SEMA SIMAI Springer Series 34

Carlos Parés
Manuel J. Castro
Tomás Morales de Luna
María Luz Muñoz-Ruiz *Editors*

Hyperbolic Problems: Theory, Numerics, Applications. Volume I

HYP2022, Málaga, Spain, June 20–24, 2022

SEMA
SIMAI
SOCIETÀ ITALIANA DI MATEMATICA APPLICATA E INDUSTRIALE
Springer

SEMA SIMAI Springer Series

Volume 34

As of 2013, the SIMAI Springer Series opens to SEMA in order to publish a joint series aiming to publish advanced textbooks, research-level monographs and collected works that focus on applications of mathematics to social and industrial problems, including biology, medicine, engineering, environment and finance. Mathematical and numerical modeling is playing a crucial role in the solution of the complex and interrelated problems faced nowadays not only by researchers operating in the field of basic sciences, but also in more directly applied and industrial sectors. This series is meant to host selected contributions focusing on the relevance of mathematics in real life applications and to provide useful reference material to students, academic and industrial researchers at an international level. Interdisciplinary contributions, showing a fruitful collaboration of mathematicians with researchers of other fields to address complex applications, are welcomed in this series.

THE SERIES IS INDEXED IN SCOPUS

Carlos Parés · Manuel J. Castro ·
Tomás Morales de Luna · María Luz Muñoz-Ruiz
Editors

Hyperbolic Problems: Theory, Numerics, Applications. Volume I

HYP2022, Málaga, Spain, June 20–24, 2022

Springer

Editors
Carlos Parés
Departamento de Análisis Matemático,
Estadística e I.O., Matemática Aplicada,
Facultad de Ciencias
Universidad de Málaga
Málaga, Spain

Manuel J. Castro
Departamento de Análisis Matemático,
Estadística e I.O., Matemática Aplicada,
Facultad de Ciencias
Universidad de Málaga
Málaga, Spain

Tomás Morales de Luna
Departamento de Análisis Matemático,
Estadística e I.O., Matemática Aplicada,
Facultad de Ciencias
Universidad de Málaga
Málaga, Spain

María Luz Muñoz-Ruiz
Departamento de Análisis Matemático,
Estadística e I.O., Matemática Aplicada,
Facultad de Ciencias
Universidad de Málaga
Málaga, Spain

ISSN 2199-3041 ISSN 2199-305X (electronic)
SEMA SIMAI Springer Series
ISBN 978-3-031-55259-5 ISBN 978-3-031-55260-1 (eBook)
https://doi.org/10.1007/978-3-031-55260-1

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

To the memory of Prof. Antonio Valle Sánchez (1930–2012), founder of the EDANYA group.

Preface

The present volume contains selected papers issuing from the *XVIII International Conference on Hyperbolic Problems: Theory, Numerics, and Applications* (HYP2022) which was held during June 20–24, 2022, in Málaga (Spain). The conference proceedings have been divided into two volumes: this one collects some of the invited talks and the contributions focusing on theoretical aspects.

This series of conferences constitute an international event of reference in the field of Hyperbolic Partial Differential Equations. Their objective is to bring together scientists with interests in the theoretical, applied, and computational aspects of hyperbolic partial differential equations (systems of hyperbolic conservation laws, wave equations, etc.) and of related mathematical models (PDEs of mixed type, kinetic equations, nonlocal or/and discrete models, etc.). The first conference was held in 1986 in St. Etienne (France) and has been organized since then biennially at different locations. The last few meetings were held at 2018 Penn State (USA), 2016 Aachen (Germany), 2014 Rio de Janeiro (Brazil), 2012 Padua (Italy), 2010 Beijing (China), 2008 College Park (USA), and 2006 Lyon (France).

The eighteenth edition of this series of conferences should have been held in June 2020 in Málaga, but the situation due to the COVID-19 pandemic led the organizing and scientific committees to postpone the conference. In order to avoid a four-year period without any activity related to the HYP series, an online activity, the *HYP2020/21 day*, took place in July 2, 2021. This event included talks by the first *Peter Lax Awardee*, Jacob Bedrossian (University of Maryland), and the first *James Glimm Lecturer*, Constantine Dafermos (Brown University). These special lectures, which will be part of the program in every future edition of the HYP series, were instituted by the scientific committee to distinguish respectively a young researcher (at most 10 years after the Ph.D.) and a senior one for their contributions to the field of hyperbolic PDEs. The names of these two distinguished talks honor the fundamental ideas and contributions of two outstanding researchers, Peter Lax and James Glimm, who were present at the HYP2020/21 day, which makes this event unforgettable for all the attendees. The program of the HYP2020/21 day was completed with two talks given by Min Tang (Shanghai Jiaotong University) and Manuel J. Castro (University of Málaga).

The second Peter Lax Awardee and James Glimm Lecturer were Maria Colombo (EPFL, Switzerland) and Benoît Perthame (Sorbonne-Université, France), respectively. Besides their distinguished lectures, the program of HYP2022 included 5 plenary talks by speakers Eduard Feireisl (Institute of Mathematics Prague, Czech Republic), Jan S. Hesthaven (EPFL Lausanne, Switzerland), Denis Serre (ENS Lyon, France), Eleuterio F. Toro (U. Trento, Italy), and Tong Yang (Hong Kong PolyU). It also included 8 invited talks by Benjamin Gess (U. Bielefeld, Germany), Kenneth H. Karlsen (U. Oslo, Norway), Qin Li (U. Wisconsin-Madison, USA), Raphaël Loubére (U. Bordeaux, France), Giovanni Russo (U. Catania, Italy), Konstantina Trivisa (U. Maryland, USA), Emil Wiedemann (U. Ulm, Germany), and Yao Yao (Georgia Tech., USA). Finally, 190 contributed talks (66 of them by Ph.D. students) were given and 19 posters were presented. Despite mobility restrictions in place in June 2022 due to the pandemic and international conflicts, which made it impossible for many colleagues to travel to Málaga, 287 researchers from 27 different countries attended the conference.

One of the main goals of HYP2022 was to promote the attendance of Ph.D. students, many of whom had never before had the opportunity of attending an international conference in person due to COVID-19. This goal was largely achieved: 86 attendees were Ph.D. students. Among the measures taken to stimulate their participation, more than 30 grants that covered the registration and accommodation fees were given (with priority for female students) and a recognition to the best presentations by Ph.D. students in the three fields of the conference, the Springer Awards, was given. The awardees in the fields Theory, Numerics, and Applications were respectively William Golding (University of Texas at Austin, USA), Alessia del Grosso (Université de Versailles Saint-Quentin-en-Yvelines, France), and Kathrin Hellmuth (University of Würzburg, Germany). The awardees received a certificate and a book voucher from Dr. Francesca Bonadei, Executive Editor of Springer.

The conference proceedings contain 69 chapters, 64 of which correspond to contributed talks or posters. The 29 chapters of this volume are grouped in two categories: Plenary Talks (5 chapters) and Theory (24 chapters). The chapters of the first category correspond to the talks by E. Feireisl, K. Karlsen, D. Serre, and E. F. Toro in HYP2022, and the one by M. J. Castro in HYP2020/21 day.

We would like to address our warmest thanks and gratitude to all who have made this book possible: first of all, to all the speakers of HYP2020/21 day and HYP2022 for their valuable contributions and, very especially, to those who accepted our invitation to contribute to this volume. This book has undergone a rigorous peer-review process: we are grateful for the work of the anonymous referees who, in a disinterested way, have helped the authors to improve the quality of their manuscripts. We would also like to thank the members of the scientific committee for their support and help in the speaker selection and those of the organizing committee for ensuring the smooth running of the event. We would like to thank the sponsors, without whom HYP2022 would not have been possible: we are really grateful to the University of Málaga and the Sociedad Española de Matemática Aplicada (SEMA). The financial support of the Office of Naval Research (ONR) of the United States allowed us to increase the number of grants for Ph.D. students: we thank Dr. Reza Malek-Medani for his interest and help. We also thank the Springer staff for their help and support during the edition process, and especially Dr. Francesca Bonadei. Finally, we are very grateful to the Editorial Board of the SEMA/SIMAI Springer series for having accepted this volume and to the Editor-in-charge, María Elena Vázquez Cendón, for her helpful comments.

Málaga, Spain
July 2023

Carlos Parés
Manuel J. Castro
Tomás Morales de Luna
María Luz Muñoz-Ruiz

Contents

Invited Talks

Implicit Exactly Well-Balanced Finite Volume Schemes for Balance Laws with Singular Source Terms

Manuel J. Castro, Irene Gómez-Bueno, and Carlos Parés

Abstract In previous works, the authors have proposed a methodology to design explicit well-balanced high-order numerical methods for 1d systems of balance laws with singular source terms. These methods rely on the combination of the Generalized Hydrostatic Reconstruction (GHR) technique and a well-balanced reconstruction operator. In this work, first these two ingredients are recalled as well as the family of semi-discrete in space well-balanced high-order methods obtained by applying them. Then the extension to implicit time discretizations are discussed. In particular, two new strategies to design exactly well-balanced implicit finite volume schemes for systems with singular source terms are introduced. Finally, some numerical tests for the Burgers' equation with singular source term are considered to check and compare the resulting implicit numerical methods.

Keywords Balance Laws · Singular source terms · Implicit finite volume method · Exactly well-balanced method

1 Introduction

We consider 1d systems of balance laws of the form

$$U_t + F(U)_x = S(U)H_x, \tag{1}$$

M. J. Castro (✉) · C. Parés
Facultad de Ciencias, University of Málaga, Málaga, Spain
e-mail: mjcastro@uma.es

C. Parés
e-mail: pares@uma.es

I. Gómez-Bueno
Complejo Tecnológico, E.T.S.I. de Telecomunicación, University of Málaga, Málaga, Spain
e-mail: igomezbueno@uma.es

C. Parés et al. (eds.), *Hyperbolic Problems: Theory, Numerics, Applications. Volume I*,
SEMA SIMAI Springer Series 34, https://doi.org/10.1007/978-3-031-55260-1_1

where $U(x, t)$ takes value in $\Omega \subset \mathbb{R}^N$, $F : \Omega \rightarrow \mathbb{R}^N$ is the flux function; $S : \Omega \rightarrow \mathbb{R}^N$; and H is a known function from $\mathbb{R} \rightarrow \mathbb{R}$. In this work we suppose that H is a.e. differentiable and has finite isolated jump discontinuities. For simplicity we suppose that H is piecewise C^1 with only a jump discontinuity located at x^*. Moreover, we suppose that the system is hyperbolic, i.e. the Jacobian $J(U)$ of the flux function is assumed to have N different real eigenvalues.

In real-world scenarios, balance laws with singular source terms often arise when the variable represented by H presents abrupt changes or discontinuities in regions with small lengths compared to the space step. In such cases, approximating the variable with a discontinuous function H can be a useful alternative to refine the mesh or to employ adaptive meshes. For example, in the case of shallow water equations, where H denotes the bottom depth, this situation can occur when the bottom features abrupt variations or discontinuities such as steps or junctions. Similar problems appear in blood flow problems (see [20] and the references therein).

When H is discontinuous, the solution U is expected to also be discontinuous, and the source term $S(U)H_x$ cannot be defined within the distributional framework. It becomes a nonconservative product whose mathematical definition has to be specified. Several mathematical theories exist to give a sense to these products. In the theory developed in [17], nonconservative products are interpreted as Borel measures that depend on a family of paths. This choice of paths can be arbitrary, but it must be consistent with the physics of the problem since the Rankine-Hugoniot conditions, and hence the definition of a weak solution, depend on this selection. Although selecting appropriate paths for general nonconservative systems can be challenging, there is a natural choice of paths for systems of balance laws with singular source terms. This choice is related to the stationary solutions of a regularized system, as discussed in [6], and can be interpreted in terms of preserving the Riemann invariant of a linearly degenerate characteristic field.

Note that system (1) has non trivial stationary solutions that satisfy the ODE system

$$F(U)_x = S(U)H_x \tag{2}$$

in the areas where H is smooth, and at x^*, the discontinuity linking the limit states (U^-, U^+) has to be admissible according to the selected family of paths.

The main objective of this work is to present a general framework to construct high-order implicit finite-volume schemes that exactly preserves all steady states, or at least a relevant family of them, when H is a non-smooth function. To do this, the ideas described in [11] for continuous H will be extended. In this reference, semi-discrete high-order finite-volume scheme for system (1) of the form:

$$\frac{d\bar{U}_i}{dt} = -\frac{1}{\Delta x}\left(F_{i+1/2} - F_{i-1/2} - \int_{x_{i-1/2}}^{x_{i+1/2}} S(P_i^t(x))H_x(x)\,dx\right), \tag{3}$$

were considered. Here,

- a mesh composed by cells $[x_{i-1/2}, x_{i+1/2}]$ whose length, Δx, is supposed to be constant for simplicity is considered;
- $\bar{U}_i(t)$ is the approximation given by the numerical method of the average of the exact solution at the ith cell, $[x_{i-1/2}, x_{i+1/2}]$ at time t, i.e.

$$\bar{U}_i(t) \cong \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} U(x,t)\, dx;$$

- $P_i^t(x)$ is the approximation of the solution at the ith cell given by a high-order reconstruction operator from the sequence of cell values $\{\bar{U}_i(t)\}$, i.e.

$$P_i^t(x) = P_i(x; \{\bar{U}_j(t)\}_{j \in \mathcal{S}_i}),$$

  where $\mathcal{S}_i$ is the set of indexes of the stencil of the ith cell;
- $F_{i+1/2} = \mathbb{F}(U_{i+1/2}^{t,-}, U_{i+1/2}^{t,+})$ where $\mathbb{F}$ is a consistent numerical flux, and

$$U_{i+1/2}^{t,-} = P_i^t(x_{i+1/2}), \quad U_{i+1/2}^{t,+} = P_{i+1}^t(x_{i+1/2}),$$

  are the reconstructed states at the intercells.

If H has discontinuities, the scheme must be modified in order to take into account the contribution of the source term $S(U)H_x$ at the discontinuity points. Following [6], the Generalized Hydrostatic Reconstruction technique introduced in [5] will be used to develop numerical methods that accurately maintain the admissible jumps at the discontinuity points of H.

Although, the derivation of (exactly) well-balanced schemes is a very active field, see for example some Refs. [1–3, 7, 9, 10, 12, 14–16, 18, 20], … and the references therein, there are not many examples in the literature devoted to the derivation of high-order exactly well-balanced schemes for singular source terms for general 1D balance laws. Up to our knowledge, the first time such general framework was presented was in [6], where high-order explicit finite-volume method was proposed. The aim of this work is to extend this previous work, to implicit high-order methods combining the Generalized Hydrostatic Reconstruction and the well-balanced reconstruction operators proposed in [6] with the methodology introduced in [11] to derive implicit high-order well-balanced methods for smooth steady states.

The organization of the article is as follows: in Sect. 2 the definition of the non-coservative products given by the source term when H is discontinuous is discussed. The Generalized Hydrostatic Reconstruction technique is recalled and a family of semi-discrete high-order well-balanced numerical methods based on this technique is presented. Section 3 is devoted to describe the implicit time discretization of the methods: the strategy introduced in [11] will be followed. Finally, in Sect. 4, some numerical tests are presented to show the ability of the schemes to approximate small perturbations around non-smooth steady states for the Burgers' equation with non-smooth source term.

2 Semi-discrete Exactly Well-Balanced Method for Balance Laws with Singular Source Terms

As pointed out in the introduction, we consider here 1D balance laws (1) with non-smooth H. We refer to [6] for a detailed description of the content of this section. Notice that, if a solution U of (1) is discontinuous at a discontinuity point of H (as it can be expected), the source term $S(U)H_x$ cannot be defined within the distributional framework: the source term becomes a nonconservative product whose meaning has to be specified. The theory developed in [17] allows one to define it as a Borel measure whose definition depends on the choice of a family of paths, i.e. a Lispchitz-continuous map

$$s \in [0,1] \mapsto \Phi(W^-, W^+; s)$$

that links W^- and W^+,

$$\Phi(W^-, W^+; s) = \begin{bmatrix} \Phi_U(W^-, W^+; s) \\ \Phi_H(W^-, W^+; s) \end{bmatrix} \in \Omega \times \mathbb{R},$$

for $W^\pm = [U^\pm, H^\pm]^T \in \Omega \times \mathbf{R}$, $s \in [0,1]$ such that

$$\Phi(W^-, W^+; 0) = W^-, \quad \Phi(W^-, W^+; 1) = W^+,$$

$$\Phi(W, W; s) = W, \quad \forall s \in [0,1],$$

where we have used the notation

$$W = \begin{bmatrix} U \\ H \end{bmatrix} \in \Omega \times \mathbb{R}$$

to shorten the expressions.

Let us suppose for simplicity that U and H are piecewise C^1 functions with only a jump discontinuity located at x^*. Once the family of paths has been chosen, the nonconservative product is defined as the measure $[S(U)H_x]_\Phi$ whose action over continuous functions of compact support is as follows:

$$\langle [S(U)H_x]_\Phi, \varphi \rangle = \int_{\mathbb{R}} S(U(x))H_x(x)\varphi(x)\,dx + \varphi(x^*) \int_0^1 S(\Phi_U(W^-, W^+; s))\partial_s \Phi_H(W^-, W^+; s)\,ds,$$

where $U^\pm$, $H^\pm$ are, respectively, the right and left limits of U and H at x^*. In other words:

$$[S(U)H_x]_\Phi = S(U)H_x + \left(\int_0^1 S(\Phi_U(W^-, W^+; s))\partial_s \Phi_H(W^-, W^+; s)\, ds\right)\delta_{x=x^*},$$

where $\delta_{x=x^*}$ represents the Dirac measure placed at x^*.

Once the family of paths has been chosen, a function U is said to be a weak solution of (1) if it satisfies

$$U_t + F(U)_x = [S(U)H_x]_\Phi$$

in the sense of measures, what implies, in particular, that the discontinuities of U either propagate through the regions where H is continuous at a speed σ satisfying the usual Rankine-Hugoniot condition

$$[F(U)] = \sigma[U],$$

with

$$[F(U)] = F(U^+) - F(U^-), \quad [U] = U^+ - U^-,$$

or they are stationary at a discontinuity point of H and satisfy the jump condition

$$[F(U)] = \int_0^1 S(\Phi_U(W^-, W^+; s))\partial_s \Phi_H(W^-, W^+; s)\, ds. \tag{4}$$

As pointed out in the introduction, (3) must be modified in order to take into account the singularites of H. Indeed once the notion of weak solution has been fixed through the choice of the family of paths, the numerical method (3) has to be adapted in order to take into account the Dirac measures produced by the source term at the discontinuities of H. To do that, let us assume that the mesh is designed so that all the discontinuity points of H are placed at an intercell. Then, we consider a path-conservative approximation of the source term (see [19]):

$$\frac{d\bar{U}_i}{dt} = -\frac{1}{\Delta x}\Big(F_{i+1/2} - F_{i-1/2} - S^+_{i-1/2} - S^-_{i+1/2} - \int_{x_{i-1/2}}^{x_{i+1/2}} S(P_i^t(x))H_x(x)\, dx\Big), \tag{5}$$

where $F_{i+1/2}$ is again the numerical flux and $S^\pm_{i+1/2}$ are such that:

$$S^-_{i+1/2} + S^+_{i+1/2} = \int_0^1 S(\Phi_U(W^{t,-}_{i+1/2}, W^{t,+}_{i+1/2}; s))\partial_s \Phi_H(W^{t,-}_{i+1/2}, W^{t,+}_{i+1/2}; s)\, ds; \tag{6}$$

$$S^\pm_{i+1/2} = 0 \text{ if } H^-_{i+1/2} = H^+_{i+1/2}; \tag{7}$$

where $H^\pm_{i+1/2}$ represent the limits of H at the left and at the right of $x_{i+1/2}$ and

$$W_{i+1/2}^{t,\pm} = \begin{bmatrix} U_{i+1/2}^{t,\pm} \\ H_{i+1/2}^{\pm} \end{bmatrix}.$$

Here, $U_{i+1/2}^{t,\pm}$ represent again the reconstructed states.

As in the case of smooth H, the reconstruction operators play a crucial role in the smooth areas. As in [6], we consider here exactly well-balanced reconstruction operators, that is

Definition 1 Given a stationary solution U of (1), the reconstruction operator is said to be exactly well-balanced for U if

$$P_i(x; \{\bar{U}_j\}_{j \in \mathcal{S}_i}) = U(x), \quad \forall x \in [x_{i-1/2}, x_{i+1/2}],\ \forall i, \tag{8}$$

where

$$\bar{U}_j = \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} U(x)\, dx, \quad \forall j.$$

The same definition could be stated if $\bar{U}_j$ is computed with a suitable quadrature formula.

Generally, a standard reconstruction operator is not guaranteed to be well-balanced because the functions P_i are typically computed using interpolation techniques within a specific class of functions (e.g., polynomials, hyperbolas), and the stationary solutions may not belong to that class. However, the strategy introduced in [4] can be used to design well-balanced reconstructions operator on the basis of a standard one.

Well-balanced reconstruction procedure: Given a family of cell values $\{\bar{U}_i\}$, to compute the reconstruction P_i at the cell $[x_{i-1/2}, x_{i+1/2}]$:

1. Look for the stationary solution $U_i^*(x)$ defined in stencil of cell I_i $(\cup I_j,\ j \in \mathcal{S}_i)$, such that:
$$\frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} U_i^*(x)\, dx = \bar{U}_i. \tag{9}$$

2. Compute the fluctuations $\{\bar{V}_j\}_{j \in \mathcal{S}_i}$:
$$\bar{V}_j = \bar{U}_j - \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} U_i^*(x)\, dx,\ \ j \in \mathcal{S}_i.$$

3. Apply a standard reconstruction operator to the fluctuations $\{\bar{V}_j\}_{j \in \mathcal{S}_i}$:
$$Q_i(x) = Q_i(x; \{\bar{V}_j\}_{j \in \mathcal{S}_i}).$$

4. Define
$$P_i(x) = U_i^*(x) + Q_i(x).$$

One can verify that the reconstruction operator P_i satisfies the well-balanced property for all stationary solutions if the reconstruction operator $Q_i(x)$ is exact for the null function. Moreover, it is conservative, i.e.

$$\frac{1}{\Delta x}\int_{x_{i-1/2}}^{x_{i+1/2}} P_i(x)\,dx = \bar{U}_i, \quad \forall i, \tag{10}$$

provided that $Q_i(x)$ is conservative, and it is high-order accurate provided that the stationary solutions are smooth.

As in Definition 1, the exact integral in (9) can be replaced by the approximation given by the quadrature formula selected to approximate cell averages.

Notice that, in general, a nonlinear problem has to be solved at the first stage of the reconstruction operator (9). If the $N \times N$ ODE system satisfied by the stationary solutions of (1)

$$F(U)_x = S(U)H_x, \tag{11}$$

has a N-parametric general solution, then (9) is a $N \times N$ nonlinear system. Nevertheless, the system may not have a unique solution:

- If (9) has no solution or if it has one that is not defined in the stencil

  $$\cup_{j\in\mathcal{S}_i}\,[x_{j-1/2}, x_{j+1/2}],$$

  then $U_i^* \equiv 0$ is chosen in the first stage and the reconstruction operator reduces to the standard one. Notice that if the cell values are the averages of a stationary solution, then (9) has always at least one solution (the stationary solution itself) and thus this choice does not affect the well-balancedness of the reconstruction operator.
- If (9) has more than one solution, a criterion to select one of them is needed: see, for instance, [8] where a well-balanced reconstruction operator for the shallow water equation has been introduced.
- Note that the computation of the steady states in (9) has to take into account the discontinuities of H.

In summary, to fully specify the numerical method (5), several choices must be made: a consistent numerical flux, a reconstruction operator, a family of paths, and a path-conservative discretization of the source term.

It can be noted that selecting a different family of paths can result in different admissible jumps of a weak solution at a discontinuity point of H. In general, choosing a consistent family of paths can be a challenging task for nonconservative systems, but for this specific case, there is a natural criterion. A pair of states (U^-, U^+) can be connected by an admissible stationary jump at a discontinuity point x^* of H if and only if there exists a solution of the ODE system

$$\frac{d}{dH}F(\mathcal{V}) = S(\mathcal{V}) \tag{12}$$

such that

$$\mathcal{V}(H^{\pm}) = U^{\pm}. \tag{13}$$

In this case, the path linking W^- to W^+ is

$$\Phi(W^-, W^+; s) = \begin{bmatrix} \mathcal{V}(H^- + s[H]) \\ H^- + s[H] \end{bmatrix}, \quad s \in [0, 1]. \tag{14}$$

Observe that H is used to represent the independent variable of system (12). This variable is used to link the limit values of the function $H(x)$ at the left and right of x^*, $H^{\pm}$: that's the reason of using the same symbol, although it may be ambiguous.

Remark 1 If the eigenvalues of $J(U^-)$ do not vanish, the jump condition (12)–(13) is equivalent to state that the solution of the Cauchy problem

$$\begin{cases} \dfrac{d\mathcal{V}}{dH} = J^{-1}(\mathcal{V})S(\mathcal{V}), \\ \mathcal{V}(H^-) = U^- \end{cases} \tag{15}$$

is defined in H^+ and satisfies $\mathcal{V}(H^+) = U^+$. In this case, the set of right states that can be linked to a left state $W^- = [U^-, H^-]^T$ through an admissible stationary discontinuity at x^* is the integral curve of the ODE system passing by U^- at H^-. If one of the eigenvalues of $J(U^-)$ vanishes (i.e. if the problem is resonant), the Cauchy problem consisting of the ODE system (12) with initial condition $\mathcal{V}(H^-) = U^-$ may have no solution or to have more than one. In the latter case, a criterion is required to decide what are the admissible discontinuities to be preserved by the numerical method.

The admissibility criterion based on this family of paths is mathematically natural because the stationary jumps at a discontinuity point of H can be understood as contact discontinuities associated to a linearly degenerate field of an equivalent system. In this sense, the criterion is equivalent to requiring that a Riemann invariant is conserved across such jumps (see [6] for more details).

Let us check that this admissibility criterion leads to pairs that satisfy the jump condition (4):

$$\begin{aligned} \int_0^1 S(\Phi_U(W^-, W^+; s))\partial_s \Phi_H(W^-, W^+; s)\, ds &= \int_{H^-}^{H^+} S(\mathcal{V}(\sigma))\, dH \\ &= \int_{H^-}^{H^+} \frac{d}{dH} F(\mathcal{V})\, d\sigma = [F(U)]. \end{aligned}$$

Remark 2 If $\mathcal{V}(H)$ is a solution of (12), then $U(x) = \mathcal{V}(H(x))$ is a stationary solution of (1): observe that (2) is trivially satisfied in smooth regions and (12)–(13) in discontinuities.

The chosen admissibility criterion determines the path (14) linking pairs of states that can be the limits of an admissible jump at the discontinuity points of H.

Nevertheless, in order to design the numerical method, the path linking two arbitrary states has to be chosen. We consider here the family of paths corresponding to the so-called Generalized Hydrostatic Reconstruction technique (see [5]): in order to define the path linking $W^{\pm} = [U^{\pm}, H^{\pm}]^T$, first an intermediate value H^0 is chosen such that $H_0 = H^- = H^+$ if $H^- = H^+$. Next, we solve (if possible) (12) with initial condition $\mathcal{V}(H^-) = U^-$ (resp. $\mathcal{V}(H^+) = U^+$) and denote by $\mathcal{V}^-(H)$ (resp. $\mathcal{V}^+(H)$) the corresponding solution of the Cauchy problem. Then, the chosen path is a parameterization in [0, 1] of the union of the following arcs:

- $H \in [H^-, H^0] \mapsto [\mathcal{V}^-(H), H]^T$;
- the straight segment linking $[\mathcal{V}^-(H^0), H^0]^T$ and $[\mathcal{V}^+(H^0), H^0]^T$;
- $H \in [H^0, H^+] \mapsto [\mathcal{V}^+(H), H]^T$.

Notice that, if (13) is satisfied, then $\mathcal{V}^- = \mathcal{V}^+ = \mathcal{V}$ and the path reduces to (14).

Remark 3 The intermediate value H^0 is a degree of freedom of the family of paths. This choice can affect the properties of the numerical method, such as its ability to preserve positivity. If H^0 is equal to either H^- or H^+, then the path consists of only an arc of the integral curve and a straight segment, resulting in a simpler shape.

Coming back to the general case, the family of paths is used now to compute the source term as follows:

$$\begin{aligned}
&\int_0^1 S(\Phi_U(W^-, W^+; s))\partial s \Phi_H(W^-, W^+; s))\, ds \\
&= \int_{H^-}^{H^0} S(\mathcal{V}^-(H))\, dH + \int_{H^0}^{H^+} S(\mathcal{V}^+(H))\, dH \\
&= \int_{H^-}^{H^0} \frac{d}{dH} F(\mathcal{V}^-(H))\, dH + \int_{H^0}^{H^+} \frac{d}{dH} F(\mathcal{V}^+(H))\, dH \\
&= F(\mathcal{V}^-(H^0)) - F(U^-) + F(U^+) - F(\mathcal{V}^+(H^0)),
\end{aligned}$$

where the fact that $\mathcal{V}^{\pm}$ are solutions of (12) have been used. Notice that the straight segment does not contribute to the source term, as $\Phi_H(W^-, W^+; s) = H_0$ in this piece of the path. This computation suggests the following numerical approximation of the source term:

$$S^+_{i+1/2} = F(U^{t,+}_{i+1/2}) - F(\mathcal{V}^{t,+}_{i+1/2}(H^0_{i+1/2})), \tag{16}$$

$$S^-_{i+1/2} = F(\mathcal{V}^{t,-}_{i+1/2}(H^0_{i+1/2})) - F(U^{t,-}_{i+1/2}), \tag{17}$$

where $H^0_{i+1/2}$ is the intermediate value between $H^-_{i+1/2}$ and $H^+_{i+1/2}$, and $\mathcal{V}^{t,\pm}_{i+1/2}(H)$ represents the solution of (12) with initial conditions

$$\mathcal{V}(H^{\pm}_{i+1/2}) = U^{t,\pm}_{i+1/2}.$$

The consistency requirements (6)–(7) can be easily checked.

Taking into account this approximation of the source term, we consider the family of methods:

$$\begin{aligned}\frac{d\bar{U}_i}{dt} = -\frac{1}{\Delta x}\Big(& F_{i+1/2} - F_{i-1/2} \\ & - F(U^{t,+}_{i-1/2}) + F(\mathcal{V}^{t,+}_{i-1/2}(H^0_{i-1/2})) - F(\mathcal{V}^{t,-}_{i+1/2}(H^0_{i+1/2})) + F(U^{t,-}_{i+1/2}) \\ & - \int_{x_{i-1/2}}^{x_{i+1/2}} S(P_i^t(x)) H_x(x)\, dx\Big),\end{aligned} \tag{18}$$

where

$$F_{i+1/2} = \mathbb{F}(\mathcal{V}^{t,-}_{i+1/2}(H^0_{i+1/2}), \mathcal{V}^{t,+}_{i+1/2}(H^0_{i+1/2})).$$

Notice that, if H is continuous at $x_{i-1/2}$ and $x_{i+1/2}$, then

$$\mathcal{V}^{t,+}_{i-1/2}(H^0_{i-1/2}) = U^{t,+}_{i-1/2}, \quad \mathcal{V}^{t,-}_{i+1/2}(H^0_{i+1/2}) = U^{t,-}_{i+1/2},$$

and (18) reduces to (3).

Now, the following result can be stated (see [6]):

Theorem 1 *If the reconstruction operator is exactly well-balanced for a stationary solution U, then the numerical method* (18) *is also exactly well-balanced for U.*

As in the smooth case, special care must be taken with the quadrature formulas. Here, we follow the same idea introduced in [6] and the numerical scheme is then written as follows:

$$\begin{aligned}\frac{d\bar{U}_i}{dt} = -\frac{1}{\Delta x}\Big(& F_{i+1/2} - F_{i-1/2} - F(U_i^{t,*}(x^-_{i+1/2})) + F(U_i^{t,*}(x^+_{i-1/2})) \\ & - F(U^{t,+}_{i-1/2}) + F(\mathcal{V}^{t,+}_{i-1/2}(H^0_{i-1/2})) - F(\mathcal{V}^{t,-}_{i+1/2}(H^0_{i+1/2})) + F(U^{t,-}_{i+1/2}) \\ & - \Delta x \sum_{k=0}^{M} \alpha_k^i \left(S(P_i^t(x_k^i)) - S(U_i^{t,*}(x_k^i))\right) H_x(x_k^i)\Big),\end{aligned} \tag{19}$$

where and α_k^i and x_k^i are, respectively, the weights and nodes of the quadrature formula and

$$U_i^{t,*}(x^{\mp}_{i\pm 1/2}) = \lim_{x \to x^{\mp}_{i\pm 1/2}} U_i^{t,*}(x),$$

with $U_i^{t,*}$ is the stationary solution computed in (9).

In the particular case of a second-order method based on the well-balanced MUSCL operator using the mid-point rule the method writes as follows:

$$\frac{d\bar{U}_i}{dt} = -\frac{1}{\Delta x}\Big(F_{i+1/2} - F_{i-1/2} + F(U_i^{t,*}(x_{i-1/2}^+)) - F(U_i^{t,*}(x_{i+1/2}^-)) \\ - F(U_{i-1/2}^{t,+}) + F(\mathcal{V}_{i-1/2}^{t,+}(H_{i-1/2}^0)) - F(\mathcal{V}_{i+1/2}^{t,-}(H_{i+1/2}^0)) + F(U_{i+1/2}^{t,-})\Big), \tag{20}$$

since $S(P_i^t(x_i)) - S(U_i^{t,*}(x_i)) = 0$. Finally, the first order method using the mid-point rule reduces to

$$\frac{d\bar{U}_i}{dt} = -\frac{1}{\Delta x}\Big(F_{i+1/2} - F_{i-1/2} + F(\mathcal{V}_{i-1/2}^{t,+}(H_{i-1/2}^0)) - F(\mathcal{V}_{i+1/2}^{t,-}(H_{i+1/2}^0))\Big), \tag{21}$$

since $P_i^t = U_i^{t,*}$ and thus

$$U_{i+1/2}^{t,-} = U_i^{t,*}(x_{i+1/2}^-), \quad U_{i-1/2}^{t,+} = U_i^{t,*}(x_{i-1/2}^+).$$

Remark 4 It is well-known that the path-conservative formalism is not enough to guarantee convergence of the numerical solutions to the desired weak solutions: the numerical diffusion and/or dispersion have to be controlled due to the small-scale sensitivity of the shocks in nonconservative systems (as it happens with conservative methods for problems where small-scale dependent shocks appear). However, the family of numerical methods considered in this paper overcomes this limitation by exactly preserving the stationary jumps associated with the discontinuities of the source terms. Therefore, these methods are expected to converge to the correct weak solutions, as shown by numerical experiments in Sect. 4.

3 Time Integration

Fully explicit numerical methods for solving (1) are easily obtained by applying an ODE solver to the semi-discrete methods (19): the TVD-RK methods introduced in [13] are a sensible choice.

Although in principle implicit or semi-implicit high-order well-balanced methods can be obtained as well by applying implicit ODE solvers to (19), in practice the well-balanced reconstruction of the unknown solution U^{n+1} and the computation of the paths may lead to complex nonlinear systems that are costly to solve. To illustrate this, let us discretize in time the first-order semidiscrete method (21) using forward Euler:

$$\bar{U}_i^{n+1} = \bar{U}_i^n - \frac{\Delta t}{\Delta x}\left(F_{i+1/2}^{n+1} - F_{i-1/2}^{n+1} + F(\mathcal{V}_{i-1/2}^{n+1,+}(H_{i-1/2}^0)) - F(\mathcal{V}_{i+1/2}^{n+1,-}(H_{i-1/2}^0))\right), \tag{22}$$

with

$$F_{i+1/2}^{n+1} = \mathbb{F}(\mathcal{V}_{i+1/2}^{n+1,-}(H_{i+1/2}^0), \mathcal{V}_{i+1/2}^{n+1,+}(H_{i+1/2}^0)),$$

and $\mathcal{V}^{n+1,\pm}_{i+1/2}$ is the solution of (12) at $H^0_{i+1/2}$ with initial conditions $\mathcal{V}(H^{\pm}_{i+1/2}) = U^{n+1,\pm}_{i+1/2}$, respectively. Moreover $U^{n+1,\pm}_{i+1/2}$ are the well-balanced reconstructed states at $x^{\pm}_{i+1/2}$ that depends on $\bar{U}^{n+1}_i$ and $\bar{U}^{n+1}_{i+1}$ respectively through the computation of stationary solutions satisfying (9). Therefore, if the nonlinear system (22) is solved by using some iterative algorithm, at every stage ODE systems (12) with conditions (13) have to be solved at every intercell and ODE systems (11) with condition (9) have to be solved at every cell, what can be too costly.

To avoid this difficulty the strategy proposed in [11] is followed here: a solution of the ODE system (19) of the form $\bar{U}_i(t) = \bar{U}^n_i + \bar{U}^f_i(t)$ is sought in $[t^n, t^{n+1}]$ and, besides the standard reconstruction operator Q, an easier non well-balanced operator $\tilde{Q}$ will be used to reconstruct the perturbations $\bar{U}^f_i$. Namely, once the approximations at time t^n, $\{\bar{U}^n_i\}$, have been computed, in order to update them we proceed as follows:

- First, the well-balanced reconstruction procedure is applied to $\{\bar{U}^n_i\}$ to obtain:

$$P^n_i(x) = U^{*,n}_i(x) + Q_i(x; \{\bar{V}^n_j\}_{j\in\mathcal{S}_i}),$$

  where $U^{*,n}_i(x)$ is the stationary solution found at the first step of the reconstruction procedure at the ith cell.
- Next we consider the following ODE system in the time interval $[t^n, t^{n+1}]$:

$$\begin{aligned}\frac{d\bar{U}^f_i}{dt} = -\frac{1}{\Delta x}\Big(& F_{i+1/2}(t) - F_{i-1/2}(t) + F(U^{*,n}_i(x^+_{i-1/2})) - F(U^{*,n}_i(x^-_{i+1/2})) \\ & - F(U^{t,+}_{i-1/2}(t)) + F(\mathcal{V}^{t,+}_{i-1/2}(H^0_{i-1/2})) - F(\mathcal{V}^{t,-}_{i+1/2}(H^0_{i+1/2})) + F(U^{t,-}_{i+1/2}) \\ & - \Delta x \sum_{k=0}^{M} \alpha^i_k \left(S(P^t_i(x^i_k)) - S(U^{n,*}_i(x^i_k)) \right) H_x(x^i_k) \Big), \end{aligned} \tag{23}$$

  with initial condition

$$\bar{U}^f_i(t^n) = 0, \quad \forall i.$$

  Here

$$P^t_i(x) = P^n_i(x) + \widetilde{Q}^t_i(x; \{\bar{U}^f_j(t)\}_{j\in\mathcal{S}_i}), \tag{24}$$

  and

$$F_{i+1/2}(t) = \mathbb{F}(\mathcal{V}^{t,-}_{i+1/2}(H^0_{i+1/2}), \mathcal{V}^{t,+}_{i+1/2}(H^0_{i+1/2})), \tag{25}$$

  where $\mathcal{V}^{t,-}_{i+1/2}(H^0_{i+1/2})$ is the solution of (12) with initial condition

$$\mathcal{V}(H^-_{i+1/2}) = P^t_i(x_{i+1/2}) := U^{t,-}_{i+1/2}.$$

  Similarly, $\mathcal{V}^{t,+}_{i+1/2}(H^0_{i+1/2})$ is the solution of (12) with initial condition

$$\mathcal{V}(H^+_{i+1/2}) = P^t_{i+1}(x_{i+1/2}) := U^{t,+}_{i+1/2}.$$

- Define:

$$\bar{U}_i^{n+1} = \bar{U}_i^n + \bar{U}_i^f(t^{n+1}). \tag{26}$$

Observe that, although

$$\bar{U}_i(t) = \bar{U}_i^n + \bar{U}_i^f(t), \quad t \in [t^n, t^{n+1}]$$

formally solves (19), the reconstruction P_i^t is not the same as the one in the previous section: while there one had

$$P_i^t(x) = U_i^{t,*}(x) + Q_i(x; \{\bar{V}_j(t)\}_{j\in\mathcal{S}_i}), \tag{27}$$

now,

$$\begin{aligned} P_i^t(x) &= U_i^{n,*}(x) + Q_i(x; \{\bar{V}_j^n\}_{j\in\mathcal{S}_i}) + \widetilde{Q}_i^t(x; \{\bar{U}_j^f(t)\}_{j\in\mathcal{S}_i}) \\ &= P_i^n(x) + \widetilde{Q}_i^t(x; \{\bar{U}_j^f(t)\}_{j\in\mathcal{S}_i}). \end{aligned} \tag{28}$$

The main differences are the following:

- while the stationary solution $U_i^{t,*}$ is used (27), the stationary solution $U_i^{n,*}$ is used in (28) for every t;
- the reconstruction operator $\widetilde{Q}_i$ will be in practice easier and cheaper to compute than Q_i: in particular, the smoothness indicators obtained to compute Q_i at time t^n may be used to compute $\widetilde{Q}_i$. We shall require that $\tilde{Q}_i$ is exact for the null function and its order of accuracy is p.

The following result holds.

Theorem 2 *If the reconstruction operator is exactly well-balanced for a stationary solution U, then the numerical method (23) with (25)–(26) is exactly well-balanced for U.*

Proof Let us consider the vector $\{\bar{U}_i^0\}$ of the cell-averages of a stationary solution U^* of (1) computed with a given quadrature formula and let us write (23) as follows:

$$\frac{d\bar{U}_i^f}{dt} = -\mathcal{L}(\bar{U}^0, \bar{U}^f), \tag{29}$$

where

$$\begin{aligned} \mathcal{L}(\bar{U}^0, \bar{U}^f) = \frac{1}{\Delta x}\Big(& F_{i+1/2}(t) - F_{i-1/2}(t) + F(U_i^{0,*}(x_{i-1/2}^+)) - F(U_i^{0,*}(x_{i+1/2}^-)) \\ & - F(U_{i-1/2}^{t,+}(t)) + F(\mathcal{V}_{i-1/2}^{t,+}(H_{i-1/2}^0)) - F(\mathcal{V}_{i+1/2}^{t,-}(H_{i+1/2}^0)) + F(U_{i+1/2}^{t,-}) \\ & - \Delta x \sum_{k=0}^{M} \alpha_k^i \left(S(P_i^t(x_k^i)) - S(U_i^{0,*}(x_k^i)) \right) H_x(x_k^i) \Big). \end{aligned} \tag{30}$$

We are going to check that $\{\bar{U}_i^f = 0\}$ is a critical point of (29), that is

$$\mathcal{L}(\bar{U}^0, 0) = 0.$$

Taking into account the definition $P_i^t(x)$ and that the reconstruction operator is well-balanced, $P_i^t(x) = U^*(x) = U_i^{0,*}(x)$, thus, $\mathcal{L}(\bar{U}^0, 0)$ reduces to

$$\mathcal{L}(\bar{U}^0, 0) = \frac{1}{\Delta x}\Big(F_{i+1/2}(0) - F_{i-1/2}(0) + F(\mathcal{V}_{i-1/2}^{0,+}(H_{i-1/2}^0)) - F(\mathcal{V}_{i+1/2}^{0,-}(H_{i+1/2}^0))\Big).$$

Let us check that at every intercell, one has

$$\mathcal{V}_{i+1/2}^{0,-}(H_{i+1/2}^0) = \mathcal{V}_{i+1/2}^{0,+}(H_{i+1/2}^0).$$

If H is continuous at $x_{i+1/2}$ this is trivial; if not, since the reconstruction operator is well-balanced for U one has

$$U_{i+1/2}^{0,\pm} = U^*(x_{i+1/2}^{\pm}),$$

and, due to the admissibility criterion, $U^*(x_{i+1/2}^-)$, $U^*(x_{i+1/2}^+)$ have to be in the same integral curve of (12). As a consequence, assuming that $H_{i+1/2}^0$ belongs to the interval where the maximal solutions of (12)–(13) are defined, we have

$$\mathcal{V}_{i+1/2}^{0,-}(H_{i+1/2}^0) = \mathcal{V}_{i+1/2}^{0,+}(H_{i+1/2}^0),$$

and thus, $\mathcal{L}(\bar{U}^0, 0) = 0$. □

To fully discretize the method, an implicit solver will be applied now to (29) and the initial value $\{U_i^f = 0\}$ will guarantee the convergence in one single iteration if the initial condition corresponds to the cell averages of a stationary solution.

The application of an implicit RK solver to (29) will lead to nonlinear systems at every stage that will be solved with an iterative algorithm. Now, the choice of the reconstruction operator makes that ODE systems (11) with condition (9) have only to be solved at time t^n and not in every iteration. Nevertheless, ODE systems (12) with conditions (13) have yet to be solved at every intercell in every iteration. What we propose then is the following reformulation of (23) in which $\mathcal{V}_{i\pm1/2}^{\pm}$ are replaced by suitable approximations $\widetilde{\mathcal{V}}_{i\pm1/2}^{\pm}$ easier to compute and preserving the well-balanced character of the method:

$$
\begin{aligned}
\frac{d\bar{U}_i^f}{dt} = -\frac{1}{\Delta x}\Big(& \widetilde{F}_{i+1/2}(t) - \widetilde{F}_{i-1/2}(t) + F(U_i^{n,*}(x_{i-1/2}^+)) - F(U_i^{n,*}(x_{i+1/2}^-)) \\
& - F(U_{i-1/2}^{t,+}(t)) + F(\widetilde{\mathcal{V}}_{i-1/2}^{t,+}(H_{i-1/2}^0)) - F(\widetilde{\mathcal{V}}_{i+1/2}^{t,-}(H_{i+1/2}^0)) + F(U_{i+1/2}^{t,-}) \\
& - \Delta x \sum_{k=0}^{M} \alpha_k^i \left(S(P_i^t(x_k^i)) - S(U_i^{n,*}(x_k^i)) \right) H_x(x_k^i) \Big),
\end{aligned} \tag{31}
$$

with

$$
\widetilde{F}_{i+1/2}(t) = \mathbb{F}(\widetilde{\mathcal{V}}_{i+1/2}^{t,-}(H_{i+1/2}^0), \widetilde{\mathcal{V}}_{i+1/2}^{t,+}(H_{i+1/2}^0)), \tag{32}
$$

where $\widetilde{\mathcal{V}}_{i+1/2}^{t,-}(H_{i+1/2}^0)$ is defined as follows:

$$
\begin{aligned}
\widetilde{\mathcal{V}}_{i+1/2}^{t,-} = {} & U_{i+1/2}^{t,f,-} + \mathcal{V}_{i+1/2}^{n,-}(H_{i+1/2}^0) \\
& + \left(J^{-1}(U_{i+1/2}^{t,-}) S(U_{i+1/2}^{t,-}) - J^{-1}(U_{i+1/2}^{n,-}) S(U_{i+1/2}^{n,-}) \right) \Delta H_{i+1/2}^-,
\end{aligned} \tag{33}
$$

with $U_{i+1/2}^{t,f,-} = \widetilde{Q}_i^t(x_{i+1/2})$ and $\mathcal{V}_{i+1/2}^{n,-}(H_{i+1/2}^0)$ is the solution of (12) with initial condition $\mathcal{V}(H_{i+1/2}^-) = U_{i+1/2}^{n,-}$ and $\Delta H_{i+1/2}^- = H_{i+1/2}^0 - H_{i+1/2}^-$.

Similarly, $\widetilde{\mathcal{V}}_{i+1/2}^{t,+}$ is defined as

$$
\begin{aligned}
\widetilde{\mathcal{V}}_{i+1/2}^{t,+} = {} & U_{i+1/2}^{t,f,+} + \mathcal{V}_{i+1/2}^{n,+}(H_{i+1/2}^0) \\
& - \left(J^{-1}(U_{i+1/2}^{t,+}) S(U_{i+1/2}^{t,+}) - J^{-1}(U_{i+1/2}^{n,+}) S(U_{i+1/2}^{n,+}) \right) \Delta H_{i+1/2}^+,
\end{aligned} \tag{34}
$$

with $U_{i+1/2}^{t,f,+} = \widetilde{Q}_{i+1}^t(x_{i+1/2})$ and $\mathcal{V}_{i+1/2}^{n,+}(H_{i+1/2}^0)$ is the solution of (12) with initial condition $\mathcal{V}(H_{i+1/2}^+) = U_{i+1/2}^{n,+}$ and $\Delta H_{i+1/2}^+ = H_{i+1/2}^+ - H_{i+1/2}^0$.

These approximations are obtained as follows:

$$
\begin{aligned}
\mathcal{V}_{i+1/2}^{t,\pm} &= U_{i+1/2}^{t,\pm} + \int_{H_{i+1/2}^{\pm}}^{H_{i+1/2}^0} \frac{d}{dH} \mathcal{V}_{i+1/2}^{t,\pm}(H)\, dH \\
&= U_{i+1/2}^{t,\pm} + \mathcal{V}_{i+1/2}^{n,\pm} - U_{i+1/2}^{n,\pm} + \int_{H_{i+1/2}^{\pm}}^{H_{i+1/2}^0} \frac{d}{dH} \left(\mathcal{V}_{i+1/2}^{t,\pm} - \mathcal{V}_{i+1/2}^{n,\pm} \right)(H)\, dH \\
&= U_{i+1/2}^{t,f,\pm} + \mathcal{V}_{i+1/2}^{n,\pm} + \int_{H_{i+1/2}^{\pm}}^{H_{i+1/2}^0} \frac{d}{d\sigma} \left(\mathcal{V}_{i+1/2}^{t,\pm} - \mathcal{V}_{i+1/2}^{n,\pm} \right)(H)\, dH \\
&\approx U_{i+1/2}^{t,f,\pm} + \mathcal{V}_{i+1/2}^{n,\pm} \mp \left(J^{-1}(U_{i+1/2}^{t,\pm}) S(U_{i+1/2}^{t,\pm}) - J^{-1}(U_{i+1/2}^{n,\pm}) S(U_{i+1/2}^{n,\pm}) \right) \Delta H_{i+1/2}^{\pm},
\end{aligned}
$$

where we have denoted $\mathcal{V}_{i+1/2}^{n,\pm} = \mathcal{V}_{i+1/2}^{n,\pm}(H_{i+1/2}^0)$ and $\mathcal{V}_{i+1/2}^{t,\pm} = \mathcal{V}_{i+1/2}^{t,\pm}(H_{i+1/2}^0)$ to shorten the notation. In the previous expressions we have used that $\mathcal{V}_{i+1/2}^{n,-}(H_{i+1/2}^0)$ is the solution of (12) with initial condition $\mathcal{V}(H_{i+1/2}^-) = U_{i+1/2}^{n,-}$ and that (12) is equivalent to (15).

The following result holds:

Theorem 3 *If the reconstruction operator is exactly well-balanced for a stationary solution U, then the numerical method (31) with (25)–(26) is also exactly well-balanced for U.*

Proof The proof is similar, taking into account that (33) reduces to

$$\widetilde{\mathcal{V}}^{0,\pm}_{i+1/2}(H^0_{i+1/2}) = \mathcal{V}^{0,\pm}_{i+1/2}(H^0_{i+1/2})$$

when the initial condition is a stationary solution. □

4 Numerical Tests

In this section we consider the Burgers' equation with source term

$$\partial_t U + \partial_x \left(\frac{U^2}{2} \right) = U^2\, H'(x), \tag{35}$$

where $H(x)$ is given by

$$H(x) = \begin{cases} -0.2x & \text{if } x < 0. \\ -0.2x - 0.5 & \text{if } x > 0. \end{cases}$$

That is, $F(U) = \frac{U^2}{2}$ and $S(U) = U^2$. The stationary solutions are given by

$$U^*(x) = C_0 \exp(H(x)),\ \ C_0 \in \mathbb{R}.$$

We consider here first- and second-order methods in space and time. Second-order in space is achieved by using the MUSCL reconstruction operator (see [21]). Time integration is performed by using forward Euler and the second-order Runge-Kutta method with Butcher tableau

$$\begin{array}{c|cc} \gamma & \gamma & 0 \\ 1 & 1-\gamma & \gamma \\ \hline & 1-\gamma & \gamma, \end{array} \tag{36}$$

where $\gamma = 1 - \frac{1}{\sqrt{2}}$ and the integrals are approximated by the mid-point rule.

As illustration, we write here the first order numerical scheme (23):

$$\begin{aligned}\bar{U}_i^{n+1,f} &= -\frac{\Delta t}{\Delta x}\Big(F_{i+1/2}^{n+1} - F_{i-1/2}^{n+1} + F(U_i^{n,*}(x_{i-1/2}^{+})) - F(U_i^{n,*}(x_{i+1/2}^{-}))\Big) \\ &\quad + \frac{\Delta t}{\Delta x}\Big(F(\mathcal{V}_{i+1/2}^{n+1,-}) - F(U_{i+1/2}^{n+1,-}) - F(\mathcal{V}_{i-1/2}^{n+1,+}) + F(U_{i-1/2}^{n+1,+})\Big) \\ &\quad + \Delta t\left(S(P_i^{n+1}(x_i)) - S(U_i^{n,*}(x_i))\right) H_x(x_i), \\ \bar{U}_i^{n+1} &= \bar{U}_i^n + \bar{U}_i^{n+1,f}.\end{aligned} \tag{37}$$

Here

$$U_i^{*,n}(x) = \bar{U}_i^n \exp\left(H(x) - H(x_i)\right) \text{ and } P_i^{n+1}(x) = U_i^{n,*}(x) + \bar{U}_i^{n+1,f}, \tag{38}$$

$H_{i+1/2}^0 = \max\left(H_{i+1/2}^-, H_{i+1/2}^+\right)$ and $\mathcal{V}_{i+1/2}^{n+1,\pm} = \mathcal{V}_{i+1/2}^{n+1,\pm}(H_{i+1/2}^0)$ are defined as follows:

$$\mathcal{V}_{i+1/2}^{n+1,\pm}(H_{i+1/2}^0) = U_{i+1/2}^{n+1,\pm} \exp\left(H_{i+1/2}^0 - H_{i+1/2}^{\pm}\right), \tag{39}$$

where

$$U_{i+1/2}^{n+1,-} = P_i^{n+1}(x_{i+1/2}) \text{ and } U_{i+1/2}^{n+1,+} = P_{i+1}^{n+1}(x_{i+1/2}). \tag{40}$$

Finally,

$$F_{i+1/2}^{n+1} = \mathbb{F}(\mathcal{V}_{i+1/2}^{n+1,-}(H_{i+1/2}^0), \mathcal{V}_{i+1/2}^{n+1,+}(H_{i+1/2}^0)), \tag{41}$$

with

$$\mathbb{F}(U^-, U^+) = \frac{F(U^-) + F(U^+)}{2} - \frac{\gamma}{2}\left(U^+ - U^-\right)$$

where $\gamma = \max_i |\bar{U}_i^n|$.

Similarly, the first-order numerical scheme (31) writes the same except that $\mathcal{V}_{i+1/2}^{n+1,\pm}$ is replaced now for $\widetilde{\mathcal{V}}_{i+1/2}^{n+1,\pm}$ that is defined in this particular case as

$$\widetilde{\mathcal{V}}_{i+1/2}^{n+1,\pm}(H_{i+1/2}^0) = \mathcal{V}_{i+1/2}^{n,\pm}(H_{i+1/2}^0) + U_{i+1/2}^{n+1,f,\pm}(1 + \Delta H_{i+1/2}^{\pm}), \tag{42}$$

where

$$U_{i+1/2}^{n+1,f,-} = \bar{U}_i^{n+1,f}, \quad U_{i+1/2}^{n+1,f,+} = \bar{U}_{i+1}^{n+1,f}.$$

The second-order method writes similarly, except that a second-order exactly well-balanced reconstruction operator is used. In particular here we use the one proposed in [11], that gives a 5-point stencil implicit algorithm. As in [6], we have also considered a non-well balanced path-conservative numerical method based on the family of straight segments. In what follows, we will denote by

- WB_o1_Exact and WB_o2_Exact the first- and second-order numerical scheme (23), respectively,