HIGHER ORDER METHOD FOR THE NUMERICAL SOLUTION OF THE COMPRESSIBLE EULER EQUATIONS

MILOSĽAV FEISTAUER† AND VÁCLAV KUČERA‡

Abstract. This paper is concerned with a numerical technique for the solution of inviscid compressible flow with a wide range of Mach numbers. It is based on the use of the discontinuous Galerkin finite element method applied to the Euler equations written in the conservative form, a semi-implicit time discretization and characteristics-based boundary conditions, which are transparent for acoustic phenomena. For transonic flows, additional shock capturing terms are added in order to avoid the Gibbs phenomenon near shock waves.

Key words. compressible inviscid flow, Euler equations, discontinuous Galerkin finite element method, semi-implicit time discretization, low Mach number flow, transonic flow

AMS subject classifications. 65M60, 76B99, 76H05

1. Introduction. In the numerical solution of compressible flow, it is necessary to overcome a number of obstacles. Let us mention the necessity to resolve accurately shock waves, contact discontinuities and (in viscous flow) boundary layers, wakes and their interaction. Some of these phenomena are connected with the simulation of high speed flow with high Mach numbers. However, it appears that the solution of low Mach number flow is also rather difficult. This is caused by the stiff behaviour of numerical schemes and acoustic phenomena appearing in low Mach number flows at incompressible limit. In this case, standard finite volume schemes fail. This led to the development of special finite volume techniques allowing the simulation of compressible flow at incompressible limit, which is based on modifications of the Euler or Navier-Stokes equations. (See, e.g. [13], [15], [18, Chapter 14], or [14, Chapter 5].)

Here we are concerned with the development of an efficient, robust and accurate method allowing the solution of compressible flow with a wide range of the Mach number without any modification of the governing equations. This technique is based on the discontinuous Galerkin finite element method (DGFEM), which can be considered as a generalization of the finite volume as well as finite element methods, using advantages of both these techniques. It employs piecewise polynomial approximations without any requirement on the continuity on interfaces between neighbouring elements. (For various applications of the DGFEM to compressible flow, see e.g. [1], [2], [3], [4], [11], [16]. Theory of the DGFEM applied to nonlinear nonstationary convection diffusion problems can be found in [5], [6] and [8].) The discontinuous Galerkin space semidiscretization is combined with a semi-implicit time discretization and a special treatment of boundary conditions in inviscid convective terms. In this way we obtain a numerical scheme requiring the solution of only one linear system on each time level.

The computational results show that the presented method is applicable to the numerical solution of inviscid compressible high-speed flow as well as flow with a very low...
Mach number at incompressible limit.

2. Continuous problem. For simplicity of the treatment we shall consider two-
dimensional flow, but the method can be applied to 3D flow as well. The system of the
Euler equations describing 2D inviscid flow can be written in the form

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

(2.1)

where $\Omega \subset \mathbb{R}^2$ is a bounded domain occupied by gas, $T > 0$ is the length of a time
interval,

$$w = (w_1, \ldots, w_4)^T = (\rho, \rho v_1, \rho v_2, E)^T$$

(2.2)

is the so-called state vector and

$$f_s(w) = (\rho v_s, \rho v_s v_1 + \delta_{s1} p, \rho v_s v_2 + \delta_{s2} p, (E + p) v_s)^T$$

(2.3)

are the inviscid (Euler) fluxes of the quantity $w$ in the directions $x_s$, $s = 1, 2$. We use the
following notation: $\rho$ – density, $p$ – pressure, $E$ – total energy, $v = (v_1, v_2)$ – velocity, $\delta_{sk}$ – Kronecker symbol. The equation of state implies that

$$p = (\gamma - 1) (E - \rho |v|^2 / 2).$$

(2.4)

Here $\gamma > 1$ is the Poisson adiabatic constant. The system (2.1)–(2.4) is diagonally
hyperbolic. It is equipped with the initial condition

$$w(x, 0) = w^0(x), \quad x \in \Omega,$$

(2.5)

and the boundary conditions, which are treated in Section 4. We define the matrix

$$P(w, n) := \sum_{s=1}^{2} A_s(w) n_s,$$

(2.6)

where $n = (n_1, n_2) \in \mathbb{R}^2$, $n_1^2 + n_2^2 = 1$ and

$$A_s(w) = \frac{D f_s(w)}{D w}, \quad s = 1, 2,$$

(2.7)

are the Jacobi matrices of the mappings $f_s$. It is possible to show that $f_s$, $s = 1, 2$, are homogenous mappings of order one, which implies that

$$f_s(w) = A_s(w) w, \quad s = 1, 2.$$

(2.8)

3. Discretization. Here we describe the construction of the discrete problem.

3.1. Space semidiscretization by the DGFEM. Let $\Omega_h$ be a polygonal approx-
imation of $\Omega$. By $T_h$ we denote a partition of $\Omega_h$ consisting of various types of convex
elements $K_i \in T_h$, $i \in I$ ($I \subset \mathbb{Z}^+ = \{0, 1, 2, \ldots\}$ is a suitable index set), e.g., triangles, quadrilaterals or in general convex polygons. (Let us note that in [6] and [8] it was shown that in the DGFEM also general nonconvex star-shaped polygonal elements can be used.) By $\Gamma_{ij}$ we denote a common edge between two neighbouring elements $K_i$ and $K_j$. Moreover, we set $s(i) = \{j \in I; K_j$ is a neighbour of $K_i\}$. The boundary $\partial \Omega_h$ is formed by a finite number of faces of elements $K_i$ adjacent to $\partial \Omega_h$. We denote all these
boundary faces by \( S_j \), where \( j \in I_b \subset \mathbb{Z}^+ = \{-1, -2, \ldots \} \). Now we set \( \gamma(i) = \{ j \in I_b; S_j \text{ is a face of } K_i \in T_h \} \) and \( \Gamma_{ij} = S_j \) for \( K_i \in T_h \) such that \( S_j \subset \partial K_i, j \in I_b \). For \( K_i \) not containing any boundary face \( S_j \) we set \( \gamma(i) = 0 \). 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

\[
\partial K_i = \bigcup_{j \in s(i)} \Gamma_{ij}, \quad \partial K_i \cap \partial \Omega_h = \bigcup_{j \in \gamma(i)} \Gamma_{ij}, \tag{3.1}
\]

The symbol \( \mathbf{n}_{ij} = ((n_{ij})_1, (n_{ij})_2) \) will denote the unit outer normal to \( \partial K_i \) on the face \( \Gamma_{ij} \). By \( h_{K_i} \) and \( |K_i| \) we shall denote the diameter and the area, respectively, of an element \( K_i \in T_h \).

The approximate solution will be sought at each time instant \( t \) as an element of the finite-dimensional space

\[
S_h = S^{r-1}(\Omega_h, T_h) = \{ v; v|_K \in P^r(K) \text{ } \forall K \in T_h \}^4,
\]

where \( r \geq 0 \) is an integer and \( P^r(K) \) denotes the space of all polynomials on \( K \) of degree \( \leq r \). Functions \( v \in S_h \) are in general discontinuous on interfaces \( \Gamma_{ij} \).

By \( v|_{\Gamma_{ij}} \) and \( v|_{\Gamma_1} \), we denote the values of \( v \) on \( \Gamma_{ij} \) considered from the interior and the exterior of \( K_i \), respectively.

In order to derive the discrete problem, we multiply (2.1) by a test function \( \varphi \in S_h \), integrate over any element \( K_i \), \( i \in I \), apply Green’s theorem and sum over all \( i \in I \). Then we approximate fluxes through the faces \( \Gamma_{ij} \) with the aid of a numerical flux \( \mathbf{H} = \mathbf{H}(\mathbf{u}, \mathbf{w}, \mathbf{n}) \) in the form

\[
\int_{\Gamma_{ij}} \sum_{s=1}^2 f_s(\mathbf{w}(t))(n_{ij})_s \cdot \varphi \, dS \approx \int_{\Gamma_{ij}} \mathbf{H}(\mathbf{w}_h(t)|_{\Gamma_{ij}}, \mathbf{w}_h(t)|_{\Gamma_{ij}}, \mathbf{n}_{ij}) \cdot \varphi \, dS.
\]

If we introduce the forms

\[
(w_h, \varphi_h)_h = \int_{\Omega_h} w_h \cdot \varphi_h \, dx, \quad \tilde{b}_h(w_h, \varphi_h) = \sigma_1 + \sigma_2, \tag{3.2}
\]

where

\[
\sigma_1 = - \sum_{K \in T_h} \int_{\Omega_h} \sum_{s=1}^2 f_s(w_h) \cdot \frac{\partial \varphi_h}{\partial x_s} \, dx,
\]

\[
\sigma_2 = \sum_{K \in T_h} \sum_{i \in I} \int_{\Gamma_{ij}} \mathbf{H}(w_h|_{\Gamma_{ij}}, w_h|_{\Gamma_{ij}}, \mathbf{n}_{ij}) \cdot \varphi_h \, dS,
\]

we can define an approximate solution of (2.1) as a function \( w_h \) satisfying the conditions

a) \( w_h \in C^1([0, T]; S_h) \),

b) \( \frac{d}{dt} (w_h(t), \varphi_h)_h + \tilde{b}_h(w_h(t), \varphi_h) = 0, \quad \forall \varphi_h \in S_h, \text{ } \forall t \in (0, T) \),

c) \( w_h(0) = \Pi_h \mathbf{w}^0 \),

where \( \Pi_h \mathbf{w}^0 \) is the \( L^2 \)-projection of \( \mathbf{w}^0 \) from the initial condition (2.5) on the space \( S_h \). If we set \( r = 0 \), then we obviously obtain the finite volume method.
3.2. Time discretization. Relation (3.4), b) represents a system of ordinary differential equations which can be solved by a suitable numerical method. Usually, Runge-Kutta schemes are applied. However, they are conditionally stable and the time step is strongly limited by the CFL-stability condition. Therefore, we develop a semi-implicit time discretization, which is unconditionally stable and requires the solution of a linear system on each time level. This is carried out with the aid of a suitable partial linearization of the form \( \dot{b}_h \). In what follows, we consider a partition \( 0 = t_0 < t_1 < t_2 \ldots \) of the time interval \( (0, T) \) and set \( \tau_k = t_{k+1} - t_k \). We use the notation \( w_h^k \) for the approximation of \( w_h(t_k) \).

On the basis of relation (2.8) and the use of the Vijayasundaram numerical flux, similarly as in [4], we construct the form

\[
\begin{align*}
b_h(w_h^k, w_h^{k+1}, \varphi_h) &= -\sum_{K \in T_h} \int_K \sum_{s=1}^2 A_h(w_h^k(x)) w_h^{k+1}(x) \cdot \frac{\partial \varphi_h(x)}{\partial x_s} \, dx \\
&\quad + \sum_{K, \in T_h} \sum_{j \in S(i)} \int_{\Gamma_{ij}} [P^+ ((w_h^k)_{ij}, n_{ij}) w_h^{k+1}|_{\Gamma_{ij}} + P^- ((w_h^k)_{ij}, n_{ij}) w_h^{k+1}|_{\Gamma_{ij}}] \cdot \varphi_h \, dS,
\end{align*}
\]

(3.5)

which is linear with respect to the second and third variable. We use the notation \( (w_h^k)_{ij} = (w_h|_{\Gamma_{ij}} + w_h|_{\Gamma_{ij}})/2 \). Further, \( P^\pm = P^\pm(w, n) \) represents positive/negative part of the matrix \( P \) defined on the basis of its diagonalization (see, e.g. [10, Section 3.1]):

\[
P = TDT^{-1}, \quad D = \text{diag} (\lambda_1, \ldots, \lambda_4),
\]

(3.6)

where \( \lambda_1, \ldots, \lambda_4 \) are the eigenvalues of \( P \). Then we set

\[
D^\pm = \text{diag} (\lambda_1^\pm, \ldots, \lambda_4^\pm),
\]

(3.7)

\[
P^\pm = T D^\pm T^{-1},
\]

where \( \lambda^+ = \max\{\lambda, 0\} \) and \( \lambda^- = \min\{\lambda, 0\} \).

On the basis of the above considerations we obtain the following semi-implicit scheme: For each \( k \geq 0 \) find \( w_h^{k+1} \) such that

\[
\begin{align*}
a) \quad & w_h^{k+1} \in S_h, \\
b) \quad & \left( \frac{w_h^{k+1} - w_h^k}{\tau_k}, \varphi_h \right)_h + b_h(w_h^k, w_h^{k+1}, \varphi_h) = 0, \quad \forall \varphi_h \in S_h, \quad k = 0, 1, \ldots, \\
c) \quad & w_h^0 = \Pi_h w^0.
\end{align*}
\]

(3.8)

This is a first order accurate scheme in time. It is also possible to construct a semi-implicit two step second order time discretization (see [4]). The linear algebraic system equivalent to (3.8), b) is solved by the GMRES method with a block diagonal preconditioning.

In order to obtain an accurate solution near curved boundaries, we use higher order isoparametric elements as in [1] or [3].

4. Boundary conditions. If \( \Gamma_{ij} \subset \partial \Omega_h \), i.e. \( j \in \gamma(i) \), it is necessary to specify the boundary state \( w|_{\Gamma_{ij}} \) appearing in the numerical flux \( H \) in the definition of the inviscid form \( b_h \). The appropriate treatment of boundary conditions plays a crucial role in the solution of low Mach number flows.

On a fixed impermeable wall we employ a standard approach using the condition \( v \cdot n = 0 \) and extrapolating the pressure. On the inlet and outlet it is necessary to use
nonreflecting boundary conditions transparent for acoustic effects coming from inside of $\Omega$. Therefore, characteristics-based boundary conditions are used.

Using the rotational invariance, we transform the Euler equations to the coordinates $\tilde{x}_1$, parallel with the normal direction $n$ to the boundary, and $\tilde{x}_2$, tangential to the boundary, neglect the influence of the states on elements that are not adjacent to $\Gamma_{ij}$ and linearize the resulting system around the state $q_{ij} = Q(n_{ij})w |_{\Gamma_{ij}}$, where

$$Q(n_{ij}) = \begin{pmatrix}
1, & 0, & 0, & 0 \\
0, & (n_{ij})_1, & (n_{ij})_2, & 0 \\
0, & -(n_{ij})_2, & (n_{ij})_1, & 0 \\
0, & 0, & 0, & 1
\end{pmatrix} \quad (4.1)$$

is the rotational matrix. Then we obtain the linear system

$$\frac{\partial q}{\partial t} + A_1(q_{ij}) \frac{\partial q}{\partial \tilde{x}_1} = 0, \quad (4.2)$$

for the vector-valued function $q = Q(n_{ij})w$, considered in the set $(-\infty, 0) \times (0, \infty)$ and equipped with the initial and boundary conditions

$$q(\tilde{x}_1, 0) = q_{ij}, \quad \tilde{x}_1 \in (-\infty, 0),$$

$$q(0, t) = q_{ji}, \quad t > 0. \quad (4.3)$$

The goal is to choose $q_{ji}$ in such a way that this initial-boundary value problem is well posed, i.e. has a unique solution. The method of characteristics leads to the following process:

Let us put $q^*_{ji} = Q(n_{ij})w^*_{ji}$, where $w^*_{ji}$ is a prescribed boundary state at the inlet or outlet. We calculate the eigenvectors $r_s$ corresponding to the eigenvalues $\lambda_s$, $s = 1, \ldots, 4$, of the matrix $A_1(q_{ij})$, arrange them as columns in the matrix $T$ and calculate $T^{-1}$ (explicit formulae can be found in [10, Section 3.1]). Now we set

$$\alpha = T^{-1}q_{ij}, \quad \beta = T^{-1}q^*_{ji}. \quad (4.4)$$

and define the state $q_{ji}$ by the relations

$$q_{ji} := \sum_{s=1}^4 \gamma_s r_s, \quad \gamma_s = \begin{cases} 
\alpha_s, & \lambda_s \geq 0, \\
\beta_s, & \lambda_s < 0.
\end{cases} \quad (4.5)$$

Finally, the sought boundary state $w |_{\Gamma_{ji}}$ is defined as

$$w |_{\Gamma_{ji}} = w_{ji} = Q^{-1}(n_{ij})q_{ji}. \quad (4.6)$$

5. Shock capturing. In the case of high speed flow, it is necessary to avoid the Gibbs phenomenon manifested by spurious overshoots and undershoots in computed quantities near discontinuities (shock waves, contact discontinuities). These phenomena do not occur in low Mach number regimes, but in transonic flow they cause instabilities in the numerical solution.

We avoid the Gibbs phenomenon by introducing a suitable stabilization terms, motivated by [7] and [12]. First we introduce the discontinuity indicator $g(i)$ proposed in [7] and defined by

$$g(i) = \int_{\partial K_i} [\rho h]^2 dS/(h_{K_i} |K_i|^{3/4}), \quad K_i \in T_h. \quad (5.1)$$
By \( [u]_{\Gamma_{ij}} = u_{ij} - u_{ji} \) we denote the jump on \( \Gamma_{ij} \) of a function \( u \in S_h \). Further, we define the discrete indicator
\[
G(i) = \begin{cases} 
0, & g(i) < 1, \\
1, & g(i) \geq 1, \\
K_i \in T_h.
\end{cases}
\] (5.2)

Now, to the left-hand side of (3.8), b) we add the artificial viscosity form
\[
\tilde{\beta}_h(w, \varphi) = \nu_1 \sum_{i \in I} h_{K_i} G(i) \int_{K_i} \nabla w \cdot \nabla \varphi \, dx,
\] (5.3)
where \( \nu_1 \approx 1 \). The stabilization form \( \tilde{\beta}_h \) is treated semi-implicitly with \( G(i) = G^k(i) \) computed from \( w^k_h \). Therefore, we write
\[
\beta_h(w^k_h, w^{k+1}_h, \varphi) = \nu_1 \sum_{i \in I} h_{K_i} G^k(i) \int_{K_i} \nabla w^{k+1}_h \cdot \nabla \varphi \, dx.
\] (5.4)

This form limits the order of accuracy only on elements lying in a small neighbourhood of a discontinuity. However, it appears that this is insufficient on strongly refined grids. Therefore, we propose to augment the left-hand side of (3.8), b) by adding the form
\[
\tilde{J}_h(w, \varphi) = \nu_2 \sum_{i \in I} \sum_{j \in s(i)} \frac{1}{2} (G(i) + G(j)) \int_{\Gamma_{ij}} [w] \cdot [\varphi] \, dS,
\] (5.5)
where \( \nu_2 \approx 1 \). In this way we penalize inter-element jumps in the vicinity of the shock wave. This form is again treated semi-implicitly, similarly as in (5.4). We set
\[
J_h(w^k_h, w^{k+1}_h, \varphi) = \nu_2 \sum_{i \in I} \sum_{j \in s(i)} \frac{1}{2} (G^k(i) + G^k(j)) \int_{\Gamma_{ij}} [w^{k+1}_h] \cdot [\varphi] \, dS.
\] (5.6)

Thus, the resulting scheme reads:

\begin{enumerate}
\item \( w^{k+1}_h \in S_h \),
\item \( \left( \frac{w^{k+1}_h - w^k_h}{\tau_k}, \varphi_h \right)_h + b_h(w^k_h, w^{k+1}_h, \varphi_h) + \beta_h(w^k_h, w^{k+1}_h, \varphi_h) + J_h(w^k_h, w^{k+1}_h, \varphi_h) = 0, \forall \varphi_h \in S_h, \ k = 0, 1, \ldots, \)
\item \( w^0_h = \Pi_h w^0 \).
\end{enumerate}

This method successfully overcomes problems with the Gibbs phenomenon in the context of the semi-implicit scheme. It is important that \( G(i) \) vanishes in regions where the solution is regular. Therefore, the scheme does not produce any nonphysical entropy in these regions (See Fig. 6.3).

6. Numerical examples. In order to show the robustness of the described technique with respect to the Mach number, we present computational results obtained for two types of compressible flow.
6.1. Low Mach number flow. First, the semi-implicit scheme (5.7) was applied to the solution of stationary inviscid low Mach number flow past a circular cylinder with the far field velocity parallel to the axis \( x_1 \) and the Mach number \( M_\infty = 10^{-4} \). The computational domain has the form of a square with sides of the length equal to 20 diameters of the cylinder. We show here details of the flow in the vicinity of the cylinder. FIG. 6.1 shows isolines of the absolute value of the velocity for the compressible flow computed by scheme (5.7) with piecewise quadratic elements (i.e. \( r = 2 \)), on a coarse mesh formed by 361 elements and on a fine mesh with 8790 elements, compared with the exact solution of incompressible flow (computed by the method of complex functions – see [9, Section 2.2.35]). The steady-state solution is obtained with the aid of the time stabilization for \( t \to \infty \).

![Velocity isolines for the approximate solution of compressible flow – coarse mesh (top left), fine mesh (top right), compared with the exact solution of incompressible flow (bottom).](image)

6.2. Transonic flow. The performance of shock capturing terms from Section 5 is tested on the GAMM channel with a 10% circular bump and the inlet Mach number
equal to 0.67. The method (5.7) with piecewise quadratic elements is used and time stabilization for $t \to \infty$ is applied for obtaining the steady-state solution. In this case a conspicuous shock wave is developed. In Fig. 6.2 the density distribution along the lower wall is shown. We see a well resolved discontinuity due to the shock wave. Moreover, we can see the so-called Zierep singularity (a small local maximum behind the shock) proving a good quality of the obtained numerical solution. The artificial viscosity forms $\beta_h$ and $J_h$ given by (5.4) and (5.6) eliminate the Gibbs phenomenon. Fig. 6.3 shows the entropy isolines. One can see that the entropy is produced only on the shock wave. This is caused by the use of the forms $\beta_h$ and $J_h$, which are active only on elements lying in a small neighbourhood of discontinuities and do not influence the solution in areas, where the exact solution is regular.

![Fig. 6.2. Transonic flow through the GAMM channel, density distribution on the lower wall.](image1)

![Fig. 6.3. Transonic flow through the GAMM channel, entropy isolines.](image2)

7. **Conclusion.** In this paper we have presented a new method for the numerical solution of the Euler equations describing inviscid compressible flow. The method allows the simulation of compressible flow with a wide range of Mach numbers – from very small values in the case of flows at incompressible limit, up to large Mach numbers for high speed transonic flows. Numerical experiments prove that the method is unconditionally stable. There are several important ingredients making the method robust with respect to the Mach number, without the necessity to modify the Euler equations:

- discontinuous Galerkin space discretization,
– semi-implicit time stepping,
– characteristic treatment of boundary conditions,
– suitable limiting of order of accuracy in the vicinity of discontinuities in order to avoid the Gibbs phenomenon,
– the use of isoparametric finite elements at curved parts of the boundary.

Our further goal is the extension of the presented technique to compressible viscous flow described by the full system of the compressible Navier-Stokes equations.

REFERENCES