Outflow positivity limiting for hyperbolic systems

E. Alec Johnson

Department of Mathematics
KU Leuven, Belgium

Oct 18, 2012

Presented at RWTH Aachen University, Aachen, Germany

[work with James Rossmanith, Iowa State University, USA]

[first presented on Sept. 13 at NUMDIFF-13, Halle(Saale), Germany]

Outline

1. Problem
2. Recipes
3. Why the recipes work
4. Implementation

Hyperbolic conservation laws

Assume a hyperbolic conservation law of the form

$$\partial_t \underline{u}(t, \mathbf{x}) + \nabla \cdot \underline{\mathbf{f}}(t, \mathbf{x}, \underline{u}) = 0.$$

Physical solutions can develop shocks but always satisfy the integral form

$$\frac{\mathrm{d}}{\mathrm{d}_t} \int_K \underline{u} + \oint_{\partial K} \hat{\mathbf{n}} \cdot \underline{\mathbf{f}} = 0 \quad \text{for any region } K. \tag{1}$$

Assume that physical solutions remain in a ***convex cone*** $\mathcal{P}$.

Positivity limiters yield a numerical method that:

1. satisfies a *discrete local conservation law* like (1),
2. is *high-order accurate* for smooth solutions, and
3. satisfies a *discrete local positivity condition* $⨍_K \underline{u} \in \mathcal{P}$,

where for the discrete versions the region K is a ***mesh cell***.

Benefits:

1. Conservation: correct shock speeds
2. Accuracy: physicality
3. Positivity: stability

Simplification: scalar case

In the scalar case, we want to enforce $u \geq 0$ for a scalar conservation law:

$$\partial_t u(t, \mathbf{x}) + \nabla \cdot \mathbf{f}(t, \mathbf{x}, u) = 0.$$

Observation: In each time step the outflow from each cell must be less than its initial content.

This leads to a simple framework we call **outflow positivity limiting**.

What about the systems case?

The systems case reduces to the scalar case!

- Any convex cone $\mathcal{P}$ is an intersection of half-spaces.
- Each half-space is the set on which a linear functional Λ is positive.
- Composing Λ with a hyperbolic system yields a scalar PDE:

$$\partial_t(\Lambda \underline{u})(t, \mathbf{x}) + \nabla \cdot (\Lambda \underline{\mathbf{f}})(t, \mathbf{x}, \underline{u}) = 0.$$

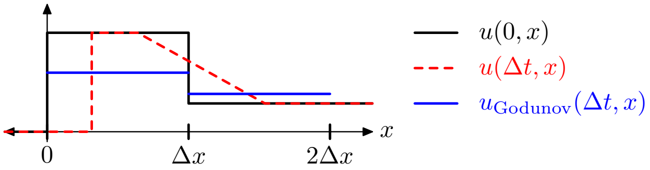
Outline

Godunov method

- Define the **speed cap** λ to be an upper bound on twice the signal speed;
- Define the **scale-invariant speed cap** to be $\bar{\lambda} := \frac{\lambda A}{V}$;
- Define the **positivity CFL number** of a time step Δt to be $\mathrm{CFL_{pos}} := \Delta t \bar{\lambda}$ (e.g. $\frac{2\Delta t \lambda}{\Delta x}$).[1]

Godunov maintains positivity of the cell average by repeating the following sequence:

1. Reset the solution to the cell average.
2. "Physically evolve" for a time step for which $\mathrm{CFL_{pos}}$ is at most 1.
3. Find the exact flux of the Riemann problem at each interface.



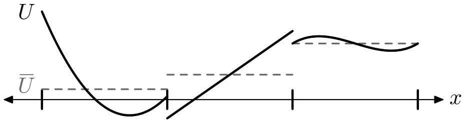
Success:

- conservative
- solution remains positive (because physical solutions remain positive).
- positivity-preserving time step is maximum attained by explicit methods.

Problem: *Solution is only first-order accurate.*

[1] An extra factor of two is needed because we assume a spatially varying flux function.

DG/WENO updates the average in each mesh cell K using a piecewise continuous solution representation U which is a high-order polynomial in each mesh cell:



$$\frac{\mathrm{d}}{\mathrm{d}t}\int_K U + \oint_{\partial K}^{\mathrm{Q}} h(U^-, U^+) = 0 \quad \text{(method-of-lines ODE)},$$

$$\int_K U^{n+1} = \int_K U - \Delta t \oint_{\partial K}^{\mathrm{Q}} h \quad \text{(Euler step)},$$

where $U|_K \in \mathcal{V}$ (a finite-dimensional polynomial representation space), $\oint_{\partial K}^{\mathrm{Q}} h \approx \oint_{\partial K} h$, and $h(U,U) = \hat{\mathbf{n}} \cdot \mathbf{f}(U)$.

Success:

- conservative
- stable if $\frac{\Delta t \lambda_{\max} A}{2DV} \leq \mathrm{CFL}_{\mathrm{stable}}$

Problem: ***Cell averages can become negative***, causing instability.

How do cell averages go negative?

Pick a mesh cell K. Look at evolution of the cell integral.

Definitions

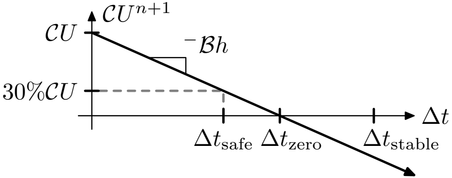
- $\mathcal{C}U = \int_K U$: **cell content**
- $\mathcal{B} = \oint_{\partial K}^{\mathrm{Q}}$: **boundary sum**
- $h(U^-, U^+, \hat{\mathbf{n}})$: **numerical outgoing flux**

Method of Lines says: $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{C}U = -\mathcal{B}h$.

For an Euler step of length Δt, the cell average changes linearly with time:

$$\boxed{\mathcal{C}U^{n+1} = \mathcal{C}U - \Delta t \mathcal{B}h}$$

Outflow capping guarantees positivity by directly limiting outflow:



Capping the outflow at $\alpha_z = 70\%$ gives:

$$\Delta t_{\mathrm{zero}}^{-1} = \frac{\mathcal{B}h}{\mathcal{C}U},$$

$$\Delta t_{\mathrm{safe}}^{-1} = \max\left\{(\alpha_z \Delta t_{\mathrm{zero}})^{-1}, \Delta t_{\mathrm{stable}}^{-1}\right\}$$

DG/WENO method with **outflow capping**

Success:

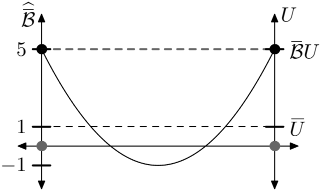
- (conservative)
- (stable if $\frac{\Delta t \lambda_{\mathsf{max}} A}{2DV} \le \mathrm{CFL}_{\mathrm{stable}}$)
- positivity-preserving

Problem: ***safe time step can go to zero,*** *causing the simulation to grind to a halt.*

Why?

Boundary crowding

One culprit is **boundary crowding:**



Update of cell comes from:

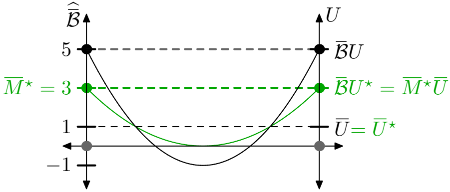
1. initial cell average
2. values at boundary nodes (proportional to potential outflow).

If the solution is not constrained by a cell positivity condition, then the ratio of cell outflow rate to cell average can become arbitrarily large.

Key definitions:

- $\overline{\mathcal{B}}U$: **boundary average**
- $\overline{\mathcal{C}}U = \overline{U}$: **cell average**
- $\widehat{\overline{\mathcal{B}}}(U) := \frac{\overline{\mathcal{B}}U}{\overline{\mathcal{C}}U}$: **boundary crowding** (ratio of boundary average to cell average)
- $U^\star$: positive solution with maximum boundary crowding
- $M^\star$: maximum boundary crowding for a positive solution

One culprit is **boundary crowding:**



Update of cell comes from:

1. initial cell average
2. values at boundary nodes (proportional to potential outflow).

If the solution is not constrained by a cell positivity condition, then the ratio of cell outflow rate to cell average can become arbitrarily large.

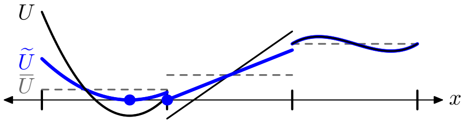
Key definitions:

- $\overline{\mathcal{B}}U$: **boundary average**
- $\overline{\mathcal{C}}U = \overline{U}$: **cell average**
- $\widehat{\overline{\mathcal{B}}}(U) := \frac{\overline{\mathcal{B}}U}{\overline{\mathcal{C}}U}$: **boundary crowding** (ratio of boundary average to cell average)
- $U^\star$: positive solution with maximum boundary crowding
- $M^\star$: maximum boundary crowding for a positive solution

Idea: ***cap the boundary crowding*** by $M^\star$.

Let's look at three ways...

Liu and Osher damp the deviation from the cell average just enough to enforce positivity at **every point**:

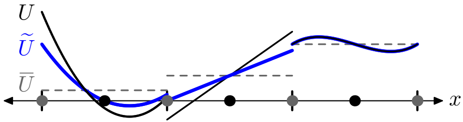


Success:

- (conservative)
- preserves high-order accuracy
- preserves positivity of the cell average if $\mathrm{CFL}_{\mathrm{pos}} \leq \overline{W}^{\star} := 1/\overline{M}^{\star}$.
- Essentially guarantees the largest positivity-preserving time step we can hope for.

Problem: ***hard to find the minimum of a high-order polynomial***, *especially in multiple dimensions.*

Zhang and Shu damp the deviation from the cell average just enough to enforce positivity at **positivity points**:



- Let X be the set of positivity points.
- Let $\overline{M}_X$ be the maximum boundary crowding over solutions positive at each positivity point.

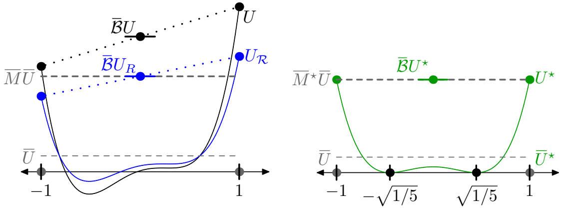
Success:

- (conservative)
- (preserves high-order accuracy)
- preserves positivity if $\mathrm{CFL}_{\mathrm{pos}} \leq \overline{W}_X := 1/\overline{M}_X$
- $\overline{W}_X = \overline{W}^\star$ if X is the set of **optimal positivity points**.
- $\overline{W}_X > 0$ if X is rich enough (to be capable of representing the solution and if the cell average can be represented as a strictly positive combination of values from X).

Problem: ***what are the positivity points?***

- $\overline{W}_X < \overline{W}^\star$ unless the optimal positivity points are included.
- The optimal points can increase computational expense by requiring evaluation at points not otherwise used in the DG scheme, particularly for high-order-accurate solutions.
- Optimal points for boxes and simplices are unknown for high-order polynomials.

We damp the deviation from the cell average just enough to cap the boundary average $\overline{\mathcal{B}}U$ at a multiple $\overline{M}$ of the cell average $\overline{U}$.



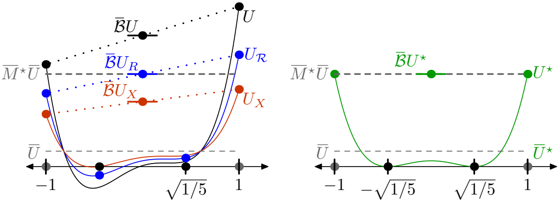
Success:

- (conservative)
- preserves positivity for an Euler step if $\mathrm{CFL_{pos}} \le \overline{W} := 1/\overline{M}$.
- preserves high-order accuracy if $\overline{M} \ge \overline{M}^\star$.
- cheap: no need to evaluate at points not otherwise used in the DG scheme.

Problem: ***what is the maximum boundary average?***

- need an upper bound $\overline{M}$ on $\overline{M}^\star$.

We damp the deviation from the cell average just enough to cap the boundary average $\overline{\mathcal{B}}U$ at a multiple $\overline{M}$ of the cell average $\overline{U}$.



Success:

- (conservative)
- preserves positivity for an Euler step if $\mathrm{CFL}_{\mathrm{pos}} \leq \overline{W} := 1/\overline{M}$.
- preserves high-order accuracy if $\overline{M} \geq \overline{M}^\star$.
- cheap: no need to evaluate at points not otherwise used in the DG scheme.

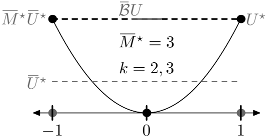
Problem: ***what is the maximum boundary average?***

- need an upper bound $\overline{M}$ on $\overline{M}^\star$.

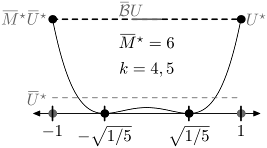
Boundary crowding: optimal positivity points and $\overline{M}^\star$ values

For an interval:

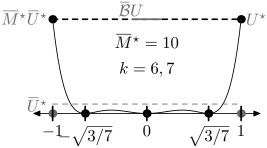
cubic polynomial space:



quintic polynomial space:



heptic polynomial space:



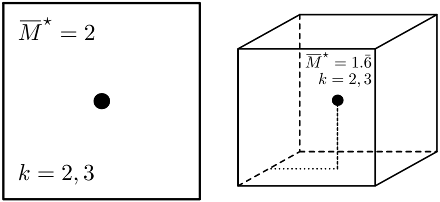
- The optimal boundary crowding cap is $\overline{M}^\star = (n+1)(n+2)/2$, where the representation space is the polynomials of degree at most $k = 2n$ or $k = 2n+1$.
- The polynomial $U^\star$ which maximizes the boundary crowding is zero at the *optimal interior points*.
- Enforcing positivity at the optimal interior points enforces the optimal boundary crowding cap.
- The *optimal positivity points* are the *Gauss-Lobatto* quadrature points (for a correct rule with the fewest points). See [ZhangShu10].

What about regular polytopes?...

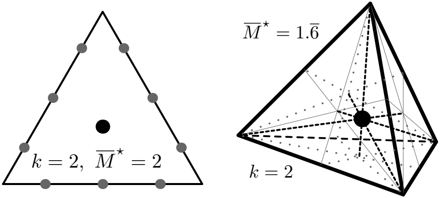
(Recall definitions for a mesh cell:)

- $\widehat{\overline{\mathcal{B}}}(U) := \frac{\overline{\mathcal{B}U}}{\overline{U}}$: **boundary crowding:** (ratio of boundary average to cell average)
- $\overline{M}^\star$: **optimal boundary crowding cap**: maximum boundary crowding of a positive solution.
- $U^\star$: **boundary crowding maximizer**: positive solution which maximizes $\widehat{\overline{\mathcal{B}}}(U)$

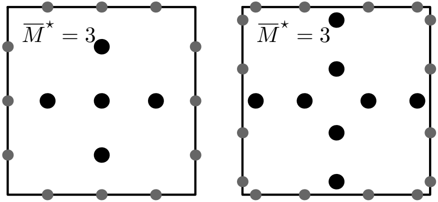
Optimal positivity points and $\overline{M}^{\star}$ values (key results)



For quadratic and cubic polynomials in a box, the set of optimal positivity points is simply the cell center.

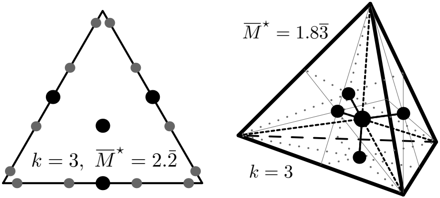


For quadratic polynomials in a simplex, the set of optimal positivity points is simply the cell center.



$\mathcal{V} = \mathbb{P}^2(x) \cdot \mathbb{P}^2(y)$ $\qquad$ $\mathcal{V} = \mathbb{P}^3(x) \cdot \mathbb{P}^3(y)$

For tensor product polynomial spaces in a box, [ZhangShu10] gives the optimal positivity points.



For cubic polynomials in a simplex, the optimal positivity points are the cell center and face centers.

Definitions for common canonical mesh cells and representation spaces

What about polynomial representations higher than third-order?

- **Representation space:**

  $\mathbb{P}^k_D := \{\text{polynomials in } D \text{ variables of degree at most } k\}$.

- **Canonical mesh cells:**

$$
\begin{aligned}
\overline{M}^k_{[0,1]} &:= \overline{M}^\star \text{ for } \mathbb{P}^k_D \text{ for a } \textbf{unit interval} \\
\overline{M}^k_{\bigcirc^D} &:= \overline{M}^\star \text{ for } \mathbb{P}^k_D \text{ for a } \textbf{sphere} \\
\overline{M}^k_{[0,1]^D} &:= \overline{M}^\star \text{ for } \mathbb{P}^k_D \text{ for a } \textbf{box} \\
\overline{M}^k_{\triangle^D} &:= \overline{M}^\star \text{ for } \mathbb{P}^k_D \text{ for a } \textbf{simplex}.
\end{aligned}
$$

Bounds on $\overline{M}^\star$ for boxes (key results)

For high-order polynomials we have calculated bounding intervals containing $\overline{M}^\star$ for boxes:

k (polynomial order) :	$0,1$	$2,3$	$4,5$	$6,7$	$8,9$	$10,11$
$n = \lfloor k/2 \rfloor$:	0	1	2	3	4	5
$m = \lfloor n/2 \rfloor$:	0	0	1	1	2	2
$\overline{M}^k_{[0,1]} = \frac{(n+1)(n+2)}{2}$ [ZhSh10]	1	3	6	10	15	21
$\overline{M}^k_{\bigcirc^2} = (m+1) \cdot \lfloor \frac{n+3}{2} \rfloor$	1	2	4	6	9	12
$\overline{M}^k_{\bigcirc^3} = \frac{m+1}{3} \cdot \left(3 + 2\lfloor \frac{n+1}{2} \rfloor\right)$	1	$1.\bar{6}$	$3.\bar{3}$	$4.\bar{6}$	7	9
$\overline{M}^k_{[0,1]^2}$	1	2	$[3.5, 4]$	$[5.5, 6]$	$[8, 9]$	$[11, 12]$
$\overline{M}^k_{[0,1]^3}$	1	$1.\bar{6}$	$[2.\bar{6}, 3.\bar{3}]$	$[4, 4.\bar{6}]$	$[5.\bar{6}, 7]$	$[7.\bar{6}, 9]$

More generally, we have the bounds

$$\Omega(k^2) = \overline{M}^{k,-}_{[0,1]^D} := \frac{1+(D-1)\overline{M}^{k,-}_{[0,1]^{D-1}}}{D} \le \overline{M}^{2n}_{[0,1]^D} = \overline{M}^{2n+1}_{[0,1]^D} \le \overline{M}^{2n}_{\bigcirc^D} = \overline{M}^{2n+1}_{\bigcirc^D} = \frac{(\lfloor \frac{n}{2} \rfloor + 1)(2\lfloor \frac{n+1}{2} \rfloor + D)}{D} = \Omega(k^2);$$

note that $\overline{M}^{k,-}_{[0,1]^1} = \overline{M}^k_{[0,1]} = \frac{(n+1)(n+2)}{2}$.

Bounds on $\overline{M}^\star$ for simplices (key results)

For high-order polynomials we have calculated bounding intervals containing $\overline{M}^\star$ for simplices:

k (polynomial order) :	0	1	2	3	4	5	6	7
$n = \lfloor k/2 \rfloor$:	0	0	1	1	2	2	3	3
$\overline{M}^k_{\triangle^1} = \overline{M}^k_{[0,1]} = \frac{(n+1)(n+2)}{2}$	1	1	3	3	6	6	10	10
$\overline{M}^k_{\triangle^2}\left(\leq \overline{M}^k_{[0,1]}, \text{see [ZXS12]}\right)$	1	1	2	$2.\bar{2}$	$[3.4, 6]$	$[3.5, 6]$	$[5+\frac{1}{7}, 10]$	$[5.25, 10]$
$\overline{M}^k_{\triangle^3}\left(\leq \frac{n+1}{3}(\lfloor \frac{k+1}{2} \rfloor + 3)\right)$	1	1	$1.\bar{6}$	$1.8\bar{3}$	$[2.5\bar{6}, 5]$	$[2.\bar{6}, 6]$	$[3+\frac{9}{14}, 8]$	$[3.75, 9.\bar{3}]$

More generally, we have the bounds

$$\Omega(k^2) = \overline{M}^{k,-}_{\triangle^D} := \left(\frac{D+k}{D}\right)\left(\frac{1 + D\frac{D-1}{D+k-1}\overline{M}^{k,-}_{\triangle^{D-1}}}{1+D}\right) \leq \overline{M}^k_{\triangle^D} \leq \frac{(\lfloor \frac{k}{2} \rfloor + 1)(\lfloor \frac{k+1}{2} \rfloor + D)}{D} = \Omega(k^2);$$

note that $\overline{M}^{k,-}_{\triangle^1} = \overline{M}^k_{[0,1]} = \frac{(n+1)(n+2)}{2}$.

Outline

Outflow positivity limiting

Question: Why can these positivity limiting frameworks guarantee a minimum positivity-preserving time step?

Answer: Because ***capping the boundary average caps the maximum rate of outflow*** of material from the cell.

Outflow positivity limiting caps the amount of material that flows out of the mesh cell by the amount of material in the mesh cell:

1. Outflow rate determines Euler time step that violates positivity:
$$1/\Delta t_{\mathrm{zero}} = \frac{\mathcal{B}h}{\mathcal{C}U} = \frac{\overline{\mathcal{B}}\bar{h}}{\overline{\mathcal{C}}U}, \qquad \text{where } \bar{h} := \frac{hA}{V}.$$
2. Wave speed times solution caps outflow rate at a node:
$$h(U^-, U^+) \leq \lambda U^-.$$
3. To cap cell outflow rate, cap boundary average and wave speeds:
$$\overline{\mathcal{B}}\bar{h}(U^-, U^+) \leq \bar{\lambda}\overline{\mathcal{B}}U^-.$$

Theorem

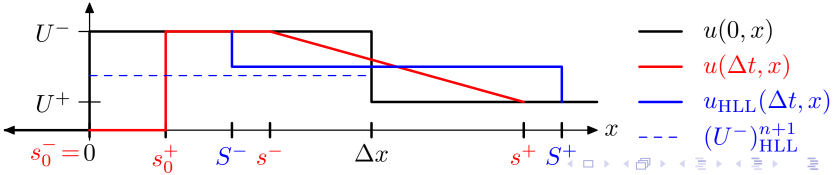
Let $U^-, U^+ \ge 0$. Let the speed cap $\lambda > 0$ be an upper bound on the sum of the right-going signal speed for the Riemann problem with states $(0, U^-)$ and the left-going signal speed for the Riemann problem with states (U^-, U^+). Let h be a numerical flux function which preserves positivity for a step of the Godunov method if the time step is short enough that signals do not cross.

Then $h(U^-, U^+) \le \lambda U^-$.

Proof: Consider the 1D problem

$$\partial_t u + \partial_x f(u) = 0, \qquad u(0,x) = \begin{cases} 0 & \text{if } x < 0, \\ U^- & \text{if } 0 < x < \Delta x, \\ U^+ & \text{if } \Delta x < x, \end{cases}$$

where $U^-, U^+ \ge 0$, $\Delta x := \lambda \Delta t$, and $f := u \mapsto \hat{\mathbf{n}} \cdot \mathbf{f}(t, \mathbf{x}, u)$. Suppose that for any $U^-, U^+ \ge 0$ an Euler update maintains positivity: $(U^-)^{n+1} = U^- - \frac{\Delta t}{\Delta x}\left[h(U^-, U^+) - h(0, U^-)\right] \ge 0$. Since material cannot flow out of a vacuum, $h(0, U^-) \le 0$. Therefore, $h(U^-, U^+) \le \frac{\Delta x}{\Delta t} U^- = \lambda U^-$, as desired.



Definitions for outflow analysis:

We define all quantities in the coordinates of a single *canonical* mesh cell K (so results go through for isoparametric mesh cells).

$\mathcal{B}U := \oint_{\partial K}^{\mathrm{Q}} U^-$	boundary integral quadrature
$\mathcal{C}U := \int_K U$	cell integral
λ	twice cap on speeds at boundary nodes
h	numerical flux ($h \approx \mathbf{f} \cdot \hat{\mathbf{n}}$)

Retentional positivity limiting

Theorem (Retentional positivity guarantees a positivity-preserving time step.)

Given:

$$\begin{aligned} h &\le \lambda U^- && \textit{(numerical flux is bounded via the speed cap } \lambda\textit{)}, \\ \lambda \Delta t &\le W && \textit{(step length is bounded via } W\textit{), and} \\ W\mathcal{B}U &\le \mathcal{C}U && \textit{(boundary crowding is capped by } M := W^{-1}\textit{).} \end{aligned}$$

Then: The loss is at most the cell content.

Proof:

$$\text{loss} = \Delta t \mathcal{B} h \le \lambda \Delta t \mathcal{B} U \le W\mathcal{B}U \le \mathcal{C}U. \quad \square$$

Remarks:

- Define the **retentional** $\mathcal{R} := \mathcal{C}U - W\mathcal{B}U$
  - $\mathcal{R}$ equals the material retained if the maximum possible loss occurs.
  - $\mathcal{R}$ is a linear functional.
- Enforcing positivity of $\mathcal{R}$ does not compromise accuracy:
  - if $\mathcal{R}$ is positive for all U positive in K, i.e.,
  - if $M \ge \frac{\mathcal{B}U}{\mathcal{C}U}$ for all nonzero solutions U positive in K.

Retentional positivity limiting (invariant version)

Theorem (Retentional positivity guarantees a positivity-preserving time step.)

Given:

$$\begin{aligned} \bar{h} \le \bar{\lambda} U^{-} & \quad \textit{(numerical flux is bounded via the speed cap } \bar{\lambda}\textit{)}, \\ \bar{\lambda}\Delta t \le \overline{W} & \quad \textit{(step length is bounded via } \overline{W}\textit{), and} \\ \overline{W}\overline{\mathcal{B}}U \le \bar{\mathcal{C}}U & \quad \textit{(boundary crowding is capped by } \overline{M} := \overline{W}^{-1}\textit{).} \end{aligned}$$

Then: The loss is at most the cell content.

Proof:

$$\text{loss} = \Delta t \overline{\mathcal{B}} \bar{h} \le \bar{\lambda} \Delta t \overline{\mathcal{B}} U \le \overline{W}\overline{\mathcal{B}}U \le \bar{\mathcal{C}}U. \quad \square$$

Remarks:

- Define the **retentional** $\overline{\mathcal{R}} := \bar{\mathcal{C}}U - \overline{W}\overline{\mathcal{B}}U$
  - $\overline{\mathcal{R}}$ equals the material retained if the maximum possible loss occurs.
  - $\overline{\mathcal{R}}$ is a linear functional.
- Enforcing positivity of $\overline{\mathcal{R}}$ does not compromise accuracy:
  - if $\overline{\mathcal{R}}$ is positive for all U positive in K, i.e.,
  - if $\overline{M} \ge \frac{\overline{\mathcal{B}}U}{\bar{\mathcal{C}}U}$ for all nonzero solutions U positive in K.

Definitions (scale-invariant):

Results are independent of the scale of the canonical mesh cell if we use scale-invariant definitions:

$\overline{\mathcal{B}}U := A^{-1}\mathcal{B}U$	**boundary averaging quadrature**
$\overline{\mathcal{C}}U := V^{-1}\mathcal{C}U$	cell average
$\bar{\lambda} := \frac{\lambda A}{V}$	scaled speed cap
$\bar{h} := \frac{hA}{V}$	scaled numerical flux
V	volume of cell
A	area of boundary

Definitions (affine-invariant):

Results are invariant under affine transformations of the canonical mesh cell if we use affine-invariant definitions:

$\bar{\mathcal{B}}U$	***arithmetic average over all faces of averaging quadrature on each face***
$\bar{\mathcal{C}}U$	cell average
$\bar{\lambda}_e := \frac{\lambda N \mathrm{d}A_e}{V}$	scaled speed cap at node $\mathbf{x}_e$
$\bar{h}_e := \frac{h N \mathrm{d}A_e}{V}$	scaled numerical flux at node $\mathbf{x}_e$
V	volume of cell K
$\mathrm{d}A_e$	area of face of node $\mathbf{x}_e$
N	number of faces

These definitions are designed to agree with the scale-invariant definitions for a regular polytope.

Outline

Systems case reduces to the scalar case

In the systems case,

$$\partial_t \underline{u}(t,\mathbf{x}) + \nabla \cdot \underline{\mathbf{f}}(t,\mathbf{x},\underline{u}) = \underline{0},$$

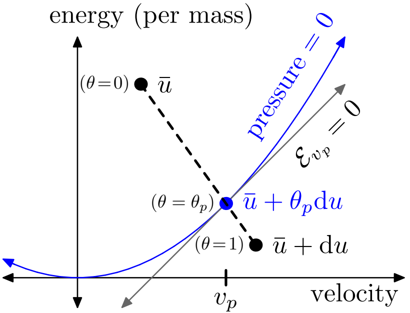
assume that positivity of the state $\underline{u}$ is defined by positivity of $\Lambda\underline{u}$ for all Λ in a collection of linear functionals. (Equivalently, the set of positive states $\mathcal{P}$ is a convex cone.) All statements made so far go through to the systems case after applying any Λ to both sides of the equation or inequality.

Caveats:

1. ***Wave speeds may need to be desingularized.*** Outflow rate limiting requires a finite cap on wave speeds. Enforcing positivity of the depth (shallow water) or density (gas dynamics) generically results in near-infinite wave speeds. A fix is to calculate fluxes with remapped states that diminish fluid speed and temperature to physically justified maxima.
2. ***Positivity indicators are needed.*** For gas dynamics, Λ comes from an infinite collection of linear functionals: the energy density in all possible reference frames. To test positivity computationally, use a *finite* set of *state positivity indicators*: density (linear) and pressure (concave).
3. ***Accuracy can be lost if the pressure is not strictly bounded away from zero.*** For gas dynamics, positivity limiting can diminish the order of accuracy. For example, enforcing positivity of the pressure can result in arbitrarily large damping of density variation.

Positivity limiting for systems

Positivity limiting requires determining *where a linear path in state space intersects the boundary of positive states*:

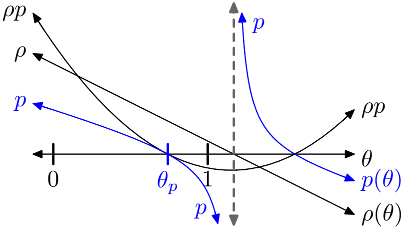


Need positivity of energy in every reference frame (an infinite collection of linear functionals).

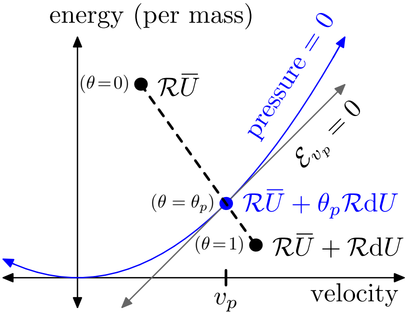
Instead use a *finite* set of nonlinear positivity indicators.

Gas-dynamic positivity indicators:

- ρ is linear
- p is concave if $\rho > 0$.
- $\frac{\rho p}{\gamma - 1} = \rho\mathcal{E} - \mathbf{m}^2/2$ is quadratic.



Positivity limiting requires determining *where a linear path in state space intersects the boundary of positive states*:

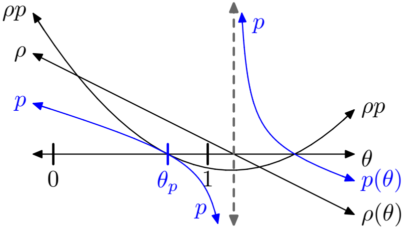


Need positivity of energy in every reference frame (an infinite collection of linear functionals).

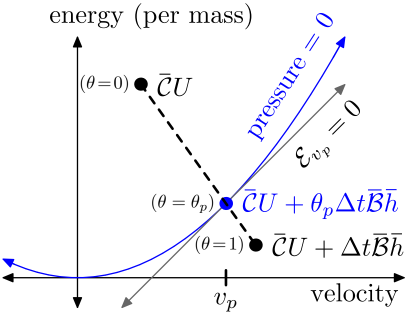
Instead use a *finite* set of nonlinear positivity indicators.

Gas-dynamic positivity indicators:

- ρ is linear
- p is concave if $\rho > 0$.
- $\frac{\rho p}{\gamma-1} = \rho\mathcal{E} - \mathbf{m}^2/2$ is quadratic.



Positivity limiting requires determining *where a linear path in state space intersects the boundary of positive states*:

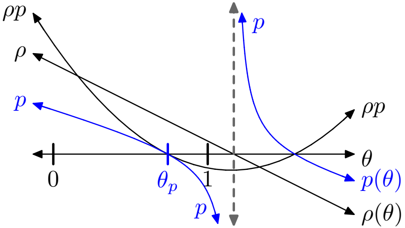


Need positivity of energy in every reference frame (an infinite collection of linear functionals).

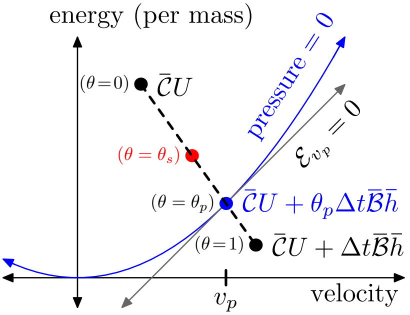
Instead use a *finite* set of nonlinear positivity indicators.

Gas-dynamic positivity indicators:

- ρ is linear
- p is concave if $\rho > 0$.
- $\frac{\rho p}{\gamma - 1} = \rho\mathcal{E} - \mathbf{m}^2/2$ is quadratic.



Positivity limiting requires determining *where a linear path in state space intersects the boundary of positive states*:

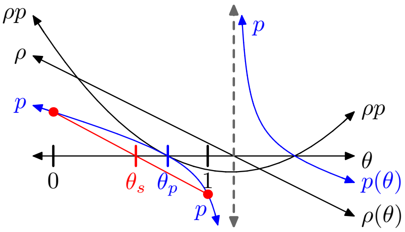


Need positivity of energy in every reference frame (an infinite collection of linear functionals).

Instead use a *finite* set of nonlinear positivity indicators.

Gas-dynamic positivity indicators:

- ρ is linear
- p is concave if $\rho > 0$.
- $\frac{\rho p}{\gamma - 1} = \rho\mathcal{E} - \mathbf{m}^2/2$ is quadratic.



Prescriptions for implementation: efficiency and stability

- ***Enforce positivity at boundary nodes.***
  In addition to enforcing positivity of the retentional, enforce positivity at the boundary quadrature points so that numerical fluxes are computed using positive states.

- ***Enforce positivity at nodal points to efficiently improve stability.***
  Since values used in volume and boundary quadratures must be calculated (or available) anyway, positivity of these values can be efficiently enforced, with the added benefit that positivity of all states actually used in computation is assured.

- ***Check positivity efficiently.***
  Use interval arithmetic to inexpensively confirm cell positivity in the vast majority of mesh cells. For nodal DG, choose positivity points to be nodal points to avoid additional computational expense. Precompute the boundary average of each polynomial basis function to efficiently evaluate the retentional.

- ***Cap outflow to estimate an optimal time step.***
  One can directly calculate the maximum stable time step for which an Euler step maintains positivity of the cell average. For multistage and local time stepping, one can maintain and iteratively adjust an estimate of a safe time step that is both stable and positivity-preserving.

Prescriptions for implementation: canonical regular polytope

- ***Use an affine-invariant definition of the boundary average.***
  The values of $\overline{M}^{\star}$ given here assume a regular simplex or box. If using a canonical simplex which is not regular, then for the retentional $\overline{\mathcal{R}} := \overline{M}\overline{\mathcal{C}} - \overline{\mathcal{B}}$, define $\overline{\mathcal{B}}$ to be the arithmetic average over all faces of the average on each face.

- ***Work in canonical coordinates.***
  When working with isoparametric mesh cells, work in canonical coordinates, where the mesh cell is a simplex or box and the representation space is $\mathbb{P}_D^k$ or a tensor product polynomial space, so that the optimal positivity points and retentional given here can be used without modification.

Prescriptions for implementation: shallow water and gas dynamics

- ***Use wave speed desingularization for systems, especially for shallow water.***
  When enforcing positivity of depth (shallow water) or density (gas dynamics), fluxes need to be calculated with remapped states in order to desingularize wave speeds. In primitive variables one can simply damp temperature and fluid speed to enforce a physically justified cap estimated e.g. via a first-order solver invoked where desingularization may be needed.

- ***Enforce positivity of the density, then the pressure.***
  For gas dynamics, first apply linear damping to enforce positivity of the density. Then apply linear damping to the state to enforce positivity of the pressure p. Note that, as a function of the state variables, $\rho p = (\gamma - 1)\left(\rho\mathcal{E} - \mathbf{M}^2/2\right)$ is quadratic and p is concave (if $\rho > 0$), allowing one to enforce positivity of p by solving a quadratic or linear polynomial equation. See [ZhangShu10].

- ***Pad inequalities.***
  To guard against error in machine arithmetic, enforce $\rho \geq \epsilon_\rho$ or $p \geq \epsilon_p$ for some small $\epsilon_\rho > 0$ and $\epsilon_p > 0$.

Prescriptions for implementation: higher-moment positivity limiting

- ***Use linear damping to maintain positive-definiteness of the pressure tensor.***

  Positive-definiteness of the pressure tensor can be enforced in a manner similar to positivity of the pressure: First enforce positivity of the density. Then enforce positivity of the pressure tensor. Use that: (1) as a function of the state variables, $\rho\mathbb{P} = \rho\mathbb{E} - \mathbf{MM}$ is quadratic and $\hat{\mathbf{n}} \cdot \mathbb{P} \cdot \hat{\mathbf{n}}$ is concave for any $\hat{\mathbf{n}}$ (if $\rho > 0$); and (2) $\mathbb{P} > 0$ iff $\operatorname{tr}\mathbb{P} > 0$, $\operatorname{tr}\operatorname{adj}\mathbb{P} > 0$, and $\det\mathbb{P} > 0$.

- ***Use remapping to project states into the the interior of the domain of hyperbolicity or realizability.***

  The invariant domain of some 13-moment models is not convex. Use positivity limiters to keep the solution *physically* realizable (a convex superset). Then damp heat flux to project the cell average and other states into the invariant domain while preserving mass, momentum, and energy. (This does not violate physical constraints, since collisions do not conserve heat flux.)

  Singularities may occur on the boundary of the domain of hyperbolicity or realizability. So characterize the singular states on the boundary via scalars that become infinite there and use justified caps on these scalars to project into the interior of the domain of positive states.

Conclusions

- Outflow positivity limiting allows cheap, positivity-preserving, high-order-accurate DG/WENO.
- Outflow positivity limiting consists of three essential parts:
  1. Cap *outflow*
     (maintains positivity),
  2. Cap *boundary averages*
     (guarantees a minimum positivity-preserving time step), and
  3. Cap *wave speeds*
     (guarantees a minimum stable time step).
- Given a speed cap, boundary average capping can guarantee the same minimum positivity-preserving time step as does enforcing positivity everywhere in each mesh cell.
- Positivity limiting makes higher-moment gas models robust.

References

[JoRo12] E.A. Johnson and J.A. Rossmanith, *Outflow positivity limiting for hyperbolic conservation laws. Part I: framework and recipe*, in preparation.

[Jo12] E.A. Johnson, *Outflow positivity limiting for hyperbolic conservation laws. Part II: analysis and extension*, in preparation.

[WZSN11] Cheng Wang, Xiangxiong Zhang, Chi-Wang Shu, Jianguo Ning, *Robust high order discontinuous Galerkin schemes for two-dimensional gaseous detonations*, J. Comp. Phys. 231 (2012) 653–665.

[LiuOsher96] X.-D. Liu and S. Osher, *Nonoscillatory high order accurate self-similar maximum principle satisfying shock capturing schemes*, SIAM J. Numer. Anal. 33 (1996), no. 2, 760–779.

[ZhangShu10] X. Zhang and C.-W. Shu, *On positivity preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes*, J. Comp. Phys. 229 (2010), 8918–8934.

[ZXS12] X. Zhang, Y. Xia, and C.-W. Shu, *Maximum-principle-satisfying and positivity-preserving high order discontinuous Galerkin schemes for conservation laws on triangular meshes*, J. Sci. Comput. 50 (2012), no. 1, 29-62.