Error estimates for $\mathcal{Q}_1$ isoparametric elements satisfying a weak angle condition

Gabriel Acosta * and Ricardo G. Durán *

December 13, 1999

Abstract. We prove optimal order error estimates in the H^1 norm for the $\mathcal{Q}_1$ isoparametric interpolation on convex quadrilateral elements under rather weak hypotheses improving previously known results.

Choose one diagonal and divide the element into two triangles. We show that, if the chosen diagonal is the longest one then, the constant in the error estimate depends only on the maximum angle of the two triangles. Otherwise, the constant depends on that maximum angle and on the ratio between the two diagonals. In particular, we obtain the optimal order error estimate under the maximum angle condition as in the case of triangular elements.

Consequently, the error estimate is uniformly valid for a rather general class of degenerate quadrilaterals.

Keywords. Lagrange interpolation, isoparametric finite elements, quadrilateral elements, anisotropic elements, maximum angle condition.

AMS Subject Classification. 65N15, 65N30.

1 Introduction

The aim of this paper is to prove optimal error estimates in the H^1 norm for the $\mathcal{Q}_1$ isoparametric Lagrange interpolation for rather general elements including degenerate or anisotropic quadrilaterals.

Calling Q the Lagrange interpolation operator and using the standard notation for Sobolev norms and seminorms, we look for geometric conditions under which the estimate

$$\|u - Qu\|_{0,K} + h|u - Qu|_{1,K} \leq Ch^2|u|_{2,K}$$

holds uniformly for a class of convex quadrilateral elements K where h is the diameter of K or, in other words, we want to analyze the dependence of the constant C on the geometry of the element.

Let us first mention that the optimal error estimate in the L^2 norm holds with a constant independent of the geometry of the element. Indeed, as will be shown in our main theorem, this follows in a simple way from the results in [9].

Therefore, we will concentrate on the estimate

*Departamento de Matemática, Facultad de Ciencias Exactas y Naturales, 1428 Buenos Aires, Argentina. Supported by Universidad de Buenos Aires under grant TX048 and by ANPCyT under grant PICT No. 03-00000-00137. The second author is a member of CONICET, Argentina.

$$|u - Qu|_{1,K} \leq Ch|u|_{2,K} \tag{1.1}$$

This error estimate for $\mathcal{Q}_1$ isoparametric finite elements was first obtained by Ciarlet and Raviart [6]. Denoting with $\underline{h}$ the length of the shortest side of K, they proved that, if

$$h/\underline{h} \leq \mu_1$$

and,

$$|\cos\theta| \leq \mu_2 < 1$$

for all angle θ of K then, (1.1) holds with a constant C depending only on μ_1 and μ_2. These hypotheses exclude degenerate elements.

Error estimates for degenerate elements have deserved much attention since the works by Babuška and Aziz [5] and Jamet [8] which showed that, for triangular elements, the usual regularity assumption (bounded ratio between outer and inner diameter) is not needed. Indeed, the above quoted papers proved that, for triangles, the constant C in the estimate (1.1) depends only on the maximum angle of the element. This result for triangles has been generalized in several directions including higher degree elements, higher dimensions and different kinds of interpolations [1, 2, 3, 4, 7, 11, 13].

For quadrilaterals, the term "degenerate" may refer to elements which are close to being a triangle or to narrow elements, which are of interest in anisotropic problems.

In the first case Jamet [9] improved the results of [6] proving that the error estimate (1.1) holds with a constant depending only on σ defined by

$$h/\rho \leq \sigma \tag{1.2}$$

where ρ denotes the diameter of the maximum circle contained in K. Under this hypothesis, the quadrilateral can degenerate into a triangle but it can not become too narrow.

On the other hand, for narrow or anisotropic elements, the estimate (1.1) has been obtained by Zenisek and Vanmaele [15, 16] and by Apel [2] under hypotheses which imply in particular that the elements are far away from triangles. In [15, 16] the authors also required that the two longest sides of the element be opposite and almost parallel and the constant C obtained by them in (1.1) depends on an angle which, in some cases, is the minimum angle of the element. As we will show, this dependence can be removed. In [2] estimates taking into account the anisotropy of the element are obtained for quadrilaterals which are small perturbations of a rectangle. If h_1 and h_2 are the element sizes in the directions of x_1 and x_2 respectively, the estimate obtained in [2] for the case of $\mathcal{Q}_1$ isoparametric elements reads as follows:

$$|u - Qu|_{1,K} \leq C\Big\{h_1\Big|\frac{\partial}{\partial x_1}\nabla u\Big|_{0,K} + h_2\Big|\frac{\partial}{\partial x_2}\nabla u\Big|_{0,K}\Big\}. \tag{1.3}$$

In view of the known results for the case of triangular elements it is natural to ask whether the estimate (1.1) holds with a constant C depending only on the maximum angle of the quadrilateral K. A positive answer will follow from our results. However, in view of the results in [9], it is clear that the maximum angle condition is not necessary. We will prove (1.1) under a weaker assumption. The results of [9, 15, 16] mentioned above for both kinds of degenerate elements follow from our result and, as we will show, there are examples covered by our theory which are not included in the cases considered in those papers. The class of elements considered in [2] is also strictly included in ours although our estimate is weaker than that in [2] for the elements considered there.

Our argument is based on two key points. First, we introduce an appropriate affine change of variables which reduces the problem to a reference family of elements depending on some parameters. Second, in order to prove the result for these reference elements, we analyze the error between Qu and the linear interpolation of u defined by three vertices of K, giving a very precise estimate for this term.

The rest of the paper is as follows. In Section 2 we introduce the sufficient condition for (1.1) that we will call "*regular decomposition property*", we analyze the relation with previously used hypotheses and present several examples in order to motivate our condition. In Section 3 we introduce an appropriate change of variables which reduces the problem to a family of reference configurations. Finally, Section 4 deals with the optimal error estimate.

2 The regular decomposition property

The object of this section is to introduce the condition that will be shown to be sufficient for (1.1). In order to motivate our condition we analyze first several examples.

We will use the following notation. For a general convex quadrilateral K, M_i will denote its vertices in anticlockwise order. Given $a, b, p, q > 0$, $K(a, b, p, q)$ will represent a convex quadrilateral with vertices at $\overline{M}_1 = (0, 0)$, $\overline{M}_2 = (a, 0)$, $\overline{M}_3 = (p, q)$ and $\overline{M}_4 = (0, b)$. In particular $\hat{K} = K(1, 1, 1, 1)$ is the reference unit square and its vertices will be denoted with $\hat{M}_i$. We will use the variable $\overline{x}$ on $K(a, b, p, q)$, $\hat{x}$ on $\hat{K}$ and x on K. In order to define the isoparametric elements on K, let $F_K : \hat{K} \to K$ be the transformation $F_K(\hat{x}) = \sum_{i=1}^{4} M_i \hat{\phi}_i(\hat{x})$ where $\hat{\phi}_i$ is the bilinear basis function associated with the vertex $\hat{M}_i$, i.e., $\hat{\phi}_i(\hat{M}_j) = \delta_i^j$. Now, the basis functions on K , no longer bilinears in general, are defined by $\phi_i(x) = \hat{\phi}_i(F_K^{-1}(x))$ and the $\mathcal{Q}_1$ isoparametric interpolation operator Q on K is defined by

$$Qu(x) = \hat{Q}\hat{u}(\hat{x})$$

where $x = F_K(\hat{x})$ and $\hat{Q}$ is the bilinear Lagrange interpolation of $\hat{u} = u \circ F_K$ on $\hat{K}$.

In order to introduce in a natural way our geometrical sufficient condition for the estimate (1.1) we will consider some examples.

We will use the following definition:

Definition 2.1 *We say that a quadrilateral K (resp. a triangle T) satisfies the maximum angle condition with a constant $\psi < \pi$, or shortly $MAC(\psi)$, if the angles of K (resp. T) are less than or equal to ψ.*

Since (1.1) holds uniformly for quadrilaterals which degenerate into a "regular" triangle (i.e., a triangle such that (1.2) holds) one could think that this is also true for quadrilaterals which degenerate into triangles satisfying $MAC(\psi)$. However, the following example shows that this is not true: on $K = K(1, a, a, a)$ (see Figure 1 below) take $u = x_1^2$. Then, it is not difficult to find the explicit expression for the interpolation $Qu = \frac{x_1(1+(a^2-1)x_2/a)}{1+x_2(a-1)/a}$ and so, a straightforward computation shows that $\|\frac{\partial(u-Qu)}{\partial x_2}\|_0^2 \geq Ca\ln(a^{-1})$ for a small. Now, since the diameter of K is one and $\|D^2u\|_0^2 \leq Ca$, the estimate (1.1) does not hold with a constant independent of a.

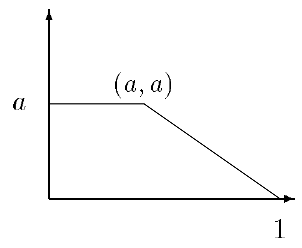

Figure 1

On the other hand, as mentioned in the introduction, to impose a maximum angle condition on the quadrilaterals would be a too strong hypothesis excluding elements degenerating into regular triangles. A weaker condition would be to divide the element by one of its diagonals into two triangles and to impose the maximum angle condition for one of them. The example given above shows that this is not a sufficient condition for (1.1). In fact, dividing the element in Figure 1 by any of its diagonals one of the triangles obtained has angles bounded by $\pi/2$.

What happens if we impose the maximum angle condition for both triangles? Again, the same example shows that this is not a sufficient condition. Indeed, dividing now the element along the diagonal joining the vertices at $(0, 0)$ and (a, a), it is easy to see that both triangles satisfy the maximum angle condition with a constant ψ independent of a.

Observe that in this example, the ratio between the two diagonals is not bounded uniformly in a.

Based on these considerations, we introduce the following geometric condition which, as we will show, is sufficient for (1.1).

We require that the two triangles obtained satisfy the maximum angle condition whenever the element is divided along its longest diagonal or by any of its diagonals if they are comparable, in other words, the constant in the error estimate will depend on the maximum angles of the two triangles but also, on the ratio between the diagonals. This is formalized in the following definition. We will call d_1 and d_2 the diagonals of K.

Definition 2.2 *Let K be a convex quadrilateral. We say that K satisfies the* regular decomposition property *with constants $N \in \mathbb{R}$ and $0 < \psi < \pi$, or shortly $RDP(N, \psi)$, if we can divide K into two triangles along one of its diagonals, that will be called always d_1, in such a way that $|d_2|/|d_1| \leq N$ and both triangles satisfy $MAC(\psi)$.*

Remark 2.1 *If a quadrilateral K satisfies $MAC(\psi)$ then, it also satisfies $RDP(1, \psi)$. In fact, this can be easily seen by dividing K by its longest diagonal. On the other hand the converse is not true, as shown by the example in Figure 2,* a).

Remark 2.2 *Let K be a convex quadrilateral satisfying the Jamet's regularity assumption (1.2) then, K satisfies $RDP(\sigma, \psi)$, with $\psi = \psi(\sigma) < \pi$. In fact, the angles of K are bounded from below by a positive constant $\delta(\sigma)$ (see [9]) and therefore there is at most one of the angles of K which is not bounded by $\pi - \delta(\sigma)$ and so, dividing K by the diagonal containing the vertex corresponding to this angle it follows easily that K satisfies $RDP(\sigma, \pi - \delta(\sigma))$.*

Remark 2.3 *It is not difficult to check that if K satisfies the hypotheses given in [15, 16] then, K satisfies $MAC(\psi)$ for certain $\psi < \pi$ whenever K degenerates in the sense of $h/\rho \to \infty$ (which* *is the only interesting case, because otherwise we fall into the case of the preceding remark) and so, by Remark 2.1 it satisfies also* $RDP(1, \psi)$.

Remark 2.4 *The case of anisotropic elements (i.e.: elements with different sizes in each direction) has also been considered in [2]. Since the elements considered there are small perturbations of rectangles, it is not difficult to see that they satisfy* $MAC(\psi)$ *and so* $RDP(1, \psi)$.

Remark 2.5 *One may ask whether or not the regular decomposition property is a weaker hypothesis than those used before by the quoted authors. The answer is positive as shown by the example* $K = (1, a, a^2, a + a^2)$, *(see Figure 2b)). As is easily seen, (1.2) does not hold with* σ *independent of* a *and on the other hand, for* a *small, the longest sides are not opposite as assumed in [15, 16]. Also, it is not difficult to see that* K *is not a perturbation of a rectangle as required in [2]. However, the longest diagonal joining the vertices* $(0, a)$ *and* $(1, 0)$ *divides this element in two triangles satisfying* $MAC(\psi)$ *with a constant* $\psi < \pi$ *independent of* a.

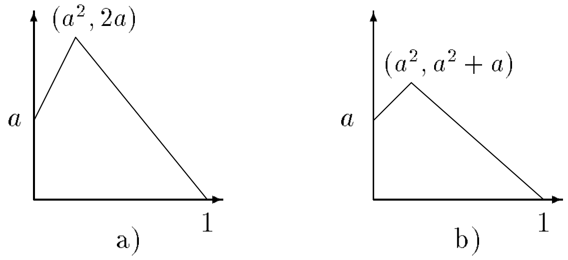

Figure 2

3 The change of variables

The first step of our argument is to reduce the problem to a family of reference configurations. In order to do that, we will introduce in this section an appropriate affine change of variables. First we give some technical lemmas.

Lemma 3.1 *Let* K *be a quadrilateral satisfying* $RDP(N, \psi)$. *We have*

1. *If* β *is any of the angles of* K *opposite to* d_1 *then, there exists* $\delta = \delta(N, \psi)$ *independent of* β *such that*
$$0 < \delta \leq \beta \tag{3.1}$$
and,

2. *If* s *is the shortest side of* K *and* α *is the angle between* s *and* d_1 *then,*
$$(\pi - \psi)/2 \leq \alpha. \tag{3.2}$$

Proof. *1)* Let T_1 and T_2 be the two triangles obtained after dividing K by the diagonal d_1. Then, T_1 and T_2 satisfy $MAC(\psi)$ and $|d_2|/|d_1| \leq N$. Assume that β is an angle of T_2 (as in Figure 3). If β is not the minimum angle of T_2 then, (3.1) holds with $\delta = (\pi - \psi)/2$ because T_2 satisfies

$MAC(\psi)$. Assume now that β is the minimum angle of T_2 and call γ_1 and γ_2 the other two angles of this triangle. Using again $MAC(\psi)$ for T_2 we have that $\gamma_i \geq (\pi - \psi)/2$ for $i = 1, 2$. Now, the diagonal d_2 divides T_2 into two triangles $\tilde{T}_2$ and $\overline{T}_2$, β into β_1 and β_2 and, d_1 into $\tilde{d}_1$ and $\overline{d}_1$. We choose the notation in such a way that $|\tilde{d}_1| \geq |d_1|/2$ and $\tilde{T}_2$ has $\tilde{d}_1$ as one of its sides and β_1 and γ_1 as two of its angles. Then, if we call $\tilde{d}_2$ the side of $\tilde{T}_2$ opposite to γ_1 we have

$$\beta \geq \beta_1 \geq \sin \beta_1 \geq \frac{|\tilde{d}_1|}{|\tilde{d}_2|} \sin \gamma_1 \geq \frac{1}{2} \frac{|d_1|}{|d_2|} \sin \gamma_1 \geq \frac{1}{2N} \sin \gamma_1$$

and so, (3.1) holds since $(\pi - \psi)/2 \leq \gamma_1 \leq \psi$.

2) Assume, without loss of generality, that s is one of the sides of T_1. To prove (3.2) it is enough to observe that α is not the minimum angle of T_1 since it is adjacent to the shortest side, which is s or d_1. □

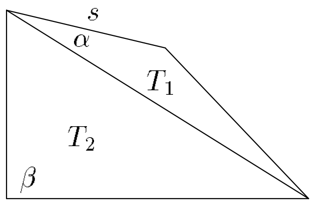

Figure 3

In the following lemma we recall a result concerning the change of an angle after a linear transformation. The proof is elementary and we omit it (see [1]).

Lemma 3.2 *Let L be the linear transformation associated with a matrix B. Given two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$, let α_1 be the angle between them and α_2 be the angle between $L(\mathbf{v}_1)$ and $L(\mathbf{v}_2)$. Calling $\kappa(B)$ the condition number of B we have*

$$\frac{2}{\kappa(B)\pi}\alpha_1 \leq \alpha_2 \leq \pi\Big(1 - \frac{2}{\pi\kappa(B)}\Big) + \frac{2}{\kappa(B)\pi}\alpha_1. \tag{3.3}$$

In the following lemma we introduce the affine change of variables which transforms a given element into a reference one. We will use the notation of the previous lemmas.

Lemma 3.3 *Let K be a quadrilateral of diameter h satisfying $RDP(N, \psi)$. Then, there exist $a, b, p, q > 0$, an affine transformation $L\overline{x} = B\overline{x} + P$ such that $L(K(a, b, p, q)) = K$ and constants $C = C(N, \psi)$, $\overline{N} = \overline{N}(N, \psi)$, $\overline{\psi} = \overline{\psi}(N, \psi)$ and $\overline{\delta} = \overline{\delta}(N, \psi)$ such that*

1. $\|B\| \leq C$ *and,* $\|B^{-1}\| \leq C$, *in particular,* $\kappa(B) \leq C^2$.
2. *$K(a, b, p, q)$ satisfies $RDP(\overline{N}, \overline{\psi})$, taking as the dividing diagonal $\overline{d}_1 = L^{-1}(d_1)$.*
3. *If ℓ is the shortest side of K and $\overline{\alpha}$ is the angle between $\overline{d}_1 = L^{-1}(d_1)$ and $\overline{\ell} = L^{-1}(\ell)$ then,*

$$0 < \overline{\delta} < \overline{\alpha} < \overline{\psi} < \pi.$$

4. If $\overline{s}$ is the shortest side of $K(a,b,p,q)$ then,

$$|\overline{\ell}| \leq C^2 |\overline{s}|.$$

5. If $\overline{h}$ is the diameter of $K(a,b,p,q)$ then,

$$C^{-1} h \leq \overline{h} \leq C h.$$

Proof. We will construct $K(a,b,p,q)$ and L explicitly. Since K satisfies $RDP(N,\psi)$ we can divide it along d_1 into T_1 and T_2 which have all their angles bounded by ψ. We choose the notation in such a way that the shortest side ℓ of K is one of the sides of T_1.

Let us call β the angle of T_2 opposite to d_1. After a rigid movement we can assume that the vertex corresponding to β is placed at the origin and call it M_1, numbering the other vertices M_i in anticlockwise order. We can also assume that the side opposite to ℓ lies on the x axis and we call a the length of this side (see Figure 4).

Now, let ℓ_{14} be the side with vertices at M_1 and M_4 and define $b = |\ell_{14}| \sin\beta$. Then, we have that M_4 is placed at $(b \operatorname{cotg}\beta, b)$.

We define L as the linear transformation associated with the matrix

$$B = \begin{pmatrix} 1 & \operatorname{cotg}\beta \\ 0 & 1 \end{pmatrix}$$

and p, q such that $L(K(a,b,p,q)) = K$.

We now prove the five properties stated in the Lemma.

1) An elementary estimate yields $\|B\| \leq \sqrt{2}/\sin\beta$ and $\|B^{-1}\| \leq \sqrt{2}/\sin\beta$. Now, since K satisfies $RDP(N,\psi)$ we know that $\beta \leq \psi$ and also, from Lemma 3.1 that, $0 < \delta = \delta(N,\psi) \leq \beta$. Therefore, *1)* holds with $C = \sqrt{2}/\min\{\sin\delta, \sin\psi\}$.

2) Calling $\overline{d}_i = L^{-1}(d_i)$, and recalling that from $RDP(N,\psi)$we have $|d_2|/|d_1| \leq N$ we obtain

$$|\overline{d}_2|/|\overline{d}_1| \leq \kappa(B)|d_2|/|d_1| \leq \kappa(B)N$$

and so, we can take $\overline{N} = \kappa(B)N$.

On the other hand, in view of Lemma 3.2, the angles of the triangles obtained by dividing $K(a,b,p,q)$ along $\overline{d}_1$ are bounded by $\pi(1 - 2/\pi\kappa(B)) + (2/\pi\kappa(B))\psi$ and so, *2)* follows from *1)*.

3) It follows immediately from *1)* combined with part *2)* of Lemma 3.1 and Lemma 3.2.

4) Let ℓ be the shortest side of K, $\overline{\ell} = L^{-1}(\ell)$ and $\overline{s} = L^{-1}(s)$ be the shortest side of $K(a,b,p,q)$ Then,

$$|\overline{\ell}| \leq \|B^{-1}\||\ell| \leq \|B^{-1}\||s| \leq \|B^{-1}\|\|B\||\overline{s}|$$

and so *4)* follows from *1)*.

5) Follows immediately from *1)*. □

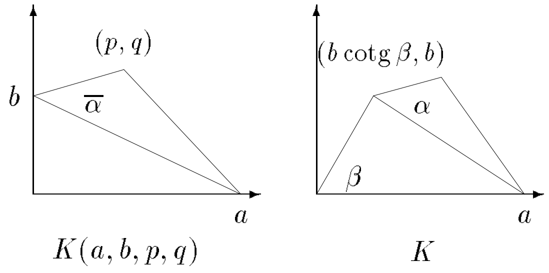

Figure 4

The following lemma shows that it is enough to prove the error estimate (1.1) for the domains $K(a,b,p,q)$.

Lemma 3.4 *Given a quadrilateral K satisfying $RDP(N,\psi)$ let L and $\overline{K} = K(a,b,p,q)$ be the affine transformation and the reference element given in Lemma 3.3. If $\overline{Q}$ is the isoparametric interpolation on $\overline{K}$ and $\overline{u} = u \circ L$ then, there are positive constants C_1 and C_2 depending only on N and ψ such that*

$$C_1 |\overline{u} - \overline{Q}\overline{u}|_{1,\overline{K}} \leq |u - Qu|_{1,K} \leq C_2 |\overline{u} - \overline{Q}\overline{u}|_{1,\overline{K}}$$

and,

$$C_1 |\overline{u}|_{2,\overline{K}} \leq |u|_{2,K} \leq C_2 |\overline{u}|_{2,\overline{K}}.$$

Proof. By definition of the isoparametric interpolation we have

$$Qu(x) = \hat{Q}\hat{u}(F_K^{-1}(x)) \qquad \text{and,} \qquad \overline{Q}\overline{u}(\overline{x}) = \hat{Q}\hat{\overline{u}}(F_{\overline{K}}^{-1}(\overline{x}))$$

Then, the lemma follows easily by observing that, since L is an affine transformation, $F_K = L \circ F_{\overline{K}}$ and so $\overline{Q}\overline{u}(\overline{x}) = Qu(x)$ and using the bound $C(N,\psi)$ for $\|B\|$ and $\|B^{-1}\|$ given in Lemma 3.3 and the fact that $\det B = 1$. □

4 Error estimate

In this section we prove the optimal order error estimate for elements satisfying $RDP(N,\psi)$. In view of the results of the previous section it is enough to analyze the case of the reference configurations $K(a,b,p,q)$ satisfying the conditions of Lemma 3.3. To simplify notation we drop the bars used in the previous section and write $K = K(a,b,p,q)$ an element with diameter h, diagonals d_1 and d_2 and vertices M_i. We call T_1 and T_2 the two triangles obtained after dividing K along d_1. With the notation of Figure 5 below, we assume that there exist positive constants N_1, N_2, N_3 and $\overline{\psi} < \pi$ such that K satisfies the following four hypotheses

(H1) $|d_2|/|d_1| \leq N_1$

(H2) $1/\sin\alpha \leq N_2$

(H3) $|\ell| \leq N_3|s|$ where s is the shortest side of K

(H4) $\theta \leq \overline{\psi}$, for all angle θ of T_1 and T_2.

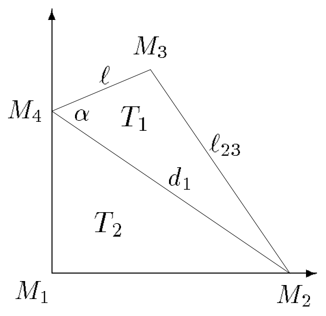

Figure 5

We recall that, from Lemma 3.3, any quadrilateral satisfying $RDP(N, \psi)$ can be transformed by an affine transformation into some $K(a, b, p, q)$ satisfying H1, H2, H3 and H4 with constants depending only on N and ψ.

As a consequence of H3 we have that

$$p/a \leq N_3 \qquad \text{and,} \qquad q/b \leq N_3 + 1. \tag{4.1}$$

Indeed, $p/a \leq |\ell|/a$ and $|q - b|/b \leq |\ell|/b$ and so (4.1) follows from H3. To simplify notation we define $N_4 = N_3 + 1$.

On the other hand, from H1 and H3 it follows that

$$h \leq N_3(1 + N_1)|d_1|. \tag{4.2}$$

Indeed, if h is the length of one of the diagonals then, it follows from H1 that $h \leq \max\{1, N_1\}|d_1|$ and so (4.2) holds because $N_3 \geq 1$. Otherwise, h agrees with the length of one of the sides of K and so $h = |\ell|$ or $h = |\ell_{23}|$, where ℓ_{23} is the side joining M_2 and M_3, because the lengths of the other two sides are bounded by $|d_1|$. Now, in view of H3, $|\ell| \leq N_3|\ell_{23}|$ and therefore, it is enough to see that $|\ell_{23}| \leq (1 + N_1)|d_1|$. But, from the triangle inequality, $|\ell_{23}| \leq a + |d_2| \leq |d_1| + |d_2| \leq (1 + N_1)|d_1|$ and therefore (4.2) holds.

In what follows we will use the standard notation (x, y) for a variable in $\mathbb{R}^2$ (instead of $x \in \mathbb{R}^2$ as in previous sections).

In order to prove the estimate we decompose the error as in [9] in the following way:

Let Π be the $\mathcal{P}_1$-Lagrange interpolation operator associated with the vertices $M_1 = (0, 0)$, $M_2 = (a, 0)$ and $M_4 = (0, b)$, i.e., Πu is an affine function which agrees with u at these three vertices.

Then, we have

$$|u - Qu|_{1,K} \leq |u - \Pi u|_{1,K} + |\Pi u - Qu|_{1,K}$$

but, since $\Pi u - Qu$ is in the finite element isoparametric space and vanishes at M_1, M_2 and M_4 it follows that

$$(\Pi u - Qu)(x) = (\Pi u - u)(M_3)\phi_3(x)$$

where ϕ_3 is the basis function corresponding to M_3. Therefore,

$$|u - Qu|_{1,K} \le |u - \Pi u|_{1,K} + |(\Pi u - u)(M_3)||\phi_3|_{1,K} \tag{4.3}$$

and so, it is enough to estimate the two terms on the right hand side and this is the object of the rest of the paper. We start with the terms $|u - \Pi u|_{1,K}$ and $|(\Pi u - u)(M_3)|$ and leave the estimate of $|\phi_3|_{1,K}$, which is the most technical one, for the end.

In order to bound $|u - \Pi u|_{1,K}$ and $|(\Pi u - u)(M_3)|$ we will use a trace theorem in the form given in [14] which gives a sharp dependence of the constant on the geometry of the triangle. We state it in the following lemma and refer to [14] for the proof.

Lemma 4.1 *Let T be a triangle with diameter h_T and e be any of its sides. Then,*

$$\|u\|_{0,e} \le \left(2\frac{|e|}{|T|}\right)^{1/2} \{\|u\|_{0,T} + h_T|u|_{1,T}\}.$$

In the next lemma we give an estimate for $|(u - \Pi u)(M_3)|$ in terms of $|u - \Pi u|_{1,K}$. We will use the notation of Figure 5.

Lemma 4.2 *If $K = K(a, b, p, q)$ is convex and satisfies the hypotheses* H1, H2 and H3 *then,*

$$|(u - \Pi u)(M_3)| \le 2(N_2N_3(1 + N_1))^{1/2} \left(\frac{|\ell|}{h}\right)^{1/2} \{|u - \Pi u|_{1,T_1} + h|u|_{2,T_1}\}. \tag{4.4}$$

Proof. Let us denote with ∂_ℓ the derivative in the direction of ℓ. Using that $(u - \Pi u)(M_4) = 0$, the Schwartz inequality on ℓ and the trace theorem in the form given in Lemma 4.1 we have

$$|(u - \Pi u)(M_3)| = \Big|\int_\ell \partial_\ell(u - \Pi u)\Big| \le |\ell|^{1/2}\|\partial_\ell(u - \Pi u)\|_{0,\ell}$$
$$\le |\ell| \left(\frac{2}{|T_1|}\right)^{1/2} \{|u - \Pi u|_{1,T_1} + h_{T_1}|u|_{2,T_1}\}$$

and writing $|T_1| = |\ell||d_1|\sin\alpha/2$, it follows from (4.2) and H2 that

$$|\ell| \left(\frac{2}{|T_1|}\right)^{1/2} \le 2(N_2N_3(1 + N_1))^{1/2} \left(\frac{|\ell|}{h}\right)^{1/2}$$

and the proof concludes by observing that $h_{T_1} \le h$. $\square$

The following lemma gives an estimate for the error $|u - \Pi u|_{1,K}$. The argument is essentially the same used for the estimate on triangles (see for example [10]) although we work directly on the element instead of on a reference one (note however, that a reference triangle is used to prove the trace theorem given in Lemma 4.1).

Lemma 4.3 *If $K = K(a, b, p, q)$ is convex then,*

$$|u - \Pi u|_{1,K} \le \left(1 + \frac{2}{\pi}\right)\left(\frac{2|K|}{|T_2|}\right)^{1/2} h|u|_{2,K} \tag{4.5}$$

and consequently, if K satisfies (4.1) then,

$$|u - \Pi u|_{1,K} \le 2\Big(1 + \frac{2}{\pi}\Big)N_4h|u|_{2,K}. \tag{4.6}$$

Proof. Consider for example $v = \frac{\partial}{\partial x}(u - \Pi u)$ and we want to show that

$$\|v\|_{0,K} \leq Ch|v|_{1,K}.$$

Let v_K be the mean value of v on K. Then, since K is convex, it is known that [12]

$$\|v - v_K\|_{0,K} \leq \frac{h}{\pi}|v|_{1,K} \tag{4.7}$$

therefore, it remains to bound $\|v_K\|_{0,K}$. Using Lemma 4.1, the Schwartz inequality and the fact that the integral between 0 and a of $v(x,0)$ vanishes, we obtain

$$\begin{aligned}\|v_K\|_{0,K} = |v_K||K|^{1/2} &= \frac{|K|^{1/2}}{a}\Big|\int_0^a (v - v_K)(x,0)dx\Big| \\ &\leq \left(\frac{2|K|}{|T_2|}\right)^{1/2}\{\|v - v_K\|_{0,K} + h|v|_{1,K}\}\end{aligned}$$

and the bound for $\frac{\partial}{\partial x}(u - \Pi u)$ is obtained by using again (4.7). In an analogous way we can estimate the derivative with respect to y and so (4.5) holds.

Now, it is easy to see that (4.1) yields $(2|K|/|T_2|)^{1/2} \leq 2N_4$ and consequently (4.6) follows immediately from (4.5). □

Our next goal is to obtain a sharp estimate for $|\phi_3|_{1,K}$ for the reference configurations $K = K(a,b,p,q)$. In order to do that we need to analyze the jacobian of the $\mathcal{Q}_1$ transformation $F_K : [0,1]^2 \to K$ defined as

$$F_K(\hat{x},\hat{y}) = (a\hat{x}(1-\hat{y}) + p\hat{x}\hat{y}\,, b\hat{y}(1-\hat{x}) + q\hat{x}\hat{y}) = (x,y). \tag{4.8}$$

Then,

$$DF_K(\hat{x},\hat{y}) = \begin{pmatrix} a + \hat{y}(p-a) & \hat{x}(p-a) \\ \hat{y}(q-b) & b + \hat{x}(q-b) \end{pmatrix}$$

and,

$$J_K := \det DF_K(\hat{x},\hat{y}) = ab(1 + \hat{x}(q/b - 1) + \hat{y}(p/a - 1)). \tag{4.9}$$

Observe first that, if K is convex then, $J_K > 0$. Indeed, since J_K is an affine function it is enough to verify that it is positive at the vertices of the reference square. The positivity at M_1, M_2 and M_4 is trivial and, since K is convex, (p,q) is over the segment joining M_2 and M_4 and so, it follows that

$$J_K(1,1) = q/b + p/a - 1 > 0. \tag{4.10}$$

Now, since $\frac{\partial\hat{\phi}_3}{\partial\hat{x}} = \hat{y}$ and $\frac{\partial\hat{\phi}_3}{\partial\hat{y}} = \hat{x}$, a simple computation yields

$$\left(\frac{\partial\phi_3}{\partial x} \circ F_K\right)(\hat{x},\hat{y}) = b\hat{y}/J_K(\hat{x},\hat{y})$$

and,

$$\left(\frac{\partial\phi_3}{\partial y} \circ F_K\right)(\hat{x},\hat{y}) = a\hat{x}/J_K(\hat{x},\hat{y}).$$

Therefore, by (4.9) and a change of variables we get

$$\left\|\frac{\partial\phi_3}{\partial x}\right\|_{0,K}^2 \le \frac{b}{a}\int_0^1\int_0^1 \frac{1}{1+\hat{x}(q/b-1)+\hat{y}(p/a-1)}\,d\hat{x}\,d\hat{y} \tag{4.11}$$

and analogously,

$$\left\|\frac{\partial\phi_3}{\partial y}\right\|_{0,K}^2 \le \frac{a}{b}\int_0^1\int_0^1 \frac{1}{1+\hat{x}(q/b-1)+\hat{y}(p/a-1)}\,d\hat{x}\,d\hat{y} \tag{4.12}$$

and so, we need to bound the integral

$$I(a,b,p,q) := \int_0^1\int_0^1 \frac{1}{1+\hat{x}(q/b-1)+\hat{y}(p/a-1)}\,d\hat{x}\,d\hat{y} \tag{4.13}$$

for positive a, b, p and q satisfying (4.10). This is the object of the following lemma.

Lemma 4.4 *Let $a,b,p,q>0$ such that $q/b+p/a-1>0$ then, the integral $I=I(a,b,p,q)$ defined in (4.13) satisfies*

1. *If $q/b\le 1$ and $p/a\le 1$ then, $I\le \frac{1}{q/b+p/a-1}$.*
2. *If $q/b\le 1$ and $p/a>1$ then, $I\le \min\{\frac{(p/a)log\,(p/a)+1}{(1-q/b)(p/a-1)},\frac{b}{q}\}$.*
3. *If $q/b>1$ and $p/a\le 1$ then, $I\le \min\{\frac{(q/b)log\,(q/b)+1}{(1-p/a)(q/b-1)},\frac{a}{p}\}$.*
4. *If $q/b>1$ and $p/a>1$* then, *$I\le 1$.*

where in 2) it is understood that the minimum is b/q when $q/b=1$ and analogously in 3).

Proof. *1)* Since $0\le\hat{x},\hat{y}\le 1$, it follows that

$$\frac{1}{1+\hat{x}(q/b-1)+\hat{y}(p/a-1)} \le \frac{1}{q/b+p/a-1}$$

and so, *1)* holds.

2) Integrating in the $\hat{x}$ variable we have

$$I = \frac{1}{1-q/b}\int_0^1\{\log(1+\hat{y}(p/a-1))-\log(q/b+\hat{y}(p/a-1))\}\,d\hat{y}$$
$$= \frac{1}{(1-q/b)(p/a-1)}\Big\{\int_1^{p/a}\log z\,dz - \int_{q/b}^{q/b+p/a-1}\log z\,dz\Big\} \le \frac{(p/a)\log(p/a)+1}{(1-q/b)(p/a-1)}$$

where we have used that $\int_{q/b}^{q/b+p/a-1}(-\log z)\,dz \le \int_0^1(-\log z)\,dz\le 1$ because $q/b\le 1$.

On the other hand, using $p/a>1$, $q/b\le 1$ and the fact that $t\log t/(t-1)\le 1$ for $0\le t\le 1$ it follows that

$$I\le\int_0^1\int_0^1\frac{1}{1+\hat{x}(q/b-1)}\,d\hat{x}\,d\hat{y} = \frac{\log(q/b)}{q/b-1}\le\frac{b}{q}$$

which concludes the proof of *2)*.

3) It follows like part *2)* interchanging the roles of a and p with b and q respectively.

4) Under these hypotheses it follows immediately that $1 + \hat{x}(q/b - 1) + \hat{y}(p/a - 1) \geq 1$ and so *4)* holds. □

The following lemma will allow us to bound $|\phi_3|_{1,K}$ in terms of $\overline{\psi}$ and N_i, $i = 1, 2, 3, 4$ for $K(a, b, p, q)$ satisfying the hypotheses H1, H2, H3, H4. To simplify notation we introduce $N_5 = \max\{N_2, 1/\sin((\pi + \overline{\psi})/2)\}$.

Lemma 4.5 *If* $K = K(a, b, p, q)$ *is convex and satisfies the hypotheses* H1, H2, H3, H4 *then,*

1. $\max\{|\ell|/a, |\ell|/b\} \leq N_2(q/b + p/a - 1)$.
2. *If* $q/b \leq 1$ *then* $a/b \leq N_2$.
3. *If* $q/b \leq 1/2$ *and* $p/a > 1$ *then* $b/a \leq 2N_3$.
4. *If* $q/b > 1$ *and* $a/b \leq tg((\pi - \overline{\psi})/2)$ *then* $|\ell| \leq N_5 p$.
5. *If* $q/b > 1$ *and* $b/a \leq tg(\alpha/2)$ *then* $|\ell| \leq 2N_2(q - b)$.
6. $\min\{1/|p - a|, 1/q\} \leq \sqrt{2}N_3/|\ell|$.
7. $\min\{1/|q - b|, 1/p\} \leq \sqrt{2}/|\ell|$.

Proof. *1)* Calling $y(x) = -(b/a)(x - a)$ the equation of the straight line passing through M_2 and M_4 and y^{-1} its inverse, we have

$$\frac{q - y(p)}{b} = \frac{p - y^{-1}(q)}{a} = \frac{q}{b} + \frac{p}{a} - 1 \tag{4.14}$$

Now, an elementary geometrical analysis yields $|\ell| \sin\alpha \leq q - y(p)$ and $|\ell| \sin\alpha \leq p - y^{-1}(q)$ and so, *1)* follows from (4.14) and H2.

2) Calling β the angle between d_1 and the segment joining M_4 and (a, b) we have that, since $q \leq b$ then, $\alpha \leq \beta$. So, $b/a = \mathrm{tg}\beta \geq \mathrm{tg}\alpha$ and therefore, using again H2 we obtain *2)*.

3) Under these hypotheses we have that $|\ell| \geq b - q \geq b/2$. Then $b/a \leq 2|\ell|/a$ and therefore, *3)* follows from H3.

4) Calling now γ the angle between d_1 and the segment joining M_1 and M_4 it is easy to see that, under these hypotheses, $\gamma \leq (\pi - \overline{\psi})/2$. Now, using H4 we have

$$\alpha \leq \alpha + \gamma \leq (\pi + \overline{\psi})/2$$

and, since $|\ell|/p = 1/\sin(\alpha + \gamma)$, *4)* follows from H2 and the fact that $\overline{\psi} < \pi$.

5) With β as in *2)* we have in this case, that $\beta \leq \alpha$. Then, $\mathrm{tg}\beta = b/a \leq \mathrm{tg}(\alpha/2)$ and so, $\beta \leq \alpha/2$. Therefore, $\alpha - \beta \geq \alpha/2$ and *5)* follows by observing that

$$q - b = |\ell| \sin(\alpha - \beta) \geq |\ell| \sin(\alpha/2) \geq |\ell|(\sin\alpha)/2$$

6) Let us call η the interior angle of K at the vertex M_2 and ℓ_{23} the side joining M_2 and M_3. Then, we have

$$|p - a|/|\ell_{23}| = \cos\eta \qquad \text{and,} \qquad q/|\ell_{23}| = \sin\eta$$

and so, *6)* follows from H3.

7) Follows by an analysis similar to that given in *6)*. □

Now, we are ready to obtain the bound for $|\phi_3|_{1,K}$ in terms of geometric properties of the reference configuration. This is done in the next lemma where we use the notation of Figure 5.

Lemma 4.6 *If* $K = K(a,b,p,q)$ *is convex and satisfies* H1, H2, H3, H4 *then, there exists a constant* C *depending only on* $\overline{\psi}$ *and* N_i, $i = 1,2,3$ *such that*

$$|\phi_3|_{1,K} \leq C\Big(\frac{h}{|\ell|}\Big)^{1/2}. \tag{4.15}$$

Proof. We consider four cases as in Lemma 4.4.
1) If $q/b \leq 1$ and $p/a \leq 1$ then, from (4.11), part *1)* of Lemma 4.4 and part *1)* of Lemma 4.5 it follows that

$$\Big\|\frac{\partial\phi_3}{\partial x}\Big\|^2_{0,K} \leq \frac{b}{a}\frac{1}{(q/b + p/a - 1)} \leq N_2\frac{b}{|\ell|} \leq N_2\frac{h}{|\ell|}.$$

Analogously, using now (4.12), we obtain

$$\Big\|\frac{\partial\phi_3}{\partial y}\Big\|^2_{0,K} \leq N_2\frac{h}{|\ell|}.$$

2) Assume now that $q/b \leq 1$ and $p/a > 1$. Using again (4.11) and (4.12) but combined now with part *2)* of Lemma 4.4 and (4.1) we obtain

$$\Big\|\frac{\partial\phi_3}{\partial x}\Big\|^2_{0,K} \leq \frac{b}{a}\min\Big\{\frac{N_4\log N_4 + 1}{(1-q/b)(p/a-1)}, \frac{b}{q}\Big\} \tag{4.16}$$

and,

$$\Big\|\frac{\partial\phi_3}{\partial y}\Big\|^2_{0,K} \leq \frac{a}{b}\min\Big\{\frac{N_4\log N_4 + 1}{(1-q/b)(p/a-1)}, \frac{b}{q}\Big\}. \tag{4.17}$$

Now, if $b/q \leq 2$ the proof concludes by using (4.16), (4.17) and H3. Otherwise, $q/b < 1/2$ and so $1/(1-q/b) < 2$. On the other hand, from part *3)* of Lemma 4.5 we have $b/a \leq 2N_3$ and therefore, using again (4.16) combined with part *6)* of Lemma 4.5 we obtain

$$\Big\|\frac{\partial\phi_3}{\partial x}\Big\|^2_{0,K} \leq 2N_3\min\Big\{\frac{2(N_4\log N_4 + 1)}{p/a - 1}, \frac{b}{q}\Big\}$$

$$\leq 4N_3(N_4\log N_4 + 1)h\min\Big\{\frac{1}{p-a}, \frac{1}{q}\Big\} \leq 4\sqrt{2}N_3^2(N_4\log N_4 + 1)\frac{h}{|\ell|}.$$

The bound for $\|\frac{\partial\phi_3}{\partial y}\|^2_{0,K}$ follows in a similar way from (4.17) and part *2)* of Lemma 4.5.

3) Consider now the case $q/b > 1$ and $p/a \leq 1$. Once again we use (4.11) and (4.12) combined now with part *3)* of Lemma 4.4 to obtain

$$\Big\|\frac{\partial\phi_3}{\partial x}\Big\|^2_{0,K} \leq \frac{b}{a}\min\Big\{\frac{N_4\log N_4 + 1}{(q/b-1)(1-p/a)}, \frac{a}{p}\Big\} \tag{4.18}$$

and,

$$\Big\|\frac{\partial\phi_3}{\partial y}\Big\|^2_{0,K} \le \frac{a}{b}\min\Big\{\frac{N_4\log N_4+1}{(q/b-1)(1-p/a)},\frac{a}{p}\Big\}. \tag{4.19}$$

However, we can not proceed exactly as in the previous case because now we do not have, as before, that a/b is bounded by above and below.

Assume first that $b/a < 1/\mathrm{tg}((\pi-\overline{\psi})/2)$. Then, as in part *2)*, we can assume that $p/a \le 1/2$ (otherwise, the estimate follows easily from (4.18) and (4.19)). In this case we have

$$\begin{aligned}\Big\|\frac{\partial\phi_3}{\partial x}\Big\|^2_{0,K} &\le \frac{2}{\mathrm{tg}((\pi-\overline{\psi})/2)}(N_4\log N_4+1)\min\Big\{\frac{b}{q-b},\frac{a}{p}\Big\}\\ &\le \frac{2}{\mathrm{tg}((\pi-\overline{\psi})/2)}(N_4\log N_4+1)h\min\Big\{\frac{1}{q-b},\frac{1}{p}\Big\}\end{aligned}$$

and so, the result follows from part *7)* of Lemma 4.5.

On the other hand, if $b/a \ge 1/\mathrm{tg}((\pi-\overline{\psi})/2)$ we can use part *4)* of Lemma 4.5 and (4.18) to get

$$\Big\|\frac{\partial\phi_3}{\partial x}\Big\|^2_{0,K} \le \frac{b}{p} \le N_5\frac{b}{|\ell|} \le N_5\frac{h}{|\ell|}.$$

Now, in order to bound the derivative with respect to y we consider again two cases: $a/b < 1/\mathrm{tg}(\alpha/2)$ and $a/b \ge 1/\mathrm{tg}(\alpha/2)$. In the first case we bound $\|\frac{\partial\phi_3}{\partial y}\|^2_{0,K}$ proceeding as before by using part *7)* of Lemma 4.5.

In the second case, we use (4.19) to obtain

$$\Big\|\frac{\partial\phi_3}{\partial y}\Big\|^2_{0,K} \le 2\frac{a}{b}(N_4\log N_4+1)\frac{b}{q-b}$$

and the proof concludes by using part *5)* of Lemma 4.5.

4) Finally, in the case $q/b > 1$ and $p/a > 1$ the result follows trivially from part *4)* of Lemma 4.4 and H3 by using again (4.11) and (4.12). □

Collecting all the lemmas we obtain our main theorem which gives the optimal error estimate for convex quadrilaterals.

Theorem 4.1 *Let K be a convex quadrilateral with diameter h. There exists a constant C_0 independent of K such that*

$$\|u-Qu\|_{0,K} \le C_0h^2|u|_{2,K} \tag{4.20}$$

and, if K satisfies $RDP(N,\psi)$ then, there exists a constant $C = C(N,\psi)$ such that

$$|u-Qu|_{1,K} \le Ch|u|_{2,K}. \tag{4.21}$$

Proof. It is always possible to choose two adjacent sides of K, ℓ_1 and ℓ_2, such that K is contained in the parallelogram defined by these two sides. Observe that this parallelogram has a diameter of the same order than that of K. Now, let L be the affine transformation taking ℓ_1 into the segment joining $(0,0)$ and $(1,0)$ and ℓ_2 into the segment joining $(0,0)$ and $(0,1)$ and call $\tilde{K} = L(K)$ (this transformation was used also in [15]). It is easy to see that $\tilde{K}$ has outer and inner diameters of order one and so, it satisfies (1.2) with a constant σ independent of K. Therefore, the optimal error estimate in L^2 on $\tilde{K}$ follows from [9] and (4.20) is obtained by changing variables.

On the other hand, from Lemma 3.3 we know that K can be transformed into a convex $K(a,b,p,q)$ satisfying H1, H2, H3, H4 with constants $\overline{\psi}$ and N_i, $i = 1,2,3$ depending only on N and ψ and with a diameter h equivalent to that of K. Moreover, from Lemma 3.4 we have that the error estimate on K follows from that in $K(a,b,p,q)$. Therefore, it is enough to prove the error estimate for these reference configurations with a constant depending only on $\overline{\psi}$ and N_i.

Now, the inequality (4.21) follows from (4.3) combined with (4.6), (4.4) and (4.15). $\square$

5 Conclusions

For $\mathcal{Q}_1$ isoparametric elements we have found geometrical hypotheses under which the optimal order error estimate in the H_1 norm holds uniformly on a family of elements. These hypotheses are simple to verify and weaker than previously used ones.

An interesting question is whether our sufficient conditions are also necessary in the sense that, given a family of elements which do not satisfy the regular decomposition property with uniform N and ψ then, the error can not be uniformly bounded. We believe that this is true but further research is needed in order to give an answer to this question.

References

[1] Acosta, G., Durán, R. G. (1999): The maximum angle condition for mixed and nonconforming elements: Application to the Stokes equations, SIAM J. Numer. Anal. **37**, 18-36.

[2] Apel, T. (1998): Anisotropic interpolation error estimates for isoparametric quadrilateral finite elements, Computing **60**, 157-174.

[3] Apel, T. (1999): Anisotropic finite elements: Local estimates and applications. Advances in Numerical Mathematics, B. G. Teubner, Stuttgart, Leipzig.

[4] Apel, T., Dobrowolski, M. (1992): Anisotropic interpolation with applications to the finite element method, Computing **47**, 277-293.

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

[6] Ciarlet, P. G., Raviart, P. A. (1972): Interpolation theory over curved elements, with applications to finite elements methods, Comp. Meth. Appl. Mech. Eng.**1**, 217-249.

[7] Durán, R. G. (1999): Error estimates for narrow 3-d finite elements, Math. Comp. **68**, 187-199.

[8] Jamet, P. (1976): Estimations d'erreur pour des éléments finis droits presque dégénérés, RAIRO Anal. Numér. **10**,46-61.

[9] Jamet, P. (1977): Estimation of the interpolation error for quadrilateral finite elements which can degenerate into triangles, SIAM J. Numer. Anal. **14**, 925-930.

[10] Krízek, M. (1991): On semiregular families of triangulations and linear interpolation, Appl. Math., **36**, 223-232.

[11] Krízek, M. (1992): On the maximum angle condition for linear tetrahedral elements, SIAM J. Numer. Anal. **29**, 513-520.

[12] Payne, L. E., Weinberger, H. F. (1960): An optimal Poincaré inequality for convex domains, Arch. Rat. Mech. Anal. **5**, 286-292.

[13] Shenk, N. Al (1994): Uniform error estimates for certain narrow Lagrange finite elements, Math. Comp. **63**, 105-119.

[14] Verfürth, R. (1997): Error estimates for some quasi-interpolation operator, Preprint 227, Fakultät für Mathematik der Ruhr-Universität Bochum.

[15] Zenisek, A., Vanmaele, M. (1995): The interpolation theorem for narrow quadrilateral isoparametric finite elements, Numer. Math. **72**, 123-141.

[16] Zenisek, A., Vanmaele, M. (1995): Applicability of the Bramble Hilbert lemma in interpolation problems of narrow quadrilateral isoparametric finite elements, J. Comp. Appl. Math. **63**, 109-122.