LOCAL A POSTERIORI ERROR ESTIMATOR BASED ON THE HYPERCIRCLE METHOD

Tomáš Vejchodský
Mathematical Institute, Academy of Sciences
Žitná 25, CZ–115 67 Prague 1, Czech Republic
e-mail: vejchod@math.cas.cz

Key words: a posteriori error estimates, method of hypercircle, equilibrated residual method, finite element method, linear elliptic problem.

Abstract. The error of the finite element solution of linear elliptic problems can be estimated a posteriori by the classical hypercircle method. This method gives accurate and guaranteed upper bound of the error measured in the energy norm. The disadvantage is that a global dual problem has to be solved, which is quite time-consuming. Combining the hypercircle method with the equilibrated residual method, we obtain locally computable guaranteed upper bound. The computer implementation of this a posteriori error estimator is also discussed.
1 INTRODUCTION

The way how to construct a guaranteed and locally computable a posteriori error estimate for linear elliptic problems is shown in this article. By the word guaranteed we mean that the computed estimator is proven to be an upper bound of the energy norm of error. These proofs are based on the assumption that all integrals are evaluated exactly. For locally computable estimators there exist fast algorithms for their evaluation. The proposed a posteriori error estimator combines the equilibrated residual method and the method of hypercircle. The idea to combine these two methods is not new. It has already been suggested in [8], which is a fundamental paper for the equilibrated residual method. However, the a posteriori error estimator proposed in the above mentioned paper is not completely explicitly computable in 2D and, moreover, the data are assumed to be piecewise constant there. The aim of the current paper is to show how to construct completely computable a posteriori error estimator for arbitrary data.

The origin of equilibration residual method goes back to works [8], [4], and [3]. The detailed description of this method can be found in books [1] and [2]. On the other hand, the hypercircle method is much older and has many application. Let us mention only the fundamental book [11].

The current article is organized as follows. The linear elliptic model problem is described in Section 2. The equilibrated residual method and the method of hypercircle are briefly introduced in Sections 3 and 4. A posteriori error estimator based on the combination of these two methods is presented in Section 5. Finally, experimental comparison of the mentioned estimators is shown in Section 6.

2 MODEL PROBLEM AND NOTATION

The a posteriori error estimators in Sections 3, 4, and 5 are explained on a simple model problem. We consider the following classical formulation of the linear elliptic problem in two dimensions

\[-\nabla \cdot (A \nabla \bar{u}) = f \quad \text{in } \Omega,\]
\[\bar{u} = g_D \quad \text{on } \Gamma_D,\]
\[(A \nabla \bar{u}) \cdot \nu = g_N \quad \text{on } \Gamma_N,\]

where $\Omega \subset \mathbb{R}^2$ is a polygonal domain with Lipschitz boundary $\partial \Omega$, $\nu$ is an outer unit normal to $\partial \Omega$, $\overline{\Gamma_D} \cup \overline{\Gamma_N} = \partial \Omega$, $\Gamma_D \cap \Gamma_N = \emptyset$, $\Gamma_N$ has only a finite number of components and the dot symbolizes the Euclidean inner product.

For the purpose of the weak formulation, let us assume that the matrix $A$ is symmetric and uniformly positive definite with entries in $L^\infty(\Omega)$ and that the function $g_D$ is also defined in the interior of $\Omega$ such that $g_D \in H^1(\Omega)$. The weak solution $\bar{u} \in H^1(\Omega)$ of (2.1) is given as a sum $\bar{u} = u + g_D$, where $u \in V = \{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_D\}$ satisfies

\[(A \nabla u, \nabla v) = (f, v) - (A \nabla g_D, \nabla v) + \langle g_N, v \rangle \quad \forall v \in V.\]
The right-hand side $f$ and the Neumann boundary condition $g_N$ are assumed to be in $L^2(\Omega)$ and $L^2(\Gamma_N)$, respectively. The following notation is used

$$(A\nabla u, \nabla v) = \int_{\Omega} (A\nabla u) \cdot \nabla v \, dx, \quad u, v \in V,$$

$$(f, v) = \int_{\Omega} fv \, dx, \quad f, v \in L^2(\Omega),$$

$$\langle g_N, v \rangle = \int_{\Gamma_N} g_N v \, ds, \quad g_N, v \in L^2(\Gamma_N).$$

Note that the $L^2$-inner product for vector functions from $[L^2(\Omega)]^2$ and for scalar functions from $L^2(\Omega)$ is denoted by the same symbol.

Let us consider an approximate solution $\bar{u}_h = u_h + g_D$ of problem (2.2), where $\bar{u}_h \in H^1(\Omega)$ and $u_h \in V_h$. The finite element space $V_h \subset V$ is based on a triangulation $T_h$ and consists of continuous and piecewise polynomial functions of degree $p$. The finite element solution $u_h \in V_h$ is defined by

$$(A\nabla u_h, \nabla v_h) = (f, v_h) - (A\nabla g_D, \nabla v_h) + \langle g_N, v_h \rangle \quad \forall v_h \in V_h. \quad (2.3)$$

The well known residual equation follows from (2.2):

$$(A\nabla e, \nabla v) = \mathcal{R}(v) \quad \forall v \in V, \quad (2.4)$$

where $e = u - u_h = \bar{u} - \bar{u}_h$ is the error of the approximate solution and the right-hand side is called residuum and is given by

$$\mathcal{R}(v) = (f, v) - (A\nabla \bar{u}_h, \nabla v) + \langle g_N, v \rangle \quad \forall v \in V.$$

We have the $L^2$-norm $\|v\|_{0,\Omega}^2 = (v, v)$ and the energy norm $\|v\|^2 = (A\nabla v, \nabla v)$ for all $v \in V$.

### 3 THE EQUILIBRATED RESIDUAL METHOD

The equilibrated residual method is motivated by the observation that the gradient of the exact solution $\nabla u$ is continuous over the domain $\Omega$ but the gradient of numerical solution $\nabla u_h$ has discontinuities over the edges of elements. The goal is to postprocess the fluxes of numerical solution on edges of the triangulation to obtain continuous fluxes.

We present only a short overview of the main features of this method here. For a more detailed description the book [1] is recommended.

The idea of the equilibrated residual method is to split the residuum $\mathcal{R}(v)$ into contributions from individual elements:

$$\mathcal{R}(v) = \sum_{K \in T_h} \mathcal{R}_K^{EQ}(v|_K) \quad \forall v \in V. \quad (3.1)$$
The local equilibrated residuum $\mathcal{R}_K^{EQ}(v)$ is given by

$$\mathcal{R}_K^{EQ}(v) = (f, v)_K - (A\nabla \bar{u}_h, \nabla v)_K + \langle g_K, v \rangle_{\partial K} \quad \forall v \in V(K),$$

where $V(K) = \{ v \in H^1(K) : v = 0 \text{ on } \Gamma_D \}$, and

$$(A\nabla u, \nabla v)_K = \int_K (A\nabla u) \cdot \nabla v \, dx, \quad u, v \in H^1(K),$$

$$(f, v)_K = \int_K f v \, dx, \quad f, v \in L^2(K),$$

$$\langle g_K, v \rangle_{\partial K} = \int_{\partial K} g_K v \, ds, \quad g_K, v \in L^2(\partial K).$$

The key role in the definition of the equilibrated residuum play the boundary fluxes $g_K$. These boundary fluxes approximate the actual fluxes of the true solution on the element boundaries

$$g_K \approx \nabla u \cdot \nu_K \text{ on } \partial K,$$

where $\nu_K$ denotes the outer unit normal to $\partial K$. The boundary fluxes are constructed as polynomials of degree $p$ on edges of the element $K$ which are not on $\Gamma_N$ and they satisfy the condition

$$g_K + g_{K^*} = 0 \text{ on } \partial K \cap \partial K^*, \quad (3.2)$$

where $K$ and $K^*$ are two adjacent elements, and

$$g_K = g_N \text{ on } \partial K \cap \partial \Gamma_N. \quad (3.3)$$

It is easily seen that conditions (3.2) and (3.3) imply equality (3.1).

The main advantage of the concept of boundary fluxes is that they can be computed from the numerical solution by solving local systems of linear equations on patches of elements. Thus, they can be determined quickly. Moreover, they can be established such that they satisfy the so-called $p$-th order equilibration condition:

$$\mathcal{R}_K^{EQ}(\theta_K) = (f, \theta_K)_K - (A\nabla \bar{u}_h, \nabla \theta_K)_K + \langle g_K, \theta_K \rangle_{\partial K} = 0 \quad (3.4)$$

for all finite element basis functions $\theta_K$ on the element $K$. In other words, equation (3.4) is satisfied for all polynomials of degree $p$ on the element $K$. The procedure of computing boundary fluxes $g_K$ is not easy and it is not the aim to describe it in this paper. The reader can find this procedure, e.g., in [1].

Now, having the residuum split according to (3.1), we define the solution $\Phi_K \in V(K)$ of the local residual problem

$$(A\nabla \Phi_K, \nabla v)_K = (f, v)_K - (A\nabla \bar{u}_h, \nabla v)_K + \langle g_K, v \rangle_{\partial K} \quad \forall v \in V(K). \quad (3.5)$$
Note that this problem corresponds to the local Neumann problem with boundary conditions given by $g_K$. This Neumann problem has a solution only if

$$ (f, 1)_K + \langle g_K, 1 \rangle_{\partial K} = 0, $$

but this is satisfied thanks to equilibration condition (3.4). The solution of this local problem is unique up to an additional constant, which is not important, since we are interested only in the gradient of $\Phi_K$, see (3.7) below.

Let us derive a guaranteed upper bound employing the local solutions $\Phi_K$ of (3.5). First, we rewrite residual equation (2.4):

$$ (\mathcal{A}\nabla e, \nabla v) = \sum_{K \in T_h} \mathcal{R}_{K}^{\text{EQ}}(v) = \sum_{K \in T_h} (\mathcal{A}\nabla \Phi_K, \nabla v)_K \quad \forall v \in V. $$

The next step is to employ two times the Cauchy-Schwarz inequality:

$$ |(\mathcal{A}\nabla e, \nabla v)| \leq \sum_{K \in T_h} \|\Phi_K\|_K \|v\|_K \leq \left( \sum_{K \in T_h} \|\Phi_K\|_K^2 \right)^{1/2} \|v\|, \quad (3.6) $$

where $\|v\|_K^2 = (\mathcal{A}\nabla v, \nabla v)_K$ denotes the local energy norm. Finally, inequality (3.6) implies

$$ \|e\| = \sup_{0 \neq v \in V} \frac{|(\mathcal{A}\nabla e, \nabla v)|}{\|v\|} \leq \left( \sum_{K \in T_h} \|\Phi_K\|_K^2 \right)^{1/2}. \quad (3.7) $$

Thus, the local solutions $\Phi_K$ provide guaranteed upper bound of the energy norm of the error. The trouble with the error estimate (3.7) is that the functions $\Phi_K$ as solutions of infinitely dimensional problems (3.5) are not computable. In practice, we usually use a higher order finite element approximation of $\Phi_K$ in (3.7). Therefore, the practical implementation of the a posteriori error estimator (3.7) is not a guaranteed upper bound.

4 THE METHOD OF HYPERCIRCLE

Let us derive a guaranteed and computable upper bound employing the method of hypercircle. The standard procedure in the method of hypercircle is the construction of the dual problem in the sense of the calculus of variations, see for example [9], [10], [5], [6], etc. We derive the a posteriori error estimator directly without the definition of the dual problem.

Recall that $H(\text{div}, \Omega) \subset [L^2(\Omega)]^2$ is the well-known space of functions with divergence in $L^2(\Omega)$. Substituting $v = e = \bar{u} - \bar{u}_h$ into weak formulation (2.2), we get:

$$ - (\mathcal{A}\nabla \bar{u}, \nabla e) = - (f, e) - \langle g_N, e \rangle. \quad (4.1) $$
Introducing a norm \( \|q\|_{A^{-1}, \Omega}^2 = (A^{-1}q, q) \) for \( q \in H(\text{div}, \Omega) \) and employing (4.1), we can compute
\[
\|q - A\nabla \bar{u}_h\|_{A^{-1}, \Omega}^2 = (A^{-1}q - \nabla \bar{u} - \nabla \bar{u}_h + \nabla \bar{u}, q - A\nabla \bar{u} - A\nabla \bar{u}_h + A\nabla \bar{u})
= \|q - A\nabla \bar{u}\|_{A^{-1}, \Omega}^2 + 2 (q - A\nabla \bar{u}, \nabla \bar{u} - \nabla \bar{u}_h) + \|\bar{u} - \bar{u}_h\|^2
= \|q - A\nabla \bar{u}\|_{A^{-1}, \Omega}^2 + 2 (q - A\nabla \bar{u}, \nabla \bar{u} - \nabla \bar{u}_h) + \|\bar{u} - \bar{u}_h\|^2
\]
for arbitrary \( q \in H(\text{div}, \Omega) \). Defining the following space
\[
Q(f, g_N) = \{ q \in H(\text{div}, \Omega) : (q, \nabla v) = (f, v) + \langle g_N, v \rangle \quad \forall v \in V \},
\]
we immediately conclude that
\[
\|q - A\nabla \bar{u}_h\|_{A^{-1}, \Omega}^2 = \|q - A\nabla \bar{u}\|_{A^{-1}, \Omega}^2 + \|\bar{u} - \bar{u}_h\|^2 \quad \forall q \in Q(f, g_N).
\]
Thus, any \( q \in Q(f, g_N) \) gives the guaranteed upper bound
\[
\|e\|^2 = \|\bar{u} - \bar{u}_h\|^2 \leq \|q - A\nabla \bar{u}_h\|_{A^{-1}, \Omega}^2. \tag{4.2}
\]
Note that this estimator is exact if \( q = A\nabla \bar{u} \), but it is unreachable, in general. The question is, how to find a suitable function \( q \in Q(f, g_N) \), which would produce a tight upper bound in (4.2). The crucial ingredient for the answer is the structure of \( Q(f, g_N) \), which is described in paper [7]. We will follow the idea described therein.

The first observation is that
\[
Q(f, g_N) = \bar{p} + Q(0, 0),
\]
where \( \bar{p} \in Q(f, g_N) \) is arbitrary but fixed and
\[
Q(0, 0) = \{ q \in H(\text{div}, \Omega) : (q, \nabla v) = 0 \quad \forall v \in V \}.
\]
The second fact, proven in [7], is that
\[
Q(0, 0) = \text{curl } W,
\]
where
\[
W = \{ v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_N \}.
\]
Note that the operator \( \text{curl} \) is defined by \( \text{curl} = (\partial/\partial x_2, -\partial/\partial x_1)^\top \). The way how to construct \( \bar{p} \) is presented in [7], too:
\[
\bar{p} = F + \text{curl } w, \tag{4.3}
\]
where
\[
F(x_1, x_2) = \left( -\int_0^{x_1} f(s, x_2) \, ds, 0 \right)^\top \tag{4.4}
\]
and \( w \in H^1(\Omega) \) is an arbitrary function satisfying
\[
\text{curl} \ w \cdot \nu = \nabla w \cdot \tau = g_N - F \cdot \nu \text{ on } \Gamma_N,
\]
where \( \tau = (-\nu_2, \nu_1) \) is a unit tangent vector to \( \Gamma_N \). We see that the tangent derivative of \( w \) is given by \( g_N - F \cdot \nu \). Thus, the values of \( w \) on \( \partial K \) are given by the primitive function to \( g_N - F \cdot \nu \).

To conclude these steps we can rewrite estimate (4.2) as follows
\[
\| \bar{u} - \bar{u}_h \|^2 \leq \| \bar{p} + \text{curl} \ y - A \nabla \bar{u}_h \|^2_{A^{-1}, \Omega} \quad \forall y \in W.
\]
Now, it is natural to replace \( W \) by a finite dimensional subspace \( W_h \subset W \) in order to obtain a computable estimate. The easiest way is to construct \( W_h \) as a finite element space based on the same triangulation with the same polynomial degree as \( V_h \). The optimal choice \( y_h \in W_h \), which minimizes the right-hand side of (4.6) over \( W_h \) satisfies
\[
(A^{-1} \text{curl} y_h, \text{curl} v_h) = (\nabla \bar{u}_h - A^{-1} \bar{p}, \text{curl} v_h) \quad \forall v_h \in W_h.
\]
Note that the idea to minimize the right-hand side of (4.6) over a finite dimensional space is used, e.g., in [12].

Now, we summarize how to construct the computable guaranteed upper bound of the energy norm of the error \( e = \bar{u} - \bar{u}_h \) by the method of hypercircle. First, construct the function \( \bar{p} \) according to (4.3), (4.4), and (4.5). Then, find solution \( y_h \in W_h \) of the finite dimensional problem (4.7) and compute the norm \( \| \bar{p} + \text{curl} y_h - A \nabla \bar{u}_h \|^2_{A^{-1}, \Omega} \), which is the guaranteed upper bound – see (4.6).

The disadvantage of this procedure is that problem (4.7) is global and its solution is time-consuming for large problems.

5 THE COMBINED METHOD

In order to obtain the locally computable guaranteed upper bound, we combine the local but not guaranteed equilibrated residual method with the guaranteed but not local hypercircle method. The idea is straightforward – compute the boundary fluxes \( g_K \) by the equilibrated residual method and then apply the method of hypercircle to local residual problem (3.5). To do that let us rewrite the equality (3.5) with \( v = \Phi_K \):
\[
-(A \nabla \Phi_K, \nabla \Phi_K)_K = -(f, \Phi_K)_K + (A \nabla \bar{u}_h, \nabla \Phi_K)_K - (g_K, \Phi_K)_{\partial K}.
\]

Now, let us compute the following norm for any \( q \in H(\text{div}, K) \):
\[
\|q\|^2_{A^{-1}, K} = \|q - A \nabla \Phi_K\|^2_{A^{-1}, K} + 2(q - A \nabla \Phi_K, \nabla \Phi_K)_K + \|\Phi_K\|^2_K
\]
\[
= 2(q, \nabla \Phi_K)_K - 2(f, \Phi_K)_K + 2(A \nabla \bar{u}_h, \nabla \Phi_K)_K - 2(g_K, \Phi_K)_{\partial K}
\]
\[
+ \|q - A \nabla \Phi_K\|^2_{A^{-1}, K} + \|\Phi_K\|^2_K,
\]
where (5.1) was utilized. By analogy with the previous section we define the following space

\[ Q_K(f, g_K, \bar{u}_h) = \{ q \in H(\text{div}, K) : (q, \nabla v)_K = (f, v)_K - (\mathcal{A} \nabla \bar{u}_h, \nabla v)_K + (g_K, v)_{\partial K} \quad \forall v \in V(K) \}. \]

The structure of the space \( Q_K(f, g_K, \bar{u}_h) \) is the same as the structure of \( Q(f, g_N) \) in the previous section. Namely,

\[ Q_K(f, g_K, \bar{u}_h) = \bar{p}_K + Q_K(0, 0, 0) = \bar{p}_K + \text{curl} W(K), \]

where \( \bar{p}_K \) is arbitrary but fixed element of \( Q_K(f, g_K, \bar{u}_h) \) and

\[ Q_K(0, 0, 0) = \{ q \in H(\text{div}, K) : (q, \nabla v)_K = 0 \quad \forall v \in V(K) \} = \text{curl} W(K), \]

\[ W(K) = \{ v \in H^1(K) : v = 0 \text{ on } \partial K \setminus \Gamma_D \}. \]

From (5.2) we immediately obtain

\[ \| q \|_{A^{-1}, K}^2 = \| q - \mathcal{A} \nabla \Phi_K \|_{A^{-1}, K}^2 + \| \Phi_K \|_K^2 \quad \forall q \in Q_K(f, g_K, \bar{u}_h). \]

Consequently,

\[ \| \Phi_K \|_K \leq \| q \|_{A^{-1}, K} \quad \forall q \in Q_K(f, g_K, \bar{u}_h). \tag{5.3} \]

Finally, using (3.7), (5.3), and the structure of the space \( Q_K(f, g_K, \bar{u}_h) \), we conclude that

\[ \| e \|^2 \leq \sum_{K \in T_h} \| \Phi_K \|_K^2 \leq \sum_{K \in T_h} \| q \|_{A^{-1}, K}^2 = \sum_{K \in T_h} \| \bar{p}_K + \text{curl} y_K \|_{A^{-1}, K}^2 \tag{5.4} \]

holds for all \( y_K \in W(K) \) with \( q = \bar{p}_K + \text{curl} y_K \).

In order to evaluate the upper bound (5.4) we consider a finite dimensional subspace \( W_h(K) \subset W(K) \). The subspace \( W_h(K) \) may consist, for example, of polynomials of degree \( p + 1 \) on the triangle \( K \). The optimal choice for \( y_K \in W_h(K) \), which minimizes the right-hand side of (5.4) over \( W_h(K) \), satisfies

\[ (A^{-1} \text{curl} y_K, \text{curl} v)_K = - (A^{-1} \bar{p}_K, \text{curl} v)_K \quad \forall v \in W_h(K). \tag{5.5} \]

The only remaining and crucial question is how to find the vector \( \bar{p}_K \in Q_K(f, g_K, \bar{u}_h) \) efficiently. The clue is given in the previous section, see (4.3), (4.4), and (4.5):

\[ \bar{p}_K = \mathbf{F} + \text{curl} w_K - \mathcal{A} \nabla \bar{u}_h, \tag{5.6} \]

where \( \mathcal{A} \nabla \bar{u}_h \) is known, \( \mathbf{F} \) is given by (4.4) and \( w_K \in H^1(K) \) has to satisfy

\[ \text{curl} w_K \cdot \nu_K = \frac{\partial w_K}{\partial \tau_K} = g_K - \mathbf{F} \cdot \nu_K \quad \text{on } \partial K \setminus \Gamma_D, \tag{5.7} \]
where the tangent vector is given by \( \tau_K = (-\nu_{K,2}, \nu_{K,1})^\top \). Notice that the values of \( w_K \) on the boundary \( \partial K \) are given by the primitive function to \( g_K - \mathbf{F} \cdot \nu_K \). For example, if we consider the triangle \( K \) with vertices \( A, B, C \), then the values \( w_K(x) \) for \( x \in \partial K \) are given by

\[
 w_K(x) = \begin{cases}
 w_K(A) + \int_A^x (g_K - \mathbf{F} \cdot \nu_K) \, ds & \text{for } x \in \overline{AB}, \\
 w_K(B) + \int_B^x (g_K - \mathbf{F} \cdot \nu_K) \, ds & \text{for } x \in \overline{BC}, \\
 w_K(C) + \int_C^x (g_K - \mathbf{F} \cdot \nu_K) \, ds & \text{for } x \in \overline{CA},
\end{cases}
\]

where \( \overline{AB}, \overline{BC}, \) and \( \overline{CA} \) denote the edges of the triangle, the integration is along these edges and

\[
 w_K(A) \in \mathbb{R} \text{ is arbitrary},
\]

\[
 w_K(B) = w_K(A) + \int_A^B (g_K - \mathbf{F} \cdot \nu_K) \, ds,
\]

\[
 w_K(C) = w_K(B) + \int_B^C (g_K - \mathbf{F} \cdot \nu_K) \, ds.
\]

Note that the particular value of \( w_K(A) \) is irrelevant, because we are interested only in \( \operatorname{curl} w_K \). The constants \( w_K(B) \) and \( w_K(C) \) are chosen such that \( w_K \) is continuous in the points \( B \) and \( C \). The important fact is that the function \( w_K \) is continuous also in the point \( A \). Denoting by \( [w_K(A)] \) a possible jump in values of \( w_K \) in the point \( A \), we find that it is zero, because

\[
[w_K(A)] = \int_{\partial K} \frac{\partial w_K}{\partial \tau_K} = \int_{\partial K} (g_K - \mathbf{F} \cdot \nu_K) \, ds = \int_{\partial K} g_K \, ds + \int_K f \, dx = 0,
\]

where the last equality follows from equilibration condition (3.4) with \( \theta_K = 1 \).

Thus, the values of \( w_K \) are continuous on the boundary of \( \partial K \) and it is possible to prolong them into the interior of \( K \) such that the prolongation lies in \( H^1(K) \). There are infinitely many possibilities how to do this prolongation, but we need the one which leads to the tight upper bound.

We suggest the prolongation illustrated in Figure 1. Consider the triangle \( K \) with vertices \( A, B, C \) and continuous function \( \omega \) defined on \( \partial K \), i.e., \( \omega \in C^0(\partial K) \). Let us define the function \( \tilde{\omega} \in C^0(\overline{K}) \) in the following way. First, construct four auxiliary functions:

- The function \( \tilde{\ell} \) is a linear function on \( \overline{K} \) such that \( \tilde{\ell}(A) = \omega(A), \tilde{\ell}(B) = \omega(B), \) and \( \tilde{\ell}(C) = \omega(C), \) i.e.,

  \[
  \tilde{\ell} = \omega(A) \varphi_A + \omega(B) \varphi_B + \omega(C) \varphi_C,
  \]

  where \( \varphi_A, \varphi_B, \) and \( \varphi_C \) are standard linear finite element basis functions on the triangle \( K \).
• The function \( \tilde{z}_3 \in C^0(K) \) is zero on \( \partial K \setminus AB \) and is defined by

\[
\tilde{z}_3(X) = \omega(X) - \tilde{\ell}(X) \quad \text{for } X \in AB,
\]

\[
\tilde{z}_3(X) = 0 \quad \text{for } X \in BC \cup CA,
\]

\[
\tilde{z}_3(X) = \left( \omega(D_3) - \tilde{\ell}(D_3) \right) \frac{|XE_3|}{|D_3E_3|} \quad \text{for } X \in K,
\]

where \( |XE_3| \) denotes the distance between the points \( X \) and \( E_3 \). The point \( D_3 \) is the intersection of the edge \( AB \) and the line which goes through the point \( X \) and is parallel to the edge \( CA \). The point \( E_3 \) is the intersection of this line with the edge \( BC \) – see Figure 2. Note that \( \tilde{z}_3 \) is linear on the line \( D_3E_3 \).

• The functions \( \tilde{z}_1 \in C^0(K) \), which is zero on \( \partial K \setminus BC \), and \( \tilde{z}_2 \in C^0(K) \), which is zero on \( \partial K \setminus CA \), are defined by analogy with the definition of \( \tilde{z}_3 \).

The function \( \tilde{\omega} \in C^0(K) \) is defined by

\[
\tilde{\omega}(X) = \tilde{\ell}(X) + \tilde{z}_1(X) + \tilde{z}_2(X) + \tilde{z}_3(X), \quad X \in K.
\] (5.8)

Notice that \( \tilde{\omega}(X) = \omega(X) \) on \( \partial K \). This function \( \tilde{\omega} \) will be called the prolongation of \( \omega \) into interior of \( K \), or simply the prolongation of \( \omega \).

The described prolongation has several important properties, which are introduced in the following lemmas. Denote by \( P^p(\Theta) \) the space of polynomials of degree \( p \) defined on a set \( \Theta \).

Figure 1: The left figure shows the splitting of \( w \) into the linear part \( \ell \) and to the rest parts \( z_1, z_2 \) and \( z_3 \) corresponding to edges of triangle \( K \). The right figure illustrates the prolongation of \( z_1, z_2 \) and \( z_3 \) into the interior of \( K \).
Lemma 5.1. Consider a triangle $K$ and $\omega \in C^0(\partial K)$. Moreover, let $\omega|_{\gamma} \in P^p(\gamma)$ for all edges $\gamma$ of the triangle $K$ and for arbitrary $p \in \mathbb{N}$. Then the prolongation $\tilde{\omega}$ of the function $\omega$ into the interior of $K$ given by (5.8) is a polynomial of degree $p$ in $K$, i.e., $\tilde{\omega} \in P^p(K)$.

Proof. Let us assume without loss of generality that $K$ is a reference triangle with the vertices $A = [0, 0]$, $B = [1, 0]$, and $C = [0, 1]$. It is enough to prove that $\tilde{z}_1$, $\tilde{z}_2$, and $\tilde{z}_3$ in (5.8) are polynomials of degree $p$. Consider, therefore, a function $z_3 \in C^0(\partial K)$, which is zero on $\overline{BC}$ and $\overline{CA}$ and which is a polynomial of degree $p$ on $\overline{AB}$. Note that the functions

$$\varphi_n^{1D}(x) = x^n(1 - x), \quad n = 1, 2, \ldots, p - 1,$$

form a basis of the space $P^p([0, 1])$ of all polynomials on $[0, 1]$ with zeroes at 0 and 1. Therefore, we can uniquely express the function $z_3$ on interval $[0, 1]$ as a linear combination of these basis functions:

$$z_3 = \sum_{n=1}^{p-1} c_n \varphi_n^{1D} \text{ on } \overline{AB} = [0, 1].$$

Now, we can consider the unique prolongation

$$\tilde{z}_3 = \sum_{n=1}^{p-1} c_n \varphi_n^{2D} \text{ in } K,$$

where

$$\varphi_n^{2D}(x_1, x_2) = x_1^n(1 - x_1 - x_2), \quad n = 1, 2, \ldots, p - 1,$$

are some of the standard finite element basis functions on the reference triangle. Now, consider lines parallel with edge $\overline{CA}$, i.e., lines described by the equality $x_1 = k$, $k \in \mathbb{R}$. All basis functions $\varphi_n^{2D}(x_1, x_2)$, $n = 1, 2, \ldots, p - 1$, are linear on these lines:

$$\varphi_n^{2D}(k, x_2) = k^n(1 - k - x_2), \quad n = 1, 2, \ldots, p - 1.$$
Therefore, also the function \( \hat{z}_3 \) is linear on lines parallel with edge \( \overline{CA} \), which is exactly the way, how the prolongation \( \hat{z}_3 \) described above is constructed. Hence, the prolongations \( \hat{z}_3 \) and \( \hat{z}_3 \) of the function \( z_3 \) coincide and we conclude that \( \hat{z}_3 \) is a polynomial of degree \( p \) in \( K \). Analogical reasoning can be done for \( z_1 \) and \( z_2 \), too. \( \square \)

Lemma 5.2. Consider a triangle \( K \) with vertices \( A, B, C \), function \( \omega \in C^0(\partial K) \) and its prolongation \( \hat{\omega} \in C^0(K) \) defined by (5.8). If there exists a finite tangent derivative \( \partial \omega / \partial \tau_K \) on all edges of the triangle \( K \) then the derivatives of the function \( \hat{z}_3 \), see (5.8), in the directions \( \overrightarrow{ED} \) and \( \overrightarrow{XB} \) are given at any interior point \( X = (x_1, x_2) \in K \) by

\[
\frac{\partial \hat{z}_3(X)}{\partial E_3D_3} = \frac{\hat{z}_3(D_3)}{|D_3E_3|}, \quad \frac{\partial \hat{z}_3(X)}{\partial XB} = \frac{\hat{z}_3(D_3)}{|XB|} \left[ \frac{|XE_3|}{|D_3E_3|} \right] \alpha,
\]

where

\[
\alpha = \frac{(B_1 - x_1)(A_1 - C_1) - (B_2 - x_2)(A_2 - C_2)}{(B_1 - A_1)(A_1 - C_1) - (B_2 - A_2)(A_2 - C_2)}.
\]

Proof. Figure 3 illustrates the used notations. The derivative in the directions \( \overrightarrow{ED} \) is easy to calculate. The derivative in the directions \( \overrightarrow{XB} \) is given by

\[
\lim_{r \to 0} \frac{\hat{z}_3(X + r(B - X)) - \hat{z}_3(X)}{r |BX|} = \lim_{r \to 0} \frac{\hat{z}_3(X) - \hat{z}_3(X)}{r |BX|} = \lim_{r \to 0} \frac{\hat{z}_3(D_3) - z_3(D_3)}{r |BX|} |XE_3| \bigg/ |D_3E_3|,
\]

where \( \hat{X} = X + r(B - X) \),

\[
\hat{z}_3(X) = \hat{z}_3(D_3) \frac{|XE_3|}{|D_3E_3|}, \quad \hat{z}_3(\hat{X}) = \hat{z}_3(\hat{D}_3) \frac{|\hat{X}E_3|}{|\hat{D}_3E_3|}.
\]
and
\[
\frac{|XE_3|}{|D_3E_3|} = \frac{\hat{X}\hat{E}_3}{\hat{D}_3\hat{E}_3},
\]
which holds thanks to the similarity of appropriate triangles. The rest of the proof is an exercise in analytical geometry. \qed

Note that the derivatives of \(\hat{z}_1\) and \(\hat{z}_2\) can be evaluated by analogy with Lemma 5.2.

Let us summarize the way how to compute the guaranteed and locally computable upper bound. First, compute boundary fluxes \(g_K\) using residual equilibration method. Then, construct for all triangles \(K\) in \(T_h\) vector \(\vec{p}_K\) by (5.6), where \(F\) is given by (4.4), \(A\nabla\bar{u}_h\) is known and \(w_K\) is constructed according to (5.7). The values of \(w_K\) on \(\partial K\) are given by the primitive function to \(g_K - F \cdot \nu\). These values are prolonged into the interior of \(K\) according to (5.8) and this prolongation is denoted also by \(w_K\). Important is that the values of \(\text{curl} w_K\) are easily computable from values of \(w_K\) on \(\partial K\) and from \(\partial w_K/\partial \tau_K = g_K - F \cdot \nu\) on \(\partial K\) thanks to Lemma 5.2. The next step is to find a solution \(y_K \in W_h(K)\) of finite dimensional local problem (5.5). The final step is to evaluate estimate (5.4).

Thanks to prolongation (5.8) we can prove the following lemma.

**Lemma 5.3.** Let the finite element solution \(u_h \in V_h\) be exact, i.e., \(u_h = u\) and let the matrix \(A\) be constant. If the vector \(\vec{p}_K \in Q_K(f, g_K, \bar{u}_h)\) is constructed as described above then the error estimator (5.4) is exact, i.e., \(\vec{p}_K + \text{curl} y_K = 0\).

**Proof.** Consider a triangle \(K \in T_h\). The right-hand side \(f\) has to be a polynomial of degree at most \(p - 2\) on \(K\), since \(u_h = u\) is a polynomial of degree at most \(p\) and \(A\) is constant. Therefore, \(F\) is a polynomial of degree at most \(p - 1\). The boundary fluxes \(g_K\) are constructed as polynomials of degree \(p\) on edges of \(K\) and they coincide with the actual fluxes of the exact solution:
\[
g_K = \nabla\bar{u} \cdot \nu_K \text{ on } \partial K.
\]
By the way, this fact implies that \(\Phi_K = 0\), see (3.5). Moreover, it implies that
\[
Q_K(f, g_K, \bar{u}_h) = \{q \in H(\text{div}, K) : (q, \nabla v)_K = 0 \quad \forall v \in V(K)\} = \text{curl} W(K).
\]
Thus, the function \(w_K\) is a polynomial of degree at most \(p + 1\) on each edge of \(K\) and Lemma 5.1 implies that the prolongation of these values into the interior of \(K\) is also a polynomial of degree at most \(p + 1\).

Hence, \(\vec{p}_K\) is a polynomial of degree at most \(p\) and that \(\vec{p}_K \in \text{curl} W(K)\). Consequently, \(\vec{p}_K \in \text{curl} W_h(K)\) and projection (5.5) gives \(\text{curl} y_K = -\vec{p}_K\). \qed
6 NUMERICAL EXPERIMENTS

The practical behaviour of the combined method is compared with the equilibrated residual method and with the method of hypercircle. The quadratic finite element method is chosen. In case of equilibrated residual method, the local residual problem (3.5) is solved on the space $V_h(K)$ of cubic polynomials on triangle $K$. In case of the combined method, local problem (5.5) is solved on the space $W_h(K)$ of cubic polynomials, too. The difference between local problems for equilibrated residual and combined methods is in boundary conditions and in the number of degrees of freedom. For example, if we consider the interior element $K$, then $\dim V_h(K) = 10$ and $\dim W_h(K) = 1$. Thus, the combined method performs faster in the solution of local problems. Anyway, the computation of boundary fluxes $g_K$ have to be done for both methods and it takes considerable amount of time.

The global problem (4.7) in classical method of hypercircle is solved by the quadratic finite elements. Note that the spaces $V_h$ and $W_h$ differ due to different boundary conditions for problems (2.3) and (4.7). The function $w$, see (4.3), is constructed in the similar way as function $w_K$ in Section 5.

Example 6.1. Consider problem (2.1) with the following data: $\Omega = [-1, 1]^2$, $\Gamma_D = \partial \Omega$, $\Gamma_N = \emptyset$, $A$ is the identity matrix, $f(x_1, x_2) = 2(2 - x_1^2 - x_2^2)$, $g_D = 0$. The exact solution corresponding to this problem is $u(x_1, x_2) = (x_1^2 - 1)(x_2^2 - 1)$. Table 1 shows the effectivity indices obtained by the equilibrated residual method, by the method of hypercircle and by the combined method on the regular meshes. These first five meshes are shown in Figure 4.

<table>
<thead>
<tr>
<th>$N_{\text{tri}}$</th>
<th>equilibrated residual method</th>
<th>method of hypercircle</th>
<th>combined method</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.43</td>
<td>1.11</td>
<td>1.06</td>
</tr>
<tr>
<td>4</td>
<td>1.23</td>
<td>1.25</td>
<td>1.01</td>
</tr>
<tr>
<td>8</td>
<td>1.34</td>
<td>1.20</td>
<td>1.00</td>
</tr>
<tr>
<td>16</td>
<td>1.30</td>
<td>1.30</td>
<td>1.16</td>
</tr>
<tr>
<td>32</td>
<td>1.39</td>
<td>1.23</td>
<td>1.29</td>
</tr>
<tr>
<td>64</td>
<td>1.32</td>
<td>1.32</td>
<td>1.27</td>
</tr>
<tr>
<td>128</td>
<td>1.41</td>
<td>1.25</td>
<td>1.52</td>
</tr>
<tr>
<td>256</td>
<td>1.33</td>
<td>1.34</td>
<td>1.33</td>
</tr>
<tr>
<td>512</td>
<td>1.41</td>
<td>1.25</td>
<td>1.64</td>
</tr>
<tr>
<td>1024</td>
<td>1.33</td>
<td>1.34</td>
<td>1.36</td>
</tr>
</tbody>
</table>

Table 1: Effectivity indices computed by the equilibrated residual method, by the method of hypercircle and by the combined method on regular triangulations of the domain with $N_{\text{tri}}$ triangles.
7 CONCLUSIONS

The equilibrated residual method and the classical method of hypercircle have been discussed in this paper. It was shown how to obtain a posteriori upper bounds of energy norm of error by both of these methods. However, the equilibrated residual method gives the upper bound which is not computable, since it involves solutions of infinitely dimensional local problems. In practice, approximations of these local problems are used and, therefore, the computable version is not guaranteed to be an upper bound.

On the other hand, the method of hypercircle produces the guaranteed and computable upper bound, but its evaluation needs the solution of a global problem. Thus, computation of the upper bound by the method of hypercircle is not fast.

The combination of the equilibrated residual method and the method of hypercircle has been suggested in this paper. This combined method provides the guaranteed and locally computable upper bound. Numerical experiments indicate that the combined method performs very well. It is fast and efficient. The efficiency of the combined method, measured by the effectivity index, is comparable with the equilibrated residual method as well as with the method of hypercircle, which both provide tight bounds of the error.

Acknowledgement

This work was supported by grant no. 201/04/P021 of the Grant Agency of the Czech Republic. This support is gratefully acknowledged.
REFERENCES


