Analysis and development of finite volume methods for the new generation of cubed sphere dynamical cores for the atmosphere

Luan da Fonseca Santos

Thesis presented to the
Institute of Mathematics and Statistics
of the University of São Paulo
in partial fulfillment
of the requirements
for the degree of
Doctor of Science

Program: Applied Mathematics
Advisor: Prof. Pedro da Silva Peixoto

During the development of this work the author was supported by CAPES and FAPESP (grant number 20/10280-4)

São Paulo
July, 2023

Analysis and development of finite volume methods for the new generation of cubed sphere dynamical cores for the atmosphere

Luan da Fonseca Santos

This is the original version of the thesis prepared by candidate Luan da Fonseca Santos, as submitted to the Examining Committee.

Education is what remains after one has forgotten everything he learned in school.
— Albert Einstein

Acknowledgements

TBW

Resumo

Luan da Fonseca Santos. **Análise e desenvolvimento de métodos de volumes finitos para modelos da nova geração da dinâmica atmosférica baseados na esfera cubada**. Tese (Doutorado). Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2023.

TBW

Palavras-chave: Núcleo dinâmico da atmosfera, esfera cubada, volumes finitos, dimension splitting, ponto de partida, corretor de massa.

Abstract

Luan da Fonseca Santos. **Analysis and development of finite volume methods for the new generation of cubed sphere dynamical cores for the atmosphere**. Thesis (Doctorate). Institute of Mathematics and Statistics, University of São Paulo, São Paulo, 2023.

TBW

Keywords: Dynamical core, cubed-sphere, finite-volume, dimension splitting, departure point, mass fixer.

Contents

Chapter 1

Introduction

1.1 Background

Chapter 2

One-dimensional finite-volume methods

The aim of this chapter is to provide a detailed description of one-dimensional (1D) finite-volume (FV) schemes within a Semi-Lagrangian (SL) framework, specifically applied to the 1D advection equation. These schemes are also known as flux-form Semi-Lagrangian schemes, and they allow for time steps beyond the Courant-Friedrichs-Lewy (CFL) condition while preserving the total mass. FV-SL schemes have been explored in the literature since the work of LeVeque (1985), which extended the finite-volume schemes from Godunov (1959) to accommodate larger time steps. This approach has been further investigated in the literature (c.f, e.g. . Leonard et al. (1996) and Lin and Rood (1996)). We are going to focus on the linear advection equation because in FV3, the horizontal dynamics are solved by using flux advection operators to compute the fluid density, absolute vorticity, and the kinetic energy (L. Harris et al., 2021; L. M. Harris & Lin, 2013; Lin & Rood, 1997; Putman, 2007). The boundary conditions are assumed to be periodic for simplicity.

To introduce the FV-SL schemes, we begin by discretizing the spatial and temporal domains into uniform grids. Subsequently, the FV-SL schemes involve three steps. The first step involves computing the departure points of the spatial grid edges. The second step, known as reconstruction, utilizes the grid cell average values to determine a piecewise function within each cell. This piecewise function approximates the values of the advected quantity and ensures the preservation of its local mass within each grid cell. The third step involves updating the fluxes at the grid edges by integrating the reconstruction function over a domain that extends from the departure point of the grid edge to the grid edge itself.

The first step of FV-SL schemes can be accomplished by integrating an ordinary differential equation backward in time. The second step is performed using the Piecewise-Parabolic Method (PPM) proposed by Colella and Woodward (1984). As the name suggests, PPM employs piecewise-parabolic functions. The third and final step is computed easily, as the reconstruction functions consist of parabolas that preserve the local mass.

It is worth noting that the reconstruction function can be constructed using functions other than parabolas. In fact, PPM can be seen as an extension of the Piecewise-Linear method proposed by Van Leer (1977), which, in turn, was inspired by the Piecewise-Constant method introduced by Godunov (1959). Additionally, other schemes inspired by PPM have been proposed in the literature utilizing higher-order polynomials, such as quartic polynomials (White & Adcroft, 2008). For a comprehensive review of general piecewise-polynomial reconstruction, we recommend referring to the technical report by Engwirda and Kelley (2016), Lauritzen et al. (2011), and the references therein.

The PPM approach has become popular in the literature for gas dynamics simulations, astrophysical phenomena modeling (Woodward, 1986), and later on atmospheric simulations (Carpenter et al., 1990). Indeed, PPM has been implemented in the FV3 dynamical core on its latitude-longitude grid (Lin, 2004) and cubed-sphere (Putman & Lin, 2007) versions. Although many other shapes for the basis functions and higher-order schemes are available in the literature, L. Harris et al. (2021) points out that the PPM scheme suits the needs of FV3 well. It is a flexible method that can be modified to ensure low diffusivity or shape preservation, for example. Additionally, a finite-volume numerical method usually requires monotonicity constraints, which, according to Godunov's order barrier theorem (Wesseling, 2001), limit the order of convergence to at most 1. Therefore, a higher-order scheme needs to strike a well-balanced trade-off between increasing computational cost and potential benefits.

This chapter begins with a basic review of one-dimensional advection equation in the integral form in Section 2.1. In Section 2.2, we establish the framework for general one-dimensional finite-volume Semi-Lagrangian schemes. Section 2.3 presents methods for computing the departure point. The PPM reconstruction is described in Section 2.4, while Subsection 2.4.2 introduces different approaches to ensure the monotonicity of parabolas. Section **??** focuses on the description and investigation of the PPM flux computation. Section **??** presents numerical results using the PPM scheme for the advection equation. Finally, Section **??** presents some concluding remarks. The application of PPM to solve two-dimensional problems will be addressed in Chapter **??**.

2.1 One-dimensional advection equation in integral form

2.1.1 Notation

Before introducing the FV-SL schemes, let us establish some notation by introducing the concepts of a Δx-grid, a Δt-temporal grid, and the $(\Delta x, \Delta t, \lambda)$-discretization, as well as the concept of grid function/winds. In this chapter, we will use the notation $\Omega = [a, b]$ to represent the interval under consideration, and ν to represent a non-negative integer indicating the number of ghost cell layers in each boundary. We also use the notations $\mathbb{R}^N_\nu := \mathbb{R}^{N+2\nu}$ and $\mathbb{R}^{N+1}_\nu := \mathbb{R}^{N+1+2\nu}$.

Definition 2.1 (Δx-grid)**.** *For a given interval Ω and a positive real number Δx such that $\Delta x = (b-a)/N$ for some positive integer N, we say that $\Omega_{\Delta x} = \{X_i\}_{i=-\nu+1}^{N+\nu}$ is a Δx-grid for Ω if*

$$X_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] = [a + (i-1)\Delta x, a + i\Delta x],$$

and $\Delta x = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$. *Each* X_i *is referred to as a control volume or cell, and* $x_{i-\frac{1}{2}}$ *and* $x_{i+\frac{1}{2}}$ *are the edges of the control volume* X_i. *The cell centroid is defined by*

$$x_i = \frac{1}{2}(x_{i+\frac{1}{2}} + x_{i-\frac{1}{2}}), \quad \forall i = -\nu+1, \dots, N+\nu,$$

and Δx *is the cell length.*

Remark 2.1. *If* $1 \leq i \leq N$, *we refer to* i *as an interior index; otherwise,* i *is considered a ghost cell index and we say the* X_i *is a ghost cell.*

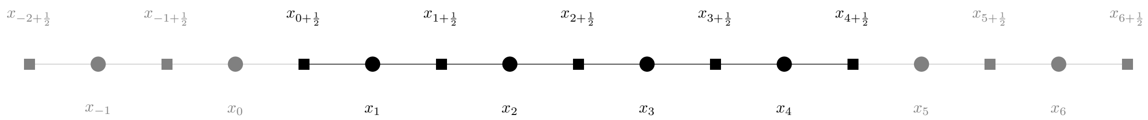

Figure 2.1: *Illustration of a* Δx*-grid with* $N = 4$ *cells in its interior (in black) and* $\nu = 2$ *ghost cell layers (in gray). The edges are denoted by squares and the cell centroids are denoted using circles.*

Definition 2.2 (Δt-temporal grid)**.** *For a given interval* $[0, T]$ *and a positive real number* Δt *such that* $\Delta t = T/N_T$ *for some positive integer* N_T, *we say that* $T_{\Delta T} = \{T_n\}_{n=0}^{N_T}$ *a* Δt*-temporal grid for* $[0, T]$ *if*

$$T_n = [t^n, t^{n+1}], \quad t^n = n\Delta t, \quad \Delta t = \frac{T}{N_T}, \quad \forall n = 0, \dots, N_T.$$

In this context, we also define $t^{n+\frac{1}{2}} = \frac{t^n + t^{n+1}}{2}$.

Definition 2.3 ($(\Delta x, \Delta t, \lambda)$-discretization)**.** *Given* $\Omega \times [0, T]$ *and positive real numbers* Δx *and* Δt, *we say that* $(\Omega_{\Delta x}, T_{\Delta t})$ *is a* $(\Delta x, \Delta t, \lambda)$*-discretization of* $\Omega \times [0, T]$ *if* $\Omega_{\Delta x}$ *is a* Δx*-grid for* Ω, $T_{\Delta t}$ *is a* Δt*-temporal grid for* $[0, T]$, *and* $\frac{\Delta t}{\Delta x} = \lambda$.

Remark 2.2. *Whenever we refer to a* Δx*-grid, a* Δt*-temporal grid, or a* $(\Delta x, \Delta t, \lambda)$*-discretization,* X_i, N, t^n, *and* N_T *are assumed to be implicitly defined.*

Next, we introduce the definitions of grid functions at cell centroids and edges.

Definition 2.4 (Δx-grid function)**.** *For a* Δx*-grid, we say that* Q *is a* Δx*-grid function if* $Q = (Q_{-\nu+1}, \dots, Q_{N+\nu}) \in \mathbb{R}^N_\nu$.

Definition 2.5 (Δx-grid wind)**.** *For a* Δx*-grid, we say that* u *is a* Δx*-grid wind if* $u = (u_{-\nu+\frac{1}{2}}, \dots, u_{N+\nu+\frac{1}{2}}) \in \mathbb{R}^{N+1}_\nu$.

The definition of a Δx-grid wind is based on the Arakawa grids (Arakawa & Lamb, 1977). Considering functions $q, u : \Omega \times [0, T] \to \mathbb{R}$ and a $(\Delta x, \Delta t, \lambda)$-discretization of $\Omega \times [0, T]$, we introduce the grid functions $q^n \in \mathbb{R}^N_\nu$ and $u^n \in \mathbb{R}^{N+1}_\nu$ where $q_i^n \approx q(x_i, t^n)$ and $u^n_{i+\frac{1}{2}} = u(x_{i+\frac{1}{2}}, t^n)$. The grid function q^n approximate the discrete values of q at cell centroids and u^n represents the values of u approximates edges, both for each time level t^n (Figure 2.2).

In this Chapter, our focus lies on periodic grid functions. We define a Δx-grid function

Q as periodic if it satisfies the following conditions:

$$
\begin{aligned}
Q_i &= Q_{N+i}, \quad i = -\nu + 1, \dots, 0, \\
Q_i &= Q_{i-N}, \quad i = N + 1, \dots, N + \nu.
\end{aligned}
$$

Similarly, we define a Δx-grid wind as periodic if it meets the following requirements:

$$
\begin{aligned}
u_{i-\frac{1}{2}} &= u_{N+i+\frac{1}{2}}, \quad i = -\nu, \dots, -1, \\
u_{i+\frac{1}{2}} &= u_{i+\frac{1}{2}-N}, \quad i = N + 1, \dots, N + \nu.
\end{aligned}
$$

We use the notation $\mathbb{P}_\nu^N$ and $\mathbb{P}_\nu^{N+1}$ to represent the spaces of periodic Δx-grid functions and winds, respectively.

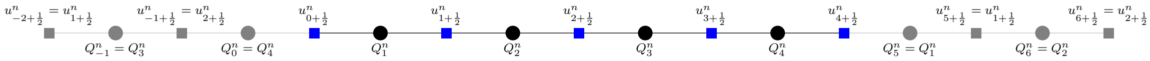

Figure 2.2: *Illustration of Δx-grid function Q (black circles) and a Δx-grid wind u (blue squares) and its ghost cell values (in gray) assuming periodicity.*

Given $Q \in \mathbb{P}_\nu^N$, we define the p-norm as

$$
\|Q\|_{p,\Delta x} = \begin{cases} \left(\sum_{i=1}^N |Q_i|^p \right)^{\frac{1}{p}} & \text{if } 1 \le p < \infty, \\ \max_{i=1,\dots,N} |Q_i| & \text{otherwise} , \end{cases} \tag{2.1}
$$

which is indeed a norm for periodic grid functions. Using a similar notation as in Engwirda and Kelley (2016), we define the stencil and a grid function evaluated on a stencil as follows.

Definition 2.6 (Stencil)**.** *For a Δx-grid, and each $i = 0, \dots, N$, we define a stencil as a set of the form $S_{i+\frac{1}{2}} = \{i - r + 1, \dots, i - 1, i, i + 1, \dots, i + s\} \subset \{-\nu + 1, \dots, N + \nu\}$.*

Definition 2.7 (Grid function restricted to a stencil)**.** *For a Δx-grid, a stencil $S_{i+\frac{1}{2}}$, and a Δx-grid function Q, we define $Q(S_{i+\frac{1}{2}}) = (Q_k)_{k \in S_{i+\frac{1}{2}}}$.*

These definitions provide the necessary notation for describing grid functions and their evaluations on stencils. To achieve a more compact notation in some situations, we introduce the centered difference notation:

$$
\delta_x g(x_i, t) = g(x_{i+\frac{1}{2}}, t) - g(x_{i-\frac{1}{2}}, t), \tag{2.2}
$$

for any function $g : \Omega \times [0, T] \to \mathbb{R}$. Additionally, we introduce the average value of q in the i-th control volume at time t, denoted as $Q_i(t)$, defined by:

$$
Q_i(t) = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} q(x, t)\, dx. \tag{2.3}
$$

Moreover, we define the Δx-grid function of average values as $Q(t) = (Q_i(t))_{i=-\nu+1}^{N+\nu}$. Here, $Q_i(t)$ represents the average value of q in the i-th control volume at time t.

For the consideration of periodic boundary conditions, we can define spaces of periodic functions over the interval Ω as follows:

$$\mathcal{S}_P(\Omega) = \{q : \mathbb{R} \times [0, +\infty[\to \mathbb{R} : q(x + b - a, t) = q(x, t), \quad \forall x \in \mathbb{R}, \quad t \geq 0\}.$$

Similarly, the space of k-times periodically differentiable functions $C_P^k(\Omega)$ can be defined as:

$$C_P^k(\Omega) = \mathcal{S}_P(\Omega) \cap C^k(\mathbb{R} \times [0, \infty[),$$

where $C^k(\mathbb{R}\times[0, +\infty[)$ denotes the space of functions that are k times continuously differentiable in both the spatial and temporal variables. In summary, $\mathcal{S}_P(\Omega)$ represents the space of periodic functions, and $C_P^k(\Omega)$ represents the space of k-times periodically differentiable functions over the interval Ω subject to periodic boundary conditions.

2.1.2 The 1D advection equation

In this section, we will derive the integral form of the 1D advection equation with periodic boundary conditions over the interval Ω. What is going to be presented here follows LeVeque (1990, 2002) closely. The advection equation with periodic boundary conditions in its differential form is given by:

$$\begin{cases} [\partial_t q + \partial_x(uq)](x, t) = 0, \quad \forall (x, t) \in \mathbb{R}\times]0, +\infty[, \\ q(a, t) = q(b, t), \quad \forall t \geq 0, \\ q_0(x) = q(x, 0), \quad \forall x \in \Omega. \end{cases} \tag{2.4}$$

Here, $q \in C_P^1(\Omega)$ represents the advected quantity, and $u \in C_P^1(\Omega)$ represents the velocity. We will focus on Equation (2.4) over the domain $D = \Omega \times [0, T]$, where $T > 0$ is a finite time. A strong or classical solution to the advection equation is defined as a function $q \in C_P^1(\Omega)$ and satisfies Equation (2.4). In order to deduce the integral form of Equation (2.4), we consider $[x_1, x_2] \times [t_1, t_2] \subset D$. Integrating Equation (2.5) over $[x_1, x_2]$, we obtain:

$$\frac{d}{dt} \int_{x_1}^{x_2} q(x, t)\, dx = -((uq)(x_2, t) - (uq)(x_1, t)), \tag{2.5}$$

and integrating Equation (2.5) over $[t_1, t_2]$, we get

$$\int_{x_1}^{x_2} q(x, t_2)\, dx = \int_{x_1}^{x_2} q(x, t_1) - \left(\int_{t_1}^{t_2} (uq)(x_2, t)\, dt - \int_{t_1}^{t_2} (uq)(x_1, t)\, dt \right). \tag{2.6}$$

The presented problem, Problem 2.1, aims to find a solution, called weak solution, to the advection equation in its integral form, considering the given initial condition q_0 and velocity function u.

Problem 2.1. *Given an initial condition q_0 and a velocity function u we would like to find a weak solution q of the advection equation in the integral form:*

$$\int_{x_1}^{x_2} q(x, t_2)\, dx = \int_{x_1}^{x_2} q(x, t_1)\, dx + \int_{t_1}^{t_2} (uq)(x_1, t)\, dt - \int_{t_1}^{t_2} (uq)(x_2, t)\, dt,$$

$\forall[x_1, x_2] \times [t_1, t_2] \subset \Omega \times [0, T]$, *and* $q(x, 0) = q_0(x)$, $\forall x \in \Omega$, $q(a, t) = q(b, t)$, $\forall t \in [0, T]$.

We point out that, for Problem 2.1, the total mass in Ω at time t defined by:

$$M_{[a,b]}(t) = \int_a^b q(x, t)\, dx,$$

remains constant over time, i.e.,

$$M_{[a,b]}(t) = M_{[a,b]}(0), \quad \forall t \in [0, T].$$

This conservation of total mass property is highly desirable for numerical schemes aiming to approximate general conservation law solutions accurately.

Applying the steps from Equation (2.4) to Equation (2.6) in reverse order, one can verify that if q is a weak solution and $q \in C_P^1(\Omega)$, then it satisfies Equation (2.4). Therefore, Equation (2.4) and Problem (2.1) are equivalent when $q \in C_P^1(\Omega)$. However, Problem (2.1) can be formulated for functions that are not C^1 and have discontinuities. In fact, Problem (2.1) only requires that q and uq are locally integrable.

It is worth noting that Equation (2.6) holds for all x_1, x_2, t_1, and t_2 such that $[x_1, x_2] \times [t_1, t_2] \subset D$. Therefore, let us consider a $(\Delta x, \Delta t, \lambda)$-discretization of D and rewrite Equation (2.6) in terms of this discretization. By replacing t_1, t_2, x_1, and x_2 with $t^n, t^{n+1}, x_{i-\frac{1}{2}}$, and $x_{i+\frac{1}{2}}$, respectively, in Equation (2.6), we obtain:

$$Q_i(t^{n+1}) = Q_i(t^n) - \frac{1}{\Delta x}\left(\int_{t^n}^{t^{n+1}} (uq)(x_{i+\frac{1}{2}}, t)\, dt - \int_{t^n}^{t^{n+1}} (uq)(x_{i-\frac{1}{2}}, t)\, dt \right), \quad \forall i = 1, \dots, N, \quad \forall n = 0, \dots, N_T - 1. \tag{2.7}$$

To achieve a more compact notation, we use the centered difference notation and then Equation (2.7) can be rewritten as:

$$Q_i(t^{n+1}) = Q_i(t^n) - \frac{1}{\Delta x}\delta_x \left(\int_{t^n}^{t^{n+1}} (uq)(x_i, t)\, dt \right), \quad \forall i = 1, \dots, N, \quad \forall n = 0, \dots, N_T - 1. \tag{2.8}$$

Now we can define a discretized version of Problem 2.1 as Problem 2.2.

Problem 2.2. *Let us consider the framework of Problem 2.1 and a* $(\Delta x, \Delta t, \lambda)$*-discretization of* $\Omega \times [0, T]$. *Since we are operating within the framework of Problem 2.1, the following relationship holds:*

$$Q_i(t^{n+1}) = Q_i(t^n) - \lambda \delta_x \left(\frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} (uq)(x_i, t)\, dt \right), \quad \forall i = 1, \dots, N, \quad \forall n = 0, \dots, N_T - 1, \tag{2.9}$$

where $Q_i(t) = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} q(x, t)\, dx$. *Our objective now is to determine the values* $Q_i(t^n)$, $\forall i = 1, \dots, N$, $\forall n = 0, \dots, N_T - 1$, *given the initial values* $Q_i(0)$, $\forall i = 1, \dots N$. *In other words, we aim to find the average values of* q *in each control volume* X_i *at the specified time instances.*

It is important to note that no approximations have been made in problems (2.1) and

(2.2). In Equation (2.9), we divided and multiplied by Δt to interpret $\frac{1}{\Delta t}\int_{t^n}^{t^{n+1}}(uq)(x_{i\pm\frac{1}{2}},t)\,dt$ as a time-averaged flux. This interpretation is useful for deriving finite-volume schemes.

In Problem 2.2, we need to approximate the time-averaged flux at the cell edges $x_{i\pm\frac{1}{2}}$ to derive a finite-volume scheme. This flux, in principle, requires knowledge of q over the entire interval $[t^n, t^{n+1}]$. To overcome this, we can express the temporal integral as a spatial integral at time t^n. This approach avoids the need for information about q throughout the entire interval $[t^n, t^{n+1}]$. Furthermore, this spatial integral domain is closely related to the definition of the departure point.

To introduce the definition of departure point, for each $s \in [t^n, t^{n+1}]$, we consider the following Cauchy problem backward in time:

$$\begin{cases} \partial_t x^d_{i+\frac{1}{2}}(t,s) = u(x^d_{i+\frac{1}{2}}(t,s),t), \quad t \in [t^n, s] \\ x^d_{i+\frac{1}{2}}(s,s) = x_{i+\frac{1}{2}}. \end{cases} \tag{2.10}$$

The point $x^d_{i+\frac{1}{2}}(t^n,s)$ is called departure point at time t^n of the point $x_{i+\frac{1}{2}}$ at time s. In Figure 2.3 we illustrate the departure point idea.

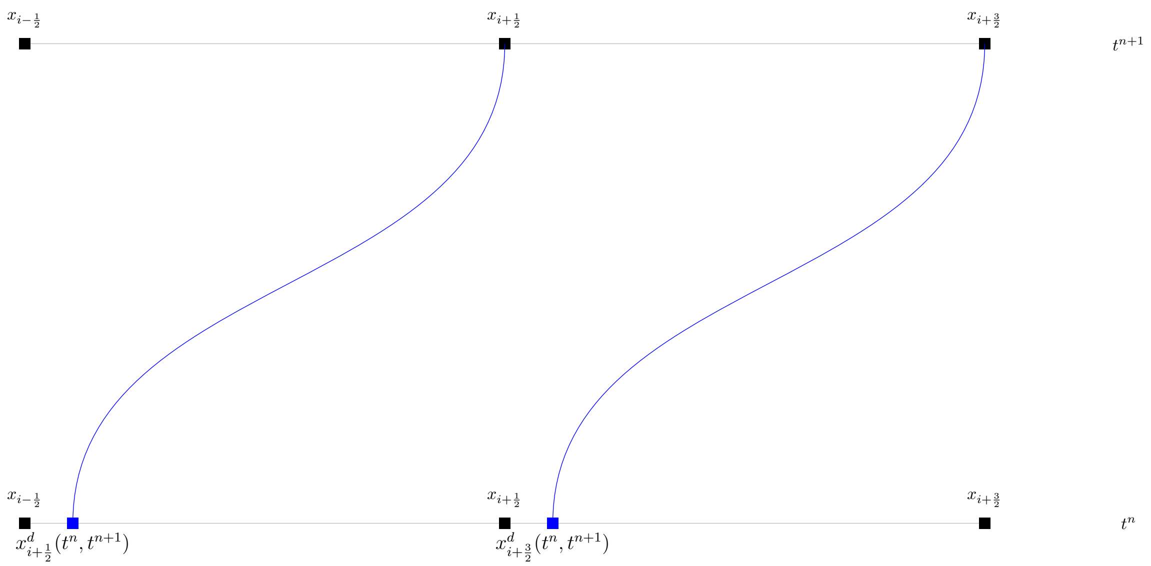

Figure 2.3: *Illustration of the departure point of the cell edges from time t^{n+1} to t^n.*

Integrating Equation (2.10) over the interval $[t,s]$, we get:

$$x^d_{i+\frac{1}{2}}(t,s) = x_{i+\frac{1}{2}} - \int_t^s u(x^d_{i+\frac{1}{2}}(\theta,s),\theta)\,d\theta. \tag{2.11}$$

In the following Proposition, we show how the time-averaged flux is related to a spatial integral over a interval depending on departure points.

Proposition 2.1. *Assume the framework of Problem 2.2. If q and u are C^1 functions, then:*

$$\int_{t^n}^{t^{n+1}}(uq)(x_{i+\frac{1}{2}},s)\,ds = \int_{x^d_{i+\frac{1}{2}}(t^n,t^{n+1})}^{x_{i+\frac{1}{2}}} q(x,t^n)\,dx \tag{2.12}$$

Proof. Using the Leibniz rule for integration (Theorem A.2 with $f(s,\theta) = u\big(x^d_{i+\frac{1}{2}}(\theta,s)\big)$), in Equation (2.11), it follows that:

$$\begin{aligned}\partial_s x^d_{i+\frac{1}{2}}(t,s) &= -\left(u(x_{i+\frac{1}{2}},s) + \int_t^s \frac{du}{ds}\big(x^d_{i+\frac{1}{2}}(\theta,s),\theta\big)\,d\theta \right) \\ &= -u(x_{i+\frac{1}{2}},s) - \int_t^s \partial_s u\big(x^d_{i+\frac{1}{2}}(\theta,s),\theta\big)\partial_s x^d_{i+\frac{1}{2}}(\theta,s)\,d\theta.\end{aligned} \tag{2.13}$$

Taking the derivative with respect to t of Equation (2.13), we have:

$$\partial_t\partial_s x^d_{i+\frac{1}{2}}(t,s) = \partial_x u\big(x^d_{i+\frac{1}{2}}(t,s),t\big)\partial_s x^d_{i+\frac{1}{2}}(t,s). \tag{2.14}$$

Using standard ordinary differential equations techniques (ODE), we get that $x^d_{i+\frac{1}{2}}$ that solves Equations (2.13) and (2.14) is given by:

$$\partial_s x^d_{i+\frac{1}{2}}(t,s) = -\exp\left(\int_t^s \partial_x u\big(x^d_{i+\frac{1}{2}}(\theta,s),\theta\big)\,d\theta \right) u(x_{i+\frac{1}{2}},s). \tag{2.15}$$

Computing q on the trajectory give by $x^d_{i+\frac{1}{2}}(t,s)$ and taking its time derivative, we obtain:

$$\begin{aligned}\frac{d}{dt}q\big(x^d_{i+\frac{1}{2}}(t,s),t\big) &= \partial_t q\big(x^d_{i+\frac{1}{2}}(t,s),t\big) + (u\partial_x q)\big(x^d_{i+\frac{1}{2}}(t,s),t\big) \\ &= -\partial_x u\big(x^d_{i+\frac{1}{2}}(t,s),t\big) q\big(x^d_{i+\frac{1}{2}}(t,s),t\big),\end{aligned} \tag{2.16}$$

where we used that q satisfies the linear advection equation on its differential (2.4) form and that $x^d_{i+\frac{1}{2}}(t,s)$ solves Equation (2.10). Using again standard ODE techniques, we get that q that solves Equation (2.16) is given by:

$$q\big(x^d_{i+\frac{1}{2}}(t,s),t\big) = \exp\left(-\int_t^s \partial_x u\big(x^d_{i+\frac{1}{2}}(\theta,s),\theta\big)\,d\theta \right) q(x_{i+\frac{1}{2}},s). \tag{2.17}$$

Notice that if u does not depend on x, then q is constant along the trajectory $x^d_{i+\frac{1}{2}}(t,s)$.

Let us consider the mapping $s \in [t^n, t^{n+1}] \to x^d_{i+\frac{1}{2}}(t^n,s)$. Integrating q over all departure points at time t^n from $x_{i+\frac{1}{2}}$ at time s, we have

$$\int_{x^d_{i+\frac{1}{2}}(t^n,t^n)=x_{i+\frac{1}{2}}}^{x^d_{i+\frac{1}{2}}(t^n,t^{n+1})} q(x,t^n)\,dx = \int_{t^n}^{t^{n+1}} q\big(x^d_{i+\frac{1}{2}}(t^n,s),t^n\big)\partial_s x^d_{i+\frac{1}{2}}(t^n,s)\,ds, \tag{2.18}$$

where we are just using the variable change integration formula. Then, it follows from Equations (2.15) and (2.17) with $t = t^n$ that:

$$\int_{x_{i+\frac{1}{2}}}^{x^d_{i+\frac{1}{2}}(t^n,t^{n+1})} q(x,t^n)\,dx = -\int_{t^n}^{t^{n+1}} (uq)(x_{i+\frac{1}{2}},s)\,ds, \tag{2.19}$$

which is the desired formula. □

With the aid of Proposition 2.1, we can rewrite Problem 2.2 in terms of the departure point, avoiding the need for knowledge about q over the entire interval $[t^n, t^{n+1}]$. This is described in Problem 2.3:

Problem 2.3. *Assume the framework of Problem 2.1 and a $(\Delta x, \Delta t, \lambda)$-discretization of $\Omega \times [0, T]$. Since we are in the framework of Problem 2.1, it follows that:*

$$Q_i(t^{n+1}) = Q_i(t^n) - \lambda\left(\frac{1}{\Delta t}\int_{X(t^n,t^{n+1};x_{i+\frac{1}{2}})}^{x_{i+\frac{1}{2}}} q(x,t^n)\,dx - \frac{1}{\Delta t}\int_{x^d_{i-\frac{1}{2}}(t^n,t^{n+1})}^{x_{i-\frac{1}{2}}} q(x,t^n)\,dx\right), \qquad \forall i = 1,\dots,N, \quad \forall n = 0,\dots,N_T-1, \tag{2.20}$$

where $Q_i(t) = \frac{1}{\Delta x}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} q(x,t)\,dx$. Our problem now consists of finding the values $Q_i(t^n)$, $\forall i = 1,\dots,N$, $\forall n = 0,\dots,N_T-1$, given the initial values $Q_i(0)$, $\forall i = 1,\dots N$. In other words, we would like to find the average values of q in each control volume X_i at the considered time instants.

At each time step t^n, we compute the values of $Q_i(t^{n+1})$ based on $Q_i(t^n)$ and the integrals of $q(x,t^n)$ over specific intervals. These intervals are defined by the departure points $X(t^n, t^{n+1}; x_{i+\frac{1}{2}})$ and $X(t^n, t^{n+1}; x_{i-\frac{1}{2}})$. To perform the computations, we need to determine the departure points from the edges of all control volumes and calculate the required integrals. This idea serves as the motivation for defining finite-volume Semi-Lagrangian schemes. These schemes involve estimating the departure points and reconstructing the function q at time t^n using its average values $Q_i(t^n)$, which enables us to compute the necessary integrals.

2.2 The finite-volume Semi-Lagrangian approach

Finally, we define the 1D FV-SL scheme problem as follows in Problem 2.3.

Problem 2.4 (1D FV-SL scheme)**.** *Assume the framework defined in Problem 2.3. The finite-volume Semi-Lagrangian approach of Problem 2.3 consists of finding a scheme of the form:*

$$Q_i^{n+1} = Q_i^n - \lambda(F^n_{i+\frac{1}{2}} - F^n_{i-\frac{1}{2}}), \quad \forall i = 1,\dots,N, \quad \forall n = 0,\dots,N_T-1, \tag{2.21}$$

where $Q^n \in \mathbb{P}^N_\nu$ is intended to be an approximation of $Q(t^n) \in \mathbb{P}^N_\nu$ in some sense. We define $Q_i^0 = Q_i(0)$ or $Q_i^0 = q_i^0$. The terms $F^n_{i\pm\frac{1}{2}}$ are known as numerical flux and are given by

$$F^n_{i\pm\frac{1}{2}} = \frac{1}{\Delta t}\int_{\tilde{x}^n_{i\pm\frac{1}{2}}}^{x_{i\pm\frac{1}{2}}} \tilde{q}(x; Q^n)\,dx, \tag{2.22}$$

where $\tilde{x}^n_{i\pm\frac{1}{2}}$ is an estimate of the departure point $x^d_{i-\frac{1}{2}}(t^n, t^{n+1})$, and $\tilde{q}$ is a reconstruction function for q built with the values Q^n. Thus, $F^n_{i\pm\frac{1}{2}}$ approximates $\frac{1}{\Delta t}\int_{x^d_{i\pm\frac{1}{2}}(t^n,t^{n+1})}^{x_{i\pm\frac{1}{2}}} q(x,t^n)\,dx$.

For a 1D FV-SL the discrete total mass at the time-step n is given by

$$M^n = \Delta x \sum_{i=1}^{N} Q_i^n. \tag{2.23}$$

Therefore, the discrete total mass is constant for a 1D-FV scheme, which follows from a straightforward computation:

$$M^{n+1} = \Delta x \sum_{i=1}^{N} Q_i^{n+1} = M^n - \Delta t \sum_{i=1}^{N} (F^n_{i+\frac{1}{2}} - F^n_{i-\frac{1}{2}}) = M^n - \Delta t (F^n_{N+\frac{1}{2}} - F^n_{\frac{1}{2}}) = M^n,$$

where we are using that $F^n_{N+\frac{1}{2}} = F^n_{\frac{1}{2}}$, since we are assuming periodic boundary conditions.

We would like to highlight an important relationship between the average values of q and its values at the cell centroids. In Problem 2.4, we mentioned that the initial condition can be represented as q_i^0 instead of $Q_i(0)$. Moreover, when analyzing the convergence of a FV-SL scheme, it is useful to compare Q_i^n with q_i^n since computing $Q_i(t^n)$ requires evaluating an analytical integral, which can be challenging in certain cases. In Proposition 2.2, we provide a simple proof that q_i^n approximates $Q_i(t^n)$ with second-order error when q is twice continuously differentiable.

Proposition 2.2. *If $q \in C_P^2(\Omega)$, then $Q_i(t^n) - q_i^n = C_1 \Delta x^2$, where $C_1 = \frac{1}{24}\frac{\partial^2 q}{\partial x^2}(\eta, t^n)$, $\eta \in X_i$.*

Proof. Just apply Theorem A.4 for the function $q(x, t^n)$. □

Hence, 1D FV-SL schemes may be conceptualized as schemes that update the centroid values. The Problem of the convergence of 1D FV-SL schemes is addressed in Section A.3. Now we are going to address the problem of the departure point estimation and the reconstruction problem.

2.3 Departure point computation

2.3.1 RK1 scheme

Equation (2.11) enables us to compute or estimate the departure point. For instance, if u is constant, the departure point at time t^n for the point $x_{i+\frac{1}{2}}$ at time t^{n+1} is given by:

$$x^d_{i+\frac{1}{2}}(t^n, t^{n+1}) = x_{i+\frac{1}{2}} - u\Delta t. \tag{2.24}$$

In general, the estimated departure point, denoted by $\tilde{x}^n_{i+\frac{1}{2}}$, takes the form:

$$\tilde{x}^n_{i+\frac{1}{2}} = x_{i+\frac{1}{2}} - \tilde{u}^n_{i+\frac{1}{2}} \Delta t, \tag{2.25}$$

where $\tilde{u}^n_{i+\frac{1}{2}}$ represents the time-averaged wind and approximates:

$$\frac{1}{\Delta t}\int_{t^n}^{t^{n+1}} u(x^d_{i+\frac{1}{2}}(\theta, t^{n+1}), \theta)\, d\theta. \tag{2.26}$$

The departure point $\tilde{x}^n_{i+\frac{1}{2}}$ is said to be p-order accurate if there exists a constant C that does not depend on Δt, such that:

$$x^d_{i+\frac{1}{2}}(t^n, t^{n+1}) - \tilde{x}^n_{i+\frac{1}{2}} = O(\Delta t^p). \tag{2.27}$$

One possible way of estimating the time-averaged wind is by using:

$$\tilde{u}^n_{i+\frac{1}{2}} = u^{n+\frac{1}{2}}_{i+\frac{1}{2}}, \tag{2.28}$$

as in FV3 papers (Lin & Rood, 1996; Putman & Lin, 2007). For simplicity, in this chapter, we shall assume that the wind is known for all time instants needed. This scheme will be referred to as **RK1**. In FV3, the wind is at time level $n + \frac{1}{2}$ is obtained by solving the horizontal dynamics on a C-grid as an intermediate step (Lin, 2004; Lin & Rood, 1997). Our objective now is to determine the value of p in Equation (2.27) in the following proposition. It is useful to introduce the concept of a material derivative beforehand:

$$\frac{Dh}{Dt} = \frac{\partial h}{\partial t} + u\frac{\partial h}{\partial x},$$

where h is a function belonging to C^1.

Proposition 2.3. *If $u \in C^1$ and the time-averaged wind is computed using Equation (2.28), then the departure point from Equation (2.25) satisfies:*

$$x^d_{i+\frac{1}{2}}(t^n, t^{n+1}) - \tilde{x}^n_{i+\frac{1}{2}} = O(\Delta t^2), \tag{2.29}$$

for a constant C that depends on u.

Proof. Using the midpoint rule (Theorem A.4) for the function $f(t) = u\big(x^d_{i+\frac{1}{2}}(t, t^{n+1}), t\big)$ in Equation (2.11), we obtain:

$$x^d_{i+\frac{1}{2}}(t^n, t^{n+1}) = x_{i+\frac{1}{2}} - u\Big(x^d_{i+\frac{1}{2}}(t^{n+\frac{1}{2}}, t^{n+1}), t^{n+\frac{1}{2}}\Big)\Delta t - \frac{1}{24}\frac{D^2 u}{Dt^2}\Big(x^d_{i+\frac{1}{2}}(\theta_1, t^{n+1}), \theta_1\Big)\Delta t^2, \tag{2.30}$$

for $\theta_1 \in [t^n, t^{n+1}]$. Now observe that, from the intermediate value theorem for integrals and Equation (2.11), we have

$$x^d_{i+\frac{1}{2}}(t^{n+\frac{1}{2}}, t^{n+1}) = x_{i+\frac{1}{2}} - \frac{\Delta t}{2}u\Big(x^d_{i+\frac{1}{2}}(\theta_2, t^{n+1}), \theta_2\Big)$$

for $\theta_2 \in [t^{n+\frac{1}{2}}, t^{n+1}]$. Combining this with a Taylor's expansion of $u\Big(x^d_{i+\frac{1}{2}}(t, t^{n+1}), t^{n+\frac{1}{2}}\Big)$ for

$t = t^{n+\frac{1}{2}}$ we have:

$$u\left(x^d_{i+\frac{1}{2}}(t^{n+\frac{1}{2}}, t^{n+1}), t^{n+\frac{1}{2}}\right) = u^{n+\frac{1}{2}}_{i+\frac{1}{2}} - \left(u\frac{\partial u}{\partial x}\right)\left(x_{i+\frac{1}{2}}(\theta_3, t^{n+1}), t^{n+\frac{1}{2}}\right)\right)u\left(x^d_{i+\frac{1}{2}}(\theta_2, t^{n+1}), \theta_2\right)\frac{\Delta t^2}{2}, \tag{2.31}$$

for $\theta_3 \in [t^n, t^{n+1}]$. Substituting Equation (2.31) into Equation (2.30), we obtain the desired estimate. □

2.3.2 RK2 scheme

Before presenting a higher-order estimate for the departure point, let us recall the definition of the CFL number.

Definition 2.8. *For Problem 2.4, the CFL number at an edge $x_{i+\frac{1}{2}}$ and at a time level t^n is defined by*

$$c^n_{i+\frac{1}{2}} = \frac{\Delta t}{\Delta x} u^n_{i+\frac{1}{2}}. \tag{2.32}$$

The CFL number is the maximum of the values $c^n_{i+\frac{1}{2}}$. The problem of estimating the departure point is very common in Semi-Lagrangian schemes, which are quite popular in atmospheric modeling. For a review of departure point calculation methods, we refer to Tumolo (2011, Chapter 3) and the references therein. There are different approaches to compute the departure point, such as integrating the ODE from Equation 2.1 using different time integrators (D. Durran, 2011) backward in time. The Runge-Kutta methods are a possible choice to compute the departure point (*cf. e.g.* Guo et al. (2014), Lu et al. (2022)). In this work, we shall consider a second-order Runge-Kutta method to compute the departure point, which we express in terms of $\tilde{u}^n_{i+\frac{1}{2}}$ using the following equations (D. R. Durran, 2010):

$$\begin{aligned} \tilde{x}^{n+\frac{1}{2}}_{i+\frac{1}{2}} &= x_{i+\frac{1}{2}} - u^n_{i+\frac{1}{2}}\frac{\Delta t}{2} = x_{i+\frac{1}{2}} - c^n_{i+\frac{1}{2}}\frac{\Delta x}{2}, \\ \tilde{u}^n_{i+\frac{1}{2}} &= u\left(\tilde{x}^{n+\frac{1}{2}}_{i+\frac{1}{2}}, t^n + \frac{\Delta t}{2}\right). \end{aligned} \tag{2.33}$$

Notice that this scheme requires values of u at points that are not grid points, both in space. We overcome this using linear interpolation in space:

$$\tilde{u}^n_{i+\frac{1}{2}} = \begin{cases} \left(1-\alpha^n_{i+\frac{1}{2}}\right)u^{n+\frac{1}{2}}_{i+\frac{1}{2}-k} + \alpha^n_{i+\frac{1}{2}}u^{n+\frac{1}{2}}_{i-\frac{1}{2}-k} & \text{if } u^n_{i+\frac{1}{2}} \geq 0, \\ \alpha^n_{i+\frac{1}{2}}u^{n+\frac{1}{2}}_{i+\frac{3}{2}-k} + \left(1-\alpha^n_{i+\frac{1}{2}}\right)u^{n+\frac{1}{2}}_{i+\frac{1}{2}-k} & \text{if } u^n_{i+\frac{1}{2}} < 0, \end{cases} \tag{2.34}$$

where $\frac{c^n_{i+\frac{1}{2}}}{2} = \alpha^n_{i+\frac{1}{2}} + k$, $k = \lfloor \frac{c^n_{i+\frac{1}{2}}}{2} \rfloor$, $\alpha^n_{i+\frac{1}{2}} \in [0, 1[$, and $\lfloor \cdot \rfloor$ is the floor function. This scheme leads to a third-order error in the departure point estimate (see *e.g.* D. R. Durran (2010, Section 7.1.2)). This scheme shall be referred to as **RK2**. Notice that for this scheme, we need ghost values for the velocity, depending on how large the CFL number is. In particular, if the CFL number is less than 2, then $k = 0$ and we need the ghost values $u^n_{-1+\frac{1}{2}}$ and $u^n_{N+\frac{3}{2}}$.

2.4 Reconstruction

In this section, we will review the Piecewise-Parabolic Method (PPM). The analysis of its accuracy will be presented in Section A.6. PPM was originally proposed by Colella and Woodward (1984) for gas dynamic simulations, and its applicability to atmospheric simulations has been demonstrated by Carpenter et al. (1990). This method is based on utilizing parabolas to reconstruct the function using its average values, ensuring both mass conservation and monotonicity. PPM is an extension of the Piecewise-Linear Method introduced by Van Leer (1977), and it is implemented in the FV3 model using the dimension splitting method developed by Lin and Rood (1996).

Let's consider a function q defined in $\Omega = [a, b]$ and a Δx-grid covering Ω. We assume that we are given the average values $Q_i = \frac{1}{\Delta x}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} q(x)\,dx$ for each control volume X_i, where $i = 1, \dots, N$. In this context, it is convenient to define the Δx-grid function $Q \in \mathbb{P}_\nu^N$ with the entries given by Q_i. To facilitate the discussion, we introduce the indicator function $\chi_i(x)$ for each control volume X_i, defined as:

$$\chi_i(x) = \begin{cases} 1 & \text{if } x \in X_i, \\ 0 & \text{otherwise.} \end{cases}$$

Drawing inspiration from Stoer and Bulirsch (2002, Chapter 1), we consider a family of functions $\Phi(\xi; \mu)$ defined for $\xi \in [0, 1]$, depending on a parameter $\mu = (\mu_0, \mu_1, \dots, \mu_d) \in \mathbb{R}^{d+1}$. The reconstruction problem involves finding a piecewise function:

$$\tilde{q}(x; Q) = \sum_{i=1}^{N} \chi_i(x) q_i(x; Q), \tag{2.35}$$

where $q_i(x; Q) = \Phi\big(\frac{x - x_{i-\frac{1}{2}}}{\Delta x}; \alpha_i\big)$ and $\alpha_i = (\alpha_{i0}, \alpha_{i1}, \dots \alpha_{id}) \in \mathbb{R}^{d+1}$. It is required that:

$$\frac{1}{\Delta x}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \tilde{q}(x; Q)\,dx = \frac{1}{\Delta x}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} q_i(x; Q)\,dx = \int_0^1 \Phi(\xi; \alpha_i)\,d\xi = Q_i,$$

which means that $q_i(x; Q)$ preserves the mass within each control volume X_i.

Notice that, given $q_i(x; Q) = \Phi\big(\frac{x - x_{i-\frac{1}{2}}}{\Delta x}; \alpha_i\big)$, it is reasonable to expect that $\Phi(0; \alpha_i)$ approximates $q_i(x_{i-\frac{1}{2}})$ and $\Phi(1; \alpha_i)$ approximates $q_i(x_{i+\frac{1}{2}})$. Additionally, if both q and Φ are sufficiently differentiable, $\Phi^{(l)}(0; \alpha_i)$ should approximate $(\Delta x)^l q^{(l)}(x_{i-\frac{1}{2}})$ and $\Phi^{(l)}(1; \alpha_i)$ should approximate $(\Delta x)^l q^{(l)}(x_{i+\frac{1}{2}})$, provided these derivatives exist.

One approach to estimating these values at the edges $x_{i+\frac{1}{2}}$ using the average values Q is by employing a reconstruction method based on primitive functions (LeVeque, 2002, Chapter 17). It is worth noting that if we define:

$$Q(x) = \int_a^x q(\xi)\,d\xi, \tag{2.36}$$

we have $Q^{(l)}(x) = q^{(l-1)}(x)$. Specifically, $Q^{(l)}(x_{i+\frac{1}{2}}) = q^{(l-1)}(x_{i+\frac{1}{2}})$ and $Q(x_{i+\frac{1}{2}}) = \Delta x \sum_{k=1}^{i} Q_k$, for all $i = 0, \ldots, N$. Therefore, we can employ finite-difference schemes to estimate $q^{(l-1)}(x_{i+\frac{1}{2}})$ using the Δx-grid function Q, given that it is assumed to be known.

Let us assume that the l-th derivative of Q at $x_{i+\frac{1}{2}}$ is approximated using a stencil $\mathcal{S}^{(l)}_{i+\frac{1}{2}}$ and weights $\beta^{(l)}_{k,i}$, where $k \in \mathcal{S}^{(l)}_{i+\frac{1}{2}}$. When d is odd, we can seek a parameter $\alpha_i \in \mathbb{R}^{d+1}$ that ensures mass conservation and approximates q and its derivatives at the edges by solving the following system:

$$\begin{cases} \int_0^1 \Phi(\xi;\alpha_i)\, d\xi & = Q_i, \\ \Phi^{(l)}(0;\alpha_i) & = (\Delta x)^l \sum_{k\in\mathcal{S}^{(l)}_{i-\frac{1}{2}}} \beta^{(l)}_{k,i} Q_k, \quad \text{for } l = 0, \ldots, d-1. \end{cases} \tag{2.37}$$

If d is even, similarly we look for a parameter $\alpha_i \in \mathbb{R}^{d+1}$ that solves:

$$\begin{cases} \int_0^1 \Phi(\xi;\alpha_i)\, d\xi & = Q_i, \\ \Phi^{(l)}(0;\alpha_i) & = (\Delta x)^l \sum_{k\in\mathcal{S}^{(l)}_{i-\frac{1}{2}}} \beta^{(l)}_{k,i} Q_k, \quad \text{for } l = 0, \ldots, \frac{d}{2}-1, \\ \Phi^{(l)}(1;\alpha_i) & = (\Delta x)^l \sum_{k\in\mathcal{S}^{(l)}_{i+\frac{1}{2}}} \beta^{(l)}_{k,i} Q_k, \quad \text{for } l = 0, \ldots, \frac{d}{2}-1. \end{cases} \tag{2.38}$$

The reconstruction problem becomes linear when $\Phi(\xi;\mu)$ can be expressed as:

$$\Phi(\xi;\mu) = \sum_{k=0}^{d} \mu_k \Phi_k(\xi),$$

where Φ_k are functions defined on $[0,1]$. In this case, Equation (2.37) and Equation (2.38) form $(d+1)\times(d+1)$ linear systems. It is common to assume that the Φ_k's are linearly independent. Therefore, we have described a method that allows us to reconstruct a function from its average values, preserving its mass in each control volume, and approximating q at the edges. This method works for functions Φ_k as long as they are sufficiently differentiable. For example, choosing $d = 0$ and $\Phi_0(\xi) = 1$ gives us piecewise constant functions, as used in Godunov (1959). If we choose $d = 1$, $\Phi_0(\xi) = 1$, and $\Phi_1(\xi) = \xi$, we obtain a piecewise linear reconstruction, similar to Van Leer (1977). For polynomial reconstruction schemes, we refer to Engwirda and Kelley (2016) and the references therein.

2.4.1 The Piecewise-Parabolic Method

Hereafter, we are going the focus on the piecewise parabolic method from Colella and Woodward (1984) that uses $d = 2$, $\Phi_0(\xi) = 1$, $\Phi_1(\xi) = \xi$, $\Phi_1(\xi) = (1-\xi)\xi$. In order to follow the notation from Colella and Woodward (1984), we write $\alpha_{0i} = q_{L,i}$, $\alpha_{1i} = \Delta q_i$ and $\alpha_{2i} = q_{6,i}$. Therefore, each q_i may be expressed as:

$$q_i(x;Q) = q_{L,i} + \Delta q_i z_i(x) + q_{6,i} z_i(x)(1 - z_i(x)), \quad \text{where } z_i(x) = \frac{x - x_{i-\frac{1}{2}}}{\Delta x}, \quad x \in X_i, \tag{2.39}$$

where the values $q_{L,i}$, Δq_i and $q_{6,i}$ will be specified latter. Note that each z_i is just a normalization function that maps X_i onto $[0,1]$. It is easy to see that $\lim_{x\to x^+_{i-\frac{1}{2}}} q_i(x;Q) = q_{L,i}$. If we define $q_{R,i} = \lim_{x \to x^-_{i+\frac{1}{2}}} q_i(x; Q)$, then we have:

$$\Delta q_i = q_{R,i} - q_{L,i}. \tag{2.40}$$

The average value of q_i is given by:

$$\frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} q_i(x; Q)\, dx = \frac{(q_{L,i} + q_{R,i})}{2} + \frac{q_{6,i}}{6}. \tag{2.41}$$

Under the hypothesis of mass conservation, we have:

$$q_{6,i} = 6\left(Q_i - \frac{(q_{L,i} + q_{R,i})}{2} \right). \tag{2.42}$$

Therefore, we have found the parameters Δq_i and $q_{6,i}$ as functions of the parameters $q_{L,i}$ and $q_{R,i}$, such that the parabola q_i from (2.35) guarantees mass conservation. To completely determine the parabola q_i, we need to set the values $q_{L,i}$ and $q_{R,i}$, which, as we have seen, represent the limits of q_i when x tends to the left and right boundaries of X_i, respectively. Hence, it is natural to seek for $q_{L,i}$ as an approximation of $q(x_{i-\frac{1}{2}})$ and $q_{R,i}$ as an approximation of $q(x_{i+\frac{1}{2}})$. As we mentioned before in after introducing Equation (2.36), this is achieved using finite-differences. An explicit expression for the approximation of $q(x_{i-\frac{1}{2}})$, denoted by $q_{i+\frac{1}{2}}$, is given by (Colella & Woodward, 1984):

$$q_{i+\frac{1}{2}} = \frac{1}{2}\left(Q_{i+1} + Q_i \right) - \frac{1}{6}\left(\delta Q_{i+1} - \delta Q_i \right), \tag{2.43}$$

where δQ_i is the average slope in the i-th control-volume:

$$\delta Q_i = \frac{1}{2}\left(Q_{i+1} - Q_{i-1} \right). \tag{2.44}$$

We notice that Formula (2.44) may be rewritten more explicitly as:

$$q_{i+\frac{1}{2}} = \frac{7}{12}\left(Q_{i+1} + Q_i \right) - \frac{1}{12}\left(Q_{i+2} + Q_{i-1} \right). \tag{2.45}$$

The Formula (2.45) is fourth-order accurate if q is at least C^4 (Colella & Woodward, 1984). Indeed, we prove this later in Proposition A.1. An explicit expression for the values of $q_{R,i}$ and $q_{L,i}$ are given by:

$$q_{R,i} = q_{i+\frac{1}{2}} = \frac{7}{12}\left(Q_{i+1} + Q_i \right) - \frac{1}{12}\left(Q_{i+2} + Q_{i-1} \right), \tag{2.46}$$

$$q_{L,i} = q_{i-\frac{1}{2}} = \frac{7}{12}\left(Q_i + Q_{i-1} \right) - \frac{1}{12}\left(Q_{i+1} + Q_{i-2} \right). \tag{2.47}$$

During this work, we refer to this PPM scheme as **hord0**. This name is justified because in FV3, the 1D advection solver input is named “hord”.

2.4.2 Monotonization

This section is dedicated to presenting possible ways of ensuring the creation of new extrema values in the PPM reconstruction. We are going to present an alternative scheme from Lin (2004), which was an attempt to reduce the diffusion of the original scheme Colella and Woodward (1984) and is currently employed in the FV3 dynamical core (L. Harris et al., 2021).

Limiter from Lin (2004) - hord8

Similarly to Colella and Woodward (1984), Lin (2004) reduces numerical oscillations in the parabolas by replacing the term δQ_i in Equation (2.43) with the values $\delta_m Q_i$ given by:

$$\delta_m Q_i = \max(|\delta Q_i|, 2\delta Q_{\min,i}, 2\delta Q_{\max,i}) \cdot \operatorname{sgn}(\delta Q_i), \tag{2.48}$$

where $\delta Q_{\min,i} = Q_i - \min(Q_{i+1}, Q_i, Q_{i-1})$ and $\delta Q_{\max,i} = \max(Q_{i+1}, Q_i, Q_{i-1}) - Q_i$. The monotonicity is achieved by the following scheme:

$$q_{L,i} \leftarrow Q_i - \max(|\delta_m Q_i|, |q_{L,i} - Q_i|) \cdot \operatorname{sgn}(\delta_m Q_i), \tag{2.49}$$

$$q_{R,i} \leftarrow Q_i - \max(|\delta_m Q_i|, |q_{R,i} - Q_i|) \cdot \operatorname{sgn}(\delta_m Q_i). \tag{2.50}$$

This scheme may be further improved to reduce the diffusion even more, as described by Lin (2004), but we are not going to assess this approach here. This scheme is referred to as **hord8**.

Appendix A

Numerical Analysis

A.1 Lagrange interpolation

Given real numbers, called nodes, $x_0 < x_1 < \ldots < x_m$, we define the k-th Lagrange polynomial by

$$L_k(x) = \prod_{j=0, j\neq k}^{m} \frac{x - x_j}{x_k - x_j}.$$

They satisfy $L_k(x_j) = \delta_{kj}$, where δ_{kj} is the Kronecker delta. Given a function f defined at the nodes x_j, its interpolating polynomial of degree m is given by:

$$P_m(x) = \sum_{k=0}^{m} f(x_k) L_k(x).$$

Indeed, this polynomial interpolates f since $P_m(x_j) = f(x_j)$. It is well known that P_m always exists and is unique. Besides that, we have the following error formula for Lagrange interpolation.

Theorem A.1. *Let $f \in C^{m+1}(\mathbb{R})$. Then, then there is ξ in the smallest interval containing $x_0, \ldots, x_m, x$ such that:*

$$f(x) - P_m(x) = \omega(x) \frac{f^{(m+1)}(\xi)}{(m+1)!}, \tag{A.1}$$

where $\omega(x) = (x - x_0)(x - x_1) \ldots (x - x_m)$.

Proof. See Stoer and Bulirsch (2002, Theorem 2.1.4.1. on p. 49). □

A.2 Numerical integration

Theorem A.2 (Leibniz integral rule)**.** *If f…*

$$\frac{d}{ds} \int_{s_0}^{s} f(s, \theta)\, d\theta = f\big(s, s\big) + \int_{s_0}^{x} \partial_x f(s, \theta)\, d\theta$$

Proof. Let us define

$$F(s) = \int_{s_0}^{s} f(s,\theta)\,d\theta.$$

Then

$$\begin{aligned}\frac{F(s+h)-F(s)}{h} &= \frac{1}{h}\int_{s_0}^{s+h} f(s+h,\theta)\,d\theta - \frac{1}{h}\int_{s_0}^{s} f(s,\theta)\,d\theta \\ &= \frac{1}{h}\left(\int_{s_0}^{s+h} f(s+h,\theta)\,d\theta - \int_{s}^{s+h} f(s+h,\theta)\,d\theta - \int_{s_0}^{s} f(s,\theta)\,d\theta + \int_{s}^{s+h} f(s+h,\theta)\,d\theta\right)\end{aligned}$$

□

The following mean value theorem for integrals is a very useful tool when working with numerical integration errors.

Theorem A.3 (Mean value theorem for integrals)**.** *If $f \in C([a,b])$, and g is a integrable function in $[a,b]$ whose sign does not change in $[a,b]$, then there exists $c \in]a,b[$ such that*

$$\int_a^b f(x)g(x)\,dx = f(c)\int_a^b g(x)\,dx.$$

Proof. See Courant and John (1999, p. 143). □

A.2.1 Midpoint rule

When considering finite-volume schemes, it is useful to compare the average value on a control volume of a function with its value at the control volume centroid. In the following theorems, for the one and two dimensional cases, respectively, we show that the value of a function at the centroid of a control volume given a second-order approximation to its average value on the control volume.

Theorem A.4. *If $f \in C^2([x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}])$, then*

$$\frac{1}{\Delta x}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} f(x)\,dx - f(x_i) = C_1\Delta x^2, \tag{A.2}$$

where C_1 is a constant that depends only on f, and $x_i = \frac{x_{i+\frac{1}{2}}+x_{i-\frac{1}{2}}}{2}$, $\Delta x = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$.

Proof. From Taylor's expansion, it follows that, for $x \in [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$, we have:

$$f(x) = f(x_i) + f(x_i)(x-x_i) + f''(\xi)\frac{(x-x_i)^2}{2}, \tag{A.3}$$

for some ξ between x and x_i. Therefore:

$$\begin{aligned}\frac{1}{\Delta x}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} f(x)\,dx - f(x_i) &= \frac{1}{\Delta x}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \left(f'(x_i)(x-x_i) + f''(\xi)\frac{(x-x_i)^2}{2}\right)dx \\ &= \frac{1}{\Delta x}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} f''(\xi)\frac{(x-x_i)^2}{2}\,dx.\end{aligned}$$

Using the mean value theorem for integrals (see Theorem A.3), we have:

$$\frac{1}{\Delta x}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} f(x)\,dx - f(x_i) = f''(\eta_i)\frac{1}{\Delta x}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \frac{(x-x_i)^2}{2}\,dx = f''(\eta_i)\frac{\Delta x^2}{24}$$

for some $\eta_i \in [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$, from which the proposition follows with

$$C_1 = \frac{1}{24}f''(\eta_i). \tag{A.4}$$

□

Theorem A.5. *If $f \in C^2([x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}])$, then*

$$\frac{1}{\Delta x \Delta y}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}}\int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} f(x,y)\,dx\,dy - f(x_i, y_j) = C\Delta x^2, \tag{A.5}$$

where C_1 is a constants that depends only on f, where we assume $x_i = \frac{x_{i+\frac{1}{2}}+x_{i-\frac{1}{2}}}{2}$, $y_i = \frac{y_{j+\frac{1}{2}}+y_{j-\frac{1}{2}}}{2}$, $\Delta x = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$, $\Delta y = y_{j+\frac{1}{2}} - y_{j-\frac{1}{2}}$ and $\Delta x = \Delta y$.

Proof. Applying Theorem A.4 in the y direction, we have

$$\int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} f(x,y)\,dy = \Delta y f(x, y_j) + \frac{\Delta y^3}{24}\partial_y^2 f(x, \eta_j),$$

for $\eta_j \in [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}]$. Hence:

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}}\int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} f(x,y)\,dx\,dy = \Delta y\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} f(x,y_j)\,dx + \frac{\Delta y^3}{24}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \partial_y^2 f(x,\eta_j)\,dx.$$

Applying Theorem A.4 in the x direction for $y = y_j$, we get

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} f(x, y_j)\,dx = \Delta x f(x_i, y_j) + \frac{\Delta x^3}{24}\partial_x^2 f(\xi_i, y_j)\,dx,$$

for $\xi_i \in [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$. From this, we obtain

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}}\int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} f(x,y)\,dx\,dy = \Delta x\Delta y f(x_i, y_j) + \frac{\Delta x^3}{24}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \partial_x^2 f(\xi_i, y_j)\,dx + \frac{\Delta y^3}{24}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \partial_y^2 f(x, \eta_j)\,dx.$$

Using Theorem A.3, we obtain the desired formula:

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} f(x,y)\,dx\,dy = \Delta x \Delta y f(x_i, y_j) + \frac{\Delta x^2}{24}\Delta x \Delta y \partial_x^2 f(v_i, y_j) + \frac{\Delta y^2}{24}\Delta x \Delta y \partial_y^2 f(\theta_i, \eta_j),$$

where $v_i, \theta_i \in [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$, recalling that $\Delta x = \Delta y$. □

Corollary A.1. *If $f \in C^2([a,b] \times [c,d])$, and $[a,b] \times [c,d]$ is written as the union of the uniformed-spaces control volumes $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}]$, $i, j = 1, \dots, N$, with lengths $\Delta x = \Delta y$, we have*

$$\int_a^b \int_c^d f(x,y)\,dx\,dy - \sum_{i,j=1}^{N} f(x_i, y_j)\Delta x \Delta y = C_1 \Delta x^2, \tag{A.6}$$

where C_1 depends only on f.

Proof. Using Theorem A.5, we have:

$$\begin{aligned}\frac{1}{\Delta x \Delta y}\int_a^b \int_c^d f(x,y)\,dx\,dy &= \frac{1}{\Delta x \Delta y}\sum_{i,j=1}^{N} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} f(x,y)\,dx\,dy \\ &= \sum_{i,j=1}^{N} f(x_i, y_j) + \frac{\Delta x^2}{24}\sum_{i,j=1}^{N}\Big(\partial_x^2 f(v_i, y_j) + \partial_y^2 f(v_i, y_j)\Big).\end{aligned}$$

We notice that

$$\Delta x \Delta y \sum_{i,j=1}^{N}\Big(\partial_x^2 f(v_i, y_j) + \partial_y^2 f(v_i, y_j)\Big) = \frac{(b-a)(d-c)}{N^2}\sum_{i,j=1}^{N}\Big(\partial_x^2 f(v_i, y_j) + \partial_y^2 f(v_i, y_j)\Big)$$

and we also point that from the inequality

$$\min\big(\partial_x^2 f + \partial_y^2 f\big)(x,y) \le \frac{1}{N^2}\sum_{i,j=1}^{N}\Big(\partial_x^2 f(v_i, y_j) + \partial_y^2 f(v_i, y_j)\Big) \le \max\big(\partial_x^2 f + \partial_y^2 f\big)(x,y),$$

and with the aid of the intermediate value theorem, we have

$$\frac{1}{N^2}\sum_{i,j=1}^{N}\Big(\partial_x^2 f(v_i, y_j) + \partial_y^2 f(v_i, y_j)\Big) = \big(\partial_x^2 f + \partial_y^2 f\big)(\overline{x}, \overline{y})$$

for some $(\overline{x}, \overline{y}) \in [a,b] \times [c,d]$ from which the claim follows. □

A.3 Convergence of 1D FV-SL schemes

A.3.1 Consistency and convergence

Hereafter, we are going to use the notations introduced in Section 2.1.1. To move towards the convergence of 1D-FV schemes, for Problem 2.4 we introduce the local truncation error (LTE hereafter) τ_i^n following LeVeque (2002):

$$Q_i(t^{n+1}) = Q_i(t^n) - \lambda\left(F^n_{i+\frac{1}{2}}(Q(t^n), \tilde{u}^n_{i+\frac{1}{2}}) - F^n_{i-\frac{1}{2}}(Q(t^n), \tilde{u}^n_{i-\frac{1}{2}})\right) + \Delta t \tau_i^n. \tag{A.7}$$

We the define $\tau^n \in \mathbb{P}_\nu^N$, which represent the LTEs at the time-step n. Notice the LTE is obtained by replacing the exact solution in Equation (2.21). Since $Q_i(t^n)$ is the exact solution of Equation (2.9), the LTE may be rewritten as

$$\begin{aligned} \tau_i^n = \frac{1}{\Delta x}\Bigg[&\left(\frac{1}{\Delta t}\int_{t^n}^{t^{n+1}} (uq)(x_{i+\frac{1}{2}}, t)\, dt - F^n_{i+\frac{1}{2}}(Q(t^n), \tilde{u}^n_{i+\frac{1}{2}})\right) + \\ &\left(\frac{1}{\Delta t}\int_{t^n}^{t^{n+1}} (uq)(x_{i-\frac{1}{2}}, t)\, dt - F^n_{i-\frac{1}{2}}(Q(t^n), \tilde{u}^n_{i-\frac{1}{2}})\right)\Bigg]. \end{aligned} \tag{A.8}$$

The LTE gives a measure of how well the 1D-FV scheme approximates the integral form of the considered conservation law. Another interpretation of the LTE is that the LTE gives the error obtained after applying the scheme for a single time-step using the exact solution. Now we can define consistency.

Definition A.1 (Consistency)**.** *Let us consider the framework of Problem 2.4. A 1D-FV scheme is said to be consistency in the p-norm if for any sequence of $(\Delta x^{(k)}, \Delta t^{(k)}, \lambda)$-discretizations, $k \in \mathbb{N}$, with $\lim_{k\to\infty} \Delta x^{(k)} = \lim_{k\to\infty} \Delta t^{(k)} = 0$, we have:*

$$\lim_{k\to\infty}\left[\max_{1\leq n\leq N_T^{(k)}} \|\tau^n\|_{p,\Delta x^{(k)}}\right] = 0,$$

and it is said to be consistent with order P in the $p-$norm if

$$\max_{1\leq n\leq N_T^{(k)}} \|\tau^n\|_{p,\Delta x^{(k)}} = O(\Delta x^P).$$

From Equation (A.8), it follows that we basically need to ensure that the numerical flux function $\mathcal{F}^n_{i+\frac{1}{2}}$ converges to the time-averaged flux at edges when $\Delta x \to 0$ in order to guarantee consistency.

At last, we define the point-wise error at time-step n by:

$$E_i^n = Q_i(t^n) - Q_i^n, \quad i = 1, \dots, N,$$

and we define the vector of errors by $E^n \in \mathbb{P}_\nu^N$ with entries E_i^n.

Definition A.2 (Convergence)**.** *Let us consider the framework of Problem 2.4. A 1D-FV scheme is said to be convergent in the p-norm if for any sequence of $(\Delta x^{(k)}, \Delta t^{(k)}, \lambda)$-*

discretizations, $k \in \mathbb{N}$*, with* $\lim_{k\to\infty} \Delta x^{(k)} = \lim_{k\to\infty} \Delta t^{(k)} = 0$*, we have:*

$$\lim_{k\to\infty} \left[\max_{1\le n\le N_T^{(k)}} \|E^n\|_{p,\Delta x^{(k)}} \right] = 0,$$

and it is said to converge with order P *in the* $p-$*norm if*

$$\max_{1\le n\le N_T^{(k)}} \|E^n\|_{p,\Delta x^{(k)}} = O(\Delta x^P).$$

Subtracting Equation (2.21) from Equation (A.7) we get the following equation for the error:

$$\begin{aligned} E_i^{n+1} = E_i^n - \lambda \Bigg[& \left(F_{i+\frac{1}{2}}^n(Q(t^n), \tilde{u}_{i+\frac{1}{2}}^n) - F_{i+\frac{1}{2}}^n(Q^n, \tilde{u}_{i+\frac{1}{2}}^n) \right) \\ & - \left(F_{i-\frac{1}{2}}^n(Q(t^n), \tilde{u}_{i-\frac{1}{2}}^n) - F_{i-\frac{1}{2}}^n(Q^n, \tilde{u}_{i-\frac{1}{2}}^n) \right) \Bigg] + \tau_i^n \Delta t. \end{aligned} \tag{A.9}$$

Notice that if $q, u \in C^3$, we can rewrite Equation (A.8) as:

$$\tau_i^n = \left[\frac{1}{\Delta x \Delta t} \int_{t^n}^{t^{n+1}} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \frac{\partial(uq)}{\partial x}(x,t)\, dx\, dt - \left(\frac{F_{i+\frac{1}{2}}^n(Q(t^n), \tilde{u}_{i-\frac{1}{2}}^n) - F_{i-\frac{1}{2}}^n(Q(t^n), \tilde{u}_{i-\frac{1}{2}}^n)}{\Delta x} \right) \right].$$

Using the midpoint rule for integration (Theorem A.4) and the mean value theorem for integrals (Theorem A.3), we have:

$$\begin{aligned} \tau_i^n &= \left[\frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} \left(\frac{\partial(uq)}{\partial x}(x_i,t) + \frac{\Delta x^2}{24} \frac{\partial(uq)}{\partial x}(\xi,t) \right) dt - \left(\frac{F_{i+\frac{1}{2}}^n(Q(t^n), \tilde{u}_{i+\frac{1}{2}}^n) - F_{i-\frac{1}{2}}^n(Q(t^n), \tilde{u}_{i-\frac{1}{2}}^n)}{\Delta x} \right) \right] \\ &= \left[\frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} \frac{\partial(uq)}{\partial x}(x_i,t)\, dt - \left(\frac{F_{i+\frac{1}{2}}^n(Q(t^n), \tilde{u}_{i+\frac{1}{2}}^n) - F_{i-\frac{1}{2}}^n(Q(t^n), \tilde{u}_{i-\frac{1}{2}}^n)}{\Delta x} \right) \right] + \frac{\Delta x^2}{24} \frac{\partial^3(uq)}{\partial x^3}(\xi, \bar{t}), \end{aligned} \tag{A.10}$$

for $\xi \in X_i$ and $\bar{t} \in [t^n, t^{n+1}]$. Therefore, if $q, u \in C^3$ the scheme is consistent, if and only if, $\frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} \frac{\partial(uq)}{\partial x}(x_i,t)\, dt$ is approximated by $\frac{F_{i+\frac{1}{2}}^n(Q(t^n), \tilde{u}_{i+\frac{1}{2}}^n)) - F_{i-\frac{1}{2}}^n(Q(t^n), \tilde{u}_{i-\frac{1}{2}}^n)}{\Delta x}$. This shall be very useful when we consider two-dimensional schemes, where we are going to use the discrete operators to estimate the divergence of velocity fields.

A.3.2 Stability

In order to define the concept of stability, it is useful to introduce an operator representation of 1D-FV schemes. In the context of Problem 2.4, we define the operators $\mathcal{H}_{\Delta x,n} : \mathbb{P}_\nu^N \to \mathbb{P}_\nu^N$ whose i-th entry is given by:

$$[\mathcal{H}_{\Delta x,n}(Q)]_i = Q_i - \lambda \left(F_{i+\frac{1}{2}}^n(Q, \tilde{u}_{i+\frac{1}{2}}^n) - F_{i-\frac{1}{2}}^n(Q, \tilde{u}_{i-\frac{1}{2}}^n) \right), \tag{A.11}$$

for $i = 1, \dots, N$, $n = 0, \dots, N_T - 1$. Notice that the dependence on n is due to the velocity that may be allowed to vary with time. As it is usual, we are assuming periodicity in the entries of Q when we apply the operator $\mathcal{H}_{\Delta x,n}$. Thus, Equation (2.21) may be rewritten in a vector form by

$$Q^{n+1} = \mathcal{H}_{\Delta x,n}(Q^n),$$

and Equation (A.7) in a vector form reads

$$Q(t^{n+1}) = \mathcal{H}_{\Delta x,n}(Q(t^n)) + \Delta t \tau^n,$$

and the error equation (A.9) is given by

$$E^{n+1} = \mathcal{H}_{\Delta x,n}(Q(t^n)) - \mathcal{H}_{\Delta x,n}(Q^n) + \Delta t \tau^n. \tag{A.12}$$

The stability theory focus on uniformly bounding the norm of $\mathcal{H}_{\Delta x,n}(Q(t^n)) - \mathcal{H}_{\Delta x,n}(Q^n)$ (LeVeque, 2002). We define stability as follows.

Definition A.3 (Stability). *In the context of Problem 2.4, a 1D-FV scheme is stable in the $p-$norm if for any $(\Delta x, \Delta t, \lambda)$-discretization of $[a,b] \times [0,T]$ we have:*

$$\|\mathcal{H}_{\Delta x,n}(Q) - \mathcal{H}_{\Delta x,n}(P)\|_{p,\Delta x} \le (1 + \alpha \Delta t)\|Q - P\|_{p,\Delta x}, \tag{A.13}$$

for all $Q, P \in \mathbb{R}^N_\nu$ and α is a constant that does not depend neither on Δx nor on Δt.

Assuming that the scheme is stable in the $p-$norm, then it follows from Equation (A.12) that:

$$\begin{aligned}
\|E^{n+1}\|_{p,\Delta x} &\le \|\mathcal{H}_{\Delta x,n}(Q(t^n)) - \mathcal{H}_{\Delta x,n}(Q^n)\|_{p,\Delta x} + \Delta t \max_{n=1,\dots,N_T} \|\tau^n\|_{p,\Delta x} \\
&\le (1 + \alpha \Delta t)\|E^n\|_{p,\Delta x} + \Delta t \max_{n=1,\dots,N_T} \|\tau^n\|_{p,\Delta x} \\
&\le (1 + \alpha \Delta t)^n \|E^0\|_{p,\Delta x} + \Delta t \max_{n=1,\dots,N_T} \|\tau^n\|_{p,\Delta x} \sum_{k=0}^{n-1} (1 + \alpha \Delta t)^k \\
&\le e^{\alpha T}(\|E^0\|_{p,\Delta x} + T \max_{n=1,\dots,N_T} \|\tau^n\|_{p,\Delta x}),
\end{aligned} \tag{A.14}$$

where we used $n\Delta t \le T$, $T = N\Delta t$ and the inequality $e^t > 1 + t$. When computing the initial average values using the value at the cell centroid, the initial error E^0 converges to zero provided q is twice continuously differentiable by Proposition 2.2. Therefore, it follows that if the scheme is stable and consistent then it is convergent. Furthermore, if it is stable and consistent with order P, then the convergence order is at least equal to $\min\{P, 2\}$. In the case where both the conservation law and $\mathcal{H}_{\Delta x,n}$ are linear, this result is a particular case of the Lax-Ritchmyer stability and the convergence is guaranteed by the Lax equivalence theorem (LeVeque, 2002). In this Chapter, we are interested only in the linear advection equation. However, as pointed in Section **??**, the operator $\mathcal{H}_{\Delta x,n}$ may become non-linear when monotonicity constraints are activated.

Notice that, if $\mathcal{H}_{\Delta x,n}$ is linear, then stability is equivalent to require that

$$\|\mathcal{H}_{\Delta x,n}\|_{p,\Delta x} \le 1 + \alpha \Delta t,$$

where

$$\|\mathcal{H}_{\Delta x,n}\|_{p,\Delta x} = \sup_{Q\in\mathbb{R}^{\Delta x}} \frac{\|\mathcal{H}_{\Delta x,n}(Q)\|_{p,\Delta x}}{\|Q\|_{p,\Delta x}},$$

is the operator p-norm.

For linear operators, we may use the discrete Fourier transform (Trefethen, 2000) to estimate the 2-norm of $\mathcal{H}_{\Delta x,n}$. This approach is known as Von Neumann stability analysis. We define the nodes $\theta_i = i\frac{2\pi}{N}$, $i = 1, \dots, N$, $\Delta\theta = \frac{2\pi}{N}$, $\theta = (\theta_1, \theta_2, \dots, \theta_N)$. The imaginary unit is denoted by $\imath$. We define $\mathbb{C}_\nu^N$ similarly as $\mathbb{P}_\nu^N$. The Fourier modes $e^{\imath k\theta}] \in \mathbb{C}_\nu^N$ for $k = 1, \dots, N$, have entries are given by:

$$[e^{\imath k\theta}]_i = e^{\imath k\theta_i}, \quad \text{for } i = 1, \dots, N.$$

Each k is referred to wavenumber and θ_k is called dimensionless wavenumber. The Fourier modes form an orthogonal basis of $\mathbb{C}_\nu^N$ with respect to the inner product

$$\langle Q, P\rangle = \frac{1}{N}\sum_{i=1}^{N} Q_i\overline{P_i},$$

for $P, Q \in \mathbb{C}_\nu^N$ and $\overline{z}$ denotes the complex conjugate of z. Given $Q \in \mathbb{P}_\nu^N$, we may may express it in terms of the Fourier modes

$$Q = \sum_{k=1}^{N} a_k e^{\imath k\theta},$$

where $a_k \in \mathbb{C}$. The 2-norm of Q is then given by:

$$\|Q\|_{2,\Delta x} = \sqrt{N\sum_{k=1}^{N}|a_k|^2}.$$

The idea of Von Neumann stability analysis is to apply the operator $\mathcal{H}_{\Delta x,n}$ on each Fourier mode and analyze how it modifies its amplitude. For ease of analysis, we assume that the velocity is constant, which implies that the operator $\mathcal{H}_{\Delta x,n}$ has constant coefficients and does not depend on n. For the general case, where the velocity is not constant, the stability can be ensured using the frozen coefficients method (Strikwerda, 2004, p. 59). This method boils down to performing multiple times the stability analysis with a constant velocity being equal to each one of the possible values of the velocity on the grid. If the scheme is stable for all the possible constant velocities, then stability is ensured. Since the operator is supposed to be linear with constant coefficients and we are assuming periodic boundaries conditions, we may write:

$$\mathcal{H}_{\Delta x,n}(e^{\imath k\theta}) = \rho(k)e^{\imath k\theta},$$

where the term $\rho(k)$ is called amplification factor and it is an eigenvalue of $\mathcal{H}_{\Delta x,n}$. The norm of $\mathcal{H}_{\Delta x,n}(Q)$ is bounded by:

$$\|\mathcal{H}_{\Delta x,n}(Q)\|_{2,\Delta x}^2 = N\sum_{k=1}^{N}|a_k|^2|\rho(k)|^2 \leq \max_{k=1,\dots,N}|\rho(k)|^2\|Q\|_{2,\Delta x}^2.$$

Therefore:

$$\|\mathcal{H}_{\Delta x,n}\|_{2,\Delta x} \le \max_{k=1,\dots,N} |\rho(k)|.$$

If we show that $\max_{k=1,\dots,N} |\rho(k)| \le 1 + \alpha\Delta t$, with α independent of Δt, N and n, then we ensure the stability of $\mathcal{H}_{\Delta x,n}$.

A.3.3 Flux accuracy analysis

With the PPM operator, we can compute the amplification factor by applying it on each Fourier mode considering the PPM and the hybrid PPM schemes, both without monotonization. We assume a constant velocity equal to one and $N = 100$ (number of control volumes). In Figure A.1 we show the amplification factor for both PPM and hybrid PPM schemes considering different CFL numbers. We can observe that both schemes damp most of the Fourier modes for larger k, regardless of the CFL number. Besides that, the hybrid scheme is more effective when reducing the Fourier modes amplitude. We point out that both schemes are exact when the CFL number is equal to 1. From this analysis, we can conclude that the PPM and hybrid PPM schemes satisfy the Von Neumann stability criteria when the CFL restriction is respected. For an analysis of stability for larger time-steps, we refer to Lauritzen (2007).

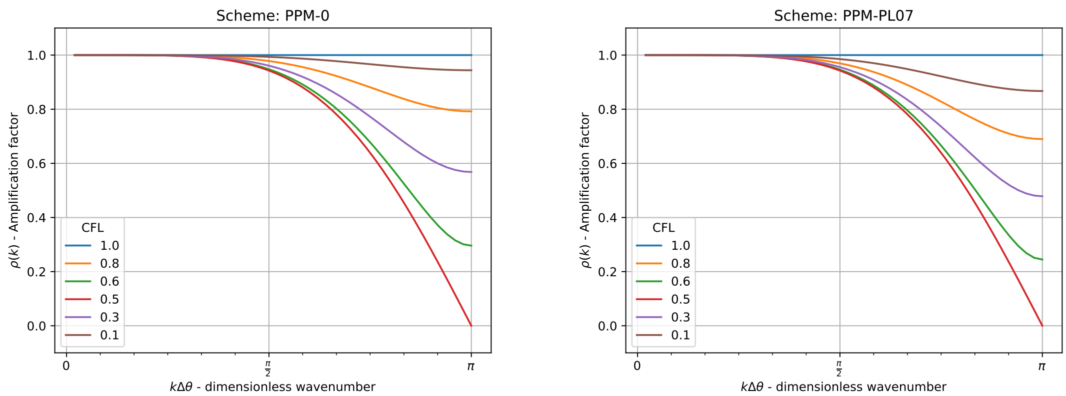

Figure A.1: *Amplification factor for the PPM (left) and hybrid PPM (right) schemes for different CFL numbers.*

A.4 Convergence, consistency and stability of 2D-FV schemes

The notions of convergence, consistency and stability for a 2D-FV schemes are straightforward from these notions for 1D-FV schemes (see Subsections A.3.1 and A.3.2). Indeed, in the context of Problem **??**, we define the operators $\mathcal{H}_{\Delta x,\Delta y,n} : \mathbb{R}^{N\times M} \to \mathbb{R}^{N\times M}$ whose (i, j) entry is given by:

$$[\mathcal{H}_{\Delta x,\Delta y,n}(Q)]_{ij} = Q_{ij} - \Delta t \mathbb{D}^n_{ij}$$

for $i = 1, \dots, N$, $j = 1, \dots, M$, $n = 0, \dots, N_T - 1$. The 2D-FV is then expressed as

$$Q^{n+1} = \mathcal{H}_{\Delta x, \Delta y, n}(Q^n).$$

The local error truncation $\tau^n \in \mathbb{R}^{N \times M}$ is given by

$$Q(t^{n+1}) = \mathcal{H}_{\Delta x, \Delta y, n}(Q(t^n)) + \Delta t \tau^n.$$

The error equation is given by

$$E^{n+1} = \mathcal{H}_{\Delta x, \Delta y, n}(Q(t^n)) - \mathcal{H}_{\Delta x, \Delta y, n}(Q^n) + \Delta t \tau^n. \tag{A.15}$$

The stability in the p-norm is defined as in the 1D case.

Definition A.4. *A 2D-FV scheme is stable in the $p-$norm if*

$$\|\mathcal{H}_{\Delta x, \Delta y, n}(Q) - \mathcal{H}_{\Delta x, \Delta y, n}(P)\|_{p, \Delta x \times \Delta y} \leq (1 + \alpha \Delta t)\|Q - P\|_{p, \Delta x \times \Delta y}, \tag{A.16}$$

for all $Q, P \in \mathbb{R}^{N \times M}$ and α is a constant that does not depend neither on Δx, Δy, Δt nor on n.

If a 2D-FV scheme is stable in the $p-$norm, similarly to Equation (A.14) we have:

$$\|E^{n+1}\|_{p, \Delta x \times \Delta y} \leq e^{\alpha T}(\|E^0\|_{p, \Delta x \times \Delta y} + T \max_{n=1,\dots,N_T} \|\tau^n\|_{p, \Delta x \times \Delta y}).$$

Again, we point out that from Proposition ??, we have that the initial error E^0 shall be second-order accurate. Consistency is defined as in Definition A.1 and convergence is defined as in Definition A.2.

The Von Neumann analysis can be applied when $\mathcal{H}_{\Delta x, \Delta y, n}$ is linear, since we are considering periodic boundary conditions. The idea is the same as in the one-dimensional case, we just apply the operator $\mathcal{H}_{\Delta x, \Delta y, n}$ on the Fourier modes to obtain the amplification factor. We introduce the nodes $\theta_i = i\frac{2\pi}{N}$, $i = 1, \dots, N$, $\Delta\theta = \frac{2\pi}{N}$, $\theta_i = (\theta_1, \theta_2, \dots, \theta_N)$, $\phi_j = j\frac{2\pi}{M}$, $j = 1, \dots, M$, $\Delta\phi = \frac{2\pi}{M}$, $\phi = (\phi_1, \phi_2, \dots, \phi_M)$. For $k_1 = 1, \dots, N$, $k_2 = 1, \dots, M$, the two-dimensional Fourier mode $\boldsymbol{k} = (k_1, k_2)$ from $\mathbb{C}^{N \times M}$ has its (i, j) entry given by $[e^{\imath \boldsymbol{k}\theta}]_{ij} = e^{\imath k_1 \theta_i} e^{\imath k_2 \phi_j}$. For an analysis of stability for the dimension splitting method, we refer to Lauritzen (2007) and Lin and Rood (1996).

Notice that if $q, u, v \in C^3$, we can rewrite the LTE as:

$$\tau^n_{ij} = \left[\frac{1}{\Delta x \Delta y \Delta t} \int_{t^n}^{t^{n+1}} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} \nabla \cdot (\boldsymbol{u}q)(x, y, t)\, dy\, dx\, dt - \mathbb{D}^n_{ij}\right].$$

Using the midpoint rule for integration (Theorem A.5), the mean value theorem for integrals (Theorem A.3) and recalling the discrete divergence (Definition ??), we have:

$$\tau^n_{ij} = \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} \nabla \cdot (\boldsymbol{u}q)(x_i, y_j, t)\, dt - \mathbb{D}^n_{ij} + O(\Delta x^2) + O(\Delta y^2). \tag{A.17}$$

Therefore, in order to investigate the consistency, we may compare how well the discrete divergence approximates the divergence.

A.5 Finite-difference estimates

This Section aims to prove all finite-difference error estimations used throughout this appendix. All the proves are very simple and consist of applying Taylor's expansions, as it is usual when computing the accuracy order of many numerical schemes.

Lemma A.1. *Let $F \in C^5(\mathbb{R})$, $x_0 \in \mathbb{R}$ and $h > 0$. Then, the following identity holds:*

$$F'(x_0) = \frac{4}{3}\left(\frac{F(x_0+h)-F(x_0-h)}{2h}\right) - \frac{1}{3}\left(\frac{F(x_0+2h)-F(x_0-2h)}{4h}\right) + C_1 h^4, \tag{A.18}$$

where C_1 is a constant that depends only on F and h.

Proof. Given $\delta \in]0, 2h]$, then $x_0 + \delta \in]x_0, x_0 + 2h]$ and $x_0 - \delta \in]x_0 - 2h, x_0]$. Then, we get using the Taylor expansion of F:

$$F(x_0+\delta) = F(x_0)+F'(x_0)\delta + F^{(2)}(x_0)\frac{\delta^2}{2} + F^{(3)}(x_0)\frac{\delta^3}{3!} + F^{(4)}(x_0)\frac{\delta^3}{4!} + F^{(5)}(\theta_\delta)\frac{\delta^5}{5!} \quad \theta_\delta \in [x_0, x_0+\delta],$$
$$F(x_0-\delta) = F(x_0)-F'(x_0)\delta + F^{(2)}(x_0)\frac{\delta^2}{2} - F^{(3)}(x_0)\frac{\delta^3}{3!} + F^{(4)}(x_0)\frac{\delta^4}{4!} - F^{(5)}(\theta_{-\delta})\frac{\delta^5}{5!}, \quad \theta_{-\delta} \in [x_0-\delta, x_0].$$

Thus:

$$\frac{F(x_0+\delta)-F(x_0-\delta)}{2\delta} = F'(x_0) + F^{(3)}(x_0)\frac{\delta^2}{3!} + \left(F^{(5)}(\theta_\delta) + F^5(\theta_{-\delta})\right)\frac{\delta^4}{2\cdot 5!}, \tag{A.19}$$

Applying Equation (A.19) for $\delta = h$ and $\delta = 2h$, we get, respectively:

$$\frac{F(x_0+h)-F(x_0-h)}{2h} = F'(x_0)+F^{(3)}(x_0)\frac{h^2}{3!}+\left(F^{(5)}(\theta_h)+F^{5)}(\theta_{-h})\right)\frac{h^4}{2\cdot 5!}, \quad \theta_h \in [x_0, x_0+h], \quad \theta_{-h} \in [x_0-h, x_0], \tag{A.20}$$

and

$$\frac{F(x_0+2h)-F(x_0-2h)}{4h} = F'(x_0) + F^{(3)}(x_0)\frac{4h^2}{3!} + \left(F^{(5)}(\theta_{2h}) + F^{(5)}(\theta_{-2h})\right)\frac{16h^4}{2\cdot 5!}, \tag{A.21}$$
$$\theta_{2h} \in [x_0, x_0+2h], \quad \theta_{-2h} \in [x_0-2h, x_0].$$

Using Equations (A.20) and (A.21), we obtain:

$$\frac{4}{3}\left(\frac{F(x_0+h)-F(x_0-h)}{2h}\right) = \frac{4}{3}F'(x_0) + F^{(3)}(x_0)\frac{4h^2}{3\cdot 3!} + \left(F^{(5)}(\theta_h) + F^{(5)}(\theta_{-h})\right)\frac{h^4}{2\cdot 5!}, \tag{A.22}$$

$$\frac{1}{3}\left(\frac{F(x_0+2h)-F(x_0-2h)}{4h}\right) = \frac{1}{3}F'(x_0) + F^{(3)}(x_0)\frac{4h^2}{3\cdot 3!} + \left(F^{(5)}(\theta_{2h}) + F^{(5)}(\theta_{-2h})\right)\frac{16h^4}{3\cdot 2\cdot 5!} \tag{A.23}$$

Subtracting Equation (A.23) from Equation (A.22) we get the desired Equation (A.18) with

$$C_1 = \frac{1}{720}\left(3F^{(5)}(\theta_h) + 3F^{(5)}(\theta_{-h}) - 16F^{(5)}(\theta_{2h}) - 16F^{(5)}(\theta_{-2h})\right), \tag{A.24}$$

where $\theta_h \in [x_0, x_0+h], \theta_{-h} \in [x_0-h, x_0], \theta_{2h} \in [x_0, x_0+2h], \theta_{-2h} \in [x_0-2h, x_0]$. Using the intermediate value theorem, we can express C_1 in a more compact way as

$$C_1 = \frac{1}{720}\left(6F^{(5)}(\eta_1) - 32F^{(5)}(\eta_2)\right), \tag{A.25}$$

where $\eta_1, \eta_2 \in [x_0-2h, x_0+2h]$, which concludes the proof. □

Lemma A.2. *Let $F \in C^4(\mathbb{R})$, $x_0 \in \mathbb{R}$ and $h > 0$. Then, the following identity holds:*

$$F''(x_0) = \frac{-2F(x_0-2h) + 15F(x_0-h) - 28F(x_0) + 20F(x_0+h) - 6F(x_0+2h) + F(x_0+3h)}{6h^2} + C_2h^2, \tag{A.26}$$

where C_2 is a constant that depends only on F and h.

Proof. From the Taylor's expansion, we have:

$$\begin{aligned}
F(x_0-2h) &= F(x_0) - 2F'(x_0)h + 2F^{(2)}(x_0)h^2 - \frac{8}{6}F^{(3)}(x_0)h^3 + \frac{16}{24}F^{(4)}(\theta_{-2h})h^4,\\
F(x_0-h) &= F(x_0) - F'(x_0)h + \frac{1}{2}F^{(2)}(x_0)h^2 - \frac{1}{6}F^{(3)}(x_0)h^3 + \frac{1}{24}F^{(4)}(\theta_{-h})h^4,\\
F(x_0+h) &= F(x_0) + F'(x_0)h + \frac{1}{2}F^{(2)}(x_0)h^2 + \frac{1}{6}F^{(3)}(x_0)h^3 + \frac{1}{24}F^{(4)}(\theta_h)h^4,\\
F(x_0+2h) &= F(x_0) + 2F'(x_0)h + 2F^{(2)}(x_0)h^2 + \frac{8}{6}F^{(3)}(x_0)h^3 + \frac{16}{24}F^{(4)}(\theta_{2h})h^4,\\
F(x_0+3h) &= F(x_0) + 3F'(x_0)h + \frac{9}{2}F^{(2)}(x_0)h^2 + \frac{27}{6}F^{(3)}(x_0)h^3 + \frac{81}{24}F^{(4)}(\theta_{3h})h^4,
\end{aligned}$$

where $\theta_{-2h} \in [x_0-2h, x_0-h], \theta_{-h} \in [x_0-h, x_0], \theta_h \in [x_0, x_0+h], \theta_{2h} \in [x_0+h, x_0+2h], \theta_{3h} \in [x_0+2h, x_0+3h]$. Multiplying these equations by their respective coefficients given in Equation (A.26), one get:

$$
\begin{aligned}
-2F(x_0-2h) &= -2F(x_0) + 4F'(x_0)h - 4F^{(2)}(x_0)h^2 + \frac{16}{6}F^{(3)}(x_0)h^3 - \frac{32}{24}F^{(4)}(\theta_{-2h})h^4, \\
15F(x_0-h) &= 15F(x_0) - 15F'(x_0)h + \frac{15}{2}F^{(2)}(x_0)h^2 - \frac{15}{6}F^{(3)}(x_0)h^3 + \frac{15}{24}F^{(4)}(\theta_{-h})h^4, \\
-28F(x_0) &= -28F(x_0), \\
20F(x_0+h) &= 20F(x_0) + 20F'(x_0)h + 10F^{(2)}(x_0)h^2 + \frac{20}{6}F^{(3)}(x_0)h^3 + \frac{20}{24}F^{(4)}(\theta_h)h^4, \\
-6F(x_0+2h) &= -6F(x_0) - 12F'(x_0)h - 12F^{(2)}(x_0)h^2 - 8F^{(3)}(x_0)h^3 - \frac{96}{24}F^{(4)}(\theta_{2h})h^4, \\
F(x_0+3h) &= F(x_0) + 3F'(x_0)h + \frac{9}{2}F^{(2)}(x_0)h^2 + \frac{27}{6}F^{(3)}(x_0)h^3 + \frac{81}{24}F^{(4)}(\theta_{3h})h^4.
\end{aligned}
$$

Summing all these equations, we get the desired Formula (A.26) with C_2 given by:

$$
C_2 = \frac{1}{24}\left(32F^{(4)}(\theta_{-2h}) - 15F^{(4)}(\theta_{-h}) - 20F^{(4)}(\theta_h) + 96F^{(4)}(\theta_{2h}) - 81F^{(4)}(\theta_{3h})\right). \tag{A.27}
$$

Using the intermediate value theorem, we can express C_2 in a more compact way as

$$
C_2 = \frac{1}{24}\left(128F^{(5)}(\eta_1) - 116F^{(5)}(\eta_2)\right), \tag{A.28}
$$

where $\eta_1, \eta_2 \in [x_0-2h, x_0+3h]$, which concludes the proof. □

Lemma A.3. *Let $F \in C^4(\mathbb{R})$, $x_0 \in \mathbb{R}$ and $h > 0$. Then, the following identity holds:*

$$
F^{(3)}(x_0) = \frac{F(x_0-2h) - 7F(x_0-h) + 16F(x_0) - 16F(x_0+h) + 7F(x_0+2h) - F(x_0+3h)}{2h^3} + C_3 h, \tag{A.29}
$$

where C_3 is a constant that depends only on F and h.

Proof. From the Taylor's expansion, we have:

$$
\begin{aligned}
F(x_0-2h) &= F(x_0) - 2F'(x_0)h + 2F^{(2)}(x_0)h^2 - \frac{8}{6}F^{(3)}(x_0)h^3 + \frac{16}{24}F^{(4)}(\theta_{-2h})h^4, \\
F(x_0-h) &= F(x_0) - F'(x_0)h + \frac{1}{2}F^{(2)}(x_0)h^2 - \frac{1}{6}F^{(3)}(x_0)h^3 + \frac{1}{24}F^{(4)}(\theta_{-h})h^4, \\
F(x_0+h) &= F(x_0) + F'(x_0)h + \frac{1}{2}F^{(2)}(x_0)h^2 + \frac{1}{6}F^{(3)}(x_0)h^3 + \frac{1}{24}F^{(4)}(\theta_h)h^4, \\
F(x_0+2h) &= F(x_0) + 2F'(x_0)h + 2F^{(2)}(x_0)h^2 + \frac{8}{6}F^{(3)}(x_0)h^3 + \frac{16}{24}F^{(4)}(\theta_{2h})h^4, \\
F(x_0+3h) &= F(x_0) + 3F'(x_0)h + \frac{9}{2}F^{(2)}(x_0)h^2 + \frac{27}{6}F^{(3)}(x_0)h^3 + \frac{81}{24}F^{(4)}(\theta_{3h})h^4,
\end{aligned}
$$

where $\theta_{-2h} \in [x_0-2h, x_0-h]$, $\theta_{-h} \in [x_0-h, x_0]$, $\theta_h \in [x_0, x_0+h]$, $\theta_{2h} \in [x_0+h, x_0+2h]$, $\theta_{3h} \in$

$[x_0 + 2h, x_0 + 3h]$. Multiplying these equations by their respective coefficients given in Equation (A.29), one get:

$$
\begin{aligned}
F(x_0 - 2h) &= F(x_0) - 2F'(x_0)h + \frac{4}{2}F^{(2)}(x_0)h^2 - \frac{8}{6}F^{(3)}(x_0)h^3 + \frac{16}{24}F^{(4)}(\theta_{-2h})h^4,\\
-7F(x_0 - h) &= -7F(x_0) + 7F'(x_0)h - \frac{7}{2}F^{(2)}(x_0)h^2 + \frac{7}{6}F^{(3)}(x_0)h^3 - \frac{7}{24}F^{(4)}(\theta_{-h})h^4,\\
16F(x_0) &= 16F(x_0),\\
-16F(x_0 + h) &= -16F(x_0) - 16F'(x_0)h - \frac{16}{2}F^{(2)}(x_0)h^2 - \frac{16}{6}F^{(3)}(x_0)h^3 - \frac{16}{24}F^{(4)}(\theta_h)h^4,\\
7F(x_0 + 2h) &= 7F(x_0) + 14F'(x_0)h + \frac{28}{2}F^{(2)}(x_0)h^2 + \frac{56}{6}F^{(3)}(x_0)h^3 + \frac{112}{24}F^{(4)}(\theta_{2h})h^4,\\
-F(x_0 + 3h) &= -F(x_0) - 3F'(x_0)h - \frac{9}{2}F^{(2)}(x_0)h^2 - \frac{27}{6}F^{(3)}(x_0)h^3 - \frac{81}{24}F^{(4)}(\theta_{3h})h^4.
\end{aligned}
$$

Summing all these equations, we have:

$$F(x_0 - 2h) - 7F(x_0 - h) + 16F(x_0) - 16F(x_0 + h) + 7F(x_0 + 2h) - F(x_0 + 3h) = 2F^{(3)}(x_0)h^3 - 2C_3h^4,$$

we get the desired Formula (A.29) with C_3 given by:

$$C_3 = \frac{1}{48}\Bigg(- 16F^{(4)}(\theta_{-2h}) + 7F^{(4)}(\theta_{-h}) + 16F^{(4)}(\theta_h) - 112F^{(4)}(\theta_{2h}) + 81F^{(4)}(\theta_{3h})\Bigg). \quad \text{(A.30)}$$

Using the intermediate value theorem, we can express C_3 in a more compact way as

$$C_3 = \frac{1}{48}\Bigg(104F^{(5)}(\eta_1) - 128F^{(5)}(\eta_2)\Bigg), \quad \text{(A.31)}$$

where $\eta_1, \eta_2 \in [x_0 - 2h, x_0 + 3h]$, which concludes the proof. □

A.6 PPM reconstruction accuracy analysis

In this Section, we are going to investigate the accuracy of the PPM reconstruction process. As we pointed out in Section 2.4.1, the approximation of q at the control volumes edges given by Equation (2.45) is fourth-order accurate when $q \in C^4(\mathbb{R})$. This is proved as a Corollary of the following Proposition A.1.

Proposition A.1. *Let $q \in C^4(\mathbb{R})$, $\overline{x} \in \mathbb{R}$ and $h > 0$. Then, the following identity holds:*

$$q(\overline{x}) = \frac{7}{12}\left(\frac{1}{h}\int_{\overline{x}}^{\overline{x}+h} q(x)\,dx + \frac{1}{h}\int_{\overline{x}-h}^{\overline{x}} q(x)\,dx\right) - \frac{1}{12}\left(\frac{1}{h}\int_{\overline{x}+h}^{\overline{x}+2h} q(x)\,dx + \frac{1}{h}\int_{\overline{x}-2h}^{\overline{x}-h} q(x)\,dx\right) + C_1h^4, \quad \text{(A.32)}$$

where C_1 is a constant that depends on q and h.

Proof. We define $Q(x) = \int_a^x q(\xi)\, d\xi$ for fixed $a \in \mathbb{R}$ as in Equation (2.36). It follows that:

$$\int_{\overline{x}}^{\overline{x}+h} q(\xi)\, d\xi + \int_{\overline{x}-h}^{\overline{x}} q(\xi)\, d\xi = Q(\overline{x}+h) - Q(\overline{x}-h),$$
$$\int_{\overline{x}+h}^{\overline{x}+2h} q(\xi)\, d\xi + \int_{\overline{x}-2h}^{\overline{x}-h} q(\xi)\, d\xi = Q(\overline{x}+2h) - Q(\overline{x}-2h) - (Q(\overline{x}+h) - Q(\overline{x}-h)).$$

Using these identities, Equation (A.32) may be rewritten as:

$$q(\overline{x}) = \frac{4}{3}\left(\frac{Q(\overline{x}+h) - Q(\overline{x}-h)}{2h}\right) - \frac{1}{3}\left(\frac{Q(\overline{x}+2h) - Q(\overline{x}-2h)}{4h}\right) + C_1 h^4, \tag{A.33}$$

which consists of finite-difference approximations. Thus, Equation (A.32) follows from Lemma A.1 with:

$$C_1 = C_1(\mu_1, \mu_2) = \frac{1}{720}\left(6q^{(4)}(\mu 1) - 32q^{(4)}(\mu_2)\right), \tag{A.34}$$

where $\mu_1, \mu_2 \in [\overline{x} - 2h, \overline{x} + 2h]$, which concludes the proof. □

Corollary A.2. *It follows from Proposition A.1 with $\overline{x} = x_{i+\frac{1}{2}}$ and $h = \Delta x$ that $q_{i+\frac{1}{2}}$ given by Equation (2.45) satisfies:*

$$q(x_{i+\frac{1}{2}}) - q_{i+\frac{1}{2}} = C_1 \Delta x^4, \tag{A.35}$$

with C_1 given by Equation (A.34), whenever $q \in C^4(\mathbb{R})$.

The parabolic function from (2.39) given with coefficients specified before approximates q with order 3 when $q \in C^4(\mathbb{R})$. In order to check this, for $x \in X_i$ we rewrite Equation (2.39) as:

$$q_i(x; Q) = q_{L,i} + \frac{(\Delta q_i + q_{6,i})}{\Delta x}(x - x_{i-\frac{1}{2}}) - \frac{q_{6,i}}{\Delta x^2}(x - x_{i-\frac{1}{2}})^2 \tag{A.36}$$

and we write q using its Taylor expansion assuming $q \in C^4(\mathbb{R})$:

$$q(x) = q(x_{i-\frac{1}{2}}) + q'(x_{i-\frac{1}{2}})(x - x_{i-\frac{1}{2}}) + \frac{q''(x_{i-\frac{1}{2}})}{2}(x - x_{i-\frac{1}{2}})^2 + \frac{q^{(3)}(\theta_i)}{6}(x - x_{i-\frac{1}{2}})^3, \tag{A.37}$$

where $\theta_i \in X_i$. Comparing Equation (A.36) with Equation (A.37), it is reasonable to seek to some bound to the expressions:

$$q'(x_{i-\frac{1}{2}}) - \frac{(\Delta q_i + q_{6,i})}{\Delta x}, \tag{A.38}$$

and:

$$\frac{q''(x_{i-\frac{1}{2}})}{2} - \left(-\frac{q_{6,i}}{\Delta x^2}\right). \tag{A.39}$$

We have seen that term $q_{L,i}$ gives a fourth-order approximation to $q(x_{i-\frac{1}{2}})$. The Corollary A.3 shall prove that the term (A.38) has a bound proportional to Δx^2, and the Corollary A.4 shall prove that the term (A.39) is bounded by a constant times Δx.

Before proving the desired bounds, it is useful to rewrite some terms explicitly as functions of the values of the Δx-grid function Q. Combining Equation (2.42) with Equations (2.46) and (2.47), we may write $q_{6,i}$ as:

$$q_{6,i} = \frac{1}{4}\Big(Q_{i-2} - 6Q_{i-1} + 10Q_i - 6Q_{i+1} + Q_{i+2} \Big). \tag{A.40}$$

Recalling the definition of Δq_i from Equation (2.40), and applying Equations (2.46) and (2.47), we may express Δq_i as:

$$\Delta q_i = \frac{1}{12}\Big(Q_{i-2} - 8Q_{i-1} + 8Q_{i+1} - Q_{i+2} \Big). \tag{A.41}$$

Finally, we combine Equations (A.40) and (A.41) and write their sum as:

$$\frac{(\Delta q_i + q_{6,i})}{\Delta x} = \frac{2Q_{i-2} - 13Q_{i-1} + 15Q_i - 5Q_{i+1} + Q_{i+2}}{6\Delta x}. \tag{A.42}$$

The next Proposition A.2 proves that Equation (A.42) approximates $q'(x_{i-\frac{1}{2}})$ with order 2.

Proposition A.2. *Let $q \in C^3(\mathbb{R})$, $\overline{x} \in \mathbb{R}$ and $h > 0$. Then, the following identity holds:*

$$\begin{aligned} q'(\overline{x}) = \frac{1}{6h}\bigg(\frac{2}{h}\int_{\overline{x}-2h}^{\overline{x}-h} q(x)\,dx - \frac{13}{h}\int_{\overline{x}-h}^{\overline{x}} q(x)\,dx + \frac{15}{h}\int_{\overline{x}}^{\overline{x}+h} q(x)\,dx \\ -\frac{5}{h}\int_{\overline{x}+h}^{\overline{x}+2h} q(x)\,dx + \frac{1}{h}\int_{\overline{x}+2h}^{\overline{x}+3h} q(x)\,dx \bigg) + C_2h^2, \end{aligned} \tag{A.43}$$

where C_2 is a constant that depends on q and h.

Proof. We consider again $Q(x) = \int_a^x q(\xi)\,d\xi$ for $a \in \mathbb{R}$ fixed as in Equation (2.36). Like in Proposition A.2, we have:

$$\begin{aligned} &\frac{1}{6h}\bigg(\frac{2}{h}\int_{\overline{x}-2h}^{\overline{x}-h} q(x)\,dx - \frac{13}{h}\int_{\overline{x}-h}^{\overline{x}} q(x)\,dx + \frac{15}{h}\int_{\overline{x}}^{\overline{x}+h} q(x)\,dx - \frac{5}{h}\int_{\overline{x}+h}^{\overline{x}+2h} q(x)\,dx + \frac{1}{h}\int_{\overline{x}+2h}^{\overline{x}+3h} q(x)\,dx \bigg) \\ &= \frac{1}{6h}\bigg(\frac{2}{h}\Big(Q(\overline{x}-h) - Q(\overline{x}-2h)\Big) - \frac{13}{h}\Big(Q(\overline{x}) - Q(\overline{x}-h)\Big) + \frac{15}{h}\Big(Q(\overline{x}+h) - Q(\overline{x})\Big) \\ &\qquad -\frac{5}{h}\Big(Q(\overline{x}+2h) - Q(\overline{x}+h)\Big) + \frac{1}{h}\Big(Q(\overline{x}+3h) - Q(\overline{x}+2h)\Big) \bigg) \\ &= \frac{1}{6h^2}\bigg(-2Q(\overline{x}-2h) + 15Q(\overline{x}-h) - 28Q(\overline{x}) + 20Q(\overline{x}+h) - 6Q(\overline{x}+2h) + Q(\overline{x}+3h) \bigg), \end{aligned}$$

which consists of the finite-difference scheme from Lemma A.2. Therefore, Equation (A.43) follows from Lemma A.2 with:

$$C_2 = C_2(\mu_1, \mu_2) = \frac{1}{24}\Big(128q^{(3)}(\mu_1) - 116q^{(3)}(\mu_2) \Big), \tag{A.44}$$

where $\mu_1, \mu_2 \in [x_0 - 2h, x_0 + 3h]$, which concludes the proof. $\square$

Corollary A.3. *It follows from Proposition A.2 with $\overline{x} = x_{i-\frac{1}{2}}$ and $h = \Delta x$ that Δq_i given by Equation (A.41) and $q_{6,i}$ given by Equation (A.40) satisfy:*

$$q'(x_{i-\frac{1}{2}}) - \frac{(\Delta q_i + q_{6,i})}{\Delta x} = C_2 \Delta x^2, \tag{A.45}$$

with C_2 given by Equation (A.44), whenever $q \in C^3(\mathbb{R})$.

Now, we analyse the following expression:

$$-\frac{2q_{6,i}}{\Delta x^2} = -\frac{1}{2\Delta x^2}\Big(Q_{i-2} - 6Q_{i-1} + 10Q_i - 6Q_{i+1} + Q_{i+2} \Big). \tag{A.46}$$

deduced from Equation (A.40) and we prove in Proposition A.3 that Equation (A.46) approximates $q''(x_{i-\frac{1}{2}})$ with order 1.

Proposition A.3. *Let $q \in C^3(\mathbb{R})$, $\overline{x} \in \mathbb{R}$ and $h > 0$. Then, the following identity holds:*

$$\begin{aligned} q''(\overline{x}) = \frac{1}{2h^2}\Bigg(&-\frac{1}{h}\int_{\overline{x}-2h}^{\overline{x}-h} q(x)\,dx + \frac{6}{h}\int_{\overline{x}-h}^{\overline{x}} q(x)\,dx - \frac{10}{h}\int_{\overline{x}}^{\overline{x}+h} q(x)\,dx \\ &+\frac{6}{h}\int_{\overline{x}+h}^{\overline{x}+2h} q(x)\,dx - \frac{1}{h}\int_{\overline{x}+2h}^{\overline{x}+3h} q(x)\,dx \Bigg) + C_3 h, \end{aligned} \tag{A.47}$$

where C_3 is a constant that depends on q and h.

Proof. Similarly to Proposition A.2 using the same function Q, we have:

$$\begin{aligned} &\frac{1}{2h^2}\Bigg(-\frac{1}{h}\int_{\overline{x}-2h}^{\overline{x}-h} q(x)\,dx + \frac{6}{h}\int_{\overline{x}-h}^{\overline{x}} q(x)\,dx - \frac{10}{h}\int_{\overline{x}}^{\overline{x}+h} q(x)\,dx + \frac{6}{h}\int_{\overline{x}+h}^{\overline{x}+2h} q(x)\,dx - \frac{1}{h}\int_{\overline{x}+2h}^{\overline{x}+3h} q(x)\,dx \Bigg) \\ &\quad = \frac{1}{2h^2}\Bigg(-\frac{1}{h}\big(Q(\overline{x}-h) - Q(\overline{x}-2h)\big) + \frac{6}{h}\big(Q(\overline{x}) - Q(\overline{x}-h)\big) - \frac{10}{h}\big(Q(\overline{x}+h) - Q(\overline{x})\big) \\ &\qquad + \frac{6}{h}\big(Q(\overline{x}+2h) - Q(\overline{x}+h)\big) - \frac{1}{h}\big(Q(\overline{x}+3h) - Q(\overline{x}+2h)\big)\Bigg) \\ &\quad = \frac{1}{2h^3}\Big(Q(\overline{x}-2h) - 7Q(\overline{x}-h) + 16Q(\overline{x}) - 16Q(\overline{x}+h) + 7Q(\overline{x}+2h) - Q(\overline{x}+3h) \Big), \end{aligned}$$

which consists of the finite-difference scheme from Lemma A.3. Therefore, Equation (A.47) follows from Lemma A.3 with:

$$C_3 = C_3(\mu_1, \mu_2) = \frac{1}{48}\Big(104 q^{(3)}(\mu_1) - 128 q^{(3)}(\mu_2) \Big), \tag{A.48}$$

where $\mu_1, \mu_2 \in [x_0 - 2h, x_0 + 3h]$, which concludes the proof. $\square$

Corollary A.4. *It follows from Proposition A.3 with $\overline{x} = x_{i-\frac{1}{2}}$ and $h = \Delta x$ that $q_{6,i}$ given by*

Equation (2.45) satisfies:

$$q''(x_{i-\frac{1}{2}}) - \left(-\frac{2q_{6,i}}{\Delta x^2} \right) = C_3 \Delta x, \tag{A.49}$$

with C_3 given by Equation (A.48), whenever $q \in C^3(\mathbb{R})$.

With the aid of Corollaries A.2, A.3, and A.4, we are able to prove that the PPM reconstruction approximates q with order 3. Indeed, we prove this on the follow up Proposition A.4.

Proposition A.4. *Let $q \in C^4([a,b])$. Then, the Piecewise-Parabolic function given by Equation (2.39) with the parameters $q_{R,i}$ and $q_{L,i}$ obeying Equations (2.46) and (2.47) gives a third-order approximation to q on the control volume X_i. Namely, there exist constants M_1 and M_2 such that*

$$|q(x) - q_i(x; Q)| \le M_1 \Delta x^4 + M_2 \Delta x^3, \quad \forall x \in X_i.$$

Proof. For $x \in X_i$, from Equations (A.37) and (A.36), we have:

$$\begin{aligned} q(x) - q_i(x; Q) = {} & (q'(x_{i-\frac{i}{2}}) - q_{L,i}) + \left(q'(x_{i-\frac{1}{2}}) - \frac{(\Delta q_i + q_{6,i})}{\Delta x} \right)(x - x_{i-\frac{1}{2}}) \\ & + \left(\frac{q''(x_{i-\frac{1}{2}})}{2} + \frac{q_{6,i}}{\Delta x^2} \right)(x - x_{i-\frac{1}{2}})^2 + \frac{q^{(3)}(\theta_i)}{6}(x - x_{i-\frac{1}{2}})^3. \end{aligned}$$

Using this fact with Corollaries A.2, A.3, and A.4, we have:

$$q(x) - q_i(x; Q) = C_1 \Delta x^4 + C_2 \Delta x^2 (x - x_{i-\frac{1}{2}}) + \frac{C_3}{2} \Delta x (x - x_{i-\frac{1}{2}})^2 + C_4 (x - x_{i-\frac{1}{2}})^3,$$

where C_1, C_2 and C_3 are given by Equations (A.34), (A.44) and (A.48), respectively, and

$$C_4 = C_4(\theta_i) = \frac{q^{(3)}(\theta_i)}{6}. \tag{A.50}$$

For $x \in X_i$, we have $|x - x_{i-\frac{i}{2}}| \le \Delta x$, thus:

$$|q(x) - q_i(x; Q)| \le M_1 \Delta x^4 + M_2 \Delta x^3,$$

where

$$M_1 = \frac{38}{720} \sup_{\xi \in [a,b]} |q^{(4)}(\xi)|,$$

$$M_2 = \left(\frac{244}{24} + \frac{232}{96} + \frac{1}{6} \right) \sup_{\xi \in [a,b]} |q^{(3)}(\xi)| = \frac{143}{12} \sup_{\xi \in [a,b]} |q^{(3)}(\xi)|,$$

which concludes the proof. $\square$

Appendix B

Code avaibility

The codes needed for this work have been built openly at GitHub. The PPM implementation for the one-dimensional advection equation used in Chapter 2 is available at https://github.com/luanfs/FV3_adv_1D.

References

Arakawa, A., & Lamb, V. R. (1977). Computational design of the basic dynamical processes of the ucla general circulation model. In *General circulation models of the atmosphere* (pp. 173–265, Vol. 17). Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-12-460817-7.50009-4 (cit. on p. 5).

Carpenter, R. L., Droegemeier, K. K., Woodward, P. R., & Hane, C. E. (1990). Application of the piecewise parabolic method (ppm) to meteorological modeling. *Monthly Weather Review*, *118*(3), 586–612. https://doi.org/10.1175/1520-0493(1990)118<0586:AOTPPM>2.0.CO;2 (cit. on pp. 4, 15).

Colella, P., & Woodward, P. R. (1984). The piecewise parabolic method (ppm) for gas-dynamical simulations. *Journal of Computational Physics*, *54*(1), 174–201. https://doi.org/https://doi.org/10.1016/0021-9991(84)90143-8 (cit. on pp. 3, 15–18).

Courant, R., & John, F. (1999). In *Introduction to calculus and analysis i*. Springer Berlin, Heidelberg. https://doi.org/https://doi.org/10.1007/978-3-642-58604-0 (cit. on p. 20).

Durran, D. (2011). Time discretization: Some basic approaches. In *Numerical techniques for global atmospheric models* (pp. 75–104). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-11640-7_5 (cit. on p. 14).

Durran, D. R. (2010). Semi-lagrangian methods. In *Numerical methods for fluid dynamics: With applications to geophysics* (pp. 357–391). Springer New York. https://doi.org/10.1007/978-1-4419-6412-0_7 (cit. on p. 14).

Engwirda, D., & Kelley, M. (2016). A weno-type slope-limiter for a family of piecewise polynomial methods. https://doi.org/10.48550/ARXIV.1606.08188 (cit. on pp. 4, 6, 16).

Godunov, S. (1959). A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. *Mat. Sb.*, *47(89):3*, 271–306 (cit. on pp. 3, 4, 16).

Guo, W., Nair, R. D., & Qiu, J.-M. (2014). A conservative semi-lagrangian discontinuous galerkin scheme on the cubed sphere. *Monthly Weather Review*, *142*(1), 457–475. https://doi.org/10.1175/MWR-D-13-00048.1 (cit. on p. 14).

Harris, L., Chen, X., Putman, W., Zhou, L., & Chen, J.-H. (2021). A scientific description of the gfdl finite-volume cubed-sphere dynamical core. *Series : NOAA technical memorandum OAR GFDL ; 2021-001*. https://doi.org/https://doi.org/10.25923/6nhs-5897 (cit. on pp. 3, 4, 18).

Harris, L. M., & Lin, S.-J. (2013). A two-way nested global-regional dynamical core on the cubed-sphere grid. *Monthly Weather Review*, *141*(1), 283–306. https://doi.org/10.1175/MWR-D-11-00201.1 (cit. on p. 3).

Lauritzen, P. H., Ullrich, P. A., & Nair, R. D. (2011). Atmospheric transport schemes: Desirable properties and a semi-lagrangian view on finite-volume discretizations. In P. Lauritzen, C. Jablonowski, M. Taylor, & R. Nair (Eds.), *Numerical techniques for global atmospheric models* (pp. 185–250). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-11640-7_8 (cit. on p. 4).

Lauritzen, P. H. (2007). A stability analysis of finite-volume advection schemes permitting long time steps. *Monthly Weather Review*, *135*(7), 2658–2673. https://doi.org/https://doi.org/10.1175/MWR3425.1 (cit. on pp. 27, 28).

Leonard, B. P., Lock, A. P., & MacVean, M. K. (1996). Conservative explicit unrestricted-time-step multidimensional constancy-preserving advection schemes. *Monthly Weather Review*, *124*(11), 2588–2606. https://doi.org/https://doi.org/10.1175/1520-0493(1996)124<2588:CEUTSM>2.0.CO;2 (cit. on p. 3).

LeVeque, R. J. (1985). A large time step generalization of godunov's method for systems of conservation laws. *SIAM Journal on Numerical Analysis*, *22*(6), 1051–1073. https://doi.org/10.1137/0722063 (cit. on p. 3).

LeVeque, R. J. (1990). *Numerical methods for conservation laws*. Birkhäuser Basel. https://doi.org/10.1007/978-3-0348-5116-9 (cit. on p. 7).

LeVeque, R. J. (2002). *Finite volume methods for hyperbolic problems*. Cambridge University Press. https://doi.org/10.1017/CBO9780511791253 (cit. on pp. 7, 15, 23, 25).

Lin, S.-J. (2004). A "vertically lagrangian" finite-volume dynamical core for global models. *Monthly Weather Review*, *132*(10), 2293–2307. https://doi.org/10.1175/1520-0493(2004)132<2293:AVLFDC>2.0.CO;2 (cit. on pp. 4, 13, 18).

Lin, S.-J., & Rood, R. B. (1996). Multidimensional flux-form semi-lagrangian transport schemes. *Monthly Weather Review*, *124*(9), 2046–2070. https://doi.org/10.1175/1520-0493(1996)124<2046:MFFSLT>2.0.CO;2 (cit. on pp. 3, 13, 15, 28).

Lin, S.-J., & Rood, R. B. (1997). An explicit flux-form semi-lagrangian shallow-water model on the sphere. *Quarterly Journal of the Royal Meteorological Society*, *123*(544), 2477–2498. https://doi.org/https://doi.org/10.1002/qj.49712354416 (cit. on pp. 3, 13).

Lu, F., Zhang, F., Wang, T., Tian, G., & Wu, F. (2022). High-order semi-lagrangian schemes for the transport equation on icosahedron spherical grids. *Atmosphere*, *13*(11). https://doi.org/10.3390/atmos13111807 (cit. on p. 14).

Putman, W. M. (2007). *Development of the finite-volume dynamical core on the cubed-sphere* [Doctoral dissertation, Florida State University]. Florida, US. http://purl.flvc.org/fsu/fd/FSU_migr_etd-0511 (cit. on p. 3).

Putman, W. M., & Lin, S.-J. (2007). Finite-volume transport on various cubed-sphere grids. *Journal of Computational Physics*, *227*(1), 55–78. https://doi.org/https://doi.org/10.1016/j.jcp.2007.07.022 (cit. on pp. 4, 13).

Stoer, J., & Bulirsch, R. (2002). In *Introduction to numerical analysis*. Springer New York, NY. https://doi.org/https://doi.org/10.1007/978-0-387-21738-3 (cit. on pp. 15, 19).

Strikwerda, J. C. (2004). *Finite difference schemes and partial differential equations, second edition*. Society for Industrial; Applied Mathematics. https://doi.org/10.1137/1.9780898717938 (cit. on p. 26).

Trefethen, L. N. (2000). *Spectral methods in matlab*. Society for Industrial; Applied Mathematics. https://doi.org/10.1137/1.9780898719598 (cit. on p. 26).

Tumolo, G. (2011). *A semi-implicit, semi-lagrangian, p-adaptative discontinuous galerkin method for the rotating shallow-water equations: Analysis and numerical experiments* [Doctoral dissertation, University of Trieste]. https://core.ac.uk/download/pdf/41173373.pdf (cit. on p. 14).

Van Leer, B. (1977). Towards the ultimate conservative difference scheme. iv. a new approach to numerical convection. *Journal of Computational Physics*, *23*(3), 276–299. https://doi.org/https://doi.org/10.1016/0021-9991(77)90095-X (cit. on pp. 4, 15, 16).

Wesseling, P. (2001). Scalar conservation laws. In *Principles of computational fluid dynamics* (pp. 339–396). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-05146-3_9 (cit. on p. 4).

White, L., & Adcroft, A. (2008). A high-order finite volume remapping scheme for nonuniform grids: The piecewise quartic method (pqm). *Journal of Computational Physics*, *227*(15), 7394–7422. https://doi.org/https://doi.org/10.1016/j.jcp.2008.04.026 (cit. on p. 4).

Woodward, P. R. (1986). Piecewise-parabolic methods for astrophysical fluid dynamics. In K.-H. A. Winkler & M. L. Norman (Eds.), *Astrophysical radiation hydrodynamics* (pp. 245–326). Springer Netherlands. https://doi.org/10.1007/978-94-009-4754-2_8 (cit. on p. 4).