Vít Dolejší; Miloslav Feistauer; Christoph Schwab

On discontinuous Galerkin methods for nonlinear convection-diffusion problems and compressible flow

Mathematica Bohemica, Vol. 127 (2002), No. 2, 163–179

Persistent URL: http://dml.cz/dmlcz/134171

127 (2002) MATHEMATICA BOHEMICA No. 2, 163–179

Proceedings of EQUADIFF 10, Prague, August 27–31, 2001

ON DISCONTINUOUS GALERKIN METHODS FOR NONLINEAR CONVECTION-DIFFUSION PROBLEMS AND COMPRESSIBLE FLOW

V. Dolejší, M. Feistauer, Praha, C. Schwab, Zürich

Abstract. The paper is concerned with the discontinuous Galerkin finite element method for the numerical solution of nonlinear conservation laws and nonlinear convection-diffusion problems with emphasis on applications to the simulation of compressible flows. We discuss two versions of this method: (a) Finite volume discontinuous Galerkin method, which is a generalization of the combined finite volume—finite element method. Its advantage is the use of only one mesh (in contrast to the combined finite volume—finite element schemes). However, it is of the first order only. (b) Pure discontinuous Galerkin finite element method of higher order combined with a technique avoiding spurious oscillations in the vicinity of shock waves.

Keywords: discontinuous Galerkin finite element method, numerical flux, conservation laws, convection-diffusion problems, limiting of order of accuracy, numerical solution of compressible Euler equations

MSC 2000: 65M15, 76M10, 76M12

Introduction

Our goal is to develop a sufficiently accurate and robust method for the numerical solution of nonlinear conservation laws, nonlinear convection-diffusion problems and compressible flows. In principle, all numerical methods for the solution of partial differential equations can be applied to the problems mentioned. The most popular ones are now the finite element (FE) and finite volume (FV) methods. The finite volume schemes are suitable for the discretization of conservation laws, whereas the FE methods are mainly used for diffusion problems. In order to exploit advantages

This work was supported under the grant No. 101/01/0938 of the Czech Grant Agency and by the Grant MSM 113200007. The authors acknowledge this support.

of both these techniques, combined FV-FE methods for the solution of convection-diffusion problems and compressible viscous flows were developed. For analysis and applications see, e.g., [1], [10]–[14]. These methods give good results in many cases of technically relevant problems in complex domains. However, their drawback is the necessity to construct two mutually associated meshes, which is rather complicated particularly in 3D ([15]).

A generalization of both the FV and FE methods is the discontinuous Galerkin finite element (DG FE) method. It uses only one mesh and allows higher order of accuracy. (For a survey about DG FE methods see [3] or [4].) However, in regions where the solution has discontinuities or steep gradients, the so-called spurious oscillations appear in the numerical solution obtained by the DG FE method. In this paper we describe two methods how to avoid this undesirable phenomenon. The first possibility uses an FV approximation of the convective terms applied in the framework of the DG FE method via averaging. This method requires only one mesh, but its order of accuracy is one. The second method is based on a new type of limiting of the order of accuracy in the vicinity of discontinuities or steep gradients. In contrast to [3], where the author introduces a slope limiter quite in analogy with the FV MUSCL type schemes, we propose a different new method based on a suitable identification of a discontinuity and the decrease of the order of the method to one in a narrow neighbourhood of the discontinuity. This numerical technique is applied to the solution of the inviscid compressible high-speed flow described by the Euler equations.

1. DG FE method for a nonstationary nonlinear convection-diffusion problem

1.1. Continuous problem. Let $\Omega \subset \mathbb{R}^2$ be a bounded polygonal domain and $T > 0$. We set $Q_T = \Omega \times (0,T)$ and denote by $\partial\Omega$ the boundary of Ω. We consider the following model *initial-boundary value problem*: Find $u\colon Q_T \to \mathbb{R}$ such that

$$\frac{\partial u}{\partial t} + \sum_{s=1}^{2} \frac{\partial f_s(u)}{\partial x_s} = \varepsilon \Delta u + g \quad \text{in } Q_T, \tag{1.1}$$

$$u|_{\partial\Omega\times(0,T)} = u_D, \tag{1.2}$$

$$u(x,0) = u^0(x), \quad x \in \Omega. \tag{1.3}$$

We suppose that $f_s \in C^1(\mathbb{R})$ and the data are sufficiently regular so that we can assume the existence of a *strong solution* u satisfying (1.1)–(1.3) pointwise (almost everywhere):

$$u \in L^2(0,T;H^2(\Omega)), \quad \partial u/\partial t \in L^2(0,T;H^1(\Omega)). \tag{1.4}$$

We use the standard notation for function spaces: $H^k(\Omega)$ = Sobolev space, $L^2(0,T; X)$ = Bochner space of square integrable functions on $(0,T)$ with values in a Banach space X, $C^1(0,T;X)$ = space of continuously differentiable mappings in $(0,T)$ with values in X.

1.2. Discretization. Let $\mathcal{T}_h$ $(h > 0)$ denote a partition of the closure $\overline{\Omega}$ of the domain Ω into a finite number of closed convex polygons K with mutually disjoint interiors. We call $\mathcal{T}_h$ a triangulation of $\overline{\Omega}$, but do not require the usual conforming properties from the finite element method. We usually choose $K \in \mathcal{T}_h$ as triangles or quadrilaterals but we can allow even more general convex elements.

We set $h_K = \operatorname{diam}(K)$, $h = \max_{K\in\mathcal{T}_h} h_K$, ϱ_K = radius of the largest ball inscribed into K. All elements of $\mathcal{T}_h$ will be numbered so that $\mathcal{T}_h = \{K_i\}_{i\in I}$, where $I \subset Z^+ = \{1,2,\ldots\}$ is a suitable index set. If two elements K_i, $K_j \in \mathcal{T}_h$ contain a nonempty open face which is a part of a straight line, we call them *neighbours*. In this case we set $\Gamma_{ij} = \partial K_i \cap \partial K_j$ and assume that the whole set Γ_{ij} is a part of a straight line. For $i \in I$ we set $s(i) = \{j \in I; K_j \text{ is a neighbour of } K_i\}$.

The boundary $\partial\Omega$ is formed by a finite number of faces of elements K_i adjacent to $\partial\Omega$. We denote all these boundary faces by S_j, where $j \in I_b \subset Z^- = \{-1,-2,\ldots\}$, and set $\gamma(i) = \{j \in I_b; S_j \text{ is a face of } K_i\}, \Gamma_{ij} = S_j$ for $K_i \in \mathcal{T}_h$ such that $S_j \subset \partial K_i$, $j \in I_b$. For K_i not containing any boundary face S_j we set $\gamma(i) = \emptyset$. Obviously, $s(i) \cap \gamma(i) = \emptyset$ for all $i \in I$. Now, if we write $S(i) = s(i) \cup \gamma(i)$, we have

$$\partial K_i = \bigcup_{j\in S(i)} \Gamma_{ij}, \qquad \partial K_i \cap \partial\Omega = \bigcup_{j\in\gamma(i)} \Gamma_{ij}. \tag{1.5}$$

Furthermore, we use the following notation: $\boldsymbol{n}_{ij} = ((n_{ij})_1, (n_{ij})_2)$ = unit outer normal to ∂K_i on the face Γ_{ij}, $|\Gamma_{ij}|$=length of Γ_{ij}. By $|K|$ we denote the two-dimensional Lebesgue measure of $K \in \mathcal{T}_h$.

Over the triangulation $\mathcal{T}_h$ we define the *broken Sobolev space*

$$H^k(\Omega,\mathcal{T}_h) = \{v; v|_K \in H^k(K)\ \forall\, K \in \mathcal{T}_h\} \tag{1.6}$$

and for $v \in H^1(\Omega,\mathcal{T}_h)$ we introduce the following notation: $v|_{\Gamma_{ij}}$ = the trace of $v|_{K_i}$ on $\Gamma_{ij}, v|_{\Gamma_{ji}}$ = the trace of $v|_{K_j}$on $\Gamma_{ji} = \Gamma_{ij}, \langle v\rangle_{\Gamma_{ij}} = \frac{1}{2}\left(v|_{\Gamma_{ij}} + v|_{\Gamma_{ji}}\right)$, $[v]_{\Gamma_{ij}} = v|_{\Gamma_{ij}} - v|_{\Gamma_{ji}}$. Obviously, $\langle v\rangle_{\Gamma_{ij}} = \langle v\rangle_{\Gamma_{ji}}$ but $[v]_{\Gamma_{ij}} = -[v]_{\Gamma_{ji}}$ and $[v]_{\Gamma_{ij}}\boldsymbol{n}_{ij} = [v]_{\Gamma_{ji}}\boldsymbol{n}_{ji}$.

The approximate solution of problem (1.1)–(1.3) is sought in the space of discontinuous piecewise polynomial functions

$$S_h = S^{p,-1}(\Omega,\mathcal{T}_h) = \{v; v|_K \in P_p(K)\ \forall\, K \in \mathcal{T}_h\}, \tag{1.7}$$

where $P_p(K)$ denotes the space of all polynomials on K of degree $\leqslant p$.

In order to derive the discrete problem, we multiply equation (1.1) by any $v \in S_h$, integrate over $K \in \mathcal{T}_h$, apply Green's theorem and sum over all $K \in \mathcal{T}_h$. Moreover, we use the relations $[u]|_{\Gamma_{ij}} = 0$, $\langle \nabla u \rangle_{\Gamma_{ij}} = \nabla u|_{\Gamma_{ij}} = \nabla u|_{\Gamma_{ji}}$ and add to the identity thus obtained some terms which mutually cancel.

The flux $\int_{\Gamma_{ij}} f_s(u)\, n_s\, v \,\mathrm{d}S$ is approximated with the aid of the *numerical flux* $H = H(\alpha, \beta, \boldsymbol{n})$:

$$\int_{\Gamma_{ij}} \sum_{s=1}^{2} f_s(u)\, n_s\, v|_{\Gamma_{ij}} \,\mathrm{d}S \approx \int_{\Gamma_{ij}} H\left(u|_{\Gamma_{ij}}, u|_{\Gamma_{ji}}, \boldsymbol{n}_{ij}\right) v|_{\Gamma_{ij}} \,\mathrm{d}S \tag{1.8}$$

and the approximate convective form is defined as

$$\begin{aligned} \tilde{b}_h(u_h, v_h) = &\sum_{i\in I} \sum_{j \in S(i)} \int_{\Gamma_{ij}} H\left(u_h|_{\Gamma_{ij}}, u_h|_{\Gamma_{ji}}, \boldsymbol{n}_{ij}\right) v_h|_{\Gamma_{ij}} \,\mathrm{d}S \\ &- \sum_{i\in I} \int_{K_i} \sum_{s=1}^{2} f_s(u_h) \frac{\partial v_h}{\partial x_s} \,\mathrm{d}x, \quad u_h,\, v_h \in S_h. \end{aligned} \tag{1.9}$$

If $\Gamma_{ij} \subset \partial\Omega$, we use the Dirichlet boundary condition (1.2) in order to specify $u_h|_{\Gamma_{ji}}$.

Now, for $u_h, v_h \in S_h$ we set

$$\begin{aligned} a_h(u_h, v_h) = &\,\varepsilon \sum_{i\in I} \int_{K_i} \nabla u_h \cdot \nabla v_h \,\mathrm{d}x \\ &- \varepsilon \sum_{i\in I} \sum_{\substack{j\in s(i)\\ j<i}} \int_{\Gamma_{ij}} \langle \nabla u_h \rangle \cdot \boldsymbol{n}_{ij}\, [v_h] \,\mathrm{d}S \\ &+ \varepsilon \sum_{i\in I} \sum_{\substack{j\in s(i)\\ j<i}} \int_{\Gamma_{ij}} \langle \nabla v_h \rangle \cdot \boldsymbol{n}_{ij}\, [u_h] \,\mathrm{d}S \\ &- \varepsilon \sum_{i\in I} \sum_{j\in \gamma(i)} \int_{\Gamma_{ij}} \nabla u_h \cdot \boldsymbol{n}_{ij}\, v_h \,\mathrm{d}S \\ &+ \varepsilon \sum_{i\in I} \sum_{j\in \gamma(i)} \int_{\Gamma_{ij}} \nabla v_h \cdot \boldsymbol{n}_{ij}\, u_h \,\mathrm{d}S \\ &\text{(diffusion terms)}, \end{aligned} \tag{1.10}$$

$$\begin{aligned} J_h^\sigma(u_h, v_h) = &\sum_{i\in I} \sum_{j\in s(i)} \int_{\Gamma_{ij}} \sigma [u_h]\, [v_h] \,\mathrm{d}S + \sum_{i\in I} \sum_{j\in \gamma(i)} \int_{\Gamma_{ij}} \sigma\, u_h\, v_h \,\mathrm{d}S \\ &\text{(stabilization jump terms)}, \end{aligned} \tag{1.11}$$

$$(1.12)\qquad \ell_h(v_h)\,(t) = \int_\Omega g(t)\,v_h\,\mathrm{d}x + \varepsilon \sum_{i\in I}\sum_{j\in\gamma(i)} \int_{\Gamma_{ij}} \nabla v_h \cdot \boldsymbol{n}_{ij} u_D(t)\,\mathrm{d}S$$
$$+ \varepsilon \sum_{i\in I}\sum_{j\in\gamma(i)} \int_{\Gamma_{ij}} \sigma\, u_D(t)\, v_h \,\mathrm{d}S,$$
$$(1.13)\qquad (\alpha,\beta) = \int_\Omega \alpha\beta\,\mathrm{d}x.$$

Here σ is a weight function such that $\sigma|_{\Gamma_{ij}} = 1/|\Gamma_{ij}|$. An *approximate solution* is defined as a function u_h satisfying the conditions

$$(1.14)\qquad \begin{aligned} &\text{(a)}\quad u_h \in C^1([0,T], S_h),\\ &\text{(b)}\quad \left(\frac{\partial u_h(t)}{\partial t}, v_h\right) + \tilde{b}_h(u_h(t), v_h) + a_h(u_h(t), v_h) + \varepsilon J_h^\sigma(u_h(t), v_h)\\ &\qquad\qquad = \ell_h(v_h)(t), \qquad \forall\, v_h \in S_h \quad \forall\, t \in (0,T),\\ &\text{(c)}\quad u_h(0) = u_h^0, \end{aligned}$$

where u_h^0 is an S_h-approximation of u^0 (e.g., L^2-projection). If $\varepsilon = 0$, it is necessary to use boundary conditions suitable for hyperbolic equations. (See Section 5.)

We have carried out the semidiscretization in space (called the method of lines) leading to a system of ordinary differential equations. In practical computations, the full discretization is carried out. We can use, e.g., the explicit Euler or Runge-Kutta schemes. Semiimplicit or fully implicit time discretization leads to large nonlinear algebraic systems which must be solved iteratively. Moreover, the integrals are evaluated with the aid of quadrature formulae. Let us note that the form a_h is a variant of the DG FE approximation of the diffusion terms proposed in [18].

We assume that the numerical flux has the following properties:

(1) $H(\alpha,\beta,\boldsymbol{n})$ is defined and (locally) Lipschitz-continuous on $\mathbb{R}^2 \times B_1$, where $B_1 = \{\boldsymbol{n} \in \mathbb{R}^2; |\boldsymbol{n}| = 1\}$

(2) $H(\alpha,\beta,\boldsymbol{n})$ is *consistent*:

$$(1.15)\qquad H(\alpha,\alpha,\boldsymbol{n}) = \sum_{s=1}^{2} f_s(\alpha)\,n_s, \quad \alpha\in\mathbb{R},\ \boldsymbol{n} = (n_1,n_2) \in B_1,$$

(3) $H(\alpha,\beta,\boldsymbol{n})$ is *conservative*:

$$(1.16)\qquad H(\alpha,\beta,\boldsymbol{n}) = -H(\beta,\alpha,-\boldsymbol{n}), \quad \alpha,\,\beta\in\mathbb{R},\ \boldsymbol{n}\in B_1.$$

The above described process yields a higher order scheme using only one (in general unstructured) mesh. Its disadvantage are spurious oscillations in approximate solutions which appear in areas with steep gradients in the case of small diffusion terms (or discontinuities if $\varepsilon = 0$). We will discuss two methods how to avoid this problem.

2. Finite volume discontinuous Galerkin method

The first method for avoiding spurious oscillations in the DG FE solution is based on a modification of the convective form with the aid of the FV approach and element averaging. Therefore, we speak of the finite volume discontinuous Galerkin method (FV DG).

In (1.7) we put $p = 1$, i.e., we use piecewise linear elements, and introduce a modification b_h of the form $\tilde{b}_h$ defined in the following way. By π_0 we denote the L^2-projection of functions $v \in L^2(\Omega)$ to the space $\mathcal{S}_h^{0,-1}(\Omega, \mathcal{T}_h)$ of piecewise constant functions: $\pi_0 v|_K = \int_K v \,\mathrm{d}x/|K|$ for $K \in \mathcal{T}_h$. Then, instead of $\tilde{b}_h$ we use in (1.14), (b) the approximate convective form

$$
\begin{aligned}
b_h(u_h, v_h) = \sum_{i\in I} \pi_0 v_h|_{K_i} \Big\{ &\sum_{j\in s(i)} H\left(\pi_0 u_h|_{K_i}, \pi_0 u_h|_{K_j}, \boldsymbol{n}_{ij}\right) |\Gamma_{ij}| \\
&+ \sum_{j\in\gamma(i)} H\left(\pi_0 u_h|_{K_i}, \pi_0 u_h|_{K_i}, \boldsymbol{n}_{ij}\right) |\Gamma_{ij}| \Big\}.
\end{aligned}
\tag{2.1}
$$

In this case, the boundary values are realized in the form b_h by extrapolation.

In order to derive *error estimates*, we introduce the following *assumptions*:

Let us consider a system $\{\mathcal{T}_h\}_{h\in(0,h_0)}$, $h_0 > 0$, of partitions of the domain Ω ($\mathcal{T}_h = \{K_i\}_{i\in I_h}$, $I_h \subset Z^+$, but for simplicity we write again I instead of I_h) and assume that it has the following properties:

(A1) There exists a constant $C_1 > 0$ such that

$$h_K/\varrho_K \leqslant C_1 \quad \forall\, K \in \mathcal{T}_h,\ \forall\, h \in (0, h_0). \tag{2.2}$$

(We say that the system $\{\mathcal{T}_h\}_{h\in(0,h_0)}$ is *shape regular*.)

(A2) There exists a constant C_2 such that

$$\operatorname{card} S(i) \leqslant C_2 \quad \forall\, K_i \in \mathcal{T}_h,\ \forall\, h \in (0, h_0). \tag{2.3}$$

(The number of neighbours K_j of K_i is uniformly bounded.)

(A3) There exists a constant $C_3 > 0$ such that

$$h_{K_i} \leqslant C_3 |\Gamma_{ij}|, \quad i \in I,\ j \in S(i),\ h \in (0, h_0). \tag{2.4}$$

(The length of faces between neighbouring elements does not degenerate.)

(A4) The numerical flux H is Lipschitz-continuous.

There is not enough space here for the complete derivation of error estimates, but we shall mention without proofs at least main steps and results in order to give an idea of the whole argumentation. Under the above assumptions, the following auxiliary results can be established.

Lemma 2.1 (Multiplicative trace inequality). *There exists a constant $C_4 > 0$ independent of h, K such that*

$$\|v\|^2_{L^2(\partial K)} \leqslant C_4(\|v\|_{L^2(K)}\,|v|_{H^1(K)} + h_K^{-1}\|v\|^2_{L^2(K)}), \quad K \in \mathcal{T}_h,\ v \in H^1(K),\ h \in (0, h_0). \tag{2.5}$$

Lemma 2.2. *The following estimates hold:*

$$\|\pi_0 v\|_{L^2(K)} \leqslant \|v\|_{L^2(K)}, \quad K \in \mathcal{T}_h,\ v \in L^2(K), \tag{2.6}$$

$$\|\pi_0 v\|_{L^2(\Omega)} \leqslant \|v\|_{L^2(\Omega)}, \quad v \in L^2(\Omega), \tag{2.7}$$

$$\|v - \pi_0 v\|_{L^2(K)} \leqslant \frac{h_K}{\pi}|v|_{H^1(K)}, \quad K \in \mathcal{T}_h,\ v \in H^1(K), \tag{2.8}$$

$$\|v - \pi_0 v\|_{L^2(\partial K)} \leqslant C_5\, h_K^{1/2}|v|_{H^1(K)}, \quad K \in \mathcal{T}_h,\ v \in H^1(K), \tag{2.9}$$

with a constant $C_5 > 0$ independent of $h \in (0, h_0)$ and K, v.

There exist a constant $C_6 > 0$ independent of $h \in (0, h_0)$ and v and a mapping $\Pi\colon\ H^1(\Omega, \mathcal{T}_h) \to S_h$ *such that*

$$\begin{aligned} \|\Pi v - v\|_{L^2(K)} &\leqslant C_6 h_K |v|_{H^1(K)}, \quad v \in H^1(K), \\ \|\Pi v - v\|_{L^2(K)} &\leqslant C_6 h_K^2 |v|_{H^2(K)}, \quad v \in H^2(K), \\ |\Pi v - v|_{H^1(K)} &\leqslant C_6 h_K |v|_{H^2(K)}, \quad v \in H^2(K),\ K \in \mathcal{T}_h,\ h \in (0, h_0), \end{aligned} \tag{2.10}$$

$$\begin{aligned} \|\Pi v - v\|_{L^2(\Omega)} &\leqslant C_6 h |v|_{H^1(\Omega, \mathcal{T}_h)}, \quad v \in H^1(\Omega, \mathcal{T}_h), \\ \|\Pi v - v\|_{L^2(\Omega)} &\leqslant C_6 h^2 |v|_{H^2(\Omega, \mathcal{T}_h)}, \quad v \in H^2(\Omega, \mathcal{T}_h), \\ |\Pi v - v|_{H^1(\Omega, \mathcal{T}_h)} &\leqslant C_6 h |v|_{H^2(\Omega, \mathcal{T}_h)}, \quad v \in H^2(\Omega, \mathcal{T}_h). \end{aligned} \tag{2.11}$$

Now we shall be concerned with properties of the form b_h.

Lemma 2.3. *The form b_h is Lipschitz continuous: There exists a constant $C_7 > 0$ such that*

$$|b_h(u_h, v_h) - b_h(u, v_h)| \leqslant C_7(J_h^\sigma(v_h, v_h)^{1/2} + |v_h|_{H^1(\Omega, \mathcal{T}_h)})\,\|u - u_h\|_{L^2(\Omega)}, \quad u_h \in S_h,\ u \in L^2(\Omega),\ v_h \in S_h,\ h \in (0, h_0). \tag{2.12}$$

Moreover, b_h is consistent: There exists a constant $C_8 > 0$ such that

$$\begin{aligned}|b_h(u, v_h) - b(u, v_h)| &\leqslant C_8\, h|u|_{H^1(\Omega)}\, (J_h^\sigma(v_h, v_h)^{1/2} \\ &\qquad + |v_h|_{H^1(\Omega,\mathcal{T}_h)}(\Phi(\|u\|_{L^\infty(\Omega)}) + 1)), \\ u \in H^1(\Omega) \cap L^\infty(\Omega),&\ v_h \in S_h, \quad h \in (0, h_0),\end{aligned} \tag{2.13}$$

where we define, for $M \geqslant 0$,

$$\Phi(M) = \max_{\substack{\xi\in[-M,M]\\ s=1,2}} |f_s'(\xi)| \tag{2.14}$$

and

$$b(u, v_h) = \int_\Omega \sum_{s=1}^{2} \frac{\partial f_s(u)}{\partial x_s} v_h \, \mathrm{d}x \tag{2.15}$$

is a weak form of the convective terms from the continuous problem.

Let us assume that the exact solution satisfies conditions (1.4). Then it satisfies the relation

$$\left(\frac{\partial u}{\partial t}, v_h\right) + a_h(u, v_h) + \varepsilon J_h^\sigma(u, v_h) + b(u, v_h) = \ell(v_h) \quad \forall v_h \in S_h. \tag{2.16}$$

We set

$$M = \|u\|_{L^\infty(Q_T)} \tag{2.17}$$

and

$$\xi = u_h - \Pi u, \quad \eta = \Pi u - u. \tag{2.18}$$

Then

$$u_h - u = \xi + \eta, \quad \xi(t) \in S_h, \quad \eta(t) \in H^2(\Omega, \mathcal{T}_h), \quad t \in [0, T]. \tag{2.19}$$

From the numerical scheme (1.14), where we write b_h instead of $\tilde{b}_h$, and identity (2.16), it is possible to derive the relation

$$\begin{aligned}&\left(\frac{\partial \xi}{\partial t}, \xi\right) + a_h(\xi, \xi) + \varepsilon J_h^\sigma(\xi, \xi) \\ &\qquad = b(u, \xi) - b_h(u_h, \xi) - \left(\frac{\partial \eta}{\partial t}, \xi\right) - a_h(\eta, \xi) - \varepsilon J_h^\sigma(\eta, \xi).\end{aligned} \tag{2.20}$$

In the sequel, we estimate the individual terms on the right-hand side of (2.20) and get the following results:

Lemma 2.4. *We have*

$$2\left(\frac{\partial \xi}{\partial t}, \xi\right) = \frac{\mathrm{d}}{\mathrm{d}t}\|\xi\|^2_{L^2(\Omega)}, \tag{2.21}$$

$$\left|\left(\frac{\partial \eta}{\partial t}, \xi\right)\right| \leqslant \left\|\frac{\partial \eta}{\partial t}\right\|_{L^2(\Omega)} \|\xi\|_{L^2(\Omega)}, \tag{2.22}$$

$$J_h^\sigma(\eta, \xi) \leqslant (J_h^\sigma(\eta, \eta))^{1/2}\,(J_h^\sigma(\xi, \xi))^{1/2}, \tag{2.23}$$

$$\begin{aligned} \|\eta\|_{L^2(\Omega)} &\leqslant C_9 h |u|_{H^1(\Omega)}, \\ \|\eta\|_{L^2(\Omega)} &\leqslant C_9 h^2 |u|_{H^2(\Omega)}, \\ |\eta|_{H^1(\Omega, \mathcal{T}_h)} &\leqslant C_9 h |u|_{H^2(\Omega)}, \\ \left\|\frac{\partial \eta}{\partial t}\right\|_{L^2(\Omega)} &\leqslant C_9 h \left|\frac{\partial u}{\partial t}\right|_{H^1(\Omega)}, \quad h \in (0, h_0), \end{aligned} \tag{2.24}$$

where $C_9 > 0$ is a constant independent of u and h.

Lemma 2.5. *We have*

$$\begin{aligned} |b(u, \xi) - b_h(u_h, \xi)| \leqslant C_{10}&(J_h^\sigma(\xi, \xi)^{1/2} + |\xi|_{H^1(\Omega, \mathcal{T}_h)}) \\ &\times (\|\xi\|_{L^2(\Omega)} + h^2 |u|_{H^2(\Omega)} + h|u|_{H^1(\Omega)}), \quad h \in (0, h_0), \end{aligned} \tag{2.25}$$

where $C_{10} > 0$ is a constant dependent on $M = \|u\|_{L^\infty(Q_T)}$ but independent of h and ξ. Moreover, there exists a constant $C_{11} > 0$ independent of u, h, ξ, ε such that

$$|a_h(\eta, \xi)| \leqslant C_{11} \varepsilon h |u|_{H^2(\Omega)} (J_h^\sigma(\xi, \xi)^{1/2} + |\xi|_{H^1(\Omega, \mathcal{T}_h)}), \quad h \in (0, h_0). \tag{2.26}$$

The last step in the proof of the error estimates is the application of Gronwall's lemma. This yields the m a i n r e s u l t.

Theorem 2.1. *Let assumptions* (1.15), (1.16) *and* (A1)–(A4) *be satisfied. Let u be the exact strong solution of problem* (1.1)–(1.3) *satisfying* (1.4) *and let u_h be the approximate solution defined by the FV DG modification of scheme* (1.14), *described in Section* 2. *Then the error $e_h = u_h - u$ satisfies the estimate*

$$\begin{aligned} &\sup_{t \in [0,T]} \|e_h(t)\|^2_{L^2(\Omega)} + \varepsilon \int_0^T \Big(|e_h(\vartheta)|^2_{H^1(\Omega, \mathcal{T}_h)} + J_h^\sigma(e_h(\vartheta), e_h(\vartheta))\Big)\,\mathrm{d}\vartheta \\ &\leqslant C h^2, \quad h \in (0, h_0), \end{aligned} \tag{2.27}$$

with a constant $C > 0$ independent of h.

All proofs are rather technical. They will appear in a separate paper [6].

As we see from (2.27), the FV DG method is of the first order. This was also confirmed by numerical experiments described in [6]. Unfortunately, the constant C from the above estimate depends on $\varepsilon \to 0$ in a very pessimistic way: $C \approx \exp(c/\varepsilon)$ (c is a constant independent of h and ε). This is caused by the application of Gronwall's lemma. A uniform estimate for $\varepsilon \to 0$ remains open.

The same results can be obtained for a three-dimensional problem and a problem in $\Omega \times (0,T)$ with $\Omega = (-1,1)^d$ ($d = 2,3$) and periodic boundary conditions.

3. Second-order DG FE method with order limiting

This section is concerned with a numerical technique avoiding disadvantages of both schemes discussed above: spurious oscillations in solutions obtained by the pure DG FE method (1.14) and a low order (= first order) of the FV DG scheme from Section 2.

Let us return to scheme (1.14), where we suppose that $\mathcal{T}_h$ is formed by triangles and $p = 1$. We carry out the discretization in time by the forward Euler method. To this end, we consider a partition $0 = t_0 < t_1 < t_2 < \ldots$ of the time interval $(0,T)$ and set $\tau_k = t_{k+1} - t_k$.

The *fully discrete problem* reads: starting from $u_h^0 \in S_h$, for each $k \geqslant 0$ find u_h^{k+1} such that

$$\begin{aligned} &\text{(a)} \quad u_h^{k+1} \in S_h = S_h^{1,-1}(\Omega, \mathcal{T}_h), \\ &\text{(b)} \quad (u_h^{k+1}, v_h) = (u_h^k, v_h) - \tau_k\, a_h(u_h^k, v_h) \\ &\qquad - \tau_k\, \tilde{b}_h(u_h^k, v_h) - \tau_k \varepsilon J_h^\sigma(u_h^k, v_h) + \tau_k\, \ell_h(v_h)\,(t_k) \quad \forall\, v_h \in S_h. \end{aligned} \tag{3.1}$$

In order to avoid spurious oscillations in the numerical solution, discontinuities and steep gradients of the solution are identified, and in their vicinity, the order of accuracy of the scheme is suppressed to one. On the basis of detailed numerical experiments ([7]), the following indicator of discontinuities and steep gradients has been proposed:

$$g(i) = \int_{\partial K_i} [u_h^k]^2 \,\mathrm{d}S / (h_{K_i} |K_i|^{3/4}), \quad K_i \in \mathcal{T}_h. \tag{3.2}$$

Now we define an *adaptive strategy* for an *automatic limiting* of the order of accuracy of scheme (3.1):

$$\begin{aligned} &\text{(a)} \quad u_h^{k+1} \in S_h = S_h^{1,-1}(\Omega, \mathcal{T}_h), \\ &\text{(b)} \quad (u_h^{k+1}, v_h) = (\tilde{u}_h^k, v_h) - \tau_k\, a_h(u_h^k, v_h) \\ &\qquad - \tau_k\, \tilde{b}_h(\tilde{u}_h^k, v_h) - \tau_k J_h^\sigma(u_h^k, v_h) + \tau_k\, \ell_h(v_h)\,(t_k) \quad \forall\, v_h \in S_h, \end{aligned} \tag{3.3}$$

where $\tilde{u}_h^k$ is the modification of u_h^k defined with the aid of our limiting strategy in the following way:

$$(3.4)\qquad \begin{aligned} &\text{(a)}\quad \text{Set}\quad \tilde{u}_h^k|_{K_i} := u_h^k|_{K_i} \quad \forall i \in I.\\ &\text{(b)}\quad \text{If } g(i) > 1 \text{ for some } i \in I, \text{ then set } \tilde{u}_h^k|_{K_i} := \pi_0\, u_h^k|_{K_i}. \end{aligned}$$

This means that in (3.3) the limiting (3.4) of the order of the scheme is applied to the elements lying on discontinuities (or regions with steep gradients). In other areas the second order of accuracy is preserved.

The next sections demonstrate the applicability of the schemes just developed.

4. Scalar numerical examples

Let us consider the Burgers equation

$$(4.5)\qquad \frac{\partial u}{\partial t} + \sum_{s=1}^{2} u \frac{\partial u}{\partial x_s} = \varepsilon \Delta u \quad \text{in } \Omega \times (0, T),$$

where $\Omega = (-1, 1)^2$, equipped with the initial condition

$$(4.6)\qquad u^0(x_1, x_2) = 0.25 + 0.5 \sin(\pi(x_1 + x_2)), (x_1, x_2) \in \Omega,$$

and periodic boundary conditions. This problem has a unique weak solution converging to a weak entropy solution of the inviscid Burgers equation as $\varepsilon \to 0+$. If $\varepsilon = 0$, the solution is discontinuous for $t \geqslant 0.3$. For $0 < \varepsilon \ll 1$, the solution has steep gradients (tending to discontinuities as $\varepsilon \to 0$).

This problem is solved by the numerical scheme (3.1) (adapted to the problem with periodic conditions). The numerical flux is chosen in the following way:

In Figure 1, the computational mesh used in Ω is plotted. The time step is chosen to be $\tau = 2.5 \cdot 10^{-4}$. Figure 2 shows the graph of the numerical solution at time $t = 0.45$ for $\varepsilon = 0$. It is seen here that the solution contains spurious oscillations near discontinuities. In Figure 3 we see the numerical solution of the problem with $\varepsilon = 0.002$ obtained by the FV DG method described in Section 2. In this case the exact solution differs only slightly from the solution of the problem with $\varepsilon = 0$. We see in Figure 3 that the oscillations are strongly suppressed. The best results were obtained with the aid of the method (3.3)–(3.4), as is seen from Figure 4 showing the numerical solution of the problem with $\varepsilon = 0$. In this case, the discontinuities are resolved very well. They are quite sharp without spurious oscillations.

In all computational results presented, we can notice an interesting fact that although the discontinuous approximation is used, the interelement jumps are negligible in the regions where the exact solution is regular. Conspicuous discontinuities

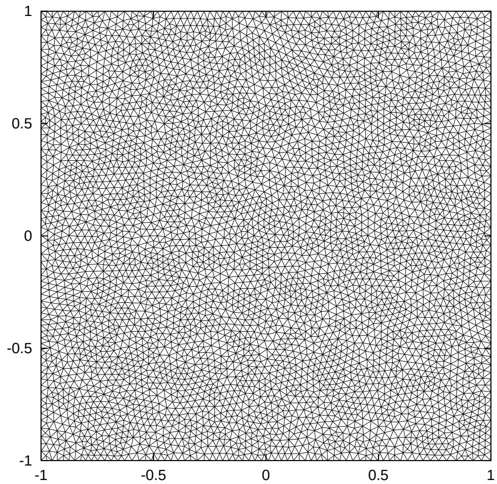

Figure 1. Triangulation used for the numerical solution of problem (4.5)–(4.6)

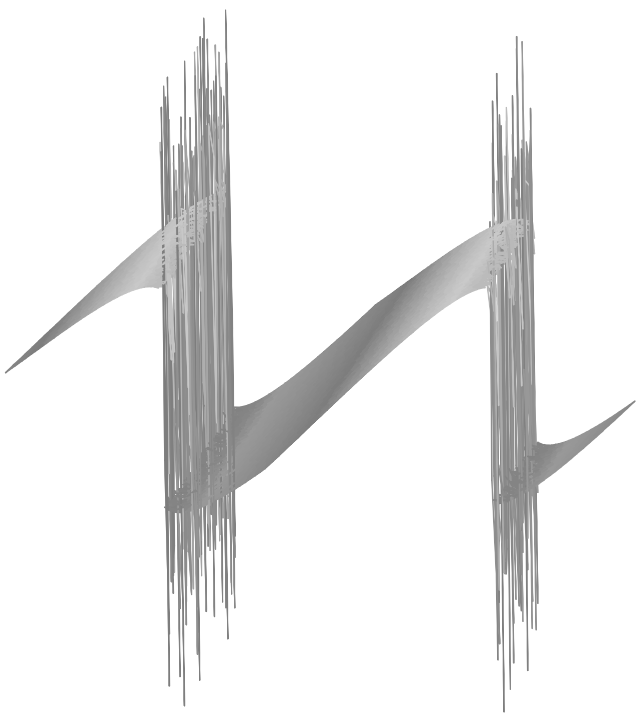

Figure 2. Numerical solution of problem (4.5)–(4.6) computed by DG FE method plotted at $t = 0.45$

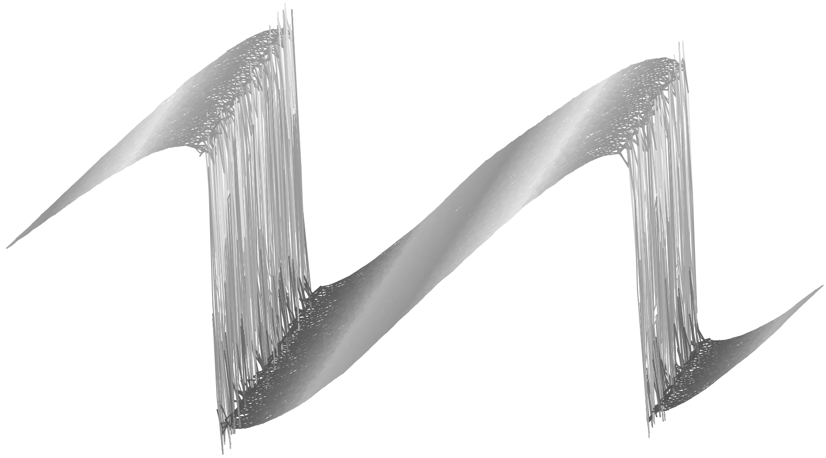

Figure 3. Numerical solution of problem (4.5)–(4.6) computed by FV DG method plotted at $t = 0.45$

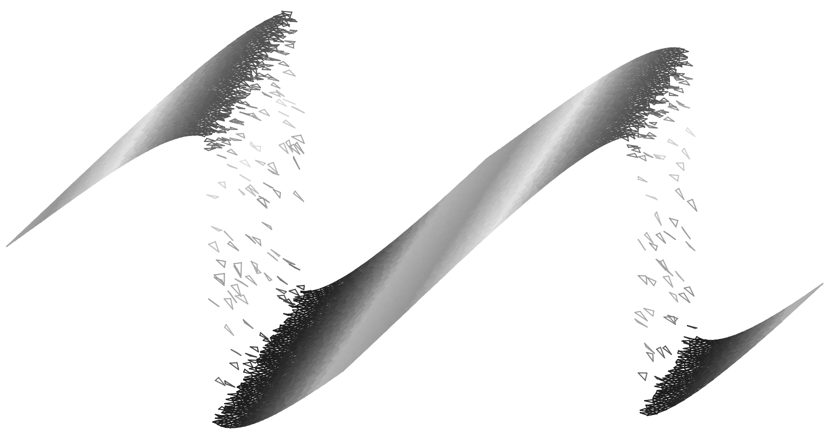

Figure 4. Numerical solution of problem (4.5)–(4.6) computed by DG FE method with limiting, plotted at $t = 0.45$

appear in the numerical solution only there where the solution is discontinuous. This indicates that the DG FE method is very suitable for the numerical solution of problem with solutions containing steep gradients or discontinuities.

5. DG FE method for the Euler equations

The system of the Euler equations describing the 2D inviscid flow can be written in the form

$$\frac{\partial \boldsymbol{w}}{\partial t} + \sum_{s=1}^{2} \frac{\partial \boldsymbol{f}_s(\boldsymbol{w})}{\partial x_s} = 0 \quad \text{in } Q_T = \Omega \times (0,T), \tag{5.1}$$

where $\Omega \subset \mathbb{R}^2$ is a bounded domain occupied by gas,

$$\boldsymbol{w} = (w_1, \ldots, w_4)^{\mathrm{T}} = (\varrho,\ \varrho v_1,\ \varrho v_2,\ e)^{\mathrm{T}} \tag{5.2}$$

is the so-called state vector and

$$\begin{aligned} \boldsymbol{f}_s(\boldsymbol{w}) &= (f_s^1(w), \ldots, f_s^4(w)) \\ &= (\varrho v_s,\ \varrho v_s v_1 + \delta_{s1} p,\ \varrho v_s v_2 + \delta_{s2} p,\ (e+p)\, v_s)^{\mathrm{T}},\ \ s = 1,2, \end{aligned} \tag{5.3}$$

are the so-called inviscid (Euler) fluxes. We use the following notation: ϱ-density, p-pressure, e-total energy, $\boldsymbol{v} = (v_1, v_2)$-velocity. The state equation implies that

$$p = (\gamma - 1)\,(e - \varrho|\boldsymbol{v}|^2/2). \tag{5.4}$$

Here $\gamma > 1$ is the Poisson adiabatic constant. The system (5.1)–(5.4) is *hyperbolic*. It is equipped with the initial condition

$$\boldsymbol{w}(\boldsymbol{x}, 0) = \boldsymbol{w}^0(x), \quad x \in \Omega, \tag{5.5}$$

and boundary conditions

$$B(\boldsymbol{w}) = 0 \quad \text{on } \partial\Omega \times (0,T), \tag{5.6}$$

chosen in such a way that problem (5.1)–(5.6) is linearly well-posed. For details see, e.g., [8] or [9].

The DG FE discretization in space combined with the forward Euler discretization in time can be written in the form (3.1) where (due to zero diffusion) $a_h = 0$, $\varepsilon = 0$, $\ell_h = 0$. In order to avoid spurious oscillations in the numerical solution, automatic adaptive limiting of order of accuracy is used, which leads to scheme (3.3)–(3.4):

$$\begin{aligned} &\text{(a)} \quad \boldsymbol{w}_h^{k+1} \in S_h := S_h^{1,-1}(\Omega, \mathcal{T}_h)^4, \\ &\text{(b)} \quad (\boldsymbol{w}_h^{k+1}, \boldsymbol{v}_h) = (\tilde{\boldsymbol{w}}_h^k, \boldsymbol{v}_h) - \tau_k\, \tilde{b}_h(\tilde{\boldsymbol{w}}_h^k, \boldsymbol{v}_h) \quad \forall \boldsymbol{v}_h \in S_h \end{aligned} \tag{5.7}$$

where

$$(5.8)\qquad \begin{array}{ll} \text{(a)} & \text{we set} \quad \tilde{\boldsymbol{w}}_h^k|_{K_i} := \boldsymbol{w}_h^k|_{K_i} \quad \forall i \in I, \\ \text{(b)} & \text{if } g(i) > 1 \text{ for some } i \in I, \text{ then we set } \tilde{\boldsymbol{w}}_h^k|_{K_i} := \pi_0\, \boldsymbol{w}_h^k|_{K_i}. \end{array}$$

The shock indicator $g(i)$ is computed by (3.2), where instead of u_h^k, the density ϱ on the k-th time level is used. The form $\tilde{b}_h$ is defined by (1.9), where H is chosen to be the well-known *Osher-Solomon* numerical flux (see [19], [9]). The Osher-Solomon boundary conditions are also described in [9].

As an example we present the inviscid flow past the NACA0012 profile with the far field Mach number $M = 0.8$, the angle of attack $\alpha = 1.25°$ and $\gamma = 1.4$. The algorithm (5.7)–(5.8) was used as an iterative time marching process with "$k \to \infty$" for obtaining the steady state solution. The computational mesh $\mathcal{T}_h$ was obtained by the anisotropic mesh adaptation (AMA) (see [5]). The stationary solution was obtained after $4.3 \cdot 10^5$ time steps (for the 7th level of mesh adaptation) when the achieved residuum was $\|\varrho^{k+1} - \varrho^k\|_{L^1(\Omega)}/\tau_k \leqslant 10^{-5}$. In Figure 5 the mesh and the Mach number isolines with well resolved sharp shock waves are plotted.

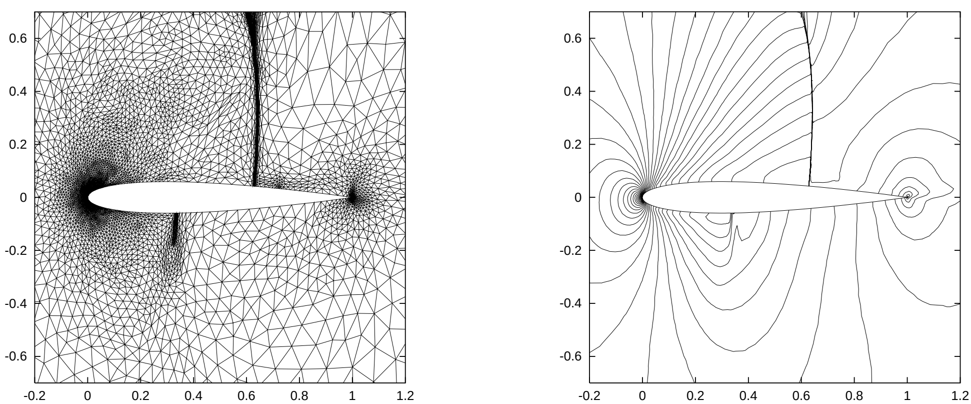

Figure 5. The final triangulation obtained by AMA (left) and the corresponding isolines of Mach number obtained by DG FE method

A series of numerical experiments have shown that in many cases the DG FE method does not give good resolution in a neighbourhood of curved parts of boundary $\partial\Omega$, if Ω is approximated by a polygonal domain. In order to get a good quality solution, it is necessary to use superparametric finite elements or a suitable numerical integration. (This will be a subject matter of a forthcoming paper. See also [2] and [17], where the subsonic flow without shock waves is solved.)

References

[1] *Ph. Angot, V. Dolejší, M. Feistauer, J. Felcman*: Analysis of a combined barycentric finite volume—nonconforming finite element method for nonlinear convection-diffusion problems. Appl. Math. *43* (1998), 263–310.

[2] *F. Bassi, S. Rebay*: A high order discontinuous Galerkin method for the numerical solution of the compressible Navier-Stokes equations. J. Comput. Phys. *131* (1997), 267–279.

[3] *B. Cockburn*: Discontinuous Galerkin methods for convection dominated problems. High-Order Methods for Computational Physics (T. J. Barth, H. Deconinck, eds.). Lecture Notes in Computational Science and Engineering 9, Springer, Berlin, 1999, pp. 69–224.

[4] *B. Cockburn, G. E. Karniadakis, C.-W. Shu*: Discontinuous Galerkin Methods. Lect. Notes Comput. Sci. Eng. 11., Springer, Berlin, 2000.

[5] *V. Dolejší*: Anisotropic mesh adaptation for finite volume and finite element methods on triangular meshes. Comput. Vis. Sci. *1* (1998), 165–178.

[6] *V. Dolejší, M. Feistauer, C. Schwab*: A finite volume discontinuous Galerkin scheme for nonlinear convection-diffusion problems. Calcolo. Preprint, Forschungsinstitut für Mathematik ETH Zürich, Januar 2001. To appear.

[7] *V. Dolejší, M. Feistauer, C. Schwab*: On some aspects of the discontinuous Galerkin finite element method for conservation laws. Mathematics and Computers in Simulation. The Preprint Series of the School of Mathematics, Charles University Prague, No. MATH-KNM-2001/5, 2001. To appear.

[8] *M. Feistauer*: Mathematical Methods in Fluid Dynamics. Longman Scientific & Technical, Harlow, 1993.

[9] *M. Feistauer*: Numerical methods for compressible flow. Mathematical Fluid Mechanics. Recent Results and Open Questions (J. Neustupa, P. Penel, eds.). Birkhäuser, Basel, 2001, pp. 105–142.

[10] *M. Feistauer, J. Felcman*: Theory and applications of numerical schemes for nonlinear convection-diffusion problems and compressible Navier-Stokes equations. The Mathematics of Finite Elements and Applications (J. R. Whiteman, ed.). Highlights, 1996, Wiley, Chichester, 1997, pp. 175–194.

[11] *M. Feistauer, J. Felcman, V. Dolejší*: Numerical solution of compressible flow through cascades of profiles. Z. Angew. Math. Mech. *76* (1996), 297–300.

[12] *M. Feistauer, J. Felcman, M. Lukáčová*: On the convergence of a combined finite volume—finite element method for nonlinear convection-diffusion problems. Numer. Methods Partial Differ. Equations *13* (1997), 163–190.

[13] *M. Feistauer, J. Felcman, M. Lukáčová, G. Warnecke*: Error estimates for a combined finite volume—finite element method for nonlinear convection-diffusion problems. SIAM J. Numer. Anal. *36* (1999), 1528–1548.

[14] *M. Feistauer, J. Slavík, P. Stupka*: On the convergence of a combined finite volume—finite element method for nonlinear convection-diffusion problems. Explicit schemes. Numer. Methods Partial Differ. Equations *15* (1999), 215–235.

[15] *J. Felcman*: On a 3D adaptation for compressible flow. Proceedings of the Conf. Finite Element Methods for Three-Dimensional Problems, Jyväskylä, June 27–July 1, 2000. To appear.

[16] *J. Fořt, J. Halama, A. Jirásek, M. Kladrubský, K. Kozel*: Numerical solution of several 2D and 3D internal and external flow problems. Numerical Modelling in Continuum Mechanics (M. Feistauer, R. Rannacher, K. Kozel, eds.). Matfyzpress, Praha, 1997, pp. 283–291.

[17] *R. Hartmann, P. Houston*: Adaptive discontinuous Galerkin finite element methods for nonlinear conservation laws. Preprint 2001–20, Mai 2001, SFB 359, Universität Heidelberg.
[18] *J. T. Oden, I. Babuška, C. E. Baumann*: A discontinuous *hp* finite element method for diffusion problems. J. Comput. Phys. *146* (1998), 491–519.
[19] *S. P. Spekreijse*: Multigrid Solution of the Steady Euler Equations. Centrum voor Wiskunde en Informatica, Amsterdam, 1987.

Authors' addresses: *V. Dolejší*, *M. Feistauer*, Faculty of Mathematics and Physics, Charles University Prague, Sokolovská 83, 186 75 Praha 8, Czech Republic, e-mail: dolejsi@karlin.mff.cuni.cz, feist@karlin.mff.cuni.cz; *C. Schwab*, Seminar for Applied Mathematics, ETH Zürich, Raemistrasse 101, CH-8092 Zürich, Switzerland, e-mail: schwab@math.ethz.ch.