Residual distribution and applications to kinetic models

Davide Torlo

Dipartimento di Matematica “Guido Castelnuovo”
Università di Roma Sapienza

Kassel

6 November 2025

joint work with Mario Ricchiuto and Rémi Abgrall

Outline

Outline

Residual Distribution[1]

- High order
- FE based
- Compact stencil
- Can recast some other FV, FE, GF, DG schemes[2]

Finite Element Setting

$$\partial_t u+\nabla_x \cdot f(u) = S(u)$$

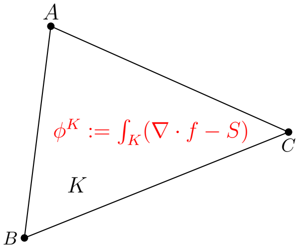
$$V_h = \{u \in \mathcal{C}^0(\Omega_h),\, u|_K \in \mathbb{P}^p,\, \forall K \in \Omega_h\}$$

$$u(x) = \sum_{\sigma\in D_h} u_\sigma \varphi_\sigma(x) = \sum_{K\in\Omega_h}\sum_{\sigma\in K} u_\sigma\varphi_\sigma(x)|_K$$

[1] R. Abgrall. Some remarks about conservation for residual distribution schemes. Computational Methods in Applied Mathematics, 2018. DOI: https://doi.org/10.1515/cmam-2017-0056.

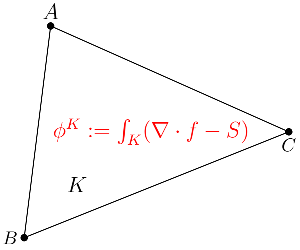
Residual Distribution - Spatial Discretization

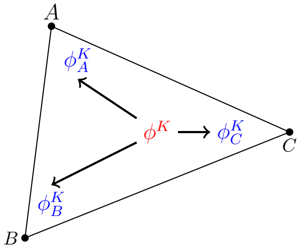
1. Define $\forall K \in \Omega_h$ a fluctuation term (total residual) $\phi^K = \int_K \nabla \cdot f(u) - S(u)dx$



Residual Distribution - Spatial Discretization

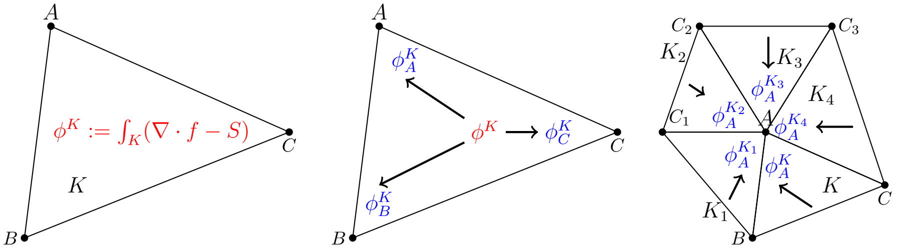
1. Define $\forall K \in \Omega_h$ a fluctuation term (total residual) $\phi^K = \int_K \nabla \cdot f(u) - S(u) dx$
2. Define nodal residuals ϕ_σ^K $\forall \sigma \in K$: $\phi^K = \sum_{\sigma \in K} \phi_\sigma^K, \quad \forall K \in \Omega_h.$





Residual Distribution - Spatial Discretization

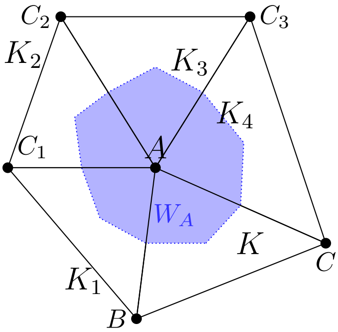
1. Define $\forall K \in \Omega_h$ a fluctuation term (total residual) $\phi^K = \int_K \nabla \cdot f(u) - S(u) dx$
2. Define nodal residuals ϕ^K_σ $\forall \sigma \in K : \phi^K = \sum_{\sigma \in K} \phi^K_\sigma, \quad \forall K \in \Omega_h.$
3. The resulting scheme is $W_\sigma \partial_t u_\sigma + \sum_{K|\sigma \in K} \phi^K_\sigma = 0, \quad \forall \sigma \in D_h$, with $W_\sigma = \sum_{K|\sigma\in K} \int_K \varphi_\sigma(x) dx.$

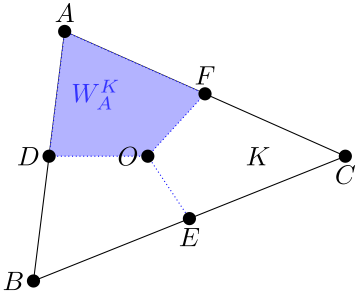


Residual distribution mass interpretation

What are we actually evolving in time? $\sum_{K|\sigma\in K}\int_K \varphi_\sigma(x)dx\, u_\sigma =: \sum_{K|\sigma\in K} W_\sigma^K u_\sigma =: W_\sigma u_\sigma$

$$\partial_t u_\sigma + \frac{1}{W_\sigma}\sum_{K|\sigma\in K}\phi_\sigma^K(u) = 0.$$





Examples of residual distribution splittings

Central scheme

$$\phi_\sigma^K := \int_K \varphi_\sigma \left(\nabla \cdot f(u) - S(u)\right) dx$$

Rusanov scheme

$$\phi_\sigma^K := \int_K \varphi_\sigma \left(\nabla \cdot f(u) - S(u)\right) dx + \alpha \left(u_\sigma - \bar{u}^K\right)$$

SUPG scheme

$$\phi_\sigma^K := \int_K \varphi_\sigma \left(\nabla \cdot f(u) - S(u)\right) dx$$
$$+ h\theta \int_K \nabla \varphi_\sigma J_f(u) \tau \left(\nabla \cdot f(u) - S(u)\right) dx$$

Jump stabilization

$$\phi_\sigma^K := \int_K \varphi_\sigma \left(\nabla \cdot f(u) - S(u)\right) dx$$
$$+ h^2 \theta \int_{\partial K} [\![\nabla \varphi_\sigma]\!][\![\nabla u]\!] d\gamma$$

Discontinuous Galerkin

$$\phi_\sigma^K := \int_K -\nabla \varphi_\sigma \cdot f(u) - \varphi_\sigma S(u) dx$$
$$+ \int_{\partial K} \varphi_\sigma \hat{f}(u_h, u_h^+) \cdot \mathbf{n} d\gamma$$

Examples of residual distribution splittings

Central scheme

$$\phi_\sigma^K := \int_K \varphi_\sigma \left(\nabla \cdot f(u) - S(u)\right) dx$$

Rusanov scheme

$$\phi_\sigma^K := \int_K \varphi_\sigma \left(\nabla \cdot f(u) - S(u)\right) dx + \alpha \left(u_\sigma - \bar{u}^K\right)$$

SUPG scheme

$$\phi_\sigma^K := \int_K \varphi_\sigma \left(\nabla \cdot f(u) - S(u)\right) dx$$
$$+ h\theta \int_K \nabla \varphi_\sigma J_f(u) \tau \left(\nabla \cdot f(u) - S(u)\right) dx$$

Jump stabilization

$$\phi_\sigma^K := \int_K \varphi_\sigma \left(\nabla \cdot f(u) - S(u)\right) dx$$
$$+ h^2 \theta \int_{\partial K} [\![\nabla \varphi_\sigma]\!] [\![\nabla u]\!] d\gamma$$

Discontinuous Galerkin

$$\phi_\sigma^K := \int_K -\nabla \varphi_\sigma \cdot f(u) - \varphi_\sigma S(u) dx$$
$$+ \int_{\partial K} \varphi_\sigma \hat{f}(u_h, u_h^+) \cdot \mathbf{n} d\gamma$$

Examples of residual distribution splittings

Central scheme

$$\phi_\sigma^K := \int_K \varphi_\sigma \left(\nabla \cdot f(u) - S(u)\right) dx$$

Rusanov scheme

$$\phi_\sigma^K := \int_K \varphi_\sigma \left(\nabla \cdot f(u) - S(u)\right) dx + \alpha \left(u_\sigma - \bar{u}^K\right)$$

SUPG scheme

$$\phi_\sigma^K := \int_K \varphi_\sigma \left(\nabla \cdot f(u) - S(u)\right) dx$$
$$+h\theta \int_K \nabla\varphi_\sigma J_f(u)\tau \left(\nabla \cdot f(u) - S(u)\right) dx$$

Jump stabilization

$$\phi_\sigma^K := \int_K \varphi_\sigma \left(\nabla \cdot f(u) - S(u)\right) dx$$
$$+h^2\theta \int_{\partial K} [\![\nabla\varphi_\sigma]\!][\![\nabla u]\!] d\gamma$$

Discontinuous Galerkin

$$\phi_\sigma^K := \int_K -\nabla\varphi_\sigma \cdot f(u) - \varphi_\sigma S(u) dx$$
$$+ \int_{\partial K} \varphi_\sigma \hat{f}(u_h, u_h^+) \cdot \mathbf{n} d\gamma$$

Examples of residual distribution splittings

Central scheme

$$\phi_\sigma^K := \int_K \varphi_\sigma \left(\nabla \cdot f(u) - S(u)\right) dx$$

Rusanov scheme

$$\phi_\sigma^K := \int_K \varphi_\sigma \left(\nabla \cdot f(u) - S(u)\right) dx + \alpha \left(u_\sigma - \bar{u}^K\right)$$

SUPG scheme

$$\phi_\sigma^K := \int_K \varphi_\sigma \left(\nabla \cdot f(u) - S(u)\right) dx$$
$$+h\theta \int_K \nabla\varphi_\sigma J_f(u)\tau \left(\nabla \cdot f(u) - S(u)\right) dx$$

Jump stabilization

$$\phi_\sigma^K := \int_K \varphi_\sigma \left(\nabla \cdot f(u) - S(u)\right) dx$$
$$+h^2\theta \int_{\partial K} [\![\nabla\varphi_\sigma]\!][\![\nabla u]\!] d\gamma$$

Discontinuous Galerkin

$$\phi_\sigma^K := \int_K -\nabla\varphi_\sigma \cdot f(u) - \varphi_\sigma S(u) dx$$
$$+ \int_{\partial K} \varphi_\sigma \hat{f}(u_h, u_h^+) \cdot \mathbf{n} d\gamma$$

Examples of residual distribution splittings

Central scheme

$$\phi_\sigma^K := \int_K \varphi_\sigma \left(\nabla \cdot f(u) - S(u)\right) dx$$

Rusanov scheme

$$\phi_\sigma^K := \int_K \varphi_\sigma \left(\nabla \cdot f(u) - S(u)\right) dx + \alpha \left(u_\sigma - \bar{u}^K\right)$$

SUPG scheme

$$\phi_\sigma^K := \int_K \varphi_\sigma \left(\nabla \cdot f(u) - S(u)\right) dx$$
$$+h\theta \int_K \nabla \varphi_\sigma J_f(u) \tau \left(\nabla \cdot f(u) - S(u)\right) dx$$

Jump stabilization

$$\phi_\sigma^K := \int_K \varphi_\sigma \left(\nabla \cdot f(u) - S(u)\right) dx$$
$$+h^2\theta \int_{\partial K} [\![\nabla \varphi_\sigma]\!][\![\nabla u]\!] d\gamma$$

Discontinuous Galerkin

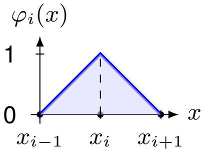
$$\phi_\sigma^K := \int_K -\nabla \varphi_\sigma \cdot f(u) - \varphi_\sigma S(u) dx$$
$$+ \int_{\partial K} \varphi_\sigma \hat{f}(u_h, u_h^+) \cdot \mathbf{n} d\gamma$$

Outline

Examples of Rusanov in residual distribution in 1D with $\mathbb{P}^1$

1D case

$$
\begin{aligned}
&W_\sigma = \Delta x\\
&\phi_\sigma^K := \int_K \varphi_\sigma \partial_x f(u) dx + \max_{j\in K} \rho(\partial_u f(u_j))(u_\sigma - \bar{u}^K)\\
&\partial_t u_i + \frac{1}{\Delta x}\left(\phi_i^{i+\frac{1}{2}} + \phi_i^{i-\frac{1}{2}}\right) = 0\\
&\partial_t u_i + \frac{1}{\Delta x}\left(\int_{x_i}^{x_{i+1}} \varphi_i \partial_x f(u)\, dx + \max_{j\in\{i,i+1\}} \rho(\partial_u f(u_j))(u_i - \frac{u_i+u_{i+1}}{2})\right)\\
&\quad + \frac{1}{\Delta x}\left(\int_{x_{i-1}}^{x_i} \varphi_i \partial_x f(u)\, dx + \max_{j\in\{i,i-1\}} \rho(\partial_u f(u_j))(u_i - \frac{u_i+u_{i-1}}{2})\right) = 0
\end{aligned}
$$



Examples of Rusanov in residual distribution in 1D with $\mathbb{P}^1$

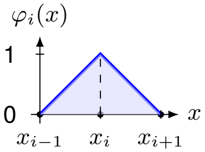
1D case

$$W_\sigma = \Delta x$$

$$\phi_\sigma^K := \int_K \varphi_\sigma \partial_x f(u) dx + \max_{j\in K} \rho(\partial_u f(u_j))(u_\sigma - \bar{u}^K)$$

$$\partial_t u_i + \frac{1}{\Delta x}\left(\phi_i^{i+\frac{1}{2}} + \phi_i^{i-\frac{1}{2}}\right) = 0$$

$$\partial_t u_i + \frac{1}{\Delta x}\left(\frac{f(u_{i+1})-f(u_i)}{2} + \max_{j\in\{i,i+1\}} \rho(\partial_u f(u_j))(u_i - \frac{u_i+u_{i+1}}{2})\right)$$
$$+ \frac{1}{\Delta x}\left(\frac{f(u_i)-f(u_{i-1})}{2} + \max_{j\in\{i,i-1\}} \rho(\partial_u f(u_j))(u_i - \frac{u_i+u_{i-1}}{2})\right) = 0$$



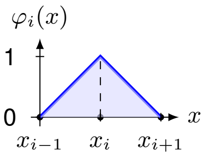
Examples of Rusanov in residual distribution in 1D with $\mathbb{P}^1$

1D case

$$
\begin{aligned}
W_\sigma &= \Delta x\\
\phi_\sigma^K &:= \int_K \varphi_\sigma \partial_x f(u) dx + \max_{j\in K} \rho(\partial_u f(u_j))(u_\sigma - \bar{u}^K)
\end{aligned}
$$

$$
\partial_t u_i + \frac{1}{\Delta x}\left(\phi_i^{i+\frac{1}{2}} + \phi_i^{i-\frac{1}{2}} \right) = 0
$$

$$
\begin{aligned}
\partial_t u_i &+ \frac{1}{\Delta x}\left(\frac{f(u_{i+1})-f(u_i)}{2} + \max_{j\in\{i,i+1\}} \rho(\partial_u f(u_j)) \frac{u_i - u_{i+1}}{2} \right) + \frac{f(u_i)}{\Delta x}\\
&+ \frac{1}{\Delta x}\left(\frac{f(u_i)-f(u_{i-1})}{2} + \max_{j\in\{i,i-1\}} \rho(\partial_u f(u_j)) \frac{u_i - u_{i-1}}{2} \right) - \frac{f(u_i)}{\Delta x} = 0
\end{aligned}
$$



Examples of Rusanov in residual distribution in 1D with $\mathbb{P}^1$

1D case

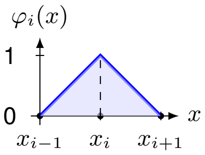
$$W_\sigma = \Delta x$$

$$\phi_\sigma^K := \int_K \varphi_\sigma \partial_x f(u) dx + \max_{j\in K} \rho(\partial_u f(u_j))(u_\sigma - \bar{u}^K)$$

$$\partial_t u_i + \frac{1}{\Delta x}\left(\phi_i^{i+\frac{1}{2}} + \phi_i^{i-\frac{1}{2}}\right) = 0$$

$$\partial_t u_i + \frac{1}{\Delta x}\left(\frac{f(u_{i+1}) - f(u_i)}{2} + \max_{j\in\{i,i+1\}} \rho(\partial_u f(u_j))\frac{u_i - u_{i+1}}{2}\right) + \frac{f(u_i)}{\Delta x}$$
$$+ \frac{1}{\Delta x}\left(\frac{f(u_i) - f(u_{i-1})}{2} + \max_{j\in\{i,i-1\}} \rho(\partial_u f(u_j))\frac{u_i - u_{i-1}}{2}\right) - \frac{f(u_i)}{\Delta x} = 0$$

$$\partial_t u_i + \frac{1}{\Delta x}\left(\frac{f(u_{i+1}) + f(u_i)}{2} - \max_{j\in\{i,i+1\}} \rho(\partial_u f(u_j))\frac{u_{i+1} - u_i}{2}\right)$$
$$- \frac{1}{\Delta x}\left(\frac{f(u_i) + f(u_{i-1})}{2} - \max_{j\in\{i,i-1\}} \rho(\partial_u f(u_j))\frac{u_i - u_{i-1}}{2}\right) = 0$$



General conservative FV into a RD $\mathbb{P}^1$ scheme on triangles

In Finite Volume scheme

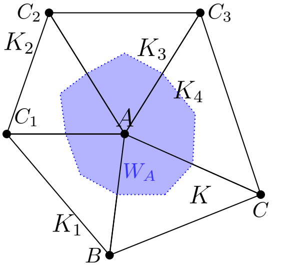
$$\partial_t u_\sigma + \frac{1}{W_\sigma} \sum_{\substack{K|\sigma\in K\\ \sigma'\in K|\sigma'\neq\sigma}} \hat{f}_{\mathbf{n}^K_{\sigma\sigma'}}(u_\sigma, u_{\sigma'}) = 0$$

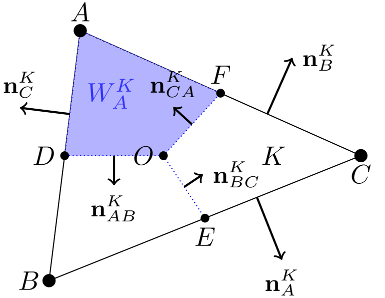
a numerical flux is defined such that **conservation** is preserved

$$\hat{f}_{\mathbf{n}}(u^+, u^-) = -\hat{f}_{-\mathbf{n}}(u^-, u^+)$$

with consistency property

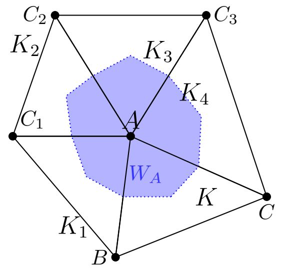
$$\hat{f}_{\mathbf{n}}(u,u) = f(u)\cdot \mathbf{n}.$$





General conservative FV into a RD $\mathbb{P}^1$ scheme on triangles

$$\partial_t u_\sigma + \frac{1}{W_\sigma} \sum_{\substack{K|\sigma\in K\\ \sigma'\in K|\sigma'\neq\sigma}} \hat{f}_{\mathbf{n}_{\sigma\sigma'}}(u_\sigma,u_{\sigma'})=0$$



Properties

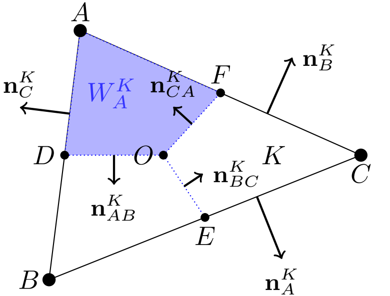
$$|\mathbf{n}_A^K| = |BC|$$
$$\mathbf{n}^K_{\sigma\sigma'} = -\mathbf{n}^K_{\sigma'\sigma}$$
$$\mathbf{n}_A^K = 2(\mathbf{n}^K_{AB}+\mathbf{n}^K_{AC})$$

$$\mathbf{n}^K_{AB}+\mathbf{n}^K_{BC}+\mathbf{n}^K_{CA}=0$$
$$\sum_{\sigma\in K}\mathbf{n}^K_\sigma = 0$$
$$\sum_{\substack{K|\sigma\in K\\ \sigma'\in K|\sigma'\neq\sigma}} \mathbf{n}^K_{\sigma\sigma'}=0$$

$$\sum_{\substack{K|\sigma\in K\\ \sigma'\in K|\sigma'\neq\sigma}} \hat{f}_{\mathbf{n}^K_{\sigma\sigma'}}(u_\sigma,u_{\sigma'}) = \sum_{\substack{K|\sigma\in K\\ \sigma'\in K|\sigma'\neq\sigma}} \hat{f}_{\mathbf{n}^K_{\sigma\sigma'}}(u_\sigma,u_{\sigma'}) - \sum_{\substack{K|\sigma\in K\\ \sigma'\in K|\sigma'\neq\sigma}} \mathbf{n}^K_{\sigma\sigma'}\cdot f(u_\sigma)$$

$$= \sum_{\substack{K|\sigma\in K\\ \sigma'\in K|\sigma'\neq\sigma}} \left(\hat{f}_{\mathbf{n}^K_{\sigma\sigma'}}(u_\sigma,u_{\sigma'}) - f(u_\sigma)\cdot\mathbf{n}^K_{\sigma\sigma'}\right)$$

$$\phi^K_\sigma(u) := \sum_{\sigma'\in K|\sigma'\neq\sigma} \left(\hat{f}_{\mathbf{n}^K_{\sigma\sigma'}}(u_\sigma,u_{\sigma'}) - f(u_\sigma)\cdot\mathbf{n}^K_{\sigma\sigma'}\right)$$



General conservative FV into a RD $\mathbb{P}^1$ scheme on triangles

Properties/Definitions

$$\phi_\sigma^K := \sum_{\substack{\sigma' \in K| \\ \sigma' \neq \sigma}} \left(\hat{f}_{\mathbf{n}_{\sigma\sigma'}^K}(u_\sigma, u_{\sigma'}) - f(u_\sigma) \cdot \mathbf{n}_{\sigma\sigma'}^K \right)$$

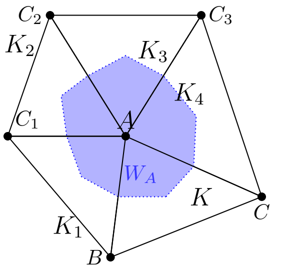
$$\int_K \nabla \varphi_\sigma dx = -\frac{1}{2}\mathbf{n}_\sigma^K$$

$$\hat{f}_{\mathbf{n}}(u^+, u^-) = -\hat{f}_{-\mathbf{n}}(u^-, u^+)$$

$$|\mathbf{n}_A^K| = |BC|$$

$$\mathbf{n}_{\sigma\sigma'}^K = -\mathbf{n}_{\sigma'\sigma}^K$$

$$\mathbf{n}_A^K = 2(\mathbf{n}_{AB}^K + \mathbf{n}_{AC}^K)$$

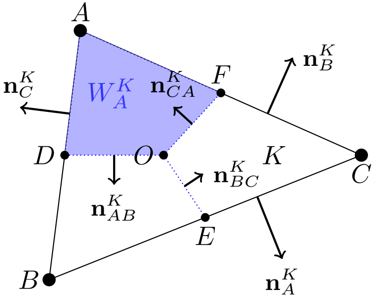


Conservation in RD

$$\sum_{\sigma \in K} \phi_\sigma^K(u) = \sum_{\sigma \in K} \sum_{\sigma' \in K|\sigma' \neq \sigma} \left(\hat{f}_{\mathbf{n}_{\sigma\sigma'}^K}(u_\sigma, u_{\sigma'}) - f(u_\sigma) \cdot \mathbf{n}_{\sigma\sigma'}^K \right)$$

$$= -\sum_{\sigma \in K} f(u_\sigma) \cdot \frac{\mathbf{n}_\sigma^K}{2} = \sum_{\sigma \in K} \int_K f(u_\sigma) \nabla \varphi_\sigma dx$$

$$= \int_K \nabla f^h(u) dx.$$



Summary FV is RD

FV to RD

- Start from a FV on dual mesh
- Define nodal residuals as $\phi_\sigma^K := \sum_{\substack{\sigma' \in K| \\ \sigma' \neq \sigma}} \left(\hat{f}_{\mathbf{n}^K_{\sigma\sigma'}}(u_\sigma, u_{\sigma'}) - f(u_\sigma) \cdot \mathbf{n}^K_{\sigma\sigma'} \right)$
- Conservation property holds $\sum_{\sigma \in K} \phi_\sigma^K(u) = \int_K \nabla f^h(u) dx = \phi^K(u)$

Outline

What are we looking for?

We start from residual distribution

$$\partial_t u_\sigma + \frac{1}{W_\sigma} \sum_{K|\sigma\in K} \phi_\sigma^K(u) = 0$$

$$\sum_{\sigma\in K} \phi_\sigma^K(u) = \int_K \nabla f^h(u) - S(u) dx.$$

Goal: FV formulation

$$\partial_t u_\sigma + \frac{1}{W_\sigma} \sum_{\sigma'\neq\sigma} \hat{f}_{\sigma,\sigma'}(\mathbf{u}) = 0$$

$$\hat{f}_{\sigma,\sigma'}(u,u,u,\dots,u) = f(u)\cdot \mathbf{n}_{\sigma,\sigma'}.$$

Where $\hat{f}_{\sigma,\sigma'}(\mathbf{u}) = \hat{f}_{\sigma,\sigma'}(u_{\sigma_1}, u_{\sigma_2}, u_{\sigma_3}, \dots, u_{\sigma_{\#K}})$ is a multidimensional numerical flux that might depend on all DOFs in K.

What we have to do

Disaggregate the nodal residuals into numerical fluxes. We have many more unknowns than equations! “Easy task”, but we should check that the problem is solvable!

For every element K, we want to find

$$\hat{f}_{\sigma,\sigma'} = -\hat{f}_{\sigma',\sigma}$$

for all connected σ, σ' such that

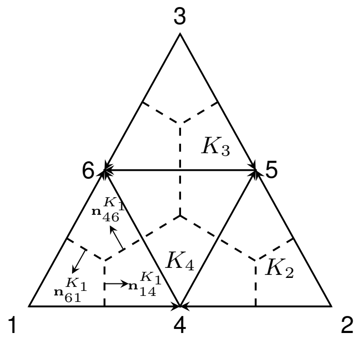
$$\sum_{\sigma'\neq\sigma} \hat{f}_{\mathbf{n}^K_{\sigma,\sigma'}}(\mathbf{u}) = \phi_\sigma(u) := \sum_{K|\sigma\in K} \phi^K_\sigma(u).$$

Hence, we define an order of the connections $[\sigma, \sigma']$ so that

$$\begin{cases} \hat{f}^K_{\sigma,\sigma'} = \hat{f}^K_{\{\sigma,\sigma'\}} & \text{if } [\sigma,\sigma'] \text{ is ordered,}\\ \hat{f}^K_{\sigma,\sigma'} = -\hat{f}^K_{\{\sigma,\sigma'\}} & \text{if } [\sigma',\sigma] \text{ is ordered.}\end{cases}$$

E.g. $\hat{f}_{\mathbf{n}^{K_1}_{1,4}} = -\hat{f}_{\mathbf{n}^{K_1}_{4,1}} = \hat{f}^{K_1}_{\{1,4\}}$ as $[1, 4]$ is ordered.

$\mathbb{P}^2$ element



Note that here all the 6 DoFs contribute to each nodal residual, even if far away, hence also to the numerical fluxes.

RD to FV

This can be recast into a linear algebra problem

$$\sum_{\sigma'} \varepsilon_{\sigma,\sigma'} \hat{f}_{\{\sigma,\sigma'\}} = \phi_\sigma(u) := \sum_{K|\sigma\in K} \phi_\sigma^K(u)$$

where

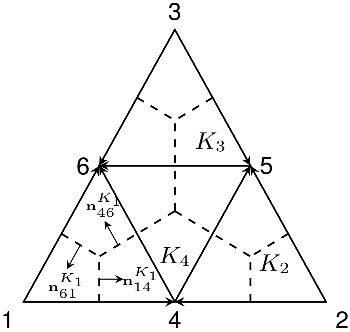
$$\varepsilon_{\sigma,\sigma'} = \begin{cases} 1 & \text{if } [\sigma,\sigma'] \text{ is ordered,} \\ -1 & \text{if } [\sigma',\sigma] \text{ is ordered,} \\ 0 & \text{if } \sigma,\sigma' \text{ not connected.} \end{cases}$$

Unknowns are only direct connected edges $[\sigma,\sigma']$. In matrix form is

$$A\hat{\mathbf{f}} = \phi,$$

with A the matrix of the incidence matrix of the graph which is rectangular of size $(\#\text{DoFs}) \times (\#\text{edges})$.

$\mathbb{P}^2$ element



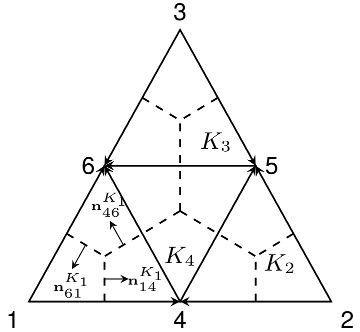
Note that here all the 6 DoFs contribute to each nodal residual, even if far away, hence also to the numerical fluxes.

$$A\hat{\mathbf{f}} = \phi$$

Consider $L = AA^T$, the graph Laplacian matrix, which is square of size $(\#\text{DoFs}) \times (\#\text{DoFs})$.

- L has rank $(\#\text{DoFs}) - 1$ as the kernel is spanned by the vector of all ones $\mathbf{1}$ and its image is span$(\mathbf{1})^\perp$.
- Define L^{-1} using the restriction of L : span$(\mathbf{1})^\perp \to$ span$(\mathbf{1})^\perp$.
- $\hat{\mathbf{f}} = A^T L^{-1}\phi$ is a solution of the problem, as ϕ is in the image of A due to conservation.

$\mathbb{P}^2$ element



Note that here all the 6 DoFs contribute to each nodal residual, even if far away, hence also to the numerical fluxes.

Outline

Kinetic Models

Kinetic relaxation models by D. Aregba-Driollet and R. Natalini[2].

Hyperbolic limit equation

$$u_t + \sum_{d=1}^{D} \partial_{x_d} A_d(u) = 0, \quad u : \Omega \to \mathbb{R}^K.$$

$$\begin{tikzcd}
\Omega \subset \mathbb{R}^D \arrow[r, "u"] \arrow[dr, bend right, "f"'] & \mathbb{R}^K \arrow[loop right, "A_d"] \arrow[d, bend right, "M"'] \\
& \mathbb{R}^L \arrow[loop right, "\Lambda_d"] \arrow[u, bend right, "P"']
\end{tikzcd}$$

Relaxation system

$$f_t^\varepsilon + \sum_{d=1}^{D} \Lambda_d \partial_{x_d} f^\varepsilon = \frac{1}{\varepsilon}\left(M(Pf^\varepsilon) - f^\varepsilon\right), \quad f^\varepsilon : \Omega \to \mathbb{R}^L$$

$$Pf^\varepsilon \to u, \quad P(M(u)) = u, \quad P\Lambda_d M(u) = A_d(u).$$

[2] D. Aregba-Driollet and R. Natalini. Discrete kinetic schemes for multidimensional systems of conservation laws. SIAM J. Numer. Anal., 37(6):1973–2004, 2000.

Scalar 1D example: Jin–Xin system

Let $u:\mathbb{R}\to\mathbb{R}$ and $A(u)=a(u)$ with $a:\mathbb{R}\to\mathbb{R}$.
Limit equation

(1) $$u_t + a(u)_x = 0.$$

Now, let $f=(f_1,f_2)$ with $Pf=f_1$, $\Lambda=\begin{pmatrix}0 & 1\\ \lambda^2 & 0\end{pmatrix}$, $M_1(u)=u$, $M_2(u)=a(u)$.

So that

$$f_t^\varepsilon + \sum_{d=1}^{D}\Lambda_d\partial_{x_d}f^\varepsilon = \begin{pmatrix}f_1\\ f_2\end{pmatrix}_t + \partial_x\begin{pmatrix}f_2\\ \lambda^2 f_1\end{pmatrix} = \frac{1}{\varepsilon}\left(M(Pf^\varepsilon)-f^\varepsilon\right) = \begin{pmatrix}\frac{f_1-f_1}{\varepsilon}\\ \frac{a(f_1)-f_2}{\varepsilon}\end{pmatrix}$$

$$\begin{cases}\partial_t f_1 + \partial_x f_2 = 0\\ \partial_t f_2 + \lambda^2\partial_x f_1 = \frac{a(f_1)-f_2}{\varepsilon}\end{cases}$$

Whitham's subcharacteristic conditions: Jin–Xin

$$\begin{cases} \partial_t f_1 + \partial_x f_2 = 0 \\ \partial_t f_2 + \lambda^2 \partial_x f_1 = \frac{a(f_1)-f_2}{\varepsilon} \end{cases}$$

$$f_2 = a(f_1) - \varepsilon \left(\partial_t f_2 + \lambda^2 \partial_x f_1 \right)$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(\partial_t f_2 + \lambda^2 \partial_x f_1 \right) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(a'(f_1) \partial_t f_1 + \lambda^2 \partial_x f_1 \right) + \mathcal{O}(\varepsilon^2) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(-a'(f_1) \partial_x f_2 + \lambda^2 \partial_x f_1 \right) + \mathcal{O}(\varepsilon^2) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(-a'(f_1) a'(f_1) \partial_x f_1 + \lambda^2 \partial_x f_1 \right) + \mathcal{O}(\varepsilon^2) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \left(-a'(f_1)^2 + \lambda^2 \right) \partial_x^2 f_1 + \mathcal{O}(\varepsilon^2) = 0$$

So the stability condition to be dissipative is

$$\lambda^2 \geq a'(f_1)^2 \quad \Leftrightarrow \quad \lambda \geq |a'(u)|, \quad \forall u \in \mathcal{R}. \tag{2}$$

Whitham's subcharacteristic conditions: Jin–Xin

$$\begin{cases} \partial_t f_1 + \partial_x f_2 = 0 \\ \partial_t f_2 + \lambda^2 \partial_x f_1 = \frac{a(f_1)-f_2}{\varepsilon} \end{cases}$$

$$f_2 = a(f_1) - \varepsilon \left(\partial_t f_2 + \lambda^2 \partial_x f_1\right)$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(\partial_t f_2 + \lambda^2 \partial_x f_1\right) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(a'(f_1)\partial_t f_1 + \lambda^2 \partial_x f_1\right) + \mathcal{O}(\varepsilon^2) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(-a'(f_1)\partial_x f_2 + \lambda^2 \partial_x f_1\right) + \mathcal{O}(\varepsilon^2) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(-a'(f_1)a'(f_1)\partial_x f_1 + \lambda^2 \partial_x f_1\right) + \mathcal{O}(\varepsilon^2) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \left(-a'(f_1)^2 + \lambda^2\right) \partial_x^2 f_1 + \mathcal{O}(\varepsilon^2) = 0$$

So the stability condition to be dissipative is

$$\lambda^2 \geq a'(f_1)^2 \quad \Leftrightarrow \quad \lambda \geq |a'(u)|, \quad \forall u \in \mathcal{R}. \tag{2}$$

Whitham's subcharacteristic conditions: Jin–Xin

$$\begin{cases} \partial_t f_1 + \partial_x f_2 = 0 \\ \partial_t f_2 + \lambda^2 \partial_x f_1 = \frac{a(f_1)-f_2}{\varepsilon} \end{cases}$$

$$f_2 = a(f_1) - \varepsilon \left(\partial_t f_2 + \lambda^2 \partial_x f_1 \right)$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(\partial_t f_2 + \lambda^2 \partial_x f_1 \right) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(a'(f_1) \partial_t f_1 + \lambda^2 \partial_x f_1 \right) + \mathcal{O}(\varepsilon^2) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(-a'(f_1) \partial_x f_2 + \lambda^2 \partial_x f_1 \right) + \mathcal{O}(\varepsilon^2) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(-a'(f_1) a'(f_1) \partial_x f_1 + \lambda^2 \partial_x f_1 \right) + \mathcal{O}(\varepsilon^2) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \left(-a'(f_1)^2 + \lambda^2 \right) \partial_x^2 f_1 + \mathcal{O}(\varepsilon^2) = 0$$

So the stability condition to be dissipative is

$$\lambda^2 \geq a'(f_1)^2 \quad \Leftrightarrow \quad \lambda \geq |a'(u)|, \quad \forall u \in \mathcal{R}. \tag{2}$$

Whitham's subcharacteristic conditions: Jin–Xin

$$\begin{cases}\partial_t f_1 + \partial_x f_2 = 0\\ \partial_t f_2 + \lambda^2 \partial_x f_1 = \frac{a(f_1)-f_2}{\varepsilon}\end{cases}$$

$$f_2 = a(f_1) - \varepsilon \left(\partial_t f_2 + \lambda^2 \partial_x f_1\right)$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(\partial_t f_2 + \lambda^2 \partial_x f_1\right) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(a'(f_1)\partial_t f_1 + \lambda^2 \partial_x f_1\right) + \mathcal{O}(\varepsilon^2) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(-a'(f_1)\partial_x f_2 + \lambda^2 \partial_x f_1\right) + \mathcal{O}(\varepsilon^2) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(-a'(f_1)a'(f_1)\partial_x f_1 + \lambda^2 \partial_x f_1\right) + \mathcal{O}(\varepsilon^2) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \left(-a'(f_1)^2 + \lambda^2\right) \partial_x^2 f_1 + \mathcal{O}(\varepsilon^2) = 0$$

So the stability condition to be dissipative is

$$\lambda^2 \geq a'(f_1)^2 \quad \Leftrightarrow \quad \lambda \geq |a'(u)|, \quad \forall u \in \mathcal{R}. \tag{2}$$

Whitham's subcharacteristic conditions: Jin–Xin

$$\begin{cases} \partial_t f_1 + \partial_x f_2 = 0 \\ \partial_t f_2 + \lambda^2 \partial_x f_1 = \frac{a(f_1) - f_2}{\varepsilon} \end{cases}$$

$$f_2 = a(f_1) - \varepsilon \left(\partial_t f_2 + \lambda^2 \partial_x f_1\right)$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(\partial_t f_2 + \lambda^2 \partial_x f_1\right) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(a'(f_1) \partial_t f_1 + \lambda^2 \partial_x f_1\right) + \mathcal{O}(\varepsilon^2) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(-a'(f_1) \partial_x f_2 + \lambda^2 \partial_x f_1\right) + \mathcal{O}(\varepsilon^2) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(-a'(f_1) a'(f_1) \partial_x f_1 + \lambda^2 \partial_x f_1\right) + \mathcal{O}(\varepsilon^2) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \left(-a'(f_1)^2 + \lambda^2\right) \partial_x^2 f_1 + \mathcal{O}(\varepsilon^2) = 0$$

So the stability condition to be dissipative is

$$\lambda^2 \geq a'(f_1)^2 \quad \Leftrightarrow \quad \lambda \geq |a'(u)|, \quad \forall u \in \mathcal{R}. \tag{2}$$

Whitham's subcharacteristic conditions: Jin–Xin

$$\begin{cases} \partial_t f_1 + \partial_x f_2 = 0 \\ \partial_t f_2 + \lambda^2 \partial_x f_1 = \frac{a(f_1)-f_2}{\varepsilon} \end{cases}$$

$$f_2 = a(f_1) - \varepsilon \left(\partial_t f_2 + \lambda^2 \partial_x f_1\right)$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(\partial_t f_2 + \lambda^2 \partial_x f_1\right) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(a'(f_1)\partial_t f_1 + \lambda^2 \partial_x f_1\right) + \mathcal{O}(\varepsilon^2) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(-a'(f_1)\partial_x f_2 + \lambda^2 \partial_x f_1\right) + \mathcal{O}(\varepsilon^2) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \partial_x \left(-a'(f_1)a'(f_1)\partial_x f_1 + \lambda^2 \partial_x f_1\right) + \mathcal{O}(\varepsilon^2) = 0$$

$$\partial_t f_1 + \partial_x a(f_1) - \varepsilon \left(-a'(f_1)^2 + \lambda^2\right) \partial_x^2 f_1 + \mathcal{O}(\varepsilon^2) = 0$$

So the stability condition to be dissipative is

$$\lambda^2 \geq a'(f_1)^2 \quad \Leftrightarrow \quad \lambda \geq |a'(u)|, \quad \forall u \in \mathcal{R}. \tag{2}$$

Chapman–Enskog

Relaxation system

$$f_t^\varepsilon + \sum_{d=1}^D \Lambda_d \partial_{x_d} f^\varepsilon = \frac{1}{\varepsilon}\left(M(Pf^\varepsilon) - f^\varepsilon\right),$$
$$P(M(u)) = u, \quad P\Lambda_d M(u) = A_d(u).$$

Define $u^\varepsilon = Pf^\varepsilon$, $v_d^\varepsilon = P\Lambda_d f^\varepsilon$

$$\begin{cases} \partial_t u^\varepsilon + \sum_{j=1}^D \partial_{x_j} v_j^\varepsilon = 0 \\ \partial_t v_d^\varepsilon + \sum_{j=1}^D \partial_{x_j} (P\Lambda_d \Lambda_j f^\varepsilon) = \frac{1}{\varepsilon}(A_d(u^\varepsilon) - v_d^\varepsilon), \end{cases}$$

Chapman–Enskog

Relaxation system

$$f_t^\varepsilon + \sum_{d=1}^D \Lambda_d \partial_{x_d} f^\varepsilon = \frac{1}{\varepsilon}\left(M(Pf^\varepsilon) - f^\varepsilon\right),$$
$$P(M(u)) = u, \quad P\Lambda_d M(u) = A_d(u).$$

Define $u^\varepsilon = Pf^\varepsilon$, $v_d^\varepsilon = P\Lambda_d f^\varepsilon$

$$\begin{cases} \partial_t u^\varepsilon + \sum_{j=1}^D \partial_{x_j} v_j^\varepsilon = 0 \\ \partial_t v_d^\varepsilon + \sum_{j=1}^D \partial_{x_j}(P\Lambda_d\Lambda_j f^\varepsilon) = \frac{1}{\varepsilon}(A_d(u^\varepsilon) - v_d^\varepsilon), \end{cases}$$

Chapman–Enskog: Taylor expansion in ε

$$v_d^\varepsilon = A_d(u^\varepsilon) - \varepsilon\left(\partial_t v_d^\varepsilon + \sum_{j=1}^D \partial_{x_j}(P\Lambda_d\Lambda_j M(u^\varepsilon))\right) + \mathcal{O}(\varepsilon^2),$$

$$\partial_t u^\varepsilon + \sum_{d=1}^D \partial_{x_d} A_d(u^\varepsilon) = \varepsilon \sum_{d=1}^D \partial_{x_d}\left(\sum_{j=1}^D B_{dj}(u^\varepsilon)\partial_{x_j} u^\varepsilon\right) + \mathcal{O}(\varepsilon^2)$$

with $B_{dj}(u) := P\Lambda_d\Lambda_j M'(u) - A_d'(u)A_j'(u) \in \mathbb{R}^{S\times S}$, $\forall\, d,j = 1,\dots,D$.

Whitham's condition[3]

Chapman Enskog

$$\partial_t u^\varepsilon + \sum_{d=1}^{D} \partial_{x_d} A_d(u^\varepsilon) = \varepsilon \sum_{d=1}^{D} \partial_{x_d} \left(\sum_{j=1}^{D} B_{dj}(u^\varepsilon) \partial_{x_j} u^\varepsilon \right) + \mathcal{O}(\varepsilon^2).$$

Right hand side must be diffusive.

Whitham's subcharacteristic condition

$$B_{jd} := P\Lambda_d \Lambda_j M'(u) - A_d'(u) A_j'(u), \qquad \sum_{j,d=1}^{D} (B_{dj}\xi_j, \xi_d) \geq 0.$$

[3] D. Aregba-Driollet and R. Natalini. Discrete kinetic schemes for multidimensional systems of conservation laws. SIAM J. Numer. Anal., 37(6):1973–2004, 2000.

Kinetic model

Kinetic Model

$$
f^\varepsilon_t + \sum_{d=1}^D \Lambda_d \partial_{x_d} f^\varepsilon = \frac{1}{\varepsilon}\left(M(Pf^\varepsilon) - f^\varepsilon\right),
$$

$$
P(M(u)) = u, \quad P\Lambda_d M(u) = A_d(u).
$$

We have to find M, P, Λ that respect previous conditions.

$$
\begin{tikzcd}
\Omega \subset \mathbb{R}^D \arrow[r, "u"] \arrow[dr, bend right, "f"'] & \mathbb{R}^K \arrow[loop right, "A_d"] \arrow[d, bend right, "M"'] \\
 & \mathbb{R}^L \arrow[u, bend right, "P"'] \arrow[loop right, "\Lambda_d"]
\end{tikzcd}
$$

Model choices

- $L = N \times K$ with $P = (I_K, \dots, I_K)$ N blocks of identity matrices in $\mathbb{R}^K$. (see Lattice Boltzmann) $\Longrightarrow f_n \in \mathbb{R}^K$ with $n = 1, \dots, N$
- $\Lambda_d = diag(\Lambda^{(d)}_1, \dots, \Lambda^{(d)}_N) \qquad \Lambda^{(d)}_n = \lambda^{(d)}_n I_K,$ for $\lambda^{(d)}_n \in \mathbb{R}$.

With this formalism we can rewrite the model as

$$
\begin{cases}
\partial_t f^\varepsilon_n + \sum_{d=1}^D \Lambda^{(d)}_n \partial_{x_d} f^\varepsilon_n = \frac{1}{\varepsilon}\left(M_n(u^\varepsilon) - f^\varepsilon_n\right), & \forall n = 1, \dots, N \\
u^\varepsilon = \sum_{n=1}^N f^\varepsilon_n
\end{cases}
. \tag{3}
$$

Kinetic model – DRM

Diagonal Relaxation Method (DRM)

- $N = D + 1$
- Take $\lambda > 0$ and
$$\Lambda_j^{(d)} = \begin{cases} -\lambda I_K & j = d \\ \lambda I_K & j = D+1 \\ 0 & \text{else} \end{cases}.$$
- $$\begin{cases} M_{D+1}(u) = \left(u + \frac{1}{\lambda}\sum_{d=1}^{D} A_d(u)\right)/(D+1) \\ M_j(u) = -\frac{1}{\lambda}A_j(u) + M_{D+1}(u) \end{cases}$$

Important: we have to choose λ according to Whitham's subcharacteristic condition.

Example of DMR model

$u : \Omega \subset \mathbb{R} \to \mathbb{R}, \quad D = 1, N = 2, \quad f : \mathbb{R} \to \mathbb{R}^2$

Limit equation

$$u_t + a(u)_x = 0 \tag{4}$$

$$\Lambda = \begin{pmatrix} -\lambda & 0 \\ 0 & \lambda \end{pmatrix}, \quad M(u) = \begin{pmatrix} \frac{u}{2} - \frac{a(u)}{2\lambda} \\ \frac{u}{2} + \frac{a(u)}{2\lambda} \end{pmatrix}, \quad Pf = f_1 + f_2 \tag{5}$$

Kinetic model is

$$\begin{cases} \partial_t f_1 - \lambda \partial_x f_1 = \frac{1}{\epsilon}\left(\frac{f_1+f_2}{2} - \frac{a(f_1+f_2)}{2\lambda} - f_1 \right) \\ \partial_t f_2 + \lambda \partial_x f_2 = \frac{1}{\epsilon}\left(\frac{f_1+f_2}{2} + \frac{a(f_1+f_2)}{2\lambda} - f_2 \right) \end{cases} \tag{6}$$

Outline

IMEX discretization - Kinetic model

Stiff source term $\Rightarrow$ unstable for $\varepsilon \ll \Delta t \Rightarrow$ IMEX approach:
IMplicit for stiff source term, EXplicit for advection term and bathymetry source

$$\frac{f^{n+1,\varepsilon} - f^{n,\varepsilon}}{\Delta t} + \Lambda \partial_x f^{n,\varepsilon} = \frac{1}{\varepsilon}\left(M(Pf^{n+1,\varepsilon}) - f^{n+1,\varepsilon} \right). \tag{7}$$

IMEX discretization - Kinetic model

Stiff source term $\Rightarrow$ unstable for $\varepsilon \ll \Delta t \Rightarrow$ IMEX approach:
IMplicit for stiff source term, EXplicit for advection term and bathymetry source

$$\frac{f^{n+1,\varepsilon} - f^{n,\varepsilon}}{\Delta t} + \Lambda \partial_x f^{n,\varepsilon} = \frac{1}{\varepsilon}\left(M(Pf^{n+1,\varepsilon}) - f^{n+1,\varepsilon} \right). \tag{7}$$

How to treat non-linear implicit functions?
Recall: $PM(u) = u$ and $Pf^{\varepsilon} = u^{\varepsilon}$, so

$$\frac{u^{n+1,\varepsilon} - u^{n,\varepsilon}}{\Delta t} + P\Lambda \partial_x f^{n,\varepsilon} = 0. \tag{8}$$

Find $u^{n+1,\varepsilon} = Pf^{n+1,\varepsilon}$ and substitute it in the Maxwellian in (7).

IMEX discretization - Kinetic model

Stiff source term $\Rightarrow$ unstable for $\varepsilon \ll \Delta t \Rightarrow$ IMEX approach:
IMplicit for stiff source term, EXplicit for advection term and bathymetry source

$$\frac{f^{n+1,\varepsilon}-f^{n,\varepsilon}}{\Delta t}+\Lambda\partial_x f^{n,\varepsilon}=\frac{1}{\varepsilon}\left(M(Pf^{n+1,\varepsilon})-f^{n+1,\varepsilon}\right). \tag{7}$$

How to treat non-linear implicit functions?
Recall: $PM(u)=u$ and $Pf^\varepsilon=u^\varepsilon$, so

$$\frac{u^{n+1,\varepsilon}-u^{n,\varepsilon}}{\Delta t}+P\Lambda\partial_x f^{n,\varepsilon}=0. \tag{8}$$

Find $u^{n+1,\varepsilon}=Pf^{n+1,\varepsilon}$ and substitute it in the Maxwellian in (7).

IMEX discretization - Kinetic model

Stiff source term $\Rightarrow$ unstable for $\varepsilon \ll \Delta t \Rightarrow$ IMEX approach:
IMplicit for stiff source term, EXplicit for advection term and bathymetry source

$$\frac{f^{n+1,\varepsilon} - f^{n,\varepsilon}}{\Delta t} + \Lambda \partial_x f^{n,\varepsilon} = \frac{1}{\varepsilon}\left(M(Pf^{n+1,\varepsilon}) - f^{n+1,\varepsilon}\right). \tag{7}$$

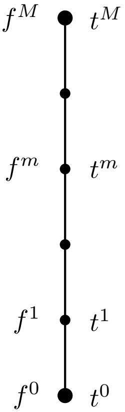
How to treat non-linear implicit functions?
Recall: $PM(u) = u$ and $Pf^\varepsilon = u^\varepsilon$, so

$$\frac{u^{n+1,\varepsilon} - u^{n,\varepsilon}}{\Delta t} + P\Lambda \partial_x f^{n,\varepsilon} = 0. \tag{8}$$

Find $u^{n+1,\varepsilon} = Pf^{n+1,\varepsilon}$ and substitute it in the Maxwellian in (7).

- IMEX formulation is first order accurate =: $\mathcal{L}^1$
- IMEX formulation is asymptotic preserving (AP) (as $\varepsilon \to 0$ we recast limit hyperbolic)

Deferred Correction[4]



[4] A. Dutt, L. Greengard, and V. Rokhlin. BIT Numerical Mathematics, 40(2):241–266, 2000.

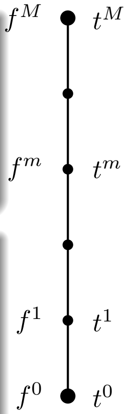
Deferred Correction[4]

$\mathcal{L}^1(f) = 0$

- IMEX
- First order accurate
- Mass lumping
- Computationally explicit

$\mathcal{L}^2(f) = 0$

- Order $M + 1$
- Quadrature in timestep
- Nonlinearly implicit
- Implicit Runge–Kutta



[4] A. Dutt, L. Greengard, and V. Rokhlin. BIT Numerical Mathematics, 40(2):241–266, 2000.

Deferred Correction[4]

How to combine two methods keeping the accuracy of the second and the stability and simplicity of the first one?

$$f^{0,(k)} := f(t^n), \quad k = 0, \dots, K,$$

$$f^{m,(0)} := f(t^n), \quad m = 1, \dots, M$$

$$\mathcal{L}^1(\underline{f}^{(k)}) = \mathcal{L}^1(\underline{f}^{(k-1)}) - \mathcal{L}^2(\underline{f}^{(k-1)}), k \leq K.$$

DeC Theorem

- $\mathcal{L}^1$ coercive
- $\mathcal{L}^1 - \mathcal{L}^2$ Lipschitz

DeC order accuracy $\min(K, M+1)$.

AP Theorem

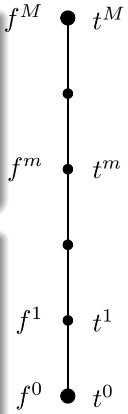
$\mathcal{L}^1$ AP $\Longrightarrow$ DeC AP

$\mathcal{L}^1(f) = 0$

- IMEX
- First order accurate
- Mass lumping
- Computationally explicit

$\mathcal{L}^2(f) = 0$

- Order $M + 1$
- Quadrature in timestep
- Nonlinearly implicit
- Implicit Runge–Kutta



[4] A. Dutt, L. Greengard, and V. Rokhlin. BIT Numerical Mathematics, 40(2):241–266, 2000.

Deferred Correction[5]

How to combine two methods keeping the accuracy of the second and the stability and simplicity of the first one?

$$f^{0,(k)} := f(t^n), \quad k = 0, \dots, K,$$
$$f^{m,(0)} := f(t^n), \quad m = 1, \dots, M$$
$$\mathcal{L}^1(\underline{f}^{(k)}) = \mathcal{L}^1(\underline{f}^{(k-1)}) - \mathcal{L}^2(\underline{f}^{(k-1)}), k \leq K.$$

DeC Theorem

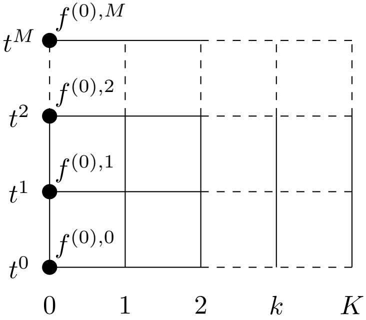
- $\mathcal{L}^1$ coercive
- $\mathcal{L}^1 - \mathcal{L}^2$ Lipschitz

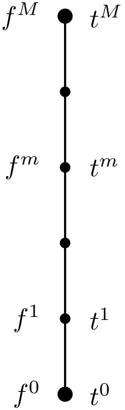
DeC order accuracy $\min(K, M+1)$.

AP Theorem[4]

$\mathcal{L}^1$ AP $\Longrightarrow$ DeC AP

- $\mathcal{L}^1(\underline{f}) = 0$, first order accuracy, easily invertible.
- $\mathcal{L}^2(\underline{f}) = 0$, high order $M+1$.





[5] R. Abgrall, and D.T.. SIAM Journal on Scientific Computing, 42(3):B816–B845, 2020.

Deferred Correction[5]

How to combine two methods keeping the accuracy of the second and the stability and simplicity of the first one?

$$f^{0,(k)} := f(t^n), \quad k = 0, \dots, K,$$
$$f^{m,(0)} := f(t^n), \quad m = 1, \dots, M$$
$$\mathcal{L}^1(\underline{f}^{(k)}) = \mathcal{L}^1(\underline{f}^{(k-1)}) - \mathcal{L}^2(\underline{f}^{(k-1)}), k \leq K.$$

DeC Theorem

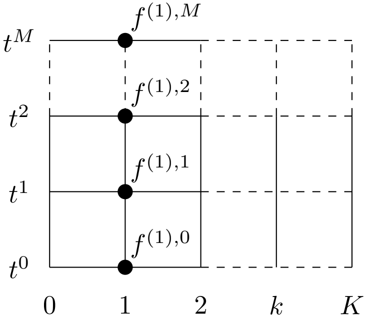
- $\mathcal{L}^1$ coercive
- $\mathcal{L}^1 - \mathcal{L}^2$ Lipschitz

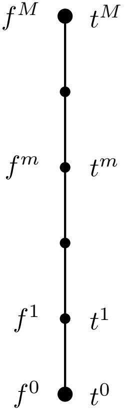
DeC order accuracy $\min(K, M+1)$.

AP Theorem[4]

$\mathcal{L}^1$ AP $\Longrightarrow$ DeC AP

- $\mathcal{L}^1(\underline{f}) = 0$, first order accuracy, easily invertible.
- $\mathcal{L}^2(\underline{f}) = 0$, high order $M+1$.





[5] R. Abgrall, and D.T.. SIAM Journal on Scientific Computing, 42(3):B816–B845, 2020.

Deferred Correction[5]

How to combine two methods keeping the accuracy of the second and the stability and simplicity of the first one?

$$f^{0,(k)} := f(t^n), \quad k = 0, \dots, K,$$
$$f^{m,(0)} := f(t^n), \quad m = 1, \dots, M$$
$$\mathcal{L}^1(\underline{f}^{(k)}) = \mathcal{L}^1(\underline{f}^{(k-1)}) - \mathcal{L}^2(\underline{f}^{(k-1)}), k \le K.$$

DeC Theorem

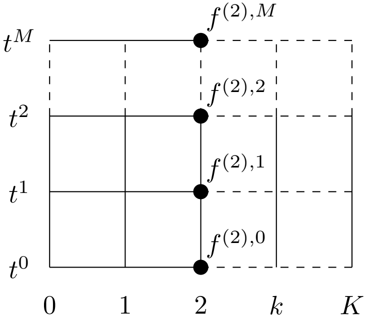
- $\mathcal{L}^1$ coercive
- $\mathcal{L}^1 - \mathcal{L}^2$ Lipschitz

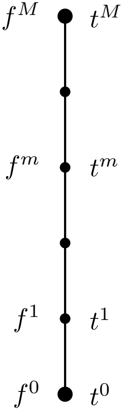
DeC order accuracy $\min(K, M+1)$.

AP Theorem[4]

$\mathcal{L}^1$ AP $\Longrightarrow$ DeC AP

- $\mathcal{L}^1(\underline{f}) = 0$, first order accuracy, easily invertible.
- $\mathcal{L}^2(\underline{f}) = 0$, high order $M+1$.





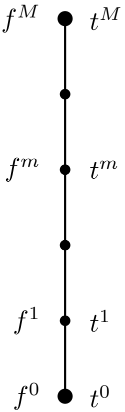
[5]R. Abgrall, and D.T.. SIAM Journal on Scientific Computing, 42(3):B816–B845, 2020.

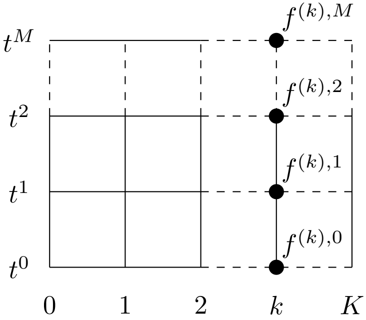
Deferred Correction[5]

How to combine two methods keeping the accuracy of the second and the stability and simplicity of the first one?

$$f^{0,(k)} := f(t^n), \quad k = 0, \dots, K,$$
$$f^{m,(0)} := f(t^n), \quad m = 1, \dots, M$$
$$\mathcal{L}^1(\underline{f}^{(k)}) = \mathcal{L}^1(\underline{f}^{(k-1)}) - \mathcal{L}^2(\underline{f}^{(k-1)}), k \le K.$$

- $\mathcal{L}^1(\underline{f}) = 0$, first order accuracy, easily invertible.
- $\mathcal{L}^2(\underline{f}) = 0$, high order $M+1$.





DeC Theorem

- $\mathcal{L}^1$ coercive
- $\mathcal{L}^1 - \mathcal{L}^2$ Lipschitz

DeC order accuracy $\min(K, M+1)$.

AP Theorem[4]

$\mathcal{L}^1$ AP $\Longrightarrow$ DeC AP

[5] R. Abgrall, and D.T.. SIAM Journal on Scientific Computing, 42(3):B816–B845, 2020.

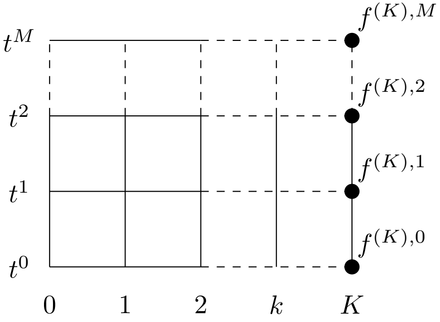
How to combine two methods keeping the accuracy of the second and the stability and simplicity of the first one?

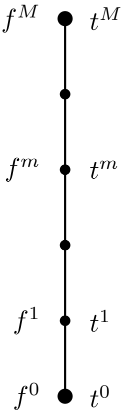
$$f^{0,(k)} := f(t^n), \quad k = 0, \dots, K,$$

$$f^{m,(0)} := f(t^n), \quad m = 1, \dots, M$$

$$\mathcal{L}^1(\underline{f}^{(k)}) = \mathcal{L}^1(\underline{f}^{(k-1)}) - \mathcal{L}^2(\underline{f}^{(k-1)}), k \leq K.$$

- $\mathcal{L}^1(\underline{f}) = 0$, first order accuracy, easily invertible.
- $\mathcal{L}^2(\underline{f}) = 0$, high order $M+1$.





DeC Theorem

- $\mathcal{L}^1$ coercive
- $\mathcal{L}^1 - \mathcal{L}^2$ Lipschitz

DeC order accuracy $\min(K, M+1)$.

AP Theorem[4]

$\mathcal{L}^1$ AP $\Longrightarrow$ DeC AP

[5] R. Abgrall, and D.T.. SIAM Journal on Scientific Computing, 42(3):B816–B845, 2020.

RK vs DeC

Explicit DeC can be rewritten into Explicit Runge Kutta stages with $(r-1)^2+1$ stages (with a correction due to the lumping of the mass matrix)

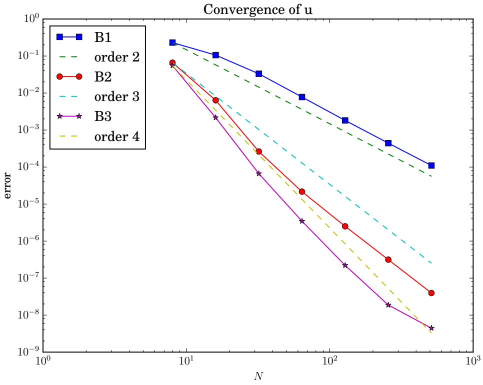
	Runge Kutta	Deferred Correction
Coefficients	Specific $\forall$ order	General algorithm
Stages	$r \leq s < r^2$	$s=(r-1)^2+1$ or $(r-1\|\|r)$
Mass matrix	Full	Lumped

Outline

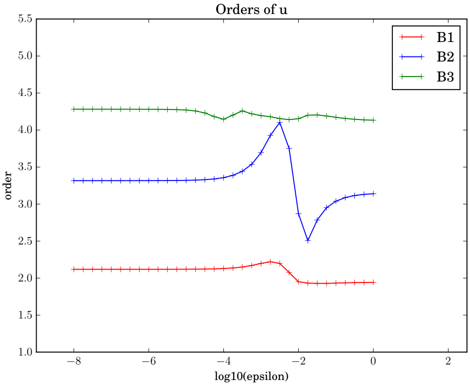
Numerical tests: Linear advection for convergence

$$u_t + u_x = 0, \quad x \in [0,1], \quad t \in [0,T], T = 0.12, \quad u_0(x) = e^{-80(x-0.4)^2},$$

outflow BC, $\lambda = 1.5$, $\varepsilon = 10^{-10}$, $\theta_1 = 1$, $\theta_2 = 5$ (derivative stabilization).



(a) Scalar 1D convergence



(b) Order varying relaxation parameter

Numerical tests: Euler equation

Next simulations will be over the Euler equation

$$\begin{pmatrix} \rho \\ \rho v \\ E \end{pmatrix}_t + \begin{pmatrix} \rho v \\ \rho v^2 + p \\ (E+p)v \end{pmatrix}_x = 0, \qquad x \in [0,1],\ t \in [0,T] \tag{9}$$

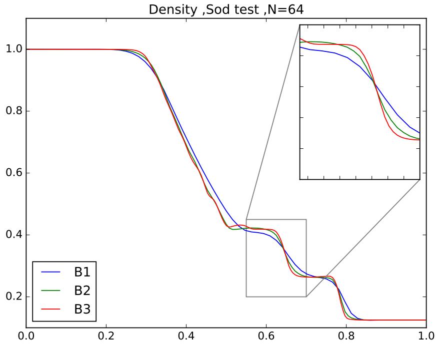
ρ is the density, v the speed, p the pressure and E the total energy. The system is closed by the equation of state

$$E = \frac{p}{\gamma - 1} + \frac{1}{2}\rho v^2. \tag{10}$$

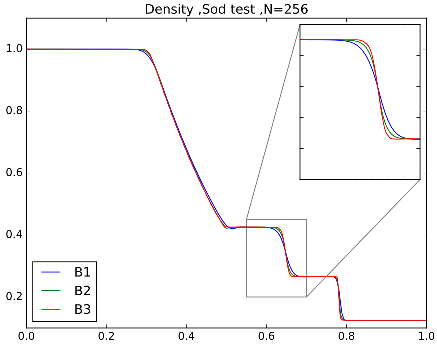
Numerical tests: Sod shock test

$\gamma = 1.4$, $T = 0.16$, outflow BC, $\varepsilon = 10^{-9}$, $\lambda = 2$, CFL = 0.2.
For $\mathbb{B}^1$ $\theta_1 = 1$, for $\mathbb{B}^2$ $\theta_1 = 1$, $\theta_2 = 0.5$, for $\mathbb{B}^3$ $\theta_1 = 2.5$, $\theta_2 = 4$.

$$\rho_0 = \mathbb{1}_{[0,0.5]}(x) + 0.1\mathbb{1}_{[0.5,1]}(x), \quad v_0 = 0, \quad p_0 = \mathbb{1}_{[0,0.5]}(x) + 0.125\mathbb{1}_{[0.5,1]}(x).$$



(a) $N = 64$



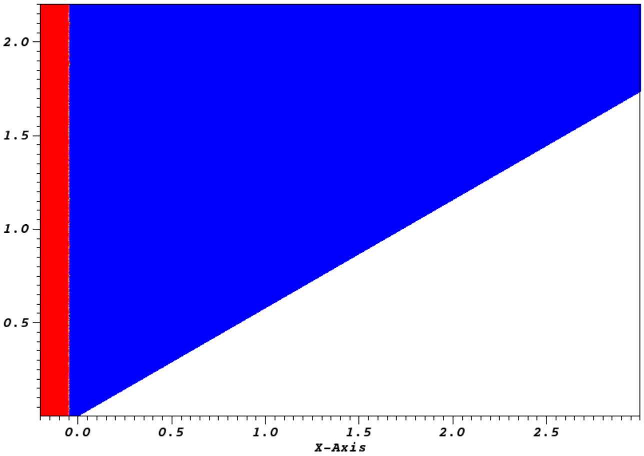
(b) $N = 256$

Double mach reflection test: initial conditions

$$\begin{pmatrix} \rho_0 \\ u_0 \\ v_0 \\ p_0 \end{pmatrix} = \begin{pmatrix} 8 \\ 8.25 \\ 0 \\ 116.5 \end{pmatrix} \text{ if } x \leq -0.05$$

$$\begin{pmatrix} \rho_0 \\ u_0 \\ v_0 \\ p_0 \end{pmatrix} = \begin{pmatrix} 1.4 \\ 0 \\ 0 \\ 1 \end{pmatrix} \text{ if } x > -0.05.$$

$T = 0.2$, $\varepsilon = 10^{-9}$, $\lambda = 15$, CFL = 0.1, $N = 19248$ triangular elements.
For $\mathbb{B}^1$ $\theta_1 = 0.1$, for $\mathbb{B}^2$ $\theta_1 = 0.01$, $\theta_2 = 0.0001$, for $\mathbb{B}^3$ $\theta_1 = 0.005$, $\theta_2 = 0.0001$.



Numerical tests 2D: DMR test

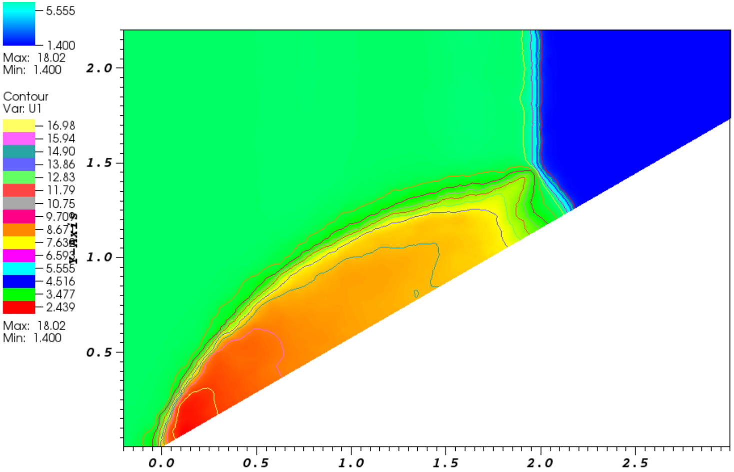


Figure: Density of DMR test $\mathbb{B}^1$

Numerical tests 2D: DMR test

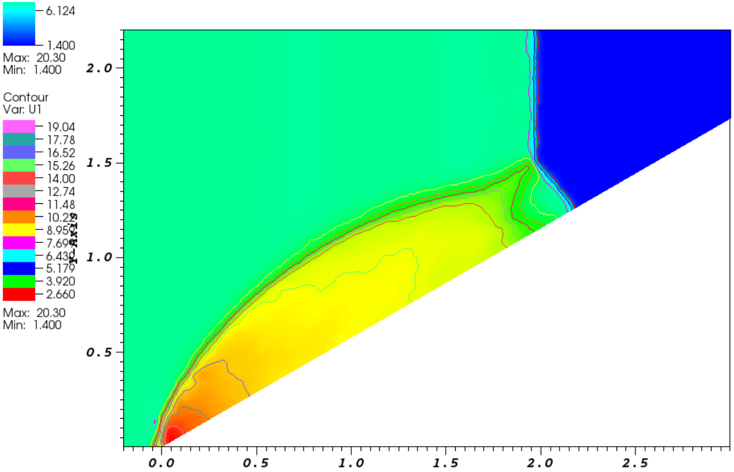


Figure: Density of DMR test $\mathbb{B}^2$

Numerical tests 2D: DMR test

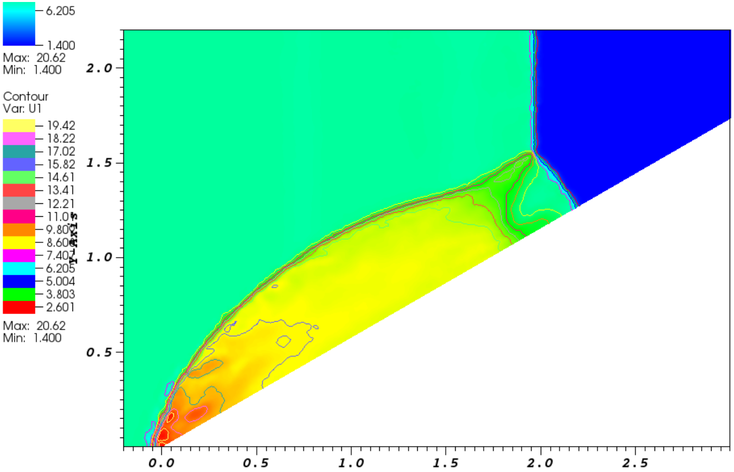


Figure: Density of DMR test $\mathbb{B}^3$

Outline

Conclusion and perspective

Summary

- Residual distribution: framework to define many method
- Residual distribution: connection to finite volume
- Kinetic Models for hyperbolic problems
- Chapman–Enskog: diffusion with Whitham's subcharacteristic condition
- Discretization:
  - IMEX in time
  - Residual distribution
  - Deferred Correction

Perspective

- MOOD
- Entropy stability
- More robust RD choices

Similar projects

- Kinetic models for SW with Source terms
- Well balancing

IMEX DeC RD – Bibliography

Residual Distribution

- R. Abgrall. Some remarks about conservation for residual distribution schemes. Computational Methods in Applied Mathematics, 2018.
- R. Abgrall, and M. Ricchiuto. "Hyperbolic balance laws: residual distribution, local and global fluxes." Numerical Fluid Dynamics: Methods and Computations (2022): 177-222.

Deferred Correction

- A. Dutt, L. Greengard, and V. Rokhlin. Spectral Deferred Correction Methods for Ordinary Differential Equations. BIT Numerical Mathematics, 40(2):241–266, 2000.
- R. Abgrall. High Order Schemes for Hyperbolic Problems Using Globally Continuous Approximation and Avoiding Mass Matrices. Journal of Scientific Computing, 73(2):461–494, 2017.

Kinetic Model

- D. Aregba-Driollet, and R. Natalini. Discrete kinetic schemes for multidimensional systems of conservation laws. SIAM J. Numer. Anal., 37(6):1973–2004, 2000.
- R. Abgrall, and D. Torlo. High Order Asymptotic Preserving Deferred Correction Implicit-Explicit Schemes for Kinetic Models. SIAM Journal on Scientific Computing, 42(3):B816–B845, 2020.

Thank you for the attention!

Shallow water equations

Modify the kinetic relaxation models by D. Aregba-Driollet and R. Natalini
Hyperbolic limit equation is

$$
\begin{tikzcd}
\Omega \subset \mathbb{R}^D \arrow[r, "u"] \arrow[dr, bend right, "f"'] & \mathbb{R}^K \arrow[loop right, "{A_d, \color{red}{S}}"] \arrow[d, bend right, "M"'] \\
& \mathbb{R}^L \arrow[u, bend right, "P"'] \arrow[loop right, "\Lambda_d"]
\end{tikzcd}
$$

$$u_t + \sum_{d=1}^{D} \partial_{x_d} A_d(u) \color{red}{+ S(u)} = 0, \quad u : \Omega \to \mathbb{R}^K$$

$$\begin{cases} h_t + (hv)_x = 0 \\ (hv)_t + (hv^2 + \frac{g}{2}h^2)_x \color{red}{+ ghb_x} = 0 \end{cases}$$

Relaxation system

$$f_t^\varepsilon + \sum_{d=1}^{D} \Lambda_d \partial_{x_d} f^\varepsilon = \frac{1}{\varepsilon}\left(M(Pf^\varepsilon) - f^\varepsilon\right), \quad f^\varepsilon : \Omega \to \mathbb{R}^L$$

$$Pf^\varepsilon \to u, \quad P(M(u)) = u, \quad P\Lambda_d M(u) = A_d(u).$$

Shallow water equations

Modify the kinetic relaxation models by D. Aregba-Driollet and R. Natalini
Hyperbolic limit equation is

$$u_t + \sum_{d=1}^{D} \partial_{x_d} A_d(u) + S(u) = 0, \quad u : \Omega \to \mathbb{R}^K$$

$$\begin{cases} h_t + (hv)_x = 0 \\ (hv)_t + (hv^2 + \frac{g}{2}(h^2 - b^2))_x + g(h+b)b_x = 0 \end{cases}$$

$$\begin{tikzcd} \Omega \subset \mathbb{R}^D \arrow[r, "u"] \arrow[dr, bend right, "f"'] & \mathbb{R}^K \arrow[d, bend right, "M"'] \arrow[loop right, "{A_d, S}"] \\ & \mathbb{R}^L \arrow[u, bend right, "P"'] \arrow[loop right, "\Lambda_d"] \end{tikzcd}$$

Relaxation system

$$f_t^\varepsilon + \sum_{d=1}^{D} \Lambda_d \partial_{x_d} f^\varepsilon = \frac{1}{\varepsilon} \left(M(Pf^\varepsilon) - f^\varepsilon \right), \quad f^\varepsilon : \Omega \to \mathbb{R}^L$$

$$Pf^\varepsilon \to u, \quad P(M(u)) = u, \quad P\Lambda_d M(u) = A_d(u).$$

Shallow water equations

Modify the kinetic relaxation models by D. Aregba-Driollet and R. Natalini
Hyperbolic limit equation is

$$
\begin{tikzcd}
\Omega \subset \mathbb{R}^D \arrow[r, "u"] \arrow[dr, bend right, "f"'] & \mathbb{R}^K \arrow[loop right, "{A_d, S}"] \arrow[d, bend right, "M"'] \\
 & \mathbb{R}^L \arrow[loop right, "{\Lambda_d, \tilde{S}}"] \arrow[u, bend right, "P"']
\end{tikzcd}
$$

$$u_t + \sum_{d=1}^{D} \partial_{x_d} A_d(u) + S(u) = 0, \quad u : \Omega \to \mathbb{R}^K$$

$$\begin{cases} h_t + (hv)_x = 0 \\ (hv)_t + (hv^2 + \frac{g}{2}(h^2 - b^2))_x + g(h+b)b_x = 0 \end{cases}$$

Relaxation system

$$f_t^\varepsilon + \sum_{d=1}^{D} \Lambda_d \partial_{x_d} f^\varepsilon + \tilde{S}(f) = \frac{1}{\varepsilon}\left(M(Pf^\varepsilon) - f^\varepsilon\right), \quad f^\varepsilon : \Omega \to \mathbb{R}^L, \quad \tilde{S}(f) := \begin{pmatrix} S(f_1) \\ \cdots \\ S(f_N) \end{pmatrix},$$

$$Pf^\varepsilon \to u, \quad P(M(u)) = u, \quad P\Lambda_d M(u) = A_d(u), \quad P\tilde{S}(f) = S(Pf), \quad P\Lambda_d \tilde{S}(f) = S(P\Lambda f).$$

Other properties

- Asymptotic preserving: Chapman–Enskog
- Well balancedness: lake at rest steady state preservation
  - Choice of a different form of the SW equation, so that the discretizations of the flux and the source match when $v = 0$
- Depth non-negativity
  - Wet and dry elements
  - Hybrid elements -> Modify the bathymetry to have positive DoFs

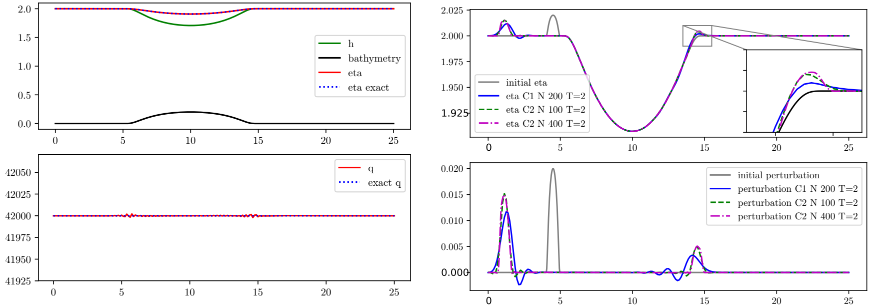


Figure: Subcritical flow: simulation $\mathbb{C}^2$, $N = 100$ (left) perturbation propagating (right)

$$b(x) = \begin{cases} 0.2 \exp\left(\frac{((x-10)/5)^2}{1-((x-10)/5)^2}\right), & \text{if } x \in B_5(10), \\ 0, & \text{else.} \end{cases}$$

$$h^\varepsilon(0,x) = 2 - b(x) \qquad q^\varepsilon(0,t) = 4.42$$

$$q^\varepsilon(0,x) = 4.42 \qquad h^\varepsilon(25,t) = 2$$

$$f^\varepsilon(0,x) = M(u^\varepsilon(0,x)) \qquad T = 100$$

$\lambda = 6.5, \quad \varepsilon = 10^{-14},$

Simulations: Subcritical Flow

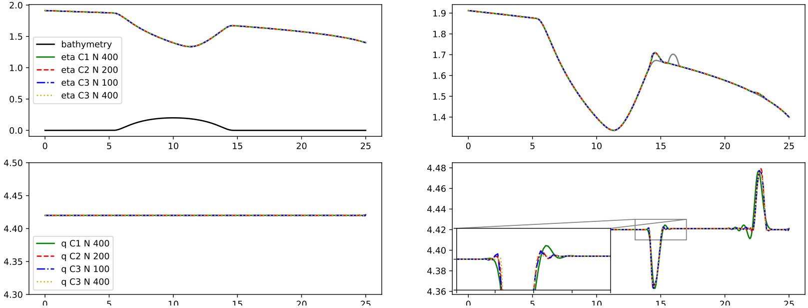


Figure: Subcritical flow: simulation with friction $\mathbb{C}^2$, $N = 100$ (left) perturbation propagating (right)

$$b(x) = \begin{cases} 0.2\exp\left(\frac{((x-10)/5)^2}{1-((x-10)/5)^2}\right), & \text{if } x \in B_5(10), \\ 0, & \text{else.} \end{cases}$$

$$h^\varepsilon(0,x) = 2 - b(x) \qquad q^\varepsilon(0,t) = 4.42$$

$$q^\varepsilon(0,x) = 4.42 \qquad h^\varepsilon(25,t) = 2$$

$$\lambda = 6.5, \quad \varepsilon = 10^{-14}, \qquad f^\varepsilon(0,x) = M(u^\varepsilon(0,x)) \qquad T = 100$$

Simulations: Subcritical Flow

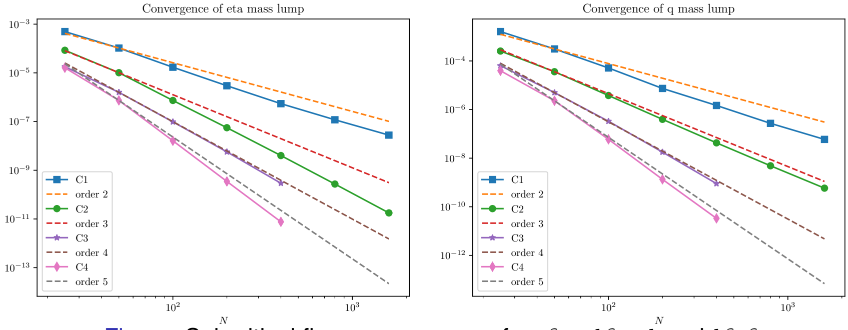


Figure: Subcritical flow: convergence for $\eta^\varepsilon = h^\varepsilon + b$ and $h^\varepsilon v^\varepsilon$

$$b(x) = \begin{cases} 0.2 \exp\left(\frac{((x-10)/5)^2}{1-((x-10)/5)^2} \right), & \text{if } x \in B_5(10), \\ 0, & \text{else.} \end{cases}$$

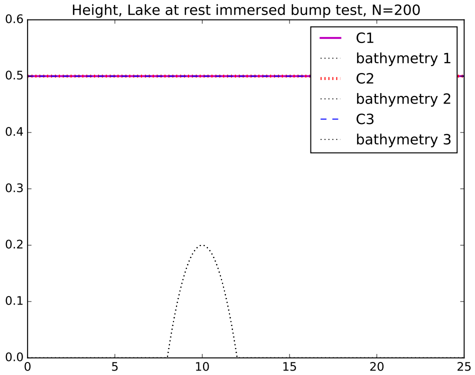
$$\lambda = 6.5, \quad \varepsilon = 10^{-14},$$

$$h^\varepsilon(0,x) = 2 - b(x) \qquad q^\varepsilon(0,t) = 4.42$$

$$q^\varepsilon(0,x) = 4.42 \qquad h^\varepsilon(25,t) = 2$$

$$f^\varepsilon(0,x) = M(u^\varepsilon(0,x)) \qquad T = 100$$

Simulations: lake at rest



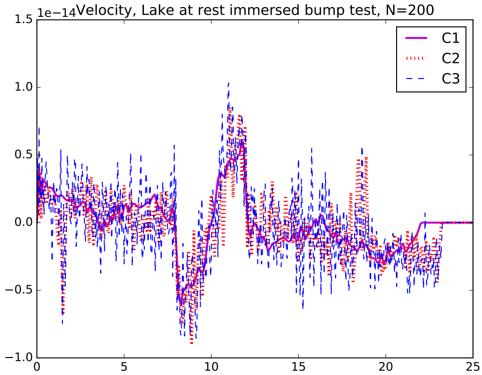


Figure: Lake at rest with immersed bump test: η^ε and v^ϵ with $N = 200$

$$b(x) = (0.2 - 0.05(x-10)^2)\mathbb{1}_{\{8<x<12\}} \qquad q^\varepsilon(0,t) = 0 \qquad q - q^{ex} = \mathcal{O}(N_t \varepsilon)$$

$$\eta^\varepsilon(0,x) = 0.5 \qquad q^\varepsilon(25,t) = 0 \qquad T = 3$$

$$q^\varepsilon(0,x) = 0 \qquad \lambda = 2 \qquad \varepsilon = 10^{-14}$$

Simulations: lake at rest

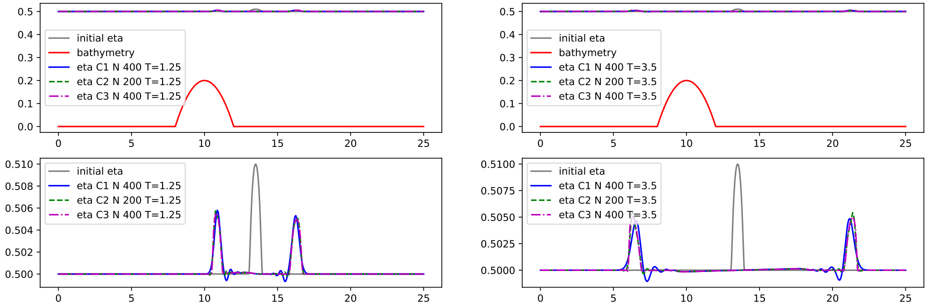


Figure: Lake at rest with immersed bump perturbed: η^ε and v^ϵ with $N = 200$

$$b(x) = (0.2 - 0.05(x-10)^2)\mathbb{1}_{\{8<x<12\}} \qquad q^\varepsilon(0,t) = 0 \qquad q - q^{ex} = \mathcal{O}(N_t \varepsilon)$$

$$\eta^\varepsilon(0,x) = 0.5 \qquad q^\varepsilon(25,t) = 0 \qquad T = 3$$

$$q^\varepsilon(0,x) = 0 \qquad \lambda = 2 \qquad \varepsilon = 10^{-14}$$

Simulations: wet and dry lake at rest

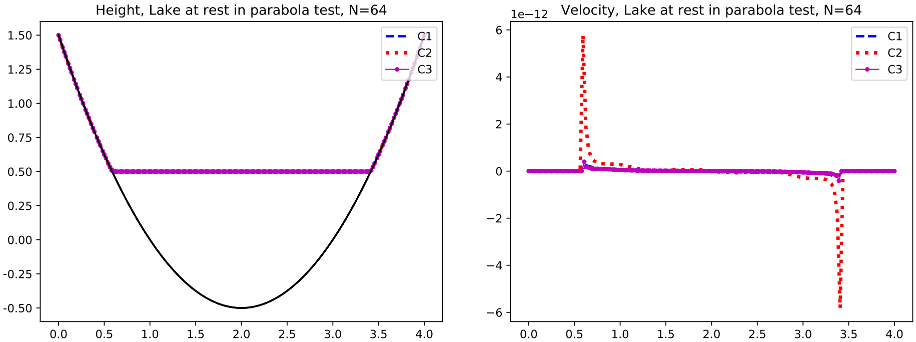


Figure: Lake at rest in parabola test: η^ε and v^ε with $N = 64$

$$b(x) = (x-2)^2 - 0.5 \qquad q - q^{ex} = \mathcal{O}(N_t \varepsilon)$$

$$\eta^\varepsilon(0,x) = \max(0.5, b(x)) \qquad T = 3$$

$$\lambda = 4 \qquad \varepsilon = 10^{-14}$$

Simulations: Thacker's Oscillations

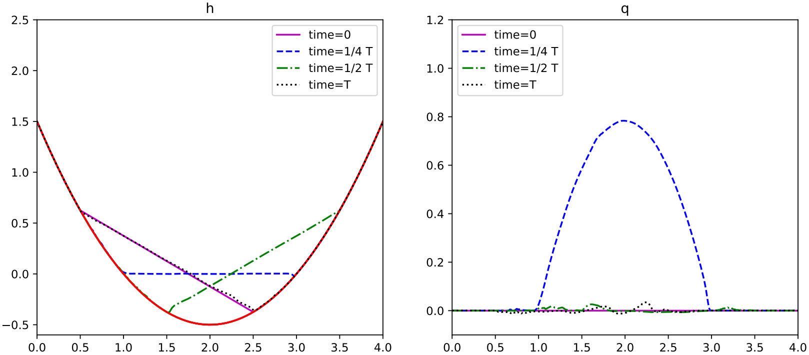


Figure: Thacker oscillations in parabola test: η^ε and $h^\varepsilon v^\varepsilon$ with $\mathbb{C}^1$ and $N = 300$

$$b(x) = (x-2)^2 - 0.5$$

$$\eta^\varepsilon(0,x) = \max(-0.5x + 0.875, b(x))$$

$$\lambda = 6.5$$

period $= 2.0606$

$$T = 5 \cdot 2.0606$$

$$\varepsilon = 10^{-14}$$

Simulations: Thacker's Oscillations

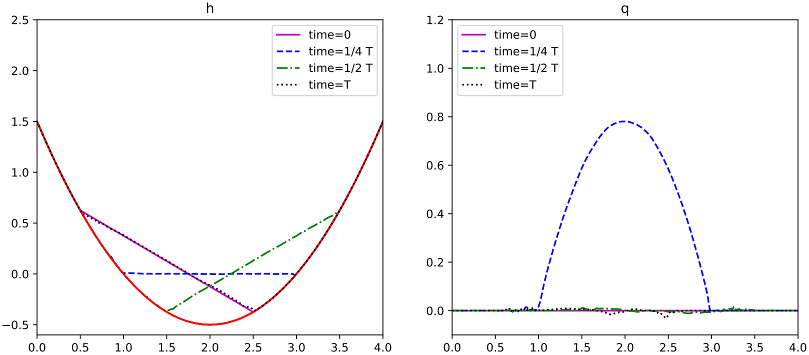


Figure: Thacker oscillations in parabola test: η^ε and $h^\varepsilon v^\varepsilon$ with $\mathbb{C}^2$ and $N = 300$

$$b(x) = (x-2)^2 - 0.5 \qquad \text{period} = 2.0606$$

$$\eta^\varepsilon(0,x) = \max(-0.5x + 0.875, b(x)) \qquad T = 5 \cdot 2.0606$$

$$\lambda = 6.5 \qquad \varepsilon = 10^{-14}$$

Simulations: Thacker's Oscillations

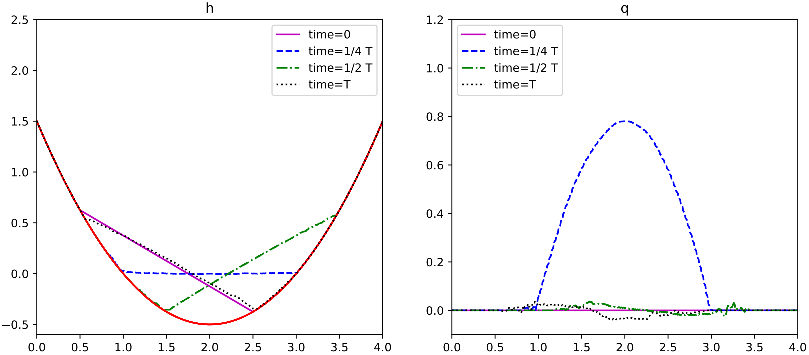


Figure: Thacker oscillations in parabola test: η^ε and $h^\varepsilon v^\varepsilon$ with $\mathbb{C}^3$ and $N = 150$

$$b(x) = (x-2)^2 - 0.5$$

$$\eta^\varepsilon(0,x) = \max(-0.5x + 0.875, b(x))$$

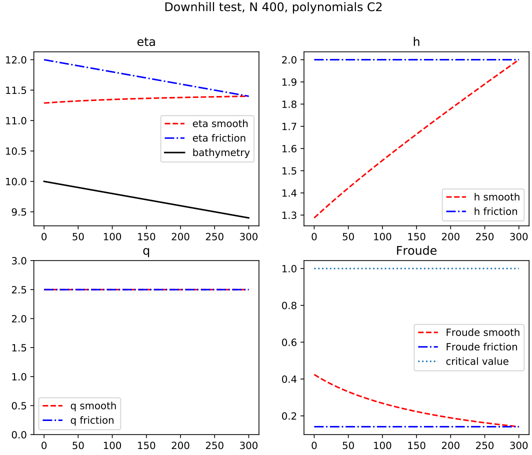
$$\lambda = 6.5$$

period $= 2.0606$

$$T = 5 \cdot 2.0606$$

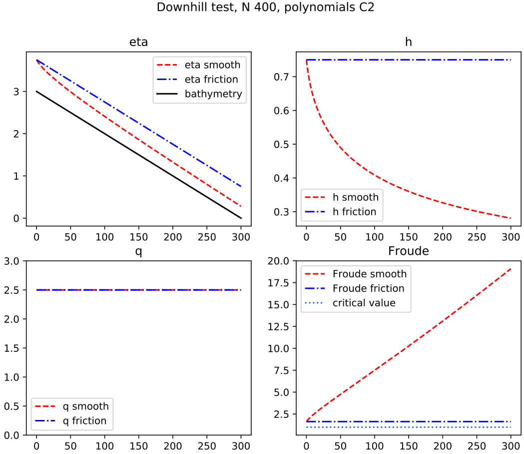
$$\varepsilon = 10^{-14}$$

Inclined plane with friction



$$\partial_x b(x) \equiv 0.002$$
$$\eta^\varepsilon(0,x) = 2$$
$$q^\varepsilon(0,x) = 2.75$$
$$q^\varepsilon(0,t) = 2.5$$
$$h^\varepsilon(300,t) = 2$$
$$\lambda = 22$$
$$n = 0.2515597$$
$$T = 1000$$
$$\varepsilon = 10^{-14}$$

Figure: Subcritical flow: η^ε and $h^\varepsilon v^\varepsilon$ with $\mathbb{C}^2$ and $N = 400$



$$\partial_x b(x) \equiv 0.01$$
$$\eta^\varepsilon(0,x) = 0.75$$
$$q^\varepsilon(0,x) = 2.75$$
$$h^\varepsilon(0,t) = 0.75$$
$$q^\varepsilon(0,t) = 2.5$$
$$\lambda = 22$$
$$n = 0.067820251$$
$$T = 1000$$
$$\varepsilon = 10^{-14}$$

Figure: Supercritical flow: η^ε and $h^\varepsilon v^\varepsilon$ with $\mathbb{C}^2$ and $N = 400$

Relaxation as Diffusion

Relaxation as diffusion

- Smooth problems
- Set $\varepsilon \approx \Delta t^{p+1}$
- Relaxation term is diffusive in Chapman–Eskong
- Pure Galerkin discretization

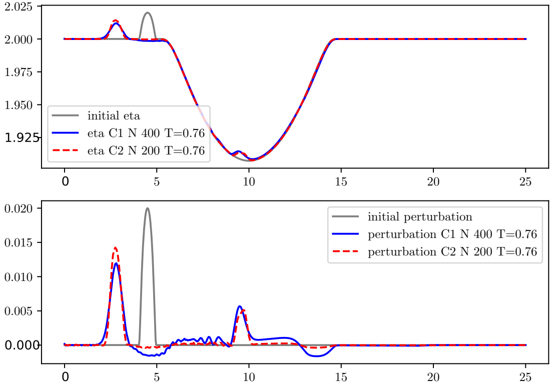


Figure: Subcritical flow $\mathbb{C}^1$ N=400, $\mathbb{C}^2$ N=200. η^ε and $h_p^\varepsilon - h^\varepsilon$ $T = 0.76$.

Relaxation as Diffusion

Relaxation as diffusion

- Smooth problems
- Set $\varepsilon \approx \Delta t^{p+1}$
- Relaxation term is diffusive in Chapman–Eskong
- Pure Galerkin discretization

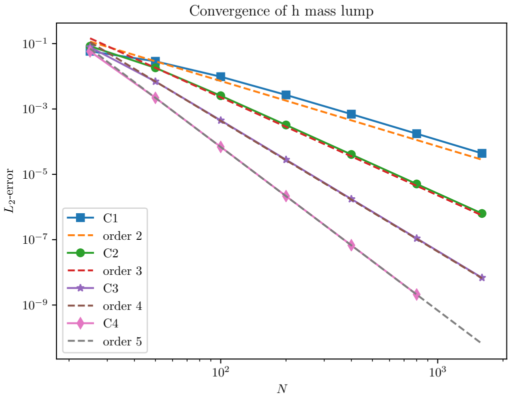


Figure: Subcritical flow $\mathbb{C}^1$ N=200, $\mathbb{C}^2$ N=100 and N=400 Perturbations of bump test cases with *cubature* elements: top to down η^ε, $h_p^\varepsilon - h^\varepsilon$ and q^ε; left $T = 0.76$ right $T = 2$.

Residual Distribution - Choice of the scheme

(11) $$\phi_\sigma^{K,LxF}(U_h) = \int_K \varphi_\sigma \left(\nabla \cdot A(U_h) - S(U_h)\right) dx + \alpha_K (U_\sigma - \overline{U}_h^K),$$

where $\overline{U}_h^K$ is the average of U_h over the cell K and α_K is defined as

(12) $$\alpha_K = \max_{e \text{ edge } \in K} \left(\rho_S \left(\nabla A(U_h) \cdot \mathbf{n}_e\right)\right),$$

ρ_S is the spectral radius.
For monotonicity near strong discontinuities, PSI limiter:

(13) $$\beta_\sigma^K(U_h) = \max\left(\frac{\Phi_\sigma^{K,LxF}}{\Phi^K}, 0\right) \left(\sum_{j\in K} \max\left(\frac{\Phi_j^{K,LxF}}{\Phi^K}, 0\right)\right)^{-1}$$

Residual Distribution - Choice of the scheme

Blending between LxF and PSI:

(14)
$$\phi_\sigma^{*,K} = (1-\Theta)\beta_\sigma^K \phi_\sigma^K + \Theta \Phi_\sigma^{K,LxF},$$
$$\Theta = \frac{|\Phi^K|}{\sum_{j\in K} |\Phi_j^{K,LxF}|}.$$

Nodal residual is finally given by

(15)
$$\phi_\sigma^K = \phi_\sigma^{*,K} + \sum_{e|\text{edge of } K} \theta h_e^2 \int_e [\nabla U_h]\cdot[\nabla \varphi_\sigma] d\Gamma.$$

DeC – Proof

Proof.

Let U^* be the solution of $\mathcal{L}^2(U^*)=0$. We know that $\mathcal{L}^1(U^*)=\mathcal{L}^1(U^*)-\mathcal{L}^2(U^*)$, so that

$$\begin{aligned}
\mathcal{L}^1(U^{(k+1)})-\mathcal{L}^1(U^*) &= \left(\mathcal{L}^1(U^{(k)})-\mathcal{L}^2(U^{(k)})\right)-\left(\mathcal{L}^1(U^*)-\mathcal{L}^2(U^*)\right)\\
&= \left(\mathcal{L}^1(U^{(k)})-\mathcal{L}^1(U^*)\right)-\left(\mathcal{L}^2(U^{(k)})-\mathcal{L}^2(U^*)\right)\\
\alpha_1||U^{(k+1)}-U^*|| \leq & ||\mathcal{L}^1(U^{(k+1)})-\mathcal{L}^1(U^*)|| =\\
=& ||\mathcal{L}^1(U^{(k)})-\mathcal{L}^2(U^{(k)})-(\mathcal{L}^1(U^*)-\mathcal{L}^2(U^*))|| \leq\\
\leq & \alpha_2\Delta||U^{(k)}-U^*||.\\
||U^{(k+1)}-U^*|| \leq & \left(\frac{\alpha_2}{\alpha_1}\Delta\right)||U^{(k)}-U^*|| \leq \left(\frac{\alpha_2}{\alpha_1}\Delta\right)^{k+1}||U^{(0)}-U^*||.
\end{aligned}$$

After K iteration we have an error at most of $\eta^K \cdot ||U^{(0)}-U^*||$. $\square$

Friction - Model

$$
\partial_t u(x,t) + \partial_x A(u(x,t)) + \\
S_b(u(x,t)) + S_f(u(x,t)) = 0,
$$

$$
S_f(u) := \begin{pmatrix} 0 \\ c_f(h,q)q \end{pmatrix}.
$$

Manning's law

$$
c_f(h,q) = \frac{n^2\|q\|}{h^{10/3}},
$$

with n being Manning's coefficient.

$$
\partial_t f^\varepsilon + \Lambda \partial_x f^\varepsilon + \tilde{S}_b(f^\varepsilon) + \tilde{S}_f(f^\varepsilon) \\
= \frac{M(Pf^\varepsilon) - f^\varepsilon}{\varepsilon},
$$

with

$$
\tilde{S}_f(f^\varepsilon) := \begin{pmatrix} 0 \\ c_f(h^\varepsilon, q^\varepsilon)q_1 \\ 0 \\ c_f(h^\varepsilon, q^\varepsilon)q_2 \end{pmatrix},
$$

so that the projection of this source term is equal to the friction in the SW equations, i.e., $P\tilde{S}_f(f^\varepsilon) = S_f(Pf^\varepsilon)$, and it verifies also $P\Lambda\tilde{S}_f(f) = S_f(P\Lambda f)$.

Friction - Implicit Discretization

Implicit Friction, without nonlinear solver.

- limit equation: $P\mathcal{L}_1$, where $h^{\varepsilon,n+1}$ explicit and $q^{\varepsilon,n+1}$

$$|K_\sigma| \left(q_\sigma^{\varepsilon,n,m} - q_\sigma^{\varepsilon,n,0}\right) + \Delta t \beta^m \sum_{K|\sigma\in K} P\Phi_K^{\sigma,ex}(f^{n,0}) + \Delta t \beta^m |K_\sigma| S_{f,q}(u_\sigma^{\varepsilon,n,m}) = 0,$$

- $\mathcal{E}_{q,\sigma}^n$ = all the explicit terms

$$q_\sigma^{\varepsilon,n,m}\left(1 + \Delta t \beta^m \frac{n^2 |q_\sigma^{\varepsilon,n,m}|}{(h_\sigma^{\varepsilon,n,m})^{10/3}}\right) = \mathcal{E}_{q,\sigma}^n.$$

- $\Delta t, n, h_\sigma^{\varepsilon,n,m} > 0$ known, solve for the absolute value of $q_\sigma^{\varepsilon,n,m}$

$$|q_\sigma^{\varepsilon,n,m}|\left(1 + \Delta t \beta^m \frac{n^2 |q_\sigma^{\varepsilon,n,m}|}{(h_\sigma^{\varepsilon,n,m})^{10/3}}\right) = |\mathcal{E}_{q,\sigma}^n|,$$

- Only 1 positive solution $|q_\sigma^{\varepsilon,n,m}| \Longrightarrow q_\sigma^{\varepsilon,n,m}$.

Friction - Implicit Discretization

- Only 1 positive solution $|q_\sigma^{\varepsilon,n,m}| \implies q_\sigma^{\varepsilon,n,m}$.
- Kinetic model

$$|K_\sigma|(f_\sigma^{n,m} - f_\sigma^{n,0}) + \Delta t \beta^m \left(\sum_{K|\sigma\in K} \Phi_K^{\sigma,ex}(f^{n,0}) + |K_\sigma| \left(\tilde{S}_f(f_\sigma^{n,m}) + \frac{u_\sigma^{n,m} - M(Pf_\sigma^{n,m})}{\varepsilon} \right) \right) = 0,$$

- Friction coefficient $c_{f,\sigma}^{n,m} := c_f(h_\sigma^{\varepsilon,n,m}, q_\sigma^{\varepsilon,n,m})$ known, $\mathcal{E}_\sigma^n$ all the explicit terms,

$$\begin{pmatrix} h_{1,\sigma}^{n,m}\left(1+\frac{\Delta t\beta^m}{\varepsilon}\right) \\ q_{1,\sigma}^{n,m}\left(1+\frac{\Delta t\beta^m}{\varepsilon}+\Delta t\beta^m c_{f,\sigma}^{n,m}\right) \\ h_{2,\sigma}^{n,m}\left(1+\frac{\Delta t\beta^m}{\varepsilon}\right) \\ q_{2,\sigma}^{n,m}\left(1+\frac{\Delta t\beta^m}{\varepsilon}+\Delta t\beta^m c_{f,\sigma}^{n,m}\right) \end{pmatrix} - \Delta t\beta^m \frac{M(u_\sigma^{\varepsilon,n,m})}{\varepsilon} + \mathcal{E}_\sigma^n = 0.$$

Again, also this final step can be computed explicitly and, hence, all the variables $f_\sigma^{n,m}$ can be obtained efficiently.

Global flux (GF) formulation

Global flux:

- Cheng et al.,Journal of Scientific Computing, 80 (2019)
- Liu et al., SIAM Journal on Scientific Computing, 42 (2020)

Main Idea

$$
\begin{aligned}
&u_t + \partial_x A(u) = S(u)\\
&u_t + \partial_x G(u) = 0\\
&\partial_x G(u) = \partial_x A(u) - S(u)\\
&G(u(x),x) = \int_{x_0}^{x} \partial_x A(u) - S(u,s)ds
\end{aligned}
$$

Global flux (GF) formulation

Global flux:
- Cheng et al.,Journal of Scientific Computing, 80 (2019)
- Liu et al., SIAM Journal on Scientific Computing, 42 (2020)

Main Idea

$$u_t + \partial_x A(u) = S(u)$$
$$u_t + \partial_x G(u) = 0$$
$$\partial_x G(u) = \partial_x A(u) - S(u)$$
$$G(u(x),x) = \int_{x_0}^{x} \partial_x A(u) - S(u,s)ds$$

RD as GF

$$\text{RD} \begin{cases} K = [a,b] \text{ with } p+1 \text{ degrees of freedom} \\ \phi_i^K = \int_K \omega_i (\partial_x A(u) + S(u))\mathrm{d}x,\ 0 \le i \le p, \\ \sum_{i \in K} \omega_i \equiv 1 \\ \partial_t u_i = -\frac{1}{\Delta x} \sum_K \phi_i^K, \end{cases}$$

- $S_h(u) \in \mathbb{P}^{p-1}(K)$,
- $\Sigma(x) := \int_{x_0}^{x} S_h(u(s))\mathrm{d}s \in \mathbb{P}^p(K)$,
- $\partial_x \Sigma(x) = S_h(u(x))$,
- $G(u) := A(u) - \Sigma(u)$.

Equivalence between RD and GF

Search: Global numerical flux $\hat{G}_{i+\frac{1}{2}}$ for $i = 0, \dots, p-1$

$$\begin{cases} \phi_i^K = \hat{G}_{i+\frac{1}{2}} - \hat{G}_{i-\frac{1}{2}}, & i = 1, \dots, p-1, \\ \phi_0^K = \hat{G}_{\frac{1}{2}} - G(a), \\ \phi_p^K = G(b) - \hat{G}_{p-\frac{1}{2}}. \end{cases} \tag{16}$$

Solution for numerical global flux

- $p+1$ equations p unknowns
- One linear dependent equation as $\phi^K = G(b) - G(a)$
- Explicit solution $\hat{G}_{\frac{1}{2}} = G(a) + \phi_0^K$; $\hat{G}_{i+\frac{1}{2}} = \hat{G}_{i-\frac{1}{2}} + \phi_i^K$.

Update formula

$$\partial_t u_i = -\frac{1}{\Delta x} \sum_K \phi_i^K$$

$$\partial_t u_i = -\frac{\hat{G}_{i+\frac{1}{2}} - \hat{G}_{i-\frac{1}{2}}}{\Delta x}$$

Consistency of RD GF

$$G(a) = \hat{G}_{i+\frac{1}{2}} = G(b),\ \forall i$$
$$\Longleftrightarrow$$
$$\phi_i^K = 0,\ \forall i$$