Numerical Schemes for the Barenblatt Model of Non-Equilibrium Two Phase Flow in Porous Media

Denise Aregba-Driollet *

Gabriella Bretti † Roberto Natalini ‡

Abstract

We introduce some numerical approximations to a quasilinear problem, proposed by G. I. Barenblatt to describe non-equilibrium two-phase fluid flow in permeable porous media, with the application to the secondary oil recovery from natural reservoirs. Taking into account the theoretical results of global existence and uniqueness, approximated solutions are computed by three numerical schemes, namely, the Diagonal First Order schemes (DFO and DFO2) and the Diagonal Second Order scheme (DSO). For DFO scheme convergence is proved. The schemes' behaviour is analysed and discussed through some numerical experiments.

Key Words: scalar conservation laws, two-phase flow.
AMS Subject Classifications: Primary 65M06. Secondary 76S05, 35L65.

*Mathématiques Appliquées de Bordeaux, Université Bordeaux 1, 351 Cours de la Liberation, F-33405 Talence, France; E-mail: aregba@math.u-bordeaux.fr

†Department of Information Engineering and Applied Mathematics of the University of Salerno, via Ponte don Melillo, 84084 Fisciano (SA), Italy E-mail: g.bretti@iac.cnr.it

‡Istituto per le Applicazioni del Calcolo "M. Picone", Viale del Policlinico 137, 00161 - Roma, Italy; E-mail: r.natalini@iac.cnr.it.

1 Introduction

In this paper we deal with a Barenblatt model for non-equilibrium two-phase flow in a porous medium, where nonconvex flux function arises. This model has its main application to the simulation of the "secondary recovery" in oil reservoirs. An investigation of intermediate asymptotic solutions, namely, the travelling waves, for this problem and a study of the capillary-imbibition problem is in [3].
The structure of an oil or gas deposit is considerably complicated. A *porous medium* consists of a huge number of randomly located grains of various shapes and sizes. This complex and irregular structure of the pore space make it impossible to use the ordinary methods of hydrodynamics, as it corresponds to solve equations of viscous fluid motion in an aggregate region of all the pores. The seepage theory is based on the assumption that the porous medium and the fluid in it form a continuum. In this framework, the features of a porous medium are described by a set of geometrical characteristics, with the use of a reduced number of averaged properties. In order to describe better two-phase fluid flows in permeable porous media from the physical viewpoint, Barenblatt and his co-authors [1, 2, 4] developed a non-equilibrium model, which in the one space dimensional case reads:

$$\tau \partial^2_{xt} f(u) + \partial_x f(u) + \partial_t u = 0 \qquad \text{in } Q = (0, L) \times (0, T), \tag{1.1}$$

$$u(x, 0) = u_0(x) \qquad \text{in } (0, L), \tag{1.2}$$

$$u(0, t) = u_1(t) \qquad \text{in } (0, T), \tag{1.3}$$

where $\tau > 0$ is a relaxation time parameter and f, $u_0 = u_0(x)$, $u_1 = u_1(t)$ are given functions; the variables x and t are space and time coordinates.
From an analytical point of view, the problem of existence and uniqueness of solutions of (1.1)-(1.3) has been studied by Natalini and Tesei [7]. The main difficulty was that this quasilinear Goursat problem presents new features with respect to standard Goursat problems. The loss of hyperbolicity, due to the fact that f' vanishes at some points, determines a problem that is strictly illposed. Hence, global solutions for general boundary and initial data may not exist. In [7] it was proven that, if the initial and boundary data lie into an *admissibility* region, the so-called hyperbolicity region, a comparison principle holds and solutions of the problem must take values in the same region at any time and then exist globally. Under these assumptions we want to find numerical schemes that mimic the analytical behaviour of solutions, thus preserving monotonicity. We propose three numerical methods: two explicit, namely the Diagonal First Order (DFO) and its improved version (DFO2), and one implicit, namely the Diagonal Second Order (DSO). The main result is convergence of DFO proved in Section 5 through consistency and stability. First, we examine the monotonicity property on the single cell, then we extend it all over the domain, thus

obtaining the stability of the scheme. From consistency investigation we deduce also that DSO scheme is a second order method, while the DFO scheme is first order accurate. A study is carried out of the coefficients of the DSO scheme in order to ensure the monotonicity property at least on the corner $x = t = 0$ of the first cell. Notice that, in general, this property may fail in the other cells.
The organization of the paper is the following. In Section 2 we describe the Barenblatt model, while Section 3 recalls the analytical results of global existence and uniqueness. In Section 4 we give a description of the numerical schemes. Sections 5 and 6 are devoted to the analysis of the schemes. In particular, convergence is proved only for DFO scheme. Although for both DFO2 and DSO we cannot ensure the conservation of monotonicity all over the domain, in our tests such schemes do not produce non-existence problems. In order to show the behaviour of our approximations some numerical tests are presented and discussed in Section 7.

2 The physical background

For a full description of the model we refer to the book [2], see also [8] for a standard reference in the field. Reservoir simulation is the process of inferring the behaviour of a real reservoir from the performance of a model of the reservoir itself. By seepage flow we mean the simultaneous flow in porous media of some immiscible fluids, that in petroleum reservoirs are water and oil. The fundamental model for this problem, where non-equilibrium effects are not considered, is the Buckley-Leverett equation [5], that will be expressed in terms of the generalized Darcy's law for two-phase flow. Start with the equations of Muskat and Meres [6] and Leverett

$$W_i = -\left(\frac{kf_i(v)}{\mu_i}\right)\nabla p_i, \quad i = 1, 2, \tag{2.1}$$

$$p_2 - p_1 = p_c(v) = \alpha\left(\frac{m}{k}\right)^{1/2} J(v), \tag{2.2}$$

which represent the seepage laws of each of the fluids and properly describe the flow of a two-phase fluid in the presence of equilibrium of the phases in the pores, where W_i is the flow velocity of each phase, v is the water saturation (part of the porous volume occupied by water) in a mixture of oil and water, f_i are dimensionless quantities known as relative permeabilities, μ_i are the viscosities of the fluids, k is the permeability of the medium, p_c the capillary pressure; the quantities p_1 and p_2 are the pressures, respectively, in the wetting fluid (water) and in the nonwetting one, $J(v)$ is the Leverett function, α is the interfacial tension between the two fluids and m the porosity of the medium.
The equations of mass conservation for the two phases, under the assumption of incompressibility of the fluids and of the porous medium, can be written as

$$m\partial_t v + \nabla \cdot W_1 = 0, \quad m\partial_t v - \nabla \cdot W_2 = 0, \tag{2.3}$$

and together with (2.1)-(2.2) constitute a close system for a two-phase incompressible fluid.
To give a convenient reformulation of this system, let us denote by

$$W = W_1 + W_2$$

the total flow velocity of both phases, and set

$$\Phi(v) = f_1(v) + \frac{\mu_1}{\mu_2} f_2(v), \;\; f(v) = \frac{f_1(v)}{\Phi(v)}.$$

Therefore the Darcy's law for the bulk flow velocity is

$$W = -\left(\frac{k\Phi(v)}{\mu_1}\right) \nabla P, \tag{2.4}$$

where P is the mean pressure, which, for incompressible fluids, is

$$P = p_1 f(v) + p_2[1 - f(v)] - \int_v^1 p_c(v) f'(v) dv. \tag{2.5}$$

Thus, we obtain the system

$$\nabla \cdot [\Phi(v) \nabla P] \;\; = \;\; 0, \tag{2.6}$$

$$\partial_t v - \nabla \cdot \left[\left(\frac{k f_1(v)}{m\mu_1}\right) \nabla P\right] \;\; = \;\; a^2 \triangle \Psi(v), \tag{2.7}$$

where

$$\Psi(v) = -\int_0^v J'(v) f_2(v) f(v) dv \text{ and } a^2 = \frac{\alpha}{\mu_2}\left(\frac{k}{m}\right)^{1/2}.$$

In one dimension (2.6) reads

$$\partial_x P = \frac{C}{\Phi(v)}. \tag{2.8}$$

So, neglecting the capillary effects ($a = 0$), (2.7) becomes

$$\partial_t v - \partial_x \left[\frac{kC f_1(v)}{m\mu_1(f_1(v) + \frac{\mu_1}{\mu_2} f_2(v))}\right] = 0. \tag{2.9}$$

This equation is derived under the assumption of stationarity during the saturation variation. However, in natural aquifers of oil and gas reservoirs usually occur important nonequilibrium phenomena determined by the delay of establishing the equilibrium phase distribution after a saturation change. The time required for the equilibrium, namely τ, characteristic for the given medium and fluid pair, can be very long (months, years), hence it considerably influences the reservoir performance. A simple model, which takes into account the presence of nonequilibrium effects, was proposed by Barenblatt and coauthors in [1, 4]. In this model, the generalized Darcy's law (2.1) is substituted by:

$$W_i = -\left(\frac{k}{\mu_i}\right) f_i(u)\nabla p_i, \tag{2.10}$$

where u is the *effective saturation*, which is related to the *actual saturation* v by the relation:

$$u = v + \tau\partial_t v, \tag{2.11}$$

that takes into account the delay in reaching the equilibrium for the water phase. This model is based on the following assumptions: τ is constant and the pressure difference between the phases $p_2 - p_1$ satisfies the relation

$$p_2 - p_1 = p_c(u), \tag{2.12}$$

with p_c the equilibrium capillary pressure function. The conservation of mass equations (2.3) and relations (2.10), (2.11), and (2.12), can be written as a system of equations

$$\nabla\cdot[\Phi(u)\nabla P] = 0, \tag{2.13}$$
$$\partial_t v + \nabla\cdot[f(u)W] = 0. \tag{2.14}$$

Then, using in (2.11) the relation (2.4), with u replacing v, and the equation (2.14), we obtain, again neglecting the capillarity effects,

$$v = u - \frac{\tau}{m}\nabla\cdot\left[f(u)\frac{k\Phi(u)}{\mu_1}\nabla P\right]. \tag{2.15}$$

Now, using the expression of v founded in this last formula into the equation (2.14) yields

$$m\partial_t u - \tau\partial_t\left[\nabla\cdot\left(f(u)\frac{k\Phi(u)}{\mu_1}\nabla P\right)\right] - \nabla\cdot\left[f(u)\frac{k\Phi(u)}{\mu_1}\nabla P\right] = 0, \tag{2.16}$$

that in one dimension, recalling (2.8), reads:

$$m\partial_t u - \tau\frac{kC}{\mu_1}\partial^2_{xt}f(u) - \frac{kC}{\mu_1}\partial_x f(u) = 0. \tag{2.17}$$

In the sequel we shall consider a rescaled version of this equation.

3 Analytical framework

In this section we recall some analytical results proved in [7]. Let us study the following quasilinear Goursat problem:

$$\tau\partial^2_{xt}f(u) + \partial_x f(u) + \partial_t u = 0 \qquad \text{in } Q = (0,L)\times(0,T), \tag{3.1}$$

$$u(x,0) = u_0(x) \qquad \text{in } (0,L), \tag{3.2}$$

$$u(0,t) = u_1(t) \qquad \text{in } (0,T), \tag{3.3}$$

where $\tau > 0$ is a constant and f, $u_0 = u_0(x)$, $u_1 = u_1(t)$ are given functions; the variables x and t are properly renormalized space and time coordinates. Here $v \in [0,1]$ is the water saturation and the flux function $f = f(v)$ is a universal nondecreasing function with both domain and range in $[0,1]$. From the experiments f is s-shaped and satisfies:

$$\begin{aligned} f(0) = &\quad 0, \quad f(1) = \quad 1, \\ f'(0) = &\quad 0, \quad f'(1) = \quad 0. \end{aligned} \tag{3.4}$$

In the standard models, $f'(u) > 0$ for $0 < u < 1$. However, the case when $f(u) \equiv 0$ in the interval $[0, u_0]$ and $f(u) \equiv 1$ in $[u_1, 1]$ $(0 < u_0 < u_1 < 1)$ can be considered along the same lines.
In order to give a well-posed problem, it is necessary to supplement equation (3.1) by appropriate initial-boundary conditions. In the following we are interested in global solutions to the Goursat problem (3.1)-(3.3), namely, in solutions defined in all of Q. Global existence and uniqueness of solutions to (3.1)-(3.3) hold under the following assumptions on the functions f, u_0, u_1:

(A_0) $f \in C^2(\mathbb{R})$;

(A_1) $u_0 \in Lip([0,L]), u_1 \in L^\infty((0,T))$ and

$$\lim_{t \to 0^+} \frac{1}{t} \int_0^t |u_1(\theta) - u_0(0)| d\theta = 0;$$

(A_2) there exists an open interval $I \subseteq \mathbb{R}$ such that $f'(u) > 0$ for any $u \in I$;

(A_3) there exists a compact subset $K \subset I$ such that $\{u_0 + \tau[f(u_0)]'\}(x) \in K$ for almost every $x \in (0,L)$, $u_0(x), u_1(t) \in K$, for almost every $t \in (0,T)$, with τ fixed.

About assumption (A_2), observe that the derivative of the flux function f vanishes at the endpoints of the interval I; this is exactly as in the case of Buckley-Leverett model, where $I = (0,1)$ (see [4]). It is easily seen that the hyperbolic character of (3.1) is lost whenever the derivative of f vanishes, therefore, to have global solutions, we need to take admissible data. This means that both the initial value $v_0 := u_0 + \tau[f(u_0)]'$ of the actual saturation and the boundary value u_1 of the effective saturation must lie inside the hyperbolicity region. As a direct consequence of comparison results (see Theorem 3.2), admissible solutions of problem (3.1)-(3.3), under the natural assumptions (A_0)-(A_3), always take values in the same region, hence solutions of this kind exist globally.

3.1 Existence and comparison results

Let us consider initial and boundary data satisfying assumption (A_3). Denoting by Q_δ $(\delta > 0)$ any connected open subset of Q of the following type:

$$Q_\delta := \bigcup_{R \in \mathcal{R}_\delta} R,$$

where $\mathcal{R}_\delta$ is an arbitrary family of rectangles $R := (0, x) \times (0, t) \subseteq Q$ such that $(0, L) \times (0, \delta) \in \mathcal{R}_\delta$, $(0, \delta) \times (0, T) \in \mathcal{R}_\delta$, we have the following

Definition 3.1 *Let assumptions (A_0)-(A_1) be satisfied. By an* admissible solution to problem *(3.1)-(3.3)* in a region Q_δ *we mean a locally bounded function u such that $\partial_x f(u) \in L^\infty(Q_\delta)$ and moreover:*

(i) for any $\phi \in C_0^\infty(Q_\delta), \phi \geq 0$,

$$\iint_{Q_\delta} \{[u + \tau \partial_x f(u)] \partial_t \phi + f(u) \partial_x \phi\} \,\mathrm{d}\xi \,\mathrm{d}\theta = 0; \tag{3.5}$$

(ii) for any interval $(x_1, x_2) \subseteq (0, L)$,

$$\lim_{t \to 0^+} \frac{1}{t} \int_0^t \int_{x_1}^{x_2} \{|u(\xi, \theta) - u_0(\xi)|$$

$$+ \tau |\partial_x f(u)(\xi, \theta) - \partial_x f(u_0)(\xi)|\} \,\mathrm{d}\xi \,\mathrm{d}\theta = 0; \tag{3.6}$$

(iii) for any interval $(t_1, t_2) \subseteq (0, T)$,

$$\lim_{x \to 0^+} \frac{1}{x} \int_0^x \int_{t_1}^{t_2} |u(\xi, \theta) - u_1(\theta)| \,\mathrm{d}\xi \,\mathrm{d}\theta = 0. \tag{3.7}$$

In the following we only consider admissible solutions. The main global existence and uniqueness result for solutions of (3.1)-(3.3) proved in [7] is the following.

Theorem 3.1 *Let assumptions (A_0)-(A_3) be satisfied. Then there exists a unique global bounded solution u to problem (3.1)-(3.3). Moreover, if $u_0 \in C^1((0, L))$, $u_1 \in C^1((0, T))$ and $u_0(0) = u_1(0)$, then $u \in C^1(Q)$ and $\partial^2_{xt} u \in C(Q)$.*

One of the main tools to prove this statement, is a convenient associate problem. Let I be the open interval given by assumption (A_2). Then for every $z \in f(I)$ we can define the inverse function $f^{-1}(z)$. In order to obtain the mentioned results, it is useful to rewrite (3.1) as a system for the new unknowns

$$v := u + \tau \partial_x f(u), \qquad z := f(u), \tag{3.8}$$

the former of which is the actual saturation. As long as $z \in f(I)$

$$\begin{cases} \partial_t v &= -\frac{1}{\tau}\Big(v - f^{-1}(z)\Big), \\ \partial_x z &= \frac{1}{\tau}\Big(v - f^{-1}(z)\Big), \end{cases} \tag{3.9}$$

$$v(\cdot,0) = v_0(x) \quad \text{in } (0,L), \tag{3.10}$$

$$z(0,\cdot) = z_1(t) \quad \text{in } (0,T). \tag{3.11}$$

Concerning initial and boundary values of problem (3.9)-(3.11) we assume:

(A_1') $v_0 \in L^\infty((0,L))$, $z_1 \in L^\infty((0,T))$ and there exists a constant $\gamma \in \mathbb{R}$ such that
$$\lim_{t\to 0^+} \frac{1}{t}\int_0^t |z_1(\theta) - \gamma| d\theta = 0.$$

We can define admissible solutions to problem (3.9)-(3.11) similarly to Definition 3.1, see [7]. Under some suitable and natural conditions it is possible to prove the equivalence between (3.1)-(3.3) and (3.9)-(3.11). The next comparison result, proved in [7], for the solutions of problem (3.9)-(3.11) plays a key role in the above analysis also in view of the numerical results.

Theorem 3.2 *Let assumptions* (A_0), (A_1') *and* (A_2) *be satisfied. Let* $(v,z), (\tilde{v},\tilde{z})$ *be solutions to problem (3.9)-(3.11) in a region* Q_δ *with data* $(v_0,z_1), (\tilde{v}_0,\tilde{z}_1)$ *respectively. Let* $z(x,t), \tilde{z}(x,t) \in f(I)$ *for almost every* $(x,t) \in Q_\delta$. *Then for almost every* $(x,t) \in Q_\delta$,

$$\begin{aligned} &\textstyle\int_0^x [v-\tilde{v}]_+(\xi,t)d\xi + \int_0^t [z-\tilde{z}]_+(x,\theta)d\theta \\ &\le \textstyle\int_0^x [v_0-\tilde{v}_0]_+(\xi)d\xi + \int_0^t [z_1-\tilde{z}_1]_+(\theta)d\theta. \end{aligned} \tag{3.12}$$

In particular, if $v_0 \le \tilde{v}_0$ *a.e. in* $(0,L)$ *and* $z_1 \le \tilde{z}_1$ *a.e. in* $(0,T)$, *then*

$$v \le \tilde{v}, \quad z \le \tilde{z} \text{ a.e. in } Q_\delta.$$

Finally, there exists one unique solution of problem (3.9)-(3.11) in Q_δ.

It has been proved in [7] that the above results for problem (3.9)-(3.11) have a counterpart for (3.1)-(3.3). Under the hypothesis of Theorem 3.2, setting $u := f^{-1}(z)$ and $\tilde{u} := f^{-1}(\tilde{z})$, the functions u and $\tilde{u}$ are solutions of problem (3.1)-(3.3); hence a comparison principle follows for this problem. In particular, in contrast with the general situation for quasilinear Goursat problems, we have the result below:

Corollary 3.3 *Let the assumptions of Theorem 3.1 and Theroem 3.2 be satisfied. Then there exists at most one solution to problem (3.1)-(3.3) in* Q_δ.

3.2 An example of non-existence

A solution to problem (3.1)-(3.3) need not be global, if the derivative of the flux function vanishes on the range of the solution itself; this is shown by the example below, taken from [7].

Example 3.1 *Let $f(u) = u^2$; seek positive travelling wave solutions of (3.1) of the form*

$$u(x,t) = \psi(x + \alpha t), \tag{3.13}$$

for some functions ψ and $\alpha \in (0, \infty)$. Then function ψ must satisfy the equation

$$\psi' = \frac{C - \psi^2 - \alpha\psi}{2\alpha\tau\psi}, \tag{3.14}$$

with

$$C = -\tau\alpha(\psi^2(0))' + \psi^2(0) + \alpha\psi(0).$$

If $C < 0$ and $\psi(0) > 0$, there exists a real value $\bar{\xi}$ such that $\psi(\bar{\xi})$ and $(\psi^2)'(\bar{\xi}) = \frac{C}{\alpha\tau} < 0$. Observe that the function $\psi = \psi(\xi)$ is only defined on the interval $[0, \bar{\xi}]$, nor can it be continued in any sense as a solution of (3.14) for $\xi > \bar{\xi}$. Consider now problem (3.1)-(3.3) with $f(u) = u^2$ and data

$$u_0(x) := \psi(x), \quad u_1(t) := \psi(\alpha t). \tag{3.15}$$

Fix L, T such that $L \in (0, \bar{\xi})$, $T \in (0, \frac{\bar{\xi}}{\alpha})$ and $L + \alpha T > \bar{\xi}$. Then the function $u(x,t) := \psi(x + \alpha t)$ solves problem (3.1)-(3.3) in the region

$$\tilde{Q}_1 := \{(x,t) \in Q; \alpha t + x \leq \bar{\xi}\}.$$

This solution vanishes on the straight line $\alpha t + x = \bar{\xi}$; moreover, $\partial_t(u^2) = \alpha C < 0$ on the same line. It follows that no continuation, even in the distributional sense, is possible beyond this line for u as a solution of problem (3.1)-(3.3) with data given by (3.15).

Observe for further reference that in Example 3.1 the first requirement of assumption (A_3) is not satisfied (although the second is).

3.2.1 Relaxation results

Here we present the results of [7] about the behaviour of solutions to (3.1)-(3.3) as the relaxation time goes to zero. Consider the initial-boundary value problem:

$$\partial_t w + \partial_x f(w) = 0 \quad \text{in } Q, \tag{3.16}$$

$$w(\cdot, 0) = w_0 \quad \text{in } (0, L), \tag{3.17}$$

$$w(0, \cdot) = w_1 \quad \text{in } (0, T). \tag{3.18}$$

The following Theorem was proved in [7].

Theorem 3.4 *Let the assumptions of Theorem 3.1 be satisfied; moreover, let* $u_0' \in BV((0,L))$, $u_1 \in BV((0,T))$. *Let* u^τ *denote the unique global solution of (3.1)-(3.3) in* Q, (v^τ, z^τ) *the corresponding global solution of (3.9)-(3.11). Then as* $\tau \to 0^+$

$$v^\tau \to \bar{u} \text{ in } C([0,T]; L^1((0,L))), \tag{3.19}$$

$$u^\tau \to \bar{u} \text{ in } C([0,T]; L^1((0,L))), \tag{3.20}$$

$$z^\tau \to f(\bar{u}) \text{ in } C([0,L]; L^1((0,T))), \tag{3.21}$$

where $\bar{u}$ *indicates the unique entropy solution of problem (3.16)-(3.18) with data* $w_0 = u_0$, $w_1 = u_1$.

This shows that Barenblatt problem ($\tau > 0$) represents only a non singular perturbation of the original Buckley-Leverett equation ($\tau = 0$). This remark will be explorated for the numerical schemes.

4 Numerical schemes

In this section we present three numerical schemes for the approximation of the solutions of the Barenblatt problem (3.1)-(3.3). In particular, the Diagonal First Order (DFO and DFO2) and the Diagonal Second Order (DSO) schemes are inspired by the diagonal problem (3.9) and they are derived rewriting the equation (3.1) in the unknowns v, z introduced in (3.8). Therefore in all cases we discretize the system (3.9). The main difference between these schemes is represented by the fact that both DFO and DFO2 are finite volume methods, while DSO is a finite difference scheme. In the sequel the inverse of f, which exists since $f' > 0$, will be indicated by $g(z) = f^{-1}(z)$ and we restrict ourselves to the study of the schemes on a uniform grid, that is:

- in space: $(0, L) = \cup_{0 \le j \le J-1} [j\Delta x, (j+1)\Delta x]$,
- in time: $(0, T) = \cup_{0 \le n \le N-1} [n\Delta t, (n+1)\Delta t]$;

the coefficients $\Delta x, \Delta t$ represent space and time steps, and we call:

$$x_j = j\Delta x, \quad t_n = n\Delta t, \quad I_j = [x_j, x_{j+1}[.$$

Recalling that

$$\begin{cases} v = u + \tau \partial_x f(u), \\ z = f(u), \end{cases}$$

the initial-boundary data for $\tau > 0$ are:

$$\begin{cases} v(x,0) = u_0(x) + \tau f'(u_0(x)) u_0'(x), \quad x > 0, \\ z(0,t) = f(u_1(t)), \quad t > 0. \end{cases} \tag{4.1}$$

As we look for a scheme owning a good relaxation limit, it is convenient to first approximate u_0 and u_1 and then to put

$$\begin{cases} v_j^0 &= u_j^0 + \tau f'(u_j^0) d_j, \\ z_0^n &= f(u_0^n). \end{cases}$$

Here d_j is an approximation of u_0' on I_j. For example:

$$d_j = \frac{u_{j+1}^0 - u_{j-1}^0}{2\Delta x} \text{ for } j = 1, \dots, J-2,$$
$$d_0 = \frac{u_1^0 - u_0^0}{\Delta x}, \quad d_{j-1} = \frac{u_{j-1}^0 - u_{j-2}^0}{\Delta x}.$$

4.1 Diagonal First Order method

Let us consider the problem (3.9)-(3.10)-(3.11). We can rewrite (3.9) as

$$\begin{cases} \partial_t v = -\frac{1}{\tau}(v - g(z)), \\ \partial_x z + \partial_t v = 0, \end{cases} \tag{4.2}$$

where $g = f^{-1}$. Inspired by reconstruction-transport-projection methods for conservation laws, we define (v, z) as follows:

$$v_h(x,t) = \sum_{j=0}^{J-1} \sum_{n=0}^{N-1} \left[v_j^n + \frac{t - t_n}{\Delta t}(v_j^{n+1} - v_j^n) \right] \chi_{I_j}(x) \chi_{[t_n, t_{n+1}[}(t), \tag{4.3}$$

$$z_h(x,t) = \sum_{j=0}^{J-1} \sum_{n=0}^{N-1} \left[z_j^n + \frac{x - x_j}{\Delta x}(z_{j+1}^n - z_j^n) \right] \chi_{I_j}(x) \chi_{[t_n, t_{n+1}[}(t). \tag{4.4}$$

Function v_h is continuous with respect to t, z_h is continuous with respect to x. The exact solution satisfies

$$v(x, t_{n+1}) = e^{-\frac{\Delta t}{\tau}} v(x, t_n) + \frac{1}{\tau} \int_{t_n}^{t_{n+1}} e^{-\frac{t_{n+1}-s}{\tau}} g(z(x,s)) ds, \tag{4.5}$$

$$\int_{t_n}^{t_{n+1}} [z(x_{j+1}, s) - z(x_j, s)] ds + \int_{x_j}^{x_{j+1}} [v(x, t_{n+1}) - v(x, t_n)] dx = 0. \tag{4.6}$$

We look for (v_h, z_h) such that (v_h, z_h) verify:

$$\begin{cases} \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} v_h(x, t_{n+1}) dx = \frac{1}{\Delta x} \exp\left(-\frac{\Delta t}{\tau}\right) \int_{x_j}^{x_{j+1}} v_h(x, t_n) \\ + \frac{1}{\Delta x} \frac{1}{\tau} \int_{x_j}^{x_{j+1}} \int_{t_n}^{t_{n+1}} \exp\left(-\frac{t_{n+1}-s}{\tau}\right) g(z_h(x,s)) ds dx, \\ \int_{t_n}^{t_{n+1}} [z_h(x_{j+1}, s) - z_h(x_j, s)] ds + \int_{x_j}^{x_{j+1}} [v_h(x, t_{n+1}) - v_h(x, t_n)] dx = 0. \end{cases} \tag{4.7}$$

Recalling that on I_j v_h does not depend on x, the first equation in (4.7) is

$$\begin{aligned} v_j^{n+1} &= v_j^n e^{-\frac{\Delta t}{\tau}} + \frac{1}{\tau \Delta x} \int_{x_j}^{x_{j+1}} \int_{t_n}^{t_{n+1}} e^{-\frac{t_{n+1}-t}{\tau}} \\ & g\left(z_j(t) + \frac{x - x_j}{\Delta x}(z_{j+1}(t) - z_j(t)) \right) dt dx. \end{aligned} \tag{4.8}$$

We approximate (4.8) by the formula:

$$v_j^{n+1} = v_j^n e^{-\frac{\Delta t}{\tau}} + (1 - e^{-\frac{\Delta t}{\tau}}) g(z_j^n).$$

Recalling that z_h does not depend on t, the second equation in (4.7) is

$$\Delta t(z_{j+1}^n - z_j^n) + \Delta x(v_j^{n+1} - v_j^n) = 0,$$

so that we obtain

$$z_{j+1}^n = z_j^n - \frac{\Delta x}{\Delta t}(v_j^{n+1} - v_j^n). \tag{4.9}$$

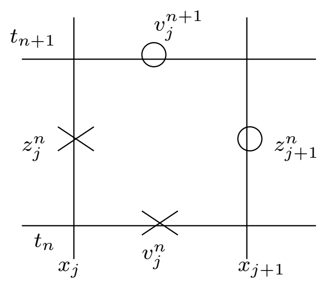

Figure 1: The crosses indicate the data points, z_j^n, v_j^n, while the circles indicate the unknown points, z_{j+1}^n, v_j^{n+1}.

Therefore we have the following explicit scheme:

$$\begin{cases} v_j^{n+1} = v_j^n e^{-\frac{\Delta t}{\tau}} + (1 - e^{-\frac{\Delta t}{\tau}}) g(z_j^n) \\ z_{j+1}^n = z_j^n - \nu(v_j^{n+1} - v_j^n) \end{cases} \tag{4.10}$$

for $n = 0, \ldots, N-1$ and $j = 0, \ldots, J-1$, with $\nu = \frac{\Delta x}{\Delta t}$. Here we approximate u_0 and u_1 by

$$\begin{aligned} u_j^0 &= \frac{1}{\Delta x} \int_{I_j} u_0(x) dx, \quad j = 0, \ldots, J-1, \\ u_0^n &= \frac{1}{\Delta t} \int_{J_n} u_1(t) dt, \quad n = 0, \ldots, N-1. \end{aligned}$$

4.1.1 Higher order for DFO method

We look for an approximation scheme of higher order respect to DFO method. As before, the conservation equation is approximated by (4.9). Then we search for a second order approximation of v_j^n, that is a second order scheme for the equation (4.5). In (4.4) we replace z_j^n by the piecewise linear function

$$z_j(t) = z_j^n + \sigma_j^n (t - t_{n+1/2}) \quad \text{for } t \in [t_n, t_{n+1}[, \quad j \geq 0,$$

where $t_{n+1/2} = (t_n + t_{n+1})/2$ and σ_j^n is a constant which can be defined as

$$\begin{aligned} \sigma_j^n &= minmod \left(\frac{z_j^{n+1} - z_j^n}{\Delta t}, \frac{z_j^n - z_j^{n-1}}{\Delta t} \right), \text{ for } n = 1, \dots, N-2, \\ \sigma_j^0 &= \sigma_j^{N-1} = 0. \end{aligned}$$

As for DFO, we have the equation (4.8) and the last term is approximated by:

$$I = \frac{1}{\tau} \int_{t_{n+1}}^{t_n} e^{-\frac{t_{n+1}-t}{\tau}} g(z_j(t)) dt.$$

Now we approach $g(z_j(t))$ by:

$$L_j(t) = -\frac{g_j^+ - g_j^-}{\Delta t}(t_{n+1} - t) + g_j^+, \quad t \in [t_n, t_{n+1}[,$$

with

$$g_j^+ = g\left(z_j^n + \frac{\Delta t}{2}\sigma_j^n\right), \quad g_j^- = g\left(z_j^n - \frac{\Delta t}{2}\sigma_j^n\right).$$

Then, computing exactly the integral

$$\frac{1}{\tau} \int_{t_{n+1}}^{t_n} e^{-\frac{t_{n+1}-t}{\tau}} L_j(t) dt = -\frac{\tau}{\Delta t}(g_j^+ - g_j^-)(1 - e^{-\frac{\Delta t}{\tau}}) - g_j^- e^{-\frac{\Delta t}{\tau}} + g_j^+,$$

we obtain the scheme:

$$\begin{cases} v_j^{n+1} = (v_j^n - g_j^-) e^{-\frac{\Delta t}{\tau}} - \frac{\tau}{\Delta t}(g_j^+ - g_j^-)(1 - e^{-\frac{\Delta t}{\tau}}) + g_j^+, \\ z_{j+1}^n = z_j^n - \nu(v_j^{n+1} - v_j^n), \end{cases} \tag{4.11}$$

for $n = 0, \dots, N-1$ and $j = 0, \dots, J-1$.

4.1.2 Relaxation limits

Here we keep Δx and Δt fixed and we make $\tau \to 0^+$. The scheme (4.10) has the limit:

$$\begin{cases} v_j^{n+1} &= g(z_j^n) \\ z_{j+1}^n &= z_j^n - \nu(v_j^{n+1} - v_j^n). \end{cases} \tag{4.12}$$

The data own the limit:

$$\begin{cases} v_j^0 &= u_j^0, \quad j = 0, \dots, J-1, \\ z_0^n &= f(u_1^n) \quad n = 0, \dots, N-1, \end{cases} \tag{4.13}$$

hence the second equation in (4.12) reads:

$$z_{j+1}^0 = z_j^0 - \nu[g(z_j^0) - g(u_j^0)], \quad j = 0, \dots, J-2,$$

and, for $n \geq 1$:

$$z_{j+1}^n = z_j^n - \nu[g(z_j^n) - g(z_j^{n-1})], \quad j = 0, \dots, J-2.$$

We recognize the upwind scheme applied to the hyperbolic in x conservation law:

$$\begin{cases} \partial_x z + \partial_t g(z) &= 0, \quad (x,t) \in \mathbb{R}^+ \times \mathbb{R}^+, \\ z(0,t) &= f(u_1), \\ z(x,0) &= f(u_0). \end{cases} \tag{4.14}$$

This scheme is stable and consistent since $g'(z) > 0$.
The relaxation limit of scheme (4.11) is:

$$\begin{cases} v_j^{n+1} &= g(z_j^n + \frac{\Delta t}{2}\sigma_j^n), \\ \\ z_{j+1}^n &= z_j^n - \nu(v_j^{n+1} - v_j^n). \end{cases} \tag{4.15}$$

4.2 Diagonal Second Order method

Whereas the DFO is a finite volume method which considers the average of v on $[x_j, x_{j+1}]$ and the average of z on $[t_n, t_{n+1}]$ as unknowns, the DSO is a finite difference method with unknowns on the nodes. If the data are smooth, we take

$$\begin{cases} v_j^0 &= v_0(x_j), \quad j = 0, \dots, J, \\ \\ z_0^n &= z_1(t_n), \quad n = 0, \dots, N. \end{cases} \tag{4.16}$$

Otherwise we can consider averages around the nodes. This is a higher order method. Although we are not able to prove convergence results for this scheme, it is quite interesting since it is second order accurate. Recall that f is strictly increasing and $z \in f(I)$.
The second equation in (3.9) cannot be solved using the exact formula, since the differential equation considered is not linear. Hence we rewrite the mentioned equation as:

$$z_x = \frac{1}{\tau}(v - g(z)) \tag{4.17}$$

and put the equation (4.17) in the integral form

$$z_{j+1}^n = z_j^n + \frac{1}{\tau}\int_{x_j}^{x_{j+1}} (v(x,t_n) - g(z(x,t_n)))dx.$$

Then we use a quadrature formula, namely the trapezoidal rule, to approximate the integral and we obtain the implicit equation:

$$z_{j+1}^n = z_j^n + \frac{\Delta x}{2\tau}[v_{j+1}^n + v_j^n - g(z_j^n) - g(z_{j+1}^n)]. \tag{4.18}$$

The first equation in (3.9) can be solved using the exact formula for a first order linear differential equation

$$v(x, t+\Delta t) = v(x,t)e^{\int_t^{t+\Delta t} -\frac{1}{\tau}d\sigma} + \frac{1}{\tau}\int_t^{t+\Delta t} e^{-\frac{t+\Delta t - s}{\tau}} g(z(x,s))ds. \tag{4.19}$$

Now, with the trapezoidal rule we obtain the second equation of the scheme

$$v_j^{n+1} = \theta v_j^n + \bar{\alpha}[g(z_j^{n+1}) + \theta g(z_j^n)] \tag{4.20}$$

and the third one

$$v_{j+1}^{n+1} = \theta v_{j+1}^n + \bar{\alpha}[g(z_{j+1}^{n+1}) + \theta g(z_{j+1}^n)], \tag{4.21}$$

where $\theta = e^{-\frac{\Delta t}{\tau}}$ and $\bar{\alpha} = \frac{\Delta t}{2\tau}$.
By integrating the conservation law $v_t + z_x = 0$ and applying again the quadrature rule, we find the equation

$$z_{j+1}^{n+1} = z_j^n + z_j^{n+1} - z_{j+1}^n + \frac{\Delta x}{\Delta t}[v_j^n + v_{j+1}^n - v_j^{n+1} - v_{j+1}^{n+1}]. \tag{4.22}$$

Finally, the implicit scheme reads

$$\begin{cases} z_{j+1}^n = z_j^n + \frac{\Delta x}{2\tau}[v_{j+1}^n + v_j^n - g(z_j^n) - g(z_{j+1}^n)], \\ v_j^{n+1} = \theta v_j^n + \bar{\alpha}[g(z_j^{n+1}) + \theta g(z_j^n)], \\ v_{j+1}^{n+1} = \theta v_{j+1}^n + \bar{\alpha}[g(z_{j+1}^{n+1}) + \theta g(z_{j+1}^n)], \\ z_{j+1}^{n+1} = z_j^n + z_j^{n+1} - z_{j+1}^n + \nu[v_j^n + v_{j+1}^n - v_j^{n+1} - v_{j+1}^{n+1}], \end{cases} \tag{4.23}$$

for $n = 0, \dots, N-1$ and $j = 0, \dots, J-1$, setting again $\nu = \frac{\Delta x}{\Delta t}$. In order to solve the implicit equations of the scheme we proceed as follows. Let us consider the equation (4.18) and set

$$\tilde{f}_1 = z_j^n + \frac{\Delta x}{2\tau}[v_{j+1}^n + v_j^n - g(z_j^n)].$$

We need to invert the equation

$$\tilde{f}_1 = h(z_{j+1}^n),$$

where the function h is

$$h(z) = z + \frac{\Delta x}{2\tau} g(z) \tag{4.24}$$

and for the invertibility we need:

$$1 + \frac{\Delta x}{2\tau} g'(z) > C > 0. \tag{4.25}$$

This condition is always verified, as we supposed from the beginning $g' > 0$. In order to ensure that $h^{-1}(f)$ exists, we also need f to take values in $[0,1]$. This is not obvious, but numerically it is satisfied. Using for instance Newton's method we can compute the first equation of the scheme:

$$z_{j+1}^n = h^{-1}\left(z_j^n + \frac{\Delta x}{2\tau}[v_{j+1}^n + v_j^n - g(z_j^n)]\right). \tag{4.26}$$

Then, using (4.20)-(4.21)-(4.26) in (4.22) and setting

$$\tilde{f}_2 = z_j^{n+1} - \nu\bar{\alpha}g(z_j^{n+1}) + \left(\nu - \nu\theta - \frac{\Delta x}{2\tau}\right)v_j^n + \left(\nu - \nu\theta - \frac{\Delta x}{2\tau}\right)v_{j+1}^n$$
$$+\left(\frac{\Delta x}{2\tau} - \nu\bar{\alpha}\theta\right)g(z_j^n) + \left(\frac{\Delta x}{2\tau} - \nu\bar{\alpha}\theta\right)g(z_{j+1}^n),$$

the equation to be inverted is

$$\tilde{f}_2 = h(z_{j+1}^{n+1}),$$

with h as in (4.24). The invertibility condition is again (4.25) and by Newton's method the fourth equation of the scheme is computed:

$$\begin{aligned} z_{j+1}^{n+1} &= h^{-1}\left(z_j^{n+1} - \nu\bar{\alpha}g(z_j^{n+1}) + \nu(1-\theta-\bar{\alpha})(v_j^n + v_{j+1}^n)\right. \\ &\quad \left.+\nu\bar{\alpha}(1-\theta)g(z_j^n) + \nu\bar{\alpha}(1-\theta)g(z_{j+1}^n)\right). \end{aligned} \tag{4.27}$$

Finally, let us detail the equations of the scheme. At the first cell ($j = 0, n = 0$) the scheme is:

$$z_1^0 = z_0^0 + \nu\bar{\alpha}[v_1^0 + v_0^0 - g(z_0^0) - g(z_1^0)], \tag{4.28}$$
$$v_0^1 = \theta v_0^0 + \bar{\alpha}[g(z_0^1) + \theta g(z_0^0)], \tag{4.29}$$
$$v_1^1 = \theta v_1^0 + \bar{\alpha}[g(z_1^1) + \theta g(z_1^0)], \tag{4.30}$$
$$\begin{aligned} z_1^1 &= z_0^1 - \nu\bar{\alpha}g(z_0^1) + \nu(1-\theta-\bar{\alpha})v_0^0 + \nu(1-\theta-\bar{\alpha})v_1^0 \\ &\quad +\nu\bar{\alpha}(1-\theta)g(z_0^0) + \nu\bar{\alpha}(1-\theta)g(z_1^0) - \nu\bar{\alpha}g(z_1^1). \end{aligned} \tag{4.31}$$

At the boundary cells, corresponding to ($j = 0, n \geq 1$), the scheme reads:

$$v_0^{n+1} = \theta v_0^n + \bar{\alpha}[g(z_0^{n+1}) + \theta g(z_0^n)], \tag{4.32}$$
$$v_1^{n+1} = \theta v_1^n + \bar{\alpha}[g(z_1^{n+1}) + \theta g(z_1^n)], \tag{4.33}$$
$$\begin{aligned} z_1^{n+1} &= z_0^n + z_0^{n+1} - z_1^n + \nu(1-\theta)v_0^n + \nu(1-\theta)v_1^n - \nu\bar{\alpha}\theta g(z_1^n) \\ &\quad -\nu\bar{\alpha}g(z_0^{n+1}) - \nu\bar{\alpha}\theta g(z_0^n) - \nu\bar{\alpha}g(z_1^{n+1}). \end{aligned} \tag{4.34}$$

For ($j \geq 1, n = 0$) we have:

$$z_{j+1}^0 = z_j^0 + \nu\bar{\alpha}[v_{j+1}^0 + v_j^0 - g(z_j^0) - g(z_{j+1}^0)], \tag{4.35}$$
$$v_{j+1}^1 = \theta v_{j+1}^0 + \bar{\alpha}[g(z_{j+1}^1) + \theta g(z_{j+1}^0)], \tag{4.36}$$
$$\begin{aligned} z_{j+1}^1 &= z_j^1 + \nu(1-\bar{\alpha})v_j^0 + \nu(1-\theta-\bar{\alpha})v_{j+1}^0 + \nu\bar{\alpha}(1-\theta)g(z_{j+1}^0) \\ &\quad +\nu\bar{\alpha}g(z_j^0) - \nu\bar{\alpha}g(z_{j+1}^1) - \nu v_j^1. \end{aligned} \tag{4.37}$$

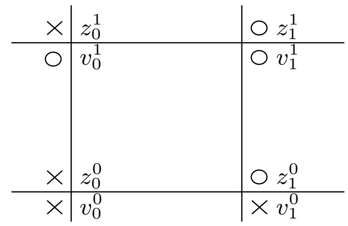

Figure 2: At the first cell: the crosses indicate the four data points, $z_0^0, z_0^1, v_0^0, v_1^0$, while the circles indicate the unknown points, $z_1^0, z_1^1, v_0^1, v_1^1$.

At a generic cell of the grid, corresponding to $(j \geq 1, n \geq 1)$, the scheme is composed by only two equations:

$$\begin{aligned} z_{j+1}^{n+1} &= z_j^n + z_j^{n+1} - z_{j+1}^n + \nu v_j^n + \nu(1-\theta) v_{j+1}^n - \nu v_j^{n+1} \\ &\quad -\nu \bar{\alpha} \theta g(z_{j+1}^n) - \nu \bar{\alpha} g(z_{j+1}^{n+1}), \end{aligned} \tag{4.38}$$

$$v_{j+1}^{n+1} = \theta v_{j+1}^n + \bar{\alpha}[g(z_{j+1}^{n+1}) + \theta g(z_{j+1}^n)]. \tag{4.39}$$

5 Consistency

Since we have regular solutions, the analysis of consistency is rigorous as long as solutions exist.

5.1 DFO scheme

Theorem 5.1 (Consistency) *The scheme (4.10) is first order accurate.*

Proof. The errors between the exact solutions computed on the grid and the approximated solutions, respectively $v(x_j, t_{n+1}), z(x_{j+1}, t_n)$ and v_j^{n+1}, z_{j+1}^n, are:

$$\begin{aligned} \epsilon_1 &= v(x_j, t_{n+1}) - v_j^{n+1}, \\ \epsilon_2 &= z(x_{j+1}, t_n) - z_{j+1}^n. \end{aligned}$$

The Taylor expansion of z_{j+1}^n about the point (x_j, t_n) is given by

$$z(x_{j+1}, t_n) = z_j^n + \Delta x \partial_x z(x_j, t_n) + \frac{\Delta x^2}{2} \partial_{xx} z(x_j, t_n) + O(\Delta x^3),$$

while the second equation of the scheme is

$$\begin{aligned} z_{j+1}^n &= z_j^n - \frac{\Delta x}{\Delta t}(v_j^{n+1} - v_j^n) \\ &= z_j^n - \frac{\Delta x}{\Delta t}(v_j^{n+1} - v(x_j, t_{n+1}) + v(x_j, t_{n+1}) - v_j^n) \\ &= z_j^n - \frac{\Delta x}{\Delta t}(\epsilon_1 + \Delta t \partial_t v(x_j, t_n) + O(\Delta t^2)). \end{aligned}$$

Then, recalling that $\partial_t v(x_j, t_n) = \partial_x z(x_j, t_n)$, we have

$$\epsilon_2 = \frac{\Delta x^2}{2}\partial_{xx} z(x_j, t_n) + O(\Delta x^3) + \frac{\Delta x}{\Delta t}\epsilon_1 + O(\Delta t \Delta x).$$

Therefore it suffices to prove that $\epsilon_1 = \Delta t[O(\Delta t) + O(\Delta x)]$. Let us now compute ϵ_1. The Taylor expansion about the point (x_j, t_n) of v_j^{n+1} is

$$\begin{aligned} v(x_j, t^{n+1}) &= v_j^n + \Delta t \partial_t v(x_j, t_n) + O(\Delta t^2) \\ &= v_j^n + \frac{\Delta t}{\tau}[-v_j^n + g(z_j^n)] + O(\Delta t^2), \end{aligned}$$

then we obtain

$$\begin{aligned} \epsilon_1 &= v(x_j, t_{n+1}) - v_j^{n+1} \\ &= v_j^n(1-\theta) + \frac{\Delta t}{\tau}[-v_j^n + g(z_j^n)] - (1-\theta)g(z_j^n) + O(\Delta t^2) \\ &= [-v_j^n + g(z_j^n)]\left(-1 + \theta + \frac{\Delta t}{\tau}\right) + O(\Delta t^2) \\ &= O\left(\frac{\Delta t^2}{\tau^2}\right) + O(\Delta t^2), \end{aligned}$$

since $-1+\theta \sim -\frac{\Delta t}{\tau} + O(\frac{\Delta t^2}{\tau^2})$ and $(-v_j^n + g(z_j^n))$ is a fixed function independent of Δt. Hence

$$\epsilon_2 = \Delta x(O(\Delta x) + O(\Delta t/\tau) + O(\Delta t))$$

and we can conclude that the scheme is first order accurate. This represents the exact rate of convergence of the scheme. □

5.2 DSO scheme

In the sequel we present the study of consistency for the DSO scheme (4.23). It results to be second order accurate and it depends on τ. In the analysis of consistency we refer to the different expressions of the scheme.

Theorem 5.2 (Consistency) *The scheme (4.23) is second order accurate.*

Proof. In order to show consistency of the method we need to consider all the different expressions of the scheme depending on which point on the grid we are solving the problem.

Lemma 5.3 (Consistency in case $j \geq 1, n \geq 1$) *Here the scheme is composed by the equations (4.39) and (4.38). The order of accuracy of the method is $O(k^3/\tau^2)$.*

Proof. The errors between the exact solutions computed on the grid and the approximated solutions, respectively $v(x_{j+1}, t_{n+1}), z(x_{j+1}, t_{n+1})$ and $v_{j+1}^{n+1}, z_{j+1}^{n+1}$, are:

$$\begin{aligned} \epsilon_1 &= v(x_{j+1}, t_{n+1}) - v_{j+1}^{n+1}, \\ \epsilon_2 &= z(x_{j+1}, t_{n+1}) - z_{j+1}^{n+1}. \end{aligned}$$

Let us write the following expansions about the point (x_j, t_n):

$$v_{j+1}^n = v_j^n + hv_x + \frac{h^2}{2}v_{xx} + O(h^3), \tag{5.1}$$

$$z_{j+1}^n = z_j^n + hz_x + \frac{h^2}{2}z_{xx} + O(h^3), \tag{5.2}$$

$$g(z_{j+1}^n) = g(z_j^n) + g'(z_j^n)(z_{j+1}^n - z_j^n) + g''(z_j^n)\frac{(z_{j+1}^n - z_j^n)^2}{2} + O(k^3), \tag{5.3}$$

$$v_j^{n+1} = v_j^n + kv_t + \frac{k^2}{2}v_{tt} + O(k^3), \tag{5.4}$$

$$z_j^{n+1} = z_j^n + kz_t + \frac{k^2}{2}z_{tt} + O(k^3), \tag{5.5}$$

$$g'(z_{j+1}^n) = g'(z_j^n) + g''(z_j^n)(z_{j+1}^n - z_j^n) + O(h^3),$$

$$\begin{aligned} g(z_{j+1}^{n+1}) &= g(z_j^n) + \frac{h}{\tau}g'(z_j^n)(v_j^n - g(z_j^n)) - \frac{h^2}{2\tau^2}(g'(z_j^n))^2(v_j^n - g(z_j^n)) \\ &\quad + \frac{h^2}{2\tau}g'(z_j^n)v_x + kg'(z_j^n)[z_t + hz_{tx} + O(h^2)] + \frac{k^2}{2}g'(z_j^n)z_{tt} + O(k^3). \end{aligned} \tag{5.6}$$

Consistency for equation (4.21) The Taylor expansion of v_{j+1}^{n+1} about (x_j, t_n) is given by

$$\begin{aligned} v(x_{j+1}, t_{n+1}) &= v_j^n\left[1 - \frac{k}{\tau} + \frac{k^2}{2\tau^2} + g'(z_j^n)\frac{hk}{\tau^2}\right] + v_x\left(h - \frac{hk}{\tau}\right) \\ &\quad + v_{xx}\frac{h^2}{2} + g(z_j^n)\left(\frac{k}{\tau} - \frac{k^2}{2\tau^2}\right) + g'(z_j^n)\left(z_t\frac{k^2}{2\tau} - g(z_j^n)\frac{hk}{\tau^2}\right) + O(k^3). \end{aligned} \tag{5.7}$$

The error is

$$\epsilon_1 = v(x_{j+1}, t^{n+1}) - \theta v_{j+1}^n - \bar{\alpha}g(z_{j+1}^{n+1}) - \bar{\alpha}\theta g(z_{j+1}^n)$$

and, putting the expansions (5.2)-(5.4)-(5.6)-(5.7) in the expression above, one has

$$\epsilon_1 = v_j^n g'(z_j^n)\frac{hk}{\tau^2} + g(z_j^n)\frac{k}{\tau} + g'(z_j^n)z_t\frac{k^2}{2\tau} - g'(z_j^n)g(z_j^n)\frac{hk}{\tau^2} - g(z_j^n)\frac{k}{2\tau}$$

$$-g'(z_j^n)(v_j^n - g(z_j^n))\frac{hk}{2\tau^2} - \frac{k}{2\tau}\left[g(z_j^n) + g'(z_j^n)(v_j^n - g(z_j^n))\frac{h}{\tau} + g'(z_j^n)z_t k\right] + O(k^3/\tau^2),$$

with $g(z_j^n), g'(z_j^n), g''(z_j^n)$ some fixed functions independent of k and the derivatives computed at (x_j, t_n). Since the low order terms cancel, we have

$$\epsilon_1 = O\left(\frac{\Delta t^3}{\tau^2}\right).$$

Consistency for equation (4.38)
Expanding in Taylor series $z(x_{j+1}, t_{n+1})$ about (x_j, t_n), we have

$$\begin{aligned} z(x_{j+1}, t_{n+1}) &= z_j^n + \frac{h}{\tau}(v_j^n - g(z_j^n)) + \frac{h^2}{2\tau} v_x - \frac{h^2}{2\tau^2} g'(z_j^n)(v_j^n - g(z_j^n)) \quad (5.8) \\ &\quad + k z_t - g'(z_j^n) z_t \frac{hk}{\tau} + g(z_j^n)\frac{hk}{\tau^2} - v_j^n \frac{hk}{\tau^2} + z_{tt}\frac{k^2}{2} + O(k^3). \end{aligned}$$

The error is given by

$$\begin{aligned} \epsilon_2 &= z(x_{j+1}, t_{n+1}) - z_j^n - z_j^{n+1} + z_{j+1}^n - \nu v_j^n - \nu(1-\theta) v_{j+1}^n + \nu v_j^{n+1} \\ &\quad + \nu\bar{\alpha}\theta g(z_{j+1}^n) + \nu\bar{\alpha} g(z_{j+1}^{n+1}) \end{aligned}$$

and now, using the expansions (5.1)-(5.2)-(5.3)-(5.4)-(5.5)-(5.6)-(5.8):

$$\begin{aligned} &= z_j^n + v_j^n\left(\frac{h}{\tau} - \frac{h^2}{2\tau^2} g'(z_j^n) - \frac{hk}{\tau^2}\right) - \frac{h}{\tau} g(z_j^n) + \frac{h^2}{2\tau} v_x + \frac{h^2}{2\tau^2} g'(z_j^n) g(z_j^n) \\ &\quad + k z_t + \frac{k^2}{2} z_{tt} - z_j^n - k z_t - \frac{k^2}{2} z_{tt} - g'(z_j^n) z_t \frac{hk}{\tau} + \frac{hk}{\tau^2} g(z_j^n) \\ &\quad + \frac{h}{2\tau} g(z_j^n) + g'(z_j^n) z_t \frac{hk}{2\tau} - v_j^n\left(\frac{h}{\tau} - \frac{hk}{\tau^2}\right) - \frac{h^2}{2\tau} v_x \left(1 - \frac{k}{\tau}\right) \\ &\quad - \frac{hk}{\tau^2} g(z_j^n) + \frac{h}{2\tau} g(z_j^n) + \frac{h^2}{2\tau^2} g'(z_j^n)(v_j^n - g(z_j^n)) + \frac{hk}{2\tau} g'(z_j^n) z_t + O(k^3) \\ &= O\left(\frac{k^3}{\tau^2}\right), \end{aligned}$$

with $g(z_j^n), g'(z_j^n), g''(z_j^n)$ some fixed functions independent of k and the derivatives computed at (x_j, t_n). As the low order terms cancel, ϵ_2 results $O(\Delta t^3/\tau^2)$, when $k \to 0$. Thus we can conclude that the scheme is of second order. □

Then, reasoning as above in the other cases, we obtain that the scheme is consistent of order two all over the domain. This ends the proof of the Theorem. □

6 Monotonicity

6.1 DFO scheme

Let us prove monotonicity on the domain for the DFO scheme. The exact solution satisfies the relation (3.12) and the analogue in the discrete case is

$$\sum_{j=0}^{J-1}\int_{x_j}^{x_{j+1}}[v_h-\tilde v_h]_+(x,t_N)dx+\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}}[z_h-\tilde z_h]_+(x_I,t)dt \tag{6.1}$$

$$\le \sum_{j=0}^{J-1}\int_{x_j}^{x_{j+1}}[v_{h,0}(x)-\tilde v_{h,0}(x)]_+dx+\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}}[z_{h,1}(t)-\tilde z_{h,1}(t)]_+dt.$$

By using (4.3)-(4.4) we obtain the following relation:

$$\sum_{j=0}^{J-1}\Delta x[v_j^N-\tilde v_j^N]_+ +\sum_{n=0}^{N-1}\Delta t[z_J^n-\tilde z_J^n]_+ \tag{6.2}$$

$$\le \sum_{j=0}^{J-1}\Delta x[v_j^0-\tilde v_j^0]_+ +\sum_{n=0}^{N-1}\Delta t[z_0^n-\tilde z_0^n]_+.$$

Denoting

$$A_j^n=\Delta x[v_j^n-\tilde v_j^n]_+,\quad B_j^n=\Delta t[z_j^n-\tilde z_j^n]_+$$

we can rewrite (6.2) in the form

$$\sum_{j=0}^{J-1}A_j^N+\sum_{n=0}^{N-1}B_J^n\le\sum_{j=0}^{J-1}A_j^0+\sum_{n=0}^{N-1}B_0^n. \tag{6.3}$$

Theorem 6.1 *The scheme (4.10) verifies the monotonicity property (6.2), under the condition*

$$\Delta x\le\frac{\Delta t}{\sup_{\zeta\in I}g'(\zeta)}.$$

Proof. In order to prove this result we need the following Lemma.

Lemma 6.2 *Monotonicity on the single cell:*

$$A_j^{n+1}+B_{j+1}^n\le A_j^n+B_j^n,\quad\forall j,n.$$

Proof. Let us introduce the following notation:

$$\gamma=v_j^n-\tilde v_j^n,\quad \gamma'=v_j^{n+1}-\tilde v_j^{n+1},$$

$$\delta=z_j^n-\tilde z_j^n,\quad \delta'=z_{j+1}^n-\tilde z_{j+1}^n.$$

Let us rewrite the first equation in (4.10) in the form:

$$v_j^{n+1} = \theta v_j^n + (1-\theta)g(z_j^n)$$

then we have

$$\begin{aligned} \gamma' &= \theta\gamma + (1-\theta)[g(z_j^n) - g(\tilde{z}_j^n)] \\ &= \theta\gamma + (1-\theta)g'(\zeta)\delta. \end{aligned} \tag{6.4}$$

Note that

$$\gamma' - \gamma = (v_j^{n+1} - v_j^n) - (\tilde{v}_j^{n+1} - \tilde{v}_j^n)$$

and

$$\delta' - \delta = (z_{j+1}^n - z_j^n) - (\tilde{z}_{j+1}^n - \tilde{z}_j^n).$$

From the second equation in (4.10) one can write

$$\begin{cases} z_{j+1}^n - z_j^n &= -\nu(v_j^{n+1} - v_j^n), \\ \tilde{z}_{j+1}^n - \tilde{z}_j^n &= -\nu(\tilde{v}_j^{n+1} - \tilde{v}_j^n). \end{cases}$$

Thus we have

$$\begin{aligned} \delta' &= \delta - \nu(\gamma' - \gamma) \\ &= \delta - \nu(\theta - 1)(\gamma - g'(\zeta)\delta) \\ &= \delta[1 - \nu(1-\theta)g'(\zeta)] + \gamma\nu(1-\theta). \end{aligned} \tag{6.5}$$

Since coefficients in (6.4) are non-negative, we only need to impose $\forall \zeta \in f(I)$:

$$1 - \nu(1-\theta)g'(\zeta) \geq 0 \Longleftrightarrow \frac{\Delta x}{\Delta t}g'(\zeta) \leq 1,$$

in order to have non-negative coefficients in (6.5).

Remark 6.3 *Recall that given $a, b \in \mathbb{R}$, $a \geq 0, b \geq 0$, the following yields*

$$[ax + by]_+ \leq a[x]_+ + b[y]_+. \tag{6.6}$$

Due to the positiveness of the coefficients under the condition written above, is possible to apply the rule in remark (6.3):

$$\begin{aligned} \Delta x[\gamma']_+ + \Delta t[\delta']_+ &\leq \Delta x[\theta\gamma + (1-\theta)g'(\zeta)\delta]_+ + \Delta t[\nu(1-\theta)\gamma + (1 - \nu(1-\theta)g'(\zeta))\delta]_+ \\ &\leq \Delta x[\gamma]_+ + \Delta t[\delta]_+ \end{aligned}$$

and the proof is completed. □

Now we are able to extend the result stated in the lemma above. In particular, for the properties of the telescopic sums we have

$$\sum_{j=0}^{J-1} A_j^N + \sum_{n=0}^{N-1} B_J^n =$$

$$= \sum_{j=0}^{J-1}\left[\sum_{n=0}^{N-1}(A_j^{n+1}-A_j^n)+A_j^0\right]+\sum_{n=0}^{N-1}\left[\sum_{j=0}^{J-1}(B_{j+1}^n-B_j^n)+B_0^n\right]$$

$$= \sum_{j=0}^{J-1}\sum_{n=0}^{N-1}\left[(A_j^{n+1}+B_{j+1}^n)-(A_j^n+B_j^n)\right]+\sum_{j=0}^{J-1}A_j^0+\sum_{n=0}^{N-1}B_0^n$$

and, by lemma 6.2, we obtain

$$\leq \sum_{j=0}^{J-1}A_j^0+\sum_{n=0}^{N-1}B_0^n,$$

which ends the proof of monotonicity on the whole grid. □

6.2 DSO scheme

Although it is not possible to obtain non-negative coefficients for the scheme on the generic cell of the domain, we want to ensure positiveness of the coefficients at least on the first cell. In subsection 6.3 an example of non-monotone behaviour of the scheme is presented.

Proposition 6.4 *Monotonicity on the first cell ($j=0, n=0$) holds under the following assumptions:*

$$\frac{\Delta x}{2\tau}g'(x)\leq 1, \quad \forall x\in f(I), \tag{6.7}$$

$$1-\bar{\alpha}>0 \Longleftrightarrow \frac{\Delta t}{2\tau}\leq 1. \tag{6.8}$$

Proof. The discretization of the relation (3.12) is the comparison condition:

$$\nu[v_j^{n+1}-\tilde{v}_j^{n+1}]_+ +\nu[v_{j+1}^{n+1}-\tilde{v}_{j+1}^{n+1}]_+ +[z_{j+1}^n-\tilde{z}_{j+1}^n]_+ +[z_{j+1}^{n+1}-\tilde{z}_{j+1}^{n+1}]_+ \tag{6.9}$$
$$\leq \nu[v_j^n-\tilde{v}_j^n]_+ +\nu[v_{j+1}^n-\tilde{v}_{j+1}^n]_+ +[z_j^n-\tilde{z}_j^n]_+ +[z_j^{n+1}-\tilde{z}_j^{n+1}]_+,$$

where $\nu=\frac{\Delta x}{\Delta t}$. Notice that at the first step the given data are $v_0^0, v_1^0, z_0^0, z_0^1$. One has

$$z_1^0-\tilde{z}_1^0 = \frac{1}{1+\nu\bar{\alpha}g'(\nu)}[(z_0^0-\tilde{z}_0^0)(1-\nu\bar{\alpha}g'(\xi))+\nu\bar{\alpha}((v_1^0-\tilde{v}_1^0)+(v_0^0-\tilde{v}_0^0))], \tag{6.10}$$

$$v_0^1-\tilde{v}_0^1 = \theta(v_0^0-\tilde{v}_0^0)+\bar{\alpha}g'(\zeta)(z_0^1-\tilde{z}_0^1)+\bar{\alpha}\theta g'(\xi)(z_0^0-\tilde{z}_0^0), \tag{6.11}$$

$$v_1^1-\tilde{v}_1^1 = \theta(v_1^0-\tilde{v}_1^0)+\bar{\alpha}g'(\omega)(z_1^1-\tilde{z}_1^1)+\bar{\alpha}\theta g'(\nu)(z_1^0-\tilde{z}_1^0), \tag{6.12}$$

$$z_1^1-\tilde{z}_1^1 = \frac{1}{1+\nu\bar{\alpha}g'(\omega)}[(z_0^1-\tilde{z}_0^1)(1-\nu\bar{\alpha}g'(\zeta))+(v_0^0-\tilde{v}_0^0)\nu(1-\theta-\bar{\alpha}) \tag{6.13}$$
$$+(v_1^0-\tilde{v}_1^0)\nu(1-\theta-\bar{\alpha})+(z_1^0-\tilde{z}_1^0)+(z_0^0-\tilde{z}_0^0)\nu\bar{\alpha}(1-\theta)(g'(\xi)+g'(\nu))].$$

Substituting in the first hand side of (6.9) the equations (6.11)-(6.12)-(6.13), it follows

$$\begin{aligned}
&\nu[v_0^1 - \tilde{v}_0^1]_+ + \nu[v_1^1 - \tilde{v}_1^1]_+ + [z_1^0 - \tilde{z}_1^0]_+ + [z_1^1 - \tilde{z}_1^1]_+ \\
&\leq [z_1^0 - \tilde{z}_1^0]_+(1 + \nu\bar{\alpha}g'(\nu)) + [z_0^1 - \tilde{z}_0^1)]_+ + \nu(1-\bar{\alpha})[v_0^0 - \tilde{v}_0^0]_+ \\
&\quad + \nu(1-\bar{\alpha})[v_1^0 - \tilde{v}_1^0]_+ + \nu\bar{\alpha}g'(\xi)[z_0^0 - \tilde{z}_0^0]_+
\end{aligned}$$

and, using (6.10), we obtain (6.9). In order to ensure the positiveness of coefficients, we need to ask

$$\frac{\Delta x}{2\tau} g'(x) \leq 1, \quad \forall x \in f(I)$$

and

$$1 - \bar{\alpha} > 0 \Longleftrightarrow \frac{\Delta t}{2\tau} \leq 1.$$

On the other hand

$$\nu(1 - \theta - \bar{\alpha}) = \nu\left(1 - e^{-\frac{k}{\tau}} - \frac{k}{2\tau}\right) > 0, \tag{6.14}$$

with $\nu > 0$, $\theta \in (0,1)$. Under assumptions (6.7) and (6.8), it was possible to apply the rule in remark 6.3. □

6.3 Non-monotonicity

Since DSO scheme is of second order, it is not possible to obtain a monotone scheme. Indeed it is possible to show that for some initial and boundary data, the scheme does not respect the comparison property. As an example of this behaviour, consider the linearization of problem (3.9) with $g(z) = z$. Recalling in this case all nonnegative data are admissible, we take the following data:

$$\begin{aligned}
v_0(x) &= 0 \quad x \in [0,1], \\
z_1(t) &= \begin{cases} 0 & \text{if } t \in [0,1] - \{0.1\}, \\ 1 & \text{if } t = 0.1\ . \end{cases}
\end{aligned}$$

After a finite time, one can observe that the approximated solution produced by DSO scheme at some points becomes negative, while the solution obtained by DFO scheme is always non-negative. In particular, if we fix $\tau = 1$ and $\Delta x = \Delta t = 0.1$ the DSO solution assumes negative values for $t = 0.3$, as showed by Fig. 3.

7 Numerical experiments

Since we do not have the exact solution of the problem (3.9)-(3.11) for $\tau \neq 0$, we compute the order γ of the numerical method by the formula:

$$\gamma = \log_2\left(\frac{e(h_1)}{e(h_2)}\right). \tag{7.1}$$

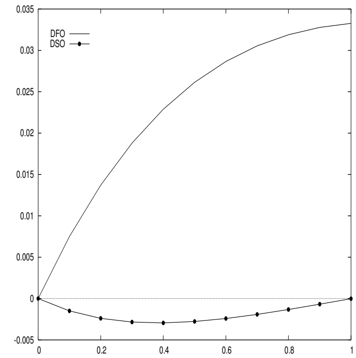

Figure 3: Approximated solutions to linearized problem: DFO depicted with line, DSO with linepoints, and x-zero axis with dotted line, $t = 0.3$, $\Delta x = \Delta t = 0.1$.

In particular, for DFO and DFO2 schemes the L^1-errors are computed by:

$$e(h_p) = h \sum_{j=0,\dots,Np} \left| w_j^M \left(\frac{h}{p}\right) - \frac{w_{2j-1}^M \left(\frac{h}{2p}\right) + w_{2j}^M \left(\frac{h}{2p}\right)}{2} \right|, \quad p = 1, 2, \tag{7.2}$$

while for DSO scheme we define the L^1-errors as:

$$e(h_p) = h \sum_{j=0,\dots,Np} \left| w_j^M \left(\frac{h}{p}\right) - w_{2j}^M \left(\frac{h}{2p}\right) \right|, \quad p = 1, 2 \tag{7.3}$$

where $w_j^M(h)$ denotes the numerical solution obtained with the space step equal to h, computed at $(x_j,\ t_M = T)$. We can compare the approximated solution to the problem (3.9)-(3.11) with the 'exact' solution of the Buckley-Leverett equation, corresponding to $\tau = 0$, computed by a numerical method as follows

$$E = h \sum_{j=0,\dots,N} \mid w_j^M(h) - u_j^M(h) \mid, \tag{7.4}$$

where $w_j^M(h)$ denotes the numerical solution obtained with the space step discretization equal to h, calculated in x_j at the final time $t_M = T$, while $u_j^M(h)$ is the 'exact' solution computed at (x_j, t_n). In our tests we consider the following non-convex flux function

$$f(u) = \frac{u^2}{u^2 + (1-u)^2}, \tag{7.5}$$

with the derivative

$$f'(u) = \frac{2u(1-u)}{[u^2 + (1-u)^2]^2}, \tag{7.6}$$

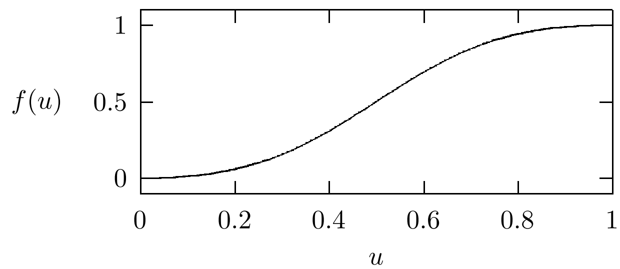

Figure 4: The flux function.

while the inverse of f is

$$g(z) = \frac{z - \sqrt{z - z^2}}{2z - 1} \qquad z \in [0,1] - 1/2. \tag{7.7}$$

7.1 Tests

Notice that when $\tau \to 0$ it is not always possible to respect the CFL condition contemporaneously for DFO and DSO, since DSO, which is not a relaxing scheme, needs $dx, dt \to 0$ and this fact will be indicated in the following tables by the symbol "-".

Test 1.

Here we take the following data:

$$\begin{cases} u(x,0) = u_0(x) = e^{-(x+0.1)} & x \in (0,1), \\ u(0,t) = u_1(t) = e^{-(t+0.1)} & t \in (0,1). \end{cases} \tag{7.8}$$

The admissibility condition for $x, t \in (0,1)$ reads

$$v_0(x) = e^{-(x+0.1)} \left[1 - \tau \frac{2e^{-(x+0.1)}(1 - e^{-(x+0.1)})}{[e^{-2(x+0.1)} + (1 - e^{-(x+0.1)})^2]^2} \right] \in (0,1), \tag{7.9}$$

$$u_1(t) = e^{-(t+0.1)} \in (0,1), \; t \in (0,1). \tag{7.10}$$

With some calculations, we find that the condition (7.10) is verified for $\tau \le 0.5$, thus it is violated when $\tau = 1$, as showed by the graph in Fig. 5. The condition (7.10) is satisfied by $u_1(t)$ since $e^{-(t+0.1)} \in [e^{-1.1}, e^{-0.1}] \subset [0,1]$.

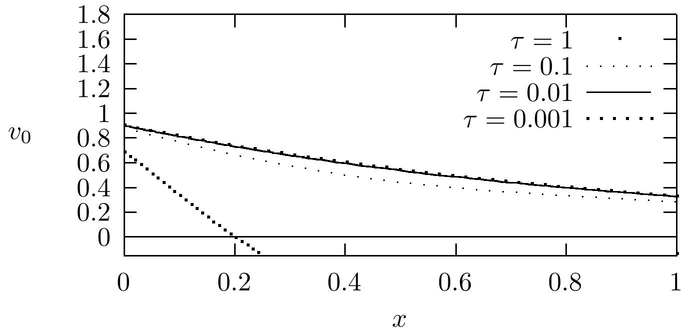

Figure 5: Graphs of function v_0 of Test 1 for τ varying.

$\tau = 0.1$						
	DFO		DFO2		DSO	
h	γ	L^1 Error	γ	L^1 Error	γ	L^1 Error
0.01	0.6	0.12878000E-03	0.8	0.22343431E-03	2.0	0.762310E-02
0.005	0.5	0.87851242E-04	1.0	0.11427204E-03	2.0	0.189550E-02
0.0025	0.8	0.50412662E-04	1.0	0.57303320E-04	2.0	0.472673E-03
0.00125	0.9	0.26875591E-04	1.0	0.28631491E-04	2.0	0.118005E-03
0.000625	0.9	0.13859800E-04	1.0	0.14302812E-04	2.0	0.294936E-04
0.0003125	1.0	0.70359236E-05	1.0	0.71471737E-05	2.0	0.737372E-05
0.00015625	1.0	0.35445293E-05	1.0	0.35724035E-05	2.0	0.184315E-05
0.000078125	1.0	0.17789140E-05	1.0	0.17858902E-05	2.0	0.460754E-06

Table T1.1: Convergence order γ and errors for the solution v computed by DFO, DFO2 and DSO, $\tau = 0.1, T = 1$.

$\tau = 0.01$						
	DFO		DFO2		DSO	
h	γ	L^1 Error	γ	L^1 Error	γ	L^1 Error
0.01	1.4	0.33081025E-02	2.0	0.96199172E-03	-	-
0.005	2.2	0.71372169E-03	1.2	0.42420943E-03	-	-
0.0025	3.9	0.48577606E-04	0.7	0.25861695E-03	2.0	0.631228E-01
0.00125	-0.3	0.58806154E-04	0.9	0.13695731E-03	2.0	0.158128E-01
0.000625	0.2	0.49848056E-04	1.0	0.69611105E-04	2.0	0.394114E-02
0.0003125	0.7	0.30020995E-04	1.0	0.34985373E-04	2.0	0.984088E-03
0.00015625	0.9	0.16280818E-04	1.0	0.17524594E-04	2.0	0.245912E-03
0.000078125	0.9	0.84573779E-05	1.0	0.87686394E-05	2.0	0.615861E-04

Table T1.2: Convergence order γ and errors for the solution v computed by DFO, DFO2 and DSO, $\tau = 0.01$, $T = 1$.

$\tau = 0.001$						
	DFO		DFO2		DSO	
h	γ	L^1 Error	γ	L^1 Error	γ	L^1 Error
0.01	0.4	0.86465068E-02	0.9	0.36504893E-02	-	-
0.005	0.8	0.48198721E-02	1.0	0.18205429E-02	-	-
0.0025	1.1	0.22723070E-02	1.2	0.79856440E-03	-	-
0.00125	1.6	0.75197835E-03	1.8	0.22513393E-03	-	-
0.000625	2.1	0.17008319E-03	1.9	0.60597812E-04	-	-
0.0003125	3.0	0.21811797E-04	0.6	0.39068259E-04	-	-
0.00015625	1.9	0.58893305E-05	0.8	0.21596323E-04	2.0	0.247399E-01
0.000078125	-0.3	0.71020382E-05	0.9	0.11144682E-04	2.0	0.618932E-02

TABLE T1.3: Convergence order γ and errors for the solution v computed by DFO, DFO2 and DSO, $\tau = 0.001$, $T = 1$.

In Fig. 6 we compare on the left DFO2 scheme with the 'exact' solution of the Buckley-Leverett equation ($\tau = 0$) computed by Godunov scheme for data (7.8), while on the right we represent the solutions obtained by DFO and DFO2. In Fig. 7 we compare the solutions obtained by DFO2 and DSO.

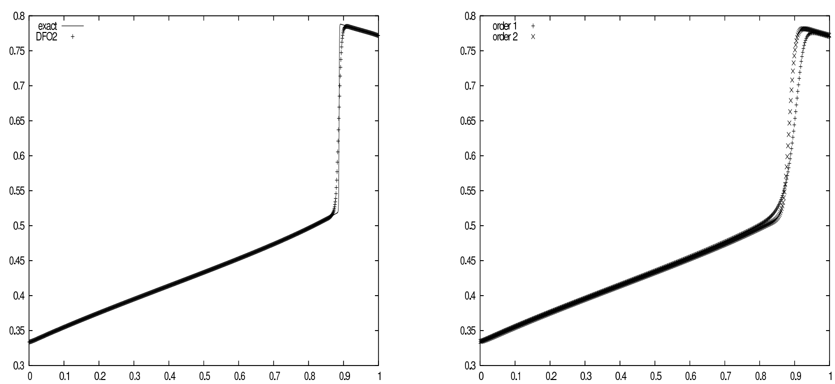

Figure 6: On the left: comparison between the approximated solution obtained by DFO2 and the 'exact' solution, $\tau = 0$, $\Delta x = 0.001, \Delta t = 0.005, T = 1$; on the right: approximated solution v computed by DFO and DFO2, $\tau = 0$, $\Delta x = 0.002, \Delta t = 0.01, T = 1$.

As for DSO scheme a relaxed scheme does not exist, in the table BL1 we present a study of the convergence of DSO for $\tau \to 0$ towards the exact solution of the Buckley-Leverett equation ($\tau = 0$).

$\tau = 0.1$	
	DSO
h	L^1 Error
0.0025	6.85030E-02
0.00125	6.84052E-02
0.000625	6.83540E-02
0.0003125	6.83295E-02
0.00015625	6.83277E-02
0.000078125	6.83218E-02
$\tau = 0.01$	
	DSO
h	L^1 Error
0.0025	1.86688E-02
0.00125	1.84146E-02
0.000625	1.83674E-02
0.0003125	1.83588E-02
0.00015625	1.83729E-02
0.000078125	1.83759E-02
$\tau = 0.001$	
	DSO
h	L^1 Error
0.0025	8.39278E-02
0.00125	2.08374E-02
0.000625	5.75331E-03
0.0003125	2.83554E-03
0.00015625	2.21510E-03
0.000078125	2.08652E-03

TABLE BL1: Difference between the 'exact' solution and the approximated solution computed by DSO for $T = 1$.

From the orders and errors tables DFO is first order accurate, while DFO2 is better than first order, although we cannot ensure monotonicity for it. DSO scheme is of second order. They all converge to the exact solution when τ decreases to zero, thus they work in practice better than the consistency analysis shows. In particular, when $\tau \to 0$ both DFO and DFO2 behave better than DSO.

Test 2. Let us consider the initial and boundary conditions:

$$\begin{cases} u(x,0) = u_0(x) = \frac{9}{10} - \frac{1}{2}x & x \in (0,1), \\ u(0,t) = u_1(t) = \frac{9}{10} - \frac{1}{2}t & t \in (0,1). \end{cases} \tag{7.11}$$

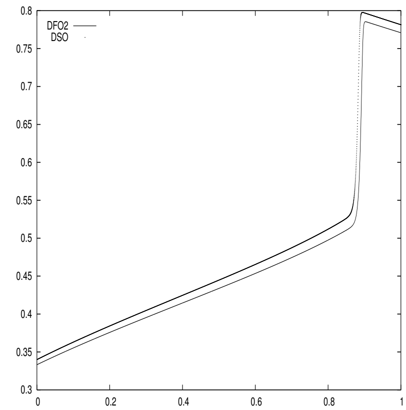

Figure 7: Approximated solution v computed by DFO2 and DSO, $\tau = 0.001$, $\Delta x = 0.0002, \Delta t = 0.0005$, $T = 1$.

The admissibility condition for $x, t \in (0, 1)$ reads

$$v_0(x) = \left(\frac{9}{10} - \frac{1}{2}x\right)\left[1 - \tau\frac{\left(\frac{1}{10} + \frac{1}{2}x\right)}{\left(\frac{x^2}{2} - \frac{4}{5}x + \frac{41}{50}\right)^2}\right] \in (0, 1), \qquad (7.12)$$

$$u_1(t) = \frac{9}{10} - \frac{1}{2}t \in (0, 1),\ t \in (0, 1). \qquad (7.13)$$

With some calculations, we find that the condition (7.12) is verified for $\tau \leq 0.45$, thus when $\tau = 1$ it is violated, as showed by Figure 8. The condition (7.13) is satisfied, since $\frac{9}{10} - \frac{1}{2}t \in \left(\frac{4}{10}, \frac{9}{10}\right) \subset [0, 1]$. In the following tables we compare the behaviour of the two schemes for τ varying in $[0, 1]$.

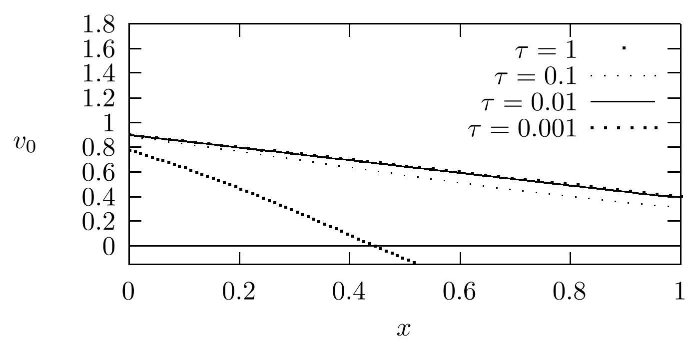

Figure 8: Graphs of function v_0 of Test 2 for τ varying.

$\tau = 0.1$				
	DFO2		DSO	
h	γ	L^1 Error	γ	L^1 Error
0.01	0.8	0.31343823E-03	2.0	0.844026E-02
0.005	0.9	0.16314647E-03	2.0	0.209845E-02
0.0025	1.0	0.82548649E-04	2.0	0.523050E-03
0.00125	1.0	0.41436804E-04	2.0	0.130591E-03
0.000625	1.0	0.20748697E-04	2.0	0.326289E-04
0.0003125	1.0	0.10380622E-04	2.0	0.815479E-05
0.00015625	1.0	0.51917159E-05	2.0	0.203840E-05
0.000078125	1.0	0.25961887E-05	2.0	0.509564E-06

TABLE T2.1: Errors for the solution of computed by DFO2 scheme and DSO scheme, $\tau = 0.1$, $T = 1$.

$\tau = 0.01$				
	DFO2		DSO	
h	γ	L^1 Error	γ	L^1 Error
0.01	1.2	0.85925126E-03	-	-
0.005	1.1	0.35131635E-03	-	-
0.0025	0.6	0.22452432E-03	2.0	0.601647E-01
0.00125	0.9	0.12212060E-03	2.0	0.148328E-01
0.000625	0.9	0.62846593E-04	2.0	0.369472E-02
0.0003125	1.0	0.31771608E-04	2.0	0.922347E-03
0.00015625	1.0	0.15959929E-04	2.0	0.230490E-03
0.000078125	1.0	0.79968398E-05	2.0	0.576149E-04

TABLE T2.2: Errors for the solution of problem (3.9)-(3.10)-(3.11) computed by DFO2 scheme and DSO scheme, $\tau = 0.01$, $T = 1$.

$\tau = 0.001$				
	DFO2		DSO	
h	γ	L^1 Error	γ	L^1 Error
0.01	1.1	0.26771951E-02	-	-
0.005	1.0	0.13210746E-02	-	-
0.0025	1.2	0.58466618E-03	-	-
0.00125	1.7	0.17434895E-03	-	-
0.000625	1.8	0.48961305E-04	-	-
0.0003125	0.7	0.30324544E-04	-	-
0.00015625	0.8	0.16806233E-04	2.0	0.232882E-01
0.000078125	0.9	0.87026361E-05	2.0	0.578744E-02

TABLE T2.3: Errors for the solution of problem (3.9)-(3.10)-(3.11) computed by DFO2 scheme and DSO scheme, $\tau = 0.001$, $T = 1$.

In Fig. 9 we compare DFO2 scheme with the 'exact' solution of the Buckley-Leverett equation ($\tau = 0$) computed by Godunov scheme for data (7.11) with a very small space step. Fig. 10 represents a comparison between DFO2 and DSO.

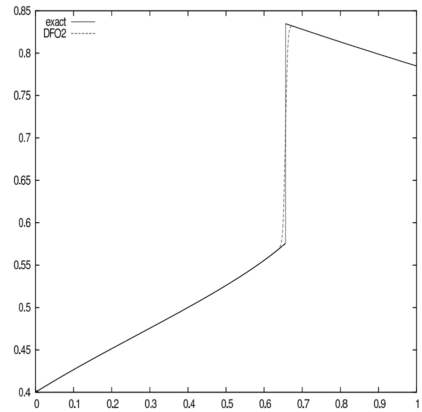

Figure 9: Comparison between DFO2 approximated solution and the 'exact' solution, $\tau = 0$, $\Delta x = 0.000625, \Delta t = 0.005, T = 1$.

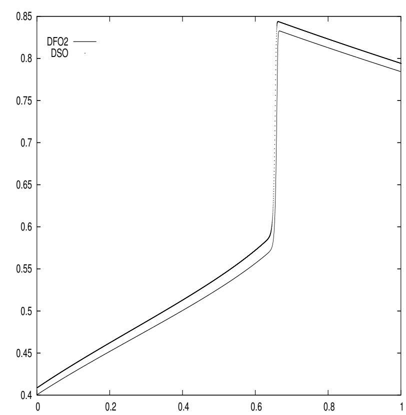

Figure 10: Approximated solution v computed by DFO2 and DSO schemes, $\tau = 0.001$, $\Delta x = 0.0002, \Delta t = 0.0005$, $T = 1$.

As for DSO scheme a relaxed scheme does not exist, in the table BL2 we study the convergence of DSO for $\tau \to 0$ towards the exact solution of the Buckley-Leverett equation ($\tau = 0$).

$\tau = 0.1$	
	DSO
h	L^1 Error
0.0025	8.50799E-02
0.00125	8.51405E-02
0.000625	8.51350E-02
0.0003125	8.51322E-02
0.00015625	8.51308E-02
0.000078125	8.51300E-02
$\tau = 0.01$	
	DSO
h	L^1 Error
0.0025	1.34838E-02
0.00125	1.35028E-02
0.000625	1.35060E-02
0.0003125	1.35216E-02
0.00015625	1.35307E-02
0.000078125	1.35356E-02
$\tau = 0.001$	
	DSO
h	L^1 Error
0.0025	7.99631E-02
0.00125	1.95989E-02
0.000625	5.35356E-03
0.0003125	2.25047E-03
0.00015625	1.54187E-03
0.000078125	1.47532E-03

TABLE BL2: Difference between the 'exact' solution and the approximated solution computed by DSO for $T = 1$.

Both DFO2 and DSO converge to the solution also for very small values of τ. In particular, when $\tau \to 0$ DFO2 behaves better than DSO.

7.2 Counterexample

Here we refer to the Example 3.1 presented in Section 3. Setting

$$\psi(0) = 0.9, \quad \psi'(0) = -1,$$

and fixing $\alpha = \tau = 1$, we get $C = -0.09$. Since $C < 0$ and $\psi(0) > 0$ there exists

$$\bar{\xi} = \int_0^{\psi(0)} \frac{2\alpha\tau s}{s^2 + \alpha s - C} ds > 0$$

such that $\psi(\bar{\xi}) = 0$ and we can approximate this value by Simpson's quadrature rule. Then we can solve the following equation

$$\psi' = \frac{C - \psi^2 - \alpha\psi}{2\alpha\tau\psi}$$

using the numerical method of Runge-Kutta of fourth order. Taking in (3.9)-(3.11) the flux function $f(u) = u^2$, and the initial and boundary conditions below:

$$v_0(x) = \{\psi(x) + \tau[f(\psi(x))]'\}(x), \quad x \in [0, 0.9] \tag{7.14}$$
$$z_1(t) = f(\psi(t)) \quad t \in [0, 0.9], \tag{7.15}$$

we obtain a solution which does not exist in corrispondence of the line $\bar{\xi} = t + x$. From the comparison between the "exact" solution and the approximated solution we can observe that DFO scheme gives a good reconstruction of the solution up to $\bar{\xi}$, since the exact solution cannot be extended beyond that point, see Fig. 11.

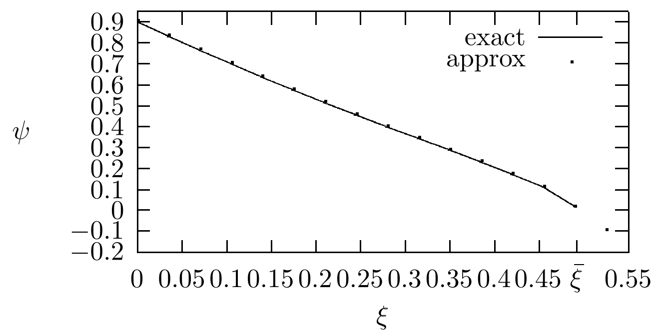

Figure 11: Comparison between the 'exact' function $\psi(2x)$ (computed by Runge-Kutta method) and the approximated function $u(x, x) = \psi(2x)$ computed by DFO scheme, $\Delta x = \Delta t = 0.035$.

8 Conclusions

In this paper we presented some numerical methods to deal with the approximation of these non-equilibrium problems. Such schemes were inspired by the diagonal problem and they follow the analytical structure of solutions. For DFO we were able to prove monotonicity. For all the methods we studied the behaviour when $\tau \to 0$ and we found that in this case both DFO and DFO2 perform better than DSO, since it is not a relaxing scheme.

References

[1] G. I. Barenblatt, Flow of two immiscible fluids in homogeneus porous medium (in Russian), *Izv. AN SSSR, Mekh. Zhidk. i Gaza* **5** (1971), 144-151.

[2] G. I. Barenblatt, V. M. Entov and V. M. Ryzhik., Theory of fluid flows through natural rocks, Kluwer Academic Publishers, Dordrecht (1990).

[3] G. I. Barenblatt, J. Garcia-Azorero, A. De Pablo and J. L. Vazquez, Mathematical Model of the Non-Equilibrium Water-Oil Displacement in Porous Strata, *Applicable Analysis* **65** (1997), 19-45.

[4] G. I. Barenblatt, A. P. Vinnitchenko, Non equilibrium seepage of immiscible fluids (in Russian), *Uspekhi Mekhaniki* **3** (1980), 35-50.

[5] S. E. Buckley, M. C. Leverett, Mechanism of fluid displacement in sands, *Trans. AIME* **142** (1942), 107-116.

[6] M. Muskat and M. Meres, The flow of heterogeneous fluids, *Physics*, **7** (1936), 346–363.

[7] R. Natalini, A. Tesei, On the Barenblatt model for non-equilibrium two phase flow in porous media, *Arch. Rational Mech. Anal.* **150** (1999), 349-367.

[8] D. W. Peaceman, Fundamentals of numerical reservoir simulation, *Developments in petroleum science* **6**, ed. Elsevier, Amsterdam (1977).