Numer. Math. (2002) 91: 1–12
Digital Object Identifier (DOI) 10.1007/s002110100321

Numerische
Mathematik

Small data oscillation implies the saturation assumption

Willy Dörfler[1], **Ricardo H. Nochetto**[2]

[1] FB Mathematik, Universität Kaiserslautern, Postfach 3049, 67653 Kaiserslautern, Germany; e-mail: doerfler@mathematik.uni-kl.de
[2] Department of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA; e-mail: rhn@math.umd.edu

Received November 17, 2000 / Published online July 25, 2001 –

Summary. The saturation assumption asserts that the best approximation error in H_0^1 with piecewise quadratic finite elements is strictly smaller than that of piecewise linear finite elements. We establish a link between this assumption and the oscillation of $f = -\Delta u$, and prove that small oscillation relative to the best error with piecewise linears implies the saturation assumption. We also show that this condition is necessary, and asymptotically valid provided $f \in L^2$.

Mathematics Subject Classification (1991): 65N15, 65N30, 65N50

1. Introduction

The *saturation assumption* is widely used in a posteriori error analysis of finite element methods [1, Ch. 5], [2], [3]. It asserts, in its simplest form, that the best approximation error $\|\nabla(u-u_2)\|_\Omega$ of a function $u \in H_0^1(\Omega)$ with quadratic finite elements is strictly smaller than that with linear finite elements $\|\nabla(u-u_1)\|_\Omega$, namely,

$$\|\nabla(u-u_2)\|_\Omega \le \alpha \, \|\nabla(u-u_1)\|_\Omega \tag{1.1}$$

for a suitable constant $\alpha \in (0,1)$. Throughout this paper we assume that Ω is a bounded polyhedral domain in $\mathbb{R}^d$, which coincides with its finite element decomposition; we also use the notation $\|v\|_\omega^2 := \int_\omega |v|^2$ for $\omega \subset \Omega$. If $\mathfrak{T}$ denotes a *graded* shape-regular partition of Ω of size $h_\mathfrak{T}$, and $\mathfrak{U}_\mathfrak{T}^1, \mathfrak{U}_\mathfrak{T}^2$

Partially supported by NSF Grant DMS-9971450.

Correspondence to: W. Dörfler

stand for the finite element subspaces of $H^1_0(\Omega)$ consisting of piecewise linear and piecewise quadratic functions, respectively, then u_1, u_2 are the Ritz projections of u onto such spaces, that is

$$u_k \in \mathfrak{U}^k_{\mathfrak{T}}: \quad \int_\Omega \nabla(u-u_k)\cdot\nabla\phi = 0, \quad \forall\phi\in\mathfrak{U}^k_{\mathfrak{T}}. \tag{1.2}$$

The assumption (1.1) becomes true, as $h_{\mathfrak{T}} \to 0$, for functions $u \in W^s_p(\Omega)$ with $s-2 > d\max(1/p-1/2,0)$ in dimension d as a consequence of standard interpolation theory, provided a non-degeneracy condition holds (see Sect. 4). However, since $W^s_p(\Omega) \subset H^{2+\varepsilon}(\Omega)$ with $\varepsilon > 0$, such a regularity is never present for elliptic problems with singularities for which adaptive mesh refinement is required. Even though it is generally believed to be valid asymptotically, (1.1) is not known to hold under any reasonable assumptions on the underlying function u which still allow for singularities. If $\mathfrak{U}^2_{\mathfrak{T}}$ is enriched with cubic bubbles, and $f = -\Delta u$ is piecewise constant over $\mathfrak{T}$, then (1.1) is shown in [3] as a by-product of the equivalence between residual and hierarchical error estimators. However, the proof of [3] is indirect and does not shed light either on the size of α nor on existence of α under more realistic conditions on f. Moreover, a simple counting argument reveals that (1.1) cannot be valid in general on any refinement level [3, Proposition 2.2], but does not explain what makes (1.1) fail.

In this paper we discuss the validity of (1.1) and disclose a close relation with data oscillation. Since (1.1) is about approximability in the $H^1_0(\Omega)$-norm, it is natural to consider $f = -\Delta u$ as datum in this discussion. We further assume

$$f \in L_2(\Omega). \tag{1.3}$$

For each interior node x_i of $\mathfrak{T}$, we have a canonical basis function $\phi_i \in \mathfrak{U}^1_{\mathfrak{T}}$ and corresponding star $\omega_i := \operatorname{supp}(\phi_i)$. We denote by $f_i := |\omega_i|^{-1}\int_{\omega_i} f$ the mean value of f in ω_i, and define *data oscillation* to be the quantity

$$\operatorname{osc}(f,\mathfrak{T}) := \Big(\sum_i \|h(f-f_i)\|^2_{\omega_i}\Big)^{1/2}. \tag{1.4}$$

We prove in Sect. 3 the following sufficient condition for the validity of (1.1):

Theorem 1.1. *There exists a constant* $0 < \mu < 1$ *solely depending on shape regularity of* $\mathfrak{T}$*, but independent of* u *and* f*, such that if*

$$\operatorname{osc}(f,\mathfrak{T}) \le \mu\,\|\nabla(u-u_1)\|_\Omega \tag{1.5}$$

holds, then (1.1) *is valid with* $\alpha := (1-\mu^2)^{1/2}$.

For simplicity, we prove Theorem 1.1 in two dimensions, but comment in Remark 3.5 about the minor modifications for higher dimensions. We note that (1.5) is asymptotically valid as $h_{\mathfrak{T}} \downarrow 0$, and we explore this matter in detail in Sect. 4. Alternatively, (1.5) can be replaced by the more practical condition (3.9) which does not involve u directly (see Remark 3.4). In Sect. 2 we exhibit an elementary example showing that (1.5) is a necessary condition for (1.1). This also demonstrates that (1.1) may not in general be valid in the preasymptotic regime, when the oscillations of f are not yet well resolved by $\mathfrak{T}$.

2. Counterexamples

The purpose of this section is twofold. We give an explicit example violating (1.1) and the same time argue about its connection with data oscillation. Let the domain be the square $\Omega = (0,1)^2$, the partition $\mathfrak{T}$ have one interior node, and the forcing function f be *piecewise constant* (see Fig. 1). By symmetry, it easily follows that

$$\text{(2.1)} \qquad \int_{\Omega} f\phi = 0 \quad \forall \phi \in \mathfrak{U}^1_{\mathfrak{T}}, \mathfrak{U}^2_{\mathfrak{T}} \qquad \Longrightarrow \qquad u_1 = u_2 = 0,$$

which violates (1.1). We realize that the oscillations of f at the star level, in the sense of (1.4), are responsible for this outcome. Since the counterexample is rather elementary, we conclude that we could not expect in general the saturation assumption to be valid in the *preasymptotic regime*, that is whenever data oscillation has not yet been resolved by the mesh $\mathfrak{T}$. But we might still expect that once $\mathfrak{T}$ becomes fine enough to detect the structure of f, then (1.1) should hold.

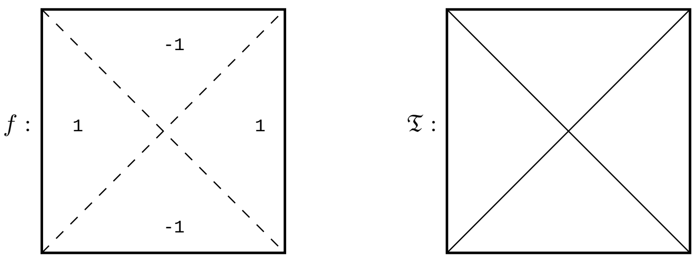

Fig. 1. Function f and mesh $\mathfrak{T}$ of counterexample 1

In Sect. 3 we establish this conjecture upon quantifying the size of data oscillation. In Sect. 4 we prove that (1.1) is always valid in the *asymptotic regime* provided $f \in L_2(\Omega)$ and u satisfies a non-degeneracy condition.

The notion of data oscillation (1.4) is not completely local since the stars ω_i overlap slightly. The counterexample of Fig. 1 shows that it is *impossible* to reduce this notion to the element level since the element oscillation $f - f_T$ of f is zero; here $f_T := |T|^{-1} \int_T f$ stands for the mean value of f in the element $T \in \mathfrak{T}$. One may wonder whether enriching the space $\mathfrak{U}^2_{\mathfrak{T}}$ with cubic bubbles, as in [3], and redefining

$$\operatorname{osc}_2(f, \mathfrak{T}) := \Big(\sum_{T \in \mathfrak{T}} \|h(f - f_T)\|_T^2 \Big)^{1/2}, \tag{2.2}$$

might lead to a statement similar to Theorem 1.1. To explore this idea consider the counterexample of Fig. 2 with enriched space $\mathfrak{V}^2_{\mathfrak{T}}$, for which (2.1) still holds due to the symmetry of f. We realize again the link between size of (2.2) and (1.1), which is further investigated in Remark 3.6.

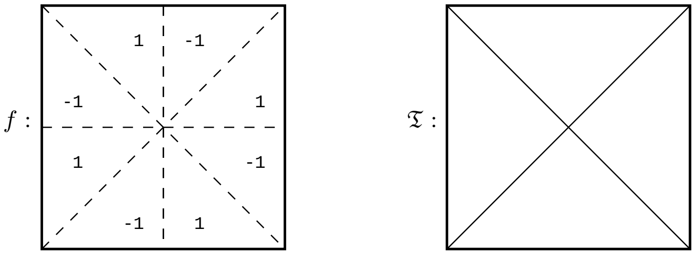

Fig. 2. Function f and mesh $\mathfrak{T}$ of counterexample 2

3. Validity of the saturation assumption

In this section we prove the main result of this paper, namely that *small data oscillation implies the saturation assumption*. For simplicity, the result is derived in two space dimensions but it is valid in any dimension.

Even though the saturation assumption is a basic issue in approximation theory, the technique used here comes from a posteriori error analysis. In fact, it consists of exploiting orthogonality to relate the errors $\|\nabla(u - u_2)\|_\Omega$ and $\|\nabla(u - u_1)\|_\Omega$ via $\|\nabla(u_2 - u_1)\|_\Omega$, and then showing a lower bound for $\|\nabla(u_2 - u_1)\|_\Omega$ in terms of $\|\nabla(u - u_1)\|_\Omega$. This entails deriving precise expressions for interior and jump residuals on *stars*. We split the argument into several steps.

3.1. Orthogonality

Since u_2 is the orthogonal projection of u onto $\mathfrak{U}^2_{\mathfrak{T}}$ with the scalar product of $H^1_0(\Omega)$, we deduce that $\int_\Omega \nabla(u - u_2) \cdot \nabla v = 0$ for all $v \in \mathfrak{U}^2_{\mathfrak{T}}$. Therefore

$u_2 - u_1 \in \mathfrak{U}^2_{\mathfrak{T}}$ is perpendicular to $u - u_2$, and $u - u_1 = (u - u_2) + (u_2 - u_1)$ satisfies the orthogonality (Pythagoras) relation

$$\|\nabla(u-u_1)\|^2_{\Omega} = \|\nabla(u-u_2)\|^2_{\Omega} + \|\nabla(u_2-u_1)\|^2_{\Omega}. \tag{3.1}$$

We conclude that to prove the saturation assumption (1.1), it suffices to establish a lower bound of $\|\nabla(u_2-u_1)\|_{\Omega}$ in terms of $\|\nabla(u-u_1)\|_{\Omega}$. This is possible at the expense of an additional term involving the oscillation of f, and is shown below. But before we prove an elementary upper bound of $\|\nabla(u-u_1)\|_{\Omega}$ in terms of residual estimators.

3.2. *Upper a posteriori bound*

We intend to express the usual upper a posteriori bound with the interior residual accumulated by stars instead of by elements. The following estimate is well-known [1, Ch. 2]

$$\|\nabla(u-u_1)\|^2_{\Omega} \le C_1\Big(\sum_{S\in\mathfrak{S}} \|h^{1/2}J\|^2_S + \sum_{T\in\mathfrak{T}} \|hf\|^2_T\Big).$$

Here, $\mathfrak{S}$ denotes the set of all interior sides. We recall that ω_i denotes an interior star and f_i indicates the mean value of f in ω_i. Since any element of $\mathfrak{T}$ belongs at most to 3 stars in $\mathbb{R}^2$, we can replace the interior residual by

$$\sum_{T\in\mathfrak{T}} \|hf\|^2_T \le C\sum_i \|hf\|^2_{\omega_i} \le C\Big(\sum_i \|hf_i\|^2_{\omega_i} + \sum_i \|h(f-f_i)\|^2_{\omega_i}\Big).$$

The above two estimates together give rise to the modified upper bound

$$\begin{aligned}\|\nabla(u-u_1)\|^2_{\Omega} &\le \\ &C_2\Big(\sum_{S\in\mathfrak{S}} \|h^{1/2}J\|^2_S + \sum_i \|hf_i\|^2_{\omega_i} + \sum_i \|h(f-f_i)\|^2_{\omega_i}\Big).\end{aligned} \tag{3.2}$$

We conclude that to prove (1.1) we need to bound the jump and interior residuals, the first two terms on the right hand side of (3.2), by $\|\nabla(u_2-u_1)\|_{\Omega}$.

3.3. *Local estimate of jump residuals*

For any side S let ω_S be the union of the elements that meet at S and let $\phi_S \in \mathfrak{U}^2_{\mathfrak{T}}$ be the canonical functions with degree of freedom at the midpoint of S. Using Simpson's rule, integration by parts, and the equation for u_2 (in this order), we obtain for any interior side S

$$\begin{aligned}\frac{2}{3}h_S J_S &= \int_S J_S\phi_S = -\int_{\omega_S} \nabla u_1\cdot\nabla\phi_S \\ &= \int_{\omega_S} \Big(\nabla(u_2-u_1)\cdot\nabla\phi_S - f_i\phi_S + (f_i - f)\phi_S\Big);\end{aligned} \tag{3.3}$$

here it is crucial that J_S is constant. Consequently, squaring and recalling that

$$\|\nabla\phi_S\|_{\omega_S} \le C, \qquad \|\phi_S\|_{\omega_S}, \|\phi_i\|_{\omega_i} \le Ch_i, \tag{3.4}$$

as well as $h_S\|J_S\|_S^2 = \|h^{1/2}J\|_S^2$, we end up with the local estimate

$$\|h^{1/2}J\|_S^2 \le C_3\Big(\|\nabla(u_2-u_1)\|_{\omega_S}^2 + \|hf_i\|_{\omega_S}^2 + \|h(f-f_i)\|_{\omega_S}^2\Big). \tag{3.5}$$

3.4. Local estimate of interior residuals

This is the key estimate which relates $\|hf_i\|_{\omega_i}$ with $\|\nabla(u_2-u_1)\|_{\omega_i}$. The proof entails deriving a sharp relation for the jump residuals in terms of piecewise linear and piecewise quadratic functions. To this end, it is essential to work on stars because on these sets there is a natural relation between interior and jump residuals for piecewise linear finite elements. A similar idea was used in [8] to remove the saturation assumption.

Let x_i be an interior node, $\phi_i \in \mathfrak{U}_{\mathfrak{T}}^1$ be the corresponding nodal basis function and ω_i be its star. Denote by $\mathfrak{S}_i$ the set of interior sides S in ω_i. Since ϕ_i is piecewise linear, the trapezoidal rule combined with the equation for u_1 implies

$$\begin{aligned}\frac{1}{2}\sum_{S\in\mathfrak{S}_i} h_S J_S &= \sum_{S\in\mathfrak{S}_i}\int_S J_S\phi_i = -\int_{\omega_i}\nabla u_1\cdot\nabla\phi_i \\ &= -\int_{\omega_i} f\phi_i = -\frac{1}{3}f_i|\omega_i| + \int_{\omega_i}(f_i-f)\phi_i.\end{aligned} \tag{3.6}$$

We now repeat this calculation for quadratics, namely $\phi_S \in \mathfrak{U}_{\mathfrak{T}}^2$. Adding (3.3) for all sides S in $\mathfrak{S}_i$, we obtain

$$\begin{aligned}\frac{2}{3}\sum_{S\in\mathfrak{S}_i} h_S J_S = -f_i &\sum_{S\in\mathfrak{S}_i}\int_{\omega_S}\phi_S \\ &+ \sum_{S\in\mathfrak{S}_i}\int_{\omega_S}\Big(\nabla(u_2-u_1)\cdot\nabla\phi_S + (f_i-f)\phi_S\Big).\end{aligned}$$

Since

$$\sum_{S\in\mathfrak{S}_i}\int_{\omega_S}\phi_S = \frac{1}{3}\sum_{S\in\mathfrak{S}_i}|\omega_S| = \frac{2}{3}|\omega_i|$$

we readily deduce

$$\frac{2}{3}\sum_{S\in\mathfrak{S}_i} h_S J_S = -\frac{2}{3} f_i|\omega_i| - \sum_{S\in\mathfrak{S}_i}\int_{\omega_S}\Big(\nabla(u_1-u_2)\cdot\nabla\phi_S + (f-f_i)\phi_S\Big). \tag{3.7}$$

Examining (3.6) and (3.7) reveals the main idea of the proof: the jump residual can be eliminated, thereby giving an expression for the interior residual in terms of $\nabla(u_2-u_1)$ and data oscillation. In fact, we get

$$f_i|\omega_i| = \frac{9}{2}\sum_{S\in\mathfrak{S}_i}\int_{\omega_S}\Big(\nabla(u_2-u_1)\cdot\nabla\phi_S + (f_i-f)\phi_S\Big) + 6\int_{\omega_i}(f-f_i)\phi_i.$$

Since each ω_S is only counted twice in the above sum, in light of (3.4) we obtain the crucial local upper bound

$$\|hf_i\|_{\omega_i} \le C|f_i||\omega_i| \le C_5\Big(\|\nabla(u_2-u_1)\|_{\omega_i} + \|h(f-f_i)\|_{\omega_i}\Big). \tag{3.8}$$

Proof of Theorem 1.1 We now collect the above estimates. We first note that combining (3.5) with (3.8) yields the local estimate

$$\sum_{S\in\mathfrak{S}_i}\|h^{1/2}J\|_S^2 + \|hf_i\|_{\omega_i}^2 \le C\Big(\|\nabla(u_2-u_1)\|_{\omega_i}^2 + \|h(f-f_i)\|_{\omega_i}^2\Big).$$

Inserting this bound into (3.2), and using the finite overlapping property of stars, gives

$$\|\nabla(u-u_1)\|_\Omega^2 \le C_6\Big(\|\nabla(u_2-u_1)\|_\Omega^2 + \operatorname{osc}(f,\mathfrak{T})^2\Big).$$

In view of (3.1) and the small oscillation assumption $\operatorname{osc}(f,\mathfrak{T}) \le \mu\|\nabla(u-u_1)\|_\Omega$, we thus end up with

$$\begin{aligned}\|\nabla(u-u_2)\|_\Omega^2 &\le (1-C_6^{-1})\,\|\nabla(u-u_1)\|_\Omega^2 + \operatorname{osc}(f,\mathfrak{T})^2\\ &\le (1-C_6^{-1}+\mu^2)\,\|\nabla(u-u_1)\|_\Omega^2.\end{aligned}$$

The asserted estimate (1.1) follows from choosing $\mu^2 = (2C_6)^{-1}$ and $\alpha^2 = 1-\mu^2$. Note that both μ and α are quantities which solely depend on shape regularity of $\mathfrak{T}$; the argument carries over even for highly refined meshes though.

Remark 3.1. It is worth stressing once more that the chief idea of the proof is to work on stars, which are viewed as basic cells for piecewise linear approximation. This gives rise to the link between interior and jump residual of (3.6), which would not be possible otherwise.

Remark 3.2. Suppose that quadratics over a mesh $\mathfrak{T}_H$ with mesh-size H are replaced by linears over a uniformly refined mesh $\mathfrak{T}_{H/2}$ obtained from $\mathfrak{T}_H$ via two bisections. The number of degrees of freedom and their location is the same for both spaces $\mathfrak{U}_H^2$ and $\mathfrak{U}_{H/2}^1$. The question thus arises whether or not $\mathfrak{U}_H^2$ could be replaced by $\mathfrak{U}_{H/2}^1$ in the above construction and argument. The example of Fig. 3, introduced in [7], shows that the answer is in general negative.

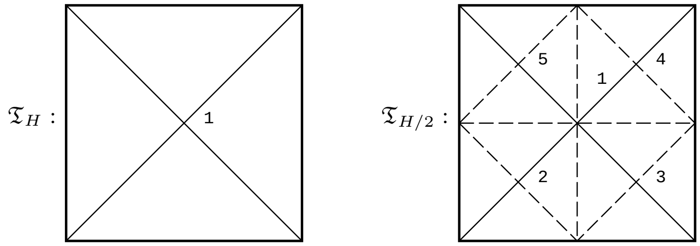

Fig. 3. Example with $f = 1$ and $u_H = u_{H/2}$.

Let $\mathfrak{T}_H$ and $\mathfrak{T}_{H/2}$ be the uniform meshes depicted in Fig. 3, let $\phi_1 \in \mathfrak{U}_H^1$ be the canonical basis function over $\mathfrak{T}_H$, and let $f = 1$. It is easy to see that

$$u_H = u_{H/2} = \phi_1/12,$$

whence

$$\|\nabla(u - u_H)\|_\Omega = \|\nabla(u - u_{H/2})\|_\Omega$$

which violates (1.1). We may thus wonder what goes wrong in the above argument which seems to extend to this situation as well. What happens is that in trying to eliminate the jump residuals from (3.6) and (3.7), the interior residuals also cancel out, thereby providing no useful information. *Quadratics do indeed encode finer information than refined linears.*

Remark 3.3. Suppose that $\mathfrak{T}_{H/2}$, obtained from $\mathfrak{T}_H$ by two bisections, is replaced by a *red refinement* $\mathfrak{T}'_{H/2}$ of all triangles around x_1 [9, Ch. 4] (see left picture in Fig. 4). Then the inequality (1.1) can be established for $u_1 = u_H$ and $u_2 = u'_{H/2}$ with the given technique. The same happens for three consecutive bisections $\mathfrak{T}''_{H/2}$ of $\mathfrak{T}_H$ because the basis functions of

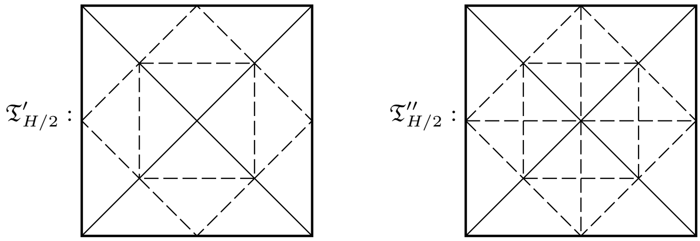

Fig. 4. Meshes produced by red refinement of $\mathfrak{T}_H$ and three bisections of $\mathfrak{T}_H$.

$\mathfrak{T}'_{H/2}$ are contained in the resulting (and richer) finite element space of $\mathfrak{T}''_{H/2}$ (see right picture in Fig. 4).

Remark 3.4. It is perhaps useful in the context of a posteriori error estimation, to express data oscillation fineness (1.5) in a computable fashion, that is without referring to the unknown function u. We thus propose the following computable alternative to (1.5):

$$\operatorname{osc}(f, \mathfrak{T}) \leq \hat{\mu}\, \|h^{1/2} J\|_{\Omega}. \tag{3.9}$$

To see that this implies (1.5), we resort to the lower bound for the error

$$\|h^{1/2} J\|_{\Omega} \leq \hat{C}_1\, \|\nabla(u - u_1)\|_{\Omega} + \hat{C}_2 \operatorname{osc}(f, \mathfrak{T}),$$

which simply arises from dropping the interior residual in the usual lower bound [9]. Hence

$$\operatorname{osc}(f, \mathfrak{T}) \leq \frac{\hat{\mu}\hat{C}_1}{(1 - \hat{\mu}\hat{C}_2)}\, \|\nabla(u - u_1)\|_{\Omega} = \mu\, \|\nabla(u - u_1)\|_{\Omega},$$

for a suitably small value of $\hat{\mu}$, still solely depending on mesh geometry.

Remark 3.5. To extend the proof of the main result to dimension $d > 2$, a few minor modifications are necessary: we must check how (3.6) and (3.7) change. Let ω_i be the star corresponding to an interior node x_i, $T \in \mathfrak{T}$ a d-simplex contained in ω_i, and S be a (closed) side of T containing the node x_i. Instead of (3.6), we now have

$$\frac{1}{d} \sum_{S \in \mathfrak{S}_i} |S| J_S = -\frac{1}{d+1} f_i |\omega_i| + \int_{\omega_i} (f_i - f)\phi_i. \tag{3.10}$$

On the other hand, to deal with quadratics we first recall the quadrature rule over T which uses the $d+1$ vertices v_i of T and the $d(d+1)/2$ midpoints of edges e_i of T as quadrature points:

$$\int_T \phi \approx \frac{|T|}{(d+1)(d+2)} \Big(\sum_{i=1}^{d+1} (2-d)\phi(v_i) + \sum_{i=1}^{d(d+1)/2} 4\phi(e_i) \Big).$$

This formula is exact for quadratics. Let ϕ_S be the quadratic function which is 1 at the $d-1$ midpoints of (closed) edges of S containing x_i, and vanishes at the vertices of T as well as the remaining midpoints of edges of T. We note that ϕ_S vanishes on $\partial\omega_i$ and thus $\phi_S \in \mathfrak{U}^2_{\mathfrak{T}}$. The quadrature rule gives rise to the following substitute for (3.7)

$$\frac{4(d-1)}{d(d+1)} \sum_{S\in\mathfrak{S}_i} |S| J_S = -\frac{8(d-1)}{(d+1)(d+2)} f_i|\omega_i| \tag{3.11}$$
$$+ \sum_{S\in\mathfrak{S}_i} \int_{\omega_S} \Big(\nabla(u_2-u_1)\cdot\nabla\phi_S + (f_i - f)\phi_S\Big).$$

Since we can still eliminate the jump residual between (3.10) and (3.11) and obtain a representation formula for the interior residual, for all $d>2$, then the proof continues as above.

Remark 3.6. Let $\mathfrak{V}^2_{\mathfrak{T}}$ denote the space of piecewise quadratic polynomials $\mathfrak{U}^2_{\mathfrak{T}}$ enriched with cubic bubbles $\{b_T\}_{T\in\mathfrak{T}}$; recall that $b_T = \lambda_1\lambda_2\lambda_3$ is the product of the three barycentric coordinates of $T\in\mathfrak{T}$. Let $v_2\in\mathfrak{V}^2_{\mathfrak{T}}$ be the finite element solution.

The presence of the additional bubble degree of freedom per element simplifies the above argument to a large extend. First, we observe that we have

$$f_T\frac{|T|}{5!} = \int_T f_T b_T = \int_T \nabla(v_2-u_1)\cdot\nabla b_T + \int_T (f_T - f) b_T, \tag{3.12}$$

because $\int_T \nabla u_1\cdot\nabla b_T = -\int_T \Delta u_1 b_T = 0$. This implies, instead of (3.8),

$$\|hf_T\|_T^2 \le C_5'\Big(\|\nabla(v_2-u_1)\|_T^2 + \|h(f-f_T)\|_T^2\Big). \tag{3.13}$$

On the other hand, if $\omega_S = T_1\cup T_2$, then (3.3) becomes

$$\frac{2}{3}h_S J_S = \int_{\omega_S}\nabla(v_2-u_1)\cdot\nabla\phi_S - \sum_{i=1}^{2}\int_{T_i}\Big(f_{T_i}\phi_S - (f_{T_i}-f)\phi_S\Big). \tag{3.14}$$

Since $\int_{T_i} f_{T_i}\phi_S = f_{T_i}\frac{|T_i|}{3}$, we can use (3.12) to replace the middle term in (3.14) in terms of $\nabla(v_2-u_1)$ plus data oscillation. Consequently, instead of (3.5), we obtain

$$\|h^{1/2}J\|_S^2 \le C_3'\Big(\|\nabla(v_2-u_1)\|_{\omega_S}^2 + \sum_{i=1}^{2}\|h(f-f_{T_i})\|_{T_i}^2\Big). \tag{3.15}$$

If $\operatorname{osc}_2(f,\mathfrak{T})$ is now defined as in (2.2), then combining (3.13) with (3.15) we deduce the fundamental estimate

$$\|\nabla(u-u_1)\|_\Omega^2 \le C_6'\Big(\|\nabla(v_2-u_1)\|_\Omega^2 + \operatorname{osc}_2(f,\mathfrak{T})^2\Big),$$

and the argument proceeds as in the proof of Theorem 1.1. We have thus derived the following more local version of Theorem 1.1: *there exists* $\mu<1$ *such that*

$$\begin{aligned}&\operatorname{osc}_2(f,\mathfrak{T}) \le \mu\,\|\nabla(u-u_1)\|_\Omega \quad \Rightarrow\\ &\|\nabla(u-v_2)\|_\Omega \le (1-\mu^2)^{1/2}\,\|\nabla(u-u_1)\|_\Omega.\end{aligned}$$

4. Asymptotics

We finally consider the generic situation in which both u and a sequence of meshes $\mathfrak{T}$ satisfy the *non-degeneracy* property: *there exists a constant* $A>0$ *independent of* $\mathfrak{T}$ *such that*

$$\|\nabla(u-u^1_{\mathfrak{T}})\|_\Omega \ge A h_{\mathfrak{T}}, \tag{4.1}$$

where $u^1_{\mathfrak{T}}$ is the Ritz projection onto $\mathfrak{U}^1_{\mathfrak{T}}$ and $h_{\mathfrak{T}}$ is the largest mesh-size of $\mathfrak{T}$. This is guaranteed, for instance, if $|D^2u(x)|\ge C>0$ for all x in a fixed region ω of Ω, where the local mesh-size is of order $h_{\mathfrak{T}}$, namely $h_T\ge Ch_{\mathfrak{T}}$; in particular, this is valid provided $\pm f(x)\ge C>0$ for all $x\in\omega$. Therefore, (4.1) is not a very restrictive condition in practice.

We first show that, as asserted in the introduction, the saturation assumption (1.1) is valid as $h_{\mathfrak{T}}\downarrow 0$ provided $u\in W_p^s(\Omega)$ with $s-2>d\max(1/p-1/2,0)$. Since

$$t := \min\big(s-1-d\max(1/p-1/2,0),\ 2\big) > 1,$$

standard approximation theory in Sobolev spaces [4, Theorem 16.2], together with (4.1), yields

$$\begin{aligned}\|\nabla(u-u_2)\|_{L_2(\Omega)} &\le C\|h^{\mathrm{t}}D^s u\|_{L_p(\Omega)} \le Ch_{\mathfrak{T}}^{\mathrm{t}}\\ &\le Ch_{\mathfrak{T}}^{t-1}\|\nabla(u-u_1)\|_{L_2(\Omega)}.\end{aligned}$$

We next prove the asymptotic validity of (1.5). To this end, we use a simple density argument for $f\in L_2(\Omega)$. Given $\varepsilon>0$, let ϕ be a smooth approximation of f satisfying $\|f-\phi\|_\Omega\le\varepsilon$. Since the mean value f_i satisfies $\|f_i\|_{\omega_i}\le\|f\|_{\omega_i}$ for all interior stars ω_i, we have

$$\begin{aligned}\|f-f_i\|_{\omega_i} &\le \|f-\phi\|_{\omega_i} + \|\phi-\phi_i\|_{\omega_i} + \|(\phi-f)_i\|_{\omega_i}\\ &\le 2\|f-\phi\|_{\omega_i} + Ch_i\|\nabla\phi\|_{\omega_i},\end{aligned}$$

whence, making use of the finite overlapping property of stars,

$$\operatorname{osc}(f,\mathfrak{T})^2 \le h_{\mathfrak{T}}^2 \sum_i \|f - f_i\|_{\omega_i}^2 \le C h_{\mathfrak{T}}^2(\varepsilon + h_{\mathfrak{T}}) = o(h_{\mathfrak{T}}^2). \tag{4.2}$$

Combining (4.1) with (4.2), we deduce that (1.5) is valid for all $f \in L_2(\Omega)$ provided the mesh-size $h_{\mathfrak{T}} \le h_*$ is sufficiently small. The size of the threshold h_* depends on both u and f. This implies the saturation assumption (1.1).

For $f \in L_2(\Omega)$, however, the ratio $o(h_{\mathfrak{T}})/h_{\mathfrak{T}}$ may tend to zero extremely slowly for the validity of (1.4) in practice. If $f \in H^s(\Omega)$ with $0 < s \le 1$, then $\operatorname{osc}(f,\mathfrak{T}) \le C h_{\mathfrak{T}}^{1+s}$ and the asymptotic regime could be reached with practical meshes. Note that $s < 1/2$ allows for discontinuous right-hand sides f.

Acknowledgements. This work was developed while the authors were visiting the Mathematical Sciences Research Institute (MSRI, Berkeley) to participate on the workshop on *A Posteriori Error Estimation and Adaptive Approaches in the Finite Element Method*. Both authors thank the MSRI for its support and hospitality, as well as for its exciting research atmosphere.

References

1. M. Ainsworth, T. Oden: A Posteriori Error Estimation in Finite Element Analysis. New York: John Wiley 2000
2. R.E. Bank, A. Weiser: Some a posteriori error estimators for elliptic partial differential equations. Math. Comp. **44**, 285–301 (1985)
3. F. A. Bornemann, B. Erdmann, R. Kornhuber: A posteriori error estimates for elliptic problems in two and three space dimensions. SIAM J. Numer. Anal. **33**, 1188–1204 (1996)
4. P.G. Ciarlet, J.L. Lions: Handbook of Numerical Analysis, vol. 2. Amsterdam: North-Holland 1989
5. P. Clément: Approximation by finite element functions using local regularizations. RAIRO Anal. Numér. **2**, 77–84 (1975)
6. W. Dörfler: A convergent adaptive algorithm for Poisson's equation. SIAM J. Numer. Anal. **33**, 1106–1124 (1996)
7. P. Morin, R.H. Nochetto, K. Siebert: Data oscillation and convergence of adaptive FEM. SIAM J. Numer. Anal. **38**, 466–488 (2000)
8. R.H. Nochetto: Removing the saturation assumption in a posteriori error analysis. Istit. Lombardo Sci. Lett. Rend. A **127**, 67–82 (1993)
9. R. Verfürth: A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques. Chichester: Wiley-Teubner 1996