

$$(iii) \qquad \| \tilde{\pi}_{\tilde{K}} \tilde{\mathbf{v}} \|_{L^2(\tilde{K})} \leq C \| \tilde{\mathbf{v}} \|_{r+1,\tilde{K}}$$

.

To establish (i), enough to show that if $\tilde{\mathbf{v}} \in \mathcal{P}_r(\tilde{K}, \mathbb{R}^2)$, then $\hat{\mathbf{v}} \in \hat{\mathbf{RT}}_r = \hat{\mathbf{V}}$. Writing $\mathbf{F}$ in form $(a_1+b_1\hat{x}+c_1\hat{y}+d_1\hat{x}\hat{y}, a_2+b_2\hat{x}+c_2\hat{y}+d_2\hat{x}\hat{y})$, straightforward calculation shows

$$\hat{v}_1 = DF_{22}v_1 - DF_{12}v_2$$

$$= (c_2 + d_2\hat{x})v_1 + (-c_1 - d_1\hat{x})v_2,$$

$$\hat{v}_2 = -DF_{21}v_1 + DF_{11}v_2$$

$$= (-b_2 - d_2\hat{y})v_1 + (b_1 + d_1\hat{y})v_2.$$

From these formulas, easy to see that if $\tilde{\mathbf{v}} \in \mathcal{P}_r(\tilde{K}, \mathbb{R}^2)$ (and hence $\mathbf{v} \in \mathcal{P}_r(K, \mathbb{R}^2)$), then $\hat{\mathbf{v}} \in \mathcal{P}_{r+1,r} \times \mathcal{P}_{r,r+1} = \hat{\mathbf{RT}}_r$.

Using this result, analogue of (ii) follows from the exactness of interpolant on $\hat{K}$ for $\hat{\mathbf{v}} \in \hat{\mathbf{RT}}_r$ and identity

$$\tilde{\mathbf{v}}(\tilde{x}) - \tilde{\pi}_{\tilde{K}}\tilde{\mathbf{v}}(\tilde{x}) = \mathbf{P}_{G_{\tilde{K}}}[\hat{\mathbf{v}}(\hat{x}) - \hat{\pi}\,\hat{\mathbf{v}}(\hat{x})],$$

The analogue of (iii) follows a similar argument to the scalar case. Although change of variables complicated by introduction of Piola transform, constants in estimates will only depend on shape-regularity constant.

Remark: To obtain approximation theory estimates in $H(\operatorname{div}, \Omega)$, one also needs estimates for $\|\operatorname{div} \mathbf{v} - \operatorname{div} \pi \mathbf{v}\|_{L^2(K)}$. Unfortunately, on quadrilaterals, Raviart-Thomas elements give an approximation of $\operatorname{div} \mathbf{v}$ that is suboptimal by one power of h.

Example 3: Argyris triangular element

Used for approximation of plate problems.

Degrees of freedom: value of function and its derivatives up to order two at triangle vertices and value of normal derivative at midpoints of triangle edges.

$\mathcal{P}_5$ on each triangle and resulting finite element space $\in C^1(\Omega) \cap H^2(\Omega)$.

Problem: Normal derivative not preserved under affine mappings – so not an affine family of finite elements.

Proof given in Ciarlet compares approximation error by Argyris element to approximation error by Hermite triangle of type (5) – replaces normal derivative degrees of freedom at midpoints of edges by $\nabla u(b_i) \cdot (a_i - b_i)$, where a_i and b_i denote triangle vertices and edge midpoints, respectively.

Latter is an affine family and so its approximation theory error derived by standard method. Then use ad hoc method to bound error between interpolation operators in two finite element spaces.

Instead, consider previous ideas. In this case, can not build up space from reference element, but simply use degrees of freedom and interpolation operator π_K defined directly on K.

Leads directly to definition of function $\tilde{v} : \tilde{K} \to \mathbb{R}$ defined by $\tilde{v}(\tilde{x}) = v(x)$ and an interpolation operator $\tilde{\pi}_{\tilde{K}}$ satisfying $\tilde{\pi}_{\tilde{K}}\tilde{v}(\tilde{x}) = \pi_K v(x)$. Since the triangles $\tilde{K}$ and K are similar, the normal derivative degrees of freedom are preserved under the simple mapping $\mathbf{M}_K$. Get for shape-regular triangulations:

Theorem: For all $v \in H^6(K)$ and $0 \leq m \leq 2$,

$$||v - \pi_K v||_{H^m(K)} \leq Ch^{6-m}|v|_{6,K},$$

where C depends only on the shape-regularity constant.

Proof: Need only establish (i)-(iii) for $r = 5$. In this case, (i) and (ii) are trivial since $V_F(K)$ is defined directly as $\mathcal{P}_5(K)$ and the interpolation operator, defined directly on the degrees of freedom on K, is exact for functions in $\mathcal{P}_5(K)$. This implies analogous properties for $V_G(\tilde{K})$, since the two spaces are related by simple mapping $\mathbf{M}_K$.

In particular, although normal derivative degree of freedom not preserved for general affine map, it is preserved for mapping $\mathbf{M}_K$.

Since $\tilde{K}$ has diameter $O(1)$, (iii) follows in standard way by the Sobolev embedding theorem. Although constant will in general depend on $\tilde{K}$, for shape-regular meshes, it depends only on shape-regularity constant.

Example 4: Arnold-Winther mixed elasticity element

Approximate symmetric stress tensor by finite element defined locally on each triangle K by

$$\boldsymbol{\Sigma}_K = \{\, \tau \in \mathcal{P}_3(\tau, \mathbb{S}) |\ \mathsf{div}\, \tau \in \mathbf{RM}(K)\},$$

where $\mathcal{P}_3(\tau, \mathbb{S})$ denotes symmetric 2×2 matrices with entries in $\mathcal{P}_3$ and $\mathbf{RM}(K)$ denotes the space of infinitesimal rigid motions on K, i.e., functions of the form $(a - by, c + bx)$ (for constant a, b, c).

Dimension of $\boldsymbol{\Sigma}_K$ is 21 and a unisolvent set of degrees of freedom given by values of three components of $\tau(x)$ at each vertex x of K (9 degrees of freedom) and values of moments of degree 0 and 1 of two normal components of τ on each edge e of K (12 degrees of freedom).

Standard approach to deriving error estimates: map back to a reference element.

In mixed formulation approach, seek symmetric stress tensor in space $H(\operatorname{div}, \Omega)$, so natural mapping is matrix Piola transform.

Defined for $\mathbf{F}(\hat{x}) = \mathcal{B}\hat{x} + \mathbf{b}$ by $\tau(x) = \mathcal{B}\,\hat{\tau}(\hat{x})\mathcal{B}^T$.

Simple calculation: $\operatorname{div} \tau(x) = \mathcal{B}\ \widehat{\operatorname{div}}\, \hat{\tau}(\hat{x})$.

Problem: condition $\operatorname{div} \tau \in \mathbf{RM}(K)$ not preserved by Piola transform associated to mapping $\mathcal{B}$.

Approach fails. If instead, we use family of reference elements, then only mappings $\mathcal{B}$ of form $h_K I$ enter into calculations. These preserve condition $\operatorname{div} \tau \in \mathbf{RM}(K)$.

Theorem: For all $\tau \in H^2(K, \mathbf{S})$,

$$||\tau - \pi_K \tau||_{L^2(K)} \le C h^2 |\tau|_{2,K},$$

where C depends only on the shape-regularity constant.

Proof: Only new element is to check which degree polynomials are included in space $\mathbf{\Sigma}_K$ and hence in corresponding space on $\tilde{K}$. A straightforward computation shows that $\mathcal{P}_1(K,\mathbb{S}) \subsetneq \mathbf{\Sigma}_K \subsetneq \mathcal{P}_2(K,\mathbb{S})$.

Remark: For plate application, also need estimate

$$||\tau - \pi_K \tau||_{L^2(K)} \le Ch|\tau|_{1,K}.$$

Cannot be obtained with this interpolant because of vertex values of τ. Get estimate with modified interpolant.