Πανεπιστημιο Κρητης

Σχολη Θετικων και Τεχνολογικων Επιστημων

Τμημα Εφαρμοσμενων Μαθηματικων

Διδακτορικη Διατριβη

# Σχηματα Πεπερασμενων Διαφορων με Αυτοματη Αναδιαμεριση για Υπερβολικους Νομους Διατηρησης

Νικολαος Σφακιανακης

Εποπτεύων Καθηγητής:

Καθ. Χαραλαμπος Μακριδακης
Τμ. Εφαρμοσμενων Μαθηματικων, Πανεπιστημιο Κρητης

Ηρακλειο, Φεβρουαριος 2009

University of Crete

Faculty of Science and Engineering

Department of Applied Mathematics

Doctoral Thesis

# Finite Difference schemes on non-Uniform meshes for Hyperbolic Conservation Laws

Nikolaos Sfakianakis

Thesis advisor:
Prof. Charalambos Makridakis
Dept. Applied Mathematics, University of Crete

Heraklion, February 2009

*to my father,*
*whose questions are the hardest*

# Acknowledgements

I would like to express my deepest appreciation to the professors of the department of Mathematics and the department of Applied Mathematics of University of Crete for their constant support throughout the years of my studies. They showed me the way of tackling with the problems of Mathematics and for this, I will always be gratetful.

In addition I would like to thank Prof. Benoît Perthame of the University Pierre et Marie Curie Paris, and Prof. Christian Schmeiser of the University of Vienna who both assisted greatly in the final stages of this work.

I wish to thank my family, my parents Iosif and Sevasti, and especially my sisters Rodoula and Anna for they always believed in me and always encouraged me. They were always there for me, in good and in bad times, and I will never forget it.

Despoina Triantafyllidou deserves special mention. Her dedication and persistent confidence in me, was in times, the sole reason that I kept on trying. I would never had succeeded in completing this work without her support.

I would like to express my gratitude to my thesis advisor, Prof. Charalambos Makridakis, for he served constantly as a role model to me, both as a mathematician and as a person. I will always cherish the small and big moments that he shared with me. Thank you.

This work was financially supported by the Maria M. Manasaki Foundation and by the Greek State Scholarship Foundation (IKY) during my Ph.D studies in the University of Crete. It was also supported by the Marie Curie research training networks *"Modelling, mathematical methods and computer simulations of tumour growth and therapy"* and *"DEASE: Differential Equations with Applications in Science and Engineering"* during my studies at École Normal Supérieure in Paris and at the University of Vienna, respectively.

# Περίληψη

Στην εργασία αυτή μελετάμε Σχήματα Πεπερασμένων διαφορών σε μονοδιάστατα, μη ομοιόμορφα, αναδρομικά οριζόμενα πλέγματα. Συνδυάζουμε τις βασικές ιδιότητες προσέγγισης συναρτήσεων σε μη-ομοιόμορφα πλέγματα, με την ανακατασκευή του πλέγματος, τη χωρική ανανέωση της διακριτής λύσης πάνω στο νέο πλέγμα, με την χρονική ανανέωση ανανέωση της λύσης, χρησιμοποιώντας σχήματα Πεπερασμένων Διαφορών σχεδιασμένα ειδικά για μη-ομοιόμορφα πλέγματα. Τα βήματα αυτά ορίζουν το Βασικό Αναδρομικό Σχήμα. Επιπλέον αναλύουμε τις ιδιότητες του Βασικού Αναδρομικού Σχήματος ως ον αφορά στην Συνολική του Κύμανση και παρέχουμε τα θεωρητικά αποτελέσματα της δουλειάς αυτής.

Αναλυτικότερα: μελετάμε τις βασικές ιδέες των προσεγγίσεων Πεπερασμένων Διαφορών σε μη-ομοιόμορφα πλέγματα. Αναλύουμε τις ιδιότητες τους και συγκρίνουμε τα ποιοτικά χαρακτηριστικά τους με τις αντίστοιχες προσεγγίσεις σε ομοιόμορφα πλέγματα.

Έπειτα, παρουσιάζουμε την μέθοδο ανακατασκευής του πλέγματος, που χρησιμοποιούμε στην εργασία αυτή. Εξηγούμε το τρόπο με τον οποίο κατασκευάζουμε το νέο μη-ομοιόμορφο πλέγμα, βασιζόμενοι σε γεωμετρικές ιδιότητες της αριθμητικής λύσης και στο υπάρχουν μη-ομοιόμορφο πλέγμα. Περιγράφουμε τα συναρτησοειδή που είναι υπεύθυνα για την ανακατασκευή αυτή και παρουσιάζουμε τις ιδιότητες τους. Σχέσεις με άλλες μεθόδους ανακατασκευής πλέγματος δίνονται υπό την μορφή αναφορών. Έπειτα παρουσιάζουμε την την διαδικασία με την οποία η αριθμητική λύση επαναπροσδιορίζεται στο νέο μη-ομοιόμορφο πλέγμα. Αναλύουμε χαρακτηριστικές ιδιότητες όπως η διατήρη της μάζας και η αρχή μεγίστου.

Προχωράμε, έπειτα, στο βήμα της χρονική ανανέωση του Βασικού Αναδρομικού Σχήματος. Αναφέρουμε μερικά γνωστά και μερικά νέα αριθμητικά σχήματα, όλα σχεδιασμένα ειδικά για μη-ομοιόμορφα πλέγματα. Παρατηρούμε ότι κάποια εξάυτών είναι ταυτόσημα όταν το πλέγμα είναι ομοιόμορφο. Αναλύουμε μερικές απο τις ιδιότητες τους όπως: Συνέπεια, Ευστάθεια, Ακρίβεια χρησιμοποιώντας την Δραστική Εξίσωση του σχήματος ως βασικό εργαλείο. Κατά τη διάρκεια της ανάλυσης αυτής ανακαλύπτουμε ότι το κλασσικό κριτήριο συνέπειας για σχήματα Πεπερασμένων Διαφορών δεν είναι ικανή συνθήκη για να εξασφαλίσει την συνέπεια του σχήματος όταν το πλέγμα είναι μη-ομοιόμορφο. Για το λόγο αυτό προτείνουμε μία γενίκευση του κριτηρίου αυτού ως ικανή συνθήκη για την συνέπεια του σχήματος, εφαρμόσιμη και στην περίπτωση του ομοιόμορφο όσο και στην περίπτωση του μη-ομοιόμορφο πλέγματος. Έπειτα παρουσιάζουμε τα αποτελέσματα μιας σειράς αριθμητικών

πειραμάτων, όπου και συγκρίνουμε τις ιδιότητες ευστάθειας και ακρίβειας σχημάτων πάνω από ομοιόμορφο και μη-ομοιόμορφο πλέγμα. Τέλος, μελετάμε την Συνολική Κύμανση του Βασικού Αναδρομικού Σχήματος όταν χρησιμοποιούνται αριθμητικά σχήματα που παράγουν ταλαντώσεις. Αποδεικνύουμε ότι κάτω από συγκεκριμένες συνθήκες η Συνολική Κύμανση είναι φραγμένη και επιπλέον (κάτω από αυστηρότερες συνθήκες) ότι η αύξηση της Κύμανσης εξάιτίας των ταλαντώσεων μειώνεται με το χρόνο.

Η εργασία αυτή οδήγησε στην κατάθεση τριών ερευνητικών άρθρων σε επιστημονικά περιοδικά.

# Abstract

In this work we consider Finite Difference numerical schemes over 1 dimensional non-uniform, adaptively redefined meshes. We combine the basic properties of function approximation over non-uniform mesh with a mesh reconstruction, the spatial solution update over the new mash and with the time evolution with Finite Difference schemes designed for non-uniform meshes. All these steps constitute the Basic Adaptive Scheme. We moreover analyse the Total Variation properties of the Basic Adaptive Scheme and provide the theoretical results of this work.

In more details: we investigate the basic notions of Finite Difference approximation over non-uniform meshes. We discuss their properties and compare their qualitative characteristics with the respective approximations over uniform meshes.

We then discuss the mesh reconstruction procedure that we use throughout this work. We explain the way the new non-uniform mesh is constructed based on geometric properties of the numerical solution itself, and on the already existing non-uniform mesh. We describe the functionals responsible for this mesh reconstruction and present their properties. Relations with other non-uniform mesh methods are provided in the form of references. Afterwards, we present the process with which the numerical solution is updated/redefined over the new non-uniform mesh. Characteristic properties like, conservation of mass and maximum principle during this process are discussed and analysed.

Next, we move to the time evolution part of the Basic Adaptive Scheme. We discuss some known and some new numerical schemes, both designed for non-uniform meshes. We notice that some of them reduce to the same numerical scheme when the mesh is uniform; hence we name the numerical schemes under consideration according to their uniform counterparts. We analyse some of their properties like consistency, stability and order of accuracy using, mainly, their modified equations as our tool. Through this process we discover that the usual consistency criterion for Finite Difference scheme is not sufficient when the mesh is non-uniform; hence we provide a generalisation of the consistency criterion valid also for non-uniform meshes. Then a series of numerical tests is conducted. Comparisons between the non-uniform vs uniform mesh case exhibit both the stabilisation properties of the Basic Adaptive Scheme and the higher accuracy that

can be achieved when non-uniform mesh is used. These tests are conducted using the numerical schemes that were previously discussed as well as elaborate Entropy Conservative numerical schemes.

Finally, we discuss the Total Variation of Basic Adaptive Schemes when oscillatory (either dispersive or anti-diffusive) Finite Difference schemes are used for the time evolution step. We prove under specific assumptions that the Total Variation of such schemes remains bounded, and even more (under more strict assumptions) that the increase of their Total Variation decreases with time.

This work has led to the submission of three journal essays.

# Contents

# Introduction

Conservation Laws is a very important class of Evolutionary Partial Differential Equations. They describe phenomena where one or more measurable properties of an isolated system does not change as the system evolves. They arise in a wide range of sciences from Particle Physics and Quantum Mechanics and Gas Dynamics to Traffic Flow, Biology and Economics.

Through our study we consider the scalar Conservation Law (CL) in one space dimension,

$$\frac{\partial}{\partial t}u(x,t) + \frac{\partial}{\partial x}f(u(x,t)) = 0, \quad x \in \mathbb{R},\ t \in [0,T]$$

where $u : \mathbb{R} \times [0,T] \to \mathbb{R}$ is the Conserved quantity/property and $f : \mathbb{R} \to \mathbb{R}$ is the Flux function. The flux function $f$ is considered to be smooth, convex, with $f(0) = 0$.

We are interested in numerical approximations of the solutions of the scalar CL. To this end discrete approximations have to be considered for both the spatial and temporal derivatives. Regarding the spatial derivatives we note that there are two main ways of discretization, i.e over a Uniform Grid where the distances of the mesh points is constant or over a Non-Uniform Grid where the distances of the mesh points is not constant.

In this work we focus on methods where the spatial discretisations are taken non-uniform and vary with time. In summary such a method is decomposed in the following steps:

**Definition** (Basic Adaptive Scheme)**.** Given approximations $U^n = \{u_1^n, \ldots, u_N^n\}$, at the time step $t = t^n$ which belong to a finite dimensional space $V^n$

1. (Mesh Reconstruction)
   Choose the next space $V^{n+1}$

2. (Solution Update)
   Project $U^n$ to the new space $V^{n+1}$ to get $\hat{U}^n$,

3. (Time Evolution)
   Use $\hat{U}^n$ as starting value to perform the evolution step in $V_{n+1}$ resulting the new approximation $U^{n+1}$.

The Basic Adaptive Scheme (BAS) describes the simplest scenario for adaptivity in evolution problems.

One of the main features in our work is to analyse the steps of the BAS and study the effect they have on the numerical approximation of solution of scalar Conservation

Laws. To justify the work that follows we just mention that this kind of procedures -when specific Mesh Reconstruction (Step 1) takes place- exhibit strong stabilisation properties.

We start in the first Chapter by restating the BAS in the frame of Finite Differences. We describe in details the use and the necessity of the steps of BAS in the case of non-uniform grids. We moreover discuss the derivative approximations on non-uniform meshes and compare them with the respective ones for the uniform mesh case.

In the second Chapter we start the discussion and analysis of the BAS steps. We describe in details the Mesh Reconstruction part (Step 1) and the several ways the Solution Update part (Step 2) can be implemented. In this part we also discuss 3 different kind of grids namely Cell centered, Vertex centered and Finite Element grids and for the moment we note that their differences are amplified in the non-uniform case.

In the third Chapter we visit numerical schemes that are responsible for the Time Evolution part (Step 3) of the BAS. We discuss known and new schemes over uniform mesh and their non-uniform counterparts. We also analyse their behaviour using as a main tool their Modified Equations. We confirm that the usual Consistency Criterion for Conservative schemes is not sufficient for the non-uniform case and provide a generalisation of this criterion valid for both uniform and non-uniform grids. We moreover discuss the class of Entropy Conservative schemes and provide evidence of the stabilisation properties of the BAS. We also provide numerical tests on the numerical schemes examined in this Chapter for both uniform and non-uniform grids.

The literature on the subject and the analysis performed in the previous Chapter points towards the fact that the BAS possesses stabilisation properties not emanating from the Evolutionary part (Step 3). So in the fourth Chapter we study the stabilisation property of the BAS. We set specific requirements on the Mesh Reconstruction and the Time Evolution and we provide an explanation regarding the suppression of oscillations that the BAS exhibits. We prove that under these requirements the Total Variation of oscillatory schemes (either dispersive or anti-diffusive) is kept bounded and under more strict requirements the Total Variation increase of oscillatory schemes diminishes. So we conclude in a Total Variation Increase Diminishing (TVID) behaviour on behalf of the BAS. We also provide numerical results supporting the previous discussion.

# Chapter 1

# Non-uniform Meshes & Derivative Approximations

As mentioned in the introductory Chapter, we are interested in Finite Difference approximations of the numerical solutions over non-uniform grids -in the space variable. Since the problems that we study are evolutionary, we shall moreover consider meshes that vary with time.

The procedures that we will study, take into account the non-uniformity of the mesh, the adaptation of the mesh and the time evolution of the numerical approximations. These comprise the *Basic Adaptive Scheme* (BAS) which will be described in the first section of this Chapter.

Moreover we will examine in this Chapter, the ways derivatives of functions can be approximated by Finite Differences over non-uniform meshes. The results will be compared with the respective approximations obtained by Finite Differences on uniform meshes.

## 1.1 *Basic Adaptive Scheme*

In contrast to the uniform mesh case, in the non-uniform, the scheme consists of three intermediate procedures in each time step. In this section we present them in order to get a broad picture of the work that will follow.

The problem that we are interested is the scalar Conservation Law in one space dimension,

$$u_t + f(u)_x = 0, \quad x \in [a, b],\ t \in [0, T],$$

with initial data having compact support in a much smaller interval.

We discretise in time with the finite sequence $\{0 = t^0 < \cdots < t^n < \cdots < t^K = T\}$ and in every time step $t = t^n$ we consider a *non-uniform mesh*

$$M_x^n = \{a = x_1^n < \cdots < x_N^n = b\}$$

with *variable* space steps $h_i^n = x_{i+1}^n - x_i^n$ for $i = 1, \ldots, N-1$. Moreover we consider approximations $U^n$ (point or average) of the exact solution $u$ and we denote,

$$U^n = \{u_1^n, \ldots, u_N^n\}$$

The BAS is divided in three steps,

1. (Mesh Reconstruction) Given the mesh $M_x^n$ and the approximation $U^n$ we construct

a *new mesh*

$$M_x^{n+1} = \{a = x_1^{n+1} < \cdots < x_N^{n+1} = b\}$$

based on *information* attained by the discrete approximation $U^n$. This construction is achieved with the use of the Adaptive Mesh Reconstruction procedure that we shall present in a following Chapter.

2. (Solution Update) We note at this point that the discrete approximations $U^n$ are related to the mesh $M_x^n$. Similarly we need new approximations $\hat{U}^n$ that will be related to the new mesh $M_x^{n+1}$. To this end we utilise $U^n$ to construct a Piecewise Constant or Piecewise Linear function -defined over the whole domain of our study- $V^n(x)$ that interpolates the discrete approximations $U^n$, that is

$$V^n(x_i) = U_i^n, \quad \text{for } i = 1, \ldots, N.$$

We can now evaluate the updated approximations

$$\hat{U}^n = \{\hat{u}_1^n, \ldots, \hat{u}_N^n\}$$

defined in the new mesh $M_x^{n+1}$ by either interpolation or by a mass conservation procedure.

3. (Time Evolution) The final step is to evolve in time the updated approximations $\hat{U}^n$, to get the new approximations $U^{n+1}$ for the time step $t = t^{n+1}$. This part is achieved by the use of the Numerical Scheme we have selected.

The procedure/algorithm we described is repeated in the next time step with new mesh $M_x^{n+1}$ and new discrete approximations $U^{n+1}$.

Since we shall often refer to the BAS, it will be useful to rewrite it in a more compact form:

**Definition 1.1.1** (BAS)**.** Given mesh $M_x^n = \{a = x_1^n < \cdots < x_N^n = b\}$ and approximations $U^n = \{u_1^n, \ldots, u_N^n\}$,

1. (Mesh Reconstruction)
   Given $M_x^n$ and $U^n$ construct new mesh $M_x^{n+1} = \{a = x_1^{n+1} < \cdots < x_N^{n+1} = b\}$

2. (Solution Update)
   Given $M_x^n$, $U^n$ and $M_x^{n+1}$

2a. construct the function $V^n(x)$ with $V^n(x_i^n) = u_i^n$

2b. compute/update approximations $\hat{U}^n = \{\hat{u}_i^n, \ldots, \hat{u}_N^n\}$

3. (Time Evolution)
Given $M_x^{n+1}$, $\hat{U}^n$ march in time to compute $U^{n+1} = \{u_1^{n+1}, \ldots, u_N^{n+1}\}$

*Remark* 1.1.1. A first comment is that in the uniform mesh case the mesh does not change with time (does not adapt) so there is no need for construction of a new mesh $M_x^{n+1}$ (no Step 1) hence no need for a solution update (no Step 2). In this case the BAS (1.1.1) reduces to just the time evolution (Step 3) which is the usual uniform numerical scheme. We shall see in later Chapters that the extra steps of the non-uniform BAS complicate significantly both the computation and the analysis of the numerical approximations and their properties.

The rest of this Chapter along with Chapters 3 and 4 deals with the analysis of the Steps of the BAS. The first issue to deal with is the nature of the non-uniform spatial meshes and the way derivatives of smooth functions can be approximated over these meshes.

## 1.2 Approximations on non-uniform mesh

The rest of this Chapter is devoted to the study of non-uniform meshes -described in the BAS (1.1.1) as $M_x^n$ and $M_x^{n+1}$. More specifically we are interested in the way derivatives can be approximated using a non-uniform mesh. We shall focus on 3-point or 3-cell partitions of the domain and the tools that we shall use are the method of Variable Coefficients for the construction of the approximations and Taylor expansions for the discussion of the order of accuracy.

Two sections comprise the rest of this Chapter. In the first we consider Point Values and we construct approximations for the first and second order derivatives on uniform and non-uniform meshes. We compare the approximations and discuss the order of accuracy for both cases. In the second section we consider Cell Averages where again we discuss and compare the derivative approximations on both uniform and non-uniform meshes.

## 1.3 Three-point derivative approximation

We consider a non-uniform 3-point mesh i.e.,

$$M_x = \{a < b < c\}$$

and we denote by $h_1 = b - a$ and $h_2 = c - b$ the variable space steps. We assume throughout this section that $h_1, h_2 > 0$.

In addition we consider a sufficiently smooth function $u(\cdot)$ that attains the values

$$\{u(a), u(b), u(c)\}$$

over the mesh $M_x$. It is this function whose derivatives at the point $x = b$ we wish to approximate, i.e., we need discrete approximations for the derivatives $u'(b)$ and $u''(b)$.

This is achieved by the method of Variable Coefficients, that is, we need to find coefficients $\alpha$, $\beta$ and $\gamma$ depending only on $h_1$ and $h_2$ -and not on $u$- so as the quantity

$$A = \alpha u(a) + \beta u(b) + \gamma u(c) \tag{1.1}$$

to approximate either $u'(b)$ or $u''(b)$. The coefficients $\alpha$, $\beta$, $\gamma$ can be regarded as the *weights* of the values $u(a)$, $u(b)$, $u(c)$ in the approximation of these derivatives. The derivative approximations that we shall construct on non-uniform meshes $h_1 \neq h_2$ should coincide with the existing approximations whenever the mesh is uniform, $h_1 = h_2$. We refer to Fornberg [10] for a thorough discussion on derivative approximations on non-uniform meshes.

*Remark* 1.3.1. As a preliminary remark we note that in the uniform mesh case the order of accuracy of the derivatives approximations is (in general) higher than in the non-uniform mesh case. The reason for this accuracy decrease is the lack of symmetry of the non-uniform meshes.

To start with the approximations, we state again that we assume that the function $u(\cdot)$ is sufficiently smooth and we expand the values $u(a)$ and $u(c)$ in Taylor series about

the point $x = b$,

$$u(a) = u(b) - h_1 u'(b) + \frac{h_1^2}{2}u''(b) - \frac{h_1^3}{6}u'''(b) + \frac{h_1^4}{24}u^{(4)}(b) + \mathcal{O}(h_1^5)$$
$$u(c) = u(b) + h_2 u'(b) + \frac{h_2^2}{2}u''(b) + \frac{h_2^3}{6}u'''(b) + \frac{h_2^4}{24}u^{(4)}(b) + \mathcal{O}(h_2^5)$$

where $h_1 = b - a$ and $h_2 = c - b$. One more assumption is that the nodes $a$, $b$, $c$ are *close* in the sense $h_1, h_2 << 1$ and that the fifth order derivatives of $u(\cdot)$ are uniformly bounded. We substitute these Taylor expansions in the quantity $A$ Rel.(1.1), which recasts into

$$\begin{aligned} A = (\alpha+\beta+\gamma)u(b) + (-\alpha h_1 + \gamma h_2)\,u'(b) + \left(\alpha\frac{h_1^2}{2} + \gamma\frac{h_2^2}{2}\right)u''(b) \\ + \left(-\alpha\frac{h_1^3}{6} + \gamma\frac{h_2^3}{6}\right)u'''(b) + \left(\alpha\frac{h_1^4}{24} + \gamma\frac{h_2^4}{24}\right)u^{(4)}(b) + \mathcal{O}(h_1^5, h_2^5) \end{aligned} \tag{1.2}$$

We note that linear combinations of the weights $\alpha$, $\beta$ and $\gamma$ constitute the coefficients for the derivatives in the new quantity $A$, Rel.(1.2). We denote by $\sigma_0, \ldots, \sigma_4$ the coefficients of the derivatives, $u(b), \ldots, u^{(4)}(b)$ -respectively- and we set the following system,

$$(\Sigma): \begin{cases} \sigma_0 &= \alpha+\beta+\gamma & \text{coefficient of} & u(b) \\ \sigma_1 &= -\alpha h_1 + \gamma h_2 & \text{coefficient of} & u'(b) \\ \sigma_2 &= \alpha\frac{h_1^2}{2} + \gamma\frac{h_2^2}{2} & \text{coefficient of} & u''(b) \\ \sigma_3 &= -\alpha\frac{h_1^3}{6} + \gamma\frac{h_2^3}{6} & \text{coefficient of} & u'''(b) \\ \sigma_4 &= \alpha\frac{h_1^4}{24} + \gamma\frac{h_2^4}{24} & \text{coefficient of} & u^{(4)}(b) \end{cases} \tag{1.3}$$

As we already stated, the objective is to choose the coefficients $\alpha$, $\beta$, $\gamma$ in such a way that the quantity $A$, Rel.(1.2) approximates either $u'(b)$ or $u''(b)$. The general idea is to set the coefficient of the derivative we wish to approximate equal to 1 and as many as possible of the other coefficients $\sigma_i$, $i = 0, \ldots, 4$ equal to 0. In each case a linear system -with respect to $\alpha$, $\beta$, $\gamma$- will occur.

Before we start with the approximate derivative construction some remarks are in order,

**Proposition 1.3.1** (For both uniform and non-uniform meshes)**.** *It is impossible to eliminate the first and the second derivative, at the same time, from the derivative approximation A Rel.(1.2)*

*The same is true for the second and the third as well as for the third and the fourth derivatives.*

*Proof.* To eliminate the first and the second derivative from the approximation $A$, we have to set both $\sigma_1 = 0$ and $\sigma_2 = 0$, that is

$$\begin{cases} -\alpha h_1 + \gamma h_2 & = 0 \\ \alpha \frac{h_1^2}{2} + \gamma \frac{h_2^2}{2} & = 0, \end{cases}$$

which yields $\alpha = \gamma = 0$. Hence the initial form Rel.(1.1) of the approximation $A$ reduces to

$$A = \beta u(b)$$

which does not constitute a derivative approximation.

The same result follows if we try to eliminate the second and the third derivative from $A$ or the third and the fourth derivatives. □

The following proposition discusses the most common drawback one faces when dealing with non-uniform meshes. The usual cancellation of the higher derivatives in centered Taylor expansions that takes place on uniform meshes fails in the case of non-uniform meshes.

**Proposition 1.3.2** (For non-uniform meshes)**.** *In a non-uniform mesh i.e.,., $h_1 \neq h_2$ it is impossible to eliminate the first and the third derivative, at the same time, from the derivative approximation $A$. The same is true for the second and the fourth derivative.*

*Proof.* To eliminate the first and the third derivative from the approximation $A$, we have to set both $\sigma_1 = 0$ and $\sigma_3 = 0$, that is

$$\begin{cases} -\alpha h_1 + \gamma h_2 & = 0 \\ -\alpha \frac{h_1^3}{6} + \gamma \frac{h_2^3}{6} & = 0, \end{cases}$$

this yields $\alpha = \gamma = 0$ since we have assumed that the mesh is non-uniform. So, the initial form of the approximation $A$ Rel.(1.1) reduces to,

$$A = \beta u(b)$$

which is not a derivative approximation.

Similarly we prove that the we cannot eliminate the second and the fourth derivative from the approximation $A$ at the same time. □

In contrast to the non-uniform mesh case, in the uniform we have elimination of higher derivatives without extra requirements.

**Proposition 1.3.3** (For uniform meshes)**.** *In a uniform mesh, i.e., $h_1 = h_2 = h$ the first and the third derivative are eliminated, at the same time, from the derivative approximation $A$. The same is true for the second and the fourth derivative.*

*Proof.* In the case of a uniform mesh i.e., $h_1 = h_2 = h$, the derivative approximation $A$ reads

$$A = (\alpha+\beta+\gamma)u(b) + (-\alpha+\gamma)\,hu'(b) + (\alpha+\gamma)\,\frac{h^2}{2}u''(b) \\ + (-\alpha+\gamma)\,\frac{h^3}{6}u'''(b) + (\alpha+\gamma)\,\frac{h^4}{24}u^{(4)}(b) + \mathcal{O}(h^5)$$

and the respective coefficient system $(\Sigma)$ becomes,

$$(\Sigma): \begin{cases} \sigma_0 &= \alpha+\beta+\gamma \\ \sigma_1 &= (-\alpha+\gamma)h \\ \sigma_2 &= (\alpha+\gamma)\frac{h^2}{2} \\ \sigma_3 &= (-\alpha+\gamma)\frac{h^3}{6} \\ \sigma_4 &= (\alpha+\gamma)\frac{h^4}{24} \end{cases}$$

It is now obvious that elimination of the first derivative which happens by setting $\sigma_1 = 0$ results to $\alpha = \gamma$ hence the coefficient $\sigma_3$ of the third derivative yields $\sigma_3 = 0$.

The reverse is also true, as is the case of the second and the fourth derivative. □

After these remarks we can continue by approximating selected derivatives of a sufficiently smooth function $u(\cdot)$.

### 1.3.1 Approximation of the first derivative

On approximating the first derivative, we are looking for coefficients $\alpha$, $\beta$, $\gamma$ such that $A \approx u'(b)$. To this end we try to eliminate as many terms of the derivative approximation $A$, Rel.(1.2), as we can while keeping the first derivative.

By Prop.(1.3.1) it is not possible to eliminate at the same time the second and the third derivative, so the best we can do (given that the system $(\Sigma)$ is of 3 unknowns) is

to eliminate the 0-th and the second while keeping the first derivative. So we solve the system $\sigma_0 = 0$, $\sigma_1 = 1$, $\sigma_2 = 0$, that is

$$\begin{cases} \alpha + \beta + \gamma &= 0 \\ -\alpha h_1 + \gamma h_2 &= 1 \\ \alpha \frac{h_1^2}{2} + \gamma \frac{h_2^2}{2} &= 0, \end{cases}$$

which yields

$$\alpha = -\frac{h_2}{h_1(h_1+h_2)}, \quad \gamma = \frac{h_1}{h_2(h_1+h_2)}, \quad \beta = -\alpha - \gamma.$$

For this solution we notice that $\sigma_3 = \frac{h_1 h_2}{6}$ and so the derivative approximation $A$, Rel.(1.2) recasts as follows

$$\begin{aligned} A &= -\frac{h_2}{h_1(h_1+h_2)}u(a) + \left(\frac{h_2}{h_1(h_1+h_2)} - \frac{h_1}{h_2(h_1+h_2)}\right)u(b) + \frac{h_1}{h_2(h_1+h_2)}u(c) \\ &\approx u'(b) + \frac{h_1 h_2}{6}u'''(b). \end{aligned}$$

This means that the approximation $A$ of the first derivative $u'(b)$ we just described is of second order accuracy, in the sense $\mathcal{O}(h_1 h_2)$. This is the best we can do with a 3-point scheme since higher accuracy would demand elimination of more derivative coefficients, which as we explained is impossible.

*Remark* 1.3.2. For comparison purposes we solve the same system i.e., $\sigma_0 = 0$, $\sigma_1 = 1$, $\sigma_2 = 0$ for the uniform mesh case $h_1 = h_2 = h$ as well. The solution this time is

$$\alpha = -\frac{1}{2h}, \quad \gamma = \frac{1}{2h}, \quad \beta = -\alpha - \gamma = 0,$$

which yields $\sigma_3 = \frac{h^2}{6}$ and the derivative approximation Rel.(1.2) recasts into

$$A = -\frac{1}{2h}u(a) + \frac{1}{2h}u(c) \approx u'(b) + \frac{h^2}{6}u'''(b),$$

that is, the quantity $A$ provides a second order approximation to the first derivative $u'(b)$.

Comparing the uniform with the non-uniform case we see that in order to get a second order approximation of the first derivative in the uniform case it suffices to utilise just two values, i.e., $u(a)$, $u(c)$, whereas in the non-uniform one we need all the three values $u(a)$, $u(b)$, $u(c)$. In other words, because of the loss of the symmetry of the mesh

we are forced to use more information (3 points instead of 2) in order to maintain second order approximation in the non-uniform mesh case.

### 1.3.2 Approximation of the second derivative

On approximating the second derivative, we are looking for coefficients $\alpha$, $\beta$, $\gamma$ such that $A \approx u''(b)$. To this end we eliminate the 0-th and the first derivative and we keep the second, we do not hope for more eliminations since the system $(\Sigma)$ is of 3 unknowns. So we solve the system $\sigma_0 = 0$, $\sigma_1 = 0$, $\sigma_2 = 1$,

$$\begin{cases} \alpha + \beta + \gamma &= 0 \\ -\alpha h_1 + \gamma h_2 &= 0 \\ \alpha \frac{h_1^2}{2} + \gamma \frac{h_2^2}{2} &= 1, \end{cases}$$

which yields

$$\alpha = \frac{2}{h_1(h_1+h_2)}, \quad \gamma = \frac{2}{h_2(h_1+h_2)}, \quad \beta = -\alpha - \gamma.$$

For this solution we note that $\sigma_3 = \frac{h2-h1}{3}$ and $\sigma_4 = \frac{h_1^2 - h_1 h_2 + h_2^2}{12}$ and so the derivative approximation $A$, Rel.(1.2), reads as follows

$$\begin{aligned} A &= \frac{2}{h_1(h_1+h_2)}u(a) + \left(-\frac{2}{h_1(h_1+h_2)} - \frac{2}{h_2(h_1+h_2)}\right)u(b) + \frac{2}{h_2(h_1+h_2)}u(c) \\ &\approx u''(b) + \frac{h_2 - h_1}{3}u'''(b) + \frac{h_1^2 - h_1 h_2 + h_2^2}{12}u^{(4)}(b). \end{aligned}$$

This means that the approximation $A$ of the second derivative $u''(b)$ we just described is of first order accuracy.

*Remark* 1.3.3. For comparison purposes we solve the same system $\sigma_0 = 0$, $\sigma_1 = 0$, $\sigma_2 = 1$ for the uniform case, which yields

$$\alpha = \frac{1}{h^2}, \quad \gamma = \frac{1}{h^2}, \quad \beta = -\frac{2}{h^2}$$

and hence $\sigma_3 = 0$ and $\sigma_4 = \frac{h^2}{12}$. So the derivative approximation $A$, Rel.(1.2), reads

$$A = \frac{1}{h^2}u(a) - \frac{2}{h^2}u(b) + \frac{1}{h^2}u(c) \approx u''(b) + \frac{h^2}{12}u^{(4)}(b)$$

that is, the quantity $A$ provides a second order approximation to the derivative $u''(b)$. We again notice that the loss of symmetry of the non-uniform mesh is responsible for

the decrease of the order of accuracy of the derivative.

The last remark introduces the following proposition,

**Proposition 1.3.4.** *With a 3-point non-uniform mesh ($h_1 \neq h_2$), the approximation of the second derivative can be at most of second order.*

*Proof.* In order to construct a second order approximation of the second derivative in a 3-point mesh, one has to solve the system $\sigma_0 = 0$, $\sigma_1 = 0$, $\sigma_2 = 1$ and $\sigma_3 = 0$.

According to the Remark (1.3.2) the solution of the system $\sigma_1 = 0$ and $\sigma_3 = 0$ yields $\gamma = 0$ and $\alpha = 0$. Also, along with $\sigma_0 = 0$, we get $\alpha = \beta = \gamma = 0$, and so the approximation in the case of a non-uniform mesh shall be $A = 0$, which of course does not constitute a derivative approximation. □

# Chapter 2

# Adaptive Mesh Reconstruction & Solution Update

We start this Chapter by restating the BAS, whose steps we need to analyse,

**Definition** (BAS)**.** Given mesh $M_x^n = \{a = x_1^n < \cdots < x_N^n = b\}$ and approximations $U^n = \{u_1^n, \ldots, u_N^n\}$,

1. (Mesh Reconstruction)
   Given $M_x^n$ and $U^n$ construct new mesh $M_x^{n+1} = \{a = x_1^{n+1} < \cdots < x_N^{n+1} = b\}$

2. (Solution Update)
   Given $M_x^n$, $U^n$ and $M_x^{n+1}$

   2a. construct the function $V^n(x)$ with $V^n(x_i^n) = u_i^n$

   2b. compute/update approximations $\hat{U}^n = \{\hat{u}_i^n, \ldots, \hat{u}_N^n\}$

3. (Time Evolution)
   Given $M_x^{n+1}$, $\hat{U}^n$ march in time to compute $U^{n+1} = \{u_1^{n+1}, \ldots, u_N^{n+1}\}$

So far we have discussed the initiation of the BAS, that is the properties of the underlying meshes (uniform and non-uniform) and discrete approximations of derivatives. In this Chapter we will start with the study of the Steps that comprise the BAS. First we will discuss the Mesh Reconstruction procedure that is responsible for the Step 1 of the BAS and we shall continue with the Solution Update procedure, Step 2 of the BAS.

In more detail, in the first section of this Chapter we discuss the basic ingredients of the Mesh Reconstruction procedure, which is responsible for the first Step of the BAS. In every time step $t = t^n$ the geometry of the discrete approximation $U^n$ is analysed with the use of two auxiliary functions the *Estimator* and the *Monitor* function. The accumulated information of the Monitor function is used to define a new mesh $M_x^{n+1}$ via an equi-distribution principle. In every step we count the number of the necessary computations.

In the second section we examine the different ways a discrete approximation $U^n$, defined over a finite set of nodes $M_x^n$, can be extended over the whole spatial domain to give rise to the function $V^n(x)$ -that is Step 2a. of the BAS. We moreover discuss the ways, a solution approximation $V^n(x)$ along with the new mesh $M_x^{n+1}$ yields the updated discrete approximations $\hat{U}^n$ which are now related to the new mesh $M_x^{n+1}$ -that is Step 2b. of the BAS.

## 2.1 Mesh Reconstruction procedure

We present in this section the Mesh Reconstruction procedure that sponsors the construction and manipulation of the *adaptive*, *non-uniform* mesh. The term *non-uniform* stands for a varying spatial step and the term *adaptive* implies that the mesh varies with time, according to the Step 1 of BAS.

The use of non-uniform adaptively redefined meshes has been studied in the past, we mention for instance the work of Harten and Hyman[12], Dorfi and Drury[9] and Tang and Tang[28], among others.

The approach that we will follow was first introduced by Arvanitis, Katsaounis and Makridakis[3] and by Arvanitis in his PhD Thesis[2]. This approach differs from the previous ones, in the sense that it utilises geometric information attained from the numerical solution and redistributes a fixed number of nodes according to an equi-distribution principle. Properties of the Mesh Reconstruction procedure have been studied in a series of papers [5], [1], [6], [4], [24].

The Mesh Reconstruction procedure is actually a way of relocating the nodes of the mesh according to the geometric *information* contained in the discrete numerical solution. The basic idea of the Mesh Reconstruction procedure is simple and geometric

*in areas where the numerical solution is smoother/flatter we need less nodes, in the contrary in areas where the numerical solution is less smooth/flat more nodes are in order*

The key aspect of this procedure is the way that we measure the geometric *information* of the discrete function. This is accomplished by using two auxiliary functions, the *Estimator* and the *Monitor* function. The estimator function measures the *geometric information* of the discrete function and the monitor function redistributes the nodes according to the information measured by the estimator function.

Some examples of geometric estimator functions we have used are the arclength estimator, the gradient estimator and the curvature estimator. Following [3] and [6] throughout this work we shall use the curvature as our estimator.

### 2.1.1 The curvature estimator function

To start with, we consider a smooth function $u$.

**Definition 2.1.1** (Curvature of a smooth function $u$)**.** Let $u$ be a smooth function. The curvature estimator $K_u$ of $u$ at every point $x$ of the domain is given by

$$K_u(x) = \frac{|u''(x)|}{(1 + (u'(x))^2)^{3/2}}$$

We refer to the section (A) of the Appendix for a discussion on the derivation of the previous definition.

We want to measure the curvature of the discrete approximations $U^n$. These values constitute a finite sequence of numbers so a discrete analog of the curvature function defined in Def.(2.1.1) is needed.

To gain a discrete analog of the curvature estimator $K_u$ we just have to discretize the derivatives that appear in the smooth estimator (2.1.1). Several choices are possible, but the one that we use throughout this work is the following,

**Definition 2.1.2** (Discrete curvature of discrete approximations $U^n$)**.** For the discrete approximation $U^n = \{u_i^n, i = 1, \ldots, N\}$ defined over the mesh $M_x^n = \{x_i^n, i = 0, \ldots, N\}$ we define the discrete curvature $K_i^{dscr}$ by

$$K_i^{dscr} = \frac{\frac{2}{x_{i+1}-x_{i-1}} \left| \frac{u_i^n - u_{i-1}^n}{x_i^n - x_{i-1}^n} - \frac{u_{i+1}^n - u_i^n}{x_{i+1}^n - x_i^n} \right|}{\left( \left( 1 + \left( \frac{u_i^n - u_{i-1}^n}{x_i^n - x_{i-1}^n} \right)^2 \right) \left( 1 + \left( \frac{u_{i+1}^n - u_i^n}{x_{i+1}^n - x_i^n} \right)^2 \right) \left( 1 + \left( \frac{u_{i+1}^n - u_{i-1}^n}{x_{i+1}^n - x_{i-1}^n} \right)^2 \right) \right)^{1/2}}$$

for every $i = 1, \ldots, N$.

By performing a point by point evaluation of the discrete curvature on the approximations $U^n = \{u_i^n, i = 1, \ldots, N\}$ we construct a finite sequence,

$$K_{U^n}^{dscr} = \left\{ (x_1^n, K_1^{dscr}), \ldots, (x_N^n, K_N^{dscr}) \right\}$$

that contains the measured geometric (curvature in this case) information of the discrete approximation $U^n$.

We continue with a proposition that relates the discrete with the smooth curvature estimator,

**Proposition 2.1.1.** *For a smooth function $u$ such that $u(x_i, t^n) = u_i^n$, the discrete curvature, defined in Def.(2.1.2) approximates the smooth curvature, defined in Def.(2.1.1).*

*Proof.* We start by performing Taylor expansions (up to the fourth derivative) on the values $u_{i+1}^n$ and $u_{i-1}^n$, hence the discrete curvature (2.1.2) yields

$$\frac{|u'' - \frac{h_1-h_2}{3}u''' + \frac{h_1^2-h_1h_2+h_2^2}{12}u''''|}{\left(1+\left(-u' + \frac{h_1-h_2}{2}u'' + \frac{h_1^2-h_1h_2+h_2^2}{6}u'''\right)^2\right)^{1/2}} \cdot$$

$$\frac{1}{\left(\left(1+\left(-u' + \frac{h_1}{2}u'' - \frac{h_1^2}{6}u''' + \frac{h_1^3}{24}u''''\right)^2\right)\left(1+\left(u' + \frac{h_2}{2}u'' - \frac{h_2^2}{6}u''' + \frac{h_2^3}{24}u''''\right)^2\right)\right)^{1/2}},$$

where to simplify the notation, we have set $h_1 = x_i - x_{i-1}$ and $h_2 = x_{i+1} - x_i$ and where all functions are evaluated at the point $x = x_i$. By assuming now that the higher derivatives of $u$, i.e., $u''$, $u'''$, $u''''$ are bounded and that both $h_1$ and $h_2$ are of order $\mathcal{O}(h)$ the previous fraction can be written in the form

$$\frac{|u'' + \mathcal{O}(h)|}{(1+u'^{\,2})^{3/2} + \mathcal{O}(h)}$$

or, since the $\mathcal{O}(h)$ function at the denominator is positive,

$$\frac{|u''|}{(1+u'^{\,2})^{3/2}} + \mathcal{O}(h)$$

So, the discrete curvature $K_i^{dscr}$ defined in Def.(2.1.2) constitutes a first order approximation of the smooth curvature $K_u(x_i^n)$ defined in Def.(2.1.1) □

*Remark* 2.1.1. Of course, we can construct second order approximations of the smooth curvature estimator by performing different discretizations in the derivatives that appear in the smooth estimator $K_u(x)$.

*Remark* 2.1.2. We note here that the proposed Mesh Reconstruction procedure produces positive discrete curvature values,

$$K_i^{dscr} \geq 0 \quad \text{for every } i$$

*Remark* 2.1.3 (A possible scenario of a linear Approximate Solution). In this case $K_i^{dscr} = 0$ for every $i$ which -as we shall see- causes problems in the definition of the Monitor function. To avoid such case we provide each node with a minimum amount of discrete

curvature *information* $\varepsilon > 0$, so from now on we assume that

$$K_i^{dscr} \geq \varepsilon, \quad \text{for every } i$$

In practice $\varepsilon$ is chosen to be a very small positive number.

*Remark* 2.1.4. Another possible scenario is that of abrupt changes in the values of the discrete estimator. This is the case if a numerical solution approximates a discontinuity, see for example Figure 2.1. In this case the last node before the jump -let it be $x_i^n$- measures huge value of discrete curvature $K_i^{dscr} >> 1$, whereas its left neighbour $x_{i-1}^n$ measures $K_{i-1}^{dscr} = \varepsilon$.

To get a smoothly varying mesh, such abrupt information changes should be avoided, so we raise every $K_i^{dscr}$ value in a predefined power $pw$ with $0 < pw < 1$. This will result in smoother discrete estimator values and hence in smoother discrete estimator function. So from now on we assume that

$$K_i^{dscr} = (K_i^{dscr})^{pw}.$$

A typically value of $pw$ is 0.9.

**Proposition 2.1.2.** *The computational cost for the sequence* $\{K_i^{dscr},\ i = 0, \ldots, N\}$ *is* $\mathcal{O}(N)$.

Before we move on to the monitor function, we construct the final version of the estimator function, that is we define it over the whole domain of dependence.

Let a discrete approximation $U^n = \{u_i^n, i = 1, \ldots, N\}$ be defined over the mesh $M_x^n = \{x_i^n, i = 1, \ldots, N\}$ with discrete curvature $K_{U^n}^{dscr} = \{(x_i^n, K_i^{dscr}), i = 1, \ldots, N\}$ as defined in Def.(2.1.2) and altered in remarks Rem.(2.1.3) and Rem.(2.1.4).

**Definition 2.1.3** (Estimator Function)**.** We define the final curvature estimator $K_{U^n}(\cdot)$ to be the continuous piecewise linear function that interpolates the values $K_{U^n}^{dscr}$,

$$K_{U^n}(x_i^n) = K_i^{dscr}$$

Graphical examples of the final estimator functions are exhibited in Figure (2.1).

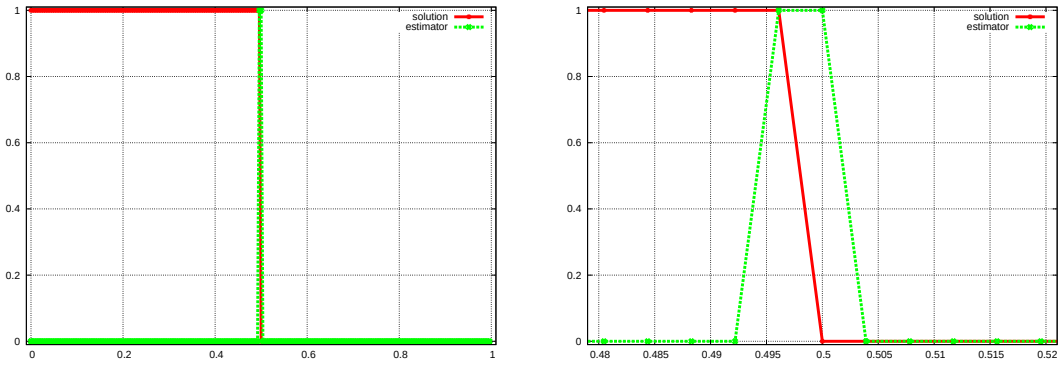



Figure 2.1: A typical estimator function is depicted along with the respective numerical solution. The right graph is a focused version of the left one.

### 2.1.2 The monitor function

The estimator function, we presented in the previous paragraph, measures the geometric information of the discrete approximation $U^n$. In order to construct a new mesh we need another function that will exploit this information. This is the role of the Monitor function and in simple words it is the definite integral of the final estimator function $K_{U^n}$ which we constructed in the previous paragraph.

The monitor function has -as the estimator function had- both a discrete and a continuous version. The construction of the monitor function starts from its discrete analog, that is we first evaluate the discrete monitor function in every old node $x_i^n$ and then we construct the continuous case by linear interpolation of the discrete monitor values.

To start with, we integrate the piecewise linear function $K_{U^n}(x)$ to find the respective value of the discrete monitor function in every node $x_i^n$,

**Definition 2.1.4** (Discrete monitor function)**.** For every $i$ we define the discrete monitor function values

$$M_i^{dscr} = \int_0^{x_i^n} K_{U^n}(x)dx. \tag{2.1}$$

This results in a sequence of discrete values

$$\left\{(x_i^n, M_i^{dscr}), i = 0, \ldots, N)\right\}.$$

Obviously,

$$M_{i+1}^{dscr} - M_i^{dscr} = \int_{x_i^n}^{x_{i+1}^n} K_{U^n}(x)dx$$

and since $K_{U^n}(x) > 0$ for every $x$ in the domain, the sequence $M_i^{dscr}$ is positive and strictly increasing with respect to $i$.

Before we move to the construction of the continuous monitor function, we state the following proposition that counts the number of operations needed for the evaluation of the discrete monitor function. The proof is obvious and is omitted.

**Proposition 2.1.3.** *The computational cost of evaluating the sequence* $\{M_i^{dscr},\ i = 0, \ldots, N\}$ *using the trapezoidal rule in the Equation (2.1) is* $3N$

That was the discrete version of the monitor function. To get the continuous version we simply interpolate in a piecewise linear manner the values $M_i^{dscr}$ of the discrete monitor function.

**Definition 2.1.5** (Continuous monitor function)**.** The continuous monitor function $M_{U^n}(x)$ is the piecewise linear interpolant of the points $\left\{(x_i^n, M_i^{dscr}), i = 0, \ldots, N)\right\}$ i.e.,

$$M_{U^n}(x_i^n) = M_i^{dscr}$$

We note that $M_{U^n}(x)$ is obviously continuous -as piecewise linear interpolant- , positive and strictly increasing -since the sequence $M_i^{dscr}$ is strictly increasing with respect to $i$, so it attains its maximum at the right end of our domain $M_{U^n}(x_N^n)$.

We shall only need these properties of the continuous monitor function so we do not write an explicit formula. Examples of monitor functions are exhibited in Figure (2.2).

We are now ready to compute the new mesh $M_x^{n+1}$.

### 2.1.3 Mesh reconstruction

What we now need is a new set of nodes, $M_x^{n+1} = \{x_i^{n+1}, i = 0, \ldots, N\}$, with $x_0^{n+1} = 0$. The exact position of the new nodes is adapted by the information contained in the monitor function $M_{U^n}(x)$. There are several ways to do so, but we propose that the new nodes are such that they equi-distribute the total information of the monitor function.

*Remark* 2.1.5. The principle of equi-distribution of the nodes is not new in the literature of moving meshes, we mention for instance the works of Dorfi and Drury [9] and Tang and Tang [28].

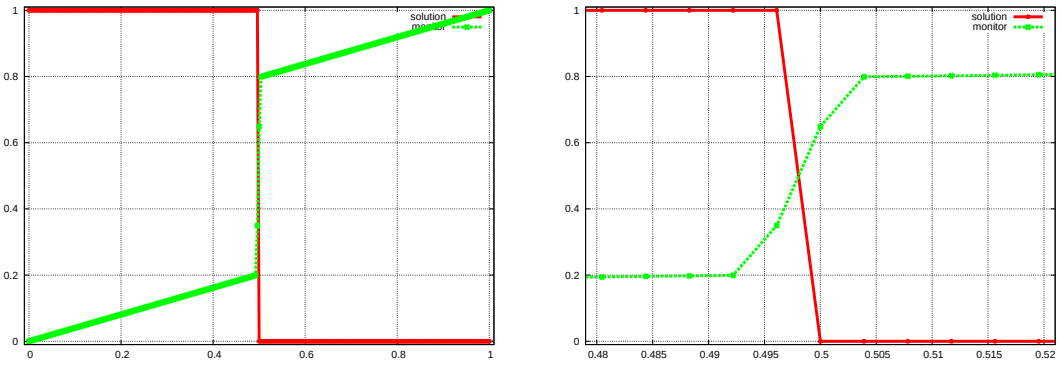



Figure 2.2: A typical monitor function is depicted along with the respective numerical solution. The right graph is a focused version of the left one.

To equi-distribute the total gathered information one has to solve -recursively with respect to $x_{i+1}^{n+1}$- the system,

$$\begin{cases} x_0^{n+1} = x_0^n, \\ M(x_{i+1}^{n+1}) - M(x_i^{n+1}) = \frac{1}{N} M_N^{dscr}, \quad i = 0, \dots, N-1. \end{cases} \tag{2.2}$$

It is obvious that the last new node $x_N^{n+1}$ coincides with the old end of the interval $x_N^n$. The system (2.2) implies an inversion of the function $M_{U^n}(x)$. To this end we have to notice that this is possible since $M_{U^n}(x)$ is a strictly increasing function and also that the computational cost for the inversion is minimum since $M_{U^n}(x)$ is piecewise linear. Hence, we rewrite the system (2.2) in the following equivalent form,

$$\begin{cases} M_i^{dscr} = \int_0^{x_i^n} K_{U^n}(x)dx, \\ M_{U^n}(x_i^{n+1}) = \frac{i}{N} M_N^{dscr}, \quad i = 0, \dots, N, \end{cases} \tag{2.3}$$

where we need to solve with respect to $x_i^{n+1}$ for $i = 0, \dots, N$. This is done in the following way,

*Construction.* For every $i = 0, \dots, N$ we find the unique interval $[x_k^n, x_{k+1}^n)$ such that

$$M_{U^n}(x_k^n) \leq \frac{i}{N} M_N^{dscr} \leq M_{U^n}(x_{k+1}^n)$$

by a comparison of $\frac{i}{N} M_N^{dscr}$ with the sequence $M_{U^n}(x_i^n)$, $i = 0, \dots, N$. This is the interval where the new node $x_i^{n+1}$ should be placed. For the exact position of this

new node we exploit the fact that $M_{U^n}(x)$ is piecewise linear (by construction) in every interval $[x_i^n, x_{i+1}^n)$, $i = 0, \ldots, N-1$. The solution of the second equation of (2.3) is given by

$$x_i^{new} = x_k^n + (x_{k+1}^n - x_k^n)\frac{\frac{i}{N}M_N^{dscr} - M_{U^n}(x_k^n)}{M_{U^n}(x_{k+1}^n) - M_{U^n}(x_k^n)}. \tag{2.4}$$

**Proposition 2.1.4.** *For a given piecewise linear monitor function $M_{U^n}(x)$ as previously described, the computational cost of computing the whole sequence $\{x_i^{n+1},\ i = 0, \ldots, N\}$ is $4N$ multiplications and at most $N$ comparisons.*

*Proof.* The number $4N$ of the multiplications is obvious. Regarding the comparisons, let us enumerate the intervals defined by the old mesh $x_i^n$ and denote by $p_i$ the serial number of the old interval $[x_k^n, x_{k+1}^n)$ in which the new node $x_i^{n+1}$ shall be placed. Let us also denote by $k_i$ the number of comparisons needed for this decision. Since the new nodes $x_i^{n+1}$ constitute an increasing sequence with respect to $i = 0, \ldots, N$ we can write the following relations that connect the position of every new node with the number of comparisons needed,

$$\begin{aligned} p_2 &= k_2 \\ p_3 &= p_2 + k_3 \\ p_4 &= p_3 + k_4 \\ &\vdots \\ p_{N-1} &= p_{N-2} + k_{N-1}. \end{aligned}$$

Immediate summation gives us that

$$\sum_{i=2}^{N-1} k_i = p_{N-1} \leq N$$

So the overall cost of finding the new nodes -for a given piecewise linear monitor function $M_{U^n}(x)$ as previously defined- is $\mathcal{O}(N)$ □

*Remark* 2.1.6. Let us note that the number of nodes remains the same during the reconstruction of the mesh. Let us also note that the previous procedure is not related to the numerical scheme that we use for the evolution part of the problem. So, continuous

repetitions of the relocation procedure do not affect the evolution part of the algorithm.

## 2.2 Solution Update

First we restate the BAS that comprises the steps that we need to analyse,

**Definition** (BAS)**.** Given mesh $M_x^n = \{a = x_1^n < \cdots < x_N^n = b\}$ and approximations $U^n = \{u_1^n, \ldots, u_N^n\}$,

1. (Mesh Reconstruction)
   Given $M_x^n$ and $U^n$ construct new mesh $M_x^{n+1} = \{a = x_1^{n+1} < \cdots < x_N^{n+1} = b\}$

2. (Solution Update)
   Given $M_x^n$, $U^n$ and $M_x^{n+1}$

   2a. construct the function $V^n(x)$ with $V^n(x_i^n) = u_i^n$

   2b. compute/update approximations $\hat{U}^n = \{\hat{u}_i^n, \ldots, \hat{u}_N^n\}$

3. (Time Evolution)
   Given $M_x^{n+1}$, $\hat{U}^n$ march in time to compute $U^{n+1} = \{u_1^{n+1}, \ldots, u_N^{n+1}\}$

The work in this section focusses in the second Step of the BAS. We discuss the construction of the solution approximations $V^n(x)$ and the computation of the updated discrete approximations $\hat{U}^n$.

To start with, we note that we have at our disposal the old mesh $M_x^n = \{x_i^n, i = 0, \ldots, N\}$ on every node of which we know the old discrete approximations $U^n = \{u_i^n, i = 1, \ldots, N\}$. We also have the new mesh $M_x^{n+1} = \{x_i^{n+1}, i = 0, \ldots, N\}$ on every node of which we want to define/update new discrete approximations $\hat{U}^n = \{\hat{u}_i^n, i = 1, \ldots, N\}$.

The new mesh $M_x^{n+1}$ does not coincide -in general- with the old one $M_x^n$, so in order to find the updated approximations $\hat{U}^n$ we first need to construct a function defined over the whole domain, the Solution Approximation function $V^n(x)$. This explains the necessity of the Step 2a. of the BAS.

The construction of the approximate solution $V^n(x)$, i.e., the extension of the old discrete approximations $U^n$ to the whole domain, can be done by several ways. We mention for instance that it can be achieved by piecewise constants or by interpolating by continuous/discontinuous piecewise linears or even by continuous/discontinuous piecewise higher order polynomials etc.

The choice in general depends on the nature of the problem under consideration, on the approximation method (finite differences, finite elements), on a desired property ($L^1$ conservation, maximum principle) etc. This step, as Chapter 5 will indicate, is of crucial importance. Therefore, we will present in detail the following alternatives,

- Vertex centered grids and conservative piecewise constant reconstruction
- Cell centered grids and conservative piecewise constant reconstruction
- Finite Element grids and conservative piecewise linear reconstruction
- Finite Element grids and piecewise linear reconstruction with interpolation

### 2.2.1 Vertex centered grids - Piecewise constant - Conservative

In this paragraph we consider Piecewise Constant Solution Approximations $U^n(x)$ over a Vertex Centered computational grid. That is,

**Definition 2.2.1** (Vertex centered computational grids)**.** For a given mesh

$$M_x^n = \{x_i^n, i = 0, \dots, N\}$$

of the domain we define the midpoints $x_{i+1/2} = \frac{x_i + x_{i+1}}{2}$ for every $i = 0, \dots, N-1$. We also define the computational cells $C_i^n$,

$$C_i^n = (x_{i-1/2}^n, x_{i+1/2}^n) \quad \text{with} \quad |C_i^n| = \Delta x_i^n = x_{i+1/2}^n - x_{i-1/2}^n,$$

which constitute a *vertex-centered* partition of the domain.

*Remark* 2.2.1. On a bounded domain $[a, b]$ the first node is $x_0^n = a$ and the last node is $x_N^n = b$. The first cell is $C_0^n = [x_0^n, x_{1/2}^n]$ and the last cell is $C_N^n = (x_{N-1/2}^n, x_N^n]$, i.e., the first and the last subintervals are half cells.

*Remark* 2.2.2. The title "Vertex Centered grids" is chosen since the cells that are constructed have vertices that are centered with respect to the nodes of the mesh.

The vertex centered cells we defined are non-overlapping; moreover, for every $i = 0, \dots, N$, there is an one-to-one relation,

$$C_i^n \longleftrightarrow x_i^n \longleftrightarrow u_i^n,$$

which gives rise to the following definition of a piecewise constant function,

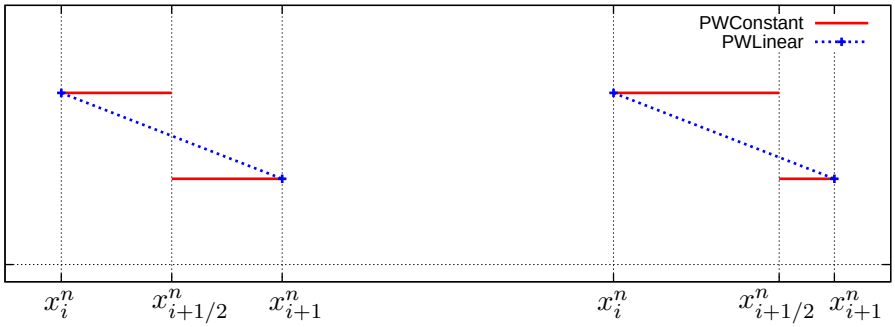



Figure 2.3: Left part: Vertex centered grids. The areas of the piecewise linear and piecewise constant as defined in Par.(2.2.1) coincide over $(x_i^n, x_{i+1}^n)$.
Right part: Cell centered grids. The areas of the piecewise linear and piecewise constant as defined in Par.(2.2.2) *do not* coincide over $(x_i^n, x_{i+1}^n)$.

**Definition 2.2.2** (Piecewise constant approximate solution $V^n(x)$)**.** Given a partition of the domain with vertex centered cells $C_i^n$ -as defined in Def.(2.2.1)- and a finite sequence of discrete approximations $U^n = \{u_i^n, i = 0, \dots N\}$, we define the approximate solution $V^n(x)$ to be the *piecewise constant function*,

$$V^n(x) = u_i^n \quad \text{for} \quad x \in C_i^n.$$

This completes the discussion on Step 2a. of the BAS. The following remark is in order,

*Remark* 2.2.3. If we had performed piecewise linear construction of the approximate solution $V_{\text{lin}}^n(x)$ -linear in the segments $(x_i^n, x_{i+1}^n)$ with $V_{\text{lin}}^n(x_i^n) = u_i^n$ and $V_{\text{lin}}^n(x_{i+1}^n) = u_{i+1}^n$ - instead of the piecewise constant $V^n(x)$ -in the cells $C_i^n = (x_{i-1/2}^n, x_{i+1/2}^n)$- the linear approximate solution $V_{lin}^n(x)$ would have had the same mass as the piecewise constant solution approximation $V^n(x)$ in the segment $(x_i^n, x_{i+1}^n)$, i.e.,

$$\int_{x_i^n}^{x_{i+1}^n} V_{\text{lin}}^n(x)dx = \int_{x_i^n}^{x_{i+1}^n} V^n(x)dx.$$

This fact is depicted in Figure 2.3 and can easily be proven since the end point $x_{i+1/2}$ of the cell $C_i^n$ is the midpoint of the interval $(x_i, x_{i+1})$.

For the Step 2b. -that is update of the discrete approximations- we incorporate the new mesh $M_x^{n+1} = \{x_i^{n+1}, i = 0, \dots, N\}$ over every node of which the new point

approximations $\hat{U}^n = \{\hat{u}_i^n, i = 0, \ldots, N\}$ have to be defined.

The new mesh $M_x^{n+1}$ gives rise to new vertex centered computational cells $C_i^{n+1}$ as in Def.(2.2.1) and the new discrete approximations $\hat{U}^n$ will give rise to a new piecewise constant approximate solution $\hat{V}^n(x)$ as in Def.(2.2.2).

In this paragraph the reconstruction procedure will be -as in [1]- conservative, so we must select each new discrete approximation $\hat{u}_i^n$, $i = 1, \ldots, N$, in such a way that the mass of the new function $\hat{V}^n(x)$ over the new cell $C_i^{n+1}$ coincides with the mass of the old function $V^n(x)$ over the same cell $C_i^{n+1}$, i.e.,

$$\int_{C_i^{n+1}} \hat{V}^n(x)dx = \int_{C_i^{n+1}} V^n(x)dx.$$

To this end we give the following definition for the new discrete approximations $\hat{U}^n = \{\hat{u}_i^n,\ i = 1, \ldots, N\}$ and then we prove that the reconstruction procedure is in fact conservative.

**Definition 2.2.3** (Conservative construction of $\hat{U}^n$)**.** We take into account the configuration of the new nodes $x_i^{n+1}$ with respect to the old cells $C_i^n$. So we have the following cases,

- If the configuration of the new nodes $x_{i-1}^{n+1}$, $x_i^{n+1}$, $x_{i+1}^{n+1}$ is such that

$$x_{i-1/2}^{n+1} \in C_k^n \quad \text{and} \quad x_{i+1/2}^{n+1} \in C_l^n, \quad \text{with } k < l,$$

  we define

$$\hat{u}_i^n = \frac{1}{\Delta x_i^{n+1}} \left( (x_{k+1/2}^n - x_{i-1/2}^{n+1})u_k^n + \sum_{j=k+1}^{l-1} \Delta x_j^n u_j^n + (x_{i+1/2}^{n+1} - x_{l-1/2}^n)u_l^n \right).$$

- If the configuration of the new nodes $x_{i-1}^{n+1}$, $x_i^{n+1}$, $x_{i+1}^{n+1}$ is such that

$$x_{i-1/2}^{n+1}, x_{i+1/2}^{n+1} \in C_k^n,$$

  we define

$$\hat{u}_i^n = u_k^n.$$

We refer to Figure 2.4 where a graphical explanation of the reconstruction is pre-

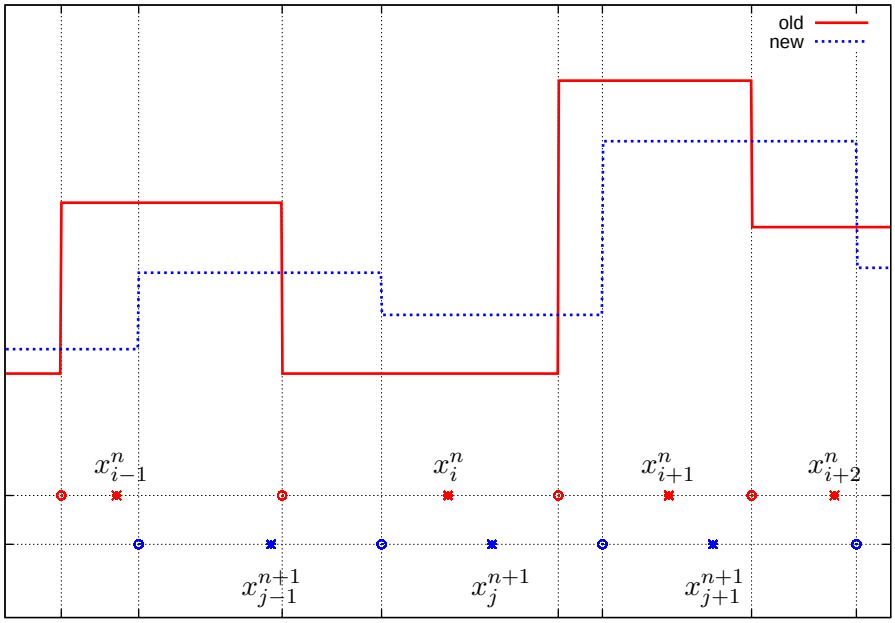



Figure 2.4: Vertex Centered grids & Conservative Piecewise Constant reconstruction. This graph depicts the old and the new mesh points $\{x_i^n, i\}$ and $\{x_j^{n+1}, j\}$ along with their middle points (with circles) that define the cells $C_i^n$ and $C_j^{n+1}$. It also depicts the old function $V^n(x)$ and its reconstruction $\hat{V}^n(x)$, both being piecewise constants over the respective cells. This reconstruction is conservative and respects the maximum principle.

sented. We now prove that the construction just described is in fact conservative.

**Proposition 2.2.1.** *The construction just defined is conservative.*

*Proof.* The area defined by the new approximation $\hat{u}^n$, over the new cell $C_i^{n+1}$, is

$$\hat{u}_i^n \Delta x_i^{n+1} = (x^n_{k+1/2} - x^{n+1}_{i-1/2})u_k^n + \sum_{j=k+1}^{l-1} \Delta x_j^n u_j^n + \underbrace{(x^{n+1}_{i+1/2} - x^n_{l-1/2})u_l^n}_{A}.$$

By writing down the following term

$$\hat{u}_{i+1}^n \Delta x_{i+1}^{n+1} = \underbrace{(x^n_{l+1/2} - x^{n+1}_{i+1/2})u_l^n}_{B} + \sum_{j=l+1}^{m-1} \Delta x_j^n u_j^n + (x^{n+1}_{i+3/2} - x^n_{m-1/2})u_m^n$$

and summing with the previous one, we notice that the terms $A$ and $B$ merge with the

sums and we end up in the right-hand side

$$(x^n_{k+1/2} - x^{n+1}_{i-1/2})u^n_k + \sum_{j=k+1}^{m-1} \Delta x^n_j u^n_j + (x^{n+1}_{i+3/2} - x^n_{m-1/2})u^n_m,$$

which is sufficient for conservation. □

We can also prove that this construction of $\hat{U}^n$ respects the Maximum Principle,

**Proposition 2.2.2** (Maximum principle)**.** *The absolute values of the new point approximations $\hat{u}^n_i$ are bounded by the maximum of the old point approximations $u^n_i$,*

$$\max_i |\hat{u}^n_i| \leq \max_i |u^n_i|$$

*Proof.* In the case where the configuration of the new nodes $x^{n+1}_{i-1}$, $x^{n+1}_i$, $x^{n+1}_{i+1}$ is such that

$$x^{n+1}_{i-1/2} \in C^n_k \quad \text{and} \quad x^{n+1}_{i+1/2} \in C^n_l, \quad \text{with } k < l,$$

the new approximation $\hat{u}^n_i$ is defined as

$$\hat{u}^n_i = \frac{1}{\Delta x^{n+1}_i} \left( (x^n_{k+1/2} - x^{n+1}_{i-1/2})u^n_k + \sum_{j=k+1}^{l-1} \Delta x^n_j u^n_j + (x^{n+1}_{i+1/2} - x^n_{l-1/2})u^n_l \right).$$

Since the intervals in the enumerator of the right-hand side sum up to $\Delta x^{n+1}_i$ we immediately bound as follows

$$|\hat{u}^n_i| \leq \max\{|u^n_j|, j = k, \ldots, l\}$$

In the case where the configuration of the new nodes $x^{n+1}_{i-1}$, $x^{n+1}_i$, $x^{n+1}_{i+1}$ is such that

$$x^{n+1}_{i-1/2}, x^{n+1}_{i+1/2} \in C^n_k$$

the new approximation $\hat{u}^n_i$ is defined as

$$\hat{u}^n_i = u^n_k$$

so again the bound

$$|\hat{u}^n_i| \leq \max\{|u^n_j|, j = k, \ldots, l\}$$

is valid □

We close this paragraph, by counting the operations needed for the proposed reconstruction procedure,

**Proposition 2.2.3.** *The computational cost for the sequence* $\{\hat{u}_i^n, i = 1, \dots, N\}$ *is* $\mathcal{O}(N)$.

*Proof.* The result is obvious, assuming that we compute the sum $\sum_{j=1}^{N} \Delta x_j^n u_j^n$ in advance. □

### 2.2.2 Cell centered grids - Piecewise constant - Conservative

In this paragraph we consider Piecewise Constant Solution Approximations $V^n(x)$ over a Cell Centered computational grid. That is,

**Definition 2.2.4** (Cell centered computational grids)**.** For a non-uniform discretization $\{0 = x_{-1/2}^n < \cdots < x_{N-1/2}^n = 1\}$ of the domain, we define the midpoints $x_i = \frac{x_{i+1/2}+x_{i-1/2}}{2}$ which constitute the mesh $M_x^n$

$$M_x = \{x_i^n, i = 1, \dots, N-1\}$$

and the cells $C_i^n$

$$C_i = (x_{i-1/2}^n, x_{i+1/2}^n) \quad \text{with} \quad |C_i^n| = \Delta x_i^n$$

which constitute a *cell-centered* partition of the domain.

*Remark* 2.2.4. The title "Cell Centered grids" is chosen since the cells that are constructed are centered with respect to the nodes of the mesh.

*Remark* 2.2.5. In the previous Definition the mesh $M_x^n$ follows from the non-uniform discretization $\{0 = x_{-1/2}^n < \cdots < x_{N-1/2}^n = 1\}$ since an attempt to start with the mesh $M_x^n = \{x_i^n, i\}$ and define cells $C_i^n = (x_{i-1/2}^n, x_{i+1/2}^n)$ with midpoints $x_i^n$ can easily result in gaps in the domain. Such an example is depicted if Figure 2.5.

We do not encounter this problem in the uniform mesh case due to the following equivalence relation

$$\left\{ x_{i+1/2} = \frac{x_i^n + x_{i+1}^n}{2} = x_i^n + \frac{h}{2}, i \right\} \Longleftrightarrow \left\{ x_i^n = \frac{x_{i-1/2}^n + x_{i+1/2}^n}{2} = x_{i-1/2}^n + \frac{h}{2}, i \right\}$$

where $h$ is the uniform mesh step.

*Remark* 2.2.6. Another point of interest with the use of cell centered computational grids is that the endpoints of the full domain are not mesh points. Although we do not

treat boundary conditions in this work we note that the cell centered grids are better for treating certain boundary conditions.

The cells of the cell centered grid we described are non overlapping, so there is a one-to-one relation

$$C_i^n \longleftrightarrow x_i^n \longleftrightarrow u_i^n,$$

which provides enough justification to define piecewise constant function $V^n(x)$,

**Definition** (Piecewise constant approximate $V^n(x)$)**.** Given a partition of the domain with vertex centered cells $C_i^n$ -as defined in Def.(2.2.4)- and a finite sequence of discrete approximations $U^n = \{u_i^n, i = 1, \dots N-1\}$, we define the Approximate Solution $V^n(x)$ to be the piecewise constant function,

$$V^n(x) = u_i^n \quad \text{for} \quad x \in C_i^n.$$

This completes the step 2a. of the BAS. The following remark is in order,

*Remark* 2.2.7. If we had performed piecewise linear construction of the approximate solution $V_{lin}^n(x)$ -linear in the segments $(x_i^n, x_{i+1}^n)$ with $V_{lin}^n(x_i^n) = u_i^n$ and $V_{lin}^n(x_{i+1}^n) = u_{i+1}^n$ - instead of the piecewise constant $V^n(x)$ -in the cells $C_i^n = (x_{i-1/2}^n, x_{i+1/2}^n)$- the linear approximate solution $V_{lin}^n(x)$ would *not* have had (in general) the same mass as the piecewise constant solution approximation $V^n(x)$ in the segment $(x_i^n, x_{i+1}^n)$, i.e.,

$$\int_{x_i^n}^{x_{i+1}^n} V_{lin}^n(x)dx \neq \int_{x_i^n}^{x_{i+1}^n} V^n(x)dx.$$

This fact is depicted in Figure 2.3 and can easily be proven since the end point $x_{i+1/2}$ of the cell $C_i^n$ is *not* the midpoint of the interval $(x_i, x_{i+1})$.

For the Step 2b. of the BAS -that is the update of the discrete approximations- we incorporate in our discussion the new partition $\{0 = x_{-1/2}^{n+1} < \cdots < x_{N-1/2}^{n+1} = 1\}$ which gives rise to new cells $C_i^{n+1}$ and new mesh $M_x^{n+1} = \{x_i^{n+1}, i = 0, \dots, N-1\}$. We want to update the discrete approximations $U^n$ -associated with old mesh $M_x^n$- to new discrete approximations $\hat{U}^n$ -associated with the new mesh $M_x^{n+1}$.

We need the reconstruction procedure to be conservative, so by repeating the arguments of the previous paragraph we define $\hat{u}_i^n$ as

**Definition 2.2.5** (Conservative construction of $\hat{U}^n$)**.** We take into account the configuration of the new nodes $x_i^{n+1}$ with respect to the old cells $C_i^n$. So we have the following

cases,

- If the configuration of the new nodes $x_{i-1}^{n+1}$, $x_i^{n+1}$, $x_{i+1}^{n+1}$ is such that

$$x_{i-1/2}^{n+1} \in C_k^n \quad \text{and} \quad x_{i+1/2}^{n+1} \in C_l^n, \quad \text{with } k < l,$$

  we define

$$\hat{u}_i^n = \frac{1}{\Delta x_i^{n+1}} \left( (x_{k+1/2}^n - x_{i-1/2}^{n+1}) u_k^n + \sum_{j=k+1}^{l-1} \Delta x_j^n u_j^n + (x_{i+1/2}^{n+1} - x_{l-1/2}^n) u_l^n \right).$$

- If the configuration of the new nodes $x_{i-1}^{n+1}$, $x_i^{n+1}$, $x_{i+1}^{n+1}$ is such that

$$x_{i-1/2}^{n+1}, x_{i+1/2}^{n+1} \in C_k^n,$$

  we define

$$\hat{u}_i^n = u_k^n.$$

Figure 2.5 gives a graphical explanation of the reconstruction just described.

**Proposition 2.2.4.** *The construction just defined is conservative.*

*Proof.* The proof is the same as in the Vertex Centered - Piecewise Constant case. □

Again, the reconstruction procedure we described yields the maximum principle,

**Proposition 2.2.5** (Maximum principle)**.** *The absolute values of the new point approximations $\hat{u}_i^n$ are bounded by the maximum of the old point approximations $u_i^n$*

$$\max_i |\hat{u}_i^n| \leq \max_i |u_i^n|.$$

*Proof.* The proof is the same as in the Vertex Centered - Piecewise Constant case, hence it is omitted. □

We close this paragraph, by counting the operations needed for the proposed reconstruction procedure,

**Proposition 2.2.6.** *The computational cost for the sequence $\{\hat{u}_i^n, i = 1, \ldots, N\}$ is $\mathcal{O}(N)$.*

*Proof.* The result is obvious, assuming that we compute the sum $\sum_{j=1}^N \Delta x_j^n u_j^n$ in advance. □

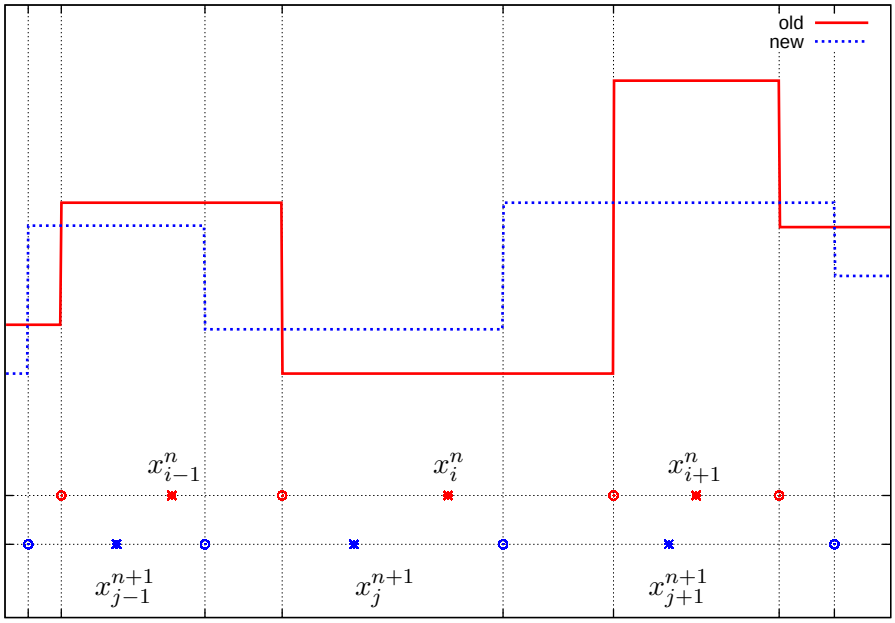



Figure 2.5: Cell Centered grids & Conservative Piecewise Constant reconstruction. This graph depicts the old and the new mesh points $\{x^n_i, i\}$ and $\{x^{n+1}_j, j\}$ being the middle points of $x^n_{i\pm1/2}$ and $x^n_{j\pm1/2}$ (denoted with circles) that define the cells $C^n_i$ and $C^{n+1}_j$. It also depicts the old function $V^n(x)$ and its reconstruction $\hat{V}^n(x)$, both being piecewise constant over the respective cells. This reconstruction is conservative and respects the maximum principle.

### 2.2.3 Finite Element grids - Piecewise Linear - Conservative

We consider the following partition that defines cells over the domain as follows,

**Definition 2.2.6** (Finite Element grid)**.** For a given mesh $M^n_x = \{x_i, i = 0, \ldots, N\}$ we define the cells

$$C^n_i = (x^n_i, x^n_{i+1}) \quad \text{with} \quad \Delta x^n_i = x^n_{i+1} - x^n_i,$$

which constitute a *Finite Element* partition of the domain.

For this partition we define the following piecewise linear function over the whole domain,

**Definition 2.2.7** (Piecewise linear approximate solution)**.** Given a partition of the domain with Finite Element cells $C^n_i$ -as defined in Def.(2.2.6)- and a finite sequence of

discrete approximations $U^n = \{u_i^n, i = 1, \dots N-1\}$, we define the Approximate Solution $V^n(x)$ to be the piecewise linear function

$$V^n(x) = u_i^n + \frac{u_{i+1}^n - u_i^n}{\Delta x_i^n}(x - x_i^n) \quad \text{for} \quad x \in C_i^n.$$

We now incorporate in our discussion the new mesh $M_x^{n+1} = \{x_i^{n+1}, i = 0, \dots, N\}$ over every node of which the new approximations $\hat{U}^n = \{\hat{u}_i^n, i = 0, \dots, N\}$ have to be defined. This new mesh $M_x^{n+1}$ gives rise to new computational cells $C_i^{n+1}$ and the new discrete approximations $\hat{U}^n$ give rise to new piecewise linear function $\hat{V}^n(x)$.

We need the reconstruction procedure to be conservative, so by repeating the arguments of the previous paragraph we define $\hat{u}_i^n$ as

**Definition 2.2.8** (Conservative construction of $\hat{U}^n$)**.** We take into account the configuration of the new nodes $x_i^{n+1}$ with respect to the old cells $C_i^n$. So we have the following cases,

- If the configuration of the new nodes $x_{i-1}^{n+1}$ and $x_i^{n+1}$ is such that

  $$x_{i-1}^{n+1} \in C_k^n \quad \text{and } x_i^{n+1} \in C_l^n, \quad \text{with } k < l,$$

  we define

  $$\hat{u}_i^n = -\hat{u}_{i-1}^n + \frac{2}{\Delta x_i^{n+1}}\Big((x_{k+1}^n - x_{i-1}^{n+1})\frac{\bar{u}_{i-1}^{n+1} + u_{k+1}^n}{2} + \sum_{j=k+1}^{l-1} \Delta x_j^n \frac{u_{j+1}^n + u_j^n}{2} + (x_i^{n+1} - x_l^n)\frac{u_l^n + \bar{u}_i^{n+1}}{2}\Big),$$

  where $\bar{u}_{i-1}^{n+1}$ and $\bar{u}_i^{n+1}$ are respectively the interpolated values in the new nodes $x_{i-1}^{n+1}$ and $x_i^{n+1}$ of the old piecewise linear function $V^n(x)$.

- If moreover the configuration of the new nodes $x_{i-1}^{n+1}$ and $x_i^{n+1}$ is such that

  $$x_{i-1}^{n+1}, x_i^{n+1} \in C_k^n,$$

  we define

  $$\hat{u}_i^n = -\hat{u}_{i-1}^n + \bar{u}_{i-1}^{n+1} + \bar{u}_i^{n+1},$$

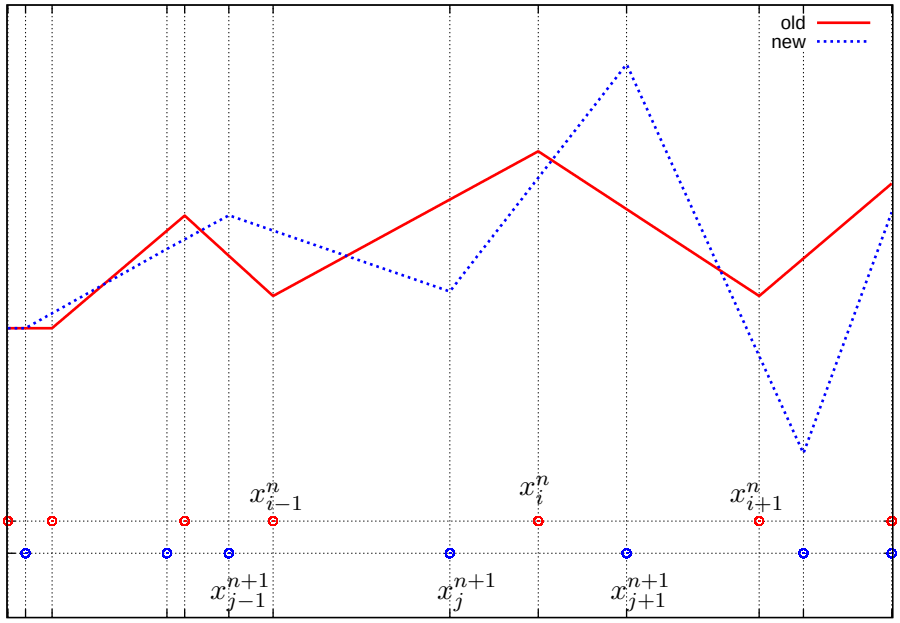



Figure 2.6: Finite Element grid & Conservative Piecewise Linear reconstruction. This graphs depicts the old and the new mesh along with the old and the new piecewise linear functions $V^n(x)$. This reconstruction is conservative but it *does not* respect the maximum principle.

where again $\bar{u}_{i-1}^{n+1}$ and $\bar{u}_i^{n+1}$ are respectively the interpolated values of the new nodes $x_{i-1}^{n+1}$ and $x_i^{n+1}$ over the old piecewise linear solution approximation $V^n(x)$.

We refer to Figure 2.6 for a graphical explanation of this definition.

**Proposition 2.2.7.** *The construction just defined is conservative.*

*Proof.* The area defined by the new approximation $\hat{u}^n$ over the cell $C_i^{n+1}$ is

$$\frac{\hat{u}_i^n + \hat{u}_{i-1}^n}{2}\Delta x_i^{n+1} = (x_{k+1}^n - x_{i-1}^{n+1})\frac{\bar{u}_{i-1}^{n+1} + u_{k+1}^n}{2} + \sum_{j=k+1}^{l-1} \Delta x_j^n \frac{u_{j+1}^n + u_j^n}{2} + \underbrace{(x_i^{n+1} - x_l^n)\frac{u_l^n + \bar{u}_i^{n+1}}{2}}_{A}$$

We write the following term

$$\frac{\hat{u}_{i+1}^n + \hat{u}_i^n}{2}\Delta x_{i+1}^{n+1} = \underbrace{(x_{l+1}^n - x_i^{n+1})\frac{\bar{u}_i^{n+1} + u_{l+1}^n}{2}}_{B} + \sum_{j=l+1}^{m-1} \Delta x_j^n \frac{u_{j+1}^n + u_j^n}{2} + (x_{i+1}^{n+1} - x_m^n)\frac{u_m^n + \bar{u}_{i+1}^{n+1}}{2}$$

and sum with the previous one, we notice that the terms $A$ and $B$ merge with the sums and we end up in the following right-hand side

$$(x_{k+1}^n - x_{i-1}^{n+1})\frac{\bar{u}_{i-1}^{n+1} + u_{k+1}^n}{2} + \sum_{j=k+1}^{m-1} \Delta x_j^n \frac{u_{j+1}^n + u_j^n}{2} + (x_{i+1}^{n+1} - x_m^n)\frac{u_m^n + \bar{u}_{i+1}^{n+1}}{2}$$

which is sufficient for conservation □

*Remark* 2.2.8. We note here that the conservative construction of $\hat{u}^n$, we just described, *does not* respect the maximum principle as can be seen in Figure 2.6.

As with the previous reconstructions, we count the number of operations this reconstruction needs,

**Proposition 2.2.8.** *The computational cost for the sequence* $\{\hat{u}_i^n, i = 1, \ldots, N\}$ *is* $\mathcal{O}(N)$.

*Proof.* The result is obvious, assuming that we compute the sum $\sum_{j=0}^{N-1} \Delta x_j^n \frac{u_{j+1}^n + u_j^n}{2}$ in advance. □

### 2.2.4 Finite Element grids - Piecewise Linear - Interpolation

We again consider the following partition that defines cells over the domain as follows,

**Definition** (Finite Element grid)**.** For a given mesh $M_x^n = \{x_i, i = 0, \ldots, N\}$ we define the cells

$$C_i^n = (x_i^n, x_{i+1}^n) \quad \text{with} \quad \Delta x_i^n = x_{i+1}^n - x_i^n,$$

which constitute a *Finite Element* partition of the domain.

For this partition we define the following piecewise linear solution approximation over the whole domain,

**Definition** (Piecewise linear approximate solution)**.** Given a partition of the domain with Finite Element cells $C_i^n$ -as just defined- and a finite sequence of discrete approximations $U^n = \{u_i^n, i = 1, \dots N-1\}$, we define the Approximate Solution $V^n(x)$ to be the piecewise linear function

$$V^n(x) = u_i^n + \frac{u_{i+1}^n - u_i^n}{\Delta x_i^n}(x - x_i^n) \quad \text{for} \quad x \in C_i^n.$$

We now incorporate in our discussion the new mesh $M_x^{n+1} = \{x_i^{n+1}, i = 0, \dots, N\}$ over every node of which the new approximations $\hat{U}^n = \{\hat{u}_i^n, i = 0, \dots, N\}$ have to be defined. This new mesh $M_x^{n+1}$ gives rise to new computational cells $C_i^{n+1}$ and the new discrete approximations $\hat{U}^n$ give rise to new piecewise linear solution approximation $\hat{V}^n(x)$.

**Definition 2.2.9** (Interpolation construction of $\hat{U}^n$)**.** Let $x_i^{n+1} \in [x_j^n, x_{j+1}^n)$. We define

$$\hat{u}_i^n = u_j^n + \frac{u_{j+1}^n - u_j^n}{x_{j+1}^n - x_j^n}(x_i^{n+1} - x_j^n)$$

to be the interpolated value of the new node $x_i^{n+1}$ over the old piecewise linear approximate solution $V^n(x)$ -as given in the previous definition.

*Remark* 2.2.9. We note here that the interpolation construction of $\hat{U}^n$ we just described *is not* conservative as can be seen in Figure 2.7.

Again, the reconstruction procedure we described yields the maximum principle,

**Proposition 2.2.9** (Maximum principle)**.** *The absolute values of the new point approximations $\hat{u}_i^n$ are bounded by the maximum of the old point approximations $u_i^n$*

$$\max_i |\hat{u}_i^n| \leq \max_i |u_i^n|$$

*Proof.* The updated values $\hat{u}_i^n$ are given by

$$\hat{u}_i^n = u_j^n + \frac{u_{j+1}^n - u_j^n}{x_{j+1}^n - x_j^n}(x_i^{n+1} - x_j^n),$$

or as a convex combination of $u_j^n$ and $u_{j+1}^n$

$$\hat{u}_i^n = \frac{x_{j+1}^n - x_i^{n+1}}{x_{j+1}^n - x_j^n} u_j^n + \frac{x_i^{n+1} - x_j^n}{x_{j+1}^n - x_j^n} u_{j+1}^n.$$

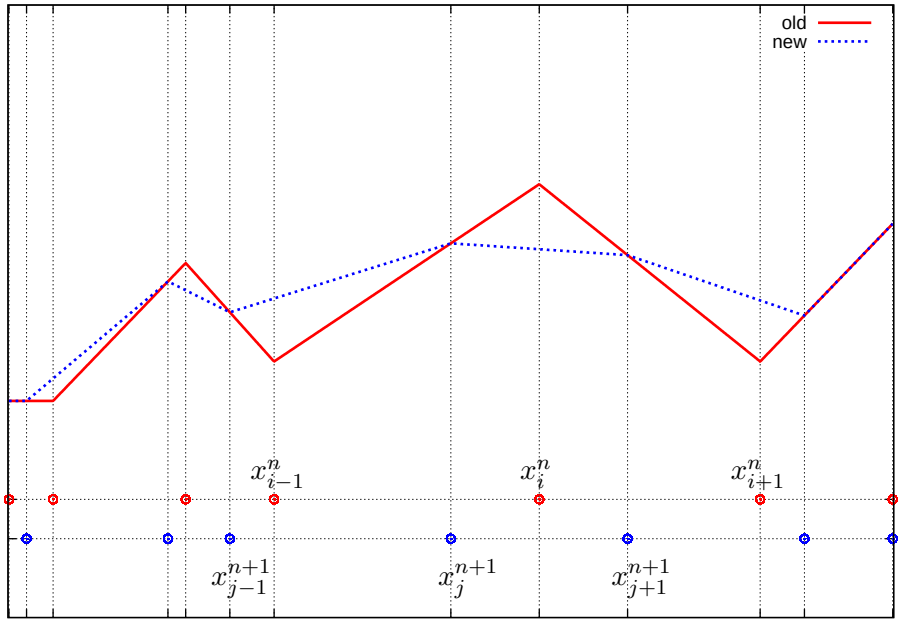



Figure 2.7: Finite Element grid & Piecewise Linear reconstruction with Interpolation. This graphs depicts the old and the new mesh along with the old and the new piecewise linear functions $V^n(x)$. This reconstruction *is not* conservative but respects the maximum principle.

The result follows since $x_i^{n+1} \in [x_j^n, x_{j+1}^n)$. □

**Proposition 2.2.10.** *The computational cost for the sequence* $\{\hat{u}_i^n, i = 1, \ldots, N\}$ *is* $\mathcal{O}(N)$.

*Proof.* Obvious since each $\hat{u}_i^n$ needs 2 multiplications and moreover the overall number of comparisons is of order $\mathcal{O}(N)$. □

# Chapter 3

# Numerical Schemes on Non-Uniform Meshes. The Evolution step

As with the previous Chapters, we continue by studying another aspect of the BAS which we restate,

**Definition** (BAS)**.** Given mesh $M_x^n = \{a = x_1^n < \cdots < x_N^n = b\}$ and approximations $U^n = \{u_1^n, \ldots, u_N^n\}$,

1. (Mesh Reconstruction)
   Given $M_x^n$ and $U^n$ construct new mesh $M_x^{n+1} = \{a = x_1^{n+1} < \cdots < x_N^{n+1} = b\}$

2. (Solution Update)
   Given $M_x^n$, $U^n$ and $M_x^{n+1}$

   2a. construct the function $V^n(x)$ with $V^n(x_i^n) = u_i^n$

   2b. compute/update approximations $\hat{U}^n = \{\hat{u}_i^n, \ldots, \hat{u}_N^n\}$

3. (Time Evolution)
   Given $M_x^{n+1}$, $\hat{U}^n$ march in time to compute $U^{n+1} = \{u_1^{n+1}, \ldots, u_N^{n+1}\}$

So far we have discussed the initiation of the BAS, that is the properties of the underlying meshes (uniform and non-uniform) and discrete approximations of derivatives. Moreover, we have studied the Adaptive Mesh Reconstruction procedure that sponsors Step 1 of the BAS and the Solution Update procedure, Step 2 of the BAS.

In this Chapter we focus on the Step 3 of the BAS, we discuss the Time Evolution of the discrete numerical approximations. More specifically this Chapter is devoted to the construction of Finite Difference numerical schemes, valid for non-uniform meshes, and analyses their basic properties, i.e., consistency, stability and conservation.

As stated in the BAS, in each time step $t = t^n$ the numerical schemes utilise the new mesh $M_x^{n+1}$ and the updated discrete approximations $\hat{U}^n$ to compute the discrete numerical approximations $U^{n+1}$ for the next time step $t = t^{n+1}$. The new discrete approximations $U^{n+1}$ along with the new mesh $M_x^{n+1}$ will serve as initiation conditions for the repetition of the algorithm/BAS.

We shall restrict our study in three-point numerical schemes and for each one we shall state both the uniform and the non-uniform version. We will start with schemes that are of first order accuracy -in their uniform mesh version- then we continue with schemes that are of second order accuracy -in their uniform mesh version. Among others we study the unstable Centered scheme, a pure second order (on non-uniform mesh) and the class of Entropy Conservative schemes.

The main tool of our analysis is the *Modified Equation*; this was first introduced by Warming and Hyett [31] and refers to the actual PDEs, which is solved numerically by the application of a given difference method to an initial value problem. We shall compute and discuss the Modified Equations for each one of the non-uniform mesh schemes we shall propose. The conclusions of this analysis include consistency, stability and order of accuracy of the schemes and we refer to [31], [13] and [14] for their connections with the Modified Equation.

Through the analysis of the proposed non-uniform numerical schemes we noticed that some of them were consistent with the underlying Conservation Law according to the usual Flux Consistency Criterion -i.e., $F(u,u) = f(u)$ where $f$ is the continuous flux and $F$ the numerical one- but they failed to be consistent according to the Modified Equation approach. This led us to reconsider the Flux consistency Criterion for the case of non-uniform meshes. We found out that in the non-uniform mesh case the usual Flux consistency Criterion is not sufficient. We propose a sufficient extension of this criterion later in this Chapter.

As we have already stated, the schemes we shall discuss aim to resolve the scalar Conservation Law for a single space variable,

$$u_t(x,t) + f(u(x,t))_x = 0, \quad x \in \mathbb{R}, \quad t \in [0,T].$$

## 3.1 First order schemes

We start the description of the schemes with the ones that are of first order accuracy whenever the mesh is uniform. In this category the easiest scheme one can consider is the Upwind scheme.

### 3.1.1 Upwind - Finite Element grid

In this approach we consider the non-uniform mesh

$$M_x = \{x_i, i \in \mathbb{Z}\}$$

that defines a partition of the domain in cells

$$C_i = (x_{i-1}, x_i] \quad \text{with} \quad h_i = x_i - x_{i-1}.$$

For this description of the grid and for the case $f' > 0$ the upwind scheme reads,

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{h_i}(f(u_i^n) - f(u_{i-1}^n)). \tag{3.1}$$

*Remark* 3.1.1. It is obvious that for the uniform mesh case this scheme recasts to the usual Upwind scheme

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x}(f(u_i^n) - f(u_{i-1}^n)),$$

where $\Delta x$ is the uniform space step.

**Proposition 3.1.1** (Modified equation for the scheme (3.1))**.** *The modified equation - truncated to third order derivatives- of the Upwind scheme (3.1) for the Linear Transport equation $f(u) = au$ is*

$$u_t + au_x = a\frac{h_i - a\Delta t}{2}u_{xx} - \frac{2a^3\Delta t^2 - 3a^2\Delta t h_i + ah_i^2}{6}u_{xxx}. \tag{3.2}$$

*Proof.* For the construction of the Modified Equation we assume that there is a smooth function $u$ that satisfied the difference equation scheme (3.1) at every point $(x_i, t^n)$. That is $u(x_i, t^n) = u_i^n$, for every $i$ and $n$, and the scheme (3.1) reads

$$u(x_i, t^{n+1}) = u(x_i, t^n) - \frac{\Delta t}{h_i}\left(u(x_i, t^n) - u(x_{i-1}, t^n)\right). \tag{3.3}$$

Now we expand $u(x_i, t^{n+1})$ and $u(x_{i-1}, t^n)$ in Taylor series about the point $(x_i, t^n)$

$$\begin{aligned} u(x_i, t^{n+1}) =& u(x_i, t^n) + \Delta t u_t(x_i, t^n) + \frac{\Delta t^2}{2}u_{tt}(x_i, t^n) + \frac{\Delta t^3}{6}u_{ttt}(x_i, t^n) + \mathcal{O}(\Delta t^4), \\ u(x_{i-1}, t^n) =& u(x_i, t^n) - h_i u_x(x_i, t^n) + \frac{h_i^2}{2}u_{xx}(x_i, t^n) - \frac{h_i^3}{6}u_{xxx}(x_i, t^n) + \mathcal{O}(h_i^4). \end{aligned}$$

We substitute the respective values at the smooth version of the Upwind scheme (3.3), we divide by $\Delta t$ and end up with the *half-way modified equation*

$$u_t + au_x = -\frac{\Delta t}{2}u_{tt} + a\frac{h_i}{2}u_{xx} - \frac{\Delta t^2}{6}u_{ttt} - a\frac{h_i^2}{6}u_{xxx} + hot, \tag{3.4}$$

where *hot* stands for higher than third order terms.

We call this equation *half-way modified equation* since it includes $t$ derivatives $u_{tt}$ and $u_{ttt}$ in the right-hand side. We need to replace them by $x$ derivatives to get the *full modified equation* (up to order 3). To do so we assume that $u$ is a solution of the

half-way modified Eq.(3.4) and not of the underlying Transport equation (this would yield to truncation error analysis).

To this end we perform the following derivations first for second order derivatives,

$$\text{Eq.(3.4)} \xrightarrow{\partial t} u_{tt} = -au_{xt} - \frac{\Delta t}{2}u_{ttt} + a\frac{h_i}{2}u_{xxt} + hot$$
$$\text{Eq.(3.4)} \xrightarrow{\partial x} u_{xt} = -au_{xx} - \frac{\Delta t}{2}u_{xtt} + a\frac{h_i}{2}u_{xxx} + hot$$

and now for third order derivatives

$$\text{Eq.(3.4)} \xrightarrow{\partial tt} u_{ttt} = -au_{xtt} + hot$$
$$\text{Eq.(3.4)} \xrightarrow{\partial xt} u_{xtt} = -au_{xxt} + hot$$
$$\text{Eq.(3.4)} \xrightarrow{\partial xx} u_{xxt} = -au_{xxx} + hot.$$

We now substitute the $t$ derivative $u_{tt}$ and $u_{ttt}$ in the half-way modified Eq.(3.4) and after some algebra we get the full modified equation (up to the third derivative),

$$u_t + au_x = a\frac{h_i - a\Delta t}{2}u_{xx} - \frac{2a^3\Delta t^2 - 3a^2\Delta t h_i + ah_i^2}{6}u_{xxx}.$$

□

Following [31] and [14] on the analysis of the Modified Equation we have,

**Corollary 3.1.1.1** (Consistency)**.** *This scheme is consistent with the underlying Transport equation in the sense that as the mesh is refined* $\max_i h_i, \Delta t \to 0$ *the right-hand side of the Modified Eq.(3.2) converges to 0.*

**Corollary 3.1.1.2** (Order of accuracy)**.** *This scheme constitutes a first order approximation of the underlying Transport equation.*

**Corollary 3.1.1.3** (Stability)**.** *The usual CFL condition* $\Delta t < \min_i h_i$ *results in a positive sign in the coefficient of the diffusion term* $u_{xx}$ *in the right-hand side of the Modified Eq.(3.2). The Modified Equation provides an implicit way to check whether the scheme is conditionally stable.*

**Corollary 3.1.1.4** (Conservation)**.** *This scheme is already in conservative form, i.e., scheme (3.1) with numerical flux function* $F_{i+1/2} = f(u_i)$.

### 3.1.2 Lax-Friedrichs

The Lax-Friedrichs (LxF) numerical scheme for a uniform mesh is

$$u_i^{n+1} = \frac{u_{i-1}^n + u_{i+1}^n}{2} - \Delta t \frac{f(u_{i+1}^n) - f(u_{i-1}^n)}{2\Delta x}$$

or in conservative form

$$\begin{aligned} u_i^{n+1} &= u_i^n - \frac{\Delta t}{\Delta x}\left(F_{i+1/2} - F_{i-1/2}\right) \\ F_{i+1/2} &= \frac{\Delta x}{2\Delta t}(u_i^n - u_{i+1}^n) + \frac{1}{2}\left(f(u_i^n) + f(u_{i+1}^n)\right). \end{aligned}$$

Its extension for non-uniform meshes is not straightforward. The main reason is that different (and equivalent in the uniform mesh case) approaches to the construction of the scheme yield different schemes in the non-uniform case. In this section we shall consider and analyse several options.

#### 3.1.2.1 LxF - Finite Element grid - Approach 1

In this approach we consider the non-uniform mesh

$$M_x = \{x_i, i \in \mathbb{Z}\}$$

that defines a partition of the domain in cells

$$C_i = (x_{i-1}, x_i] \quad \text{with} \quad h_i = x_i - x_{i-1}.$$

We start our study with the LxF scheme for the Linear Transport equation in one space dimension,

$$u_t + a u_x = 0, \quad x \in [0,1], t \in [0,T].$$

The LxF scheme for this problem can be geometrically explained by intersecting the characteristic line, passing through $(x_i, t^{n+1})$, with the line $t = t^n$. We shall use this approach to construct a generalisation of the LxF for non-uniform meshes. We also refer to Figure 3.1 for a graphical explanation of the construction that follows.

*Construction* (Geometric LxF). Consider the points $A = (x_{i-1}, u_{i-1}^n)$ and $B = (x_{i+1}, u_{i+1}^n)$.

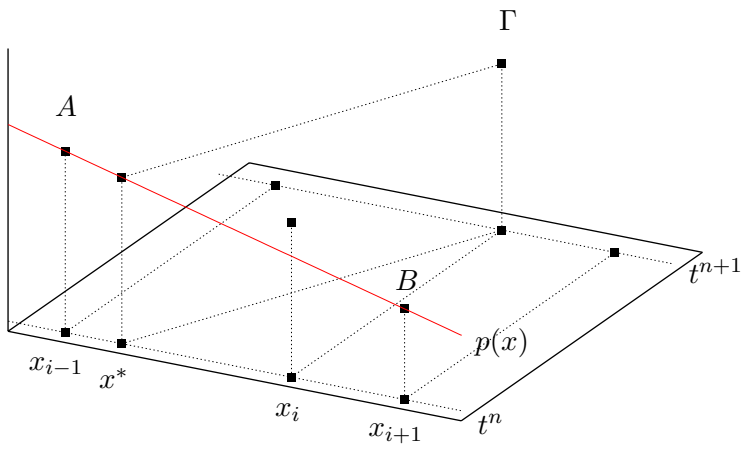



Figure 3.1: Geometric construction of the LxF scheme. The value $u_i^{n+1}$ at the point $\Gamma$ is defined as $u_i^{n+1} = p(x^*)$, where $x^*$ is the intersection point of the characteristic line -passing through $(x_i, t^{n+1})$- with the interval $[x_{i-1}, x_{i+1}]$ at time $t^n$. Obviously no overshoot is possible since $p(x^*) \leq \max\{p(x_{i-1}), p(x_{i+1})\}$.

The unique straight line that interpolates these points is given by

$$p(x) = \frac{u_{i+1}^n - u_{i-1}^n}{x_{i+1} - x_{i-1}}(x - x_{i-1}) + u_{i-1}^n.$$

The slope of the characteristic line passing through the point $\Gamma = (x_i, u_i^{n+1})$ is $\frac{1}{a} \in \overline{\mathbb{R}}\backslash\{0\}$. This characteristic line intersects -for $\Delta t$ sufficiently small- the interval $[x_{i-1}, x_{i+1}]$ at time $t^n$ in a point $x^*$ and hence

$$\frac{1}{a} = \frac{\Delta t}{x_i - x^*}$$

or

$$x^* = x_i - a\Delta t.$$

It is known that the solution is constant along the characteristic lines, so we define

$$u_i^{n+1} = p(x^*) = \frac{u_{i+1}^n - u_{i-1}^n}{x_{i+1} - x_{i-1}}(x^* - x_{i-1}) + u_{i-1}^n.$$

We now substitute $x^* = x_i - a\Delta t$ and the numerical scheme takes the form

$$u_i^{n+1} = \frac{u_{i+1}^n - u_{i-1}^n}{x_{i+1} - x_{i-1}}(x_i - x_{i-1} - a\Delta t) + u_{i-1}^n.$$

If we moreover set $h_i = x_i - x_{i-1}$ and $h_{i+1} = x_{i+1} - x_i$ the previous scheme takes the form

$$u_i^{n+1} = \frac{h_{i+1}}{h_i + h_{i+1}}u_{i-1}^n + \frac{h_i}{h_i + h_{i+1}}u_{i+1}^n - \frac{a\Delta t}{h_i + h_{i+1}}(u_{i+1}^n - u_{i-1}^n). \tag{3.5}$$

*Remark* 3.1.2. The scheme (3.5) recasts in the case of a uniform mesh $h_i = \Delta x$, for every $i$, into

$$u_i^{n+1} = \frac{u_{i-1}^n + u_{i+1}^n}{2} - \frac{a\Delta t}{2\Delta x}(u_{i+1}^n - u_{i-1}^n), \tag{3.6}$$

which is the usual LxF scheme for the Transport equation with $f(u) = au$.

*Remark* 3.1.3. Comparing the last terms of the uniform scheme (3.6) and the non-uniform mesh scheme (3.5) version,

$$\frac{u_{i+1} - u_{i-1}}{h_i + h_{i+1}} \quad \text{and} \quad \frac{u_{i+1} - u_{i-1}}{2h},$$

we note that they both constitute approximations of $u_x(x_i)$. In the uniform case this approximation is second order accurate while in the non-uniform mesh case is only first order accurate. The lack of symmetry of the non-uniform mesh is responsible for this loss of accuracy. We shall improve this approximation in the next paragraph, where we shall propose a scheme in which the respective derivative approximation will be of second order accuracy.

We now discuss the Modified Equation for the scheme (3.5).

**Proposition 3.1.2** (Modified equation of scheme (3.5))**.** *The modified equation for the scheme (3.5) is,*

$$\begin{aligned} u_t + au_x =& \frac{(h_i - a\Delta t)(h_{i+1} - a\Delta t)}{2\Delta t}u_{xx} \\ &+ \frac{(h_i - a\Delta t)(h_{i+1} - a\Delta t)(2a\Delta t - h_i + h_{i+1})}{6\Delta t}u_{xxx}. \end{aligned} \tag{3.7}$$

*Proof.* Let $u$ be a smooth function that satisfies the difference scheme (3.5) at every

point $(x_i, t^n)$, in which case it reads

$$u(x_i, t^{n+1}) = \frac{h_{i+1}}{h_i + h_{i+1}} u(x_{i-1}, t^n) + \frac{h_i}{h_i + h_{i+1}} u(x_{i+1}, t^n) - \frac{a\Delta t}{h_i + h_{i+1}} \left( u(x_{i+1}, t^n) - u(x_{i-1}, t^n) \right).$$

We now expand $u(x_i, t^{n+1})$, $u(x_{i-1}, t^n)$, $u(x_{i+i}, t^n)$ in Taylor series about the point $(x_i, t^n)$,

$$\begin{aligned}
u(x_i, t^{n+1}) =& u(x_i, t^n) + \Delta t u_t(x_i, t^n) + \frac{\Delta t^2}{2} u_{tt}(x_i, t^n) + \frac{\Delta t^3}{6} u_{ttt}(x_i, t^n) + \mathcal{O}(\Delta t^4) \\
u(x_{i-1}, t^n) =& u(x_i, t^n) - h_i u_x(x_i, t^n) + \frac{h_i^2}{2} u_{xx}(x_i, t^n) - \frac{h_i^3}{6} u_{xxx}(x_i, t^n) + \mathcal{O}(h_i^4) \\
u(x_{i+1}, t^n) =& u(x_i, t^n) + h_{i+1} u_x(x_i, t^n) + \frac{h_{i+1}^2}{2} u_{xx}(x_i, t^n) + \frac{h_{i+1}^3}{6} u_{xxx}(x_i, t^n) + \mathcal{O}(h_{i+1}^4).
\end{aligned}$$

We replace them in the previous relation, we divide by $\Delta t$ and after some algebraic manipulations we end up with the half-way modified equation

$$\begin{aligned}
&u_t + a u_x + \frac{\Delta t}{2} u_{tt} - \frac{a\Delta t h_i - a\Delta t h_{i+1} + h_i h_{i+1}}{2\Delta t} u_{xx} \\
&+ \frac{\Delta t^2}{6} u_{ttt} - \frac{-a\Delta t h_{i+1}^2 - a\Delta t h_i^2 + a\Delta t h_i h_{i+1} + h_i h_{i+1}^2 - h_i^2 h_{i+1}}{6\Delta t} u_{xxx} = hot.
\end{aligned} \tag{3.8}$$

Here all terms are evaluated at the space-time point $(x_i, t^n)$ and *hot* stands for higher than 3-order terms.

We need to replace the $t$ derivatives $u_{tt}$ and $u_{ttt}$ by $x$ derivatives to get the full modified equation. To do so we assume that $u$ is a solution of the half way modified Eq.(3.8). To this end we perform the following derivations for the second order derivatives,

$$\begin{aligned}
\text{Eq.(3.8)} \xrightarrow{\partial t} u_{tt} &= -a u_{xt} - \frac{\Delta t}{2} u_{ttt} + \frac{-a\Delta t h_{i+1} + a\Delta t h_i + h_i h_{i+1}}{2\Delta t} u_{xxt} + hot \\
\text{Eq.(3.8)} \xrightarrow{\partial x} u_{xt} &= -a u_{xx} - \frac{\Delta t}{2} u_{xtt} + \frac{-a\Delta t h_{i+1} + a\Delta t h_i + h_i h_{i+1}}{2\Delta t} u_{xxx} + hot
\end{aligned}$$

and now for third order derivatives

$$\begin{aligned}
\text{Eq.(3.8)} &\xrightarrow{\partial ttt} u_{ttt} = -au_{xtt} + hot \\
\text{Eq.(3.8)} &\xrightarrow{\partial xtt} u_{xtt} = -au_{xxt} + hot \\
\text{Eq.(3.8)} &\xrightarrow{\partial xxt} u_{xxt} = -au_{xxx} + hot.
\end{aligned}$$

We now replace them in the half way modified Eq.(3.8) and after some algebra we get the full modified equation

$$\begin{aligned}
u_t + au_x =& \frac{-a^2\Delta t^2 + h_i h_{i+1} - a\Delta t h_{i+1} + a\Delta t h_i}{2\Delta t} u_{xx} \\
&+ \left( \frac{4a\Delta t h_i h_{i+1} - 2a^3\Delta t^3 + 3a^2\Delta t^2 h_i - 3a^2\Delta t^2 h_{i+1}}{6\Delta t} \right. \\
&\left. + \frac{-a\Delta t h_{i+1}^2 - a\Delta t h_i^2 + h_i h_{i+1}^2 - h_i^2 h_{i+1}}{6\Delta t} \right) u_{xxx} + hot
\end{aligned}$$

or after some algebra and by discarding *hot*

$$\begin{aligned}
u_t + au_x =& \frac{(h_i - a\Delta t)(h_{i+1} - a\Delta t)}{2\Delta t} u_{xx} \\
&+ \frac{(h_i - a\Delta t)(h_{i+1} - a\Delta t)(2a\Delta t - h_i + h_{i+1})}{6\Delta t} u_{xxx}
\end{aligned}$$

and this completes the proof of the proposition. □

We then have,

**Corollary 3.1.2.1** (Consistency)**.** *This scheme is consistent in the sense that as the mesh is refined* $\max_i h_i \to 0$ *and* $\Delta t \to 0$ *the modified equation converges to the transport equation.*

**Corollary 3.1.2.2** (Order of accuracy)**.** *The scheme constitutes a first order approximation of the transport equation, and a second order approximation of the respective diffusion problem.*

**Corollary 3.1.2.3** (Stability)**.** *The usual CFL condition* $a\Delta t < \min_i h_i$ *results in a positive sign in the coefficient of the diffusion term* $u_{xx}$ *in the right-hand side of the Modified Eq.(3.7).*

**Corollary 3.1.2.4** (Conservation)**.** *The scheme (3.5) cannot be written in conservative form.*

*Proof.* We rewrite the scheme in the following form

$$u_i^{n+1} = u_i^n - \frac{2\Delta t}{h_i + h_{i+1}} \Bigg( \frac{h_i}{2\Delta t}(u_i^n - u_{i+1}^n) + \frac{a}{2}(u_i^n + u_{i+1}^n) \\ - \frac{h_{i+1}}{2\Delta t}(u_{i-1}^n - u_i^n) - \frac{a}{2}(u_{i-1}^n + u_i^n) \Bigg).$$

This scheme is 3-point so the term inside the parenthesis should be decomposed in a forward (depending on $u_i^n$ and $u_{i+1}^n$) and a backward (depending on $u_{i-1}^n$ and $u_i^n$) part. So, we set $F_i^R = \frac{h_i}{2\Delta t}(u_i^n - u_{i+1}^n) + \frac{a}{2}(u_i^n + u_{i+1}^n)$ and $F_i^L = \frac{h_{i+1}}{2\Delta t}(u_{i-1}^n - u_i^n) + \frac{a}{2}(u_{i-1}^n + u_i^n)$. It is now obvious that $F_{i+1}^L \neq F_i^R$ because of the loss of the symmetry of the mesh. Hence the scheme cannot be written in conservative form. □

Closing this paragraph we state a non-linear version of the scheme (3.5)

$$u_i^{n+1} = \frac{h_{i+1}}{h_i + h_{i+1}} u_{i-1} + \frac{h_i}{h_i + h_{i+1}} u_{i+1} - \frac{\Delta t}{h_i + h_{i+1}}(f(u_{i+1}) - f(u_{i-1})),$$

where $f$ is the flux function of the Conservation Law $u_t + f(u)_x = 0$.

#### 3.1.2.2 LxF - Finite Element grid - Approach 2

In the previous paragraph we devised the numerical scheme (3.5)

$$u_i^{n+1} = \frac{h_{i+1}}{h_i + h_{i+1}} u_{i-1}^n + \frac{h_i}{h_i + h_{i+1}} u_{i+1}^n - \frac{a\Delta t}{h_i + h_{i+1}}(u_{i+1}^n - u_{i-1}^n)$$

for non-uniform meshes, that reduces to the LxF scheme on the uniform case. We also noted in Remark (3.1.3) that the fraction

$$\frac{u_{i+1} - u_{i-1}}{h_i + h_{i-1}}$$

constitutes a first order accurate approximation of the derivative $u_x(x_i)$ and it increases to second order accuracy in the uniform mesh case.

In this paragraph we intent to construct another scheme, where the respective approximation will be of second order accuracy.

*Remark* 3.1.4. As we saw in Chapter 2, a second order approximation of the first derivative on a 3-point non-uniform mesh is given by

$$-\frac{h_{i+1}}{h_i(h_i + h_{i+1})} u_{i-1} + \left( \frac{h_{i+1}}{h_i(h_i + h_{i+1})} - \frac{h_i}{h_{i+1}(h_i + h_{i+1})} \right) u_i + \frac{h_i}{h_{i+1}(h_i + h_{i+1})} u_{i+1}.$$

We combine the last remark with the LxF generalisation Scheme (3.5) of the previous section, that is we substitute the first order derivative approximation

$$\frac{u_{i+1}-u_{i-1}}{h_i+h_{i-1}}$$

by the respective second order

$$-\frac{h_{i+1}}{h_i(h_i+h_{i+1})}u_{i-1}+\left(\frac{h_{i+1}}{h_i(h_i+h_{i+1})}-\frac{h_i}{h_{i+1}(h_i+h_{i+1})}\right)u_i+\frac{h_i}{h_{i+1}(h_i+h_{i+1})}u_{i+1}$$

and we propose another generalization of the LxF scheme for non-uniform meshes for the linear transport equation, i.e.,

$$\begin{aligned}(3.9)\qquad u_i^{n+1}=&\frac{h_{i+1}}{h_i+h_{i+1}}u_{i-1}+\frac{h_i}{h_i+h_{i+1}}u_{i+1}\\&-\frac{a\Delta t}{h_i+h_{i+1}}\left(-\frac{h_{i+1}}{h_i}u_{i-1}+\left(\frac{h_{i+1}}{h_i}-\frac{h_i}{h_{i+1}}\right)u_i+\frac{h_i}{h_{i+1}}u_{i+1}\right).\end{aligned}$$

This scheme can be written in the case of non-linear flux $f$ as follows

$$\begin{aligned}u_i^{n+1}=&\frac{h_{i+1}}{h_i+h_{i+1}}u_{i-1}+\frac{h_i}{h_i+h_{i+1}}u_{i+1}\\&-\frac{\Delta t}{h_i+h_{i+1}}\left(-\frac{h_{i+1}}{h_i}f(u_{i-1})+\left(\frac{h_{i+1}}{h_i}-\frac{h_i}{h_{i+1}}\right)f(u_i)+\frac{h_i}{h_{i+1}}f(u_{i+1})\right)\end{aligned}$$

*Remark* 3.1.5. One can easily check that this scheme reduces to the usual LxF scheme whenever the mesh is uniform.

We will analyse the scheme we just proposed -in the linear case $f(u)=au$- and compare it with the scheme (3.5) we constructed in the previous paragraph. We shall again follow the modified equation approach,

**Proposition 3.1.3** (Modified Equation for the scheme (3.9))**.** *The modified equation of the scheme (3.9) is*

$$(3.10)\quad u_t+au_x=\frac{h_ih_{i+1}-a^2\Delta t^2}{2\Delta t}u_{xx}+\frac{2ah_ih_{i+1}\Delta t-h_i^2h_{i+1}+h_ih_{i+1}^2-2a^3\Delta t^3}{6\Delta t}u_{xxx}.$$

*Proof.* We proceed as in the previous paragraph, we expand in Taylor series up to third

order derivatives with respect to the space-time point $(x_i, t^n)$,

$$\begin{aligned}
u(x_i, t^{n+1}) =& u(x_i, t^n) + \Delta t u_t(x_i, t^n) + \frac{\Delta t^2}{2} u_{tt}(x_i, t^n) + \frac{\Delta t^3}{6} u_{ttt}(x_i, t^n) + \mathcal{O}(\Delta t^4) \\
u(x_{i-1}, t^n) =& u(x_i, t^n) - h_i u_x(x_i, t^n) + \frac{h_i^2}{2} u_{xx}(x_i, t^n) - \frac{h_i^3}{6} u_{xxx}(x_i, t^n) + \mathcal{O}(\Delta h_i^4) \\
u(x_{i+1}, t^n) =& u(x_i, t^n) + h_{i+1} u_x(x_i, t^n) + \frac{h_{i+1}^2}{2} u_{xx}(x_i, t^n) + \frac{h_{i+1}^3}{6} u_{xxx}(x_i, t^n) + \mathcal{O}(\Delta h_{i+1}^4)
\end{aligned}$$

We replace these expansions in the scheme (3.9), and divide by $\Delta t$,

$$\begin{aligned}
u_t + a u_x + \frac{\Delta t}{2} u_{tt} - \frac{h_i h_{i+1}}{2\Delta t} u_{xx} \\
+ \frac{\Delta t^2}{6} u_{ttt} + \frac{h_i h_{i+1}^2 - a h_i h_{i+1} \Delta t - h_i^2 h_{i+1}}{6\Delta t} u_{xxx} = 0,
\end{aligned} \tag{3.11}$$

this is the half-way modified equation. We now replace the $t$ derivatives $u_{tt}$ and $u_{ttt}$ by $x$ derivatives. To this end we repeat the computations of the previous paragraph, first for second order derivatives,

$$\begin{aligned}
\text{Eq}(3.11) \xrightarrow{\partial t} u_{tt} &= -a u_{xt} - \frac{\Delta t}{2} u_{ttt} + \frac{h_i h_{i+1}}{2\Delta t} u_{xxt} \\
\text{Eq}(3.11) \xrightarrow{\partial x} u_{xt} &= -a u_{xx} - \frac{\Delta t}{2} u_{xtt} + \frac{h_i h_{i+1}}{2\Delta t} u_{xxx}
\end{aligned}$$

and now for third order derivatives,

$$\begin{aligned}
\text{Eq}(3.11) \xrightarrow{\partial tt} u_{ttt} &= -a u_{xtt} \\
\text{Eq}(3.11) \xrightarrow{\partial tx} u_{ttx} &= -a u_{xxt} \\
\text{Eq}(3.11) \xrightarrow{\partial xx} u_{xxt} &= -a u_{xxx},
\end{aligned}$$

where the derivatives are up to third order, since this is the order of the modified equation that we want to achieve. Now, we replace the $t$ derivatives in Rel.(3.11) by the previous evaluations

$$u_t + a u_x = \frac{h_i h_{i+1} - a^2 \Delta t^2}{2\Delta t} u_{xx} + \frac{2a h_i h_{i+1} \Delta t - h_i^2 h_{i+1} + h_i h_{i+1}^2 - 2a^3 \Delta t^3}{6\Delta t} u_{xxx}$$

and this completes the proposition. □

Regarding the properties of this scheme we state,

**Corollary 3.1.2.5** (Consistency)**.** *This scheme is consistent in the sense that as the mesh is refined, i.e.,* $\max_i h_i, \Delta t \to 0$*, the right-hand side of the Modified Equation (3.10) converges to 0.*

**Corollary 3.1.2.6** (Order of accuracy)**.** *This scheme constitutes a first order approximation of the Linear Transport equation.*

**Corollary 3.1.2.7** (Stability)**.** *The usual CFL condition* $a\Delta t < \min_i h_i$ *results in a positive sign in the coefficient of the diffusion term* $u_{xx}$ *in the right-hand side of the Modified Equation (3.10).*

**Corollary 3.1.2.8** (Conservation)**.** *The scheme (3.9) cannot be written in conservative form.*

Before closing this paragraph, we recall that our objective was to improve the numerical scheme (3.5) of the previous paragraph by improving the spatial derivative approximation of the scheme. To this end we constructed numerical scheme (3.9) which we analysed. We just need some comparative results of the two schemes. This is done in the following remark,

*Remark* 3.1.6. We state again the modified equation of the Approach-1 of the non-uniform LxF, that is Eq.(3.7)

$$\begin{aligned} u_t + a u_x =& \frac{(h_i - a\Delta t)(h_{i+1} - a\Delta t)}{2\Delta t} u_{xx} \\ &+ \frac{(h_i - a\Delta t)(h_{i+1} - a\Delta t)(2a\Delta t - h_i + h_{i+1})}{6\Delta t} u_{xxx} \end{aligned}$$

and the modified equation of the Approach-2, i.e., Eq.(3.10),

$$u_t + a u_x = \frac{h_i h_{i+1} - a^2\Delta t^2}{2\Delta t} u_{xx} + \frac{2a h_i h_{i+1}\Delta t - h_i^2 h_{i+1} + h_i h_{i+1}^2 - 2a^3\Delta t^3}{6\Delta t} u_{xxx}.$$

Focusing on the coefficients of the diffusion terms, we see that they both are positive under the CFL condition $a\Delta t < \min_i h_i$, they both are of first order with respect to the mesh ratio and their difference is

$$\frac{(h_i - a\Delta t)(h_{i+1} - a\Delta t)}{2\Delta t} - \frac{h_i h_{i+1} - a^2\Delta t^2}{2\Delta t} = \frac{2a^2\Delta t - a\Delta t h_i - a\Delta t h_{i+1}}{2\Delta t},$$

which can be written as

$$a\frac{a\Delta t - h_i + a\Delta t - h_{i+1}}{2} < 0,$$

where the inequality is justified by the CFL condition. This means that the first approach scheme (3.5) is less diffusive than the second one (3.9) even though the spatial derivative approximation in the second approach is of higher order of accuracy.

So far, our approaches for constructing generalisations of the LxF scheme for non-uniform meshes, lack of conservation. So we continue with the construction of schemes, for non-uniform meshes, that are conservative. As we shall see achieving conservation on non-uniform meshes comes with the loss of either stability or consistency. Although these schemes lack of such important properties, we shall study them both for their heuristic interest and as an introduction to the capabilities of the Mesh Relocation procedure that we will study in the following Chapter.

#### 3.1.2.3 Generalised LxF - Cell centered grid

In this approach we consider the non-uniform discretization of the domain in cells

$$C_i = (x_{i-1/2}, x_{i+1/2}) \quad \text{with} \quad |C_i| = h_i.$$

The mesh $M_x = \{x_i, i \in \mathbb{Z}\}$ consists of the middle points,

$$x_i = \frac{x_{i+1/2} + x_{i-1/2}}{2}; \quad \text{hence} \quad x_i - x_{i-1} = \frac{h_i + h_{i-1}}{2}.$$

For this description of the grid we consider the following scheme,

$$u_i^{n+1} = \underbrace{\frac{h_{i-1}}{2h_i}u_{i-1} + \frac{h_{i+1}}{2h_i}u_{i+1}}_{A} - \frac{\Delta t}{2h_i}\big(f(u_{i+1}) - f(u_{i-1})\big),$$

where the term $A$ is a non-uniform analog of the semi-sum term $\frac{u_{i-1}+u_{i+1}}{2}$ that appears in the uniform LxF scheme in the sense that $A$ reduces to $\frac{u_{i-1}+u_{i+1}}{2}$ in the case of a uniform mesh. Moreover, the term $A$ satisfies

$$Ah_i = \frac{1}{2}(h_{i-1}u_{i-1} + h_{i+1}u_{i+1}),$$

which in the case of piecewise constant numerical approximations is the mean value of the area of the approximate solution over the cells $C_{i-1}$ and $C_{i+1}$.

The previous numerical scheme can be rewritten in the following conservative form

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{h_i}\Big( \frac{h_i}{2\Delta t}u_i^n - \frac{h_{i+1}}{2\Delta t}u_{i+1}^n + \frac{1}{2}f(u_{i+1}^n) + \frac{1}{2}f(u_i^n) \\ - \frac{h_{i-1}}{2\Delta t}u_{i-1}^n + \frac{h_i}{2\Delta t}u_i^n - \frac{1}{2}f(u_i^n) - \frac{1}{2}f(u_{i-1}^n)\Big)$$

or

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{h_i}\big(F_{i+1/2} - F_{i-1/2}\big)$$

with

$$F_{i+1/2} = \frac{1}{2\Delta t}(h_i u_i^n - h_{i+1}u_{i+1}^n) + \frac{1}{2}\big(f(u_{i+1}^n) + f(u_i^n)\big).$$

*Remark* 3.1.7. In the uniform mesh case, $h_i = \Delta x$ for every $i$, this numerical flux recasts into

$$F_{i+1/2} = \frac{\Delta x}{2\Delta t}(u_i^n - u_{i+1}^n) + \frac{1}{2}\big(f(u_{i+1}^n) + f(u_i^n)\big),$$

which is the uniform LxF flux.

In the non-uniform case, though, it yields the following Modified Equation for the Linear Transport case, $f(u) = u$

$$u_t + u_x = \frac{h_{i-1} - 2h_i + h_{i+1}}{2\Delta t h_i}u \\ + \frac{h_{i+1}(h_i + h_{i+1}) - h_{i-1}(h_{i-1} + h_i) - \Delta t(h_{i-1} - 2h_i + h_{i+1})}{4\Delta t h_i}u_x + hot,$$

where *hot* stands for higher than second order terms. We notice that this scheme is not consistent, it is not even able to discard $u_i$.

As a conclusion, in this approach we witness the loss of everything (except conservation) due to our naive approach to generalise the LxF scheme.

#### 3.1.2.4 Generalised LxF - Vertex centered grid

In this approach we consider the non-uniform mesh

$$M_x = \{x_i, i \in \mathbb{Z}\} \quad \text{with} \quad h_i = x_i - x_{i-1}.$$

The middle points $x_{i-1/2} = \frac{x_{i-1}+x_i}{2}$ define a partition of the domain in cells

$$C_i = (x_{i-1/2}, x_{i+1/2}) \quad \text{with} \quad |C_i| = \frac{h_i + h_{i+1}}{2}.$$

For this description of the grid we shall discuss the following scheme,

$$u_i^{n+1} = u_i^n - \frac{2\Delta t}{h_i + h_{i+1}} \left( F_{i+1/2} - F_{i-1/2} \right)$$

with

$$F_{i+1/2} = \frac{h_{i+1}}{2\Delta t}(u_i^n - u_{i+1}^n) + \frac{1}{2}\big(f(u_i^n) + f(u_{i+1}^n)\big).$$

*Remark* 3.1.8. This is the typical Finite Volume LxF scheme used on non-uniform grids.

The modified equation for the linear case $f(u) = u$ follows by the usual procedure and reads

$$u_t + u_x = \frac{h_{i+1} - h_i}{\Delta t} u_x + \frac{(h_{i+1} - h_i)(\Delta t - h_{i+1}) + (h_i - \Delta t)(h_i + \Delta t)}{2\Delta t} u_{xx}.$$

We make two notes regarding the consistency of this scheme,

- This scheme is not consistent in the sense of the modified equation, since the right-hand side contains the term $u_x$ -with 0-th order coefficient.
- This scheme is consistent in the sense of the flux, that is $F_{i+1/2}(u, u) = f(u)$.

So in this Approach we witness that the Consistency Criterion of the numerical flux, i.e.,

$$\textit{if } F_{i+1/2} = f(u) \textit{ whenever } u_i = u_{i+1} = u \textit{ then the scheme consistent}$$

fails in the non-uniform mesh case. This criterion is sufficient for consistency on uniform meshes and we refer to Thomas [30] Chapter 9 for the proof, we just note here that the proof depends heavily on the uniformity of the mesh. We moreover refer to the last section of this Chapter where a generalisation of the consistency criterion -valid for non-uniform meshes- is proposed.

*Remark* 3.1.9. For comparison purposes we write the same scheme in the uniform mesh case, $h_i = \Delta x$ for every $i$,

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x} \left( F_{i+1/2} - F_{i-1/2} \right)$$

with

$$F_{i+1/2} = \frac{1}{2}\big(f(u_i) + f(u_{i+1})\big) - \frac{\Delta x}{2\Delta t}(u_{i+1} - u_i),$$

which is the uniform LxF scheme.

### 3.1.3 First order scheme - Vertex centered

In the previous section we discussed several possible non-uniform generalisations of the uniform LxF scheme. In this section we discuss a central first order scheme that is not a LxF generalisation.

We consider the non-uniform mesh

$$M_x = \{x_i, i \in \mathbb{Z}\} \quad \text{with} \quad h_i = x_i - x_{i-1}.$$

The middle points $x_{i-1/2} = \frac{x_{i-1}+x_i}{2}$ define a partition of the domain in cells

$$C_i = (x_{i-1/2}, x_{i+1/2}) \quad \text{with} \quad |C_i| = \frac{h_i + h_{i+1}}{2}.$$

For this description of the grid we consider the scheme

$$u_i^{n+1} = u_i^n - \frac{2\Delta t}{h_i + h_{i+1}} \left(F_{i+1/2} - F_{i-1/2}\right) \tag{3.12}$$

with

$$F_{i+1/2} = \frac{1}{2}\big(f(u_i) + f(u_{i+1})\big) + \frac{\Delta t}{2h_{i+1}}(u_i - u_{i+1}).$$

We note that this scheme is flux consistent, i.e., $F_{i+1/2}(u, u) = f(u)$.

**Proposition 3.1.4** (Modified equation for the scheme (3.12))**.** *The modified equation of the scheme (3.12) in the case of a linear flux $f(u) = au$ is*

$$\begin{aligned} u_t + au_x = &- \frac{a(h_i - h_{i+1}) + \Delta t(a^2 - 1)}{2} u_{xx} \\ &- \frac{ah_{i+1}^2 + (-3a + 2a^3)\Delta t^2 + (h_i - h_{i+1})(-3a^2\Delta t + ah_i + \Delta t)}{6} u_{xxx}. \end{aligned}$$

*Proof.* The proof follows from the same procedure as the ones performed in the previous schemes and is omitted. □

We then have,

**Corollary 3.1.3.1** (Consistency)**.** *This scheme is consistent in the sense that the right-hand side of the Modified Equation converges to 0 as the mesh is refined* $\max h_i, \Delta t \to 0$.

**Corollary 3.1.3.2** (Order of accuracy)**.** *This scheme constitutes a first order approximation of the Linear Transport equation.*

**Corollary 3.1.3.3** (Stability)**.** *This scheme is not stable in the sense that if* $0 < a < 1$ *and the mesh is decompressed* $h_i < h_{i+1}$*, the coefficient of the diffusion term* $u_{xx}$ *is negative.*

**Corollary 3.1.3.4** (Conservation)**.** *This scheme is obviously in conservative form.*

*Remark* 3.1.10. In the uniform mesh case, $h_i = \Delta x$ for every $i$, this scheme reduces to

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x}\left(F_{i+1/2} - F_{i-1/2}\right)$$

with numerical flux $F_{i+1/2} = F(u_i, u_{i+1})$ given by

$$F_{i+1/2} = \frac{1}{2}\big(f(u_i) + f(u_{i+1})\big) + \frac{\Delta t}{2\Delta x}(u_i - u_{i+1}).$$

It is an easy matter to see that the modified equation for this scheme is

$$u_t + au_x = \frac{\Delta t}{2}(a^2 - 1)u_{xx} - \frac{a\Delta x^2 + (-3a + 2a^3)\Delta t^2}{6}u_{xxx}.$$

For comparison purposes we state the LxF numerical flux

$$F_{i+1/2}^{LxF} = \frac{1}{2}\big(f(u_i) + f(u_{i+1})\big) + \frac{\Delta x}{2\Delta t}(u_i - u_{i+1}).$$

### 3.1.4 Another first order - Cell centered

In this approach we consider the non-uniform discretization of the domain in cells

$$C_i = (x_{i-1/2}, x_{i+1/2}) \quad \text{with} \quad |C_i| = h_i.$$

The mesh $M_x = \{x_i, i \in \mathbb{Z}\}$ consists of the middle points,

$$x_i = \frac{x_{i+1/2} + x_{i-1/2}}{2}; \quad \text{hence} \quad x_i - x_{i-1} = \frac{h_i + h_{i-1}}{2}.$$

For this description of the grid we propose the following scheme

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{h_i}(F_{i+1/2} - F_{i-1/2}),$$

where

$$F_{i+1/2} = \frac{h_{i+1}}{h_i + h_{i+1}}f_i + \frac{h_i}{h_i + h_{i+1}}f_{i+1} - \frac{2\Delta t}{h_i + h_{i+1}}(u_{i+1} - u_i).$$

*Remark* 3.1.11. We note that this scheme is consistent in the sense of the flux, that is $F_{i+1/2} = f(u)$ whenever $u_i = u_{i+1} = u$.

The modified equation (up to second order) for the linear flux case $f(u) = au$ is

$$u_t + au_x = \frac{-ah_i(h_{i+1} - h_{i-1}) + 2\Delta t(h_{i+1} + h_{i-1}) - 4(a^2-1)h_i\Delta t}{8h_i} u_{xx}.$$

**Corollary 3.1.4.1** (Consistency)**.** *This scheme is consistent in the sense that the right-hand side of the Modified Equation converges to 0 as the mesh is refined* $\max h_i, \Delta t \to 0$.

**Corollary 3.1.4.2** (Order of accuracy)**.** *This scheme constitutes a first order approximation of the Linear Transport equation.*

**Corollary 3.1.4.3** (Stability)**.** *Assuming a mesh smoothness variation* $m < \frac{h_{i\pm i}}{h_i} < M$ *the scheme is unstable in areas of decompression if* $a > \sqrt{M+1}$.

*Proof.* We examine the sign of the numerator of the diffusion coefficient

$$-ah_i(h_{i+1} - h_{i-1}) + 2\Delta t(h_{i+1} + h_{i-1}) - 4(a^2-1)h_i\Delta t,$$

which, if negative, yields

$$2\Delta t(h_{i-1} + h_{i+1}) < h_i \left(a(h_{i+1} - h_{i-1}) + 4(a^2-1)\Delta t\right)$$

or

$$\frac{h_{i-1} + h_{i+1}}{h_i} < \frac{a}{2}\frac{h_{i+1} - h_{i-1}}{\Delta t} + 2(a^2-1).$$

- If the mesh is uniform, then $m = M = 1$ and the previous relation is valid (thus the scheme is unstable) whenever $a > \sqrt{2}$.

- If the mesh is non-uniform, the previous relation is satisfied if

$$2M < \frac{a}{2}\frac{h_{i+1} - h_{i-1}}{\Delta t} + 2(a^2-1).$$

  So, in areas of mesh decompression $h_{i+1} > h_{i-1}$, a sufficient condition for instability is $a > \sqrt{M+1}$.

And this completes the proof of the corollary. □

### 3.1.5 Unstable centered, FTCS - Vertex centered

In this approach we consider the non-uniform mesh

$$M_x = \{x_i, i \in \mathbb{Z}\} \quad \text{with } h_i = x_i - x_{i-1}.$$

The middle points $x_{i-1/2} = \frac{x_{i-1}+x_i}{2}$ define a partition of the domain in cells

$$C_i = (x_{i-1/2}, x_{i+1/2}) \quad \text{with} \quad |C_i| = \frac{h_i + h_{i+1}}{2}.$$

For this description of the grid we discuss the known to be *unstable* Forward in Time Centered in Space (FTCS) scheme

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{h_i + h_{i+1}}\big(f(u_{i+1}) - f(u_{i-1}\big).$$

This scheme can be written in conservative form as follows

$$u_i^{n+1} = u_i^n - \frac{2\Delta t}{h_i + h_{i+1}}(F_{i+1/2} - F_{i-1/2})$$

with

$$F_{i+1/2} = \frac{f(u_i) + f(u_{i+1})}{2}.$$

We note that this scheme is consistent in the sense of the flux, that is $F_{i+1/2} = f(u)$ whenever $u_i = u_{i+1} = u$.

Regarding the modified equation for $f(u) = au$ we perform the usual procedure, which yields

$$u_t + au_x = a\frac{h_i - h_{i+1} - a\Delta t}{2}u_{xx} + \frac{a(h_i - 6a\Delta t)(h_{i+1} - h_i) - 2a^3\Delta t^2 - ah_{i+1}^2}{6}u_{xxx}.$$

**Corollary 3.1.5.1** (Consistency)**.** *This scheme is consistent in the sense that the right-hand side of the Modified Equation converges to 0 as the mesh is refined* $\max h_i, \Delta t \to 0$.

**Corollary 3.1.5.2** (Order of accuracy)**.** *This scheme constitutes a first order approximation of the Transport equation.*

**Corollary 3.1.5.3** (Stability)**.** *This scheme is unstable in areas of mesh decompression* $h_i < h_{i+1}$ *and conditionally stable under the severe restriction* $a\Delta t < h_i - h_{i+1}$ *in areas of mesh compression* $h_i > h_{i+1}$ *(this restriction is severe in the sense that the difference*

*$h_i - h_{i+1}$ can be arbitrarily small).*

**Corollary 3.1.5.4** (Conservation)**.** *This scheme is obviously in conservative form.*

*Remark* 3.1.12. For comparison purposes we state the modified equation of the same scheme for the case of a uniform mesh, i.e., $h_i = h$ for every $i$,

$$u_t + u_x = -\frac{\Delta t}{2} u_{xx} + \frac{4\Delta t^2 - h_i^2}{6} u_{xxx},$$

where the coefficient of the diffusive term $u_{xx}$ is negative for every space step $h$ and every time step $\Delta t$. Hence this scheme (uniform mesh) is *unstable*.

In this paragraph we studied a scheme that is consistent, conservative and only locally stable (under severe assumptions).We will use this scheme to study the stabilisation properties of the BAS (1.1.1) and more specifically of the Adaptive Mesh Reconstruction procedure. In fact, despite of the instability of the evolution step, we will see that the BAS will be stable.

## 3.2 Second order schemes

We continue with the study of the non-uniform versions/extensions of numerical schemes that are of second order accuracy on uniform meshes. We will see that it is not possible to maintain second order accuracy along with consistency and conservation, at least for 3-point or 3-cell schemes.

### 3.2.1 Richtmyer 2-step Lax-Wendroff - Cell centered

In this approach we consider the non-uniform discretization of the domain in cells

$$C_i = (x_{i-1/2}, x_{i+1/2}) \quad \text{with} \quad |C_i| = h_i.$$

The mesh $M_x = \{x_i, i \in \mathbb{Z}\}$ consists of the middle points,

$$x_i = \frac{x_{i+1/2} + x_{i-1/2}}{2}; \quad \text{hence} \quad x_i - x_{i-1} = \frac{h_i + h_{i-1}}{2}.$$

For this description of the grid we propose the following numerical scheme as the generalisation on non-uniform meshes of the Richtmyer 2-step Lax-Wendorff numerical

scheme,

$$\hat{u}_{i+1/2} = \frac{h_{i+1}}{h_i + h_{i+1}} u_i + \frac{h_i}{h_i + h_{i+1}} u_{i+1} - \frac{\Delta t}{h_i + h_{i+1}} (f_{i+1} - f_i)$$
$$u_i^{n+1} = u_i^n - \frac{\Delta t}{h_i} \big( f(\hat{u}_{i+1/2}) - f(\hat{u}_{i-1/2}) \big)$$

or, in conservative form

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{h_i} \big( F_{i+1/2} - F_{i-1/2} \big)$$

with

$$F_{i+1/2} = f(\hat{u}_{i+1/2}) = f\Big( \frac{h_{i+1}}{h_i + h_{i+1}} u_i + \frac{h_i}{h_i + h_{i+1}} u_{i+1} - \frac{\Delta t}{h_i + h_{i+1}} (f_{i+1} - f_i) \Big).$$

We note that this scheme is consistent in the sense of the flux, that is $F_{i+1/2} = f(u)$ whenever $u_i = u_{i+1} = u$.

Following the same procedure as with the previous schemes, we conclude to the modified equation of this scheme for $f(u) = au$,

$$\begin{aligned} u_t + a u_x = {} & \frac{a h_{i-1} h_i - a h_i h_{i+1} + a^2 \Delta t h_{i+1} + a^2 \Delta t h_{i-1} - 2 a^2 \Delta t h_i}{8 h_i} u_{xx} \\ & - \Bigg( \frac{2 a h_i^3 + 2 a h_{i-1} h_i^2 + 4 a^2 \Delta t h_i h_{i+1} - 44 a^3 \Delta t^2 h_i + 2 a h_i^2 h_{i+1} + a h_i h_{i-1}^2 + a h_{i-1}^2 h_i}{48 h_i} \\ & \quad + \frac{-4 a^2 \Delta t h_i h_{i-1} - a^2 \Delta t h_{i+1}^2 + a^2 \Delta t h_{i-1}^2 - 6 a^3 \Delta t^2 h_{i-1} - 6 a^3 \Delta t^2 h_{i-1}}{48 h_i} \Bigg) u_{xxx}. \end{aligned}$$

We then have,

**Corollary 3.2.1.1** (Consistency)**.** *The scheme is consistent in the sense that it approximates the transport equation as the mesh is refined,* $\max h_i \to 0$ *and* $\Delta t \to 0$.

**Corollary 3.2.1.2** (Order of accuracy)**.** *The scheme constitutes a first order approximation to the transport equation.*

**Corollary 3.2.1.3** (Stability)**.** *This scheme is unstable whenever the mesh configuration is* $h_{i-1} < h_{i+1} < h_i$.

*Proof.* The coefficient of the diffusion term $u_{xx}$ of the Modified equation

$$\frac{ah_{i-1}h_i - ah_ih_{i+1} + a^2\Delta t h_{i+1} + a^2\Delta t h_{i-1} - 2a^2\Delta t h_i}{8h_i}$$

reads, for $a = 1$,

$$\frac{h_{i-1} - h_{i+1}}{8} + \frac{\Delta t(h_{i+1} + h_{i-1} - 2h_i)}{8h_i},$$

which is negative for the mesh configuration $h_{i-1} < h_{i+1} < h_i$. □

**Corollary 3.2.1.4** (Conservation)**.** *This scheme is obviously in conservative form.*

*Remark* 3.2.1. This scheme reduces to the usual Richtmyer 2-step Lax-Wendroff whenever the mesh is uniform. In this case the modified equation recasts into

$$u_t + u_x = -\frac{h^2 - 7\Delta t^2}{6}u_{xxx},$$

which constitutes a second order approximation of the underlying transport equation $u_t + u_x = 0$. The increase of the accuracy in the uniform case is due to the fact that the coefficient of the diffusion term $u_{xx}$ discards -from the non-uniform modified equation- due to the summetry of the mesh.

### 3.2.2 MacCormack - Cell centered

In this approach we consider the non-uniform discretization of the domain in cells

$$C_i = (x_{i-1/2}, x_{i+1/2}) \quad \text{with} \quad |C_i| = h_i.$$

The mesh $M_x = \{x_i, i \in \mathbb{Z}\}$ consists of the middle points

$$x_i = \frac{x_{i+1/2} + x_{i-1/2}}{2}; \quad \text{hence} \quad x_i - x_{i-1} = \frac{h_i + h_{i-1}}{2}.$$

For this description of the grid we propose the following scheme as the generalisation of the MacCormack,

$$\begin{aligned} u_i^* &= u_i^n - \frac{\Delta t}{h_i}(f(u_{i+1}) - f(u_i)) \\ u_i^{**} &= u_i^n - \frac{\Delta t}{h_i}(f(u_i^*) - f(u_{i-1}^*)) \\ u_i^{n+1} &= \frac{u_i^* + u_i^{**}}{2} \end{aligned}$$

or, in conservative form,

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{h_i}\big(F_{i+1/2} - F_{i-1/2}\big)$$

with

$$F_{i+1/2} = \frac{f(u_{i+1}^n) + f(u_i^*)}{2}.$$

Following the previous paragraphs we conclude to the modified equation for the case of a linear flux $f(u) = u$,

$$u_t - \frac{-4h_ih_{i-1}h_{i+1} - 4h_ih_{i-1}^2 + 4dth_{i-1}h_{i+1} - 4dth_i^2 - 8h_i^2h_{i-1}}{16h_{i-1}h_i^2}u_x = \mathcal{O}(h)u_{xx},$$

where we have omitted the diffusion coefficient in the right-hand side, because of the size of the expression.

*Remark* 3.2.2. The modified equation, clearly states that this scheme is not consistent with the original transport equation $u_t + u_x = 0$.

For a better approach we consider a Vertex centered description for the MacCormack scheme,

**MacCormack - Vertex centered** In this approach we consider the non-uniform mesh

$$M_x = \{x_i, i \in \mathbb{Z}\} \quad \text{with } h_i = x_i - x_{i-1}.$$

The middle points $x_{i-1/2} = \frac{x_{i-1}+x_i}{2}$ define a partition of the domain in cells

$$C_i = (x_{i-1/2}, x_{i+1/2}) \quad \text{with} \quad |C_i| = \frac{h_i + h_{i+1}}{2}.$$

For this description of the grid we propose the following scheme as the generalisation of the MacCormack,

$$\begin{aligned} u_i^* &= u_i^n - \frac{2\Delta t}{h_i + h_{i+1}}(f(u_{i+1}) - f(u_i)) \\ u_i^{**} &= u_i^* - \frac{2\Delta t}{h_{i-1} + h_i}(f(u_i^*) - f(u_{i-1}^*)) \\ u_i^{n+1} &= \frac{u_i^n + u_i^{**}}{2}. \end{aligned}$$

For the stability analysis we consider the linear case $f(u) = au$ and by collecting all the

terms in the right-hand side, expanding the terms $u^n_{i\pm1}$ in Taylor series we end up with the following modified equation,

$$u_t + u_x = \frac{ah_{i-1}h_i - ah_{i+1}h_{i-1} - ah_{i+1}h_i + 2a^2\Delta t h_{i+1} + ah_{i-1}^2 - 2a^2\Delta t h_{i-1}}{8(h_{i-1}+h_i)}u_{xx}$$

or, after some algebra,

$$u_t + au_x = a\frac{(h_{i-1}-h_{i+1})(h_{i-1}+h_i-2a\Delta t)}{8(h_{i-1}+h_i)}u_{xx},$$

where we have omitted the third derivative because of the size of its coefficient expression.

**Corollary 3.2.2.1** (Consistency)**.** *The scheme is consistent in the sense that the right-hand side of the Modified Equation converges to 0 as the mesh is refined,* $\max h_i, \Delta t \to 0$.

**Corollary 3.2.2.2** (Order of accuracy)**.** *The scheme constitutes a first order approximation of the transport equation.*

**Corollary 3.2.2.3** (Stability)**.** *The scheme is unstable in areas of decompression and conditionally stable in areas of compression.*

*Proof.* With the usual CFL condition $\Delta t < \min h_i$ the term $h_{i-1} + h_i - 2\Delta t$ of the diffusion coefficient is kept positive. The other term $h_{i-1} - h_{i+1}$ changes sign according to the compression/decompression of the mesh. More specifically it is negative whenever $h_{i-1} < h_{i+1}$ -which occurs on decompression. So the scheme is unstable in areas of decompression and conditionally stable in areas of compression. □

**Corollary 3.2.2.4** (Conservation)**.** *This scheme cannot be written in conservative form.*

*Remark* 3.2.3. In the uniform mesh case, the scheme recasts to the usual MacCormack scheme [21] and the coefficient of second derivative in the Modified Equation vanishes. Hence this scheme is second order accurate approximation to the Transport equation.

### 3.2.3 Generalized LxW - Pure second order

This paragraph is part of a join work with Ch. Makridakis and S. Noelle [22]

The scheme we have studied so far are of first order accuracy on their non-uniform mesh versions even if their uniform mesh version where of second order. In this paragraph we construct a second order accurate scheme, on non-uniform mesh. This is achieved on the expense of conservation.

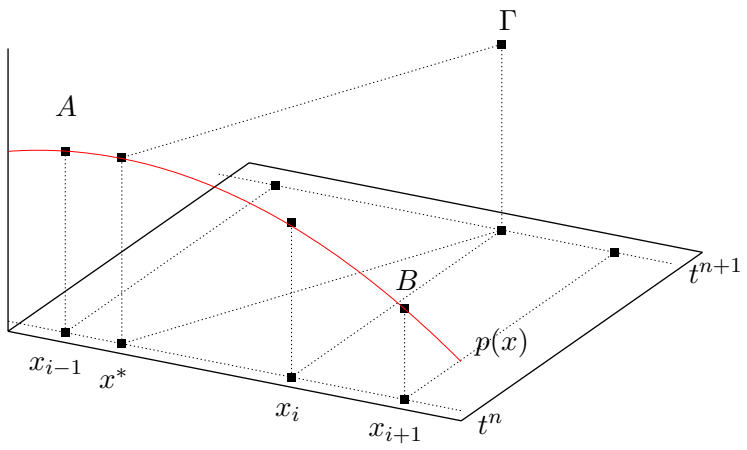



Figure 3.2: Geometric construction of a second order, LxW type scheme. The value $u_i^{n+1}$ at point $\Gamma$ is defined as $u_i^{n+1} = p(x^*)$, where $x^*$ is the intersection point of the characteristic line -emanating from $(x_i, t^{n+1})$- with the interval $[x_{i-1}, x_{i+1}]$ at time $t^n$. The polynomial $p(x)$ interpolates the points $A$, $B$ and $(x_i, u_i^n)$, it is of second degree so it is possible to attain its maximum in the interval $(x_{i-1}, x_{i+1})$ and not at the end points $x_{i-1}$, $x_{i+1}$, hence overshoot is possible. This explains geometrically the oscillations that second order numerical schemes produce.

We start by considering the transport equation

$$u_t + au_x = 0, \quad x \in [0,1], t \in [0,T].$$

We recall that the LxF scheme can be geometrically constructed using the linear interpolant of the points $(x_{i-1}, u_{i-1})$ and $(x_{i+1}, u_{i+1})$. It is known also that the LxW scheme can be constructed using the quadratic interpolant of the points $(x_{i-1}, u_{i-1})$, $(x_i, u_i)$ and $(x_{i+1}, u_{i+1})$. In the construction that follows we utilise this approach to construct a second order accurate scheme for non-uniform meshes.

*Construction.* We refer to Figure 3.2 for a graphical explanation of the construction.

Consider the point values $A = (x_{i-1}, u_{i-1}^n)$, $B = (x_i, u_i^n)$, $\Gamma = (x_{i+1}, u_{i+1}^n)$. The

unique second order polynomial interpolating these points is given by

$$p(x) = \frac{(x-x_i)(x-x_{i+1})}{(x_{i-1}-x_i)(x_{i-1}-x_{i+1})}u_{i-1}^n + \frac{(x-x_{i-1})(x-x_{i+1})}{(x_i-x_{i-1})(x_i-x_{i+1})}u_i^n + \\ + \frac{(x-x_{i-1})(x-x_i)}{(x_{i+1}-x_{i-1})(x_{i+1}-x_i)}u_{i+1}^n.$$

We write $h_i = x_i - x_{i-1}$ and $h_{i+1} = x_{i+1} - x_i$, so the previous polynomial reads

$$p(x) = \frac{(x-x_i)(x-x_i-h_{i+1})}{h_i(h_i+h_{i+1})}u_{i-1}^n - \frac{(x-x_i+h_i)(x-x_i-h_{i+1})}{h_i h_{i+1}}u_i^n \\ + \frac{(x-x_i+h_i)(x-x_i)}{(h_i+h_{i+1})h_{i+1}}u_{i+1}^n.$$

To evaluate the numerical solution at the node $x_i$ at the time $t^{n+1}$, we travel backwards in time along the respective characteristic until we intersect time $t = t^n$. This will result in a point $x$ which -for small enough $\Delta t$- shall belong to the interval $[x_{i-1}, x_{i+1}]$. Since along the characteristics the solution is constant, the new nodal value $u_i^{n+1}$ will be equal to the value of the polynomial $p(x)$, where $x$ is the point previously described.

Now $1/a \in [-\infty, +\infty] - \{0\}$ is the slope of the characteristic line passing through the point $x_i$ at the time step $n+1$. By the slope of the characteristic we have that

$$\frac{1}{a} = \frac{\Delta t}{x_i - x} \quad \text{and so } x - x_i = -a\Delta t;$$

hence the resulting value of the polynomial at the point $x = x_i - a\Delta t$ will provide us with the desired $u_i^{n+1}$ value. So

$$\begin{aligned} u_i^{n+1} &= p(x_i^n - a\Delta t) \\ &= \frac{a\Delta t(a\Delta t + h_{i+1})}{h_i(h_i+h_{i+1})}u_{i-1}^n + \frac{(-a\Delta t + h_i)(a\Delta t + h_{i+1})}{h_i h_{i+1}}u_i^n - \frac{a\Delta t(-a\Delta t + h_i)}{(h_i+h_{i+1})h_{i+1}}u_{i+1}^n \end{aligned}$$

or

$$u_i^{n+1} = \frac{a\Delta t(a\Delta t + h_{i+1})}{h_i(h_i+h_{i+1})}u_{i-1}^n + \frac{(-a\Delta t + h_i)(a\Delta t + h_{i+1})}{h_i h_{i+1}}u_i^n \\ + \frac{a\Delta t(a\Delta t - h_i)}{(h_i+h_{i+1})h_{i+1}}u_{i+1}^n. \tag{3.13}$$

By setting the coefficient of $u_{i-1}^n$, $u_i^n$ and $u_{i+1}^n$ as $\alpha$, $\beta$, $\gamma$, respectively, i.e.,

$$\begin{aligned}\alpha &= \frac{a\Delta t(a\Delta t + h_{i+1})}{h_i(h_i + h_{i+1})}\\ \beta &= \frac{(-a\Delta t + h_i)(a\Delta t + h_{i+1})}{h_i h_{i+1}}\\ \gamma &= \frac{a\Delta t(a\Delta t - h_i)}{(h_i + h_{i+1})h_{i+1}}\end{aligned}$$

it is an easy task to verify that $\alpha + \beta + \gamma = 1$ and $0 < \alpha \leq 1$ and $-1 \leq \gamma \leq 0$ and $0 < \beta \leq 2$ as long as the CFL-like requirements $a\Delta t < h_i$ and $a\Delta t < h_{i+1}$ are satisfied.

*Remark* 3.2.4. It is an easy matter to check that this scheme results in the usual LxW scheme whenever the mesh is uniform. So, in that sense, this scheme is considered to be a generalisation of the usual LxW scheme.

We move on to the analysis of this numerical scheme and by following the procedure presented in previous paragraphs we construct the modified equation of this scheme,

**Proposition 3.2.1** (Modified equation for the scheme (3.13))**.** *The modified equation of the scheme (3.13) is*

$$\begin{aligned}u_t + au_x =& \frac{a(a\Delta t - h_i)(a\Delta t + h_{i+1})}{6}u_{xxx}\\ &+ \frac{a(h_i - a\Delta t)(a\Delta t + h_{i+1})(h_i - 3a\Delta t - h_{i+1})}{24}u_{xxxx}.\end{aligned} \tag{3.14}$$

*Proof.* As with the previous cases, for the construction of the Modified Equation we need to consider a sufficiently smooth function $u$ that satisfies the numerical scheme (3.13) exactly at every node $(x_i, t^n)$. That is $u(x_i, t^n) = u_i^n$ for every $i$ and $n$ and the scheme (3.13) reads,

$$\begin{aligned}u(x_i, t^{n+1}) =& \frac{a\Delta t(a\Delta t + h_{i+1})}{h_i(h_i + h_{i+1})}u(x_{i-1}, t^n) + \frac{(-a\Delta t + h_i)(a\Delta t + h_{i+1})}{h_i h_{i+1}}u(x_i, t^n)\\ &- \frac{a\Delta t(-a\Delta t + h_i)}{(h_i + h_{i+1})h_{i+1}}u(x_{i+1}, t^n).\end{aligned} \tag{3.15}$$

We expand in Taylor series up to order 4 with respect to the node $(x_i, t^n)$,

$$\begin{aligned}
u(x_i, t^{n+1}) =& u(x_i, t^n) + \Delta t u_t(x_i, t^n) + \frac{\Delta t^2}{2} u_{tt}(x_i, t^n) + \frac{\Delta t^3}{6} u_{ttt}(x_i, t^n) + \frac{\Delta t^4}{24} u_{tttt}(x_i, t^n) \\
u(x_{i-1}, t^n) =& u(x_i, t^n) - h_i u_x(x_i, t^n) + \frac{h_i^2}{2} u_{xx}(x_i, t^n) - \frac{h_i^3}{6} u_{xxx}(x_i, t^n) + \frac{h_i^4}{24} u_{xxxx}(x_i, t^n) \\
u(x_{i+1}, t^n) =& u(x_i, t^n) + h_{i+1} u_x(x_i, t^n) + \frac{h_{i+1}^2}{2} u_{xx}(x_i, t^n) + \frac{h_{i+1}^3}{6} u_{xxx}(x_i, t^n) + \frac{h_{i+1}^4}{24} u_{xxxx}(x_i, t^n).
\end{aligned}$$

We replace these expansion into equation (3.15), divide by $\Delta t$, simplify, gather to the left hand side, to get the half-way modified equation -we omit $(x_i, t^n)$ for the sake of the presentation,

$$\begin{aligned}
& u_t + a u_x + \frac{\Delta t}{2} u_{tt} - \frac{a^2 \Delta t}{2} u_{xx} \\
& \quad + \frac{\Delta t^2}{6} u_{ttt} + \frac{a}{6}(-a\Delta t h_{i+1} + h_i h_{i+1} + a\Delta t h_i) u_{xxx} \\
& \quad + \frac{\Delta t^3}{24} u_{tttt} + \frac{a}{24}(h_i h_{i+1}^2 - a\Delta t h_{i+1}^2 - h_i^2 h_{i+1} + a\Delta t h_i h_{i+1} - a\Delta t h_i^2) u_{xxxx} = 0.
\end{aligned} \tag{3.16}$$

We need the full modified equation so we repeat the work done in the previous paragraphs in order to replace the derivatives with respect to $t$:

$$\begin{aligned}
\text{Eq(3.16)} \xrightarrow{\partial t} u_{tt} = & -a u_{xt} - \frac{\Delta t}{2} u_{ttt} + \frac{a^2 \Delta t}{2} u_{xxt} - \frac{\Delta t^2}{6} u_{tttt} \\
& - \frac{a}{6}(-a\Delta t h_{i+1} + h_i h_{i+1} + a\Delta t h_i) u_{xxxt} + O(\Delta t^3)
\end{aligned}$$

and

$$\begin{aligned}
\text{Eq(3.16)} \xrightarrow{\partial x} u_{xt} = & -a u_{xx} - \frac{\Delta t}{2} u_{xtt} + \frac{a^2 \Delta t}{2} u_{xxx} - \frac{\Delta t^2}{6} u_{xttt} \\
& - \frac{a}{6}(-a\Delta t h_{i+1} + h_i h_{i+1} + a\Delta t h_i) u_{xxxx} + O(\Delta t^3);
\end{aligned}$$

now for the third order derivatives,

$$\begin{aligned}
\text{Eq(3.16)} \xrightarrow{\partial tt} u_{ttt} &= -a u_{xtt} - \frac{\Delta t}{2} u_{tttt} + \frac{a^2 \Delta t}{2} u_{xxtt} + O(\Delta t^2) \\
\text{Eq(3.16)} \xrightarrow{\partial xt} u_{xtt} &= -a u_{xxt} - \frac{\Delta t}{2} u_{xttt} + \frac{a^2 \Delta t}{2} u_{xxxt} + O(\Delta t^2) \\
\text{Eq(3.16)} \xrightarrow{\partial xx} u_{xxt} &= -a u_{xxx} - \frac{\Delta t}{2} u_{xxtt} + \frac{a^2 \Delta t}{2} u_{xxxx} + O(\Delta t^2).
\end{aligned}$$

Finally for the fourth order,

$$\begin{aligned}
\text{Eq}(3.16) &\xrightarrow{\partial xtt} u_{xttt} = -au_{xxtt} + O(\Delta t) \\
\text{Eq}(3.16) &\xrightarrow{\partial xxt} u_{xxtt} = -au_{xxxt} + O(\Delta t) \\
\text{Eq}(3.16) &\xrightarrow{\partial xxx} u_{xxxt} = -au_{xxxx} + O(\Delta t).
\end{aligned}$$

We now substitute the previous derivatives in the half-way modified equation Eq.(3.16) and after some algebra, we end up with the following full modified equation

$$\begin{aligned}
u_t + au_x =& \frac{a(a\Delta t - h_i)(a\Delta t + h_{i+1})}{6} u_{xxx} \\
&+ \frac{a(h_i - a\Delta t)(a\Delta t + h_{i+1})(h_i - 3a\Delta t - h_{i+1})}{24} u_{xxxx}
\end{aligned} \tag{3.17}$$

-up to fourth derivative- of our numerical scheme (3.13), valid for both uniform and non-uniform meshes. □

*Remark* 3.2.5. The modified equation (3.14) reveals that our numerical scheme provides with solutions that are second order accurate approximations of the initial Transport equation, third order accurate approximations of a Dispersion equation and fourth order accurate approximations of a Dispersion-Diffusion equation (higher order diffusion).

We then have,

**Corollary 3.2.3.1** (Consistency)**.** *The scheme (3.13) is consistent in the sense that it approximates the transport equation as the mesh is refined,* $\max h_i \to 0$ *and* $\Delta t \to 0$.

**Corollary 3.2.3.2** (Order of accuracy)**.** *The scheme (3.13) constitutes a second order approximation of the transport equation.*

**Corollary 3.2.3.3** (Stability)**.** *Under the CFL requirement* $a\Delta t < \min h_i$ *this scheme is unstable in areas of mesh decompression with* $h_{i+1} < h_i - 3a\Delta t$.

*Proof.* This result is obvious since, for stability, the diffusion coefficient (4th order diffusion) should be negative

$$\frac{a(h_i - a\Delta t)(a\Delta t + h_{i+1})(h_i - 3a\Delta t - h_{i+1})}{24} \leq 0.$$

This is not the case, if $h_{i+1} < h_i - 3a\Delta t$. □

**Corollary 3.2.3.4** (Conservation). *The scheme (3.13) cannot be written in conservative form.*

*Remark* 3.2.6. Whenever the mesh is uniform, the modified equation (3.14) results to the usual modified equation of the LxW scheme ([20]); thus it serves as a generalisation of the LxW scheme for uniform meshes. To expand more on this subject let us set $h_i = h_{i+1} = h$ and recast the modified equation (3.14) into

$$\begin{aligned} u_t + u_x =& \frac{a}{6}(a^2\Delta t^2 - h^2)u_{xxx} + \frac{3a^2\Delta t}{24}(a^2\Delta t^2 - h^2)u_{xxxx} \\ =& \frac{a}{6}h^2\left(\left(\frac{a\Delta t}{h}\right)^2 - 1\right)u_{xxx} + \frac{a}{8}h^3\frac{a\Delta t}{h}\left(\left(\frac{a\Delta t}{h}\right)^2 - 1\right)u_{xxxx} \\ =& \frac{a}{6}h^2\left(\nu^2 - 1\right)u_{xxx} + \frac{a}{8}h^3\nu\left(\nu^2 - 1\right)u_{xxxx} \end{aligned}$$

where $\nu = \frac{a\Delta t}{h}$ is the Courant number. One can easily see that for $0 < \nu < 1$ the coefficient of the 4th derivative has the correct sign (negative).

By keeping $h$ constant and letting $\Delta t \to 0$ we have $\nu \to 0$. In this regime we notice that the coefficient of $u_{xxx}$ (dispersion) increases in absolute value and the coefficient of $u_{xxxx}$ (diffusion) decreases in magnitude. This possibly explains the increase of oscillations that we notice in uniform meshes, as we decrease the time step.

Moreover, while keeping $\nu$ constant and decreasing $h$, the diffusion coefficient decreases faster -like $h^3$- than the dispersion coefficient -like $h^2$. The significant presence of the dispersion term explains the growth of oscillations we observe numerically as we refine both the space and the time steps.

## 3.3 Consistency criterion extended

Throughout the analysis of the previous schemes, consistency was studied in terms of the modified equation. Moreover for every conservative scheme we examined the Numerical Flux Consistency Criterion and in one case we demonstrated that it is not sufficient for non-uniform mesh cases. That is even though the usual Flux Consistency criterion stated that the scheme should be consistent with the underlying Differential Equation the modified equation showed that it was not. We refer to [15] Lemma 3.2.8 for another example of this criterion failure.

This last remark gives rise to the question of extending the flux Consistency Criterion to include non-uniform meshes and schemes. This is exactly the purpose of this

paragraph.

We start by stating the usual/uniform flux Consistency Criterion and then we propose extensions for conservative numerical schemes for both Cell centered and Vertex centered grids. We also prove that these extensions are sufficient to achieve consistency with the underlying PDE

$$u_t + f(u)_x = 0.$$

### 3.3.1 Uniform mesh case

Let a 3-point explicit numerical scheme on a uniform mesh be given in conservative form

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x}\left(F_{i+1/2} - F_{i-1/2}\right),$$

where $\Delta x$ is the uniform space step and $F_{i+1/2} = F(u_i^n, u_{i+1}^n)$ with $F(\cdot,\cdot)$ Lipschitz continuous numerical flux function.

**Definition 3.3.1** (Consistency)**.** The finite difference scheme

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x}\left(F_{i+1/2} - F_{i-1/2}\right)$$

is pointwise consistent with the conservation law $u_t + f(u)_x = 0$ at the point $(x,t)$ if the local truncation error $\tau_i^n$ defined via

$$\tau_i^n \Delta t = u(x_i, t^n + \Delta t) - u(x_i, t^n) + \frac{\Delta t}{\Delta x}\left(F\left(u(x_i,t^n), u(x_{i+1},t^n)\right) - F(u(x_i,t^n), u(x_{i+1},t^n))\right)$$

where $u$ is a solution to the conservation law, converges when then mesh is refined

$$\tau_i^n \to 0 \quad \text{as } \Delta t, \Delta x \to 0 \text{ and } (i\Delta x, n\Delta t) \to (x,t).$$

With respect to this consistency definition we state the following criterion,

**Proposition 3.3.1** (Consistency criterion for uniform meshes)**.** *If $F(u,u) = f(u)$, the finite difference scheme $u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x}\left(F_{i+1/2} - F_{i-1/2}\right)$ is consistent with the conservation law $u_t + f(u)_x = 0$.*

*Proof.* For the proof we refer to Thomas [30], Chapter 9. We only mention here that the proof depends heavily on the uniformity of the mesh and that the actual requirement is $\partial_x F(\bar{u},\bar{u}) = \partial_x f(\bar{u})$. □

We continue by studying the extension of the Consistency criterion for the case of cell centered non-uniform grids,

### 3.3.2 Non-uniform Cell centered grids

We consider the non-uniform discretization of the domain in cells

$$C_i = (x_{i-1/2}, x_{i+1/2}) \quad \text{with} \quad |C_i| = h_i.$$

The mesh $M_x = \{x_i, i \in \mathbb{Z}\}$ consists of the middle points

$$x_i = \frac{x_{i+1/2} + x_{i-1/2}}{2}; \quad \text{hence} \quad x_i - x_{i-1} = \frac{h_i + h_{i-1}}{2}.$$

Let also a 3-point explicit numerical scheme on a non-uniform mesh be given in conservative form

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{h_i}\left(F_{i+1/2} - F_{i-1/2}\right).$$

The numerical flux depends (generally) on the local size of the mesh, hence a better description of the non-uniform numerical scheme is

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{h_i}\left(F(u_i, u_{i+1}, h_i, h_{i+1}) - F(u_{i-1}, u_i, h_{i-1}, h_i)\right); \tag{3.18}$$

to simplify the notation we set

$$F^{i+1/2}(u_i, u_{i+1}) = F(u_i, u_{i+1}, h_i, h_{i+1})$$
$$F^{i-1/2}(u_{i-1}, u_i) = F(u_{i-1}, u_i, h_{i-1}, h_i),$$

where the superscripts $^{i+1/2}$ and $^{i-1/2}$ are used to denote the dependence of the flux on the respective mesh steps. With this notation the numerical scheme reads,

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{h_i}\left(F^{i+1/2}(u_i, u_{i+1}) - F^{i-1/2}(u_{i-1}, u_i)\right); \tag{3.19}$$

*Remark* 3.3.1. We note that whenever the mesh is uniform $h_{i-1} = h_i = h_{i+1} = \Delta x$,

$$F^{i+1/2}(u, v) = F^{i-1/2}(u, v) = F(u, v, \Delta x, \Delta x),$$

which justifies the notion of extension of the flux.

Regarding the consistency of this scheme with the underlying original conservation

law

$$u_t + f(u)_x = 0$$

we state the following proposition

**Theorem 3.3.1** (Consistency criterion for non uniform meshes on cell centered grids)**.** *The non-uniform 3-point, conservative numerical scheme (3.19), is consistent with the original underlying conservation law, if for every value $u_i^n$ the numerical flux functions $F^{i+1/2}$ and $F^{i-1/2}$ satisfy the following relations,*

$$F^{i+1/2}(u_i^n, u_i^n) = F^{i-1/2}(u_i^n, u_i^n)$$
$$\frac{h_{i-1}+h_i}{2h_i} F_1^{i-1/2}(u_i^n, u_i^n) + \frac{h_i + h_{i+1}}{2h_i} F_2^{i+1/2}(u_i^n, u_i^n) = f'(u_i^n),$$

*where the subscripts $_1$ and $_2$ in the numerical fluxes $F^{i+1/2}$ and $F^{i-1/2}$ denote the partial derivatives with respect to the first or the second variable, respectively.*

*Remark* 3.3.2. This criterion constitutes a generalisation of the uniform mesh case criterion, i.e., $F(u,u) = f(u)$ in the sense that, whenever $h_i = \Delta x$ for every $i$, it reads

$$\frac{\Delta x + \Delta x}{2\Delta x} F_1^{i-1/2}(u_i^n, u_i^n) + \frac{\Delta x + \Delta x}{2\Delta x} F_2^{i+1/2}(u_i^n, u_i^n) = f'(u_i^n)$$

or

$$F_1^{i-1/2}(u_i^n, u_i^n) + F_2^{i+1/2}(u_i^n, u_i^n) = f'(u_i^n);$$

hence

$$\frac{\partial}{\partial u} F(u,u) = \frac{\partial}{\partial u} f(u)$$

or

$$F(u,u) = f(u) + \text{constant}$$

but the conservation law $u_t + f(u)_x = 0$ is invariant under translation of the flux function $f$; hence we conclude that $F(u,u) = f(u)$.

*Proof of the Theorem.* We start by performing Truncation error analysis on the non-uniform numerical scheme, for $tr = \tau_i^n \Delta t$

$$tr = u_i^{n+1} - u_i^n + \frac{\Delta t}{h_i}\left(F^{i+1/2}(u_i^n, u_{i+1}^n) - F^{i-1/2}(u_{i-1}^n, u_i^n)\right).$$

We now substitute the terms $u_i^{n+1}$, $F^{i+1/2}(u_i^n, u_{i+1}^n)$, $F^{i-1/2}(u_{i-1}^n, u_i^n)$ by their respective Taylor expansions about the point $u_i^n$ to get

$$tr = u_i^n + \Delta t(u_t)_i^n + \mathcal{O}(\Delta t^2) - u_i^n + \frac{\Delta t}{h_i}\Big(F^{i+1/2}(u_i^n, u_i^n) + F_2^{i+1/2}(u_i^n, u_i^n)(u_{i+1}^n - u_i^n) + \dots \\ - F^{i-1/2}(u_i^n, u_i^n) - F_1^{i-1/2}(u_i^n, u_i^n)(u_{i-1}^n - u_i^n) - \dots\Big);$$

by expanding now the terms $u_{i+1}^n$, $u_{i-1}^n$ in Taylor series around the point $u_i^n$, the truncation term $tr$ takes the form

$$tr = \Delta t(u_t)_i^n + \mathcal{O}(\Delta t^2) + \frac{\Delta t}{h_i}\left(F_2^{i+1/2}(u_i^n, u_i^n)\left(\frac{h_i + h_{i+1}}{2}(u_x)_i^n + \mathcal{O}(h^2)\right) + \dots \right. \\ \left. - F_1^{i-1/2}(u_i^n, u_i^n)\left(-\frac{h_{i-1} + h_i}{2}(u_x)_i^n + \mathcal{O}(h^2)\right) - \dots\right),$$

where the terms $\mathcal{O}(h^2)$ are used as abbreviations for higher than second order terms with respect to the step sizes $h_{i-1}$, $h_i$, $h_{i+1}$. We also exploited the first requirement of the proposition, $F^{i+1/2}(u_i^n, u_i^n) = F^{i-1/2}(u_i^n, u_i^n)$.

This last relation can also be written as follows

$$tr = \Delta t(u_t)_i^n + \frac{\Delta t}{h_i}\Big(F_2^{i+1/2}(u_i^n, u_i^n)\frac{h_i + h_{i+1}}{2}(u_x)_i^n + F_1^{i-1/2}(u_i^n, u_i^n)\frac{h_{i-1} + h_i}{2}(u_x)_i^n\Big) \\ + \mathcal{O}(\Delta t^2) + \mathcal{O}\left(\frac{\Delta t}{h_i}h^2\right)$$

and finally recasts into

$$tr = \left(u_t + \Big(F_2^{i+1/2}(u_i^n, u_i^n)\frac{h_i + h_{i+1}}{2h_i} + F_1^{i-1/2}(u_i^n, u_i^n)\frac{h_{i-1} + h_i}{2h_i}\Big)(u_x)_i^n\right)\Delta t \\ + \mathcal{O}(\Delta t^2) + \mathcal{O}\left(\frac{\Delta t}{h_i}h^2\right).$$

We close the proof utilising the second requirement of the proposition, namely $\frac{h_{i-1}+h_i}{2h_i}F_1^{i-1/2}(u_i^n, u_i^n) +$

$\frac{h_i+h_{i+1}}{2h_i} F_2^{i+1/2}(u_i^n, u_i^n) = f'(u_i^n)$ and so the truncation term $tr$ reads

$$\begin{aligned} tr &= \left( (u_t)_i^n + f'(u_i^n)(u_x)_i^n \right) \Delta t + \mathcal{O}(\Delta t^2) + \mathcal{O}(\frac{\Delta t}{h_i} h^2) \\ &= \mathcal{O}(\Delta t^2) + \mathcal{O}(\frac{\Delta t}{h_i} h^2), \end{aligned}$$

and finally the truncation error yields

$$\tau_i^n = \mathcal{O}(\Delta t) + \mathcal{O}(\frac{h^2}{h_i})$$

so the non-uniform numerical scheme is consistent with the original conservation law. □

We continue with the extension of the consistency criterion for the case of vertex centered grids.

### 3.3.3 Vertex centered grids

In this approach we consider the non-uniform mesh

$$M_x = \{x_i, i \in \mathbb{Z}\} \quad \text{with} \quad h_i = x_i - x_{i-1}.$$

The middle points $x_{i-1/2} = \frac{x_{i-1}+x_i}{2}$ define a partition of the domain in cells

$$C_i = (x_{i-1/2}, x_{i+1/2}) \quad \text{with} \quad |C_i| = \frac{h_i + h_{i+1}}{2}.$$

Let also a 3-point explicit numerical scheme on a non-uniform mesh be given in conservative form

$$u_i^{n+1} = u_i^n - \frac{2\Delta t}{h_i + h_{i+1}} \left( F^{i+1/2}(u_i^n, u_{i+1}^n) - F^{i-1/2}(u_{i-1}^n, u_i^n) \right),$$

where the numerical flux functions $F^{i+1/2}$ and $F^{i-1/2}$ depend also on the local mesh sizes.

As in the previous paragraph, the uniform mesh consistency criterion can be extended as follows,

**Theorem 3.3.2** (Consistency criterion for non uniform meshes with vertex centered grids)**.** *The non-uniform 3-point, conservative numerical scheme, is consistent with the*

*original conservation law if for every value $u_i^n$ the numerical flux functions $F^{i+1/2}$ and $F^{i-1/2}$ satisfy the following relation*

$$F^{i+1/2}(u_i^n, u_i^n) = F^{i-1/2}(u_i^n, u_i^n)$$
$$\frac{2h_i}{h_i + h_{i+1}} F_1^{i-1/2}(u_i^n, u_i^n) + \frac{2h_{i+1}}{h_i + h_{i+1}} F_2^{i+1/2}(u_i^n, u_i^n) = f'(u_i^n),$$

*where the subscripts $_1$ and $_2$ in the numerical fluxes $F^{i+1/2}$ and $F^{i-1/2}$ denote the partial derivatives with respect to the first or the second variable, respectively.*

*Proof.* The proof is similar to the Cell centered case and is omitted. □

### 3.3.4 Numerical fluxes revisited

In this paragraph we examine the non-uniform consistency criteria we developed in the previous paragraph. The numerical schemes we devised in the first section of this Chapter have already been examined for their consistency in terms of the modified equation. We now repeat the tests in terms of the extended consistency criteria we just devised. We first examine the conservative schemes acting over Vertex centered grids and then the conservative schemes acting over Cell centered grids.

*Remark* 3.3.3. To simplify the notation we further set, for the rest of this paragraph

$$F^R(u, v) = F^{i+1/2}(u, v) = F(u, v, h_i, h_{i+1})$$
$$F^L(u, v) = F^{i-1/2}(u, v) = F(u, v, h_{i-1}, h_i)$$

#### 3.3.4.1 Vertex centered grids

In this approach we consider the non-uniform mesh

$$M_x = \{x_i, i \in \mathbb{Z}\} \quad \text{with} \quad h_i = x_i - x_{i-1}.$$

The middle points $x_{i-1/2} = \frac{x_{i-1}+x_i}{2}$ define a partition of the domain in cells

$$C_i = (x_{i-1/2}, x_{i+1/2}) \quad \text{with} \quad |C_i| = \frac{h_i + h_{i+1}}{2}.$$

For this description, the schemes attain the following conservative form

$$u_i^{n+1} = u_i^n - \frac{2\Delta t}{h_i + h_{i+1}} \left( F_{i+1/2}^R - F_{i-1/2}^L \right).$$

We restate the consistency requirements for the Vertex centered grid configuration as established in Th.(3.3.2)

1. $F^R(u_i^n, u_i^n) = F^L(u_i^n, u_i^n)$
2. $\frac{2h_i}{h_i+h_{i+1}} F_1^L(u_i^n, u_i^n) + \frac{2h_{i+1}}{h_i+h_{i+1}} F_2^R(u_i^n, u_i^n) = f'(u_i)$.

**First order scheme - Vertex centered** The numerical flux for this conservative scheme is

$$F^R(u_i, u_{i+1}) = \frac{1}{2}\big(f(u_i) + f(u_{i+1})\big) + \frac{\Delta t}{2h_{i+1}}(u_i - u_{i+1})$$
$$F^L(u_{i-1}, u_i) = \frac{1}{2}\big(f(u_{i-1}) + f(u_i)\big) + \frac{\Delta t}{2h_i}(u_{i-1} - u_i).$$

Now for the consistency criteria,

1. Obviously, $F^R(u_i, u_i) = F^L(u_i, u_i) = f(u_i)$ and
2. Now for the second criterion,

$$\frac{2h_i}{h_i + h_{i+1}} F_1^L(u_i^n, u_i^n) + \frac{2h_{i+1}}{h_i + h_{i+1}} F_2^R(u_i^n, u_i^n)$$

or

$$\frac{2h_i}{h_i + h_{i+1}} \left(\frac{1}{2} f'(u_i) + \frac{\Delta t}{2h_i}\right) + \frac{2h_{i+1}}{h_i + h_{i+1}} \left(\frac{1}{2} f'(u_i) - \frac{\Delta t}{2h_{i+1}}\right)$$

or

$$\left(\frac{h_i}{h_i + h_{i+1}} + \frac{h_{i+1}}{h_i + h_{i+1}}\right) f'(u_i) + \frac{\Delta t}{h_i + h_{i+1}} - \frac{\Delta t}{h_i + h_{i+1}},$$

which yields

$$f'(u_i).$$

So both consistency criteria are satisfied, hence the scheme is consistent with the underlying conservation law, $u_t + f(u)_x = 0$. This confirms the modified equation analysis we performed in the respective scheme for the linear flux $f(u) = u$.

**Generalised LxF - Vertex centered** The numerical flux for this conservative scheme is

$$F^R(u_i, u_{i+1}) = \frac{1}{2}\big(f(u_i) + f(u_{i+1})\big) - \frac{h_{i+1}}{2\Delta t}(u_{i+1} - u_i)$$
$$F^L(u_{i-1}, u_i) = \frac{1}{2}\big(f(u_{i-1}) + f(u_i)\big) - \frac{h_i}{2\Delta t}(u_i - u_{i-1}).$$

Now for the consistency criteria,

1. Obviously, $F^R(u_i, u_i) = F^L(u_i, u_i) = f(u_i)$ and
2. Now for the second criterion,

$$\frac{2h_i}{h_i + h_{i+1}} F_1^L(u_i^n, u_i^n) + \frac{2h_{i+1}}{h_i + h_{i+1}} F_2^R(u_i^n, u_i^n)$$

or

$$\frac{2h_i}{h_i + h_{i+1}} \left( \frac{1}{2} f'(u_i) + \frac{h_i}{2\Delta t} \right) + \frac{2h_{i+1}}{h_i + h_{i+1}} \left( \frac{1}{2} f'(u_i) - \frac{h_{i+1}}{2\Delta t} \right)$$

or

$$\left( \frac{h_i}{h_i + h_{i+1}} + \frac{h_{i+1}}{h_i + h_{i+1}} \right) f'(u_i) + \frac{h_i^2}{\Delta t(h_i + h_{i+1})} - \frac{h_{i+1}^2}{\Delta t(h_i + h_{i+1})},$$

which yields

$$f'(u_i) + \frac{h_i - h_{i+1}}{\Delta t}.$$

So the second consistency criterion failed, hence the scheme is not consistent with the underlying conservation law, $u_t + f(u)_x = 0$. This confirms the modified equation analysis we performed in the respective scheme for the linear flux $f(u) = u$.

*Remark* 3.3.4. We note that the usual consistency flux criterion (now on valid only for uniform meshes) is satisfied since $F_{i+1/2}(u_i, u_i) = f(u_i)$.

**Unstable centered, FTCS - Vertex centered** The numerical flux for this conservative scheme is

$$F^R(u_i, u_{i+1}) = \frac{f(u_i) + f(u_{i+1})}{2}$$
$$F^L(u_{i-1}, u_i) = \frac{f(u_{i-1}) + f(u_i)}{2}$$

Now for the consistency criteria,

1. Obviously, $F^R(u_i, u_i) = F^L(u_i, u_i) = f(u_i)$ and
2. Now for the second criterion,

$$\frac{2h_i}{h_i + h_{i+1}} F_1^L(u_i^n, u_i^n) + \frac{2h_{i+1}}{h_i + h_{i+1}} F_2^R(u_i^n, u_i^n)$$

or

$$\frac{2h_i}{h_i + h_{i+1}} \left( \frac{1}{2} f'(u_i) \right) + \frac{2h_{i+1}}{h_i + h_{i+1}} \left( \frac{1}{2} f'(u_i) \right)$$

or

$$\left(\frac{h_i}{h_i + h_{i+1}} + \frac{h_{i+1}}{h_i + h_{i+1}}\right) f'(u_i),$$

which yields

$$f'(u_i).$$

So both consistency criteria are satisfied, hence the scheme is consistent with the underlying conservation law, $u_t + f(u)_x = 0$. This confirms the modified equation analysis we performed in the respective scheme for the linear flux $f(u) = u$.

#### 3.3.4.2 Cell centered grids

In this approach we consider the non-uniform discretization of the domain in cells

$$C_i = (x_{i-1/2}, x_{i+1/2}) \quad \text{with} \quad |C_i| = h_i.$$

The mesh $M_x = \{x_i, i \in \mathbb{Z}\}$ consists of the middle points

$$x_i = \frac{x_{i+1/2} + x_{i-1/2}}{2}; \quad \text{hence} \quad x_i - x_{i-1} = \frac{h_i + h_{i-1}}{2}.$$

For this description, the schemes attain the following conservative form

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{h_i}\left(F^R_{i+1/2} - F^L_{i-1/2}\right).$$

We restate the consistency requirements for cell centered grid configuration as established in Th.(3.3.1)

1. $F^R(u_i^n, u_i^n) = F^L(u_i^n, u_i^n)$
2. $\frac{h_{i-1}+h_i}{2h_i} F_1^L(u_i^n, u_i^n) + \frac{h_i+h_{i+1}}{2h_i} F_2^R(u_i^n, u_i^n) = f'(u_i)$.

**Generalised LxF - Cell centered** The numerical flux for this conservative scheme is

$$F^R(u_i, u_{i+1}) = -\frac{1}{2\Delta t}(h_{i+1}u_{i+1} - h_i u_i) + \frac{1}{2}\big(f(u_{i+1}) + f(u_i)\big)$$
$$F^L(u_{i-1}, u_i) = -\frac{1}{2\Delta t}(h_i u_i - h_{i-1}u_{i-1}) + \frac{1}{2}\big(f(u_i) + f(u_{i-1})\big).$$

Now for the consistency criteria,

1. Obviously, $F^R(u_i,u_i) = -\frac{h_{i+1}-h_i}{2\Delta t}u_i + f(u_i)$ and $F^L(u_i,u_i) = -\frac{h_i-h_{i-1}}{2\Delta t}u_i + f(u_i)$.

So the first consistency criterion failed; hence the scheme is not consistent with the underlying conservation law, $u_t+f(u)_x = 0$. This confirms the modified equation analysis we performed in the respective scheme for the linear flux $f(u) = u$.

*Remark* 3.3.5. We note that the usual consistency flux criterion (valid only for uniform meshes) also fails since $F_{i+1/2}(u_i,u_i) = -\frac{h_{i+1}-h_i}{2\Delta t}u_i + f(u_i)$.

**Another first order - Cell centered** The numerical flux for this conservative scheme is

$$F^R(u_i,u_{i+1}) = \frac{h_{i+1}}{h_i+h_{i+1}}f_i + \frac{h_i}{h_i+h_{i+1}}f_{i+1} - \frac{2\Delta t}{h_i+h_{i+1}}(u_{i+1}-u_i)$$
$$F^L(u_{i-1},u_i) = \frac{h_i}{h_{i-1}+h_i}f_{i-1} + \frac{h_{i-1}}{h_{i-1}+h_i}f_i - \frac{2\Delta t}{h_{i-1}+h_i}(u_i-u_{i-1}).$$

Now for the consistency criteria,

1. Obviously, $F^R(u_i,u_i) = F^L(u_i,u_i) = f(u_i)$.

2. Now for the second criterion,

$$\frac{h_{i-1}+h_i}{2h_i}F_1^L(u_i^n,u_i^n) + \frac{h_i+h_{i+1}}{2h_i}F_2^R(u_i^n,u_i^n)$$

or

$$\frac{h_{i-1}+h_i}{2h_i}\left(\frac{h_i}{h_{i-1}+h_i}f'(u_i) + \frac{2\Delta t}{h_{i-1}+h_i}\right) + \frac{h_i+h_{i+1}}{2h_i}\left(\frac{h_i}{h_i+h_{i+1}}f'(u_i) - \frac{2\Delta t}{h_i+h_{i+1}}\right)$$

or

$$\frac{1}{2}f'(u_i) + \frac{\Delta t}{h_i} + \frac{1}{2}f'(u_i) - \frac{\Delta t}{h_i},$$

which yields

$$f'(u_i).$$

So both consistency criteria are satisfied, hence the scheme is consistent with the underlying conservation law, $u_t + f(u)_x = 0$. This confirms the modified equation analysis we performed in the respective scheme for the linear flux $f(u) = u$.

**Richtmyer 2-step Lax-Wendroff - Cell centered** The numerical flux for this conservative scheme is

$$F^R(u_i,u_{i+1}) = f\Big(\frac{h_{i+1}}{h_i+h_{i+1}}u_i + \frac{h_i}{h_i+h_{i+1}}u_{i+1} - \frac{\Delta t}{h_i+h_{i+1}}(f_{i+1}-f_i)\Big)$$
$$F^L(u_{i-1},u_i) = f\Big(\frac{h_i}{h_{i-1}+h_i}u_{i-1} + \frac{h_{i-1}}{h_{i-1}+h_i}u_i - \frac{\Delta t}{h_{i-1}+h_i}(f_i-f_{i-1})\Big).$$

Now for the consistency criteria,

1. Obviously, $F^R(u_i,u_i) = F^L(u_i,u_i) = f(u_i)$.

2. Now for the second criterion,

$$\frac{h_{i-1}+h_i}{2h_i}F_1^L(u_i^n,u_i^n) + \frac{h_i+h_{i+1}}{2h_i}F_2^R(u_i^n,u_i^n)$$

or

$$\frac{h_{i-1}+h_i}{2h_i}\left(\frac{h_i}{h_{i-1}+h_i} + \frac{\Delta t f'(u_i)}{h_{i-1}+h_i}\right)f'(u_i) + \frac{h_i+h_{i+1}}{2h_i}\left(\frac{h_i}{h_i+h_{i+1}} - \frac{\Delta t f'(u_i)}{h_i+h_{i+1}}\right)f'(u_i)$$

or

$$\left(\frac{1}{2} + \frac{\Delta t f'(u_i)}{2h_i}\right)f'(u_i) + \left(\frac{1}{2} - \frac{\Delta t f'(u_i)}{2h_i}\right)f'(u_i),$$

which yields

$$f'(u_i).$$

So both consistency criteria are satisfied, hence the scheme is consistent with the underlying conservation law, $u_t + f(u)_x = 0$. This confirms the modified equation analysis we performed in the respective scheme for the linear flux $f(u) = u$.

**MacCormack - Cell centered** The numerical flux for this conservative scheme is

$$F^R(u_i,u_{i+1}) = \frac{f(u_{i+1}) + f\left(u_i - \frac{\Delta t}{h_i}(f(u_{i+1}) - f(u_i))\right)}{2}$$
$$F^L(u_{i-1},u_i) = \frac{f(u_i) + f\left(u_{i-1} - \frac{\Delta t}{h_{i-1}}(f(u_i) - f(u_{i-1}))\right)}{2}.$$

Now for the consistency criteria,

1. Obviously, $F^R(u_i,u_i) = F^L(u_i,u_i) = f(u_i)$.

2. Now for the second criterion,

$$\frac{h_{i-1}+h_i}{2h_i}F_1^L(u_i^n,u_i^n)+\frac{h_i+h_{i+1}}{2h_i}F_2^R(u_i^n,u_i^n)$$

or

$$\frac{h_{i-1}+h_i}{2h_i}\left(1+\frac{\Delta t}{h_{i-1}}f'(u_i)\right)\frac{1}{2}f'(u_i)+\frac{h_i+h_{i+1}}{2h_i}\left(1-\frac{\Delta t}{h_i}f'(u_i)\right)\frac{1}{2}f'(u_i),$$

which yields, after some algebra

$$\left(h_{i-1}+2h_i+h_{i+1}+\Delta t f'(u_i)\left(\frac{h_i}{h_{i-1}}-\frac{h_{i+1}}{h_i}\right)\right)\frac{f'(u_i)}{4h_i}.$$

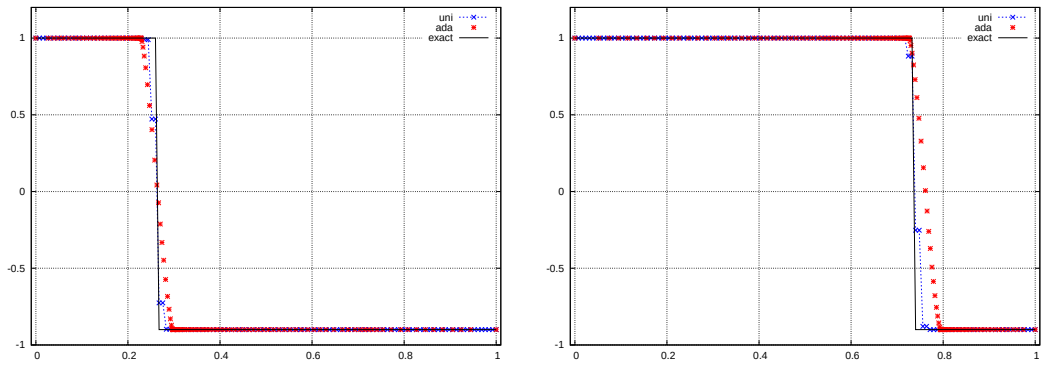
So the second consistency criterion failed, hence the scheme is not consistent with the underlying conservation law, $u_t+f(u)_x=0$. This confirms the modified equation analysis we performed in the respective scheme for the linear flux $f(u)=u$.

*Remark* 3.3.6. We note that the usual consistency flux criterion (valid only for uniform meshes) is satisfied since $F_{i+1/2}(u_i,u_i)=f(u_i)$.

## 3.4 Numerical tests

In this section we present numerical tests on some of the previously described and analysed schemes.

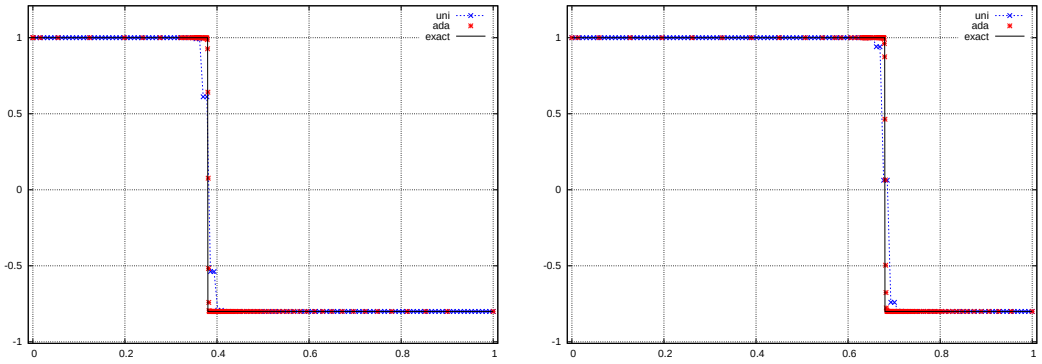
**LxF-Approach 1**




LxF - Approach 1 on non-uniform Finite Element grid with conservative reconstruction and LxF on uniform grid. The problem is a Burgers equation with a single shock. The non-uniform mesh scheme is not conservative which explains the wrong propagation

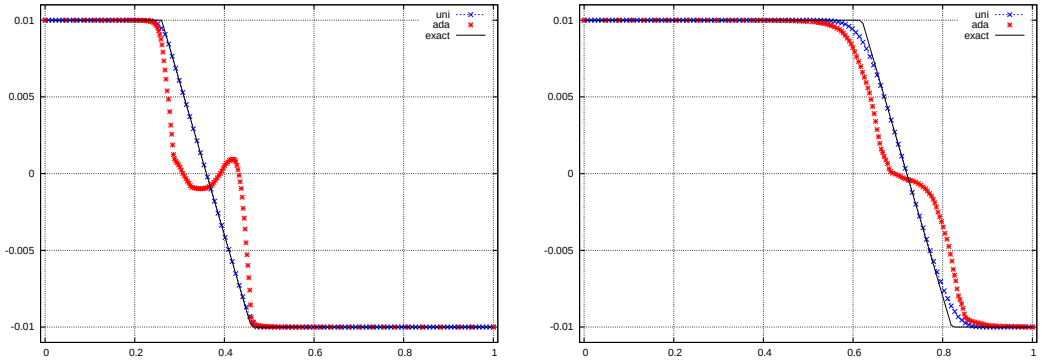
speed of the shock. Left graph $t = 1.2$, right graph $t = 9.6$, $N = 128$ number of mesh points with $CFL = 0.9$ and $pw = 0.11$

**Generalised LxF - Vertex centered**




Generalised LxF - Vertex Centered (non-uniform with conservative reconstruction) vs LxF (uniform)

The first problem is a Burgers equation with a single shock. Left graph $t = 1.8$, right graph $t = 4.8$, $N = 128$ number of points with $CFL = 0.9$, $pw = 0.11$. The non-uniform case resolves the shock much better than the uniform one.



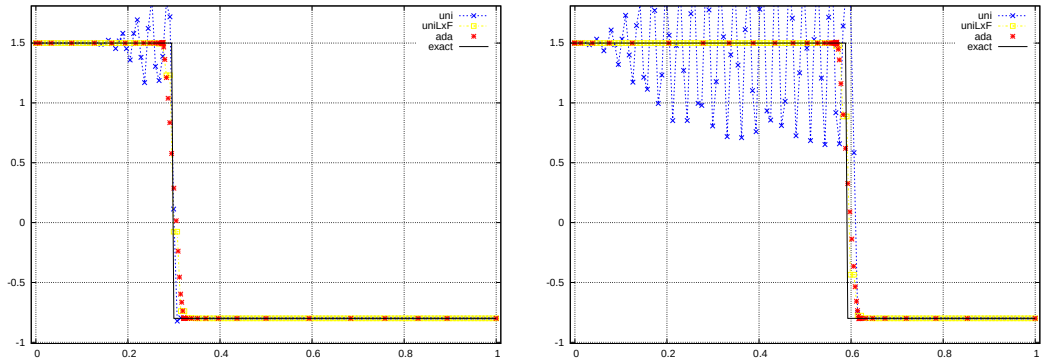

Generalised LxF - Vertex Centered (non-uniform with conservative reconstruction) vs LxF (uniform).

The second problem is a Linear Transport with a slow slope. Left graph $t = 0.06$, right graph $t = 0.42$, $N = 128$, $CFL = 0.9$, $pw = 0.09$. The non-uniform does not perform well on this problem. This is because this scheme is not consistent (in the non-uniform case). The modified equation for this scheme reads,

$$u_t + u_x = \frac{h_{i+1} - h_i}{\Delta t} u_x + \frac{(h_{i+1} - h_i)(\Delta t - h_{i+1}) + (h_i - \Delta t)(h_i + \Delta t)}{2\Delta t} u_{xx}$$

and we note the coefficient $\frac{h_{i+1}-h_i}{\Delta t}$ of $u_x$ in the right side. The inconsistency is visible in the Transport case since in the Burgers equation the intervals $h_{i+1}$, $h_i$ are compressed more in areas where $u_x$ has a significant value.

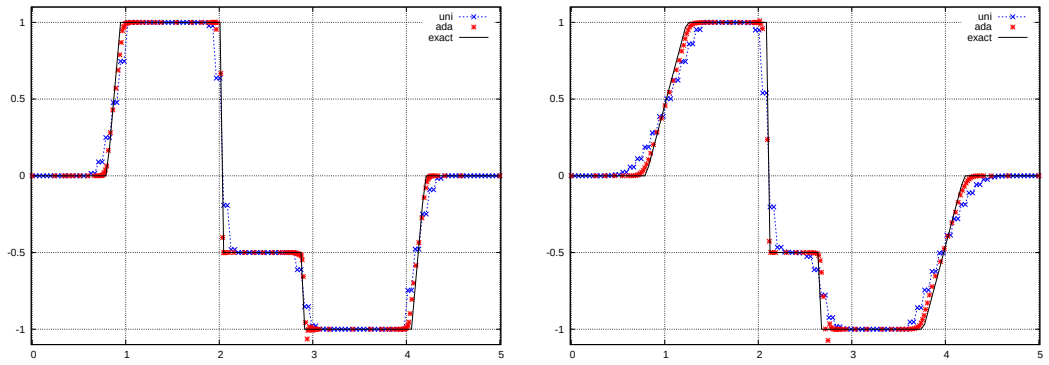
### First Order - Vertex Centered




First Order - Vertex centered (non-uniform with conservative reconstruction) vs First Order (uniform grid).
This problems is a Burgers Equation with a single shock. Left graph $t = 0.28$, right graph $t = 1.12$, $N = 128$, $CFL = 0.9$, $pw = 0.11$. The modified equation for the Transport problem $f(u) = au$ reads,

$$\begin{aligned} u_t + au_x =& \frac{a(h_i - h_{i+1}) + \Delta t(a^2 - 1)}{2} u_{xx} \\ & - \frac{ah_{i+1}^2 + (-3a + 2a^3)\Delta t^2 + (h_i - h_{i+1})(-3a^2\Delta t + ah_i + \Delta t)}{6} u_{xxx} \end{aligned}$$

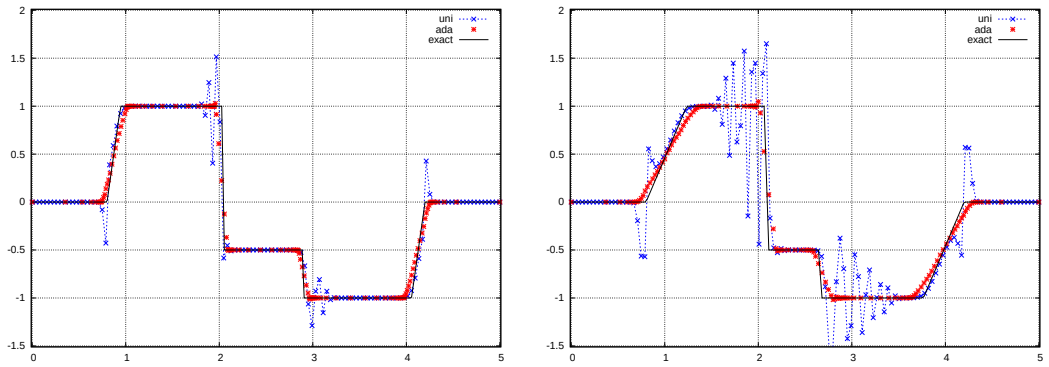
Even though both cases -uniform and non-uniform- are not stable for local speeds $a > 1$ the non-uniform case performs very well by suppressing the oscillations. On the other hand the uniform case exhibits wild oscillations. This is a manifestation of the stabilisation property of the BAS.




First Order - Vertex centered (non-uniform with conservative reconstruction) vs classic LxF (uniform).

This problem is a Burgers equation with a combination of 2 shock and 2 rarefactions. Left graph $t = 0.145$, right graph $t = 0.435$, $N = 128$, $CFL = 0.9$, $pw = 0.09$. Even though the non-uniform scheme is not stable it out-performs the uniform LxF by resolving the solution much better.

**Unstable centered, FTCS - Vertex Centered**




Unstable centered, FTCS - Vertex centered (non-uniform with conservative reconstruction) vs FTCS (uniform).

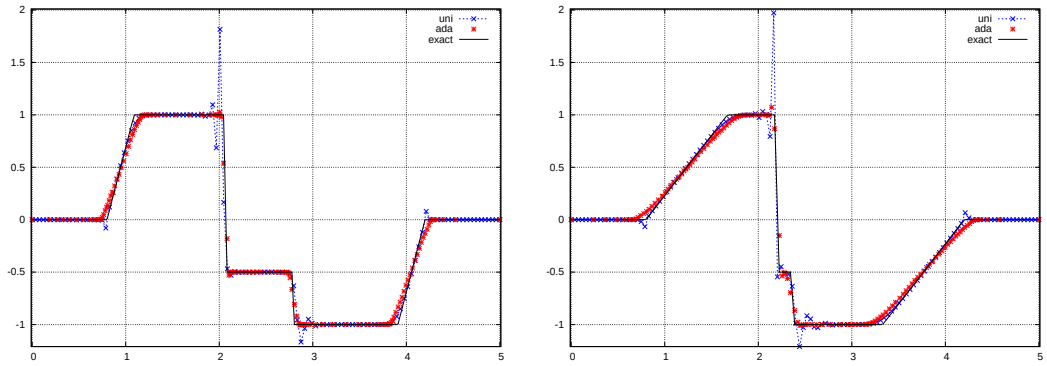
This problem is a Burgers equation with a combination of 2 shocks and 2 rarefactions. Left graph $t = 0.145$, right graph $t = 0.435$, $N = 128$, $CFL = 0.5$, $pw = 0.09$. The modified equation of this scheme for the case of a linear flux $f(u) = au$ reads,

$$u_t + au_x = a\frac{h_i - h_{i+1} - a\Delta t}{2}u_{xx} + \frac{a(h_i - 6a\Delta t)(h_{i+1} - h_i) - 2a^3\Delta t^2 - ah_{i+1}^2}{6}u_{xxx}$$

Even though both cases -uniform and non-uniform- are not not stable the non-uniform FTCS performs very well by taming the oscillations. Again the stabilisation properties

of the BAS are obvious.

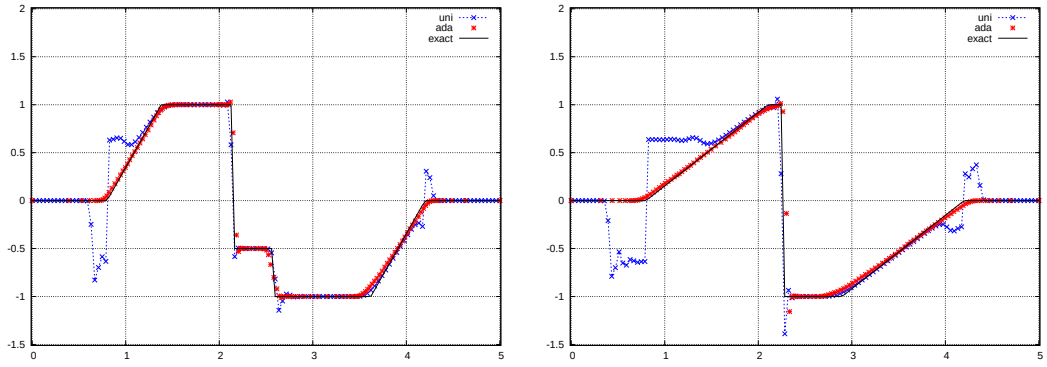
**Richtmyer 2-step LxW - Cell Centered**




Richtmyer 2-step LxW - Cell centered (non-uniform with conservative reconstruction) vs Richtmyer (uniform).

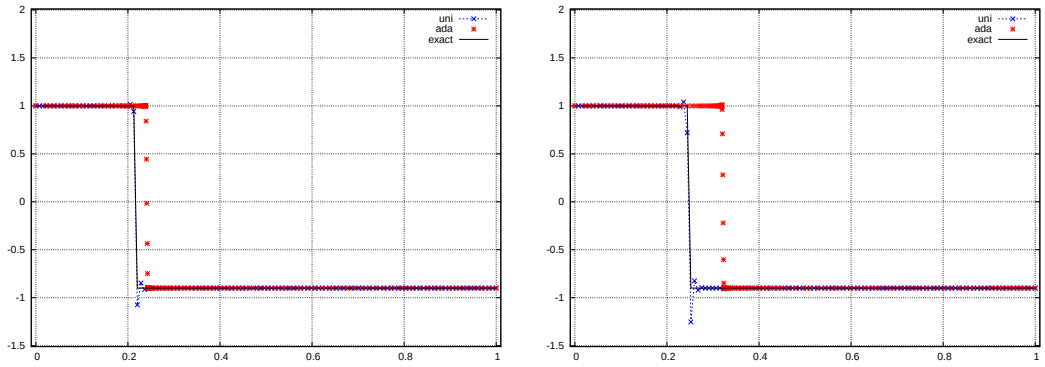
This problem is a Burgers equation with a combination of 2 shocks and 2 rarefactions. Left graph $t = 0.29$, right graph $t = 0.87$, $N = 128$, $CFL = 0.9$, $pw = 0.099$. The uniform Richtmyer produces oscillations due to its dispersive nature, on the non-uniform case the oscillations are suppressed. Once again the stabilisation properties of the BAS are depicted.

**MacCormack - Vertex Centered**



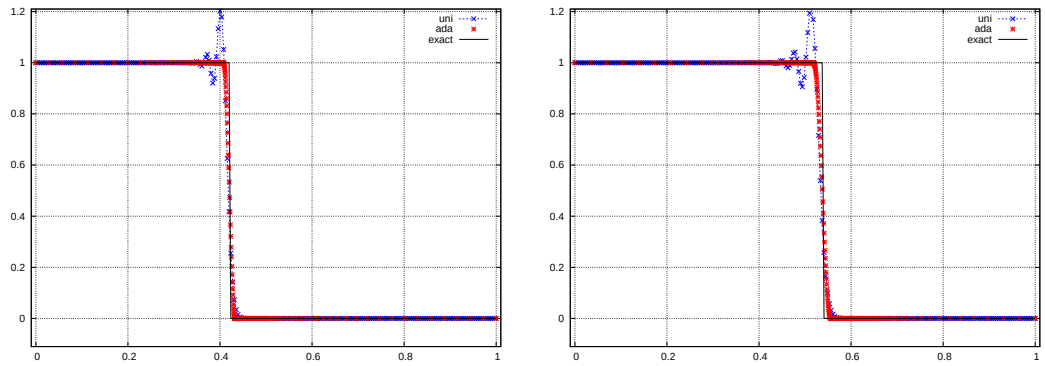

MacCormack - Vertex centered (non-uniform with conservative reconstruction) vs MacCormack (uniform).

This problem is a Burgers equation with a combination of 2 shocks and 2 rarefactions. Left graph $t = 0.58$, right graph $t = 1.305$, $N = 128$, $CFL = 0.9$, $pw = 0.07$. The uniform MacCormack fails to produce the correct rarefactions, the non-uniform though performs very well exhibiting once again the stabilisation properties of the BAS.




MacCormack - Vertex centered (non-uniform with conservative reconstruction) vs MacCormack (uniform).
This problem is a Burgers equation with a single shock. Left graph $t = 0.3$, right graph $t = 0.9$, $N = 128$, $CFL = 0.9$, $pw = 0.07$. The non-uniform case fails on the speed of propagation of the shock, this is due to the lack of conservation of the numerical scheme. This example shows that the BAS -even with conservative reconstruction- is not able to correct conservation problems of the numerical schemes.

**Generalized LxW - Pure second order**




Generalised LxW - Pure second order (non-uniform with conservative reconstruction) vs LxW (uniform).
The problem is a Transport equation with a single shock. Left graph $t = 0.117$, right graph $t = 0.234$, $N = 256$, $CFL = 0.5$, $pw = 0.06$.
The $CFL = 0.5$ was chosen to exhibit the oscillations produced by the uniform LxW schemes. With the same $CFL$ the BAS tames the oscillations, this is another manifestation of the stabilisation property of the BAS.

**Summary** Summarising on the numerical test section we can conclude that the BAS :

- Is not able to correct neither inconsistency problems nor conservation problems of the numerical scheme

- It is able to correct instability problems caused either by the possible anti-diffusive or the dispersive nature of the scheme. The stabilisation property of the BAS is investigated in the following Chapter.

## 3.5 Entropy conservative schemes on non-uniform meshes

This section is part of a joint work with Ch. Arvanitis and Ch. Makridakis [4]

The *Entropy Conservative Schemes* were first introduced by Tadmor [26], [27] and further studied by Lefloch and Rhode [19]. They are semi-discrete schemes that satisfy an exact entropy equality instead of the usual inequality. These schemes are interesting on their own right for they appear in the context of zero dispersion limits, complete integrable systems and computation of non-classical shocks. They are also important as a building block for the construction of Entropy Stable schemes [26]

These schemes, when explicit time discretisation is chosen, become entropy unstable and produce oscillations [27]. Our purpose in this section is to examine the effect the BAS has when combined with (explicitly discretised in time) Entropy Conservative 3-point and 5-point schemes.

To start with, we consider the Conservation Law

$$u_t + f(u)_x = 0, \quad (x,t) \in \mathbb{R} \times [0, +\infty),$$

where $f$ is a smooth flux function. We also consider a semi-discrete, consistent with the conservation law and conservative scheme

$$\frac{d}{dt}u_i(t) = -\frac{1}{\Delta x_i}\left(g_{i+\frac{1}{2}} - g_{i-\frac{1}{2}}\right),$$

where $u_i(t)$ denotes the discrete solution along the line $(x_i, t)$, $\Delta x_i = \frac{1}{2}(x_{i+1} - x_{i-1})$ is the variable mesh step in a Vertex centered description of the grid and $g_{i+\frac{1}{2}}$ a Lipschitz continuous numerical flux consistent with the differential flux $f$.

### 3.5.1 Semi-Discrete Entropy Conservative Scheme

This paragraph is a summary of the one dimensional part of entropy schemes of [26] and [27] and discusses the construction of the Entropy Conservative schemes. We will restrict our attention to the construction of 3-point second order accurate schemes.

The construction starts by assuming that the one dimensional conservation law

$$u_t + f(u)_x = 0$$

is equipped with a convex Entropy function $U(u)$ along with an Entropy Flux function $F$, which satisfies

$$F'(u) = U'(u)f'(u).$$

The Entropy function $U$ provides the new variables - the Entropy Variables,

$$v(u) = U'(u),$$

and due to the convexity of $U$ the mapping $u(v)$ is 1-1 and serves as a change of variables $u = u(v)$. Hence the initial conservation law yields

$$\frac{\partial}{\partial_t}u(v) + \frac{\partial}{\partial_x}g(v) = 0, \quad g(v) = f\big(u(v)\big).$$

The potential functions that follow play an essential role in the construction and analysis of the Entropy conservative Schemes, namely the Entropy Potential, $\phi(v)$ is defined as

$$u(v) = \frac{d}{dv}\phi(v) \Longrightarrow \phi(v) = vu(v) - U\big(u(v)\big), \tag{3.20}$$

and the Entropy Flux Potential $\psi(v)$ is defined as

$$g(v) = \frac{d}{dv}\psi(v) \Longrightarrow \psi(v) = vg(v) - F\big(u(v)\big). \tag{3.21}$$

We can now introduce the Entropy Conservative Numerical Flux

$$g_{i+\frac{1}{2}} = \int_{\xi=0}^{1} g\big(v_i + \xi(v_{i+1} - v_i)\big)d\xi = \frac{\psi(v_{i+1}) - \psi(v_i)}{v_{i+1} - v_i},$$

which provides the Entropy Conservative Centered Semi-Discrete numerical scheme:

$$\frac{d}{dt}u_i(t) = -\frac{1}{\Delta x_i}\left(g_{i+\frac{1}{2}} - g_{i-\frac{1}{2}}\right).$$

It is proven in [26] Theorem 4.1 that this scheme is in fact entropy conservative.

As very elegantly described by Tadmor [26] conservative schemes with more numerical viscosity than the Entropy Conservative ones are Entropy Stable (Theorem 5.2). Moreover conservative schemes containing more viscosity than an Entropy Stable scheme are also Entropy Stable (Theorem 5.3). So there is a type of ordering in the class of Entropy Conservative/Stable schemes with the Entropy conservative ones as marginal cases.

### 3.5.2 Time discretisation and numerical tests

The Entropy Conservative Schemes of the previous paragraph are semi-discrete. We refer to Tadmor [27] to note that implicit time discretization enforces Entropy stability, on the contrary explicit time discretization leads to entropy production.

So, in the case of explicit time discretization a balance has to be kept between the temporal entropy production and the spatial entropy dissipation. The Entropy Conservative schemes though, do not lead to spatial entropy dissipation. Hence by combining the BAS with an Entropy Conservative scheme for the evolution step we can test the entropy dissipation properties of the Mesh Reconstruction (Step 1) and the Solution Update (Step 2) of the BAS. As stated previously in the uniform mesh case spurious oscillations are produced, on the other hand the non-uniform adaptive mesh case is clean of oscillations, exhibiting that enough entropy is dissipated by the AMR step of the BAS.

We shall provide 4 examples where the entropy dissipation of the BAS is exhibited. The first 3 are 3-point entropy conservative schemes. We follow the work of Tadmor [26] and [27] up to the point of constructing a semi-discrete entropy conservative scheme and then we discretise explicitly in time.

The final example is a 5-point entropy conservative semi-discrete scheme emanating from the work of LeFloch and Rhode [19] and having the property of producing non classical shocks. The time discretization in this example is a 4th order Runge-Kutta. This example is more interesting since the oscillations that are generated (in the uniform mesh) at the shock position and progress with a pattern.

In all the examples that we studied we noticed that the adaptive mesh selection eliminates the oscillations, with minimal CFL condition. In the more interesting examples

we exhibit the numerical tests using both uniform and adaptive mesh selection. For the rest we just present the adaptive mesh case, since it is well known that the relevant uniform mesh schemes produce large oscillation rather fast.

#### 3.5.2.1 Problem 1

First example is the Inviscid Burgers equation equiped with entropy function $U(u) = -\ln u$. We follow the steps stated in the introduction up to the point of constructing the relevant semi-discrete entropy conservative scheme. We then select the temporal discretization and create the fully-discrete entropy conservative scheme. The work cited in Tadmor [26] and [27] is used as was presented previously.

The model is the Inviscid Burgers equation,

$$u_t + \Big(\frac{1}{2}u^2\Big)_x = 0$$

In this example we want to conserve the entropy $U(u) = -\ln u$ along with the entropy flux $F$ defined via,

$$F'(u) = U'(u)f'(u) \Rightarrow F(u) = -u.$$

By the convexity of the entropy function $U(u)$ we perform the following change of variables

$$v(u) = U'(u) \Rightarrow v(u) = -\frac{1}{u} \Rightarrow u(v) = -\frac{1}{v},$$

the entropy variable flux reads

$$g(v) = f\big(u(v)\big) \Rightarrow g(v) = -\frac{1}{2v^2},$$

and the model equation recasts,

$$\frac{\partial}{\partial_t}u(v) + \frac{\partial}{\partial_x}g(v) = 0.$$

We shall also use the entropy flux potential $\psi(v)$,

$$\psi(v) = vg(v) - F\big(u(v)\big) \Rightarrow \psi(v) = -\frac{1}{2v},$$

and the entropy conservative flux shall be,

$$g_{i+\frac{1}{2}} = \frac{\psi(v_{i+1}) - \psi(v_i)}{v_{i+1} - v_i} \Rightarrow g_{i+\frac{1}{2}} = \frac{1}{2}\frac{1}{v_{i+1}v_i} = \frac{1}{2}u_{i+1}u_i.$$

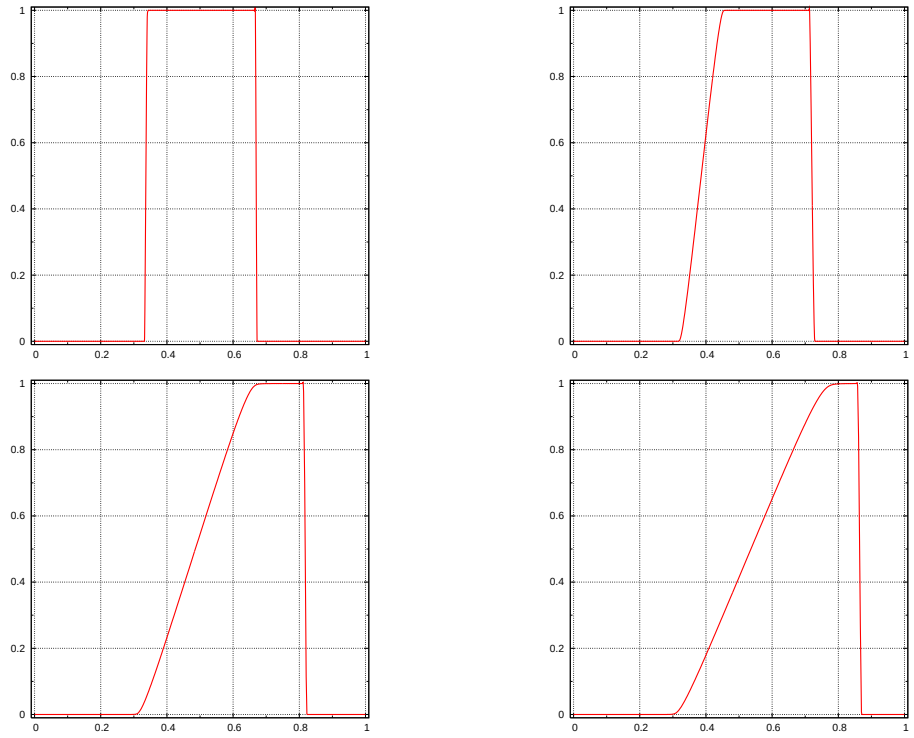


Figure 3.3: Problem 1: Burgers Equation and entropy $U(u) = -\ln u$ and box initial conditions. Exhibiting time steps, $t = 0$, $t = 0.10$ $t = 0.30$ $t = 0.40$. Showing only the adaptive case since the uniform exhibits large scale oscillations. 250 nodes and CFL 0.9

We employ the last result to write the *Entropy Conservative Centered Semi-Discrete numerical scheme*,

$$\frac{d}{dt}u_i(t) = -\frac{1}{\Delta x_i}(g_{i+\frac{1}{2}} - g_{i-\frac{1}{2}}) \Longrightarrow \frac{d}{dt}u_i(t) = -u_i(t)\frac{u_{i+1}(t) - u_{i-1}(t)}{x_{i+1} - x_{i-1}}$$

Finally we discretise explicitly in time and the fully discrete scheme is:

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{x_{i+1} - x_{i-1}}u_i^n(u_{i+1}^n - u_{i-1}^n)$$

We simply note that the grid is in a Vertex centered configuration and we refer to Figure (3.3) for a graphical presentation of this scheme in the non-uniform adaptive case.

#### 3.5.2.2 Problem 2

For the second example we use the same method, 3-point entropy conservative scheme with adaptive mesh selection and explicit time discretization. The convex entropy in this example is $U(u) = e^u$. The work cited in [26] and [27] is again used as was previously described.

In this example the model equation is,

$$u_t + (e^u)_x = 0,$$

The entropy that we want conserved is $U(u) = e^u$ with entropy flux

$$F'(u) = U'(u)f'(u) \Rightarrow F(u) = \frac{1}{2}e^{2u}.$$

Due to the convexity of the entropy $U$ we perform the following change of variables,

$$v(u) = U'(u) \Rightarrow v(u) = e^u \Rightarrow u(v) = \ln v$$

Hence the entropy variables flux reads

$$g(v) = f\big(u(v)\big) \Rightarrow g(v) = v,$$

which provides us with the new model equation,

$$\frac{\partial}{\partial_t}u(v) + \frac{\partial}{\partial_x}g(v) = 0.$$

We define the entropy flux potential $\psi(v)$ via,

$$\psi(v) = vg(v) - F\big(u(v)\big) \Rightarrow \psi(v) = \frac{1}{2}v^2.$$

So the entropy conservative flux is written,

$$g_{i+\frac{1}{2}} = \frac{\psi(v_{i+1}) - \psi(v_i)}{v_{i+1} - v_i} \Rightarrow g_{i+\frac{1}{2}} = \frac{1}{2}(e^{u_{i+1}} + e^{u_i}),$$

and the *Entropy Conservative Centered Semi-Discrete numerical scheme* reads,

$$\frac{d}{dt}u_i(t) = -\frac{1}{\Delta x_i}(g_{i+\frac{1}{2}} - g_{i-\frac{1}{2}}) \Rightarrow \frac{d}{dt}u_i(t) = -\frac{e^{u_{i+1}(t)} - e^{u_{i-1}(t)}}{x_{i+1} - x_{i-1}}.$$

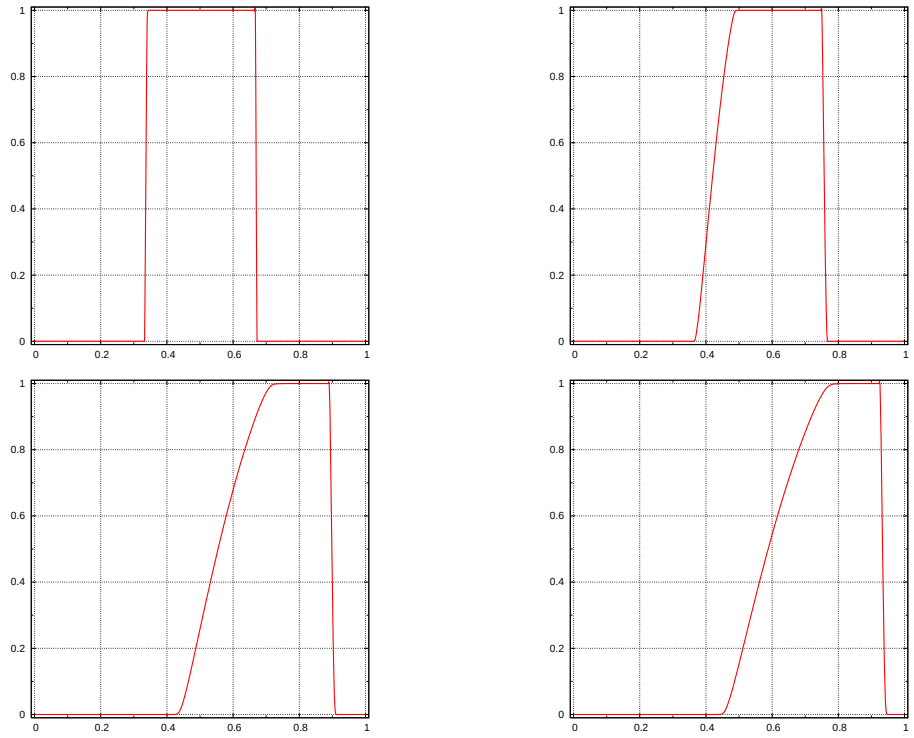


Figure 3.4: Problem 2: $u_t + (e^u)_x = 0$ and entropy $U(u) = e^u$ with box initial conditions. Exhibiting time steps $t = 0$, $t = 0.05$ $t = 0.13$ $t = 0.15$ only for the adaptive case since the uniform exhibits large scale oscillations.

Finally we discretise explicitly in time to get the fully discrete scheme,

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{x_{i+1} - x_{i-1}} \cdot (e^{u_{i+1}^n} - e^{u_{i-1}^n})$$

We refer to Figure (3.4) for a graphical presentation of this scheme in the non-uniform adaptive case.

#### 3.5.2.3 Problem 3

Again a 3-point entropy conservative scheme. This example is again Burgers equation but this time the entropy is $U(u) = \int^u f(s)ds$, which is convex due to the sign of $u$, $u \geq 0$. The choice of the specific entropy function leads to entropy variables flux $g(v) = v$. The work cited in [26] and [27] is again used as was previously described.

The model we use is the inviscid Burgers equation:

$$u_t + \Big(\frac{1}{2}u^2\Big)_x = 0,$$

The entropy we want to conserve in this case is

$$U(u) = \int^u f(s)ds \Rightarrow U(u) = \frac{1}{6}u^3,$$

with entropy flux defined via,

$$F'(u) = U'(u)f'(u) \Rightarrow F(u) = \frac{1}{8}u^4.$$

The convexity of $U$ provides us with the new variables $v$,

$$v(u) = U'(u) \Rightarrow v(u) = \frac{1}{2}u^2 \Rightarrow u(v) = \sqrt{2v}$$

The entropy variables flux is given by,

$$g(v) = f\big(u(v)\big) \Rightarrow g(v) = v,$$

and the new model equation is written as,

$$\frac{\partial}{\partial_t}u(v) + \frac{\partial}{\partial_x}g(v) = 0.$$

We move on to the entropy flux potential $\psi(v)$,

$$\psi(v) = vg(v) - F\big(u(v)\big) \Rightarrow \psi(v) = \frac{1}{2}v^2,$$

which provides us with the entropy conservative flux,

$$g_{i+\frac{1}{2}} = \frac{\psi(v_{i+1}) - \psi(v_i)}{v_{i+1} - v_i} \Rightarrow g_{i+\frac{1}{2}} = \frac{1}{4}(u_{i+1}^2 + u_i^2).$$

Hence the *Entropy Conservative Centered Semi-Discrete numerical scheme* reads

$$\frac{d}{dt}u_i(t) = -\frac{1}{\Delta x_i}(g_{i+\frac{1}{2}} - g_{i-\frac{1}{2}}) \Rightarrow \frac{d}{dt}u_i(t) = -\frac{1}{4}\frac{u_{i+1}^2 - u_{i-1}^2}{x_{i+1} - x_{i-1}}.$$

To fully discretize the previous equation we once again use Forward Euler time discretiza-

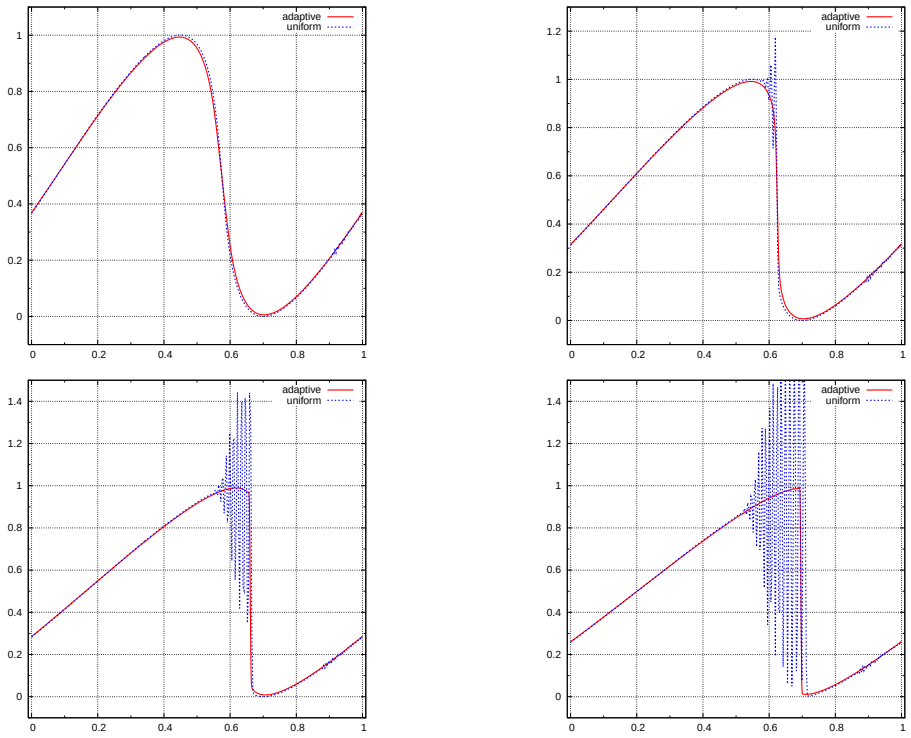



Figure 3.5: Problem 3: Burgers and entropy $U(u) = \frac{1}{6}u^3$ and $u_0(x) = 0.5\sin(2\pi(x + 0.05)) + 0.5$ initial conditions. $t = 0.5$, $t = 0.7$ $t = 0.85$ $t = 1.0$. 300 nodes 4000 time steps

tion and the fully discrete numerical schemes assumes the form,

$$u_i^{n+1} = u_i^n - 0.25\frac{\Delta t}{x_{i+1} - x_{i-1}}(u_{i+1} - u_{i-1})(u_{i+1} + u_{i-1}).$$

We refer to Figure (3.5) for a graphical presentation of this scheme in the non-uniform adaptive case.

#### 3.5.2.4 Problem 4

The previous examples were 3-point entropy conservative centered schemes based mainly in the work cited in [26], [27]. We now implement a 5-point entropy conservative scheme cited in [19].

We consider the one dimensional model:

$$u_t + (u^3)_x = 0$$

with entropy function

$$U(u) = \int^u f(s)ds \Rightarrow U(u) = \frac{1}{4}u^4.$$

The entropy flux in this case is given by the relation:

$$F'(u) = U'(u)f'(u) \Rightarrow F(u) = \frac{1}{2}u^6.$$

Due to the convexity of the entropy function we define the new variables, entropy variables $v$ as

$$v(u) = U'(u) \Rightarrow v(u) = u^3 \Rightarrow u(v) = \sqrt[3]{v}$$

Applying this change of variables we gain the new entropy variables flux:

$$g(v) = f\big(u(v)\big) \Rightarrow g(v) = v.$$

Up to this point we have used the work cited [26], [27]. We now include the work of [19] for the construction of a 5-point semi and fully-discrete entropy conservative scheme. Main ingredient is the construction of the Entropy Conservative Flux $g^*_{i+1/2}$. LeFloch and Rohde (Theorem 3.1) state that the numerical scheme

$$\frac{d}{dt}u_i(t) = -\frac{1}{\Delta x}(g^*_{i+1/2} - g^*_{i-1/2}), \quad g^*_{i+1/2} = g^*(v_{i-1}, v_i, v_{i+1}, v_{i+2}),$$

with entropy conservative flux $g^*(v_{i-1}, v_i, v_{i+1}, v_{i+2})$ defined by

$$g^*(v_{i-1}, v_i, v_{i+1}, v_{i+2}) = \int_0^1 g(v_i + s(v_{i+1} - v_i))ds$$

$$-\frac{1}{12}\Big((v_{i+2} - v_{i+1}) \cdot B^*(v_i, v_{i+1}, v_{i+2}) - (v_i - v_{i-1}) \cdot B^*(v_{i-1}, v_i, v_{i+1})\Big),$$

is entropy conservative and third order accurate provided,

$$B^*(v, v, v) = B(v) = Dg(v).$$

The choices $B^* = 0$ or $B^* = B$, where $B(v) = Dg(v)$, produce classical shock waves, while other choices such as $B^* = 5B$ produce non-classical shocks. This is exactly the case of the entropy conservative scheme to be exhibited in this example. Entropy

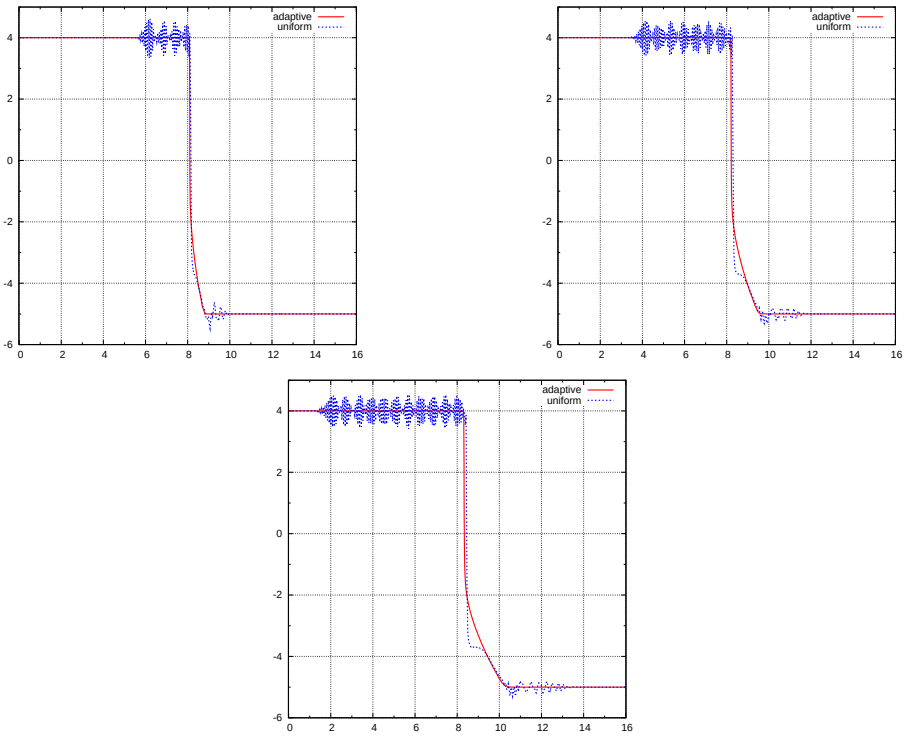



Figure 3.6: Problem 4: Burgers and entropy $U(u) = \frac{1}{6}u^3$ and jump $u_l = 4$, $u_r = -5$ initial conditions. $t = 0.01$, $t = 0.02$ $t = 0.03$. The oscillations appear to have a pattern, for this responsible is the factor 5 that appears in the entropy conservative flux $g^*$ and generates non-classical shocks

Conservative Semi-Discrete numerical scheme:

$$\frac{d}{dt}u_i(t) = -\frac{1}{2\Delta x_i}\Big(g(v_{i+1}) - g(v_{i-1})\Big) - \frac{5}{12\Delta x_i}\Big(-g(v_{i+2}) + 2g(v_{i+1}) - 2g(v_{i-1}) + g(v_{i-2})\Big)$$

The numerical tests were performed using 400 mesh points and Runge-Kutta 4th order temporal discretization for a comparison of *uniform mesh vs adaptive mesh* selection. Join graphs at each set of graphs expose the differences of the two cases -as in the previous example. In this example initial conditions are $u_l = 4$, $u_r = -5$.

We refer to Figure (3.6) for a graphical presentation of this scheme in the non-uniform adaptive case.

### 3.5.3 Conclusions

It is known that Entropy Conservative Schemes are not used for computation of entropy solutions but are used in [27] as a tool to construct entropy dissipative schemes for scalar and systems of Conservation Laws, so we conclude with the following remarks regarding the use of the BAS with Entropy conservative schemes for the evolution step,

1. The entropy conservative schemes, when combined with adaptive mesh selection converge to the entropy solution without oscillation.

2. Optimal CFL condition can be accomplished even with explicit time discretization

3. These schemes are used as a basis for the construction of numerical schemes for non-classical shock computation. Our results show that even in this case these schemes combined with appropriate mesh selection converge to the classical shock solution. It is important therefore to note:

    - the stabilization mechanism of mesh selection
    - such schemes should be used with great care if one wants to compute non classical shock behaviour

# Chapter 4

# Total Variation Bound

This Chapter is part of the work [24].

In this Chapter we consider numerical scheme that produce oscillations either due to their dispersive nature (like Richtmyer, MacCormack) or due to their anti-diffusive nature (like the unstable centered scheme FTCS). We prove that a proper use of non-uniform adaptively redefined meshes is capable of controlling the extremes in the sense that the Total Variation Increase due to oscillations is kept bounded. We shall moreover prove, under specific assumptions on the reconstructed meshes, that the increase of the Total Variation decreases with time.

To connect the work of this Chapter with the previous ones, we start by restating the BAS that is responsible for the overall treatment of non-uniform meshes and numerical schemes,

**Definition** (BAS)**.** Given mesh $M_x^n = \{a = x_1^n < \cdots < x_N^n = b\}$ and approximations $U^n = \{u_1^n, \ldots, u_N^n\}$,

1. (Mesh Reconstruction)
   Given $M_x^n$ and $U^n$ construct new mesh $M_x^{n+1} = \{a = x_1^{n+1} < \cdots < x_N^{n+1} = b\}$

2. (Solution Update)
   Given $M_x^n$, $U^n$ and $M_x^{n+1}$

   2a. construct the function $V^n(x)$ with $V^n(x_i^n) = u_i^n$

   2b. compute/update approximations $\hat{U}^n = \{\hat{u}_i^n, \ldots, \hat{u}_N^n\}$

3. (Time Evolution)
   Given $M_x^{n+1}$, $\hat{U}^n$ march in time to compute $U^{n+1} = \{u_1^{n+1}, \ldots, u_N^{n+1}\}$

We have already stated that the most important element of the BAS is the Solution Update procedure (Step 2) of the BAS. The effect of this Step is the core of the work in the current chapter.

Two basic parts consist the overall phenomenon. The first is Time Evolution for which responsible is the respective non-uniform numerical scheme and the second one is the mesh reconstruction procedure. Both parts affect the numerical approximation and more specifically the local extremes. We intend to couple these two parts and study their connection with the oscillations.

To this end, we start by stating the requirements that we place on the numerical schemes and on the mesh reconstruction procedure. We continue by discussing the

creation and propagation of oscillations -at the level of the local extremes- and present the way that the mesh reconstruction procedure affects the magnitude of the local extremes. Based on these properties we prove that the Total Variation Increase due to oscillations is kept bounded and, moreover we prove that under specific assumption it decreases.

In the numerical test section, we consider some known oscillatory numerical schemes. We prove that these schemes satisfy the posed requirements and we provide comparative numerical results on uniform and non-uniform adaptive meshes. The results of the theoretical analysis of the previous sections are depicted in these tests.

## 4.1 Requirements

This section is devoted to the requirements that we place on the numerical scheme and on the non-uniform mesh. They are restrictions on the time evolution and the mesh reconstruction steps of the BAS. We also provide the basic coupling of these requirements, which lies in the core of the analysis that we will perform in the next section.

We start with the following remark that explains the importance of the coupling of the two phenomena.

*Remark* 4.1.1. The effect of time evolution -Step 3. of the BAS- can be studied with several means such as the respective Modified Equation or the Von Neumann Analysis (at least for Linear problems on uniform meshes). In the contrary, the effect of the mesh reconstruction and the solution update procedure -Steps 1. and 2. of the BAS- cannot be discussed by these methods since they take place between two adjacent time steps. But still, they are part of the overall phenomenon hence their effect has to studied.

So, we start with the requirement that we place on the numerical scheme.

*Requirement* 1 (Evolution requirement). There exists a global constant $C$ that bounds the increase/change of every value -from $u_i^n$ to $u_i^{n+1}$- with respect to its neighbours $u_{i-1}^n$ and $u_{i+1}^n$, namely

$$|u_i^{n+1} - u_i^n| \leq C \max\left\{|u_{i+1}^n - u_i^n|, |u_i^n - u_{i-1}^n|\right\} \tag{4.1}$$

We shall prove -later in this Chapter- that this requirement is satisfied by the oscillatory numerical schemes that we consider. We moreover shall prove that it is satisfied by a wider range of numerical schemes, the Conservative schemes.

We move on, to the second requirement which is placed on the reconstruction of the mesh. Let $N$ be the constant number of nodes that we use and $M_x^{old} = \{x_i^{old}, i =$

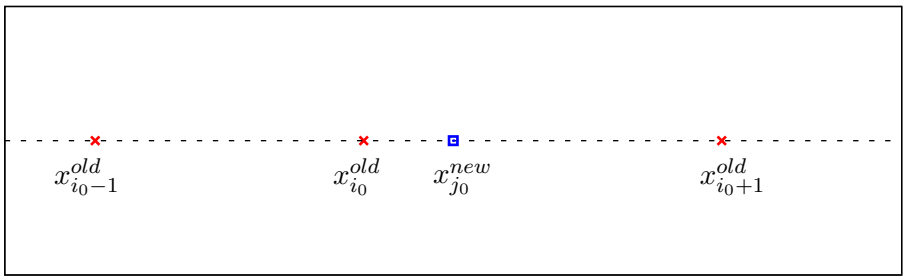



Figure 4.1: The new node $x^{new}_{j_0}$ avoid the places of the extreme $x^{old}_{i_0}$ and if the $\lambda$-rule requirement is satisfied then $x^{old}_{i_0+1} - x^{new}_{j_0} \leq \lambda(x^{old}_{i_0+1} - x^{old}_{i_0})$

$1, \ldots, N\}$ and $M^{new}_x = \{x^{new}_j, j = 1, \ldots, N\}$ the partitions of the mesh before and after the mesh reconstruction procedure respectively. Let also, $x^{old}_{i_0}$ be the position of a local extreme on the old mesh and $x^{new}_{j_0}$ the node of the new mesh that is closest to $x^{old}_{i_0}$. We assume that the mesh is smoothly varying with time in the sense $x^{new}_{j_0} \in [x^{new}_{i_0-1}, x^{new}_{i_0+1}]$.

For this configuration the second requirement reads,

*Requirement* 2 ($\lambda$-rule requirement for piecewise linears and interpolation). There exist a global constant $\lambda$ such that for every pair $x^{old}_{i_0}$ - $x^{new}_{j_0}$ (as just described) the following hold,

- If $x^{old}_{i_0-1} \leq x^{new}_{j_0} \leq x^{old}_{i_0}$ then

$$x^{old}_{i_0} - x^{new}_{j_0} \geq (1-\lambda)(x^{old}_{i_0} - x^{old}_{i_0-1})$$

  hence

$$x^{new}_{j_0} - x^{old}_{i_0-1} \leq \lambda(x^{old}_{i_0} - x^{old}_{i_0-1})$$

- If $x^{old}_{i_0} \leq x^{new}_{j_0} \leq x^{old}_{i_0+1}$ then

$$x^{new}_{j_0} - x^{old}_{i_0} \geq (1-\lambda)(x^{old}_{i_0+1} - x^{old}_{i_0})$$

  hence

$$x^{old}_{i_0+1} - x^{new}_{j_0} \leq \lambda(x^{old}_{i_0+1} - x^{old}_{i_0})$$

We refer to Figure 4.1 for a graphical presentation of the mesh configuration.

*Remark* 4.1.2. The meaning of the $\lambda$-rule requirement is that the new nodes avoid the places of the extremes, by a prescribed percentage $1-\lambda$ of the respective interval.

The $\lambda$-rule Requirement is placed on the mesh but is addressed to piecewise linear functions. The following discusses their relation,

*Remark* 4.1.3 (Application in piecewise linear functions). Assume that $u$ is a piecewise linear function that oscillates as depicted in Fig.(4.2). Assume moreover that the new nodes respect the $\lambda$-rule Req.(2) at the extremes in the respective subintervals. Let also $y=\upsilon$ be a horizontal line that separates the extremes.

According to Fig.(4.2), the node $x$ is the new node placed closer than all the other new nodes to the extreme old position $a$. Moreover $x\in[a,b]$ and by the $\lambda$-rule requirement

$$x-a\geq(1-\lambda)(b-a)$$

hence

$$\frac{b-x}{b-a}\leq\lambda$$

Now, since $u$ is linear in the interval $[a,b]$

$$\frac{u(b)-u(x)}{u(b)-u(a)}=\frac{b-x}{b-a}\leq\lambda$$

by the monotonicity of $u$ in the interval $[a,b]$ the previous relation recasts into

$$u(x)-u(b)\leq\lambda(u(a)-u(b))$$

since $0<\lambda<1$ and $u(b)<\upsilon$ the previous relation reads

$$u(x)\leq\lambda u(a)+(1-\lambda)u(b)\leq\lambda u(a)+(1-\lambda)\upsilon\Rightarrow u(x)-\upsilon\leq\lambda(u(a)-\upsilon)$$

For the rest of the extremes of Fig.(4.2), that is $u(b)$ and $u(c)$ and for the respective new nodes $y$, $z$ that are closest to these extremes, we can similarly prove that,

$$|u(y)-\upsilon|\leq\lambda|u(b)-\upsilon|\quad\text{ and }\quad|u(z)-\upsilon|\leq\lambda|u(c)-\upsilon|$$

The meaning of this remark is that for piecewise linear functions if a new node respects the $\lambda$-rule in a specific interval, the same should be true for the interpolated value of this new node with respect to the variation of the function in this interval.

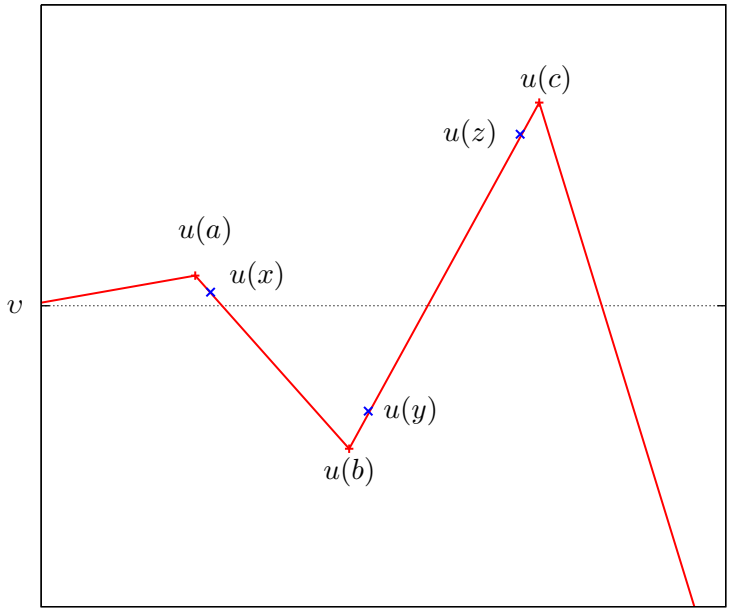



Figure 4.2: This figure depicts the application of the $\lambda$-rule in the case of a piecewise linear function. The places of the new nodes are depicted along with the old extremes.

As stated earlier, we need to study the Evolution part and the Reconstruction part both separately and together. For the separate analysis the requirements we have stated so far are sufficient, but for the joined analysis one more requirement is needed. This third requirement couples the previous ones, in the sense that relates the constants $C$ and $\lambda$ of the Evolution Req.(1) and the $\lambda$-rule Req.(2) respectively. This requirement emerged during the analysis of the problem and the development of the proofs but we state it here for the sake of completeness.

*Requirement* 3 (Coupling of the Evolution and the $\lambda$-rule Requirement). The constants $C$ and $\lambda$ of the Evolution Req.(1) and the $\lambda$-rule Req.(2) should be connected via the following relation

$$\lambda + 3\lambda C < 1 \tag{4.2}$$

These requirements and the introductory discussion are sufficient to continue with the way oscillations are created and transport.

## 4.2 Analysis

This section is devoted to the study of the extremes that oscillatory schemes produce. In the first paragraph, *Time Evolution* we discuss the creation and evolution of the extremes due to the numerical scheme and the mesh relocation procedure. We also devise recursive relations that incorporate the effect that both evolution and mesh relocation have on the extremes. In the second paragraph, *Extremes* we analyse the extremes via solving the recursive relations that describe them. We moreover prove uniform, with respect to the time step $k$, bounds on the magnitude of the extremes. In the third and fourth paragraph, *Variation* and *Variation-Revisited* we pass from the extremes and their magnitudes to the Total Variation Increase they define and the respective bounds.

### 4.2.1 Time Evolution

In this paragraph we discuss the creation and evolution of the extremes in a time step by time step manner. In every time step we shall discuss their temporal evolution (due to the numerical scheme) and their spatial modification (due to the node relocation and solution update procedure). We moreover provide recursive relations that couple the effect of both time evolution and node relocation on the magnitude of the extremes.

So we start with a jump initial condition which we discretize over a non-uniform (in the general case) mesh. In the description that follows we have split every step into two sub-steps. The first is the time evolution, which is due to the numerical scheme and is governed by the Evolution Req.(1) and the spatial modification, which is due to the mesh relocation and the solution update procedure and is governed by the $\lambda$-rule Req.(2).

1-st step We refer to Fig.(4.3) for a graphical description of the following configuration. The first nodal point located at the top of the shock that is $u_i^0$ will evolve according to the Evolution Req.(1), which reads

$$|u_i^1 - u_i^0| \leq C \max\left\{|u_i^0 - u_{i-1}^0|, |u_i^0 - u_{i+1}^0|\right\}$$

Since we consider jump initial conditions, it is obvious that

$$|u_i^0 - u_{i-1}^0| = 0 \quad \text{and} \quad |u_i^0 - u_{i+1}^0| \leq TV(u^0)$$

We denote $\hat{a}_1 = |u_i^0 - u_{i+1}^0|$ which describes the vertical distance of the value $u_i^0$ from it's neighbour $u_{i+1}^0$. Moreover we define $a_1 = C\hat{a}_1$ (in order to simplify the

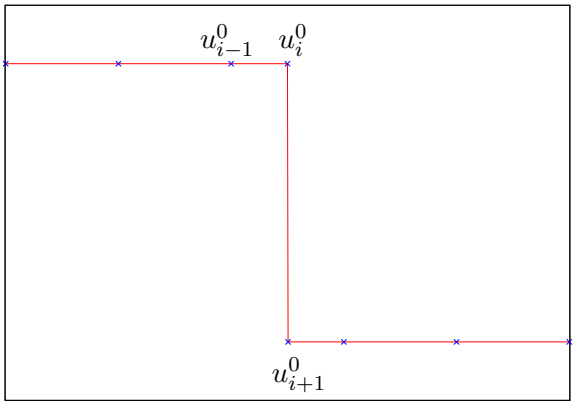



Figure 4.3: This is the initial condition. In this configuration we set $\hat{a}_1 = |u_i^0 - u_{i+1}^0|$

work that follows), so the Evolution Req.(1) for the value $u_i^0$ reads,

$$|u_i^1 - u_i^0| \leq C \max\{|u_i^0 - u_{i-1}^0|, |u_i^0 - u_{i+1}^0|\} \leq C|u_i^0 - u_{i+1}^0| = C\hat{a}_1 = a_1$$

To introduce the notation for the continuation of this work we define $E_1^{1/2}$ to be the magnitude of this extreme, hence

$$E_1^{1/2} = |u_i^1 - u_i^0| = a_1$$

To explain the symbolism, we use the letter $E$ because we refer to the magnitude of extremes, the subscript $_1$ states that we refer to 1-st extreme and the superscript $^{1/2}$ states that we have moved from the time step $k = 0$, with the use of the numerical scheme but the node relocation has not taken place yet.

We now relocate the nodes according to the mesh reconstruction procedure and because of the $\lambda$-rule Req.(2) the new extreme will be of magnitude $E_1^1$ (full superscript is used since the relocation has taken place) bounded by

$$E_1^1 \leq= a_1$$

Fig.(4.4) depicts the situation at the head of the shock at the end of the 1-st step.

*Remark* 4.2.1. This 1-st extreme shall "pollute" its neighbour by provoking the

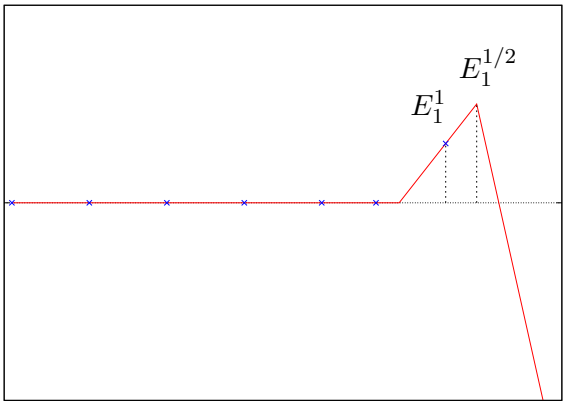



Figure 4.4: The resulting situation at the head of the shock at the end of the 1-st time step. Only one extreme exists in this time step and it is of magnitude $E_1^1$. The numerical solution -before the remeshing procedure takes place- is depicted red, the new nodes -that occur after the remeshing procedure- are depicted in blue. The new node -that is closest to the old extreme- avoids the extreme by the $\lambda$-rule resulting in a magnitude bounded by $E_1^1 \leq \lambda a_1$. The reconstruction of the numerical solution over the new mesh results in the clipping of the magnitude of the extreme according to the $\lambda$-rule.

appearance of a 2-nd extreme of the opposite direction.

*Remark* 4.2.2. ($E_m^k$ Notation) We denote by $E_m^{k+1}$ (full superscript) the bound on the magnitude of $m$-th extreme at the end of the $k$-th time step, that is after the time evolution and the mesh reconstruction procedure and by $E_m^{k-1/2}$ (half superscript) the bound on the magnitude of the $m$-th extreme at the $k$-th time step, after time evolution (due to the numerical scheme) and before the mesh reconstruction procedure.

2-nd step At the end of the previous step, we had only one extreme of magnitude $E_1^1$ bounded as $E_1^1 \leq \lambda a_1$. Due to the temporal evolution -numerical scheme- we expect the 1-st extreme to evolve to a new value, we also expect the creation of a 2-nd extreme at the left side of the 1-st extreme. We will study each extreme separately.

1-st Extr. The Evolution Req.(1) dictates that this extreme shall evolve according to

$$|u_i^{1+1/2} - u_i^1| \leq C \max\left\{|u_i^1 - u_{i-1}^1|, |u_i^1 - u_{i+1}^1|\right\}$$

where we use half superscript in $u$ since the relocation procedure has not taken

place yet. From the previous time step we have that,

$$|u_i^1 - u_{i-1}^1| \leq E_1^1 \quad \text{and} \quad |u_i^1 - u_{i+1}^1| \leq 2E_1^1 + \hat{a}_2$$

To justify the second inequality we return at the end of the time step $k = 1$ and notice that the node $i+1$ is placed along the shock, which -by symmetry- is of variation $E_1^1 + TV(u^0) + E_1^1$. So the Evolution Req.(1) for the 1-st Extreme reads,

$$|u_i^{1+1/2} - u_i^1| \leq C(2E_1^1 + \hat{a}_2) = 2CE_1^1 + a_2$$

where we have defined $a_2 = C\hat{a}_2$. If now we set $\upsilon$ to be the level from which we measure the magnitudes of the extremes, the previous bound recasts,

$$|(u_i^{1+1/2} - \upsilon) - (u_i^1 - \upsilon)| \leq 2CE_1^1 + a_2$$

By setting $E_1^{1+1/2} = u_i^{1+1/2} - \upsilon$ and since $E_1^1 = u_i^1 - \upsilon$ we deduce that the magnitude of the 1-st extreme will be bounded as

$$E_1^{1+1/2} \leq E_1^1 + 2CE_1^1 + a_2$$

Now the relocation procedure takes place and the $\lambda$-rule Req.(2) dictates that the magnitude of the 1-st extreme at the end of this step shall be bounded as follows,

$$E_1^2 = \lambda(E_1^1 + 2CE_1^1 + a_2)$$

2-nd Extr. The Evolution Req.(1) dictates that this extreme shall be created and controlled as,

$$|u_{i-1}^{1+1/2} - u_{i-1}^1| \leq C \max\left\{|u_{i-1}^1 - u_{i-2}^1|, |u_{i-1}^1 - u_i^1|\right\}$$

where again half superscript is used on $u_{i-1}^{1+1/2}$ since the relocation procedure has not taken place yet. From the previous time step we know that

$$|u_{i-1}^1 - u_{i-2}^1| = 0 \quad \text{and} \quad |u_{i-1}^1 - u_i^1| \leq E_1^1$$

So the Evolution Req.(1) recasts, for the 2-nd extreme as follows,

$$|u_{i-1}^{1+1/2} - u_{i-1}^{1}| \leq CE_1^1$$

or by noting that $u_{i-1}^1 = \upsilon$ is the level from which we measure the magnitudes of the extremes, the previous bound recasts

$$E_2^{1+1/2} \leq CE_1^1$$

where, as we explained earlier half superscript is used because the relocation procedure has not taken place yet.

Now the relocation procedure takes place and the $\lambda$-rule Req.(2) dictates that the magnitude of the 2-nd extreme at the end of this step shall be bounded as follows,

$$E_2^2 = \lambda CE_1^1$$

So at the end of the 2-nd step the bounds on the existing extremes are as follows,

$$E_1^2 = \lambda(E_1^1 + 2CE_1^1 + a_2), \quad E_2^2 = \lambda CE_1^1$$

Fig.(4.5) depicts the situation at the head of the shock at the end of the 2-nd step.

*Remark* 4.2.3. The 2-nd extreme shall provoke the appearance of a new extreme of the opposite direction. This is the pollution process.

3-rd step At the end of the previous step we had two extremes with magnitudes $E_1^2$ and $E_2^2$. In this step we expect them to evolve to new values $E_1^3$ and $E_2^3$, we also expect a new extreme to appear, namely $E_3^3$.

1-st Extr. Following the discussion of the previous steps, we note that the evolution of the 1-st extreme will be governed by the Evolution Req.(1), so

$$|u_i^{2+1/2} - u_i^2| \leq C \max\left\{|u_i^2 - u_{i-1}^2|, |u_i^2 - u_{i+1}^2|\right\}$$

where from the previous time steps we note

$$|u_i^2 - u_{i-1}^2| \leq E_1^2 + E_2^2 \quad \text{and} \quad |u_i^2 - u_{i+1}^2| \leq 2E_1^2 + \hat{a}_3$$

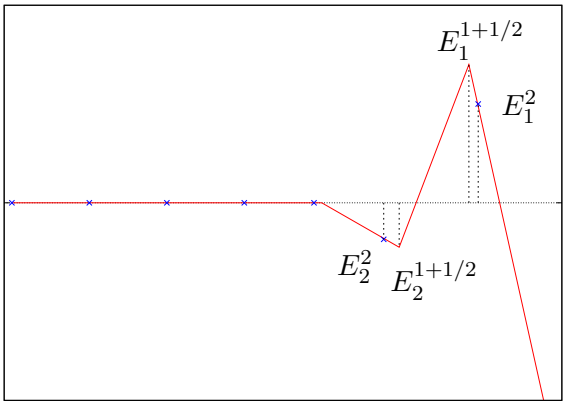



Figure 4.5: The resulting situation at the end of the 2-nd time step. Two extremes of magnitudes $E_1^2$ and $E_2^2$ exist in this time step. The numerical solution before the remeshing procedure is depicted in red and the new nodes -after the remeshing procedure are depicted in blue. The new nodes -that are closest to the previous extremes- avoid the extremes by the $\lambda$-rule

We also note that from the previous time step the bound of $E_1^2 \leq \lambda(E_1^1 + 2CE_1^1 + a_2)$ is obviously larger than the bound of $E_2^2 \leq \lambda C E_1^1$ hence the Evolution Req.(1) for the 1-st extreme reads as follows (after the subtraction of $v$),

$$E_1^{2+1/2} \leq E_1^2 + 2CE_1^2 + a_3$$

Now the node relocation procedure takes place and the new magnitude of the 1-st extreme shall be bounded by

$$E_1^3 \leq \lambda(E_1^2 + 2CE_1^2 + a_3)$$

2-nd Extr. Using similar arguments as before, the Evolution Req.(1) dictates the evolution of the 2-nd extreme as follows,

$$|u_{i-1}^{2+1/2} - u_{i-1}^2| \leq \max\left\{|u_{i-1}^2 - u_{i-2}^2|, |u_{i-1}^2 - u_i^2|\right\}$$

From the previous time step we note

$$|u_{i-1}^2 - u_{i-2}^2| \leq E_2^2 \quad \text{and} \quad |u_{i-1}^2 - u_i^2| \leq E_2^2 + E_1^2$$

hence the Evolution Req.(1) for the 2-nd extreme yields,

$$E_2^{2+1/2} \leq E_2^2 + C(E_2^2 + E_1^2)$$

Now, relocation takes place and the new 2-nd extreme shall be of magnitude bounded by

$$E_2^3 = \lambda(E_2^2 + C(E_2^2 + E_1^2))$$

3-rd Extr. Repeating the work we did for the 2-nd extreme in the previous time step, the magnitude of the 3-rd extreme after both the time evolution and the node relocation procedure will be bounded

$$E_3^3 \leq \lambda C E_2^2$$

So at the end of the 3-rd step the bounds on the magnitudes of the existing extremes are as follows,

$$\begin{aligned}
&E_1^3 \leq \lambda(E_1^2 + 2CE_1^2 + a_3) \leq \lambda^3(1+2C)^2 a_1 + \lambda^2(1+2C)a_2 + \lambda a_3,\\
&E_2^3 \leq \lambda(E_2^2 + C(E_2^2 + E_1^2)) \leq \lambda^3 2C(1+2C)a_1 + \lambda^2 C a_2,\\
&E_3^3 \leq \lambda C E_2^2 \leq \lambda^3 C^2 a_1
\end{aligned}$$

Fig.(4.6) depicts the situation at the end of the 3-rd step.

A synopsis of the results is depicted in the following,

| | $m=1$ | $m=2$ | $m=3$ | $m=4$ |
|---|---|---|---|---|
| $k=1$ | $\lambda a_1$ | 0 | 0 | 0 |
| $k=2$ | $\lambda^2(1+2C)a_1 + \lambda a_2$ | $\lambda^2 C a_1$ | 0 | 0 |
| $k=3$ | $\lambda^3(1+2C)^2 a_1 + \lambda^2(1+2C)a_2 + \lambda a_3$ | $\lambda^3 2C(1+2C)a_1 + \lambda^2 C a_2$ | $\lambda^3 C^2 a_1$ | 0 |

For the sake of completeness we define the increases $a_i$ that we used throughout the previous paragraph. For this we first analyse the variation of the shock at the $k$-th time step. It consists of three parts, the oscillatory part at the top of the shock with magnitude $E_1^k$, the main part of the shock which is of variation $TV(u^0)$ and the oscillatory part at the foot of the shock being of magnitude $E_1^k$.

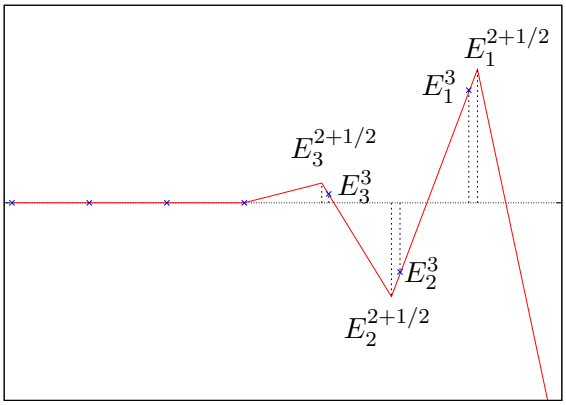



Figure 4.6: The resulting situation at the end of the 3-rd time step. Three extremes of magnitude $E_1^3$, $E_2^3$, $E_3^3$ exist in this time step.

**Definition 4.2.1** (Definition of the $a_i$ increases)**.** Let $u_i^k$ be the value at the top of the shock. The node $x_{i+1}^k$ is located along the shock, in one of the three parts that consist the shock.

We define

$$\hat{a}_k = \left(|u_i^k - u_{i+1}^k| - 2E_1^k\right)_+$$

where the subscript $_+$ denotes the positive part. Moreover we define $a_k = C\hat{a}_k$.

*Remark* 4.2.4. By definition, $\hat{a}_k$ describes the possibly more that $2E_1^k$ distance $|u_i^k - u_{i+1}^k|$. That is if the distance $|u_i^k - u_{i+1}^k| < 2E_1^k$ then $\hat{a}_k = 0$ and hence $a_k = 0$.

To comment more on the increases $a_k$ we say that they are the main reason that the evolution description is successful. We can see this from the 1-st step of the analysis, where the magnitude of the 1-st extreme was $E_1^1 = \lambda a_1$.

We can generalise the situation in the $k$-th time step as follows,

$k$-th step To generalise the description we presented in the previous time steps, we introduce here recursive relations for the general extreme $m$ at the general time step $k$,

$$\begin{cases} E_m^k = \lambda\left(E_m^{k-1} + C\cdot(E_m^{k-1} + E_{m-1}^{k-1})\right), & \text{for the general extreme } m > 1 \\ E_1^k = \lambda\left(E_1^{k-1} + 2C\cdot E_1^{k-1} + a_k\right), & \text{for the first extreme } m = 1 \end{cases} \tag{4.3}$$

*Remark* 4.2.5. We add at least $2CE_1^{k-1}$ in the increase of the 1-st extreme even if the

actual increase of the highest node is less. This increases the magnitude of the 1-st extreme but at the same time simplifies the presentation and the route of the proof. More precise increases result in sharper final bounds.

In analysing these recursive relations, we see that for the evolution of the extreme $E_m^{k-1}$ to $E_m^k$ we take into account the neighbouring extreme in the right hand side, $E_{m-1}^{k-1}$. To justify such a choice, we have to prove that the bounds on the magnitudes of the extremes $E_m^k$ constitute a decreasing sequence with respect to $m = 1, \ldots$ for every step $k$. That is $E_1^k > E_2^k > E_3^k > \cdots$. This is accomplished in the following lemma,

**Lemma 4.2.1.** *For every step $k$ the magnitudes of the bounds of the extremes are given by Rel.(4.3) and they are in a decreasing order w.r.t $m = 1, 2, 3, \ldots$*

*Proof.* By induction. Let's assume that in the step $k$ the extremes are given by Rel.(4.3) and that they are in a decreasing order, that is $E_{m+1}^k \leq E_m^k$ for every $m = 1, \cdots$.

We shall first prove that the recursive relations Rel.(4.3) are valid for $k+1$,

1-st Extr. The Evolution req.(1) for the 1-st extreme dictates that

$$E_1^{k+1/2} \leq E_1^k + C \max\left\{E_1^k + E_2^k, 2E_1^k + \hat{a}_{k+1}\right\}$$

by the induction hypothesis $E_2^k \leq E_1^k$ so,

$$E_1^{k+1/2} \leq E_1^k + 2CE_1^k + a_{k+1}$$

where $a_{k+1} = C\hat{a}_{k+1}$. Now the relocation part takes place, hence

$$E_1^{k+1} \leq \lambda(E_1^k + 2CE_1^k + a_{k+1})$$

So the 2-nd part of Rel.(4.3) valid.

$m$-th Exr. The Evolution req.(1) for the $m$-th extreme dictates that

$$E_m^{k+1/2} \leq E_m^k + C \max\left\{E_m^k + E_{m+1}^k, E_m^k + E_{m-1}^k\right\}$$

by the induction hypothesis $E_{m+1}^k \leq E_{m-1}^k$ so,

$$E_m^{k+1/2} \leq E_m^k + C(E_m^k + E_{m-1}^k)$$

Now the relocation part takes place, hence

$$E_m^{k+1} \leq \lambda \left( E_m^k + C(E_m^k + E_{m-1}^k) \right)$$

So the 1-st part of Rel.(4.3) is valid.

We shall now prove that the monotonicity of the extremes is preserved, that is $E_{m+1}^{k+1} \leq E_m^{k+1}$ for every $m = 1, 2, 3, \ldots$

$m = 1$ The recursive relations Rel.(4.3) read for $E_1^{k+1}$ and $E_2^{k+1}$ as follows,

$$E_2^{k+1} = \lambda(E_2^k + CE_2^k + CE_1^k),$$
$$E_1^{k+1} = \lambda(E_1^k + CE_1^k + CE_1^k + a_{k+1})$$

Utilising the induction hypotheses and either one of the lower bound assumptions of $a_{k+1}$ the result $E_2^{k+1} < E_1^{k+1}$ is immediate.

$m > 1$ The recursive relations Rel.(4.3) read for $E_m^{k+1}$ and $E_{m+1}^{k+1}$ as follows,

$$E_{m+1}^{k+1} = \lambda(E_{m+1}^k + CE_{m+1}^k + CE_m^k)$$
$$E_m^{k+1} = \lambda(E_m^k + CE_m^k + CE_{m-1}^k)$$

The induction hypotheses states that $E_{m+1}^k \leq E_m^k \leq E_{m-1}^k$, so immediately we conclude that $E_{m+1}^{k+1} \leq E_m^{k+1}$.

So, it is proven that in order to bound the new magnitude of every extreme we could use the recursive relations Rel.(4.3). □

*Remark* 4.2.6. The previous lemma does not imply that the magnitudes of the actual extremes are in a decreasing order, merely that the bounds $E_m^k$ of the actual extremes are in a decreasing order.

Having devised recursive evolution relations for the extremes, that is Rel.(4.3), that comprise the effect of both the numerical scheme and the relocation procedure we continue with the study of the bounds of their magnitudes.

### 4.2.2 Extremes

In this paragraph we solve the recursive relation Rel.(4.3) for every extreme $m$. The resulting form is valid form every $m = 1, 2, 3, \ldots$ so the two recursive relations of Rel.(4.3)

collapse into one non-recursive relation. We moreover provide uniform -with respect to the time step $k$- bounds on the magnitude of the extremes.

We start by providing bounds on the extremes of $E_m^k$ with respect to the sequence of increases $a_i$.

**Lemma 4.2.2** (Magnitude of the 1-st extreme)**.** *The magnitude of the first extreme in the $k$-th time step is bounded by,*

$$E_1^k \leq \lambda \sum_{j=1}^{k} \lambda^{k-j}(1+2C)^{k-j} a_j,$$

*or, by setting $l = k - j$*

$$E_1^k \leq \lambda \sum_{l=0}^{k-1} \lambda^{l}(1+2C)^{l} a_{k-l} \tag{4.4}$$

*Proof.* By induction. We note from the previous discussion that

$$E_1^1 \leq \lambda a_1 = \lambda \sum_{j=1}^{1} \lambda^{k-j}(1+2C)^{k-j} a_j$$

For the induction hypothesis we assume that the magnitude of the 1-st extreme is bounded in the $k$-th time step as

$$E_1^k \leq \lambda \sum_{l=0}^{k-1} \lambda^{l}(1+2C)^{l} a_{k-l}$$

Using the evolution relation (4.3) of the 1-st extreme, that is

$$E_1^{k+1} \leq \lambda \left( E_1^k + 2C E_1^k + a_{k+1} \right)$$

we can bound its magnitude in the $k+1$ time step,

$$E_1^{k+1} \leq \lambda \left( (1+2C) E_1^k + a_{k+1} \right)$$

The right hand side recast -by the induction hypothesis- as follows,

$$\begin{aligned} E_1^{k+1} &\leq \lambda \left( (1+2C)\lambda \sum_{j=1}^{k} \lambda^{k-j}(1+2C)^{k-j} a_j + a_{k+1} \right) \\ &\leq \lambda \left( \sum_{j=1}^{k} \lambda^{k+1-j}(1+2C)^{k+1-j} a_j + \lambda^{k+1-(k+1)}(1+2C)^{k+1-(k+1)} a_{k+1} \right) \\ &\leq \lambda \sum_{j=1}^{k+1} \lambda^{k+1-j}(1+2C)^{k+1-j} a_j \end{aligned}$$

This completes the proof regarding the bound of the magnitude of the 1-st extreme. □

We now need a similar bound on the magnitude of the 2-nd extreme,

**Lemma 4.2.3** (Magnitude of the 2-nd extreme)**.** *The magnitude of the second extreme in the k-th time step is bounded by,*

$$E_2^k \leq \lambda^2 C \sum_{j=1}^{k-1} \binom{k-j}{k-j-1} \lambda^{k-j-1}(1+2C)^{k-j-1} a_j,$$

*or by setting* $l = k - j$

$$E_2^k \leq \lambda^2 C \sum_{l=1}^{k-1} \binom{l}{l-1} \lambda^{l-1}(1+2C)^{l-1} a_{k-l} \tag{4.5}$$

*Proof.* Proof by induction using relations (4.3), (4.4), (4.5) and the fact $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$. We note from the previous discussion that

$$E_2^2 \leq \lambda^2 C a_1 = \lambda^2 C \sum_{j=1}^{2-1} \binom{2-j}{2-j-1} \lambda^{2-j-1}(1+2C)^{2-j-1} a_j$$

For the induction hypothesis we assume that

$$E_2^k \leq \lambda^2 C \sum_{j=1}^{k-1} \binom{k-j}{k-j-1} \lambda^{k-j-1}(1+2C)^{k-j-1} a_j$$

and for the induction step we have the following

$$E_2^{k+1} = \lambda\left(E_2^k + C(E_2^k + E_1^k)\right) = \lambda\left((1+C)E_2^k + CE_1^k\right)$$
$$\leq \lambda\left(\lambda^2 C(1+C)\sum_{l=1}^{k-1}\binom{l}{l-1}\lambda^{l-1}(1+2C)^{l-1}a_{k-l} + \lambda C\sum_{l=0}^{k-1}\lambda^l(1+2C)^l a_{k-l}\right)$$

Where in the last step we utilised the induction hypothesis. Now, since $1+C \leq 1+2C$ the bound recasts

$$\begin{aligned}
E_2^{k+1} &\leq \lambda\left(\lambda C\sum_{l=1}^{k-1}\binom{l}{l-1}\lambda^l(1+2C)^l a_{k-l} + \lambda C\sum_{l=0}^{k-1}\lambda^l(1+2C)^l a_{k-l}\right)\\
&= \lambda^2 C\left(\sum_{l=1}^{k-1}\left(\binom{l}{l-1}+1\right)\lambda^l(1+2C)^l a_{k-l} + a_k\right)\\
&= \lambda^2 C\left(\sum_{l=1}^{k-1}\left(\binom{l}{l-1}+\binom{l}{l}\right)\lambda^l(1+2C)^l a_{k-l} + a_k\right)\\
&= \lambda^2 C\left(\sum_{l=1}^{k-1}\binom{l+1}{l}\lambda^l(1+2C)^l a_{k-l} + a_k\right)\\
&= \lambda^2 C\left(\sum_{l=1}^{k-1}\binom{l+1}{l}\lambda^l(1+2C)^l a_{k-l} + \binom{0+1}{0}\lambda^0(1+2C)^0 a_{k-0}\right)\\
&= \lambda^2 C\sum_{l=0}^{k-1}\binom{l+1}{l}\lambda^l(1+2C)^l a_{k-l}
\end{aligned}$$

Finally we set $\mu = l+1$ and the bound on the magnitude of the 2-nd extreme reads,

$$E_2^{k+1} \leq \lambda^2 C\sum_{\mu=1}^{(k+1)-1}\binom{\mu}{\mu-1}\lambda^{\mu-1}(1+2C)^{\mu-1}a_{k+1-\mu} \tag{4.6}$$

and this completes the proof regarding the magnitude of the 2-nd extreme. □

Similarly we prove that the magnitude of the 3-rd extreme $k$-th time step is bounded by,

$$E_3^k \leq \lambda^3 C^2\sum_{j=1}^{k-1}\binom{k-j}{k-j-2}\lambda^{k-j-2}(1+2C)^{k-j-2}a_j,$$

or by setting $l = k - j$,

$$E_3^k \leq \lambda^3 C^2 \sum_{l=2}^{k-1} \binom{l}{l-2} \lambda^{l-2} (1+2C)^{l-2} a_{k-l}$$

We can generalise the previous lemmas, in a compact form for the $m$-th extreme in the $k$-th time step.

**Lemma 4.2.4** (Magnitude of the $m$-th extreme)**.** *The magnitude of the $m$-th extreme in the $k$-th time step is bounded by,*

$$E_m^k \leq \lambda^m C^{m-1} \sum_{j=1}^{k-1} \binom{k-j}{k-j-m+1} \lambda^{k-j-m+1} (1+2C)^{k-j-m+1} a_j,$$

*or by setting $l = k - j$,*

$$E_m^k \leq \lambda^m C^{m-1} \sum_{l=m-1}^{k-1} \binom{l}{l-m+1} \lambda^{l-m+1} (1+2C)^{l-m+1} a_{k-l} \tag{4.7}$$

*Proof.* The proof is exactly the same as in the 2-nd extreme and is omitted. □

The last lemmas provided bounds on the magnitudes of the extremes. In the following remark we merge these bounds in a single relation valid for every $m = 1, 2, 3, \ldots$.

*Remark* 4.2.7. The bound we have extracted for the $m$-th extreme at the $k$-th time step, that is Rel.(4.7):

$$E_m^k \leq \lambda^m C^{m-1} \sum_{l=m-1}^{k-1} \binom{l}{l-m+1} \lambda^{l-m+1} (1+2C)^{l-m+1} a_{k-l}$$

is valid for every $m = 1, 2, 3, \ldots$, -not just for $m > 1$- since, for $m = 1$ the bound we extracted for the 1-st extreme at the $k$-th time step, that is Rel.(4.4),

$$E_1^k \leq \lambda \sum_{l=0}^{k-1} \lambda^l (1+2C)^l a_{k-l}$$

can be written in the form

$$E_1^k \leq \lambda^1 C^{1-1} \sum_{l=1-1}^{k-1} \binom{l}{l-1+1} \lambda^{l-1+1} (1+2C)^{l-1+1} a_{k-l}$$

So far we have described the creation and evolution of the extremes. We provided Recursive relations (4.3) that connect the magnitudes of extremes, we have solved the recursions and merged the magnitudes of the extremes into a single relation (4.7).

Now, we note that the bounds on the magnitudes of the extremes, that we have extracted, depend on the time step $k$. We shall bound the magnitudes of the extremes uniformly with respect to the time step $k$. This will allow us to provide the final proof regarding the total variation increase due to oscillations.

**Lemma 4.2.5** (Uniform -with respect to the time step $k$- bound on the extremes)**.** *If there is a constant $M > 0$ such that $a_i \leq CM$ for every $i = 0, \cdots \infty$ and if $\lambda + 2\lambda C < 1$ then every extreme $m$ is uniformly -with respect to the time step $k$- bounded as,*

$$E_m^k \leq M \left( \frac{\lambda C}{1 - \lambda - 2\lambda C} \right)^m \tag{4.8}$$

*Proof.* The magnitude of the bound of the $m$-th extreme, $m = 1, 2, 3, \ldots$ at the $k$-th time step, with $m \leq k$, is given by the Rel.(4.7),

$$E_m^k \leq \lambda^m C^{m-1} \sum_{l=m-1}^{k-1} \binom{l}{l-m+1} \lambda^{l-m+1} (1+2C)^{l-m+1} a_{k-l}$$

Since the increase $a_i$ are uniformly bounded, $a_i \leq CM$ (we refer to the definition of the increases $a_i$ Def.(4.2.1)) we can bound the extremes as,

$$E_m^k \leq \lambda^m C^m M \sum_{l=m-1}^{k-1} \binom{l}{l-m+1} \lambda^{l-m+1} (1+2C)^{l-m+1}.$$

Setting $\nu = l - m + 1$ the previous relation reads,

$$E_m^k \leq \lambda^m C^m M \sum_{\nu=0}^{k-m} \binom{\nu+m-1}{\nu} \lambda^\nu (1+2C)^\nu.$$

All the terms inside the sum are positive, hence the right hand side of the previous relation is increasing with respect to $k$. Hence it can be bounded for $k = \infty$ as follows,

$$E_m^k \leq \lambda^m C^m M \sum_{\nu=0}^{\infty} \binom{\nu+m-1}{\nu} \lambda^\nu (1+2C)^\nu$$

or

$$E_m^k \leq \lambda^m C^m M \sum_{\nu=0}^{\infty} \binom{\nu+m-1}{\nu} (\lambda + 2\lambda C)^\nu$$

For the convergence of the previous infinite sum we recall at this point the power series expansion (for the proof of which we refer to the Appendix B)

$$\sum_{\nu=0}^{\infty} \binom{\nu+m-1}{\nu} t^\nu = \frac{1}{(1-t)^m}, \qquad \text{whenever } |t| < 1,$$

and since $\lambda + 2\lambda C < 1$ the last bound on $E_m^k$ reads as follows,

$$\begin{aligned} E_m^k &\leq \lambda^m C^m M \frac{1}{(1-\lambda-2\lambda C)^m} \\ &= M \left( \frac{\lambda C}{1-\lambda-2\lambda C} \right)^m \end{aligned}$$

Which proves the assertion of the lemma. □

*Remark* 4.2.8. If moreover we assume $\lambda + 3\lambda C < 1$ -instead of $\lambda + 2\lambda C < 1$- then the sequence of bounds on the extremes $\{E_m^k\}$ is decreasing with respect to $m$, since they can be written as,

$$E_m^k \leq M \left( \frac{\lambda C}{1-\lambda-2\lambda C} \right)^m$$

hence

$$\lim_{m\to\infty} E_m^k \leq \lim_{m\to\infty} M \left( \frac{\lambda C}{1-\lambda-2\lambda C} \right)^m = 0$$

since $E_m^k \geq 0$ and the fraction $\frac{\lambda C}{1-\lambda-2\lambda C} < 1$ because $\lambda + 3\lambda C < 1$.

We are ready now to measure the total variation increase due to the oscillations.

### 4.2.3 Variation

In the previous lemma we proved that each extreme separately is of bounded magnitude, uniformly with respect to the time steps $k$. The next Theorem is the basic one and states that in addition to the magnitude of the extremes, also the sum of the extremes is also bounded uniformly with respect to the time step $k$. This constitutes our main result, and it is one step before the final Total Variation Increase bound.

**Theorem 4.2.1** (Main Result)**.** *We assume that the requirements Req.(1) and Req.(2) are satisfied for $\lambda$ such that $\lambda + 3\lambda C < 1$ by the numerical scheme and the mesh. We*

*more over assume that the sequence* $\{a_i, i = 1, \infty\}$ *is uniformly bounded* $a_i \leq CM$*. Then the sum of the magnitudes of the extremes is uniformly -with respect to the time step k- bounded as follows,*

$$\sum_{m=1}^{k} E_m^k \leq M \frac{1 - \lambda - 2\lambda C}{1 - \lambda - 3\lambda C}$$

*Proof.* We shall utilise relation (4.8), which is valid since the requirements of the relevant lemma are satisfied. At the end of the $k$-th step we have $k$ extremes $E_1^k, E_2^k, \ldots, E_k^k$. The sum -with respect to $m$- of their magnitudes can be bounded as,

$$\begin{aligned}\sum_{m=1}^{k} E_m^k &\leq M \sum_{m=1}^{k} \left(\frac{\lambda C}{1 - \lambda - 2\lambda C}\right)^m \leq M \sum_{m=1}^{\infty} \left(\frac{\lambda C}{1 - \lambda - 2\lambda C}\right)^m = M \frac{1}{1 - \frac{\lambda C}{1 - \lambda - 2\lambda C}} \\ &\leq M \frac{1 - \lambda - 2\lambda C}{1 - \lambda - 3\lambda C}\end{aligned}$$

where the second inequality and the only equality are valid since $\lambda + 3\lambda C < 1$. □

Summarising and concluding we can state the following theorem, which constitute our target result on the Total Variation Increase.

**Theorem 4.2.2** (Total Variation Increase Bound)**.** *Given the requirements of the previous Theorem, the Total Variation increase due to the oscillations is bounded and given by*

$$\text{TVI} \leq 2M \frac{1 - \lambda - 2\lambda C}{1 - \lambda - 3\lambda C} \tag{4.9}$$

*Proof.* The variation of the oscillatory part is bounded by twice the magnitude of the extremes. So,

$$\text{TVI} \leq 2M \sum_{m=1}^{\infty} \left(\frac{\lambda C}{1 - \lambda - 2\lambda C}\right)^m \leq 2M \frac{1 - \lambda - 2\lambda C}{1 - \lambda - 3\lambda C}$$

where the last inequality results from the previous Theorem (Main Result). □

Although we proved our main result, we can gain better insight if we study the effect of each increase factor -separately- to the total variation. For this reason we include the following paragraph.

### 4.2.4 Variation-Revisited

We shall follow now another approach that will provide us a with further insight of the "pollution" process and with a sharper bound on the Total Variation Increase.

This approach differs from the previous one in the sense that instead of adding directly the magnitudes of the extremes $E_m^k$, we compute the contributions of the increase terms $a_i$, $i = 1, 2, 3, \ldots$ in the each one of the extremes $E_m^k$ separately. Then we add this contributions with respect to $a_i$.

$a_1$ cont. The contribution of $a_1$ in the $k$-th step,
In the $k$-th time step there exist $k$ extremes and the increase factor $a_1$ is present in each one of these extremes. So we extract the contribution of $a_1$ in all the extremes that are produced during this procedure up to the $k$-th time step.

The contribution of $a_1$ in the 1-st extreme in the $k$-th time step is given by the relation (4.4) and reads as

$$\lambda\lambda^{k-1}(1+2C)^{k-1}$$

and in the general extreme $m$ the contribution of $a_1$ is

$$\lambda^m C^{m-1}\binom{k-1}{k-m}\lambda^k(1+2C)^{k-m}$$

Summing these contributions with respect to $m$ we end up with the total contribution of $a_1$ in the $k$-th time step,

$$\begin{aligned} I_{a_1}^k &= \sum_{m=1}^{k} \lambda^m C^{m-1}\binom{k-1}{k-m}\lambda^{k-m}(1+2C)^{k-m}a_1 \\ &= \lambda^k \sum_{m=1}^{k} \binom{k-1}{k-m} C^{m-1}(1+2C)^{k-m}a_1 \\ (\text{for } \nu = k-m) \quad &= \lambda^k \sum_{\nu=0}^{k-1} \binom{k-1}{\nu} C^{k-1-\nu}(1+2C)^{\nu}a_1 \\ &= \lambda^k(1+3C)^{k-1}a_1 = \lambda(\lambda+3\lambda C)^{k-1}a_1 \end{aligned}$$

$a_2$ cont. The contribution of $a_2$ in the $k$-th step,
Similarly we notice that in the $k$-th time step the increase factor $a_2$ contributes in

all the extremes except the last one, $m = k$ and its total contributions is

$$I^k_{a_2} = \lambda(\lambda + 3\lambda C)^{k-2} a_2$$

$a_m$ cont. The contribution of $a_m$ in the $k$-th step,

More generally, in the $k$-th time step the increase factor $a_m$ contributes in all but $m-1$ extremes (the last $m-1$) in the $k$-th time step ($k \geq m$), and its total contribution is

$$I^k_{a_m} = \lambda(\lambda + 3\lambda C)^{k-m} a_m$$

*Remark* 4.2.9. We note here that as long as $\lambda < \frac{1}{1+3C}$ i.e $\lambda + 3\lambda C < 1$ each one of these contributions converges to 0 as $k \to \infty$,

$$\lim_{k\to\infty} I^k_{a_m} = \lim_{k\to\infty} \lambda(\lambda + 3\lambda C)^{k-m} a_m = 0 \quad \text{if} \quad \lambda < \frac{1}{1+3C} \tag{4.10}$$

This is the very essence of the $\lambda$-rule affect. Namely a remeshing procedure which respects the $\lambda$-rule requirement (2) at the extremes is able to limit the increase of the variation due to each $a_i$ -eventually kill it- and hence provide us with a control over the total variation of the scheme.

To finalize this second approach to the Total Variation Increase due to oscillations we continue by summing the contributions of all the $a_i$'s in the $k$-th time step. This will result in half the Total Variation Increase due to oscillations in the $k$-th time step.

By the previous talk th following corollary is obvious,

**Corollary 4.2.4.1** (Total contribution in th $k$-th step)**.** *In the $k$-th time step we have contribution by $a_1, a_2, \ldots, a_k$, with sum,*

$$I^k_{tot} = \sum_{m=1}^{k} I^k_{a_m} = \lambda \sum_{m=1}^{k} (\lambda + 3\lambda C)^{k-m} a_m \tag{4.11}$$

**Corollary 4.2.4.2** (Result 1)**.** *If we assume that the sequence $a_i$, is bounded i.e there exists $M > 0$ such that $a_i \leq CM$ for all $i = 1, \ldots, \infty$ and that $\lambda + 3\lambda C < 1$ then the total contribution in the $k$-th time step reads,*

$$\text{TVI} \leq \frac{2\lambda C}{1 - (\lambda + 3\lambda C)} M \tag{4.12}$$

*Proof.* Since the increase factors $a_i$ are uniformly bounded as $a_i \leq CM$, their total contribution given in Rel.(4.11) reads

$$I_{tot}^k \leq \lambda CM \sum_{m=1}^{k} (\lambda + 3\lambda C)^{k-m} \overset{n=k-m}{=} \lambda CM \sum_{n=0}^{k-1} (\lambda + 3\lambda C)^n$$
$$= \lambda CM \frac{1 - (\lambda + 3\lambda C)^k}{1 - (\lambda + 3\lambda C)}$$

the previous sequence, in the right hand side, is increasing with respect to the time step $k$, so by taking the limit as $k \to \infty$ we deduce an uniform -with respect to $k$- bound on the total contribution

$$I_{tot}^{\infty} = \lim_{k\to\infty} I_{tot}^k \leq \lambda CM \lim_{k\to\infty} \frac{1 - (\lambda + 3\lambda C)^k}{1 - (\lambda + 3\lambda C)} = \frac{\lambda C}{1 - (\lambda + 3\lambda C)} M$$

This is exactly the result we were looking for since now, the Total Variation Increase due to oscillations is bounded

$$\mathrm{TVI} \leq 2 \cdot I_{tot}^{\infty} \leq \frac{2\lambda C}{1 - (\lambda + 3\lambda C)} M$$

□

**Corollary 4.2.4.3** (Result 2)**.** *If we assume that the sequence $a_i$, $i = 1 \ldots \infty$ is uniformly bounded as $a_i \leq CM = C\mathrm{TV}(u_0)$ then the previous bound on the total variation increase becomes,*

$$2 \cdot I_{tot}^{\infty} \leq \frac{2\lambda C}{1 - (\lambda + 3\lambda C)} \mathrm{TV}(u_0)$$

*Remark* 4.2.10. This result is even better than the previous one given Rel.(4.12) since the bound that provides on the increase of the Total Variation due to oscillations is directly related to the variation of the initial condition $u_0$.

With more delicate assumptions on the increases $a_i$ we get the following Corollary,

**Corollary 4.2.4.4** (Result 3)**.** *If we assume that the sum of all increases $\sum_{i=0}^{\infty} a_i$ is finite $\sum_{i=0}^{\infty} a_i = A < \infty$ then the Total Variation Increase due to oscillations diminishes with respect to the time step $k$.*

*Proof.* We note that the following sums converge

$$\sum_{i=0}^{\infty}(\lambda+3\lambda C)^i=\frac{1}{1-(\lambda+3\lambda C)}<\infty$$

$$\sum_{i=0}^{\infty}a_i$$

Moreover the sum

$$\sum_{k=1}^{\infty}I_{tot}^k=\sum_{k=1}^{\infty}\left(\lambda\sum_{j=1}^{k}(\lambda+3\lambda C)^{k-j}a_j\right)=\lambda\sum_{k=1}^{\infty}\left(\sum_{j=1}^{k}(\lambda+3\lambda C)^{k-j}a_j\right)$$

constitutes the sum of the terms of the Cauchy product of the series $\sum_{i=1}^{\infty}a_i$ and $\sum_{i=1}^{\infty}(\lambda+3\lambda C)^i$. So we deduce that the following sum also converges,

$$\sum_{k=1}^{\infty}I_{tot}^k<\infty$$

and hence the sequence $I_{tot}^k$ must converge to 0, so

$$I_{tot}^{\infty}=\lim_{k\to\infty}I_{tot}^k=0$$

□

*Remark* 4.2.11. This last remark states that if the sum $\sum a_i$ is finite then the overall increase of the variation due to the oscillations produced by the $a_i$'s diminishes with respect to the time step $k$. This result is a sort of asymptotic TVD behaviour.

Moreover we note that the requirement $\sum a_i<\infty$ is supported numerically since the 1-st node on the shock at the right hand side of the 1-st extreme, is always very close to the 1-st extreme. Hence by the definition of the increase factors Def.(4.2.1), $\hat{a}_k=0$ and so $a_k=0$.

We have devised two different bounds concerning the Total Variation Increase due to oscillations. The first was be immediate summation of the magnitudes of the extremes and resulted in the bound Rel.(4.9). The second one came by investigating the contributions of the increase factors and resulted in the bound Rel.(4.12).

### 4.2.5 Comparison of the two bounds

We start this paragraph by restating the two bounds on the Total Variation. The first given in Rel.(4.9)

$$B_1 = 2M\frac{1-\lambda-2\lambda C}{1-\lambda-3\lambda C}$$

and the 2-nd given in Rel.(4.12)

$$B_2 = \frac{2\lambda C}{1-(\lambda+3\lambda C)}M$$

Before the comparison, a comment on the nature of the second bound

*Remark* 4.2.12. The second bound Rel.(4.12) is an increasing function of $\lambda$, so by decreasing $\lambda$ we can decrease the bound on the increase of the total variation due to oscillations.

This remark is also valid but not immediate for the 1-st bound Rel.(4.9).

To compare the two bounds one can examine their ratio, that is the fraction of the bound Rel.(4.12) over the bound of Rel.(4.9),

**Proposition 4.2.1.** *(Comparison of the bounds $B_2$ and $B_1$) If $\lambda+3\lambda C<1$ then $\frac{B_2}{B_1}<1$. If moreover $\lambda+4\lambda C<1$ then $\frac{B_2}{B_1}<\frac{1}{2}$*

*Proof.* The ratio of the bounds $B_2$ and $B_1$ is

$$\frac{B_2}{B_1} = \frac{\frac{2\lambda CM}{1-\lambda-3\lambda C}}{2M\frac{1-\lambda-2\lambda C}{1-\lambda-3\lambda C}} = \frac{\lambda C}{1-\lambda-2\lambda C}$$

If $\lambda+3\lambda C<1$ then $\frac{B_2}{B_1}<1$ since $\lambda C<1-\lambda-2\lambda C$, hence the later bound $B_2$ is sharper. If moreover $\lambda+4\lambda C<1$ instead of $\lambda+3\lambda C<1$ then $\frac{B_2}{B_1}<\frac{1}{2}$ since $2\lambda C<1-\lambda-2\lambda C$. □

The meaning of this proposition is that by a careful selection of the respect factor $\lambda$ the bounds on the increase of the variation in the second approach can be significantly better than that of the first approach.

## 4.3 Numerical tests

In this section we present some numerical experiments that exhibit and support the theory we have established in this Chapter.

As stated at the introduction of this Chapter the numerical schemes that we shall discuss are oscillatory, either due to their dispersive or to their anti-diffusive nature.

Each one of these numerical schemes has already been presented in Chapter 3, where we refer for discussion on their properties. Here we shall only restate their description for non-uniform meshes and prove that they satisfy the Evolution Requirement Req.(1), which we restate

*Requirement* (Evolution requirement). There exists a global constant $C$ that bounds the increase/change of every value from $u_i^n$ to $u_i^{n+1}$ with respect to its neighbours $u_{i-1}^n$ and $u_{i+1}^n$, namely,

$$|u_i^{n+1} - u_i^n| \leq C \max\left\{|u_{i+1}^n - u_i^n|, |u_i^n - u_{i-1}^n|\right\} \tag{4.13}$$

The problems that we shall deal with, are the Transport Equation

$$u_t + u_x = 0, \quad x \in [0,1],\ t \in [0,1]$$

and the inviscid Burgers Equation

$$u_t + \left(\frac{u^2}{2}\right)_x = 0, \quad x \in [0,1],\ t \in [0,1]$$

both with jump initial conditions

$$u_0(x) = \mathcal{X}_{[0,1/2]}(x), \quad x \in [0,1]$$

### 4.3.1 Richtmyer 2-step Lax-Wendroff

This scheme is 2-nd order accurate and of dispersive nature hence it produces oscillations.

In this approach we consider the non-uniform cell centered discretization of the domain in cells

$$C_i = (x_{i-1/2}, x_{i+1/2}) \quad \text{with} \quad |C_i| = h_i$$

The mesh $M_x = \{x_i, i \in \mathbb{Z}\}$ consists of the middle points,

$$x_i = \frac{x_{i+1/2} + x_{i-1/2}}{2} \quad \text{hence} \quad x_i - x_{i-1} = \frac{h_i + h_{i-1}}{2}$$

For this description of the grid we propose the following numerical scheme as the generalisation on non-uniform meshes of the Richtmyer 2-step Lax-Wendorff numerical scheme,

$$\hat{u}_{i+1/2} = \frac{h_{i+1}}{h_i + h_{i+1}} u_i + \frac{h_i}{h_i + h_{i+1}} u_{i+1} - \frac{\Delta t}{h_i + h_{i+1}} (f_{i+1} - f_i)$$
$$u_i^{n+1} = u_i^n - \frac{\Delta t}{h_i} \big( f(\hat{u}_{i+1/2}) - f(\hat{u}_{i-1/2}) \big)$$

or

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{h_i} \big( F_{i+1/2} - F_{i-1/2} \big)$$

with

$$F_{i+1/2} = f(\hat{u}_{i+1/2}) = f\Big( \frac{h_{i+1}}{h_i + h_{i+1}} u_i + \frac{h_i}{h_i + h_{i+1}} u_{i+1} - \frac{\Delta t}{h_i + h_{i+1}} (f_{i+1} - f_i) \Big)$$

We need to bound the difference

$$\begin{aligned} |u_i^{n+1} - u_i^n| &\leq \frac{\Delta t}{h_i} \left| f(\hat{u}_{i+1/2}) - f(\hat{u}_{i-1/2}) \right| \\ &\leq \frac{\Delta t}{h_i} \max |f'| |\hat{u}_{i+1/2} - \hat{u}_{i-1/2}| \\ &\leq CFL |\hat{u}_{i+1/2} - \hat{u}_{i-1/2}| \end{aligned}$$

and now the difference

$$\begin{aligned} |\hat{u}_{i+1/2} - \hat{u}_{i-1/2}| = \Bigg| & \frac{h_{i+1}}{h_i + h_{i+1}} u_i + \frac{h_i}{h_i + h_{i+1}} u_{i+1} - \frac{\Delta t}{h_i + h_{i+1}} \left( f(u_{i+1}) - f(u_i) \right) \\ & - \frac{h_i}{h_{i-1} + h_i} u_{i-1} - \frac{h_{i-1}}{h_{i-1} + h_i} u_i + \frac{\Delta t}{h_{i-1} + h_i} \left( f(u_i) - f(u_{i-1}) \right) \Bigg| \end{aligned}$$

by setting $\frac{h_{i+1}}{h_i + h_{i+1}} = \mu_1$ and $\frac{h_i}{h_{i-1} + h_i} u_{i-1} = \mu_2$ the previous bound recasts to

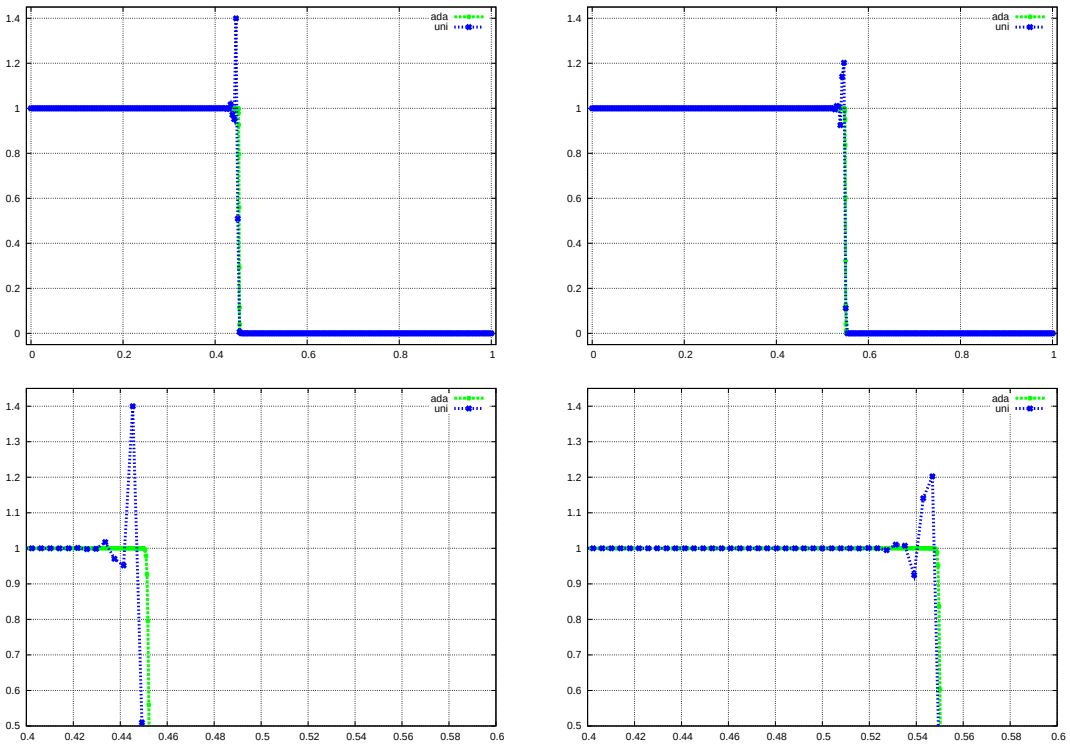



Figure 4.7: Two instances of the numerical solutions of the inviscid Burgers equation under the two step Richtmyer Lax Wendroff. The uniform mesh case exhibits oscillations where as the adaptive case is almost clear. The details of these graphs are uni: CFL=0.9 timestep$\approx 10^{-3}$ and ada: CFL=0.1 timestep$\approx 3 \cdot 10^{-5}$. The target time of this tets is 0.907

$$
\begin{aligned}
|\hat{u}_{i+1/2} - \hat{u}_{i-1/2}| = \Big| & \mu_1 u_i + (1-\mu_1)u_{i+1} - \frac{\Delta t}{h_i + h_{i+1}} \left(f(u_{i+1}) - f(u_i)\right) \\
& - \mu_2 u_{i-1} - (1-\mu_2)u_i + \frac{\Delta t}{h_{i-1} + h_i} \left(f(u_i) - f(u_{i-1})\right) \Big| \\
\leq\ & \mu_1 |u_i - u_{i+1}| + \mu_2 |u_i - u_{i-1}| + |u_{i+1} - u_i| \\
& + \frac{\Delta t}{2 \min h_i} \max |f'| |u_{i+1} - u_i| + \frac{\Delta t}{2 \min h_i} \max |f'| |u_i - u_{i-1}| \\
\leq\ & (1 + \mu_1 + \mu_2 + CFL) \max\{|u_{i-1} - u_i|, |u_i - u_{i+1}|\} \\
\leq\ & (3 + CFL) \max\{|u_{i-1} - u_i|, |u_i - u_{i+1}|\}
\end{aligned}
$$

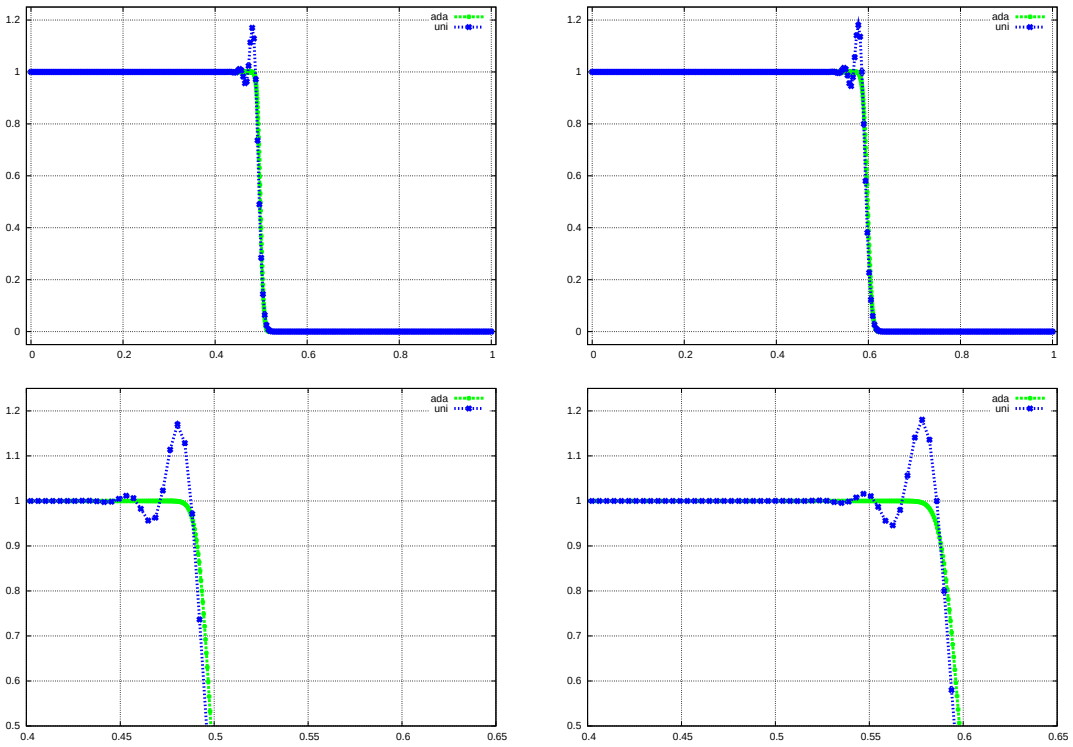



Figure 4.8: Transport equation with $a = 1$, under the two step Richtmyer Lax Wendroff. Again oscillations appear in the uniform mesh case, whereas the adaptive is almost clear. The details are uni:CFL=0.9, timestep$\approx 2 \cdot 10^{-3}$, ada: CFL 0.5, timestep $2 \cdot 10^{-4}$. The target time is 0.907

where the last inequality is valid since $0 < \mu_1, \mu_2 \leq 1$. So the overall bound reads,

$$\begin{aligned} |u_i^{n+1} - u_i^n| &\leq CFL(3 + CFL) \max\left\{|u_{i+1}^n - u_i^n|, |u_i^n - u_{i-1}^n|\right\} \\ &\leq 4CFL \max\left\{|u_{i+1}^n - u_i^n|, |u_i^n - u_{i-1}^n|\right\} \end{aligned}$$

The constant $C$ in this case is chosen to be $C = 4CFL$, for this choice the evolution requirement is satisfied.

### 4.3.2 MacCormack

In this approach we consider the non-uniform cell centered discretization of the domain in cells

$$C_i = (x_{i-1/2}, x_{i+1/2}) \quad \text{with} \quad |C_i| = h_i$$

The mesh $M_x = \{x_i, i \in \mathbb{Z}\}$ consists of the middle points,

$$x_i = \frac{x_{i+1/2} + x_{i-1/2}}{2} \quad \text{hence} \quad x_i - x_{i-1} = \frac{h_i + h_{i-1}}{2}$$

For this description of the grid we propose the following scheme as the generalisation of the MacCormack,

$$\begin{aligned} u_i^* &= u_i^n - \frac{2\Delta t}{h_i + h_{i+1}}(f(u_{i+1}) - f(u_i)) \\ u_i^{**} &= u_i^* - \frac{2\Delta t}{h_{i-1} + h_i}(f(u_i^*) - f(u_{i-1}^*)) \\ u_i^{n+1} &= \frac{u_i + u_i^{**}}{2} \end{aligned}$$

We first rewrite the scheme in the following form, for $f_i^* = f(u_i^*)$ and $f_i = f(u_i^n)$

$$\begin{aligned} u_i^{n+1} &= \frac{u_i^n}{2} + \frac{u_i^* - \frac{2\Delta t}{h_{i-1}+h_i}(f_i^* - f_{i-1}^*)}{2} \\ &= \frac{u_i^n}{2} + \frac{u_i^n - \frac{2\Delta t}{h_i+h_{i+1}}(f_{i+1} - f_i) - \frac{2\Delta t}{h_{i-1}+h_i}(f_i^* - f_{i-1}^*)}{2} \\ &= u_i^n + -\frac{\Delta t}{h_i + h_{i+1}}(f_{i+1} - f_i) - \frac{\Delta t}{h_{i-1} + h_i}(f_i^* - f_{i-1}^*) \end{aligned}$$

So, to prove the Evolution Requirement for this scheme we need to bound

$$|u_i^{n+1} - u_i^n| = | - \frac{\Delta t}{h_i + h_{i+1}}(f_{i+1} - f_i) - \frac{\Delta t}{h_{i-1} + h_i}(f_i^* - f_{i-1}^*)|$$

which reads,

$$\begin{aligned} |u_i^{n+1} - u_i^n| &\leq \frac{CFL}{2}\left(|u_{i+1}^n - u_i^n| + |u_i^* - u_{i-1}^*|\right) \\ &\leq \frac{CFL}{2}\left(|u_{i+1}^n - u_i^n| + |u_i^n - u_{i-1}^n| + CFL|u_i^n - u_{i-1}^n| + CFL|u_{i+1}^n - u_i^n|\right) \\ &\leq CFL(1 + CFL)\max\left\{|u_{i+1}^n - u_i^n|, |u_i^n - u_{i-1}^n|\right\} \\ &\leq 2CFL\max\left\{|u_{i+1}^n - u_i^n|, |u_i^n - u_{i-1}^n|\right\} \end{aligned}$$

Where the last inequality results by the usual assumption $CFL \leq 1$. So, the constant $C$ in this case is chosen to be $C = 2CFL$ and for this choice the Evolution Requirement is satisfied.

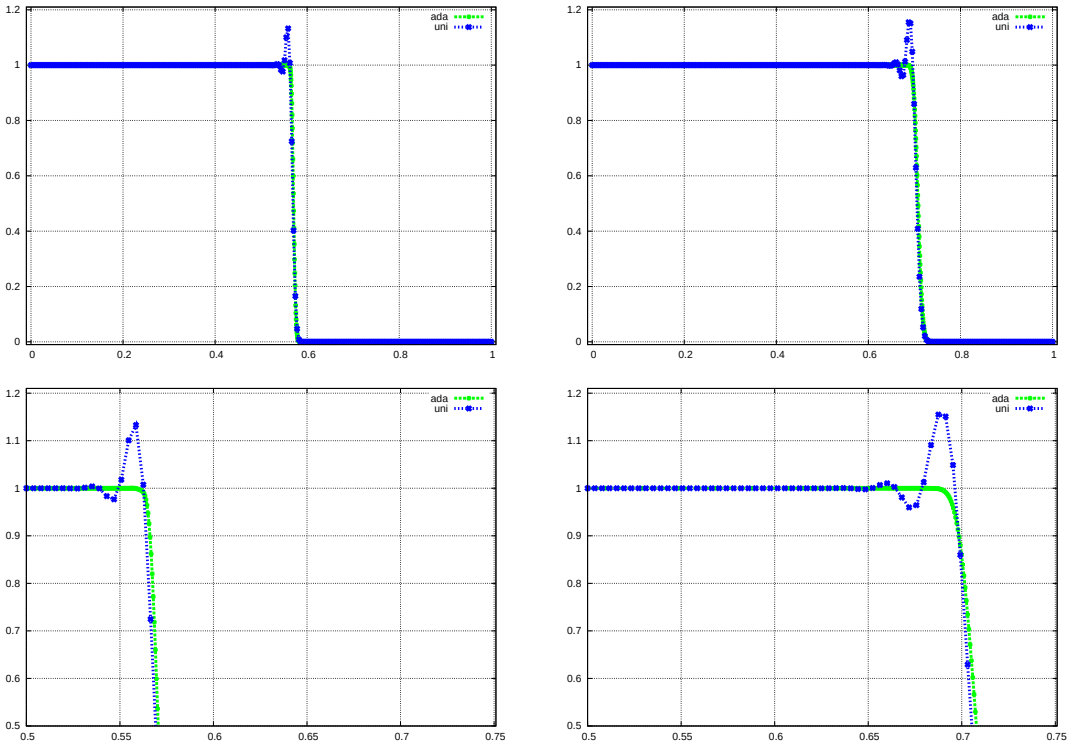



Figure 4.9: Transport equation with $a = 1$, under the MacCormack. Again oscillations are apparent in the uniform case whereas the adaptive is almost clean. The details of this test are uni:CFL=0.8 and timestep$\approx 3 \cdot 10^{-3}$ ada:CFL=0.5 timestep$\approx 4 \cdot 10^{-4}$

### 4.3.3 Noelle - Pure 2-nd order

In this approach we consider the non-uniform mesh

$$M_x = \{x_i, i \in \mathbb{Z}\} \quad \text{with} \quad h_i = x_i - x_{i-1}$$

The middle points $x_{i-1/2} = \frac{x_{i-1}+x_i}{2}$ define a partition of the domain in cells

$$C_i = (x_{i-1/2}, x_{i+1/2}) \quad \text{with} \quad |C_i| = \frac{h_i + h_{i+1}}{2}$$

The scheme reads for the linear case $f(u) = u$,

$$u_i^{n+1} = \frac{\Delta t(\Delta t + h_{i+1})}{h_i(h_i + h_{i+1})} u_{i-1}^n + \frac{(-\Delta t + h_i)(\Delta t + h_{i+1})}{h_i h_{i+1}} u_i^n + \frac{\Delta t(\Delta t - h_i)}{(h_i + h_{i+1})h_{i+1}} u_{i+1}^n$$

By naming the coefficients of $u_{i-1}$ and $u_{i+1}$ as $\alpha$ and $\gamma$ respectively, that is

$$\alpha = \frac{\Delta t(\Delta t + h_{i+1})}{h_i(h_i + h_{i+1})}$$
$$\gamma = \frac{\Delta t(\Delta t - h_i)}{(h_i + h_{i+1})h_{i+1}}$$

and noting that the coefficient of $u_i^n$ can be written as $1 - \alpha - \gamma$, the scheme recast into,

$$u_i^{n+1} = \alpha u_{i-1} + (1 - \alpha - \gamma)u_i + \gamma u_{i+1}$$

so,

$$u_i^{n+1} - u_i^n = \alpha(u_{i-1} - u_i) + \gamma(u_{i+1} - u_i)$$

hence,

$$|u_i^{n+1} - u_i^n| \leq (|\alpha| + |\gamma|) \max\{|u_{i-1} - u_i|, |u_{i+1} - u_i|\}$$

Now, by the definitions of $\alpha$ and $\gamma$ we can bound $|\alpha| + |\gamma|$ as follows,

$$\begin{aligned}
|\alpha| + |\gamma| &= \frac{\Delta t^2 + \Delta t h_{i+1}}{h_i(h_i + h_{i+1})} + \frac{\Delta t h_i - \Delta t^2}{h_{i+1}(h_i + h_{i+1})} \\
&\leq \frac{\Delta t}{h_i}\frac{\Delta t}{h_i + h_{i+1}} + \frac{\Delta t}{h_i}\frac{h_{i+1}}{h_i + h_{i+1}} + \frac{\Delta t}{h_{i+1}}\frac{h_i}{h_i + h_{i+1}} \\
&\leq CFL \cdot \frac{CFL}{2} + CFL \cdot 1 + CFL \cdot 1 = CFL(\frac{CFL}{2} + 2) \leq \frac{5}{2}CFL
\end{aligned}$$

Finally,

$$|u_i^{n+1} - u_i^n| = \frac{5}{2}CFL \max\{|u_{i-1} - u_i|, |u_{i+1} - u_i|\}$$

So, the constant $C$ in this case is chosen to be $C = \frac{5}{2}CFL$ and for this choice the Evolution Requirement is satisfied.

### 4.3.4 Unstable Centered - FTCS

This scheme produces oscillations due to its anti-diffusive nature, which property is also responsible for the instability of the scheme.

In this approach we consider the non-uniform mesh

$$M_x = \{x_i, i \in \mathbb{Z}\} \quad \text{with } h_i = x_i - x_{i-1}$$

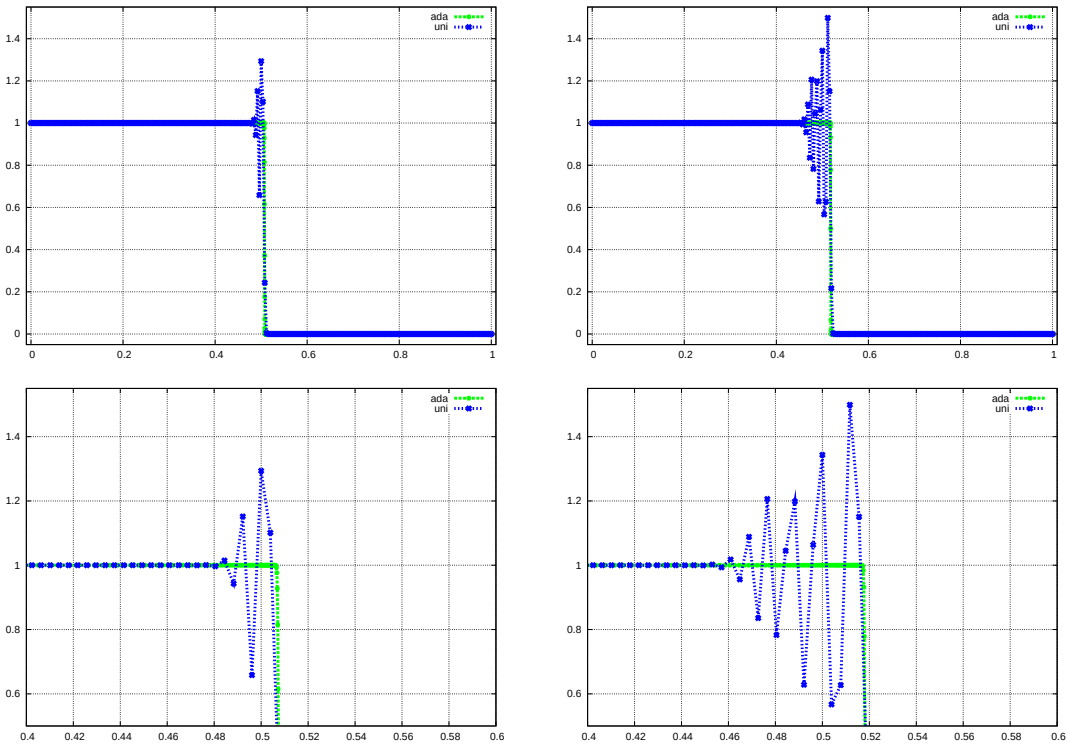



Figure 4.10: Inviscid Burgers equation, under the unstable FTCS. The oscillations in the uniform case are worse than with the previous dispersive schemes. This is because of anti-diffusive nature of FTCS. Interestingly the adaptive case is almost clean. The details of this test are uni:CFL=0.1 and timestep$\approx 2 \cdot 10^{-4}$ ada:CFL=0.1 timestep$\approx 2 \cdot 10^{-5}$

The middle points $x_{i-1/2} = \frac{x_{i-1}+x_i}{2}$ define a partition of the domain in cells,

$$C_i = (x_{i-1/2}, x_{i+1/2}) \quad \text{with} \quad |C_i| = \frac{h_i + h_{i+1}}{2}$$

For this description of the grid we discuss the known to be *unstable* Forward in Time Centered in Space (FTCS) scheme

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{h_i + h_{i+1}} \big( f(u_{i+1}) - f(u_{i-1}) \big)$$

This scheme can be written in conservative form as follows,

$$u_i^{n+1} = u_i^n - \frac{2\Delta t}{h_i + h_{i+1}} (F_{i+1/2} - F_{i-1/2})$$

with

$$F_{i+1/2} = \frac{f(u_i) + f(u_{i+1})}{2}$$

Easily we deduce that

$$\begin{aligned} |u_i^{n+1} - u_i^n| &\leq \frac{\Delta t}{2 \min h_i} \max |f'| |u_{i+1}^n - u_{i-1}^n| \leq \frac{CFL}{2} \left( |u_{i+1}^n - u_i^n| + |u_i^n - u_{i-1}^n| \right) \\ &\leq CFL \max \left\{ |u_{i+1}^n - u_i^n|, |u_i^n - u_{i-1}^n| \right\} \end{aligned}$$

The constant $C$ in this case is chosen to be $C = CFL$, for this choice the Evolution requirement is satisfied.

# Appendix A

# Curvature for plane curves

Let a smooth plane curve be $C(t) = (x(t), y(t))$. Its curvature $\kappa$ is given by

$$\kappa = \frac{d\phi}{ds}, \tag{A.1}$$

where $s$ is the arclength and $\phi$ is the tangential angle. Expanding more on (A.1) one can write

$$\kappa = \frac{\frac{d\phi}{dt}}{\frac{ds}{dt}} = \frac{\frac{d\phi}{dt}}{\sqrt{x'^{\,2} + y'^{\,2}}} \tag{A.2}$$

where the last equality is justified by the definition of the arclength, i.e

$$s(t) = \int_0^t \sqrt{x'(\tau)^2 + y'(\tau)^2} d\tau$$

Moreover one evaluates,

$$\tan\phi = \frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{y'}{x'} \tag{A.3}$$

so,

$$\frac{d\tan\phi}{dt} = \frac{y''x' - y'x''}{x'^{\,2}} \tag{A.4}$$

hence,

$$\frac{d\tan\phi}{dt} = \frac{1}{1+\tan^2\phi}\frac{d\phi}{dt} \tag{A.5}$$

Now, by combining equations (A.2), (A.4), (A.5) we end with the following expression on the curvature of plane curves

$$\kappa = \frac{|x'y'' - x''y'|}{(x'^2 + y'^2)^{3/2}} \tag{A.6}$$

which for the case of the curve $(x, f(x))$ defined by the function $f$ can be written as

$$\kappa = \frac{|f''(x)|}{(1+f'(x)^2)^{3/2}}$$

# Appendix B

# Power Series Expansion

**Lemma B.0.1.** *For $|t| < 1$ the following power series expansion holds*

$$\sum_{\nu=0}^{\infty} \binom{\nu+m-1}{\nu} t^{\nu} = \frac{1}{(1-t)^m}$$

*Proof.* We prove this lemma by means of induction

($m = 1$) We start with the obvious remark that the previous expansion is valid for $m = 1$, since

$$\frac{1}{1-t} = (1 + t + t^2 + \cdots) = \sum_{\nu=0}^{\infty} t^k = \sum_{\nu=0}^{\infty} \binom{\nu+1-1}{\nu} t^k$$

($m = k$) We continue by making the induction hypothesis that the expansion is valid for $m = k$, i.e

$$\sum_{\nu=0}^{\infty} \binom{\nu+k-1}{\nu} t^{\nu} = \frac{1}{(1-t)^k}$$

($m = k+1$) We shall prove that the induction is valid for $m = k+1$, that is

$$\sum_{\nu=0}^{\infty} \binom{\nu+k}{\nu} t^{\nu} = \frac{1}{(1-t)^{k+1}}$$

By the induction hypothesis the power series expansion of $\frac{1}{1-t}$ and $\frac{1}{(1-t)^k}$ converge, and hence can be multiplied as follows

$$\frac{1}{1-t} \cdot \frac{1}{(1-t)^k} = \sum_{\nu_1=0}^{\infty} t^{\nu_1} \cdot \sum_{\nu_2=0}^{\infty} \binom{\nu_2+k-1}{\nu} t^{\nu_2} = \sum_{\nu_1=0}^{\infty} \sum_{\nu_2=0}^{\infty} \left( \binom{\nu_1+k-1}{\nu_1} \cdot 1 \right) t^{\nu_1+\nu_2}$$

where the last equality is valid by the product of convergent power series. The last term can be written, for $\nu = \nu_1 + \nu_2$ as follows

$$\sum_{\nu=0}^{\infty} \left( \sum_{\nu_1=0}^{\nu} \binom{\nu_1 + k - 1}{\nu_1} \right) t^{\nu}$$

which now recasts since $\binom{n}{k} = \binom{n}{n-k}$ for every $n \geq k$ as follows

$$\sum_{\nu=0}^{\infty} \left( \sum_{\nu_1=0}^{\nu} \binom{\nu_1 + k - 1}{k - 1} \right) t^{\nu}$$

or by setting $j = \nu_1 + k - 1$

$$\sum_{\nu=0}^{\infty} \left( \sum_{j=k-1}^{\nu+k-1} \binom{j}{k-1} \right) t^{\nu}$$

which recasts since $\sum_{j=k}^{n} \binom{j}{k} = \binom{n+1}{k+1}$ for every $n \geq k$ as follows

$$\sum_{\nu=0}^{\infty} \binom{\nu + k}{k} t^{\nu}$$

or by utilising again the identity $\binom{n}{k} = \binom{n}{n-k}$,

$$\sum_{\nu=0}^{\infty} \binom{\nu + k}{\nu} t^{\nu}$$

Which completes the proof of the lemma. □

# Bibliography


[1] A. I. Arvanitis, Ch.and Delis. Behavior of finite volume schemes for hyperbolic conservation laws on adaptive redistributed spatial grids. *SIAM J. Sci. Comput.*, 28:1927–1956, 2006.

[2] Ch. Arvanitis. *Finite Elements for Hyperbolic Conservation Laws: New methods and computational techniques*. PhD thesis, University of Crete, 2002.

[3] Ch. Arvanitis, Th. Katsaounis, and Ch. Makridakis. Adaptive finite element relaxation schemes for hyperbolic conservation laws. *Math. Model. Anal. Numer.*, 35:17–33, 2001.

[4] Ch. Arvanitis, Ch. Makridakis, and N. Sfakianakis. Entropy conservative schemes and adaptive mesh selection for hyperbolic conservation laws. *preprint*, 2006.

[5] Ch. Arvanitis, Ch. Makridakis, and A. Tzavaras. Stability and convergence of a class of finite element schemes for hyperbolic systems of conservation laws. *SIAM J. Numer. Anal.*, 42:1357–1393, 2004.

[6] Chr. Arvanitis. Mesh redistribution strategies and finite element method schemes for hyperbolic conservation laws. *J. Sci. Computing*, 34:1–25, 2008.

[7] R. Courant and K. Friedrichs. *Supersonic Flow and Shock Waves*. Springer, 1948.

[8] R. Courant, K. Friedrichs, and H. Lewy. On the partial difference equations of mathematical physics. *IBM Journal*, pages 215–234, 1967. English translation of the 1928 German original.

[9] E. Dorfi and L. Drury. Simple adaptive grids for 1d initial value problems. *J. Computational Physics*, 69:175–195, 1987.

[10] B. Fornberg. Generation of finite difference formulas on arbitrary spaced grids. *Mathematics of Computations*, 51:699–706, 1988.

[11] E. Godlewski and P. A. Raviart. *Hyperbolic Systems of Conservation Laws*. Ellipses, 1990.

[12] A. Harten and J. Hyman. Self adjusting grid methods for one-dimensional hyperbolic conservation laws. *J. Comput. Physics*, 50:235–269, 1983.

[13] G. W. Hedstrom. Models of difference schemes for $u_t + u_x = 0$ by partial differential equations. *J. Comput. Physics*, 50:235–269, 1983.

[14] C. W. Hirt. Heuristic stability theory for finite difference schemes. *J. Comput. Physics*, 2:339–355, 1968.

[15] D. Kroener. *Numerical schemes for Conservation Laws*. Wiley Teubner, 1997.

[16] P. D. Lax. Weak solutions of nonlinear hyperbolic equations and their numerical computation. *Comm. pure and applied mathematics*, 7:159–193, 1954.

[17] P. D. Lax and R. Richtmyer. Survey of the stability of linear finite difference equations. *Comm. Pure Appl. Math.*, 9:267–293, 1956.

[18] P. D. Lax and B. Wendroff. Systems of conservation laws. *Comm. Pure Appl. Math.*, 13:217–237, 1960.

[19] Ph. LeFloch and Ch. Rohde. High-order schemes, entropy inequalities and nonclassical shocks. *J. Numerical Analysis SIAM*, pages 2023–2060, 2000.

[20] R. LeVeque. *Numerical methods for Conservation Laws*. Birkhauser Verlag, second edition, 1992.

[21] R. LeVeque. *Finite volume methods for hyperbolic problems*. Cambridge texts in applied mathematics, first edition, 2002.

[22] Ch. Makridakis, Noelle S., and N. Sfakianakis. Adaptive mesh reconstruction and 2-nd order schemes. *preprint*, 2008.

[23] O. Oleinik. Discontinuous solutions of non-linear differential equations. *AMS Transactions Series 2*, pages 95–172, 1963.

[24] N. Sfakianakis. Adaptive mesh reconstruction, total variation bound. *preprint*, 2008.

[25] J. Smoller. *Shock Waves and Reaction-Diffusion Equations*. Springer-Verlag, second edition, 1991.

[26] E. Tadmor. The numerical viscosity of entropy stable schemes for systems of conservation laws. *Mathematics of Computations*, pages 91–103, 1987.

[27] E. Tadmor. Entropy stability theory for difference approximations of nonlinear conservation laws and related time dependent problems. *Acta Numerica*, pages 451–512, 2003.

[28] H. Tang and T. Tang. Adaptive mesh methods for one- and two-dimensional hyperbolic conservation laws. *SIAM J. Numerical Analysis*, 41:487–515, 2003.

[29] J.W Thomas. *Numerical partial differential equations - Finite difference methods*. Springer, 1995.

[30] J.W Thomas. *Numerical partial differential equations - Conservation laws and elliptic equations*. Springer, 1999.

[31] R. F. Warming and B. J. Hyett. The modified equation approach to the stability and accuracy of finite difference methods. *J. Comp. Physics*, 14:159–179, 1974.