An Iterative FEM-FCT Algorithm for the Compressible Euler Equations

M. Möller, D. Kuzmin, and S. Turek

Institute of Applied Mathematics,
University of Dortmund, Germany

- State of the art: discretization techniques
- Discrete upwinding for scalar equations
- Iterative defect correction scheme
- Generalized FEM-FCT formulation

- Matrix assembly for the Euler equations
- Construction of artificial viscosities
- Solution strategies for coupled equations
- Implementation of boundary conditions
Algebraic Flux Correction of FCT-type

Scalar conservation law \[ \frac{\partial u}{\partial t} + \nabla \cdot (\mathbf{v} u) = 0 \text{ in } \Omega \]

1. Linear high-order scheme (e.g. Galerkin FEM)

\[ M_C \frac{u^{n+1} - u^n}{\Delta t} = \theta K u^{n+1} + (1 - \theta) K u^n, \quad \exists j \neq i : k_{ij} < 0 \]

2. Linear low-order scheme \[ L = K + D \]

\[ M_L \frac{u^{n+1} - u^n}{\Delta t} = \theta L u^{n+1} + (1 - \theta) L u^n, \quad l_{ij} \geq 0, \forall j \neq i \]

3. Nonlinear high-resolution scheme \[ f_i = \sum_{j \neq i} f^a_{ij} \]

\[ M_L \frac{u^{n+1} - u^n}{\Delta t} = \theta L u^{n+1} + (1 - \theta) L u^n + f(u^{n+1}, u^n) \]

Equivalent representation \[ A u^{n+1} = B(\bar{u})\tilde{u}, \text{ where} \]

\[ A = M_L - \theta \Delta t L \text{ is an } M\text{-matrix and } b_{ij} \geq 0, \forall i, j \]
Algebraic Design Criteria

Algebraic Constraint I: (semi-discrete level)

\[
\frac{du_i}{dt} = \sum_{j \neq i} \sigma_{ij} (u_j - u_i), \quad \sigma_{ij} \geq 0
\]

\[
u_i = \left\{ \max_j \right\} u_j \Rightarrow u_j - u_i \left\{ \leq \right\} 0 \Rightarrow \frac{du_i}{dt} \left\{ \leq \right\} 0
\]

In 1D, Jameson’s (1993) LED property is equivalent to Harten’s (1983) TVD conditions.

Algebraic Constraint II: (fully discrete level)

A discrete scheme of the form \(Au^{n+1} = Bu^n, \quad u^n \geq 0\) is positivity-preserving if \(A\) is an \(M\)-matrix \((a_{ij} \leq 0, \forall j \neq i, \quad A^{-1} \geq 0)\) and all entries of \(B\) are non-negative.

Tool:

Discrete diffusion operators \(D = \{d_{ij}\}\),

where \(d_{ij} = d_{ji}\) and \(\sum_j d_{ij} = \sum_i d_{ij} = 0\).

Flux decomposition of diffusive terms into antisymmetric fluxes

\[
(Du)_i = \sum_j d_{ij} u_j = \sum_{j \neq i} d_{ij} (u_j - u_i) = \sum_j f_{ij}, \quad \text{where} \quad f_{ji} = -f_{ij}.
\]
High- and Low-Order Schemes

Difference between high- and low-order scheme

\[ P(u) = [M_L - M_C] \frac{du}{dt} - \underbrace{[L - K]}_{D} u \]

Flux decomposition of raw antidiffusion

\[ P_i = \sum_{j \neq i} f_{ij}, \quad f_{ij} = -[m_{ij} \frac{d}{dt} + d_{ij}](u_j - u_i), \quad f_{ji} = -f_{ij} \]

Nonlinear system for an implicit time discretization (standard \( \theta \)-scheme)

\[ M_L \frac{u^{n+1} - u^n}{\Delta t} = \theta Lu^{n+1} + (1 - \theta)Lu^n + P(u^{n+1}, u^n) \]

Successive approximation

\[ Au^{(m+1)} = b^{(m+1)} \quad u^{(0)} = u^n, \quad m = 0, 1, 2, \ldots \]

Preconditioner \( A = M_L - \theta \Delta tL \) and load-vector \( b^{(m+1)} = b^n + P(u^{(m)}, u^n) \)

Low-order contribution

\[ b^n = [M_L + (1 - \theta)\Delta tL]u^n \]

Raw antidiffusion

\[ P_i^{(m)} = \sum_{j \neq i} f_{ij}^{(m)} \quad \text{with antisymmetric fluxes} \]

\[ f_{ij}^{(m)} = [m_{ij} - (1 - \theta)\Delta t d_{ij}^{(m)}](u_j^n - u_i^n) - [m_{ij} + \theta \Delta t d_{ij}^{(m)}](u_j^{(m)} - u_i^{(m)}) = -f_{ji}^{(m)} \]
Basic FEM-FCT Algorithm

\[ Au^{(m+1)} = b^{(m+1)} \]

By construction, \( A = M_L - \theta \Delta t L \) is an M-matrix which is easy to ‘invert’ and satisfies \textit{Algebraic Constraint II}.

\textbf{Strategy:} Multiply the antidiffusive fluxes by ‘some’ correction factors \( \alpha_{ij} \in [0, 1] \) so that there exists a matrix \( B(\tilde{u}) \geq 0 \) and a positivity-preserving solution \( \tilde{u} \) such that \( b^{(m+1)} = B(\tilde{u})\tilde{u} \).

\text{Positivity transfer cycle} \quad u^n \geq 0 \ \Rightarrow \ \tilde{u} \geq 0 \ \Rightarrow \ u^{(m+1)} = A^{-1}B(\tilde{u})\tilde{u} \geq 0

\text{Auxiliary solution to the explicit subproblem} \quad M_L\tilde{u}^n = b^n

proves to be positivity-preserving for \( \Delta t \leq \frac{1}{1-\theta} \min_i \left\{ -m_i / l_{ii} | l_{ii} < 0 \right\} , \quad 0 \leq \theta < 1 \)

\text{Correction factors} \quad \alpha_{ij}^{(m)} = \alpha_{ij}(\tilde{u}^n, f_{ij}^{(m)}), \quad 0 \leq \alpha_{ij} \leq 1 \quad \text{Zalesak’s limiter (1979)}

\text{Modified right-hand side} \quad \begin{align*}
b_i^{(m+1)} &= b_i^n + \sum_{j \neq i} \alpha_{ij}^{(m)} f_{ij}^{(m)}
\end{align*}

\textbf{Remark:} The admissible percentage of the antidiffusive flux depends on the magnitude of the time step \( \Delta t \).
Iterative FEM-FCT Algorithm

**Strategy:** Build accepted antidiffusion into $\tilde{u}$ and limit only the rejected portion.

**Variable** auxiliary solution

$$M_L \tilde{u}^{(m)} = b^{(m)}, \quad b^{(0)} = b^n$$

Rejected antidiffusive fluxes

$$\Delta f_{ij}^{(m)} = f_{ij}^{(m)} - g_{ij}^{(m)}, \quad \Delta f_{ij}^{(0)} = f_{ij}^{(0)}$$

Correction factors

$$\alpha_{ij}^{(m)} = \alpha_{ij}(\tilde{u}^{(m)}, \Delta f_{ij}^{(m)}), \quad 0 \leq \alpha_{ij}^{(m)} \leq 1$$

Accepted antidiffusive fluxes

$$g_{ij}^{(m+1)} = g_{ij}^{(m)} + \alpha_{ij}^{(m)} \Delta f_{ij}^{(m)}, \quad g_{ij}^{(0)} = 0$$

Modified right-hand side

$$b_i^{(m+1)} = b_i^{(m)} + \sum_{j \neq i} \alpha_{ij}^{(m)} \Delta f_{ij}^{(m)}$$

Consistency:

$$\alpha_{ij}^{(m)} \equiv 1 \quad \Rightarrow \quad b_i^{(m+1)} = b_i^n + \sum_{j \neq i} \left( g_{ij}^{(m)} + \Delta f_{ij}^{(m)} \right) = b_i^n + \sum_{j \neq i} f_{ij}^{(m)}$$

**Remark:** The iterative flux limiter makes it possible to ‘recycle’ the rejected antidiffusion.
Zalesak’s Flux Limiter

1. Positive/negative contributions

\[ P_i = P_i^+ + P_i^- \]  
where  
\[ P_i^\pm = \sum_{j \neq i}^{\max} \min \{0, f_{ij}\} \]

2. Maximum/minimum increment

\[ Q_i^\pm = \max_{\min} \Delta u_{ij}^\pm \]  
where  
\[ \Delta u_{ij}^\pm = \max_{\min} \{0, \tilde{u}_j - \tilde{u}_i\} \]

3. Nodal correction factors

\[ R_i^\pm = \begin{cases} 
\min\{1, m_i Q_i^\pm / P_i^\pm\}, & \text{if } P_i^\pm \neq 0 \\
1, & \text{if } P_i^\pm = 0 
\end{cases} \]

4. Final correction factors

\[ \alpha_{ij} = \begin{cases} 
\min\{R_i^+, R_j^-\}, & \text{if } f_{ij} \geq 0 \\
\min\{R_j^+, R_i^-\}, & \text{if } f_{ij} < 0 
\end{cases} \]

Remark: This choice of the correction factors guarantees that

\[ \tilde{u}_i^{(m)} + Q_i^- \leq \tilde{u}_i^{(m+1)} \leq \tilde{u}_i^{(m)} + Q_i^+ \]

so that no enhancement of local extrema takes place.
Solid Body Rotation

exact solution / initial data

iterative limiter

Crank-Nicolson time-stepping, \( \Delta t = 10^{-3} \), 16,384 \( Q_1 \) elements, at \( t = 2\pi \)

high-order vs. low-order scheme
Swirling Flow Problem

16,384 $Q_1$ elements

32,768 $P_1$ elements

Crank-Nicolson time-stepping, $\Delta t = 10^{-3}$, at $t = 2.5$

Initial data

Velocity field
Compressible Euler Equations

Divergence form
\[
\frac{\partial U}{\partial t} + \nabla \cdot F = 0
\]
where
\[
\nabla \cdot F = \sum_{d=1}^{3} \frac{\partial F^d}{\partial x_d}
\]

Conservative variables and fluxes
\[
U = \begin{bmatrix}
\rho \\
\rho v \\
\rho E
\end{bmatrix}, \quad F = (F^1, F^2, F^3) = \begin{bmatrix}
\rho v \\
\rho v \otimes v + pI \\
\rho H v
\end{bmatrix}
\]
\[
H = E + \frac{p}{\rho} \quad \gamma = c_p/c_v
\]

Equation of state
\[
p = (\gamma - 1)\rho(E - 0.5|v|^2) \quad \text{for a polytropic gas}
\]

Quasi-linear form
\[
\frac{\partial U}{\partial t} + A \cdot \nabla U = 0
\]
where
\[
A \cdot \nabla U = \sum_{d=1}^{3} A^d \frac{\partial U}{\partial x_d}
\]

Jacobian matrices
\[
A = (A^1, A^2, A^3), \quad F^d = A^d U, \quad A^d = \frac{\partial F^d}{\partial U}, \quad d = 1, 2, 3
\]
**High-Order Scheme**

Group FEM formulation

\[
M_C \frac{dU_d}{dt} = K U
\]

\[
(KU)_i = - \sum_{j \neq i} c_{ij} \cdot F_j = - \sum_{j \neq i} c_{ij} \cdot (F_j - F_i)
\]

due to the fact that \( \sum_{j} \varphi_j \equiv 1 \Rightarrow c_{ii} = - \sum_{j \neq i} c_{ij} \)

Roe averaging

\[
F_j - F_i = \hat{A}_{ij} (U_j - U_i)
\]

where

\[
\hat{A}_{ij} = A(\hat{\rho}_{ij}, \hat{v}_{ij}, \hat{H}_{ij})
\]

\[
\begin{align*}
\hat{\rho}_{ij} &= \sqrt{\rho_i \rho_j}, \\
\hat{v}_{ij} &= \frac{\sqrt{\rho_i v_i} + \sqrt{\rho_j v_j}}{\sqrt{\rho_i} + \sqrt{\rho_j}}, \\
\hat{H}_{ij} &= \frac{\sqrt{\rho_i H_i} + \sqrt{\rho_j H_j}}{\sqrt{\rho_i} + \sqrt{\rho_j}}
\end{align*}
\]

Quasi-linear formulation

\[
(KU)_i = - \sum_{j \neq i} c_{ij} \cdot \hat{A}_{ij} (U_j - U_i) = \sum_{j \neq i} (A_{ij} + B_{ij})(U_j - U_i)
\]

Cumulative Roe matrices

\[
A_{ij} = a_{ij} \cdot \hat{A}_{ij} = -A_{ji}, \quad a_{ij} = 0.5 \left( c_{ij} - c_{ji} \right)
\]

\[
B_{ij} = b_{ij} \cdot \hat{A}_{ij} = B_{ji}, \quad b_{ij} = 0.5 \left( c_{ij} + c_{ji} \right)
\]

Contribution of edge \( \vec{i}j \)

\[
(A_{ij} + B_{ij})(U_i - U_j) \longrightarrow (KU)_i
\]

\[
(A_{ij} - B_{ij})(U_i - U_j) \longrightarrow (KU)_j
\]
Galerkin Matrix Assembly

Edge contribution to the operator $K$

\[
K_{ii} = A_{ij} + B_{ij} \quad K_{ij} = -A_{ij} - B_{ij} \\
K_{ji} = A_{ij} - B_{ij} \quad K_{jj} = -A_{ij} + B_{ij}
\]

Edge contribution to the operator $L$

\[
L_{ii} = A_{ij} - D_{ij} \quad L_{ij} = -A_{ij} + D_{ij} \\
L_{ji} = A_{ij} + D_{ij} \quad L_{jj} = -A_{ij} - D_{ij}
\]

Raw antidiffusive flux

\[
F_{ij} = -[M_{ij} \frac{d}{dt} + D_{ij} + B_{ij}](U_j - U_i), \quad F_{ji} = -F_{ij}
\]

where $M_{ij} = m_{ij}I$ is a block of $M_C$ and $D_{ij}$ is the tensorial artificial diffusion.

Remark: Depending on the solution strategy (segregated/coupled), only ‘a few’ blocks of the global matrix need to be assembled and stored.
Design of Artificial Viscosities

**LED principle for systems:** (semi-discrete level)

Render all off-diagonal matrix blocks $L_{ij}$ positive semi-definite

Characteristic decomposition

$$A_{ij} = |a_{ij}| R_{ij} \Lambda_{ij} R_{ij}^{-1} \quad |a_{ij}| = \sqrt{a_{ij} \cdot a_{ij}}$$

where $R_{ij}$ is the matrix of right eigenvectors and the eigenvalues of $A_{ij}$ are given by

$$\Lambda_{ij} = \text{diag}\{\hat{v}_{ij} - \hat{c}_{ij}, \hat{v}_{ij}, \hat{v}_{ij}, \hat{v}_{ij} + \hat{c}_{ij}\}$$

Characteristic velocities

$$\hat{v}_{ij} = \frac{a_{ij}}{|a_{ij}|} \cdot \hat{v}_{ij}, \quad \hat{c}_{ij} = \sqrt{(\gamma - 1)(\hat{H}_{ij} - 0.5 |\hat{v}_{ij}|^2)}$$

**System upwinding** (expensive)

$$D_{ij} = |A_{ij}| = |a_{ij}| R_{ij} \Lambda_{ij} R_{ij}^{-1}$$

Generalization of Roe’s approximate Riemann solver (1981) to multidimensions

**Scalar dissipation** (efficient)

$$D_{ij} = d_{ij} I \quad \text{where} \quad d_{ij} = |a_{ij}| \max_i |\lambda_i|$$

Optimal choice for FCT since artificial diffusion is removed by the flux limiter
Iterative Defect Correction

(Preconditioned) defect correction

\[ U^{(m+1)} = U^{(m)} + [A(U^{(m)})]^{-1} R^{(m)}, \quad m = 0, 1, \ldots \]

\[ R^{(m)} = B^n - A(U^{(m)}) U^{(m)} + F(U^{(m)}, U^n) \]

Practical implementation

\[ A(U^{(m)}) \Delta U^{(m+1)} = R^{(m)}, \quad m = 0, 1, \ldots \]

\[ U^{(m+1)} = U^{(m)} + \Delta U^{(m+1)}, \quad U^{(0)} = U^n \]

How to solve this system?

\[
\begin{bmatrix}
A_{11}^{(m)} & A_{12}^{(m)} & A_{13}^{(m)} & A_{14}^{(m)} & A_{15}^{(m)} \\
A_{21}^{(m)} & A_{22}^{(m)} & A_{23}^{(m)} & A_{24}^{(m)} & A_{25}^{(m)} \\
A_{31}^{(m)} & A_{32}^{(m)} & A_{33}^{(m)} & A_{34}^{(m)} & A_{35}^{(m)} \\
A_{41}^{(m)} & A_{42}^{(m)} & A_{43}^{(m)} & A_{44}^{(m)} & A_{45}^{(m)} \\
A_{51}^{(m)} & A_{52}^{(m)} & A_{53}^{(m)} & A_{54}^{(m)} & A_{55}^{(m)}
\end{bmatrix}
\begin{bmatrix}
\Delta U_1^{(m+1)} \\
\Delta U_2^{(m+1)} \\
\Delta U_3^{(m+1)} \\
\Delta U_4^{(m+1)} \\
\Delta U_5^{(m+1)}
\end{bmatrix}
= \begin{bmatrix}
R_1^{(m)} \\
R_2^{(m)} \\
R_3^{(m)} \\
R_4^{(m)} \\
R_5^{(m)}
\end{bmatrix}
\]

Block-Jacobi method

\[ A_{kl}^{(m)} := \delta_{kl} A_{kl}^{(m)} \]

\[ A_{kk}^{(m)} \Delta U_k^{(m+1)} = R_k^{(m)}, \quad k = 1, \ldots, 5 \]

\[ U_k^{(m+1)} = U_k^{(m)} + \Delta U_k^{(m+1)}, \quad U_k^{(0)} = U_k^n \]

⊕ only 5 blocks need to be assembled and stored
⊕ equations can be solved separately or in parallel
⊕ poor/no convergence for large time steps due to increasing stiffness of the equations for Mach numbers near 0 or 1 → local preconditioning
Implementation of Characteristic Boundary Conditions

Numbers of PBC vs. NBC \( N_v = N_p + N_n \) depend on the local Mach number \( M = |v_n|/c \)

For all \( x_i \in \Gamma \) do

1. Nullify \( a_{ij}^{kl} := 0 \) \( \forall j \neq i, \forall l \neq k \)
   and update \( U_i^* = U_i^{(m)} + \text{diag}\{A_{ii}^{-1}\}R_i^{(m)} \)

2. Transform \( U_i^* \) into \( W_i^* \) and apply PBC
   for the incoming Riemann invariants

3. Transform \( W_i^{**} \) to \( U_i^{**} \) and nullify \( R_i^{(m)} := 0 \)

<table>
<thead>
<tr>
<th>( N_p/N_n )</th>
<th>1D</th>
<th>2D</th>
<th>3D</th>
</tr>
</thead>
<tbody>
<tr>
<td>subsonic inflow</td>
<td>2/1</td>
<td>3/1</td>
<td>4/1</td>
</tr>
<tr>
<td>subsonic outflow</td>
<td>3/0</td>
<td>4/0</td>
<td>5/0</td>
</tr>
<tr>
<td>supersonic inflow</td>
<td>1/2</td>
<td>1/4</td>
<td>1/4</td>
</tr>
<tr>
<td>supersonic outflow</td>
<td>0/3</td>
<td>0/4</td>
<td>0/5</td>
</tr>
</tbody>
</table>

Variable transformations

\( U_i^{(m)} \to U_i^* \to W_i^* \to U_i^{**} \to W_i^{**} \)

⊕ Values \( U_i^{**} \) represent Dirichlet boundary conditions for the end-of-step solution \( U_i^{(m+1)} \)
⊕ No ad hoc extrapolation of data from the interior
⊕ Easy to implement as a ‘black-box’ module
Shock Tube Problem

Crank-Nicolson time-stepping, Δt = 10^{-3}, 16,384 \ Q_1 elements at t = 0.231
Radially Symmetric Riemann Problem

Crank-Nicolson time-stepping, $\Delta t = 10^{-3}$, 16,384 $Q_1$ elements at $t = 0.13$
Compression Corner \( M_\infty = 2.5, \theta = 15^\circ \)

Low-order method, scalar dissipation

FEM-FCT, \( \alpha = \min\{\alpha_\rho, \alpha_E\} \)

Backward Euler time-stepping, \( \Delta t = 10^{-2}, 16,384 \, Q_1 \) elements
Compression Corner \( M_\infty = 2.5, \theta = 15^\circ \)

Low-order method, scalar dissipation

FEM-FCT, \( \alpha = \min\{\alpha_\rho, \alpha_E\} \)

Backward Euler time-stepping, \( \Delta t = 10^{-2} \), 10,016 \( Q_1 \) elements
**Prandtl-Meyer Expansion**

$M_\infty = 2.5, \theta = 15^\circ$

- Low-order method, scalar dissipation
- FEM-FCT, $\alpha = \min\{\alpha_\rho, \alpha_E\}$

Backward Euler time-stepping, $\Delta t = 10^{-2}$, 16,384 $Q_1$ elements
Algebraic Flux Correction: FCT vs. TVD

• Both node-oriented flux limiters act at the algebraic level which makes them applicable to ‘arbitrary’ discretizations in time (explicit/implicit) and space (FD/FV/FE)

• Fully discrete FCT-type schemes, which are positivity preserving by construction, are to be recommended for time dependent problems

• TVD-type methods are preferable for the treatment of stationary flows. They are derived at the semi-discrete level but preserve positivity only upon convergence

Outstanding tasks

• Extension of the methodology to higher-order finite elements

• Robust and efficient iterative solvers for nonsymmetric algebraic systems

• Understanding of clipping and terracing phenomena

• Adaptive grid refinement for flux limiting schemes