eman ta zabal zazu

Universidad del País Vasco
Euskal Herriko Unibertsitatea

Large-time behavior of some numerical schemes: application to the sonic-boom phenomenon

Doktorego Tesia / Tesis Doctoral

Egilea / Autor:
Alejandro Pozo Pazos

Zuzendaria / Director:
Enrique Zuazua Iriondo

2014

Doctoral Thesis

Large-time behavior of some numerical schemes: application to the sonic-boom phenomenon

---

Doktorego Tesia

Zenbait eskema numerikoren konportaera denboraldi luzeetan: soinu-eztandaren fenomenorako erabilera

---

Tesis Doctoral

Comportamiento en tiempos grandes de algunos esquemas numéricos: aplicación al fenómeno de la explosión sónica

---

Author / Egilea / Autor: Alejandro Pozo Pazos
Advisor / Zuzendaria / Director: Enrique Zuazua Iriondo

Bilbao, 2014

Acknowledgements

This thesis is supported by the Basque Government (PREDOC Program 2012, 2013 and 2014). My research has also been covered by the Advanced Grant NUMERIWAVES / FP7-246775 grant of the European Research Council Executive Agency and the MTM2011-29306-C02-00 grant of the Spanish Ministry of Economy and Competitiveness MINECO. Also by the Basque Government through the BERC 2014-2017 program and by the Spanish Ministry of Economy and Competitiveness MINECO: BCAM Severo Ochoa excellence accreditation SEV-2013- 0323.

Abstract

In this thesis we highlight the necessity of employing numerical schemes that preserve the large-time dynamical properties of the continuous system. We focus on Burgers-like equations, which are well known to develop N-waves as intermediate or asymptotic profiles. As we show, not only forward simulations can be distorted; the efficiency of solutions of optimization and inverse design problems might be affected too. In particular, we apply our results to the case of the prediction and control of the sonic-boom phenomenon, modeled by the augmented Burgers equation.

Laburpena

Tesi honetan sistema jarraiaren denbora luzeko ezaugarri dinamikoak mantentzen duten eskema numerikoen erabileraren beharra azpimarratzen dugu. Burgers erakoak diren ekuazioetan jartzen dugu arreta. Izan ere, hauetan "N" itxura duten bitarteko ebazpenak edo profil asintotikoak agertzen dira. Erakusten dugun moduan, aurrerantzako simulazioak ez ezik, optimizazio eta alderantzizko diseinu arazoetarako ebazpenenen eraginkortasuna ere kaltetu daiteke. Gure emaitzak bereziki Burgers ekuazio handituak modelatzen duen soinu-eztandaren fenomenoaren iragarpenaren eta kontrolaren kasuan erabiltzen ditugu.

Resumen

En esta tesis recalcamos la necesidad de emplear esquemas numéricos que preserven las propiedades dinámicas en tiempos grandes del sistema continuo. Nos centramos en ecuaciones del tipo Burgers, que son conocidas por desarrollar perfiles con forma de "N" como soluciones intermedias o asintóticas. Tal y como mostramos, no solo las simulaciones hacia adelante pueden ser distorsionadas; la eficacia de las soluciones para problemas de optimización y diseño inverso también puede verse afectada. En particular, aplicamos nuestros resultados al caso de la predicción y control del fenómeno de la explosión sónica, modelada por la ecuación de Burgers aumentada.

Contents

List of Figures

Chapter 1

Introduction

Nowadays, one of the major concerns in aeronautics research is related to the control and reduction of the noise generated by aircrafts. Indeed, one of the main goals in this broad and important area of industrial and commercial activity is to meet stringent noise and design constraints for supersonic airplanes. In particular, the minimization of the sonic-boom generated by these aircrafts is the key point for succeeding in the development of efficient civilian supersonic transports [3].

As we shall see, this type of issues requires numerical tools that can handle with evolution problems in large-time horizons. In this thesis, we highlight the necessity of employing numerical approximation schemes that preserve the large-time dynamical properties of the continuous system. We apply this to the particular case of the prediction and control of the sonic-boom.

1.1 Sonic-boom minimization

When flying above the speed of sound, supersonic airplanes create pressure disturbances in the atmosphere, resulting both from the displacement of the air when passing by and from the generation of the aerodynamic lift, necessary to follow the desired path. Many of these disturbances, which imply a significant amount of acoustic energy, reach the ground level, resulting on the so-called sonic-boom. This phenomenon is perceived on the ground as two subsequent loud bangs with a short time lapse in between. This can turn out to be harmful for humans and building structures and has been one of the main obstacles when it comes to designing supersonic aircrafts.

The first manned airplane that surpassed the speed of sound was the Bell X-1, in 1947 [24, Chapter 10]. Afterwards, in the 1960s and 1970s efforts were focused on building this kind of transport for civilian uses. At the beginning, two projects were able to succeed: the Concorde, by the French and English, and the Tupolev-144, by the Soviets. Two

other American projects, the Lockheed L-2000 and the Boeing 2707, were canceled due to the restrictions imposed by the US Congress regarding environmental and acoustical reasons. This prohibition did not allow to reach supersonic speeds over populated areas (in practice, everywhere but the oceans), so all economic prospects were destroyed. In the end, all civilian supersonic transport projects were condemned to failure.

Therefore, throughout the last decades of the 20th century, it became clear that sonic-boom-reduction technologies had to be applied necessarily, in order to achieve the goal of building efficient aircrafts that were able to deal with this noise issue. In fact, the DARPA/NASA/Northrop-Grumman Shaped Sonic Boom Demonstrator (SSBD) project [79] showed that the sonic-boom could be partially mitigated by tailoring the shape of the aircraft, confirming it experimentally for the first time.

It was well known that the sound signature in the near field evolves into an N-wave (see Figure 1.1) if the time of propagation is long enough. The N-wave refers to the shape that results from the collapse of the multiple shocks into a leading and a trailing shock separated by a nearly linear pressure expansion. At the beginning, it was thought that this N-wave was unavoidable. But McLean proved that it was possible to tailor the near field signature so that the N-wave was not achieved by the time the signature had reached the ground level [74]. Nevertheless, as mentioned above, this was not empirically verified until the tests done within the SSBD project in 2003.

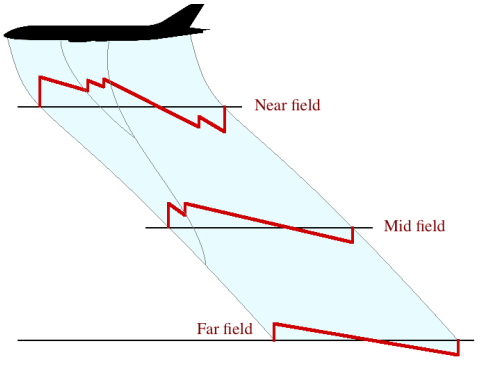


FIGURE 1.1: Diagram of the propagation of the sonic-boom. Shocks created in the near field collapse into two single shocks, forming an N-wave in the far-field.

Historically, linear theory was used to model the sonic-boom propagation, following the works by Hayes [43] and Whitham [93] in the late 1940s and early 1950s. Most of the analytical and numerical research was done assuming that simplification. More recently, new nonlinear physical models have been developed to improve the existing propagation methods. Two such methods are worth to mention. On the one hand, Ozcer [78] proposed the use of the full potential equation in the region between the near field and the ground. On the other hand, following the work by Cleveland [20], Rallabhandi used

an augmented Burgers equation, within the context of ray-tracing/geometrical acoustics to propagate the source signatures to the ground, including viscous effects that lead to nonzero thickness shock discontinuities [85].

In this thesis we will focus on the latter. The augmented Burgers (ABE) equation takes into account nonlinear effects, such as waveform steepening and variable-speed wave propagation, as well as molecular relaxation phenomena, ray tube spreading and atmospheric stratification. The terminology is not consistent; depending on the source, it can be found also as extended Burgers equation [20]. The equation is given by:

$$\frac{\partial P}{\partial \sigma} = P\frac{\partial P}{\partial \tau} + \frac{1}{\Gamma}\frac{\partial^2 P}{\partial \tau^2} + \sum_{\nu} C_\nu \frac{\frac{\partial^2 P}{\partial \tau^2}}{1+\theta_\nu \frac{\partial}{\partial \tau}} - \frac{1}{2G}\frac{\partial G}{\partial \sigma}P + \frac{1}{2\rho_0 c_0}\frac{\partial(\rho_0 c_0)}{\partial \sigma}P, \tag{1.1}$$

were $P = P(\sigma, \tau)$ is the dimensionless perturbation of the pressure distribution. The distance of the propagation σ and time of the perturbation τ are also dimensionless. The operator appearing in the summation corresponding to the molecular relaxations is defined as follows:

$$\frac{\theta_\nu}{1+\theta_\nu \frac{\partial}{\partial \tau}} f(\tau) = \int_{-\infty}^{\tau} e^{(\xi-\tau)/\tau_\nu} f(\xi) d\xi.$$

Function G denotes the ray-tube area. The atmosphere conditions are given by its density ρ_0, the ambient speed of sound c_0, a thermo-viscous parameter Γ and a dimensionless relaxation time θ_ν and dimensionless dispersion parameter C_ν for each relaxation mode (typically, one corresponding to the relaxation produced by oxygen and another one for nitrogen).

In that case, the sonic-boom minimization problem consists on, given a desired ground signature P^* and the duration of the propagation Σ –which is closely related to the altitude of the flight–, recovering the near-field signature that better reproduces it. This is commonly formulated from the optimal control point of view, through a least square approach [86] using the following functional:

$$\mathcal{S}(P_0) = \frac{1}{2}\int_{\mathbb{R}} \big(P(\Sigma, \tau) - P^*(\tau)\big)^2 d\tau. \tag{1.2}$$

Here P_0 lies in a set of admissible near-field signatures –which is usually constrained by the design variables associated to the geometry of the aircraft [76, 86]– and P is the solution of (1.1) with $P(0, \tau) = P_0(\tau)$ for all $\tau \in \mathbb{R}$. The perceived loudness (PLdB) or the shock over-pressures [76] are other functionals that are usually considered for minimization too.

In this sonic-boom minimization problem the time scales are very different: the perturbation of the pressure takes place for less than half a second, while the propagation can last up to a minute, depending on the flight conditions [85]. As we shall see, this

makes the computational treatment a hard task and motivates the study of the large-time asymptotic behavior of the numerical schemes.

1.2 Mathematical framework

In this section we gather some of the mathematical tools that appear along this thesis. Since the equation for the sonic boom is a Burgers-like equation, we first gather some properties concerning the behavior of solutions to the Burgers equations. Then, this equation being a scalar conservation law, we recall the classical way of discretizing this type of equations through finite volumes. We conclude with a brief introduction to line search methods and, in particular, to gradient descent methods, widely used for minimizing functionals.

1.2.1 Burgers equation and its large-time dynamics

The augmented Burgers equation (1.1) is one of the many variations of the Burgers equation. Together with linear theory, Burgers-type equations have been one of the main tools to model the propagation of finite-amplitude plane waves. The classical viscous Burgers equation was first considered for wave propagation in a lossy medium. Successive generalizations included other effects such as geometrical spreading and inhomogeneous mediums (generalized Burgers equation [16, 31, 69]) or relaxation processes (augmented Burgers equation [80]). As explained in the previous section, the augmented Burgers equation has been recently used to model the propagation of the sonic-boom produced by supersonic aircrafts [85], taking into account all those phenomena mentioned above.

The inviscid version of the Burgers equation is a scalar conservation law, given by

$$u_t + \left(\frac{u^2}{2}\right)_x = 0, \quad x \in \mathbb{R}, t > 0, \tag{1.3}$$

while the viscous version is a one-dimensional convection-diffusion equation:

$$u^\nu_t + \left(\frac{(u^\nu)^2}{2}\right)_x = \nu u^\nu_{xx}, \quad x \in \mathbb{R}, t > 0, \tag{1.4}$$

where $\nu > 0$. This last equation was introduced by Bateman [8] and Burgers [15], who first considered it as a simplification of the Navier-Stokes equations. It is the simplest model that gathers nonlinear propagation and diffusive effects [94]. As a matter of fact, Hopf [45] and Cole [22] found that the general solution could be obtained explicitly, through the so-called Hopf-Cole transformation.

It is well known that the asymptotic behavior of the solutions of the inviscid Burgers equation (1.3) is of self-similar nature (see [72] and the references therein). Indeed, as $t \to \infty$ solutions develop an N-wave behavior, conserving the mass of the initial datum that is invariant under the evolution. Note however that the mass does not suffice to identify the asymptotic self-similar profile [73], which belongs to a two-parameter family of solutions. These parameters correspond to two invariants of the system, namely the positive and the negative masses:

$$p = -\min_{x\in\mathbb{R}} \int_{-\infty}^{x} u^0(z)dz \quad \text{and} \quad q = \max_{x\in\mathbb{R}} \int_{x}^{\infty} u^0(z)dz. \tag{1.5}$$

In particular, generically, the N-wave corresponding to solutions emanating from changing sign initial data changes sign.

The asymptotic behavior differs significantly for the viscous version (1.4). Indeed, for $\nu > 0$ these problems are of parabolic nature and, as $t \to \infty$, the solutions behave in a self-similar way as a viscous profile of constant sign that is fully determined by the conserved mass (see [45]):

$$M = q - p = \int_{\mathbb{R}} u^0(x)dx.$$

Figure 1.2 compares the two asymptotic profiles for the same parameters.

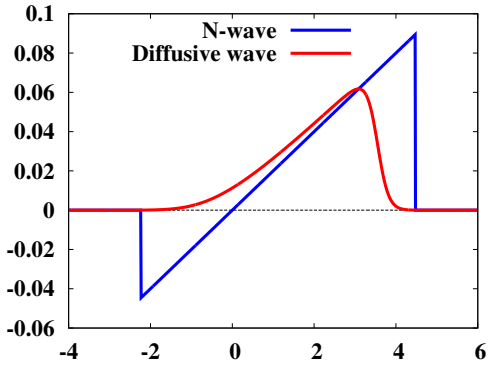


FIGURE 1.2: Asymptotic profiles for the parabolic $\nu > 0$ (red) and hyperbolic $\nu = 0$ (blue) cases. Here the total mass is $M = 0.15$ and the negative and positive masses, $p = 0.05$ and $q = 0.2$ respectively.

Of course, for finite time, the solutions of the viscous model (1.4) are well known to converge to the entropy solutions of the hyperbolic scalar conservation law (1.3) as $\nu \to 0$ (e.g. [94]). But, as shown above, this limit can not be made uniform as time tends to infinity. Indeed, roughly, we could say that the vanishing viscosity and large-time limits do not commute and that, accordingly, the following two limits yield to different results:

$$\lim_{t\to\infty}\lim_{\nu\to 0} u^\nu(x,t) \qquad \text{and} \qquad \lim_{\nu\to 0}\lim_{t\to\infty} u^\nu(x,t). \tag{1.6}$$

While the first limit leads to the two-parameter hyperbolic N-waves, possibly changing sign, the second one leads to a more restrictive class of asymptotic profiles, corresponding to the N-waves of constant sign. This issue has been precisely analyzed, for instance, in [59, 60]. In particular, in [60] the authors describe the transition from the N-wave shape —the asymptotic profile of the inviscid equation— to the diffusion wave —the asymptotic profile in the viscous equation—.

The main results obtained in this thesis deal with this dichotomy at the numerical level. In fact, the same can occur when approximating the hyperbolic equation (1.3) by numerical schemes. This is not so surprising since, as it is well known, convergent numerical schemes introduce some degree of numerical viscosity (e.g. [38, 67]). As we shall prove, this may affect severely the efficiency of the numerical methods for optimization problems in large-time horizons, even in more complex Burgers-type equations.

1.2.2 Numerics for scalar conservation laws

The inviscid Burgers equation is a one-dimensional equation that belongs to the class of conservation laws. These equations take the general form

$$\begin{cases} \mathbf{u}_t + \operatorname{div}\big(f(\mathbf{u})\big) = 0, & (t,x) \in (0,\infty)\times\mathbb{R}^d, \\ \mathbf{u}(0,x) = \mathbf{u}_0(x), & x \in \mathbb{R}^d. \end{cases} \tag{1.7}$$

Here $\mathbf{u} : [0,\infty)\times\mathbb{R}^d \to \mathbb{R}^m$ is an m-dimensional vector of state variables –for instance, mass, momentum, energy...– and $F : \mathbb{R}^m \to \mathbb{R}^m$ is the flux of the system. Clearly, for the Burgers equation we have $m = d = 1$ and $f(u) = u^2/2$. We refer to [65, 94] for an introduction and the basic theory of nonlinear conservation laws and to [30] for the scalar case.

This type of equations are very practical to model scientific and engineering problems that involve wave motion or advective transport of substances and, hence, the conservation of some quantities [66]. Gas dynamics (Euler equations), plasma in a fusion reactor (magnetohydrodynamics equations) or flow in a porous material (Buckley-Leverett equation), to name but a few, are examples in which conservation laws naturally arise.

Despite the simple appearance of (1.7), there are significant difficulties associated with their solutions that need to be handled carefully when developing numerical methods. The formation of shocks is specially delicate. Indeed, classical finite difference approaches work well when solutions are smooth but fail to approximate discontinuous ones. Therefore, additional efforts are required to overcome this issue.

In this thesis we focus on finite volume methods, which are based on the integral form of (1.7) rather than the differential equation, and one-dimensional scalar conservation laws. Instead of computing pointwise approximations at grid points, these numerical methods break the spatial domain into cells and try to approximate the average of the continuous solution in each of the cells. The key point consists on defining reasonable numerical flux functions to determine the flux trough the edges of the cells. We refer to [67] for an introduction about these techniques. Other methods, such as finite element and spectral ones, have been also applied to conservation laws, but these are not discussed here.

Nevertheless, many finite volume methods are close to finite difference ones, in the sense that similar notation can be assumed, identifying the cell average with the value at the middle point of the cell. In the following, we shall adopt this point of view (the same as in [38, Chapter III]). Given some cell size $\Delta x > 0$ and time step $\Delta t > 0$, we consider u_j^n to be an approximation of $u(n\Delta t, j\Delta x)$, obtained by a finite volume method that approximates equation (1.7):

$$u_j^{n+1} = u_n^j - \frac{\Delta t}{\Delta x}\left(g_{j+1/2}^n - g_{j-1/2}^n\right), \qquad j \in \mathbb{Z}, n > 0. \tag{1.8}$$

Here $g_{j+1/2}^n = g(u_{j-k+1}, \ldots, u_{j+k})$ is the numerical flux, an approximation of the continuous flux $f(u)$ by a continuous function $g : \mathbb{R}^{2k} \to \mathbb{R}$. We refer to [38] for a review on the most classical first and second order methods and to [67] for an introduction to high-resolution methods.

1.2.3 Optimization by line search methods

Optimization is crucial in our world: people optimize, nature optimizes. There are plenty of situations in which an objective that one would like to attain can be identified. From the mathematical point of view, this corresponds to minimizing a functional (maximizing it is equivalent to minimizing its inverse), usually constrained by some restrictions:

$$\min_{z\in\Omega} f(z) \quad \text{subject to} \quad \begin{cases} a(z) = 0, \\ b(z) \geq 0. \end{cases}$$

Generally, this type of problems cannot be solved *by hand* and, hence, it requires effective algorithms to compute approximations of their solutions. The most extended ones are the iterative methods, which seek to construct a sequence $\{z_n\}_{n\geq 0}$ of elements of the set Ω such that $\lim_{n\to\infty} z_n = \bar{z}$, $\bar{z}$ being a exact solution of the optimization problem.

In this thesis we mainly consider line search methods and, in particular, gradient descent methods. Roughly, the key idea behind these techniques is the same that we would use to climb down a mountain: look for a descent direction in the surroundings of the point you are standing at and take a step in that way (see Figure 1.3).

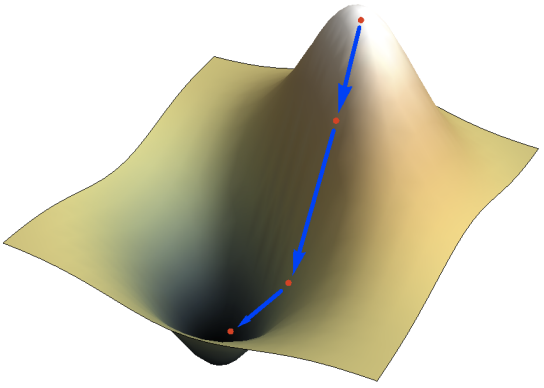


FIGURE 1.3: Minimizing a functional by a gradient descent method.

In the line search strategy, to construct a new approximation z_{n+1} from the current z_n, we shall find a direction d_n at z_n in which the functional decreases. Then, one has to decide the size of the step ε_n to take in that direction. Thus, from an initial guess z_0, one iteratively computes

$$z_{n+1} = z_n + \varepsilon_n d_n, \quad n \in \mathbb{N}.$$

The most common method, the steepest descent, considers the gradient of the functional as the descending direction $d_n = -\nabla f(x_n)$ and chooses the step-size according to the minimum of the functional along the line passing through z_n in the direction d_n. Since we are taking the gradient in a negative direction, $\mathcal{J}(z_{n+1}) < \mathcal{J}(z_n)$ for ε_n small enough, except for the case in which one reaches a minimum in a finite number of iterations: $\nabla\mathcal{J}(z_n) = 0$. Observe that, in general, as the minimum is approached, the gradient tends to zero and, hence, the sequence is expected to converge to a (possibly local) minimizer. We refer to [19] for the basic theory on gradient descent methods.

Gradient descent methods, such as the steepest descent method or conjugate gradient method, are attractive due to their algorithm simplicity. They can even be slightly modified to improve the convergence rate (see, for instance, the works by U. Ascher and his collaborators [4, 91]). Moreover, they do not require calculating second order derivatives. This makes them be a suitable approach for optimizing large scale problems, where the cost of computing the Hessian matrix and solving the corresponding linear system becomes prohibitive. Nevertheless, based on the idea of line search, many

other methods –such as Newton, quasi-Newton...– have been developed since the mid-twentieth century. We refer to [77] for an introduction on line search strategies and their implementation, as well as other optimization tools of the class of trust-region methods.

1.3 Contents of the thesis

In this section we briefly introduce the problems studied in this thesis. Four main subjects are covered:

1. **Vanishing viscosity in numerics:** In Chapter 2, we analyze the large-time asymptotic behavior of the discrete solutions of numerical approximation schemes for scalar conservation laws and, in particular, for the inviscid Burgers equation. We prove that, at the numerical level, the large-time dynamics depends on the amount of numerical viscosity introduced by the scheme. The results of this chapter are based on the article [84], in collaboration with L. I. Ignat and E. Zuazua.

2. **Optimal control in large-time horizons:** In Chapter 3, we discuss various numerical methods for the inverse design of the Burgers equation in large-time horizons. We highlight that convergence in the classical sense of numerical analysis does not suffice. Numerical schemes can alter the dynamics of the underlying continuous system and, hence, affect the efficiency of the numerical optimization tools. This chapter is based on the work [81], in collaboration with N. Allahverdi and E. Zuazua.

3. **Large-time behavior preserving numerics for ABE:** In Chapter 4, we study the well-posedness of the Cauchy problem for the constant-parameter version of the augmented Burgers equation and its large-time dynamics. Then, we propose a semi-discrete numerical scheme, which preserves this asymptotic behavior by introducing two correcting factors in the discretization of the non-local term. The results of this chapter are based on the article [82], in collaboration with L. I. Ignat.

4. **Operator splitting for ABE:** In Chapter 5, we consider an operator-splitting method for the augmented Burgers equation. We prove the first order convergence and obtain the first term in the asymptotic expansion. This chapter is based on the work [83], in collaboration with L. I. Ignat.

In what follows, we describe the most important aspects of each of these topics, the obtained results and the methods we have developed.

Remark 1.1*: Here and subsequently, for a sequence* $v = \{v_j\}_{j\in\mathbb{Z}}$ *and* $p \in [1, \infty)$*, we use the following discrete norms:*

$$\|v\|_{p,\Delta} = \Big(\Delta x \sum_{j\in\mathbb{Z}} |v_j|^p\Big)^{1/p}, \qquad \|v\|_{\infty,\Delta} = \max_{j\in\mathbb{Z}} |v_j|, \qquad TV(v) = \sum_{j\in\mathbb{Z}} |v_{j+1} - v_j|.$$

Moreover, for any $p \in [1, \infty]$*, we denote* $\|\cdot\|_p$ *the usual norm in the* $L^p(\mathbb{R})$ *space.*

1.3.1 Chapter 2: Vanishing viscosity in numerics

As mentioned in the previous section, there is a significant difference concerning the large-time behavior between the viscous and the inviscid Burgers equation. The main result of Chapter 2 states that the same can occur when approximating the hyperbolic equation (1.3) by numerical schemes [84]. This is not so surprising since, as it is well known, convergent numerical schemes introduce some degree of numerical viscosity [38]. Our analysis allows classifying numerical schemes in those that, as time tends to infinity, introduce a negligible amount of numerical viscosity –and therefore lead to the correct asymptotic behavior described by the N-waves– and those that introduce too much numerical viscosity –thus, leading to viscous self-similar profiles–.

As we shall see, Engquist-Osher [26] and Godunov [39] schemes belong to the first category while the classical Lax-Friedrichs scheme [63] fits in the second one. Summarizing, we can say that the solutions of the Engquist-Osher and Godunov schemes, for a fixed mesh, capture the hyperbolic dynamics of the continuous systems. On the contrary, the Lax-Friedrichs scheme, because of the excess of numerical viscosity, leads to the wrong asymptotic behavior, of a viscous nature and not of a hyperbolic one. Note that, although analysis is limited to the Burgers equation, the same conclusions are also to be taken into account when numerically approximating viscous conservation laws where the amount of asymptotic effective viscosity as t tends to infinity may very significantly depend on the nature of the numerical scheme under consideration.

The main goal of Chapter 2 is to analyze the asymptotic behavior as $n \to \infty$ of these discrete solutions for Δx and Δt fixed. Of course, we are interested in numerical schemes that are well known to converge to the entropy solution of (1.7) and with mesh-size parameters satisfying the corresponding CFL condition. We consider three-point conservative numerical schemes to approximate equation (1.7). In particular, we take $k = 1$ and $f(u) = u^2/2$ in (1.8).

Our analysis is mainly concerned with the numerical schemes of Lax-Friedrichs, Engquist-Osher and Godunov, which are of a conservative nature. They are well known to converge to the entropy solution of (1.7) under suitable CFL conditions and to satisfy

the so-called one-sided Lipschitz condition (see [12] and the references therein) that is required to establish, in particular, decay properties as the discrete time tends to infinity.

Our results, corresponding to the L^1-setting, exhibit a significant difference with respect to previous works regarding conservative monotone schemes. In [41], the author analyses the large-time behavior of these schemes in the context of rarefaction waves, thus rather corresponding to a L^∞-setting. Our case can be formally understood as the limit one in which both values at $\pm\infty$ vanish and, hence, reveals the second term in the asymptotic expansion of solutions. We show that, in this framework, the extra viscosity added by the schemes has to be handled carefully to detect the asymptotic behavior as time tends to infinity of discrete solutions.

To be more precise, let us consider $\{u_j^0\}_{j\in\mathbb{Z}}$ an approximation of the initial data; for instance,

$$u_j^0 = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} u_0(x)dx, \quad x_{j+1/2} = (j+\frac{1}{2})\Delta x. \tag{1.9}$$

We introduce the piecewise constant function u_Δ defined almost everywhere in $[0,\infty)\times\mathbb{R}$ by

$$u_\Delta(t,x) = u_j^n, \quad x_{j-1/2} < x < x_{j+1/2},\, t_n \le t < t_{n+1}, \tag{1.10}$$

where $t_n = n\Delta t$ and u_j^n is computed by (1.8).

The following theorem, focused on the Burgers equation, is the main result of Chapter 2.

Theorem 1.1. *Let $u_0 \in L^1(\mathbb{R})$ and choose mesh-size parameters Δx and Δt satisfying the CFL condition $\lambda\|u^n\|_{\infty,\Delta} \le 1$, $\lambda = \Delta t/\Delta x$. Let u_Δ be the corresponding solution of the discrete scheme* (1.8) *for the hyperbolic Burgers conservation law* (1.3)*. Then, for any $p \in [1,\infty)$, the following holds*

$$\lim_{t\to\infty} t^{\frac{1}{2}(1-\frac{1}{p})}\|u_\Delta(t) - w(t)\|_p = 0, \tag{1.11}$$

where the profile w is as follows:

1. *for the Lax-Friedrichs scheme, $w = w_{M_\Delta}$ is the unique solution of the continuous viscous Burgers equation*

$$\begin{cases} w_t + \left(\frac{w^2}{2}\right)_x = \frac{(\Delta x)^2}{2\Delta t} w_{xx}, & x\in\mathbb{R}, t>0, \\ w(0) = M_\Delta \delta_0, \end{cases} \tag{1.12}$$

with $M_\Delta = \int_{\mathbb{R}} u_\Delta^0$.

2. for Engquist-Osher and Godunov schemes, $w = w_{p_\Delta, q_\Delta}$ is the unique solution of the hyperbolic Burgers equation

$$\begin{cases} w_t + \left(\frac{w^2}{2}\right)_x = 0, \quad x \in \mathbb{R}, t > 0, \\ \\ w(0) = M_\Delta \delta_0, \quad \lim\limits_{t\to 0} \displaystyle\int_{-\infty}^{x} w(t,z)dz = \begin{cases} 0, & x < 0, \\ -p_\Delta, & x = 0, \\ q_\Delta - p_\Delta, & x > 0, \end{cases} \end{cases} \tag{1.13}$$

with $M_\Delta = \int_{\mathbb{R}} u_\Delta^0$ and

$$p_\Delta = -\min_{x\in\mathbb{R}} \int_{-\infty}^{x} u_\Delta^0(z)dz \quad \text{and} \quad q_\Delta = \max_{x\in\mathbb{R}} \int_{x}^{\infty} u_\Delta^0(z)dz.$$

The initial data in the above equations (1.12) and (1.13) have to be understood in the sense of convergence of bounded measures. We refer to [28] and [73] for a precise definition.

It is well known [29, 30] that the above profiles are explicitly given by

$$w_{M_\Delta}(x,t) = -\frac{2\sqrt{\nu}}{t^{1/2}} \exp\left(-\frac{x^2}{4\nu t}\right) \left[C_{M_\Delta} + \int_{-\infty}^{x/\sqrt{\nu t}} \exp\left(-\frac{s^2}{4}\right) ds\right]^{-1}, \tag{1.14}$$

where $\nu = \Delta x^2/(2\Delta t)$ and C_{M_Δ} is such that the mass of the solution w_{M_Δ} is M_Δ, and

$$w_{p_\Delta, q_\Delta}(x,t) = \begin{cases} \frac{x}{t}, & -\sqrt{2p_\Delta t} < x < \sqrt{2q_\Delta t}, \\ 0, & \text{elsewhere.} \end{cases} \tag{1.15}$$

Note that the viscous profiles (1.14) are fully determined by the total mass, which is conserved under the dynamics under consideration both in the time-continuous and time-discrete case. By the contrary, the N-wave profiles (1.15) are uniquely determined by the two parameters (p, q) of invariants that are constant along the continuous and discrete dynamics. The quantity $q_\Delta - p_\Delta$ is precisely M_Δ, the mass of u_Δ^0.

The difference among them can be observed in Figure 1.2, for instance, where we have taken $t = 50$, $\Delta x = 1/200$, $\Delta t = 1/200$, $M_\Delta = 0.15$, $p_\Delta = 0.05$ and $q_\Delta = 0.2$.

In order to prove Theorem 1.1, we use scaling arguments, similar to those applied in the proofs of the continuous analogues. We also introduce similarity variables, which are also a standard tool in the analysis of asymptotic behavior of partial differential equations. This will allow us to observe some phenomena in a clearer manner.

1.3.2 Chapter 3: Optimal control in large-time horizons

In Chapter 3 we analyze the numerical approximation of the inverse design problem for the Burgers equation, both in the viscous (1.4) and in the inviscid (1.3) case. As it happens in the case of the sonic-boom minimization problem, given a time $T > 0$ and a target function u^*, the aim is to identify the initial datum u_0 so that the solution, at time $t = T$, reaches the target u^* or gets as close as possible to it.

Essentially, the question consists in solving the Burgers equation backwards, a problem that is ill-posed. In the viscous case $\nu > 0$, this is due to the intrinsic strong time-irreversibility of the parabolic Burgers equation that is enhanced by the nonlinear phenomena governing the hyperbolic dynamics. In the inviscid hyperbolic case, the nonlinearity of the model that produces, in particular, the emergence of shock discontinuities, makes the problem to be ill-posed too, having, in some cases, multiple solutions.

We formulate the problem from the point of view of optimal control. Using a least square approach, we consider the minimization of the following functional:

$$\mathcal{J}(u_0) = \frac{1}{2} \int_{\mathbb{R}} (u(x,T) - u^*(x))^2 \, dx, \tag{1.16}$$

where u is the solution of the Burgers equation and the initial data u_0 lies in a suitable functional class, for instance $L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$.

This optimal control problem, a simplified version of the complete sonic-boom model, arises naturally. One of the key ingredients is that the time horizon under consideration $[0, T]$, for practical purposes, needs to be long. As we shall see, this makes the choice of the numerical scheme approximating the PDE to be a sensitive issue, since schemes that do not yield the correct long-time dynamics are incapable of providing an accurate approximation of the optimal control. As we mentioned before, the asymptotic parabolic/hyperbolic dichotomy for large times needs to be handled carefully. In particular, the excess of numerical viscosity may destroy the long-time hyperbolic dynamics to make it parabolic. The analysis in Chapter 2 is carried out in the hyperbolic contest. Nevertheless, we show that this pathology may arise also for the viscous Burgers equation when the numerical viscosity dominates the physical one. The main aim of this chapter is to accurately describe the various phenomena that overlap when handling numerically this problem.

In this work, we emphasize the consequences of this fact at the level of inverse design. This is done, mainly, by means of a gradient descent method, using the adjoint system methodology. We also use IPOPT, an open-source software package for nonlinear optimization [92], to support our results. Note, however, that the large-time behavior

dichotomy can be extended to other numerical methods. We refer to [19, 77] and the references therein for the basics in this broad area of numerical optimization.

This optimal control problem, without the requirement of the large-time horizon, has already been addressed by several authors in the past, both from the continuous point of view and the discrete one. As we already pointed out, the interplay of discretization and optimization makes this topic really challenging. In the more general context of scalar conservation laws, it is worth to mention [7, 10, 55, 90]. In fact, there has been an extensive research by F. James and his collaborators regarding the adjoint system and its discretization (see, for instance, [9, 40, 56]). Certainly, the presence of shocks in this type of problems is not a vain issue. In [35, 36] the authors analyze the pointwise convergence of the linearized and adjoint approximations for discontinuous solutions in a discretize-then-optimize approach. In particular, they already point out the importance of controlling the diffusion in order to obtain convergence. Nevertheless, our approach shows that, in practice, their proposal might be insufficient when considering a large T.

Moreover, the interplay between the optimize-then-discretize and the discretize-then-optimize approaches was analyzed in [89]. In Chapter 3 we opt for the latter, which does not take into account the derivatives with respect to shocks. We avoid discontinuities in the solutions to minimize their impact on the optimization process and focus only on the large-time effect, so the sensitivities of the shocks do not play an important role in our case. Of course, note that this does not mean that large-time effects should not been taken into account in the presence of shocks.

Our results constitute an interesting warning about the necessity of employing numerical approximation schemes, capable of mimicking the continuous large-time dynamical properties of the system, when addressing inverse design and optimal control problems in long-time horizons. This was previously observed when dealing with control problems for wave propagation [27, 95]. It is interesting to see that the same pathologies persist for the apparently more robust problem of inverse design for inviscid and viscous flows.

1.3.3 Chapter 4: Large-time behavior preserving numerics for the augmented Burgers equation

Chapter 4 is devoted to the augmented Burgers equation (1.1) with constant parameters and a unique relaxation process. We focus on the following equation:

$$\begin{cases} u_t = uu_x + \nu\, u_{xx} + c\, K_\theta * u_{xx}, & (t,x) \in (0,\infty)\times\mathbb{R}, \\ u(0,x) = u^0(x), & x \in \mathbb{R}, \end{cases} \tag{1.17}$$

where $*$ denotes the convolution in the x variable, the parameters ν, c, θ are positive and

$$K_\theta(z) = \begin{cases} \frac{1}{\theta} e^{-z/\theta}, & z > 0, \\ 0, & \text{elsewhere.} \end{cases} \tag{1.18}$$

The same pathologies analyzed in Chapters 2 and 3 arise in the context of the augmented Burgers equation. Small values for ν and c in (1.17) require a similar treatment from the numerical point of view, as if the equation was a hyperbolic conservation law. Therefore, in Chapter 4 we study the asymptotic behavior of the solutions to (1.17) as $t \to \infty$ and we build a semi-discrete numerical scheme that preserves this behavior.

In what concerns the large-time behavior of solutions of system (1.17), the main result is stated in the following theorem.

Theorem 1.2. *Let $u^0 \in L^1(\mathbb{R})$. For any $p \in [1, \infty]$, the solution u to (1.17) satisfies*

$$t^{\frac{1}{2}(1-\frac{1}{p})} \|u(t) - u_M(t)\|_p \longrightarrow 0, \quad \text{as } t \to \infty,$$

where $u_M(t,x)$ is the solution of the following equation:

$$\begin{cases} u_t = uu_x + (\nu + c)u_{xx}, & x \in \mathbb{R}, t > 0, \\ u(0) = M\delta_0. \end{cases}$$

Here δ_0 denotes the Dirac delta at the origin and M is the mass of the initial data, $M = \int_{\mathbb{R}} u^0(x)dx$.

Note that u_M is explicitly defined in (1.14), simply taking $\nu + c$ instead of ν as the viscosity parameter. In fact, this shows that u_M is of the form $t^{-\frac{1}{2}} f_M\left(\frac{x}{\sqrt{t}}\right)$ for some function f_M and, hence, self-similar.

This will be particularly important at the numerical level. On the one hand, when choosing the numerical flux to discretize the nonlinearity, we need to handle carefully the numerical viscosity that is introduced. On the other hand, we need to treat the truncation of the integral term in such a manner that we do not introduce undesired pathologies in the long-time behavior of the numerical solutions. In this sense, we propose two correcting factors.

Let us denote by u_Δ an approximation to the solution u of (1.17). We define this piecewise constant in space function as follows:

$$u_\Delta(t,x) = u_j(t), \quad x \in (x_{j-1/2}, x_{j+1/2}), t \geq 0, \tag{1.19}$$

where $x_{j+1/2} = (j + \frac{1}{2})\Delta x$, for all $j \in \mathbb{Z}$, and $\Delta x > 0$ is a given mesh-size. We will also denote by $x_j = j\Delta x$ the intermediate points of the spatial cells. For each $j \in \mathbb{Z}$ we need to compute a function $u_j(t)$ that approximates the value of the solution in the cell. Taking into account the issues enumerated above, we choose the following discretization of (1.17): the Engquist-Osher scheme for the flux, centered finite differences for the laplacian and the composite rectangle rule for the integral:

$$\begin{cases} u_j'(t) = \frac{g_{j+1/2}(t) - g_{j-1/2}(t)}{\Delta x} + \nu \frac{u_{j-1}(t) - 2u_j(t) + u_{j+1}(t)}{\Delta x^2} \\ \qquad\qquad + \frac{c}{\theta^2} \left(\sum_{m=1}^{N} \omega_m^\theta u_{j-m}(t) - F_0^{\Delta,\theta} u_j(t) + F_1^{\Delta,\theta} \theta \frac{u_{j+1}(t) - u_j(t)}{\Delta x} \right), \quad j \in \mathbb{Z}, t \geq 0, \\ u_j(0) = \dfrac{1}{\Delta x} \displaystyle\int_{x_{j-1/2}}^{x_{j+1/2}} u^0(x)dx, \qquad\qquad j \in \mathbb{Z}, \end{cases} \tag{1.20}$$

where

$$\omega_m^\theta = e^{-m\Delta x/\theta} \left(e^{\Delta x/\theta} - 1 \right), \quad m = 1, \ldots, N, \tag{1.21}$$

and

$$g_{j+1/2}(t) = \frac{u_j(t)\big(u_j(t) - |u_j(t)|\big)}{4} + \frac{u_{j+1}(t)\big(u_{j+1}(t) + |u_{j+1}(t)|\big)}{4}, \quad j \in \mathbb{Z}, t \geq 0.$$

The parameter $N = N(\Delta x) \in \mathbb{N}$ denotes the number of nodes considered in the quadrature formula of the integral. The correcting factors $F_0^{\Delta,\theta}$ and $F_1^{\Delta,\theta}$ in front of the approximations of u and u_x, given by

$$F_0^{\Delta,\theta} = \sum_{m=1}^{N} \omega_m^\theta \quad \text{and} \quad F_1^{\Delta,\theta} = \frac{\Delta x}{\theta} \sum_{m=1}^{N} m\omega_m^\theta, \tag{1.22}$$

will handle, from the asymptotic behavior point of view, the correct truncation of the nonlocal term.

Finally, for Δx fixed we study the asymptotic behavior as $t \to \infty$ of these semi-discrete solutions u_Δ.

Theorem 1.3. *Let $u_0 \in L^1(\mathbb{R})$, $\Delta x > 0$ and u_Δ be the corresponding solution of the semi-discrete scheme* (1.20) *for the augmented Burgers equation* (1.17)*. For any $p \in [1, \infty]$, the following holds*

$$t^{\frac{1}{2}(1-\frac{1}{p})} \|u_\Delta(t) - u_M(t)\|_p \longrightarrow 0, \quad \text{as } t \to \infty, \tag{1.23}$$

where $u_M(t, x)$ is the unique solution of the following viscous Burgers equation:

$$\begin{cases} v_t = vv_x + (\nu + c\, F_2^{\Delta,\theta}) v_{xx}, \quad x \in \mathbb{R}, t > 0, \\ v(x, 0) = M\delta_0. \end{cases}$$

Here, $M = \int_{\mathbb{R}} u^0(x)dx$ is the mass of the initial data and

$$F_2^{\Delta,\theta} = \frac{\Delta x^2}{2\theta^2}\left(\sum_{m=1}^{N} m(m-1)\omega_m^\theta\right).$$

For a fixed $\theta > 0$, let us observe that if N is taken such that $N \to \infty$ and $N\Delta x \to \infty$ when $\Delta x \to 0$, then $F_2^{\Delta,\theta} \to 1$. This is, precisely, the value that we should expect from the continuous model.

1.3.4 Chapter 5: Operator splitting for ABE

Solving (1.20) can be computationally expensive if N is big. Thus, in Chapter 5 we set up the framework to empower the use of splitting methods to solve the augmented Burgers equation, focusing on the large-time behavior. As a matter of fact, let us mention that these techniques have already been used in [20, 85, 86] in the context of the sonic-boom phenomenon. Moreover, various versions of this method have been developed, for instance, for the nonlinear Schrödinger, Korteweg-de Vries, Boltzmann... (see [44] and the references therein) and, more recently, for the nonlocal Fowler equation [11].

The basic idea behind operator splitting methods is that the overall evolution operator can be formally written as a sum of evolution operators for each term appearing in the model. In other words, one separates the complete system into a set of simpler sub-equations, for which more practical algorithms are available. Once we solve successively these sub-equations, their solutions are put together to compute the complete solution of the model. We refer to [44] for a more detailed introduction on operator splitting methods.

We introduce the following Trotter formula for the augmented Burgers equation (1.17). For the sake of clarity, we analyze the case $\nu = c = \theta = 1$, but the obtained results can be easily extended to the general case. Let X^t be the evolution operator associated with

$$\begin{cases} v_t = K * v - v + v_x, & (t,x) \in (0,\infty)\times\mathbb{R}, \\ v(t=0,x) = v_0(x), & x \in \mathbb{R}, \end{cases}$$

and Y^t the one corresponding to

$$\begin{cases} w_t - \left(\frac{w^2}{2}\right)_x = w_{xx}, & (t,x) \in (0,\infty)\times\mathbb{R}, \\ w(t=0,x) = w_0(x), & x \in \mathbb{R}. \end{cases}$$

We consider the flow Z^t defined by

$$Z^t = X^t Y^t. \tag{1.24}$$

The aim is to approximate the solution u of (1.17) by

$$(Z^{\Delta t})^n u_0 = (X^{\Delta t} Y^{\Delta t})^n u_0,$$

where $\Delta t > 0$ and $n \in \mathbb{N} \cup \{0\}$. For the sake of simplicity on the notation, from now on we denote $Z^{n\Delta t} = (Z^{\Delta t})^n$. Let us recall that $t_n = n\Delta t$ and $t_{n+1/2} = (n + \frac{1}{2})\Delta t$ for every $n \in \mathbb{N} \cup \{0\}$.

The first result of Chapter 5 is the following one. It confirms that the splitting (1.24) is first-order accurate for sufficiently smooth solutions.

Theorem 1.4. *Let $r \in \{1, 2\}$. For all $u_0 \in H^r(\mathbb{R})$ and for all $T > 0$, there exist positive constants c_1, c_2 and Δt_0 such that, for all $\Delta t \in (0, \Delta t_0)$ and for all $n \in \mathbb{N}$ such that $0 \leq n\Delta t \leq T$,*

$$\left\| Z^{n\Delta t} u_0 - u(n\Delta t) \right\|_2 \leq c_1 (\Delta t)^{r/2}$$

and

$$\left\| Z^{n\Delta t} u_0 \right\|_{H^r(\mathbb{R})} \leq c_2.$$

Here, c_1, c_2 and Δt_0 depend on T and on $\|u_0\|_{H^r(\mathbb{R})}$.

Following similar techniques as in Chapter 4, we obtain the first term in the asymptotic expansion of the solution given by the splitting operator Z^t. Let us define the function u^Δ by

$$u^\Delta(t, x) = \begin{cases} Y^{2(t-t_n)} Z^{n\Delta t} u_0(x), & t \in (t_n, t_{n+1/2}),\ x \in \mathbb{R}, \\ X^{2(t-t_{n+1/2})} Y^{\Delta t} Z^{n\Delta t} u_0(x), & t \in (t_{n+1/2}, t_{n+1}),\ x \in \mathbb{R}. \end{cases}$$

It follows that u^Δ satisfies

$$\begin{cases} u^\Delta_t = 2(Y^{2(t-t_n)} Z^{n\Delta t} u_0)_t = 2u^\Delta_{xx} + \left((u^\Delta)^2\right)_x, & t \in (t_n, t_{n+1/2}), \\ u^\Delta_t = 2(K * u^\Delta - u^\Delta + u^\Delta_x), & t \in (t_{n+1/2}, t_{n+1}), \\ u^\Delta(t_n) = Z^{n\Delta t} u_0, \\ u^\Delta(t_{n+1/2}) = Y^{\Delta t} Z^{n\Delta t} u_0. \end{cases}$$

Let us denote $I^\Delta_n = (t_n, t_{n+1/2})$. If we define the function $\psi^{\Delta t}(t) = \sum_{n \geq 0} \chi_{I^\Delta_n}(t)$ (whose value is 1 if t is in $(t_n, t_{n+1/2})$ and 0 elsewhere), it is clear that the previous system can

be rewritten as

$$\begin{cases} u_t^\Delta = 2\psi^{\Delta t}(t)u_{xx}^\Delta + \psi^{\Delta t}(t)((u^\Delta)^2)_x \\ \qquad\qquad +2\left(1-\psi^{\Delta t}(t)\right)(K*u^\Delta - u^\Delta + u_x^\Delta), \quad t>0, \\ u^\Delta(0) = u_0. \end{cases} \tag{1.25}$$

It can be shown that $\psi^{\Delta t}(t) \rightharpoonup \frac{1}{2}$ (see [23]). Moreover, by construction and the properties of X^t and Y^t, we have that u^Δ defined by (1.25) satisfies $u^\Delta \in C([0,\infty); L^1(\mathbb{R}))$. Thus, formally we have that $u^\Delta \to u$ as $\Delta t \to 0$, where u satisfies (1.17). The behavior of u^Δ as $t \to \infty$ follows from a scaling argument and it is stated in the theorem below.

Theorem 1.5. *For any $u_0 \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ and $p \in [1,\infty)$*

$$\left\|u^\Delta(t)\right\|_p \le C_p t^{-\frac{1}{2}(1-\frac{1}{p})}\|u_0\|_1, \quad n \ge 1.$$

Moreover, for all $p \in [1,\infty)$,

$$t^{\frac{1}{2}(1-\frac{1}{p})}\left\|u^\Delta(t) - u_M(t)\right\|_p \to \infty, \quad \text{as } t \to \infty,$$

where $u_M(t,x) = t^{-1/2} f_M(x/\sqrt{t})$ is the self-similar profile of the following viscous Burgers equation

$$\begin{cases} u_t = \left(\frac{u^2}{2}\right)_x + 2u_{xx}, & (t,x) \in (0,\infty)\times\mathbb{R}, \\ u(0,x) = M\delta_0, & x \in \mathbb{R}. \end{cases}$$

Note that this is, precisely, the large-time asymptotic behavior of the solution of (1.17), as shown in Chapter 4.

1. Kapitulua

Sarrera

Gaur egun, aeronautika ikerketan dagoen ardurarik handienetariko bat hegazkinek egindako zarata kontrolatzera eta murriztera zuzenduta dago. Izan ere, industria-arlo garrantzitsu honetan dagoen helburu nagusienetako bat abioi supersonikoek sortzen duten burrunba gogorrari aurre egitea da. Bereziki, ontzi hauek egindako soinu-eztandaren gutxitzea da garraio zibil supersonikoak garatzea lortzeko giltza [3].

Tesi honetan ikusiko dugunez, mota horretako arazoek iraupen luzeko eboluzio problemak ondo ebatzi ditzaketen tresna numerikoen beharra dute. Sistema jarraiaren denboraldi luzeko ezaugarri dinamikoak mantentzen dituzten hurbilketa eskema numerikoen premia nabarmenduko dugu. Aplikazio gisa, soinu-eztandaren iragarpen eta kontrolaren kasuan erabiliko dugu hori.

1.1 Soinu-eztandaren minimizatzea

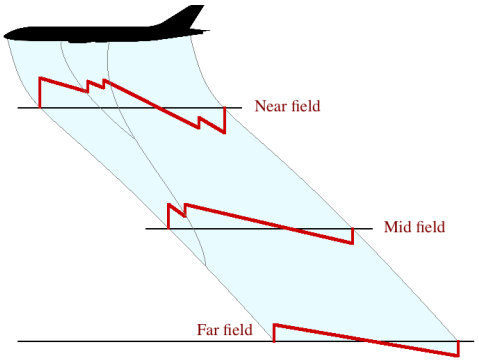
Soinuaren abiadura gainditzean, hegazkin supersonikoek atmosferan presio perturbazioak sortzen dituzte, airea lekualdatzeak eta euste aerodinamikoak eraginda. Energia akustiko itzela duten perturbazio gehienak lurzorura heltzen dira, soinu-eztanda agertaraziz. Gizakiok fenomeno hau danbada jarrai bi balira bezala entzuten dugu. Beraz, guretzako ez ezik, eraikinentzat ere oso kaltegarria izan daiteke. Horrregatik, hegazkin supersonikoak diseinatzeko orduan hori izan da oztoporik nagusienetakoa.

Soinuaren abiadura gainditu zuen lehenengo hegazkina Bell X-1 izan zen 1947 urtean [24, 10. Kapitulua]. Geroxeago, 60 eta 70 hamarkadetan garraio mota horri erabilera zibila emateko lehenengo ahaleginak hasi ziren. Bi proiektu baino ez zuten arrakasta izatea lortu: frantziar eta britainiarren Concordeak eta sobietarren Tupolev-144ak. Beste proiektu bi, estatubatuarrak biak (Lockheed L-2000 eta Boeing 2707), bidean geratu ziren EEBB-ko kongresuak ezarri zituen ingurune-murrizketengatik. Debeku hura zela eta, gune jendetsuen gainetik (praktikan edonondik ozeanen gainetik izan ezik) abiadura

supersonikoz hegan egitea ezinezkoa zen. Beraz, aukera ekonomiko guztiak bertan behera geratu ziren eta, azkenean, garraio zibil supersonikoak eraikitzeko proeiektu guztiek porrot egin zuten.

Horrezkero, XX. mendeko azken hamarkadetan soinu-eztanda murrizteko teknologiak erabili behar zirela argi geratu zen, zarata arazo hori gainditzeko gai diren hegazkin eraginkorrak eraikitzearren. Izan ere, DARPA/NASA/Northrop-Grumman Shaped Sonic Boom Demonstrator (SSBD) proiektuak [79] hegazkinaren forma eraldatuz soinu-eztanda leuntzea posible zela erakutsi zuen esperimentalki lehen aldiz.

Eremu hurbileko sinadura N-uhin bat bilakatzen dela ezaguna da (ikusi 1.1 Irudia). Hegazkinek sortzen dituzten astinaldi anitzak elkartu egiten dira, presioaren hedapen lineal batek banandutako bi talka-uhin soilik geratuz; hortik dator N-uhina izena (*N-wave* ingelesez). Hasieran, N-uhina saihetsezina zela uste izan zen. Baina McLean-ek eremu hurbileko sinadura moldatuz lurzoruan N-uhinarik ez heltzea posiblea zela frogatu zuen [74]. Halere, lehen esan dugunaren arabera, horren ebidentzia empirikoa SSBD proeiektuak eman zuen lehendabiziz, 2003 urtean.



IRUDIA 1.1: Soinu-eztandaren hedapenaren diagrama. Eremu hurbilean sortutako astinaldiak elkartu egiten dira, N-uhina eratuz.

Historikoki, teoria lineala erabili ohi zen soinu-eztandaren hedapenaren eredua egiteko, 40 eta 50 hamarkadetako Hayes [43] eta Whitham-en [93] lanei jarraipena emanez. Beraz, ikerketa analitiko eta numeriko gehiena arlo horretan izan zen. Oraindik orain, eredu ez-lineal berriak garatzen hasi dira emaitza zehatzagoak lortzeko. Horrelako bi metodo aipatzea komenia da. Alde batetik, Ozcer-ek ekuazio potenzial osoa erabiltzea proposatu zuen [78]. Bestetik, Rallahbandi-k efektu biskosoak kontuan hartzen dituen Burgers ekuazio handitu bat aurkitu zuen (ikusi [20] eta duela gutxiko [85]).

Tesi honetan azkenengoan jarriko dugu arreta. Burgers ekuazio handituak (hemendik

aurrera, ABE ingeleseko siglei jarraituz) efektu ez-linealak –adibidez, uhinaren makurtzea eta hedapen-abiadura aldakorra– ez ezik, lasaikuntza molekularra, izpi-hodien zabalketa eta atmosferaren geruzak ere kontuan hartzen ditu. Ekuazioa hurrengoa da:

$$\frac{\partial P}{\partial \sigma} = P\frac{\partial P}{\partial \tau} + \frac{1}{\Gamma}\frac{\partial^2 P}{\partial \tau^2} + \sum_{\nu} C_\nu \frac{\frac{\partial^2 P}{\partial \tau^2}}{1+\theta_\nu \frac{\partial}{\partial \tau}} - \frac{1}{2G}\frac{\partial G}{\partial \sigma}P + \frac{1}{2\rho_0 c_0}\frac{\partial(\rho_0 c_0)}{\partial \sigma}P, \tag{1.1}$$

non $P = P(\sigma, \tau)$ dimentsiorik gabeko presio banaketaren perturbazioa den. Hedapenaren distantzia σ eta perturbazioaren iraupena τ ere dimentsiorik gabekoak dira. Lasaikuntza molekularrei dagokien batuketan agertzen den eragilea horrela definiturik dago:

$$\frac{\theta_\nu}{1+\theta_n u\frac{\partial}{\partial \tau}} f(\tau) = \int_{-\infty}^{\tau} e^{(\xi-\tau)/\tau_\nu} f(\xi) d\xi.$$

G funtzioak izpi-hodiaren azalera adierazten du. Atmosferaren baldintzak ρ_0 dentsitateak, c_0 soinuaren abiadurak, Γ parametro termo-biskoso batek eta lasaikuntza mota bakoitzeko (normalean, oxigenorako eta nitrogenorako) dimentsiorik gabeko θ_ν lasaikuntza denborek eta C_ν dispertsio parametroek ezartzen dituzte.

Kasu honetan, soinu-eztandaren minimizazio arazoa zertan datza: lurzoruan lortu nahi den P^* sinadura eta hedapenaren Σ iraupena –hegaldiaren altuera erlazionaturik dagoena– emanik, hobeto erreproduzitzen duen eremu hurbileko sinadura berreskuratu. Hau formulatzeko ohizko era kontrol optimoa da, karratu txikienen metodoaren bidez [86], hurrengo funtzionala erabiliz:

$$\mathcal{S}(P_0) = \frac{1}{2}\int_{\mathbb{R}} \big(P(\Sigma,\tau) - P^*(\tau)\big)^2 d\tau. \tag{1.2}$$

Hemen P_0 hasierako datuak onargarriak diren eremu hurbileko sinaduren multzo batean daude (adibidez, hegazkinaren geometriak ezarritako diseinu aldagaiek definituriko multzoan [76, 86]) eta, bestetik, P (1.1) Ekuazioan $P(0,\tau) = P_0(\tau)$ hartuz lortzen den ebazpena da. Hautemandako danbatekoa (PLdB) edo astinaldien gain-presioa dira minimizatzeko kontuan har daitezkeen beste funtzional bi [76].

Soinu-eztandaren minimizazio arazo honetan denbora-eskalak oso ezberdinak dira: presioaren perturbazioa segundu erdi baino gutxiago batean gertatzen den bitartean, uhinaren hedapena ia minutu oso bat iraun daiteke, hegaldiaren baldintzen arabera [85]. Tesi honetan frogatuko dugun arabera, horrek tratamendu konputazionala zaildu egiten du eta, beraz, eskema numerikoen denboraldi luzeetan daukaten konportaera aztertzea ezinbestekoa da.

1.2 Esparru matematikoa

Atal honetan tesian zehar agertuko diren zenbait tresna matematikoak bildu ditugu. Soinu-eztandaren eredua Burgers motako ekuazio bat denez, Burgers ekuazioaren ebazpenen konportaeraren hainbat ezaugarri aurkeztuko ditugu. Gainera, ekuazio hau kontserbazio-lege bat denez, ekuazio mota horiek bolumen finitoen bidez diskretizatzeko era klasikoa ere gogoratuko dugu. Azkenik, bilaketa lineala erabiltzen duten metodoen sarrera egingo dugu, bereziki gradiente metodoetan arreta jarriz.

1.2.1 Burgers ekuazioa eta denboraldi luzeko dinamika

Burgers ekuazio handitua (1.1) Burgers ekuazioaren aldakuntza anitzetako bat da. Teoria linealarekin batera, Burgers motako ekuazioak anplitude mugatuko uhinen hedapena imitatzeko tresnarik inportanteenetariko bat izan dira. Burgers ekuazio biskoso klasikoa ingurune disipakor batean gertatzen den uhin hedapena aztertzeko hartu zen kontuan lehendabiziz. Ondoz ondoko orokortzeetan beste zenbait efektu sartu ziren; esate baterako, zabalpen geometrikoa eta ingurune hetereogeneoak (Burgers ekuazio orokortua [16, 31, 69]) edota lasaikuntza prozesuak (Burgers ekuazio zabaldua [80]). Aurreko atalean azaldutakoaren arabera, Burgers ekuazio handituak, aipatutako fenomeno guzti horiek kontuan izanik, hegazkin supersonikoek sortutako soinu-eztandaren hedapena modelatzeko erabilitzen hasi dira [85].

Burgers ekuazio ez-biskosoa hurrengo forma duen kontserbazio-legea da:

$$u_t + \left(\frac{u^2}{2}\right)_x = 0, \quad x \in \mathbb{R}, t > 0. \tag{1.3}$$

Bertsio biskosoa, ordea, dimentsio bateko konbekzio-difusio ekuazioa da:

$$u^\nu_t + \left(\frac{(u^\nu)^2}{2}\right)_x = \nu u^\nu_{xx}, \quad x \in \mathbb{R}, t > 0, \tag{1.4}$$

non $\nu > 0$ den. Ekuazio hau Bateman-ek [8] eta Burgers-ek [15] ondorioztatu zuten. Gainera, azken honek Navier-Stokes ekuazioen sinplifikaziotzat hartu zuen lehena izan zen. Burgers ekuazio biskosoa hedapen ez-lineala eta difusio-efektua kontuan hartzen dituen eredurik sinpleena da [94]. Izan ere, Hopf-ek [45] eta Cole-k [22] ebazpenaren formula esplizitua aurkitu zuten, Hopf-Cole delako transformazioa erabiliz.

Ezaguna da (1.3) Burgers ekuazio ez-biskosoaren ebazpenen konportaera asintotikoa auto-antzekoa dela (begira [72] eta bertako erreferentziak). Izan ere, $t \to \infty$ heinean, ebazpenek N-uhinaren konportaera garatzen dute, hasierako datuaren masa kontserbatuz. Halere, kontuan izan masa soilik ez dela nahikoa auto-antzeko profil asintotikoa

zein den jakiteko [73]. Horretarako bi parametro behar dira, masa positiboa eta masa negatiboa alegia:

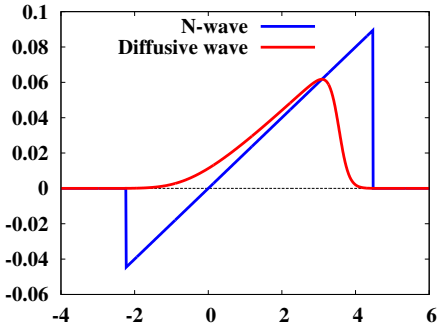
$$p = -\min_{x\in\mathbb{R}} \int_{-\infty}^{x} u^0(z)dz \quad \text{eta} \quad q = \max_{x\in\mathbb{R}} \int_{x}^{\infty} u^0(z)dz. \tag{1.5}$$

Orokorrean, seinua aldatzen duten hasierako datuek seinua aldatzen duen N-uhina garatuko dute.

Konportaera hori ez dator bat (1.4) kasu biskosoarenarekin. Izan ere, $\nu > 0$ hartuz, problema horiek parabolikoak dira eta, $t \to \infty$ heinean, masa kontserbatzen duen seinu konstanteko profil biskoso auto-antzeko baterantz jotzen dute [45]:

$$M = q - p = \int_{\mathbb{R}} u^0(x)dx.$$

1.2 Irudiak bi profilak konparatzen ditu, parametro berdinak hartu ezkero.



IRUDIA 1.2: $\nu > 0$ kasu parabolikoan (gorriz) eta $\nu = 0$ hiperbolikoan (urdinez) agertzen diren profil asintotikoak. Hemen $M = 0.15$, $p = 0.05$ eta $q = 0.2$ hartu dugu.

Jakina, denbora finituan, (1.4) eredu biskosoaren ebazpenek (1.3) kontserbazio-lege hiperbolikoko ebazpen entropikoetarantz konbergitzen dute, $\nu \to 0$ ahala (begira, adibidez, [94]). Baina, limite hori ezin da uniformeki hartu denbora infinitorantz joatean. Izan ere, biskositate desagerkorra eta denboraldi luzeko limiteak ezin daitezkeela trukatu esan genezake. Hau dan, hurrengo bi limiteek emaitza desberdinak dituzte:

$$\lim_{t\to\infty}\lim_{\nu\to 0} u^\nu(x,t) \qquad \text{eta} \qquad \lim_{\nu\to 0}\lim_{t\to\infty} u^\nu(x,t). \tag{1.6}$$

Arazo hau [59, 60] artikuluetan aztertuak izan dira. Egileek, [60] artikuluan bereziki, N-uhinetik difusio uhinerako iragatea deskribatzen dute.

Tesi honetan lortutako emaitzek dikotomia horri egiten diote aurre maila numerikoan. Izan ere, gauza bera gerta liteke (1.3) Ekuazio hiperbolikoaren ebazpena eskema numerikoen bidez hurbiltzean. Halere, hau ez luke harrigarria izan behar, eskema numerikoek biskositate numerikoa sartzen baitute sisteman (begira, adibidez, [38, 67]).

Frogatuko dugun arabera, optimizazio problemetan ere eragin handia izan liteke, baita Burgers motako beste eredu konplexuagoetan.

1.2.2 Kontserbazio-lege eskalarretako tresna numerikoak

Burgers ekuazio ez-biskosoa dimentsio bakarreko kontserbazio-legea da. Ekuazio hauek hurrengo itxura orokorra hartzen dute:

$$\begin{cases} \mathbf{u}_t + \operatorname{div}\big(f(\mathbf{u})\big) = 0, & (t,x) \in (0,\infty) \times \mathbb{R}^d, \\ \mathbf{u}(0,x) = \mathbf{u}_0(x), & x \in \mathbb{R}^d. \end{cases} \tag{1.7}$$

Hemen $\mathbf{u} : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^m$ m dimentsioko egoera-aldagaiez osaturiko bektorea da –adibidez, masa, momentua, energia...– eta $F : \mathbb{R}^m \to \mathbb{R}^m$, sistemaren fluxua. Argi dago Burgers ekuazioaren kasuan $m = d = 1$ eta $f(u) = u^2/2$ dugula. Begira [65, 94] kontserbazio-lege ez-linealen oinarrizko teoria aurkitzeko eta [30] kasu eskalarra aztertzeko.

Ekuazio mota hau oso erabilgarria da uhin higidura edota substantzien garraio adbektiboa (eta, beraz, zenbait kopururen kontserbazioa) berekin dakarten zientzia eta ingenieritza-problemak modelatzeko orduan [66]. Gasen dinamika (Euler ekuazioak), fusio-erreaktore bateko plasma (magnetohidrodinamikaren ekuazioak) edo material porotsu batean fluxua (Buckley-Leverett ekuazioa) eredu hiru besterik ez dira.

Nahiz eta (1.7) sinple irudi, zailtasun handiak ager daitezke ebazpenak hurbiltzeko metodo numerikoak garatzeko orduan. Talka-uhinen sortzea gai labana da benetan. Izan ere, fiderentzia finitu klasikoak ondo dabiltza ebazpenak leunak direnean, baina etenak hurbiltzean huts egiten dute. Beraz, ahalegin gehigarria behar da arazo hori menderatzeko.

Tesi honetan bolumen finitoen metodoetan jartzen dugu arreta, dimentsio bakarreko kontserbazio-lege eskalarretarakoetan bereziki. Hauek (1.7) Ekuazioaren forma integralean oinarriturik daude. Sareko nodo bakoitzeko balioa aurkitzen saiatu beharrean, bolumen finitoen metodoek eremu espaziala gelaxkatan banatu eta hauetan ebazpen jarraituak duen batez-besteko balioa hurbiltzen dute. Puntu nagusia gelaxken arteko fluxua zehazteko fluxu numeriko dezentea definitzea da. Begira [67] teknika hauei buruzko sarrera sakonagoa izateko. Beste metodorik ere posible da, elementu finituak edo espektralak esaterako, baina ez ditugu kontuan izango lan honetan.

Hala eta guztiz ere, bolumen finitoen hainbat metodo diferentzia finitoen antzekoak dira; hau da, gelaxken batez-besteko balioa euren erdiko puntuaren balioarekin topatuz gero, antzeko notazioa erabil daiteke. Hemendik aurrera, [38] liburuan bezala, ikuspuntu

hau jarraituko dugu. $\Delta x > 0$ gelaxka tamaina eta $\Delta t > 0$ denbora pausua emanik, u_j^n $u(n\Delta t, j\Delta x)$ balioaren hurbilketatzat hartuko dugu, (1.7) Ekuazioaren ebazpena hurbiltzen duen bolumen finitoen metodoaz lortuta:

$$u_j^{n+1} = u_n^j - \frac{\Delta t}{\Delta x}\left(g_{j+1/2}^n - g_{j-1/2}^n\right), \qquad j \in \mathbb{Z}, n > 0. \tag{1.8}$$

Hemen $g_{j+1/2}^n = g(u_{j-k+1}, \ldots, u_{j+k})$ fluxu numerikoa da, $f(u)$ fluxu jarraituaren hurbilketa $g : \mathbb{R}^{2k} \to \mathbb{R}$ funtzio baten bidez. Begira [38] lehenengo eta bigarren ordeneko metodorik klasikoen ikerlanerako eta [67], bereizmen altuko metodoei buruzko sarrerarako.

1.2.3 Lerro-bilaketa metodoen bidezko optimizazioa

Gure munduan optimizazioa funtsezkoa da: jendeak optimizatzen du, naturak optimizatzen du. Lortu nahi den helburu bat aurki dezakegun egoera ugari daude. Matematikaren ikuspuntutik, hori funtzional bat minimizatzearekin (edo maximizatzearekin) bat dator, askotan zenbait murrizketaz mugaturik.

$$\min_{z\in\Omega} f(z), \quad \text{non} \quad \begin{cases} a(z) = 0, \\ b(z) \geq 0. \end{cases}$$

Orokorrean, problema mota hau ezin da *eskuz* ebatzi eta, hortaz, bere ebazpenen hurbilketak kalkulatzeko algoritmo eraginkorrak behar dira. Zabalduenak metodo iteratiboak dira. Hauek Ω multzoko elementuez $\{z_n\}_{n\geq 0}$ segida eraikitzen dute, non

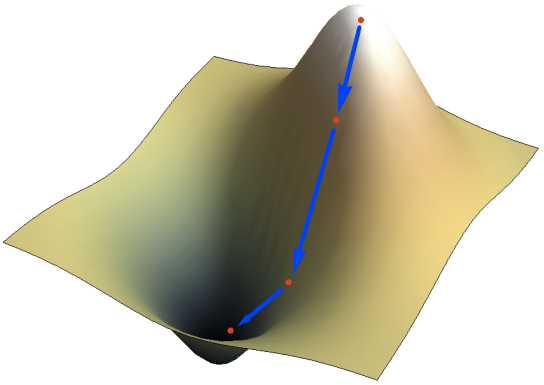
$$\lim_{n\to\infty} z_n = \bar{z}$$

den, $\bar{z}$ optimizazio problemaren ebazpen zehatza izanik.

Tesi honetan lerro-bilaketa metodoak izango ditugu kontuan eta, bereziki, gradiente jaiste-metodoak. Gutxi gorabehera, teknika hauen ideia menditik jaisteko erabiliko genukeen ideia bera da: gure inguruan beherantza doan norabide bat bilatu eta bide horretan pausu bat eman (begira 1.3 Irudia).

Lerro-bilaketa estrategian, z_n ebazpenetik z_{n+1} hurbilketa berria eraikitzeko, z_n puntuan hasten den eta beherantza doan d_n norabidea bilatu behar dugu. Ondoren, norabide horretan hartuko dugun pausuaren ε_n tamaina aukeratuko dugu. Horrezkero, hasierako z_0 susmo batetik, iteratiboki kalkulatuko ditugu hurrengo hurbilketak:

$$z_{n+1} = z_n + \varepsilon_n d_n, \quad n \in \mathbb{N}.$$



IRUDIA 1.3: Gradiente jaiste-metodo bat erabiliz minimizatu.

Metodorik ezagunenak, jaisterik aldapatsuenak, $d_n = -\nabla f(x_n)$ funtzionalaren grandientea hartzen du jaiste-norabidetzat. Pausuaren tamaina z_n puntutik hasita d_n norabide horretan doan lerroan funtzionalaren minimoa ematen duena hartuko dugu. Gradientea norabide negatiboan hartuko dugunez, $\mathcal{J}(z_{n+1}) < \mathcal{J}(z_n)$ ε_n behar bezain txikia bada. Salbuespena iterazio kopuru finituan minimo bat lortzea izango liteke, hau da, $\nabla\mathcal{J}(z_n) = 0$. Orokorrean, minimo baterantz hurbiltzen garen heinean, gradientea zerorantza jotzen du eta, hortaz, segida (agian lokala den) minimo baterantz konbergitzea espero dugu. Begira [19] gradiente jaiste-metodoen oinarrizko teoria aurkitzeko.

Gradiente jaiste-metodoak, jaisterik aldapatsuena edo gradiente konjugatua esaterako, oso erakargarriak dira algoritmo errastasuna dela eta. Aldaketa txikiak eginez ere konbergentzia abiadura hobetu daiteke (begira, adibidez, U. Ascher eta bere kolaboratzaileen lanak [4, 91]). Gainera, metodo hauek ez dute bigarren ordeneko deribatuen beharrik. Beraz, oso egokiak dira eskala handiko optimizazio problemei aurre egiteko. Hala eta guztiz ere, lerro-bilaketaren ideian oinarriturik, beste hainbat metodo garatu dira –adibidez, Newton, quasi-Newton...– hogeigarren mendeko erdialdetik aurrera. Begira [77] lerro-bilaketa estrategien sarrerara eta beren gauzatze konputazionalari buruzko informazioa lortzeko, baita konfiantza-eremuetan oinarrituriko beste optimizazio tresnei buruzkoa ere.

1.3 Tesiaren edukia

Azken atal honetan tesi honetan ikertutako arazoak laburbilduko ditugu. Lau gai nagusi aztertu ditugu:

1. **Biskositate numeriko desagerkorra:** 2. Kapituluan Burgers ekuazio ez-biskosorako eskema numerikoekin lortzen diren ebazpen diskertuen astialdi luzetan duten konportaera aztertuko dugu. Bereziki, eskemek sisteman sartzen duten

biskositate numerikoak dinamika nola eragiten duten frogatuko dugu. Kapituluan lortutako emaitzak L. I. Ignat eta E. Zuazua-rekin idatzitako [84] artikuluan oinarriturik daude.

2. **Kontrol optimoa denbora luzetan:** 3. Kapituluan Burgers ekuazioaren alderantzizko diseinurako zenbait metodo numerikoei buruz dihardugu. Bereziki, kontrola denbora luzean gertatzen den egoera aztertuko dugu, konbergentzia klasikoa soilik ez dela nahikoa azpimarratzen dugu. Eskema numerikoek sistema jarraituaren dinamika alda dezakete eta, beraz, optimizazio numeriko tresnen errendimendua kaltetu. Kapitulu honek N. Allahverdi eta E. Zuazua-ren lankidegoaz egindako [81] artikuluan hartzen du oinarria.

3. **ABE-aren denboraldi luzeko konportaera mantentzen duen eskema numerikoa:** 4. Kapituluan, Burgers ekuazio handituaren parametro konstanteko bertsioarentzako Cauchy arazoa aztertzen dugu, baita denboraldi luzeko dinamika ere. Gero, konportaera hori mantentzen duen eskema numeriko erdi-diskretu bat proposatzen dugu, atal ez-lokalaren diskretizazioa ondo kudeatzeko bi zuzentze-faktore erabiliz. Kapituluan lortutako emaitzak L. I. Ignat-en laguntzaz idatzitako [84] artikuluan oinarriturik daude.

4. **ABE-rako eragile-banaketa:** 5. Kapituluan Burgers ekuazio handiturako eragile-banaketa metodoa aztertuko dugu. Lehen ordeneko konbergentzia frogatu eta hedapen asintotikoaren lehenengo terminoa lortuko dugu. Kapitulu honek L. I. Ignat-en lankidegoaz egindako [83] artikuluan hartzen du oinarria.

Orain gai horien alderdirik garrantzitsuenak deskribatuko ditugu, lortutako emaitzetan eta garatutako metodoetan arreta jarriz.

Oharra 1.1*: Hemendik aurrera, $v = \{v_j\}_{j\in\mathbb{Z}}$ seguida eta $p \in [1,\infty)$ bakoitzeko, hurrengo norma diskretuak erabiliko ditugu:*

$$\|v\|_{p,\Delta} = \Big(\Delta x \sum_{j\in\mathbb{Z}} |v_j|^p\Big)^{1/p}, \qquad \|v\|_{\infty,\Delta} = \max_{j\in\mathbb{Z}} |v_j|, \qquad TV(v) = \sum_{j\in\mathbb{Z}} |v_{j+1} - v_j|.$$

Gainera, $L^p(\mathbb{R})$ espazioko ohizko norma $\|\cdot\|_p$ erabiliz adieraziko dugu.

1.3.1 2. Kapitulua: Biskositate numeriko desagerkorra

Aurreko atalean aipatu dugun arabera, Burgers ekuazio ez-biskosoa eta biskosoaren denboraldi luzeko konportaerak oso ezberdinak dira euren arten. 2. Kapituluko emaitza nagusiak (1.3) Ekuazio hiperbolikoa eskema numerikoak erabiliz hurbiltzerakoan gauza bera gertatzen dela dio [84]. Halere, hau ez luke harrigarria izan behar, konbergitzen

duten eskema numerikoek biskositate numeriko apur bat sartzen baitute sisteman [38]. Gure analisiak eskemak sailkatzeko aukera ematen du: alde batetik, biskositate numeriko hutsala sartzen dutenak –eta, beraz, N-uhina lortzen dutenak– eta, bestetik, biskositate numeriko gehiegi sartzen dutenak –hortaz, auto-antzekoak diren profiletara heltzen direnak–. Aztertuko ditugun eskemen kasuan, Engquist-Osher [26] eta Godunov [39] eskemak lehengo mailan dauden bitartean, Lax-Friedrichs eskema [63] bigarrenean dago. Hala eta guztiz ere, gure analisia Burgers ekuazio ez-biskosora mugaturik egon arren, biskositate fisiko txikia duten beste kontserbazio-legeetara zabaldu daitekeela aipatzekoa da.

2. Kapituluko helburua, behin sarea definitzen duten Δx eta Δt parametroak finkatuz, ebazpen diskretuen $n \to \infty$ doaneko konportaera aztertzea da. Jakina, (1.3) Ekuazioaren ebazpen entropikorantz konbergitzen duten eta CFL baldintza betetzen duten eskema numerikoetan interesaturik gaude soilik. Bereziki, lehen aipatutako hiru puntuko metodo kontserbakorrak hausnartuko ditugu. Izan ere, hirurak alde bateko Lipschitz baldintza (OSLC ingelesez) betetzeagatik ezagunak dira (begira [12] eta bertan dauden erreferentziak). Baldintza hori ezinbestekoa da beherakuntza ezaugarriak ezartzeko.

Zehatzago izatearren, demagun hasierako datuaren hurrengo hurbiltzea hartzen dugula:

$$u_j^0 = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} u_0(x)dx, \quad x_{j+1/2} = (j + \frac{1}{2})\Delta x. \tag{1.9}$$

Hortik u_Δ zatikako funtzioa konstante definituko dugu:

$$u_\Delta(t, x) = u_j^n, \quad x_{j-1/2} < x < x_{j+1/2},\, t_n \le t < t_{n+1}, \tag{1.10}$$

non $t_n = n\Delta t$ eta u_j^n (1.8) erabiliz kalkulatzen den.

Hurrengo teorema, (1.3) Burgers ekuazioa kontuan hartzen duena, 2. Kapituluko emaitza nagusia da.

Teorema 1.1. *Demagun $u_0 \in L^1(\mathbb{R})$ eta $\Delta x, \Delta t > 0$ CFL baldintza betetzen dutela, $\lambda \|u^n\|_{\infty,\Delta} \le 1$ alegia ($\lambda = \Delta t/\Delta x$). Demagun u_Δ (1.8) ematen duen (1.3) Ekuazioaren ebazpenaren hurbilketa dela. Orduan, edozein $p \in [1, \infty)$-rako, hurrengoa bete egiten da:*

$$\lim_{t\to\infty} t^{\frac{1}{2}(1-\frac{1}{p})} \|u_\Delta(t) - w(t)\|_p = 0, \tag{1.11}$$

non w profila horrelakoa den:

1. *Lax-Friedrichs eskemarekin, $w = w_{M_\Delta}$ hurrengo Burgers ekuazio biskosoaren ebazpen bakarra da:*

$$\begin{cases} w_t + \left(\frac{w^2}{2}\right)_x = \frac{(\Delta x)^2}{2\Delta t} w_{xx}, \quad x \in \mathbb{R}, t > 0, \\ w(0) = M_\Delta \delta_0, \end{cases} \tag{1.12}$$

$M_\Delta = \int_{\mathbb{R}} u_\Delta^0$ *izanik.*

2. *Engquist-Osher eta Godunov eskemekin* $w = w_{p_\Delta, q_\Delta}$ *hurrengo Burgers ekuazio hiperbolikoaren ebazpen bakarra da:*

$$\begin{cases} w_t + \left(\frac{w^2}{2}\right)_x = 0, \quad x \in \mathbb{R}, t > 0, \\ \\ w(0) = M_\Delta \delta_0, \quad \lim\limits_{t \to 0} \int_{-\infty}^{x} w(t,z) dz = \begin{cases} 0, & x < 0, \\ -p_\Delta, & x = 0, \\ q_\Delta - p_\Delta, & x > 0, \end{cases} \end{cases} \tag{1.13}$$

$M_\Delta = \int_{\mathbb{R}} u_\Delta^0$ *eta*

$$p_\Delta = -\min_{x \in \mathbb{R}} \int_{-\infty}^{x} u_\Delta^0(z) dz \quad \text{eta} \quad q_\Delta = \max_{x \in \mathbb{R}} \int_{x}^{\infty} u_\Delta^0(z) dz.$$

izanik.

Aurreko (1.12) eta (1.13) Ekuazioen hasierako datua neurri mugatuen konbergentziaren zentsuan ulertu beharra dago. [28] eta [73] kontsultatu definizio zehatzagoa izateko.

Edozein kasutan, goian emandako profilaz esplizituki idatzi daitezkeela ezaguna da [29, 30]:

$$w_{M_\Delta}(x,t) = -\frac{2\sqrt{\nu}}{t^{1/2}} \exp\left(-\frac{x^2}{4\nu t}\right) \left[C_{M_\Delta} + \int_{-\infty}^{x/\sqrt{\nu t}} \exp\left(-\frac{s^2}{4}\right) ds \right]^{-1}, \tag{1.14}$$

non $\nu = \Delta x^2/(2\Delta t)$ eta C_{M_Δ} w_{M_Δ}-ren masa M_Δ izateko hartzen den, eta

$$w_{p_\Delta, q_\Delta}(x,t) = \begin{cases} \frac{x}{t}, & -\sqrt{2p_\Delta t} < x < \sqrt{2q_\Delta t}, \\ 0, & \text{beste edonon.} \end{cases} \tag{1.15}$$

(1.14) profila hasierako datuaren masak zehazturik dago. Ostera, (1.15) N-uhinak denboran zehar konstante mantentzen diren bi parametro behar ditu, p eta q. Bien arteko diferentzia 1.2 Irudian ikus daiteke. Bertan, $t = 50$, $\Delta x = 1/200$, $\Delta t = 1/200$, $M_\Delta = 0.15$, $p_\Delta = 0.05$ eta $q_\Delta = 0.2$ hartu dugu.

1.1 Teorema frogatzeko, eskala argudioak erabiliko ditugu, kasu jarraiaren parekoak. Gainera, deribatu partzialetako ekuazioen analisirako tresna arrunta diren aldagai autoantzekoez baliatuko gara zenbait fenomeno argiago irudikatzeko.

1.3.2 3. Kapitulua: Kontrol optimoa denboraldi luzetan

3. Kapituluan Burgers ekuazioa dakarren alderantzizko diseinurako hurbilketa numerikoa aztertzen dugu, bai (1.4) kasu biskosoan baita (1.3) ez-biskosoan ere. Soinu-eztandaren ereduan gertatzen den bezala, $T > 0$ denbora eta u^* jomuga funtzio emanda, u_0 hasierako datua aurkitzea da helburua, ekuazioaren ebazpenak $t = T$ denboran u^* ahalik eta ondoen hurbil dezan.

Izatez, Burgers ekuazioa atzerantz ebazten datza problema, txarto jarrita dagoen arazoa izanik. $\nu > 0$ kasu biskosoan, Burgers ekuazio parabolikoaren berezko denbora-itzuliezintasunak eraginda dago. Kasu ez-biskoso hiperbolikoan, berriz, ereduaren ez-linealtasunak talka-uhinen agerraldia sortzen du eta, beraz, ebazpen anitzak dira posible.

Arazoa kontrol optimoaren ikuspuntutik formulatu dugu. Karratu txikienen metodoa jarraituz, hurrengo funtzionala minimizatzen saiatuko gara:

$$\mathcal{J}(u_0) = \frac{1}{2} \int_{\mathbb{R}} \left(u(x,T) - u^*(x)\right)^2 dx, \tag{1.16}$$

non u Burgers ekuazioaren ebazpena den eta non u_0 hasierako datua espazio funtzional aproposean dagoen, adibidez, $L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$.

Izan ere, arazo hau, soinu-eztanda eredu osoaren sinplifikazioa, naturalki agertzen da. Asmo praktikoagatik, osagairik nagusienetariko bat $[0,T]$ denboraldiak luze izan behar duela da. Beraz, deribatu partzialetako ekuazioa (DPE) hurbiltzeko eskema numerikoaren hautaketa kontu korapilatsua bihurtzen da, denboraldi luzeko dinamika emateko gaitasuna ez duten eskemek kontrol optimoaren hurbilketa okerra lortzen baitute. Aurreko atalean aipatu dugun arabera, paraboliko/hiperboliko dikotomia asintotikoa arreta handiz maneiatu beharra dago. Kasu hiperbolikoan gehiegizko biskositate numerikoak dinamika hiperbolikoa hondatu eta paraboliko bihurrarazi dezake, 2. Kapituluan erakusten dugun arabera, eta Burgers ekuazio biskosoan antzeko gauza gerta daitekeela frogatuko dugu.

Lan honetan egoera horrek alderantzizko diseinuaren arloan izan ditzakeen ondorioak azpimarratuko ditugu. Horretarako, adjuntuaren metodoan oinarrituriko gradiente jaiste-metodo bat eta IPOPT (optimizazio ezlineala egiteko iturburu-kode irekia

[92]) erabiliko ditugu. Halere, kontuan izan aipatutako dikotomia beste tresna numerikoetara zabal daitekeela. [19, 77] kontsultatu optimizazio numerikoaren arlo honi buruzko funtsezko teoria aurkitzeko.

Kontrol optimo arazo hau, denboraldi luzearen eskakizuna barik, egile askok ikasia izan da, bai maila jarraituaren ikuspuntutik baita diskretuarenetik ere. Aipatu dugun arabera, diskretizazioaren eta optimizazioaren arteko elkarrekintzak oso sustagarri bihurtzen du gai hau. Kontserbazio-lege eskalarren testuinguruan, [7, 10, 55, 90] lanak aipatekoak dira. F. James eta bere kolaboratzaileek luze jardun dute sistema adjuntuaren eta bere diskretizazioaren inguruan ere [9, 40, 56]. Izan ere, talka-uhinen agertzea ez da kontu makala problema hauetan. Esate baterako, [35, 36] artikuluetan egileek sistema linealizatuaren eta sistema adjuntuaren hurbilketen puntukako konbergentzia aztertzen dute hasierako datua jarraitua ez denean. Bereziki, biskositate numerikoak dakarren arazoa aipatzen dute dagoeneko. Hala eta guztiz ere, gure emaitzek T denbora handia denean euren ekarpena eskasa izan daitekeela erakusten dute.

Gainera, optimizatu-gero-diskretizatu eta diskretizatu-gero-optimizatu ikuspuntuak [89]-n aztertuak izan dira. 3. Kapituluan bigarrena erabiliko dugu. Izan ere, etenak dituzten ebazpenak saihestuko ditugu, optimizazio prozesuan duten inpaktua murriztearren eta denboraldi luzearen ondorioan soilik arreta jartzearren. Horregatik, gure kasuan talka-uhinen sentikortasunak ez dauka garrantzi handirik.

3. Kapituluko emaitzek abisu garrantzitsua osatzen dute, denboraldi luzeko dinamika jarraituaren ezaugarriak imitatzeko gai diren eskema numerikoen beharraz ohartaraziz. Hau uhinen hedapenaren kontrol problemetan nabaria da [27, 95]. Sendoago dirudien fluido biskoso eta ez-biskosoen alderantzizko diseinuan patologia berbera jarraitzea oso interesgarria da.

1.3.3 4. Kapitulua: ABE-aren denboraldi luzeko konportaera mantentzen duen eskema numerikoa

4. Kapitulua parametro konstanteak eta lasaikuntza prozesu bakarra duen Burger ekuazio handituaz doa. Hurrengo ekuazioan jartzen dugu arreta:

$$\begin{cases} u_t = uu_x + \nu\, u_{xx} + c\, K_\theta * u_{xx}, & (t,x) \in (0,\infty)\times\mathbb{R}, \\ u(0,x) = u^0(x), & x \in \mathbb{R}, \end{cases} \tag{1.17}$$

non $* x$ aldagaiarekiko konboluzioa adierazten duen, ν, c, θ parametroak positiboak diren eta

$$K_\theta(z) = \begin{cases} \frac{1}{\theta} e^{-z/\theta}, & z > 0, \\ 0, & \text{beste edonon.} \end{cases} \tag{1.18}$$

2. eta 3. Kapituluetan aztertutako patologiak Burgers ekuazio handituaren testuinguruan ere agertzen dira. (1.17) Ekuazioko ν eta c parametroentzako balio txikiek kontserbazio-lege hiperbolikoaren tratamendu bera behar dute, numerikoaren ikuspuntutik begiratuta. Hortaz, 4. Kapituluan (1.17) Ekuazioaren ebazpenen konportaera asintotikoa, $t \to \infty$ ahala, ikertuko dugu eta konportaera hori gordetzen duen eskema numeriko erdi-diskretua eraikiko dugu.

(1.17) Sistemako ebazpenen denboraldi luzeko konportaerari dagokionez, emaitzarik nagusiena hurrengo teorema da.

Teorema 1.2. *Demagun $u^0 \in L^1(\mathbb{R})$ dela. Orduan, edozein $p \in [1, \infty]$ hartuz,* (1.17) *Ekuazioaren u ebazpenak*

$$t^{\frac{1}{2}(1-\frac{1}{p})} \|u(t) - u_M(t)\|_p \longrightarrow 0, \quad t \to \infty \text{ ahala,}$$

betetzen du, non $u_M(t,x)$ hurrengo ekuazioaren ebazpena den:

$$\begin{cases} u_t = uu_x + (\nu + c)u_{xx}, & x \in \mathbb{R}, t > 0, \\ u(0) = M\delta_0. \end{cases}$$

Hemen δ_0 jatorriko Dirac-en delta adierazten du eta M, hasierako datuaren masa, $M = \int_{\mathbb{R}} u^0(x)dx$ alegia.

Kontuan izan u_M (1.14) formulan esplizituki definiturik dagoela, biskositate parametrotzat $\nu + c$ hartuz ν-ren ordez. Izan ere, honek u_M $t^{-\frac{1}{2}} f_M\left(\frac{x}{\sqrt{t}}\right)$ itxura daukala erakusten du, nolabaiteko f_M funtzio batekin, eta, beraz, auto-antzekoa.

Guzti hau maila numerikoan garrantzi handia dauka. Alde batetik, sartzen dugun biskositate numerikoa arretaz maneiatu behar dugu, ez-linealtasuna diskretizatzeko fluxu numerikoa aukeratzerakoan. Bestetik, atal integrala etentzeko orduan, konportaera asintotikoan opa ez diren patologiak ekiditu behar dira. Horretarako bi zuzentze-faktore proposatzen ditugu.

Demagun u_Δ (1.17) Ekuazioaren u ebazpenaren hurbilketa bat dela. Espazioan zatikako funtzio hau horrelaxe definituko dugu:

$$u_\Delta(t,x) = u_j(t), \quad x \in (x_{j-1/2}, x_{j+1/2}), t \geq 0, \tag{1.19}$$

non $x_{j+1/2} = (j+\frac{1}{2})\Delta x$ den, $j \in \mathbb{Z}$ guztientzako, eta $\Delta x > 0$ emandako mailaren tamaina den. Gelaxken erdiko puntuak $x_j = j\Delta x$ adieraziko ditugu. $j \in \mathbb{Z}$ bakoitzeko, gelaxkako ebazpenaren balioa hurbiltzen duen $u_j(t)$ funtzio bat aurkitu behar dugu. Orain arte aipatutako arazoak kontuan izanik, (1.17) Ekuazioaren hurrengo diskretizazioa aukeratu dugu: Engquist-Osher eskema fluxurako, diferentzia finitu zentratuak laplacetarrarako eta laukizuzenaren erregela konposatua integralerako:

$$\begin{cases} u_j'(t) = \frac{g_{j+1/2}(t)-g_{j-1/2}(t)}{\Delta x} + \nu\frac{u_{j-1}(t)-2u_j(t)+u_{j+1}(t)}{\Delta x^2} \\ \qquad\qquad + \frac{c}{\theta^2}\left(\sum_{m=1}^{N}\omega_m^\theta u_{j-m}(t) - F_0^{\Delta,\theta}u_j(t) + F_1^{\Delta,\theta}\theta\frac{u_{j+1}(t)-u_j(t)}{\Delta x}\right), \quad j\in\mathbb{Z}, t\geq 0, \\ u_j(0) = \dfrac{1}{\Delta x}\displaystyle\int_{x_{j-1/2}}^{x_{j+1/2}} u^0(x)dx, \qquad\qquad j\in\mathbb{Z}, \end{cases} \tag{1.20}$$

non

$$\omega_m^\theta = e^{-m\Delta x/\theta}\left(e^{\Delta x/\theta}-1\right), \quad m=1,\dots,N, \tag{1.21}$$

eta

$$g_{j+1/2}(t) = \frac{u_j(t)\big(u_j(t)-|u_j(t)|\big)}{4} + \frac{u_{j+1}(t)\big(u_{j+1}(t)+|u_{j+1}(t)|\big)}{4}, \quad j\in\mathbb{Z}, t\geq 0.$$

$N = N(\Delta x) \in \mathbb{N}$ zenbakiak integralaren koadratura formularen nodo kopurua adierazten du. u eta u_x hurbilketen aurrean agertzen diren $F_0^{\Delta,\theta}$ eta $F_1^{\Delta,\theta}$ zuzentze-faktoreak,

$$F_0^{\Delta,\theta} = \sum_{m=1}^{N}\omega_m^\theta \quad \text{eta} \quad F_1^{\Delta,\theta} = \frac{\Delta x}{\theta}\sum_{m=1}^{N} m\omega_m^\theta \tag{1.22}$$

emanda, atal ez-lokalaren etendura zuzena egiteaz arduratzen dira, konportaera asintotikoaren ikuspuntutik.

Azkenik, Δx finkorako, u_Δ ebazpen semi-diskretu hauen konpotaera asintotikoa, $t \to \infty$ ahala, aztertuko dugu.

Teorema 1.3. *Demagun $u_0 \in L^1(\mathbb{R})$ eta $\Delta x > 0$ direla eta u_Δ (1.20) eskemak emandako* (1.17) *Burgers ekuazio handiatuaren ebazpena hurbiltzeari dagokion ebazpena dela. Orduan, edozein $p \in [1,\infty]$ hartuz, hurrengoa betetzen da:*

$$t^{\frac{1}{2}(1-\frac{1}{p})}\|u_\Delta(t) - u_M(t)\|_p \longrightarrow 0, \quad t\to\infty \text{ ahala}, \tag{1.23}$$

non $u_M(t,x)$ hurrengo Burgers ekuazioaren ebazpen bakarra den:

$$\begin{cases} v_t = vv_x + (\nu + c\,F_2^{\Delta,\theta})v_{xx}, \quad x\in\mathbb{R}, t>0, \\ v(x,0) = M\delta_0. \end{cases}$$

Hemen $M = \int_{\mathbb{R}} u^0(x)dx$ *hasierako datuaren masa da eta*

$$F_2^{\Delta,\theta} = \frac{\Delta x^2}{2\theta^2}\left(\sum_{m=1}^{N} m(m-1)\omega_m^\theta\right).$$

Kontuan izan, edozein $\theta > 0$ finkorako, $\Delta x \to 0$ hartzean $N \to \infty$ eta $N\Delta x \to \infty$ betetzen duen N hartzen badugula, orduan $F_2^{\Delta,\theta} \to 1$. Hori da, hain zuzen ere, eredu jarraituak daukan balioa.

1.3.4 5. Kapitulua: ABE-rako eragile-banaketa

(1.20) kalkulatzea oso garestia izan daiteke, ikuspuntu konputazionalaren aldetik, N handia bada. Beraz, 5. Kapituluan Burgers ekuazio handitua ebazteko eragile-banaketa erabili ahal izateko esparrua zehaztuko dugu, denboraldi luzeko konportaeran azpimarratuz. Horiek horrela, teknika hau soinu-eztandaren testuinguruan dagoeneko erabilia izan dela aipatzekoa da [20, 85, 86]. Gainera, metodo honen hainbat bertsio garatu dira, esate baterako, Schrödinger ekuazio ez-linealerako, Korteweg-de Vries ekuaziorako, Boltzmann ekuaziorako... (begira [44] eta bertan agertzen diren erreferentziak) eta, berriki, Fowler ekuazio ez-lokalerako [11].

Eragile-banaketaren oinarrizko ideia eboluzio eragile orokorra, formalki, beste eboluzio eragileren batuketa bezala idaztean datza. Beste hitzetan, sistema osoa sinpleagoak diren zenbait azpi-ekuaziotan banatu egiten da, algoritmo praktikoagoak erabiltzera bideratuz. Behin azpi-ekuazio horiek ebatzita, soluzioak bateratu egiten dira eredu osoaren ebazpena lortzeko. Begira [44] metodo hauen sarrera sakonagorako.

(1.17) Burgers ekuazio handiturako hurrengo Trotter formula erabiliko dugu. Argiago izateko, $\nu = c = \theta = 1$ kasua soilik aztertuko dugu, baina lortuko ditugun emaitzak erraz zabaldu daiteke kasu orokorrera. Demagun X^t

$$\begin{cases} v_t = K * v - v + v_x, & (t,x) \in (0,\infty)\times\mathbb{R}, \\ v(t=0,x) = v_0(x), & x \in \mathbb{R}, \end{cases}$$

sistemari dagokion eboluzio eragilea dela eta, Y^t

$$\begin{cases} w_t - \left(\frac{w^2}{2}\right)_x = w_{xx}, & (t,x) \in (0,\infty)\times\mathbb{R}, \\ w(t=0,x) = w_0(x), & x \in \mathbb{R}. \end{cases}$$

sistemari dagokiona. Hortaz, Z^t fluxua horrelaxe definituko dugu:

$$Z^t = X^t Y^t. \tag{1.24}$$

Kapituluaren helburua

$$(Z^{\Delta t})^n u_0 = (X^{\Delta t} Y^{\Delta t})^n u_0,$$

erabiliz (1.17) ekuazioaren u ebazpena hurbiltzea da, non $\Delta t > 0$ eta $n \in \mathbb{N} \cup \{0\}$. Notazio sinplea erabiltzearren, hemendik aurrera $Z^{n\Delta t} = (Z^{\Delta t})^n$ idatziko dugu. Gogoratu $t_n = n\Delta t$ eta $t_{n+1/2} = (n + \frac{1}{2})\Delta t$, $n \in \mathbb{N} \cup \{0\}$ bakoitzeko, hartzen dugula beti.

5. Kapituluko lehenengo emaitza hurrengo teorema da. Leun nahikoak diren ebazpenentzako (1.24) eragile-banaketa lehenengo ordenekoa dela egiaztatzen du.

Teorema 1.4. *Demagun $r \in \{1, 2\}$ dela. Orduan, $u_0 \in H^r(\mathbb{R})$ guztietarako eta $T > 0$ guztietarako, c_1, c_2 eta Δt_0 konstante positiboak existitzen dira, hurrengo hau betetzen dutena*

$$\left\| Z^{n\Delta t} u_0 - u(n\Delta t) \right\|_2 \leq c_1 (\Delta t)^{r/2}$$

eta

$$\left\| Z^{n\Delta t} u_0 \right\|_{H^r(\mathbb{R})} \leq c_2,$$

$\Delta t \in (0, \Delta t_0)$ guztietarako eta $0 \leq n\Delta t \leq T$ betetzen duten $n \in \mathbb{N}$ guztietarako. Hemen c_1, c_2 eta Δt_0 konstanteek T eta $\|u_0\|_{H^r(\mathbb{R})}$-rekiko dependentzia dute.

4. Kapituluko antzeko teknikak erabiliz, Z^t eragileak emandako ebazpenaren hedapen asintotikoaren lehenengo terminoa ere lortuko dugu. Horretarako, u^Δ horrela definituko dugu:

$$u^\Delta(t,x) = \begin{cases} Y^{2(t-t_n)} Z^{n\Delta t} u_0(x), & t \in (t_n, t_{n+1/2}),\ x \in \mathbb{R}, \\ X^{2(t-t_{n+1/2})} Y^{\Delta t} Z^{n\Delta t} u_0(x), & t \in (t_{n+1/2}, t_{n+1}),\ x \in \mathbb{R}. \end{cases}$$

Beraz, u^Δ hurrengo sistema beteko du:

$$\begin{cases} u_t^\Delta = 2(Y^{2(t-t_n)} Z^{n\Delta t} u_0)_t = 2u_{xx}^\Delta + \left((u^\Delta)^2\right)_x, & t \in (t_n, t_{n+1/2}), \\ u_t^\Delta = 2(K * u^\Delta - u^\Delta + u_x^\Delta), & t \in (t_{n+1/2}, t_{n+1}), \\ u^\Delta(t_n) = Z^{n\Delta t} u_0, \\ u^\Delta(t_{n+1/2}) = Y^{\Delta t} Z^{n\Delta t} u_0. \end{cases}$$

Orain, $I_n^\Delta = (t_n, t_{n+1/2})$ hartzen badugu, $\psi^{\Delta t}(t) = \sum_{n \geq 0} \chi_{I_n^\Delta}(t)$ funtzioa defini dezakegu (t $(t_n, t_{n+1/2})$ tartean badago, 1 balioko du eta, ostera, ez badago, 0). Argi dago, orduan,

aurreko sistema eraldatu dezakegula,

$$\begin{cases} u_t^\Delta = 2\psi^{\Delta t}(t)u_{xx}^\Delta + \psi^{\Delta t}(t)((u^\Delta)^2)_x \\ \qquad\qquad +2\left(1-\psi^{\Delta t}(t)\right)(K*u^\Delta - u^\Delta + u_x^\Delta), \quad t>0, \\ u^\Delta(0) = u_0, \end{cases} \tag{1.25}$$

lortuz. [23] artikuluan $\psi^{\Delta t}(t) \rightharpoonup \frac{1}{2}$ frogatzen da. Bestalde, X^t eta Y^t eragileen ezaugarriak kontuan izanik, badakigu (1.25) ekuazioak definitzen duen u^Δ funtzioak $u^\Delta \in C([0,\infty);L^1(\mathbb{R}))$ betetzen duela. Horregatik, formalki $u^\Delta \to u$ $\Delta t \to 0$ ahala, non u (1.17) betetzen duen. $t\to\infty$ heineko u^Δ ebazpenaren konportaera, hurrengo teoremak adierazita, eskala argudio bat jarraituz lortuko dugu.

Teorema 1.5. *Edozein $u_0 \in L^1(\mathbb{R})\cap L^\infty(\mathbb{R})$ eta $p\in[1,\infty)$ ahrtuz,*

$$\left\|u^\Delta(t)\right\|_p \le C_p t^{-\frac{1}{2}(1-\frac{1}{p})}\|u_0\|_1, \quad n\ge 1.$$

Gainera, edozein $p\in[1,\infty)$ baliorako,

$$t^{\frac{1}{2}(1-\frac{1}{p})}\left\|u^\Delta(t)-u_M(t)\right\|_p \to \infty, \quad t\to\infty \text{ ahala,}$$

non $u_M(t,x) = t^{-1/2}f_M(x/\sqrt{t})$ hurrengo Burgers ekuazio biskosoaren profil auto-antzekoa den

$$\begin{cases} u_t = \left(\frac{u^2}{2}\right)_x + 2u_{xx}, & (t,x)\in(0,\infty)\times\mathbb{R}, \\ u(0,x) = M\delta_0, & x\in\mathbb{R}. \end{cases}$$

Kontuan izan hori dela, izan ere, (1.17) Ekuazioaren ebazpenaren denboraldi luzeko konportaera, 4. Kapituluan erakutsi dugun arabera.

Capítulo 1

Introducción

Hoy en día, una de las mayores preocupaciones en la investigación aeronáutica es el control y la reducción del sonido generado por las aeronaves. De hecho, uno de los principales objetivos de esta vasta e importante área de actividad industrial y comercial consiste en cumplir con las estrictas restricciones de ruido y diseño para aviones supersónicos. En particular, la minimización de la explosión sónica generada por estas aeronaves es el punto clave para tener éxito en el desarrollo de transporte supersónico civil eficiente [3].

Como veremos, este tipo de asuntos requiere herramientas numéricas que funcionen bien con problemas de evolución con horizontes temporales lejanos. En esta tesis enfatizaremos la necesidad de emplear esquemas de aproximación numérica que preserven las propiedades dinámicas del sistema continuo en tiempos grandes. Aplicaremos esto en el caso particular de la predicción y el control de la explosión sónica.

1.1 Minimización de la explosión sónica

Al volar por encima de la velocidad del sonido, los aviones supersónicos crean una perturbación de la presión en la atmósfera, como resultado del desplazamiento del aire al atravesarlo y de la generación de la sustentación, necesaria para seguir el trayecto deseado. Muchas de estas perturbaciones, que conllevan una significante cantidad de energía acústica, alcanzan el nivel del suelo, dando como resultado la llamada explosión sónica. Este fenómeno se percibe como dos grandes estallidos consecutivos separados por un breve instante de tiempo. Esto puede resultar perjudicial para el ser humano y para las estructuras de los edificios, por lo que ha sido uno de los mayores obstáculos a la hora de diseñar aeronaves supersónicas.

El primer avión pilotado por una persona que superó la barrera del sonido fue el Bell X-1, en 1947 [24, Chapter 10]. Después de eso, en las décadas de los 60 y 70, los esfuerzos

se centraron en construir este tipo de transporte para uso civil. Al principio, fueron dos los proyectos que tuvieron éxito: el Concorde, construido por franceses e ingleses, y el Tupolev-144, por los soviéticos. Otros dos proyectos estadounidenses, el Lockheed L-2000 y el Boeing 2707, se cancelaron debido a las restricciones medioambientales y acústicas impuestas por el Congreso de los Estados Unidos. Esta prohibición no permitía alcanzar velocidades supersónicas al sobrevolar áreas urbanas (en la práctica, cualquier lugar salvo el océano). Esto hizo que todas las perspectivas económicas fueran en balde, condenando al fracaso a todos los proyectos de transporte civil supersónico.

Así pues, durante las últimas décadas del siglo XX, quedó claro que era necesario aplicar tecnologías de reducción de la explosión sónica, de cara a alcanzar el objetivo de construir aeronaves eficientes que evitaran ese problema de ruido. De hecho, el proyecto DARPA/NASA/Northrop-Grumman Shaped Sonic Boom Demonstrator (SSBD) demostró que la explosión sónica podía ser mitigada parcialmente mediante la adaptación de la forma de la aeronave [79], siendo la primera vez que se confirmaba experimentalmente.

Ya entonces era bien sabido que la firma de sonido en el campo cercano evoluciona hacia una onda-N (véase Imagen 1.1). La onda-N (textitN-wave en terminología inglesa) se refiere a la forma que resulta del colapso de las múltiples ondas de choque, formando un choque frontal y otro trasero separados por una expansión casi lineal de la presión. Al principio se creía que esta onda-N era inevitable. Pero McLean probó que es posible modificar la firma sonora del campo cercano de manera que no se haya alcanzado la onda-N para cuando el sonido llega al nivel del suelo [74]. No obstante, tal y como hemos mencionado anteriormente, esto no se verificó empíricamente hasta las pruebas realizadas en el proyecto SSBD en 2003.

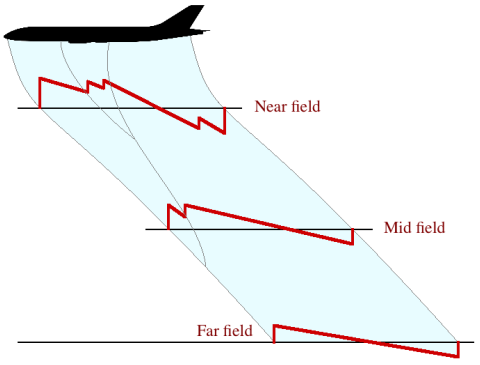


Imagen 1.1: Diagrama de la propagación de la explosión sónica. Los choques creados en el campo cercano colapsan, formando una onda-N en el campo lejano, con dos choques únicamente.

Históricamente, la teoría lineal fue la más usada para modelar la propagación de la explosión sónica, siguiendo los trabajos de Hayes [43] y Whitham [93] a finales de los 40 y principios de los 50. Gran parte de la investigación analítica y numérica se realizó asumiendo tal simplificación. Sin embargo, recientemente, se han propuesto nuevos modelos no lineales para mejorar los métodos existentes. Cabe destacar dos de estos métodos. Por un lado, Ozcer [78] propuso el uso de la ecuación potencial completa en la región entre el campo cercano y el suelo. Por otro, en el contexto de trazado de rayos/acústica geométrica, Rallabhandi empleó una ecuación de Burgers aumentada (ABE, por sus siglas en inglés) para propagar la firma fuente hacia el suelo, incluyendo efectos viscosos que conducen a discontinuidades de choque de grosor no nulo [85].

En esta tesis nos centraremos en el segundo modelo. La ecuación de Burgers aumentada tiene en cuenta efectos no lineales, tales como el aumento de la pendiente de la onda y la propagación con velocidad variable, así como fenómenos de relajación molecular, la extensión en tubos de rayos y la estratificación atmosférica. La terminología no es consistente; dependiendo de la fuente, también puede encontrarse como ecuación de Burgers extendida [20]. La ecuación viene dada por:

$$\frac{\partial P}{\partial \sigma} = P\frac{\partial P}{\partial \tau} + \frac{1}{\Gamma}\frac{\partial^2 P}{\partial \tau^2} + \sum_{\nu} C_\nu \frac{\frac{\partial^2 P}{\partial \tau^2}}{1+\theta_\nu \frac{\partial}{\partial \tau}} - \frac{1}{2G}\frac{\partial G}{\partial \sigma}P + \frac{1}{2\rho_0 c_0}\frac{\partial(\rho_0 c_0)}{\partial \sigma}P, \tag{1.1}$$

donde $P = P(\sigma, \tau)$ es la perturbación adimensional de la distribución de presiones. La distancia de propagación σ y el tiempo de propagación τ también son adimensionales. El operador que aparece en el sumatorio correspondiente a las relajaciones moleculares se define así:

$$\frac{\theta_\nu}{1+\theta_\nu \frac{\partial}{\partial \tau}} f(\tau) = \int_{-\infty}^{\tau} e^{(\xi-\tau)/\tau_\nu} f(\xi) d\xi.$$

La función G indica el área del tubo de rayos. Las condiciones de la atmósfera están dadas por su densidad ρ_0, la velocidad del sonido c_0, un parámetro termo-viscoso Γ y un tiempo de relajación adimensional θ_ν y un parámetro de dispersión adimensional $\mathbb{C}_\nu$ para cada modo de relajación (habitualmente, uno para el oxígeno y otro para el nitrógeno).

En ese caso, el problema de la minimización de la explosión sónica consiste en, dada la firma sonora P^* deseada en el nivel del suelo y la duración de la propagación Σ –que está estrechamente relacionada con la altura del vuelo–, recuperar la firma sonora del campo cercano que mejor la reproduzca. Esto se suele formular desde el punto de vista del control óptimo, por medio de una aproximación por mínimos cuadrados [86], usando el siguiente funcional:

$$\mathcal{S}(P_0) = \frac{1}{2}\int_{\mathbb{R}} \big(P(\Sigma, \tau) - P^*(\tau)\big)^2 d\tau. \tag{1.2}$$

Aquí P_0 se toma en un conjunto de firmas del campo cercano admisibles –que está usualmente restringido por las variables de diseño asociadas a la geometría de la aeronave [76, 86]– y P es la solución de (1.1) con $P(0,\tau) = P_0(\tau)$ para todo $\tau \in \mathbb{R}$. El volumen percibido (PLdB) o los choques de sobre-presión [76] son otros posibles funcionales que también suelen considerarse.

En el problema de la minimización del sonic boom las escalas de tiempo son muy distintas: la perturbación de la presión tiene lugar durante menos de un segundo, mientras que la propagación puede durar hasta un minuto, dependiendo de las condiciones del vuelo [85]. Como veremos más adelante, esto hace que su tratamiento computacional sea una ardua tarea y sirve de motivación para el estudio del comportamiento asintótico en tiempos grandes de los esquemas numéricos.

1.2 Contexto matemático

En esta sección recogemos algunas de las herramientas matemáticas que aparecerán a lo largo de esta tesis. Como la ecuación de la explosión sónica es una ecuación de tipo Burgers, primero presentaremos algunas propiedades sobre el comportamiento de las soluciones de la ecuación de Burgers. Después, siendo esta una ley de conservación escalar, recordaremos el método clásico de discretizar este tipo de ecuaciones mediante volúmenes finitos. Concluiremos con una breve introducción a los métodos de búsqueda lineal y, en particular, a los métodos del gradiente descendente, ampliamente empleados para la minimización de funcionales.

1.2.1 Ecuación de Burgers y su dinámica en tiempos grandes

La ecuación de Burgers aumentada (1.1) es una de las muchas variaciones existentes de la ecuación de Burgers. Junto con la teoría lineal, las ecuaciones de tipo Burgers han constituido una de las principales herramientas para modelar la propagación de ondas planas de amplitud finita. La clásica ecuación de Burgers viscosa fue considerada por primera vez para la propagación de ondas en medios disipativos. Sucesivas generalizaciones incluyeron otros efectos tales como la expansión geométrica y medios no homogéneos (ecuación de Burgers generalizada [16, 31, 69]) o procesos de relajación (ecuación de Burgers aumentada [80]). Tal y como hemos explicado en la sección anterior, la ecuación de Burgers aumentada ha sido usada recientemente para modelar la propagación de la explosión sónica producida por las aeronaves supersónicas [85], teniendo en cuenta todos los fenómenos mencionados.

La versión no viscosa de la ecuación de Burgers es una ley de conservación escalar, dada por

$$u_t + \left(\frac{u^2}{2}\right)_x = 0, \quad x \in \mathbb{R}, t > 0, \tag{1.3}$$

mientras que la versión viscosa es una ecuación de convección-difusión unidimensional:

$$u^\nu_t + \left(\frac{(u^\nu)^2}{2}\right)_x = \nu u^\nu_{xx}, \quad x \in \mathbb{R}, t > 0, \tag{1.4}$$

donde $\nu > 0$. Esta última ecuación fue introducida por Bateman [8] y Burgers [15], quien fue el primero en considerarla como una simplificación de las ecuaciones de Navier-Stokes. Es el modelo más simple que reúne propagación no lineal y efectos difusivos [94]. A pesar de la no linealidad, Hopf [45] y Cole [22] hallaron, de manera independiente, que su solución puede ser calculada de manera explícita, por medio de la llamada transformación de Hopf-Cole.

Es bien sabido que el comportamiento asintótico de las soluciones de la ecuación de Burgers no viscosa (1.3) son de naturaleza auto-semejante (véase [72] y las referencias ahí incluidas). De hecho, cuando $t \to \infty$, las soluciones desarrollan un comportamiento de tipo onda-N, conservando la masa del dato inicial, que no varía con la evolución. Sin embargo, nótese que la masa no es suficiente para identificar el perfil asintótico auto-semejante [73], que pertenece a una familia de soluciones biparamétrica. Estos parámetros se corresponden con dos invariantes del sistema; concretamente, las masas positiva y negativa:

$$p = -\min_{x\in\mathbb{R}} \int_{-\infty}^{x} u^0(z)dz \quad \text{y} \quad q = \max_{x\in\mathbb{R}} \int_{x}^{\infty} u^0(z)dz. \tag{1.5}$$

En particular, las ondas-N correspondientes a soluciones que surgen de datos iniciales que cambian de signo también cambian de signo.

El comportamiento asintótico de la versión viscosa (1.4) difiere notablemente. De hecho, para $\nu > 0$ estos problemas son de naturaleza parabólica y, cuando $t \to \infty$, las soluciones se comportan de una manera auto-semejante como un perfil viscoso de signo constante que está completamente determinado por la masa conservada (véase [45]):

$$M = q - p = \int_{\mathbb{R}} u^0(x)dx.$$

En la Imágen 1.2 se comparan los dos perfiles asintóticos para los mismos parámetros de los datos iniciales.

Naturalmente, en tiempo finito, las soluciones del modelo viscoso (1.4) convergen a las soluciones entrópicas de la ley de conservación escalar hiperbólica (1.3) cuando $\nu \to 0$

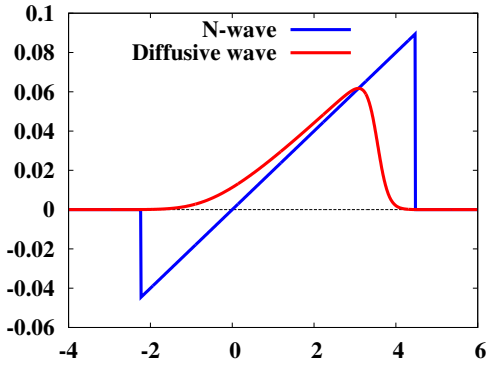


IMAGEN 1.2: Perfiles asintóticos para el caso parabólico $\nu > 0$ (rojo) y el hiperbólico $\nu = 0$ (azul). Aquí la masa total es $M = 0.15$, y las masas negativa y positiva, $p = 0.05$ y $q = 0.2$ respectivamente.

(véase, por ejemplo, [94]). Pero, tal y como hemos mencionado antes, este límite no es uniforme cuando el tiempo tiende a infinito. De hecho, podríamos decir que el límite de viscosidad evanescente y el de tiempo grande no conmutan, por lo que, en consecuencia, los siguientes límites producen resultados diferentes:

$$\lim_{t\to\infty}\lim_{\nu\to 0} u^\nu(x,t) \qquad \text{y} \qquad \lim_{\nu\to 0}\lim_{t\to\infty} u^\nu(x,t). \tag{1.6}$$

Mientras el primero conduce a las ondas-N biparamétricas, que posiblemente cambien de signo, el segundo lleva a una clase más restrictiva de perfiles asintóticos, correspondientes a las ondas-N de signo constante. Esta cuestión ha sido analizada de manera precisa, por ejemplo, en [59, 60]. En particular, en [60] los autores describen la transición de una onda-N – el perfil asintótico de la ecuación no viscosa– a la onda difusiva –el perfil asintótico de la ecuación viscosa–.

Los principales resultados de esta tesis comprenden esta dicotomía a nivel numérico. De hecho, lo mismo puede ocurrir al aproximar la ecuación hiperbólica (1.3) mediante esquemas numéricos. Esto no debería sorprender, dado que, como es bien sabido, los esquemas numéricos que convergen introducen cierta cantidad de viscosidad numérica (ver, por ejemplo, [38, 67]). Tal y como probaremos, esto puede afectar severamente la eficiencia de los métodos de optimización para problemas con horizontes de tiempo lejanos, incluso en ecuaciones de tipo Burgers más complejas.

1.2.2 Esquemas numéricos para leyes de conservación escalares

La ecuación de Burgers no viscosa es una ecuación unidimensional que pertenece a la clase de leyes de conservación. Estas ecuaciones toman la forma general

$$\begin{cases} \mathbf{u}_t + \operatorname{div}\big(f(\mathbf{u})\big) = 0, & (t,x) \in (0,\infty) \times \mathbb{R}^d, \\ \mathbf{u}(0,x) = \mathbf{u}_0(x), & x \in \mathbb{R}^d. \end{cases} \tag{1.7}$$

Aquí $\mathbf{u} : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^m$ es un vector de variables de estado m-dimensional –por ejemplo, masa, momento, energía...– y $F : \mathbb{R}^m \to \mathbb{R}^m$ es el flujo del sistema. Claramente, para la ecuación de Burgers tenemos $m = d = 1$ y $f(u) = u^2/2$. Recomendamos [65, 94] para una introducción y la teoría básica sobre leyes de conservación no lineales y [30] para el caso escalar.

Este tipo de ecuaciones son muy prácticas para modelar problemas de carácter científico o ingenieril que incluyan movimiento de ondas o transporte advectivo de sustancias y, consecuentemente, la conservación de ciertas magnitudes [66]. La dinámica de los gases, (ecuaciones de Euler), el plasma de un reactor de fusión (ecuaciones de la magnetohidrodinámica) o el flujo en un material poroso (ecuacion de Buckley-Leverett) son algunos ejemplos en los que las leyes de conservación aparecen de manera natural.

A pesar de la apariencia simple de (1.7), hay dificultades significativas asociadas a sus soluciones que necesitan ser tratadas cuidadosamente al desarrollar métodos numéricos. La formación de choques es especialmente delicada. De hecho, las clásicas aproximaciones de diferencias finitas funcionan bien cuando las soluciones son suaves, pero fallan al aproximar las que son discontinuas. Por tanto, se requieren esfuerzos adicionales para sobreponerse a este problema.

En esta tesis nos centraremos en los métodos de volúmenes finitos, que se basan en la forma integral de (1.7), y en las leyes de conservación escalares unidimensionales. En vez de calcular aproximaciones puntuales en los nodos de la malla, estos métodos numéricos dividen el dominio espacial en celdas y tratan de aproximar la media de la solución continua en cada una de ellas. El punto clave consiste en definir flujos numéricos razonables para determinar el flujo en los bordes de las celdas. Nos remitimos a [67] para una introducción sobre estas técnicas. También se han empleado otros métodos, tales como los elementos finitos y los espectrales, para la simulación de leyes de conservación, pero no lo analizaremos aquí.

En cualquier caso, muchos métodos de volúmenes finitos pueden ser tratados como los de diferencias finitas, en el sentido de que se puede adoptar la misma notación identificando las medias de las celdas con el valor en el punto medio de estas. A partir de ahora asumiremos este punto de vista (el mismo que en [38, Capítulo III]). Dados un tamaño de celda $\Delta x > 0$ y un paso de tiempo $\Delta t > 0$, consideraremos que u_j^n es una aproximación de $u(n\Delta t, j\Delta x)$, obtenida por medio de un método de volúmenes finitos

que aproxima la ecuación (1.7):

$$u_j^{n+1} = u_n^j - \frac{\Delta t}{\Delta x}\left(g_{j+1/2}^n - g_{j-1/2}^n\right), \qquad j \in \mathbb{Z}, n > 0. \tag{1.8}$$

Aquí $g_{j+1/2}^n = g(u_{j-k+1}, \ldots, u_{j+k})$ es el flujo numérico, una aproximación del flujo continuo $f(u)$ mediante una función continua $g : \mathbb{R}^{2k} \to \mathbb{R}$. Consúltese [38] para una reseña sobre los métodos de primer y segundo orden más clásicos y [67], para una introducción sobre métodos de alta resolución.

1.2.3 Optimización mediante búsqueda lineal

La optimización es crucial en nuestro mundo: las personas optimizan, la naturaleza optimiza. Hay numerosas situaciones en las que se puede identificar un objetivo que uno querría conseguir. Desde el punto de vista matemático, esto se corresponde con minimizar un funcional (obsérvese que maximizar un funcional equivale a minimizar su opuesto o inverso), normalmente condicionado por varias restricciones:

$$\min_{z\in\Omega} f(z) \quad \text{sujeto a} \quad \begin{cases} a(z) = 0, \\ b(z) \geq 0. \end{cases}$$

Generalmente, este tipo de problemas no puede ser resuelto *a mano* y, por tanto, se requiere de algoritmos efectivos para calcular aproximaciones de sus soluciones. Los más extendidos son los métodos iterativos, que buscan construir una sucesión $\{z_n\}_{n\geq 0}$ de elementos del conjunto Ω tales que $\lim_{n\to\infty} z_n = \bar{z}$, siendo $\bar{z}$ una solución exacta del problema de optimización.

En esta tesis consideramos, principalmente, los métodos de búsqueda lineal y, en particular, los métodos de gradiente descendente. La idea básica tras estas técnicas es la misma que la que uno usaría para bajar de una montaña: buscar una dirección descendente en los alrededores del punto en que nos encontramos y dar un paso en ese sentido (véase la Imagen 1.3).

En la estrategia de búsqueda lineal, para construir una nueva aproximación z_{n+1} desde la actual z_n, deberemos encontrar una dirección d_n en z_n en la que el funcional decrezca. Después, tendremos que decidir el tamaño del paso ε_n que daremos en esa dirección. Por tanto, a partir de una suposición inicial z_0, calcularemos de manera iterativa

$$z_{n+1} = z_n + \varepsilon_n d_n, \quad n \in \mathbb{N}.$$

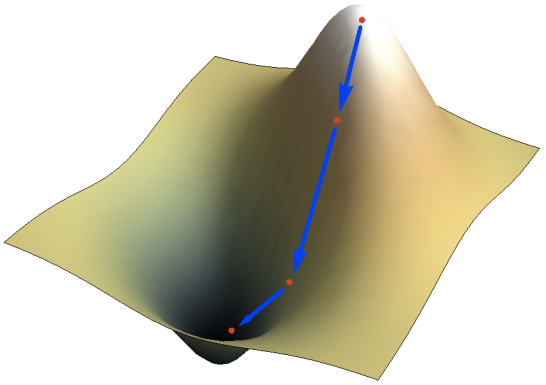


IMAGEN 1.3: Minimizar un funcional por medio de un método de gradiente descendente.

El método más común, el de descenso de pendiente máxima, considera el gradiente del funcional como la dirección de descenso $d_n = -\nabla f(x_n)$ y elige como tamaño del paso aquel que minimiza el funcional a lo largo de la línea que pasa por z_n en la dirección d_n. Dado que estamos tomando el gradiente en el sentido negativo, $\mathcal{J}(z_{n+1}) < \mathcal{J}(z_n)$ para un ε_n suficientemente pequeño, excepto en el caso en el que se alcance el mínimo en un número finito de iteraciones: $\nabla\mathcal{J}(z_n) = 0$. Obsérvese que, en general, a medida que nos acercamos al mínimo, el gradiente tiende a cero y, por tanto, podemos esperar que la sucesión converja a un mínimo (posiblemente local). Nos referiremos a [19] para la teoría básica sobre métodos de gradiente descendente.

Estos métodos (como el anteriormente mencionado descenso de pendiente máxima o el método del gradiente conjugado) son atractivos por su simplicidad algorítmica. Incluso se pueden modificar ligeramente para mejorar la tasa de convergencia (véase, por ejemplo, los trabajos de U. Ascher y sus colaboradores [4, 91]). Además, no necesitan que se calculen derivadas de segundo orden, lo que los hace apropiados para problemas de optimización de gran escala, en los que el coste de calcular la matriz hessiana y resolver los correspondientes sistemas lineales es prohibitivo. No obstante, basados en la idea de la búsqueda lineal, se han desarrollado muchos otros métodos –como el de Newton, el cuasi-Newton...– desde mediados del siglo XX. Consúltese [77] para una introducción en las estrategias de búsqueda lineal y su implementación, así como otras herramientas de optimización de la clase de métodos de región de confianza.

1.3 Contenidos de la tesis

En esta sección describimos brevemente los problemas estudiados en esta tesis. Se tratan cuatro temas:

1. **Viscosidad evanescente en esquemas numéricos:** En el Capítulo 2, analizamos el comportamiento asintótico en tiempos grandes de diversos esquemas de aproximación numérica para leyes de conservación escalares y, en particular, para la ecuación de Burgers no viscosa. Demostramos que, a nivel numérico, la dinámica en tiempos grandes depende de la cantidad de viscosidad numérica introducida por el esquema. Los resultados de este capítulo están basados en el artículo [84], en colaboración con L. I. Ignat y E. Zuazua.

2. **Control óptimo en horizontes lejanos:** En el Capítulo 3, tratamos con métodos numéricos para el diseño inverso de la ecuación de Burgers en horizontes lejanos. Hacemos hincapié en que la convergencia en el sentido clásico del análisis numérico no es suficiente. Los esquemas numéricos pueden alterar la dinámica del sistema continuo subyacente y, por tanto, afectar a la eficiencia de las herramientas de optimización numéricas. Este capítulo se basa en el artículo [81], en colaboración con N. Allahverdi y E. Zuazua.

3. **Esquemas numéricos que preservan el comportamiento en tiempos grandes de la ABE:** En el Capítulo 4, estudiamos el problema de Cauchy para la versión con parámetros constantes de la ecuación de Burgers aumentada y su dinámica en tiempos grandes. Después, proponemos un esquema numérico semidiscreto que preserva ese comportamiento asintótico mediante la introducción de dos factores correctores en la discretización del término no local. Los resultados de este capítulo se basan en el artículo [82], en colaboración con L. I. Ignat.

4. **Separación de operadores para la ABE:** En el Capítulo 5, consideramos un método de separación de operadores para la ecuación de Burgers aumentada. Probamos la convergencia de primer orden y obtenemos el primer término de la expansión asintótica. Este capítulo está basado en el artículo [83], en colaboración con L. I. Ignat.

Seguidamente, describimos los aspectos más importantes de cada uno de estos temas, los resultados obtenidos y los métodos que para tal efecto hemos desarrollado.

Observación 1.1*: De aquí en adelante para $p \in [1,\infty)$ y una sucesión $v = \{v_j\}_{j\in\mathbb{Z}}$ cualquiera, usaremos las siguientes normas discretas:*

$$\|v\|_{p,\Delta} = \Big(\Delta x \sum_{j\in\mathbb{Z}} |v_j|^p\Big)^{1/p}, \qquad \|v\|_{\infty,\Delta} = \max_{j\in\mathbb{Z}} |v_j|, \qquad TV(v) = \sum_{j\in\mathbb{Z}} |v_{j+1} - v_j|.$$

Además, denotaremos $\|\cdot\|_p$ la norma usual del espacio $L^p(\mathbb{R})$.

1.3.1 Capítulo 2: Viscosidad evanescente en esquemas numéricos

Tal y como mencionamos en la sección anterior, hay una diferencia significativa en lo que respecta al comportamiento en tiempos grandes de la ecuación de Burgers viscosa y la no viscosa. El principal resultado del Capítulo 2 afirma que lo mismo puede suceder a la hora de aproximar la ecuación hiperbólica (1.3) mediante esquemas numéricos [84]. Esto no es sorprendente, ya que, como es bien sabido, los esquemas numéricos convergentes introducen cierto grado de viscosidad numérica [38]. Nuestro análisis permite clasificar los esquemas numéricos en aquellos que, cuando el tiempo tiende a infinito, introducen una cantidad de viscosidad numérica insignificante –y que, por tanto, conducen al comportamiento asintótico correcto descrito por las ondas-N– y aquellos que introducen demasiada viscosidad numérica –dirigiéndose, por ello, a perfiles auto-semejantes viscosos–.

Como veremos, los esquemas de Engquist-Osher [26] y Godunov [39] pertenecen a la primera categoría, mientras que el clásico esquema de Lax-Friedrichs [63] encaja en el segundo. Resumiendo, podemos decir que las soluciones de los esquemas de Engquist-Osher y Godunov, para una malla fijada, capturan la dinámica hiperbólica del sistema continuo. Por el contrario, el esquema de Lax-Friedrichs scheme, debido a la cantidad excesiva de viscosidad numérica, lleva a un comportamiento asintótico incorrecto, de naturaleza viscosa y no hiperbólica. Nótese que, a pesar de que nuestro análisis se reduce a la ecuación de Burgers no viscosa, se pueden generalizar las mismas conclusiones a la aproximación numérica de leyes de conservación viscosas en las que la cantidad de viscosidad efectiva, cuando t tiende a infinito, dependa significativamente de la naturaleza del esquema numérico considerado.

El principal objetivo del Capítulo 2 es analizar, una vez fijados Δx y Δt, el comportamiento asintótico de estas soluciones discretas cuando $n \to \infty$. Por supuesto, estamos interesados en esquemas numéricos que convergen a la solución entrópica de (1.7) y en parámetros de la malla que satisfagan la correspondiente condición de CFL. Consideraremos esquemas numéricos conservativos de tres puntos para aproximar (1.7). En particular, tomamos $k = 1$ y $f(u) = u^2/2$ en (1.8).

Nuestro análisis se centra, sobre todo, en los esquemas numéricos de Lax-Friedrichs, Engquist-Osher y Godunov, que son conservativos. Además de converger a la solución entrópica de (1.7) bajo unas condiciones de CFL adecuadas, son conocidos por verificar la condición de Lipschitz unilateral (véase [12] y las referencias incluidas ahí), que es necesaria para establecer las propiedades de decaimiento cuando el tiempo discreto tiende a infinito.

Nuestros resultados, correspondientes a un marco L^1, exhiben una diferencia significativa respecto a trabajos anteriores relacionados con esquemas conservativos monótonos. En [41] el autor analiza el comportamiento en tiempos grandes de estos esquemas en el contexto de ondas de rarefacción, más cercano a un marco L^∞. Nuestro caso se puede entender formalmente como el límite en el que los valores en $\pm\infty$ se desvanecen y dan lugar al segundo término de la expansión asintótica de las soluciones. Probamos que en ese contexto, la viscosidad numérica adicional añadida por los esquemas ha de ser cuidadosamente tratada, a fin de poder detectar correctamente el comportamiento asintótico de las soluciones discretas cuando el tiempo tiende a infinito.

Para ser más precisos, consideremos que $\{u_j^0\}_{j\in\mathbb{Z}}$ es una aproximación del dato inicial; por ejemplo,

$$u_j^0 = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} u_0(x)dx, \quad x_{j+1/2} = (j+\frac{1}{2})\Delta x. \tag{1.9}$$

Definimos también una función contante a trozos u_Δ, que toma valores en $[0,\infty)\times\mathbb{R}$, como

$$u_\Delta(t,x) = u_j^n, \quad x_{j-1/2} < x < x_{j+1/2},\, t_n \le t < t_{n+1}, \tag{1.10}$$

donde $t_n = n\Delta t$ y u_j^n se calcula mediante (1.8).

El siguiente teorema, centrado en la ecuación de Burgers, es el principal resultado del Capítulo 2.

Teorema 1.1. *Sea $u_0 \in L^1(\mathbb{R})$ y Δx y Δt parámetros de la malla satisfaciendo la condición de CFL $\lambda\|u^n\|_{\infty,\Delta} \le 1$, $\lambda = \Delta t/\Delta x$. Sea u_Δ la correspondiente solución del esquema discreto* (1.8) *para la ecuación de Burgers no viscosa* (1.3)*. Entonces, para cualquier $p \in [1,\infty)$, se cumple lo siguiente:*

$$\lim_{t\to\infty} t^{\frac{1}{2}(1-\frac{1}{p})}\|u_\Delta(t) - w(t)\|_p = 0, \tag{1.11}$$

donde el perfil w es el siguiente:

1. *Para el esquema de Lax-Friedrichs, $w = w_{M_\Delta}$ es la única solución de la ecuación de Burgers viscosa dada por*

$$\begin{cases} w_t + \left(\frac{w^2}{2}\right)_x = \frac{(\Delta x)^2}{2\Delta t} w_{xx}, & x\in\mathbb{R}, t>0, \\ w(0) = M_\Delta \delta_0, \end{cases} \tag{1.12}$$

   donde $M_\Delta = \int_{\mathbb{R}} u_\Delta^0$.

2. *Para los esquemas de Engquist-Osher y Godunov,* $w = w_{p_\Delta, q_\Delta}$ *es la única solución de la ecuación de Burgers hiperbólica dada por*

$$\begin{cases} w_t + \left(\frac{w^2}{2}\right)_x = 0, \quad x \in \mathbb{R}, t > 0, \\ \\ w(0) = M_\Delta \delta_0, \quad \lim\limits_{t\to 0} \displaystyle\int_{-\infty}^{x} w(t,z)dz = \begin{cases} 0, & x < 0, \\ -p_\Delta, & x = 0, \\ q_\Delta - p_\Delta, & x > 0, \end{cases} \end{cases} \tag{1.13}$$

con $M_\Delta = \int_{\mathbb{R}} u_\Delta^0$ *y*

$$p_\Delta = -\min_{x\in\mathbb{R}} \int_{-\infty}^{x} u_\Delta^0(z)dz \quad \text{y} \quad q_\Delta = \max_{x\in\mathbb{R}} \int_{x}^{\infty} u_\Delta^0(z)dz.$$

El dato inicial en las ecuaciones (1.12) y (1.13) anteriores se ha de entender en el sentido de la convergencia de medidas acotadas. Véase [28] y [73] para una definición más precisa.

Los perfiles son bien conocidos [29, 30] y están dados explícitamente por:

$$w_{M_\Delta}(x,t) = -\frac{2\sqrt{\nu}}{t^{1/2}} \exp\left(-\frac{x^2}{4\nu t}\right) \left[C_{M_\Delta} + \int_{-\infty}^{x/\sqrt{\nu t}} \exp\left(-\frac{s^2}{4}\right) ds\right]^{-1}, \tag{1.14}$$

donde $\nu = \Delta x^2/(2\Delta t)$ y C_{M_Δ} es tal que la masa de la solución w_{M_Δ} es M_Δ, y

$$w_{p_\Delta, q_\Delta}(x,t) = \begin{cases} \frac{x}{t}, & -\sqrt{2p_\Delta t} < x < \sqrt{2q_\Delta t}, \\ 0, & \text{en otra parte.} \end{cases} \tag{1.15}$$

Nótese que los perfiles viscosos (1.14) están completamente determinados por la masa total, que se conserva en el tiempo tanto bajo la dinámica continua como bajo la discreta. Por el contrario, los perfiles onda-N (1.15) están unívocamente determinados por dos parámetros (p, q) de invariantes que se mantienen constantes en el tiempo, también en ambos casos. La cantidad $q_\Delta - p_\Delta$ es precisamente M_Δ, la masa de u_Δ^0.

La diferencia entre ambos perfiles puede observarse en la Imagen 1.2, donde hemos tomado $t = 50$, $\Delta x = 1/200$, $\Delta t = 1/200$, $M_\Delta = 0.15$, $p_\Delta = 0.05$ y $q_\Delta = 0.2$.

Para probar el Teorema 1.1 usamos argumentos de escala, similares a los aplicados en las demostraciones de los análogos continuos. Además, también introducimos las variables semejantes, que son una herramienta común en el análisis del comportamiento asintótico de ecuaciones en derivadas parciales. Esto nos permitirá observar de manera más clara algunos de los fenómenos mencionados.

1.3.2 Capítulo 3: Control óptimo en horizontes lejanos

En el Capítulo 3 analizamos la aproximación numérica del problema del diseño inverso para la ecuación de Burgers, tanto para el caso viscoso (1.4) como para el no viscoso (1.3). Al igual que sucede en el problema de la minimización de la explosión sónica, dados un tiempo $T > 0$ y una función objetivo u^*, el propósito es identificar el dato inicial u_0 de manera que su correspondiente solución alcance el objetivo u^* en $t = T$ o, al menos, se acerque lo máximo posible.

Esencialmente, la cuestión consiste en resolver hacia atrás la ecuación de Burgers, un problema que está mal puesto. En el caso viscoso $\nu > 0$, eso se debe a la fuerte irreversibilidad en tiempo intrínseca de la ecuación de Burgers parabólica, que se ve realzada por los fenómenos no lineales de la dinámica hiperbólica. En el caso hiperbólico no viscoso, la no linealidad del modelo, que produce la aparición de discontinuidades, hace que el problema también esté mal puesto, teniendo además múltiples soluciones en algunos casos.

Formularemos el problema desde el punto de vista del control óptimo. Usando un enfoque de mínimos cuadrados, consideramos la minimización del siguiente funcional:

$$\mathcal{J}(u_0) = \frac{1}{2} \int_{\mathbb{R}} (u(x,T) - u^*(x))^2 \, dx, \tag{1.16}$$

donde u es la solución de la ecuacion de Burgers y el dato inicial u_0 forma parte de una clase de funciones adecuada, como por ejemplo $L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$.

Por tanto, este problema de control óptimo surge de manera natural como una versión simplificada del modelo completo de la explosión sónica. Uno de los ingredientes principales es que el horizonte temporal $[0, T]$ considerado tiene que ser grande, por motivos prácticos. Como veremos, esto hace que la elección del esquema numérico que aproxime la EDP sea una cuestión delicada, dado que los esquemas que no reproducen correctamente la dinámica de tiempos grandes son incapaces de producir una aproximación precisa del control óptimo. Tal y como mencionábamos anteriormente, la dicotomía parabólica/hiperbólica asintótica ha de ser tratada cuidadosamente. En particular, el exceso de viscosidad numérica podría destruir la dinámica hiperbólica y hacerla parabólica. El análisis del capítulo 2 se lleva a cabo en un contexto puramente hiperbólico; pero dicha patología también puede aparecer en el caso de la ecuación de Burgers no viscosa cuando la viscosidad numérica domina a la física. El principal objetivo de este capítulo es describir los diversos fenómenos que se solapan a la hora de resolver numéricamente este problema de diseño inverso.

En este trabajo enfatizamos las consecuencias de este hecho a nivel del diseño inverso. Para ello empleamos un método de gradiente descendente junto con la metodología del adjunto. También utilizamos IPOPT, un paquete informático de libre distribución para optimización no lineal [92], como soporte para nuestros resultados. No obstante, nótese que la dicotomía del comportamiento en tiempos grandes puede extenderse a otros métodos numéricos. Véase [19, 77] y sus referencias para la teoría básica en esta extensa área de la optimización numérica.

Este problema de control óptimo, sin el requerimiento del horizonte temporal lejano, ya ha sido tratado en el pasado por numerosos autores, tanto desde el punto de vista continuo como del discreto. Tal y como apuntábamos anteriormente, la interacción entre discretización y optimización es un tema exigente y estimulante. En el contexto más general de las leyes de conservación escalares, cabe destacar los trabajos [7, 10, 55, 90]. De hecho, hay una extensa investigación por parte de F. James y sus colaboradores en lo que respecta al sistema adjunto y su discretización (véase, por ejemplo, [9, 40, 56]). Ciertamente, la presencia de choques en este tipo de problemas no es un asunto trivial. En [35, 36], los autores analizan la convergencia puntual de las aproximaciones del sistema lineal y el adjunto para soluciones discontinuas desde el enfoque discretizar-después-optimizar. En particular, hacen hincapié en la importancia de controlar la difusión para poder obtener la convergencia. Sin embargo, nuestros resultados muestran que, en la práctica, su propuesta podría ser insuficiente al considerar un valor grande para T.

Además, la interacción entre las estrategias optimizar-después-discretizar y discretizar-después-optimizar fue detalladamente analizada en [89]. En el Capítulo 3 optamos por la segunda. Tratamos de evitar discontinuidades en las soluciones para minimizar su impacto en el proceso de optimización y centrarnos únicamente en el efecto en tiempos grandes, por lo que la sensibilidad de los choques no juega un papel importante en nuestro caso. Por supuesto, esto no significa que los efectos en tiempos grandes no deban ser considerados en la presencia de choques.

Nuestros resultados constituyen una advertencia importante sobre la necesidad de emplear esquemas de aproximación numérica, capaces de imitar las propiedades dinámicas de tiempos grandes del sistema, a la hora de resolver problemas de diseño inverso y control óptimo en horizontes temporales lejanos. Esto ya fue observado previamente en problemas de control para la propagación de ondas [27, 95]. Es interesante ver que las mismas patologías persisten en el problema del diseño inverso para flujos viscosos y no viscosos, aparentemente más robusto.

1.3.3 Capítulo 4: Esquemas numéricos que preservan el comportamiento en tiempos grandes de la ABE

El Capítulo 4 está dedicado a la ecuación de Burgers aumentada con parámetros constantes y un único proceso de relajación molecular. Nos centramos en la siguiente ecuación:

$$\begin{cases} u_t = uu_x + \nu\, u_{xx} + c\, K_\theta * u_{xx}, & (t,x) \in (0,\infty)\times\mathbb{R}, \\ u(0,x) = u^0(x), & x \in \mathbb{R}, \end{cases} \tag{1.17}$$

donde $*$ denota la convolución en la variable x, los parámetros ν, c, θ son positivos y

$$K_\theta(z) = \begin{cases} \frac{1}{\theta} e^{-z/\theta}, & z > 0, \\ 0, & \text{en otra parte.} \end{cases} \tag{1.18}$$

Las mismas patologías que analizamos en los Capítulos 2 y 3 aparecen también en el contexto de la ecuación de Burgers aumentada. Valores pequeños de ν y c en (1.17) necesitan un tratamiento similar desde el punto de vista numérico, como si la ecuación fuera hiperbólica. Por tanto, en el Capítulo 4 estudiamos el comportamiento asintótico de las soluciones de (1.17) cuanto $t \to \infty$ y desarrollamos un esquema numérico semi-discreto que preserva tal comportamiento.

En lo que concierne al comportamiento de las soluciones del sistema (1.17) en tiempos grandes, el teorema siguiente contiene el principal resultado:

Teorema 1.2. *Sea $u^0 \in L^1(\mathbb{R})$. Para cualquier $p \in [1,\infty]$, la solución u de (1.17) satisface*

$$t^{\frac{1}{2}(1-\frac{1}{p})} \|u(t) - u_M(t)\|_p \longrightarrow 0, \quad \text{cuando } t \to \infty,$$

donde $u_M(t,x)$ es la solución de la siguiente ecuación:

$$\begin{cases} u_t = uu_x + (\nu + c)u_{xx}, & x \in \mathbb{R}, t > 0, \\ u(0) = M\delta_0. \end{cases}$$

Aquí δ_0 indica la delta de Dirac en el origen y M es la masa del dato inicial, $M = \int_{\mathbb{R}} u^0(x)dx$.

Nótese que u_M está explícitamente definido en (1.14), sin más que tomar $\nu + c$ en vez de ν como parámetro de viscosidad. De hecho, esto demuestra que u_M es de la forma $t^{-\frac{1}{2}} f_M\left(\frac{x}{\sqrt{t}}\right)$ para alguna función f_M y, por tanto, auto-semejante.

Nuevamente, esto será particularmente importante a nivel numérico. Por un lado, al elegir el flujo numérico que discretice la no linealidad, necesitamos tener cuidado con

la viscosidad numérica que introducimos. Por otro lado, hemos de tratar con cuidado el truncamiento del término integral, de manera que no introduzcamos patologías no deseadas en el comportamiento asintótico de la solución numérica. En ese sentido, nosotros proponemos dos factores correctores.

Sea u_Δ una aproximación de la solución u de (1.17). Definimos esta función constante a trozos (a diferencia del Capítulo 2, ahora solamente en el espacio) como sigue:

$$u_\Delta(t,x) = u_j(t), \quad x \in (x_{j-1/2}, x_{j+1/2}), t \geq 0, \tag{1.19}$$

donde $x_{j+1/2} = (j + \frac{1}{2})\Delta x$, para todo $j \in \mathbb{Z}$, y $\Delta x > 0$ es el tamaño de malla dado. También indicaremos con $x_j = j\Delta x$ el punto intermedio de las celdas espaciales. Para cada $j \in \mathbb{Z}$, necesitamos calcular una función $u_j(t)$ que aproxime el valor de la solución en la celda. Teniendo en cuenta la problemática descrita anteriormente, optamos por la siguiente discretización de (1.17): el esquema Engquist-Osher para el flujo, diferencias finitas centradas para el laplaciano y y la regla del rectángulo compuesta para la integral:

$$\begin{cases} u_j'(t) = \frac{g_{j+1/2}(t) - g_{j-1/2}(t)}{\Delta x} + \nu \frac{u_{j-1}(t) - 2u_j(t) + u_{j+1}(t)}{\Delta x^2} \\ \qquad\qquad + \frac{c}{\theta^2} \left(\sum_{m=1}^{N} \omega_m^\theta u_{j-m}(t) - F_0^{\Delta,\theta} u_j(t) + F_1^{\Delta,\theta} \theta \frac{u_{j+1}(t) - u_j(t)}{\Delta x} \right), \quad j \in \mathbb{Z}, t \geq 0, \\ u_j(0) = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} u^0(x) dx, \qquad\qquad j \in \mathbb{Z}, \end{cases} \tag{1.20}$$

donde

$$\omega_m^\theta = e^{-m\Delta x/\theta} \left(e^{\Delta x/\theta} - 1 \right), \quad m = 1, \ldots, N, \tag{1.21}$$

y

$$g_{j+1/2}(t) = \frac{u_j(t)\big(u_j(t) - |u_j(t)|\big)}{4} + \frac{u_{j+1}(t)\big(u_{j+1}(t) + |u_{j+1}(t)|\big)}{4}, \quad j \in \mathbb{Z}, t \geq 0.$$

El parámetro $N = N(\Delta x) \in \mathbb{N}$ señala el número de nodos considerados en la fórmula de cuadratura de la integral. Los factores correctores $F_0^{\Delta,\theta}$ y $F_1^{\Delta,\theta}$ que aparecen junto con las aproximaciones de los términos u y u_x, dados por

$$F_0^{\Delta,\theta} = \sum_{m=1}^{N} \omega_m^\theta \quad \text{y} \quad F_1^{\Delta,\theta} = \frac{\Delta x}{\theta} \sum_{m=1}^{N} m\omega_m^\theta, \tag{1.22}$$

sirven para que el truncamiento del término no local sea correcto desde el punto de vista del comportamiento en tiempos grandes.

Finalmente, para Δx fijo, estudiamos el comportamiento asintótico de estas soluciones u_Δ semi-discretas cuando $t \to \infty$.

Teorema 1.3. *Sea $u_0 \in L^1(\mathbb{R})$, $\Delta x > 0$ and u_Δ la correspondiente solución del esquema semi-discreto* (1.20) *para la ecuación de Burgers aumentada* (1.17)*. Para cualquier $p \in [1, \infty]$, se verifica lo siguiente:*

$$t^{\frac{1}{2}(1-\frac{1}{p})} \|u_\Delta(t) - u_M(t)\|_p \longrightarrow 0, \quad \text{as } t \to \infty, \tag{1.23}$$

donde $u_M(t,x)$ es la única solución de la siguiente ecuación de Burgers viscosa:

$$\begin{cases} v_t = vv_x + (\nu + c\, F_2^{\Delta,\theta}) v_{xx}, & x \in \mathbb{R}, t > 0, \\ v(x,0) = M\delta_0. \end{cases}$$

Aquí $M = \int_{\mathbb{R}} u^0(x)dx$ es la masa del dato inicial y

$$F_2^{\Delta,\theta} = \frac{\Delta x^2}{2\theta^2} \left(\sum_{m=1}^{N} m(m-1)\omega_m^\theta \right).$$

Obsérvese que, para un $\theta > 0$, $F_2^{\Delta,\theta} \to 1$ si se toma N de manera que $N \to \infty$ y $N\Delta x \to \infty$ cuando $\Delta x \to 0$. Ese es, de hecho, el valor que uno debiera esperar por el modelo continuo.

1.3.4 Capítulo 5: Separación de operadores para la ABE

Resolver (1.20) puede ser costoso computacionalmente si N es grande. Por tanto, en el Capítulo 5 establecemos el marco para fortalecer el uso de métodos de separación de operadores para resolver la ecuación de Burgers aumentada. De hecho, cabe mencionar que estas técnicas ya han sido empleadas en [20, 85, 86] en el contexto del fenómeno de la explosión sónica. Además, se han desarrollado otras versiones de estos métodos para, entre otros modelos, la ecuación de Schrödinger no lineal, la de Korteweg-de Vries, la de Boltzmann... (veáse [44] y sus referencias) y, más recientemente, para la ecuación de Fowler no local [11].

La idea básica tras estos métodos es la de escribir el operador de evolución global como suma de los operadores de evolución de cada uno de los términos que aparecen en el modelo. En otras palabras, uno puede separar el sistema completo en sub-ecuaciones más simples, para las que existen algoritmos más prácticos. Una vez resueltas estas sub-ecuaciones, se vuelven a juntar sus soluciones para calcular la solución completa del modelo en cuestión. Véase [44] para una introducción más detallada.

Nosotros presentamos la siguiente fórmula de Trotter para la ecuación de Burgers aumentada. A fin de mantener una notación clara, analizaremos únicamente el caso $\nu = c = \theta = 1$, pero los resultados obtenidos se pueden extender fácilmente al caso general. Sea X^t el operador de evolución asociado a

$$\begin{cases} v_t = K * v - v + v_x, & (t,x) \in (0,\infty) \times \mathbb{R}, \\ v(t=0,x) = v_0(x), & x \in \mathbb{R}, \end{cases}$$

e Y^t, el correspondiente a

$$\begin{cases} w_t - \left(\frac{w^2}{2}\right)_x = w_{xx}, & (t,x) \in (0,\infty) \times \mathbb{R}, \\ w(t=0,x) = w_0(x), & x \in \mathbb{R}. \end{cases}$$

Consideramos el flujo Z^t definido por

$$Z^t = X^t Y^t. \tag{1.24}$$

El objetivo es aproximar la solución u de (1.17) mediante

$$(Z^{\Delta t})^n u_0 = (X^{\Delta t} Y^{\Delta t})^n u_0,$$

donde $\Delta t > 0$ and $n \in \mathbb{N} \cup \{0\}$. Para simplificar, de aquí en adelante denotaremos $Z^{n\Delta t} = (Z^{\Delta t})^n$. Recordemos también que $t_n = n\Delta t$ and $t_{n+1/2} = (n + \frac{1}{2})\Delta t$ para cada $n \in \mathbb{N} \cup \{0\}$.

El primer resultado del Capitulo 5, que confirma que la separación de operadores (1.24) es de primer orden para soluciones suficientemente regulares, es el siguiente.

Teorema 1.4. *Sea $r \in \{1,2\}$. Para todo $u_0 \in H^r(\mathbb{R})$ y todo $T > 0$, existen constantes positivas c_1, c_2 and Δt_0 tales que, para todo $\Delta t \in (0, \Delta t_0)$ y para todo $n \in \mathbb{N}$ tal que $0 \le n\Delta t \le T$,*

$$\left\| Z^{n\Delta t} u_0 - u(n\Delta t) \right\|_2 \le c_1 (\Delta t)^{r/2}$$

y

$$\left\| Z^{n\Delta t} u_0 \right\|_{H^r(\mathbb{R})} \le c_2.$$

Aquí c_1, c_2 y Δt_0 solamente dependen de T y de $\|u_0\|_{H^r(\mathbb{R})}$.

Además, siguiendo técnicas parecidas a las empleadas en el Capítulo 4, obtenemos el primer término de la expansión asintótica de la solución dada por el operador Z^t. Definimos la función u^Δ como

$$u^{\Delta}(t,x)=\begin{cases}Y^{2(t-t_n)}Z^{n\Delta t}u_0(x), & t\in(t_n,t_{n+1/2}),\ x\in\mathbb{R},\\ X^{2(t-t_{n+1/2})}Y^{\Delta t}Z^{n\Delta t}u_0(x), & t\in(t_{n+1/2},t_{n+1}),\ x\in\mathbb{R}.\end{cases}$$

Se deduce que u^{Δ} verifica:

$$\begin{cases}u_t^{\Delta}=2(Y^{2(t-t_n)}Z^{n\Delta t}u_0)_t=2u_{xx}^{\Delta}+\left((u^{\Delta})^2\right)_x, & t\in(t_n,t_{n+1/2}),\\ u_t^{\Delta}=2(K*u^{\Delta}-u^{\Delta}+u_x^{\Delta}), & t\in(t_{n+1/2},t_{n+1}),\\ u^{\Delta}(t_n)=Z^{n\Delta t}u_0,\\ u^{\Delta}(t_{n+1/2})=Y^{\Delta t}Z^{n\Delta t}u_0.\end{cases}$$

Sea $I_n^{\Delta}=(t_n,t_{n+1/2})$. Si definimos la función $\psi^{\Delta t}(t)=\sum_{n\geq 0}\chi_{I_n^{\Delta}}(t)$ (cuyo valor es 1 si t está en $(t_n,t_{n+1/2})$ y 0 en caso contrario), queda claro que el sistema anterior puede ser escrito de la siguiente forma:

$$\begin{cases}u_t^{\Delta}=2\psi^{\Delta t}(t)u_{xx}^{\Delta}+\psi^{\Delta t}(t)((u^{\Delta})^2)_x\\ \qquad\qquad +2\left(1-\psi^{\Delta t}(t)\right)(K*u^{\Delta}-u^{\Delta}+u_x^{\Delta}), & t>0,\\ u^{\Delta}(0)=u_0.\end{cases}\tag{1.25}$$

Se puede demostrar que $\psi^{\Delta t}(t)\rightharpoonup\frac{1}{2}$ (véase [23]). Más aún, por construcción y las propiedades de X^t e Y^t, se tiene que la función u^{Δ} definida por (1.25) satisface $u^{\Delta}\in C([0,\infty);L^1(\mathbb{R}))$. Por lo tanto, tenemos que formalmente $u^{\Delta}\to u$ cuando $\Delta t\to 0$, donde u satisface (1.17). El comportamiento de u^{Δ} cuando $t\to\infty$ se deduce mediante un argumento de escala y se enuncia a continuación.

Teorema 1.5. *Para cualquier $u_0\in L^1(\mathbb{R})\cap L^{\infty}(\mathbb{R})$ y $p\in[1,\infty)$*

$$\left\|u^{\Delta}(t)\right\|_p\leq C_p t^{-\frac{1}{2}(1-\frac{1}{p})}\|u_0\|_1,\quad n\geq 1.$$

Además, para todo $p\in[1,\infty)$,

$$t^{\frac{1}{2}(1-\frac{1}{p})}\left\|u^{\Delta}(t)-u_M(t)\right\|_p\to\infty,\quad \textit{cuando } t\to\infty,$$

donde $u_M(t,x)=t^{-1/2}f_M(x/\sqrt{t})$ es el perfil auto-semejante de la siguiente ecuación de Burgers viscosa:

$$\begin{cases}u_t=\left(\frac{u^2}{2}\right)_x+2u_{xx}, & (t,x)\in(0,\infty)\times\mathbb{R},\\ u(0,x)=M\delta_0, & x\in\mathbb{R}.\end{cases}$$

Obsérvese, que este es, precisamente, el comportamiento asintótico en tiempos grandes de la solución de (1.17), tal y como probamos en el Capítulo 4.

Chapter 2

Numerical vanishing viscosity for 1-D scalar conservation laws

2.1 Motivation

This chapter is devoted to the large-time behavior of the numerical approximations of the solution for 1-D scalar hyperbolic conservation laws of the form

$$u_t + \big[f(u)\big]_x = 0, \qquad x \in \mathbb{R}, t > 0. \tag{2.1}$$

We shall mainly focus on the Burgers equation with a quadratic flux $f(u) = u^2/2$:

$$u_t + \left(\frac{u^2}{2}\right)_x = 0, \qquad x \in \mathbb{R}, t > 0, \tag{2.2}$$

but the results can be extended to other fluxes.

As we mentioned in Chapter 1, the solutions of the hyperbolic Burgers equation develop a N-wave behavior as $t \to \infty$. They conserve the mass of the initial datum, this does not suffice to identify the asymptotic profile. Indeed, this profile, of self-similar nature, belongs to a two-parameter family of solutions, these parameters corresponding to two invariants of the system: the positive and the negative masses.

The large-time behavior differs significantly for the viscous version of these models:

$$u^\nu_t + \big[f(u^\nu)\big]_x = \nu u^\nu_{xx}, \qquad x \in \mathbb{R}, t > 0, \tag{2.3}$$

and

$$u^\nu_t + \left[\frac{(u^\nu)^2}{2}\right]_x = \nu u^\nu_{xx}, \qquad x \in \mathbb{R}, t > 0. \tag{2.4}$$

In fact, for $\nu > 0$ these problems are of parabolic nature. Thus, as t tends to infinity, the solutions converge towards a viscous profile of constant sign which is fully determined by the conserved mass.

The main result of this chapter deals with a similar dichotomy that arises when approximating the hyperbolic equations (2.1) and (2.2) by numerical schemes. It is well known that convergent numerical schemes for this type of models introduce some degree of numerical viscosity. Therefore, it is not surprising that the precision of the approximation of solutions is affected in large-time horizons.

Our analysis allows classifying numerical schemes in those that, as time tends to infinity, introduce a negligible amount of numerical viscosity, and therefore lead to the correct asymptotic behavior described by the N-waves, and those that introduce too much numerical viscosity thus leading to viscous self-similar profiles. In particular, we shall focus on the Engquist-Osher and Godunov schemes, which belong to the first category, and on the classical Lax-Friedrichs scheme, which belongs to the second one.

Of course, for finite time, the solutions of the viscous models (2.3) and (2.4) are well known to converge to the entropy solutions of the hyperbolic scalar conservation laws (2.1) and (2.2), respectively, as $\nu \to 0$. But this limit can not be made uniform as time tends to infinity (see [60]). At the discrete level, an equivalent question could be laid out: do the following limits commute?

$$\lim_{t\to\infty} \lim_{\substack{\Delta x\to 0\\ \Delta t\to 0}} u_\Delta(x,t) \qquad \text{and} \qquad \lim_{\substack{\Delta x\to 0\\ \Delta t\to 0}} \lim_{t\to\infty} u_\Delta(x,t),$$

where u_Δ is defined in (1.10).

The main goal of the present chapter is to analyze the asymptotic behavior as $t \to \infty$ of these discrete solutions for Δx and Δt fixed. We state this in the following theorem:

Theorem 2.1. *Let $u_0 \in L^1(\mathbb{R})$ and choose mesh-size parameters Δx and Δt satisfying the CFL condition $\lambda\|u^n\|_{\infty,\Delta} \le 1$, $\lambda = \Delta t/\Delta x$. Let u_Δ be the corresponding solution of the discrete scheme* (1.8) *for the hyperbolic Burgers conservation law* (1.3)*. Then, for any $p \in [1,\infty)$, the following holds*

$$\lim_{t\to\infty} t^{\frac{1}{2}(1-\frac{1}{p})} \|u_\Delta(t) - w(t)\|_p = 0, \tag{2.5}$$

where the profile w is as follows:

1. *for the Lax-Friedrichs scheme, $w = w_{M_\Delta}$ is the unique solution of the continuous viscous Burgers equation*

$$\begin{cases} w_t + \left(\frac{w^2}{2}\right)_x = \frac{(\Delta x)^2}{2\Delta t} w_{xx}, \quad x \in \mathbb{R}, t > 0, \\ w(0) = M_\Delta \delta_0, \end{cases} \tag{2.6}$$

with $M_\Delta = \int_{\mathbb{R}} u_\Delta^0$.

2. *for Engquist-Osher and Godunov schemes,* $w = w_{p_\Delta, q_\Delta}$ *is the unique solution of the hyperbolic Burgers equation*

$$\begin{cases} w_t + \left(\frac{w^2}{2}\right)_x = 0, \quad x \in \mathbb{R}, t > 0, \\ w(0) = M_\Delta \delta_0, \quad \lim\limits_{t\to 0} \displaystyle\int_{-\infty}^{x} w(t,z)dz = \begin{cases} 0, & x < 0, \\ -p_\Delta, & x = 0, \\ q_\Delta - p_\Delta, & x > 0, \end{cases} \end{cases} \tag{2.7}$$

with $M_\Delta = \int_{\mathbb{R}} u_\Delta^0$ *and*

$$p_\Delta = -\min_{x\in\mathbb{R}} \int_{-\infty}^{x} u_\Delta^0(z)dz \quad \text{and} \quad q_\Delta = \max_{x\in\mathbb{R}} \int_{x}^{\infty} u_\Delta^0(z)dz.$$

We mainly focus on the numerical schemes of Lax-Friedrichs, Engquist-Osher and Godunov. They are of a conservative nature, and well-known to converge to the entropy solution of (2.1) under suitable CFL conditions. Moreover, they satisfy the so-called one-sided Lipschitz condition (OSLC) that is required to establish, in particular, decay properties as the discrete time tends to infinity.

The rest of this chapter is divided as follows: in Section 2.2 we present some classical facts about the numerical approximation of one-dimensional conservation laws and obtain preliminary results that will be used in the proof of the main results of this chapter. In Section 2.3 we prove the main result, Theorem 2.1, and we illustrate it in Section 2.4 with numerical simulations. In Section 2.5, we analyze the same issues in the similarity variables and compare the results to the approximations obtained directly from the physical ones. Finally, in Section 2.6 we discuss possible generalizations to other numerical schemes and to more general fluxes (uniformly convex or odd ones).

2.2 Preliminary results

Following [12] and [38], we recall some well-known results about numerical schemes for 1-D scalar conservation laws. We prove some new technical results that will be used in Section 2.3 in the proof of Theorem 2.1. We restrict our attention to the Burgers equation (2.2). More general results will be discussed in Section 2.5 for uniformly convex and odd fluxes.

First, given a time-step Δt and a uniform spatial grid Δ with space increment Δx, we approximate the conservation law

$$\begin{cases} u_t + \left(\frac{u^2}{2}\right)_x = 0, & x \in \mathbb{R}, t > 0, \\ u(x,0) = u_0(x), & x \in \mathbb{R}, \end{cases} \tag{2.8}$$

by an explicit difference scheme of the form

$$u_j^{n+1} = H(u_{j-k}^n, \dots, u_{j+k}^n), \quad \forall n \geq 0, j \in \mathbb{Z}, \tag{2.9}$$

where $H : \mathbb{R}^{2k+1} \to \mathbb{R}$, $k \geq 1$, is a continuous function and u_j^n denotes the approximation of the exact solution u at the node $(n\Delta t, j\Delta x)$. We assume that there exists a continuous function $g : \mathbb{R}^{2k} \to \mathbb{R}$, called numerical flux, such that

$$H(u_{-k}, \dots, u_k) = u_0 - \lambda \left[g(u_{-k+1}, \dots, u_k) - g(u_{-k}, \dots, u_{k-1}) \right], \quad \lambda = \Delta t / \Delta x,$$

so that scheme (2.9) can be put in conservation form. This means that setting $g_{j+1/2}^n = g(u_{j-k+1}, \dots, u_{j+k})$, we can rewrite scheme (2.9) as:

$$u_j^{n+1} = u_j^n - \lambda \left(g_{j+1/2}^n - g_{j-1/2}^n \right), \quad \forall n \geq 0, j \in \mathbb{Z}, \tag{2.10}$$

where $\{u_j^0\}_{j\in\mathbb{Z}}$ is an approximation of $u_0 \in L^1(\mathbb{R})$, defined, for instance, as in (1.9). It is obvious that if a scheme can be put into conservation form, then the mass of the solution is conserved in time.

We will focus our analysis on monotone schemes. We recall that a numerical scheme (2.9) is said to be monotone if function H is monotone increasing in each of its arguments.

Let us remark that any 3-point ($k = 1$) monotone scheme in conservation form satisfies that any numerical flux $g(u, v)$ is an increasing function in the first argument and decreasing in the second one. The consistency of the scheme also guarantees that

$$g(u, u) = \frac{u^2}{2}, \quad \forall u \in \mathbb{R}. \tag{2.11}$$

Now, we recall a classical result about conservative schemes. For the sake of simplicity, we denote $H_\Delta(v) = \{H(v_{j-k}, \dots, v_{j+k})\}_{j\in\mathbb{Z}}$.

Proposition 2.2 (cf. [38, Chapter 3])**.** *Let $v = \{v_j\}_{j\in\mathbb{Z}}$ and $w = \{w_j\}_{j\in\mathbb{Z}}$ be two sequences in $l^1(\mathbb{Z}) \cap l^\infty(\mathbb{Z})$. Any monotone numerical scheme* (2.9) *which can be written in conservation form satisfies the following properties:*

1. *It is a contraction for $\|\cdot\|_{1,\Delta}$, that is:*

$$\|H_\Delta(v) - H_\Delta(w)\|_{1,\Delta} \leq \|v - w\|_{1,\Delta}.$$

2. It is L^∞-stable, that is:

$$\|H_\Delta(v)\|_{\infty,\Delta} \leq \|v\|_{\infty,\Delta}.$$

3. It preserves the sign, that is, if $v \geq 0$ then $H_\Delta(v) \geq 0$.

Another important property that we need in order to prove the asymptotic behavior of the numerical scheme is the OSLC. Let us introduce

$$D^n = \sup_{j\in\mathbb{Z}} \left(\frac{u_{j+1}^n - u_j^n}{\Delta x} \right)^+$$

where $z^+ := \max\{0, z\}$.

Definition 2.3 (cf. [12])**.** *A numerical scheme* (2.9) *is said to be OSLC consistent if:*

$$D^n \leq \frac{D^0}{1 + n\Delta t D^0}, \quad n \geq 1. \tag{2.12}$$

In particular, if a numerical scheme is OSLC consistent, it satisfies that

$$\frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x} \leq \frac{2}{n\Delta t}, \quad n \geq 1. \tag{2.13}$$

We emphasize that, to the best of our knowledge, there is no general result stating whether a numerical scheme satisfies the OSLC or not. Nonetheless, there are some well-known schemes that have already been proved to be OSLC consistent (see [12]) on which we concentrate. In the sequel, we say that a scheme satisfies the OSLC when (2.13) holds.

The analysis in this chapter is limited to the following three 3-point schemes, with their numerical fluxes respectively:

1. Lax-Friedrichs
$$g^{LF}(u,v) = \frac{u^2+v^2}{4} - \frac{\Delta x}{\Delta t}\left(\frac{v-u}{2}\right), \tag{2.14}$$

2. Engquist-Osher
$$g^{EO}(u,v) = \frac{u(u+|u|)}{4} + \frac{v(v-|v|)}{4}, \tag{2.15}$$

3. Godunov
$$g^G(u,v) = \begin{cases} \min\limits_{w\in[u,v]} \frac{w^2}{2}, & \text{if } u \leq v, \\ \max\limits_{w\in[v,u]} \frac{w^2}{2}, & \text{if } v \leq u. \end{cases} \tag{2.16}$$

Proposition 2.4 (cf. [12])**.** *Assuming that the CFL condition $\lambda\|u^n\|_{\infty,\Delta} \leq 1$ is fulfilled, the Lax-Friedrichs, Engquist-Osher and Godunov schemes are monotone and OSLC consistent.*

In the case of the three numerical schemes above, thanks to the OSLC, we obtain that the l^p_Δ-norms of the solutions decay similarly as in the continuous case.

Proposition 2.5. *Let us consider a monotone conservative numerical scheme that is OSCL consistent. For any $p \in [1,\infty]$, there exists a constant $C_p > 0$ such that the following holds*

$$\|u^n\|_{p,\Delta} \le C_p (n\Delta t)^{-\frac{1}{2}(1-\frac{1}{p})} \|u^0\|_{1,\Delta}^{\frac{1}{2}(1+\frac{1}{p})}, \quad \forall n \ge 1. \tag{2.17}$$

Proof. Estimate (2.17) for $p = 1$ follows from the fact that the scheme is conservative and, for $1 < p < \infty$, it follows by applying Hölder's inequality once (2.17) is proved for $p = \infty$. Moreover, by the comparison principle, it is sufficient to consider the case of nonnegative initial data u^0.

Let us now prove (2.17) for $p = \infty$ and nonnegative initial data. By the maximum principle, u_j^n is nonnegative for all $j \in \mathbb{Z}$ and $n \ge 0$. We use now the OSLC (2.13). For all integers $m \ge 1$ we have

$$\frac{u^n_{j+2m} - u^n_j}{2\Delta x} \le \frac{2m}{n\Delta t}, \quad n \ge 1. \tag{2.18}$$

For a fixed n, let us now assume that the point j where u_j^n attains its maximum is even, the treatment of the other case being analogous,

$$u^n_{2j_0} := \max_{j\in\mathbb{Z}} u^n_j.$$

Hence, in view of (2.18) we get

$$u^n_{2j} \ge u^n_{2j_0} - 4(j_0 - j)\frac{\Delta x}{n\Delta t}, \quad \forall\, j \le j_0.$$

Let us set

$$\gamma = j_0 - \frac{n\Delta t}{4\Delta x} u^n_{2j_0}.$$

The last inequality and the mass conservation property imply that

$$\begin{aligned}\|u^0\|_{1,\Delta} = \Delta x \sum_{j\in\mathbb{Z}} u^0_j = \Delta x \sum_{j\in\mathbb{Z}} u^n_j \ge \Delta x \sum_{j\in\mathbb{Z}} u^n_{2j} \ge \Delta x \sum_{j=\lfloor\gamma\rfloor+1}^{j_0} u^n_{2j} \\ \ge \Delta x \sum_{j=\lfloor\gamma\rfloor+1}^{j_0} \Big(u^n_{2j_0} - 4(j_0-j)\frac{\Delta x}{n\Delta t}\Big) = \frac{4(\Delta x)^2}{n\Delta t} \sum_{j=\lfloor\gamma\rfloor+1}^{j_0} (j-\gamma),\end{aligned}$$

where $\lfloor\gamma\rfloor$ denotes the largest integer less than or equal to γ. Since $\gamma \le j_0$, it follows that

$$\begin{aligned}\|u^0\|_{1,\Delta} &\geq \frac{4(\Delta x)^2}{n\Delta t}\sum_{j=\lfloor\gamma\rfloor+1}^{j_0}(j-\gamma) = \frac{2(\Delta x)^2}{n\Delta t}(j_0-\lfloor\gamma\rfloor)(j_0+\lfloor\gamma\rfloor+1-2\gamma)\\ &= \frac{2(\Delta x)^2}{n\Delta t}(j_0-\gamma+\{\gamma\})(j_0-\gamma+1-\{\gamma\}) \geq \frac{2(\Delta x)^2}{n\Delta t}(j_0-\gamma)^2\\ &= \frac{2(\Delta x)^2}{n\Delta t}(\frac{n\Delta t}{4\Delta x}u^n_{2j_0})^2,\end{aligned}$$

where $\{\gamma\} = \gamma - \lfloor\gamma\rfloor \in [0,1)$. Hence, we obtain (2.17) for $p=\infty$:

$$\|u^n\|_{\infty,\Delta} \leq \frac{\sqrt{8}}{\sqrt{n\Delta t}}\|u^0\|^{1/2}_{1,\Delta}. \tag{2.19}$$

The proof is now finished. □

As in the context of the continuous hyperbolic conservation laws, the asymptotic profile of the numerical solutions needs to satisfy another property. For any initial data $u_0 \in L^1(\mathbb{R})$, the solution of (2.2) converges as the time t goes to infinity to the N-wave $w_{p,q}$, determined by two quantities,

$$p = -\min_{x\in\mathbb{R}}\int_{-\infty}^{x} u_0(y)dy, \qquad q = \max_{x\in\mathbb{R}}\int_{x}^{\infty} u_0(y)dy.$$

In fact, these parameters remain invariant for all time (e.g. [64]) and the same should be expected at the discrete level. Let us remark that the mass M of the solution of (2.2) at each instant t is $M = p+q$. We already know that the mass is also preserved at the discrete level.

Theorem 2.6. *Assume that $u_0 \in L^1(\mathbb{R})$, the CFL condition $\lambda\|u^n\|_{\infty,\Delta} \leq 1$ is fulfilled and the numerical flux of a 3-point monotone conservative scheme as in* (2.10) *satisfies*

$$g(\eta,\xi) = 0, \quad \textit{when} \quad -1/\lambda \leq \eta \leq 0 \leq \xi \leq 1/\lambda, \tag{2.20a}$$

and

$$\xi - \lambda g(\xi,-\xi) \geq 0, \quad \textit{when} \quad 0 \leq \xi \leq 1/\lambda. \tag{2.20b}$$

Then, for any $n \geq 0$ the following holds:

$$\min_{k\in\mathbb{Z}}\sum_{j=-\infty}^{k} u_j^n = \min_{k\in\mathbb{Z}}\sum_{j=-\infty}^{k} u_j^0 \quad \textit{and} \quad \max_{k\in\mathbb{Z}}\sum_{j=k}^{\infty} u_j^n = \max_{k\in\mathbb{Z}}\sum_{j=k}^{\infty} u_j^0. \tag{2.21}$$

We point out that both Engquist-Osher and Godunov schemes satisfy the hypothesis of this theorem, while Lax-Friedrichs does not. Indeed, for any η, ξ such that $-1/\lambda \leq$

$\eta \leq 0 \leq \xi \leq 1/\lambda$, we have that:

$$g^{LF}(\eta, \xi) = 0 \quad \text{if and only if} \quad \xi = \eta = 0,$$

and

$$g^{EO}(\eta, \xi) = g^{G}(\eta, \xi) = 0.$$

Moreover, for any $0 \leq \xi \leq 1/\lambda$, the following holds:

$$\xi - \lambda g^{EO}(\xi, -\xi) = \xi - \lambda \xi^2 = \xi(1 - \lambda\xi) \geq 0,$$
$$\xi - \lambda g^{G}(\xi, -\xi) = \xi - \lambda \frac{\xi^2}{2} = \xi(1 - \lambda\frac{\xi}{2}) \geq 0.$$

In the case of Engquist-Osher and Godunov schemes, property (2.21) will allow us to identify the asymptotic N-wave as in the continuous case [73]. On the contrary, for the Lax-Friedrichs scheme, the lack of the conservation of these quantities produces the loss of the N-wave shape and the appearance of the diffusive wave.

Proof. For each $n \geq 0$ we define

$$p_k^n := \sum_{j=-\infty}^{k} u_j^n \quad \text{and} \quad q_k^n := \sum_{j=k}^{\infty} u_j^n.$$

Let p^n and q^n be the corresponding minimum and maximum of $\{p_k^n\}_{k\in\mathbb{Z}}$ and $\{q_k^n\}_{k\in\mathbb{Z}}$, respectively. It is easy to check that, according to (2.10), p_k^n and q_k^n satisfy:

$$p_k^{n+1} = H_p(p_{k-1}^n, p_k^n, p_{k+1}^n) \qquad \text{and} \qquad q_k^{n+1} = H_q(q_{k-1}^n, q_k^n, q_{k+1}^n),$$

where

$$H_p(x, y, z) := y - \lambda g(y - x, z - y) \qquad \text{and} \qquad H_q(x, y, z) := y + \lambda g(x - y, y - z).$$

Let us fix $n \geq 0$ and assume that the minimum of $\{p_k^n\}_{k\in\mathbb{Z}}$ is attained at some index K, p_K^n. Then, it follows that

$$u_K^n = p_K^n - p_{K-1}^n \leq 0 \leq p_{K+1}^n - p_K^n = u_{K+1}^n$$

and the maximum of $\{q_k^n\}_{k\in Z}$ is given by q_{K+1}^n. Thus, using (2.20a), we obtain that

$$p^{n+1} \leq p_K^{n+1} = p_K^n - \lambda g(p_K^n - p_{K-1}^n, p_{K+1}^n - p_K^n) = p^n - \lambda g(p^n - p_{K-1}^n, p_{K+1}^n - p^n) = p^n$$

and

$$q^{n+1} \geq q^{n+1}_{K+1} = q^n_{K+1} + \lambda g(q^n_K - q^n_{K+1}, q^n_{K+1} - q^n_{K+2}) = q^n + \lambda g(q^n_K - q^n, q^n - q^n_{K+2}) = q^n.$$

Therefore $p^n \leq p^0$ and $q^n \geq q^0$ for all $n \geq 0$.

We will now prove the reverse inequalities $p^n \geq p^0$ and $q^n \leq q^0$, that will finish the proof. This will be done by an inductive argument. Assuming that $p^n_k \geq p^0$ for all $k \in \mathbb{Z}$ we show that $p^{n+1}_k \geq p^0$. Using the identities

$$p^{n+1}_k - p^0 = p^n_k - p^0 - \lambda g(p^n_k - p^n_{k-1}, p^n_{k+1} - p^n_k) = H_p(p^n_{k-1} - p^0, p^n_k - p^0, p^n_{k+1} - p^0)$$

and

$$q^{n+1}_k - q^0 = q^n_k - p^0 + \lambda g(q^n_k - q^n_{k-1}, q^n_{k+1} - q^n_k) = H_q(q^n_{k-1} - q^0, q^n_k - q^0, q^n_{k+1} - q^0)$$

it is enough to prove that

$$H_p(x, y, z) \geq 0 \quad \text{for all} \quad x, y, z \geq 0 \tag{2.22}$$

and

$$H_q(x, y, z) \leq 0 \quad \text{for all} \quad x, y, z \leq 0. \tag{2.23}$$

Let us first prove (2.22). Set $y - x = u$ and $z - y = v$. Since $x, z \geq 0$, we have that $y \geq u$ and $y \geq -v$. We deduce that

$$y \geq \max\{u, -v, 0\}.$$

This means that

$$H_p(x, y, z) = y - \lambda g(u, v) \geq \max\{u, -v, 0\} - \lambda g(u, v) := F(u, v)$$

By the CFL condition, it is sufficient to prove that function F is nonnegative on the set

$$\Omega = \{(u, v) \in \mathbb{R}^2 : \lambda|u| \leq 1, \lambda|v| \leq 1\}.$$

We distinguish four regions in Ω, according to the sign of u and v:

$$\Omega^+_+ = \{(u, v) \in \Omega : u, v \geq 0\}, \qquad \Omega^-_- = \{(u, v) \in \Omega : u, v \leq 0\},$$
$$\Omega^-_+ = \{(u, v) \in \Omega : u \leq 0 \leq v\}, \qquad \Omega^+_- = \{(u, v) \in \Omega : u \geq 0 \geq v\}.$$

Thus, we have explicitly:

$$F(u,v) = \begin{cases} u - \lambda g(u,v), & \text{if } (u,v) \in \Omega_+^+, \\ -v - \lambda g(u,v), & \text{if } (u,v) \in \Omega_-^-, \\ -\lambda g(u,v), & \text{if } (u,v) \in \Omega_+^-, \\ \max\{|u|,|v|\} - \lambda g(u,v), & \text{if } (u,v) \in \Omega_-^+. \end{cases}$$

The monotonicity of the numerical flux g guarantees that F is increasing on v in Ω_+^+, decreasing on u in Ω_-^-, decreasing on u and increasing on v in Ω_+^-. This shows that

$$\min_{\Omega} F = \min_{\Omega_-^+} F.$$

Using that in set Ω_-^+ function F is increasing on u if $|u| \geq |v|$ and decreasing on v if $|u| \leq |v|$, we get

$$\min_{\Omega_-^+} F \geq \min_{0 \leq \xi \leq 1/\lambda} F(\xi, -\xi) = \min_{0 \leq \xi \leq 1/\lambda} \Big(\xi - \lambda g(\xi, -\xi)\Big).$$

The right-hand side of the above inequality is nonnegative due to hypothesis (2.20b). Therefore, H_p satisfies (2.22) and, hence, $p^n = p^0$ for all $n \geq 0$. Using a similar argument, the same result is proved for q^n, i.e. that $q^n = q^0$ for all $n \geq 0$. The proof is now complete. □

To conclude this section, we present a second characterization of conservative monotone schemes, that better illustrate the artificial viscosity issue we are dealing with. The difference scheme (2.9) is said to be in viscous form if there exists a function $Q : \mathbb{R}^{2k} \to \mathbb{R}$, called coefficient of numerical viscosity, such that

$$u_j^{n+1} = u_j^n - \lambda \left[\frac{(u_{j+1}^n)^2 - (u_{j-1}^n)^2}{4} \right] + \frac{Q_{j+1/2}^n (u_{j+1}^n - u_j^n) - Q_{j-1/2}^n (u_j^n - u_{j-1}^n)}{2},$$

where

$$Q_{j+1/2}^n = Q(u_{j-k+1}^n, \ldots, u_{j+k}^n).$$

Three-point monotone schemes, for instance, can be always written in that way. For simplicity, when we treat the long time behavior of the numerical schemes, we rather prefer to put them in the following equivalent form

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + \frac{(u_{j+1}^n)^2 - (u_{j-1}^n)^2}{4\Delta x} = R(u_j^n, u_{j+1}^n) - R(u_{j-1}^n, u_j^n) \tag{2.24}$$

where R can be defined in a unique manner as

$$R(u,v) = \frac{Q(u,v)(v-u)}{2\Delta t} = \frac{1}{2\Delta x}\Big(\frac{u^2}{2} + \frac{v^2}{2} - 2g(u,v)\Big). \tag{2.25}$$

We recall that for the schemes considered in Theorem 2.4 we have

$$\begin{aligned} R^{LF}(u,v) &= \frac{v-u}{2\Delta t}, \\ R^{EO}(u,v) &= \frac{1}{4\Delta x}(v|v| - u|u|), \end{aligned} \tag{2.26}$$

$$R^{G}(u,v) = \begin{cases} \frac{1}{4\Delta x}\,\mathrm{sign}(|u|-|v|)(v^2-u^2), & v \le 0 \le u, \\ \frac{1}{4\Delta x}(v|v| - u|u|), & \text{elsewhere.} \end{cases}$$

2.3 Asymptotic behavior

This section is devoted to the proof of the main result of this chapter, stated in Theorem 2.1, which describes the asymptotic profile developed by the numerical solutions of the schemes defined in Proposition 2.4, that is, those satisfying the OSLC. Our analysis uses the method of self-similar variables, i.e., a rescaling of the solutions together with the compactness of the trajectories.

The key point in the analysis of the asymptotic behavior of the solutions of our numerical schemes is the degree of homogeneity of the term $R(u,v)$. We assume that there exists a real number α such that for any $u, v \in \mathbb{R}$ and $\mu > 0$, function R satisfies

$$R(\mu u, \mu v) = \mu^{\alpha} R(u,v). \tag{2.27}$$

From (2.26), it is clear that $\alpha^{LF} = 1$ for Lax-Friedrichs, while $\alpha^{G} = \alpha^{EO} = 2$ for Godunov and Engquist-Osher, respectively.

2.3.1 The piecewise constant solution

In order to pass to the limit when doing the scaling argument, we first need to obtain bounds on the piecewise constant function u_Δ, the piecewise constant interpolation (1.10) of $\{u_j^n\}_{j\in\mathbb{Z},n\ge 0}$ solution of scheme (2.24), in some Lebesgue spaces. Let us now apply the results of Section 2.2 to u_Δ. It follows from (2.24) that it satisfies the following equation:

$$\begin{cases} \frac{u_\Delta(t+\Delta t,x)-u_\Delta(t,x)}{\Delta t} + \frac{(u_\Delta(t,x+\Delta x))^2-(u_\Delta(t,x-\Delta x))^2}{4\Delta x} \\ \qquad = R(u_\Delta(t,x), u_\Delta(t,x+\Delta x)) \\ \qquad\quad -R(u_\Delta(t,x-\Delta x), u_\Delta(t,x)), & t \geq 0, \text{ a.e. } x \in \mathbb{R}, \\ u_\Delta(t,x) = u_\Delta^0(x), & t \in [0,\Delta t). \end{cases} \tag{2.28}$$

The following lemma gives us the first bound on the solution u_Δ. In the sequel, for any functions f and g, we will write $f \lesssim g$ if there exists a constant $C > 0$ such that $f \leq Cg$.

Lemma 2.7. *There exists a positive constant $C = C(\Delta t, \|u_0\|_{1,\Delta})$ such that the following holds*

$$\|u_\Delta(t)\|_\infty \leq \frac{C}{\sqrt{t}}, \quad \forall t > \Delta t.$$

Proof. From Proposition 2.5 we know that for any $n \geq 1$ the following holds

$$\|u^n\|_{\infty,\Delta} \leq \frac{C}{\sqrt{n\Delta t}} \|u^0\|_{1,\Delta}.$$

Let us now consider $t \in [n\Delta t, (n+1)\Delta t)$ with $n \geq 1$. Then

$$\|u_\Delta(t)\|_\infty = \|u^n\|_{\infty,\Delta} \leq \frac{C}{\sqrt{n\Delta t}} \leq \frac{2C}{\sqrt{(n+1)\Delta t}} \leq \frac{2C}{\sqrt{t}},$$

which proves the desired inequality. □

For the simplicity of presentation, from now on we will denote by $\omega(h)$ the $L^1(\mathbb{R})$-modulus of continuity of the initial data u_Δ^0:

$$\omega(h) = \int_{\mathbb{R}} |u_\Delta^0(x+h) - u_\Delta^0(x)| dx.$$

Lemma 2.8. *The solution of system* (2.28) *satisfies*

$$\int_{\mathbb{R}} |u_\Delta(t,x+h) - u_\Delta(t,x)| dx \leq \omega(h)$$

for all $h > 0$ and $t > 0$.

Proof. Let us consider $k \in \mathbb{Z}$ such that $k\Delta x \leq h < (k+1)\Delta x$. Then for any piecewise constant function v as in (1.10), we have

$$\begin{aligned}\int_{\mathbb{R}}|v(x+h)-v(x)|dx &= \sum_{j\in\mathbb{Z}}\int_{x_{j-1/2}}^{x_{j+1/2}}|v(x+h)-v(x)|dx\\ &= \sum_{j\in\mathbb{Z}}\int_{x_{j-1/2}}^{x_{j-1/2}+x_{k+1}-h}|v(x+h)-v(x)|dx+\int_{x_{j-1/2}+x_{k+1}-h}^{x_{j+1/2}}|v(x+h)-v(x)|dx\\ &= (x_{k+1}-h)\sum_{j\in\mathbb{Z}}|v_{j+k}-v_j|+(h-x_k)\sum_{j\in\mathbb{Z}}|v_{j+k+1}-v_j|.\end{aligned}$$

Applying this property to function u_Δ and using that for any $k\geq 1$ (cf. Proposition 2.2),

$$\sum_{j\in\mathbb{Z}}|u^n_{j+k}-u^n_j|\leq\sum_{j\in\mathbb{Z}}|u^0_{j+k}-u^0_j|$$

we obtain that

$$\begin{aligned}\int_{\mathbb{R}}|u_\Delta(t,x+h)-u_\Delta(t,x)|dx&\\ &\leq (x_{k+1}-h)\sum_{j\in\mathbb{Z}}|u^0_{j+k}-u^0_j|+(h-x_k)\sum_{j\in\mathbb{Z}}|u^0_{j+k+1}-u^0_j|\\ &= \omega(h).\end{aligned}$$

This proves the desired result. □

2.3.2 The rescaled solutions

Let us now introduce for any $\mu>0$ the family of rescaled solutions

$$u^\mu(t,x)=\mu u_\Delta(\mu^2 t,\mu x),\quad t\geq 0, x\in\mathbb{R}.$$

It follows that u^μ is piecewise constant on time intervals of length $\Delta t/\mu^2$ and space intervals of length $\Delta x/\mu$. Moreover, it satisfies the system

$$\begin{cases}\frac{\mu^2}{\Delta t}\Big(u^\mu(t+\frac{\Delta t}{\mu^2},x)-u^\mu(t,x)\Big)+\frac{\mu}{4\Delta x}\Big((u^\mu(t,x+\frac{\Delta x}{\mu}))^2-(u^\mu(t,x-\frac{\Delta x}{\mu}))^2\Big)\\ \quad=\mu^{1-\alpha}\Big(\mu^2R(u^\mu(t,x),u^\mu(t,x+\frac{\Delta x}{\mu}))\\ \qquad\qquad -\mu^2R(u^\mu(t,x-\frac{\Delta x}{\mu}),u^\mu(t,x))\Big), & t\geq 0,\ \text{a.e. } x\in\mathbb{R},\\ u^\mu_\Delta(t,x)=\mu u^0_\Delta(0,\mu x), & t\in[0,\Delta t/\mu^2).\end{cases}\tag{2.29}$$

The following lemmas will guarantee the convergence of the trajectories $\{u^\mu(t)\}_{\mu>0}$ as $\mu\to\infty$.

Lemma 2.9. *The solution of system* (2.29) *satisfies the following two estimates:*

1. *There exists a positive constant C independent of μ such that*

$$\|u^\mu(t)\|_\infty \le \frac{C}{\sqrt{t}}, \quad \forall t > \frac{\Delta t}{\mu^2}. \tag{2.30}$$

2. For all $h > 0$ and $t > 0$ the following holds

$$\int_{\mathbb{R}} |u^\mu(t, x+h) - u^\mu(t,x)|dx \le \omega(h).$$

Remark 2.1*: On the interval $[0, \Delta t/\mu^2]$ we have the rough estimate*

$$\|u^\mu(t)\|_\infty = \mu \|u^0_\Delta\|_\infty \le \frac{\mu}{\Delta x}\|u^0_\Delta\|_1. \tag{2.31}$$

Proof. The first estimate is a consequence of Lemma 2.7, while the second one follows from Lemma 2.8. □

Lemma 2.10. *For any $0 < t_1 < t_2$, there exists a positive constant C such that*

$$\sup_{t\in[t_1,t_2]} \int_{\mathbb{R}} |u^\mu(t+h,x) - u^\mu(t,x)|dx \le C(h^{1/3} + \frac{h^{2/3}}{\sqrt{t_1}})\|u_0\|_1 + \omega(h^{1/3})$$

holds for any $h > 0$ and $\mu > \sqrt{\frac{\Delta t}{t_1}}$.

Proof. We proceed as in [61, 62]. Since u^μ is piecewise constant in time, it is sufficient to consider the case when $t_1, t_2 \in (\Delta t/\mu^2)\mathbb{Z}$ and $h = k\Delta t/\mu^2$ with $k \in \mathbb{Z}$, $k \ge 1$. Let us set $t' = t + h - \Delta t/\mu^2$. Then, for $t \in (\Delta t/\mu^2)\mathbb{Z}$ we have

$$u^\mu(t+h) - u^\mu(t) = \frac{\mu^2}{\Delta t}\int_t^{t'} \Big(u^\mu(s + \frac{\Delta t}{\mu^2}) - u^\mu(s)\Big)ds.$$

Let us choose ϕ a smooth, bounded function on $\mathbb{R}$. Multiplying (2.29) by ϕ and integrating in time and space we get

$$\begin{aligned}
&\int_{\mathbb{R}} (u^\mu(t+h,x) - u^\mu(t,x))\phi(x)dx = I_1 + I_2 = \\
&= \frac{\mu}{4\Delta x}\int_t^{t'}\int_{\mathbb{R}} (u^\mu(s,x))^2\Big(\phi(x + \frac{\Delta x}{\mu}) - \phi(x - \frac{\Delta x}{\mu})\Big)dsdx \\
&\quad + \mu^{1-\alpha}\int_t^{t'}\int_{\mathbb{R}} \mu^2 R\Big(u^\mu(s,x), u^\mu(s, x+\frac{\Delta x}{\mu})\Big)\Big(\phi(x) - \phi(x+\frac{\Delta x}{\mu})\Big)dsdx.
\end{aligned} \tag{2.32}$$

We now evaluate I_1 and I_2. Observe that since $\mu^2 > \Delta t/t_1$ then $t \ge t_1 \ge \Delta t/\mu^2$ and estimate (2.30) applies

$$\begin{aligned}
|I_1| &\lesssim \|\phi'\|_\infty \int_t^{t'}\int_{\mathbb{R}} (u^\mu(s,x))^2 dsdx \lesssim \|\phi'\|_\infty \int_t^{t'} \|u^\mu(s)\|_\infty \int_{\mathbb{R}} |u^\mu(s,x)|dsdx \\
&\lesssim \|\phi'\|_\infty \|u_0\|_1 \int_t^{t'} \frac{ds}{\sqrt{s}} \lesssim \|\phi'\|_\infty \|u_0\|_1 \frac{h}{\sqrt{t}}.
\end{aligned} \tag{2.33}$$

In the case of I_2, we use that $R(u,v)$ satisfies $R(u,v) = (v-u)/(2\Delta t)$ for the Lax-Friedrichs scheme and $|R(u,v)| \lesssim |u|^2 + |v|^2$ for Engquist-Osher and Godunov schemes. For Lax-Friedrichs ($\alpha^{LF} = 1$) we have for all $t > 0$ that

$$\begin{aligned} I_2 &= \int_t^{t'} \int_{\mathbb{R}} \frac{\mu^2}{2\Delta t} u^\mu(s,x)\Big(\phi(x + \frac{\Delta x}{\mu}) - 2\phi(x) + \phi(x - \frac{\Delta x}{\mu})\Big) dx ds \\ &\lesssim \|\phi''\|_\infty \int_t^{t'} \int_{\mathbb{R}} |u^\mu(s,x)| dx ds \le h \|\phi''\|_\infty \|u_0\|_1. \end{aligned} \tag{2.34}$$

For Engquist-Osher and Godunov schemes ($\alpha^{EO} = \alpha^G = 2$), we have similar estimates as in the case of I_1:

$$I_2 \lesssim \|\phi'\|_\infty \int_t^{t'} \int_{\mathbb{R}} (u^\mu(s,x))^2 ds dx \lesssim \|\phi'\|_\infty \|u_0\|_1 \frac{h}{\sqrt{t}}. \tag{2.35}$$

Plugging estimates (2.33), (2.34) and (2.35) into (2.32), we obtain that

$$\int_{\mathbb{R}} (u^\mu(t+h,x) - u^\mu(t,x))\phi(x) dx \lesssim \|u_0\|_1 \Big(h\|\phi''\|_\infty + \frac{h}{\sqrt{t}} \|\phi'\|_\infty \Big).$$

Let us choose a mollifier ρ, a smooth nonnegative function supported in the interval $(-1,1)$ with unit mass, and take

$$\phi_h = h^{-1/3} \rho(h^{-1/3}) * \Big[\operatorname{sign} u^\mu(t+h) - \operatorname{sign} u^\mu(t) \Big].$$

We have that $|\phi_h'| \le h^{-1/3}$, $|\phi_h''| \le h^{-2/3}$ and

$$\int_{\mathbb{R}} (u^\mu(t+h,x) - u^\mu(t,x))\phi_h(x) dx \lesssim \|u_0\|_1 \Big(h^{1/3} + \frac{h^{2/3}}{\sqrt{t}} \Big). \tag{2.36}$$

Using that ϕ_h has unit mass and that for any $a, b \in \mathbb{R}$ we have $|a| - a \operatorname{sign}(b) \le 2|a-b|$, we get

$$\begin{aligned} &|u^\mu(t+h,x) - u^\mu(t,x)| - (u^\mu(t+h,x) - u^\mu(t,x))\phi_h(x) \\ &= \int_{\mathbb{R}} \phi_h(x-y) \Big(|u^\mu(t+h,x) - u^\mu(t,x)| \\ &\qquad - (u^\mu(t+h,x) - u^\mu(t,x)) \big(\operatorname{sign} u^\mu(t+h,y) - \operatorname{sign} u^\mu(t,y) \big) \Big) dy \\ &\le 2 \int_{\mathbb{R}} \phi_h(x-y) \big| (u^\mu(t+h,x) - u^\mu(t,x) - (u(t+h,y) - u(t,y)) \big| dy \\ &\le 2 \int_{\mathbb{R}} \phi_h(x-y) |u^\mu(t+h,x) - u^\mu(t+h,y)| dy \\ &\quad + 2 \int_{\mathbb{R}} \phi_h(x-y) |u^\mu(t,x) - u^\mu(t,y)| dy. \end{aligned}$$

Integrating the above inequality in x we obtain that

$$\begin{aligned}\int_{\mathbb{R}} &|u^\mu(t+h,x)-u^\mu(t,x)| - (u^\mu(t+h,x)-u^\mu(t,x))\phi_h(x)dx \\ &\leq 2\int_{\mathbb{R}} |u^\mu(t+h,x+h^{1/3})-u^\mu(t+h,x)|dx \\ &\qquad + 2\int_{\mathbb{R}} |u^\mu(t,x+h^{1/3})-u^\mu(t,x)|dx \\ &\leq 4\omega(h^{1/3}). \end{aligned} \tag{2.37}$$

Combining (2.36) and (2.37) we obtain the desired result. □

Lemma 2.11. *There exists a constant $C = C(\|u_0\|_1)$ such that*

$$\int_{|x|>2R} |u^\mu(t,x)|dx \leq \int_{|x|>R} |u^0_\Delta|dx + C(\frac{t}{R^2} + \frac{\mu^{-1}+t^{1/2}}{R}). \tag{2.38}$$

holds for any $t>0$, $R>0$ and $\mu>1$.

Proof. We first observe that it is sufficient to consider nonnegative initial data. Indeed, choosing $\tilde{u}^0 = |u^0|$ as initial data in the numerical scheme, we have by the maximum principle that $|u^\mu(t,x)| \leq \tilde{u}^\mu(t,x)$ where $\tilde{u}^\mu$ is the solution that corresponds to the initial data $\tilde{u}^0$. It is then sufficient to prove estimate (2.38) for nonnegative initial data and solutions.

Let us now prove (2.38) for nonnegative solutions. Since u^μ is piecewise constant in time we consider the case $t = k\Delta t/\mu^2$, $k \in \mathbb{Z}$, $k \geq 1$, the case $k=0$ being obvious. Let us choose $\rho \in C^\infty(\mathbb{R})$ such that $0 \leq \rho \leq 1$ and

$$\rho = \begin{cases} 0, & |x| \leq 1, \\ 1, & |x| \geq 2. \end{cases}$$

We set $\rho_R(x) = \rho(x/R)$. We multiply system (2.29) by ρ_R and integrate on $(0,t') \times \mathbb{R}$ where $t' = t - \Delta t/\mu^2$. The right hand side is given by

$$\begin{aligned}\frac{\mu^2}{\Delta t}\int_0^{t'}\int_{\mathbb{R}} \Big(u^\mu(s+\frac{\Delta t}{\mu^2},x) - u^\mu(s,x)\Big)\rho_R(x)dxds \\ = \int_{\mathbb{R}} u^\mu(t,x)\rho_R(x)dx - \int_{\mathbb{R}} u^\mu(0,x)\rho_R(x)dx.\end{aligned}$$

Hence

$$\begin{aligned}
&\int_{\mathbb{R}} u^\mu(t,x)\rho_R(x)dx - \int_{\mathbb{R}} u^\mu(0,x)\rho_R(x)dx \\
&\quad = \int_0^{t'}\int_{\mathbb{R}} (u^\mu(s,x))^2\Big[\rho_R(x+\frac{\Delta x}{\mu}) - \rho_R(x-\frac{\Delta x}{\mu})\Big]\frac{\mu}{4\Delta x}dxds \\
&\qquad + \mu^{1-\alpha}\int_0^{t'}\int_{\mathbb{R}} \mu^2 R\Big(u^\mu(s,x), u^\mu(s+\frac{\Delta x}{\mu})\Big)\Big[\rho_R(x) - \rho_R(x+\frac{\Delta x}{\mu})\Big]dxds. \\
&\quad = I_1 + I_2.
\end{aligned}$$

In the first case, using (2.30) and (2.31) we get

$$\begin{aligned}
I_1 &\lesssim \|\rho_R'\|_\infty \int_0^{t'}\int_{\mathbb{R}} (u^\mu(s,x))^2 dsdx \lesssim \frac{1}{R}\int_0^{t'} \|u^\mu(s)\|_\infty \int_{\mathbb{R}} |u^\mu(s,x)|dsdx \\
&\le \frac{1}{R}\|u_0\|_1\Big(\int_0^{\Delta t/\mu^2} \|u^\mu(s)\|_\infty + \int_{\Delta t/\mu^2}^{t'} \frac{1}{\sqrt{s}}ds\Big) \lesssim R^{-1}\|u_0\|_1(\frac{1}{\mu} + t^{1/2}).
\end{aligned}$$

In the case of I_2, using the same argument as in Lemma 2.10, we get

$$I_2 \lesssim \begin{cases} \frac{\mu^{-1}+t^{1/2}}{R}\|u_0\|_1, & \alpha = 2, \\ \\ \frac{t}{R^2}\|u_0\|_1, & \alpha = 1. \end{cases}$$

It follows that

$$\begin{aligned}
\int_{|x|>2R} u^\mu(t,x)dx &\le \int_{|x|>\mu R} u_\Delta^0(x)dx + C(\frac{t}{R^2} + \frac{\mu^{-1}+t^{1/2}}{R}) \\
&\le \int_{|x|>R} u_\Delta^0(x)dx + C(\frac{t}{R^2} + \frac{\mu^{-1}+t^{1/2}}{R}).
\end{aligned}$$

The proof is now complete. □

2.3.3 Passing to the limit

We are now in position to prove the main result of this chapter, stated in Theorem 2.1. The results obtained in the previous section will guarantee the compactness of the set $\{u^\mu\}_{\mu>0}$ needed to pass to the limit.

Proof of Theorem 2.1. We proceed in several steps.

Step I. Passing to the limit as $\mu \to \infty$. From Riesz-Fréchet-Kolmogorov and Arzelà-Ascoli theorems and Lemmas 2.9, 2.10 and 2.11, we infer that the set $\{u^\mu\}_{\mu>0}$ is relatively compact in $C([t_1,t_2];L^1(\mathbb{R}))$ for any $0 < t_1 < t_2$. Consequently, there exist a

subsequence, which we do not relabel, and a function $u^\infty \in C((0,\infty); L^1(\mathbb{R}))$ such that for any $0 < t_1 < t_2$

$$u^\mu \to u^\infty \text{ in } C([t_1,t_2]; L^1(\mathbb{R})) \text{ as } \mu \to \infty \tag{2.39}$$

and

$$u^\mu(t,x) \to u^\infty(t,x), \quad \text{a.e. } (t,x) \in (0,\infty)\times\mathbb{R}. \tag{2.40}$$

Using the mass conservation of u^μ we obtain that

$$\int_{\mathbb{R}} u^\infty(t,x)dx = M_\Delta = \int_{\mathbb{R}} u^0_\Delta(x)dx.$$

Moreover, the almost everywhere convergence in (2.40) shows that there is a positive constant C such that the limit function u^∞ satisfies

$$t^{\frac{1}{2}}\|u^\infty(t)\|_\infty \le C, \quad \forall t > 0. \tag{2.41}$$

We will now pass to the limit in the sense of distributions in equation (2.29). Let us multiply it by a test function $\varphi \in C_c^\infty((0,\infty)\times\mathbb{R})$ and integrate it both in space and time. The limit in the left-hand side is

$$u_t^\infty + \Big(\frac{(u^\infty)^2}{2}\Big)_x, \quad \text{in } \mathcal{D}((0,\infty)\times\mathbb{R}).$$

It remains to identify the limit for the right-hand side. Let us denote

$$I_\mu := \mu^{1-\alpha}\int_0^\infty\int_{\mathbb{R}} \Big(\mu^2 R(u^\mu(t,x), u^\mu(t,x+\frac{\Delta x}{\mu})) \\ - \mu^2 R(u^\mu(t,x-\frac{\Delta x}{\mu}), u^\mu(t,x))\Big)\varphi(t,x)dx.$$

In the case of the Lax-Friedrichs scheme, $\alpha^{LF} = 1$ and $R(u,v) = (v-u)/(2\Delta t)$. Thus

$$I_\mu = \frac{\mu^2}{2\Delta t}\int_0^\infty\int_{\mathbb{R}} u^\mu(t,x)\Big(\varphi(t,x+\frac{\Delta x}{\mu})) - 2\varphi(t,x) + \varphi(t,x-\frac{\Delta x}{\mu})\Big)dxdt \\ \longrightarrow \frac{(\Delta x)^2}{2\Delta t}\int_0^\infty\int_{\mathbb{R}} u^\infty\varphi_{xx}dxdt, \quad \text{as } \mu\to\infty.$$

Hence the limit u^∞ satisfies

$$u_t^\infty + \Big(\frac{(u^\infty)^2}{2}\Big)_x = \frac{(\Delta x)^2}{2\Delta t}u^\infty_{xx} \quad \text{in } \mathcal{D}'((0,\infty)\times\mathbb{R}). \tag{2.42}$$

In the case of Engquist-Osher and Godunov schemes, $\alpha = 2$. Using the explicit form of $R(u,v)$ given in (2.26), we obtain that

$$|R(u,v)| \lesssim ||u|u - |v|v| \le |u-v|(|u|+|v|).$$

Assume that φ is supported in the time interval $[t_1, t_2]$ with $t_1 > 0$. Then

$$\begin{aligned}
|I_\mu| &= \mu^{-1} \left| \int_{t_1}^{t_2} \int_{\mathbb{R}} \mu^2 R(u^\mu(t,x), u^\mu(t, x + \frac{\Delta x}{\mu})) \Big(\varphi(t,x)dx - \varphi(t, x + \frac{\Delta x}{\mu}) \Big) dx \right| \\
&\lesssim \Delta x \|\varphi'\|_\infty \int_{t_1}^{t_2} \int_{\mathbb{R}} |R(u^\mu(t,x), u^\mu(t, x + \frac{\Delta x}{\mu}))| dx dt \\
&\lesssim \Delta x \|\varphi'\|_\infty \int_{t_1}^{t_2} \int_{\mathbb{R}} |u^\mu(t,x) - u^\mu(t, x + \frac{\Delta x}{\mu})| \big(|u^\mu(t,x)| + |u^\mu(t, x + \frac{\Delta x}{\mu})| \big) dx dt \\
&\lesssim \Delta x \|\varphi'\|_\infty C(t_1) \max_{t \in [t_1, t_2]} \int_{\mathbb{R}} |u^\mu(t,x) - u^\mu(t, x + \frac{\Delta x}{\mu})| dx.
\end{aligned}$$

Using Lemma 2.10 we obtain that $I_\mu \to 0$ as $\mu \to \infty$. Therefore, the limit point u^∞ satisfies

$$u^\infty_t + \Big(\frac{(u^\infty)^2}{2} \Big)_x = 0 \quad \text{in } \mathcal{D}'((0,\infty) \times \mathbb{R}). \tag{2.43}$$

Let us now recall that in view of the OSLC (2.13) for any $t > \Delta t$ and a.e. $x \in \mathbb{R}$ we have

$$\frac{u_\Delta(t, x + \Delta x) - u_\Delta(t, x - \Delta x)}{2\Delta x} \le \frac{C}{t}.$$

Hence, for all $t > \Delta t/\mu^2$ we have

$$\frac{\mu^2}{2\Delta x} \Big(u^\mu(t, x + \frac{\Delta x}{\mu^2}) - u^\mu(t, x - \frac{\Delta x}{\mu^2}) \Big) \le \frac{C}{t}.$$

Letting $\mu \to \infty$, we obtain that for any $t > 0$ the limit point u^∞ satisfies

$$u^\infty_x(t) \le \frac{C}{t} \quad \text{in } \mathcal{D}'(\mathbb{R}).$$

This shows that u^∞ is an entropy solution. Note that this can also be guaranteed by the monotonicity of the numerical schemes, as monotone schemes are consistent with any entropy condition [38, Chapter 3].

Step II. Initial data. It remains to identify the behavior of u^∞ as $t \to 0$. We will prove that $u^\infty(t) \to M_\Delta \delta_0$ as $t \to 0$, in the sense of bounded measures, i.e.,

$$\lim_{t \to 0} \int_{\mathbb{R}} u^\infty(t,x) \varphi(x) dx = M_\Delta \varphi(0) \tag{2.44}$$

for every bounded continuous function φ. By a density argument it is sufficient to consider the case $\varphi \in C_c^\infty(\mathbb{R})$. Thus we will conclude that $u_\infty(0) = M_\Delta \delta_0$ in the sense of bounded measures.

Let us choose $t = k\Delta t/\mu^2$, $k \in \mathbb{Z}$, $k \geq 1$. Then for any $k \geq 1$ (for $k = 0$ it is obvious)

$$\begin{aligned}\frac{\mu^2}{\Delta t}\int_0^t \int_{\mathbb{R}} \Big(u^\mu(s+\frac{\Delta t}{\mu^2},x) - u^\mu(s,x)\Big)\varphi(x)dxds \\ = \frac{\mu^2}{\Delta t}\int_t^{t+\frac{\Delta t}{\mu^2}} \int_{\mathbb{R}} u^\mu(s,x)\varphi(x)dxds - \frac{\mu^2}{\Delta t}\int_0^{\frac{\Delta t}{\mu^2}} \int_{\mathbb{R}} u^\mu(s,x)\varphi(x)dxds \\ = \int_{\mathbb{R}} u^\mu(t,x)\varphi(x)dx - \int_{\mathbb{R}} u^\mu(0,x)\varphi(x)dx.\end{aligned}$$

Let us consider a t such that $k\Delta t/\mu^2 \leq t < (k+1)\Delta t/\mu^2$ for some $k \geq 0$. Since u^μ is piecewise constant, we have:

$$\begin{aligned}\int_{\mathbb{R}} u^\mu(t,x)\varphi(x)dx - \int_{\mathbb{R}} u^\mu(0,x)\varphi(x)dx = \int_{\mathbb{R}} \Big(u^\mu(k\frac{\Delta t}{\mu^2},x) - u^\mu(0,x)\Big)\varphi(x)dx \\ = \frac{\mu^2}{\Delta t}\int_0^{k\frac{\Delta t}{\mu^2}} \int_{\mathbb{R}} \Big(u^\mu(s+\frac{\Delta t}{\mu^2},x) - u^\mu(s,x)\Big)\varphi(x)dxds \\ = \frac{\mu}{4\Delta x}\int_0^{k\frac{\Delta t}{\mu^2}} \int_{\mathbb{R}} (u^\mu(s,x))^2\Big(\varphi(x+\frac{\Delta x}{\mu}) - \varphi(x-\frac{\Delta x}{\mu})\Big)dsdx \\ + \mu^{1-\alpha}\int_0^{k\frac{\Delta t}{\mu^2}} \mu^2 R\Big(u^\mu(s,x), u^\mu(s,x+\frac{\Delta x}{\mu})\Big)(\varphi(x) - \varphi(x+\frac{\Delta x}{\mu}))dsdx.\end{aligned}$$

Following the same steps as in Lemma 2.10 and using that $k\Delta t/\mu^2 \leq t$ we obtain that

$$\begin{aligned}\Big|\int_{\mathbb{R}} u^\mu(t,x)\varphi(x)dx - \int_{\mathbb{R}} u^\mu(0,x)\varphi(x)dx\Big| \\ \lesssim \|u_\Delta^0\|_1\big(\|\varphi'\|_\infty(\mu^{-1} + t^{1/2}) + \|\varphi''\|_\infty t\big).\end{aligned}$$

Using the definition of $u^\mu(0,x)$ and letting $\mu \to \infty$ we get

$$\Big|\int_{\mathbb{R}} u^\infty(t,x)\varphi(x)dx - \varphi(0)\Big(\int_{\mathbb{R}} u_\Delta^0(x)dx\Big)\Big| \leq C(\varphi)(t^{1/2} + t).$$

The proof of (2.44) is now complete.

Step III. Identification of the limit. In the case of the Lax-Friedrichs scheme system (2.42-2.44) has a unique solution w_M given by (1.14). Since w_M is the only possible accumulation point of $\{u^\mu\}_{\mu>0}$ in $C((0,\infty), L^1(\mathbb{R}))$ as $\mu \to \infty$, the whole family converges to w_M. Therefore:

$$\lim_{\mu\to\infty} \|u^\mu(1) - w_M(1)\|_1 = 0$$

so, setting $\mu = t^{1/2}$, we recover (2.5) for $p = 1$. Hölder's inequality and Proposition 2.5 allow us to deduce (2.5) for $p \in (1,\infty)$.

In the case of Engquist-Osher and Godunov schemes, as proved in [73], there are

infinitely many solutions $w_{p_\Delta,q_\Delta} \in C((0,\infty), L^1(\mathbb{R}))$ of system (2.43-2.44), so we have to identify the parameters p_Δ and q_Δ. As pointed out in [73], it remains to identify the limit as $t \downarrow 0$ of

$$v(t,x) = \int_{-\infty}^{x} u^\infty(t,y)dy.$$

Since u^∞ converges to $M\delta_0$, we have:

$$\lim_{t\to 0} \int_{-\infty}^{x} u^\infty(t,y)dy = 0, \ \forall x < 0 \quad \text{and} \quad \lim_{t\to 0} \int_{-\infty}^{x} u^\infty(t,y)dy = M_\Delta, \ \forall x > 0.$$

It remains to determine the above limit when $x = 0$. Note that the map $t \to v(t,0)$ is increasing when $t \downarrow 0$ so that there exists

$$-l = \lim_{t\downarrow 0} v(t,0).$$

This proves that $u^\infty = w_{l,l+M_\Delta}$. To finish the proof it remains to show that $l = p_\Delta$. According to [73], parameter l is characterized by

$$-l = \min_{x\in\mathbb{R}} \int_{-\infty}^{x} w_{l,l+M_\Delta}(t,y)dy = \min_{x\in\mathbb{R}} \int_{-\infty}^{x} u^\infty(t,y)dy.$$

So it is sufficient to prove that

$$\min_{x\in\mathbb{R}} \int_{-\infty}^{x} u^\infty(t,y)dy = \min_{x\in\mathbb{R}} \int_{-\infty}^{x} u^0_\Delta(y)dy.$$

By Theorem 2.6, we know that

$$\min_{x\in\mathbb{R}} \int_{-\infty}^{x} u^0_\Delta(y)dy = \min_{x\in\mathbb{R}} \int_{-\infty}^{x} u_\Delta(\mu^2 t, y)dy = \min_{x\in\mathbb{R}} \int_{-\infty}^{x} u^\mu(t,y)dy.$$

Since u^μ converges to u^∞ in $L^1(\mathbb{R})$, its primitive converges uniformly to the primitive of u^∞ when $\mu \to \infty$. So we have

$$\min_{x\in\mathbb{R}} \int_{-\infty}^{x} u^\infty(t,y)dy = \lim_{\mu\to\infty} \min_{x\in\mathbb{R}} \int_{-\infty}^{x} u^\mu(t,y)dy = \min_{x\in\mathbb{R}} \int_{-\infty}^{x} u^0_\Delta(y)dy = -p_\Delta.$$

We conclude that u^∞ is the unique solution w_{p_Δ,q_Δ} to (2.7) with p_Δ and q_Δ as in Theorem 2.1. Since w_{p_Δ,q_Δ} is the is the only possible accumulation point of $\{u^\mu\}$ in $C((0,\infty), L^1(\mathbb{R}))$ as $\mu \to \infty$, the whole family converges to w_{p_Δ,q_Δ}. Therefore:

$$\lim_{\mu\to\infty} \|u^\mu(1) - w_{p_\Delta,q_\Delta}(1)\|_1 = 0$$

so, setting $\mu = t^{1/2}$, we recover (2.5) for $p = 1$. Hölder's inequality and Proposition 2.5 allow us to deduce assertion (2.5) for $p \in (1,\infty)$. This completes the proof of the main result of this chapter. □

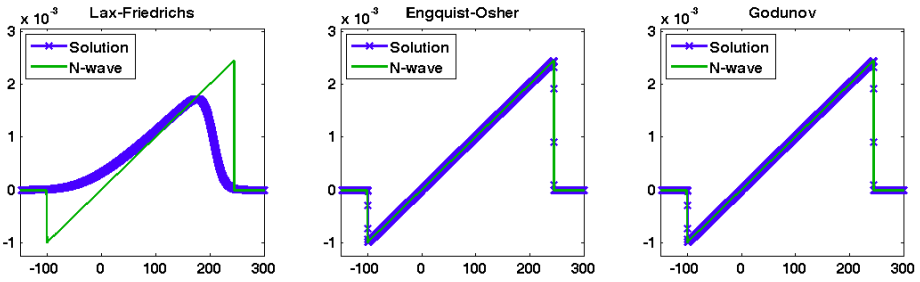


FIGURE 2.1: Solution to the Burgers equation at $t = 10^5$ using Lax-Friedrichs (left), Engquist-Osher (center) and Godunov (right) schemes. The thin line corresponds to the predicted N-wave, defined as in (1.15).

2.4 Simulations

In the following, we illustrate the main results of previous sections with some numerical simulations. Let us consider the inviscid Burgers equation (2.2) with initial data

$$u_0(x) = \begin{cases} -0.05, & x \in [-1, 0], \\ 0.15, & x \in [0, 2], \\ 0, & \text{elsewhere.} \end{cases}$$

In this case, the parameters that describe the asymptotic N-wave profile, defined in (1.15), are:

$$M = 0.25\ , \quad p = 0.05 \quad \text{and} \quad q = 0.3.$$

We focus our experiments on the schemes described in Proposition 2.4: on the one hand, Engquist-Osher and Godunov schemes, as examples of well-behaving schemes, and on the other, the Lax-Friedrichs scheme.

For the spatial domain discretization, we take $\Delta x = 0.1$ as the mesh size for the interval $[-350, 800]$. Let us remark that, in general, it is not possible to impose homogeneous Dirichlet boundary conditions on both sides of the interval (e.g. [6]). Nevertheless, due to the finite speed of propagation, we can consider a large enough domain to guarantee that the boundary conditions do not interfere on the solution. Regarding the time-step, we simply choose $\Delta t = 0.5$, that verifies the CFL condition in the three cases.

In Figure 2.1 we show the numerical solution obtained at time $t = 10^5$. It is possible to observe how the numerical viscosity of Lax-Friedrichs has dissipated the negative part of the solution. After such a long time, it only remains a diffusive positive profile, i.e., the wave described in the first case of Theorem 2.1. On the contrary, both Engquist-Osher and Godunov schemes preserve the shape of the N-wave.

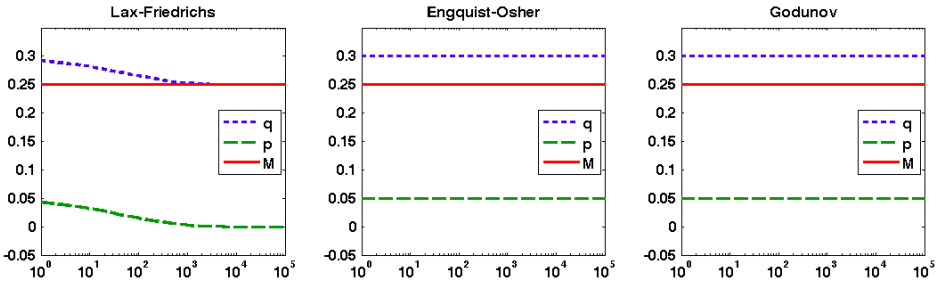


FIGURE 2.2: Evolution in time (on a logarithmic scale) of the total mass of the solution (continuous), together with the positive (dotted) and negative (dashed) masses, using Lax-Friedrichs (left), Engquist-Osher (center) and Godunov (right) schemes.

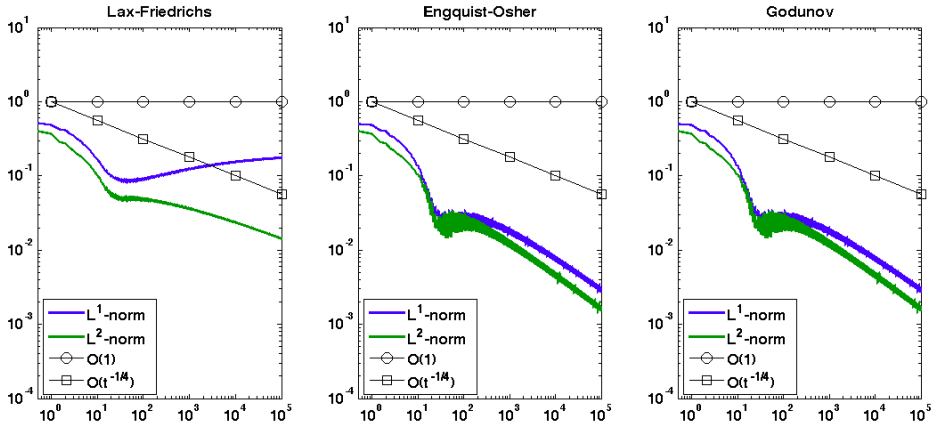


FIGURE 2.3: Evolution in time (on a logarithmic scale) of the L^1 and L^2 norms of the difference between the N-wave and the numerical solution given by Lax-Friedrichs (left), Engquist-Osher (center) and Godunov (right) schemes.

We can confirm this loss of the negative part of the solution for the Lax-Friedrichs scheme in Figure 2.2. While the mass of the solution is conserved throughout time in the three cases, the Lax-Friedrichs scheme fails to preserve, in addition, both the masses of the positive and negative parts, respectively. Let us notice that when a solution crosses the horizontal axis just once, these masses are equivalent to the parameters p and q computed at each time.

Finally, in Figure 2.3 we show the evolution of the L^1 and L^2 norms of the difference between the numerical solution and the N-wave. This confirms that, for large times, the behavior of the solutions obtained by the Engquist-Osher and Godunov schemes are the expected ones. By the contrary, the performance of the Lax-Friedrichs scheme is far far from being correct.

2.5 Similarity variables

One of the disadvantages of the numerical approach we have developed in the previous sections is that the considered computational domain has to be extremely large, in comparison to the support of the initial data because of its time-spreading. In this section we use similarity variables, a classical tool at the continuous level and that, as we shall see, at the discrete one, leads to an alternate way of understanding the large-time behavior and to a significant decrease of the computational cost. In [60] similarity variables were used to analyze the transition to the asymptotic states to estimate the time of evolution from an N-wave to the final stage of a diffusion wave for the viscous Burgers equation. As we shall see, the same phenomena occurs for numerical schemes in case its effective asymptotic numerical positive is non-negligible as it happens with the Lax-Friedrichs scheme.

Let us consider the change of variables given by:

$$s = \ln(t+1), \qquad \xi = x/\sqrt{t+1}, \qquad w(\xi, s) = \sqrt{t+1}\; u(x,t),$$

which turns (2.2) into

$$w_s + \left(\frac{1}{2}w^2 - \frac{1}{2}\xi w\right)_\xi = 0, \quad \xi \in \mathbb{R}, s > 0. \tag{2.45}$$

In this case, the asymptotic profile of the solutions is a N-wave as follows:

$$N_{p,q}(\xi) = \begin{cases} \xi, & -\sqrt{2p} < \xi < \sqrt{2q}, \\ 0, & \text{elsewhere}, \end{cases} \tag{2.46}$$

where p and q are, respectively, the negative and positive mass of the initial data.

The asymptotic profiles of the Burgers equation in the original variables become, in the similarity ones, steady state solutions. Accordingly, in the similarity variables, the asymptotic convergence towards a self-similar solution in the self-similar ones becomes, simply, the convergence towards steady-states. One further advantage of considering similarity variables is that the the support of the solutions no longer grows indefinitely and, thus, their numerical approximation is easier to handle.

2.5.1 Presentation of discrete similarity schemes

In our numerical analysis, we first need to adjust the three numerical schemes under consideration to the similarity variables. The general form of the scheme is still given

in the conservative form

$$w_j^{n+1} = w_j^n - \frac{\Delta s}{\Delta \xi}(g_{j+1/2}^n - g_{j-1/2}^n), \quad j \in \mathbb{Z}, n \geq 0. \tag{2.47}$$

From [60], the numerical flux for the Godunov scheme is given by

$$g_{j+1/2}^n = \begin{cases} I(w_{j+1}^n, \bar{\xi}), & \text{if } h(w_j^n, \bar{\xi}) + h(w_{j+1}^n, \bar{\xi}) \leq 0 \text{ and } h(w_{j+1}^n, \bar{\xi}) \leq 0, \\ I(w_j^n, \bar{\xi}), & \text{if } h(w_j^n, \bar{\xi}) + h(w_{j+1}^n, \bar{\xi}) > 0 \text{ and } h(w_j^n, \bar{\xi}) > 0, \\ -3\bar{\xi}^2/8, & \text{if } h(w_j^n, \bar{\xi}) < 0 \text{ and } h(w_{j+1}^n, \bar{\xi}) > 0, \end{cases} \tag{2.48}$$

where $\bar{\xi} = \xi_{j+1/2}$, h is the wave speed

$$h(w, \xi) = w - \xi/2$$

and $I(w, \xi)$ is as follows

$$I(w, \xi) = \frac{1}{2} w^2 (e^{\Delta s} - 1) - \xi w (e^{\Delta s/2} - 1). \tag{2.49}$$

For the Lax-Friedrichs scheme we take:

$$g_{j+1/2}^n = \frac{(w_j^n)^2 - \bar{\xi} w_j^n + (w_{j+1}^n)^2 - \bar{\xi} w_{j+1}^n}{4} - \frac{\Delta \xi}{\Delta s}\left(\frac{w_{j+1}^n - w_j^n}{2}\right), \tag{2.50}$$

while for Engquist-Osher we choose:

$$\begin{aligned} g_{j+1/2}^n = & \frac{(w_j^n - \bar{\xi}/2)(w_j^n - \bar{\xi}/2 + |w_j^n - \bar{\xi}/2|)}{4} \\ & + \frac{(w_{j+1}^n - \bar{\xi}/2)(w_{j+1}^n - \bar{\xi}/2 - |w_{j+1}^n - \bar{\xi}/2|)}{4} - \frac{\bar{\xi}^2}{8}. \end{aligned} \tag{2.51}$$

The advantage of using numerical schemes in these similarity variables is that we do not need to cover large domains, neither in time nor space, to capture the dynamics of solutions. In Figure 2.4 we transform a rectangular mesh on the space-time domain $[-3, 3] \times [0, 4]$ for (ξ, s) into the corresponding parabolic mesh on (x, t). We can observe that computations done for equation (2.45) in short periods of time (for instance, up to $s = 4$) are equivalent to large-time solutions in the original equation (2.2) ($t \approx 53$ in the example under consideration).

2.5.2 Discussion on discrete steady states

For the numerical approximation schemes above in the similarity variables we expect a similar behavior as in the continuous case. Namely, that numerical solutions as the

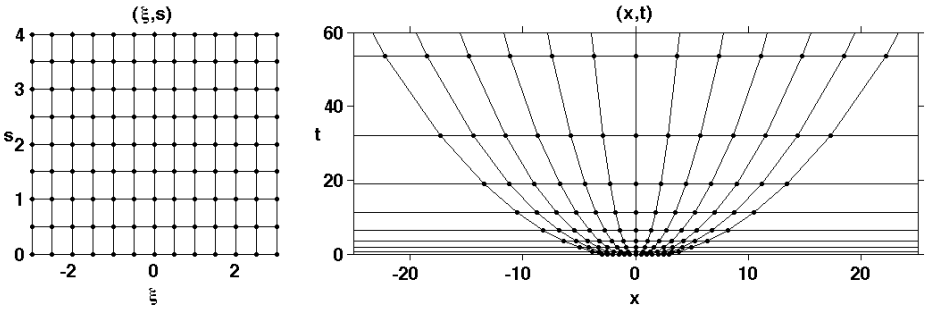


FIGURE 2.4: Comparison between the mesh on variables (ξ, s) and (x, t). The rectangular mesh $[-3, 3] \times [0, 4]$ for (ξ, s) covers a trapezoidal domain that reaches $x \in [-22, 22]$ at $t \approx 53$.

discrete self-similar time evolves converge towards numerical steady state solutions. Of course we expect this steady state solutions to converge towards the corresponding continuous ones as the mesh-size tends to zero. A complete analysis of these issues is out of the scope of this chapter. As we shall see, the numerical experiments confirm this fact establishing once more a clear distinction between the Lax-Friedrichs scheme that behaves in a parabolic manner and the two others.

To better understand the nature of the steady-state solutions, observe that those of (2.45) satisfy:

$$\Big(\frac{w^2 - \xi w}{2}\Big)_\xi = 0, \quad \xi \in \mathbb{R}, s > 0. \tag{2.52}$$

Since the solution must vanish on the tails, we deduce that

$$\frac{1}{2}w^2 - \frac{1}{2}\xi w = 0. \tag{2.53}$$

Thus, either $w = 0$ or $w = \xi$. Whether to choose one or the other is decided by using entropy conditions and the conservation of p and q, as in the case of equation (2.2) (cf. [60, 73]). The obtained profiles are, precisely, those given by (2.46).

On the other hand, the steady-state solution for the viscous version (with some viscosity $\nu > 0$) satisfies

$$-\nu w_\xi + \frac{w^2 - \xi w}{2} = 0, \quad \xi \in \mathbb{R}, s > 0 \tag{2.54}$$

which is no longer an algebraic equation, but an ODE.

Similarly, a steady-state solution $\bar{w} = \{\bar{w}_j\}_{j\in\mathbb{Z}}$ for (2.47), if it exists, must satisfy that

$$g(\bar{w}_j, \bar{w}_{j+1}, \bar{\xi}_{j+1/2}) = 0 \quad \forall j \in \mathbb{Z}.$$

In the case of the Godunov scheme (2.48), we can formally deduce from (2.49) that the asymptotic profile can only take values

$$\bar{w}_j = 0 \quad \text{or} \quad \bar{w}_j = \frac{2}{e^{\Delta s/2}+1}\bar{\xi} \quad \text{or} \quad \bar{w}_{j+1} = \frac{2}{e^{\Delta s/2}+1}\bar{\xi},$$

that is, $\bar{w}$ can just be 0, linear or a combination of both. Let us observe that the slope of the latter is not the same as the one of the continuous model, but tends to it when $\Delta s \to 0$. Note that this is compatible with a closed form of the numerical flux of Godunov where the inequalities in (2.49) are not strict.

The nature of the steady-state solution of the Engquist-Osher scheme (2.51) is slightly different. We have:

$$\frac{1}{2}\Big(\frac{(\bar{w}_j - \bar{\xi}/2)(\bar{w}_j - \bar{\xi}/2 + |\bar{w}_j - \bar{\xi}/2|)}{2} + \frac{(\bar{w}_{j+1} - \bar{\xi}/2)(\bar{w}_{j+1} - \bar{\xi}/2 - |\bar{w}_{j+1} - \bar{\xi}/2|)}{2}\Big) - \frac{\bar{\xi}^2}{8} = 0.$$

This is an upwind discretization of equation (2.53), rewritten as follows:

$$\frac{(w-\xi/2)^2}{2} - \frac{\xi^2}{8} = 0, \quad \xi \in \mathbb{R}, s > 0.$$

Therefore, the expected large-time behavior of the numerical simulation is, again, similar to the one of the continuous equation.

On the contrary, for the Lax-Friedrichs scheme (2.50) we have that:

$$-\frac{\Delta\xi}{\Delta s}\left(\frac{\bar{w}_{j+1} - \bar{w}_j}{2}\right) + \frac{1}{2}\left(\frac{(\bar{w}_j)^2 - \bar{\xi}\bar{w}_j}{2} + \frac{(\bar{w}_{j+1})^2 - \bar{\xi}\bar{w}_{j+1}}{2}\right) = 0.$$

We can distinguish two terms: the second is an average of the flux, but the first term corresponds to an artificial viscosity, as in (2.54). Thus, we can expect diffusivity at large times that distort the asymptotic N-wave.

2.5.3 Numerical example

In the following example we compare the behavior of the numerical solutions directly with the asymptotic profile of the continuous solution of (2.45). Let us choose initial data

$$w_0(x) = \begin{cases} x+10, & -12 < x < -8, \\ x, & -\sqrt{2} < x < \sqrt{6}, , \\ 0, & \text{elsewhere}, \end{cases} \tag{2.55}$$

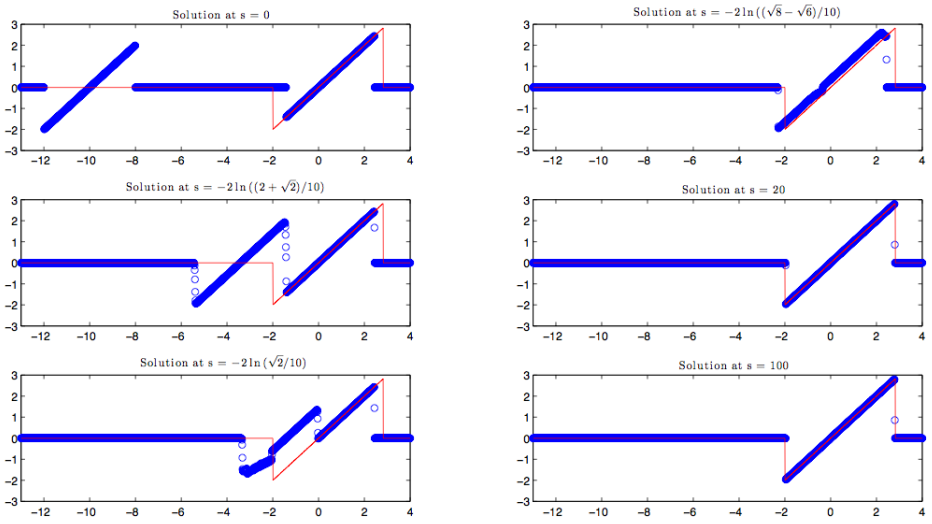


FIGURE 2.5: Convergence of the numerical solution of (2.45) using Godunov scheme (circle dots) to the asymptotic N-wave (solid line). We take $\Delta\xi = 0.01$ and $\Delta s = 0.0005$.

which corresponds to two separate N-waves. From the continuous point of view, at the beginning the first one moves towards the origin until it collides with the other one. Then they both interact, resulting on a new N-wave which is similar to the expected asymptotic profile. The same behavior should be required for the numerical schemes, but, as we show in Figures 2.5, 2.6 and 2.7, the performance may vary depending on the chosen numerical flux.

We consider the mesh size $\Delta\xi = 0.01$ and a time step $\Delta s = 0.0005$, which is small enough to satisfy the CFL condition. Let us recall that, since the support of the solution remains in a bounded interval, we can choose a small spatial domain. In the first part of the simulation, the three numerical schemes behave in the correct manner, as the two N-waves collapse into one. The first column in the figures shows this regime.

Once the unique N-wave takes form, the behavior of the schemes takes different paths. Both Godunov and Engquist-Osher schemes maintain the N-wave shape that gradually converges to the hyperbolic asymptotic profile, as we can observe in the second column of Figures 2.5 and 2.6.

Meanwhile, the artificial viscosity of the Lax-Friedrichs scheme starts becoming dominant, making the solution evolve to the parabolic diffusion wave, which is the steady state of the viscous version of (2.45) (cf. [60]).

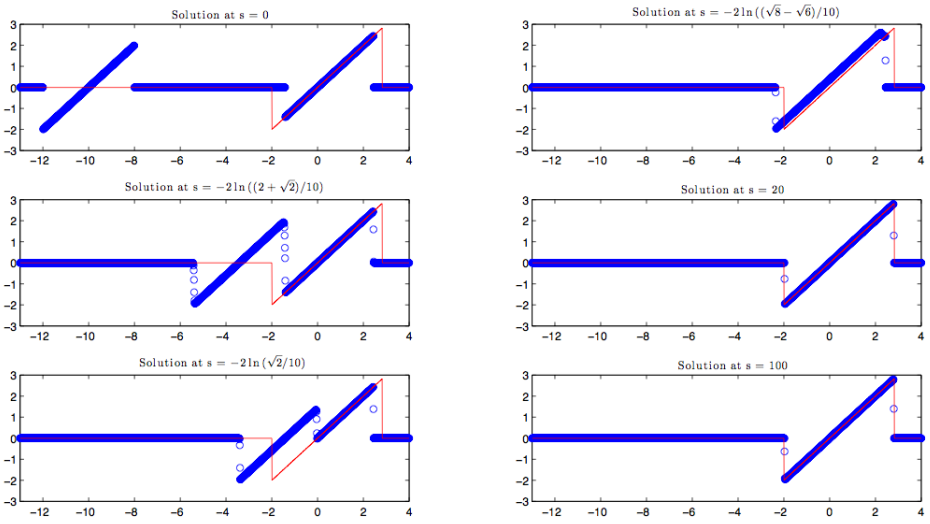


FIGURE 2.6: Convergence of the numerical solution of (2.45) using Engquist-Osher scheme (circle dots) to the asymptotic N-wave (solid line). We take $\Delta\xi = 0.01$ and $\Delta s = 0.0005$.

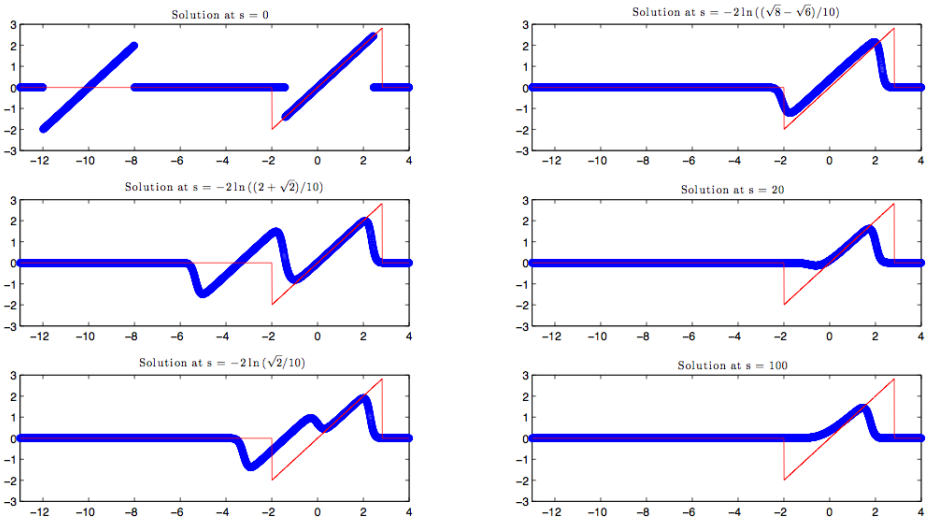


FIGURE 2.7: Numerical solution of (2.45) using the Lax-Friedrichs scheme (circle dots), taking $\Delta\xi = 0.01$ and $\Delta s = 0.0005$. The N-wave (solid line) is not reached, as it converges to the diffusion wave.

2.5.4 Computational benefits

It is important to emphasize the benefits of using similarity variables to perform long-time simulations. For instance, let us consider the following initial data:

$$u_0(x) = \begin{cases} 2, & 0 < x \leq 2, \\ -1, & -1 \leq x \leq 0, \\ 0, & \text{elsewhere.} \end{cases}$$

	Nodes	Time-steps	$\|u_\Delta - u\|_{1,\Delta}$	$\|u_\Delta - u\|_{2,\Delta}$	$\|u_\Delta - u\|_{\infty,\Delta}$
Phys. var.	501	2001	0.2140	0.1352	0.2745
Simil. var.	100	1306	0.2057	0.1136	0.2543

	Nodes	Time-steps	$\|u_\Delta - u\|_{1,\Delta}$	$\|u_\Delta - u\|_{2,\Delta}$	$\|u_\Delta - u\|_{\infty,\Delta}$
Phys. var.	5001	20001	0.0280	0.0517	0.2828
Simil. var.	750	9877	0.0276	0.0379	0.2465

TABLE 2.1: Comparison of solutions at $t = 100$. We take $\Delta x = 0.1$ (top) and $\Delta x = 0.01$ (bottom). We choose $\Delta\xi$ such that the $\|\cdot\|_{1,\Delta}$ error is similar. The time-steps are $\Delta t = \Delta x/2$ and $\Delta s = \Delta\xi/20$, respectively, enough to satisfy the CFL condition.

We compute the numerical approximation of the corresponding solution to (2.2) in two different ways, either in the original or in the self-similar variables, and compare them to the exact solution, which can be computed explicitly. First we consider scheme (2.10), based on physical variables, and then, scheme (2.47) using similarity variables. In both cases we choose the Engquist-Osher numerical flux, i.e., (2.15) and (2.51) respectively. Also, in the latter case we use piecewise constant interpolation to recover the solution in the physical space so that we can compare it to the exact one.

In Table 2.1 we compare the errors of the solutions at $t = 100$. Let us remark that, in order to avoid interferences of boundary conditions, we need to choose large enough spatial domains. We have chosen $[-20, 30]$ for the case of physical variables and $[-3, 4]$ for the other one. We consider $\Delta x = 0.1$ and $\Delta x = 0.01$ and $\Delta t = \Delta x/2$, accordingly to the CFL condition. Parameter $\Delta\xi$ is chosen such that the $\|\cdot\|_{1,\Delta}$ error made is similar to the corresponding case, while we take $\Delta s = \Delta\xi/20$. We make the same comparison in Table 2.2 for the solutions at $t = 1000$. In this case, we have taken $[-50, 100]$ and $[-3, 4]$ space intervals, respectively. The criteria for the mesh-size and time-step are the same as above.

We observe that to obtain similar accuracy, we need much less nodes and time iterations to compute the numerical solution at a given time. In fact, the results using similarity variables could be improved using a higher order reconstruction of the solution, instead of piecewise constant, when doing the change of variables to recover the physical solution.

	Nodes	Time-steps	$\|u_\Delta - u\|_{1,\Delta}$	$\|u_\Delta - u\|_{2,\Delta}$	$\|u_\Delta - u\|_{\infty,\Delta}$
Phys. var.	1501	19987	0.0867	0.0482	0.0893
Simil. var.	215	4225	0.0897	0.0332	0.0367

	Nodes	Time-steps	$\|u_\Delta - u\|_{1,\Delta}$	$\|u_\Delta - u\|_{2,\Delta}$	$\|u_\Delta - u\|_{\infty,\Delta}$
Phys. var.	15001	199867	0.0093	0.0118	0.0816
Simil. var.	2000	39459	0.0094	0.0106	0.0233

TABLE 2.2: Comparison of solutions at $t = 1000$. We take $\Delta x = 0.1$ (top) and $\Delta x = 0.01$ (bottom). We choose $\Delta\xi$ such that the $\|\cdot\|_{1,\Delta}$ error is similar. The time-steps are $\Delta t = \Delta x/2$ and $\Delta s = \Delta\xi/20$, respectively, enough to satisfy the CFL condition.

2.6 Generalizations and further comments

Until now we have considered just the Burgers equation and Lax-Friedrichs, Engquist-Osher and Godunov schemes, but the described techniques can be extended to more general types of fluxes and numerical schemes.

Regarding the latter, the main requirement is that the chosen scheme must verify the OSLC, so that the decay estimates can be used to guarantee the compactness of the rescaled solutions. The homogeneity of the dissipation, as defined at the end of Section 2.2, will indicate if the introduced artificial viscosity is strong enough to modify the asymptotic behavior of the numerical solution or if it preserves the continuous property.

As for the fluxes, it is worth mentioning that the N-wave appearing as the asymptotic profile is a common characteristic for all 1-D scalar conservation laws with uniformly convex flux, that is, for those with $f''(u) \geq \gamma > 0$, with $\gamma > 0$. For that reason, one expects to observe the same phenomena in the discrete level as the ones described in this chapter. As we said before, the OSLC will play a key role. Nevertheless, in this more general situation, obtaining the homogeneity α of the dissipation might be not so straightforward. The coefficient R of the Lax-Friedrichs scheme does not depend on f, so it will not develop the N-wave regardless the flux we consider. The analysis of Engquist-Osher and Godunov schemes is more delicate, since their coefficients of viscosity are strictly related to the flux. In any case, whenever $\alpha \geq 1$, the asymptotic profile will be the desired N-wave. The conclusion is the same for any uniformly concave flux, just by considering the reflected N-wave.

The analysis is also valid for some type of odd fluxes, those that are concave on one side of their axis of symmetry and convex in the other. Nevertheless, there will be

no difference in the asymptotic profile anymore. For instance, let us consider equation (2.1) with flux $f(u) = |u|u/2$. Then, the asymptotic behavior is the one stated by the following theorem.

Theorem 2.12. *Let $u_0 \in L^1(\mathbb{R})$ and choose mesh-size parameters Δx and Δt satisfying the CFL condition $\lambda \|u^n\|_{\infty,\Delta} \leq 1$, $\lambda = \Delta t/\Delta x$. Let u_Δ be the corresponding solution of the discrete scheme (1.8) for the hyperbolic conservation law (2.1) with flux $f(u) = |u|u/2$. Then, for any $p \in [1, \infty)$, it holds that*

$$\lim_{t\to\infty} t^{\frac{1}{2}(1-\frac{1}{p})} \|u_\Delta(t) - w(t)\|_p = 0, \tag{2.56}$$

where the profile w is as follows:

1. *for the Lax-Friedrichs scheme, $w = w_{M_\Delta}$ defined in (2.6).*
2. *for Engquist-Osher and Godunov schemes, $w = w_{0,M_\Delta}$ if $M_\Delta > 0$ or $w = w_{M_\Delta,0}$ if $M_\Delta < 0$, both given by (2.7).*

The proof of this result is analogous to the one of Theorem 2.1. In this case, the key point now is the uniqueness of solution (see [73]) of the equation

$$\begin{cases} u_t^\infty + \left(\dfrac{u^\infty|u^\infty|}{2}\right)_x = \dfrac{\Delta x^2}{2\Delta t} u_{xx}^\infty, \\ u(0) = M\delta_0, \end{cases}$$

which is the one appearing for Lax-Friedrichs, and of the equation

$$\begin{cases} u_t^\infty + \left(\dfrac{u^\infty|u^\infty|}{2}\right)_x = 0, \\ u(0) = M\delta_0, \end{cases}$$

corresponding to Engquist-Osher and Godunov. Moreover, both converge to the same continuous N-wave, which has unique sign, when $\Delta x, \Delta t \to 0$.

Chapter 3

Optimal control in large-time horizons

3.1 Introduction

In this chapter, we show the influence of the results of Chapter 2 in the context of numerical optimization. In particular, we analyze the numerical approximation of the inverse design problem for the Burgers equation, both in the viscous and in the inviscid case.

Given a time $T > 0$ and a target function u^* the aim is to identify the initial datum u_0 so that the solution, at time $t = T$, reaches the target u^* or gets as close as possible to it. We formulate the problem from the point of view of optimal control. Using a least square approach, we consider the following minimization problem:

$$\min_{u^0 \in \mathcal{U}_{ad}} \mathcal{J}(u_0) := \min_{u^0 \in \mathcal{U}_{ad}} \frac{1}{2} \int_{\mathbb{R}} \left(u(x,T) - u^*(x)\right)^2 dx, \tag{3.1}$$

subject to

$$\begin{cases} \partial_t u + \partial_x \left(\frac{u^2}{2}\right) = \nu \partial_{xx} u, & x \in \mathbb{R},\, t > 0, \\ u(x,0) = u_0(x), & x \in \mathbb{R}. \end{cases} \tag{3.2}$$

where $\mathcal{U}_{ad}$ a suitable functional class, for instance $L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$. As we shall see, for a large final time T, schemes that do not yield the correct long-time dynamics are incapable of providing an accurate approximation of minimizer. Thus, the choice of the numerical scheme approximating the PDE becomes a critical issue.

In this chapter we shall focus on this particular Burgers model. But similar issues arise in other contexts, involving different PDE, in many applications in physical sciences, engineering, economics and management disciplines. This is particularly the

case in the context of data assimilation [32] with applications in climate forecasting and hydrology modeling, or in identification of pollution source problems [68].

We consider both the viscous, $\nu > 0$, and the inviscid case, $\nu = 0$. The first one leads to a non-linear parabolic problem while the second one constitutes a nonlinear hyperbolic conservation law. The solutions to the inviscid Burgers equation ($\nu = 0$) may develop shocks and, if $u_0 \in L^1(\mathbb{R})$, as time tends to infinity, they converge towards a self-similar N-wave (c.f. [73]). Even if the scaling is the same, this behavior differs significantly from the viscous version (3.2), which is of parabolic nature (see [45]). The self-similar profiles are then diffusive, smooth and of constant sign. Nevertheless, when ν is sufficiently small and time is large (but not enough for the viscosity to be dominant), the behavior of the solutions is close to the hyperbolic case [60].

At the numerical level, when solving (3.2) with usual finite-difference conservative schemes, this asymptotic dichotomy for large time needs to be handled carefully since, in particular, the excess of numerical viscosity may destroy the long-time hyperbolic dynamics to make it parabolic. This issue was carefully analyzed in Chapter 2, where it was shown that, while the large-time behavior of the inviscid dynamics is correctly captured by means of the Enguist-Osher scheme, the Lax-Friedrichs scheme fails to do so, due to the excess of viscosity that destroys the N-wave dynamics and leads to a viscous profile. This pathology may arise also for the viscous Burgers equation when the numerical viscosity dominates the physical one.

In this work, we emphasize the consequences of this fact at the level of inverse design. This is done mainly by means of a gradient descent method, using the adjoint system methodology. Our numerical results are also tested and compared with those obtained with the optimization package IPOPT, an open-source software package for nonlinear optimization [92], to confirm our predictions and conclusions. Note, however, that the large-time behavior dichotomy can be extended to other methods.

The functional $\mathcal{J}$ under consideration, due to the quadratic nonlinearity involved in the Burgers equation, fails to be convex or, at least, there is no evidence of its convexity. Consequently, the existence of several critical points cannot be excluded, in principle. However, in view of the recent literature based on Łojaciewicz's inequality, which ensures convergence towards critical points without requiring the convexity of the functional (see, for instance, [1, 5, 42, 75]), one expects the convergence of gradient descent methods towards local minimizers. This is something we observe in our numerical simulations, but the velocity of convergence turns out to depend significantly on the numerical scheme that is employed for approximating the PDE and the descent method implemented.

In any case, gradient descent methods, such as the steepest descent method or conjugate gradient method, are attractive due to their algorithm simplicity. Descent methods do not require calculating second order derivatives. This makes them be a suitable approach for optimizing large scale problems, where the cost of computing the Hessian matrix and solving the corresponding linear system becomes prohibitive. The rate of convergence in gradient methods is related to the condition number. In any case, in this chapter we emphasize that this velocity also depends on the numerical scheme chosen to discretize the PDE.

As we shall see, while the direct implementation of a discrete approach based on the Enguist-Osher discretization leads to rather robust and satisfactory results at the level of the inverse design, the results turn out to be very sensitive to the discretization parameters when employing the Lax-Friedrichs scheme and its modified version coping with viscosity. With the latter, in some instances, the recovered inverse profile appears to be corrupted by unwanted spurious numerical high frequency components.

This chapter is organized as follows. In Section 3.2 we discuss the existence of minimizers in the continuous setting, distinguishing the viscous and inviscid case. In Section 3.3 we present the numerical tools that we use for optimization and discretization processes. Then, in Section 3.4 we solve an optimal control problem in a large period of time for Burgers equation with small viscosity using the gradient descent method and IPOPT. Section 3.5 revisits the experiments done in Section 3.4 for the case of inviscid Burgers. In Section 3.6, we modify some assumptions on the initial setting of the numerical experiments performed in previous sections to test. In Section 3.7, we visualize functional landscapes and discuss the behavior of the minimizers further. We conclude with some final remarks in Section 3.3.

3.2 The continuous inverse design problem

In this section we briefly recall some of the main features of the continuous inverse design or optimal control problem (3.1)-(3.2) distinguishing the viscous, $\nu > 0$, and the inviscid case, $\nu = 0$.

3.2.1 Viscous Burgers

In the case $\nu > 0$, (3.2) is a parabolic equation that is very well behaved. In particular, given the initial datum in $L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ (actually, $L^1(\mathbb{R})$ would suffice), equation (3.2) has a unique solution within the class $C([0,\infty); L^1(\mathbb{R})) \cap L^\infty(\mathbb{R} \times (0,\infty))$ that is smooth for $t > 0$.

The optimization problem (3.1)-(3.2) was addressed in [18] with the aim of adapting the alternating descent method introduced in [17] in the inviscid case.

In [18] it was observed that the minimization problem (3.1)-(3.2) has always a solution provided the minimization is performed within a class of a priori bounded initial data, i.e., under an additional constraint of the form

$$||u_0||_{L^1(\mathbb{R})\cap L^\infty(\mathbb{R})} \leq K. \tag{3.3}$$

Under such a restriction, the proof of the existence of the minimizer follows by the classical Direct Method of the Calculus of Variations.

As far as we know, the existence of the minimizer cannot be guaranteed in the absence of constraints on the initial data, since minimizing sequences could be unbounded. As we shall see, this issue is very closely related to the one-sided Lipschitz condition that the solutions of the viscous Burgers equation satisfy and that imposes a universal threshold within the range of the semigroup at time $t = T$.

Note however that, due to the parabolic nature of the equation and classical results of backward uniqueness [70], when the target u^* is exactly reachable, the initial datum u_0 leading to it (the minimizer of the functional $\mathcal{J}$) is unique, something that fails to be true in the inviscid case [17].

3.2.2 Inviscid Burgers

In the case $\nu = 0$, (3.2) is a nonlinear hyperbolic scalar conservation law. When the initial datum is in $L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ (actually, $L^1(\mathbb{R})$ would suffice), equation (3.2) has a unique entropy solution within the class $C([0,\infty); L^1(\mathbb{R})) \cap L^\infty(\mathbb{R} \times (0,\infty))$. But in the present case, contrarily to the viscous one, smoothness cannot be guaranteed for $t > 0$ since solutions may develop shock discontinuities.

In this hyperbolic context, the optimization problem (3.1)-(3.2) was addressed in [17] with the aim of developing specific numerical techniques to treat the cases where optimal solutions develop shock discontinuities. Proceeding as in [17], one can prove that also in the inviscid case the optimal control problem (3.1)-(3.2) has at least a minimizer, provided the initial data are restricted by a constraint of the form (3.3).

Note that one of the main differences between the viscous and the inviscid case is that, for the latter, backward uniqueness does not hold and, consequently, the minimizer is not unique in general.

3.3 Description of the numerical algorithms

In this section, we describe the discrete techniques we employ for solving the optimal control problem (3.1)-(3.2). First, we discuss two numerical discretization methods for the nonlinear PDE, namely the Engquist-Osher (EO) and the modified Lax-Friedrichs (MLF) schemes. We also briefly introduce the discrete version of the inverse design problem.

As we shall see, the wrong long time dynamics introduced by the MLF method or the correct one that EO is capable of mimicking, will be the ultimate reason for the overall performance of the optimization methods, and in particular for the spurious results that one observes in the experiments of Section 3.4 and 3.5 when employing MLF.

3.3.1 Discretization schemes and large-time behavior

To implement the methods numerically, we opt for a discretization of (3.2) using classical conservative schemes (c.f. [38]). As we shall see in the experiments, the choice of the numerical flux is critical when dealing with optimal control problems in large periods of time. Indeed, as pointed out in Chapter 2, numerical schemes that add too much viscosity into the system may lead to solutions with the wrong asymptotic behavior.

Let us denote spatial nodes $x_{j+1/2} = \Delta x(j+1/2)$, $j \in \mathbb{Z}$, and time instants $t_n = n\Delta t$, $n \in \mathbb{N} \cup \{0\}$, where $\Delta x, \Delta t > 0$ are the mesh size and time-step respectively. We approximate the solution u of (3.2) by a piecewise constant function u_Δ such that

$$u_\Delta(x,t) = u_j^n, \quad x \in [x_{j-1/2}, x_{j+1/2}),\ t \in [t_n, t_{n+1}),$$

where

$$\begin{cases} u_j^{n+1} = u_j^n - \dfrac{\Delta t}{\Delta x}(g_{j+1/2}^n - g_{j-1/2}^n) \\ \qquad\qquad + \dfrac{\nu \Delta t}{\Delta x^2}(u_{j-1}^n - 2u_j^n + u_{j+1}^n),\ j \in \mathbb{Z}, n = 0, \dots, N, \\ u_j^0 = \dfrac{1}{\Delta x}\displaystyle\int_{x_{j-1/2}}^{x_{j+1/2}} u_0(x)dx,\ j \in \mathbb{Z}. \end{cases} \tag{3.4}$$

Here, $N = \lceil T/\Delta t \rceil$ is the number of time-steps needed to reach T. We denote $g_{j+1/2}^n = g(u_j^n, u_{j+1}^n)$, where the function g is the numerical flux. In this chapter we compare two fluxes:

$$\text{Engquist-Osher (EO): } g^{EO}(v,w) = \frac{v(v+|v|)}{4} + \frac{w(w-|w|)}{4}, \tag{3.5}$$

$$\text{Modified Lax-Friedrichs (MLF): } g^{MLF}(v,w) = \frac{v^2+w^2}{4} - \frac{\Delta x}{\Delta t}\left(\frac{w-v}{4}\right). \tag{3.6}$$

These schemes are convergent in the classical sense of the numerical analysis (see, for instance, [38]), being of first order of accuracy. Nevertheless, in accordance with the classification in Chapter 2, this is not enough when dealing with large-time evolution problems. Given $\Delta x, \Delta t > 0$, convergence in the sense of asymptotic dynamics needs to be taken into account too. The choice of the numerical fluxes has been made according to the parabolic-hyperbolic dichotomy shown there.

Remark 3.1*: In the continuous case (see Figure 1.2), the changing sign N-waves are the asymptotic profiles as $t \to \infty$ if $\nu = 0$ [73] and intermediate metastable states if $\nu > 0$ [60]. In the hyperbolic regime, the key point in the identification of the asymptotic N-wave, which belongs to a two-parameter family, is the preservation of the quantities* (1.5)*. In the parabolic case, there is a period of time when the dynamics are close to the hyperbolic ones. The larger* $\min(p,q)/\nu$ *is, the longer the diffusion needs to become dominant. Therefore, this needs to be taken into account at the discrete level, since the numerical viscosity can interfere with the correct dynamics of the model.*

In our simulations, we have implemented modified Lax-Friedrichs scheme instead of the usual Lax-Friedrichs scheme. It can be shown that Lax-Friedrichs scheme has the maximum amount of numerical viscosity that is allowed by linear stability theory (e.g. [18]) and the scheme becomes unstable in the presence of any physical viscosity ($\nu > 0$). This is not the case for modified Lax-Friedrichs scheme. Nevertheless, both schemes have similar performance in terms of their large-time behavior.

It is well known (e.g. [38]) that (3.4) can be rewritten in its viscous form in the following manner:

$$\frac{u_j^{n+1}-u_j^n}{\Delta t} + \frac{(u_{j+1}^n)^2-(u_{j-1}^n)^2}{4\Delta x} = R(u_j^n,u_{j+1}^n) - R(u_{j-1}^n,u_j^n) + \frac{\nu\Delta t}{\Delta x^2}(u_{j-1}^n - 2u_j^n + u_{j+1}^n), \tag{3.7}$$

where R is uniquely defined by

$$R(u,v) = \frac{1}{2\Delta x}\Big(\frac{u^2}{2} + \frac{v^2}{2} - 2g(u,v)\Big). \tag{3.8}$$

In the case of the numerical fluxes that we consider in this chapter, we have:

$$R^{MLF}(u,v) = \frac{v-u}{4\Delta t},$$
$$R^{EO}(u,v) = \frac{1}{4\Delta x}(v|v| - u|u|).$$

In Chapter 2 we show that the large-time behavior of u_Δ depends on the degree of homogeneity of the term $R(u, v)$. In other words, let us assume that there exists $\alpha \in \mathbb{R}$ and $\mu > 0$ such that

$$R(\mu u, \mu v) = \mu^\alpha R(u, v).$$

It is clear that $\alpha = 2$ for Engquist-Osher and $\alpha = 1$ for modified Lax-Friedrichs. In the first case, the numerical scheme introduces quadratic numerical viscosity. Thus, it decays to zero as the L^∞-norm of the solution, vanishing in time and preserving the continuous dynamics. On the other hand, the numerical viscosity inherent in MLF is of the order of $\Delta x^2/\Delta t$. This means that in the case of $\nu \ll \Delta x^2/\Delta t$, the system is driven into a diffusive wave too early, and consequently continuous metastable states are not reproduced numerically.

Figure 3.1 shows this pathology of the MLF in the viscous case. It compares the solution to (3.2) with the same initial data at different times, using EO and MLF. Both numerical solutions are similar at early stages. Nevertheless, the one by MLF starts losing the N-wave shape earlier than EO and achieves the diffusive profile in shorter time.

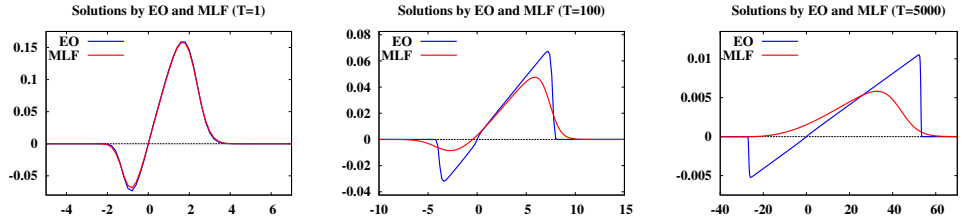


FIGURE 3.1: Solutions of (3.2) with $\nu = 10^{-6}$ at $t = 1$, $t = 100$ and $t = 5000$, using (3.4) with $\Delta x = 0.2$, $\Delta t = 0.5$ and numerical fluxes EO (blue) and MLF (red).

3.3.2 The discrete optimization problem

There are two main ways of addressing optimal inverse design problems like the one we present in this work. On the one hand, in the continuous approach, one first develops the optimization method at the continuous level and then discretizes it. On the other hand, the discrete approach consists in discretizing every element involved in the problem and then optimizing at the discrete level. See, for instance, [34] and the references therein for an introduction on both approaches in the context of design in Computational Fluid Dynamics.

In the simulations done in this work, we follow the discrete approach to optimization. The discretization of (3.2) is done according to (3.4). Regarding the discrete version of

(3.1), we consider a simple quadrature rule:

$$\mathcal{J}_\Delta(u_\Delta^0) = \frac{\Delta x}{2} \sum_{\mathbb{Z}} (u_j^N - u_j^*)^2, \tag{3.9}$$

where the target function u^* has been discretized in the same manner as the initial data u_0 in (3.4).

The class of admissible data for the discrete minimization problem (3.4)-(3.9) is an approximation of the continuous one, constituted by piecewise constant functions satisfying (3.3). In the inviscid case, no further assumptions are required by virtue of the One-Sided Lipschitz condition (OSLC) satisfied by EO and MLF numerical fluxes [12]. To the best of our knowledge, this condition is no longer available for the viscous case and, thus, an additional restriction is required; namely,

$$\|u_0\|_{L^1(\mathbb{R})\cap L^\infty(\mathbb{R})} + TV(u_0) \leq K. \tag{3.10}$$

Within that class of functions, the existence of minimizers for (3.4)-(3.9) is obvious since we are dealing with a finite-dimensional problem. The convergence of the discrete minimizers towards the continuous one is less trivial. For the inviscid Burgers equation this fact was already proved in [17], where the authors used the OSLC and a Γ-convergence argument. Regarding the viscous case, the same procedure is valid, but using BV estimates to work around the lack of OSLC.

Remark 3.2*: It is worth emphasizing that in [17] no distinction is made between EO and MLF (or any other numerical scheme, as long as it is monotone and satisfies the OSLC) in terms of their numerical viscosity. In fact, the convergence towards the continuous minimizer is true in both cases. However, as we show in this chapter, there is a significant difference in the rate of convergence, at least, in large-time problems.*

3.3.3 Optimization techniques

Finally, we present two numerical techniques for solving (3.4)-(3.9). First we recall the classical gradient descent method based on the adjoint methodology. Afterwards, we briefly describe some features of the open source software IPOPT, which uses interior point optimization.

For the sake of clarity, this section is developed without taking into account restriction (3.10). In fact, this condition is not restrictive in our numerical experiments since the minimizing sequences we obtain do not activate it. We refer to [19, 77] for a review

about classical methods for constrained problems, such as projected-gradient methods and penalty-function methods.

3.3.3.1 Gradient descent by adjoint methodology

The gradient descent method is a widely spread technique to compute minimizers. This iterative process consists in taking steps in the direction of the gradient of the functional with respect to the controls. Roughly speaking, to minimize a functional $\mathcal{J}$, starting from an initial guess z_0, one iteratively computes

$$z_{k+1} = z_k - \varepsilon_k \nabla \mathcal{J}(z_k), \quad k \in \mathbb{N},$$

where $\varepsilon_k > 0$ is the size of the step. Since we are taking the gradient in a negative direction, $\mathcal{J}(z_{k+1}) < \mathcal{J}(z_k)$ for ε_k small enough, except for the case in which one reaches a minimum in a finite number of iterations: $\nabla \mathcal{J}(z_k) = 0$. Observe that, in general, as the minimum is approached, the gradient tends to zero and, hence, the sequence is expected to converge to a (possibly local) minimizer. The convergence of the gradient descent method is well understood when dealing with quadratic coercive functionals and, more generally, for C^2 functionals exhibiting strict convexity conditions (see [19]). More recently convergence results have been proved for a much larger class of functionals satisfying the so-called Łojaciewicz inequality (see, for instance, [1]).

In our case, the implementation of gradient methods requires a sensitivity analysis of $\mathcal{J}$ with respect to the initial data u_0. This can be developed in a very classical way for the viscous Burgers equation, because of its parabolic nature and the regularity of solutions for $t > 0$. But this is a much more subtle issue in the hyperbolic inviscid frame.

Let us consider perturbations of the form

$$u_0^\varepsilon(x) = u_0(x) + \varepsilon \delta u_0(x),$$

with $\delta u_0 \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ and $\varepsilon > 0$. Then, the solution u^ε of (3.2) with initial data u_0^ε is a classical solution too. Moreover, it satisfies

$$u^\varepsilon = u + \varepsilon \delta u + o(\varepsilon),$$

in a suitable functional setting and, in particular, in $C([0,\infty); L^1(\mathbb{R})) \cap L^\infty(\mathbb{R} \times (0,\infty))$ and in classes of smooth solutions for $t > 0$. Here, of course, δu stands for the solution of the linearized system:

$$\begin{cases} \partial_t \delta u + \partial_x \left(u \delta u\right) = \nu \partial_{xx} \delta u, & x \in \mathbb{R},\, t > 0, \\ \\ \delta u(x,0) = \delta u_0(x), & x \in \mathbb{R}. \end{cases} \tag{3.11}$$

Thus, the Gateaux derivative of $\mathcal{J}$ at u_0 in the direction δu_0 is

$$\delta \mathcal{J}[\delta u_0](u_0) = \int_{\mathbb{R}} \big(u(x,T) - u^*(x)\big) \delta u(x,T) dx. \tag{3.12}$$

Now, to simplify this representation of $\delta \mathcal{J}$, we make use of the adjoint methodology. The corresponding adjoint system to (3.2) is

$$\begin{cases} -\partial_t \rho - u \partial_x \rho = \nu \partial_{xx} \rho, & x \in \mathbb{R},\, t > 0, \\ \\ \rho(x,T) = u(x,T) - u^*(x), & x \in \mathbb{R}, \end{cases} \tag{3.13}$$

so that

$$\delta \mathcal{J}[\delta u_0](u_0) = \int_{\mathbb{R}} \rho(x,0) \delta u^0(x) dx.$$

Therefore, once we have computed the adjoint initial state, this expression provides us an easy way to compute a descending direction for the continuous functional $\mathcal{J}$:

$$\delta u_0(x) = -\rho(x,0), \quad x \in \mathbb{R}.$$

The new perturbed initial data that the gradient descent method yields will be given by

$$u_0^\varepsilon(x) = u_0(x) - \varepsilon \rho(x,0).$$

The same ideas apply in the inviscid case too, but the issue becomes much more technical because of the lack of regularity of solutions. This leads to the necessity of developing suitable notions of solutions for the linearized systems (see [9, 14, 17]).

Following the discrete approach, we now reproduce the same procedure at the discrete level. The discretization of (3.1)-(3.2) is done according to (3.4)-(3.9). Now, it is easy to obtain the corresponding discrete adjoint system for (3.4):

$$\begin{cases} \rho_j^n = \rho_j^{n+1} + \frac{\Delta t}{\Delta x} \Big(\partial_v g(u_j^n, u_{j+1}^n) \left(\rho_{j+1}^{n+1} - \rho_j^{n+1} \right) \\ \qquad\qquad + \partial_w g(u_{j-1}^n, u_j^n) \left(\rho_j^{n+1} - \rho_{j-1}^{n+1} \right) \Big) \\ \qquad + \frac{\nu \Delta t}{\Delta x^2} (\rho_{j-1}^{n+1} - 2\rho_j^{n+1} + \rho_{j+1}^{n+1}), & j \in \mathbb{Z}, n = N-1 \ldots, 0, \\ \\ \rho_j^N = u_j^N - u_j^*, & j \in \mathbb{Z}, \end{cases} \tag{3.14}$$

Remark 3.3*: The same analysis as in* (3.7) *can be done for the discrete adjoint system, reordering* (3.14) *as follows:*

$$-\frac{\rho_j^{n+1}-\rho_j^n}{\Delta t}-u_j^n\frac{\rho_{j+1}^{n+1}-\rho_{j-1}^{n+1}}{2\Delta x}=\bar{R}(\rho_j^{n+1},\rho_{j+1}^{n+1},u_j^n)-\bar{R}(\rho_{j-1}^{n+1},\rho_j^{n+1},u_j^n)\\+\frac{\nu(\rho_{j-1}^{n+1}-2\rho_j^{n+1}+\rho_{j+1}^{n+1})}{\Delta x^2},$$

where

$$\bar{R}^{MLF}(\rho,\sigma,u)=\frac{\sigma-\rho}{4\Delta t},$$
$$\bar{R}^{EO}(\rho,\sigma,u)=\frac{|u|}{2\Delta x}(\sigma-\rho).$$

Of course, both systems are linear in the adjoint variables. But, as in (3.7)*, the numerical viscosity for the MFL adjoint does not depend on the forward solution and, hence, it is kept constant in time. Thus, the large-time behavior of the discrete adjoint system is affected in the same way as in* (3.4)*. On the contrary, the numerical viscosity in EO is proportional to the numerical solution u_Δ, which decays towards zero as $t\to\infty$ and, hence, vanishes in time.*

To minimize (3.9), we will take the descent direction given by:

$$\delta u_j^0=-\rho_j^0,\quad j\in\mathbb{Z}. \tag{3.15}$$

This direction is straightforwardly obtained following the same arguments as for the continuous level. Thus, the new initial data $u_\Delta^{0,\varepsilon}$ will be given by

$$u_j^{0,\varepsilon}=u_j^0-\varepsilon\rho_j^0,\quad j\in\mathbb{Z} \tag{3.16}$$

for some $\varepsilon>0$ small enough.

In conclusion, each iteration for the gradient descent method is as follows. Given an initial guess u_Δ^0 to the minimizer, we solve (3.4). Then, we use the obtained solution u_Δ^N to compute the initial adjoint state ρ_Δ^0 from (3.14). Lastly, we update the approximation of the optimal solution using the descent direction (3.15).

Solving the adjoint system (3.14) requires knowing the solution of (3.4) at every time instant. For large-time problems, storing the time-history solution can be computationally prohibitive. In our experiments we did not use further strategies, but problems like the sonic-boom minimization could require additional efforts. For the detailed implementation of the adjoint methodology including memory-saving techniques, such as backtracking, see [37].

Moreover, one cannot always expect this iterative process to finish successfully (i.e, finding an exact minimizer) in a reasonable time. When this occurs, additional stopping criteria need to be set. For instance, one could stop the iterations by specifying:

- The smallest difference between two consecutive approximations of the optimal solution.
- The smallest value that the norm of the gradient can take.
- The smallest value that ε can take without descending.
- The absolute error between the approximation and the optimal solution, if this is known a priori.

Note that all of them are related to the proximity to a minimizer, either local or global. In our experiments we opt for the third one. We refer to [77] for an extended summary about these and other method-specific stopping criteria.

3.3.3.2 The interior-point optimization method with IPOPT

IPOPT has been developed for efficient optimization of large-scale nonlinear programing. It implements interior-point or barrier methods, which provide a more attractive alternative to active set strategies in optimizing problems with a large number of inequality constraints. IPOPT can converge to an optimal solution even from infeasible starting points through using a filter line-search method. The underlying concept of filter methods is to accept trial points along the line-search if they improve the objective function or improve the constraint violation instead of a combination of these two criteria defined by a combined merit function. The filter concept excludes the possibility of the solution to alternate between two iteration points, so that one improves the objective function while the other voids constraint violations (see [92]).

The optimal control problem discussed in (3.1)-(3.2) is formulated as a minimization problem with constraint equations in the form of Partial Differential Equations (PDE). The PDE constraint is reduced to a set of nonlinear algebraic equations through (3.4). The resulting nonlinear optimization problem is solved using a primal-dual interior point algorithm implemented in the open-source IPOPT code.

Setting up a problem in IPOPT, as a general purpose optimization package, usually requires less effort and time in comparison with developing an adjoint solver for a specific equation. However, IPOPT runs are more computationally intensive due to the fact that IPOPT assembles and solves a linear system for the solution of Karush-Kuhn-Tucker

(KKT) optimality condition in each iteration if compared with less intensive backward solution of the adjoint system adopted in GDM.

Regarding the stopping criterion, the one in IPOPT is based on the norm of the residual in the optimality condition. At the optimum point, the optimality condition has no residual. Optimization algorithm in IPOPT stops when the norm of the residuals in the optimality condition gets smaller than a predefined tolerance, the default value being 10^{-8}.

3.4 Numerical experiments for the viscous Burgers equation

In this section we present a first numerical experiment showing the influence of the numerical viscosity outlined in the previous section. The main aim is to emphasize that ignoring the dynamics of the continuous model at the numerical level can produce undesired results in optimal control problems in large-time horizons.

On the one hand, we show that the gradient descent method performs successfully whenever the numerical flux and the mesh sizes are chosen appropriately. On the other hand, we present examples where diffusive patterns appear due to the artificial numerical viscosity. In those cases, we show that IPOPT may produce solutions with amplified spurious numerical oscillations.

Let us choose $T = 50$ and the following target function:

$$u^*(x) = \begin{cases} \frac{3}{2000}\Big(- e^{-(5\sqrt{20}+x)^2} + e^{-(2\sqrt{20}+x)^2} \\ \qquad +\sqrt{\pi}\, x\big(\operatorname{erf}(5\sqrt{20}-x) + \operatorname{erf}(2\sqrt{20}+x)\big)\Big), & |x-5| \le 25, \\ 0, & \text{elsewhere.} \end{cases} \tag{3.17}$$

Note that this target, being of compact support, is not exactly reachable for the viscous Burgers equation. Indeed, in view of the Hopf-Cole transformation [94, Chapter 4], it can be seen that the value of solutions at the final time cannot vanish on an open set. Observe also that this target is not too steep, so that we can avoid the numerical consequences on the behavior of the algorithms caused by quasi-shocks [18]. This will allow us to better focus on the impact of the large-time behavior on the performance of optimization techniques.

The final time T, and viscosity ν have been chosen to be self-similar with the sonic-boom propagation in the atmosphere which is characterized by low value of viscosity

present in a long time duration. In doing so, we selected the dimensionless number of $\frac{T\nu}{\rho|u|^2}$ where ρ represents density. $\rho = 1$ in Burgers equation in (3.2). In other words, we chose $T = 50$, $\nu = 10^{-4}$ such that the dimensionless number is of the same order of magnitude as that of the sonic boom propagation in the atmosphere.

Regarding the discretization parameters, in each experiment we first fix the mesh-size Δx. Then, the time-step is chosen according to the stability condition of the scheme, that is

$$\frac{\Delta t}{\Delta x} \max_{j\in\mathbb{Z}} |u_j^0| + 2\nu \frac{\Delta t}{\Delta x^2} \leq C_{sta}. \tag{3.18}$$

In the case of EO, $C_{sta} = 1$, while for MLF, $C_{sta} = 1/2$ (e.g. [18, 38]). However, in our experiments we use $C_{sta} = 1/2$ in both cases, unless otherwise specified, which allows us to better compare their performances. The computational domain is chosen large enough to avoid reflection from boundaries.

3.4.1 Solutions using GDM

We initialize the gradient descent method from the function $u_0 = 0$. For each of the numerical fluxes, we repeat the experiment using a different mesh size, namely $\Delta x = 0.1, 0.2, 0.4, 0.8$.

As we observe in Figure 3.2, with EO we are able to compute a quite satisfactory minimizer even for $\Delta x = 0.8$, which returns an optimal value below 10^{-5}. Making Δx smaller, the obtained result is even better (blue lines in Figure 3.4) but yielding always a similar minimizer. This shows the robustness of the method, the optimal solution being similar in the four cases.

Nevertheless, as we expected, the GDM does not perform so well when coupled with MLF. In Figure 3.3 we observe that large enough mesh-sizes introduce small overshootings around the most steep regions of the target function. This is due to the numerical viscosity that MLF introduces, which is proportional to $\Delta x^2/\Delta t$. When that ratio is comparable with ν, the dynamics of the numerical solution cannot preserve the positive and negative masses (1.5) along the control horizon $[0, T]$. As a consequence, the obtained initial datum requires those oscillations in order to maintain the N-wave shape after large periods of time.

Figure 3.4 shows that, furthermore, the results obtained this way are always worse than the ones obtained using EO. Note that, despite the oscillations, the optimal solution approaches the one obtained using EO as Δx gets smaller. This is in agreement with the results of Section 3.3 on the convergence of discrete minimizers towards the continuous

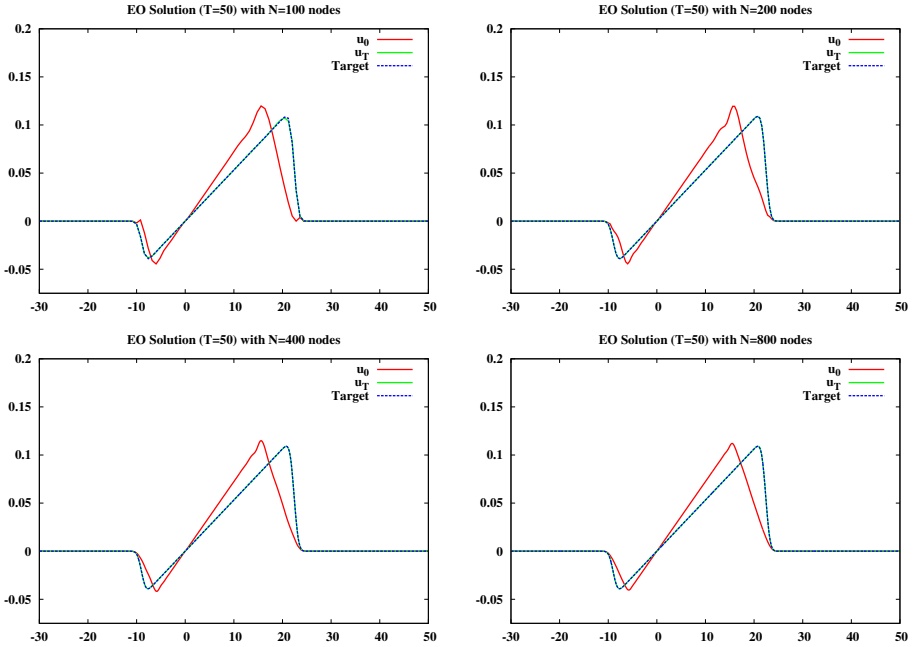


FIGURE 3.2: Optimal solutions (red) for (3.1)-(3.2) with $\nu = 10^{-4}$ and their corresponding state at time $T = 50$ (green) compared to the target (blue), using GDM+EO. From left to right and top to bottom, $\Delta x = 0.8$, 0.4, 0.2 and 0.1 respectively.

ones. But, as expected, convergence is much slower as the mesh-size decreases for the MLF.

Nevertheless, the values corresponding to MLF stay always above the ones obtained using EO after the same number of iterations. In fact, GDM+MLF meets the stopping criteria much before than GDM+EO.

3.4.2 Solutions using IPOPT

In this section, the numerical experiment for viscous Burgers is solved with IPOPT using both EO and MLF numerical fluxes introduced in equations (3.5)-(3.6). As explained in Section 3.3, the algebraic constraints consists of the discretized form of the Burgers equation shown in equation (3.4).

The optimization results obtained with IPOPT are shown in Figure 3.5 for the EO discretization and in Figure 3.6 for the MLF scheme. The initial data u^0 that is sought is shown in red solid lines, while the final solution at time T and the target function u^* are plotted respectively in green and blue lines, being indistinguishable within plotting tolerance. For each scheme, the optimization problem is solved for four different cell sizes from coarse to refined meshes corresponding to $\Delta x = 0.2, 0.1\bar{3}, 0.10, 0.08$. Both EO and MLF fluxes achieve a discrete initial condition which reduce the functional to values less than 10^{-8}.

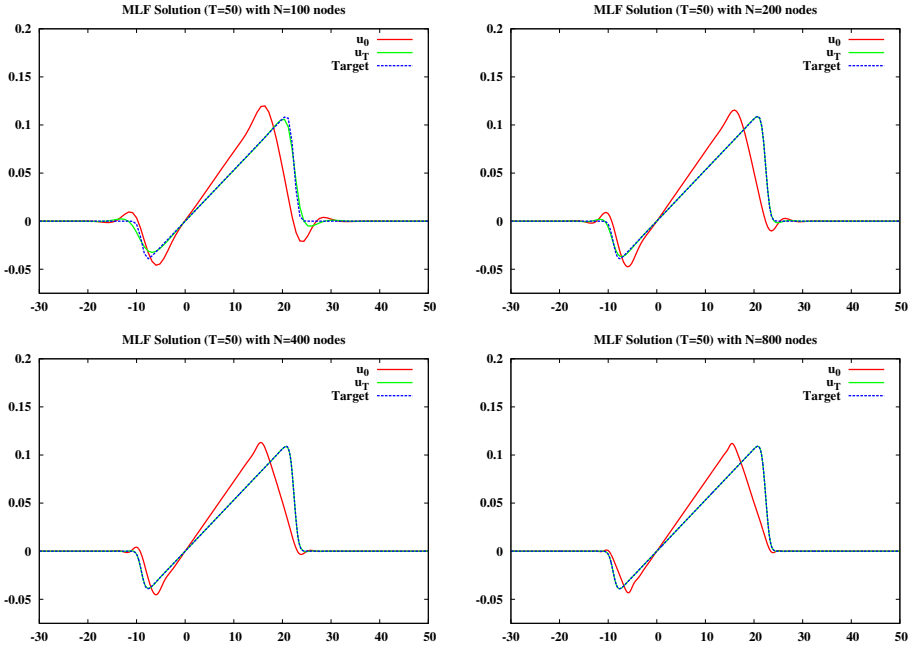


FIGURE 3.3: Optimal solutions (red) for (3.1)-(3.2) with $\nu = 10^{-4}$ and their corresponding state at time $T = 50$ (green) compared to the target (blue), using GDM+MLF. From left to right and top to bottom, $\Delta x = 0.8$, 0.4, 0.2 and 0.1 respectively.

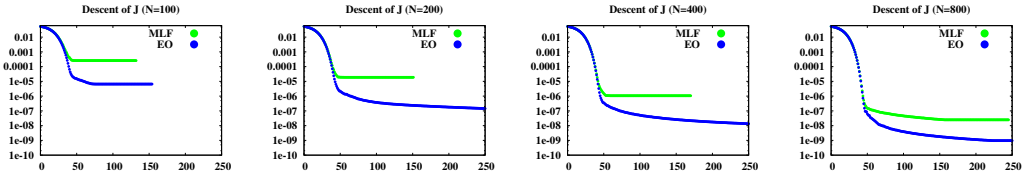


FIGURE 3.4: Descent of the functional J using GDM coupled with EO and MLF. We take, respectively, $\Delta x = 0.8$, $\Delta x = 0.4$, $\Delta x = 0.2$ and $\Delta x = 0.1$ for the case $\nu = 10^{-4}$ and $T = 50$.

The spurious oscillations in initial data obtained with MLF are due to the numerical viscosity that MLF scheme introduces. In cases where numerical viscosity in MLF is comparable with the physical viscosity, MLF discretization can alter the dynamics of the continuous system, especially in long time horizons. This fact constitutes a warning regarding the selection of cell size and time-step when MLF discretization is employed. Indeed, from the inverse problem point of view, extra numerical viscosity in MLF makes the inverse problem even more ill-posed.

Remark 3.4*: To show that oscillations in MLF results are artifacts of the numerical viscosity, the MLF initial data are evolved with a Burgers equation solver which adopts EO discretization and the result is shown in Figure 3.7 on the left. The fact that EO retains oscillations from MLF initial condition corroborates the oscillations in MLF results do not have any physical significance with regard to the Burgers equation and they are merely undesirable artifacts of numerical viscosity of MLF. On the right plot*

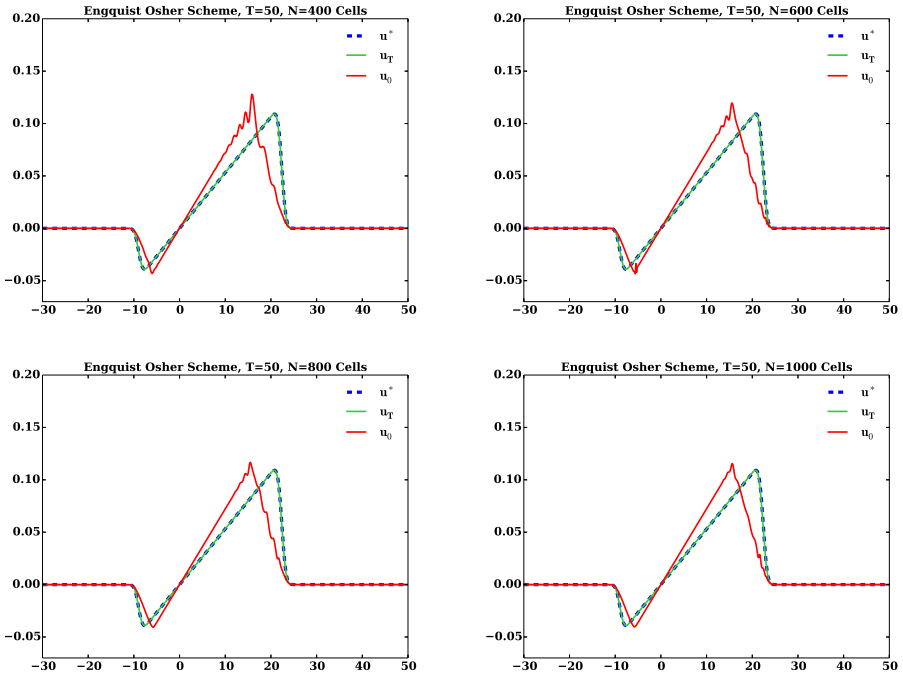


FIGURE 3.5: Optimal solutions for (3.1)-(3.2) with $\nu = 10^{-4}$ and $T = 50$, using EO flux discretization. From left to right and top to bottom, $\Delta x = 0.2, 0.1\bar{3}, 0.10, 0.08$ mesh sizes respectively.

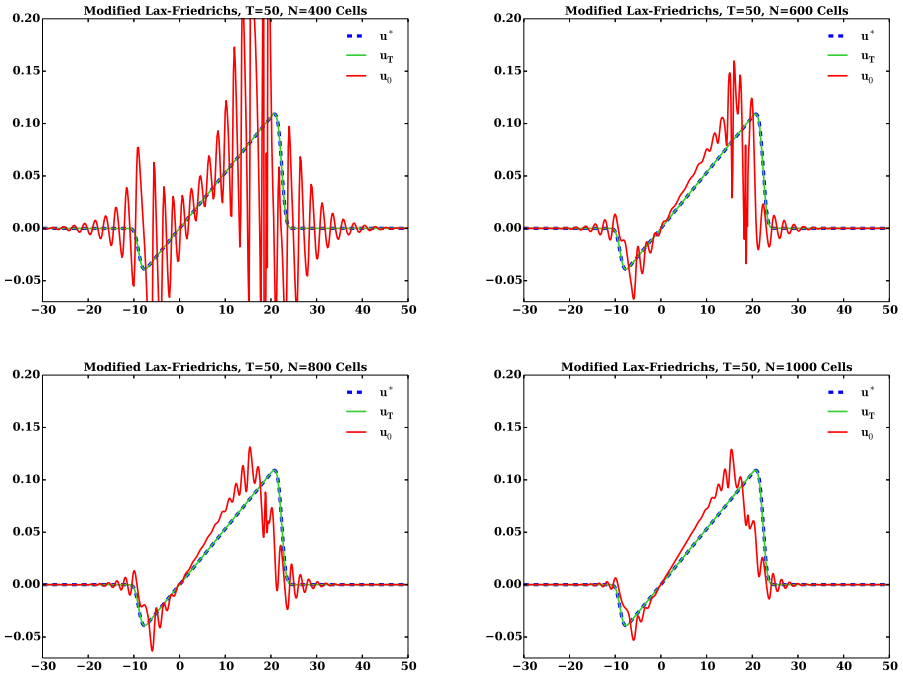


FIGURE 3.6: Optimal solutions for (3.1)-(3.2) with $\nu = 10^{-4}$ and $T = 50$, using MLF flux discretization. From left to right and top to bottom, $\Delta x = 0.2, 0.1\bar{3}, 0.10, 0.08$ mesh sizes respectively.

of Figure 3.7, the optimization results obtained from EO are evolved with MLF. As

observed, the numerical viscosity in MLF solver causes more dissipation in final solution u when it is compared with the target function u^*. In summary, the optimal initial datum obtained by one method is not necessarily the best fit for the functional if the initial datum is evolved with a different scheme. Similar issues were already detected in [33].

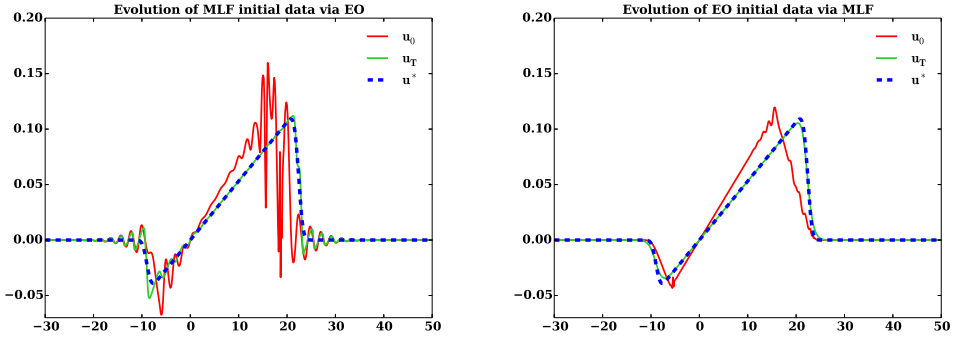


FIGURE 3.7: Optimal solutions (red lines) for (3.1)-(3.2) with $\nu = 10^{-4}$ and $T = 50$ obtained from IPOPT with MLF (left) and EO (right), using $\Delta x = 2/15$. Initial data obtained with MLF is evolved with EO discretization and vice versa, reaching solutions (green lines) which deviate from the target function (blue lines). The optimal initial data obtained for one scheme is not necessarily the best fit for the functional if it is evolved in time with a different scheme.

3.4.3 Comparing GDM and IPOPT results

Here, we compare qualitatively the optimization results of GDM and IPOPT. The results obtained from GDM and IPOPT are quite similar in the case of EO. However, in the case of MLF, results obtained from IPOPT contain more oscillations when compared with those of GDM. It is seen that even for the smallest cell size corresponding to $\Delta x = 0.08$, IPOPT still maintains oscillations. From the frequency domain perspective, as cells are refined, IPOPT will have the opportunity to include high frequency oscillations that were not resolved on coarser meshes. IPOPT via using a higher order optimization scheme is capable of including high oscillations.

In the case of GDM, which employs gradient information calculated by solving the adjoint equation, it does not allow high frequency oscillations due to the boundedness of backward adjoint solutions as we iterate in time. This observation is consistent with the ill-posedness of the backward solution for viscous processes. Lower order convergence of gradient method with adjoint methodology acts as a regularizing effect for the ill-posed problem of the numerical viscosity term introduced by MLF method.

Another interesting observation regarding the extent of oscillations observed in GDM methods will be discussed in Section 3.6 in relation with the choice of the step-size in

descent methods. It will be shown that adopting big step-size ε in the line-search method may result in oscillatory results even in GDM+EO case as shown in Figure 3.12. The aggressive optimal step-size may not guarantee a smooth descent toward the solution in the case of Burgers equation. The sensitivity of the optimization algorithm to the initialization and to the step-size have been further elaborated in Section 3.6.

3.5 Numerical experiment for inviscid Burgers equation

So far, we have focused on the viscous Burgers equation with small viscosity. In this section, we revisit the experiment shown in Section 3.4, but for the inviscid case, in which $\nu = 0$.

It is important to recall that the inviscid Burgers equation is not time reversible. Thus, in principle, not all states lead to a unique minimizer. Let us remark also that, in the absence of viscosity, shock discontinuities may arise in finite time even for smooth initial data. Nevertheless, the target function (3.17) has been chosen in order to avoid the severe influence of shocks on the optimization process to the extent possible. We refer to discussion in Section 3.6 for more details about this issue in the context of the small viscosity and quasi-shocks, which can be extended to the inviscid case.

The results obtained from the gradient descent method, using both EO and MLF, are shown in Figure 3.8. In Figure 3.9 we display the results for the same experiment using IPOPT. One can immediately notice that neglecting the physical viscosity does not cause any substantial difference in the performance of the algorithms. As in the viscous case, EO performs well, while spurious oscillations pollute MLF results.

3.6 Other variants of numerical experiments

It is well known that iterative algorithms can be sensitive to the underlying assumptions considered in their initialization. This section constitutes a warning on how results could be affected if intrinsic characteristics of the optimal control problem (mainly, those related to the PDE and its discretized version) are not taken into account. We revisit the experiment of Section 3.4, based on the viscous case, while modifying the initial settings of the numerical optimization process. All the same, the conclusions can be extended to the inviscid case too.

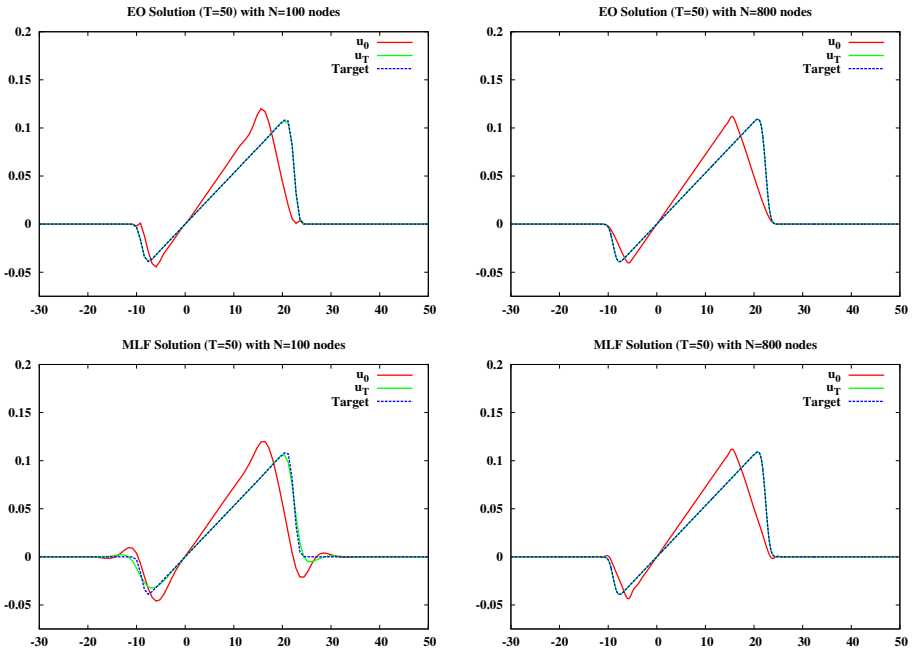


FIGURE 3.8: Initial data $u^0(x)$ for inviscid Burgers equation obtained from GDM+EO (top) and GDM+MLF (bottom). We use $\Delta x = 0.8$ (left) and $\Delta x = 0.1$ (right).

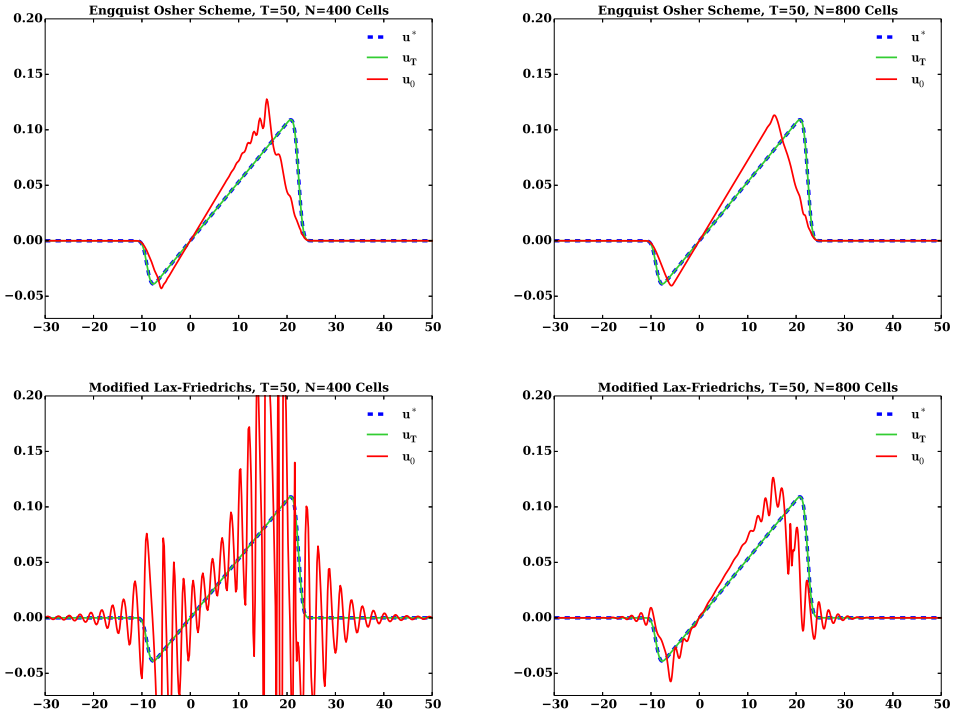


FIGURE 3.9: Initial data $u^0(x)$ for inviscid Burgers equation obtained from IPOPT to fit target function $u^*(x)$ at time $T = 50$ using EO (top) and MLF (bottom) flux discretizations. We use $\Delta x = 0.2$ (left) and $\Delta x = 0.1$ (right).

3.6.1 Sensitivity to the initialization of the optimization algorithm

In all numerical experiments presented so far, the initial data considered to initialize the optimization algorithm were taken to be identically zero. To evaluate the impact of this choice, we solve (3.4)-(3.9) for target (3.17) starting from other initial guesses: a sinusoidal function and a step function. The results obtained with GDM are shown in Figure 3.10, while the ones for IPOPT are shown in Figure 3.11.

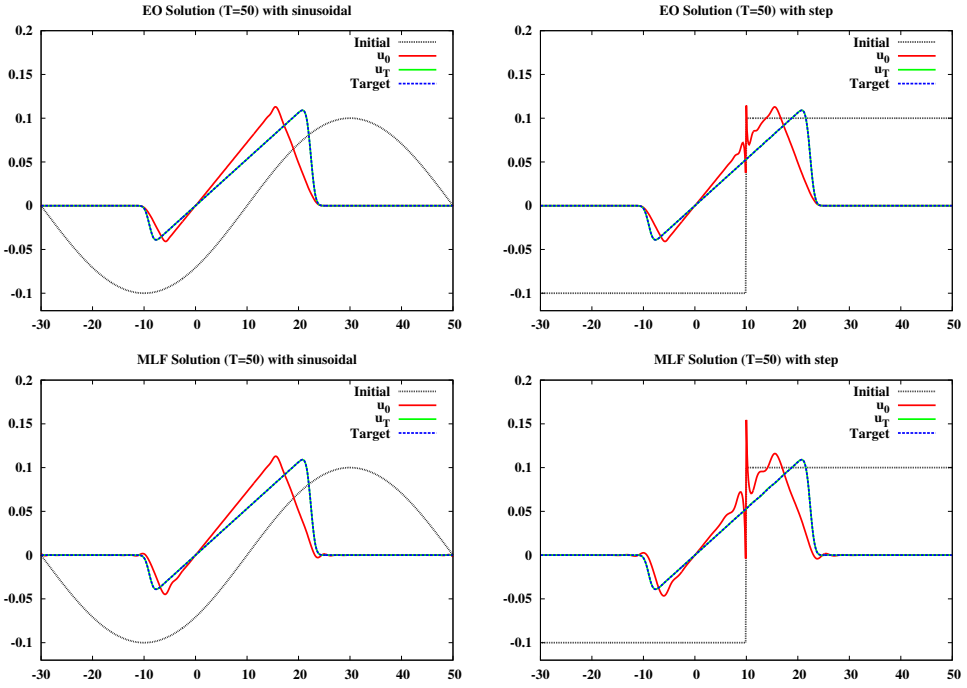


FIGURE 3.10: Optimal initial data $u^0(x)$ obtained with the GDM, using sinusoidal (left column) and step function initialization (right). We compare EO (top) and MLF (bottom) numerical fluxes. The experiment was performed for $\Delta x = 2/15$, $T = 50$ and $\nu = 10^{-4}$.

It is worth underlying the numerical results obtained by means of the step function initialization. This choice of the initialization gives rise to a spike on the optimal initial datum regardless the method we employ. As a matter of fact, the presence of discontinuities in the initial data, even in the viscous Burgers equation, is a delicate issue to deal with (e.g, see [18]). Once they are raised, they have the tendency to persist when employing discrete approaches as in here.

In terms of the value of the functional $\mathcal{J}$, the presence of this kind of spikes in the initial data does not bear any consequences, since it is attenuated during the evolution process. Let us remark that, in what concerns the inviscid dynamics, p and q in (1.5) are not very sensitive to this kind of perturbations; a conclusion that can be extrapolated to the viscous case. Nevertheless, misplaced discontinuities and, more generally, quasi-shocks affect severely the optimization process.

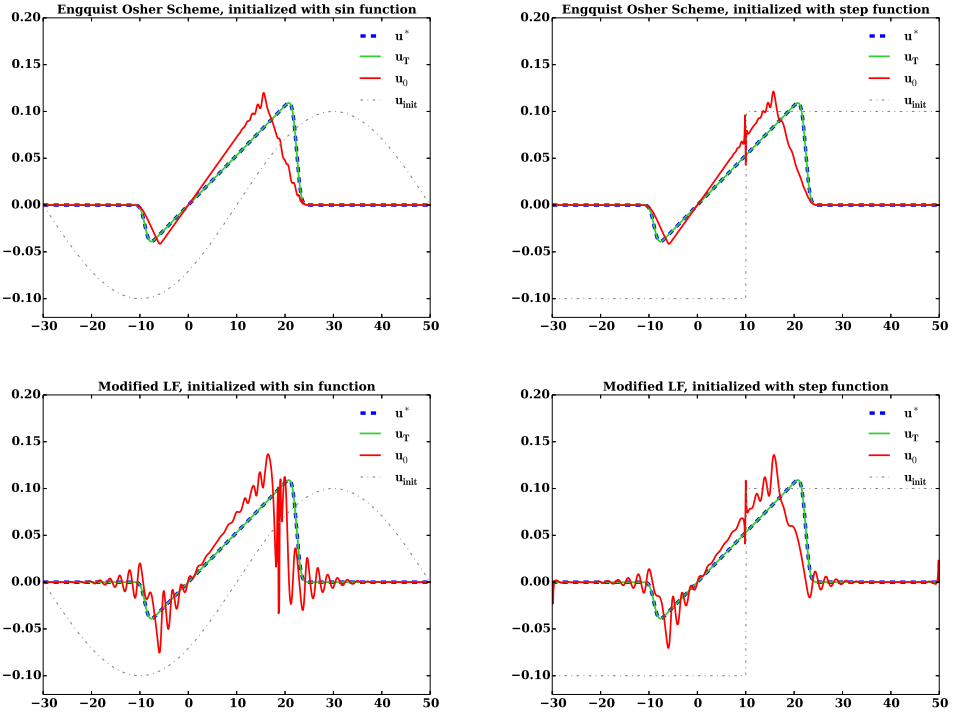


FIGURE 3.11: Optimal initial data $u^0(x)$ obtained from IPOPT to fit the target function $u^*(x)$, using sinusoidal (left column) and step function initialization (right). We compare EO (top) and MLF (bottom) numerical fluxes. The experiment was performed for $\Delta x = 2/15$, $T = 50$ and $\nu = 10^{-4}$.

3.6.2 Choice of the step-size in descent methods

Gradient descent methods based on the adjoint equation, that we introduced in Section 3.3, have several variations depending on the way of choosing the descending direction or the step-size. We refer to [19, 77] for basic reference about some of those. The main objective of these variations is, usually, to obtain a better convergence rate. In this section we want to highlight that a quick descent at the very first iterations does not imply a better global convergence. In other words, being conservative at the beginning can lead into a better global performance of the optimization process.

In the method that we introduced in Section 3.3 we did not specify the way one should choose ε in (3.16). Taking $\delta u^0 = -\rho(0)$ and a sufficiently small ε guarantees a descending step, even if it might not be optimal. In the experiments of this chapter concerning the GDM, in each iteration, we have chosen ε in the following manner:

$$\varepsilon_{new} = 0.5^r(1.2\varepsilon_{old}), \tag{3.19}$$

with the smallest $r = 0, 1, 2, \ldots$ such that

$$\mathcal{J}(u_\Delta^{0,new}) \leq \mathcal{J}(u_\Delta^{0,old}).$$

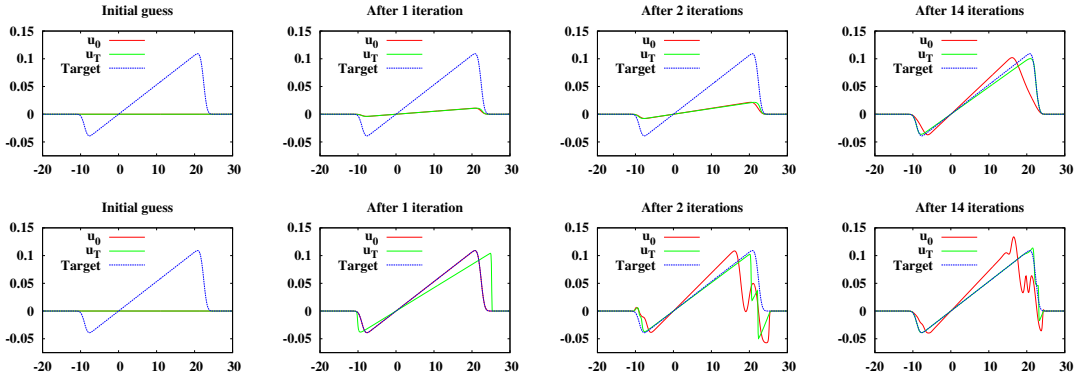


FIGURE 3.12: Initial data obtained after 1, 2 and 14 iterations of the GDM+EO method, started from $u_\Delta^0 = 0$, using $\varepsilon_0 = 0.1$ (top) and $\varepsilon_0 = 1$ (bottom) as initial step-sizes for (3.19). As shown in Figure 3.13, the descent gained in the first iteration using a bigger step is lost due to the creation of a misplaced quasi-shock.

Note that we do not claim this step-size to be the best choice for this type of problems. Due to its simplicity, it is enough to show the pathological results that numerical viscosity may produce in large-time horizon problems. Moreover, steepest descent or conjugate gradient methods should be more efficient generally, but they also undergo the same issues.

A bigger value for the initial ε_0 will imply a deeper modification of the initial data in the earlier steps and, possibly, a bigger decrease of the functional. Nevertheless, this does not guarantee a better convergence. As a matter of fact, in Figure 3.12 we present the initial data obtained in the experiment of Section 3.4 after the first iterations of the GDM+EO method with $\Delta x = 0.1$. In this case, we have chosen the step size according to (3.19), but starting from $\varepsilon_0 = 0.1$ and $\varepsilon_0 = 1$. One can clearly notice that, after the first iteration, the initial data obtained using $\varepsilon_0 = 1$ produces a closer solution to the target function.

The cost of creating a misplaced quasi-shock is high and it is reflected in the subsequent iterations. Dealing with quasi-shocks in optimal control problems like the one we present in this chapter is a very hard task. Even if solutions of (3.2) are continuos for $t > 0$, very steep regions (quasi-shocks) may arise if ν is small. From the numerical point of view, these behave almost like shocks and, thus, their presence in incorrect places is not desirable. Since the gradient descent method of Section 3.3 is designed from the point of view of smoothness of solutions, it finds difficult to bypass quasi-shock misplacements. In a figurative manner, this can be represented as in Figure 3.13. Note that taking a big step can reduce drastically the functional, but in exchange for a badly located quasi-shock. This ends up driving the descending path through flatter regions that are harder to overcome.

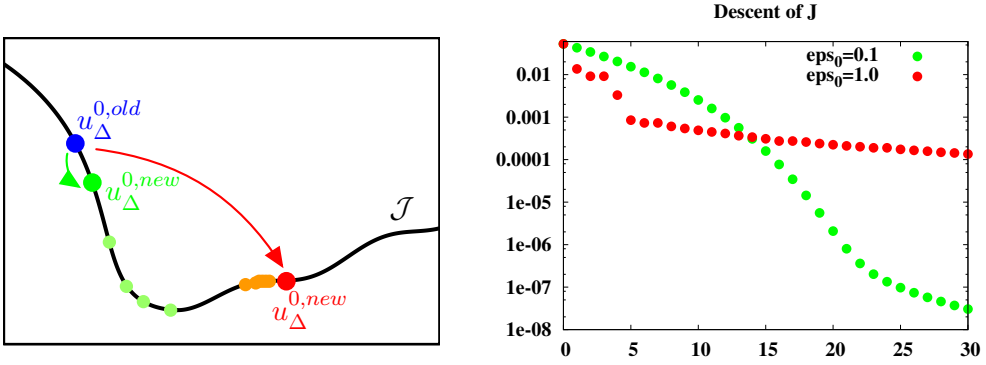


FIGURE 3.13: A big step at the beginning can take the descending path through flat regions. For instance, the creation of a quasi-shock in an incorrect place, as in Figure 3.12, can lead to a slower convergence. On the right, evolution of the functional using GDM+EO with $\varepsilon_0 = 0.1$ (green) and $\varepsilon_0 = 1$ (red) as initial step-sizes.

Remark 3.5*: Let us emphasize that this issue is not specific to the gradient descent method that we have employed. Other methods, such as the conjugate gradient, can be equally affected if quasi-shocks arise in other places but the correct ones.*

Of course, this issue could be mitigated if more complex techniques, capable of handling quasi-shocks, were used. For instance, the alternating descent method developed for the hyperbolic case in [17] and applied to the parabolic case in [18], allows to move the location of the discontinuities in the initial data, reducing the number of iterations drastically.

3.6.3 Reducing the time-step

So far, we have chosen the largest time-step Δt that (3.18) allows, reducing Δx to illustrate the convergence of the algorithms. However, one could think on taking a smaller time-step instead, choosing Δt such that

$$\Delta t < \Delta t_{max} = C_{sta} \frac{\Delta x^2}{\Delta x \max\limits_{j\in\mathbb{Z}} |u_j^0| + 2\nu}. \tag{3.20}$$

In terms of stability of the numerical schemes, this is permitted. But, as we show below in Figure 3.14, the optimization results can be severely affected by this fact.

In the case of EO, a smaller time-step does not have much influence on the results. On the contrary, MLF is strongly affected by this situation. Let us recall that the large-time asymptotic profile of the MLF is given by a viscous Burgers equation in which the viscosity parameter is proportional to $\Delta x^2/\Delta t$. Therefore, a smaller Δt implies a higher

numerical viscosity. Figure (3.14) highlights this pathology in the case of the gradient descent method.

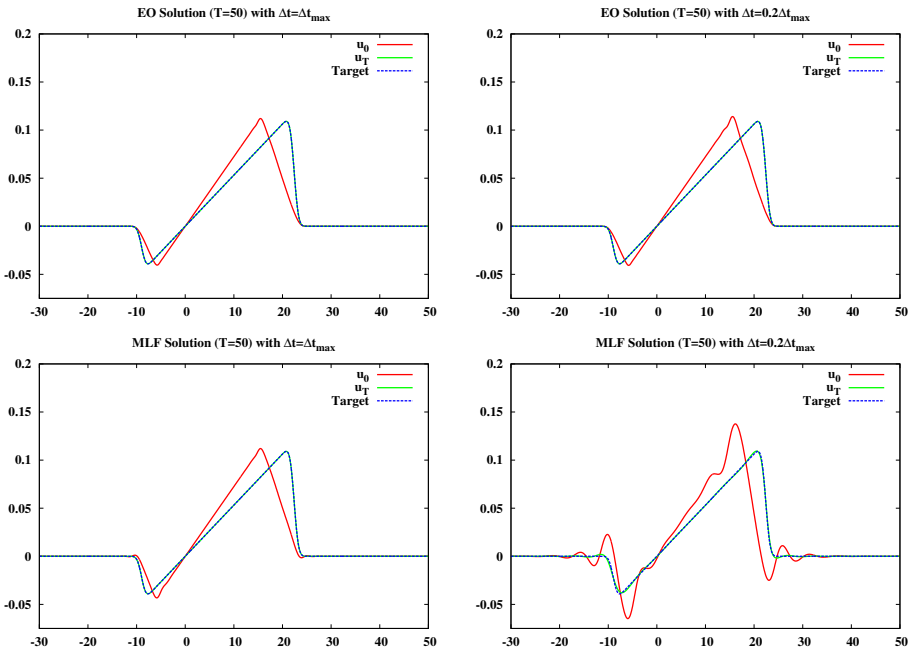


FIGURE 3.14: Initial data $u^0(x)$ for inviscid Burgers equation obtained after 300 iterations of the GDM+EO (top) and GDM+MLF (bottom). We use $\Delta x = 0.1$ and $\Delta t = \Delta t_{max}$ (left) and $\Delta t = 0.2\Delta t_{max}$ (right).

The enhancement of the numerical oscillations on the recovered initial datum when reducing Δt for MLF appears to be a purely linear issue. Actually, one can observe the same performance when solving the heat equation numerically backwards in time. Let us apply the same technique of the MLF to the heat equation:

$$v_t - \nu v_{xx} = 0 \quad \approx \quad \frac{v_j^{n+1} - \frac{v_{j-1}^n + 2v_j^n + v_{j+1}^n}{4}}{\Delta t} + \nu \frac{v_{j-1}^n - 2v_j^n + v_{j+1}^n}{\Delta x^2} = 0$$

In the Fourier space we have that

$$\hat{v}(t,\xi) = e^{-\nu t \xi^2} \hat{v}(0,\xi)$$

and

$$\hat{v}^n(\xi) = \left(1 - \left(4\nu \frac{\Delta t}{\Delta x^2} + 1\right) \sin^2\left(\frac{\xi \Delta x}{2}\right)\right)^n \hat{v}^0(\xi).$$

Hence, the inverse problem is clearly ill-posed, both from the continuous and the discrete points of view. In fact, we get, respectively,

$$\hat{v}(0,\xi) = e^{\nu T \xi^2} \hat{v}(T,\xi) \tag{3.21}$$

and

$$\hat{v}^0(\xi) = \left(1 - \left(4\nu\frac{\Delta t}{\Delta x^2} + 1\right)\sin^2\left(\frac{\xi\Delta x}{2}\right)\right)^{-N}\hat{v}^N(\xi), \tag{3.22}$$

Note that, of course, a final solution $v(x,T)$ in $L^2(\mathbb{R})$ does not guarantee that the initial data is in $L^2(\mathbb{R})$, since each frequency $\xi \in \mathbb{R}$ is weighted by an exponential term $e^{\nu T\xi^2}$. Moreover, for the corresponding initial datum to be in $L^2(\mathbb{R})$, one needs the final target to be in an exponentially weighted space in the frequency domain.

The discrete case is even more sensitive. In fact, a Taylor expansion of the weights in (3.21) and (3.22) already denotes a substantial difference in the second term:

$$e^{\nu T\xi^2} = 1 + \nu T\xi^2 + O(\xi^4),$$

while, taking $N = T/\Delta t$,

$$\left(1 - \left(4\nu\frac{\Delta t}{\Delta x^2} + 1\right)\sin^2\left(\frac{\xi\Delta x}{2}\right)\right)^{-N} = 1 + \left(\frac{\Delta x^2 T}{4\Delta t} + \nu T\right)\xi^2 + O(\xi^4).$$

Therefore, the dissipation is enhanced in the whole spectrum when making Δt smaller.

3.6.4 The set of reachable target functions

The aim of the forthcoming discussion is to informally illustrate the underlying performance of the discrete approximations with less degree of rigor in the terminology.

Let us denote by $S(t)$ the semigroup corresponding to equation (3.2). The approximated solutions given by (3.4) generate a semigroup too. Let us denote by $S_\Delta^{EO}(t)$ the one corresponding to Engquist-Osher and by $S_\Delta^{MLF}(t)$ the one for modified Lax-Friedrichs. The previous experiments exhibit that $S_\Delta^{EO}(T)[L^1(\mathbb{R})]$ approximates quite accurately the set $S(T)[L^1(\mathbb{R})]$, even for large values of T. On the other hand, MLF requires smaller values of Δx to reproduce $S(T)[L^1(\mathbb{R})]$ (see Figure 3.15). For large enough Δx, MLF looks for the closest solution within $S_\Delta^{MLF}(T)[L^1(\mathbb{R})]$, producing the spurious oscillations that we have shown.

In all previous experiments we have used target functions that can be very well approximated following the dynamics of the Burgers equation. In particular, we have chosen functions that satisfy the Oleinik condition. It is well known that solutions to (3.2) satisfy $u_x \leq 1/t$ in the sense of distributions.

If the target is not in $S(T)[L^1(\mathbb{R})]$ (for instance, it is enough to take a function u^* such that $u_x^*(T) > 1/T$ in some interval), both numerical semigroups will project the

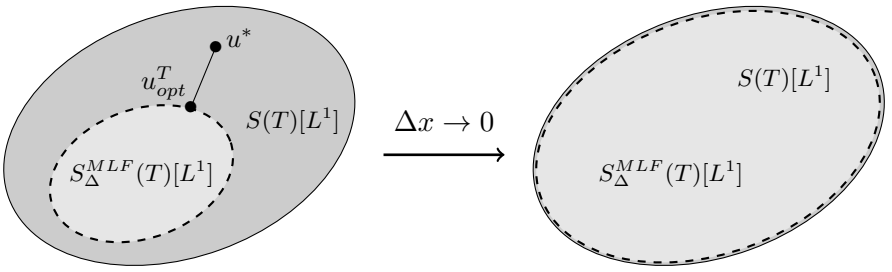


FIGURE 3.15: Set of functions that are reachable at time T by the semigroup $S(t)$ of the Burgers equation for L^1 initial data and the semigroup $S^{MLF}_\Delta(t)$ associated to the modified Lax-Friedrichs scheme. As $\Delta x \to 0$, the set $S^{MLF}_\Delta(T)[L^1(\mathbb{R})]$ gets closer to $S(T)[L^1(\mathbb{R})]$.

target onto their reachable sets (Figure 3.16). Since the set $S^{MLF}_\Delta(T)[L^1(\mathbb{R})]$ is smaller than $S(T)[L^1(\mathbb{R})]$, MLF will make use of spurious oscillations again.

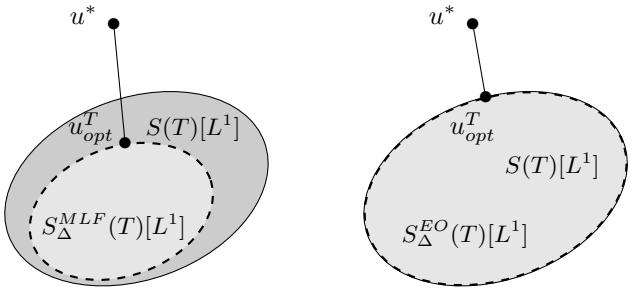


FIGURE 3.16: Projection of the target function onto the reachable sets by modified Lax-Friedrichs (left) and Engquist-Osher (right). The approximation obtained by EO is closer than the one by MLF.

Let us illustrate this with another experiment. Let us consider a new target function u^{**}, given by

$$u^{**}(x) = \begin{cases} \frac{1}{100}\Big(-e^{-(5\sqrt{20}+x)^2} + e^{-(2\sqrt{20}+x)^2} \\ \qquad +\sqrt{\pi}\, x\big(\operatorname{erf}(5\sqrt{20}-x) + \operatorname{erf}(2\sqrt{20}+x)\big)\Big), & |x-5| \le 25, \\ 0, & \text{elsewhere.} \end{cases} \tag{3.23}$$

Note that u^{**} is the same as u^*, but with a larger amplitude. In this case, we have chosen the target function such that $u^{**}_x(x) > 1/50$ in some interval. Therefore, u^{**} cannot be reached at $T = 50$ following the dynamics of the Burgers equation, not even approximately.

Indeed, we can observe in Figure 3.17 that the optimal solutions do not fit u^{**}. Let us remark that the closest function that GDM can recover at final time $T = 50$ is precisely

the one that has slope of 1/50 in the intermediate part which is the maximum slope allowed by the Oleinik condition.

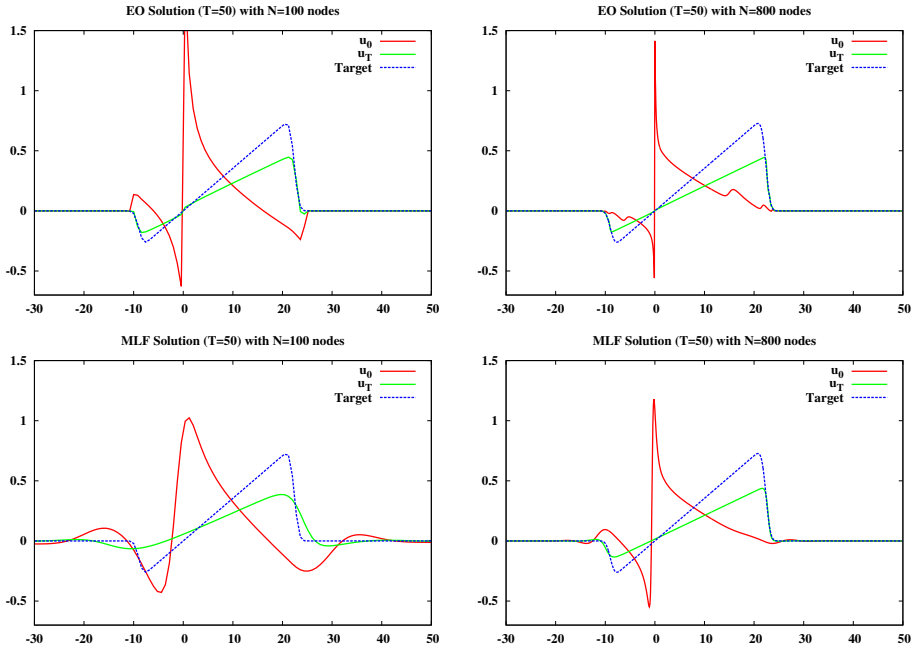


FIGURE 3.17: Initial data $u^0(x)$ obtained using both GDM+EO (top) and GDM+MLF (bottom) to solve (3.4)-(3.9) for the target function u^{**}. We take $\Delta x = 0.8$ (left) and $\Delta x = 0.1$ (right) with $\nu = 10^{-4}$.

Let us remark that, when the target is not reachable, distinct local minimizers could be obtained using the classical discrete adjoint methodology. For the sake of simplicity, let us focus on the inviscid case now. The characteristics for the inviscid Burgers equation are the same as the ones of its adjoint system. Therefore, rarefaction waves that arise from increasing discontinuities in the initial data, as in the example above, generate a region in which almost all the information of the adjoint is lost (see gray triangle in Figure 3.18). There, the adjoint solution collapses into a single point that the discrete adjoint is not capable of recovering correctly. Note the spikes appearing around the discontinuity in Figure 3.17, that arise because of this pathology. In conclusion, this problem needs to be handled carefully, since any non-entropic region (in the sense of not satisfying Oleinik inequality) creates a discontinuity in the initial data, which can hardly be moved with classical optimization tools. Nevertheless, this is not surprising, since, as pointed out in [35, 36], the adjoint in the shock region might not be well approximated if the numerical viscosity is not handled correctly.

This issue may end up forcing the gradient descent method to converge to a local minimizer, also in the low-viscous case. Let us illustrate it with the following example.

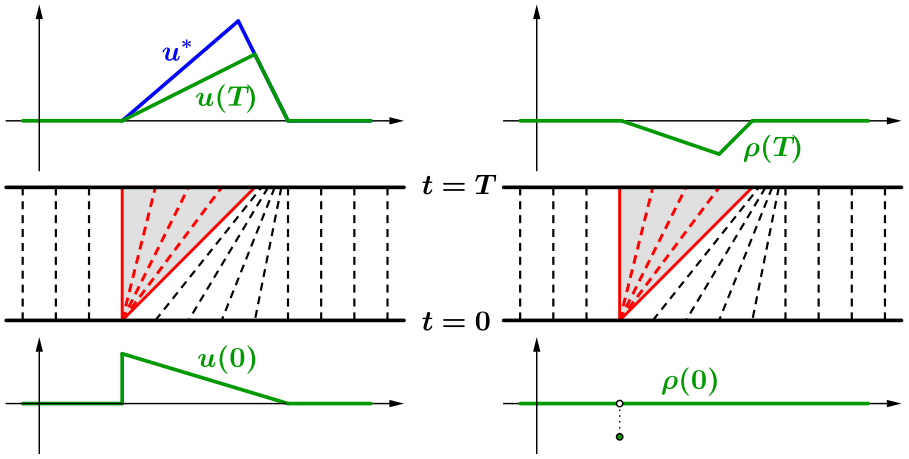


FIGURE 3.18: In the inviscid case, the solutions are propagated through characteristics (left). The adjoint has the same characteristics as the forward problem and, thus, the rarefaction produces a collapse of the adjoint state.

We consider the target function

$$u^{***}(x) = \begin{cases} \frac{3\sqrt{\pi}}{2000}\, x\big(\operatorname{erf}(5\sqrt{20} - x) + \operatorname{erf}(x)\big), & |x - 10| \leq 25, \\ 0, & \text{elsewhere.} \end{cases} \tag{3.24}$$

and a final time $T = 300$ and solve (3.4)-(3.9) with $\nu = 10^{-4}$. As we can observe in Figure 3.19, the minimizer that we obtain is different depending on the initialization chosen. If we start from $u_0 = 0$, the algorithm is not able to move the discontinuity of the initial data to the left. A similar issue occurs when starting from $u_0 = 0.12$. A better choice, but not necessarily the best one, is $u_0 = 0.06$.

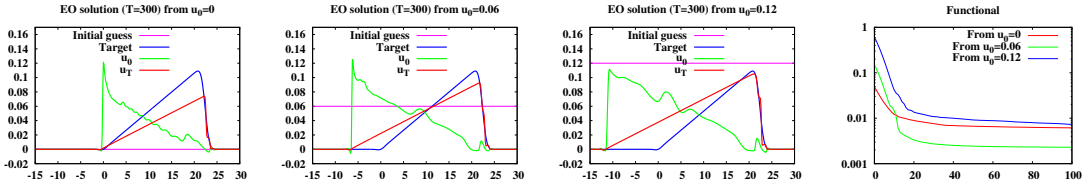


FIGURE 3.19: From left to right, minimizer obtained starting from $u_0 = 0$, $u_0 = 0.06$ and $u_0 = 0.12$, using EO with $\Delta x = 0.1$, together with the evolution of the functional. We consider target function u^{***}, $\nu = 10^{-4}$ and $T = 300$.

3.7 Functional landscapes

In this section we plot some of the landscapes of the functionals we have minimized, in order to explain the observed phenomena in the behavior of the minimizers. As observed

in the experiments performed in Section 3.4, the obtained minimizers exhibit a critical sensitivity to the EO/MLF large-time behavior dichotomy.

The difference among the obtained solutions could correspond to the presence of multiple local minimizers of the functionals under consideration. In order to shed light on this issue a simple experiment is proposed, the idea being to visualize the functional $\mathcal{J}(u_0)$ for a set of initial conditions parameterized with two degrees of freedom. The functional landscape is plotted as these parameters vary, to check the possible presence of multiple wells.

For constructing the initial data, a set of parameterized sinusoidal waves are superposed on a smooth N-wave shaped function w, obtained by subtracting and scaling two error functions. The oscillatory signals are placed only in the neighborhood of sharp slopes of the N-wave and they are parameterized with their amplitude r and frequency s as follows:

$$u^0_{r,s}(x) = \begin{cases} w(x) + r\sin(sx), & \text{near steep fronts,} \\ w(x), & \text{elsewhere,} \end{cases}$$

where

$$w(x) = h\,x\,\big(\operatorname{erf}(m(x-a)) - \operatorname{erf}(m(x-b)\big)$$

Parameters a and b are related to the location of the peaks and m to the sharp slopes they create. Also, the slope of the middle part of the N-wave is proportional to h. In any case, we will fix all of them, leaving r and s the only degrees of freedom. A sample initial condition with sinusoidal oscillations is depicted in Figure 3.20 (left).

The target function that we choose in this case is selected such that it corresponds to the solution of viscous Burgers equation at t=T starting from $u^0_{0,0}$; this means that the desired initial condition does not contain any sinusoidal oscillation. Hence, optimal solutions will be achieved for all pairs of $(r, 0)$, and $(0, s)$ where $r, s \in \mathbb{R}$ (which, indeed, correspond to the same function). Of course, the cost function value corresponding to these pairs would be exactly zero. A plot of the target function and its corresponding initial condition is shown in Figure 3.20 (right).

In our experiment, we first sample pairs of $(r, s) \in [-0.025, 0.025] \times [0, 2\pi]$. Then, we compute $u_\Delta(x, T)$ using (3.4) (with $\nu = 10^{-4}$) from initial condition $u^0_{r,s}(x)$ and evaluate the corresponding cost function $\mathcal{J}_\Delta(u^0_{r,s})$. In particular, we set w with $a = -10$, $b = 20$, $h = 0.003$ and $m = 0.3$; that is,

$$w(x) = 0.003\,x\,\big(\operatorname{erf}(0.3(x+10)) - \operatorname{erf}(0.3(x-20)\big).$$

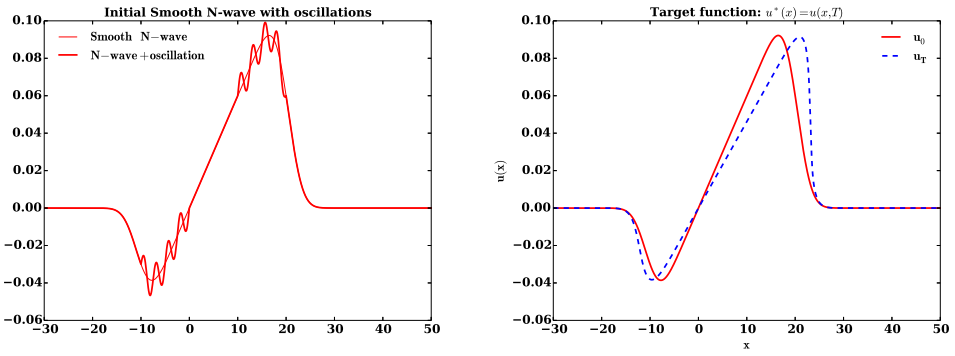


FIGURE 3.20: Initial conditions are constructed from superposing parameterized sinusoidal oscillations on a base "smooth N-wave" (left). Target function u^* is constructed from the evolution of the initial data corresponding to the N-wave and free from any oscillations. The optimal solution would be a N-wave without superposed oscillations. Final time is $T = 50$ and physical viscosity is $\nu = 10^{-4}$.

The surface plots of the cost function $\mathcal{J}(u^0_{r,s})$ for the sampled (r, s) are shown in Figure 3.21. In the top row, u_Δ is calculated using EO flux with $\Delta x = 0.2$ (left), and $\Delta x = 0.1$ (right) mesh-sizes; while in the bottom row, u_Δ is evaluated by using MLF flux.

As observed from figures, for the range of initial conditions considered, the functional landscapes obtained from EO and MLF discretizations show no presence of multiple wells. The landscape obtained from EO shows more sensitivity to the variation of the parameters. On the contrary, MLF dissipates sinusoidal signals and the landscape obtained from MLF is less sensitive to higher frequencies $s \gg 1$. Moreover, in the case of MLF the choice of Δx (and, hence, Δt) also impacts the sensitivity of the landscape.

3.8 Conclusions and perspectives

This chapter further builds on the previous work done on analyzing discretization schemes which preserve the large-time behavior. Here, we showed some of the pathologies that may arise in numerical approximation of optimal control problems in such regimes. In particular, we highlighted the performance of classical flux discretization methods and the consequences if the discretization scheme is not chosen properly. It was shown how the choice of the discretization scheme can alter the underlying dynamics of the system and its corresponding functional. Moreover, the optimal initial datum obtained by one discretization scheme may not be the best fit for the functional if it is evolved in time with a different numerical scheme.

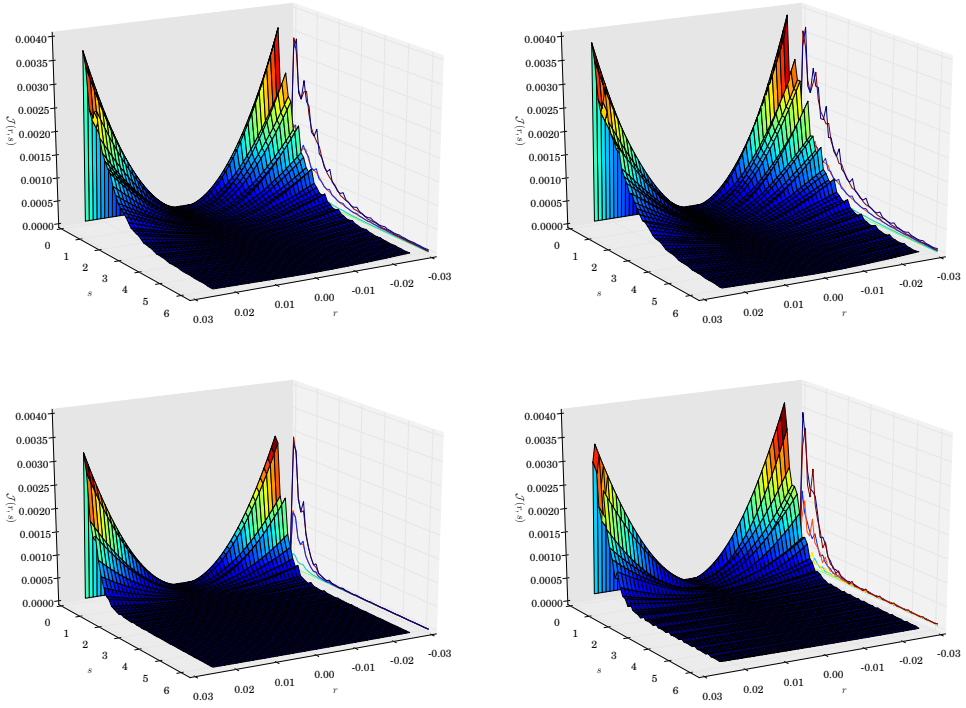


FIGURE 3.21: Functional landscape $(r, s, \mathcal{J})$; Engquist-Osher (top row) and modified Lax-Friedrichs (bottom row) fluxes are used for discretizing the viscous Burgers equation with $\Delta x = 0.2$ (left), and $\Delta x = 0.1$ (right). The landscape of MLF is less sensitive to the variation of parameters (r, s) due the dissipative nature of MLF.

In this regard, we presented several numerical experiments based on the resolution of the Burgers equation (both viscous and inviscid) for large times. We emphasize that, in this regime, the impact of the numerical viscosity is enhanced. The optimization was done employing gradient descent method (GDM) and an open source interior point optimization software (IPOPT), but the consequences can be extended to other optimization techniques.

In summary, following conclusions are made:

- The performance of MLF is not satisfactory in comparison with EO. The initial data obtained with MLF can be quite oscillatory in the vicinity of sharp slopes for large cell sizes. Also, the convergence rate of MLF is slower in comparison with EO for equal number of cells.

- The performance of EO when employed in GDM or IPOPT is satisfactory. The initial data obtained is less sensitive to the choice of Δx and Δt. Furthermore, the results are also less sensitive to the initial setting of the optimization algorithms.

- The numerical viscosity present in MLF is proportional to $\Delta x^2/\Delta t$; hence, the solution is quite dependent on the domain discretization. While coarse discretizations

result in more numerical viscosity, small time steps can increase the numerical viscosity too. This point underlines the importance of selecting a proper CFL number in the simulations done with MLF.

- Results obtained from IPOPT with MLF were more oscillatory in comparison with those of GDM method. The MLF functional shows insensitivity to the high frequency oscillations as it was illustrated in the functional landscape. It is likely to obtain two dissimilar solutions in terms of high frequency contents if we slightly perturb the convergence tolerance in the optimization algorithm. This observation is consistent with the ill-posedness of the backward solution of the diffusion process.

- The possibility of presence of multiple distinct local minima in the functional landscape, when the target is approximately reachable, is excluded based on the functional landscape visualization. However, the functional landscape shows flatness (or lack of sensitivity) in the neighborhood of minimizers in the case of MLF.

- For optimization problems with physically unreachable target functions, which they do not satisfy Oleinik inequality, the optimization solvers try to fit the target as best as possible within their reachable sets. In these situations, EO performance is again more robust compared with MLF.

Chapter 4

A large-time behavior preserving scheme for the augmented Burgers equation

4.1 Introduction and main results

In this chapter we consider the augmented Burgers equation with constant parameters and a unique relaxation process (see, for instance, [20] and the references therein for its derivation). We focus on the following equation:

$$\begin{cases} u_t = uu_x + \nu\, u_{xx} + c\, K_\theta * u_{xx}, & (t,x) \in (0,\infty)\times\mathbb{R}, \\ u(0,x) = u^0(x), & x \in \mathbb{R}, \end{cases} \tag{4.1}$$

where $*$ denotes the convolution in the x variable, the parameters ν, c, θ are positive and

$$K_\theta(z) = \begin{cases} \frac{1}{\theta} e^{-z/\theta}, & z > 0, \\ 0, & \text{elsewhere.} \end{cases} \tag{4.2}$$

As we explained in Chapter 1, the industrial application of this kind of models to the simulation and control of the sonic-boom phenomenon needs to approximate solutions for large time. Therefore, they need a good understanding of the behavior of the solutions in these extended regimes in order to be able to simulate them accurately. As we already showed in Chapters 2 and 3, this issue needs to be treated carefully. In our case, (4.1) is not a hyperbolic equation and, hence, the asymptotic profile is not an N-wave, but a diffusive wave. Nevertheless, in our simulations we show that small values for ν and c require a similar treatment from the numerical point of view, as if the equation was a hyperbolic conservation law. In fact, in those situations, the solution may develop very

steep regions (in what follows, we refer to these as quasi-shocks), which numerically behave almost like shocks.

For the sake of simplicity, the asymptotical analysis done in the first sections is focused only on the case $\nu = c = \theta = 1$, but the extension to any positive value of the parameters is immediate. We will omit the subindex θ whenever its value is one. In this case, we have that

$$K * u_{xx} = K * u - u + u_x.$$

Thus, (4.1) can be rewritten in a more suitable manner as follows:

$$\begin{cases} u_t = uu_x + u_{xx} + K * u - u + u_x, & (t,x) \in (0,\infty) \times \mathbb{R}, \\ u(t=0,x) = u^0(x), & x \in \mathbb{R}. \end{cases} \tag{4.3}$$

The main goals of the present paper are to analyze the asymptotic behavior of the solutions to (4.3) as $t \to \infty$ and to build a semi-discrete numerical scheme that preserves this behavior. In what concerns the large-time behavior of solutions of system (4.3), the main result is stated in the following theorem.

Theorem 4.1. *Let $u^0 \in L^1(\mathbb{R})$. For any $p \in [1,\infty]$, the solution u to (4.3) satisfies*

$$t^{\frac{1}{2}(1-\frac{1}{p})} \|u(t) - u_M(t)\|_p \longrightarrow 0, \quad \text{as } t \to \infty,$$

where $u_M(t,x)$ is the solution of the following equation:

$$\begin{cases} u_t = uu_x + 2u_{xx}, & x \in \mathbb{R}, t > 0, \\ u(0) = M\delta_0. \end{cases}$$

Here δ_0 denotes the Dirac delta at the origin and M is the mass of the initial data, $M = \int_{\mathbb{R}} u^0(x)dx$.

Let us recall that function u_M is defined in (1.14), simply taking $\nu + c$ instead of ν as the viscosity parameter. Moreover, we can assure that u_M is of self-similar nature.

In the cases when ν, c and θ are no longer equal to one, the asymptotic profile does not depend on θ. Moreover, the coefficient in front of the viscosity term in the equation satisfied by the profile is $\nu + c$:

$$\begin{cases} u_t = uu_x + (\nu + c)u_{xx}, & x \in \mathbb{R}, t > 0, \\ u(0) = M\delta_0. \end{cases}$$

Note that for small values of ν and c, the numerical treatment needs to be done carefully. One needs to choose a numerical flux that does not introduce too much numerical viscosity in the system. Moreover, the truncation of the integral term is also critical, since it can change the large-time dynamics of the model. For these reasons we propose the following semi-discrete scheme:

$$\begin{cases} u_j'(t) = \dfrac{g_{j+1/2}(t) - g_{j-1/2}(t)}{\Delta x} + \dfrac{u_{j-1}(t) - 2u_j(t) + u_{j+1}(t)}{\Delta x^2} \\ \qquad\qquad + \displaystyle\sum_{m=1}^{N} \omega_m u_{j-m}(t) - F_0^\Delta u_j(t) + F_1^\Delta \frac{u_{j+1}(t) - u_j(t)}{\Delta x}, \qquad j \in \mathbb{Z}, t \geq 0, \\ u_j(0) = \dfrac{1}{\Delta x} \displaystyle\int_{x_{j-1/2}}^{x_{j+1/2}} u^0(x)dx, \qquad\qquad j \in \mathbb{Z}, \end{cases} \tag{4.4}$$

where

$$\omega_m = e^{-m\Delta x}\left(e^{\Delta x} - 1\right), \quad m = 1, \dots, N, \tag{4.5}$$

and

$$g_{j+1/2}(t) = \frac{u_j(t)\big(u_j(t) - |u_j(t)|\big)}{4} + \frac{u_{j+1}(t)\big(u_{j+1}(t) + |u_{j+1}(t)|\big)}{4}, \quad j \in \mathbb{Z}, t \geq 0.$$

The parameter $N = N(\Delta x) \in \mathbb{N}$ denotes the number of nodes considered in the quadrature formula of the integral. The correcting factors F_0^Δ and F_1^Δ in front of the approximations of u and u_x, given by

$$F_0^\Delta = \sum_{m=1}^{N} \omega_m \quad \text{and} \quad F_1^\Delta = \Delta x \sum_{m=1}^{N} m\omega_m, \tag{4.6}$$

handle, from the asymptotic behavior point of view, the truncation of the nonlocal term. The approximated solution u_Δ is a piecewise constant in space function, defined as follows:

$$u_\Delta(t, x) = u_j(t), \quad x \in (x_{j-1/2}, x_{j+1/2}), t \geq 0, \tag{4.7}$$

where $x_{j+1/2} = (j + \frac{1}{2})\Delta x$, for all $j \in \mathbb{Z}$, and $\Delta x > 0$ is a given mesh-size. We will also denote by $x_j = j\Delta x$ the intermediate points of the spatial cells.

Thus, we analyze the asymptotic behavior as $t \to \infty$ of these semi-discrete solutions u_Δ for Δx fixed.

Theorem 4.2. *Let $u_0 \in L^1(\mathbb{R})$, $\Delta x > 0$ and u_Δ be the corresponding solution of the semi-discrete scheme* (4.4) *for the augmented Burgers equation* (4.3). *For any $p \in [1, \infty]$, the following holds*

$$t^{\frac{1}{2}(1-\frac{1}{p})} \|u_\Delta(t) - u_M(t)\|_p \longrightarrow 0, \quad \text{as } t \to \infty, \tag{4.8}$$

where $u_M(t,x)$ is the unique solution of the following viscous Burgers equation:

$$\begin{cases} v_t = vv_x + (1 + F_2^\Delta)v_{xx}, & x \in \mathbb{R}, t > 0, \\ v(x,0) = M\delta_0. \end{cases}$$

Here, $M = \int_{\mathbb{R}} u^0(x)dx$ is the mass of the initial data and

$$F_2^\Delta = \frac{\Delta x^2}{2}\left(\sum_{m=1}^{N} m(m-1)\omega_m\right).$$

Let us observe that if N is taken such that $N \to \infty$ and $N\Delta x \to \infty$ when $\Delta x \to 0$, then $F_2^\Delta \to 1$, which is, precisely, the value that we should expect from the continuous model. We refer to Section 1.3.3 in Chapter 1 for the general case $\nu, c, \theta > 0$. Moreover, in Appendix B we extend this result to the full augmented Burgers equation.

Let us conclude this section by adding a final comment on the time discretization, which we do not address in this paper. At the continuous/semi-discrete level, we obtain estimates on the solution that allow us to prove the compactness of a family of rescaled solutions. Then, the asymptotic behavior is obtained as in (4.8). The analogous step for the fully time-explicit discrete scheme requires further development.

The chapter is organized as follows. In Section 4.2, we deal with the well-posedness of equation (4.3) and the asymptotical behavior of its solutions, showing that $K * u_{xx}$ behaves like u_{xx} as $t \to \infty$. In Section 4.3, we focus on the semi-discrete numerical scheme (4.4), showing its convergence and analyzing for a fixed Δx the large-time behavior of the numerical solutions. To illustrate the main results of this work, we conclude with some numerical simulations in Section 4.4.

4.2 Analysis of the augmented Burgers equation

In this section we study the well-posedness of the Cauchy problem for (4.3) with initial data in $L^1(\mathbb{R})$. We also obtain estimates in the L^p-norms of its solutions. We mainly proceed as in [29] and [58].

4.2.1 Existence and uniqueness of solutions

The following theorem concerns the global existence of solutions and specifies their regularity. Let us remark that the result coincides with the one for the classical convection-diffusion equation [29].

Theorem 4.3. *For any initial data $u^0 \in L^1(\mathbb{R})$, there exists a unique solution $u \in C([0,\infty), L^1(\mathbb{R}))$ of (4.3). Moreover, it also satisfies*

$$u \in C((0,\infty), \mathcal{W}^{2,p}(\mathbb{R})) \cap C^1((0,\infty), L^p(\mathbb{R})), \quad \forall p \in (1,\infty).$$

Additionally, equation (4.3) generates a contractive semigroup in $L^1(\mathbb{R})$.

Proof. Existence in $L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$. The local existence of the solution follows by a classical Banach fixed point argument as in [29] or [48]. To extend the solution globally, we deduce a priori estimates on the $L^1(\mathbb{R})$ and $L^\infty(\mathbb{R})$ norms of the solution. Let us first focus on the L^1-norm. Multiplying (4.3) by $\operatorname{sign}(u)$ and integrating in $\mathbb{R}$, it follows that

$$\frac{d}{dt}\int_{\mathbb{R}} |u|dx \le \int_{\mathbb{R}} (K*u-u)\operatorname{sign}(u)dx \le \int_{\mathbb{R}} Kdx \int_{\mathbb{R}} |u|dx - \int_{\mathbb{R}} |u|dx \le 0 \tag{4.9}$$

and, consequently, $\|u(t)\|_1 \le \|u^0\|_1$.

To estimate the L^∞-norm similar arguments apply. We define $\mu = \|u^0\|_\infty$, multiply equation (4.3) by $\operatorname{sign}[(u-\mu)^+]$, where $z^+ := \max\{0, z\}$, and integrate it in $\mathbb{R}$. We obtain

$$\frac{d}{dt}\int_{\mathbb{R}} (u-\mu)^+ dx \le \int_{\mathbb{R}} (K*u-u+u_x)\operatorname{sign}(u-\mu)^+ dx = \int_{\mathbb{R}} (K*u-u)\operatorname{sign}(u-\mu)^+ dx.$$

Let us now recall the following result in [48, Lemma 3.1]: for any $f \in L^1(\mathbb{R})$ and H a nonnegative function with mass one, the following holds for any $\alpha \ge 0$:

$$\int_{f(x)>\alpha}\int_{\mathbb{R}} H(x-y)f(y)dydx \le \int_{f(x)>\alpha} f(x)dx.$$

Hence, $(u-\mu)^+ \le (u^0-\mu)^+ = 0$. Thus, we conclude that $u(t) \le \mu$ almost everywhere. The same argument for $(u+\mu)^-$, where $z^- := -\max\{0,-z\}$, shows that $u \ge -\mu$. Therefore, if $u^0 \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$, then $\|u(t)\|_\infty \le \|u^0\|_\infty$ for all $t > 0$. Since both L^1-norm and L^∞-norm remain bounded in time, the solution u exists globally.

To estimate the L^∞-norm similar arguments apply. We define $\mu = \|u^0\|_\infty$, multiply equation (4.3) by $\operatorname{sign}[(u-\mu)^+]$, where $z^+ := \max\{0, z\}$, and integrate it in $\mathbb{R}$. We obtain

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{R}} (u-\mu)^+ dx &\le \int_{\mathbb{R}} (K*u-u+u_x)\operatorname{sign}(u-\mu)^+ dx \\ &= \int_{\mathbb{R}} (K*(u-\mu)-(u-\mu))\operatorname{sign}(u-\mu)^+ dx \\ &\le \int_{\mathbb{R}} K*(u-\mu)^+ - \int_{\mathbb{R}} (u-\mu)^+ \le 0. \end{aligned} \tag{4.10}$$

We conclude that $(u-\mu)^+ \le (u^0-\mu)^+ = 0$ and, consequently, $u(t) \le \mu$ almost everywhere. The same argument for $(u+\mu)^-$, where $z^- := -\max\{0,-z\}$, shows that $u \ge -\mu$. Therefore, if $u^0 \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$, then $\|u(t)\|_\infty \le \|u^0\|_\infty$ for all $t > 0$. Lastly, since both L^1-norm and L^∞-norm remain bounded in time, the solution u exists globally.
Regularity. It follows from classical regularity arguments (e.g., [57]) that

$$u \in C((0,T), W^{2,p}(\mathbb{R})) \cap C^1((0,T), L^p(\mathbb{R}))$$

for every $p \in (1,\infty)$. This also holds for $T = \infty$. Let us remark that this regularity makes the integrals in the previous steps be well defined.
Uniqueness. To prove the uniqueness of solution it is enough to check that (4.3) generates a contractive semigroup in $L^1(\mathbb{R})$; that is, for any initial datum $u^0, v^0 \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$; that is,

$$\|u(t) - v(t)\|_1 \le \|u^0 - v^0\|_1, \quad \forall t > 0, \tag{4.11}$$

where u and v are the corresponding solutions. An analogous argument as in (4.9), applied to the equation verified by $u - v$, shows

$$\frac{d}{dt}\int_{\mathbb{R}} |u-v| dx \le 0,$$

hence the contraction property in $L^1(\mathbb{R})$.
Existence and uniqueness in $L^1(\mathbb{R})$. The extension of the result to a general $u^0 \in L^1(\mathbb{R})$ can be done following the same arguments as in [29]. □

4.2.2 Decay estimates and large-time behavior

Now we obtain L^p-decay rates for the solution to (4.3). These are the same as the ones for the viscous Burgers equation [29].

Theorem 4.4. *For all $p \in [1,\infty]$, there exists a constant $C = C(p) > 0$ such that*

$$\|u(t)\|_p \le C\|u^0\|_1 t^{-\frac{1}{2}(1-\frac{1}{p})}, \quad \forall t > 0, \tag{4.12}$$

for all solutions of equation (4.3) *with initial data $u^0 \in L^1(\mathbb{R})$.*

Proof. The case $p = 1$ is an immediate consequence of Theorem 4.3. In the case $p \in [2,\infty)$, we multiply equation (4.3) by $|u|^{p-2}u$ and integrate it in $\mathbb{R}$. We obtain:

$$\begin{aligned}\frac{1}{p}\frac{d}{dt}\left(\|u\|_p^p\right) &= \int_{\mathbb{R}} |u|^{p-2}uu_t dx \\ &= \int_{\mathbb{R}} |u|^p u_x dx + \int_{\mathbb{R}} |u|^{p-2}uu_{xx}dx + \int_{\mathbb{R}} |u|^{p-2}u(K*u-u+u_x)dx \\ &= -\frac{4(p-1)}{p^2}\left\|\left(|u|^{p/2}\right)_x\right\|_2^2 - \|u\|_p^p + \int_{\mathbb{R}} |u|^{p-2}u(K*u)dx. \end{aligned} \tag{4.13}$$

Let us focus on the last term, so that we can compare it with the L^p-norm of u. Young's inequality gives us that

$$\left||u(t,x)|^{p-2}u(t,x)u(t,y)\right| = |u(t,x)|^{p-1}|u(t,y)| \leq \frac{p-1}{p}|u(t,x)|^p + \frac{1}{p}|u(t,y)|^p.$$

Thus, using that K has mass one, it follows:

$$\left|\int_{\mathbb{R}} |u|^{p-2}u(K*u)dx\right| \leq \int_{\mathbb{R}}\int_{\mathbb{R}} K(x-y)|u(t,x)|^{p-1}|u(t,y)|dydx \leq \|u\|_p^p.$$

Plugging this last estimate in (4.13) we have

$$\frac{d}{dt}\left(\|u(t)\|_p^p\right) + \frac{4(p-1)}{p}\left\|\left(|u(t)|^{p/2}\right)_x\right\|_2^2 \leq 0. \tag{4.14}$$

Finally, with the same arguments as in [29] we obtain the desired estimate (4.12) for any $p \in [2,\infty)$, as well as for $p=\infty$. The case $p \in (1,2)$ follows by applying Hölder's inequality and (4.12) with $p=1$ and $p=2$. □

Similar estimates can be found for the derivative of the solution of (4.3). Let us define the re-scaled function u^λ, which will also be used in the following section to obtain the asymptotic profile. For $\lambda > 0$ we define

$$u^\lambda(t,x) = \lambda u(\lambda^2 t, \lambda x). \tag{4.15}$$

The scales are the same as for the Burgers or heat equations. Clearly, u^λ is the solution of the following equation:

$$\begin{cases} u_t^\lambda = u^\lambda\ u_x^\lambda + u_{\lambda,xx} + \lambda^2(K_\lambda * u^\lambda - u^\lambda) + \lambda u_x^\lambda, & (t,x) \in (0,\infty)\times\mathbb{R}, \\ u^\lambda(0,x) = u_0^\lambda(x) = \lambda u^0(\lambda x), & x \in \mathbb{R}, \end{cases} \tag{4.16}$$

where $K_\lambda(z) = \lambda K(\lambda z)$, $z \in \mathbb{R}$.

Theorem 4.5. *For each $p \in [1,\infty]$, there exists a constant $C = C(p) > 0$, such that the solution of equation (4.3) satisfies*

$$\|u_x(t)\|_p \leq C\|u^0\|_1 t^{-\frac{1}{2}(1-\frac{1}{p})-\frac{1}{2}}, \quad \forall t > 0. \tag{4.17}$$

Proof. First, let us denote by D_λ^t the semigroup associated to the linear problem

$$\begin{cases} v_t = \lambda^2(K_\lambda * v - v) + \lambda v_x, & (t,x) \in (0,\infty)\times\mathbb{R}, \\ v(0,x) = v^0(x), & x \in \mathbb{R}. \end{cases}$$

It is immediate that D_λ^t is stable in $L^1(\mathbb{R})$:

$$\frac{d}{dt}\int_{\mathbb{R}} |v| dx = \lambda^2 \int_{\mathbb{R}} (K_\lambda * v - v)\,\mathrm{sign}(v) dx \le 0.$$

On the other hand, for all $\tau > 0$, function u^λ solution of (4.16) verifies the following integral equation:

$$u^\lambda(t+\tau) = G(t) * D_\lambda^t u^\lambda(\tau) + \int_0^t G(t-s) * D_\lambda^s \left(\left(\frac{u^\lambda(s+\tau)^2}{2} \right)_x \right) ds,$$

where $G(t)$ is the heat kernel, given by

$$G(t,x) = (4\pi t)^{-1/2} e^{-\frac{x^2}{4t}}.$$

If we differentiate it with respect to x, we obtain:

$$u_x^\lambda(t+\tau) = G_x(t) * D_\lambda^t u^\lambda(\tau) + \int_0^t G_x(t-s) * D_\lambda^s \left(\left(\frac{u^\lambda(s+\tau)^2}{2} \right)_x \right) ds.$$

Now, let us first estimate the L^1 norm of $u_x^\lambda(t+\tau)$. Note that there exists a constant $C = C(p)$ such that

$$\|G_x(t)\|_p \le C t^{-\frac{1}{2}(1-\frac{1}{p})-\frac{1}{2}}, \quad \forall t > 0.$$

Moreover, by Theorem 4.4 we have

$$\|u^\lambda(\tau)\|_1 = \|u(\lambda^2\tau)\|_1 \le \|u^0\|_1$$

and

$$\|u^\lambda(s+\tau)\|_\infty = \lambda \|u(\lambda^2(s+\tau))\|_\infty \le C\tau^{-\frac{1}{2}} \|u^0\|_1.$$

Thus, it follows:

$$\begin{aligned} \|u_x^\lambda(t+\tau)\|_1 &\le \|G_x(t)\|_1 \|D_\lambda^t u^\lambda(\tau)\|_1 + \int_0^t \|G_x(t-s)\|_1 \left\| D_\lambda^s \left(\left(\tfrac{u^\lambda(s+\tau)^2}{2} \right)_x \right) \right\|_1 ds \\ &\le C t^{-\frac{1}{2}} \|u^\lambda(\tau)\|_1 + C \int_0^t (t-s)^{-\frac{1}{2}} \|(u^\lambda(s+\tau)^2)_x\|_1 ds \\ &\le C t^{-\frac{1}{2}} \|u^0\|_1 + C\tau^{-1/2} \|u^0\|_1 \int_0^t (t-s)^{-\frac{1}{2}} \|u_x^\lambda(s+\tau)\|_1 ds. \end{aligned}$$

Applying Gronwall's Lemma, we deduce for $t = \tau$ that

$$\|u_x^\lambda(2\tau)\|_1 \leq C_\tau, \quad \forall \lambda > 0, \tag{4.18}$$

for some constant $C_\tau > 0$ that only depends on τ and $\|u^0\|_1$. This is equivalent to (4.17) for $p = 1$.

The case $p \in (1, \infty)$ is an immediate consequence of (4.18):

$$\begin{aligned}
\|u_x^\lambda(t+\tau)\|_p &\leq \|G_x(t)\|_p \|D_\lambda^t u^\lambda(\tau)\|_1 + \int_0^t \|G_x(t-s)\|_p \left\| D_\lambda^s \left(\left(\tfrac{u^\lambda(s+\tau)^2}{2} \right)_x \right) \right\|_1 ds \\
&\leq Ct^{-\frac{1}{2}(1-\frac{1}{p})-\frac{1}{2}} \|u^\lambda(\tau)\|_1 + C \int_0^t (t-s)^{-\frac{1}{2}(1-\frac{1}{p})-\frac{1}{2}} \|(u^\lambda(s+\tau)^2)_x\|_1 ds \\
&\leq Ct^{-\frac{1}{2}(1-\frac{1}{p})-\frac{1}{2}} \|u^0\|_1 + C_\tau \int_0^t (t-s)^{-\frac{1}{2}(1-\frac{1}{p})-\frac{1}{2}} \|u_x^\lambda(s+\tau)\|_1 ds \\
&\leq Ct^{-\frac{1}{2}(1-\frac{1}{p})-\frac{1}{2}} \|u^0\|_1 + C_\tau \int_0^t (t-s)^{-\frac{1}{2}(1-\frac{1}{p})-\frac{1}{2}} ds
\end{aligned}$$

Taking $t = \tau$, we conclude that

$$\|u_x^\lambda(2\tau)\|_p \leq C_\tau, \quad \forall \lambda > 0,$$

which is equivalent to (4.17) for $p \in (1, \infty)$.

Finally, we repeat a similar argument for the case $p = \infty$:

$$\begin{aligned}
\|u_x^\lambda(t+\tau)\|_\infty &\leq \|G_x(t)\|_\infty \|D_\lambda^t u^\lambda(\tau)\|_1 + \int_0^t \|G_x(t-s)\|_q \left\| D_\lambda^s \left(\left(\tfrac{u^\lambda(s+\tau)^2}{2} \right)_x \right) \right\|_{q'} ds \\
&\leq Ct^{-1} \|u^\lambda(\tau)\|_1 + C \int_0^t (t-s)^{-\frac{1}{2}(1-\frac{1}{q})-\frac{1}{2}} \|(u^\lambda(s+\tau)^2)_x\|_{q'} ds \\
&\leq Ct^{-1} \|u^0\|_1 + C_\tau \|u^0\|_1 \int_0^t (t-s)^{-\frac{1}{2}(1-\frac{1}{q})-\frac{1}{2}} \|u_x^\lambda(s+\tau)\|_{q'} ds \\
&\leq Ct^{-1} \|u^0\|_1 + C_\tau \|u^0\|_1 \int_0^t (t-s)^{-\frac{1}{2}(1-\frac{1}{q})-\frac{1}{2}} ds
\end{aligned}$$

where $q \in (1, \infty)$ and $1/q + 1/q' = 1$. It is now enough to take $t = \tau$ to conclude the proof. □

4.2.3 Large-time behavior

The decay rates of the previous section will allow us to obtain the asymptotic profile of solutions for (4.3). The aim is to compute the limit $\lambda \to \infty$ in (4.16), which is equivalent to taking the limit $t \to \infty$ in (4.3).

Let us first observe that the estimates in Theorem 4.4 and Theorem 4.5 are also valid for u^λ defined in (4.15). The mass is conserved too. We state this in the following lemma.

Lemma 4.6. *For each $p \in [1,\infty]$, there exists a constant $C = C(p, \|u^0\|_1) > 0$ such that, for all $\lambda > 0$, the solution of (4.16) satisfies*

$$\|u^\lambda(t)\|_p \le C\|u^0\|_1 t^{-\frac{1}{2}(1-\frac{1}{p})} \quad \text{and} \quad \|u^\lambda_x(t)\|_p \le C\|u^0\|_1 t^{-\frac{1}{2}(1-\frac{1}{p})-\frac{1}{2}}, \quad \forall t > 0.$$

Moreover, the mass of u^λ is conserved in time.

Proof. We just have to use the definition of u^λ in (4.15) and apply Theorem 4.4. For all $t > 0$ and $\lambda > 0$ we have

$$\|u^\lambda(t)\|_p = \lambda^{1-\frac{1}{p}}\|u(\lambda^2 t)\|_p \le C t^{-\frac{1}{2}(1-\frac{1}{p})}.$$

Same procedure applies to u^λ_x, concerning Theorem 4.5. Regarding the last result, it is easy to see that:

$$\int_{\mathbb{R}} u^\lambda(t,x)dx = \int_{\mathbb{R}} u(\lambda^2 t, x)dx = \int_{\mathbb{R}} u^0(x)dx$$

which proves the mass conservation. □

In particular, this lemma implies that, for any finite time interval $[\tau, T]$ with $0 < \tau < T < \infty$, the set $\{u^\lambda\}_{\lambda>0}$ is uniformly bounded in $L^\infty([\tau,T], L^p(\mathbb{R}))$.

4.2.3.1 Compactness of the family $\{u^\lambda\}_{\lambda>0}$

As we said at the beginning, we would like to pass to the limit $\lambda \to \infty$. We need the following theorem due to J. Simon ([88]), as an extension of the Aubin-Lions Lemma, to assure the compactness of the set $\{u^\lambda\}_{\lambda>0}$.

Theorem 4.7 ([88, Theorem 5])**.** *Let X, Z and Y be Banach spaces satisfying $X \subset Z \subset Y$ with compact embedding $X \subset Z$. Assume, for $p \in [1,\infty]$ and $T > 0$, that F is bounded in $L^p(0,T;X)$ and $\{\partial_t f : f \in F\}$ is bounded in $L^p(0,T;Y)$. Then, F is relatively compact in $L^p(0,T;Z)$ and, in the case of $p = \infty$, also in $\mathcal{C}(0,T;Z)$.*

Applying this result we can prove the following theorem regarding the relative compactness of the set $\{u^\lambda\}_{\lambda>0}$. In the sequel, for any functions f and g, we denote $f \lesssim g$ if there exists a constant $C > 0$, not depending on the scaling parameter nor the time, such that $f \le Cg$.

Theorem 4.8. *For every $0 < \tau < T < \infty$, the set $\{u^\lambda\}_{\lambda>0} \subset C([\tau,T],L^1(\mathbb{R}))$ is relatively compact.*

Proof. Step 1. First, for any $r > 0$ we will show that it is relatively compact in $C([\tau,T],L^2(I))$, with $I=[-r,r]$. Let us consider the spaces $X = H^1(I)$, $Z = L^2(I)$ and $Y = H^{-1}(I)$. We would like to apply Theorem 4.7 to the set $F = \{u^\lambda\}_{\lambda>0}$.

From Lemma 4.6 we know that the sets $\{u^\lambda\}_{\lambda>0}$ and $\{u^\lambda_x\}_{\lambda>0}$ are bounded in $L^\infty([\tau,T],L^2(I))$. In particular, first condition on F is fulfilled. Therefore, it suffices to check that u^λ_t is bounded in $L^\infty([\tau,T],H^{-1}(I))$. Using (4.16), for every $\varphi \in \mathcal{C}^\infty_c(I)$, we have:

$$\begin{aligned}\left|\int_{\mathbb{R}} u^\lambda_t\varphi dx\right| &\le \left|\int_{\mathbb{R}} ((u^\lambda)^2)_x\varphi dx\right| + \left|\int_{\mathbb{R}} u_{\lambda,xx}\varphi dx\right| \\ &\quad + \left|\int_{\mathbb{R}} \big(\lambda^2(K_\lambda * u^\lambda - u^\lambda) + \lambda u^\lambda_x\big)\varphi dx\right| \\ &\lesssim \|\varphi_x\|_2\|u^\lambda\|_4^2 + \|\varphi_x\|_2\|u^\lambda_x\|_2 + \left|\int_{\mathbb{R}} \big(\lambda^2(K_\lambda * u^\lambda - u^\lambda) + \lambda u^\lambda_x\big)\varphi dx\right|.\end{aligned} \tag{4.19}$$

Obviously, the first and second terms on the right hand side of (4.19) are uniformly bounded in $[\tau,T]$, so let us focus on the third one:

$$\begin{aligned}\mathcal{I}_\lambda &= \left|\int_{\mathbb{R}} \big(\lambda^2(K_\lambda * u^\lambda - u^\lambda) + \lambda u^\lambda_x\big)\varphi dx\right| \\ &= \left|\int_{\mathbb{R}} \Big(\lambda^2\big(\widehat{K}(\xi/\lambda) - 1\big) + i\lambda\xi\Big)\widehat{u^\lambda}(\xi)\widehat{\varphi}(\xi)d\xi\right|.\end{aligned}$$

Let us denote

$$m_\lambda(\xi) = \lambda^2\Big(\widehat{K}(\frac{\xi}{\lambda}) - 1\Big) + i\lambda\xi.$$

We claim that

$$|m_\lambda(\xi)| \le \xi^2, \quad \forall \xi \in \mathbb{R}, \forall \lambda > 0. \tag{4.20}$$

Using the Cauchy-Schwartz inequality, we have:

$$\mathcal{I}_\lambda = \left|\int_{\mathbb{R}} m_\lambda(\xi)\widehat{u^\lambda}(\xi)\widehat{\varphi}(\xi)d\xi\right| \lesssim \|\varphi\|_{H^1(\mathbb{R})}\left\|u^\lambda\right\|_{H^1(\mathbb{R})}. \tag{4.21}$$

Hence, going back to (4.19) and replacing (4.21), we obtain

$$\left|\int_{\mathbb{R}} u^\lambda_t\varphi dx\right| \lesssim \|\varphi\|_{H^1(\mathbb{R})}\left(\|u^\lambda\|_4^2 + \left\|u^\lambda\right\|_{H^1(\mathbb{R})}\right).$$

By Lemma 4.6, all the quantities in the right-hand side are uniformly bounded in $[\tau,T]$. Consequently, the set $\{u^\lambda\}_{\lambda>0}$ is relatively compact in $C([\tau,T],L^2(I))$.

It remains to prove claim (4.20). Observe that

$$\widehat{K}(\xi) = \frac{1 - i\xi}{1 + \xi^2}, \quad \xi \in \mathbb{R}, \tag{4.22}$$

and, therefore,

$$|m_\lambda(\xi)| = \left| \lambda^2 \left(\frac{1 - i\xi/\lambda}{1 + (\xi/\lambda)^2} - 1 \right) + i\lambda\xi \right| = \frac{\lambda\xi^2}{\sqrt{\lambda^2 + \xi^2}} \leq \xi^2, \quad \forall \lambda > 0.$$

Step 2. The next step consists in proving the compactness in $C([\tau, T], L^1(I))$. Since $L^2(I)$ is continuously embedded in $L^1(I)$, the compactness in $C([\tau, T], L^2(I))$ is clearly transferred to $C([\tau, T], L^1(I))$.

Step 3. Now we need to extend the result to $C([\tau, T], L^1(\mathbb{R}))$. We do that by proving uniform, with respect to λ, estimates on the tails of u^λ.

For every $r > 0$, let us define function $\psi_r(z) = \psi(z/r)$, where ψ is a nonnegative $C^\infty(\mathbb{R})$ function such that

$$\psi(z) = \begin{cases} 0, & |z| < 1, \\ 1, & |z| > 2. \end{cases} \tag{4.23}$$

Since $\{u^\lambda\}_{\lambda>0}$ is relatively compact in $C([\tau, T], L^1(I))$, it suffices to show that

$$\sup_{t\in[\tau,T]} \|u^\lambda(t)\psi_r\|_1 \longrightarrow 0 \quad \text{as } r \to \infty, \text{ uniformly for } \lambda > 0. \tag{4.24}$$

We first observe that it is enough to consider nonnegative initial data. Indeed, the same argument as in Theorem 4.3 shows that for any $u_0, v_0 \in L^1(\mathbb{R})$ the corresponding solutions u^λ, v^λ to (4.16) satisfy $\|u_\lambda - v_\lambda\|_1 \leq \|u_0 - v_0\|_1$. As a consequence, due to Lemma 4.6 and Crandall-Tartar Lemma (see, for instance, [38, Chapter II]), we know that $u \leq v$ if $u_0 \leq v_0$. Thus, choosing $v_0 = |u_0|$ and $w_0 = -|u_0|$ as initial data implies that $|u_\lambda(t, x)| \leq |v_\lambda(t, x)| + |w_\lambda(t, x)|$, where u_λ, v_λ and w_λ are the solutions corresponding to u_0, v_0 and w_0 respectively. In conclusion, it is sufficient to show (4.24) for nonnegative initial data and solutions.

Let us assume that u_λ is a nonnegative solution. We multiply (4.16) by ψ_r and integrate it over $(0, t) \times \mathbb{R}$. We obtain:

$$\begin{aligned} \int_0^t \int_{\mathbb{R}} u_s^\lambda \psi_r dx ds = -\frac{1}{2} \int_0^t \int_{\mathbb{R}} (u^\lambda)^2 \psi_r' dx ds + \int_0^t \int_{\mathbb{R}} u^\lambda \psi_r'' dx ds \\ + \int_0^t \int_{\mathbb{R}} \Big(\lambda^2 (K_\lambda * u^\lambda - u^\lambda) + \lambda u_x^\lambda \Big) \psi_r dx ds. \end{aligned}$$

and, therefore,

$$\int_{\mathbb{R}} u^\lambda(t)\psi_r dx \leq \int_{\mathbb{R}} u_0^\lambda \psi_r dx + \frac{\|\psi'\|_\infty}{2r}\int_0^t \|u^\lambda(s)\|_2^2 ds + \frac{\|\psi''\|_\infty}{r^2}\int_0^t \|u^\lambda(s)\|_1 ds + \int_0^t \int_{\mathbb{R}} \Big(\lambda^2\big(K_\lambda * u^\lambda(s) - u^\lambda(s)\big) + \lambda u_x^\lambda(s)\Big)\psi_r dx ds. \tag{4.25}$$

We have to obtain an estimate on the last term in the integral, uniformly on λ. Let us denote

$$I = \int_{\mathbb{R}} \Big(\lambda^2\big(K_\lambda * u^\lambda(s) - u^\lambda(s)\big) + \lambda u_x^\lambda(s)\Big)\psi_r dx.$$

A change of variables and integration by parts give us that

$$\begin{aligned} I &= \lambda^2 \int_{\mathbb{R}}\int_{\mathbb{R}} K(x-y)u(\lambda^2 s, y)\psi_{\lambda r}(x) dy dx \\ &\quad - \lambda^2 \int_{\mathbb{R}} u(\lambda^2 s, x)\psi_{\lambda r}(x) dx - \lambda^2 \int_{\mathbb{R}} u(\lambda^2 s, x)\psi'_{\lambda r}(x) dx \\ &= \lambda^2 \int_{\mathbb{R}} u(\lambda^2 s, y)\Big(\int_{\mathbb{R}} K(x-y)\big(\psi_{\lambda r}(x) - \psi_{\lambda r}(y) - (x-y)\psi'_{\lambda r}(y)\big)dx\Big)dy \\ &\leq \lambda^2 \|u^0\|_1 \|\psi''_{\lambda r}\|_\infty \leq \frac{\|u_0\|_1 \|\psi''\|_\infty}{r^2} \end{aligned} \tag{4.26}$$

Remark 4.1*: Note that the first moment of K plays an important role here. The fact that*

$$\int_{\mathbb{R}} K(z)dz = \int_{\mathbb{R}} zK(z)dz = 1$$

is critical to be able to find a bound for I and, hence, to show that $K_\lambda * u_{\lambda,xx} \to u_{xx}$ as $t \to \infty$. This is also much related with the decomposition of K in Dirac delta functions as in [25]. Moreover, taking into account this identity at the numerical level will be essential to preserve the large-time behavior correctly. As a matter of fact, the correcting factors F_0^Δ and F_1^Δ are related to this observation.

So, plugging (4.26) into (4.25) and using Theorem 4.4, we get:

$$\int_{\mathbb{R}} \Big|u^\lambda(t)\Big|\,\psi_r dx \leq \int_{\mathbb{R}} |u_0|\,\psi_{\lambda r} dx + C\left(\frac{\sqrt{t}}{r} + \frac{t}{r^2}\right)$$

where $C > 0$ depends only on $\|u^0\|_1$ and $\|\psi\|_{W^{2,\infty}(\mathbb{R})}$, which are both bounded. For $\lambda > 1$, since $\psi_r(x) > \psi_{\lambda r}(x)$, we get

$$\int_{\mathbb{R}} \Big|u^\lambda(t,x)\Big|\,\psi_r(x) dx \leq \int_{\mathbb{R}} |u_0(x)|\,\psi_r(x) dx + C\left(\frac{\sqrt{t}}{r} + \frac{t}{r^2}\right),$$

which tends to zero uniformly on λ when $r \to \infty$. Thus, we proved (4.24) and, consequently, $\{u^\lambda\}_{\lambda>0}$ is relatively compact in $C([\tau, T], L^1(\mathbb{R}))$. □

Modifying slightly the previous proof, we can also conclude the following lemma, regarding the initial condition u_0^λ:

Lemma 4.9. *For every test function $\varphi \in C_c^\infty(\mathbb{R})$, there exists a positive constant $C = C(\varphi, u^0)$, such that*

$$\left| \int_{\mathbb{R}} u^\lambda(t,x)\varphi(x)dx - \int_{\mathbb{R}} u_0^\lambda(x)\varphi(x)dx \right| \leq C(t+\sqrt{t}), \quad \forall t > 0,$$

holds uniformly on $\lambda > 0$.

Proof. We multiply (4.16) by $\varphi \in C_c^\infty(\mathbb{R})$ and integrate it over $(0,t)\times\mathbb{R}$. We get:

$$\int_0^t \int_{\mathbb{R}} u_t^\lambda \varphi = \int_0^t \int_{\mathbb{R}} u^\lambda u_x^\lambda \varphi + \int_0^t \int_{\mathbb{R}} u_{\lambda,xx}\varphi + \int_0^t \int_{\mathbb{R}} \big(\lambda^2(K_\lambda * u^\lambda - u^\lambda) + \lambda u_x^\lambda\big)\varphi.$$

Integrating by parts and making use of Lemma 4.6, we have

$$\begin{aligned}\left| \int_{\mathbb{R}} u^\lambda(t)\varphi dx - \int_{\mathbb{R}} u_0^\lambda \varphi dx \right| \leq \frac{\|\varphi'\|_\infty}{2} \int_0^t \|u^\lambda(s)\|_2^2 ds + \|\varphi''\|_\infty \int_0^t \|u^\lambda(s)\|_1 ds \\ + \left| \int_0^t \int_{\mathbb{R}} \Big(\lambda^2\big(K_\lambda * u^\lambda(s) - u^\lambda(s)\big) + \lambda u_x^\lambda(s)\Big)\varphi dx ds \right|.\end{aligned}$$

To conclude the proof, it is enough to apply a similar argument as for (4.25) to get:

$$\begin{aligned}\left| \int_{\mathbb{R}} u^\lambda(t)\varphi dx - \int_{\mathbb{R}} u_0^\lambda \varphi dx \right| &= \left| \int_{\mathbb{R}} u(\lambda^2 t, x)\varphi(\frac{x}{\lambda})dx - \int_{\mathbb{R}} u^0(x)\varphi(\frac{x}{\lambda})dx \right| \\ &=\leq C(\|\varphi\|_{W^{2,\infty}(\mathbb{R})}, \|u^0\|_1)(\sqrt{t}+t).\end{aligned}$$

□

4.2.3.2 Passing to the limit

Now we have all the ingredients that we need to prove our main result on the large-time behavior of solutions to problem (4.3), stated in Theorem 4.1.

Proof of Theorem 4.1. By Theorem 4.8, we know that for every $0 < \tau < T < \infty$, the family $\{u^\lambda\}_{\lambda>0}$ is relatively compact in $C([\tau,T], L^1(\mathbb{R}))$. Consequently, there exists a subsequence of it (which we will not relabel) and a function $\bar{u} \in C((0,\infty), L^1(\mathbb{R}))$ such that

$$u^\lambda \longrightarrow \bar{u} \in C([\tau,T], L^1(\mathbb{R})), \quad \text{as } \lambda \to \infty. \tag{4.27}$$

We can also assume that $u^\lambda(t,x) \to \bar{u}(t,x)$ almost everywhere in $(0,\infty)\times\mathbb{R}$ as $\lambda \to \infty$.

Our claim is that, passing to the limit $\lambda \to \infty$, we obtain that $\bar{u}$ is a weak solution of the equation:

$$\begin{cases} \bar{u}_t = \bar{u}\bar{u}_x + 2\bar{u}_{xx}, & (t,x) \in (0,\infty) \times \mathbb{R}, \\ \bar{u}(0) = M\delta_0. \end{cases} \tag{4.28}$$

Let us multiply equation (4.16) by a test function $\phi \in C_c^\infty((0,\infty) \times \mathbb{R})$ and integrate it over $(0,\infty) \times \mathbb{R}$. We have:

$$\int_0^\infty \int_{\mathbb{R}} u_t^\lambda \phi = \int_0^\infty \int_{\mathbb{R}} u^\lambda \, u_x^\lambda \phi + \int_0^\infty \int_{\mathbb{R}} u_{\lambda,xx}\phi + \int_0^\infty \int_{\mathbb{R}} (\lambda^2(K_\lambda * u^\lambda - u^\lambda) + \lambda u_x^\lambda)\phi.$$

Using the properties of $\{u^\lambda\}_{\lambda>0}$ shown in the previous section, it is sufficient to check that

$$\lim_{\lambda\to\infty} \int_0^\infty \int_{\mathbb{R}} \Big(\lambda^2\big(K_\lambda * u^\lambda(t) - u^\lambda(t)\big) + \lambda u_x^\lambda(t)\Big)\phi(t)dxdt = \int_0^\infty \int_{\mathbb{R}} \bar{u}(t)\phi_{xx}(t)dxdt.$$

Let us focus on the integral over the spatial domain. Taking into account the definition of K_λ and that $\int_{\mathbb{R}} z^m K(z)dz = m!$ for $m \in \mathbb{N} \cup \{0\}$:

$$\begin{aligned} \mathcal{L}_\lambda(t) &= \int_{\mathbb{R}} \Big(\lambda^2\big(K_\lambda * u^\lambda(t) - u^\lambda(t)\big) + \lambda u_x^\lambda(t)\Big)\phi(t)dx \\ &= \lambda^2 \int_{\mathbb{R}} \int_{\mathbb{R}} \Big(\phi(t, x + \frac{y}{\lambda}) - \phi(t,x)\Big)K(y)u^\lambda(t,x)dydx \\ &\quad - \lambda \int_{\mathbb{R}} \int_{\mathbb{R}} \phi_x(t,x)yK(y)u^\lambda(t,x)dydx, \end{aligned} \tag{4.29}$$

Now, because of Taylor's Theorem, we know that there exists a point $\zeta \in (x, x + y/\lambda)$ such that

$$\phi(t, x + \frac{y}{\lambda}) - \phi(t,x) = \frac{y}{\lambda}\phi_x(t,x) + \frac{1}{2}\frac{y^2}{\lambda^2}\phi_{xx}(t,x) + \frac{1}{6}\frac{y^3}{\lambda^3}\phi_{xxx}(t,\zeta).$$

We introduce this in (4.29):

$$\begin{aligned} \mathcal{L}_\lambda(t) &= \frac{1}{2}\int_{\mathbb{R}} \phi_{xx}(t,x)u^\lambda(t,x)dx \int_{\mathbb{R}} y^2K(y)dy \\ &\quad + O(\|\phi_{xxx}(t)\|_\infty)\frac{1}{6\lambda}\int_{\mathbb{R}} u^\lambda(t,x)dx \int_{\mathbb{R}} y^3K(y)dy \\ &= \int_{\mathbb{R}} \phi_{xx}(t,x)u^\lambda(t,x)dx + O(\|\phi_{xxx}(t)\|_\infty)\frac{1}{\lambda}\int_{\mathbb{R}} u^\lambda(t,x)dx \\ &= \int_{\mathbb{R}} \phi_{xx}(t,x)u^\lambda(t,x)dx + \lambda^{-1}O(\|\phi_{xxx}(t)\|_\infty). \end{aligned}$$

Since $u^\lambda(t) \to \bar{u}(t)$ in $C([\tau,T], L^1(\mathbb{R}))$, we obtain that

$$\lim_{\lambda\to\infty} \int_0^\infty \mathcal{L}_\lambda(t)dt = \int_0^\infty \int_{\mathbb{R}} \phi_{xx}(t,x)\bar{u}(t,x)dxdt.$$

It follows that $\bar{u}$ satisfies

$$-\int_0^\infty \int_{\mathbb{R}} \bar{u}\phi_t = -\frac{1}{2}\int_0^\infty \int_{\mathbb{R}} \bar{u}^2\phi_x + 2\int_0^\infty \int_{\mathbb{R}} \bar{u}\phi_{xx}.$$

It remains to identify the behavior of $\bar{u}$ as $t \to 0$. From Lemma 4.9, for any $\varphi \in C_c^\infty(\mathbb{R})$ we have

$$\left| \int_{\mathbb{R}} u^\lambda(t,x)\varphi(x)dx - \int_{\mathbb{R}} u_0^\lambda(x)\varphi(x)dx \right| \leq C(t+\sqrt{t})$$

and, due to (4.27), we deduce

$$\left| \int_{\mathbb{R}} \bar{u}(t,x)\varphi(x)dx - M\varphi(0) \right| \leq C(t+\sqrt{t})$$

by letting $\lambda \to \infty$. Passing to the limit $t \to 0$ and using classical approximation arguments, we conclude that $\bar{u}(0) = M\delta_0$ in the sense of bounded measures.

Therefore, we can finally conclude that $\bar{u}$ is the unique solution u_M of (4.28), and that, indeed, the whole family $\{u^\lambda\}_{\lambda>0}$ converges to u_M in $C((0,\infty), L^1(\mathbb{R}))$. In particular, we have:

$$\lim_{\lambda\to\infty} \|u^\lambda(1) - u_M(1)\|_1 = 0.$$

Setting $\lambda = \sqrt{t}$ and using the self-similar form of u_M, explicitly defined in (1.14), we obtain that

$$\lim_{t\to\infty} \|u(t) - u_M(t)\|_1 = 0. \tag{4.30}$$

Finally, the convergence in the L^p-norms for $p \in (1,\infty)$ follows from (4.30), the decay estimate given in Lemma 4.6 for $p = \infty$ and the Hölder inequality. In fact, we have:

$$\|u(t) - u_M(t)\|_p \leq \left(\|u(t)\|_\infty + \|u_M(t)\|_\infty\right)^{1-\frac{1}{p}} \|u(t) - u_M(t)\|_1^{\frac{1}{p}} \leq o(t^{-\frac{1}{2}(1-\frac{1}{p})}). \tag{4.31}$$

In the case of the L^∞-norm, we use the decay of $u_x(t)$ given by Theorem 4.5 and the estimate $\|u_{M,x}(t)\|_2 \lesssim t^{-\frac{3}{4}}$, resulting from (1.14). Using the Gagliardo-Nirenberg-Sobolev inequality and (4.31), we obtain:

$$\|u(t) - u_M(t)\|_\infty \lesssim \left(\|u_x(t)\|_2 + \|u_{M,x}(t)\|_2\right)^{\frac{1}{2}} \|u(t) - u_M(t)\|_2^{\frac{1}{2}} \leq o(t^{-\frac{1}{2}}). \tag{4.32}$$

The proof is now finished. □

4.3 Semi-discrete scheme

In this section, we focus on the semi-discrete numerical scheme for equation (4.3), defined in (4.4). In order to prove Theorem 4.2, we need some preliminary results on the decay

of u_Δ similar to those obtained in Section 4.2 for the solution of equation (4.3). For simplicity, for every $h > 0$, we define the operators d_h^+ and d_h^- as follows:

$$d_h^+ f(x) := \frac{f(x+h) - f(x)}{h} \quad \text{and} \quad d_h^- f(x) := \frac{f(x) - f(x-h)}{h}.$$

As in the continuous case, for $\mu > 0$ we also introduce the family of rescaled solutions

$$u^\mu(t,x) = \mu u_\Delta(\mu^2 t, \mu x), \quad t \geq 0, x \in \mathbb{R}, \tag{4.33}$$

and analyze the behavior of u^μ when $\mu \to \infty$. Note that function u^μ is piecewise constant on space intervals of length $\Delta x/\mu$. Moreover, it satisfies the following system:

$$\begin{cases} u_t^\mu(t,x) = \frac{1}{4}\Big(d^+_{\Delta x/\mu}\big(u^\mu(t,x)^2\big) + d^-_{\Delta x/\mu}\big(u^\mu(t,x)^2\big)\Big) \\ \qquad\qquad +\Delta x\, d^+_{\Delta x/\mu} R\big(u^\mu(t,x-\frac{\Delta x}{\mu}), u^\mu(t,x)\big) + d^-_{\Delta x/\mu}\Big(d^+_{\Delta x/\mu} u^\mu(t,x)\Big) \\ \qquad\qquad +\mu^2 \sum_{m=1}^{N} \omega_m u^\mu(t, x - m\frac{\Delta x}{\mu}) - \mu^2 F_0^\Delta u^\mu(t,x) + \mu F_1^\Delta d^+_{\Delta x/\mu} u^\mu(t,x), \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad t > 0, \text{ a.e. } x \in \mathbb{R}, \\ u^\mu(0,x) = \mu u_\Delta^0(\mu x), \qquad\qquad\qquad\qquad\qquad\qquad \text{a.e. } x \in \mathbb{R}, \end{cases} \tag{4.34}$$

where

$$R(u,v) = \frac{1}{4\Delta x}(v|v| - u|u|). \tag{4.35}$$

Of course, the approximated solution u_Δ defined in (4.7) satisfies (4.34) for $\mu = 1$.

4.3.1 L^1-L^p estimates

We are interested in the large-time behavior of u_Δ. The following two propositions are the discrete versions of Theorems 4.4 and 4.5. The way of proceeding will be, indeed, very similar.

Proposition 4.10. *For all $p \in [1,\infty]$ there exists a constant $C = C(p) > 0$ such that:*

$$\|u^\mu(t)\|_p \leq C\|u_\Delta^0\|_1 t^{-\frac{1}{2}(1-\frac{1}{p})}, \quad \forall t > 0. \tag{4.36}$$

for all solutions of (4.34) with initial data $u_\Delta^0 \in L^1(\mathbb{R})$.

Proof. Let us consider first the case $\mu = 1$ and $p \in [2, \infty)$. We multiply (4.34) by $|u_\Delta|^{p-2}u_\Delta$ and integrate it over the whole space domain. We have:

$$\frac{1}{p}\frac{d}{dt}\|u_\Delta(t)\|_p^p \le I_1 + \int_{\mathbb{R}} d_{\Delta x}^- \left(d_{\Delta x}^+ u_\Delta(x)\right)|u_\Delta(x)|^{p-2}u_\Delta(x)dx + I_2, \tag{4.37}$$

where

$$\begin{aligned} I_1 = \frac{1}{4}\int_{\mathbb{R}} \Big(d_{\Delta x}^+\big(u_\Delta(t,x)^2\big) + d_{\Delta x}^-\big(u_\Delta(t,x)^2\big)\Big)|u_\Delta(x)|^{p-2}u_\Delta(x)dx \\ + \Delta x \int_{\mathbb{R}} d_{\Delta x}^+ R\big(u_\Delta(t,x-\Delta x), u_\Delta(t,x)\big)|u_\Delta(x)|^{p-2}u_\Delta(x)dx \end{aligned}$$

and

$$\begin{aligned} I_2 &= \int_{\mathbb{R}} \left(\sum_{m=1}^N \omega_m u_\Delta(x-m\Delta x) - F_0^\Delta u_\Delta(x)\right)|u_\Delta(x)|^{p-2}u_\Delta(x)dx \\ &\quad + \int_{\mathbb{R}} F_1^\Delta d_{\Delta x}^+ u_\Delta(x)|u_\Delta(x)|^{p-2}u_\Delta(x)dx \\ &= \sum_{m=1}^N \omega_m \left(\int_{\mathbb{R}} u_\Delta(x-m\Delta x)|u_\Delta(x)|^{p-2}u_\Delta(x)dx - \int_{\mathbb{R}}|u_\Delta(x)|^p dx\right) \\ &\quad + \frac{F_1^\Delta}{\Delta x}\left(\int_{\mathbb{R}} u_\Delta(x+\Delta x)|u_\Delta(x)|^{p-2}u_\Delta(x)dx - \int_{\mathbb{R}}|u_\Delta(x)|^p dx\right). \end{aligned}$$

On the one hand, for any $k \in \mathbb{Z}$, we know that

$$\begin{aligned} &\int_{\mathbb{R}} u_\Delta(x+k\Delta x)|u_\Delta(x)|^{p-2}u_\Delta(x)dx \\ &\qquad \le \frac{p-1}{p}\int_{\mathbb{R}}|u_\Delta(x+k\Delta x)|^p dx + \frac{1}{p}\int_{\mathbb{R}}|u_\Delta(x)|^p dx = \int_{\mathbb{R}}|u_\Delta(x)|^p dx. \end{aligned}$$

Therefore, $I_2 \le 0$.

On the other hand, for $i \in \{-1, 0, 1\}$ let us denote $U_i^\pm = \{x \in \mathbb{R} : \pm u_\Delta(x+i\Delta x) > 0\}$ and $U_i^0 = \{x \in \mathbb{R} : u_\Delta(x+i\Delta x) = 0\}$. From the definition of R in (4.35), reordering I_1 we get:

$$\begin{aligned} I_1 &= \frac{1}{4\Delta x}\int_{\mathbb{R}} \left(u_\Delta^2(x+\Delta x) + u_\Delta(x+\Delta x)|u_\Delta(x+\Delta x)|\right)|u_\Delta(x)|^{p-2}u_\Delta(x)dx \\ &\quad - \frac{1}{2\Delta x}\int_{\mathbb{R}}|u_\Delta(x)|^{p+1}dx \\ &\quad + \frac{1}{4\Delta x}\int_{\mathbb{R}} \left(u_\Delta(x-\Delta x)|u_\Delta(x-\Delta x)| - u_\Delta^2(x-\Delta x)\right)|u_\Delta(x)|^{p-2}u_\Delta(x)dx \\ &\le \frac{1}{2\Delta x}\int_{U_0^+\cap U_1^+} u_\Delta^2(x+\Delta x)|u_\Delta(x)|^{p-1}dx - \frac{1}{2\Delta x}\int_{\mathbb{R}}|u_\Delta(x)|^{p+1}dx \\ &\quad + \frac{1}{2\Delta x}\int_{U_{-1}^-\cap U_0^-} u_\Delta^2(x-\Delta x)|u_\Delta(x)|^{p-1}dx. \end{aligned}$$

Using

$$a^2|b|^{p-1} \le \frac{2}{p+1}|a|^{p+1} + \frac{p-1}{p+1}|b|^{p+1}, \quad \forall a, b \in \mathbb{R},$$

we obtain that

$$\begin{aligned} I_1 &\le \frac{1}{2\Delta x}\left(\frac{2}{p+1}\int_{U_0^+\cap U_1^+}|u_\Delta(x+\Delta x)|^{p+1}dx + \frac{p-1}{p+1}\int_{U_0^+\cap U_1^+}|u_\Delta(x)|^{p+1}dx\right) \\ &\quad - \frac{1}{2\Delta x}\int_{\mathbb{R}}|u_\Delta(x)|^{p+1}dx \\ &\quad + \frac{1}{2\Delta x}\left(\frac{2}{p+1}\int_{U_{-1}^-\cap U_0^-}|u_\Delta(x-\Delta x)|^{p+1}dx + \frac{p-1}{p+1}\int_{U_{-1}^-\cap U_0^-}|u_\Delta(x)|^{p+1}dx\right) \\ &\le \frac{1}{2\Delta x}\int_{U_0^+}|u_\Delta(x)|^{p+1}dx - \frac{1}{2\Delta x}\int_{\mathbb{R}}|u_\Delta(x)|^{p+1}dx + \frac{1}{2\Delta x}\int_{U_0^-}|u_\Delta(x)|^{p+1}dx \end{aligned}$$

and, hence, $I_1 \le 0$.

Thus, from (4.37) we deduce:

$$\begin{aligned} \frac{1}{p}\frac{d}{dt}\|u_\Delta(t)\|_p^p &\le \int_{\mathbb{R}} d_{\Delta x}^-\left(d_{\Delta x}^+ u_\Delta(x)\right)|u_\Delta(x)|^{p-2}u_\Delta(x)dx \\ &= -\frac{1}{\Delta x^2}\int_{\mathbb{R}}\left(u_\Delta(x+\Delta x)-u_\Delta(x)\right) \\ &\qquad \left(|u_\Delta(x+\Delta x)|^{p-2}u_\Delta(x+\Delta x) - |u_\Delta(x)|^{p-2}u_\Delta(x)\right)dx. \end{aligned} \tag{4.38}$$

Moreover, the inequality

$$\left||x|^{p/2}-|y|^{p/2}\right|^2 \le \frac{p^2}{4(p-1)}(x-y)\left(|x|^{p-2}x - |y|^{p-2}y\right), \quad \forall x, y \in \mathbb{R},$$

guarantees that

$$\begin{aligned} \frac{d}{dt}\|u_\Delta(t)\|_p^p &\le -\frac{4(p-1)}{p}\int_{\mathbb{R}}\left|\frac{|u_\Delta(x+\Delta x)|^{p/2}-|u_\Delta(x)|^{p/2}}{\Delta x}\right|^2 \\ &= -\frac{4(p-1)}{p}\|d_{\Delta x}^+|u_\Delta|^{p/2}\|_2^2 \le 0. \end{aligned} \tag{4.39}$$

This estimate and Lemma A.1 allow us to write

$$\frac{d}{dt}\|u_\Delta(t)\|_p^p + \frac{p-1}{p}\frac{\|u_\Delta(t)\|_p^{p(p+1)/(p-1)}}{\|u_\Delta^0(t)\|_1^{2p/(p-1)}} \le 0. \tag{4.40}$$

Following the same arguments as in [29], we conclude that for any $p \in [2, \infty)$

$$\|u_\Delta(t)\|_p \le C(p)\|u_\Delta^0\|_1\, t^{-\frac{1}{2}(1-\frac{1}{p})}, \quad \forall t > 0. \tag{4.41}$$

In the same way, the cases $p = \infty$ and $p \in (1, 2)$ are proved too.

Finally, the general case $\mu > 0$ is immediate from (4.41) and the definition of u^μ (4.33), since for any $p \in [1, \infty]$ we have

$$\|u^\mu(t)\|_p = \mu^{1-\frac{1}{p}} \|u_\Delta(\mu^2 t)\|_p \le C(p) \|u^0_\Delta\|_1 \, t^{-\frac{1}{2}(1-\frac{1}{p})}.$$

The proof is now complete. □

Now that we have estimates on the L^p-norms of the solution, we need to obtain a similar result for the discrete gradient. We proceed as in Theorem 4.5.

Proposition 4.11. *For all $p \in [1, \infty]$ there exists a constant $C = C(p, \|u^0_\Delta\|_1) > 0$ such that:*

$$\|d^+_{\Delta x} u^\mu(t)\|_p \le C t^{-\frac{1}{2}(1-\frac{1}{p})-\frac{1}{2}}, \quad \forall t > 0, \tag{4.42}$$

for all solutions of (4.34) with initial data $u^0_\Delta \in L^1(\mathbb{R})$.

Proof. Let us denote by D^t_μ the semigroup associated to

$$\begin{cases} v_t(t,x) = \mu^2 \displaystyle\sum_{m=1}^{N} \omega_m v(t, x - m\frac{\Delta x}{\mu}) - \mu^2 F^\Delta_0 v(t,x) + \mu F^\Delta_1 d^+_{\Delta x/\mu} v(t,x), \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad t > 0, \text{ a.e. } x \in \mathbb{R}, \\ v(0,x) = v^0(x), \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{a.e. } x \in \mathbb{R}. \end{cases} \tag{4.43}$$

Multiplying (4.43) by $\operatorname{sign}(v(t,x))$, integrating on $\mathbb{R}$ and using that

$$\int_{\mathbb{R}} v(x-h)\operatorname{sign}(v(x))dx \le \int_{\mathbb{R}} |v(x)|dx, \quad \forall h \in \mathbb{R},$$

one shows that D^t_μ is stable in $L^1(\mathbb{R})$.

Now, for every $\tau > 0$ and $\mu > 0$, we know that the solution of (4.34) satisfies:

$$u^\mu(t+\tau) = G^\mu_\Delta(t) * D^t_\mu u^\mu(\tau) + \int_0^t G^\mu_\Delta(t-s) * D^s_\mu \Big(H\big(u^\mu(s+\tau)\big)\Big) ds, \tag{4.44}$$

where

$$\begin{aligned} H(u^\mu(s,x)) = \frac{1}{4}\Big(& d^+_{\Delta x/\mu}\left(u^\mu(s,x)^2\right) + d^-_{\Delta x\mu}\left(u^\mu(s,x)^2\right)\Big) \\ & + \Delta x \, d^+_{\Delta x/\mu} R\big(u^\mu(s, x - \frac{\Delta x}{\mu}), u^\mu(s,x)\big) \end{aligned}$$

and G^μ_Δ is the fundamental solution of the one-dimensional semi-discrete heat equation, defined by

$$\big(G^\mu_\Delta(t)\big)_j = \frac{1}{2\pi} \int_{-\pi\mu/\Delta x}^{\pi\mu/\Delta x} e^{-\frac{4t\mu^2}{\Delta x^2}\sin^2\frac{\xi\Delta x}{2\mu}} e^{ij\xi\frac{\Delta x}{\mu}} d\xi, \quad j \in \mathbb{Z},$$

It is well known (e.g. [13]) that

$$\|G^\mu_\Delta(t)\|_p \le C(p)t^{-\frac{1}{2}(1-\frac{1}{p})}, \quad t>0,$$

and

$$\|d^+_{\Delta x/\mu}G^\mu_\Delta(t)\|_p \le C(p)t^{-\frac{1}{2}(1-\frac{1}{p})-\frac{1}{2}}, \quad t>0.$$

Now let us apply the discrete operator $d^+_{\Delta x/\mu}$ to (4.44). Then

$$d^+_{\Delta x/\mu}u^\mu(t+\tau) = d^+_{\Delta x/\mu}G^\mu_\Delta(t) * D^t_\mu u^\mu(\tau) + \int_0^t d^+_{\Delta x/\mu}G^\mu_\Delta(t-s) * D^s_\mu\Big(H\big(u^\mu(s+\tau)\big)\Big)ds. \tag{4.45}$$

Using the decay properties of G^μ_Δ, Proposition 4.10 and the L^1-stability of D^t_μ, we obtain

$$\begin{aligned}
\|d^+_{\Delta x/\mu}u^\mu(t+\tau)\|_1 &\le \left\|d^+_{\Delta x/\mu}G^\mu_\Delta(t)\right\|_1 \left\|D^t_\mu u^\mu(\tau)\right\|_1 \\
&\quad + \int_0^t \left\|d^+_{\Delta x/\mu}G^\mu_\Delta(t-s)\right\|_1 \left\|D^s_\mu\Big(H\big(u^\mu(s+\tau)\big)\Big)\right\|_1 ds \\
&\le \left\|d^+_{\Delta x/\mu}G^\mu_\Delta(t)\right\|_1 \|u^\mu(\tau)\|_1 \\
&\quad + \int_0^t \left\|d^+_{\Delta x/\mu}G^\mu_\Delta(t-s)\right\|_1 \left\|H\big(u^\mu(s+\tau)\big)\right\|_1 ds \\
&\le C\left\|u^0\right\|_1 t^{-\frac{1}{2}} + C\int_0^t (t-s)^{-\frac{1}{2}} \left\|H\big(u^\mu(s+\tau)\big)\right\|_1 ds.
\end{aligned} \tag{4.46}$$

We now prove that for any $p\in[1,\infty)$, we have

$$\|H(u^\mu(s+\tau))\|_p \le C\|u^0_\Delta\|_1\tau^{-\frac{1}{2}}\|d^+_{\Delta x/\mu}u^\mu(s+\tau)\|_p \tag{4.47}$$

Observe that, in view of Proposition 4.10, we have

$$\begin{aligned}
\left\|d^+_{\Delta x/\mu}\left(u^\mu(s+\tau)^2\right)\right\|_p &\le 2\left\|u^\mu(s+\tau)\right\|_\infty \left\|d^+_{\Delta x/\mu}u^\mu(s+\tau)\right\|_p \\
&\le C\|u^0_\Delta\|_1\tau^{-\frac{1}{2}}\left\|d^+_{\Delta x/\mu}u^\mu(s+\tau)\right\|_p.
\end{aligned}$$

A similar result holds for $d^-_{\Delta x/\mu}$ since $\left\|d^-_h f\right\|_p = \left\|d^+_h f\right\|_p$ for all $f\in L^p(\mathbb{R})$ and $h>0$. Moreover, from the definition of R in (4.35) we have:

$$\begin{aligned}
&\Delta x\left\|d^+_{\Delta x/\mu}R\big(u^\mu(s+\tau,x-\frac{\Delta x}{\mu}),u^\mu(s+\tau,x)\big)\right\|_p \\
&\qquad \le \frac{1}{2}\left\|d^+_{\Delta x/\mu}\left(u^\mu(s+\tau)|u^\mu(s+\tau)|\right)\right\|_p \\
&\qquad \le \left\|u^\mu(s+\tau)\right\|_\infty \left\|d^+_{\Delta x/\mu}u^\mu(s+\tau)\right\|_p \le C\left\|u^0_\Delta\right\|_1\tau^{-\frac{1}{2}}\left\|d^+_{\Delta x/\mu}u^\mu(s+\tau)\right\|_p,
\end{aligned}$$

where we have used Proposition 4.10 and that

$$\big|x|x| - y|y|\big| \le 2|x-y|\max\{|x|,|y|\}, \quad \forall x, y \in \mathbb{R}.$$

Therefore, introducing in (4.46) the case $p = 1$ of (4.47), we get

$$\|d^+_{\Delta x/\mu} u^\mu(t+\tau)\|_1 \le C\left\|u^0\right\|_1 t^{-\frac{1}{2}} + C_\tau \int_0^t (t-s)^{-\frac{1}{2}} \left\|d^+_{\Delta x/\mu} u^\mu(s+\tau)\right\|_1 ds,$$

for some $C_\tau > 0$ depending only on τ and $\|u^0\|_1$. Applying Gronwall's Lemma and taking $t = \tau$, we conclude that

$$\|d^+_{\Delta x/\mu}(u^\mu(2\tau))\|_1 \le C_\tau, \quad \forall \mu > 0, \tag{4.48}$$

It is enough now to use the definition of u^μ in (4.33), taking $\tau = 1/2$ and $\mu = \sqrt{t}$ to obtain

$$\|d^+_{\Delta x}(u_\Delta(t))\|_1 \le Ct^{-\frac{1}{2}}, \quad \forall t > 0,$$

that is, (4.42) for $\mu = 1$ and $p = 1$.

The case $\mu = 1$ and $p \in (1, \infty)$ is immediate from (4.45), (4.47) and (4.48). Indeed, we have

$$\begin{aligned}
\|d^+_{\Delta x/\mu} u^\mu(t+\tau)\|_p &\le \left\|d^+_{\Delta x/\mu} G^\mu_\Delta(t)\right\|_p \|u^\mu(\tau)\|_1 \\
&\quad + \int_0^t \left\|d^+_{\Delta x/\mu} G^\mu_\Delta(t-s)\right\|_p \left\|H\big(u^\mu(s+\tau)\big)\right\|_1 ds \\
&\le C\|u^0\|_1 t^{-\frac{1}{2}(1-\frac{1}{p})-\frac{1}{2}} + C_\tau \int_0^t (t-s)^{-\frac{1}{2}(1-\frac{1}{p})-\frac{1}{2}} ds.
\end{aligned}$$

with $C_\tau = C(p, \tau, \|u^0\|_1)$. Taking $t = \tau$ implies that

$$\|d^+_{\Delta x/\mu} u^\mu(2\tau)\|_p \le C_\tau, \quad \forall \mu > 0, \tag{4.49}$$

This is equivalent to (4.42) for $\mu = 1$ and $p \in (1, \infty)$.

Furthermore, repeating similar arguments, the case $\mu = 1$ and $p = \infty$ follows from (4.45) and estimates (4.47) and (4.49):

$$\begin{aligned}
\|d^+_{\Delta x/\mu} u^\mu(t+\tau)\|_\infty &\le \left\|d^+_{\Delta x/\mu} G^\mu_\Delta(t)\right\|_\infty \|u^\mu(\tau)\|_1 \\
&\quad + \int_0^t \left\|d^+_{\Delta x/\mu} G^\mu_\Delta(t-s)\right\|_q \left\|H\big(u^\mu(s+\tau)\big)\right\|_{q'} ds \\
&\le C\|u^0\|_1 t^{-1} + C_\tau \int_0^t (t-s)^{-\frac{1}{2}(1-\frac{1}{q})-\frac{1}{2}} \|d^+_{\Delta x/\mu} u^\mu(s+\tau)\|_{q'} ds \\
&\le C\|u^0\|_1 t^{-1} + C_\tau \int_0^t (t-s)^{-\frac{1}{2}(1-\frac{1}{q})-\frac{1}{2}} ds.
\end{aligned}$$

where $q \in (1,\infty)$, $1/q + 1/q' = 1$ and $C_\tau = C(q, q', \tau, \|u^0\|_1)$. It is now enough to take $t = \tau$ to conclude that

$$\|d^+_{\Delta x/\mu} u^\mu(2\tau)\|_\infty \leq C_\tau, \quad \forall \mu > 0,$$

which is equivalent to (4.42) for $\mu = 1$ and $p = \infty$.

Finally, the general case $\mu > 0$ is immediate from (4.41) and the definition of u^μ (4.33), since for any $p \in [1,\infty]$ we have

$$\|d^+_{\Delta x/\mu} u^\mu(t)\|_p = \mu^{2-\frac{1}{p}} \|d^+_{\Delta x} u_\Delta(\mu^2 t)\|_p \leq C t^{-\frac{1}{2}(1-\frac{1}{p})-\frac{1}{2}}$$

This concludes the proof. □

To end this part, let us remark that the solution u^μ of system (4.34) conserves the mass of the initial data u^0_Δ. In fact, note that it is the same as the mass of u_0, when u^0_Δ is defined as in (4.4). Moreover, we show that (4.34) defines a contractive semigroup. This will be useful to obtain the estimates for the compactness of $\{u^\mu\}_{\mu>0}$. For the sake of clarity, we prove this lemma in the final appendix.

Lemma 4.12. *For any initial data $u^0_\Delta \in L^1(\mathbb{R})$, the solution u^μ to (4.34) satisfies*

$$\int_{\mathbb{R}} u^\mu(t,x) = \int_{\mathbb{R}} u^0_\Delta(x), \quad \forall t > 0.$$

Moreover, (4.34) defines a contractive semigroup in $L^1(\mathbb{R})$.

4.3.2 Compactness of the set $\{u^\mu\}_{\mu>0}$

In this section, we prove the compactness of the trajectories of the set $\{u^\mu(t)\}_{\mu>0}$ introduced in the previous section, in order to pass to the limit $\mu \to \infty$. Unlike the continuous case, we do not have estimates of u^μ in $H^1(\mathbb{R})$, since it is piecewise constant. Nevertheless, the following lemma makes possible the use of the compact embedding of $H^s_{loc}(\mathbb{R})$ into $L^2_{loc}(\mathbb{R})$, with $0 < s < 1/2$. The proof will be given in the Appendix.

Lemma 4.13. *For any $0 < s < \frac{1}{2}$, there exists a positive constant $C = C(s)$ such that, for any mesh-size $0 < \Delta x < 1$, the following holds for all piecewise constant functions w as in (4.7):*

$$\|w\|_{H^s(\mathbb{R})} \leq C \left(\|w\|_2 + \|d^+_{\Delta x} w\|_2 \right).$$

Let us remark that, as a consequence of this lemma, Proposition 4.10 and Proposition 4.11, we obtain the time-decay estimate for u^μ in $H^s(\mathbb{R})$:

$$\|u^\mu(t)\|_{H^s(\mathbb{R})} \leq C\left(\|u^\mu(t)\|_2 + \|d^+_{\Delta x/\mu}u^\mu(t)\|_2\right) \leq C\left(t^{-\frac{1}{4}} + t^{-\frac{3}{4}}\right),\ \forall t>0, \forall \mu>0, \tag{4.50}$$

with $0 < s < 1/2$. Thus, we can use Theorem 4.7 to prove the compactness of the family $\{u^\mu\}_{\mu>0}$.

Theorem 4.14. *For every $0<\tau<T<\infty$, the family $\{u^\mu\}_{\mu>0} \subset C([\tau,T],L^1(\mathbb{R}))$ is relatively compact.*

Proof. We will proceed in two steps, analogously to Theorem 4.8.

Step 1. First we will show the result locally in $C([\tau,T],L^1(I))$, with $I=[-r,r]$ for an arbitrary $r>0$. Let us consider the spaces $X=H^s(I)$ with $s\in(0,\frac{1}{2})$, $Z=L^2(I)$ and $Y=H^{-1}(I)$.

From (4.50) we know that the set $\{u^\mu\}_{\mu>0}$ is bounded in $L^\infty([\tau,T],H^s_{loc}(\mathbb{R}))$. In particular, first condition of Theorem 4.7 is fulfilled. Thus, it suffices to check that u^μ_t is bounded in $L^\infty([\tau,T],H^{-1}(I))$. Let us multiply (4.34) by a function $\varphi\in\mathcal{C}^\infty_c(\mathbb{R})$ and integrate it over $\mathbb{R}$. Using the definition of R in (4.35), we have:

$$\begin{aligned}
&\Big|\int_{\mathbb{R}} u^\mu_t\varphi dx\Big| \\
&\leq \frac{1}{4}\left|\int_{\mathbb{R}}\left(d^+_{\Delta x/\mu}\big(u^\mu(x)^2\big)+d^-_{\Delta x/\mu}\big(u^\mu(x)^2\big)\right)\varphi(x)dx\right| \\
&\quad+\Delta x\left|\int_{\mathbb{R}} d^+_{\Delta x/\mu}R\big(u^\mu(x-\frac{\Delta x}{\mu}),u^\mu(x)\big)\varphi(x)dx\right|+\left|\int_{\mathbb{R}} d^-_{\Delta x/\mu}\big(d^+_{\Delta x/\mu}u^\mu(x)\big)\varphi(x)dx\right| \\
&\quad+\left|\int_{\mathbb{R}}\left(\mu^2\sum_{m=1}^N\omega_m u^\mu(x-m\frac{\Delta x}{\mu})-\mu^2F_0^\Delta u^\mu(x)+\mu F_1^\Delta d^+_{\Delta x/\mu}u^\mu(x)\right)\varphi(x)dx\right| \\
&\leq \frac{1}{2}\|d^+_{\Delta x}\varphi\|_2\|u^\mu\|_4^2+\frac{1}{2}\|d^+_{\Delta x}\varphi\|_2\|u^\mu\|_4^2+\|d^+_{\Delta x}\varphi\|_2\|d^+_{\Delta x}u^\mu\|_2 \\
&\quad+\left|\int_{\mathbb{R}}\left(\mu^2\sum_{m=1}^N\omega_m\big(u^\mu(x-m\frac{\Delta x}{\mu})-u^\mu(x)\big)+\mu F_1^\Delta d^+_{\Delta x/\mu}u^\mu(x)\right)\varphi(x)dx\right|.
\end{aligned}$$

Obviously, the first three terms on the right hand side of the inequality are uniformly bounded for $\mu>0$, so let us focus on the last one. Using the Fourier transform and the definition of F_0^Δ in (4.6), we have

$$\begin{aligned}
I_\mu &= \left|\int_{\mathbb{R}}\left(\mu^2\sum_{m=1}^N\omega_m\big(u^\mu(x-m\frac{\Delta x}{\mu})-u^\mu(x)\big)+\mu F_1^\Delta d^+_{\Delta x/\mu}u^\mu(x)\right)\varphi(x)dx\right| \\
&\leq \mu^2\int_{\mathbb{R}}\left|\sum_{m=1}^N\omega_m\left(e^{-im\frac{\Delta x}{\mu}\xi}-1\right)+F_1^\Delta\frac{e^{i\frac{\Delta x}{\mu}\xi}-1}{\Delta x}\right|\left|\widehat{u^\mu}(\xi)\right||\widehat{\varphi}(\xi)|\,d\xi.
\end{aligned}$$

If we take $p = a^{-\Delta x}$ and $b = e^{-i\frac{\Delta x}{\mu}\xi}$ on Lemma A.2 and use the definitions of ω_m in (4.5) and F_1^Δ in (4.6), we have:

$$\begin{aligned}
&\left| \sum_{m=1}^{N} \omega_m \left(e^{-im\frac{\Delta x}{\mu}\xi} - 1 \right) + F_1^\Delta \frac{e^{i\frac{\Delta x}{\mu}\xi} - 1}{\Delta x} \right| \\
&\qquad = \left| e^{\Delta x} - 1 \right| \left| \sum_{m=1}^{N} e^{-m\Delta x} \left(e^{-im\frac{\Delta x}{\mu}\xi} - 1 \right) + \sum_{m=1}^{N} m e^{-m\Delta x} (e^{i\frac{\Delta x}{\mu}\xi} - 1) \right| \\
&\qquad \leq \left| e^{\Delta x} - 1 \right| \left| e^{-i\frac{\Delta x}{\mu}\xi} - 1 \right|^2 \frac{e^{-\Delta x}}{(1 - e^{-\Delta x})^3} \\
&\qquad = \left| e^{-i\frac{\Delta x}{\mu}\xi} - 1 \right|^2 \frac{1}{(1 - e^{-\Delta x})^2}. \qquad (4.51)
\end{aligned}$$

Therefore, combining this with the Cauchy-Schwartz inequality, we obtain

$$I_\mu \leq \frac{\Delta x^2}{(1 - e^{-\Delta x})^2} \|d^+_{\Delta x/\mu} u^\mu\|_2 \|d^+_{\Delta x/\mu} \varphi\|_2.$$

Thus, using that $\|d^+_{\Delta x/\mu}\varphi\|_2 \leq \|\varphi'\|_2$, we get

$$\begin{aligned}
\left| \int_{\mathbb{R}} u_t^\mu(t)\varphi dx \right| &\leq \|d^+_{\Delta x/\mu}\varphi\|_2 \|u^\mu(t)\|_4^2 + \|d^+_{\Delta x/\mu}\varphi\|_2 \|d^+_{\Delta x/\mu} u^\mu(t)\|_2 \\
&\quad + \frac{\Delta x^2}{(1 - e^{-\Delta x})^2} \|d^+_{\Delta x/\mu} u^\mu(t)\|_2 \|d^+_{\Delta x/\mu}\varphi\|_2 \\
&\leq C\|\varphi\|_{H^1(\mathbb{R})} \Big(\|u^\mu(t)\|_4^2 + \|d^+_{\Delta x/\mu} u^\mu(t)\|_2 \Big).
\end{aligned}$$

for any $\varphi \in \mathcal{C}_c^\infty(I)$ and with $C > 0$ independent of μ. In view of Propositions 4.10 and 4.11, both norms of u^μ in the right-hand side are uniformly bounded in $[\tau, T]$, so u_t^μ is uniformly bounded in $L^\infty([\tau, T], H^{-1}(I))$. We conclude that the family $\{u^\mu\}_{\mu>0}$ is relatively compact in $C([\tau, T], L^2_{loc}(\mathbb{R}))$. Finally, compactness in $L^2_{loc}(\mathbb{R})$ implies compactness in $L^1_{loc}(\mathbb{R})$, so $\{u^\mu\}_{\mu>0}$ is also relatively compact in $C([\tau, T], L^1_{loc}(\mathbb{R}))$.

Step 2. Now we need to extend the result globally. Let us consider again the same function ψ_r defined in the third step of the proof of Theorem 4.8, such that $\psi_r(z) = \psi(z/r)$ with ψ given by (4.23) and $r > 0$. Since we know that $\{u^\mu\}_{\mu>0}$ is relatively compact in $C([\tau, T], L^1_{loc}(\mathbb{R}))$, it suffices to show that

$$\sup_{[\tau,T]} \|u^\mu(t)\psi_r\|_1 \longrightarrow 0 \quad \text{as } r \to \infty, \text{ uniformly on } \mu \geq 1. \qquad (4.52)$$

Note that, because of Lemma 4.12 and Crandall-Tartar Lemma [38, Chapter II], a similar argument as in Theorem 4.8 shows that it is enough to prove (4.52) for nonnegative initial data and solutions.

Thus, we focus on nonnegative solutions. Let us multiply (4.34) by ψ_r and integrate it over $(0,t)\times\mathbb{R}$. We obtain:

$$\begin{aligned}\int_{\mathbb{R}} u^\mu(t,x)\psi_r(x)dx &= \int_{\mathbb{R}} u_0^\mu(x)\psi_r(x)dx\\ &+\frac{1}{4}\int_0^t\int_{\mathbb{R}}\Big(d^+_{\Delta x/\mu}\big(u^\mu(s,x)^2\big)+d^-_{\Delta x/\mu}\big(u^\mu(s,x)^2\big)\Big)\psi_r(x)dxds\\ &+\Delta x\int_0^t\int_{\mathbb{R}} d^+_{\Delta x/\mu}\Big(R\big(u^\mu(s,x-\frac{\Delta x}{\mu}),u^\mu(s,x)\big)\Big)\psi_r(x)dxds\\ &+\int_0^t\int_{\mathbb{R}} d^-_{\Delta x/\mu}\Big(d^+_{\Delta x/\mu}\big(u^\mu(s,x)\big)\Big)\psi_r(x)dxds\\ &+\mu^2\int_0^t\int_{\mathbb{R}}\sum_{m=1}^N\omega_m\big(u^\mu(s,x-m\frac{\Delta x}{\mu})-u^\mu(s,x)\big)\psi_r(x)dxds\\ &+\mu F_1^\Delta\int_0^t\int_{\mathbb{R}} d^+_{\Delta x/\mu}u^\mu(s,x)\psi_r(x)dxds. \end{aligned}\tag{4.53}$$

We pass now the discrete derivatives to ψ_r and estimate the right-hand side using time-decay estimates from Proposition 4.10:

$$\begin{aligned}\int_{\mathbb{R}}|u^\mu(t,x)|\psi_r(x)dx &\lesssim \int_{\mathbb{R}}|u_0^\mu(x)|\psi_r(x)dx+\|\psi'\|_\infty\frac{\sqrt{t}}{r}+\|\psi''\|_\infty\frac{t}{r^2}\\ &+\mu^2\int_0^t\int_{\mathbb{R}}\sum_{m=1}^N\omega_m\big(u^\mu(s,x-m\frac{\Delta x}{\mu})-u^\mu(s,x)\big)\psi_r(x)dxds\\ &+\mu F_1^\Delta\int_0^t\int_{\mathbb{R}} d^+_{\Delta x/\mu}u^\mu(s,x)\psi_r(x)dxds. \end{aligned}\tag{4.54}$$

Let us focus on the lasts two terms, for which we have

$$\begin{aligned}&\int_{\mathbb{R}}\Big(\mu^2\sum_{m=1}^N\omega_m\big(u^\mu(s,x-m\frac{\Delta x}{\mu})-u^\mu(s,x)\big)+\mu F_1^\Delta d^+_{\Delta x/\mu}u^\mu(s,x)\Big)\psi_r(x)dx\\ &\quad=\mu^2\sum_{m=1}^N\omega_m\int_{\mathbb{R}}u^\mu(s,x)\Big(\psi_r(x+m\frac{\Delta x}{\mu})-\psi_r(x)-m\frac{\Delta x}{\mu}\,d^-_{\Delta x/\mu}(\psi_r(x))\Big)dx\\ &\quad\lesssim\left\|\mu^2\sum_{m=1}^N\omega_m\Big(\psi_r(x+m\frac{\Delta x}{\mu})-\psi_r(x)-m\frac{\Delta x}{\mu}\,d^-_{\Delta x/\mu}(\psi_r(x))\Big)\right\|_\infty\|u^\mu(s)\|_1\\ &\quad\lesssim\frac{\|\psi''\|_\infty}{r^2}\|u^0_\Delta\|_1.\end{aligned}$$

Thus, plugging this into (4.54) and using the non-negativity of the solution, we get

$$\int_{\mathbb{R}}|u^\mu(t,x)|\psi_r(x)dx\lesssim\int_{\mathbb{R}}|u_0^\mu(x)|\psi_r(x)dx+\|\psi'\|_\infty\frac{\sqrt{t}}{r}+\|\psi''\|_\infty\frac{t}{r^2},\tag{4.55}$$

which tends to 0 uniformly on $\mu>0$ when $r\to\infty$. Therefore, we proved (4.52) and, consequently, we can assure that $\{u^\mu\}_{\mu>0}$ is relatively compact in $C([\tau,T],L^1(\mathbb{R}))$. □

A slight modification of the proof of the previous theorem gives as the necessary estimate to identify the initial data, stated in the following proposition.

Proposition 4.15. *For every test function $\varphi \in C_c^\infty(\mathbb{R})$, there exists $C > 0$, independent of μ, such that*

$$\left| \int_{\mathbb{R}} u^\mu(t,x)\varphi(x)dx - \int_{\mathbb{R}} u_0^\mu(x)\varphi(x)dx \right| \leq C(t + \sqrt{t}). \tag{4.56}$$

Proof. It is enough to multiply (4.34) by $\varphi \in C_c^\infty(\mathbb{R})$ and integrate it over $(0,t) \times \mathbb{R}$. Then, integrating by parts and repeating arguments similar to the ones in the second step of the proof for Theorem 4.14, we deduce (4.56). □

4.3.3 Passing to the limit

Finally, we have everything that we need to prove our main result, stated in Theorem 4.2, regarding the large-time behavior of the approximations to the solution of problem (4.3).

Proof of Theorem 4.2. By Theorem 4.14, we know that for every $0 < \tau < T < \infty$, the family $\{u^\mu\}_{\mu>0}$ is relatively compact in $C([\tau,T], L^1(\mathbb{R}))$. Consequently, there exists a subsequence of it (which we will not relabel) and a function $\bar{u} \in C((0,\infty), L^1(\mathbb{R}))$ such that

$$u^\mu \longrightarrow \bar{u} \in C([\tau,T], L^1(\mathbb{R})), \quad \text{as } \mu \to \infty. \tag{4.57}$$

We can also assume that $u^\mu(t,x) \to \bar{u}(t,x)$ almost everywhere in $(0,\infty) \times \mathbb{R}$ as $\mu \to \infty$.

Now, we multiply equation (4.34) by a test function $\phi \in C_c^\infty((0,\infty)\times\mathbb{R})$ and integrate it over $(0,\infty) \times \mathbb{R}$. We have:

$$\begin{aligned}
&\int_0^\infty \int_{\mathbb{R}} u_t^\mu(t,x)\phi(t,x)dxdt \\
&\quad = \frac{1}{4}\int_0^\infty \int_{\mathbb{R}} \Big(d^+_{\Delta x/\mu}\big(u^\mu(t,x)^2\big) + d^-_{\Delta x/\mu}\big(u^\mu(t,x)^2\big)\Big)\phi(t,x)dxdt \\
&\qquad + \Delta x \int_0^\infty \int_{\mathbb{R}} d^+_{\Delta x/\mu} R\big(u^\mu(t,x-\frac{\Delta x}{\mu}), u^\mu(t,x)\big)\phi(t,x)dxdt \\
&\qquad + \int_0^\infty \int_{\mathbb{R}} d^-_{\Delta x/\mu}\left(d^+_{\Delta x/\mu} u^\mu(t,x)\right)\phi(t,x)dxdt \\
&\qquad + \int_0^\infty \int_{\mathbb{R}} \left(\mu^2 \sum_{m=1}^N \omega_m u^\mu(t,x-m\frac{\Delta x}{\mu}) - \mu^2 F_0^\Delta u^\mu(t,x)\right) \\
&\qquad + \mu F_1^\Delta \int_0^\infty \int_{\mathbb{R}} d^+_{\Delta x/\mu} u^\mu(t,x)\phi(t,x)dxdt
\end{aligned} \tag{4.58}$$

Our claim is that, passing to the limit $\mu \to \infty$, we obtain that $\bar{u}$ is a weak solution of the equation:

$$\begin{cases} \bar{u}_t = \bar{u}\bar{u}_x + (1 + F_2^\Delta)\bar{u}_{xx}, \quad (t,x) \in (0,\infty) \times \mathbb{R}, \\ \bar{u}(0) = M\delta_0. \end{cases} \tag{4.59}$$

All the limits in (4.58) are known (see Chapter 2) except the last term. It is sufficient to check that we can take the limit $\mu \to \infty$ in

$$\mathcal{L}^\mu(t) = \int_{\mathbb{R}} \left(\mu^2 \sum_{m=1}^{N} \omega_m u^\mu(t, x - m\frac{\Delta x}{\mu}) - \mu^2 F_0^\Delta u^\mu(t,x) \right) \phi(t,x)dx \\ + \mu F_1^\Delta \int_{\mathbb{R}} d^+_{\Delta x/\mu} u^\mu(t,x)\phi(t,x)dx.$$

First, we reorder $\mathcal{L}_\mu$:

$$\mathcal{L}^\mu(t) = \mu^2 \int_{\mathbb{R}} u^\mu(t,x) \sum_{m=1}^{N} \omega_m \left(\phi(t, x + m\frac{\Delta x}{\mu}) - \phi(t,x) - m\frac{\Delta x}{\mu} d^-_{\Delta x/\mu}(\phi(t,x)) \right) dx. \tag{4.60}$$

Now, due to Taylor's Theorem, for each $m \in \{1, \dots, N\}$

$$\phi(t, x + m\frac{\Delta x}{\mu}) - \phi(t,x) = m\frac{\Delta x}{\mu}\phi_x(t,x) + \frac{1}{2}m^2\frac{\Delta x^2}{\mu^2}\phi_{xx}(t,x) + \frac{1}{\mu^3}O(\|\phi_{xxx}(t)\|_\infty).$$

In the same way,

$$d^-_{\Delta x/\mu}(\phi(t,x)) = \phi_x(t,x) - \frac{1}{2}\frac{\Delta x}{\mu}\phi_{xx}(t,x) + \frac{1}{\mu^2}O(\|\phi_{xxx}(t)\|_\infty).$$

We combine this into (4.60) and get

$$\mathcal{L}^\mu(t) = F_2^\Delta \int_{\mathbb{R}} u^\mu(t,x)\phi_{xx}(t,x)dx + O(\|\phi_{xxx}(t)\|_\infty)\frac{1}{\mu}\int_{\mathbb{R}} u^\mu(t,x)dx. \tag{4.61}$$

Therefore, as $u^\mu \to \bar{u}$ in $C([\tau, T], L^1(\mathbb{R}))$, taking the limit $\mu \to \infty$ in (4.61), we obtain:

$$\lim_{\mu\to\infty} \int_0^\infty \mathcal{L}^\mu(t) = F_2^\Delta \int_0^\infty \int_{\mathbb{R}} \bar{u}(t,x)\phi_{xx}(t,x)dx.$$

It follows that $\bar{u}$ satisfies

$$-\int_0^\infty \int_{\mathbb{R}} \bar{u}\phi_t = -\frac{1}{2}\int_0^\infty \int_{\mathbb{R}} \bar{u}^2\phi_x + (1 + F_2^\Delta)\int_0^\infty \int_{\mathbb{R}} \bar{u}\phi_{xx},$$

so it is a weak solution of the equation in (4.59). It remains to identify the behavior of $\bar{u}$ as $t \to 0$. Due to Proposition 4.15, for any $\varphi \in C_c^\infty(\mathbb{R})$ we have

$$\left| \int_{\mathbb{R}} u^\mu(t,x)\varphi(x)dx - \int_{\mathbb{R}} u_0^\mu(x)\varphi(x)dx \right| \le C(t + \sqrt{t})$$

and from (4.57) we deduce

$$\left| \int_{\mathbb{R}} \bar{u}(t,x)\varphi(x)dx - M\varphi(0) \right| \leq C(t+\sqrt{t})$$

by letting $\mu \to \infty$. Passing to the limit $t \to 0$ and using classical approximation arguments, we deduce that $\bar{u}(0) = M\delta_0$ in the sense of bounded measures. Thus, we conclude that $\bar{u}$ is the unique solution u_M of equation (4.59), and that, in fact, the whole family $\{u^\mu\}_{\mu>0}$ converges to u_M in $C((0,\infty), L^1(\mathbb{R}))$.

Therefore, by (4.57), we have:

$$\lim_{\mu\to\infty} \|u^\mu(1) - u_M(1)\|_1 = 0$$

and, setting $\mu = \sqrt{t}$ and making use of the self-similar form of u_M (see e.g. [29]) we obtain

$$\lim_{t\to\infty} \|u_\Delta(t) - u_M(t)\|_1 = 0. \tag{4.62}$$

Finally, the convergence in the L^p-norms for $p \in (1,\infty)$ follows from (4.62), the decay estimate of Proposition 4.10 for $p = \infty$ and the Hölder inequality. In fact, we have:

$$\|u_\Delta(t) - u_M(t)\|_p \leq (\|u_\Delta(t)\|_\infty + \|u_M(t)\|_\infty)^{1-\frac{1}{p}} \|u_\Delta(t) - u_M(t)\|_1^{\frac{1}{p}} \leq o(t^{-\frac{1}{2}(1-\frac{1}{p})}).$$

Using the piecewise constant interpolation of u_M, which we denote $S(u_M)$, and (A.2), the case $p = \infty$ follows:

$$\begin{aligned}
\|u_\Delta(t) - u_M(t)\|_\infty &\leq \|u_\Delta(t) - S(u_M(t))\|_\infty + \|S(u_M(t)) - u_M(t)\|_\infty \\
&\lesssim \|u_\Delta(t) - S(u_M(t))\|_2^{\frac{1}{2}} \|d^+_{\Delta x}(u_\Delta(t) - S(u_M(t)))\|_2^{\frac{1}{2}} + \Delta x \|u_{M,x}(t)\|_\infty \\
&\lesssim (\|u_\Delta(t) - u_M(t)\|_2 + \|u_M(t) - S(u_M(t))\|_2)^{\frac{1}{2}} (\|d^+_{\Delta x} u_\Delta(t)\|_2 \\
&\qquad + \|d^+_{\Delta x} S(u_M(t))\|_2)^{\frac{1}{2}} + \Delta x \|u_{M,x}(t)\|_\infty \\
&\leq o(t^{-\frac{1}{2}} + t^{-\frac{3}{4}} + t^{-1}).
\end{aligned}$$

Now the proof is complete. □

4.3.4 Convergence of the scheme

To conclude this section, let us prove that u_Δ converges to the solution u of (4.3) as $\Delta x \to 0$.

Theorem 4.16. *Let $u^0 \in L^1(\mathbb{R})$. The set of approximated solutions $\{u_\Delta\}_{\Delta x>0}$ given by (4.4) converges in $C((0,\infty), L^1(\mathbb{R}))$ to the solution u of (4.3) as $\Delta x \to 0$.*

Proof. Following the same arguments as in Theorem 4.14, one shows that for every $0 < \tau < T < \infty$, the family $\{u_\Delta\}_{\Delta x>0} \subset C([\tau, T], L^1(\mathbb{R}))$ is relatively compact. Thus, there exists a subsequence of it (which we will not relabel) and a function $\bar{u} \in C((0,\infty), L^1(\mathbb{R}))$ such that

$$u_\Delta \longrightarrow \bar{u} \in C([\tau, T], L^1(\mathbb{R})), \quad \text{as } \Delta x \to 0. \tag{4.63}$$

We can also assume that $u_\Delta(t,x) \to \bar{u}(t,x)$ almost everywhere in $(0,\infty)\times\mathbb{R}$ as $\Delta x \to 0$.

Now, we take $\mu = 1$ in equation (4.34), multiply it by a test function $\phi \in C_c^\infty((0,\infty)\times\mathbb{R})$ and integrate it over $(0,\infty)\times\mathbb{R}$. We have:

$$\begin{aligned}
&\int_0^\infty \int_{\mathbb{R}} u_{\Delta,t}(t,x)\phi(t,x)dxdt \\
&= \frac{1}{4}\int_0^\infty\int_{\mathbb{R}} \Big(d^+_{\Delta x}\big(u_\Delta(t,x)^2\big) + d^-_{\Delta x}\big(u_\Delta(t,x)^2\big)\Big)\phi(t,x)dxdt \\
&\quad + \Delta x \int_0^\infty\int_{\mathbb{R}} d^+_{\Delta x} R\big(u_\Delta(t,x-\Delta x), u_\Delta(t,x)\big)\phi(t,x)dxdt \\
&\quad + \int_0^\infty\int_{\mathbb{R}} d^-_{\Delta x}\left(d^+_{\Delta x}u_\Delta(t,x)\right)\phi(t,x)dxdt \\
&\quad + \int_0^\infty\int_{\mathbb{R}} \left(\sum_{m=1}^N \omega_m u_\Delta(t,x-m\Delta x) - F_0^\Delta u_\Delta(t,x) + F_1^\Delta d^+_{\Delta x}u_\Delta(t,x)\right)\phi(t,x)dxdt
\end{aligned} \tag{4.64}$$

Our claim is that, passing to the limit $\Delta x \to 0$, we obtain that $\bar{u}$ is a weak solution of the equation (4.3). Thanks to (4.63) and to Propositions 4.10 and 4.11, we know that we can take all the limits in (4.64), except for the last term. Thus, it is sufficient to check that we can pass to the limit $\Delta x \to 0$ in

$$\begin{aligned}
\mathcal{L}_\Delta(t) &= \int_{\mathbb{R}} \left(\sum_{m=1}^N \omega_m u_\Delta(t,x-m\Delta x) - F_0^\Delta u_\Delta(t,x) + F_1^\Delta d^+_{\Delta x}u_\Delta(t,x)\right)\phi(t,x)dx \\
&= \int_{\mathbb{R}} u_\Delta(t,x)\left(\sum_{m=1}^N \omega_m \phi(t,x+m\Delta x) - F_0^\Delta\phi(t,x) - F_1^\Delta d^-_{\Delta x}\phi(t,x)\right)dx.
\end{aligned}$$

First, let us first observe that

$$F_0^\Delta = \sum_{m=1}^N e^{-m\Delta x}(e^{\Delta x}-1) = 1 - e^{-N\Delta x} \to 1$$

and

$$F_1^\Delta = \Delta x(e^{\Delta x}-1)\sum_{m=1}^N m e^{-m\Delta x} = \frac{\Delta x e^{-N\Delta x}(e^{(N+1)\Delta x} - e^{\Delta x}(N+1) + N)}{e^{\Delta x}-1} \to 1,$$

as long as $N = N(\Delta x)$ is taken such that $N\Delta x \to \infty$ as $\Delta x \to 0$. Moreover, using (4.22) and that

$$(e^{\Delta x} - 1)\frac{1 - e^{-N\Delta x(1-i\xi)}}{e^{(1-i\xi)\Delta x} - 1} \to \frac{1}{1 - i\xi} = \frac{1 + i\xi}{1 + \xi^2}, \quad \text{as } \Delta x \to 0$$

we obtain

$$\begin{aligned}\left| \sum_{m=1}^{N} \omega_m \phi(t, x + m\Delta x) - \tilde{K} * \phi(t, x) \right| &\leq \int_{\mathbb{R}} |\widehat{\phi}(t, \xi)| \left| \sum_{m=1}^{N} \omega_m e^{im\Delta x\xi} - \widehat{K}(-\xi) \right| d\xi \\ &= \int_{\mathbb{R}} |\widehat{\phi}(t, \xi)| \left| (e^{\Delta x} - 1)\frac{1 - e^{-N\Delta x(1-i\xi)}}{e^{(1-i\xi)\Delta x} - 1} - \frac{1 + i\xi}{1 + \xi^2} \right| d\xi \to 0,\end{aligned}$$

where $\tilde{K}(z) = K(-z)$. Therefore

$$\lim_{\Delta x \to 0} \mathcal{L}_\Delta(t) = \int_{\mathbb{R}} \bar{u}(t, x) \left(\tilde{K} * \phi(t, x) - \phi(t, x) - \phi_x(t, x) \right) dx.$$

It follows that $\bar{u}$ satisfies

$$-\int_0^\infty \int_{\mathbb{R}} \bar{u}\phi_t = -\frac{1}{2}\int_0^\infty \int_{\mathbb{R}} \bar{u}^2\phi_x + \int_0^\infty \int_{\mathbb{R}} \bar{u}\phi_{xx} + \int_0^\infty \int_{\mathbb{R}} \bar{u}(\tilde{K} * \phi - \phi - \phi_x).$$

so it is a weak solution of the equation in (4.3).

Now, it remains to identify the behavior of $\bar{u}$ as $t \to 0$. In the same way as in Proposition 4.15, we can prove that for every test function $\varphi \in C_c^\infty(\mathbb{R})$ and $\Delta x < 1$, there exists $C > 0$, independent of Δx, such that

$$\left| \int_{\mathbb{R}} u_\Delta(t, x)\varphi(x)dx - \int_{\mathbb{R}} u_\Delta^0(x)\varphi(x)dx \right| \leq C(t + \sqrt{t}).$$

and from (4.63) and the definition of u_Δ^0 in (4.7), we deduce

$$\left| \int_{\mathbb{R}} \bar{u}(t, x)\varphi(x)dx - \int_{\mathbb{R}} u^0(x)\varphi(x)dx \right| \leq C(t + \sqrt{t})$$

by letting $\Delta x \to 0$. Passing to the limit $t \to 0$ and using classical approximation arguments, we deduce that $\bar{u}(0) = u^0$ in the sense of bounded measures. Thus, we conclude that $\bar{u}$ is the unique solution u of equation (4.3), and that, in fact, the whole family $\{u_\Delta\}_{\Delta x > 0}$ converges to u in $C((0, \infty), L^1(\mathbb{R}))$. Now the proof is complete. □

4.4 Numerical experiments

The aim of this last section is to support the necessity of using large-time behavior preserving schemes for the augmented Burgers equation. On the one hand, we show the importance of a numerical flux that does not destroy the N-wave shape at the early stages. On the other, we emphasize the role of the correcting factors F_0^Δ and F_1^Δ in the truncation of the convolution.

Regarding the time discretization, we opt for the explicit Euler for its simplicity. Even if there is no guarantee that the asymptotic behavior is preserved, numerical simulations exhibit a correct performance. Thus, we consider it enough to illustrate the key points enumerated above. We need to take into account that there is a stability condition that must be satisfied to ensure the convergence. It is easy to see (e.g. [18, 38]) that a sufficient condition is that

$$\frac{\Delta t}{\Delta x}\max_j\{u_j^0\} + 2\nu\frac{\Delta t}{\Delta x^2} + c\,\Delta t\sum_{m=1}^{N}(m+1)\omega_m \leq 1 \tag{4.65}$$

Let us choose the following compactly supported initial data.

$$u_0(x) = \begin{cases} -\frac{1}{10}\sin(x/2), & x \in [-\pi, 0], \\ -\frac{1}{20}\sin(2x), & x \in [0, \pi/2], \\ 0, & \text{elsewhere} \end{cases} \tag{4.66}$$

We take a mesh size $\Delta x = 0.1$. In order to avoid boundary issues, we choose a large enough spatial domain.

In Figure 4.1 we show the solution for $\nu = 10^{-2}$, $c = 2 \times 10^{-2}$ and $\theta = 1$ at time $t = 10^4$, as well as the corresponding asymptotic profile u_M, defined in (1.14). As we can observe, the solution given by (4.4) is already quite close to u_M. However, a non-suitable viscous numerical flux like, for instance, the modified Lax-Friedrichs (e.g. [38, Chapter 3]) can definitely modify the large-time behavior of the solution. In fact, in this case a viscosity proportional to $\Delta x^2/\Delta t$ is being added to the equation of the asymptotic profile (see Chapter 2), producing a more diffused wave. Nevertheless, the discretization of the non-linear term is not the only one with the ability to perturb the dynamics of the model. Let us emphasize that an inappropriate discretization of the non-local term also leads to an incorrect asymptotic profile. Note that in Figure 4.1 the same scheme (4.4) but taking $F_0^\Delta = F_1^\Delta = 1$ produces a translated solution.

The convergence rates, given in (4.8), are shown in Figure 4.2. The graphic highlights the different performances mentioned above. In fact, the solution given by (4.4) is the only one for which the norm is converging to zero with the corresponding rates.

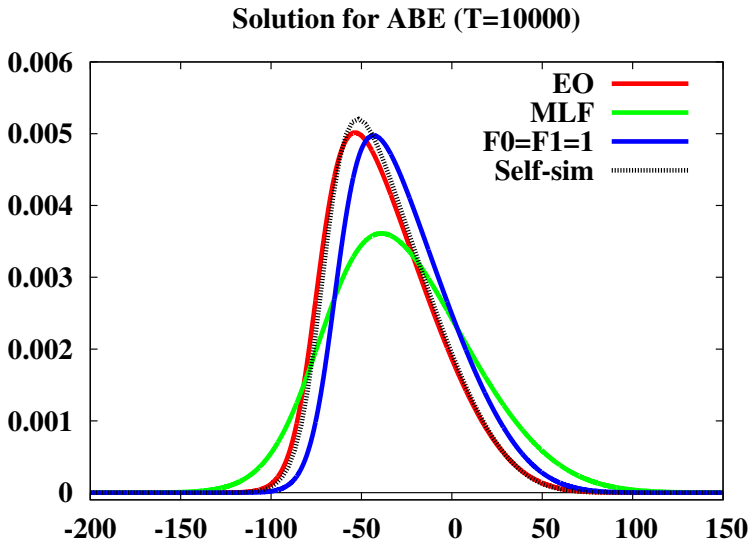


FIGURE 4.1: Solution of ABE with $\nu = 10^{-2}$, $c = 2 \times 10^{-2}$ and $\theta = 1$ at $t = 10^4$, using scheme (4.4) discretized explicitly. We use Engquist-Osher (red) and modified Lax-Friedrichs (green) numerical fluxes for the nonlinearity, as well as no correcting factors $F_0^\Delta = F_1^\Delta = 1$ (blue), comparing the solutions to the asymptotic profile (dotted).

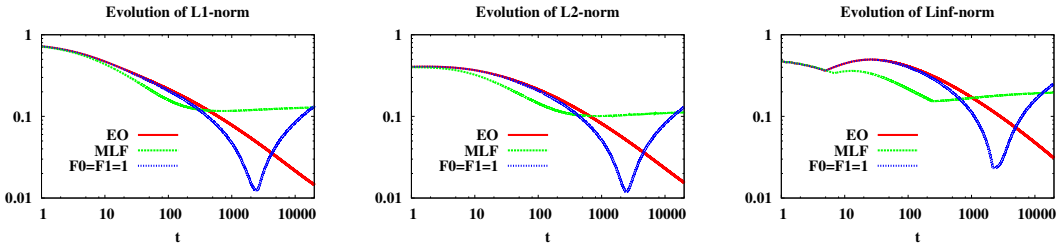


FIGURE 4.2: Evolution of the norms of the difference between the asymptotic profile and the solutions, multiplied by their corresponding rate $t^{\frac{1}{2}(1-\frac{1}{p})}$. From left to right, L^1, L^2 and L^∞. We compare (4.4) (red), LFM numerical flux (green) and $F_0^\Delta = F_1^\Delta = 1$ (blue).

To conclude, let us remark again the importance of taking a well-behaving numerical flux. In this paper we have proved that the asymptotic profile of (4.3) is a diffusive wave. Therefore, any sign-changing initial data will lose its positive or negative part, depending on the sign of its mass. As in the case of the viscous Burgers equation ([60]), simulations show that N-waves are intermediate states. Therefore, if the numerical viscosity is sufficiently large, the diffusion will become dominant much earlier than in the continuous model and destroy these profiles. For instance, let us consider the case $\nu = 10^{-4}$ and $c = 2 \times 10^{-4}$. In Figure 4.3, we can observe that at $t = 100$ the N-wave

shape is not preserved if the modified Lax-Friedrichs flux is used, while Engquist-Osher is able to keep the continuous dynamics.

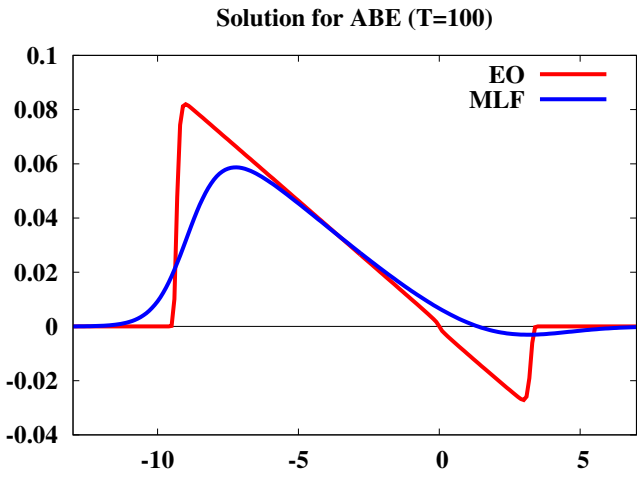


FIGURE 4.3: Solution of ABE with $\nu = 10^{-4}$, $c = 2 \times 10^{-4}$ and $\theta = 1$ at $t = 100$, using scheme (4.4) discretized explicitly. We use Engquist-Osher (red) and modified Lax-Friedrichs (blue) numerical fluxes for the nonlinearity.

Chapter 5

Operator splitting for the augmented Burgers equation

5.1 Introduction

In this chapter we analyze a splitting method for the augmented Burgers equation. We consider the following equation:

$$\begin{cases} u_t - \left(\dfrac{u^2}{2}\right)_x = u_{xx} + K * u - u + u_x, & (t,x) \in (0,\infty) \times \mathbb{R}, \\ u(0,x) = u_0(x), & x \in \mathbb{R}, \end{cases} \tag{5.1}$$

where $K(z) = e^{-z}\chi_{(0,\infty)}$. Roughly, the splitting method consists in writing the evolution operator corresponding to (5.1) as a sum of two evolution operators, X^t and Y^t, which are defined below. Then, the complete solution of the model is computed by applying both operators successively to the initial data. We refer to [44] for a more detailed introduction on operator splitting methods. Moreover, in Appendix B we use this technique to produce a numerical example for the full augmented Burgers equation.

In our case, we analyze a Trotter formula. Let Z^t be the flow given by $Z^t = X^t Y^t$, where X^t is the evolution operator associated with the following equation

$$\begin{cases} v_t = K * v - v + v_x, & (t,x) \in (0,\infty) \times \mathbb{R}, \\ v(0,x) = v_0(x), & x \in \mathbb{R}, \end{cases} \tag{5.2}$$

and Y^t, the one corresponding to the viscous Burgers equation

$$\begin{cases} w_t - \left(\frac{w^2}{2}\right)_x = w_{xx}, & (t,x) \in (0,\infty) \times \mathbb{R}, \\ w(0,x) = w_0(x), & x \in \mathbb{R}. \end{cases} \tag{5.3}$$

The aim is to approximate the solution u of (5.1) by alternating both operators; that is,

$$S^{n\Delta t}u_0 = u(n\Delta t) \approx (Z^{\Delta t})^n u_0 = (X^{\Delta t}Y^{\Delta t})^n u_0, \tag{5.4}$$

where S^t is the flow corresponding to (5.1), $\Delta t > 0$ and $n \in \mathbb{N} \cup \{0\}$. For the sake of simplicity on the notation, from now on we denote $Z^{n\Delta t} = (Z^{\Delta t})^n$. Let us recall that $t_n = n\Delta t$ and $t_{n+1/2} = (n + \frac{1}{2})\Delta t$ for every $n \in \mathbb{N} \cup \{0\}$.

Under sufficient conditions, Trotter formulas are known to be of first-order of accuracy [44]. The first result of this chapter, stated below, confirms this.

Theorem 5.1. *Let $r \in \{1, 2\}$. For all $u_0 \in H^r(\mathbb{R})$ and for all $T > 0$, there exist positive constants c_1, c_2 and Δt_0 such that, for all $\Delta t \in (0, \Delta t_0)$ and for all $n \in \mathbb{N}$ such that $0 \le n\Delta t \le T$,*

$$\left\| Z^{n\Delta t}u_0 - u(n\Delta t) \right\|_2 \le c_1 (\Delta t)^{r/2} \tag{5.5}$$

and

$$\left\| Z^{n\Delta t}u_0 \right\|_{H^r(\mathbb{R})} \le c_2. \tag{5.6}$$

Here, c_1, c_2 and Δt_0 depend on T and on $\|u_0\|_{H^r(\mathbb{R})}$.

We will give numerical evidence of this in the simulations that we carry out in Section 5.4. As an alternative to the Trotter formula, one could use the so-called Strang splitting, which is formally second-order accurate.

In order to analyze the large-time asymptotic behavior of the solution obtained by (5.4), we employ techniques similar to the ones from Chapter 4. Firstly, we define the piecewise function u^Δ by

$$u^\Delta(t,x) = \begin{cases} Y^{2(t-t_n)} Z^{n\Delta t}u_0(x), & t \in (t_n, t_{n+1/2}),\ x \in \mathbb{R}, \\ X^{2(t-t_{n+1/2})} Y^{\Delta t} Z^{n\Delta t}u_0(x), & t \in (t_{n+1/2}, t_{n+1}),\ x \in \mathbb{R}. \end{cases}$$

It is easy to check that in the time interval $[t_n, t_{n+1/2})$, u^Δ satisfies

$$\begin{cases} u^\Delta_t = 2(Y^{2(t-t_n)} Z^{n\Delta t}u_0)_t = 2u^\Delta_{xx} + \left((u^\Delta)^2\right)_x, & t \in (t_n, t_{n+1/2}), \\ u^\Delta(t_n) = Z^{n\Delta t}u_0. \end{cases}$$

Analogously, we can also show that

$$\begin{cases} u^\Delta_t = 2(K * u^\Delta - u^\Delta + u^\Delta_x), & t \in (t_{n+1/2}, t_{n+1}), \\ u^\Delta(t_{n+1/2}) = Y^{\Delta t} Z^{n\Delta t}u_0. \end{cases}$$

For the sake of clarity, let us denote $I_n^\Delta = (t_n, t_{n+1/2})$ and $\psi^{\Delta t}(t) = \sum_{n\geq 0} \chi_{I_n^\Delta}(t)$. We can now merge both equations into a single one:

$$\begin{cases} u_t^\Delta = 2\psi^{\Delta t}(t) u_{xx}^\Delta + \psi^{\Delta t}(t)((u^\Delta)^2)_x + 2\big(1-\psi^{\Delta t}(t)\big)(K * u^\Delta - u^\Delta + u_x^\Delta), \quad t>0, \\ u^\Delta(0) = u_0. \end{cases} \tag{5.7}$$

Remark 5.1*: It is simple to show that* $\psi^{\Delta t}(t) \rightharpoonup \frac{1}{2}$ *(see [23]). Thus, formally we have that* $u^\Delta \to u$ *as* $\Delta t \to 0$*, where* u *satisfies*

$$u_t = u_{xx} + \left(\frac{u^2}{2}\right) + K * u - u + u_x.$$

Note also that, by construction and the properties of X^t *and* Y^t*, we have that* $u^\Delta \in C([0,\infty); L^1(\mathbb{R}))$.

The second main result of the chapter confirms that the Trotter formula for the augmented Burgers equation preserves its large-time behavior.

Theorem 5.2. *For any* $u_0 \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ *and* $p \in [1,\infty)$

$$\big\|u^\Delta(t)\big\|_p \leq C_p t^{-\frac{1}{2}(1-\frac{1}{p})} \|u_0\|_{L^1(\mathbb{R})}, \quad t>0.$$

Moreover, for all $p \in [1,\infty)$*,*

$$t^{\frac{1}{2}(1-\frac{1}{p})} \big\|u^\Delta(t) - u_M(t)\big\|_p \to \infty, \quad \text{as } t\to\infty, \tag{5.8}$$

where $u_M(t,x) = t^{-1/2} f_M(x/\sqrt{t})$ *is the self-similar profile of the following viscous Burgers equation*

$$\begin{cases} u_t = \left(\frac{u^2}{2}\right)_x + 2u_{xx}, & (t,x) \in (0,\infty)\times\mathbb{R}, \\ u(0,x) = M\delta_0, & x\in\mathbb{R}. \end{cases} \tag{5.9}$$

The rest of the chapter is organized as follows. In Section 5.2 we show the convergence of the solution obtained with the splitting, as well as its accuracy order. Then, in Section 5.3, we obtain the decay estimates for the L^p norms of the solution and we use them to obtain the first term in the asymptotic expansion by a scaling argument. We conclude with some numerical experiments that illustrate our results in Section 5.4.

5.2 Convergence of the splitting

This section is devoted to the proof of Theorem 5.1. Thus, we first obtain some preliminary results, in order to compute local error estimates and, as a consequence, the order

of the splitting.

5.2.1 Estimates on X^t and Y^t

Let us denote D_t the kernel of the semigroup corresponding to (5.2). It follows that $X^t\varphi = D_t * \varphi$, for any $\varphi \in H^r(\mathbb{R})$ with $r \in \mathbb{N} \cup \{0\}$. In fact, X^t is stable in $H^r(\mathbb{R})$ and the local error can be estimated, at least, in terms of t.

Lemma 5.3. *Let $r \in \mathbb{N} \cup \{0\}$. For any $v_0 \in H^r(\mathbb{R})$, the following holds*

$$\|X^t v_0\|_{H^r(\mathbb{R})} \leq \|v_0\|_{H^r(\mathbb{R})}, \tag{5.10}$$

$$\|(X^t - I)v_0\|_2 \leq Ct^{\min\{1,r\}}\|v_0\|_{H^r(\mathbb{R})}, \tag{5.11}$$

for all $t \geq 0$ and a positive constant $C = C(r)$.

Proof. For the first estimate, using the Fourier transform (see Chapter 4), we have that

$$\widehat{D_t}(\xi) = e^{t(\widehat{K}(\xi)-1+i\xi)} = e^{-t\xi^2\left(\frac{1-i\xi}{1+\xi^2}\right)}. \tag{5.12}$$

Thus, it is enough to observe that $|\widehat{D_t}(\xi)| \leq 1$ for all $t \geq 0$ and use Plancherel's identity.

Regarding the local error estimate, integrating (5.2) in $(0,t)$ we obtain that

$$(X^t - I)v_0 = \int_0^t v_s(s)ds = \int_0^t \big(K * v(s) - v(s) + v_x(s)\big)ds$$

and, hence,

$$\|(X^t - I)v_0\|_2 \leq \int_0^t \|K * v(s) - v(s) + v_x(s)\|_2 ds$$

But, using Plancherel's identity and (5.12) we have

$$\begin{aligned}\|K * v - v + v_x\|_2^2 &= \int_{\mathbb{R}} \left|\frac{-\xi^2 + i\xi^3}{1+\xi^2}\right|^2 |\widehat{v}|^2 d\xi = \int_{\mathbb{R}} \frac{\xi^4}{1+\xi^2}|\widehat{v}|^2 d\xi \\ &\leq \int_{\mathbb{R}} \xi^2 \left|e^{-t(\frac{\xi^2 - i\xi^3}{1+\xi^2})}\right|^2 |\widehat{v_0}|^2 d\xi \leq \|v_0\|_{H^1(\mathbb{R})}^2.\end{aligned}$$

Therefore, we deduce that

$$\|(X^t - I)v_0\|_2 \leq t\,\|v_0\|_{H^1(\mathbb{R})}$$

Using (5.10) we deduce that the result is true for any $r = 0$. When $r > 1$, the result trivially follows from the case $r = 1$. □

Regarding the flow Y^t, all the estimates that we need can be found in [11] (similar estimates can also be obtained following [87]). Let us now recall from [11] the following result about the solution of system (5.3).

Proposition 5.4. *For any $w_0 \in H^r(\mathbb{R})$, $r \in \{1,2\}$ there exists a unique global solution of system (5.3) $w \in C([0,\infty), H^r(\mathbb{R}))$. Moreover, for any $T > 0$ there exists $C = C(T, \|w_0\|_2)$ such that*

$$\|Y^t w_0\|_{H^r(\mathbb{R})} \leq e^{Ct} \|w_0\|_{H^r(\mathbb{R})}, \quad \forall t \in [0,T].$$

5.2.2 Estimates on S^t

Let us now focus on the continuous flow S^t, for which we need to show that it is stable in $H^r(\mathbb{R})$ for any $r \in \mathbb{N} \cup \{0\}$ and uniformly Lipschitz on balls of $H^2(\mathbb{R})$.

Lemma 5.5. *Let $r \in \mathbb{N} \cup \{0\}$ and $T > 0$. For every $u_0 \in H^r(\mathbb{R})$, there exists a constant $C = C(T, \|u_0\|_{H^{r-1}(\mathbb{R})})$ such that its corresponding solution u of (5.1) satisfies*

$$\|u(t)\|_{H^r(\mathbb{R})} \leq C \|u_0\|_{H^r(\mathbb{R})}.$$

Proof. We proceed as in Chapter 4. In fact, the case $r = 0$ has been proved there already, taking $C = 1$. Differentiating the Duhamel formula in space, we know that u satisfies

$$u_x(t) = G_t * X^t(\partial_x u_0) + \frac{1}{2} \int_0^t \partial_x G_{t-s} * X^s(((u^2(s))_x) ds$$

For $r \geq 1$, using Young's inequality, Lemma 5.3 and the well-known decay estimates of the heat kernel, we infer

$$\begin{aligned} \|u_x(t)\|_{H^{r-1}(\mathbb{R})} &\leq \|G_t\|_1 \|X^t(\partial_x u_0)\|_{H^{r-1}(\mathbb{R})} + \frac{1}{2} \int_0^t \|\partial_x G_{t-s}\|_2 \|X^s(((u^2(s))_x)\|_{W^{r-1,1}(\mathbb{R})} ds \\ &\leq \|\partial_x u_0\|_{H^{r-1}(\mathbb{R})} + C \int_0^t (t-s)^{-\frac{3}{4}} \|u(s) u_x(s)\|_{W^{r-1,1}(\mathbb{R})} ds, \end{aligned}$$

where C only depends on r. Moreover, for $r = 1$, the Cauchy-Schwarz inequality and Lemma 5.3 yield

$$\begin{aligned} \|u_x(t)\|_{L^2(\mathbb{R})} &\leq \|\partial_x u_0\|_{L^2(\mathbb{R})} + C \int_0^t (t-s)^{-\frac{3}{4}} \|u(s)\|_{L^2(\mathbb{R})} \|u_x(s)\|_{L^2(\mathbb{R})} ds \\ &\leq \|\partial_x u_0\|_{L^2(\mathbb{R})} + C \|u_0\|_{L^2(\mathbb{R})} \int_0^t (t-s)^{-\frac{3}{4}} \|u_x(s)\|_{L^2(\mathbb{R})} ds. \end{aligned}$$

The fractional Gronwall Lemma from [11] implies the lemma in the case $r = 1$. Then, the case $r > 1$ easily follows by induction. □

Lemma 5.6. *Let $R, T > 0$. There exist $K = K(R,T) < \infty$ such that if*

$$\|u_0\|_{H^2(\mathbb{R})} \leq R \quad \text{and} \quad \|v_0\|_{H^2(\mathbb{R})} \leq R,$$

then

$$\|S^t u_0 - S^t v_0\|_2 \leq K\|u_0 - v_0\|_2, \quad \forall t \in [0,T].$$

Proof. Let u and v be the corresponding solutions of (5.1) for u_0 and v_0, respectively, and let us denote $w = u - v$. It is clear that

$$w_t = uu_x - vv_x + w_{xx} + K * w - w + w_x.$$

If we multiply this equation by w and integrate on $\mathbb{R}$, we get:

$$\frac{1}{2}\frac{d}{dt}\|w\|_2^2 \leq \int_{\mathbb{R}} (uu_x - vv_x)w dx = -\frac{1}{2}\int_{\mathbb{R}} (u+v)ww_x dx = \frac{1}{4}\int_{\mathbb{R}} (u_x + v_x)w^2 dx,$$

where the estimate for the non-local term was already obtained in Chapter 4. Therefore, due to the Sobolev embedding and Lemma 5.5, we conclude that

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|w(t)\|_2^2 &\leq \frac{1}{4}(\|u_x(t)\|_\infty + \|v_x(t)\|_\infty)\|w^2(t)\|_2^2 \\ &\leq C(\|u(t)\|_{H^2(\mathbb{R})} + \|v(t)\|_{H^2(\mathbb{R})})\|w^2(t)\|_2^2 \\ &\leq C(R,T)\|w^2(t)\|_2.\end{aligned}$$

The Gronwall lemma yields the result. □

5.2.3 Local error estimate

Before proving Theorem 5.1, we also need estimates on the L^2 local error for Z^t. We state this result in the following proposition.

Proposition 5.7. *Let $u_0 \in H^r(\mathbb{R})$ with $r \in \{1,2\}$. Then, for every $t \in (0,1)$ there exists a constant $C = C(\|u_0\|_2)$ such that*

$$\|u(t) - Z^t u_0\|_2 \leq C\|u_0\|_{H^r(\mathbb{R})}^2 t^{\alpha_r},$$

where $\alpha_r = \frac{3}{2}$ for $r = 1$ and $\alpha_r = 2$ for $r = 2$.

Proof. We proceed as in [11]. We know that the solution u of (5.1) satisfies

$$u(t) = D_t * G_t * u_0 + \frac{1}{2}\int_0^t D_{t-s} * G_{t-s} * \left(u^2(s)\right)_x ds,$$

where G_t is the kernel of the heat equation. On the other hand, we also have that

$$Z^t u_0 = X^t Y^t u_0 = D_t * Y^t u_0 = D_t * \left(G_t * u_0 + \frac{1}{2} \int_0^t G_{t-s} * \left((Y^s u_0)^2 \right)_x ds \right).$$

Therefore,

$$\begin{aligned} \|u(t) - Z^t u_0\|_2 &\leq \frac{1}{2} \int_0^t \Big\| D_{t-s} * G_{t-s} * \left(u^2(s) \right)_x \\ &\qquad - D_t * G_{t-s} * \left((Y^s u_0)^2 \right)_x \Big\|_2 ds \\ &\leq \frac{1}{2} \int_0^t \Big(\|R_1(s)\|_2 + \|R_2(s)\|_2 + \|R_3(s)\|_2 \Big) ds, \end{aligned} \tag{5.13}$$

where

$$\begin{aligned} R_1(s) &= D_{t-s} * G_{t-s} * \left(u^2(s) \right)_x - D_{t-s} * G_{t-s} * \left((Z^s u_0)^2 \right)_x \\ R_2(s) &= D_{t-s} * G_{t-s} * \left((Z^s u_0)^2 \right)_x - D_t * G_{t-s} * \left((Z^s u_0)^2 \right)_x \\ R_3(s) &= D_t * G_{t-s} * \left((Z^s u_0)^2 \right)_x - D_t * G_{t-s} * \left((Y^s u_0)^2 \right)_x \end{aligned}$$

Let us estimate each of these terms.

Estimate on R_1. Using Young's inequality and the decay estimates of the heat kernel, we have

$$\begin{aligned} \|R_1(s)\|_2 &\leq \|D_{t-s} * \partial_x G_{t-s}\|_2 \left\| u^2(s) - (Z^s u_0)^2 \right\|_1 \\ &\lesssim (t-s)^{-\frac{3}{4}} \|u(s) - Z^s u_0\|_2 \left(\|u(s)\|_2 + \|Z^s u_0\|_2 \right) \\ &\lesssim (t-s)^{-\frac{3}{4}} \|u(s) - Z^s u_0\|_2 \left(\|u_0\|_2 + \|Y^s u_0\|_2 \right) \end{aligned} \tag{5.14}$$

Estimate on R_2. For any $t \in (0,1)$, from Lemma 5.3 we have

$$\begin{aligned} \|R_2(s)\|_2 &= \|(D_{t-s} - D_t) * G_{t-s} * \left((Z^s u_0)^2 \right)_x \|_2 \\ &\lesssim s \|G_{t-s} * \left((Z^s u_0)^2 \right)_x \|_{H^1(\mathbb{R})} \lesssim s \|G_{t-s} * (Z^s u_0)^2\|_{H^2(\mathbb{R})} \\ &\lesssim s \|G_{t-s}\|_{W^{2-r,1}(\mathbb{R})} \|(Z^s u_0)^2\|_{H^r(\mathbb{R})} \lesssim s(t-s)^{-\frac{2-r}{2}} \|(Z^s u_0)^2\|_{H^r(\mathbb{R})}. \end{aligned}$$

Since $r > 1/2$, we have that $H^r(\mathbb{R})$ is an algebra and, hence,

$$\|R_2(s)\|_2 \lesssim s(t-s)^{-\frac{2-r}{2}} \|Z^s u_0\|^2_{H^r(\mathbb{R})} \lesssim s(t-s)^{-\frac{2-r}{2}} \|Y^s u_0\|^2_{H^r(\mathbb{R})}.$$

Integrating on the time interval $[0,t]$ we obtain

$$\begin{aligned} \int_0^t \|R_2(s)\|_2 &\lesssim \|Y^s u_0\|_{L^\infty((0,t),H^r(\mathbb{R}))} \int_0^t s(t-s)^{-\frac{2-r}{2}} ds \\ &\lesssim t^{1+\frac{r}{2}} \|Y^s u_0\|_{L^\infty((0,t),H^r(\mathbb{R}))}. \end{aligned} \tag{5.15}$$

Estimate on R_3. The way of proceeding is the same as for R_2. We have:

$$\begin{aligned}\|R_3(s)\|_2 &= \|D_t * G_{t-s} * \left((Z^s u_0)^2 - (Y^s u_0)^2\right)_x\|_2 \\ &\leq \|G_{t-s} * \left((Z^s u_0)^2 - (Y^s u_0)^2\right)\|_{H^1(\mathbb{R})} \\ &\leq \|G_{t-s}\|_{W^{2-r,1}(\mathbb{R})}\|(Z^s u_0)^2 - (Y^s u_0)^2\|_{H^{r-1}(\mathbb{R})} \\ &\lesssim (t-s)^{-\frac{2-r}{2}}\|\left((Z^s u_0)^2 - (Y^s u_0)^2\right)\|_{H^{r-1}(\mathbb{R})}\end{aligned}$$

On the one hand, since $H^1(\mathbb{R})$ is an algebra, for $r = 2$ we get:

$$\begin{aligned}\|R_3(s)\|_2 &\lesssim \|(Z^s u_0)^2 - (Y^s u_0)^2\|_{H^1(\mathbb{R})} \\ &\lesssim \|Z^s u_0 + Y^s u_0\|_{H^1(\mathbb{R})}\|Z^s u_0 - Y^s u_0\|_{H^1(\mathbb{R})} \\ &\lesssim \left(\|X^s Y^s u_0\|_{H^1(\mathbb{R})} + \|Y^s u_0\|_{H^1(\mathbb{R})}\right)\|X^s Y^s u_0 - Y^s u_0\|_{H^1(\mathbb{R})} \\ &\lesssim s\|Y^s u_0\|_{H^1(\mathbb{R})}\|Y^s u_0\|_{H^2(\mathbb{R})} \lesssim s\|Y^s u_0\|^2_{H^2(\mathbb{R})},\end{aligned}$$

where we have used Lemma 5.3. On the other hand, for the case $r = 1$, we simply have

$$\begin{aligned}\|R_3(s)\|_2 &\lesssim (t-s)^{-\frac{1}{2}}\|(Z^s u_0)^2 - (Y^s u_0)^2\|_2 \\ &\lesssim (t-s)^{-\frac{1}{2}}\left(\|Z^s u_0\|_\infty + \|Y^s u_0\|_\infty\right)\|Z^s u_0 - Y^s u_0\|_2 \\ &\lesssim s(t-s)^{-\frac{1}{2}}\left(\|Z^s u_0\|_{H^1(\mathbb{R})} + \|Y^s u_0\|_{H^1(\mathbb{R})}\right)\|Y^s u_0\|_{H^1(\mathbb{R})} \\ &\lesssim s(t-s)^{-\frac{1}{2}}\|Y^s u_0\|^2_{H^1(\mathbb{R})},\end{aligned}$$

where we have used the Sobolev embedding, as well as Lemma 5.3. Thus, setting $\alpha_r = \frac{3}{2}$ for $r = 1$ and $\alpha_r = 2$ for $r = 2$, we have:

$$\|R_3(s)\|_2 \lesssim s(t-s)^{\alpha_r - 2}\|Y^s u_0\|^2_{H^r(\mathbb{R})}$$

Integrating on the time interval $[0, t]$, we get

$$\begin{aligned}\int_0^t \|R_3(s)\|_2 &\lesssim \|Y^s u_0\|^2_{L^\infty((0,t),H^r(\mathbb{R}))}\int_0^t s(t-s)^{\alpha_r-2}ds \\ &\lesssim \|Y^s u_0\|^2_{L^\infty((0,t),H^r(\mathbb{R}))}t^{\alpha_r}. \qquad (5.16)\end{aligned}$$

Finally, plugging estimates (5.14), (5.15) and (5.16) into (5.13), we obtain

$$\begin{aligned}\|u(t) - Z^t u_0\|_2 \lesssim \frac{1}{2}\int_0^t (t-s)^{-\frac{3}{4}}\|u(s) - Z^s u_0\|_2\left(\|u_0\|_2 + \|Y^s u_0\|_2\right)ds \\ + t^{\alpha_r}\|Y^s u_0\|^2_{L^\infty((0,t),H^r(\mathbb{R}))}.\end{aligned}$$

Using Proposition 5.4 and applying Gronwall Lemma, we conclude

$$\|u(t) - Z^t u_0\|_2 \lesssim t^{\alpha_r}\|Y^s u_0\|^2_{L^\infty((0,t),H^r(\mathbb{R}))} \lesssim t^{\alpha_r}\|u_0\|^2_{H^r(\mathbb{R})}.$$

□

5.2.4 Proof of Theorem 5.1

We are now in condition to prove Theorem 5.1. It is enough to follow the same induction arguments in [11]. Let us denote $u^n = (Z^{\Delta t})^n u_0$ the semi-discrete solution and

$$u_m^n = S^{(n-m)\Delta t} u^m$$

the exact evolution of the value u^m at time $t_m = m\Delta t$ up to time $t_n = n\Delta t$.

First, from Lemma 5.5 we know that there exists $\rho > 0$, depending on $\|u_0\|_{H^r(\mathbb{R})}$ but independent of t, such that $\|u(t)\|_{H^r(\mathbb{R})} \leq \rho$ for all $t \in [0, T]$. We have to show that there exists $\gamma, \Delta t_0, c > 0$ such that if $0 < \Delta t \leq \Delta t_0$, for all $n \in \mathbb{N}$ with $n\Delta t \leq T$,

$$\|u^n\|_2 \leq 2\rho, \quad \|u^n - u(t_n)\|_2 \leq \gamma \Delta t \quad \text{and} \quad \|u^n\|_{H^r(\mathbb{R})} \leq C_T,$$

where $C_T = C(T, \|u_0\|_{H^r(\mathbb{R})} > 0$.

Obviously, this is true for $n = 0$. Let $m \geq 1$ and assume that it is also true for $0 \leq m \leq n-1$. Since $u^n = u_n^n$ and $u(t_n) = u_0^n$, we have

$$\begin{aligned} \|u^n - u(t_n)\|_2 &\leq \sum_{m=0}^{n-1} \|u_{m+1}^n - u_m^n\|_2 \\ &\leq \sum_{m=0}^{n-1} \|S^{(n-m-1)\Delta t}(Z^{\Delta t} u^m) - S^{(n-m-1)\Delta t}(S^{\Delta t} u^m)\|_2 \end{aligned} \tag{5.17}$$

Thanks to Lemma 5.6 and the induction assumption, for $m \leq n-1$ we know that

$$\begin{aligned} \|S^{\Delta t} u^m\|_2 &\leq \|S^{\Delta t} u^m - S^{\Delta t} u(t_m)\|_2 + \|S^{\Delta t} u(t_m)\|_2 \\ &\leq K\|u^m - u(t_m)\|_2 + \|u(t_m)\|_2 \leq K\gamma\Delta t + \rho, \end{aligned}$$

which is bounded by 2ρ if $0 < \Delta t \leq \Delta t_0$. Moreover,

$$\|S^{(n-m-1)\Delta t}(Z^{\Delta t} u^m) - S^{(n-m-1)\Delta t}(S^{\Delta t} u^m)\|_2 \leq K\|Z^{\Delta t} u^m - S^{\Delta t} u^m\|_2.$$

and, using the local error estimate,

$$\|S^{(n-m-1)\Delta t}(Z^{\Delta t} u^m) - S^{(n-m-1)\Delta t}(S^{\Delta t} u^m)\|_2 \leq KC\|u^m\|_{H^r(\mathbb{R})}^2 \Delta t^{\alpha_r}$$

Therefore, from (5.17) we conclude

$$\|u^n - u(t_n)\|_2 \leq KC\|u^m\|_{H^r(\mathbb{R})}^2 n\Delta t^{\alpha_r} \leq \gamma \Delta t^{\alpha_r - 1},$$

where γ is uniform in n and Δt. The last estimate of the induction follows immediately from Lemma 5.3 and Proposition 5.4.

5.3 Large-time behavior

In this section, we obtain the large-time asymptotic behavior of u^Δ. We follow the same scaling argument as in Chapter 4.

5.3.1 Time-decay estimates

First, we need to obtain decay estimates for the L^p norms of u^Δ, the solution of (5.7). We proceed as in [48, 49].

Lemma 5.8. *For all $p \in [1,\infty]$, there exists a constant $C = C(p, \|u^0\|_1, \|u^0\|_\infty) > 0$ such that*

$$\|u^\Delta(t)\|_p \le Ct^{-\frac{1}{2}(1-\frac{1}{p})}, \quad \forall t > 0,$$

for all solutions of equation (5.7) *with initial data $u_0 \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$.*

Proof. The case $p = 1$ follows immediately from the properties of X^t and Y^t. Now, let us focus on the case $p = 2$. Multiplying equation (5.7) by u^Δ and integrating on $\mathbb{R}$, we have

$$\frac{1}{2}\frac{d}{dt}\|u^\Delta\|_2^2 = -2\psi^{\Delta t}(t)\left\|u_x^\Delta\right\|_2^2 + 2(1-\psi^{\Delta t}(t))\int_{\mathbb{R}}(K * u^\Delta - u^\Delta)u^\Delta dx.$$

Now, note that for any function φ

$$\int_{\mathbb{R}}(K*\varphi-\varphi)\varphi = -\int_{\mathbb{R}}\int_{\mathbb{R}}\frac{K(x-y)+\tilde{K}(x-y)}{4}\Big(\varphi(x)-\varphi(y)\Big)^2 dxdy,$$

where $\tilde{K}(z) = K(-z)$. Let us denote $J(z) = \frac{1}{2}(K(z) + \tilde{K}(z))$. Note that J is an even function with mass one and satisfies $J(z) > 0$ in a neighborhood of the origin. Therefore,

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\int_{\mathbb{R}}|u^\Delta|^2dx =& -2\psi^{\Delta t}(t)\int_{\mathbb{R}}|u_x^\Delta|^2dx \\ &-(1-\psi^{\Delta t}(t))\int_{\mathbb{R}}\int_{\mathbb{R}}J(x-y)\Big(u^\Delta(x)-u^\Delta(y)\Big)^2dxdy.\end{aligned}$$

Since

$$\int_{\mathbb{R}}\int_{\mathbb{R}}J(x-y)\Big(\varphi(x)-\varphi(y)\Big)^2dxdy \lesssim \int_{\mathbb{R}}|\varphi_x|^2dx,$$

we deduce that

$$\frac{1}{2}\frac{d}{dt}\int_{\mathbb{R}}|u^{\Delta}|^2dx \lesssim -\int_{\mathbb{R}}\int_{\mathbb{R}}J(x-y)\Big(u^{\Delta}(x)-u^{\Delta}(y)\Big)^2dxdy.$$

Finally, following the same arguments as in [48] we conclude that

$$\|u^{\Delta}(t)\|_2 \leq C\|u_0\|_1 t^{-\frac{1}{2}(1-\frac{1}{2})}.$$

Let us remark that in that process $u^0 \in L^1(\mathbb{R})\cap L^{\infty}(\mathbb{R})$ is necessary.

The case $p \in (1,2)$ follows from the Holder inequality. Moreover, the same inductive argument in [48] are valid to get the L^p-estimate for $2 < p \leq \infty$ too. □

5.3.2 The family of rescaled solutions

Let us now introduce the family of rescaled functions. Given $\lambda > 0$, we define u^{λ} by

$$u^{\lambda}(t,x) = \lambda u^{\Delta}(\lambda^2 t, \lambda x). \tag{5.18}$$

From (5.7), it follows that

$$\begin{cases} u_t^{\lambda} = 2\psi^{\frac{\Delta t}{\lambda^2}}(t)u_{xx}^{\lambda} + \psi^{\frac{\Delta t}{\lambda^2}}(t)u^{\lambda}u_x^{\lambda} + 2\left(1-\psi^{\frac{\Delta t}{\lambda^2}}(t)\right)(\lambda^2 K_{\lambda} * u^{\lambda} - \lambda^2 u^{\lambda} + \lambda u_x^{\lambda}), \\ u_0^{\lambda}(x) = \lambda u_0(\lambda x), \end{cases} \tag{5.19}$$

where $K_{\lambda}(z) = \lambda K(\lambda z)$. Observe that

$$\psi^{\Delta t}(\lambda^2 t) = \sum_{n\geq 1}\chi_{(n\Delta t<\lambda^2 t<(n+\frac{1}{2})\Delta t)} = \sum_{n\geq 1}\chi_{(\frac{\Delta t}{\lambda^2}<t<(n+\frac{1}{2})\frac{\Delta t}{\lambda^2})} = \psi^{\frac{\Delta t}{\lambda^2}}(t)$$

Thus, as in the continuous case, note that it is immediate to extend Lemma 5.8 to the family $\{u^{\lambda}\}_{\lambda>0}$.

Lemma 5.9. *For any initial data $u_0 \in L^1(\mathbb{R})\cap L^{\infty}(\mathbb{R})$ and $p \in [1,\infty)$, there exists a constant $C = C(p, \|u^0\|_1, \|u^0\|_{\infty}) > 0$ such that*

$$\|u^{\lambda}(t)\|_p \leq Ct^{-\frac{1}{2}(1-\frac{1}{p})}, \quad \forall t > 0,$$

holds for any $\lambda > 0$.

Now we proceed as in [47] to obtain an estimate on the nonlocal term which will be equivalent to having estimates on the spatial derivative of u^{λ}.

Lemma 5.10. *For any $0 < t_1 < t_2 < \infty$, the following holds for any $\lambda 0$:*

$$\lambda^2 \int_{t_1}^{t_2} \int_{\mathbb{R}} \int_{\mathbb{R}} J_\lambda(x-y)(u^\lambda(x) - u^\lambda(y))^2 dxdydt \lesssim t_1^{-\frac{1}{2}}$$

Proof. Let us multiply (5.19) by u^λ and integrate on space and time. We get

$$\begin{aligned}
\frac{d}{dt}\int_{\mathbb{R}} |u^\lambda|^2 dx &= -2\psi^{\frac{\Delta t}{\lambda^2}}(t) \int_{\mathbb{R}} |u_x^\lambda|^2 dx + 2(1-\psi^{\frac{\Delta t}{\lambda^2}}(t)) \int_{\mathbb{R}} \lambda^2 (K_\lambda * u^\lambda - u^\lambda) u^\lambda dx \\
&= -2\psi^{\frac{\Delta t}{\lambda^2}}(t) \int_{\mathbb{R}} |u_x^\lambda|^2 dx \\
&\quad - (1-\psi^{\frac{\Delta t}{\lambda^2}}(t)) \int_{\mathbb{R}} \int_{\mathbb{R}} \lambda^2 J_\lambda(x-y)(u^\lambda(x) - u^\lambda(y))^2 dxdy.
\end{aligned}$$

Hence, integrating in time between t_1 and t_2,

$$\begin{aligned}
t_1^{-\frac{1}{2}} &\gtrsim \|u^\lambda(t_2)\|_2^2 - \|u^\lambda(t_1)\|_2^2 \\
&= 2\int_{t_1}^{t_2} \psi^{\frac{\Delta t}{\lambda^2}}(t) \int_{\mathbb{R}} |u_x^\lambda|^2 dxdt \\
&\quad + \int_{t_1}^{t^2} (1-\psi^{\frac{\Delta t}{\lambda^2}}(t)) \int_{\mathbb{R}} \int_{\mathbb{R}} \lambda^2 J_\lambda(x-y)(u^\lambda(x) - u^\lambda(y))^2 dxdydt.
\end{aligned}$$

Finally, using the fact that

$$\lambda^2 \int_{\mathbb{R}} \int_{\mathbb{R}} J_\lambda(x-y)(u^\lambda(x) - u^\lambda(y))^2 dxdy \lesssim \left(\int_{\mathbb{R}} J_\lambda(z) dz \right) \left(\int_{\mathbb{R}} |u_x^\lambda|^2 \right),$$

we conclude

$$\lambda^2 \int_{\mathbb{R}} \int_{\mathbb{R}} J_\lambda(x-y)(u^\lambda(x) - u^\lambda(y))^2 dxdy \leq t_1^{-\frac{1}{2}}.$$

□

It remains to prove a last estimate, regarding the $H^{-1}(\mathbb{R})$ norm, in order to apply Aubin-Lions Lemma and show the compactness of the set $\{u^\lambda\}_{\lambda>0}$.

Lemma 5.11. *For any $0 < t_1 < t_2 < \infty$, the following holds for any $\lambda > 0$:*

$$\|u_t^\lambda\|_{L^2(t_1,t_2;H^{-1}(\mathbb{R}))} \leq C(t_1).$$

Proof. Let us consider $\phi \in C_c^\infty(\mathbb{R}^2)$. Then

$$\begin{aligned}
< u_t^\lambda, \phi >_{-1,1} =& < 2\psi^{\frac{\Delta t}{\lambda^2}} u_{xx}^\lambda + 2(1-\psi^{\frac{\Delta t}{\lambda^2}}) u^\lambda u_x^\lambda \\
&+ 2(1-\psi^{\frac{\Delta t}{\lambda^2}}) \left(\lambda^2 \left(K_\lambda * u^\lambda - u^\lambda \right) + \lambda u_x^\lambda \right), \phi >_{-1,1}
\end{aligned}$$

$$= -2\psi^{\frac{\Delta t}{\lambda^2}} \int_{\mathbb{R}} u_x^\lambda \phi_x dx - 2(1-\psi^{\frac{\Delta t}{\lambda^2}}) \int_{\mathbb{R}} (u^\lambda)^2 \phi_x dx$$
$$+ 2(1-\psi^{\frac{\Delta t}{\lambda^2}}) \int_{\mathbb{R}} \left(\lambda^2 \left(K_\lambda * u^\lambda - u^\lambda\right) + \lambda u_x^\lambda\right) \phi dx.$$

It follows that

$$\left| \int_{t_1}^{t_2} < u_t^\lambda, \phi >_{-1,1} dt \right| \lesssim \int_{t_1}^{t_2} \psi^{\frac{\Delta t}{\lambda^2}} \left(\int_{\mathbb{R}} |u_x^\lambda|^2 \right)^{\frac{1}{2}} \|\phi\|_{H^1(\mathbb{R})} dxdt$$
$$+ \int_{t_1}^{t_2} (1-\psi^{\frac{\Delta t}{\lambda^2}}) \int_{\mathbb{R}} (u^\lambda)^2 |\phi_x| dxdt$$
$$+ \left| \int_{t_1}^{t_2} (1-\psi^{\frac{\Delta t}{\lambda^2}}) \int_{\mathbb{R}} \left(\lambda^2 \left(K_\lambda * u^\lambda - u^\lambda\right) + \lambda u_x^\lambda\right) \phi dx \right|$$
$$= \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3.$$

For $\mathcal{I}_3$ we remark that

$$\left| \int_{\mathbb{R}} \left(\lambda^2 \big(K_\lambda * u^\lambda - u^\lambda\big) + \lambda u_x^\lambda\right) \phi dx \right| = \left| \int_{\mathbb{R}} \left(\lambda^2 (\hat{K}_\lambda(\xi) - 1) + \lambda i \xi\right) \widehat{u^\lambda}\, \widehat{\phi}\, d\xi \right|$$
$$= \left| \int_{\mathbb{R}} \frac{\lambda^2 (\hat{K}_\lambda(\xi) - 1) + \lambda i \xi}{\xi} \widehat{u^\lambda}\, \xi\, \widehat{\phi}\, d\xi \right| \leq \int_{\mathbb{R}} \left| \frac{\xi \lambda}{\sqrt{\lambda^2 + \xi^2}} \right| |\widehat{u^\lambda}| |\xi| |\widehat{\phi}| d\xi$$
$$\leq \left(\int_{\mathbb{R}} \left| \frac{\xi^2 \lambda^2}{\lambda^2 + \xi^2} \right| |\widehat{u^\lambda}|^2 d\xi \right)^{\frac{1}{2}} \|\phi\|_{H^1(\mathbb{R})} \leq \left(\int_{\mathbb{R}} (u^\lambda - J_\lambda * u^\lambda) u^\lambda dx \right)^{\frac{1}{2}} \|\phi\|_{H^1(\mathbb{R})}$$
$$\leq \left(\lambda^2 \int_{\mathbb{R}} \int_{\mathbb{R}} J(x-y)(u^\lambda(x) - u^\lambda(y)) dxdy \right)^{\frac{1}{2}} \|\phi\|_{H^1(\mathbb{R})}$$

Hence,

$$\mathcal{I}_3^2 \leq \lambda^2 \|\phi\|^2_{L^2(t_1,t_2;H^1(\mathbb{R}))} \int_{t_1}^{t_2} (1-\psi^{\frac{\Delta t}{\lambda^2}}) \int_{\mathbb{R}} J_\lambda(x-y)(u^\lambda(x) - u^\lambda(y))^2 dxdy$$

Moreover, $\mathcal{I}_1$ satisfies

$$\mathcal{I}_1^2 \leq \|\phi\|^2_{L^2(t_1,t_2;H^1(\mathbb{R}))} \int_{t_1}^{t_2} \psi^{\frac{\Delta t}{\lambda^2}}(t) \int_{\mathbb{R}} (u_x^\lambda)^2 dxdt$$

Then

$$\mathcal{I}_1^2 + \mathcal{I}_3^2 \leq \|\phi\|^2_{L^2(t_1,t_2;H^1(\mathbb{R}))} \int_{t_1}^{t_2} \psi^{\frac{\Delta t}{\lambda^2}} \int_{\mathbb{R}} (u_x^\lambda)^2 dxdt$$
$$+ \|\phi\|^2_{L^2(t_1,t_2;H^1(\mathbb{R}))} \int_{t_1}^{t_2} (1-\psi^{\frac{\Delta t}{\lambda^2}}) \int_{\mathbb{R}} \int_{\mathbb{R}} J_\lambda(x-y)\big(u^\lambda(x) - u^\lambda(y)\big) dxdydt$$
$$\leq \|\phi\|^2_{L^2(t_1,t_2;H^1(\mathbb{R}))} \|u^\lambda(t_1)\|_2^2 \lesssim t_1^{-\frac{1}{2}}$$

On the other hand, for $\mathcal{I}_2$ we have

$$|\mathcal{I}_2| \leq \|(u^\lambda)^2\|^2_{L^2(t_1,t_2;L^2(\mathbb{R}))}\|\phi\|^2_{L^2(t_1,t_2;H^1(\mathbb{R}))} \leq \|u^\lambda\|^2_{L^2(t_1,t_2;L^4(\mathbb{R}))}\|\phi\|^2_{L^2(t_1,t_2;H^1(\mathbb{R}))} \quad (5.20)$$

In conclusion, $\mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3$ is uniformly bounded with respect to λ and, thus, u^λ_t is uniformly bounded in $L^2(t_1, t_2; H^{-1}(\mathbb{R}))$. □

5.3.3 Proof of Theorem 5.2

We can now prove the last main result of this chapter. Since the above estimates are obtained, the compactness tool in [46] and [47] gives us the existence of a function $\bar{u} \in L^2_{loc}((0,\infty); H^1(\mathbb{R}))$ such that

$$\begin{gathered} u^\lambda \to \bar{u}, \text{ in } L^1_{loc}((0,\infty)\times\mathbb{R}) \\ u^\lambda \to \bar{u}, \text{ a.e. on } (0,\infty)\times\mathbb{R} \\ u^\lambda(t) \to \bar{u}(t), \text{ in } L^2(\mathbb{R}) \\ (u^\lambda)^2 \rightharpoonup \bar{u}^2, \text{ in } L^p_{loc}((0,\infty)\times\mathbb{R}) \end{gathered}$$

Therefore, we have that $\{u^\lambda\}_{\lambda>0}$ is compact in $L^1(\tau, T; L^1_{loc}(\mathbb{R}))$ for any $0 < \tau < T$. Repeating the process of the third step of Theorem 4.8 in Chapter 4, where we show that

$$\int_\tau^T \|u^\lambda(t)\psi_r\|_1 dt \to 0, \quad \text{as } r \to \infty, \text{ uniformly on } \lambda \geq 1,$$

one extends the compactness to $L^1(\tau, T; L^1(\mathbb{R}))$.

Finally, it is enough to repeat the arguments in the proof of Theorem 4.1 in Chapter 4 to show that $\bar{u}$ is a weak solution of (5.9). In a similar manner, we can also identify its behavior as $t \to 0$. Regarding estimate (5.8), it follows immediately from the convergence in L^1 and the time-decay estimates of Lemma 5.8.

5.4 Numerical examples

To conclude this chapter, we show some numerical simulations that illustrate the analytical results that we have proved. First, let us remark that the use of splitting methods makes the numerical resolution of the nonlocal term easier. Note, for instance, that sub-equation (5.2) can be rewritten in the following way:

$$\begin{cases} v_t + v_{tx} = v_{xx}, & (t,x) \in (0,\infty)\times\mathbb{R}, \\ v(0,x) = v_0(x), & x \in \mathbb{R}, \end{cases}$$

Now we can simply use centered finite differences in space and Crank-Nicolson in time (e.g. [52]). In fact, this second order scheme, formally preserves the scales given by (5.18) and can be efficiently solved using any tridiagonal matrix solver. On the other hand, for sub-equation (5.3) we use Engquist-Osher numerical flux and centered finite differences for the diffusion (see Chapter 3). We take a mesh size $\Delta x = 0.1$. In order to avoid boundary issues, we choose a large enough spatial domain.

Since we do not know the exact solution of (5.1), we determine the convergence order by comparing two solutions for the same initial data but a different time step. That is, we compute $\|u_1(T) - u_2(T)\|_2$, where u_1 and u_2 are obtained using Δt and $\Delta t/2$ time-steps, respectively. In Figure 5.2 we observe that the numerical simulations at $T = 10$ corresponding to two different initial data (see Figure 5.1) are consistent with the theoretical results of Section 5.2.

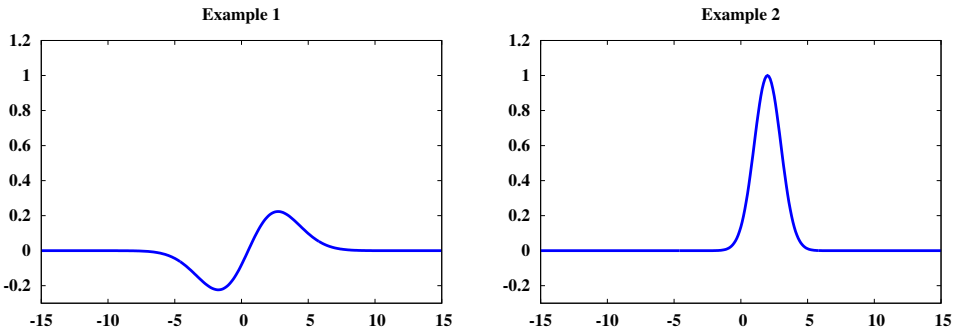


FIGURE 5.1: The two different initial data that we consider.

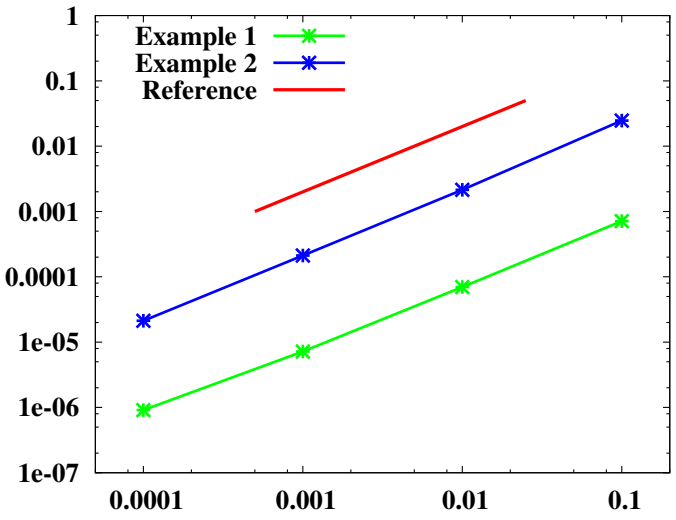


FIGURE 5.2: Accuracy order for the initial data shown in Figure 5.1.

Furthermore, in Figure 5.3 we plot a solution for $\nu = 10^{-2}$, $c = 2 \times 10^{-2}$ and $\theta = 1$ at time $t = 1.2 \times 10^4$, as well as the corresponding asymptotic profile u_M, defined in

(1.14). The convergence rates, defined in (5.8), are shown in Figure 5.3. Both graphics compare the splitting method with the discretization (4.4) of Chapter 4.

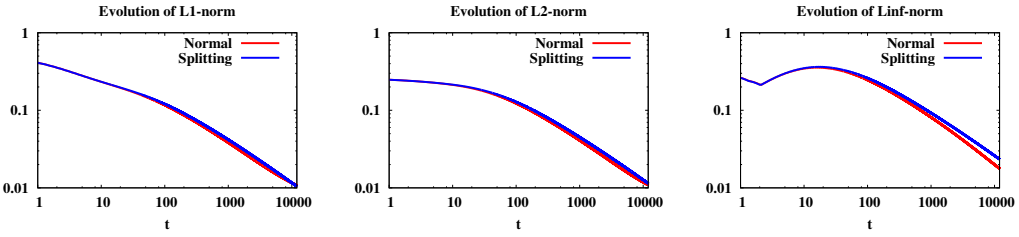


FIGURE 5.3: Evolution of the norms of the difference between the asymptotic profile and the solution, multiplied by their corresponding rate, as in (5.8). From left to right, L^1, L^2 and L^∞. We compare the splitting method (blue) and the numerical scheme (4.4) discretized explicitly (red).

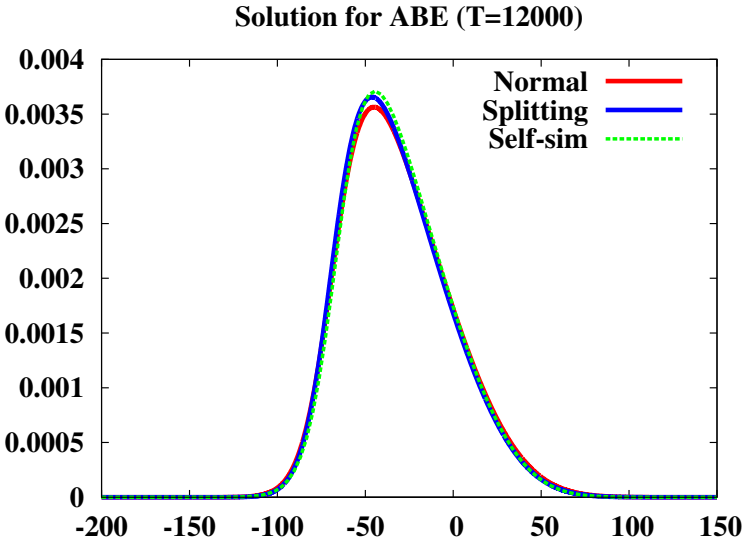


FIGURE 5.4: Solution of ABE with $\nu = 10^{-2}$, $c = 2 \times 10^{-2}$ and $\theta = 1$ at $T = 1.2 \times 10^4$. We compare the splitting method (blue) with the numerical scheme (4.4) discretized explicitly (red) and the self-similar asymptotic profile (green).

Chapter 6

Conclusions and open problems

In this Thesis we have applied theoretical results about the asymptotic behavior of solutions to improve the efficiency of numerical algorithms in problems involving large-time horizons. In particular, we have focused on the sonic boom phenomenon. We conclude with some final remarks and perspectives related to these subjects.

- In Chapter 2 we have studied the large-time behavior of some numerical schemes for the inviscid Burgers equation. Our results rely on the one-sided Lipschitz condition that the chosen numerical fluxes satisfy. In fact, it has been critical to obtain all the estimates in the L^p norms. To the best of our knowledge, there is not any general result to guarantee that a numerical scheme is consistent with the OSLC.

  Therefore, it is worth looking for an alternative proof in which this condition is not used, in order to extend the analysis to other methods. Higher order methods for conservation laws (MUSCL, ENO, WENO...) might not undergo this pathology. But it would be interesting to study, for instance, the large-time behavior of finite element methods (for example, discrete Galerkin methods [21]) or spectral ones [2]). The multidimensional case is particularly challenging and still open too.

  The extension to the L^∞ initial data framework remains open too. The existence of rarefaction and traveling waves at the continuous level is proved in [50, 51, 41]. The equivalent discrete version is already shown in [41] for rarefaction waves, but the decay rates are not optimal.

- In Chapter 3 we have applied the results of Chapter 2 to an inverse design problem involving the Burgers equation. We have observed that the efficiency of optimization algorithms can be negatively affected if the discretization scheme does not preserve the large-time dynamics of the PDE. The same analysis could be done in many other frameworks. Many other situations (parameter identification [54],

data assimilation [7], flux identification [53]... to name but a few) may be chosen to observe the importance of mimicking the dynamics of the equation.

The development of techniques that handle the presence of shocks (or quasi-shocks) in optimization problems of this type is still quite open. In this sense, the application of alternated descent methods like the one introduced in [17] can be very useful. Moreover, its extension to viscous systems proposed in [18] could be helpful in problems with non-reachable target functions. In particular, as we show at the end of Chapter 3, the classical adjoint methodology can fail when non-entropic regions appear in the target. Thus, some additional technique is required in order to avoid local minimizers.

- In Chapter 4 we have obtained the first term of the asymptotic expansion of the solution of a simplified version of the augmented Burgers equation. Moreover, we propose a semidiscrete numerical scheme that preserves this large-time behavior. Simulations suggest that the same results are true at the fully discrete level, but the proof requires further developments.

  On the other hand, results regarding non-local operators in scalar conservation laws usually involve C^1 symmetric kernels (see for instance, [46, 71]). It would be interesting to extend those results to more general kernels, like the one appearing in the augmented Burgers equation.

- In Chapter 5 we have studied the large-time behavior of solutions of the augmented Burgers equation when applying a splitting method. The analysis is done at the continuos level, but employing the same techniques as in Chapter 4 the extension to semidiscrete schemes is possible too.

  Specifically, it would be worth studying the equation

$$\begin{cases} u_t = K * u_{xx}, & (t,x) \in (0,\infty,\mathbb{R}), \\ u(0,x) = u^0(x), & x \in \mathbb{R}, \end{cases}$$

  or, equivalently,

$$\begin{cases} u_t + u_{tx} = u_{xx}, & (t,x) \in (0,\infty,\mathbb{R}), \\ u(0,x) = u^0(x), & x \in \mathbb{R}. \end{cases}$$

  It is particularly interesting the analysis at the continuous level of the second version. Firstly, it avoids integrals on infinite domains and, secondly, it could be useful to obtain the large-time behavior of solutions at the discrete level. If one opts for the first version, inspired by [11] one could use the Fast Fourier Transform to compute the convolution.

The study of the splitting method for the full augmented Burgers equation remains open as well. In this case, since a change of variable is required at the continuous level (see Appendix B), the extension may require further developments.

To sum up, we highlight that the preservation of the large time dynamical properties of the continuous system at the discrete level is critical. We have made clear that forward simulations in large-time horizons can be distorted. Moreover, the efficiency of solutions to optimization and inverse design problems might be affected too. This is particularly important in the problem of the sonic-boom minimization, where N-waves arise naturally.

Appendix A

Auxiliary results

A.1 Auxiliary results

Here we prove some of the auxiliary results that we have used along this Thesis.

Lemma A.1. *For any piecewise constant function w defined as in* (1.19) *and $\Delta x > 0$, the following holds:*

$$\|w\|_p^{p(p+1)/(p-1)} \leq 4\|w\|_1^{2p/(p-1)} \|d_{\Delta x}^+ |w|^{p/2}\|_2^2$$

for all $p \in (1, \infty)$.

Proof. First, let us define a piecewise linear function v as follows:

$$v(x) := w_j \frac{x_{j+1} - x}{\Delta x} + w_{j+1} \frac{x - x_j}{\Delta x}, \quad x \in [x_j, x_{j+1}].$$

On the one hand, we know that

$$\|v\|_\infty^2 \leq 2\|v\|_2 \|v_x\|_2. \tag{A.1}$$

On the other hand, we have that:

$$\|v\|_2^2 = \Delta x \sum_{j \in \mathbb{Z}} \int_0^1 |w_j(1-x) + w_{j+1}x|^2 \, dx \leq \frac{1}{2} \Delta x \sum_{j \in \mathbb{Z}} \left(|w_j|^2 + |w_{j+1}|^2 \right) = \|w\|_2^2.$$

Moreover, it is easy to see that $\|v_x\|_2 = \|d_{\Delta x}^+ w\|_2$. Therefore, we can obtain a similar inequality as (A.1) for w:

$$\|w\|_\infty^2 = \|v\|_\infty^2 \leq 2\|v\|_2 \|v_x\|_2 \leq 2\|w\|_2 \|d_{\Delta x}^+ w\|_2. \tag{A.2}$$

Applying this inequality to $|w|^{p/2}$, we deduce:

$$\|w\|_\infty^{2p} = \||w|^{p/2}\|_\infty^4 \leq 4\||w|^{p/2}\|_2^2 \|d_{\Delta x}^+ |w|^{p/2}\|_2^2 = 4\|w\|_p^p \|d_{\Delta x}^+ |w|^{p/2}\|_2^2.$$

Thus, combining this with

$$\|w\|_p^{2p^2/(p-1)} \leq \|w\|_\infty^{2p} \|w\|_1^{2p/(p-1)},$$

we conclude

$$\|w\|_p^{p(p+1)/(p-1)} \leq 4\|w\|_1^{2p/(p-1)} \|d_{\Delta x}^+ |w|^{p/2}\|_2^2.$$

□

Proof of Lemma 4.12. For the first assertion, we simply integrate (4.34) over the whole space domain. We observe that all terms on the right hand side vanish, so

$$\frac{d}{dt}\int_{\mathbb{R}} u^\mu(t,x)dx = 0, \quad \forall t \geq 0,$$

for all $\mu > 0$ and, hence, the mass is conserved. Using the definition of u^μ, we conclude

$$\int_{\mathbb{R}} u^\mu(t,x)dx = \int_{\mathbb{R}} u^\mu(0,x)dx = \int_{\mathbb{R}} \mu u_\Delta^0(\mu x)dx = \int_{\mathbb{R}} u_\Delta^0(x)dx.$$

For the contractivity we prove that for any $u_0, v_0 \in L^1(\mathbb{R})$, their corresponding solutions u^μ and v^μ satisfy

$$\|u^\mu - v^\mu\|_1 \leq \|u_0^\mu - v_0^\mu\|_1. \tag{A.3}$$

For the sake of clarity, let us define $w^\mu = u^\mu - v^\mu$. Clearly, w^μ verifies

$$\begin{aligned} w_t^\mu(t,x) = {} & \frac{1}{4}\Big(d_{\Delta x/\mu}^+\big(u^\mu(t,x)^2\big) + d_{\Delta x/\mu}^-\big(u^\mu(t,x)^2\big)\Big) \\ & - \frac{1}{4}\Big(d_{\Delta x/\mu}^+\big(v^\mu(t,x)^2\big) - d_{\Delta x/\mu}^-\big(v^\mu(t,x)^2\big)\Big) \\ & + \Delta x\, d_{\Delta x/\mu}^+ R\big(u^\mu(t,x-\frac{\Delta x}{\mu}), u^\mu(t,x)\big) \\ & - \Delta x\, d_{\Delta x/\mu}^+ R\big(v^\mu(t,x-\frac{\Delta x}{\mu}), v^\mu(t,x)\big) + d_{\Delta x/\mu}^-\Big(d_{\Delta x/\mu}^+ w^\mu(t,x)\Big) \\ & + \mu^2 \sum_{m=1}^N \omega_m w^\mu(t,x-m\frac{\Delta x}{\mu}) - \mu^2 F_0^\Delta w^\mu(t,x) + \mu F_1^\Delta d_{\Delta x/\mu}^+ w^\mu(t,x). \end{aligned}$$

We multiply it by $\operatorname{sign}(w^\mu)$ and integrate it on all $\mathbb{R}$. Using the definition of R in (4.35) and reordering the terms we get

$$\begin{aligned} \frac{d}{dt}\int_{\mathbb{R}} |w^\mu(x)|dx = \frac{1}{4}\int_{\mathbb{R}} d_{\Delta x/\mu}^+\big(u^\mu(x)^2 + u^\mu(x-\frac{\Delta x}{\mu})^2 + u^\mu(x)|u^\mu(x)| \\ - u^\mu(x-\frac{\Delta x}{\mu})|u^\mu(x-\frac{\Delta x}{\mu})|\big)\operatorname{sign}(w^\mu(x))dx \end{aligned} \tag{A.4}$$

$$
\begin{aligned}
&- \frac{1}{4}\int_{\mathbb{R}} d^+_{\Delta x/\mu}\big(v^\mu(x)^2 + v^\mu(x-\frac{\Delta x}{\mu})^2 + v^\mu(x)|v^\mu(x)| \\
&\qquad\qquad - v^\mu(x-\frac{\Delta x}{\mu})|v^\mu(x-\frac{\Delta x}{\mu})|\big)\,\mathrm{sign}(w^\mu(x))dx \\
&+ \int_{\mathbb{R}} d^-_{\Delta x/\mu}(d^+_{\Delta x/\mu}(w^\mu(x)))\,\mathrm{sign}(w^\mu(x))dx \\
&+ \mu^2 \sum_{m=1}^{N} \omega_m \int_{\mathbb{R}} \big(w^\mu(x-m\frac{\Delta x}{\mu}) - w^\mu(x)\big)\,\mathrm{sign}(w^\mu(x))dx \\
&+ \mu F_1^\Delta \int_{\mathbb{R}} d^+_{\Delta x/\mu} w^\mu(t,x)\,\mathrm{sign}(w^\mu(x))dx \\
&= I_1 + I_2 + I_3 + I_4 + I_5.
\end{aligned}
$$

For $i = 0, 1$, let us denote $W_i^\pm = \{x \in \mathbb{R} : \pm w^\mu(x - i\Delta x) > 0\}$ and $W_i^0 = \{x \in \mathbb{R} : w^\mu(x - i\Delta x) = 0\}$. Now we can split the domains of the integrals into several parts, according to the sign of w^μ. On the one hand, we have:

$$
\begin{aligned}
&I_1 + I_2 = \\
&= -\frac{1}{4}\int_{\mathbb{R}} \big(u^\mu(x)^2 + u^\mu(x)|u^\mu(x)|\big) d^-_{\Delta x/\mu}(\mathrm{sign}(w^\mu(x)))dx \\
&\quad - \frac{1}{4}\int_{\mathbb{R}} \big(u^\mu(x-\frac{\Delta x}{\mu})^2 - u^\mu(x-\frac{\Delta x}{\mu})|u^\mu(x-\frac{\Delta x}{\mu})|\big) d^-_{\Delta x/\mu}(\mathrm{sign}(w^\mu(x)))dx \\
&\quad + \frac{1}{4}\int_{\mathbb{R}} \big(v^\mu(x)^2 + v^\mu(x)|v^\mu(x)|\big) d^-_{\Delta x/\mu}(\mathrm{sign}(w^\mu(x)))dx \\
&\quad + \frac{1}{4}\int_{\mathbb{R}} \big(v^\mu(x-\frac{\Delta x}{\mu})^2 - v^\mu(x-\frac{\Delta x}{\mu})|v^\mu(x-\frac{\Delta x}{\mu})|\big) d^-_{\Delta x/\mu}(\mathrm{sign}(w^\mu(x)))dx \\
&= -\frac{\mu}{2\Delta x}\int_{W_0^-\cap W_1^+} \big(u^\mu(x)^2 + u^\mu(x)|u^\mu(x)| - v^\mu(x)^2 - v^\mu(x)|v^\mu(x)|\big)\,\mathrm{sign}(w^\mu(x))dx \\
&\quad - \frac{\mu}{2\Delta x}\int_{W_0^-\cap W_1^+} \big(v^\mu(x-\frac{\Delta x}{\mu})^2 - v^\mu(x-\frac{\Delta x}{\mu})|v^\mu(x-\frac{\Delta x}{\mu})| \\
&\qquad\qquad - u^\mu(x-\frac{\Delta x}{\mu})^2 + u^\mu(x-\frac{\Delta x}{\mu})|u^\mu(x-\frac{\Delta x}{\mu})|\big)\,\mathrm{sign}(w^\mu(x-\frac{\Delta x}{\mu}))dx \\
&\quad - \frac{\mu}{2\Delta x}\int_{W_0^+\cap W_1^-} \big(u^\mu(x)^2 + u^\mu(x)|u^\mu(x)| - v^\mu(x)^2 - v^\mu(x)|v^\mu(x)|\big)\,\mathrm{sign}(w^\mu(x))dx \\
&\quad - \frac{\mu}{2\Delta x}\int_{W_0^+\cap W_1^-} \big(v^\mu(x-\frac{\Delta x}{\mu})^2 - v^\mu(x-\frac{\Delta x}{\mu})|v^\mu(x-\frac{\Delta x}{\mu})| \\
&\qquad\qquad - u^\mu(x-\frac{\Delta x}{\mu})^2 + u^\mu(x-\frac{\Delta x}{\mu})|u^\mu(x-\frac{\Delta x}{\mu})|\big)\,\mathrm{sign}(w^\mu(x-\frac{\Delta x}{\mu}))dx \\
&\quad - \frac{\mu}{4\Delta x}\int_{W_1^0} \big(u^\mu(x)^2 + u^\mu(x)|u^\mu(x)| - v^\mu(x)^2 - v^\mu(x)|v^\mu(x)|\big)\,\mathrm{sign}(w^\mu(x))dx \\
&\quad - \frac{\mu}{4\Delta x}\int_{W_0^0} \big(v^\mu(x-\frac{\Delta x}{\mu})^2 - v^\mu(x-\frac{\Delta x}{\mu})|v^\mu(x-\frac{\Delta x}{\mu})| \\
&\qquad\qquad - u^\mu(x-\frac{\Delta x}{\mu})^2 + u^\mu(x-\frac{\Delta x}{\mu})|u^\mu(x-\frac{\Delta x}{\mu})|\big)\,\mathrm{sign}(w^\mu(x-\frac{\Delta x}{\mu}))dx.
\end{aligned}
$$

Using that

$$\big(b(b+|b|) - a(a+|a|)\big)\,\mathrm{sign}(b-a) \geq 0, \quad \forall a, b \in \mathbb{R},$$

and that

$$\big(a(a-|a|) - b(b-|b|)\big)\,\mathrm{sign}(b-a) \geq 0, \quad \forall a, b \in \mathbb{R},$$

we conclude that $I_1 + I_2 \leq 0$. On the other hand, since

$$\int_{\mathbb{R}} w^\mu(x - m\frac{\Delta x}{\mu})\,\mathrm{sign}(w^\mu(x))dx \leq \int_{\mathbb{R}} |w^\mu(x)|dx, \quad \forall m \in \mathbb{Z},$$

it is immediate that

$$\begin{aligned} I_3 = \frac{\mu^2}{\Delta x^2}\int_{\mathbb{R}} w^\mu(x - \frac{\Delta x}{\mu})\,\mathrm{sign}(w^\mu(x))dx - \frac{\mu^2}{\Delta x^2}\int_{\mathbb{R}} |w^\mu(x)|dx \\ + \frac{\mu^2}{\Delta x^2}\int_{\mathbb{R}} w^\mu(x + \frac{\Delta x}{\mu})\,\mathrm{sign}(w^\mu(x))dx - \frac{\mu^2}{\Delta x^2}\int_{\mathbb{R}} |w^\mu(x)|dx \leq 0. \end{aligned}$$

Moreover, for the same reason, we deduce that $I_4 \leq$ and $I_5 \leq 0$. Therefore, from (A.4) we get that

$$\frac{d}{dt}\int_{\mathbb{R}} |w^\mu(x)|dx \leq 0, \tag{A.5}$$

This guarantees the contractive property (A.3). □

Proof of Lemma 4.13. Let us consider the Fourier transform of w as

$$\widehat{w}(\xi) = \int_{\mathbb{R}} e^{-ix\xi} w(x)dx, \quad \xi \in \mathbb{R}$$

and the discrete Fourier transform of the sequence $\{w_j\}_{j\in\mathbb{Z}}$ as

$$\overline{w}(\xi) = \Delta x \sum_{j\in\mathbb{Z}} w_j e^{-ij\Delta x\xi}, \quad \xi \in [-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}].$$

It is also clear that for a piecewise constant function w defined as in (1.19)

$$\widehat{w}(\xi) = \frac{2\sin(\frac{\xi\Delta x}{2})}{\xi\Delta x}\,\overline{w}(\xi) \quad \text{and} \quad \widehat{d^+_{\Delta x} w}(\xi) = \frac{e^{i\xi\Delta x} - 1}{\Delta x}\widehat{w}(\xi).$$

Now, we know that

$$\|D^s w\|_2^2 = \int_{\mathbb{R}} |\xi|^{2s}|\widehat{w}(\xi)|^2 d\xi = \int_{-\pi/\Delta x}^{\pi/\Delta x} |\xi|^{2s}|\widehat{w}(\xi)|^2 d\xi + \sum_{j\neq 0}\int_{(2j-1)\pi/\Delta x}^{(2j+1)\pi/\Delta x} |\xi|^{2s}|\widehat{w}(\xi)|^2 d\xi. \tag{A.6}$$

For each $j \neq 0$, we have

$$\begin{aligned}
\int_{(2j-1)\pi/\Delta x}^{(2j+1)\pi/\Delta x} |\xi|^{2s}|\widehat{w}(\xi)|^2 d\xi &= \int_{(2j-1)\pi/\Delta x}^{(2j+1)\pi/\Delta x} |\xi|^{2s}|\,\overline{w}(\xi)|^2 \left|\frac{2\sin(\frac{\xi\Delta x}{2})}{\xi\Delta x}\right|^2 d\xi \\
&= \int_{-\pi/\Delta x}^{\pi/\Delta x} \left|\xi + 2j\frac{\pi}{\Delta x}\right|^{2s} |\,\overline{w}(\xi)|^2 \left|\frac{2\sin(\frac{\xi\Delta x}{2} + j\pi)}{(\xi + 2j\frac{\pi}{\Delta x})\Delta x}\right|^2 d\xi \\
&= \int_{-\pi/\Delta x}^{\pi/\Delta x} \left|\frac{2}{\Delta x}\sin(\frac{\xi\Delta x}{2})\right|^{2s} |\,\overline{w}(\xi)|^2 \left|\frac{2}{\Delta x}\sin(\frac{\xi\Delta x}{2})\right|^{2-2s} \left|\xi + 2j\frac{\pi}{\Delta x}\right|^{2s-2} d\xi \\
&\leq \int_{-\pi/\Delta x}^{\pi/\Delta x} |\xi|^{2s}|\,\overline{w}(\xi)|^2 \frac{|\xi|^{2-2s}}{\left|\xi + 2j\frac{\pi}{\Delta x}\right|^{2-2s}} d\xi \\
&\leq \frac{1}{|2|j| - 1|^{2-2s}} \int_{-\pi/\Delta x}^{\pi/\Delta x} |\xi|^{2s}|\,\overline{w}(\xi)|^2 d\xi
\end{aligned}$$

Therefore, replacing this in (A.6) and using that $0 < s < \frac{1}{2}$, we get

$$\begin{aligned}
\|D^s w\|_2^2 &\leq \int_{-\pi/\Delta x}^{\pi/\Delta x} |\xi|^{2s}|\widehat{w}(\xi)|^2 d\xi + \sum_{j\neq 0} \frac{1}{|2|j| - 1|^{2-2s}} \int_{-\pi/\Delta x}^{\pi/\Delta x} |\xi|^{2s}|\,\overline{w}(\xi)|^2 d\xi \\
&= \int_{-\pi/\Delta x}^{\pi/\Delta x} |\xi|^{2s} \left|\frac{2\sin(\frac{\xi\Delta x}{2})}{\xi\Delta x}\right|^2 |\,\overline{w}(\xi)|^2 d\xi \\
&\quad + \sum_{j\neq 0} \frac{1}{|2|j| - 1|^{2-2s}} \int_{-\pi/\Delta x}^{\pi/\Delta x} |\xi|^{2s}|\,\overline{w}(\xi)|^2 d\xi \\
&\lesssim \int_{-\pi/\Delta x}^{\pi/\Delta x} |\xi|^{2s}|\,\overline{w}(\xi)|^2 d\xi
\end{aligned}$$

On the other hand, using analogous arguments, we also have

$$\begin{aligned}
\|d_{\Delta x}^+ w\|_2^2 = \int_{\mathbb{R}} \left|\frac{e^{i\xi\Delta x} - 1}{\Delta x}\right|^2 |\widehat{w}(\xi)|^2 d\xi &= \int_{\mathbb{R}} \left|\frac{e^{i\xi\Delta x} - 1}{\Delta x}\right|^2 \left|\frac{2\sin(\frac{\xi\Delta x}{2})}{\xi\Delta x}\right|^2 |\,\overline{w}(\xi)|^2 d\xi \\
&\gtrsim \int_{-\pi/\Delta x}^{\pi/\Delta x} \left|\frac{e^{i\xi\Delta x} - 1}{\Delta x}\right|^2 |\,\overline{w}(\xi)|^2 d\xi \gtrsim \int_{-\pi/\Delta x}^{\pi/\Delta x} |\xi|^2\,|\,\overline{w}(\xi)|^2 d\xi
\end{aligned}$$

Finally, we conclude

$$\begin{aligned}
\|w\|_{H^s(\mathbb{R})}^2 &= \int_{\mathbb{R}} (1 + |\xi|^{2s})|\widehat{w}(\xi)|^2 d\xi = \int_{\mathbb{R}} |\widehat{w}(\xi)|^2 d\xi + \int_{\mathbb{R}} |\xi|^{2s}|\widehat{w}(\xi)|^2 d\xi \\
&\lesssim \int_{-\pi/\Delta x}^{\pi/\Delta x} |\,\overline{w}(\xi)|^2 d\xi + \int_{-\pi/\Delta x}^{\pi/\Delta x} |\xi|^{2s}|\,\overline{w}(\xi)|^2 d\xi \\
&\lesssim \int_{-\pi/\Delta x}^{\pi/\Delta x} (1 + |\xi|^{2s})|\,\overline{w}(\xi)|^2 d\xi \\
&\lesssim \int_{-\pi/\Delta x}^{\pi/\Delta x} (1 + |\xi|^2)|\,\overline{w}(\xi)|^2 d\xi \lesssim (\|w\|_2^2 + \|d_{\Delta x}^+ w\|_2^2).
\end{aligned}$$

□

Lemma A.2. *Given any $a \in (0,1)$ and $b \in \mathbb{C}$ with $|b| = 1$, the following inequality holds:*

$$\left| \sum_{n=1}^{N} a^n (b^n - 1) + \left(\sum_{n=1}^{N} n a^n \right) \left(\frac{1}{b} - 1 \right) \right| \leq |b-1|^2 \, \frac{a}{(1-a)^3}$$

Proof. Using that $|b| = 1$, we have:

$$\begin{aligned} \left| \sum_{n=1}^{N} a^n (b^n - 1) + \left(\sum_{n=1}^{N} n a^n \right) \left(\frac{1}{b} - 1 \right) \right| &= \left| \sum_{n=1}^{N} a^n b (b^n - 1) - \left(\sum_{n=1}^{N} n a^n \right) (b-1) \right| \\ &= |b-1| \left| \sum_{n=1}^{N} a^n \sum_{k=0}^{n-1} (b^{k+1} - 1) \right| = |b-1|^2 \left| \sum_{n=1}^{N} a^n \sum_{k=0}^{n-1} \sum_{j=0}^{k} b^j \right| \\ &\leq \frac{1}{2} |b-1|^2 \sum_{n=1}^{\infty} n(n+1) a^n = |b-1|^2 \, \frac{a}{(1-a)^3} \end{aligned}$$

□

Appendix B

Additional aspects on the augmented Burgers equation

B.1 Large-time behavior of the complete ABE

In this section we include some additional results related to the complete form of the augmented Burgers equation (1.1). Specifically, we suppose that there is ray-tube spreading and heterogeneous atmosphere (that is, non-constant density and ambient speed of sound). In this scenario, (1.1) fits in the following type of equation:

$$\begin{cases} u_t = uu_x + u_{xx} + K * u_{xx} + \big(H(t)\big)_t\, u, & (t,x) \in (0,\infty)\times\mathbb{R}, \\ u(0,x) = u^0(x), & x \in \mathbb{R}, \end{cases} \tag{B.1}$$

where $K(z) = e^{-z}\chi_{(0,\infty)}$. According to [20], if we assume cylindrical spreading, H is a function of the following form

$$H(t) = \ln\left(\frac{g(t)}{k_1(t+k_2)}\right),$$

k_1 and k_2 being positive constants and $g \in C(\mathbb{R})$ a strictly positive bounded function. Thus, the change of variable $v = ue^{-H}$ transforms (B.1) into

$$\begin{cases} v_t = e^H v v_x + v_{xx} + K * v_{xx}, & (t,x) \in (0,\infty)\times\mathbb{R}, \\ v(0,x) = u^0(x)e^{H(0)}, & x \in \mathbb{R}. \end{cases} \tag{B.2}$$

The same arguments from Chapter 4 are valid to prove the following theorem, concerning the solution of (B.2).

Theorem B.1. *For any initial data $u^0 \in L^1(\mathbb{R})$, there exists a unique solution $v \in C([0,\infty), L^1(\mathbb{R}))$ of (B.2). Moreover, for all $p \in [1,\infty]$, there exists a constant $C =$*

$C(p) > 0$ such that

$$\|v(t)\|_p \leq C\|u^0\|_1 t^{-\frac{1}{2}(1-\frac{1}{p})}, \quad \forall t > 0, \tag{B.3}$$

and

$$\|v_x(t)\|_p \leq C\|u^0\|_1 t^{-\frac{1}{2}(1-\frac{1}{p})-\frac{1}{2}}, \quad \forall t > 0. \tag{B.4}$$

The key point is that H is bounded from above for all $t \geq 0$ and that it goes to $-\infty$ as $t \to \infty$. Therefore, all the estimates obtained in Chapter 4 are still valid. For the same reason, we can repeat the scaling procedure to obtain the asymptotic profile. The only difference now is that the nonlinear term vanishes since $e^H \to 0$ as $t \to \infty$. Thus, the equation for the first term in the asymptotic expansion is the heat equation.

Theorem B.2. *Let $u^0 \in L^1(\mathbb{R})$. For any $p \in [1, \infty]$, the solution v to* (B.2) *satisfies*

$$t^{\frac{1}{2}(1-\frac{1}{p})}\|v(t) - v_M(t)\|_p \longrightarrow 0, \quad \text{as } t \to \infty,$$

where $v_M(t, x)$ is the solution of the following heat equation:

$$\begin{cases} v_t = 2v_{xx}, & x \in \mathbb{R}, t > 0, \\ v(0) = \frac{M}{H(0)}\delta_0. \end{cases} \tag{B.5}$$

Here δ_0 denotes the Dirac delta at the origin and M is the mass of the initial data, $M = \int_{\mathbb{R}} u^0(x)dx$.

It is obvious that we can now reverse the change of variable $v = u/H$ to obtain the large-time behavior of the solution to (B.1).

Corollary B.3. *Let $u^0 \in L^1(\mathbb{R})$. For any $p \in [1, \infty]$, the solution u to* (B.1) *satisfies*

$$t^{\frac{1}{2}(1-\frac{1}{p})}e^{-H(t)}\|u(t) - e^{H(t)}v_M(t)\|_p \longrightarrow 0, \quad \text{as } t \to \infty,$$

where $v_M(t, x)$ is the solution of (B.5).

B.2 Realistic simulation

We conclude this appendix with a numerical simulation of the sonic-boom minimization problem (1.1)-(1.2) in a realistic setting. One of the most common ways of computing the pressure signature in the near-field of the plane is the F-function developed by Witham (see [76] and the references therein). In Table B.1 we collect the main parameters that are required to obtain the initial near-field pressure signature. We refer to [76] for more detailed information about the way of computing it.

Mach velocity	2.7
Body length	91.46 m
Flight altitude	18288 m
Distance for near-field signature	182.92 m

TABLE B.1: Parameters of the flight (taken from [76]) to compute the F-function for the near-field signature.

Regarding the physical parameters of the atmosphere, we take $\Gamma = 8 \times 10^6$. The relaxation parameters are shown in Table B.2 and the density and speed of sound, in Figure B.1. Moreover, we consider cylindrical ray-tube spreading [20], which implies that $G/G_\sigma = 1/(2(\sigma + \sigma_0))$. Note that the initial signature is measured at some distance from the plain and, thus, $\sigma_0 > 0$.

	Oxygen	Nitrogen
C_ν	1.7×10^{-5}	1.2×10^{-4}
θ_ν	4.6×10^{-8}	1.6×10^{-6}

TABLE B.2: Parameters for the relaxation phenomena (taken from [20, 85, 86]).

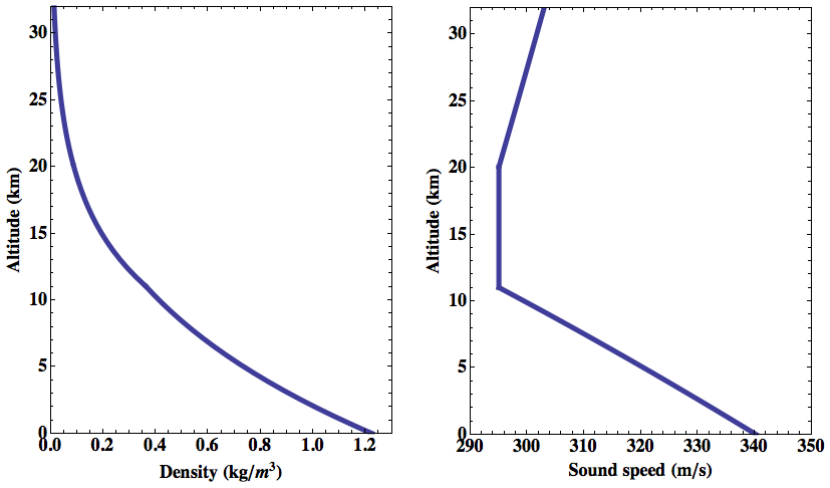


FIGURE B.1: Density of the air (left) and sound speed (right) of the atmosphere, from [20].

For the numerical approximation of (1.1), we follow [20, 86], using splitting for each of the operators appearing in the equation. However, in our case we select Engquist-Osher for the nonlinear term, which we have shown to be appropriate for large-time evolution problems. With respect to the optimization method, we opt for a conjugate gradient descent method based on the adjoint methodology. Let us remark that we pay attention to the settings of the algorithm as suggested in Chapter 3.

As a target function we use the evolution obtained from a given initial profile. In Figure B.2 we show the evolution of the functional throughout the optimization process.

The obtained minimizer is shown in Figure B.3, as well as its corresponding signature at the ground level, in comparison to the exact solution. Even if the final profile is really close to the target function, note that the initial data are very different. This is not surprising, due to the viscous nature of the model and the large-time evolution process it involves. Therefore, from an engineering point of view, additional restrictions on the initial data (such as the parametrization of an F-function) might be required to restrict the admissible near-field signature to realistic profiles.

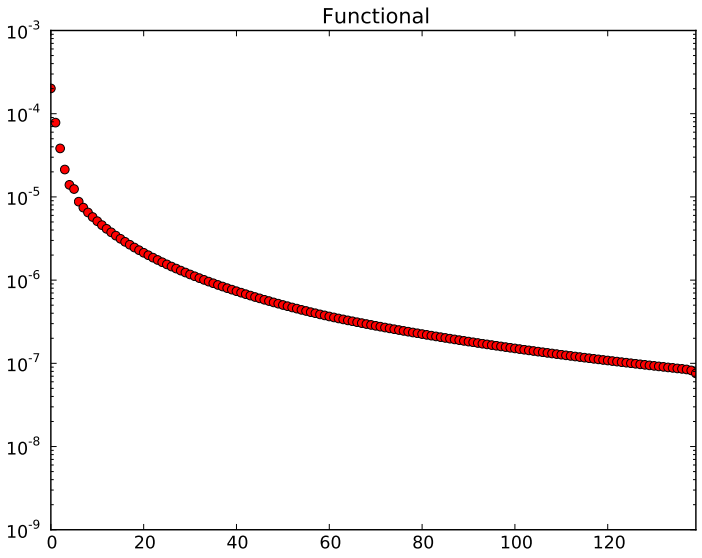


FIGURE B.2: Evolution of the functional throughout the iterative process.

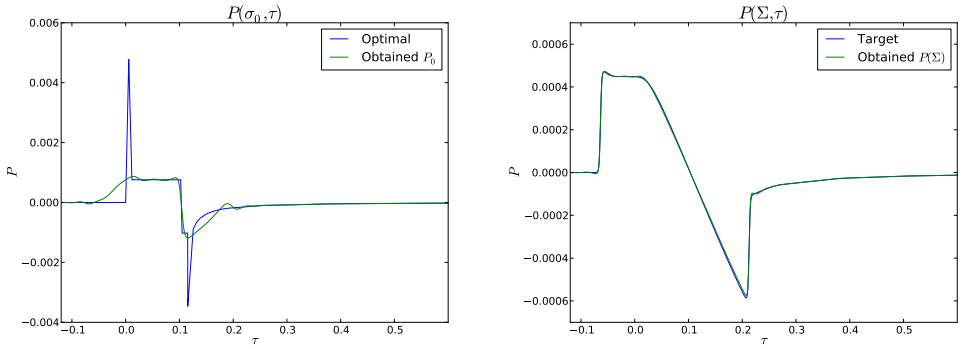


FIGURE B.3: In green, obtained minimizer (left) and the corresponding solution at the ground level (right), compared to the optimal initial data and the target function.

Nevertheless, let us point out that there is also a nonlinear issue hiding behind. It is well known that the optimal control problem that we considered in Chapter 3 has multiple minimizers in the case of the inviscid Burgers equation [17]. We have shown that, in case of having a shock in the target function, the classical adjoint methodology tends to recover a compression wave for which the shock does not arise until the final

time T. Strategies like the alternating descent method proposed in [17] create the shock at the beginning. But it is also possible to have the shock arising at some instant in between (see Figure B.4).

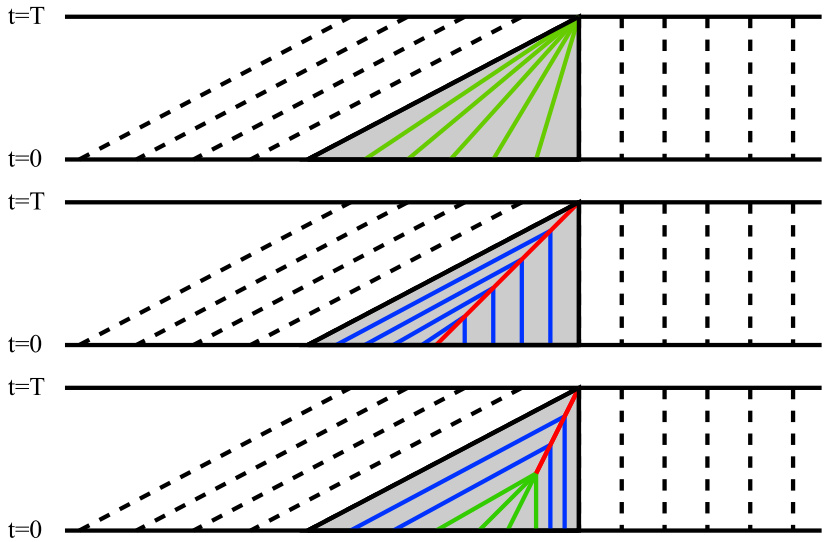


FIGURE B.4: Different ways of creating a shock at a given time: from a compression wave (top), from a shock at the initial data (middle) and from a shock that arises somewhere in between (bottom).

This can be extended, from a numerical point of view, to the low-viscous case and the quasi-shocks arising there. This is the case of the sonic boom minimization problem (1.1)-(1.2). In the example above we have obtained two very different initial data that fit (quite) well for the same target function. But, in practice, we do not know the exact solution, so the obtained approximated optimal solutions, even if mathematically correct, could be unsuitable from an engineering point of view. Even more, a parametrization of the initial data (which could partially fix the smoothing of peaks pointed out above) will not solve this issue.

Bibliography

[1] Absil, P.-A., Mahony, R., and Andrews, B. Convergence of the iterates of descent methods for analytic functions. *SIAM Journal on Optimization 16*, 2 (2005), 531–547.

[2] Aguirre, J., and Rivas, J. A spectral viscosity method based on hermite functions for nonlinear conservation laws. *SIAM Journal on Numerical Analysis 46*, 2 (2008), 1060–1078.

[3] Alonso, J. J., and Colonno, M. R. Multidisciplinary optimization with applications to sonic-boom minimization. *Annual Review of Fluid Mechanics 44* (2012), 505–526.

[4] Ascher, U. M., van den Doel, K., Huang, H., and Svaiter, B. F. Gradient descent and fast artificial time integration. *ESAIM: M2AN 43* (2009), 689–708.

[5] Attouch, H., Bolte, J., and Svaiter, B. F. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods. *Mathematical Programming 137* (2013), 91–129.

[6] Bardos, C., Leroux, A.-Y., and Nedelec, J.-C. First order quasilinear equations with boundary conditions. *Communications in Partial Differential Equations 4*, 9 (1979), 1017–1034.

[7] Bardos, C., and Pironneau, O. Data assimilation for conservation laws. *Methods and Applications of Analysis 12*, 2 (2005), 103–134.

[8] Bateman, H. Some recent researches on the motion of fluids. *Monthly Weather Review 43*, 4 (April 1915), 163–170.

[9] Bouchut, F., and James, F. One-dimensional transport equations with discontinuous coefficients. *Nonlinear Analysis: Theory, Methods & Applications 32*, 7 (1998), 891–933.

[10] Bouchut, F., and James, F. Differentiability with respect to initial data for a scalar conservation law. In *Hyperbolic Problems: Theory, Numerics, Applications* (1999), M. Fey and R. Jeltsch, Eds., vol. 129 of *International Series of Numerical Mathematics*, Springer-Verlag, pp. 113–118.

[11] Bouharguane, A., and Carles, R. Splitting methods for the nonlocal Fowler equation. *Mathematics of Computation 83* (2013), 1121–1141.

[12] Brenier, Y., and Osher, S. The discrete one-sided Lipschitz condition for convex scalar conservation laws. *SIAM Journal on Numerical Analysis 25*, 1 (February 1988), 8–23.

[13] Brenner, P., Thomée, V., and Wahlbin, L. B. *Besov Spaces and Applications to Difference Methods for Initial Value Problems*, vol. 434 of *Lecture Notes in Mathematics*. Springer-Verlag, 1975.

[14] Bressan, A., and Marson, A. A variational calculus for discontinuous solutions of systems of conservation laws. *Communications in Partial Differential Equations 20*, 9-10 (1995), 1491–1552.

[15] Burgers, J. M. Application of a model system to illustrate some points of the statistical theory of free turbulence. *Proceedings of the Royal Netherlands Academy of Sciences 43*, 1 (1940), 2–12.

[16] Carlton, T. W., and Blackstock, D. T. Propagation of plane waves of finite amplitude in inhomogeneous media with applications to vertical propagation in the ocean. Tech. Rep. ARL-TR-74-31, Applied Research laboratories, The University of Texas at Austin, 1974.

[17] Castro, C., Palacios, F., and Zuazua, E. An alternating descent method for the optimal control of the inviscid Burgers' equation in the presence of shocks. *Mathematical Models and Methods in Applied Sciences 18*, 3 (2008), 369–416.

[18] Castro, C., Palacios, F., and Zuazua, E. Optimal control and vanishing viscosity for the Burgers equation. In *Integral Methods in Science and Engineering*, C. Costanda and M. E. Pérez, Eds., vol. 2. Birkhäuser Verlag, 2010, ch. 7, pp. 65–90.

[19] Ciarlet, P. G. *Introduction to numerical linear algebra and optimisation*, vol. 2 of *Cambridge Texts in Applied Mathematics*. Cambridge University Press, 1989.

[20] Cleveland, R. O. *Propagation of sonic booms through a real, stratified atmosphere*. PhD thesis, University of Texas at Austin, 1995.

[21] Cockburn, B., Karniadakis, G. E., and Shu, C.-W. The development of discontinuous galerkin methods. In *Discontinuous Galerkin Methods: Theory, Computation and Applications* (2000), B. Cockburn, G. E. Karniadakis, and C.-W. Shu, Eds., vol. 11 of *Lecture Notes in Computational Science and Engineering*, Springer-Verlag, pp. 3–50.

[22] Cole, J. D. On a quasilinear parabolic equation occurring in aerodynamics. *Quarterly of Applied Mathematics 9* (1951), 225–236.

[23] Crandall, M. G., and Majda, A. Monotone difference approximations for scalar conservation laws. *Mathematics of Computation 34*, 149 (January 1980), 1–21.

[24] Dick, S. J., Ed. *NASA's first 50 years: Historical Perspectives* (2009), NASA.

[25] Duoandikoetxea, J., and Zuazua, E. Moments, masses de Dirac et décomposition de fonctions. *Comptes rendus de l'Académie des sciences 315*, I (1992), 693–698.

[26] Engquist, B., and Osher, S. Stable and entropy satisfying approximations for transonic flow calculations. *Mathematics of Computation 34*, 179 (January 1980), 45–75.

[27] Ervedoza, S., and Zuazua, E. *On the numerical approximation of exact controls for waves*. Springer Briefs in Mathematics. Springer-Verlag, 2013.

[28] Escobedo, M., Vazquez, J. L., and Zuazua, E. Asymptotic behavior and source type solutions for a diffusion-convection equation. *Archive for Rational Mechanics and Analysis 124* (1993), 43–65.

[29] Escobedo, M., and Zuazua, E. Large time behavior for convection-diffusion equations in $\mathbb{R}^n$. *Journal of Functional Analysis 100*, 1 (August 1991), 119–161.

[30] Evans, L. C. *Partial differential equations*, second ed., vol. 19 of *Graduate Studies in Mathematics*. American Mathematical Society, 2002.

[31] Fridman, V. E. Propagation of a strong sound wave in a plane layered medium. *Soviet Physics - Acoustics 22* (1976), 349–350.

[32] Ghil, M., and Malanotte-Rizzoli, P. *Data assimilation in meteorology and oceanography*, vol. 33 of *Advances in Geophysics*. Academic Press, 1991.

[33] Giles, M. B. Discrete adjoint approximations with shocks. In *Hyperbolic Problems: Theory, Numerics, Applications* (2003), T. Y. Hou and E. Tadmor, Eds., vol. XXVIII, Springer-Verlag, pp. 185–194.

[34] Giles, M. B., and Pierce, N. A. An introduction to the adjoint approach to design. *Turbulence and Combustion 65* (2000), 393–415.

[35] Giles, M. B., and Ulbrich, S. Convergence of linearized and adjoint approximations for discontinuous solutions of conservation laws. Part 1: linearized approximations and linearized output functionals. *SIAM Journal on Numerical Analysis 48*, 3 (2010), 882–904.

[36] Giles, M. B., and Ulbrich, S. Convergence of linearized and adjoint approximations for discontinuous solutions of conservation laws. Part 2: adjoint approximations and extensions. *SIAM Journal on Numerical Analysis 48*, 3 (2010), 905–921.

[37] Glowinski, R., Lions, J.-L., and He, J. *Exact and Approximate Controllability for Distributed Parameter Systems: A Numerical Approach*, vol. 117 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 2008.

[38] Godlewski, E., and Raviart, P.-A. *Hyperbolic systems of conservation laws*. No. 3 in Mathematiques & Applications. Ellipses, 1991.

[39] Godunov, S. K. A difference scheme for numerical computation of discontinuous solution of hydrodynamic equations. *Matematicheskii Sbornik 47*, 89 (1959), 271–306.

[40] Gosse, L., and James, F. Numerical approximations of one-dimensional linear conservation equations with discontinuous coefficients. *Mathematics of Computation 69* (2000), 987–1015.

[41] Harabetian, E. Rarefaction and large time behavior for parabolic equations and monotone schemes. *Communications in Mathematical Physics 144* (1988), 527–536.

[42] Haraux, A. Some applications of the Łojasiewicz gradient inequality. *Communications on Pure and Applied Analysis 11*, 6 (November 2012), 2417–2427.

[43] Hayes, W. D. *Linearized Supersonic Flow*. PhD thesis, California Institute of Technology, Pasadena, California, 1947.

[44] Holden, H., Karlsen, K. H., Lie, K.-A., and Risebro, N. H. *Splitting Methods for Partial Differential Equations with Rough Solutions*. EMS Series of Lectures in Mathematics. European Mathematical Society, 2010.

[45] Hopf, E. The partial differential equation $u_t + uu_x = \mu u_{xx}$. *Communications on Pure and Applied Mathematics 3*, 3 (September 1950), 201–230.

[46] Ignat, L. I., Ignat, T. I., and Stancu-Dumitru, D. A compactness tool for the analysis of nonlocal evolution equations. *Submitted* (2013).

[47] Ignat, L. I., and Pazoto, A. Large time behavior of a nonlocal diffusion-convection equation related with the gas dynamics. *Discrete and Continuous Dynamical Systems - Series A accepted* (2013).

[48] Ignat, L. I., and Rossi, J. D. A nonlocal convection-diffusion equation. *Journal of Functional Analysis 251* (2007), 399–437.

[49] Ignat, L. I., and Rossi, J. D. Decay estimates for nonlocal problems via energy methods. *Journal de Mathématiques Pures et Appliquées 92*, 2 (August 2009), 163–187.

[50] Il'in, A. M., and Oleinik, O. A. Behavior of the solutions of the Cauchy problem for certain quasilinear equations for unbounded increase of time. *Doklady Akademii Nauk SSSR 120* (1958), 25–28.

[51] Il'in, A. M., and Oleinik, O. A. Asymptotic behavior of solutions of the Cauchy problem for some quasilinear equations for large values of time. *Matematicheskii Sbornik 51(93)*, 2 (1960), 191–216.

[52] Iserles, A. *A first course in the numerical analysis of differential equations.* Cambridge University Press, 1996.

[53] James, F., and Postel, M. Numerical gradient methods for flux identification in a system of conservation laws. *Journal of Engineering Mathematics 60*, 3-4 (2008), 293–317.

[54] James, F., and Sepúlveda, M. Parameter identification for a model of chromatographic column. *Inverse Problems 6* (1994), 1299–1314.

[55] James, F., and Sepúlveda, M. Convergence results for the flux identification in a scalar conservation law. *SIAM Journal on Control and Optimization 37*, 3 (1999), 869–891.

[56] James, F., and Vauchelet, N. A remark on duality solutions for some weakly nonlinear scalar conservation laws. *Comptes rendus de l'Académie des sciences 349* (2011), 657–661.

[57] Jones, B. F. A class of singular integrals. *American Journal of Mathematics 86*, 2 (April 1964), 441–462.

[58] Karch, G., and Suzuki, K. Spikes and diffusion waves in one-dimensional model of chemotaxis. *Nonlinearity*, 23 (2010), 3119–3137.

[59] Kim, Y. J., and Ni, W. M. On the rate of convergence and asymptotic profile of solutions to the viscous Burgers equation. *Indiana University Mathematics Journal 51*, 3 (2002), 727–752.

[60] Kim, Y. J., and Tzavaras, A. E. Diffusive N-waves and metastability in the Burgers equation. *SIAM Journal on Mathematical Analysis 33*, 3 (2001), 607–633.

[61] Lattanzio, C., and Marcati, P. Global well-posedness and relaxation limits of a model for radiating gas. *Journal of Differential Equations 190*, 2 (may 2003), 439–465.

[62] Laurençot, P. Asymptotic self-similarity for a simplified model for radiating gases. *Asymptotic Analysis 42* (2005), 251–262.

[63] Lax, P. D. Weak solutions of nonlinear hyperbolic equations and their numerical approximation. *Communications on Pure and Applied Mathematics 7* (1954), 159–193.

[64] Lax, P. D. Hyperbolic systems of conservation laws II. *Communications on Pure and Applied Mathematics 10* (1957), 537–566.

[65] Lax, P. D. *Hyperbolic systems of conservation laws and the mathematical theory of shock waves*. No. 11 in CBMS-NSF Regional Conference Series in Appiled Mathematics. SIAM, 1973.

[66] LeVeque, R. J. *Numerical Methods for Conservation Laws*, second ed. Birkhäuser Verlag, 1992.

[67] LeVeque, R. J. *Finite-Volume Methods for Hyperbolic Problems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2004.

[68] Li, Y., Osher, S., and Tsai, R. Heat source identification based on l_1 constrained minimization. *Inverse Problems and Imaging 8*, 1 (2014), 199–221.

[69] Lighthill, M. J. Viscosity effects in sound waves of finite amplitude. In *Surveys in Mechanics*, G. K. Batchelor and R. M. Davies, Eds. Cambridge University Press, 1956, pp. 250–351.

[70] Lions, J.-L., and Malgrange, B. Sur l'unicité rétrograde dans les problèmes mixtes paraboliques. *Mathematica Scandinavica 8* (1960), 277–286.

[71] Liu, H., and Tadmor, E. Critical thresholds in a convolution model for nonlinear conservation laws. *SIAM Journal on Mathematical Analysis 33*, 4 (2001), 930–945.

[72] Liu, T. P. Invariants and asymptotic behavior of solutions of a conservation law. *Proceedings of the American Mathematical Society 71*, 2 (1978), 227–231.

[73] Liu, T. P., and Pierre, M. Source-solutions and asymptotic behavior in conservation laws. *Journal of Differential Equations 51* (1984), 419–441.

[74] McLean, F. E. Some nonasymptotic effects on the sonic boom of large aircraft. Tech. Rep. D-2877, NASA, June 1965.

[75] Merlet, B., and Pierre, M. Convergence to equilibrium for the backward Euler scheme and applications. *Communications on Pure and Applied Analysis 8*, 3 (2010), 685–702.

[76] Minelli, A., el Din, I. S., and Carrier, G. Advanced optimization approach for supersonic low-boom design. In *18th AIAA/CEAS Aeroacoustics Conference (33rd AIAA Aeroacoustics Conference)* (Colorado Springs, CO, June 2012).

[77] Nocedal, J., and Wright, S. J. *Numerical Optimization*, second ed. Springer Series in Operations Research and Financial Engineering. Springer-Verlag, 2006.

[78] Ozcer, I. Sonic boom prediction using Euler/full potential methodology. In *45th AIAA Aerospace Sciences Meeting and Exhibit* (2007), Aerospace Sciences Meetings, American Institute of Aeronautics and Astronautics.

[79] Pawlowski, J. W., Graham, D. H., Boccadoro, C. H., Coen, P. G., and Maglieri, D. J. Origins and overview of the shaped sonic boom demonstration program. Paper 2005-5, American Institute of Aeronautics and Astronautics, January 2005.

[80] Pierce, A. D. *Acoustics: an introduction to its physical principles and applications*. Acoustical Society of America, 1989.

[81] Pozo, A., Allahverdi, N., and Zuazua, E. Numerical aspects of large-time optimal control of Burgers equation. *Submitted* (2014).

[82] Pozo, A., and Ignat, L. I. A semi-discrete large-time behavior preserving scheme for the augmented Burgers equation. *Submitted* (2014).

[83] Pozo, A., and Ignat, L. I. A splitting method for the augmented Burgers equation. *Submitted* (2014).

[84] Pozo, A., Ignat, L. I., and Zuazua, E. Large-time asymptotics, vanishing viscosity and numerics for 1-D scalar conservation laws. *Mathematics of Computation*, to appear (2013).

[85] Rallabhandi, S. K. Advanced sonic boom prediction using augmented Burger's equation. *Journal of Aircraft 48*, 4 (2011), 1245–1253.

[86] Rallabhandi, S. K. Sonic boom adjoint methodology and its applications. In *29th AIAA Applied Aerodynamics Conference* (June 2011), American Institute of Aeronautics and Astronautics.

[87] Ribaud, F. Cauchy problem for semilinear parabolic equations with initial data in $H_p^s(\mathbb{R}^n)$ spaces. *Revista Matemática Iberoamericana 14*, 1 (1998), 1–46.

[88] Simon, J. Compact sets in the space $\mathcal{L}^p(0,T;B)$. *Annali di Matematica pura ed applicata 146*, 4 (1987), 65–96.

[89] Ulbrich, S. Optimal control of nonlinear hyperbolic conservation laws with source terms. Habilitation Thesis, Fakultät für Mathematik, Technische Universität München, June 2001.

[90] Ulbrich, S. A sensitivity and adjoint calculus for discontinuous solutions of hyperbolic conservation laws with source terms. *SIAM Journal on Control and Optimization 41*, 3 (2002), 740–797.

[91] van den Doel, K., and Ascher, U. M. The chaotic nature of faster gradient descent methods. *Journal of Scientific Computing 51*, 3 (June 2012), 560–581.

[92] Wächter, A., and Biegler, L. T. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming 106*, 1 (March 2006), 25–57.

[93] Whitham, G. B. The flow pattern of a supersonic projectile. *Communications on Pure and Applied Mathematics 5*, 3 (August 1952), 301–348.

[94] Whitham, G. B. *Linear and nonlinear waves*. John Wiley & Sons, 1974.

[95] Zuazua, E. Propagation, observation, and control of waves approximated by finite difference methods. *SIAM Review 47*, 2 (2005), 197–243.