Applications of Mathematics

Jan Brandts

A posteriori error estimation and adaptivity in the method of lines with mixed finite elements

Applications of Mathematics, Vol. 44 (1999), No. 6, 407–419

Persistent URL: http://dml.cz/dmlcz/134418

44 (1999) APPLICATIONS OF MATHEMATICS No. 6, 407–419

A POSTERIORI ERROR ESTIMATION AND ADAPTIVITY IN THE METHOD OF LINES WITH MIXED FINITE ELEMENTS

Jan H. Brandts, Praha

Abstract. We will investigate the possibility to use superconvergence results for the mixed finite element discretizations of some time-dependent partial differential equations in the construction of a posteriori error estimators. Since essentially the same approach can be followed in two space dimensions, we will, for simplicity, consider a model problem in one space dimension.

Keywords: superconvergence, method of lines, mixed finite elements, a posteriori error estimation, adaptive time-stepping, adaptive refinement

MSC 2000: 65M20, 65M15

1. Introduction

The popular method of lines for the approximation of some time-dependent partial differential equations has in the past years been thoroughly analysed in the case standard finite element methods were employed to realize the discretization in space. For the time-discretization, successful adaptive schemes have been developed that aim to keep the time-discretization error uniformly bounded. Together with quantitative error control, separately in both space and time, this leads to numerical schemes in which the space-discretization error usually is chosen to be dominant. We refer to [1, 2, 14, 15, 17] and also to the famous series of papers [8, 9, 10, 11, 12] of Eriksson and Johnson and [13] together with Larsson on the discontinuous Galerkin method for parabolic problems.

Recently, one also focusses on the use of mixed finite elements in the space discretization (Cf. [20, 23]), and apart from a priori estimates there are also superconvergence results available. We believe that these superconvergence results can be successfully used in the construction of a posteriori error estimation of the space-discretization error, which, in turn, gives qualitative information about how to direct an adaptive time-stepping procedure in order to keep the time-discretization error

small compared to the space-discretization error. The better one succeeds in this, the better the space-discretization error can be estimated, and the better the magnitude of the next time step can be determined.

In this short note we present numerical examples of the use in adaptive time-stepping of superconvergence-based a posteriori estimation of the mixed finite element space-discretization error. Moreover, we will consider an example with adaptivity in space based on the same error estimator as well. We will concentrate on one-dimensional problems only, but we note that the superconvergence results in [5] are in fact for two-dimensional Raviart-Thomas elements (Cf. [3, 4, 7, 6]) and that the theory for adaptive time-stepping in [15] is for problems of arbitrary space-dimension. Therefore we believe that the same approach can be followed in the two-dimensional setting.

2. A model problem

To explain the main principles involved, we will concentrate on a simple diffusion equation as model problem to illustrate its mixed semi-discretization and its fully discrete time-adaptive scheme. Then we will concentrate on superconvergence and a posteriori error estimation, and develop a fully discrete and fully adaptive algorithm.

2.1. Mixed formulation of a diffusion equation

Let I be the unit interval and $T > 0$. Further, a is a continuously differentiable strictly positive function, and $f \in L^2$ arbitrary. Our model problem will be to solve u from

$$u_t = (au_x)_x + f, \quad u_x = 0 \text{ on } \{0,1\} \times [0,T], \quad u = u_0 \text{ on } I \times \{0\}. \tag{1}$$

Note that we chose homogeneous Neumann boundary conditions; they allow the easiest superconvergence results. For unique solvability we will look for u with mean zero. The mixed formulation of (1) is

$$\begin{cases} p = au_x \\ u_t = p_x + f \end{cases} \text{on } I \times [0,T], \qquad \begin{matrix} p = 0 \ \text{ on } \{0,1\} \times [0,T] \\ u = u_0 \text{ on } I \times \{0\} \end{matrix} \tag{2}$$

Testing p in H_0^1 and u in $L^2/\mathbb{R}$ gives rise to the following mixed weak formulation of the problem. Solve $p \in H_0^1$ and $u \in L^2/\mathbb{R}$ from

$$\forall v \in H_0^1 \colon\ a(p,v) + b(u,v) = 0 \quad \text{and} \quad \forall w \in L^2 \colon\ b(w,p) + (f,w) = (u_t,w), \tag{3}$$

$$\text{where} \quad a(q,v) = (a^{-1}q, v) \quad \text{and} \quad b(w,q) = (w, q_x), \tag{4}$$

with the same initial condition as above.

2.2. Semi-discretization

The mixed finite element semi-discretization in space of (3–4) will be the following. Let $\Delta_h = \{0 = x_0, x_1, \ldots, x_{N+1} = 1\}$ be a partitioning of I, and let h be the length of the longest subinterval. Denote by $V = V(\Delta)$ the space of all continuous piecewise linear functions relative to the partitioning Δ that are zero on the boundary and let $W = W(\Delta)$ contain all the piecewise constants. Solve $u_h \in W/\mathbb{R}$ and $p_h \in V$ from

$$(5)\quad \forall v \in V\colon\ a(p_h, v) + b(u_h, v) = 0 \quad \text{and} \quad \forall w \in W\colon\ b(w, p_h) + (f, w) = ((u_h)_t, w).$$

As discrete initial condition we take the mixed elliptic projection u_h^0 of u_0 (Cf. [5, 23]). Writing $p_0 := a(u_0)_x$, u_h^0 is (part of) the solution of the system

$$(6)\qquad \forall v \in V\colon\ a(p_h^0, v) + b(u_h^0, v) = 0 \quad \text{and} \quad \forall w \in W\colon\ b(w, p_h^0) = b(w, p_0).$$

All this leads in the standard way to a system of ordinary differential equations for the coefficients of the nodal basisfunctions v_j, w_j of the spaces V and W. Note that V has dimension $N-1$ and W dimension N. Denoting the L^2-orthogonal projection on W by P_W define the coefficient vectors U, P and F by

$$(7)\qquad u_h = \sum_{j=1}^{N} U_j w_j, \quad p_h = \sum_{j=1}^{N-1} P_j v_j, \quad P_W f = \sum_{j=1}^{N} F_j w_j$$

and matrices A, B, C with entries

$$(8)\qquad a_{ij} = a(v_j, v_i), \quad b_{ij} = b(w_j, v_i), \quad c_{ij} = (w_j, w_i).$$

Then the vectors U and P are to be solved from

$$(9)\qquad \begin{cases} AP = -BU, \\ CU_t = B^T P + CF, \end{cases} \qquad U(0) = u_h^0.$$

2.3. Adaptive time-stepping using Euler backward method

In the system of equations (9), P can be eliminated, resulting in

$$(10)\qquad U_t + C^{-1}B^T A^{-1} BU = F, \quad U(0) = u_h^0.$$

In [15], a framework for adaptive time-stepping is given for the Euler backward method. However, their theory applies for symmetric and positive semi-definite system matrices. Our system matrix $C^{-1}B^T A^{-1}B$, though positive semi-definite, is however only symmetric if C is a constant times identity, which corresponds to a

uniform grid. We will therefore first derive some results in case the matrix C is not constant.

Lemma 2.1. *Consider the system of ordinary differential equations*

$$U_t + K^{-1}LU = F, \quad U(0) = u_h^0, \tag{11}$$

where K is symmetric positive definite and L symmetric positive semi-definite. Then the Euler backward iterates for (11) are the same as $K^{-\frac{1}{2}}$ applied to the Euler backward iterates of the system

$$V_t + K^{-\frac{1}{2}}LK^{-\frac{1}{2}}V = K^{\frac{1}{2}}F, \quad V(0) = K^{\frac{1}{2}}u_h^0, \tag{12}$$

of which the system matrix is symmetric positive definite.

P r o o f. We will prove the lemma by induction. First, we define the Euler backward sequences as follows. Denoting the time steps by Δ_n we get

$$(I + \Delta_{n+1}K^{-1}L)U^{n+1} = U^n + \Delta_{n+1}F, \quad U^0 = u_h^0, \tag{13}$$

$$(I + \Delta_{n+1}K^{-\frac{1}{2}}LK^{-\frac{1}{2}})V^{n+1} = V^n + \Delta_{n+1}K^{\frac{1}{2}}F, \quad V^0 = K^{\frac{1}{2}}u_h^0. \tag{14}$$

As induction hypothesis we take that $K^{-\frac{1}{2}}V^n = U^n$ for certain n. Rewriting (14) as

$$K^{\frac{1}{2}}(I + \Delta_{n+1}K^{-1}L)K^{-\frac{1}{2}}V^{n+1} = K^{\frac{1}{2}}(U^n + \Delta_{n+1}F) \tag{15}$$

immediately shows that $K^{-\frac{1}{2}}V^{n+1} = U^{n+1}$. Together with the initial conditions on both sequences, this completes the proof. □

Now, defining $L := B^TA^{-1}B$ and $K := C$, we see that the Euler backward discretization of (10) is of the form (13). We will now relate the results of [15] for the sequence V^n to the sequence U^n that solves our problem. We first state their mean theorem loosely, for the exact conditions and notations we refer to the original paper.

Theorem 2.2. [15] *Suppose that the time steps Δ_n are chosen such that for each pair of consecutive approximations in a fixed time domain,*

$$\|V^{n+1} - V^n\|_\infty \leqslant \delta \tag{16}$$

for some positive tolerance δ. Then the error in the approximations is uniformly bounded by

$$\|V^n - V(t_n)\|_\infty \leqslant \kappa\delta. \tag{17}$$

Here, κ is a constant only depending on the quotients of lengths of consecutive time steps.

In our example, C is a diagonal matrix with the lengths of the intervals as diagonal entries, which means that

$$\sqrt{h_{\min}}\|v\|_\infty \leqslant \|C^{\frac{1}{2}}v\|_\infty \leqslant \sqrt{h_{\max}}\|v\|_\infty. \tag{18}$$

These inequalities, in combination with Theorem 2.2 and the relation $V^n = C^{\frac{1}{2}}U^n$ can be used to show that if the time steps Δ_n are chosen such that

$$\|U^{n+1} - U^n\|_\infty \leqslant \delta, \tag{19}$$

then, writing $t_n = \sum_1^n \Delta_j$,

$$\|U^n - U(t_n)\|_\infty \leqslant \kappa\delta\sqrt{\frac{h_{\max}}{h_{\min}}}. \tag{20}$$

An algorithm for the practical determination of the time-steps is the following. We assume that it is the U^n that are calculated.

ALGORITHM 2.1: **adaptive time-stepping**

(1) Assemble the matrices A, B, C and vectors F and V^0.
(2) Choose a tolerance $\delta > 0$ and a $\tau_0 > 0$.
(3) **for** $n = 0$ **to** M
Find by trial and error the largest time-step $\Delta_{n+1} := 2^k\tau_n$ such that the calculated U^{n+1} satisfies $\psi_n := \|U^{n+1} - U^n\|_\infty < \delta$.
Set $\tau_{n+1} := \Delta_{n+1}\delta/\psi_n$.
(4) **end**

2.4. Exploiting superconvergence

In the method of lines one aims at a certain balance between the error in the semi-discretization and the additional error made in solving the semi-discretization numerically. Recall (7) and define respectively

$$u_h(t_n) = \sum_{j=1}^N U_j(t_n)w_j \quad \text{and} \quad \hat{u}_h(t_n) = \sum_{j=1}^N U_j^n w_j. \tag{21}$$

So $u_h(t_n)$ is the exact solution of the semi-discretization at time t_n and $\hat{u}_h(t_n)$ its approximation according to the adaptive time-stepping scheme of the previous section. We can use the results from this section to conclude the second inequality in (22) while a simple calculation proves the first.

$$\|u_h(t_n) - \hat{u}_h(t_n)\|_{L^2} \leqslant \|U(t_n) - U^n\|_\infty \leqslant \kappa\delta\sqrt{\frac{h_{\max}}{h_{\min}}}, \tag{22}$$

From [5, 16] we know that the exact solution u_h of the semi-discretisation (5) exhibits superconvergence in the sense that for each t

$$\|P_W u(t) - u_h(t)\|_{L^2} = \mathcal{O}(h^2) \quad \text{while} \quad \|u(t) - u_h(t)\|_{L^2} = \mathcal{O}(h) \tag{23}$$

only. This fact assures that the post-processed function $\pi(u_h)(t) \in V$ which interpolates u_h at the averages of the jumps is a higher order approximation of $u(t)$,

$$\|u(t) - \pi u_h(t)\|_{L^2} = \mathcal{O}(h^2). \tag{24}$$

Having available a higher order approximation $\pi(u_h)(t)$ than the original approximation $u_h(t)$, the error in the latter can be estimated asymptotically exact as follows

$$\text{If } \varepsilon(w) := \|(I-\pi)(w)\|_{L^2} \ (w \in W), \quad \text{then} \quad \frac{\varepsilon(u_h(t))}{\|u(t) - u_h(t)\|_{L^2}} \to 1 \quad (h \to 0). \tag{25}$$

If we make sure that we choose δ initially such that

$$0 < \delta \leqslant q\|u_0 - u_0^h\|_{L^2} \tag{26}$$

for some parameter q that is 'small enough', then the total error at time t_1 will be dominated by the semi-discretization error and the estimation $\varepsilon(\hat{u}_h(t_1))$ will probably still be good. Furthermore, observing that ε_h will typically be variable (it estimates the absolute error, and the solution will vary), we propose to take also δ variable,

$$0 < \delta_{n+1} \leqslant q\varepsilon(t_n). \tag{27}$$

This can be done without extra restrictions on the theory in [15], since each time t can be considered initial time for a diffusion problem with new initial condition. Following this approach we can aim to keep the time-discretization error relatively low, which will prevent that the quality of the space-error estimator $\varepsilon(t)$ will be rapidly lost because of an interfering time-discretization error.

We note that in case the exact error in the elliptic projection of the initial condition is hard to evaluate, one might as well use the error estimator already at $t = 0$, although especially for relatively coarse meshes the exact error is to be preferred. This leads to the following new version of the algorithm (2.1).

ALGORITHM 2.2: **improved adaptive time-stepping**

(1) Assemble the matrices A, B, C and vectors F and U^0.
(2) Choose the parameter q and an initial τ_0.
(2) Set $\delta_0 := q\|u_0 - u_h^0\|_{L^2}$.
(3) **for** $n = 0$ **to** M
    Find by trial and error the largest time-step $\Delta_{n+1} := 2^k\tau_0$ such, that the calculated U^{n+1} satisfies $\|U^{n+1} - U^n\|_\infty \leqslant \delta_n$.
    Set $\delta_n := q\varepsilon(\hat{u}_h(t_{n+1}))$.
    Set $\tau_{n+1} := \Delta_{n+1}\delta_n/\|U^{n+1} - U^n\|_\infty$.
(4) **end**

In Section 3 we will present numerical experiments for this algorithm. First however we will comment on using the space-error estimator for a different goal.

2.5. Adaptivity in space

In the previous section we used a superconvergence-based a posteriori error estimator to regulate the adaptive time-stepping in such a way, that the quality of the space-error estimator was preserved during the time-stepping as much as possible. We will now consider how to use the same estimator to find good grids in space as well.

Consider the following local error indicators. Recall that error *indicators*, as opposed to error *estimators* are not theoretically justified, although heuristicly they make sense.

$$\forall j \in \{1, \ldots, N\}\colon\ \varepsilon_j(w) := \|(I - \pi)w\|_{L^2,I_j}, \quad w \in W. \tag{28}$$

Given some tolerance σ for the relative global space-discretization error, it is in many applications of interest to aim for a uniform distribution of this error over the intervals, which is, to aim for a partitioning into subintervals, such that

$$\forall j \in \{1, \ldots, N\}\colon\ \|u(t) - u_h(t)\|_{L^2,I_j} \leqslant \sigma \mathrm{Vol}(I_j)^{\frac{1}{2}}\|u(t)\|_{L^2}. \tag{29}$$

Indeed, if this criterion is satisfied, we have

$$\|u(t) - u_h(t)\|_{L^2}^2 \leqslant \sum_{j=1}^{N} \sigma^2 \mathrm{Vol}(I_j)\|u(t)\|_{L^2}^2 = \sigma^2\|u(t)\|_{L^2}^2. \tag{30}$$

Of course, the exact errors are in general not available, and neither is the exact solution $u_h(t)$ of the semi-discretization. Therefore, we replace them by the error

indicators (28) and the fully discrete solution $\hat{u}_h(t_n)$ (even post-processed in the right-hand side) and aim for

$$\forall j \in \{1, \dots, N\}\colon\ \varepsilon_j(\hat{u}_h(t_n)) \leqslant \sigma \mathrm{Vol}(I_j)^{\frac{1}{2}} \|\pi \hat{u}_h(t_n)\|_{L^2}. \tag{31}$$

In practice one might want to relax criterion (31) because it can be very hard to realize it for *all* the subintervals. However, in Section 3 we will perform an experiment based on this criterion.

3. Numerical experiments

3.1. Adaptive time stepping with uniform grids

First we consider a numerical experiment in which we use an equidistant partitioning of the unit interval. This avoids the complications that arise as a result of the system matrix in (10) being not symmetric. We test the success of adaptive time stepping and estimating the error in space with respect to some values of the parameter q in (26). The exact solution of a simple heat equation $u_t = u_{xx}$ with homogeneous Neumann boundary conditions was

$$16u(x,t) = \cos(10\pi x)\mathrm{e}^{-100\pi^2 t} + 5\cos(6\pi x)\mathrm{e}^{-36\pi^2 t} + 10\cos(2\pi x)\mathrm{e}^{-4\pi^2 t}. \tag{32}$$

In Figure 1 below the exact solution is given for $t = 0$ (left) and $t = 0.002$ (middle) and $t = 0.01$ (right). For $t = 0$ the function is actually equal to $\cos^5(2\pi x)$.

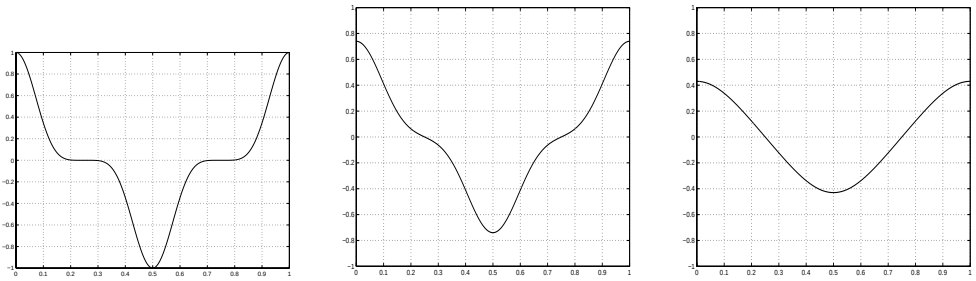


Fig. 1

We took a grid consisting of 40 intervals. This gave a relative space error in the initial function of 0.095. The parameter q was taken $q = 0.5$. So, in each time step, the time-tolerance was set to be half the estimated space-error.

In Figure 2 above we see the relative total error against time (left, ranging from 0.055 to 0.095 initially) and the global efficiency index of the error estimator against time (right, ranging from 1.004 to 1.0085).

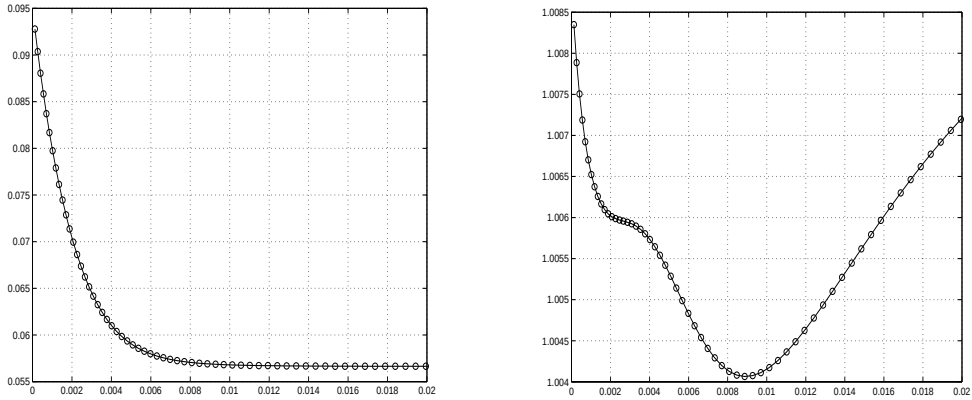


Fig. 2

Clearly, the results are excellent. The global efficiency index is very close to one, which means that both space and time-discretizations are performing very well. We noted that the lengths of the time-steps increased monotoneously.

We will now proceed to investigate the dependence on q of the relative error made in the time-discretization of the semi-discrete system. Our aim was to be able to keep it uniformly bounded as well as relatively small compared to the semi-discretization error.

3.2. Adaptive time-stepping with non-uniform grid

We will now continue to present a test for the algorithm for a non-uniform grid, which, as in the previous experiment, consisted of 40 elements. The exact solution was taken the same as in the previous experiment. To obtain an initial grid, we applied elliptic projection with superconvergence-based adaptivity, i.e., we solved equation (6) adaptively. This gave an initial non-uniform grid, which we kept constant during the time-stepping. In Figure 4 on the left-hand side, we see the approximation in the final time step, and the grid points are indicated by stars.

On the right-hand side, we again plotted for the values of $q = 0.5, 1, 2, 4$ the relative time-error, so the error made by adaptive time stepping with respect to the exact semi-discrete solution. The predicted behaviour is clearly visible in this right picture, although a bit less clear than in the uniform grid case of the previous section. Still, the time errors seem for each q uniformly bounded in the time-interval $[0, 0.01]$ in which the solution changes the most rapidly. Also, the uniform bound seems to be linear in q.

R e m a r k 3.1. Comparing the graphs in Figure 4 with the graphs in Figure 3, we moreover see that they are between one and two times higher in magnitude, which

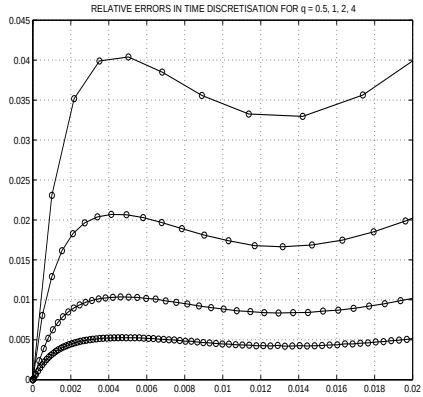
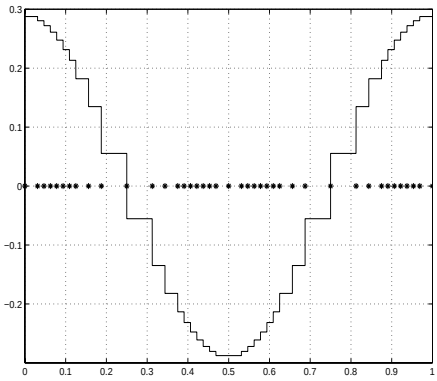


Fig. 3. For each time-step, the relative time-error is given, for $q = 0.5$ (lowest graph), $q = 1, 2$ and $q = 4$ (upper graph). As predicted, the graphs seem uniformly bounded, and their maximum linear in q. The uniform grid consisted of 40 intervals.

was predicted by the factor $\sqrt{\frac{h_{\max}}{h_{\min}}}$ in equation (20). For this specific grid, this factor is equal to two.



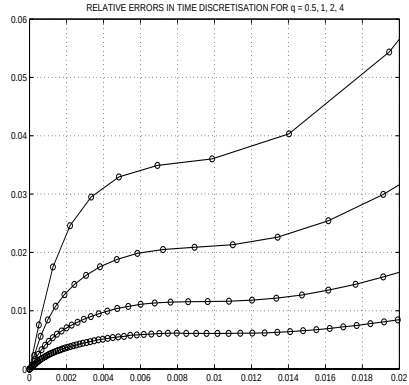


Fig. 4

In Figure 5 below we show at the left-hand side the total relative error at each of the time-steps, which shows that for the values of q considered, the time-error is relatively small compared to the error due to semi-discretization. At the right-hand side we show the performance of the error-estimator, which is in the range of ninety to ninety-three percent.

Since the range in the left picture is 0.06 to 0.095, we see that for $q = 4$ the time-error is of the same magnitude as the relative space-error, while for $q = 0.5$ it is quite smaller.

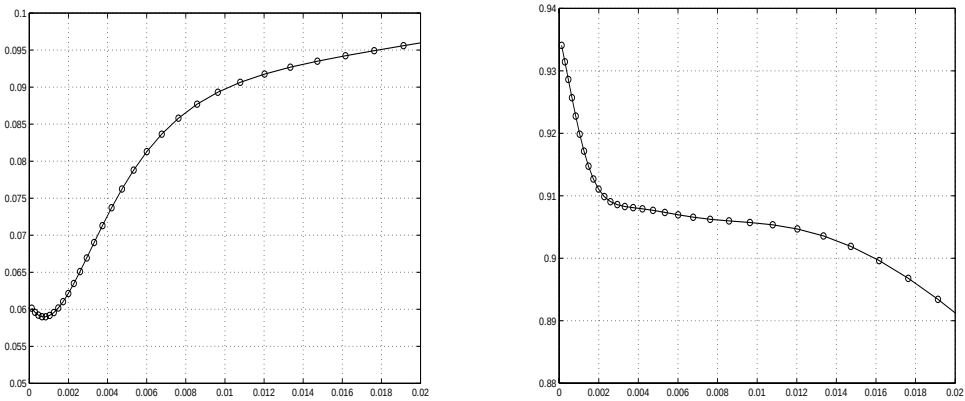


Fig. 5

R e m a r k 3.2. It can be understood that the non-uniform grid is only *initially* better than the uniform grid (with respect to the total relative errors made). Indeed, for larger time the solution changes such that it needs a more uniform grid to be well-approximated. See Figure 1.

3.3. Adaptive refinement during time-stepping

According to the previous two experiments, the ideal situation would be if adaptive grid refinement could be applied in order to follow changes in the shape of the solution. We implemented a very simple adaptive strategy which is based on halving an interval when the error in this interval becomes too large, and removing grid points in case the new solution would still satisfy a criterion based on uniform distribution of the error (Cf. (29)). Applying this strategy to our model problem gave the following results. Note that we did not calculate the relative time-discretization error because of the problem of how to obtain the exact solution (the number of grid points changed!)

On the left we see the relative total error. It jumps up due to the removal of grid points in the area where the solution was initially steep (but later not so steep anymore) and then jumps down again when grid points are added in the area where the initial solution was flat (but later not flat anymore). The effect on the error is good, and also the error estimator, of which we see the efficiency index in the right picture, behaves better than in the previous experiment. It seems that the space-adaptive approach combines the good aspects of both the previous two experiments.

Acknowledgements. The author wishes to thank Jan Zitko for pointing out an error in the first version of this manuscript, and Michal Křížek for his comments and

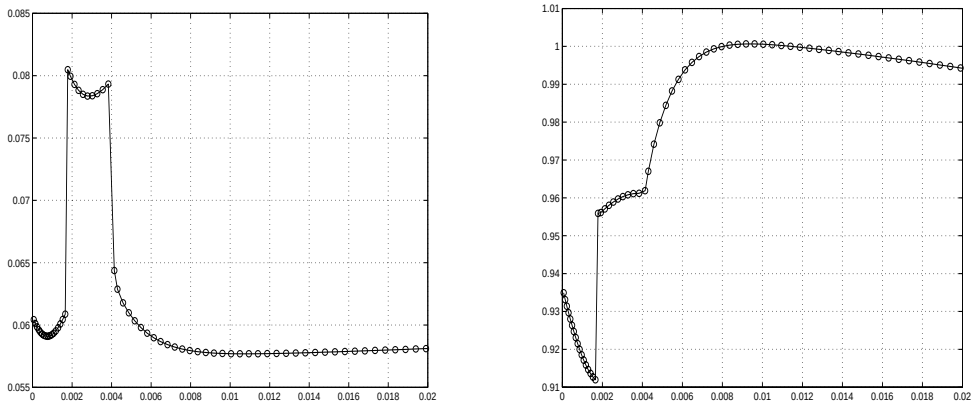


Fig. 6

corrections. Furthermore, the work was supported by grant no. 201/97/0217 of the Grant Agency of the Czech Republic and this support is gratefully acknowledged.

References

[1] *S. Adjerid, J.E. Flaherty, Y.J. Wang*: A posteriori error estimation with finite element methods of lines for one-dimensional parabolic systems. Numer. Math. *65* (1993), 1–21.

[2] *M. Berzins*: Global error estimation in the method of lines for parabolic equations. SIAM J. Sci. Stat. Comput. *9(4)* (1988), 687–703.

[3] *J.H. Brandts*: Superconvergence and a posteriori error estimation for triangular mixed finite elements. Numer. Math. *68(3)* (1994), 311–324.

[4] *J.H. Brandts*: Superconvergence for triangular order $k = 1$ Raviart-Thomas mixed finite elements and for triangular standard quadratic finite element methods. Appl. Numer. Math. (1996), to appear (accepted).

[5] *J.H. Brandts*: Superconvergence of mixed finite element semi-discretizations of two time-dependent problems. Appl. Math. *44(1)* (1999), 43–53.

[6] *J. Douglas, J.E. Roberts*: Global estimates for mixed methods for second order elliptic problems. Math. Comp. *44(169)* (1985), 39–52.

[7] *R. Durán*: Superconvergence for rectangular mixed finite elements. Numer. Math. *58* (1990), 2–15.

[8] *K. Eriksson, C. Johnson*: Adaptive finite element methods for parabolic problems I: A linear model problem. SIAM J. Numer. Anal. *28* (1991), 43–77.

[9] *K. Eriksson, C. Johnson*: Adaptive finite element methods for parabolic problems II: Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$. SIAM J. Numer. Anal. *32* (1995), 706–740.

[10] *K. Eriksson, C. Johnson*: Adaptive finite element methods for parabolic problems III: Time steps variable in space. Manuscript.

[11] *K. Eriksson, C. Johnson*: Adaptive finite element methods for parabolic problems IV: Nonlinear problems. SIAM J. Numer. Anal. *32* (1995), 1729–1749.

[12] *K. Eriksson, C. Johnson*: Adaptive finite element methods for parabolic problems V: Long-time integration. SIAM J. Numer. Anal. *32* (1995), 1750–1763.

[13] *K. Eriksson, C. Johnson, S. Larsson*: Adaptive finite element methods for parabolic problems VI: Analytic semigroups. SIAM J. Numer. Anal. *35(4)* (1998), 1315–1325.

[14] *D. Estep*: A posteriori error bounds and global error control for approximation of ordinary differential equations. SIAM J. Numer. Anal. *32(1)* (1995), 1–48.
[15] *C. Johnson, Y. Nie, V. Thomée*: An a posteriori error estimate and adaptive time step control for a backward Euler discretization of a parabolic problem. SIAM J. Numer. Anal. *27(2)* (1990), 277–291.
[16] *M. Křížek, P. Neittaanmäki, R. Stenberg (eds)*: Finite element methods: superconvergence, post-processing and a posteriori estimates. Proc. Conf. Univ. of Jyväskylä, 1996. Lecture Notes in Pure and Applied Mathematics volume 196, Marcel Dekker, New York, 1998.
[17] *J. Lawson, M. Berzins, P.M. Dew*: Balancing space and time errors in the method of lines for parabolic equations. SIAM J. Sci. Stat. Comput. *12(3)* (1991), 573–594.
[18] *P. Monk*: A comparison of three mixed methods for the time-dependent Maxwell's equations. SIAM J. Sci. Stat. Comput. *13(5)* (1992), 1097–1122.
[19] *P. Monk*: An analysis of Nédélec's method for the spatial discretization of Maxwell's equations. J. Comp. Appl. Math. *47* (1993), 101–121.
[20] *A.K. Pani*: An H^1-Galerkin mixed finite element method for parabolic partial differential equations. SIAM J. Numer. Anal. *35(2)* (1998), 712–727.
[21] *P.A. Raviart, J.M. Thomas*: A mixed finite element method for second order elliptic problems. Lecture Notes in Mathematics 606. 1977, pp. 292–315.
[22] *A.H. Schatz, V. Thomeé, W.L. Wendland (eds)*: Mathematical Theory of Finite and Boundary Element Methods. Birkhäuser Verlag, Basel, 1990.
[23] *V. Thomée*: Galerkin Finite Element Methods for Parabolic Problems. Lecture Notes in Mathematics 1054, Springer Verlag, New York, 1998.

Author's address: *Jan Brandts*, Mathematical Institute, Academy of Sciences, Žitná 25, 115 67 Praha 1, Czech Republic; *present address*: Mathematical Institute, Utrecht University, P. O. Box 80.010, 3508 TA Utrecht, The Netherlands, e-mail: `brandts@math.uu.nl`.