

# Nodal Discontinuous Galerkin Method for the Euler Equations in GR

Samuel J Dunham<sup>1</sup>, Eirik Endeve<sup>2</sup>, et al.

## Contents

<b>1</b>	<b>Discontinuous Galerkin Scheme</b>	<b>1</b>
<b>2</b>	<b>Bound-Preserving Methods Using First-Order DG Scheme</b>	<b>2</b>
2.1	Cartesian Coordinates . . . . .	2
2.1.1	Set of Admissible States . . . . .	2
2.1.2	Time-Step Derivation/CFL Condition . . . . .	2
2.2	Curvilinear Coordinates . . . . .	4
2.2.1	Set of Admissible States . . . . .	4
2.2.2	Time-Step Derivation/CFL Condition . . . . .	5
2.2.3	Numerical flux term . . . . .	6
	Time-step for Spherical-Polar Coordinates . . . . .	7
2.2.4	Source term . . . . .	7
<b>3</b>	<b>Computing the Time-Step Using Higher-Order DG Schemes</b>	<b>9</b>
<b>4</b>	<b>Recovery of Primitive Variables</b>	<b>10</b>
4.1	Upper and Lower Bounds for Pressure . . . . .	10

## 1. Discontinuous Galerkin Scheme

We assume a spacetime metric

$$ds^2 = -\alpha^2 dt^2 + \gamma_{ij} dx^i dx^j, \quad (1)$$

and consider the system of conservation laws with sources

$$\partial_t(\sqrt{\gamma} \mathbf{U}) + \sum_{i=1}^d \partial_i(\alpha \sqrt{\gamma} \mathbf{F}^i(\mathbf{U})) = \alpha \sqrt{\gamma} \mathbf{G}(\mathbf{U}), \quad (2)$$

---

<sup>2</sup>Department of Astronomy, Vanderbilt University, 6301 Stevenson Science Center, Nashville TN, 37212, USA; samuel.j.dunham@vanderbilt.edu

<sup>2</sup>Computational and Applied Mathematics Group, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6354, USA; endevee@ornl.gov

where

$$\mathbf{U} = (D, S_j, \tau)^\top = (\rho W, \rho h W^2 v_j, \rho W (h W - 1) - p)^\top, \quad (3)$$

$$\mathbf{F}^i(\mathbf{U}) = (D v^i, )^\top \quad (4)$$

## 2. Bound-Preserving Methods Using First-Order DG Scheme

### 2.1. Cartesian Coordinates

This section closely follows [Qin et al. \(2016\)](#).

#### 2.1.1. Set of Admissible States

We consider a one-dimensional system of conservation laws:

$$\partial_t \mathbf{U} + \partial_x \mathbf{F}(\mathbf{U}) = \mathbf{0}, \quad (5)$$

where  $\mathbf{U}$  is a vector of conserved variables, defined as:

$$\mathbf{U} \longrightarrow \begin{pmatrix} D \\ S \\ \tau \end{pmatrix} = \begin{pmatrix} \rho W \\ \rho h W^2 v \\ \rho W (h W - 1) - p \end{pmatrix}, \quad (6)$$

and  $\mathbf{F}(\mathbf{U})$  are the fluxes of those conserved quantities:

$$\mathbf{F}(\mathbf{U}) \longrightarrow \begin{pmatrix} \rho W v \\ \rho h W^2 v^2 + p \\ \rho h W^2 v - D v \end{pmatrix}. \quad (7)$$

The physics leads us to define a set of admissible states,  $\mathcal{G}_p$  (the subscript  $p$  stands for primitive), as:

$$\mathcal{G}_p \equiv \left\{ \mathbf{U} \mid \rho > 0, p > 0, v^2 < 1 \right\}. \quad (8)$$

It is shown in [Mignone & Bodo \(2005\)](#) that  $\mathcal{G}$  is a convex set<sup>3</sup> and can equivalently be written in terms of the conserved variables as:

$$\mathcal{G} \equiv \left\{ \mathbf{U} \mid D > 0, \tau + D > \sqrt{D^2 + S^2} \right\}. \quad (9)$$

#### 2.1.2. Time-Step Derivation/CFL Condition

For the first-order DG method using forward-Euler time-stepping, we evolve the vector of conserved variables as:

$$\overline{\mathbf{U}}_i^{n+1} = \overline{\mathbf{U}}_i^n - \eta_i \left[ \hat{\mathbf{F}}(\overline{\mathbf{U}}_i^n, \overline{\mathbf{U}}_{i+1}^n) - \hat{\mathbf{F}}(\overline{\mathbf{U}}_{i-1}^n, \overline{\mathbf{U}}_i^n) \right], \quad (10)$$

---

<sup>3</sup>Convex in the sense that if  $\mathbf{U}_1 \in \mathcal{G}$  and  $\mathbf{U}_2 \in \mathcal{G}$ , then  $\alpha_1 \mathbf{U}_1 + \alpha_2 \mathbf{U}_2 \in \mathcal{G}$ , where  $\alpha_1, \alpha_2 \in [0, 1]$  and  $\alpha_1 + \alpha_2 = 1$ .

where

$$\overline{U}_i \equiv \frac{1}{\Delta x_i} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} U_i dx, \quad (11)$$

$\eta_i \equiv \Delta t_i / \Delta x_i$ , and  $\hat{\mathbf{F}}$  is the numerical flux. In this document we use the local Lax-Friedrichs flux, defined as:

$$\hat{\mathbf{F}}(a, b) = \frac{1}{2} [\mathbf{F}(a) + \mathbf{F}(b) - \alpha_{ab}(b - a)], \quad (12)$$

where  $a$  and  $b$  represent the state of the fluid in two different elements,  $\alpha_{ab}$  is an estimate for the wave-speed:

$$\alpha_{ab} = \max[\alpha(a), \alpha(b)], \quad (13)$$

and  $\alpha$  is the largest (in absolute value) eigenvalue of the flux-Jacobian:

$$\alpha = \left\| \frac{\partial \mathbf{F}}{\partial \mathbf{U}} \right\|. \quad (14)$$

Using this we define the following variables:

$$\alpha_{i+\frac{1}{2}} = \max[\alpha(\overline{U}_i), \alpha(\overline{U}_{i+1})], \quad \alpha_{i-\frac{1}{2}} = \max[\alpha(\overline{U}_{i-1}), \alpha(\overline{U}_i)]. \quad (15)$$

Substituting (12) with (15) into (10):

$$\begin{aligned} \overline{U}_i^{n+1} &= \overline{U}_i^n - \frac{\eta_i}{2} [\mathbf{F}(\overline{U}_i^n) + \mathbf{F}(\overline{U}_{i+1}^n) - \alpha_{i+\frac{1}{2}}(\overline{U}_{i+1}^n - \overline{U}_i^n) \\ &\quad - \mathbf{F}(\overline{U}_i^n) - \mathbf{F}(\overline{U}_{i-1}^n) + \alpha_{i-\frac{1}{2}}(\overline{U}_i^n - \overline{U}_{i-1}^n)] \\ &= \left[1 - \frac{\eta_i}{2}(\alpha_{i+\frac{1}{2}} + \alpha_{i-\frac{1}{2}})\right] \overline{U}_i^n + \frac{\eta_i}{2} \alpha_{i+\frac{1}{2}} \left[\overline{U}_{i+1}^n - \frac{1}{\alpha_{i+\frac{1}{2}}} \mathbf{F}(\overline{U}_{i+1}^n)\right] \\ &\quad + \frac{\eta_i}{2} \alpha_{i-\frac{1}{2}} \left[\overline{U}_{i-1}^n + \frac{1}{\alpha_{i-\frac{1}{2}}} \mathbf{F}(\overline{U}_{i-1}^n)\right] \\ &= \left[1 - \frac{\eta_i}{2}(\alpha_{i+\frac{1}{2}} + \alpha_{i-\frac{1}{2}})\right] \overline{U}_i^n + \frac{\eta_i}{2} \alpha_{i+\frac{1}{2}} \overline{\mathbf{H}}^-(\overline{U}_{i+1}^n, \alpha_{i+\frac{1}{2}}) + \frac{\eta_i}{2} \alpha_{i-\frac{1}{2}} \overline{\mathbf{H}}^+(\overline{U}_{i-1}^n, \alpha_{i-\frac{1}{2}}), \end{aligned} \quad (16)$$

where

$$\overline{\mathbf{H}}^\pm(\overline{U}, \alpha) \equiv \overline{U} \pm \frac{1}{\alpha} \mathbf{F}(\overline{U}). \quad (17)$$

The proof that  $\overline{\mathbf{H}}^\pm \in \mathcal{G}$  is given in [Qin et al. \(2016\)](#). Therefore, we see that with a restriction on  $\alpha_{i\pm\frac{1}{2}}$  that (16) is a convex combination. The restriction is (recalling that  $\eta_i = \Delta t_i / \Delta x_i$ ):

$$1 - \frac{\eta_i}{2}(\alpha_{i+\frac{1}{2}} + \alpha_{i-\frac{1}{2}}) > 0 \implies \frac{\eta_i}{2}(\alpha_{i+\frac{1}{2}} + \alpha_{i-\frac{1}{2}}) < 1 \implies \Delta t_i < \frac{2 \Delta x_i}{\alpha_{i+\frac{1}{2}} + \alpha_{i-\frac{1}{2}}} \leq \frac{\Delta x_i}{\max(\alpha_{i\pm\frac{1}{2}})}. \quad (18)$$

We want a time-step that is the same for all elements at a given time, so we tighten the restriction to:

$$\Delta t < \min_i \left( \frac{\Delta x_i}{\max(\alpha_{i\pm\frac{1}{2}})} \right) = \frac{\Delta x}{\max_i(\alpha_{i\pm\frac{1}{2}})}, \quad (19)$$

where the equality follows for a uniform mesh, i.e.  $\Delta x_i = \Delta x \forall i$ .

## 2.2. Curvilinear Coordinates

NOTE: We assume a conformally-flat, time-independent spatial three-metric:

$$\gamma_{ij}(x^k, t) \longrightarrow \psi^4(x^k) \bar{\gamma}_{ii}(x^k), \quad (20)$$

where  $\psi(x^k)$  is the conformal factor and  $\bar{\gamma}_{ii}$  is the flat-space metric.

### 2.2.1. Set of Admissible States

We again consider a one-dimensional system of conservation laws, but this time with a curvilinear metric:

$$\partial_t(\sqrt{\gamma} \mathbf{U}) + \partial_1(\sqrt{\gamma} \mathbf{F}) = \sqrt{\gamma} \mathbf{Q}, \quad (21)$$

where  $\mathbf{U}$  is given by:

$$\mathbf{U} \longrightarrow \begin{pmatrix} D \\ S_1 \\ \tau \end{pmatrix} = \begin{pmatrix} \rho W \\ \rho h W^2 v_1 \\ \rho W(h W - 1) - p \end{pmatrix} = \begin{pmatrix} \rho W \\ \rho h W^2 \gamma_{1j} v^j \\ \rho W(h W - 1) - p \end{pmatrix}, \quad (22)$$

$\mathbf{F}(\mathbf{U})$  are the fluxes of those conserved quantities:

$$\mathbf{F}(\mathbf{U}) \longrightarrow \begin{pmatrix} D v^1 \\ S^1 v_1 + p \delta_1^1 \\ S^1 - D v^1 \end{pmatrix} = \begin{pmatrix} \rho W v^1 \\ \rho h W^2 v^1 v_1 + p \\ \rho h W^2 v^1 - D v^1 \end{pmatrix} = \begin{pmatrix} \rho W v^1 \\ \rho h W^2 \gamma_{1j} v^1 v^j + p \\ \rho h W^2 v^1 - D v^1 \end{pmatrix}, \quad (23)$$

and  $\mathbf{Q}$  is a source term:

$$\mathbf{Q} \longrightarrow \begin{pmatrix} 0 \\ \frac{1}{2} P^{kj} \partial_1 \gamma_{kj} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{2} [P^{11} \partial_1 \gamma_{11} + P^{22} \partial_1 \gamma_{22} + P^{33} \partial_1 \gamma_{33}] \\ 0 \end{pmatrix} \quad (24)$$

$$= \begin{pmatrix} 0 \\ P^{11} h_1 \partial_1 h_1 + P^{22} h_2 \partial_1 h_2 + P^{33} h_3 \partial_1 h_3 \\ 0 \end{pmatrix}, \quad (25)$$

where we have used the fact that  $\gamma_{jj} = (h_j)^2$ . The  $P^{kj}$  are components of the pressure tensor:

$$P^{kj} = S^k v^j + p \gamma^{kj} = \gamma^{k\ell} S_\ell v^j + p \gamma^{kj} = \gamma^{k\ell} S_\ell v^j + p \gamma^{k\ell} \delta_\ell^j = \gamma^{k\ell} (S_\ell v^j + p \delta_\ell^j). \quad (26)$$

Since the spatial three-metric is diagonal we must have that  $\ell = k$ . We can therefore simplify further:

$$P^{kj} = \gamma^{kk} (S_k v^j + p \delta_k^j) = \frac{1}{\gamma_{kk}} (S_k v^j + p \delta_k^j) = \frac{1}{(h_k)^2} (S_k v^j + p \delta_k^j) \quad (27)$$

For the pressure-tensor sum, we then have:

$$P^{kj} \partial_1 \gamma_{kj} = P^{kk} \partial_1 \gamma_{kk} = P^{kk} \partial_1 (h_k)^2 = 2 P^{kk} h_k \partial_1 h_k = \frac{2}{h_k} (S_k v^k + p) \partial_1 h_k. \quad (28)$$

These definitions lead us to define the same set of admissible states as before, namely:

$$\mathcal{G}_p \equiv \left\{ \mathbf{U} \mid \rho > 0, p > 0, v^2 < 1 \right\}, \quad (29)$$

the only difference being that  $v^2$  now involves the metric:

$$v^2 = v^j v_j = \gamma_{kj} v^k v^j. \quad (30)$$

Before continuing, we show that the introduction of the metric doesn't affect the translation between  $\mathcal{G}_p$  and  $\mathcal{G}_{\dots}$  (SD: I've shown this, just need to TeX it up)

### 2.2.2. Time-Step Derivation/CFL Condition

We start by integrating both sides over  $dx^1$  and dividing by the volume,  $\Delta V_i$ :

$$\frac{1}{\Delta V_i} \int_{x_{i-\frac{1}{2}}^1}^{x_{i+\frac{1}{2}}^1} \partial_t (\sqrt{\gamma} \mathbf{U}_i) dx^1 + \frac{1}{\Delta V_i} \int_{x_{i-\frac{1}{2}}^1}^{x_{i+\frac{1}{2}}^1} \partial_1 (\sqrt{\gamma} \mathbf{F}(\mathbf{U}_i)) dx^1 = \frac{1}{\Delta V_i} \int_{x_{i-\frac{1}{2}}^1}^{x_{i+\frac{1}{2}}^1} \sqrt{\gamma} \mathbf{Q}_i dx^1, \quad (31)$$

where:

$$\Delta V_i = \int_{x_{i-\frac{1}{2}}^1}^{x_{i+\frac{1}{2}}^1} dV = \int_{x_{i-\frac{1}{2}}^1}^{x_{i+\frac{1}{2}}^1} \sqrt{\gamma} dx^1. \quad (32)$$

By defining the cell-average as:

$$\overline{\mathbf{W}}_i \equiv \frac{1}{\Delta V_i} \int_{x_{i-\frac{1}{2}}^1}^{x_{i+\frac{1}{2}}^1} \mathbf{W}_i dV, \quad (33)$$

we have:

$$\frac{d\overline{\mathbf{U}}_i}{dt} + \frac{1}{\Delta V_i} \left( \sqrt{\gamma} \hat{\mathbf{F}}(\overline{\mathbf{U}}) \right) \Big|_{x_{i-\frac{1}{2}}^1}^{x_{i+\frac{1}{2}}^1} = \overline{\mathbf{Q}}_i, \quad (34)$$

or, using the common notation of the time step being represented as a superscript and the spatial element represented by a subscript:

$$\overline{\mathbf{U}}_i^{n+1} = \overline{\mathbf{U}}_i^n - \frac{\Delta t_i}{\Delta V_i} \left[ \sqrt{\gamma}_{i+\frac{1}{2}} \hat{\mathbf{F}}_{i+\frac{1}{2}}^n - \sqrt{\gamma}_{i-\frac{1}{2}} \hat{\mathbf{F}}_{i-\frac{1}{2}}^n \right] + \Delta t_i \overline{\mathbf{Q}}_i^n. \quad (35)$$

Now we define a parameter  $\varepsilon$  (Zhang & Shu (2011)):  $\varepsilon \in (0, 1)$ , such that (NOTE: Zhang & Shu (2011) set  $\varepsilon = 1/2$ ):

$$\overline{\mathbf{U}}_i^n = \varepsilon \overline{\mathbf{U}}_i^n + (1 - \varepsilon) \overline{\mathbf{U}}_i^n. \quad (36)$$

We can use the first term to balance out the term in the square brackets and the second term to balance out the source term.

So, we get:

$$\overline{\mathbf{U}}_i^{n+1} = \varepsilon \left\{ \overline{\mathbf{U}}_i^n - \frac{\Delta t_i}{\varepsilon \Delta V_i} \left[ \sqrt{\gamma}_{i+\frac{1}{2}} \hat{\mathbf{F}}_{i+\frac{1}{2}}^n - \sqrt{\gamma}_{i-\frac{1}{2}} \hat{\mathbf{F}}_{i-\frac{1}{2}}^n \right] \right\} + (1 - \varepsilon) \overline{\mathbf{U}}_i^n + \Delta t_i \overline{\mathbf{Q}}_i^n \quad (37)$$

$$= \varepsilon \left\{ \overline{\mathbf{U}}_i^n - \eta_i \left[ \sqrt{\gamma}_{i+\frac{1}{2}} \hat{\mathbf{F}}(\overline{\mathbf{U}}_{i+1}^n, \overline{\mathbf{U}}_i^n) - \sqrt{\gamma}_{i-\frac{1}{2}} \hat{\mathbf{F}}(\overline{\mathbf{U}}_i^n, \overline{\mathbf{U}}_{i-1}^n) \right] \right\} + (1 - \varepsilon) \overline{\mathbf{U}}_i^n + \Delta t_i \overline{\mathbf{Q}}_i^n \quad (38)$$

$$= \varepsilon \overline{\mathbf{H}}_1 + (1 - \varepsilon) \overline{\mathbf{H}}_2, \quad (39)$$

where

$$\overline{\mathbf{H}}_1 \equiv \overline{\mathbf{U}}_i^n - \eta_i \left[ \sqrt{\gamma}_{i+\frac{1}{2}} \hat{\mathbf{F}}(\overline{\mathbf{U}}_{i+1}^n, \overline{\mathbf{U}}_i^n) - \sqrt{\gamma}_{i-\frac{1}{2}} \hat{\mathbf{F}}(\overline{\mathbf{U}}_i^n, \overline{\mathbf{U}}_{i-1}^n) \right], \quad (40)$$

$$\overline{H}_2 \equiv \overline{U}_i^n + \frac{\Delta t_i}{1 - \varepsilon} \overline{Q}_i^n, \quad (41)$$

and

$$\eta_i \equiv \frac{\Delta t_i}{\varepsilon \Delta V_i}. \quad (42)$$

We proceed by focusing on each term individually, starting with the numerical flux term,  $\overline{H}_1$ .

### 2.2.3. Numerical flux term

We have to show that  $\overline{H}_1 \in \mathcal{G}$ . We again use the Local-Lax-Friedrichs flux, (12), yielding for  $\overline{H}_1$ :

$$\overline{U}_i^n - \frac{\eta_i}{2} \left\{ \sqrt{\gamma_{i+\frac{1}{2}}} \left[ \mathbf{F}(\overline{U}_{i+1}^n) + \mathbf{F}(\overline{U}_i^n) - \alpha_{i+\frac{1}{2}} (\overline{U}_{i+1}^n - \overline{U}_i^n) \right] \right. \quad (43)$$

$$\left. - \sqrt{\gamma_{i-\frac{1}{2}}} \left[ \mathbf{F}(\overline{U}_i^n) + \mathbf{F}(\overline{U}_{i-1}^n) - \alpha_{i-\frac{1}{2}} (\overline{U}_i^n - \overline{U}_{i-1}^n) \right] \right\} \quad (44)$$

$$= \left( 1 - \frac{1}{2} \eta_i \sqrt{\gamma_{i+\frac{1}{2}}} \alpha_{i+\frac{1}{2}} - \frac{1}{2} \eta_i \sqrt{\gamma_{i-\frac{1}{2}}} \alpha_{i-\frac{1}{2}} \right) \overline{U}_i^n \quad (45)$$

$$- \frac{1}{2} \eta_i \sqrt{\gamma_{i+\frac{1}{2}}} \mathbf{F}(\overline{U}_i^n) + \frac{1}{2} \eta_i \sqrt{\gamma_{i-\frac{1}{2}}} \mathbf{F}(\overline{U}_i^n) \quad (46)$$

$$+ \frac{1}{2} \eta_i \sqrt{\gamma_{i-\frac{1}{2}}} \alpha_{i-\frac{1}{2}} \left[ \overline{U}_{i-1}^n + \frac{1}{\alpha_{i-\frac{1}{2}}} \mathbf{F}(\overline{U}_{i-1}^n) \right] + \frac{1}{2} \eta_i \sqrt{\gamma_{i+\frac{1}{2}}} \alpha_{i+\frac{1}{2}} \left[ \overline{U}_{i+1}^n - \frac{1}{\alpha_{i+\frac{1}{2}}} \mathbf{F}(\overline{U}_{i+1}^n) \right]. \quad (47)$$

Now we add and subtract  $\frac{1}{2} \eta_i \sqrt{\gamma_{i+\frac{1}{2}}} \alpha_{i+\frac{1}{2}} \overline{U}_i^n$  and  $\frac{1}{2} \eta_i \sqrt{\gamma_{i-\frac{1}{2}}} \alpha_{i-\frac{1}{2}} \overline{U}_i^n$ , yielding:

$$\left( 1 - \eta_i \sqrt{\gamma_{i+\frac{1}{2}}} \alpha_{i+\frac{1}{2}} - \eta_i \sqrt{\gamma_{i-\frac{1}{2}}} \alpha_{i-\frac{1}{2}} \right) \overline{U}_i^n \quad (48)$$

$$+ \frac{1}{2} \eta_i \sqrt{\gamma_{i+\frac{1}{2}}} \alpha_{i+\frac{1}{2}} \left[ \overline{U}_i^n - \frac{1}{\alpha_{i+\frac{1}{2}}} \mathbf{F}(\overline{U}_i^n) \right] + \frac{1}{2} \eta_i \sqrt{\gamma_{i-\frac{1}{2}}} \alpha_{i-\frac{1}{2}} \left[ \overline{U}_i^n + \frac{1}{\alpha_{i-\frac{1}{2}}} \mathbf{F}(\overline{U}_i^n) \right] \quad (49)$$

$$+ \frac{1}{2} \eta_i \sqrt{\gamma_{i-\frac{1}{2}}} \alpha_{i-\frac{1}{2}} \left[ \overline{U}_{i-1}^n + \frac{1}{\alpha_{i-\frac{1}{2}}} \mathbf{F}(\overline{U}_{i-1}^n) \right] + \frac{1}{2} \eta_i \sqrt{\gamma_{i+\frac{1}{2}}} \alpha_{i+\frac{1}{2}} \left[ \overline{U}_{i+1}^n - \frac{1}{\alpha_{i+\frac{1}{2}}} \mathbf{F}(\overline{U}_{i+1}^n) \right]. \quad (50)$$

All of the terms in square brackets are similar to the  $\overline{H}$  quantities in [Qin et al. \(2016\)](#), and are therefore in  $\mathcal{G}$ . It can easily be seen that the sum of the coefficients is unity. The final condition is that the coefficient of  $\overline{U}_i^n > 0$ , or (recalling that  $\eta_i = \Delta t_i / (\varepsilon \Delta V_i)$ ):

$$1 - \eta_i \sqrt{\gamma_{i+\frac{1}{2}}} \alpha_{i+\frac{1}{2}} - \eta_i \sqrt{\gamma_{i-\frac{1}{2}}} \alpha_{i-\frac{1}{2}} > 0 \implies \eta_i \left( \sqrt{\gamma_{i+\frac{1}{2}}} \alpha_{i+\frac{1}{2}} + \sqrt{\gamma_{i-\frac{1}{2}}} \alpha_{i-\frac{1}{2}} \right) < 1 \quad (51)$$

$$\implies \Delta t_i < \frac{\varepsilon \Delta V_i}{\sqrt{\gamma_{i+\frac{1}{2}}} \alpha_{i+\frac{1}{2}} + \sqrt{\gamma_{i-\frac{1}{2}}} \alpha_{i-\frac{1}{2}}} \leq \frac{\varepsilon \Delta V_i}{2 \max \left( \sqrt{\gamma_{i\pm\frac{1}{2}}} \alpha_{i\pm\frac{1}{2}} \right)}. \quad (52)$$

Again we want a time-step that is the same for all elements at a given time, so:

$$\Delta t < \min_i \left( \frac{\varepsilon \Delta V_i}{2 \max \left( \sqrt{\gamma_{i\pm\frac{1}{2}}} \alpha_{i\pm\frac{1}{2}} \right)} \right). \quad (53)$$

We close the numerical flux section by writing the explicit form of the time-step for spherical-polar coordinates.

### Time-step for Spherical-Polar Coordinates

For spherical-polar coordinates in 1-D we have that  $\Delta V_i = 1/3 \left( r_{i+\frac{1}{2}}^3 - r_{i-\frac{1}{2}}^3 \right)$ , and (assuming  $\alpha_{i\pm\frac{1}{2}} = 1 \ \forall \ i$ )  $\max \left( \sqrt{\gamma_{i\pm\frac{1}{2}}} \alpha_{i\pm\frac{1}{2}} \right) = r_{i+\frac{1}{2}}^2$ , so:

$$\Delta t < \min_i \left\{ \frac{\varepsilon 1/3 \left[ r_{i+\frac{1}{2}}^3 - r_{i-\frac{1}{2}}^3 \right]}{2 r_{i+\frac{1}{2}}^2} \right\} \quad (54)$$

$$= \min_i \left\{ \frac{\varepsilon}{6} r_{i+\frac{1}{2}} \left[ 1 - \frac{r_{i-\frac{1}{2}}^3}{r_{i+\frac{1}{2}}^3} \right] \right\} \quad (55)$$

$$= \min_i \left\{ \frac{\varepsilon}{6} r_{i+\frac{1}{2}} \left[ 1 - \left( 1 - \frac{\Delta r_i}{r_{i+\frac{1}{2}}} \right)^3 \right] \right\} \quad (56)$$

$$= \min_i \left\{ \frac{\varepsilon}{6} r_{i+\frac{1}{2}} \left[ 1 - \left( 1 + \left( \frac{\Delta r_i}{r_{i+\frac{1}{2}}} \right)^2 - 2 \frac{\Delta r_i}{r_{i+\frac{1}{2}}} \right) \left( 1 - \frac{\Delta r_i}{r_{i+\frac{1}{2}}} \right) \right] \right\} \quad (57)$$

$$= \min_i \left\{ \frac{\varepsilon}{6} r_{i+\frac{1}{2}} \left[ \left( \frac{\Delta r_i}{r_{i+\frac{1}{2}}} \right)^3 - 3 \left( \frac{\Delta r_i}{r_{i+\frac{1}{2}}} \right)^2 + 3 \frac{\Delta r_i}{r_{i+\frac{1}{2}}} \right] \right\} \quad (58)$$

$$= \min_i \left\{ \frac{\varepsilon}{6} \Delta r_i \left[ \left( \frac{\Delta r_i}{r_{i+\frac{1}{2}}} \right)^2 - 3 \left( \frac{\Delta r_i}{r_{i+\frac{1}{2}}} \right) + 3 \right] \right\}. \quad (59)$$

We know that  $\Delta r_i / r_{i+\frac{1}{2}} \in [0, 1]$ ; the minimum value of the quadratic function in this domain is unity. So, we have that for spherical-polar coordinates:

$$\Delta t < \frac{\varepsilon}{6} \min (\Delta r_i). \quad (60)$$

Next we handle the source term.

#### 2.2.4. Source term

For this section we drop the subscript  $i$  and the superscript  $n$ . We have to show that  $\overline{\mathbf{H}}_2 \in \mathcal{G}$ , where

$$\overline{\mathbf{H}}_2 = \left( \overline{S}_1 + \frac{\overline{D}}{2(1-\varepsilon)} \overline{P^{kk}} \overline{\partial_1 \gamma_{kk}} \right), \quad (\overline{H}_2)_1 > 0, \quad (\overline{H}_2)_3 + (\overline{H}_2)_1 > \sqrt{(\overline{H}_2)_1 (\overline{H}_2)_1 + (\overline{H}_2)_2 (\overline{H}_2)_2}. \quad (61)$$

It is clear that the first requirement for  $\overline{H}_2$  is met, i.e.  $\overline{D} > 0$ . The second requirement is:

$$\overline{D} + \overline{\tau} > \sqrt{\overline{D}^2 + \left[ \overline{S}_1 + \frac{\Delta t}{2(1-\varepsilon)} \overline{P^{kk} \partial_1 \gamma_{kk}} \right] \left[ \overline{S}^1 + \frac{\Delta t}{2(1-\varepsilon)} \overline{P^{kk} \partial^1 \gamma_{kk}} \right]} \quad (62)$$

$$= \sqrt{\overline{D}^2 + \left[ \overline{S}_1 + \frac{\Delta t}{2(1-\varepsilon)} \overline{P^{kk} \partial_1 \gamma_{kk}} \right] \left[ \overline{S}^1 + \gamma^{11} \frac{\Delta t}{2(1-\varepsilon)} \overline{P^{kk} \partial_1 \gamma_{kk}} \right]} \quad (63)$$

$$= \sqrt{\overline{D}^2 + \overline{S}_1 \overline{S}^1 + \gamma^{11} \left[ \frac{\overline{P^{kk} \partial_1 \gamma_{kk}}}{2(1-\varepsilon)} \right]^2 (\Delta t)^2 + \frac{\overline{P^{kk} \partial_1 \gamma_{kk}}}{2(1-\varepsilon)} [\overline{S}_1 \gamma^{11} + \overline{S}^1] \Delta t} \quad (64)$$

$$= \sqrt{a (\Delta t)^2 + b \Delta t + c} \geq 0, \quad (65)$$

where

$$a = \left[ \frac{\overline{P^{kk} \partial_1 \gamma_{kk}}}{2 \sqrt{\gamma_{11}} (1-\varepsilon)} \right]^2 \geq 0 \quad (66)$$

$$b = \frac{\overline{P^{kk} \partial_1 \gamma_{kk}}}{\gamma_{11} (1-\varepsilon)} \overline{S}_1 = \frac{2 \overline{S}_1}{\sqrt{\gamma_{11}}} \sqrt{a} \quad (67)$$

$$c = \overline{D}^2 + \overline{S}_1 \overline{S}^1 \geq 0. \quad (68)$$

The condition for a quadratic function to be positive-definite is that  $a > 0$  (concave-up) and  $b^2 - 4ac < 0$  (no real roots). The first condition is clearly met. For the second condition:

$$b^2 - 4ac = \frac{4 (\overline{S}_1)^2}{\gamma_{11}} a - 4a (\overline{D}^2 + \overline{S}_1 \overline{S}^1) = 4 \overline{S}_1 \overline{S}^1 a - 4a \overline{D}^2 - 4a \overline{S}_1 \overline{S}^1 = -4a \overline{D}^2 < 0. \quad (69)$$

So, the square root is always positive.

With that we have:

$$\overline{D}^2 + \overline{\tau}^2 + 2 \overline{D} \overline{\tau} > a (\Delta t)^2 + b \Delta t + c \quad (70)$$

$$\implies a (\Delta t)^2 + b \Delta t + c' < 0, \quad (71)$$

where:

$$c' = c - \overline{D}^2 - \overline{\tau}^2 - 2 \overline{D} \overline{\tau} = \overline{S}_1 \overline{S}^1 - \overline{\tau}^2 - 2 \overline{D} \overline{\tau}. \quad (72)$$

This time we want to make sure that our function has at least one real root. We still have that  $a > 0$ , so now we check if  $b^2 - 4ac' > 0$ :

$$b^2 - 4ac' = \frac{4 (\overline{S}_1)^2}{\gamma_{11}} a - 4a (\overline{S}_1 \overline{S}^1 - \overline{\tau}^2 - 2 \overline{D} \overline{\tau}) = 4a \overline{\tau} (\overline{\tau} + 2 \overline{D}). \quad (73)$$

Since  $\overline{\tau} \geq 0$ , we must have that  $\overline{\tau} > -2 \overline{D}$ . But, from condition two for  $\overline{H}_2$  we have that  $\overline{\tau} > -\overline{D}$ , so this condition is automatically satisfied.



The solutions to this quadratic equation are:

$$\Delta t = \frac{-b}{2a} \pm \frac{1}{2a} \sqrt{b^2 - 4ac'} = \frac{-\bar{S}_1}{\sqrt{\gamma_{11}} \sqrt{a}} \pm \frac{1}{2a} \sqrt{\frac{4\bar{S}_1^2}{\gamma_{11}} a - 4ac'} \quad (74)$$

$$= \frac{-\bar{S}_1}{\sqrt{\gamma_{11}} \sqrt{a}} \pm \frac{1}{\sqrt{\gamma_{11}} \sqrt{a}} \sqrt{\bar{S}_1^2 - c' \gamma_{11}} = \frac{1}{\sqrt{\gamma_{11}} \sqrt{a}} \left[ -\bar{S}_1 \pm \sqrt{\bar{S}_1^2 - c' \gamma_{11}} \right] \quad (75)$$

$$= \frac{1}{\sqrt{\gamma_{11}} \sqrt{a}} \left[ -\bar{S}_1 \pm \sqrt{\gamma_{11} (\bar{\tau}^2 + 2\bar{D}\bar{\tau})} \right] \quad (76)$$

$$= \frac{2(1-\varepsilon)}{\bar{P}^{kk} \partial_1 \gamma_{kk}} \left[ -\bar{S}_1 \pm \sqrt{\gamma_{11} (\bar{\tau}^2 + 2\bar{D}\bar{\tau})} \right] \quad (77)$$

$$= \frac{2(1-\varepsilon)}{\bar{P}^{kk} \partial_1 \gamma_{kk}} \left[ -\bar{S}_1 \pm \sqrt{\gamma_{11} \bar{\tau} (\bar{\tau} + 2\bar{D})} \right]. \quad (78)$$

### 3. Computing the Time-Step Using Higher-Order DG Schemes

When making the jump to higher-order DG schemes, we can simply do the same as in the first-order scheme, except we compute the quantities in all of the nodal points instead of using a cell-average. This is valid because the cell-average is a convex combination...**(SD: Need to expand on this)**. The proof starts with the discretized equation valid at each quadrature point,  $q$ :

$$\mathbf{U}_q^{n+1} = \mathbf{U}_q^n + \Delta t \mathcal{L}_q^n, \quad (79)$$

where  $\mathcal{L}_q^n$  is a general form of the RHS at time  $t^n$ . If we define a vector  $\bar{\mathbf{U}} \equiv (\mathbf{U}_1, \dots, \mathbf{U}_q, \dots, \mathbf{U}_Q)^T$ , where  $Q$  is the total number of quadrature points, and  $\bar{\mathbf{W}} \equiv (\mathbf{W}_1, \dots, \mathbf{W}_q, \dots, \mathbf{W}_Q)^T$  as a vector of quadrature weights, then we can write the cell-average of  $\mathbf{U}$  as:

$$\mathbf{U}_K \equiv \bar{\mathbf{W}}^T \bar{\mathbf{U}}. \quad (80)$$

If we then compute the cell-average of the above equation, we get:

$$\mathbf{U}_K^{n+1} = \mathbf{U}_K^n + \Delta t \bar{\mathbf{W}}^T \bar{\mathcal{L}}_q^n = \bar{\mathbf{W}}^T (\bar{\mathbf{U}}^n + \Delta t \bar{\mathcal{L}}^n) \quad (81)$$

#### 4. Recovery of Primitive Variables

In order to recover the primitive from the conserved variables we need to solve the nonlinear equation:

$$f(p) = p - \bar{p}(p) = 0, \quad (82)$$

where  $\bar{p}(p)$  is the pressure as obtained via the ideal gas equation of state with an initial guess,  $p$ :

$$\bar{p} = (\Gamma - 1) \rho \epsilon, \quad (83)$$

where

$$\rho = \rho(\mathbf{U}, p), \quad \epsilon = \epsilon(\mathbf{U}, p). \quad (84)$$

In order to solve this equation we make use of the bisection method, and therefore need bounds on our initial guess for the pressure.

##### 4.1. Upper and Lower Bounds for Pressure

We obtain a lower bound for the pressure with:

$$\tau = D(hW - 1) - p \implies p = -(\tau + D) + D h W \geq -(\tau + D) + D h W \sqrt{v^i v_i} = -(\tau + D) + \sqrt{S^i S_i}. \quad (85)$$

So, since the pressure must be non-negative, we have:

$$p \geq \text{MAX} \left[ -(\tau + D) + \sqrt{S^i S_i}, \text{SqrtTiny} \right]. \quad (86)$$

For an upper bound, we first note that:

$$h = 1 + \frac{e + p}{\rho} = 1 + \frac{\Gamma}{\Gamma - 1} \frac{p}{\rho} = 1 + \frac{\Gamma}{\Gamma - 1} \frac{p W}{D}, \quad (87)$$

so,

$$\tau = D \left( W + \frac{\Gamma}{\Gamma - 1} \frac{p W^2}{D} - 1 \right) - p = D(W - 1) + p \left( \frac{\Gamma}{\Gamma - 1} W^2 - 1 \right). \quad (88)$$

So,

$$p = \frac{\tau - D(W - 1)}{\frac{\Gamma}{\Gamma - 1} W^2 - 1}. \quad (89)$$

We also have:

$$W = (1 - v^i v_i)^{-1/2} = \left( 1 - \frac{S^i S_i}{(\tau + D + p)^2} \right)^{-1/2}. \quad (90)$$

Treating  $p$  as an independent variable (SD: is this valid?), we have:

$$W \Big|_{p \rightarrow \infty} = 1, \quad (91)$$

which gives us an upper limit:

$$p \leq \frac{\Gamma - 1}{\Gamma} \tau. \quad (92)$$

Just to be safe, in the code we multiply this by two, so that:

$$p \leq 2 \frac{\Gamma - 1}{\Gamma} \tau. \quad (93)$$

## REFERENCES

- Mignone, A., & Bodo, G. 2005, *Monthly Notices of the Royal Astronomical Society*, 364, 126
- Qin, T., Shu, C.-W., & Yang, Y. 2016, *Journal of Computational Physics*, 315, 323
- Zhang, X., & Shu, C.-W. 2011, *Journal of Computational Physics*, 230, 1238