New perpective on time stepping techniques: Beyond strong stability.

J.-L. Guermond

Department of Mathematics
Texas A&M University

Oden Institute Seminar, Feb 2, 2022

Collaborators and acknowledgments

This work done in collaboration with:

- Alexandre Ern (École Nationale des Ponts & Chaussées, Paris, France)

Other collaborators

- Bennett Clayton (TAMU, TX)
- Martin Kronbichler (Uppsala, Sweden)
- Matthias Maier (TAMU, TX)
- B. Popov (TAMU, TX)
- Laura Saavedra (Universidad Politcnica de Madrid)
- Madison Sheridan (TAMU, TX)
- Ignacio Tomas (SANDIA, NM)
- Eric Tovar (TAMU, TX)

Support:

AFOSR AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
ARL
Lawrence Livermore National Laboratory
NSF
A&M

Outline

Introduction

Cauchy problem

- Cauchy problem

$$\begin{aligned} &\partial_t \mathbf{u} + \nabla\cdot(\mathbb{f}(\mathbf{u}) + \mathbb{g}(\mathbf{u}, \nabla \mathbf{u})) = \mathbf{S}(\mathbf{u}), && (\mathbf{x}, t) \in D \times \mathbb{R}_+. \\ &u(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}), && \mathbf{x} \in D. \end{aligned}$$

Cauchy problem

- Cauchy problem

$$
\begin{aligned}
&\partial_t \mathbf{u} + \nabla\cdot(\mathbb{f}(\mathbf{u}) + \mathbb{g}(\mathbf{u}, \nabla\mathbf{u})) = \mathbf{S}(\mathbf{u}), && (\mathbf{x}, t) \in D\times\mathbb{R}_+.\\
&u(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}), && \mathbf{x} \in D.
\end{aligned}
$$

- D open polyhedral domain in $\mathbb{R}^d$.

Cauchy problem

- Cauchy problem

$$
\begin{aligned}
&\partial_t \mathbf{u} + \nabla\cdot(\mathbb{f}(\mathbf{u}) + \mathbb{g}(\mathbf{u}, \nabla\mathbf{u})) = \mathbf{S}(\mathbf{u}), && (\mathbf{x}, t) \in D\times\mathbb{R}_+.\\
&u(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}), && \mathbf{x} \in D.
\end{aligned}
$$

- D open polyhedral domain in $\mathbb{R}^d$.
- Field $\mathbf{u}$ takes values in $\mathbb{R}^m$; that is, $\mathbf{u} : D\times\mathbb{R}_+ \to \mathbb{R}^m$

Cauchy problem

- Cauchy problem

$$\begin{aligned} &\partial_t \mathbf{u} + \nabla\cdot(\mathbb{f}(\mathbf{u}) + \mathbb{g}(\mathbf{u}, \nabla\mathbf{u})) = \mathbf{S}(\mathbf{u}), && (\mathbf{x}, t) \in D\times\mathbb{R}_+. \\ &u(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}), && \mathbf{x} \in D. \end{aligned}$$

- D open polyhedral domain in $\mathbb{R}^d$.
- Field $\mathbf{u}$ takes values in $\mathbb{R}^m$; that is, $\mathbf{u} : D\times\mathbb{R}_+ \to \mathbb{R}^m$
- $\mathbb{f} \in \mathcal{C}^1(\mathbb{R}^m; \mathbb{R}^{m\times d})$, the hyperbolic flux.

Cauchy problem

- Cauchy problem

$$\begin{aligned} &\partial_t \mathbf{u} + \nabla\cdot(\mathbb{f}(\mathbf{u}) + \mathbb{g}(\mathbf{u}, \nabla \mathbf{u})) = \mathbf{S}(\mathbf{u}), && (\mathbf{x}, t) \in D \times \mathbb{R}_+. \\ &u(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}), && \mathbf{x} \in D. \end{aligned}$$

- D open polyhedral domain in $\mathbb{R}^d$.
- Field $\mathbf{u}$ takes values in $\mathbb{R}^m$; that is, $\mathbf{u} : D \times \mathbb{R}_+ \to \mathbb{R}^m$
- $\mathbb{f} \in \mathcal{C}^1(\mathbb{R}^m; \mathbb{R}^{m\times d})$, the hyperbolic flux.
- $\mathbb{g} \in \mathcal{C}^1(\mathbb{R}^m \times \mathbb{R}^{m\times d}; \mathbb{R}^{m\times d})$, the diffusive/parabolic flux.

Cauchy problem

- Cauchy problem

$$\begin{aligned} &\partial_t \mathbf{u} + \nabla\cdot(\mathbb{f}(\mathbf{u}) + \mathbb{g}(\mathbf{u}, \nabla\mathbf{u})) = \mathbf{S}(\mathbf{u}), && (\mathbf{x}, t) \in D\times\mathbb{R}_+. \\ &u(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}), && \mathbf{x} \in D. \end{aligned}$$

- D open polyhedral domain in $\mathbb{R}^d$.
- Field $\mathbf{u}$ takes values in $\mathbb{R}^m$; that is, $\mathbf{u} : D\times\mathbb{R}_+ \to \mathbb{R}^m$
- $\mathbb{f} \in \mathcal{C}^1(\mathbb{R}^m; \mathbb{R}^{m\times d})$, the hyperbolic flux.
- $\mathbb{g} \in \mathcal{C}^1(\mathbb{R}^m\times\mathbb{R}^{m\times d}; \mathbb{R}^{m\times d})$, the diffusive/parabolic flux.
- $\mathbf{S} \in \mathcal{C}^1(\mathbb{R}^m; \mathbb{R}^{m\times d})$, source.

Cauchy problem

- Cauchy problem

$$\begin{aligned} &\partial_t \mathbf{u} + \nabla\cdot(\mathbb{f}(\mathbf{u}) + \mathbb{g}(\mathbf{u}, \nabla \mathbf{u})) = \mathbf{S}(\mathbf{u}), && (\mathbf{x}, t) \in D\times\mathbb{R}_+. \\ &u(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}), && \mathbf{x} \in D. \end{aligned}$$

- D open polyhedral domain in $\mathbb{R}^d$.
- Field $\mathbf{u}$ takes values in $\mathbb{R}^m$; that is, $\mathbf{u} : D\times\mathbb{R}_+ \to \mathbb{R}^m$
- $\mathbb{f} \in \mathcal{C}^1(\mathbb{R}^m; \mathbb{R}^{m\times d})$, the hyperbolic flux.
- $\mathbb{g} \in \mathcal{C}^1(\mathbb{R}^m\times\mathbb{R}^{m\times d}; \mathbb{R}^{m\times d})$, the diffusive/parabolic flux.
- $\mathbf{S} \in \mathcal{C}^1(\mathbb{R}^m; \mathbb{R}^{m\times d})$, source.
- $\mathbf{u}_0$, admissible initial data.

Cauchy problem

- Cauchy problem

$$\partial_t \mathbf{u} + \nabla\cdot(\mathbb{f}(\mathbf{u}) + \mathbb{g}(\mathbf{u}, \nabla\mathbf{u})) = \mathbf{S}(\mathbf{u}), \qquad (\mathbf{x}, t) \in D\times\mathbb{R}_+.$$
$$u(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}), \qquad \mathbf{x} \in D.$$

- D open polyhedral domain in $\mathbb{R}^d$.
- Field $\mathbf{u}$ takes values in $\mathbb{R}^m$; that is, $\mathbf{u} : D\times\mathbb{R}_+ \to \mathbb{R}^m$
- $\mathbb{f} \in \mathcal{C}^1(\mathbb{R}^m; \mathbb{R}^{m\times d})$, the hyperbolic flux.
- $\mathbb{g} \in \mathcal{C}^1(\mathbb{R}^m\times\mathbb{R}^{m\times d}; \mathbb{R}^{m\times d})$, the diffusive/parabolic flux.
- $\mathbf{S} \in \mathcal{C}^1(\mathbb{R}^m; \mathbb{R}^{m\times d})$, source.
- $\mathbf{u}_0$, admissible initial data.
- Periodic BCs or $\mathbf{u}_0$ has compact support (to simplify BCs)

Hyperbolicity (recall)

Definition

The system $\partial_t \mathbf{u} + \nabla\cdot(\mathbb{f}(\mathbf{u}) = \mathbf{0}$ is said to be hyperbolic if for all unit vector $\mathbf{n} \in \mathbb{R}^d$ and all $\mathbf{v}$ in the domain of $\mathbb{f}$

$$\boxed{\mathbb{f}'(\mathbf{v})\mathbf{n} \text{ is diagonalizable with real eigenvalues.}}$$

Example 1: Navier-Stokes

- Find $\mathbf{u} := (\rho, \mathbf{m}, E)^{\mathsf{T}}$ so that

$$\begin{aligned}
&\partial_t \rho + \nabla\cdot(\mathbf{v}\rho) = 0, \\
&\partial_t \mathbf{m} + \nabla\cdot(\mathbf{v}\otimes\mathbf{m} + p(\mathbf{u})\mathbb{I} - \mathbb{s}(\mathbf{v})) = \mathbf{0}, \\
&\partial_t E + \nabla\cdot(\mathbf{v}(E + p(\mathbf{u})) - \mathbb{s}(\mathbf{v})\cdot\mathbf{v} + \mathbf{q}(\mathbf{u})) = 0,
\end{aligned}$$

with $\mathbf{v} := \mathbf{m}/\rho$: velocity; $p(\mathbf{u})$: pressure.

Example 1: Navier-Stokes

- Find $\mathbf{u} := (\rho, \mathbf{m}, E)^{\mathsf{T}}$ so that

$$
\begin{aligned}
&\partial_t \rho + \nabla\cdot(\mathbf{v}\rho) = 0,\\
&\partial_t \mathbf{m} + \nabla\cdot(\mathbf{v}\otimes\mathbf{m} + p(\mathbf{u})\mathbb{I} - \mathbb{s}(\mathbf{v})) = \mathbf{0},\\
&\partial_t E + \nabla\cdot(\mathbf{v}(E + p(\mathbf{u})) - \mathbb{s}(\mathbf{v})\cdot\mathbf{v} + \mathbf{q}(\mathbf{u})) = 0,
\end{aligned}
$$

with $\mathbf{v} := \mathbf{m}/\rho$: velocity; $p(\mathbf{u})$: pressure.

- Fluxes

$$
\mathbb{f}(\mathbf{u}) := \begin{pmatrix} \mathbf{v}\rho \\ \mathbf{v}\otimes\mathbf{m} + p(\mathbf{u})\mathbb{I} \\ \mathbf{v}(E + p(\mathbf{u})) \end{pmatrix}, \qquad \mathbb{g}(\mathbf{u}, \nabla\mathbf{u}) := \begin{pmatrix} 0 \\ -\mathbb{s}(\mathbf{v}) \\ -\mathbb{s}(\mathbf{v})\cdot\mathbf{v} + \mathbf{q}(\mathbf{u}) \end{pmatrix}.
$$

Example 1: Navier-Stokes

- Find $\mathbf{u} := (\rho, \mathbf{m}, E)^{\mathsf{T}}$ so that

$$\begin{aligned} &\partial_t \rho + \nabla\cdot(\mathbf{v}\rho) = 0, \\ &\partial_t \mathbf{m} + \nabla\cdot(\mathbf{v}\otimes\mathbf{m} + p(\mathbf{u})\mathbb{I} - \mathbb{s}(\mathbf{v})) = \mathbf{0}, \\ &\partial_t E + \nabla\cdot(\mathbf{v}(E + p(\mathbf{u})) - \mathbb{s}(\mathbf{v})\cdot\mathbf{v} + \mathbf{q}(\mathbf{u})) = 0, \end{aligned}$$

with $\mathbf{v} := \mathbf{m}/\rho$: velocity; $p(\mathbf{u})$: pressure.

- Fluxes

$$\mathbb{f}(\mathbf{u}) := \begin{pmatrix} \mathbf{v}\rho \\ \mathbf{v}\otimes\mathbf{m} + p(\mathbf{u})\mathbb{I} \\ \mathbf{v}(E + p(\mathbf{u})) \end{pmatrix}, \qquad \mathbb{g}(\mathbf{u}, \nabla\mathbf{u}) := \begin{pmatrix} 0 \\ -\mathbb{s}(\mathbf{v}) \\ -\mathbb{s}(\mathbf{v})\cdot\mathbf{v} + \mathbf{q}(\mathbf{u}) \end{pmatrix}.$$

- Possible definitions for $\mathbb{s}$ and q:

$$\mathbb{s}(\mathbf{v}) = 2\mu\mathbb{e}(\mathbf{v}) + (\lambda - \tfrac{2}{3}\mu)(\nabla\cdot\mathbf{v})\mathbb{I}, \qquad \mathbf{q}(\mathbf{u}) = -\kappa\nabla e(\mathbf{u}).$$

Example 2: Gray radiation hydrodynamics

- Find $\mathbf{u} := (\rho, \mathbf{m}, E, \mathcal{E}_{\mathrm{R}})^{\mathsf{T}}$ so that

$$
\begin{aligned}
&\partial_t \rho + \nabla\cdot(\mathbf{v}\rho) = 0, \\
&\partial_t \mathbf{m} + \nabla\cdot(\mathbf{v}\otimes\mathbf{m} + (p(\mathbf{u}) + p_{\mathrm{R}}(\mathcal{E}_{\mathrm{R}}))\mathbb{I}) = \mathbf{0}, \\
&\partial_t E + \nabla\cdot(\mathbf{v}(E + p(\mathbf{u}) + p_{\mathrm{R}}(\mathcal{E}_{\mathrm{R}}))) - \nabla\cdot(\tfrac{c}{3\sigma_{\mathrm{t}}}\nabla\mathcal{E}_{\mathrm{R}}) = 0, \\
&\partial_t \mathcal{E}_{\mathrm{R}} + \nabla\cdot(\mathbf{v}(\mathcal{E}_{\mathrm{R}} + p_{\mathrm{R}}(\mathcal{E}_{\mathrm{R}}))) - \mathbf{v}\cdot\nabla p_{\mathrm{R}}(\mathcal{E}_{\mathrm{R}}) - \nabla\cdot(\tfrac{c}{3\sigma_{\mathrm{t}}}\nabla\mathcal{E}_{\mathrm{R}}) = \sigma_{\mathrm{a}} c(a_{\mathrm{R}} T^4 - \mathcal{E}_{\mathrm{R}}),
\end{aligned}
$$

with $\mathcal{E}_{\mathrm{R}}$: radiation energy; $p_{\mathrm{R}}(\mathcal{E}_{\mathrm{R}})$: radiation pressure; $T(\mathbf{u})$: temperature;

Example 2: Gray radiation hydrodynamics

- Find $\mathbf{u} := (\rho, \mathbf{m}, E, \mathcal{E}_{\mathrm{R}})^{\mathsf{T}}$ so that

$$
\begin{aligned}
&\partial_t \rho + \nabla\cdot(\mathbf{v}\rho) = 0,\\
&\partial_t \mathbf{m} + \nabla\cdot(\mathbf{v}\otimes\mathbf{m} + (p(\mathbf{u}) + p_{\mathrm{R}}(\mathcal{E}_{\mathrm{R}}))\mathbb{I}) = \mathbf{0},\\
&\partial_t E + \nabla\cdot(\mathbf{v}(E + p(\mathbf{u}) + p_{\mathrm{R}}(\mathcal{E}_{\mathrm{R}}))) - \nabla\cdot(\tfrac{c}{3\sigma_{\mathrm{t}}}\nabla\mathcal{E}_{\mathrm{R}}) = 0,\\
&\partial_t \mathcal{E}_{\mathrm{R}} + \nabla\cdot(\mathbf{v}(\mathcal{E}_{\mathrm{R}} + p_{\mathrm{R}}(\mathcal{E}_{\mathrm{R}}))) - \mathbf{v}\cdot\nabla p_{\mathrm{R}}(\mathcal{E}_{\mathrm{R}}) - \nabla\cdot(\tfrac{c}{3\sigma_{\mathrm{t}}}\nabla\mathcal{E}_{\mathrm{R}}) = \sigma_{\mathrm{a}} c(a_{\mathrm{R}} T^4 - \mathcal{E}_{\mathrm{R}}),
\end{aligned}
$$

with $\mathcal{E}_{\mathrm{R}}$: radiation energy; $p_{\mathrm{R}}(\mathcal{E}_{\mathrm{R}})$: radiation pressure; $T(\mathbf{u})$: temperature;

c: speed of light; σ_{a}, σ_{t}: absorption and total cross sections; $a_{\mathrm{R}} := \frac{4\sigma}{c}$ radiation constant; σ the Stefan–Boltzmann constant.

Example 2: Gray radiation hydrodynamics

- Find $\mathbf{u} := (\rho, \mathbf{m}, E, \mathcal{E}_\mathrm{R})^\mathsf{T}$ so that

$$
\begin{aligned}
&\partial_t \rho + \nabla\cdot(\mathbf{v}\rho) = 0, \\
&\partial_t \mathbf{m} + \nabla\cdot(\mathbf{v}\otimes\mathbf{m} + (p(\mathbf{u}) + p_\mathrm{R}(\mathcal{E}_\mathrm{R}))\mathbb{I}) = \mathbf{0}, \\
&\partial_t E + \nabla\cdot(\mathbf{v}(E + p(\mathbf{u}) + p_\mathrm{R}(\mathcal{E}_\mathrm{R}))) - \nabla\cdot(\tfrac{c}{3\sigma_\mathrm{t}}\nabla\mathcal{E}_\mathrm{R}) = 0, \\
&\partial_t \mathcal{E}_\mathrm{R} + \nabla\cdot(\mathbf{v}(\mathcal{E}_\mathrm{R} + p_\mathrm{R}(\mathcal{E}_\mathrm{R}))) - \mathbf{v}\cdot\nabla p_\mathrm{R}(\mathcal{E}_\mathrm{R}) - \nabla\cdot(\tfrac{c}{3\sigma_\mathrm{t}}\nabla\mathcal{E}_\mathrm{R}) = \sigma_\mathrm{a} c(a_\mathrm{R} T^4 - \mathcal{E}_\mathrm{R}),
\end{aligned}
$$

with $\mathcal{E}_\mathrm{R}$: radiation energy; $p_\mathrm{R}(\mathcal{E}_\mathrm{R})$: radiation pressure; $T(\mathbf{u})$: temperature;

c: speed of light; σ_a, σ_t: absorption and total cross sections; $a_\mathrm{R} := \frac{4\sigma}{c}$ radiation constant; σ the Stefan–Boltzmann constant.

- Possible definitions:

$$
p_\mathrm{R}(\mathcal{E}_\mathrm{R}) := \frac{1}{3}\mathcal{E}_\mathrm{R}; \qquad c_\mathrm{v} T = e(\mathbf{u}) := \tfrac{1}{\rho}(E - \tfrac{1}{2}\rho\mathbf{v}^2).
$$

Example 2: Gray radiation hydrodynamics

- Notice non conservative term: $\mathbf{v}\cdot\nabla p_{\mathrm{R}}(\mathcal{E}_{\mathrm{R}})$
- Variable $\mathcal{E}_{\mathrm{R}}$ is smooth (due to presence of $\nabla\cdot(\frac{c}{3\sigma_{\mathrm{t}}}\nabla\mathcal{E}_{\mathrm{R}})$).
- $\Longrightarrow$ field $\mathcal{E}_{\mathrm{R}}$ is smooth
  $\Longrightarrow$ legitimate to perform change of variable $\mathcal{E}_{\mathrm{R}} \to \widetilde{\mathcal{E}}_{\mathrm{R}} := \mathcal{E}_{\mathrm{R}}^{\frac{3}{4}}$.
- After some algebra $\rightsquigarrow$ Definition of conservative hyperbolic flux:

$$\boxed{\mathbb{f}(\widetilde{\mathbf{u}}) := \begin{pmatrix} \mathbf{v}\rho \\ \mathbf{v}\otimes\mathbf{m} + q(\widetilde{\mathbf{u}})\mathbb{I} \\ \mathbf{v}(E + q(\widetilde{\mathbf{u}})) \\ \mathbf{v}\widetilde{\mathcal{E}}_{\mathrm{R}} \end{pmatrix}.}$$

with $q(\widetilde{\mathbf{u}}) := p(\widetilde{\mathbf{u}}) + \frac{1}{3}\widetilde{\mathcal{E}}_{\mathrm{R}}^{\frac{4}{3}}$ (notice that $p(\widetilde{\mathbf{u}}) = p(\mathbf{u})$).

- Hyperbolic flux is made conservative.

Example 2: Gray radiation hydrodynamics

- Parabolic flux

$$\boxed{\mathrm{g}(\mathbf{u}, \nabla\mathbf{u}) := \begin{pmatrix} 0 \\ \mathbf{0} \\ -\frac{c}{3\sigma_{\mathrm{t}}}\nabla\mathcal{E}_{\mathrm{R}} \\ -\frac{c}{3\sigma_{\mathrm{t}}}\nabla\mathcal{E}_{\mathrm{R}} \end{pmatrix}.}$$

- Source

$$\boxed{\mathbf{S}(\mathbf{u}) := \begin{pmatrix} 0 \\ \mathbf{0} \\ 0 \\ \sigma_{\mathrm{a}} c(\mathcal{E}_{\mathrm{R}} - a_{\mathrm{R}} T^4) \end{pmatrix}.}$$

Existence uniqueness: Scalar equations

- Existence and uniqueness of entropy solutions well understood in any space dimension for any Lipschitz flux.
  - **Oleinik (1959)**
  - **Vol'pert (1967)**
  - **Kruzkov (1970)**

Approximation: Scalar equations

- Approximation theory for scalar conservation theory well understood in any space dimension.
- Convergence rate deduced from a posteriori estimates:
  - **Kuznecov (1976)**,
  - **Cockburn–Gremaud (1996)**,
  - **Bouchut–Perthame (1998)**,
  - **Eymard–Gallouet–Herbin (1998)**,
  - **Chainais-Hillaret (1999)**,
  - **JLG-Popov (2016)** ...

Existence uniqueness: Hyperbolic systems

- Wellposedness known only for data with small total variation in 1D (**Glimm (1965)**, **Bianchini–Bressan (2005)**).

Existence uniqueness: Hyperbolic systems

- Wellposedness known only for data with small total variation in 1D (**Glimm (1965)**, **Bianchini–Bressan (2005)**).
- Partial positive results for special systems in 1D.

Existence uniqueness: Hyperbolic systems

- Wellposedness known only for data with small total variation in 1D (**Glimm (1965)**, **Bianchini–Bressan (2005)**).
- Partial positive results for special systems in 1D.
- Negative results: Entropy conditions may not be sufficient for systems **Chiodaroli–De Lellis (2015)**.

Approximation: Hyperbolic systems

- Approximation theory for systems almost non existent in 1D:
  See Bressan's seminar, Jan 8, 2021, (Laboratoire Jacques-Louis Lions)
  https://www.ljll.math.upmc.fr/IMG/pdf/ljll210108bressan.a-2.5mo.pdf

Approximation: Hyperbolic systems

- Approximation theory for systems almost non existent in 1D: See Bressan's seminar, Jan 8, 2021, (Laboratoire Jacques-Louis Lions) `https://www.ljll.math.upmc.fr/IMG/pdf/ljll210108bressan.a-2.5mo.pdf`
- For the Lax-Friedrichs and the Godunov scheme, convergence is known only for $2{\times}2$ systems in 1D (compensated compactness argument, no convergence rate), **Ding-Chen-Luo (1989)**

Approximation: Hyperbolic systems

- Approximation theory for systems almost non existent in 1D:
  See Bressan's seminar, Jan 8, 2021, (Laboratoire Jacques-Louis Lions)
  `https://www.ljll.math.upmc.fr/IMG/pdf/ljll210108bressan.a-2.5mo.pdf`
- For the Lax-Friedrichs and the Godunov scheme, convergence is known only for 2×2 systems in 1D (compensated compactness argument, no convergence rate), **Ding-Chen-Luo (1989)**
- There are 2×2 systems in 1D for which the Godunov method yields an unbounded BV norm as the mesh size goes to zero, **Bressan-Jenssen-Baiti (2006)**.

Approximation: Hyperbolic systems

- Approximation theory for systems almost non existent in 1D:
  See Bressan's seminar, Jan 8, 2021, (Laboratoire Jacques-Louis Lions)
  `https://www.ljll.math.upmc.fr/IMG/pdf/ljll210108bressan.a-2.5mo.pdf`
- For the Lax-Friedrichs and the Godunov scheme, convergence is known only for 2×2 systems in 1D (compensated compactness argument, no convergence rate), **Ding-Chen-Luo (1989)**
- There are 2×2 systems in 1D for which the Godunov method yields an unbounded BV norm as the mesh size goes to zero, **Bressan-Jenssen-Baiti (2006)**.
- Very little hope to prove convergence of approximation techniques in two and three dimensions with realistic data. (With current mathematical knowledge.)

Questions

What can we do?

- What can Numerical Analysis do?

Questions

What can we do?

- What can Numerical Analysis do?
- Numerous computational fluid dynamics codes developed!

Questions

What can we do?

- What can Numerical Analysis do?
- Numerous computational fluid dynamics codes developed!
- Planes fly. Nuclear reactors run.

Questions

What can we do?

- What can Numerical Analysis do?
- Numerous computational fluid dynamics codes developed!
- Planes fly. Nuclear reactors run.
- Engineers do not wait for mathematicians to find answers.

Questions

Proposed strategy

- Reduce expectations.
- Try to ensure the approximation satisfies "physical bounds"
- Try to ensure the approximation complies with thermodynamics.
- Try to achieve linear complexity with respect the number of degrees of freedom.

Outline

Invariant domains

Key assumption: existence of an invariant domain

- Let $\mathbf{u}_0 \in \mathcal{D}$.
- There exists a set $\mathcal{A} \subsetneq \mathbb{R}^m$, convex and depending on $\mathbf{u}_0$, so that the "entropy" solution takes values in $\mathcal{A}$ for a.e. $\mathbf{x} \in D$ and $t > 0$.

$$\boxed{(\mathbf{u}_0(\mathbf{x}) \in \mathcal{A}, \forall \mathbf{x} \in D) \Longrightarrow (\mathbf{u}(\mathbf{x}, t) \in \mathcal{A}, \forall \mathbf{x} \in D, \forall t > 0)\,.}$$

- This is a generalization of the maximum principle.

Examples

- Scalar conservation equations

$$\mathcal{A} := [\operatorname*{ess\,inf}_{x\in\mathbb{R}} u_0(x), \operatorname*{ess\,sup}_{x\in\mathbb{R}} u_0(x)] \quad \text{is a convex subset of } \mathbb{R}$$

Examples

- Euler equations with specific entropy s

$$\mathcal{A} := \{\mathbf{u} := (\rho, \mathbf{m}, E) \in \mathbb{R}^{d+2} \mid \rho > 0, E - \frac{1}{2}\frac{\mathbf{m}^2}{\rho} > 0, s(\mathbf{u}) \geq \operatorname*{ess\,inf}_{\mathbf{x}\in D} s(\mathbf{u}_0)\}$$

- Navier-Stokes equations

$$\mathcal{A} := \{(\rho, \mathbf{m}, E) \in \mathbb{R}^{d+2} \mid \rho > 0, E - \frac{1}{2}\frac{\mathbf{m}^2}{\rho} > 0\}$$

- $\mathcal{A}$ is convex in both cases.
- Invariant domain for the Euler equations is smaller than that for the Navier-Stokes equations.

Questions

- Hyperbolic and parabolic operators may have conflicting constraints.

Questions

- Hyperbolic and parabolic operators may have conflicting constraints.
- Example 1: Navier-Stokes
  - Euler: Conserved variables are natural for solving the hyperbolic problem
  - Navier-Stokes: primitive variables (velocity, internal energy) are more appropriate for the parabolic part.
  - The invariant domain of the Euler part is smaller than the invariant domain of the parabolic part.

Questions

- Hyperbolic and parabolic operators may have conflicting constraints.
- Example 1: Navier-Stokes
  - Euler: Conserved variables are natural for solving the hyperbolic problem
  - Navier-Stokes: primitive variables (velocity, internal energy) are more appropriate for the parabolic part.
  - The invariant domain of the Euler part is smaller than the invariant domain of the parabolic part.
- Example 2: Gray radiation hydrodynamics
  - Euler: Conserved variables $(\rho, \mathbf{m}, E, \mathcal{E}_{\mathrm{R}}^{\frac{3}{4}})^{\mathsf{T}}$.
  - Parabolic part: $(T, \mathcal{E}_{\mathrm{R}})^{\mathsf{T}}$.
  - The invariant domain of the Euler part is smaller than the invariant domain of the parabolic part.

Questions

- How can one reconcile all these constraints?
- How can one construct approximation techniques in time and space that preserve invariant domains?

Outline

SSP

SSP (strong stability preserving)

- Approximate $\mathbf{u}(\mathbf{x}, t)$ in space with dofs in $\mathbb{R}^{m\times I}$.
- I : dimension of the approximation vector space (Finite elements (C^0 or dG), Finite Volume, Finite Differences, etc.).
- Let $\mathbf{F} : \mathbb{R}^{m\times I} \to \mathbb{R}^{m\times I}$ be approximation in space of $-\nabla\cdot\mathbb{f}(\mathbf{u})$. (The way this is done does not matter here.)
- Semi-discrete problem: Find $\mathbf{U} \in C^1([0, T]; \mathbb{R}^{m\times I})$ s.t.

$$\mathbb{M}\partial_t\mathbf{U} = \mathbf{F}(\mathbf{U}), \quad \mathbf{U}(0) = \mathbf{U}_0.$$

$\mathbb{M}$: mass matrix (invertible)

SSP (strong stability preserving)

- Assume $\mathbf{U}_0 \in \mathcal{A}^I$.

SSP (strong stability preserving)

- Assume $\mathbf{U}_0 \in \mathcal{A}^I$.
- How can one construct time-stepping technique that guarantee $\mathbf{U}^n \in \mathcal{A}^I$, for all $n \geq 0$?

SSP (strong stability preserving)

- Key assumption: (Forward Euler with low-order flux is invariant-domain preserving.) $\exists \Delta t^* > 0$ s.t. $\forall \Delta t \in (0, \Delta t^*)$ and $\forall \mathbf{V} \in \mathbb{R}^{m \times I}$

$$\boxed{(\mathbf{V} \in \mathcal{A}^I) \Longrightarrow (\mathbf{V} + \Delta t (\mathbb{M})^{-1} \mathbf{F}(\mathbf{V}) \in \mathcal{A}^I).}$$

$\Leftrightarrow \mathcal{A}^I$ is invariant by the forward Euler method under the CFL condition $\Delta t \in (0, \Delta t^*)$.

SSP (strong stability preserving)

- Key assumption: (Forward Euler with low-order flux is invariant-domain preserving.) $\exists \Delta t^* > 0$ s.t. $\forall \Delta t \in (0, \Delta t^*)$ and $\forall \mathbf{V} \in \mathbb{R}^{m \times I}$

$$\boxed{(\mathbf{V} \in \mathcal{A}^I) \Longrightarrow (\mathbf{V} + \Delta t(\mathbb{M})^{-1}\mathbf{F}(\mathbf{V}) \in \mathcal{A}^I).}$$

  $\Leftrightarrow \mathcal{A}^I$ is invariant by the forward Euler method under the CFL condition $\Delta t \in (0, \Delta t^*)$.

- Key idea by **Shu&Osher (1988)**

  > Use explicit Runge-Kutta methods where the final update is a convex combination of updates computed with the forward Euler method.

SSP (strong stability preserving)

- Generic form of s-stage, explicit Runge-Kutta, strong-stability-preserving methods (SSPRK)

$$\mathbf{W}^{(i)} := \sum_{k \in \{0:i-1\}} \alpha_{ik} \mathbf{W}^{(k)} + \beta_{ik} \Delta t \mathbf{F}(\mathbf{W}^{(k)}), \quad \forall i \in \{1:s\}.$$

- The update at t_{n+1} is given by $\mathbf{U}^{n+1} := \mathbf{W}^{(s)}$.
- Theory well understood now:
  - **Kraaijevanger (1991)** (amazing paper),
  - **Spiteri-Ruuth (2002)**,
  - **Ferracina-Spijker (2005)**,
  - **Higueras (2005)**.

Examples (for $\partial_t u = L(t,u)$)

- SSPRK(2,2)

α		β		γ
1		1		0
$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{2}$	1

$$\begin{aligned} w^{(1)} &:= u^n + \Delta t L(t_n, u^n), \\ w^{(2)} &:= \tfrac{1}{2} u^n + \tfrac{1}{2}\left(w^{(1)} + \Delta t L(t_n + \Delta t, w^{(1)})\right), \end{aligned}$$

- SSPRK(3,3)

α			β			γ
1			1			0
$\frac{3}{4}$	$\frac{1}{4}$		0	$\frac{1}{4}$		1
$\frac{1}{3}$	0	$\frac{2}{3}$	0	0	$\frac{2}{3}$	$\frac{1}{2}$

$$\begin{aligned} w^{(1)} &:= u^n + \Delta t L(t_n, u^n), \\ w^{(2)} &:= \tfrac{3}{4} u^n + \tfrac{1}{4}\left(w^{(1)} + \Delta t L\left(t_n + \Delta t, w^{(1)}\right)\right), \\ w^{(3)} &:= \tfrac{1}{3} u^n + \tfrac{2}{3}\left(w^{(2)} + \Delta t L\left(t_n + \tfrac{1}{2}\Delta t, w^{(2)}\right)\right), \end{aligned}$$

Examples (for $\partial_t u = L(t, u)$)

- SSPRK(4,3)

α				β				γ
1				$\frac{1}{2}$				0
0	1			0	$\frac{1}{2}$			$\frac{1}{2}$
$\frac{2}{3}$	0	$\frac{1}{3}$		0	0	$\frac{1}{6}$		1
0	0	0	1	0	0	0	$\frac{1}{2}$	$\frac{1}{2}$

$$w^{(1)} := u^n + \tfrac{1}{2}\Delta t L(t_n, u^n),$$

$$w^{(2)} := w^{(1)} + \tfrac{1}{2}\Delta t L(t_n + \tfrac{1}{2}\Delta t, w^{(1)}),$$

$$w^{(3)} := \tfrac{2}{3}u^n + \tfrac{1}{3}\big(w^{(2)} + \tfrac{1}{2}\Delta t L(t_n + \Delta t, w^{(}$$

$$w^{(4)} := w^{(3)} + \tfrac{1}{2}\Delta t L(t_n + \tfrac{1}{2}\Delta t, w^{(3)}),$$

Problems with SPPRK

Definition (Efficiency ratio)

Let $c_{\text{os}} := \inf_{i\in\{1:s\}} \inf_{k\in\mathcal{K}_i} \alpha_{ik}\beta_{ik}^{-1}$.

Proposition

Under the same CFL constraint, the number of function evaluations of SSPRK(s,p) is equal to s/c_{os} × that of the forward Euler method.

Examples

- c_{os} for SSPRK(2,2) is 1 (instead of 2 $\Rightarrow \frac{1}{2}$ efficiency).
- c_{os} for SSPRK(3,3) is 1 (instead of 3 $\Rightarrow \frac{1}{3}$ efficiency).
- c_{os} for SSPRK(4,3) is 2 (instead of 4 $\Rightarrow \frac{1}{2}$ efficiency).

Problems with SPPRK: Efficiency

- *SSPRK* methods are inefficient!
- The most popular method *SSPRK*$(3,3)$ is actually the most inefficient!

Problems with SPPRK: Accuracy

- Accuracy of *SSPRK* methods restricted to fourth-order if one insists on never stepping backward in time.

Problems with SPPRK: extensions to IMEX methods

- The *SSPRK* paradigm cannot be easily modified to accommodate implicit and explicit sub-steps.
- Two exceptions:
  - Parabolic time step restriction $\Delta t \leq ch^2$
  - Scalar conservation equations that are variations of the heat-equation.

Problems with SPPRK: extensions to IMEX methods

Example (Compressible Navier-Stokes)

- Difficulties: conflicting invariant sets and conflicting variables.

- Which invariant domain to preserve?
  - Minimum entropy principle is true for Euler.
  - Minimum entropy principle is false for NS.
- Which variable should be used?
  - "Right variable" for Euler is $\mathbf{u} = (\rho, \mathbf{m}, E)$ (conserved variables).
  - "Right variable" for NS is $(\rho, \mathbf{v}, e)$ (primitive variables).
  - Some advocate "entropy variable" and "entropy stability". Why?
- How to do the explicit-implicit time stepping?
- How linearization should be done in the implicit substeps?
  - Most "IMEX" methods cannot make the difference between conserved and primitive variables.
  - Most "IMEX" methods cannot be properly linearized and be conservative (no generic theory).
  - Difficulty can be overcome by assuming $\Delta t \leq ch^2$, **Zhang & Shu (2017)**.

Problems with SPPRK: extensions to IMEX methods

- Conjecture: There does not exist any IMEX method that is SSP for general systems. (At the exclusion of scalar conservation equations and under the proper time step restriction).
- Conclusion: One needs a new paradigm.

Outline

IDPERK

Peep under the hood of SSPRK

- The beauty of SSPRK methods is that the forward Euler sub-step is a black box.
- The black box invokes two fluxes (not just one as one might think):
  - Low-order (in space) $\mathbf{F}^{\mathsf{L}}$, low-order mass matrix $\mathbb{M}^{\mathsf{L}}$
  - High-order (in space) $\mathbf{F}^{\mathsf{H}}$, low-order mass matrix $\mathbb{M}^{\mathsf{H}}$
- Ideally, one would like to solve

$$\mathbb{M}^{\mathsf{H}}\partial_t \mathbf{U} = \mathbf{F}^{\mathsf{H}}(\mathbf{U})$$

since the space approximation is accurate, but this method violates the invariant domain property.

Peep under the hood of SSPRK

Key assumptions

- Assumption 1: (Forward Euler with low-order flux is invariant-domain preserving.) Assume $\exists \Delta t^* > 0$ so that for all $\Delta t \in (0, \Delta t^*)$ for all $\mathbf{V} \in \mathbb{R}^{m\times I}$

$$(\mathbf{V} \in \mathcal{A}^I) \Longrightarrow (\mathbf{V} + \Delta t(\mathbb{M}^{\mathrm{L}})^{-1}\mathbf{F}^{\mathrm{L}}(\mathbf{V}) \in \mathcal{A}^I).$$

- Assumption 2: There exists a nonlinear limiting operator $\ell : \mathcal{A}^I \times (\mathbb{R}^m)^I \times (\mathbb{R}^m)^I \to (\mathbb{R}^m)^I$ such that for all $(\mathbf{V}, \mathbf{\Phi}^{\mathrm{L}}, \mathbf{\Phi}^{\mathrm{H}})$

$$(\mathbf{V} + \Delta t(\mathbb{M}^{\mathrm{L}})^{-1}\mathbf{\Phi}^{\mathrm{L}} \in \mathcal{A}^I) \Longrightarrow (\ell(\mathbf{V}, \mathbf{\Phi}^{\mathrm{L}}, \mathbf{\Phi}^{\mathrm{H}}) \in \mathcal{A}^I).$$

Lemma

For all $\mathbf{V} \in \mathcal{A}^I$ *and all* $\Delta t \in (0, \Delta t^*)$, *we have*

$$\ell(\mathbf{V}, \mathbf{F}^{\mathrm{L}}(\mathbf{V}), \mathbf{F}^{\mathrm{H}}(\mathbf{V})) \in \mathcal{A}^I$$

Peep under the hood of SSPRK

- Given $\mathbf{U}^n$ in the invariant set $\mathcal{A}^I$ (approximation at time t^n),
- The forward Euler step proceeds as follows:
  - Compute low-order flux $\mathbf{F}^{\mathrm{L}}(\mathbf{U}^n)$
  - Compute high-order flux $\mathbf{F}^{\mathrm{H}}(\mathbf{U}^n)$
  - Compute update $\mathbf{U}^{n+1}$ by limiting

$$\mathbf{U}^{n+1} := \ell(\mathbf{U}^n, \mathbf{F}^{\mathrm{L}}(\mathbf{U}^n), \mathbf{F}^{\mathrm{H}}(\mathbf{U}^n)).$$

Theorem (IDP Explicit Euler)

Assume $\mathbf{U}^n \in \mathcal{A}^I$. *Then* $\mathbf{U}^{n+1} \in \mathcal{A}^I$ *for all* $\Delta t \in (0, \Delta t^*)$.

Key idea of invariant-domain-preserving ERK

- Externalize the limiting process at each RK sub-step.

Details for s-stage ERK method

- Consider Butcher tableau for s-stage method

$$
\begin{array}{c|ccccc}
c_1 & 0 & & & & \\
c_2 & a_{2,1} & 0 & & & \\
c_3 & a_{3,1} & a_{3,2} & 0 & & \\
\vdots & \vdots & & \ddots & \ddots & \\
c_s & a_{s,1} & a_{s,2} & \cdots & a_{s,s-1} & 0 \\
\hline
 & b_1 & b_2 & \cdots & b_{s-1} & b_s
\end{array}
$$

- Rename last line, set $c_1 = 0$ and $c_{s+1} = 1$.

$$
\begin{array}{c|ccccc}
\color{blue}{0} & 0 & & & & \\
c_2 & a_{2,1} & 0 & & & \\
c_3 & a_{3,1} & a_{3,2} & 0 & & \\
\vdots & \vdots & & \ddots & \ddots & \\
c_s & a_{s,1} & a_{s,2} & \cdots & a_{s,s-1} & 0 \\
\color{blue}{1} & \color{blue}{a_{s+1,1}} & \color{blue}{a_{s+1,2}} & \cdots & \color{blue}{a_{s+1,s-1}} & \color{blue}{a_{s+1,s}}
\end{array}
$$

Details

- Assume $c_k \geq 0$ for all $k \in \{1{:}s+1\}$.
- For sake of simplicity assume $c_{l-1} \leq c_l$, $\forall l \in \{2{:}s+1\}$, and set

$$l' := l - 1.$$

(Otherwise set $l' := \max\{k < l \mid c_l - c_k \geq 0\}$.)

Details

- Let $\mathbf{U}^n \in \mathcal{A}^l$.
- Set $\mathbf{U}^{n,1} := \mathbf{U}^n$.
- Loop over $l \in \{2{:}s+1\}$.
- Compute first-order update starting from $\mathbf{U}^{n,l'}$ (think of $l' = l - 1$)

$$\mathbb{M}^{\mathsf{L}}\mathbf{U}^{\mathsf{L},l} := \mathbb{M}^{\mathsf{L}}\mathbf{U}^{n,l'} + \Delta t(c_l - c_{l'})\mathbf{F}^{\mathsf{L}}(\mathbf{U}^{n,l'}).$$

- Compute high-order ERK update starting from $\mathbf{U}^n$

$$\mathbb{M}^{\mathsf{H}}\mathbf{U}^{\mathsf{H},l} := \mathbb{M}^{\mathsf{H}}\mathbf{U}^{n} + \Delta t \sum_{k\in\{1:l-1\}} a_{l,k}\mathbf{F}^{\mathsf{H}}(\mathbf{U}^{n,k}).$$

Details

- Let $\mathbf{U}^n \in \mathcal{A}^l$.
- Set $\mathbf{U}^{n,1} := \mathbf{U}^n$.
- Loop over $l \in \{2{:}s+1\}$.
- Compute first-order update starting from $\mathbf{U}^{n,l'}$ (think of $l' = l - 1$)

$$\mathbb{M}^{\mathsf{L}}\mathbf{U}^{\mathsf{L},l} := \mathbb{M}^{\mathsf{L}}\mathbf{U}^{n,l'} + \Delta t(c_l - c_{l'})\mathbf{F}^{\mathsf{L}}(\mathbf{U}^{n,l'}).$$

- Compute high-order ERK update starting from $\mathbf{U}^n$

$$\mathbb{M}^{\mathsf{H}}\mathbf{U}^{\mathsf{H},l} := \mathbb{M}^{\mathsf{H}}\mathbf{U}^n + \Delta t \sum_{k\in\{1:l-1\}} a_{l,k}\mathbf{F}^{\mathsf{H}}(\mathbf{U}^{n,k}).$$

- Incompatibility of the starting points ($\mathbf{U}^{n,l'} \neq \mathbf{U}^n$ in general).

Details

- Let $\mathbf{U}^n \in \mathcal{A}^I$.
- Set $\mathbf{U}^{n,1} := \mathbf{U}^n$.
- Loop over $l \in \{2{:}\, s+1\}$.
- Compute first-order update starting from $\mathbf{U}^{n,l'}$ (think of $l' = l - 1$)

$$\mathbb{M}^{\mathrm{L}}\mathbf{U}^{\mathrm{L},l} := \mathbb{M}^{\mathrm{L}}\mathbf{U}^{n,l'} + \Delta t(c_l - c_{l'})\mathbf{F}^{\mathrm{L}}(\mathbf{U}^{n,l'}).$$

- Compute high-order ERK update starting from $\mathbf{U}^n$

$$\mathbb{M}^{\mathrm{H}}\mathbf{U}^{\mathrm{H},l} := \mathbb{M}^{\mathrm{H}}\mathbf{U}^{n} + \Delta t \sum_{k\in\{1:\, l-1\}} a_{l,k}\mathbf{F}^{\mathrm{H}}(\mathbf{U}^{n,k}).$$

- Incompatibility of the starting points ($\mathbf{U}^{n,l'} \neq \mathbf{U}^n$ in general).
- Subtract ERK update at $t^n + c_l\Delta t$ from ERK update at $t^n + c_{l'}\Delta t$

$$\Longrightarrow \qquad \mathbb{M}^{\mathrm{H}}\mathbf{U}^{\mathrm{H},l} = \mathbb{M}^{\mathrm{H}}\mathbf{U}^{\mathrm{H},l'} + \Delta t \sum_{k\in\{1:\, l-1\}} (a_{l,k} - a_{l',k})\mathbf{F}^{\mathrm{H}}(\mathbf{U}^{n,k}).$$

Details

- Replace $\mathbf{U}^{\mathsf{H},l'}$ (which is not IDP) by $\mathbf{U}^{n,l'}$ (which is IDP by induction assumption).
- Final scheme

$$\mathbb{M}^{\mathsf{L}}\mathbf{U}^{\mathsf{L},l} := \mathbb{M}^{\mathsf{L}}\mathbf{U}^{n,l'} + \Delta t \underbrace{(c_l - c_{l'})\mathbf{F}^{\mathsf{L}}(\mathbf{U}^{n,l'})}_{\mathbf{\Phi}^{\mathsf{L}}}.$$

$$\mathbb{M}^{\mathsf{H}}\mathbf{U}^{\mathsf{H},l} := \mathbb{M}^{\mathsf{H}}\mathbf{U}^{n,l'} + \Delta t \underbrace{\sum_{k\in\{1:l-1\}} (a_{l,k} - a_{l',k})\mathbf{F}^{\mathsf{H}}(\mathbf{U}^{n,k})}_{\mathbf{\Phi}^{\mathsf{H}}}.$$

$$\mathbf{U}^{n,l} := \ell(\mathbf{U}^{n,l'}, \mathbf{\Phi}^{\mathsf{L}}, \mathbf{\Phi}^{\mathsf{H}}).$$

- Set $\mathbf{U}^{n+1} := \mathbf{U}^{n,s+1}$.

Details

Theorem

Assume that $\mathbf{U}^n \in \mathcal{A}^I$*. Then* $\mathbf{U}^{n+1} \in \mathcal{A}^I$ *for all* $\Delta t \in (0, \frac{\Delta t^*}{\max_{l\in\{2:s+1\}}(c_l - c_{l'})})$.

Corollary

The complexity of the ERK method is optimal if the points $\{c_l\}_{l\in\{1:s+1\}}$ *are equi-distributed in* $[0,1]$.

Outline

Numerical illustrations

Examples (optimal methods)

$$\begin{array}{c|cc} 0 & 0 & \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \hline 1 & 0 & 1 \end{array}$$

RK(2,2;1)

$$\begin{array}{c|ccc} 0 & 0 & & \\ \frac{1}{3} & \frac{1}{3} & 0 & \\ \frac{2}{3} & 0 & \frac{2}{3} & 0 \\ \hline 1 & \frac{1}{4} & 0 & \frac{3}{4} \end{array}$$

RK(3,3;1)

$$\begin{array}{c|cccc} 0 & 0 & & & \\ \frac{1}{4} & \frac{1}{4} & 0 & & \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & \\ \frac{3}{4} & 0 & \frac{1}{4} & \frac{1}{2} & 0 \\ \hline 1 & 0 & \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \end{array}$$

RK(4,3;1)

Examples SSPRK (sub-optimal methods)

$$
\begin{array}{c|cc}
0 & 0 & \\
1 & 1 & 0 \\
\hline
 & \frac{1}{2} & \frac{1}{2}
\end{array}
$$

SSPRK(2,2;$\frac{1}{2}$)

$$
\begin{array}{c|ccc}
0 & 0 & & \\
1 & 1 & 0 & \\
\frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 \\
\hline
 & \frac{1}{6} & \frac{1}{6} & \frac{2}{3}
\end{array}
$$

SSPRK(3,3;$\frac{1}{3}$)

Examples: popular RK4 (left) and 3/8 rule (right)

$$
\begin{array}{c|cccc}
0 & 0 & & & \\
\frac{1}{2} & \frac{1}{2} & 0 & & \\
\frac{1}{2} & 0 & \frac{1}{2} & 0 & \\
1 & 0 & 0 & 1 & 0 \\
\hline
1 & \frac{1}{6} & \frac{2}{6} & \frac{2}{6} & \frac{1}{6}
\end{array}
$$

$\mathrm{RK}(4{,}4;\frac{1}{2})$

$$
\begin{array}{c|cccc}
0 & 0 & & & \\
\frac{1}{3} & \frac{1}{3} & 0 & & \\
\frac{2}{3} & -\frac{1}{3} & 1 & 0 & \\
1 & 1 & -1 & 1 & 0 \\
\hline
1 & \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8}
\end{array}
$$

$\mathrm{RK}(4{,}4;\frac{3}{4})$

Examples RK5 methods: Equi-distributed (left), Butcher's method (right)

$$
\begin{array}{c|cccccc}
0 & 0 \\
\frac{1}{5} & \frac{1}{5} & 0 \\
\frac{2}{5} & 0 & \frac{2}{5} & 0 \\
\frac{3}{5} & \frac{3}{20} & 0 & \frac{9}{20} & 0 \\
\frac{4}{5} & \frac{4}{5} & -\frac{8}{5} & \frac{8}{5} & 0 & 0 \\
1 & -\frac{71}{4} & 40 & -\frac{75}{4} & -10 & \frac{15}{2} & 0 \\
\hline
1 & \frac{17}{144} & 0 & \frac{25}{36} & -\frac{25}{72} & \frac{25}{48} & \frac{1}{72}
\end{array}
$$

RK$(6,5;\frac{5}{6})$

$$
\begin{array}{c|cccccc}
0 & 0 \\
\frac{1}{4} & \frac{1}{4} & 0 \\
\frac{1}{4} & \frac{1}{8} & \frac{1}{8} & 0 \\
\frac{1}{2} & 0 & -\frac{1}{2} & 1 & 0 \\
\frac{3}{4} & \frac{3}{16} & 0 & 0 & \frac{9}{16} & 0 \\
1 & -\frac{3}{7} & \frac{2}{7} & \frac{12}{7} & -\frac{12}{7} & \frac{8}{7} & 0 \\
\hline
1 & \frac{7}{90} & 0 & \frac{32}{90} & \frac{12}{90} & \frac{32}{90} & \frac{7}{90}
\end{array}
$$

RK$(6,5;\frac{2}{3})$

Convergence tests

- All the tests are done with

$$\Delta t := \mathrm{CFL} \times s \times \Delta t^*,$$

- $\Rightarrow$ All the methods perform exactly the same number of time steps independently of s (i.e., number of flux evaluations is constant).

1D linear transport, 4th-order FD

- 4th-order FD in space.
- Linear transport $D = (0, 1)$

$$\partial_t u + \partial_x u = 0, \qquad u_0(x) := \begin{cases} (4\frac{(x-x_0)(x_1-x)}{x_1-x_0})^6 & x \in (x_0 := 0.1, x_1 := 0.4) \\ 0 & \textit{otherwise} \end{cases}$$

- Local maximum/minimum principle guaranteed at every grid point.
- Global maximum and minimum also exactly enforced.
- All errors computed in L^∞-norm.

1D linear transport, 4th-order FD

Table: Second-order methods (SSPRK(2,2) behaves badly).

I	CFL = 0.2				CFL = 0.25			
	RK(2,2;1)	rate	RK(2,2; $\frac{1}{2}$)	rate	RK(2,2;1)	rate	RK(2,2; $\frac{1}{2}$)	rate
50	4.72E-02	–	1.23E-01	–	4.91E-02	–	1.30E-01	–
100	2.81E-03	4.07	1.50E-02	3.03	4.51E-03	3.44	4.32E-02	1.60
200	1.16E-03	1.28	1.24E-03	3.60	2.01E-03	1.17	2.14E-03	4.34
400	3.38E-04	1.78	3.47E-04	1.84	5.41E-04	1.89	5.67E-04	1.91
800	8.79E-05	1.94	9.28E-05	1.90	1.38E-04	1.97	1.48E-04	1.94
1600	2.22E-05	1.98	2.33E-05	1.99	3.47E-05	1.99	3.78E-05	1.97
3200	5.58E-06	1.99	5.92E-06	1.98	8.73E-06	1.99	5.36E-05	-.50

1D linear transport, 4th-order FD

Table: Third-order methods (SSPRK(3,3) behaves badly).

I	CFL = 0.05						CFL = 0.25					
	RK(3,3;1)	rate	RK(3,3;$\frac{1}{3}$)	rate	RK(4,3;1)	rate	RK(3,3;1)	rate	RK(3,3;$\frac{1}{3}$)	rate	RK(4,3;1)	rate
50	5.15E-02	–	4.76E-02	–	5.15E-02	–	5.48E-02	–	1.55E-01	–	6.08E-02	–
100	5.41E-03	3.25	5.41E-03	3.14	5.41E-03	3.25	5.15E-03	3.41	6.12E-02	1.35	6.15E-03	3.31
200	3.79E-04	3.83	3.79E-04	3.83	3.79E-04	3.83	3.92E-04	3.72	1.07E-03	5.84	3.83E-04	4.01
400	2.27E-05	4.06	2.27E-05	4.06	2.27E-05	4.06	2.89E-05	3.76	2.18E-04	2.29	2.30E-05	4.06
800	1.58E-06	3.85	1.58E-06	3.85	1.58E-06	3.85	3.20E-06	3.18	6.41E-05	1.77	1.59E-06	3.85
1600	9.12E-08	4.12	1.22E-07	3.69	8.13E-08	4.28	8.23E-07	1.96	1.83E-05	1.81	8.25E-08	4.27
3200	1.52E-08	2.58	6.84E-08	0.84	5.31E-09	3.94	2.40E-07	1.78	5.39E-06	1.76	5.39E-09	3.94

1D linear transport, 4th-order FD

Table: Fourth-order methods (SSPRK(5,4) behaves badly).

I	CFL = 0.05						CFL = 0.1					
	RK(4,4;$\frac{1}{2}$)	rate	RK(4,4;$\frac{3}{4}$)	rate	RK(5,4;$\frac{1}{2}$)	rate	RK(4,4;$\frac{1}{2}$)	rate	RK(4,4;$\frac{3}{4}$)	rate	RK(5,4;$\frac{1}{2}$)	rate
50	4.32E-02	–	4.72E-02	–	4.32E-02	–	6.35E-02	–	5.18E-02	–	6.28E-02	–
100	5.41E-03	3.00	5.40E-03	3.13	5.41E-03	3.00	5.36E-03	3.57	5.20E-03	3.31	5.66E-03	3.47
200	3.79E-04	3.84	3.79E-04	3.83	3.79E-04	3.83	3.79E-04	3.82	3.79E-04	3.78	3.79E-04	3.90
400	2.27E-05	4.06	2.27E-05	4.06	2.27E-05	4.06	2.27E-05	4.06	2.59E-05	3.87	2.27E-05	4.06
800	1.58E-06	3.85	1.58E-06	3.85	1.58E-06	3.84	1.58E-06	3.84	4.05E-06	2.68	1.58E-06	3.85
1600	8.13E-08	4.28	2.88E-07	2.46	8.58E-08	4.20	8.13E-08	4.28	9.94E-07	2.03	1.13E-07	3.80
3200	5.36E-09	3.92	6.98E-08	2.04	8.95E-09	3.26	4.97E-09	4.03	2.45E-07	2.02	2.72E-08	2.06

1D linear transport, 4th-order FD

Table: Fifth-order methods (least efficient method behaves badly).

I	CFL = 0.02				CFL = 0.025			
	RK(6,5; $\frac{5}{6}$)	rate	RK(6,5; $\frac{2}{3}$)	rate	RK(6,5; $\frac{5}{6}$)	rate	RK(6,5; $\frac{2}{3}$)	rate
50	5.19E-02	–	5.19E-02	–	5.20E-02	–	5.19E-02	–
100	5.41E-03	3.26	5.41E-03	3.26	5.41E-03	3.26	5.41E-03	3.26
200	3.79E-04	3.83	3.79E-04	3.83	3.79E-04	3.84	3.79E-04	3.84
400	2.27E-05	4.06	2.27E-05	4.06	2.27E-05	4.06	2.27E-05	4.06
800	1.58E-06	3.84	1.58E-06	3.85	1.58E-06	3.85	1.58E-06	3.85
1600	8.13E-08	4.28	8.48E-08	4.22	8.24E-08	4.26	8.71E-08	4.18
3200	6.24E-09	3.70	7.10E-09	3.58	6.32E-09	3.70	1.16E-08	2.91

2D linear transport, $\mathbb{P}_1$ FE (3th-order super-convergent)

- 4th-order FD in space.
- Linear transport $D := (0,1)^2$ with $\boldsymbol{\beta} := (0.9, 1)^\mathsf{T}$

$$\partial_t u + \nabla\cdot(\boldsymbol{\beta} u) = 0, \qquad u_0(\mathbf{x}) := \begin{cases} (4\frac{(x-x_0)(x_1-x)}{x_1-x_0})^4 \times (4\frac{(y-y_0)(y_1-y)}{y_1-y_0})^4 & x \in D_0 \\ 0 & oth. \end{cases}$$

with $D_0\{x_0 \le x \le x_1,\ y_0 \le y \le y_1\}$, $x_0 = y_0 = 0.1$, $x_1 = y_1 = 0.4$.

- Local maximum/minimum principle guaranteed at every grid point.
- Global maximum and minimum also exactly enforced.
- All errors computed at $T = 0.5$ with CFL $= 0.2$

2D linear transport, $\mathbb{P}_1$ FE (3th-order super-convergent)

Table: Second- and third-order ERK methods at CFL $= 0.2$.

	I	RK(2,2;1)	rate	RK(2,2;$\frac{1}{2}$)	rate	RK(3,3;1)	rate	RK(3,3;$\frac{1}{3}$)	rate	RK(4,3;1)	rate
L^∞-norm	51^2	2.58E-02	–	2.61E-02	–	3.27E-02	–	3.33E-02	–	3.29E-02	–
	101^2	1.32E-03	4.29	1.32E-03	4.30	7.82E-04	5.39	1.00E-03	5.05	8.02E-04	5.36
	201^2	4.73E-04	1.48	4.73E-04	1.49	8.28E-05	3.24	1.09E-04	3.21	8.03E-05	3.32
	401^2	1.26E-04	1.90	1.26E-04	1.90	9.44E-06	3.13	2.41E-05	2.17	9.33E-06	3.11
	801^2	3.22E-05	1.97	3.22E-05	1.97	1.03E-06	3.19	6.46E-06	1.90	1.06E-06	3.13

Linear transport with non-smooth solutions

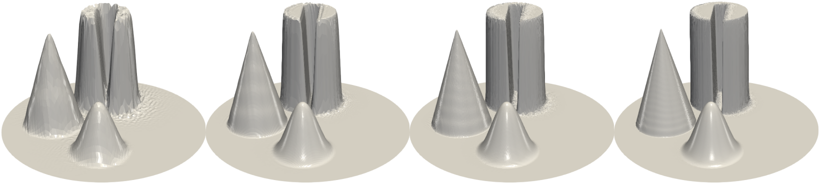

Figure: Three solids problem at $T = 1$, using RK(2,2;1) at CFL $= 0.25$. 2D $\mathbb{P}_1$ finite elements on non-uniform meshes. From left to right: $I = 6561$; $I = 24917$; $I = 98648$; $I = 389860$.

Linear transport with non-smooth solutions

Table: Three solids problem at $T = 1$ and CFL $= 0.25$. 2D $\mathbb{P}_1$ finite elements on non-uniform meshes. Relative error in the L^1-norm for methods RK(2,2;1) and RK(4,3;1).

I	RK(2,2;1)	rate	RK(4,3;1)	rate
1605	2.45E-01	–	2.49E-01	–
6561	1.28E-01	0.93	1.31E-01	0.92
24917	7.34E-02	0.81	7.49E-02	0.84
98648	4.26E-02	0.78	4.44E-02	0.76
389860	2.44E-02	0.81	2.56E-02	0.80

2D Burgers equation

2D Burgers equation in $D := (-.25, 1.75)^2$:

$$\partial_t u + \nabla\cdot(\mathbf{f}(u)) = 0, \qquad \mathbf{f}(u) := \tfrac{1}{2}(u^2, u^2)^\mathsf{T}, \qquad u(\mathbf{x}, 0) = u_0(\mathbf{x}) \text{ a.e. } \mathbf{x} \in D,$$

with the initial data

$$u_0(\mathbf{x}) := \begin{cases} 1 & \text{if } |x_1 - \frac{1}{2}| \le 1 \text{ and } |x_2 - \frac{1}{2}| \le 1 \\ -a & \text{otherwise.} \end{cases}$$

2D Burgers equation

Table: Burgers' equation. 2D $\mathbb{P}_1$ finite elements on uniform meshes. $T = 0.65$ at CFL $= 0.25$. Relative error in the L^1-norm for all the methods.

I	RK(2,2;1)	rate	RK(2,2;$\frac{1}{2}$)	rate	RK(3,3;1)	rate	RK(3,3;$\frac{1}{3}$)	rate	RK(4,3;1)	rate
51^2	7.71E-02	–	7.79E-02	–	7.71E-02	–	8.03E-02	–	7.71E-02	–
101^2	3.69E-02	1.06	3.73E-02	1.06	3.69E-02	1.06	3.85E-02	1.06	3.69E-02	1.06
201^2	2.30E-02	0.68	2.32E-02	0.68	2.30E-02	0.68	2.38E-02	0.70	2.30E-02	0.68
401^2	1.24E-02	0.90	1.24E-02	0.90	1.24E-02	0.90	1.27E-02	0.90	1.24E-02	0.90
801^2	6.47E-03	0.93	6.52E-03	0.93	6.48E-03	0.93	6.65E-03	0.93	6.47E-03	0.93

I	RK(4,4;$\frac{1}{2}$)	rate	RK(4,4;$\frac{3}{4}$)	rate	RK(5,4;0.51)	rate	RK(6,5;$\frac{5}{6}$)	rate	RK(6,5;$\frac{2}{3}$)	rate
51^2	7.94E-02	–	8.15E-02	–	7.79E-02	–	1.81E-01	–	9.29E-02	–
101^2	3.80E-02	1.06	3.89E-02	1.07	3.89E-02	1.00	8.56E-02	1.08	4.39E-02	1.08
201^2	2.36E-02	0.69	2.40E-02	0.70	2.47E-02	0.66	4.78E-02	0.84	2.72E-02	0.69
401^2	1.26E-02	0.90	1.28E-02	0.90	1.36E-02	0.86	2.38E-02	1.00	1.41E-02	0.95
801^2	6.61E-03	0.93	6.72E-03	0.94	7.11E-03	0.93	1.22E-02	0.97	7.24E-03	0.96

- Non-SSP methods converge as well as the SSP methods.

Outline

IDPMEX

The low-order, linearized update

- Let $\mathbf{F}^{\mathsf{L}}$ be low-order approximation of hyperbolic flux.
- Let $\mathbf{G}^{\mathsf{L,lin}}$ be Low-order linearized approximation of parabolic flux plus sources (i.e., approximation of $-\nabla\cdot(\mathbb{g}(\mathbf{u}, \nabla\mathbf{u})) + \mathbf{S}(\mathbf{u})$).
- Consider the low-order update (IMEX Euler)

$$\mathbb{M}^{\mathsf{L}}\mathbf{U}^{\mathsf{L},n+1} = \mathbb{M}^{\mathsf{L}}\mathbf{U}^{n} + \Delta t\mathbf{F}^{\mathsf{L}}(\mathbf{U}^{n}) + \Delta t\mathbf{G}^{\mathsf{L,lin}}(\mathbf{U}^{n}; \mathbf{U}^{\mathsf{L},n+1}).$$

The low-order, linearized update

- Assumption 1: (Forward Euler with low-order hyperbolic flux is invariant-domain preserving.) There exists $\Delta t^* > 0$ such that:
  - For every $\Delta t \in (0, \Delta t^*]$, the low-order hyperbolic flux satisfies

$$(\mathbf{V} \in \mathcal{A}^I) \Longrightarrow (\mathbf{U} := \mathbf{V} + \Delta t(\mathbb{M}^{\mathrm{L}})^{-1}\mathbf{F}^{\mathrm{L}}(\mathbf{V}) \in \mathcal{A}^I).$$

  - (Backward Euler with low-order, linearized, parabolic flux is invariant-domain preserving.) For all $\Delta t \in (0, \Delta t^*]$ and all $\mathbf{W} \in \mathcal{A}^I$, the operator $\mathbb{I} - \Delta t(\mathbb{M}^{\mathrm{L}})^{-1}\mathbf{G}^{\mathrm{L,lin}}(\mathbf{W};\cdot) : (\mathbb{R}^m)^I \to (\mathbb{R}^m)^I$ is bijective and

$$(\mathbf{V} \in \mathcal{A}^I) \Longrightarrow \left((\mathbb{I} - \Delta t(\mathbb{M}^{\mathrm{L}})^{-1}\mathbf{G}^{\mathrm{L,lin}}(\mathbf{W};\cdot))^{-1}\mathbf{V} \in \mathcal{A}^I\right).$$

Lemma (Low-order IDP Euler IMEX)

Let Assumption 1 hold. Assume that $\mathbf{U}^n \in \mathcal{A}^I$ and $\Delta t \in (0, \Delta t^*]$. Then, $\mathbf{U}^{\mathrm{L},n+1} \in \mathcal{A}^I$.

The high-order, linearized update (one Euler step)

- Assumption 2: There exists two nonlinear limiting operators ℓ^{hyp}, $\ell^{\text{par}} : \mathcal{A}^I \times (\mathbb{R}^m)^I \times (\mathbb{R}^m)^I \to (\mathbb{R}^m)^I$ s.t. for all $(\mathbf{V}, \mathbf{\Phi}^{\text{L}}, \mathbf{\Phi}^{\text{H}}) \in \mathcal{A}^I \times (\mathbb{R}^m)^I \times (\mathbb{R}^m)^I$,

$$
\begin{aligned}
(\mathbf{V} + \Delta t(\mathbb{M}^{\text{L}})^{-1}\mathbf{\Phi}^{\text{L}} \in \mathcal{A}^I) &\Longrightarrow (\ell^{\text{hyp}}(\mathbf{V}, \mathbf{\Phi}^{\text{L}}, \mathbf{\Phi}^{\text{H}}) \in \mathcal{A}^I),\\
(\mathbf{V} + \Delta t(\mathbb{M}^{\text{L}})^{-1}\mathbf{\Phi}^{\text{L}} \in \mathcal{A}^I) &\Longrightarrow (\ell^{\text{par}}(\mathbf{V}, \mathbf{\Phi}^{\text{L}}, \mathbf{\Phi}^{\text{H}}) \in \mathcal{A}^I).
\end{aligned}
$$

The high-order update (one Euler step)

- Given $\mathbf{U}^n \in \mathcal{A}^I$, the high-order update $\mathbf{U}^{n+1}$ is constructed as follows.
- **Step 1:** Compute the low-order and high-order hyperbolic updates defined by

$$\mathbb{M}^{\mathsf{L}}\mathbf{W}^{\mathsf{L},n+1} := \mathbb{M}^{\mathsf{L}}\mathbf{U}^n + \Delta t \mathbf{F}^{\mathsf{L}}(\mathbf{U}^n),$$
$$\mathbb{M}^{\mathsf{H}}\mathbf{W}^{\mathsf{H},n+1} := \mathbb{M}^{\mathsf{H}}\mathbf{U}^n + \Delta t \mathbf{F}^{\mathsf{H}}(\mathbf{U}^n).$$

- **Step 2:** Compute the hyperbolic fluxes $\mathbf{\Phi}^{\mathsf{L}}$, $\mathbf{\Phi}^{\mathsf{H}}$ (details given later) and limit

$$\mathbf{W}^{n+1} := \ell^{\mathsf{hyp}}(\mathbf{U}^n, \mathbf{\Phi}^{\mathsf{L}}, \mathbf{\Phi}^{\mathsf{H}}).$$

The high-order update (one Euler step)

- **Step 3:** Compute the low-order and high-order parabolic updates defined by

$$\mathbb{M}^{\mathsf{L}}\mathbf{U}^{\mathsf{L},n+1} - \Delta t\mathbf{G}^{\mathsf{L},\mathrm{lin}}(\mathbf{U}^{n};\mathbf{U}^{\mathsf{L},n+1}) := \mathbb{M}^{\mathsf{L}}\mathbf{W}^{n+1},$$
$$\mathbb{M}^{\mathsf{H}}\mathbf{U}^{\mathsf{H},n+1} - \Delta t\mathbf{G}^{\mathsf{H},\mathrm{lin}}(\mathbf{U}^{n};\mathbf{U}^{\mathsf{H},n+1}) := \mathbb{M}^{\mathsf{H}}\mathbf{W}^{n+1},$$

- **Step 4:** Compute the parabolic fluxes $\mathbf{\Psi}^{\mathsf{L}}$, $\mathbf{\Psi}^{\mathsf{H}}$ (details given later) and limit

$$\mathbf{U}^{n+1} := \ell^{\mathrm{par}}(\mathbf{W}^{n+1}, \mathbf{\Psi}^{\mathsf{L}}, \mathbf{\Psi}^{\mathsf{H}}).$$

Lemma (High-order IDP Euler IMEX)

Assume Assumptions 1 and 2. Assume that $\mathbf{U}^{n} \in \mathcal{A}^{I}$ and $\Delta t \in (0, \Delta t^{*}]$. Let $\mathbf{U}^{n+1}$ be defined as above. Then $\mathbf{U}^{n+1} \in \mathcal{A}^{I}$.

The high-order update (IMEX)

- Key idea: Consider low-order and high-order updates and limit.
  - Set $\mathbf{U}(t^n) = \mathbf{U}^n$ (with the induction assumption $\mathbf{U}^n \in \mathcal{A}$)
  - For $t \in (t^n, t^{n+1})$ solve

$$\mathbb{M}^{\mathsf{L}} \partial_t \mathbf{U} = \underbrace{\mathbf{F}^{\mathsf{L}}(\mathbf{U})}_{\textit{Explicit}} + \underbrace{\mathbf{G}^{\mathsf{H,lin}}(\mathbf{U}^n; \mathbf{U})}_{\textit{Implicit}},$$

$$\mathbb{M}^{\mathsf{H}} \partial_t \mathbf{U} = \underbrace{\mathbf{F}^{\mathsf{H}}(\mathbf{U}) + \mathbf{G}^{\mathsf{H}}(\mathbf{U}) - \mathbf{G}^{\mathsf{H,lin}}(\mathbf{U}^n; \mathbf{U})}_{\textit{Explicit}} + \underbrace{\mathbf{G}^{\mathsf{H,lin}}(\mathbf{U}^n; \mathbf{U})}_{\textit{Implicit}}.$$

The high-order update (IMEX)

- Explicit Butcher tableau

$$
\begin{array}{c|cccccc}
0 & 0 & & & & \\
c_2 & a^e_{2,1} & 0 & & & \\
c_3 & a^e_{3,1} & a^e_{3,2} & 0 & & \\
\vdots & \vdots & \ddots & \ddots & \ddots & \\
c_s & a^e_{s,1} & a^e_{s,2} & \cdots & a^e_{s,s-1} & 0 \\
\hline
1 & a^e_{1+1,1} & a^e_{s+1,2} & \cdots & a^e_{s+1,s-1} & a^e_{s+1,s}
\end{array}
$$

- Implicit Butcher tableau

$$
\begin{array}{c|cccccc}
0 & 0 & & & & \\
c_2 & a^i_{2,1} & a^i_{2,2} & & & \\
c_3 & a^i_{3,1} & a^i_{3,2} & a^i_{3,3} & & \\
\vdots & \vdots & \ddots & \ddots & \ddots & \\
c_s & a^i_{s,1} & a^i_{s,2} & \cdots & a^i_{s,s-1} & a^i_{s,s} \\
\hline
1 & a^i_{s+1,1} & a^i_{s+1,2} & \cdots & a^i_{s+1,s-1} & a^i_{s+1,s}
\end{array}
$$

Hyperbolic update

- Let $l \in \{2{:}\, s+1\}$
- Compute $\mathbf{W}^{\mathrm{L},l}$ and $\mathbf{W}^{\mathrm{H},l}$

$$\mathbb{M}^{\mathrm{L}}\mathbf{W}^{\mathrm{L},l} := \mathbb{M}^{\mathrm{L}}\mathbf{U}^{n,l'} + \Delta t(c_l - c_{l'})\mathbf{F}^{\mathrm{L}}(\mathbf{U}^{n,l'}),$$

$$\mathbb{M}^{\mathrm{H}}\mathbf{W}^{\mathrm{H},l} := \mathbb{M}^{\mathrm{H}}\mathbf{U}^{n,l'} + \Delta t \sum_{k\in\{1:\, l-1\}} (a^{\mathrm{e}}_{l,k} - a^{\mathrm{e}}_{l',k})\mathbf{F}^{\mathrm{H}}(\mathbf{U}^{n,k}).$$

- Use hyperbolic limiter

$$\mathbf{W}^{n,l} := \ell^{\mathrm{hyp}}(\mathbf{U}^{\mathrm{L},l}, \Phi^{\mathrm{L}}, \Phi^{\mathrm{H}}), \qquad \forall l \in \{2{:}\, s+1\}.$$

Parabolic update

- Let $l \in \{2{:}\, s+1\}$
- Compute $\mathbf{U}^{\mathsf{L},l}$ and $\mathbf{U}^{\mathrm{H},l}$

$$\mathbb{M}^{\mathsf{L}}\mathbf{U}^{\mathsf{L},l} := \mathbb{M}^{\mathsf{L}}\mathbf{W}^{n,l'} + \Delta t(c_l - c_{l'})\mathbf{G}^{\mathsf{L,lin}}(\mathbf{U}^n;\mathbf{U}^{\mathsf{L},l}),$$

$$\mathbb{M}^{\mathsf{H}}\mathbf{U}^{\mathrm{H},l} := \mathbb{M}^{\mathsf{H}}\mathbf{W}^{n,l'} + \Delta t a^{\mathsf{i}}_{l,l}\mathbf{G}^{\mathsf{H,lin}}(\mathbf{U}^n;\mathbf{U}^{\mathrm{H},l})$$

$$+ \sum_{k\in\{1:\, l-1\}} \Delta t\Big\{(a^{\mathsf{e}}_{l,k} - a^{\mathsf{e}}_{l',k})\mathbf{G}^{\mathsf{H}}(\mathbf{U}^{n,k}) + (a^{\mathsf{i}}_{l,k} - a^{\mathsf{i}}_{l',k} - a^{\mathsf{e}}_{l,k} + a^{\mathsf{e}}_{l',k})\mathbf{G}^{\mathsf{H,lin}}(\mathbf{U}^n;\mathbf{U}^{n,k})\Big\}.$$

- Notice $\Delta t(c_l - c_{l'}) > 0$, but $\Delta t a^{\mathsf{i}}_{l,l} \geq 0$ (i.e., $a^{\mathsf{i}}_{s+1,s+1} = 0$).
- Use hyperbolic limiter

$$\mathbf{U}^{n+1} := \ell^{\mathsf{hyp}}(\mathbf{W}^{\mathsf{L},l}, \Psi^{\mathsf{L}}, \Psi^{\mathsf{H}}), \qquad \forall l \in \{2{:}\, s+1\}.$$

Key result

Theorem (s-stage IDP-IMEX)

Assume Assumptions 1 and 2 and

$$\Delta t c_{\text{eff}} \leq \Delta t^*, \qquad c_{\text{eff}} := \max_{l \in \{2:s+1\}} (c_l - c_{l'})$$

If $\mathbf{U}^n \in \mathcal{A}^I$*, then* $\mathbf{U}^{\mathrm{L},n+1} \in \mathcal{A}^I$*.*

Example: Second-order

- Heun's method + Crank-Nicolson:

$$\begin{array}{c|cc} 0 & 0 & \\ 1 & 1 & 0 \\ \hline 1 & \frac{1}{2} & \frac{1}{2} \end{array} \qquad \begin{array}{c|cc} 0 & 0 & \\ 1 & \frac{1}{2} & \frac{1}{2} \\ \hline 1 & \frac{1}{2} & \frac{1}{2} \end{array}$$

- $l' = l - 1$ for all $l \in \{2{:}3\}$, and the efficiency ratio is $\frac{1}{2}$.

Example: Second-order

- Explicit and implicit midpoint rules.

$$\begin{array}{c|cc} 0 & 0 & \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \hline 1 & 0 & 1 \end{array} \qquad \begin{array}{c|cc} 0 & 0 & \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \hline 1 & 0 & 1 \end{array}$$

- $l' = l - 1$ for all $l \in \{2{:}3\}$, and the efficiency ratio is 1.

Example: Third-order

- Two-stage, third-order (A-stable) SDIRK method **Crouzeix (1975), Norsett (1974)**

$$
\begin{array}{c|ccc}
0 & 0 & & \\
\gamma & \gamma & 0 & \\
1-\gamma & \gamma-1 & 2-2\gamma & 0 \\
\hline
1 & 0 & \frac{1}{2} & \frac{1}{2}
\end{array}
\qquad
\begin{array}{c|ccc}
0 & 0 & & \\
\gamma & 0 & \gamma & \\
1-\gamma & 0 & 1-2\gamma & \gamma \\
\hline
1 & 0 & \frac{1}{2} & \frac{1}{2}
\end{array}
$$

- with $\gamma := \frac{1}{2} + \frac{1}{2\sqrt{3}} \approx 0.78867$.
- The values for I' are $(1, 1, 2)$. The efficiency ratio is $\frac{1}{3}\gamma \approx 0.26$.

Example: Third-order

- Three-stage, third-order

$$
\begin{array}{c|ccc}
0 & 0 & & \\
\frac{1}{3} & \frac{1}{3} & 0 & \\
\frac{2}{3} & 0 & \frac{2}{3} & 0 \\
\hline
1 & \frac{1}{4} & 0 & \frac{3}{4}
\end{array}
\qquad
\begin{array}{c|ccc}
0 & 0 & & \\
\frac{1}{3} & \frac{1}{3}-\gamma & \gamma & \\
\frac{2}{3} & \gamma & \frac{2}{3}-2\gamma & \gamma \\
\hline
1 & \frac{1}{4} & 0 & \frac{3}{4}
\end{array}
$$

- With $\gamma := \frac{1}{2} + \frac{1}{2\sqrt{3}} \approx 0.78867$.
- We have $l' = l - 1$ for all $l \in \{2{:}4\}$, and the efficiency is 1.

Important omitted details

- The definition of $\mathbf{G}^{\mathrm{L,lin}}$ is problem-dependent.
- Conservation
- Limiting done with the Flux Transport Correction technique **Zalezak (1979)** if the constraints are not affine
- Limiting done with convex limiting (**Guermond, Popov, Tomas (2019)**) if the constraints are not affine.

Conclusions

- Every ERK and IMEX methods can be made invariant-domain preserving.