Dynamic capillarity in porous media : mathematical analysis

DOI:
10.6100/IR721598

Document status and date:
Published: 01/01/2012

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Dynamic Capillarity in Porous Media - Mathematical Analysis

A catalogue record is available from the Eindhoven University of Technology Library
ISBN: 978-90-386-3069-4
NUR 919
Subject headings: partial differential equations; travelling wave solutions; weak solutions
2000 Mathematics Subject Classification: 35D30; 35K70; 65M12; 65M15

Dynamic Capillarity in Porous Media - Mathematical Analysis

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
rector magnificus, prof.dr.ir. C.J. van Duijn, voor een
commissie aangewezen door het College
voor Promoties in het openbaar te verdedigen
op dinsdag 31 januari 2012 om 16.00 uur

door

Yabin Fan

geboren te Henan, China

Dit proefschrift is goedgekeurd door de promotor:

prof.dr.ir. C.J. van Duijn

Copromotor:
dr. I.S. Pop

Contents

Chapter 1

Introduction

This thesis is devoted to the analysis of the pseudo-parabolic equation

$$u_t + \nabla \cdot \mathbf{F}(u, x, t) = \nabla \cdot \Big(H_1(u)\nabla u\Big) + \tau\nabla \cdot \Big(H_2(u)\nabla u_t\Big). \tag{1.0.1}$$

The particularity of this equation is in the higher order terms, where the functions H_1, H_2 may vanish at certain values of u. Note that for such equations, smooth solutions do not exist in general. In this thesis, we focus on two kinds of solutions: the travelling wave solution and the weak solution. Before explaining the results in more details, we describe the physical model that motivated this work.

1.1 Why non-equilibrium models?

Already for decades there has been considerable interest in understanding and predicting flow processes (especially of two immiscible fluids) in porous media. Such processes are encountered in many engineering applications, for instance the transport of dissolved contaminants in the subsurface system, enhanced oil recovery, paper production and so on. Generally, these are multi-phase flow systems, which include many unknowns in the modelling process.

Multi-phase flow processes are investigated theoretically and experimentally in [1], [5], [6], [8], [9], [10]. Commonly, these processes are considered under equilibrium conditions. This means that quantities such as the capillary pressure and the saturation are related to each other by formulae that are determined by experiments carried out under equilibrium conditions. However, there is also experimental evidence of transport processes where the equilibrium conditions are violated and non-equilibrium effects need to be considered. For instance in the work of DiCarlo [5], which is shown in the left picture of Fig. 1.1. There non-monotone saturation profiles are obtained during infiltration. Standard equilibrium models for porous media flow rule out such profiles. In this thesis, especially in Chapter 2, we explain the experimental results by means of a mathematical analysis, which is based on a non-equilibrium effect in the capillary pressure. A numerical example in this sense is presented in the right picture of Fig. 1.1.

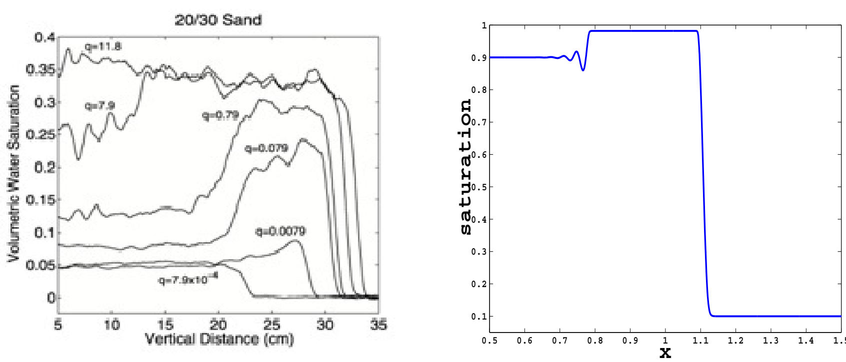


FIGURE 1.1: SATURATION OVERSHOOT PROFILES: EXPERIMENTAL RESULTS IN [5] (LEFT), A NUMERICAL EXAMPLE (RIGHT)

In the following section, we will elaborate the corresponding two-phase model.

1.2 The two-phase flow model

When modelling two-phase in porous media, different spatial scales are involved. The smallest scale is the pore scale, where one distinguishes between the fluid and the solid matrix. The next scale is the so-called representative elementary volume (REV) scale, where an averaging process has been carried out. Finally, there is the macro-scale which is encountered in the laboratory or beyond. The model that describes two-phase flow on the REV scale consists of the mass balance equations, and Darcy law for the phases, and a constitutive relationship between the phases depending on the porous medium. Before giving the model equations, we start by listing some important properties of the porous medium and the phases. For more details, we refer to [7], [12].

1.2.1 The porous medium

There is a large body of literature on multi-phase flow in porous media. Some of the main references are collected in the reference list of this chapter. Here we discuss some of the main fractures on the scale of a REV.

a. The capillary pressure

With P_n, P_w ($[kg \cdot m^{-1} \cdot s^{-2}]$) denoting the non-wetting and the wetting phase pressure, the capillary pressure is defined as

$$P_n - P_w = P_c(S_w). \tag{1.2.1}$$

Brooks and Corey developed in [2] an empirical relationship between the effective water saturation S_e and the capillary pressure P_c. They introduced

$$P_c(S_e) = P_d S_e^{-\frac{1}{\lambda}}, \text{ with } S_e = \frac{S_w - S_{wr}}{1 - S_{wr} - S_{nr}}, \tag{1.2.2}$$

where S_w, S_{wr}, S_{nr} are the wetting phase saturation, the irreducible wetting phase saturation and the residual non-wetting phase saturation, respectively. The parameter λ is related to the pore size distribution. An alternative capillary pressure model was proposed by Van Genuchten in [13] and reads

$$P_c(S_e) = \frac{1}{\alpha}\left(S_e^{-\frac{1}{m}} - 1\right)^{\frac{1}{n}}. \tag{1.2.3}$$

However, experimental evidence shows that expression (1.2.1) is only valid under equilibrium conditions when the fluids are at rest. Alternatively, Hassanizadeh and Gray suggested in [8] a relation including an extra term to account for non-equilibrium effects. They proposed

$$P_n - P_w = P_c(S_w) - \tau \frac{\partial S_w}{\partial t}, \tag{1.2.4}$$

where τ ($[kg \cdot m^{-1} \cdot s^{-1}]$) is a damping factor.

b. The relative permeability-saturation relationship

The relative permeability $k_{r\alpha}$ ([-]) is defined as the ratio of the effective permeability K_α ($[m^2]$) to the intrinsic permeability K ($[m^2]$),

$$k_{r\alpha} = \frac{K_\alpha}{K}, \quad \alpha = n, w. \tag{1.2.5}$$

Generally, $k_{r\alpha}$ is assumed to be a function of S_α. For example, in the model of Burdine (see [4]) one has

$$k_{rw} = S_e^{\frac{2+3\lambda}{\lambda}}, \tag{1.2.6}$$

for wetting phase and

$$k_{rn} = (1 - S_e)^2 \left(1 - S_e^{\frac{2+\lambda}{\lambda}}\right), \tag{1.2.7}$$

for non-wetting phase. Alternatively, Van Genuchten (see [13]) suggested

$$k_{rw} = S_e^l \left(1 - \left(1 - S_e^{\frac{1}{m}}\right)^m\right)^2, \tag{1.2.8}$$

for wetting phase and

$$k_{nw} = (1 - S_e)^l \left(1 - S_e^{\frac{1}{m}}\right)^{2m}, \tag{1.2.9}$$

for non-wetting phase. Here l is a parameter related to the tortuosity of the porous medium.

1.2.2 The model equations

Combining the equations describing mass conservation and Darcy law one finds for the two phases (w=wetting, n=non-wetting phases) the following expressions (see Bear [3], Helimg [7])

$$\phi\frac{\partial S_w}{\partial t} - \nabla\cdot\left(\frac{Kk_{rw}}{\mu_w}\nabla\Phi_w\right) = 0, \tag{1.2.10}$$

$$\phi\frac{\partial S_n}{\partial t} - \nabla\cdot\left(\frac{Kk_{rn}}{\mu_n}\nabla\Phi_n\right) = 0. \tag{1.2.11}$$

Here ϕ, S_α, K, $k_{r\alpha}$, μ_α, Φ_α ($\alpha = w, n$) denote the porosity [-], the saturation [-], the intrinsic permeability [m^2], the phase relative permeability [-], the dynamic viscosity [$kg \cdot m^{-1} \cdot s^{-1}$], the phase potential [$kg \cdot m^{-1} \cdot s^{-2}$], respectively. Additionally, Φ_w and Φ_n are given by

$$\Phi_w = P_w + \rho_w gz, \quad \Phi_n = P_n + \rho_n gz, \tag{1.2.12}$$

where P_α, ρ_α, g ($\alpha = w, n$) denote pressure [$kg \cdot m^{-1} \cdot s^{-2}$], density [$kg \cdot m^{-3}$] and gravity [$m \cdot s^{-2}$], respectively. Further z is the vertical coordinator, being position in the upward direction.

Furthermore, we assume that only two fluid phases are present in the medium, implying

$$S_w + S_n = 1. \tag{1.2.13}$$

Following Hassanizadeh and Gray [8] we use for the difference of the phase pressures

$$P_n - P_w = P_c(S_w) - \tau\frac{\partial S_w}{\partial t}. \tag{1.2.14}$$

Adding (1.2.10) and (1.2.11), using (1.2.13), one finds

$$\nabla\cdot Q = \nabla\cdot\left(\frac{Kk_{rw}}{\mu_w}\nabla\Phi_w + \frac{Kk_{rn}}{\mu_n}\nabla\Phi_n\right) = 0, \tag{1.2.15}$$

where $Q = \frac{Kk_{rw}}{\mu_w}\nabla\Phi_w + \frac{Kk_{rn}}{\mu_n}\nabla\Phi_n$ ($[m \cdot s^{-1}]$) denotes the total flow. In one spatial dimension, equation (1.2.15) implies $Q = q$ (constant w.r.t spatial parameter). Define

$$f(S_w) := \frac{\frac{k_{rw}}{\mu_w}}{\frac{k_{rw}}{\mu_w} + \frac{k_{rn}}{\mu_n}}, \quad \text{and} \quad H(S_w) := \frac{\frac{k_{rn}}{\mu_n}\frac{k_{rw}}{\mu_w}}{\frac{k_{rw}}{\mu_w} + \frac{k_{rn}}{\mu_n}}. \tag{1.2.16}$$

We have for one dimensional (vertical) flow

$$\begin{aligned} &\phi\frac{\partial S_w}{\partial t} + q\frac{\partial f(S_w)}{\partial z} + (\rho_n - \rho_w)g\frac{\partial H(S_w)}{\partial z} \\ &= -\frac{\partial}{\partial z}\left(H(S_w)\frac{\partial P_c(S_w)}{\partial z}\right) + \tau\frac{\partial}{\partial z}\left(H(S_w)\frac{\partial}{\partial z}\left(\frac{\partial S_w}{\partial t}\right)\right). \end{aligned} \tag{1.2.17}$$

We put the model in a dimensionless form by introducing the reference quantities T, L, P_d, q_c (characteristic velocity) and redefine

$$t := \frac{t}{T}, \quad z := \frac{z}{L}, \quad P_c := \frac{P_c}{P_d}, \quad q := \frac{q}{q_c}, \quad H := \mu_w H, \quad u := S_w.$$

Then

$$\begin{aligned} &\frac{\partial u}{\partial t} + \frac{q_c T}{\phi L}q\frac{\partial f(u)}{\partial z} + \frac{(\rho_n - \rho_w)gKT}{L\phi\mu_w}\frac{\partial H(u)}{\partial z} \\ &= -\frac{P_d T k}{L^2\phi\mu_w}\frac{\partial}{\partial z}\left(H(u)\frac{\partial P_c(u)}{\partial z}\right) + \frac{\tau K}{L^2\phi\mu_w}\frac{\partial}{\partial z}\left(H(u)\frac{\partial}{\partial z}\left(\frac{\partial u}{\partial t}\right)\right). \end{aligned} \tag{1.2.18}$$

Four dimensionless numbers can be identified:

$$\begin{aligned} \text{the Péclet number: } \mathrm{Pe} &= \tfrac{q_c T}{\phi L}, \\ \text{the gravity number: } \mathrm{Gr} &= \tfrac{(\rho_n - \rho_w)gKT}{L\mu_w\phi}, \\ \text{the capillary number: } \mathrm{Ca} &= \tfrac{P_d T K}{L^2\mu_w\phi}, \\ \text{the dynamic number: } \mathrm{Dy} &= \tfrac{\tau K}{L^2\mu_w\phi}. \end{aligned}$$

Here we use the notations introduced in [10], [11]. After defining these dimensionless numbers, we have

$$\frac{\partial u}{\partial t} + \mathrm{Pe}q\frac{\partial f(u)}{\partial z} + \mathrm{Gr}\frac{\partial H(u)}{\partial z} = -\mathrm{Ca}\frac{\partial}{\partial z}\left(H(u)\frac{\partial P_c(u)}{\partial z}\right) + \mathrm{Dy}\frac{\partial}{\partial z}\left(H(u)\frac{\partial}{\partial z}\left(\frac{\partial u}{\partial t}\right)\right). \tag{1.2.19}$$

Then equation (1.0.1) can be seen as a general form of equation (1.2.19) with $P_c(u) = u$. Note the 'τ' (to avoid confusion we denote τ_{new} in the following) in the equation (1.0.1) can be formulated by

$$\tau_{new} := \frac{Dy}{Ca^2} = \frac{\tau L^2 \mu_w \phi}{P_d^2 T^2 K} = \frac{\tau \mu_w q_c^2}{\mathrm{Pe}^2 P_d^2 K \phi}.$$

In this way, if we denote Ca as ε, we make a connection between equation (1.2.19) and equation (2.0.4) in Chapter 2. Further, if we choose Pe=1 and fix the quantities τ, μ_w, P_d, T, K, then τ_{new} is increasing as q_c is increasing, for more qualitative analysis, we refer to [10], [11].

1.3 Overview and main results of this thesis

In this thesis we mainly investigate travelling wave solutions and weak solutions of equation (1.0.1). We prove the existence of travelling wave solutions. These solutions depend on the left and right states as well as on the parameters of the model. Next we define a weak solution, for which the existence and uniqueness are proved under certain assumptions. Inspired by the definition of the capillary pressure, we also transform the equation (1.0.1) into a system by introducing an extra unknown p (the capillary pressure), and prove the equivalence between different formulations for the continuous and the semi-discrete cases. At the end, we give some numerical results using the numerical scheme based on the different formulations. In detail, this thesis is organized in the following way.

In Chapter 2, we investigate (1.0.1) in one spatial dimension with $H_1(u) = H_2(u) = H(u)$. The particularity of this equation is that H becomes 0 for $u \leq 0$ or $u \geq 1$, which makes the model degenerate. In this chapter, we study the existence of the travelling wave solutions, based on several parameters: the left state u_ℓ, the right state u_r and the damping coefficient τ. Two extra parameters α and β are defined, the existence of the travelling wave solutions is proved for the non-degenerate case. Further, the travelling wave solution is unique and decreasing. Inspired by the monotonicity of the solution u, we introduce $w = -u'$ and obtain one ordinary differential equation. The existence of travelling wave solutions for degenerate case is proved by a shooting method. Moreover, we show that there exists a threshold value of τ_{max}, such that whenever τ is beyond τ_{max}

classical travelling waves do not exist anymore. Therefore we introduce the definition of sharp travelling wave solutions, including a discontinuity point in the function u'. Figure 1.2 gives an example of classical (smooth) and sharp travelling wave solutions. In the last part of Chapter 2, we use a semi-implicit Euler finite volume method to solve the equation (1.0.1). The smooth and sharp travelling wave solutions are confirmed by the numerical computations.

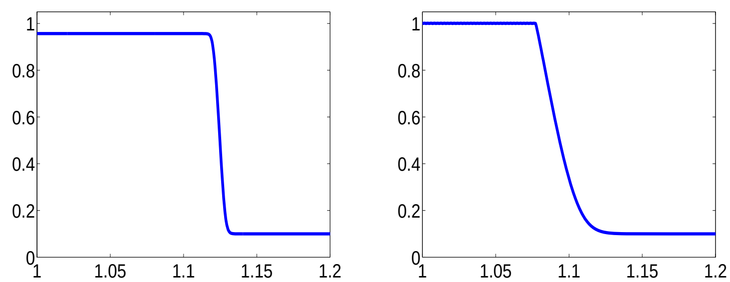


FIGURE 1.2: AN EXAMPLE OF TWO KINDS OF TRAVELLING WAVE SOLUTIONS: SMOOTH (LEFT), SHARP (RIGHT)

In Chapter 3, we deal equation (1.0.1) with $H_1(u) = H(u)$, $H_2 = 1$ and focus on weak solutions. In this case, no assumption on the sign of H is needed, enlarging the context that is usually encountered in physical models, where $H \geq 0$. The existence of weak solutions is proved by the method of Rothe, based on the Euler implicit time stepping. Further, we prove the uniqueness of weak solutions by constructing suitable Green functions as test functions. The error estimate for the time discretization is also derived. Then two numerical examples (one with positive H, one with negative H) are given to verify the theoretical results.

In Chapter 4, we consider the degenerate case $H_1(u) = H_2(u) = H(u)$, where H vanishes when u is outside $(0, 1)$. As in Chapter 3, we focus on weak solutions. To prove their existence, we adapt a regularization approach based on perturbating H by $H_\delta = H + \delta$ with $\delta > 0$, and prove the existence of weak solutions to the regularized problem firstly. Then we pass $\delta \searrow 0$, which gives the the uniform bounds for u_δ in the space H^1 as well as for $H_\delta(u_\delta)\partial_t \nabla u_\delta$. However, this is not sufficient for identifying the equation that is satisfied by the limit u, as one needs the weak convergence of $\partial_t \nabla u_\delta$. To overcome this difficulty, we decompose the spacial variable x and use the Div-Curl Lemma and the Vitali Convergence Theorem to prove the existence of weak solutions for the original problem. Finally, we prove that the solution u is essentially bounded by 0 and 1.

In Chapter 5, we continue the investigation of (1.0.1) with $H_1(u) = H_2(u) = H(u)$. However, in this chapter we assume the function H is essentially bounded by m and M ($m \leq H \leq M$). We introduce an extra unknown 'pressure p' and transform the equation (1.0.1) into different systems. The equivalence of different formulations for both continuous and semidiscrete cases is proved. Finally in Chapter 6, a fully discrete scheme to solve the equation (1.0.1) is described inspiring from the formulations introduced in Chapter 5. Several numerical results are also given and compared with the numerical results in Chapter 2.

References

[1] S. Bottero, *Advances in the Theory of Capillarity in Porous Media*, Ph.D. thesis, Utrecht University, 2009.

[2] R.H. Brooks and A.T. Corey, *Hydraulic properties of porous media*, in hydrology Papers, volume 3. Colorado State Unviersity, Fort Collins, 1964.

[3] J. Bear, *Dynamics of Fluids in Porous Media*, Elsevier, New York, 1972.

[4] N. Burdine, *Relative permeability calculations from pore size distribution data*, Transaction of the American Institute of Mining and Metallurgical Engineers, **198**(1953), 71-78.

[5] D.A. DiCarlo, *Experimental measurements of saturation overshoot on infiltration*, Water Resour. Res. **40**(2004), W04215.1-W04215.

[6] F. Kalaydijan, *Dynamic capillary pressure curve for water/oil displacement in porous media, theory vs. experiments*, Proc. SPE Conference (1992), Washington DC. SPE, Brookfield, CT.

[7] R. Helmig, *Multiphase Flow and Transport Processes in the Subsurface: a Contribution to the Modeling of Hydrosystems*, Springer, Berlin, 1997.

[8] S.M. Hassanizadeh and W.G. Gray, *Thermodynamic basis of capillary pressure in porous media*, Water Resour. Res. **29** (1993), 3389-3405.

[9] F. Kalaydijan and C.-M. Marle, *Microscopic flow and generalized Darcy′s equations*, The mathematics of oil recovery (Cambridge, 1989), 495-514, Inst. Math. Appl. Conf. Ser. New Ser., 31, Oxford Univ. Press, New York, 1992.

[10] S. Manthey, *Two-phase Flow Processes with Dynamic Effects in Porous Media-Parameter Estimation and Simulation*, Ph.D. thesis, University of Stuttgart, 2006.

[11] S. Manthey, S.M. Hassanizadeh, R. Helmig, R. Hilfer, *Dimensional analysis of two-phase flow including a rate-dependent capillary pressure-saturation relationship*, Advances in Water Resources **31**(2008), 1137-1150.

[12] G.F. Pinder and W.G. Gray, *Essentials of Multiphase Flow in Porous Media*, John Wiley & Sons, Inc., Hoboken, New Jersey, 2008.

[13] M.T.Van Genuchten, *A closed-form equation for predicting the hydraulic conductivity of unsaturated soil*, Soil Sci. Soc. Am. J., **44**(1980), 892-898.

Chapter 2

Travelling wave solutions

In this chapter, we discuss non-classical solutions of the *Buckley-Leverett (BL) equation* from the perspective of a regularization derived from two-phase flow through porous media. In particular, we study an equation arising in a model for oil recovery by water-drive in a one-dimensional horizontal flow:

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = \varepsilon \frac{\partial}{\partial x}\left(H(u)\frac{\partial p_c}{\partial x}\right). \tag{2.0.1}$$

Here u stands for water saturation, which is expected to take values in the interval [0,1]; when $u = 0$ there is no water, and when $u = 1$ there is no oil. The function p_c stands for the capillary pressure expressing the difference between phase pressures:

$$p_c = p_o - p_w, \tag{2.0.2}$$

where p_o and p_w are the oil pressure and the water pressure respectively. Typically, p_c is determined experimentally as a function of u. However, in this chapter we consider

This chapter is a collaborative work with C.J. van Duijn, L.A. Peletier, I.S. Pop the dynamic extension suggested by Hassanizadeh and Gray in [16], where the capillary pressure has a relaxation term:

$$p_c = p_c^{static} + p_c^{dynamic} = p_c^{static} + \varepsilon\tau L(u)u_t, \tag{2.0.3}$$

where p_c^{static} and $p_c^{dynamic}$ are the static, respectively the dynamic components in the capillary pressure. For simplicity we put $p_c^{static}(u) = u$ and $L(u) = 1$. In this expression, τ is the relaxation parameter and ε is a small positive constant. With these assumptions about the pressures, equation (2.0.1) becomes

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = \varepsilon\frac{\partial}{\partial x}\left\{H(u)\left(\frac{\partial u}{\partial x} + \varepsilon\tau\frac{\partial^2 u}{\partial x\partial t}\right)\right\}. \tag{2.0.4}$$

The functions, f and H in equations (2.0.1) and (2.0.4) are the water fractional flow function and the capillary induced diffusion function. They are given by

$$f(u) = \frac{\lambda_w(u)}{\lambda_w(u) + M\lambda_o(u)}, \qquad \text{and} \qquad H(u) = \frac{\lambda_w(u)\lambda_o(u)}{\lambda_w(u) + M\lambda_o(u)}, \tag{2.0.5}$$

where M is the water/oil viscosity ratio, while λ_w and λ_o are the normalized relative permeabilities. Commonly accepted in the engineering literature is the *Brooks-Corey model* in which λ_w and λ_o are defined by:

$$\lambda_w(u) = u^{p+1}, \qquad \text{and} \qquad \lambda_o(u) = (1-u)^{q+1}, \qquad p > 0,\ q > 0. \tag{2.0.6}$$

In this chapter, we restrict the analysis to the functions of the *Brooks-Corey model* in order to avoid non-essential technical difficulties. The functions H and f are then given by

$$H(u) = \frac{u^{p+1}(1-u)^{q+1}}{u^{p+1} + M(1-u)^{q+1}}, \qquad \text{and} \qquad f(u) = \frac{u^{p+1}}{u^{p+1} + M(1-u)^{q+1}}. \tag{2.0.7}$$

Their graphs are shown in Figure 2.1. Note that

$$\begin{aligned} &H(u) > 0 \quad \text{and} \quad f(u) > 0 \qquad \text{for} \quad 0 < u < 1, \quad \text{and} \\ &\qquad\qquad H(0) = 0, \quad H(1) = 0 \quad \text{and} \quad f(0) = 0. \end{aligned}$$

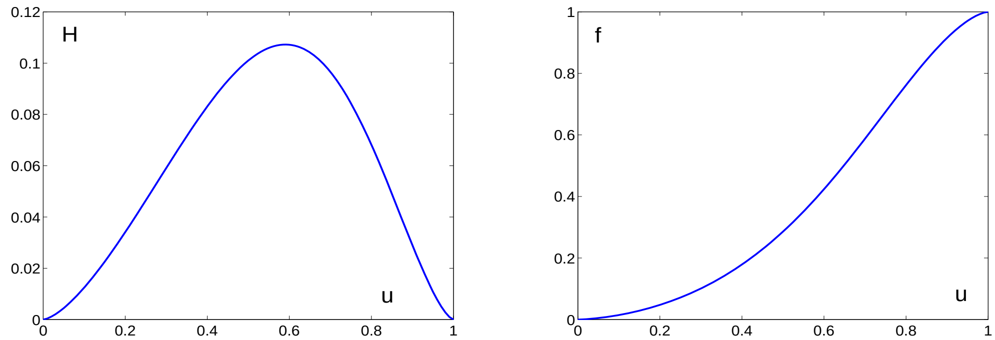


FIGURE 2.1: THE FUNCTIONS H (LEFT) AND f (RIGHT) FOR $p = 0.5$, $q = 0.5$ AND $M = 2.5$

Remark 2.0.1 *The definitions in (2.0.6) and (2.0.7) make sense only in the physically relevant regime* $0 \leq u \leq 1$. *For mathematical completeness we extend* λ_w *and* λ_o *by continuity with constant values* 0 *or* 1 *whenever* u *is outside* $[0, 1]$. *The functions* f *and* H *are extended accordingly.*

For $\varepsilon = 0$, (2.0.1) becomes the non-viscous BL equation

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0, \tag{2.0.8}$$

a hyperbolic conservation law that can be seen as the limit ($\varepsilon \to 0$) of a family of extended equations of the form

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = \mathcal{A}_\varepsilon(u), \quad \varepsilon > 0. \tag{2.0.9}$$

Here, $\mathcal{A}_\varepsilon(u)$ is a regularization term involving second or higher order derivatives. Classical entropy solutions to the BL-equation are constructed as limits of travelling wave (TW) solutions to (2.0.9), when $\mathcal{A}_\varepsilon(u)$ is defined by

$$\mathcal{A}_\varepsilon(u) = \varepsilon \frac{\partial^2 u}{\partial x^2}.$$

A non-classical regularization is given in (2.0.4), which is motivated by dynamic capillarity effects, as mentioned in (2.0.3). For the case $H = 1$, we have the following linear pseudo-parabolic regularization of the BL equation

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = \varepsilon \frac{\partial^2 u}{\partial x^2} + \varepsilon^2 \tau \frac{\partial^3 u}{\partial x^2 \partial t},$$

for which the existence of TW solutions has been studied in [11]. In the limit $\varepsilon \to 0$, these TW solutions become shocks, which are weak solutions to the non-viscous BL equation. These shocks violate the *Oleinik entropy condition*, and therefore are called non-classical. TW solutions to (2.0.9) and the relation with non-standard shock solutions to (2.0.8) are analyzed in [22], see also [2], [3], [17], [19]. The regularization there involves higher order spatial derivatives, but no mixed terms. Furthermore, non-local regularization operators and their effects on shock solutions to hyperbolic conservation laws are studied in [21], [31]. The TW approach for degenerate pseudo-parabolic problems modelling one phase flow in porous media is considered in [4], [6], [13], [28]. In a similar context, a fourth order regularization is studied in [9] .

This chapter deals with (2.0.4), which is a non-linear and degenerate regularization of (2.0.8). In the spirit of [11], we seek TW solutions connecting a left state u_ℓ to a right state u_r. Here we only consider the case $u_l > u_r$, but in the degenerate context. As will be seen below, the degeneracy can lead to the so-called "sharp TW solutions", which are non-smooth solutions, see [1], [23], [24], [32]. The case $u_l < u_r$ in the degenerate context is left for a future investigation. For linear regularization, this has been thoroughly analyzed in [11].

Qualitative properties of solutions to pseudo-parabolic problems are discussed in [20], where the small and waiting-time behavior of solutions to a Darcy type model involving a dynamic pressure saturation is analyzed. A non-linear model involving memory terms is investigated in [30]. Short time existence is obtained in [12], whereas global existence results can be found in [26], [27], [14] and [5]. We refer in particular to the existence results in [26] and [5], which can be applied directly to the present situation. Besides, in these papers it also proved that the solution is bounded essentially by the degeneracy values. This chapter is also motivated by the experimental results in [10]. We further refer to [15] for a review of experimental work on dynamic effects in the pressure-saturation relationship, and to [25] for a dimensional analysis of such models.

This chapter is organized in the following way: In Section 2.1 we investigate the non-linear, non-degenerate case where $0 < u_r < u_\ell < 1$. Then the results are similar to the ones for a linear regularization (see [11]). In particular, a monotone and continuous dependence of τ on u_ℓ is shown. Section 2.2 includes degenerate cases, but is devoted to smooth TW, defined in the classical sense. The focus lies on the case $u_r = 0$. Depending on the parameters, two situations are encountered. These are described in terms of two constants $\alpha \in (u_r, 1)$ and $\beta \in (\alpha, 1]$ that will be defined below. In the first situation, $\beta < 1$ and for any $u_\ell \in (\alpha, \beta)$, there exists a $\tau > 0$ s.t. TW solutions connecting u_l to u_r are possible. Whenever $\beta = 1$, smooth TW solutions are only possible if $\tau < \tau^*$, here τ^* is the threshold value of τ. In the limit case $\tau \nearrow \tau^*$, the corresponding u_l approaches 1. Section 2.3 continues the investigations in Section 2.2 by considering the case $\tau > \tau^*$, when smooth TW solutions are not possible. Then we consider the notion of TW in a larger sense, allowing for discontinuities in the derivatives and connecting $u_\ell = 1$ to $u_r > 0$. In the spirit of [1], [23], [24], [32], such solutions are called "sharp front solutions". These fronts are encountered due to the degeneracy at $u = 1$ and appear at the transition from regions where $u = 1$ to values of u below 1. At the end of Section 2.3 we also consider two degenerate points. Specifically, we take $u_\ell = 1$ and $u_r = 0$. Then we give a selection criterion leading to sharp TW's that are continuously differentiable whenever $u < 1$. In particular, the transition $u > 0$ to $u = 0$ is smooth and encountered at some finite coordinate. Finally, Section 2.4 presents some numerical examples to illustrate the theoretical results.

2.1 Travelling waves: the non-linear, non-degenerate case

Entropy shock solutions to BL equation are based on TW solutions to (2.0.9) and their limits as $\varepsilon \to 0$. A TW solution has the form

$$u(x,t) = u(\eta), \qquad \text{where} \qquad \eta = \frac{x - st}{\varepsilon}. \tag{2.1.1}$$

Note that the TW solution is still denoted by u to avoid unnecessary notations. Applying this into (2.0.4) we obtain

$$-su' + (f(u))' = \{H(u)(u' - \tau s u'')\}'. \tag{2.1.2}$$

With given $0 \le u_r < u_\ell \le 1$, these waves are connecting the left state u_ℓ to the right state u_r,

$$u(-\infty) = u_\ell, \qquad \text{and} \qquad u(+\infty) = u_r. \tag{2.1.3}$$

Remark 2.1.1 *Values outside* $[0,1]$ *are physically unrealistic. From the mathematical point of view, if one of the states is outside* $[0,1]$*, the problem degenerates and the solution remains constant, so no connection with different states is possible. This is why we only consider the case* $0 \le u \le 1$.

Integrating (2.1.2) over $\mathbb{R}$ and assuming that $H(u)(u' - \tau s u'')(\eta) \to 0$ as $\eta \to \pm\infty$ gives

$$-s(u_r - u_\ell) + \{f(u_r) - f(u_\ell)\} = 0.$$

This leads to the *Rankine-Hugoniot (RH) condition* relating the speed s with the limiting values u_ℓ and u_r:

$$(\mathbf{RH}) \qquad s = s(u_r, u_\ell) = \frac{f(u_r) - f(u_\ell)}{u_r - u_\ell}. \tag{2.1.4}$$

Note that since f is a strictly increasing function, it follows that $s > 0$.

Remark 2.1.2 *Instead of assuming* $H(u)(u' - \tau s u'')(\eta) \to 0$ *as* $\eta \to \pm\infty$*, we can just assume* $u'(\eta) \to 0$ *as* $\eta \to \pm\infty$*, then we have* $u''(\eta) \to 0$ *and* $H(u)(u' - \tau s u'')(\eta) \to 0$ *as* $\eta \to \pm\infty$ *automatically.*

Furthermore, integrating (2.1.2) over $(\eta, +\infty)$ and using the condition at $\eta = +\infty$, we obtain

$$s(u - u_r) - \{f(u) - f(u_r)\} = -H(u)(u' - \tau s u''). \tag{2.1.5}$$

Dividing by $H(u)$ and rearranging the terms we obtain the following equation

$$s\tau u'' - u' - g(u; u_r, u_\ell) = 0, \tag{2.1.6}$$

where the function g, depending on the parameters u_r, u_ℓ, is defined as

$$g(u; u_r, u_\ell) \overset{\text{def}}{=} \frac{1}{H(u)} \{s(u_r, u_\ell)(u - u_r) - [f(u) - f(u_r)]\}. \tag{2.1.7}$$

Clearly, in light of the *Rankine-Hugoniot condition* (2.1.4), the function $g(u) = g(u; u_r, u_\ell)$ vanishes at the limit values:

$$g(u_r; u_r, u_\ell) = 0, \qquad \text{and} \qquad g(u_\ell; u_r, u_\ell) = 0.$$

In summary, with the wave speed s defined in (2.1.4), we seek a solution u of the problem

$$(\mathrm{TW}_1)\begin{cases} s\tau u'' - u' - g(u; u_r, u_\ell) = 0, & \eta \in \mathbb{R}, \\ u(-\infty) = u_\ell, \qquad u(+\infty) = u_r. \end{cases} \tag{2.1.8}$$

For some purposes it will be convenient to reformulate problem (TW_1) into a more conventional form and introduce a new spatial variable:

$$\xi = -\frac{1}{\sqrt{s\tau}}\,\eta, \qquad u(\eta) = u(\xi) \quad \text{and} \quad c = \frac{1}{\sqrt{s\tau}}. \tag{2.1.9}$$

Then Problem (TW_1) becomes:

$$(\mathrm{TW}_2)\begin{cases} u'' + cu' - g(u; u_r, u_\ell) = 0, & \xi \in \mathbb{R}, \\ u(-\infty) = u_r, \qquad u(+\infty) = u_\ell. \end{cases} \tag{2.1.10}$$

We start with a necessary condition for the existence of travelling wave solutions with the prescribed limiting values:

Lemma 2.1.1 *A necessary condition for the existence of a travelling wave solution of Problem* $(TW_{1,2})$ *is*

$$\int_{u_r}^{u_\ell} g(u; u_r, u_\ell)\, du > 0. \tag{2.1.11}$$

Proof Multiplying the differential equation in (2.1.10) by u' and integrating over $\mathbb{R}$ yield

$$c\int_{\mathbb{R}} (u')^2 d\xi - \int_{u_r}^{u_\ell} g(u; u_r, u_\ell)\, du = 0.$$

Since $c = 1/\sqrt{s\tau} > 0$, the assertion follows. □

We now make an additional assumption about the function f:

Assumption 1: *There exists a unique* $u^* \in (0, 1)$ *such that* $f''(u^*) = 0$, $f''(u) > 0$ *for* $0 < u < u^*$ *and* $f''(u) < 0$ *for* $u^* < u < 1$.

Remark 2.1.3 *Assumption* 1 *is satisfied for the* f *with* $p > 0$ *and* $q > 0$. *The proof will be given in Appendix A.*

For classical entropy solutions of the BL equation, i.e., when $\tau = 0$, an elementary analysis shows that Problem (TW_1) has a solution if and only if f and two limit states u_ℓ and u_r satisfy the Oleinik Entropy condition:

$$(\mathbf{E}) \qquad \frac{f(u_\ell) - f(u)}{u_\ell - u} \geq \frac{f(u_\ell) - f(u_r)}{u_\ell - u_r} \qquad \text{for} \qquad u_r < u < u_\ell.$$

Since $f'(0) = 0$ and $f''(0) > 0$, the *Oleinik entropy condition* (E) is satisfied in a right-neighbourhood of u_r, provided u_r is chosen small enough. Hence, it is natural to define $\alpha(u_r) > u_r$ such that

$$\alpha(u_r) = \sup\left\{ u > u_r : f'(u) > \frac{f(u) - f(u_r)}{u - u_r} \right\}, \tag{2.1.12}$$

Because $f'(1) = 0$ and f satisfies Assumption 1, it follows that $\alpha < 1$ and that

$$f'(\alpha) = \frac{f(\alpha) - f(u_r)}{\alpha - u_r}.$$

To investigate the non-standard case $\tau > 0$, Lemma 2.1.1 suggests the second critical value $\beta(u_r)$:

$$\beta(u_r) = \sup\left\{ \alpha < u < 1 : \int_{u_r}^{u} g(t\,; u_r, u)\, dt > 0 \right\}. \tag{2.1.13}$$

Plainly, the left state u_ℓ is bounded from above by $\beta(u_r)$.

As we saw in (2.0.7), the diffusion function $H(u)$ vanishes whenever $u = 0$ or $u = 1$, so that the equation becomes degenerate at those values of u. We start by discussing the non-degenerate case and assume in the remaining of this section that $0 < u_r < u_\ell < 1$.

Lemma 2.1.2 *Suppose that* $0 < u_r < \alpha < u_\ell < \beta$. *Then the function* $g(u) \overset{\text{def}}{=} g(u; u_r, u_\ell)$ *defined in (2.1.7) has three positive zeros:*

$$\begin{cases} u = u_r, \quad u = u_m \quad \text{and} \quad u = u_\ell, \\ g'(u_r) > 0, \; g'(u_m) < 0 \;\text{and}\quad g'(u_\ell) > 0, \end{cases} \tag{2.1.14}$$

where primes denote differentiation with respect to u.

The proof is entirely elementary and we omit it. The graph of g is shown in Figure 2.2.

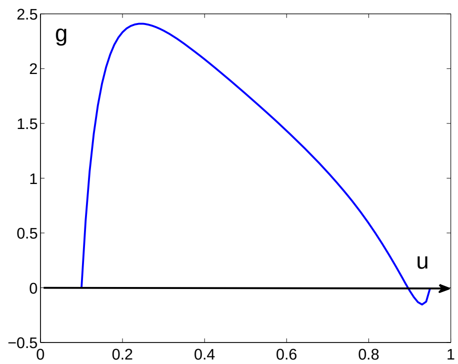


FIGURE 2.2: THE FUNCTION $g(u)$ FOR $p = 0.5$, $q = 0.5$, $M = 2$, $u_r = 0.1$, $u_\ell = 0.95$ AND HENCE, BY THE RH-CONDITION (2.1.4), $s = 1.12$

As we shall see, an important issue will be whether $\beta(u_r) < 1$ or $\beta(u_r) = 1$. In the next lemma we give a condition on the power q in the relative permeability $\lambda_o(u) = u^{q+1}$ which ensures that $\beta < 1$.

Lemma 2.1.3 *If* $q \geq 1$, *then* $\beta < 1$.

Proof Suppose to the contrary that $\beta = 1$. Then by Lemma 2.1.1,

$$\int_{u_r}^{1} g(t; u_r, 1)\, dt \geq 0. \tag{2.1.15}$$

However, an elementary computation shows that

$$g(t; u_r, 1) \sim -s(u_r, 1)(1-t)^{-q} \qquad \text{as} \quad t \to 1^-.$$

Therefore we can find a constant $C > 0$ and a $t_0 \in (u_r, 1)$ such that

$$g(t; u_r, 1) \leq -C(1-t)^{-q} \qquad \text{for all} \quad t \in [t_0, 1). \tag{2.1.16}$$

For $t_0 < u < 1$ we write

$$I(u) \stackrel{\text{def}}{=} \int_{u_r}^{u} g(t; u_r, 1)\, dt = \int_{u_r}^{t_0} g(t; u_r, 1)\, dt + \int_{t_0}^{u} g(t; u_r, 1)\, dt = I_1 + I_2(u). \tag{2.1.17}$$

From the upper bound given in (2.1.16) we have

$$I_2(u) < -C \int_{u_0}^{u} (1-t)^{-q}\, dt,$$

and hence, because $q \geq 1$, it follows that $I(u) = I_1 + I_2(u) \to -\infty$ as $u \to 1^-$. This contradicts (2.1.15) and thus proves the assertion. □

Remark 2.1.4 *The same result holds when $u_r = 0$. This means that if $q \geq 1$, the left state u_ℓ can not be 1, so degeneracy may only occur at $u = 0$.*

Whenever $\beta < 1$, we can establish the following existence theorem. Its proof goes along the lines of Lemma 4.1 in [11]. We omit the details here.

Theorem 2.1.1 *Let $u_r > 0$ be given so that α is well defined by (2.1.12), and let β be defined by (2.1.13). Then, for every $u_\ell \in (\alpha,\beta)$ there exists a unique value of $\tau > 0$ such that Problem (TW_1) admits a solution. This solution is unique, decreasing and travels with speed $s(u_\ell, u_r)$ given in (2.1.4).*

Collorary 2.1.1 *For fixed $u_r > 0$ and given $u_\ell \in (\alpha,\beta)$, Theorem 2.1.1 provides a unique $\tau = \tau(u_\ell)$. Thus, we can define the function*

$$\tau : (\alpha,\beta) \to \mathbb{R}^+.$$

Lemma 2.1.4 *The function τ is continuous and strictly increasing on (α,β).*

Proof We follow the proof of Lemma 4.2 in [11] and show first the monotonicity of τ. Taking two left states $\alpha < u_{\ell,1} < u_{\ell,2} < \beta$, define

$$s_i = \frac{f(u_{\ell,i}) - f(u_r)}{u_{\ell,i} - u_r}, \qquad \text{and} \qquad g_i(u) = \frac{1}{H(u)}\{s_i(u-u_r) - [f(u) - f(u_r)]\}, \quad i = 1,2.$$

Observe that

$$\frac{d}{du}\left(\frac{f(u) - f(u_r)}{u - u_r}\right) < 0 \qquad \text{for} \qquad \alpha < u < \beta,$$

therefore

$$s_1 > s_2 \qquad \text{and} \qquad g_1(u) > g_2(u) \qquad \text{for all} \quad u \in (u_r, u_{\ell,1}).$$

Rewriting Problem (TW_2) as the first order system

$$\begin{cases} u' = w, \\ w' = -c_i w + g_i(u), \end{cases} \tag{2.1.18}$$

we denote by Γ_i ($i = 1, 2$) the orbits emerging from the saddle $(u_r, 0)$. They do so under the angles θ_i given by

$$\theta_i = \frac{1}{2}\left(\sqrt{c_i^2 + 4g_i'(u_r)} - c_i\right), \qquad i = 1, 2.$$

Plainly,

$$g_i'(u_r) = \frac{1}{H(u_r)}\{s_i - f'(u_r)\}.$$

Define the function

$$\theta(c, s) \stackrel{\text{def}}{=} \frac{1}{2}\left(\sqrt{c^2 + 4g'(u_r)} - c\right).$$

Then

$$\frac{\partial\theta}{\partial c} < 0 \qquad \text{and} \qquad \frac{\partial\theta}{\partial s} > 0,$$

and we conclude that $c_1 > c_2$ as in [11]. This gives

$$\tau_2 s_2 > \tau_1 s_1, \quad \text{therefore} \quad \tau_2 > \frac{s_1}{s_2}\tau_1 > \tau_1,$$

as asserted. For the continuity of τ, one can follow the ideas in Lemma 4.3 in [11]. □

2.2 Smooth travelling waves for the degenerate case $u_r = 0$

In this section, we seek standard, smooth TW solutions for the degenerate case $u_r = 0$, where H vanishes, whereas g becomes unbounded.

Since $u_l > u_r$ is assumed, we seek monotone decreasing TW solutions of Problem (TW_1). If such solutions exist, a bijection $\eta \to u$, where $\eta \in \mathbb{R}$ and $u \in (u_r, u_\ell)$, can be defined. Therefore the functions $\eta : (u_r, u_\ell) \to \mathbb{R}$, as well as $w : (u_r, u_\ell) \longrightarrow \mathbb{R}$, $w(u) = -u'(\eta(u))$ can be defined. Further, since u is decreasing, we have $w > 0$ on (u_r, u_ℓ). Nevertheless, $w(u_r)$ and $w(u_\ell)$ still have to be fixed. To do so we recall that the waves sought in this section are smooth, monotone, and approaching u_ℓ and u_r asymptotically.

Therefore we have $\lim_{\eta\to\pm\infty} u'(\eta) = 0$ yielding $w(u_r) = w(u_l) = 0$. In terms of w, Problem (TW_1) introduced in (2.1.8) becomes

$$(\mathrm{TW}_3)\begin{cases} \tau s w w' + w = g(u), & \text{for } u \in (u_r, u_\ell), \\ w(u_r) = 0, w(u_\ell) = 0. \end{cases} \tag{2.2.1}$$

Note that this first order equation has two boundary conditions, which will be fixed by the parameter τ. When seeking monotone TW solutions to Problem (TW_3) we in fact seek a pair $(w, \tau) \in C^1(u_r, u_\ell) \times (0, \infty)$ for which (2.2.1) holds. Once w is known, one can obtain u from:

$$\eta(u) = \int_u^{\frac{u_\ell+u_r}{2}} \frac{dz}{w(z)}, \tag{2.2.2}$$

defining a TW satisfying $u(0) = \frac{u_\ell+u_r}{2}$. This choice of $u(0)$ is a possible normalization of the TW, any value in (u_r, u_ℓ) is possible.

We start with a non-existence result, which imposes a restriction on the value of the power p.

Theorem 2.2.1 *If $p \geq 1$ and $u_r = 0$, then no TW solutions are possible.*

Proof As in (2.1.6), whenever a TW solution exists, it satisfies the equation

$$s\tau u'' - u' - g(u; u_r, u_\ell) = 0. \tag{2.2.3}$$

The particular forms of f and H allow writing

$$s\tau u'' - u' \geq C u^{-p}, \tag{2.2.4}$$

for some $C > 0$, when u is close to 0. This shows that u can not oscillate near $u = 0$. To see this, we note that at each (local) extremum, where $u' = 0$, we have

$$u'' \geq \frac{C}{s\tau} u^{-p} > 0, \tag{2.2.5}$$

which implies that u is always convex there. Therefore u can not have a local maximum, so it is monotone.

Based on this observation we rule out the existence of a TW solution. To see this we assume that a TW solution u exists. Then up to a $\tilde{u} > 0$ small enough, u is increasing from 0 to $\tilde{u}$. This allows rewriting the TW equation into (2.2.1) with $u_r = 0$ and $w(0) = 0$, yielding after integration

$$\frac{\tau s}{2} w^2(u) + \int_0^u w(v)dv = \int_0^u g(v)dv, \tag{2.2.6}$$

for all $u \in (0, \tilde{u})$. As g becomes unbounded at $u = 0$, $\tilde{u}$ can be chosen such that $w(\tilde{u}) < g(\tilde{u})$, implying the same ordering everywhere in $(0, \tilde{u})$. This gives

$$\int_0^{\tilde{u}} g(u)du \leq \frac{\tau s}{2} g(\tilde{u})^2 + \tilde{u} g(\tilde{u}). \tag{2.2.7}$$

Since the left side is unbounded for $p \geq 1$, we obtain a contradiction. □

Remark 2.2.1 *No TW solutions exist for $u_\ell = 1$ whenever $q \geq 1$. To see this, one can follow the proof of Theorem 2.2.1.*

In view of the above, no TW solutions exist for $p \geq 1$ and $u_r = 0$, or $q \geq 1$ and $u_\ell = 1$. Therefore in this section we restrict to the cases $0 < p < 1$ and $0 < q < 1$. In the following, we present the existence

Theorem 2.2.2 *Let $0 < p < 1$, $u_r = 0$ and $\alpha < u_\ell < \beta$. Then there exists a unique τ for which Problem (TW_1) admits a solution.*

Proof The proof is divided into two steps:
Step 1: Existence. We prove that there exists a unique pair (τ, w) such that (2.2.1) is satisfied. By (2.2.2), this also provides a solution to Problem (TW_1).

Observe that, since $u_r = 0$ and $u_\ell < 1$, we have $g(u) \to +\infty$ as $u \to 0$ and $g(u_m) = g(u_\ell) = 0$ for some $u_m \in (0, u_\ell)$. Clearly, for any $\tilde{w} > 0$, there exists a unique $\tilde{u} \in (0, u_m)$ such that $g(\tilde{u}) = \tilde{w}$. For proving the theorem, we consider two cases $u > \tilde{u}$, and $u < \tilde{u}$.

We prove the existence of a pair (w, τ) by a two-way shooting method starting from the segment of the graph of the function $g(u)$ that lies between $u = 0$ and $u = u_m$. Thus, fixing $\tilde{u} \in (0, u_m)$ we first consider the initial value problem

$$\begin{cases} \tau s w w' + w = g(u), & \text{for } u > \tilde{u}, \\ w(\tilde{u}) = g(\tilde{u}). \end{cases} \tag{2.2.8}$$

We denote the solution by $w(u;\tau)$. We prove that there exists a $\tau = \tau(\tilde{u})$ such that the solution of (2.2.8) also satisfies $w(u_\ell) = 0$. To this end we define the point

$$v^+(\tau) = \sup\{\tilde{u} < u \le u_\ell \mid w(\cdot\,;\tau) > 0 \text{ on } [\tilde{u}, u)\}.$$

We make two observations:

(1) If $\tau \to \infty$ then $w'(u;\tau) \to 0$ uniformly on $[\tilde{u}, v]$ so that if τ is large enough then $w(u_\ell) > 0$ and $v^+(\tau) = u_\ell$.

(2) If $\tau \to 0$ then $w(u;\tau) - g(u) \to 0$ uniformly on $[\tilde{u}, v]$ and hence if τ is small enough, then $v^+(\tau) < u_\ell$.

Due to the continuous dependence of $w(u;\tau)$ on τ there exists a $\tau = \tau(\tilde{u})$ such that $w(\cdot\,;\tau(\tilde{u})) > 0$ on $(\tilde{u}, u_\ell)$ and $w(u_\ell;\tau(\tilde{u})) = 0$.

Next we turn to the backward problem

$$\begin{cases} \tau(\tilde{u})sww' + w = g(u), & \text{for } u < \tilde{u}, \\ w(\tilde{u}) = g(\tilde{u}), \end{cases} \tag{2.2.9}$$

where we now have fixed the value of τ to the one obtained in the forward problem. Since the starting value $\tilde{u}$ is still not determined, we now denote the solution of the backward problem in (2.2.9) by $w(u;\tilde{u})$. Proceeding as the forward problem we define the point

$$v^-(\tilde{u}) = \inf\{0 < u < \tilde{u} \mid w(\cdot\,;\tilde{u}) > 0 \text{ on } (u, \tilde{u}]\},$$

and we define three sets:

$$\begin{aligned} \mathcal{A}_+ &\overset{\text{def}}{=} \{\tilde{u} \in (0, u_m) \mid v^-(\tilde{u}) = 0 \text{ and } w(0;\tilde{u}) > 0\}, \\ \mathcal{A}_0 &\overset{\text{def}}{=} \{\tilde{u} \in (0, u_m) \mid v^-(\tilde{u}) = 0 \text{ and } w(0;\tilde{u}) = 0\}, \\ \mathcal{A}_- &\overset{\text{def}}{=} \{\tilde{u} \in (0, u_m) \mid v^-(\tilde{u}) > 0\}. \end{aligned} \tag{2.2.10}$$

Plainly, $\mathcal{A}_+$ and $\mathcal{A}_-$ are open; we shall prove that they are also nonempty. This then enables us to conclude that the set $\mathcal{A}_0$ is nonempty, and hence that there exists a value of $\tilde{u}$ and hence of τ such that Problem (TW$_3$) has a solution if $u_r = 0$. This is proved in the lemmas below.

Lemma 2.2.1 *The set $\mathcal{A}_-$ is nonempty.*

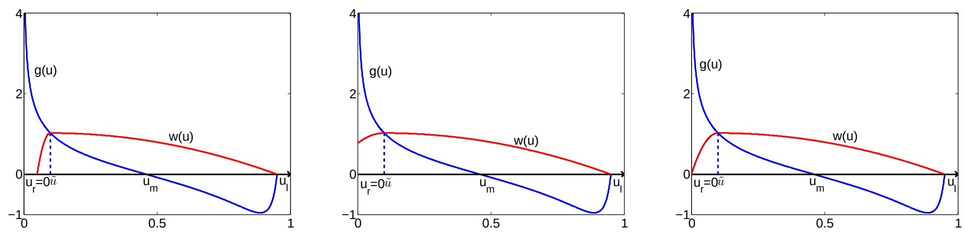


FIGURE 2.3: SOLUTIONS IN THE SET $\mathcal{A}_-$ (LEFT), THE SET $\mathcal{A}_+$ (MIDDLE) AND THE SET $\mathcal{A}_0$ (RIGHT)

Proof Suppose to the contrary that $\mathcal{A}_-$ is empty and hence $\nu^-(\tilde{u}) = 0$ for all $\tilde{u} \in (0, u_m)$. Integration of the equation in Problem (TW_3) over $(0, u_\ell)$ yields

$$-\frac{1}{2}\tau(\tilde{u})sw^2(0;\tilde{u}) + \int_0^{u_\ell} w(u;\tilde{u})du = \int_0^{u_\ell} g(u)du \stackrel{\text{def}}{=} G(u_\ell). \tag{2.2.11}$$

Note that in light of Lemma 2.1.1, $G(u_\ell) > 0$. Since $w(u;\tilde{u}) < g(\tilde{u})$, we have

$$G(u_\ell) - u_\ell g(\tilde{u}) < -\frac{1}{2}\tau(\tilde{u})sw^2(0;\tilde{u}). \tag{2.2.12}$$

Since $g(u) \to 0$ as $u \to u_m$, we can ensure that $G(u_\ell) - u_\ell g(\tilde{u}) > 0$ by choosing $\tilde{u}$ sufficiently close to u_m. In this manner we arrive at a contradiction. □

Lemma 2.2.2 *The set $\mathcal{A}_+$ is nonempty.*

Before proving this lemma, we establish two auxiliary results.

Lemma 2.2.3 *We have*

$$\tau(\tilde{u}) \to 0 \qquad \textit{and} \qquad w(u_m;\tilde{u}) \to \infty \qquad \textit{as} \qquad \tilde{u} \to 0. \tag{2.2.13}$$

Proof Integration of the equation over $(\tilde{u}, u_\ell)$ yields

$$\int_{\tilde{u}}^{u_\ell} g(u)du = -\frac{1}{2}s\tau(\tilde{u})g^2(\tilde{u}) + \int_{\tilde{u}}^{u_\ell} w(u;\tilde{u})du < -\frac{1}{2}s\tau(\tilde{u})g^2(\tilde{u}) + u_\ell g(\tilde{u}),$$

where we have used the fact that by construction $w(u;\tilde{u}) \leq g(\tilde{u})$. Thus,

$$\frac{1}{2}s\tau(\tilde{u}) < \frac{u_\ell}{g(\tilde{u})} - \frac{1}{g^2(\tilde{u})}\int_{u_m}^{u_\ell} g(u)du.$$

Since $g(\tilde{u}) \to \infty$ as $\tilde{u} \to 0$, the first assertion follows.

To prove the second assertion we claim that

$$w''(u;\tilde{u}) < 0 \qquad \text{for} \qquad u_m \leq u \leq u_e, \tag{2.2.14}$$

where u_e is the value of u where $g(u)$ reaches its minimum. Accepting this claim for the moment and observing that $w'(u_m;\tilde{u}) = 1/s\tau(\tilde{u})$, we conclude that

$$w(u_e;\tilde{u}) < w(u_m;\tilde{u}) - \frac{1}{s\tau(\tilde{u})}(u_e - u_m).$$

Since $w(u_e;\tilde{u}) > 0$ this implies that

$$w(u_m;\tilde{u}) > \frac{1}{s\tau(\tilde{u})}(u_e - u_m),$$

and hence, in light of the first result, that $w(u_m;\tilde{u}) \to \infty$ as $\tilde{u} \to 0$.

It remains to prove the claim (2.2.14). We have

$$w'' = \frac{1}{s\tau(\tilde{u})w^2}\{g'(u)w - g(u)w'\}. \tag{2.2.15}$$

Because $g' < 0$ and $g > 0$ as well as $w > 0$ and $w' < 0$ on (u_m, u_e) it follows that the right-hand side of (2.2.15) is negative and the claim is proved. □

We are now ready to prove that the set $\mathcal{A}_+$ is nonempty.

Proof of Lemma 2.2.2 We need to prove that there exists a $\tilde{u} \in (0, u_m)$ such that $\nu^-(\tilde{u}) = 0$ and $w(0;\tilde{u}) > 0$.

Suppose to the contrary that for any $\tilde{u} \in (0, u_m)$, either $\nu^-(\tilde{u}) > 0$ or $\nu^-(\tilde{u}) = 0$ and $w(0;\tilde{u}) = 0$.

To arrive at a contradiction, we integrate the equation in Problem (TW$_3$) over (ν^-, u_ℓ):

$$\frac{1}{2}\tau(\tilde{u})\{w^2(u_\ell;\tilde{u}) - w^2(\nu^-(\tilde{u});\tilde{u})\} + \int_{\nu^-(\tilde{u})}^{u_\ell} w(u;\tilde{u})\,du = \int_{\nu^-(\tilde{u})}^{u_\ell} g(u)du,$$

or, since w vanishes at u_ℓ and v^-,

$$\int_{v^-(\tilde u)}^{u_\ell} g(u)\,du = \int_{v^-(\tilde u)}^{u_\ell} w(u;\tilde u)\,du > \int_{\tilde u}^{u_m} w(u;\tilde u)\,du > (u_m - \tilde u)w(u_m;\tilde u).$$

As $\tilde u \to 0$, the right-hand term tends to infinity, whilst the left-hand term remains bounded; a contradiction. This completes the proof of Lemma 2.2.2 and thereby of the existence of a $\tilde u \in (0, u_m)$ for which the pair $(w(\cdot\,;\tilde u), \tau(\tilde u))$ is a solution of Problem (TW$_3$). $\square$

Step 2: Uniqueness. We assume the existence of (τ_1, w_1) and (τ_2, w_2) satisfying (2.2.1), where $\tau_1 > \tau_2 > 0$. Integrating (2.2.1) over $(0, u)$ yields

$$\frac{s\tau_1}{2} w_1(u)^2 + \int_0^u w_1(t)dt = \int_0^u g(t)dt = \frac{s\tau_2}{2} w_2(u)^2 + \int_0^u w_2(t)dt. \qquad (2.2.16)$$

Since $\tau_1 > \tau_2$, this gives $w_1 < w_2$ for u small enough (see the left picture of Figure 2.4). If w_1 and w_2 do not intersect inside $(0, u_\ell)$, then $w_1 < w_2$ everywhere. Taking $u = u_\ell$ in (2.2.16) gives

$$\int_0^{u_\ell} w_1(u)du = \int_0^{u_\ell} g(u)du = \int_0^{u_\ell} w_2(u)du,$$

which contradicts the ordering in the $w's$. Thus w_1 and w_2 must have at least one in-

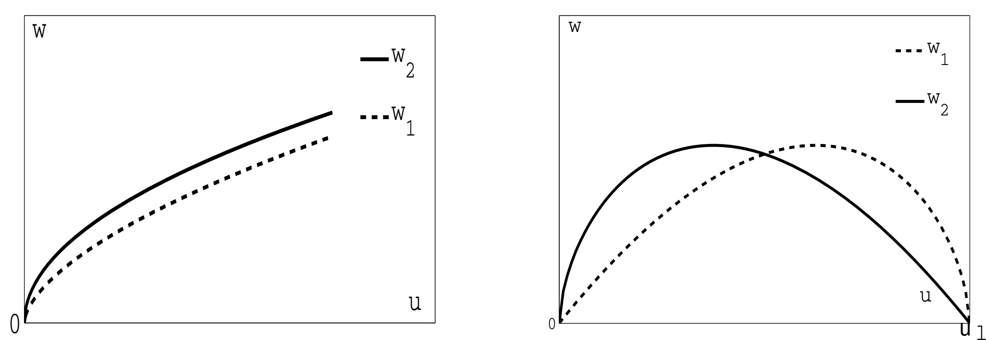


FIGURE 2.4: SKETCHED SOLUTIONS w_1 AND w_2 OF (2.2.1) FOR, RESPECTIVELY, $\tau_1 > \tau_2$: BEHAVIOUR CLOSE TO $u = 0$ (LEFT), AND GLOBAL BEHAVIOUR ASSUMING A UNIQUE INTERSECTION POINT INSIDE $(0, u_l)$ (RIGHT).

tersection point inside $(0, u_\ell)$, where w_1 and w_2 are both positive. No intersection can occur at points where w_1 or w_2 is increasing. To see this we let $u_0 \in (0, u_\ell)$ be the first intersection point, so $w_1(u_0-) < w_2(u_0-)$ implying that $w_1'(u_0) \geq w_2'(u_0)$. However, since

$$w_1'(u_0) = \frac{g(u_0) - w_1(u_0)}{s\tau_1 w_1(u_0)}, \quad \text{and} \quad w_2'(u_0) = \frac{g(u_0) - w_2(u_0)}{s\tau_2 w_2(u_0)},$$

if $w_1(u_0) < g(u_0)$, then

$$w_2'(u_0) = \frac{g(u_0) - w_2(u_0)}{s\tau_2 w_2(u_0)} = \frac{g(u_0) - w_1(u_0)}{s\tau_2 w_1(u_0)} > \frac{g(u_0) - w_1(u_0)}{s\tau_1 w_1(u_0)} = w_1'(u_0),$$

which contradicts the above. Next, if $w_1(u_0) = g(u_0)$, then $w_1'(u_0) = w_2'(u_0) = 0$, $g'(u_0) \neq 0$ and there exists $\delta > 0$ small enough such that

$$w_1'(u_0 - \delta) > w_2'(u_0 - \delta) > 0, \quad \text{and} \quad w_2(u_0 - \delta) > w_1(u_0 - \delta) > 0.$$

Hence

$$(w_1' - w_2')|_{u=u_0-\delta} = \left[\frac{1}{s\tau_1}\left(\frac{g}{w_1} - 1\right) - \frac{1}{s\tau_2}\left(\frac{g}{w_2} - 1\right)\right]\Big|_{u=u_0-\delta} > 0,$$

therefore

$$\frac{\tau_2}{\tau_1} > \frac{\frac{g}{w_2} - 1}{\frac{g}{w_1} - 1}\Big|_{u=u_0-\delta}.$$

As $\delta \to 0$, l'Hôpital's rule gives

$$\lim_{\delta\to 0} \frac{\frac{g}{w_2} - 1}{\frac{g}{w_1} - 1}\Big|_{u=u_0-\delta} = \lim_{\delta\to 0} \frac{\frac{g' w_2 - g w_2'}{w_2^2}}{\frac{g' w_1 - g w_1'}{w_1^2}}\Big|_{u=u_0-\delta} = 1,$$

thus

$$\frac{\tau_2}{\tau_1} \geq 1,$$

contradicting with $\tau_1 > \tau_2$. Thus $w_1(u_0) > g(u_0)$, implying $w_1'(u_0) < 0$ and $w_2'(u_0) < 0$.

We assume now there exist at least two intersection points inside $(0, u_\ell)$, and let u_0 and u_1 be the first two of them. We have $0 > w_1'(u_0) \geq w_2'(u_0)$ and $w_1'(u_1) \leq w_2'(u_1) < 0$. However

$$w_1'(u_1) = \frac{g(u_1) - w_1(u_1)}{s\tau_1 w_1(u_1)} = \frac{g(u_1) - w_2(u_1)}{s\tau_1 w_2(u_1)} > \frac{g(u_1) - w_2(u_1)}{s\tau_2 w_2(u_1)} = w_2'(u_1),$$

contradicting the above.

It only remains to rule out the possibility of having exactly one intersection point $u_0 \in (0, u_\ell)$ (see the right picture of Figure 2.4). Then $w_1 > w_2$ for $u \in (u_0, u_\ell)$. Since $w_1(u_\ell) = w_2(u_\ell)$ there exists u_2 close to u_ℓ such that $w_1'(u_2) < w_2'(u_2)$. However, since $\tau_1 > \tau_2$, $w_1(u_2) > w_2(u_2)$ and $g(u_2) < 0$,

$$w_1'(u_2) = \frac{g(u_2)}{s\tau_1 w_1(u_2)} - \frac{1}{s\tau_1} > \frac{g(u_2)}{s\tau_2 w_2(u_2)} - \frac{1}{s\tau_2} = w_2'(u_2),$$

which is a contradiction and shows the uniqueness. □

In the following lemma we show that $\tau(u_\ell)$ in Theorem 2.2.2 is a strictly increasing function of $u_\ell \in (\alpha, \beta)$.

Lemma 2.2.4 *Let $\alpha < u_{\ell,1} < u_{\ell,2} < \beta$. Then $\tau(u_{\ell,1}) < \tau(u_{\ell,2})$.*

Proof For convenience we write $u_i = u_{\ell,i}$ and $g_i(u) = g(u; 0, u_{\ell,i})$, $\tau(u_{\ell,i}) = \tau_i$, $s_i = \frac{f(u_{\ell,i})}{u_{\ell,i}}$ for $i = 1, 2$. Then w_i $(i = 1, 2)$ satisfies

$$\begin{cases} \tau_i s_i w_i w_i' + w_i = g_i(u), & \text{for } u \in (0, u_i), \\ w_i(0) = 0, w_i(u_i) = 0. \end{cases} \tag{2.2.17}$$

Here we prove the statement:

$$\text{Let } \tau_1 < \tau_2, \text{ then } u_1 < u_2,$$

which is equivalent to the conclusion of the lemma. The case $u_1 = u_2$ is ruled out by the uniqueness result in Theorem 2.2.2. Assuming $u_1 > u_2$ we have

$$\begin{aligned} w_1(u_2) > 0 &= w_2(u_2), \\ s_1 &< s_2, \\ g_1(u) < g_2(u) & \text{ for } u \in (0, u_2). \end{aligned} \tag{2.2.18}$$

In the following, we prove the assertion

$$s_1 \tau_1 w_1^2 \le s_2 \tau_2 w_2^2, \quad \text{for all } u \in [0, u_2]. \tag{2.2.19}$$

This gives

$$0 \le s_1 \tau_1 w_1^2(u_2) \le s_2 \tau_2 w_2^2(u_2) = 0, \tag{2.2.20}$$

implying $w_1(u_2) = 0$. This contradicts the first inequality in (2.2.18), and therefore the assumption $u_1 > u_2$ is not true.

To prove (2.2.19) we use (2.2.17) and obtain for all $u \in [0, u_2]$

$$\frac{1}{2}\left(s_1\tau_1(w_1^2)' - s_2\tau_2(w_2^2)'\right) + w_1 - w_2 = g_1 - g_2 \le 0. \tag{2.2.21}$$

With $[z]_+$ denoting the positive cut of z, we multiply (2.2.21) by $[s_1\tau_1 w_1^2 - s_2\tau_2 w_2^2]_+$ and get

$$\begin{aligned}&\frac{1}{2}\left([s_1\tau_1 w_1^2 - s_2\tau_2 w_2^2]_+^2\right)' + (w_1 - w_2)[s_1\tau_1 w_1^2 - s_2\tau_2 w_2^2]_+ \\ &= (g_1 - g_2)[s_1\tau_1 w_1^2 - s_2\tau_2 w_2^2]_+.\end{aligned} \tag{2.2.22}$$

Integrating (2.2.22) from 0 to u ($u \in [0, u_2]$) gives

$$\begin{aligned}&\frac{1}{2}[s_1\tau_1 w_1(u)^2 - s_2\tau_2 w_2(u)^2]_+^2 \\ &\qquad + \frac{1}{\sqrt{s_1\tau_1}}\int_0^u \left(\sqrt{s_1\tau_1}w_1(v) - \sqrt{s_2\tau_2}w_2(v)\right)[s_1\tau_1 w_1(v)^2 - s_2\tau_2 w_2(v)^2]_+ dv \\ &\qquad + \int_0^u \left(\sqrt{\frac{s_2\tau_2}{s_1\tau_1}} - 1\right) w_2(v)[s_1\tau_1 w_1(v)^2 - s_2\tau_2 w_2(v)^2]_+ dv \\ &= \int_0^u \left(g_1(v) - g_2(v)\right)[s_1\tau_1 w_1(v)^2 - s_2\tau_2 w_2(v)^2]_+ dv.\end{aligned} \tag{2.2.23}$$

Denoting the terms on the left side by T_1, T_2 and T_3, we observe that $T_1 \ge 0$. Further

$$T_2 = \frac{1}{\sqrt{s_1\tau_1}}\int_0^u \left(\sqrt{s_1\tau_1}w_1(v) + \sqrt{s_2\tau_2}w_2(v)\right)[\sqrt{s_1\tau_1}w_1(v) - \sqrt{s_2\tau_2}w_2(v)]_+^2 dv \ge 0. \tag{2.2.24}$$

Since $s_2\tau_2 > s_1\tau_1$ and $w_2 \ge 0$, we have $T_3 \ge 0$.

By (2.2.18) the right hand side is negative. In view of the above, this gives

$$[\sqrt{s_1\tau_1}w_1(v) - \sqrt{s_2\tau_2}w_2(v)]_+^2 = 0 \quad \text{for all } v \in [0, u_2], \tag{2.2.25}$$

implying (2.2.19). □

The proof of Theorem 2.2.2 also provides bounds for the parameter τ in terms of the function

$$G(u) = \int_0^u g(t)\, dt,$$

and the intermediate zero $u_m \in (0, u_\ell)$.

Lemma 2.2.5 *We have*

$$\tau < \frac{2}{s}\left(\frac{u_\ell}{G(u_\ell)}\right)^2 G(u_m). \tag{2.2.26}$$

Proof Recall that

$$\int_0^{u_\ell} w(u) = \int_0^{u_\ell} g(u)\,du = G(u_\ell).$$

Let $w_{\max} = \max\{w(u) : 0 < u < u_\ell\}$. Then $\int_0^{u_\ell} w(u)\,du < u_\ell w_{\max}$ and it follows that

$$w_{\max} > \frac{G(u_\ell)}{u_\ell} \stackrel{\text{def}}{=} A. \tag{2.2.27}$$

Let $w(u)$ reach its maximum value at $u = u_{\max}$. Inspection of the function $g(u)$ shows that $0 < u_{\max} < u_m$. Integrating the equation in (2.2.1) over $(0, u_{\max})$, we obtain

$$\frac{\tau s}{2} w_{\max}^2 < \int_0^{u_{\max}} g(u)\,du < \int_0^{u_m} g(u)\,du = G(u_m). \tag{2.2.28}$$

Combining (2.2.27) and (2.2.28) yields the desired upper bound of (2.2.26). □

Lemma 2.2.5 can be completed by a lower bound for τ. To this aim we introduce two particular values of u:

$\mathbf{u_A}$: We put

$$u_A \stackrel{\text{def}}{=} \sup\{u < u_m : g(u) < A = u_\ell^{-1} G(u_\ell)\}.$$

Since $w_{\max} > A$ and $w_{\max} = g(u_{\max})$ it follows that u_A exists and that $u_{\max} < u_A < u_m$.

$\mathbf{u_e}$: Note that $g(u)$ is negative and has a minimum on the interval (u_m, u_ℓ). The value of u where $g(u)$ attains its minimum value on (u_m, u_ℓ) will be denoted by u_e.

Lemma 2.2.6 *We have*

$$\tau > \frac{1}{s}\frac{(u_e - u_m)(u_m - u_A)}{G(u_\ell)}. \tag{2.2.29}$$

Proof Since $g(u_m) = 0$ it follows from the equation in (2.2.1) that $w'(u_m) = -\frac{1}{\tau s}$. Dividing the equation by w and differentiating the result, we obtain

$$w'' = \frac{1}{\tau s}\frac{g'w - gw'}{w^2}.$$

Since $g' < 0$ on (u_m, u_e) and both $g < 0$ and $w' < 0$ on this interval, it follows that $w'' < 0$ on (u_m, u_ε). Therefore

$$0 < w(u_e) < w(u_m) - \frac{1}{\tau s}(u_e - u_m).$$

Therefore

$$\tau > \frac{1}{s}\frac{u_e - u_m}{w(u_m)}. \tag{2.2.30}$$

It remains to establish an upper bound for $w(u_m)$. Observe that

$$\int_0^{u_\ell} g(u)du = \int_0^{u_\ell} w(u)du > \int_{u_{\max}}^{u_m} w(u)du > (u_m - u_{\max})w(u_m).$$

Hence,

$$w(u_m) < \frac{G(u_\ell)}{u_m - u_{\max}} < \frac{G(u_\ell)}{u_m - u_A}, \qquad \text{since} \quad u_{\max} < u_A. \tag{2.2.31}$$

Combining the inequalities (2.2.30) and (2.2.31) we arrive at the desired lower bound. □

Remark 2.2.2 *If $\beta = 1$, then the upper bound of Lemma 2.2.5 is uniformly bounded with respect to $u_\ell \in (0,1)$. This means in particular that*

$$\limsup_{u_\ell \to 1^-} \tau(u_\ell) < \infty.$$

Remark 2.2.3 *For the non-degenerate case, we can proceed in the same manner to get similar bounds.*

The proof of Theorem 2.2.2 can be extended without major differences to the case $u_\ell = 1$, gives the following result:

Theorem 2.2.3 *Let $u_r \geq 0$, $u_\ell = 1$ and assume $\int_{u_r}^1 g(t; u_r, 1)\, dt > 0$ (thus $\beta = 1$). Then there exists a unique pair (τ, w) solving (2.2.1).*

As we can see from Remark 2.2.2 and Theorem 2.2.3, with $u_r \geq 0$ and $u_\ell \in (u_r, 1]$, if $\beta = 1$ then travelling wave solutions exist only for finite τ. As mentioned in Remark 2.2.2, together with Lemma 2.2.4 this justifies the following

Definition 2.2.1 *Let $u_r \geq 0$ and assume $\int_{u_r}^1 g(u; u_r, 1) > 0$, define*

$$\tau^* = \sup\{\tau(u_\ell) : u_\ell \in (\alpha, 1)\}.$$

The value τ^* can be interpreted as the maximal value of τ for which there exists a travelling wave solution connecting u_ℓ to u_r. We further have $\tau^* = \tau(1)$.

The right state u_r is fixed in Definition 2.2.1. Note that the value τ^* depends on u_r, so the notation $\tau^*(u_r)$ makes sense. Without doing a rigorous proof on the dependency of u_r, we refer to Figure 2.5, where τ^* is computed numerically for $p = q = 0.5$, $M = 2.5$ and u_r ranges from 0.01 to 0.1. This figure suggests that τ^* is an increasing function with u_r.

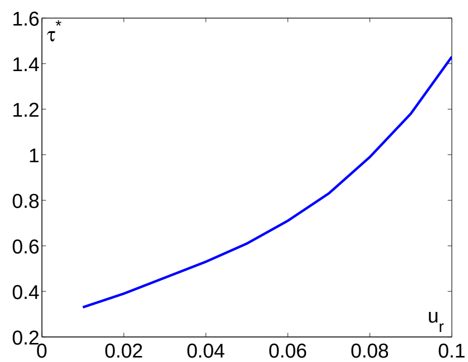


FIGURE 2.5: THE VALUES OF $\tau^* = \tau^*(u_r)$ FOR $u_r \in [0.01, 0.1]$ WHEN $p = q = 0.5, M = 2.5$

2.3 Non-smooth travelling waves

The results proven yet are obtained for smooth C^2 travelling wave solutions of (2.0.4). Following Theorem 2.2.3 (see also Remark 2.2.2), when $\beta = 1$ then there is an upper bound to the values of τ for which such solutions exist. Specifically, given a $u_r \geq 0$ there exists a critical constant $\tau^*(u_r)$, such that if $\tau < \tau^*(u_r)$ there exists a unique $u_\ell < 1$ for which a smooth travelling wave exists which connects the left state u_ℓ and the right state u_r. In the limit when $\tau = \tau^*(u_r)$, we have $u_\ell = 1$. A natural question is then: what happens when $\tau > \tau^*(u_r)$?

An indication of what to expect can be found in the graph presented in Figure 2.6, which is the numerical solution $u(x,t)$ of the Cauchy-Dirichlet Problem for the original partial differential equation (2.0.4) on a sufficiently large interval $(-1, 19)$, with initially a smooth approximation of a jump from the left state $u_\ell = 1$ to the right state $u_r = 0.1$, located at $x = 0$. At the boundaries we consider the limiting values, u_ℓ and u_r. In this computation, we took $p = q = 0.5$, $M = 2.5$, $\varepsilon = 1$ and $\tau = 2$ which exceeds the critical value $\tau^*(u_r)$. Details of the numerical method are provided in Section 2.4. Observe that $\partial u/\partial x$ becomes discontinuous at the point where the value of u first changes from $u = 1$ to $u < 1$.

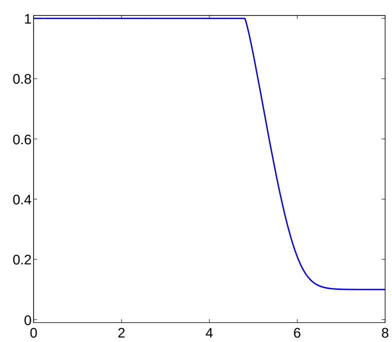


FIGURE 2.6: THE NUMERICAL SOLUTION u OF (2.0.4), COMPUTED FOR $u_\ell = 1$ AND $u_r = 0.1$. HERE $p = q = 0.5$, $M = 2.5$, $\varepsilon = 1$ AND $\tau = 2 > \tau^*(u_r)$. NOTE THE KINK APPEARS AT THE TRANSITION FROM $u = 1$ TO $u < 1$.

In what follows we consider the two cases: $u_r > 0$, when only one degeneracy value is achieved, and $u_r = 0$ - the doubly degenerate case in succession. We start by analyzing the case of a single degeneracy.

2.3.1 The case $u_r > 0$ and $u_\ell = 1$

The numerical solution in Figure 2.6 suggests a discontinuity of the first order derivative of u. Consequently, the standard notion of TW is no longer valid. This requires an extended definition that allows for non-smooth waves, akin to the so-called sharp TW introduced in [32]. To define such waves we rewrite (2.0.1) as

$$\partial_t u + \partial_x F = 0, \quad \text{where} \quad F = f(u) - \varepsilon H(u)\partial_x p_c, \quad \text{and} \quad p_c = u + \varepsilon\tau\partial_t u. \tag{2.3.1}$$

In this way we define explicitly the flux F and the phase pressure difference p_c. In terms of the travelling wave coordinate η the equations of (2.3.1) becomes:

$$-su' + F' = 0, \quad F = f(u) - H(u)p_c', \quad \text{and} \quad p_c = u - \tau su'. \tag{2.3.2}$$

Physical arguments, as well as the numerical results presented above and in Section 2.4, suggest that u and F are continuous, whereas their derivatives may become discontinuous. In terms of the travelling wave coordinate η, we denote the point where u first becomes less than 1, and the derivative is discontinuous, by η_1. In view of this, we propose the following definition:

Definition 2.3.1 *Let $u_r > 0$, $s = \big(f(1) - f(u_r)\big)/(1 - u_r)$ and let $\eta_1 \in \mathbb{R}$ be an arbitrary fixed coordinate. A triple (u, F, p_c) is a sharp travelling wave solution to (2.1.8) if $u, F \in C(\mathbb{R})$, $u', F', p_c \in C(\mathbb{R} \setminus \{\eta_1\})$, u is decreasing, and*

$$\begin{cases} su' = F', \\ H(u)p_c' = f(u) - F, \\ p_c = u - \tau su', \end{cases} \tag{2.3.3}$$

for all $\eta > \eta_1$, whereas $u(\eta) = 1$, $F(\eta) = 1$, and $p_c(\eta) = 1$ for all $\eta < \eta_1$.

Since $u(\eta)$ is monotone and bounded below, it tends to a limit u_∞ as $\eta \to \infty$. Integrating ($2.3.3_1$) over (η_1, ∞) and using the continuity of u and F at η_1 lead to $F(\infty) = 1 + s(u_\infty - 1)$, whilst taking the limit in ($2.3.3_2$) yields $F(\infty) = f(u_\infty)$. In light of the definition of the wave speed s we conclude $u_\infty = u_r$. Therefore, the TW solution introduced in Definition 2.3.1 connects the two states $u_\ell = 1$ and $u_r > 0$.

Remark 2.3.1 *Definition 2.3.1 requires the continuity of u, but not of its derivative. As suggested by ($2.3.3_3$), p_c may become discontinuous at η_1. In physical terms, η_1 separates a fully saturated region, when only one phase is present, from a partially saturated one, containing both phases. Then for $\eta < \eta_1$ we have a fully saturated region, where $u = 1$ and $p_c = 1$, implying the same value for its left limit in η_1. In the unsaturated* *region $\eta > \eta_1$ we have $p_c = u - \tau s u'$, with 1 only as an upper bound, but not necessary as a right limit in η_1. However, this does not contradict the concept of the capillary pressure, since this is defined as the difference between the pressures inside the two phases. This definition only makes sense for partially saturated regions, where $0 < u < 1$. For equilibrium models, when $\tau = 0$, the value of p_c in the fully saturated region is defined by continuity. But this approach does not include any dynamic effects, which explains the eventual discontinuity of the capillary pressure at the boundaries of partially saturated regions.*

The existence of a travelling wave when $u_r > 0$ and $\tau > \tau^*(u_r)$, and hence $u_\ell = 1$, is established by the following theorem:

Theorem 2.3.1 *Let $u_r > 0$ and $\tau > \tau^*(u_r)$, implying $u_\ell = 1$. Then equation* (2.0.4) *admits a sharp TW solution in the sense of Definition 2.3.1.*

Proof To prove this theorem we proceed as in the previous section and start by seeking a positive solution of

$$\begin{cases} \tau s(u_r, 1) w w' + w = g(u; u_r, 1), & u \in (u_r, 1], \\ w(u_r) = 0, \end{cases} \tag{2.3.4}$$

in which (see also (2.1.4) and (2.1.7))

$$s(u_r, 1) = \frac{1 - f(u_r)}{1 - u_r} \qquad \text{and} \qquad g(u; u_r, 1) = \frac{s(u_r, 1)(u-1) - [f(u) - 1]}{H(u)}. \tag{2.3.5}$$

Then, as in (2.2.2), we use w to construct a TW solution u. Note that since $\tau > \tau^*(u_r)$, we will now have $w(u_\ell) > 0$.

In the following lemma we demonstrate that Problem (2.3.4) has a unique solution.

Lemma 2.3.1 *Let $\tau > \tau^*$ and $u_r > 0$. Then there exists a unique $w \in C^1([u_r, 1])$ such that $w(u) > 0$ on $(u_r, 1]$, which satisfies (2.3.4).*

Proof Existence is proved by means of the shooting method described in Theorem 2.2.2. We omit the details here.

For the uniqueness, we assume there are two different solutions w_1, w_2,

$$\frac{s\tau}{2}(w_i^2)' + w_i = g(u), \quad i = 1, 2. \tag{2.3.6}$$

Therefore,

$$\frac{s\tau}{2}\left(w_1^2 - w_2^2\right)' + (w_1 - w_2) = 0. \tag{2.3.7}$$

With $[w_1^2 - w_2^2]_+$ denoting the positive cut, multiplying $[w_1^2 - w_2^2]_+$ gives

$$\frac{s\tau}{4}\left([w_1^2 - w_2^2]_+^2\right)' + (w_1 - w_2)[w_1^2 - w_2^2]_+ = 0. \tag{2.3.8}$$

Integrating from u_r to u,

$$\frac{s\tau}{4}[w_1^2 - w_2^2]_+^2(u) + \int_{u_r}^{u}(w_1 - w_2)[w_1^2 - w_2^2]_+ du = 0. \tag{2.3.9}$$

Since $w_{1,2}$ are both non-negative as well, therefore $[w_1^2 - w_2^2]_+ = 0$, implying $w_1 \leq w_2$. Similarly, we have $w_2 \leq w_1$. Thus $w_1 = w_2$, which provides uniqueness. □

Since $w = -u'(\eta)$, it follows that $u \in (u_r, 1)$ is defined implicitly on the interval $(\eta_1, +\infty)$ by:

$$\eta(u) = \eta_1 + \int_u^1 \frac{dz}{w(z)}, \tag{2.3.10}$$

where we have used the fact that $u(\eta_1) = 1$.

Having obtained u, the other components F and p_c are obtained in a straightforward manner by intergrating the first and the third equation of (2.3.3). □

Remark 2.3.2 *Similar arguments can be used to extend the existence and uniqueness result of Lemma 2.3.1 to the case* $u_r = 0$*. However, in the next section we will consider an alternative approach based on regularization. Specifically, we take* $u_r = \delta$ *and investigate the limit as* $\delta \searrow 0$ *of the family of solutions* w_δ *of Problem (2.3.4) of which existence and uniqueness have been established in Lemma 2.3.1. This approach yields additional properties of* w *and* u*.*

In Theorem 2.3.1 we have constructed a sharp TW solution, using ODE arguments only. It is interesting to return to the simulation shown in Figure 2.6 in which a sharp TW solution appears as the limit of the solution $u(x,t)$ of a PDE. In Figure 2.7 we show the same graph side by side with the graph of the corresponding solution $w(u)$ of Problem (2.3.4).

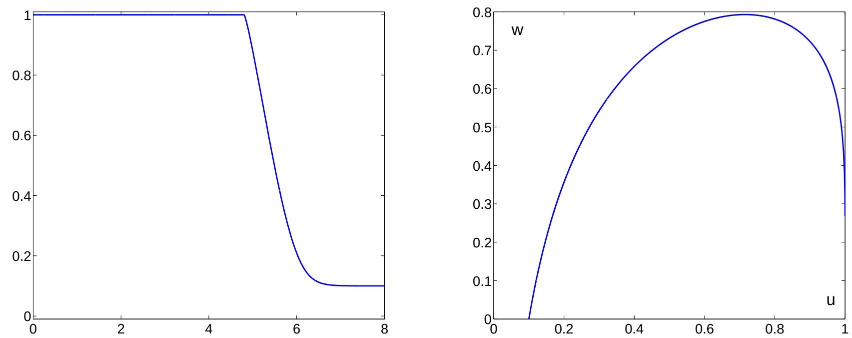


FIGURE 2.7: LEFT: THE NUMERICAL SOLUTION u OF (2.0.4) FOR $u_\ell = 1$ AND $u_r = 0.1$, YIELDING $u'(\eta_1) \approx -0.27$. RIGHT: THE CORRESPONDING w SOLVING (2.3.4), WITH $w(1) \approx 0.266$. THE CALCULATIONS ARE FOR $p = q = 0.5$, $M = 2.5$ AND $\tau = 2 > \tau^*(u_r)$.

In the left figure, the slope of u to the right of the kink was estimated numerically at -0.27, whilst in the right figure, $w(1)$ was estimated at 0.266. Because $w(1) = -u'(\eta(1)) = -u'(\eta_1)$, we can conclude that $w(1)$ agrees well with the slope of u at the kink.

2.3.2 The case $u_r = 0$ and $u_\ell = 1$

In this section we construct a sharp TW solution for the doubly degenerate problem when $u_r = 0$ and $u_\ell = 1$ and $\beta = 1$ (see (2.1.13)) and $\tau > \tau^*(0)$. As in the previous section we do this by constructing an appropriate solution w of Problem (2.3.4) as the limit of a sequence of solutions w_δ which connect $u_\ell = 1$ and $u_r = \delta$, as $\delta \to 0$. Specifically, let

$$s_\delta := \frac{1 - f(\delta)}{1 - \delta}, \quad g_\delta(u) := g(u; \delta, 1),$$

then w_δ is the solution of the problem

$$\begin{cases} \tau s_\delta w_\delta w_\delta' + w_\delta = g_\delta(u), & \delta < u \leq 1, \\ w_\delta(\delta) = 0. \end{cases} \tag{2.3.11}$$

Lemma 2.3.1 ensures its existence and uniqueness on the interval $[\delta, 1]$. Note that

$$s_\delta \searrow 1 \quad \text{and} \quad g_\delta(u) \nearrow g_0(u) = \frac{u - f(u)}{H(u)} \quad \text{as} \quad \delta \to 0, \tag{2.3.12}$$

where the convergence of $g_\delta(u) \to g_0(u)$ is pointwise on $(0, 1)$.

As seen in Remark 2.3.2, this result extends to the case $\delta = 0$, when $s = 1$ and the function $g_0(u) := g(u; 0, 1)$ has two singular points: $u = 0$ and $u = 1$. We denote this solution by w, i.e.,

$$\begin{cases} \tau w w' + w = g_0(u), \quad u \in (0, 1], \\ w(0) = 0. \end{cases} \tag{2.3.13}$$

We begin by establishing an ordering relation between w and w_δ.

Proposition 2.3.1 *Let $\tau > \tau^*(0)$, $\delta > 0$, while w and w_δ satisfy* (2.3.13) *and* (2.3.11) *respectively. Further let $\tilde{u} \in (0, 1)$ such that $w(\tilde{u}) = g(\tilde{u})$. We have:*

$$w_\delta(u) < w(u) \quad \textit{for all} \quad u \in (\delta, \tilde{u}). \tag{2.3.14}$$

Proof Assume that there exist a first point $u^* \in (\delta, \tilde{u})$ such that $w_\delta(u^*) = w(u^*)$. Since $w_\delta < w$ on (δ, u^*) it follows that $w'_\delta(u^*) \geq w'(u^*)$.

Inspection of the functions $g_\delta(u)$ and $g_0(u)$ shows that

$$\tau w(u^*) w'(u^*) + w(u^*) = g(u^*) > g_\delta(u^*) = \tau s_\delta w_\delta(u^*) w'_\delta(u^*) + w_\delta(u^*),$$

from which we conclude that $w'(u^*) > s_\delta w'_\delta(u^*)$.

Since $u^* \in (\delta, \tilde{u})$, it follows that $w'(u^*) \geq 0$ and $w'_\delta(u^*) \geq w'(u^*) \geq 0$. As $s_\delta > 1$, this implies that $w'(u^*) > w'_\delta(u^*)$, contradicting the previous inequality. □

As stated above, we consider the case $\tau > \tau^*(0)$, therefore the function w in (2.3.13) is defined on $[0, 1]$. Numerical results presented in Figure 2.5 suggest that τ^* increases with the right state u_r. Therefore taking $\tau > \tau^*(0)$ does not necessarily imply that $\tau > \tau^*(\delta)$ if $\delta > 0$. Therefore w_δ is only defined on $[\delta, c(\delta)]$ for some $c(\delta)$ defined by

$$c(\delta) = \sup\{\tilde{u} < u < 1 \mid w_\delta(u) > 0\}.$$

For practical reasons we extend w_δ by 0 on $[0, \delta]$, and on $[c(\delta), 1]$ if $c(\delta) < 1$, and investigate its behavior as $\delta \searrow 0$. We do so by considering two intervals, $[0, \tilde{u}]$ and $[\tilde{u}, 1]$.

Proposition 2.3.2 *Let w and w_δ solve (2.3.13) and (2.3.11). Along any sequence $\delta \to 0$, the functions w_δ converges point-wise to w on* $[0, \tilde{u}]$.

Proof . Integrating the equations in (2.3.13) and (2.3.11), we obtain

$$\frac{\tau}{2} w^2(u) + \int_0^u w(z)dz = \int_0^u g(z)dz,$$

and

$$\frac{s_\delta \tau}{2} w_\delta^2(u) + \int_\delta^u w_\delta(z)dz = \int_\delta^u g_\delta(z)dz.$$

We subtract the second equation from the first. Since $w > w_\delta$ on $(\delta, \tilde{u})$, we obtain

$$\begin{aligned} 0 \le \frac{\tau}{2}\{w^2(u) - w_\delta^2(u)\} &= \frac{\tau}{2}(s_\delta - 1)w_\delta^2(u) - \int_0^\delta w(z)\,dz \\ &+ \int_0^\delta g(z)dz - \int_\delta^u \{w(z) - w_\delta(z)\}dz + \int_\delta^u \{g(z) - g_\delta(z)\}\,dz. \end{aligned} \tag{2.3.15}$$

Applying the comparison from Proposition 2.3.1 enables us to simplify (2.3.15) to

$$0 \le \frac{\tau}{2}\{w^2(u) - w_\delta^2(u)\} \le \frac{\tau}{2}(s_\delta - 1)w_\delta^2(u) + \int_0^\delta g(z)\,dz + \int_\delta^u \{g(z) - g_\delta(z)\}dz, \tag{2.3.16}$$

Plainly, the first term on the right vanishes as $\delta \to 0$ because $s_\delta \searrow 1$ by (2.3.12), and the two integrals in (2.3.16) vanish since

$$\int_0^\delta g(z)dz = \int_0^\delta \{Mz^{-p} - (1 - z)^{-q}\}\,dz = \frac{M}{1-p}\delta^{1-p} + \frac{1}{1-p}\{(1-\delta)^{1-p} - 1\} \longrightarrow 0,$$

and

$$\int_\delta^u \{g(z) - g_\delta(z)\}\,dz \le C\delta^{1-p} \longrightarrow 0,$$

in which C is a positive constant.

Since w and w_δ are non-negative, we have established the point-wise convergence of w_δ towards w on the compact interval $[0, \tilde{u}]$. □

Now we consider the interval $[\tilde{u}, 1]$, where the following proposition holds.

Proposition 2.3.3 *Along any sequence $\delta \to 0$, w_δ converges point-wise to w on* $[\tilde{u}, 1]$.

Proof Let $\delta > 0$ and $u < c(\delta)$. Integrating (2.3.13) and (2.3.11) from $\tilde{u}$ to u, we have

$$\frac{\tau}{2}\Big(w^2(u) - w^2(\tilde{u})\Big) + \int_{\tilde{u}}^{u} w(z)dz = \int_{\tilde{u}}^{u} g(z)dz, \tag{2.3.17}$$

$$\frac{s_\delta \tau}{2}\Big(w_\delta^2(u) - w_\delta^2(\tilde{u})\Big) + \int_{\tilde{u}}^{u} w_\delta(z)dz = \int_{\tilde{u}}^{u} g_\delta(z)dz. \tag{2.3.18}$$

Subtracting (2.3.18) by (2.3.17), we have

$$\begin{aligned} &\frac{s_\delta \tau}{2}\Big(w_\delta^2(u) - w^2(u)\Big) \\ &= \int_{\tilde{u}}^{u} \Big(g_\delta(z) - g(z)\Big)dz - \int_{\tilde{u}}^{u} \Big(w_\delta(z) - w(z)\Big)dz + \frac{\tau}{2}\Big(s_\delta w_\delta^2(\tilde{u}) - w^2(\tilde{u})\Big) - \frac{\tau}{2}(s_\delta - 1)w^2(u) \\ &=: T_1 - T_2 + T_3 - T_4. \end{aligned} \tag{2.3.19}$$

By (2.3.12), T_1 vanishes as δ approaches 0. Furthermore, Proposition 2.3.2 gives the convergence of $w_\delta(\tilde{u})$ to $w(\tilde{u})$. Using (2.3.12) again, since w is bounded we obtain

$$T_4 = \frac{\tau}{2}(s_\delta - 1)w^2(u) \longrightarrow 0, \tag{2.3.20}$$

as well as

$$T_3 = \frac{\tau}{2}\left((s_\delta - 1)w_\delta^2(\tilde{u}) + \Big(w_\delta^2(\tilde{u}) - w^2(\tilde{u})\Big)\right) \longrightarrow 0. \tag{2.3.21}$$

Next, with $M := \max_{z\in[\tilde{u},u]} |w_\delta(z) - w(z)|$ one has

$$|T_2| = \left|\int_{\tilde{u}}^{u} \Big(w_\delta(z) - w(z)\Big)dz\right| \leq (u - \tilde{u})M. \tag{2.3.22}$$

Further, since w is decreasing on $[\tilde{u}, 1]$ we have

$$\max_{z\in[\tilde{u},u]} |w_\delta^2(z) - w^2(z)| \geq w(1)M.$$

Since $s_\delta \geq 1$, by (2.3.19) – (2.3.22)

$$\frac{\tau}{2}w(1)M \leq \frac{s_\delta \tau}{2}\Big(w_\delta^2(u) - w^2(u)\Big) \leq |T_1| + (u - \tilde{u})M + |T_3| + |T_4|, \tag{2.3.23}$$

Taking $u = c(\delta)$ and with δ small enough, from (2.3.23) we get

$$c(\delta) > \tilde{u} + \frac{\tau}{4}w(1).$$

Further, (2.3.23) also gives

$$\left(\frac{\tau}{2}w(1) - (u - \tilde{u})\right) M \leq |T_1| + |T_3| + |T_4|, \tag{2.3.24}$$

whenever δ is small enough. As $\delta \to 0$, all limits on the right side in (2.3.24) go to 0, which gives $w_\delta(u) \to w(u)$ pointwisely for $u \in [\tilde{u}, \tilde{u} + \frac{\tau}{4}w(1)]$. Let $\Delta u = \frac{\tau}{4}w(1)$. If $\tilde{u}+\Delta u \geq 1$, then the conclusion is shown. Otherwise, if $\tilde{u}+\Delta u < 1$, notice that Δu does not dependent on δ, therefore we can continue the same procedure for $u \in [\tilde{u} + \Delta u, \tilde{u} + 2\Delta u]$ and further until reaching 1. □

Together, Propositions 2.3.2 and 2.3.3, establish the following theorem:

Theorem 2.3.2 *Let $\tau > \tau^*(0)$, for any $\delta > 0$, w_δ solves (2.3.4) with $u_r = \delta$. Along any sequence $\delta \searrow 0$, the sequence $\{w_\delta\}$ approaches w solving (2.3.13). In particular, the limit w satisfies $w(1) > 0$.*

Remark 2.3.3 *Theorem 2.3.2 provides a selection criterium for the TW solution to (2.0.4), in the doubly degenerate case. Specifically, each w_δ solving (2.3.4) provides a TW solution to (2.0.4) connecting $u_\ell = 1$ to $u_r = \delta$. Letting $\delta \searrow 0$, we have seen that the limit w provides a TW solution to (2.0.4) which connects $u_\ell = 1$ to $u_r = 0$, i.e. we can view this sharp TW solution as the limit of 'regular' TW solutions each having a kink at $u = 1$.*

The selection criterion introduced above is obtained after giving up the condition $w(1) = 0$. However, one can consider a symmetric approach, namely to solve the backward problem

$$\begin{cases} \tau w w' + w = g_1(u), & u \in (u_0, 1), \\ w(1) = 0, \end{cases} \tag{2.3.25}$$

with $u_0 > 0$ being the infimum of the interval where $w > 0$, while $g_1(u) = g(u; u_0, 1)$ is defined in (2.3.5). If $u_0 = 0$, this solution w allows defining a travelling wave u connecting $u_\ell = 1$ to $u_r = 0$, but having now a kink at the transition from $u > 0$ to $u = 0$, and being continuously differentiable everywhere else.

Theorem 2.3.3 below rules out this possibility. Before stating it we observe that since g_1 becomes unbounded in $u = 1$, one can use a regularization argument again and view w as the limit along a sequence $\delta \searrow 0$ of the functions w_δ satisfying the regularized backward problem

$$\begin{cases} \tau s_\delta w_\delta w_\delta' + w_\delta = g_\delta, \quad u \in (0, 1-\delta), \\ w_\delta(1-\delta) = 0. \end{cases} \tag{2.3.26}$$

Here we consider $u_r = 0$, while

$$s_\delta := \frac{f(1-\delta)}{1-\delta}, \quad \text{and} \quad g_\delta(u) := g(u; 0, 1-\delta).$$

As before, (2.3.26) is solved backward as long as w_δ remains positive.

In case such solutions exist, their limit w would satisfy

$$\begin{cases} \tau w w' + w = g_0(u), \quad u \in (0, 1), \\ w(0) = w_0, w(1) = 0, \end{cases} \tag{2.3.27}$$

for a properly chosen $w_0 \geq 0$. Moreover, another selection criterion for the sharp travelling waves connecting $u_\ell = 1$ to $u_r = 0$ could then be defined. However, this possibility is again ruled out by Theorem 2.3.3 below, implying that $w(1) > 0$ for any initial condition $w(0) = w_0 \geq 0$, so (2.3.27) has no solution.

Theorem 2.3.3 *Let $u_r = 0$, $\tau > \tau^*(0)$, while w and $\widetilde{w}$ solve (2.3.25) with initial data w_0 and $\widetilde{w}_0$. If $\widetilde{w}_0 > w_0 \geq 0$, then $\widetilde{w} > w$ for all $u \in [0, 1]$.*

Proof . Assume w and $\widetilde{w}$ intersect. Let $\overline{u}$ be the smallest intersection point. Since $\widetilde{w}_0 > w_0$, we know $\overline{u} > 0$. We distinguish two cases:

Case 1: $\overline{u} < 1$, from

$$w'(\overline{u}) = \frac{g(\overline{u}) - w(\overline{u})}{\tau w(\overline{u})}, \quad \text{and} \quad \widetilde{w}'(\overline{u}) = \frac{g(\overline{u}) - \widetilde{w}(\overline{u})}{\tau \widetilde{w}(\overline{u})},$$

we obtain

$$w'(\overline{u}) = \widetilde{w}'(\overline{u}).$$

(i) If $w'(\overline{u}) \geq 0$, then $w'(\overline{u}+\delta) < \widetilde{w}'(\overline{u}+\delta)$, $w(\overline{u}+\delta) < \widetilde{w}(\overline{u}+\delta)$ and $g(\overline{u}+\delta) > 0$ for $\delta > 0$ small enough. However,

$$w'(\overline{u}+\delta) = \frac{g(\overline{u}+\delta)}{\tau w(\overline{u}+\delta)} - \frac{1}{\tau} > \frac{g(\overline{u}+\delta)}{\tau \widetilde{w}(\overline{u}+\delta)} - \frac{1}{\tau} = \widetilde{w}'(\overline{u}+\delta),$$

which is a contradiction.
(ii) If $w'(\overline{u}) < 0$, then $w'(\overline{u}+\delta) < \widetilde{w}'(\overline{u}+\delta)$ and $w(\overline{u}+\delta) < \widetilde{w}(\overline{u}+\delta)$ for $\delta > 0$ small enough. If $g(\overline{u}) \geq 0$, then

$$w'(\overline{u}+\delta) = \frac{g(\overline{u}+\delta)}{\tau w(\overline{u}+\delta)} - \frac{1}{\tau} \geq \frac{g(\overline{u}+\delta)}{\tau \widetilde{w}(\overline{u}+\delta)} - \frac{1}{\tau} = \widetilde{w}'(\overline{u}+\delta),$$

which is a contradiction again. If $g(\overline{u}) < 0$, we know $w'(\overline{u}-\delta) > \widetilde{w}'(\overline{u}-\delta)$ and $w(\overline{u}-\delta) < \widetilde{w}(\overline{u}-\delta)$ for $\delta > 0$ small enough. However

$$w'(\overline{u}-\delta) = \frac{g(\overline{u}-\delta)}{\tau w(\overline{u}-\delta)} - \frac{1}{\tau} < \frac{g(\overline{u}-\delta)}{\tau \widetilde{w}(\overline{u}-\delta)} - \frac{1}{\tau} = \widetilde{w}'(\overline{u}-\delta),$$

contradicting the inequalities above.

Case 2: $\overline{u} = 1$, there exists $u_0 < 1$ close enough to 1 such that $w(u_0) < \widetilde{w}(u_0)$ and $w'(u_0) > \widetilde{w}'(u_0)$. note that $w'(u)$, $\widetilde{w}'(u)$ and $g(u)$ are negative when u is close enough to 1. But

$$w'(u_0) = \frac{g(u_0)}{\tau w(u_0)} - \frac{1}{\tau} < \frac{g(u_0)}{\tau \widetilde{w}(u_0)} - \frac{1}{\tau} = \widetilde{w}'(u_0), \textit{contradiction},$$

meaning $\widetilde{w} > w$ for all $u \in [0, 1]$. □

2.3.3 Regularity and compact support

As we have seen, u has a kink at the transition from $u = 1$ to $u < 1$. In what follows we study the transition to the other degenerate value $u = 0$.

Theorem 2.3.4 *Let $u_r = 0$ and $u_\ell = 1$. If $\tau > \tau^*(0)$, then the travelling wave selected by Theorem 2.3.2 vanishes at a finite $\eta_0 \in \mathbb{R}$, and $u'(\eta_0) = 0$.*

Proof By Theorem 2.3.2, u is constructed from the limit $w = \lim_{\delta \searrow 0} w_\delta$. Integrating (2.3.11) from δ to $u > \delta$ yields

$$\frac{s_\delta \tau}{2} w_\delta(u)^2 + \int_\delta^u w_\delta(v)dv = \int_\delta^u g_\delta(v)dv \leq \int_\delta^u g_0(v)dv \leq \int_0^u g_0(v)dv.$$

From the asymptotic behavior of the function $g_0(u)$ as $u \searrow 0$, we deduce the upper bound

$$w_\delta(u)^2 \leq \frac{2}{s_\delta \tau} \int_0^u g_0(v)dv \leq \frac{2}{\tau} \int_0^u g_0(v)dv \leq \frac{2}{\tau} C_1 u^{1-p},$$

where C_1 is a positive constant which does not depend on δ. Therefore, with $C_2 = \sqrt{2C_1/\tau}$, we have $w_\delta(u) \le C_2 u^{\frac{1-p}{2}}$ and hence

$$w(u) \le C_2 u^{\frac{1-p}{2}}. \tag{2.3.28}$$

Next, we derive a lower bound for $w(u)$. We integrate equation (2.3.13) over $(0, u)$ to obtain

$$\frac{\tau}{2} w(u)^2 + \int_0^u w(v)dv = \int_0^u g(v)dv,$$

and use the upper bound (2.3.28) for $w(u)$ and the lower bound $g(u) \ge C_3 u^{-p}$ for some $C_3 > 0$ and u small enough to reduce this to

$$\frac{\tau}{2} w(u)^2 + \frac{2C_2}{3-p} u^{\frac{3-p}{2}} \ge \frac{C_3}{1-p} u^{1-p}. \tag{2.3.29}$$

Because for u sufficiently small,

$$\frac{2C_2}{3-p} u^{\frac{3-p}{2}} < \frac{C_3}{2(1-p)} u^{1-p},$$

the second term on the left in (2.3.29) can be absorbed in the term on the right, so that

$$\frac{\tau}{2} w(u)^2 \ge \frac{C_3}{2(1-p)} u^{1-p},$$

and we arrive at the desired lower bound

$$w(u) \ge C_4\, u^{\frac{1-p}{2}}, \qquad \text{where} \qquad C_4 = \sqrt{\frac{C_3}{\tau(1-p)}}, \tag{2.3.30}$$

for u small enough.

Returning to the variables u and η we conclude from (2.3.30) that

$$-u'(\eta) \ge C_4 u^{\frac{1-p}{2}}(\eta), \qquad \text{for} \qquad 0 < u < u^*, \tag{2.3.31}$$

for u^* small enough. Suppose that $u^* = u(\eta^*)$. Then, when we integrate this inequality over (η_0, η) we obtain

$$u(\eta) \le \left\{ (u^*)^{\frac{p+1}{2}} - \frac{(p+1)C_4}{2}(\eta - \eta^*) \right\}^{\frac{2}{1+p}}, \qquad \text{as long as} \quad u(\eta^*) \ge 0. \tag{2.3.32}$$

It follows from (2.3.32) that $u(\eta)$ vanishes at some point $\eta_0 < \infty$ and from the upper bound (2.3.28) for $w(u) = -u'(\eta)$ that $u'(\eta_0) = 0$. □

2.4 Numerical results

In this section, we provide some numerical experiments. We solve the full problem (2.0.4), using a semi-implicit Euler finite volume scheme. This scheme is similar to the ones investigated in [4], [8], or [18]. There a particular attention is paid to heterogeneities and the conditions at the interface between two homogeneous sub-domains. We also mention [29] for a review of different numerical methods for pseudo-parabolic equations.

We consider the problem (2.0.4) in the domain $S = R \times R^+$:

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = \varepsilon \frac{\partial}{\partial x} \left\{ H(u) \left(\frac{\partial u}{\partial x} + \varepsilon\tau \frac{\partial^2 u}{\partial x \partial t} \right) \right\}, \tag{2.4.1}$$

with initial value

$$u(x,0) = (u_B - u_r)\widetilde{H}(-x) + u_r, \tag{2.4.2}$$

where u_r is the right state, u_B is the inflow value and $\widetilde{H}(x)$ is a smooth monotone approximation of the Heaviside function H. By using $\widetilde{H}$ instead of H we avoid unnecessary technical difficulties due to discontinuities in the initial conditions. As shown in [7], if the initial data has jumps, these will persist for all $t > 0$, at the same location. This would require an adapted and more complicated numerical approach for ensuring the continuity in flux and pressure (see for example [4], Chapter 3, or [8]).

Remark 2.4.1 *We emphasize that u_B is an inflow value, which in general is not equal to the value associated to τ, $u_\ell = \bar{u}(\tau)$. This value will be an outcome of the calculations.*

Since the scaling

$$x \to \frac{x}{\varepsilon}, \quad t \to \frac{t}{\varepsilon}, \tag{2.4.3}$$

removes the parameter ε from (2.4.1), we fix $\varepsilon = 1$ here. In the absence of analytic solutions, for verifying the numerical solution we recall the transformation $w(u) = -u'(\eta(u))$, based on which a relation between τ and an admissible left state u_ℓ can be established.

As shown in Section 2.2, given $u_r \geq 0$, a value $\tau^* \in (0, \infty]$ exists such that to any $\tau < \tau^*$ a unique left state $u_\ell = u_\ell(\tau) \leq 1$ can be associated. This left state can be connected to u_r through a smooth TW solution to (2.4.1). Whenever $\tau^* < \infty$, if $\tau > \tau^*$ no smooth travelling waves are possible, but sharp ones connecting $u_\ell = 1$ to u_r, and having a kink at the point when u becomes less than 1. Figure 2.8 below presents the diagrams u_ℓ - τ for $p = q = 0.5$, $M = 2.5$, and for two values of u_r: $u_r = 0.1$ (non-degenerate), and $u_r = 0$ (degenerate). To obtain these diagrams we have solved (2.2.1) numerically with fixed u_r, but for several left states u_ℓ, providing different pairs (w, τ) such that $w(u_\ell) = 0$. We start with $u_\ell = \alpha$, which is defined in (2.1.12). In terms of hyperbolic conservation laws, the shock $\{\alpha, u_r\}$ is an admissible entropy solution to the non-viscous (BL) equation (obtained for $\varepsilon = 0$). We have $\alpha \approx 0.926$ if $u_r = 0.1$, respectively $\alpha \approx 0.936$ if $u_r = 0$. Starting with $u_\ell = \alpha$, for which a lower value $\tau = \tau_*$ is obtained, we increase u_ℓ by a small Δu_ℓ (in this case 5×10^{-4}) and determine the corresponding τ value either until $u_\ell = 1$ (yielding a finite upper limit τ^* to τ), or up to a maximal value less than one, which is attained asymptotically as $\tau \nearrow \infty$. The pairs (u_ℓ, τ) obtained in this way are included in the diagram.

Both cases considered here give $\tau^* < \infty$: $\tau^* \approx 1.37$ for $u_r = 0.1$ and $\tau^* \approx 0.22$ for $u_r = 0$. For the lower limits we get $\tau_* \approx 0.067$, respectively $\tau_* \approx 0.054$. Below we will present numerical solutions to (2.4.1) for two values of τ: $\tau_1 = 0.1$, and $\tau_2 = 2$. For both right states u_r mentioned above they satisfy $\tau_* < \tau_1 < \tau^* < \tau_2$. As resulting from the diagrams, $\tau_1 = 0.1$ is associated to the left state $u_\ell = 0.9475$ if $u_r = 0.1$, respectively to $u_\ell = 0.977$ if $u_r = 0$.

To discretize (2.4.1) we take a fixed time step $\Delta t = t_{n+1} - t_n$ and apply a semi-implict first order method:

$$\frac{u^{n+1} - u^n}{\Delta t} + \frac{d}{dx} F^n(u) = 0.$$

Here $F^n(u)$ is the time discrete flux function at $t = t_n$,

$$F^n(u) := f^n(u) - H^n(u)\left(\partial_x u^{n+1} + \tau \frac{\partial_x u^{n+1} - \partial_x u^n}{\Delta t}\right).$$

Similarly, the functions f^n and H^n are time discrete variants of f and H. These are not defined explicitly since we are interested here only in their fully discrete counterparts. Note that the scheme is explicit in the convective terms, and semi-implicit in the higher

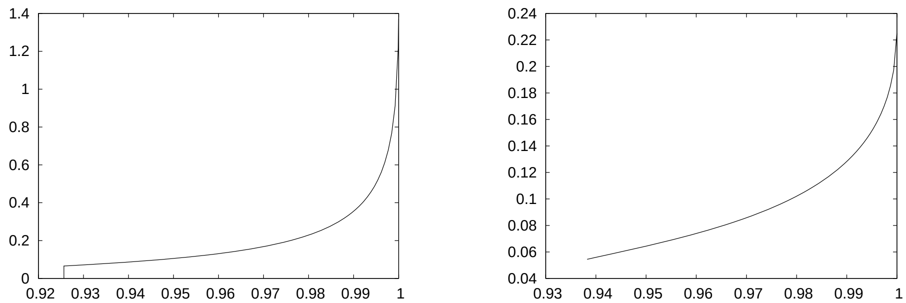


Figure 2.8: The diagrams u_ℓ - τ computed for $p = q = 0.5$, $M = 2.5$ and with $u_r = 0.1$ (left) respectively $u_r = 0$ (right). Numerically we obtain $\tau^* \approx 1.37$, respectively $\tau^* \approx 0.22$.

order ones.

For the space discretization, we use a finite volume scheme on a dual mesh. Taking a uniform grid with mesh size $\Delta x = x_n - x_{n-1}$ and defining $u_i = \frac{1}{\Delta x}\int_{i-1/2}^{i+1/2} u(x)dx$, the fully discretized equation becomes

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} + \frac{F^n(u_i, u_{i+1}) - F^n(u_{i-1}, u_i)}{\Delta x} = 0. \tag{2.4.4}$$

Here the numerical flux $F^n(u_i, u_{i+1})$ is defined by

$$F^n(u_i, u_{i+1}) = f(u_i^n) - H_{i+1/2}^n \frac{u_{i+1}^{n+1} - u_i^{n+1}}{\Delta x} - \tau H_{i+1/2}^n \frac{u_{i+1}^{n+1} - u_i^{n+1} - u_{i+1}^n + u_i^n}{\Delta x \Delta t}.$$

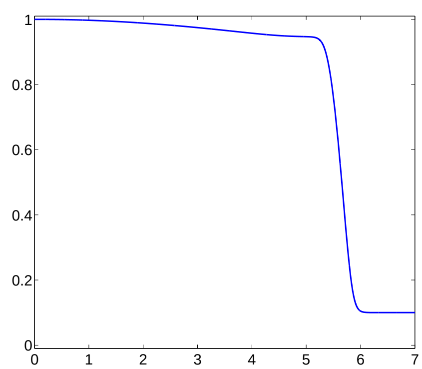
For the coefficient $H_{i+1/2}^n$, we use the arithmetic average value:

$$H_{i+1/2}^n = \frac{1}{2}\Big(H(u_i^n) + H(u_{i+1}^n)\Big).$$

This approach is important when doing calculations with degenerate outflow value, $u_r = 0$. The numerical diffusion added in this way has regularizing effects, leading to a numerical solution fulfilling the selection criterion in Remark 2.3.3.

In what follows we present the numerical solutions of (2.4.1) obtained on a spatial interval $(-1, 19)$ and at time $T = 5$. As mentioned above, we take $p = q = 0.5$, $M = 2.5$, and consider two right states, $u_r = 0.1$ and $u_r = 0$, as well as two values for τ: $\tau_1 = 0.1$ and $\tau_2 = 2$. The discretization parameters are $\Delta x = 5 \times 10^{-4}$ and $\Delta t = 10^{-4}$, providing stable numerical results. On the endpoints of the interval we take value that are compatible with the ones appearing in (2.4.2): u_B at the inflow, and u_r at the outflow. All calculations are carried out for $u_B = 1$, which is not necessarily equal to the value u_ℓ related to τ. Therefore the numerical solution of the degenerate pseudo-parabolic problem (2.4.1)–(2.4.2) does not necessarily have a TW profile, but instead will feature a "plateau" region of constant value $\bar{u}$ corresponding to u_ℓ related to τ.

The solutions presented in Figure 2.9, computed for $\tau_1 = 0.1$ are clearly presenting this situation: they both decay from 1 to the plateau value $u = \bar{u} < 1$. This value is taken over an interval that is delimited on the right by a front going down from $\bar{u}$ to u_r. This front travels with a constant speed provided by the RH condition in (2.1.4), written for the states $\bar{u}$ and u_r. A similar situation is obtained in [11] for the non-degenerate case $H = 1$. As in that paper, we associate the plateau value $\bar{u}$ with the value $u_\ell = \bar{u}(\tau)$. The



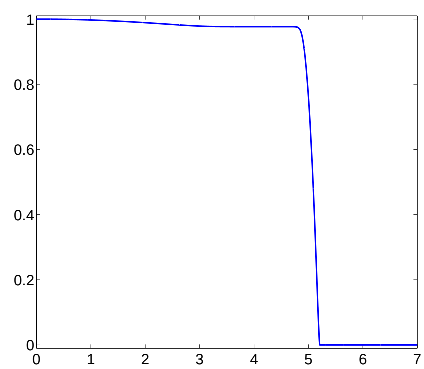


Figure 2.9: Left: graph of u for $u_r = 0.1$ and $\tau = 0.1 < \tau^*$, containing a plateau at $\bar{u} = 0.9467$. Right: graph of u for $u_r = 0$, and the plateau value $\bar{u} = 0.979$.

plateau value $\bar{u}$ exhibited by the numerical solution is $\bar{u} = 0.9467$ for $u_r = 0.1$, whereas $\bar{u} = 0.979$ for $u_r = 1$. This agrees well with the value $u_\ell = \bar{u}(\tau)$ predicted at $\tau = 0.1$ by the u_ℓ - τ diagrams discussed above. There we obtained $u_\ell = 0.9475$ if $u_r = 0.1$, respectively to $u_\ell = 0.977$ if $u_r = 0$.

The next numerical results are obtained for $\tau_2 = 2$, exceeding τ^* up to which smooth travelling waves are possible. Therefore the u_ℓ - τ diagrams are not providing any information that can be used for testing the numerical solutions. However, as discussed in

Section 2.3, waves connecting the left state $u_\ell = 1$ to u_r are still possible, but these have a discontinuous derivative (kink) at the transition point from $u = 1$ to $u < 1$. Correspondingly, the transformed w solving (2.3.4) on $(u_r, 1]$ will remain strictly positive at $u = 1$. The value $w(1)$ gives the slope of u at the right of the kink. In this case we compare this (numerical) slope to $w(1) = -u'(\eta_1 + 0)$.

The left pictures in Figures 2.10 and 2.11 are presenting the numerical results for $u_r = 0.1$, respectively $u_r = 0$. The kinks encountered at the transition from $u = 1$ to $u < 1$ are estimated to -0.27 for $u_r = 0.1$, and to -1.27 for $u_r = 0$. For w we obtain $w(1) = 0.266$ in the first case, and $w(1) = 1.266$ in the second one. The two functions $w's$ are presented in the right pictures of Figures 2.10 and 2.11.

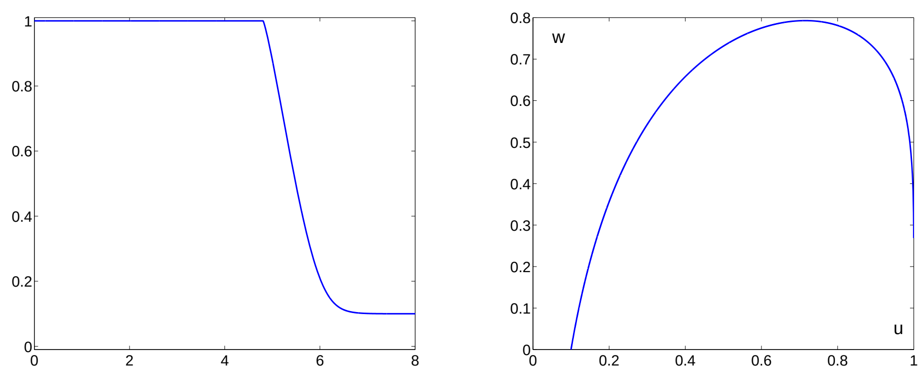


FIGURE 2.10: THE GRAPH OF u FOR $u_r = 0.1$, AND $\tau = 2 > \tau^*$, PRESENTING A KINK AT THE TRANSITION $u = 1$ TO $u < 1$ (LEFT); THE SLOPE AT THE RIGHT OF THE KINK IS $u' = -0.27$. THE CORRESPONDING w (RIGHT), WHERE $w(1) = 0.266$.

Finally, we recall that in the doubly degenerate case $u_\ell = 1$ and $u_r = 0$ the sharp waves are not unique. Theorem 2.3.2 provides a selection criterion. As follows from Theorem 2.2.8, this particular sharp wave is smooth everywhere away from the transition from $u = 1$ to $u < 1$. The smoothness includes the transition $u > 0$ to $u = 0$, which is achieved for a finite η_0. The same is featured by the numerical solution: Figure 2.12 presents two zoomed views of it. We clearly see a kink also in the left picture, whereas the transition to $u = 0$ is smooth, as displayed in the right picture.

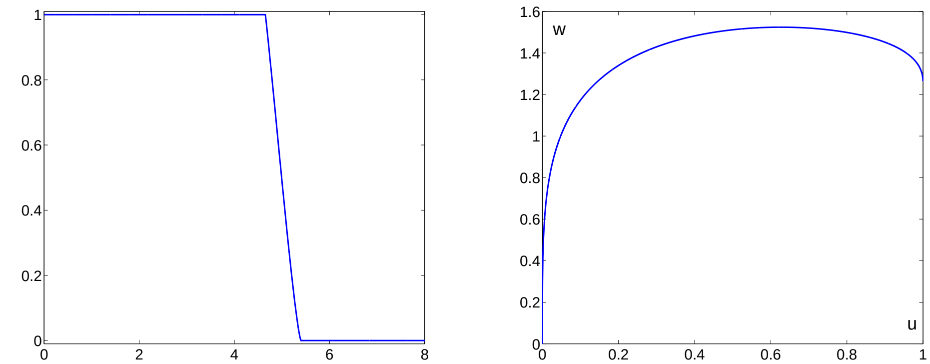


FIGURE 2.11: THE GRAPH OF u FOR $u_r = 0$, AND $\tau = 2 > \tau^*$, PRESENTING A KINK AT THE TRANSITION $u = 1$ TO $u < 1$ (LEFT); THE SLOPE AT THE RIGHT OF THE KINK IS $u' = -1.27$. THE CORRESPONDING w (RIGHT) WHERE $w(1) = 1.26$.

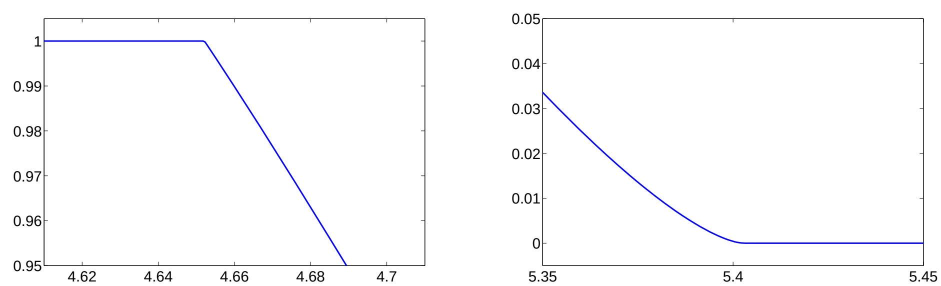


FIGURE 2.12: ZOOMED VIEW OF u FOR $u_r = 0$, $\tau = 2 > \tau^*$: A KINK APPEARS AT THE TRANSITION TO $u < 1$ (LEFT), WHEREAS THE TRANSITION TO $u = 0$ IS SMOOTH (RIGHT).

References

[1] P.W. Bates, P.C. Fife, X. Ren, X. Wang, *Traveling waves in a convolution model for phase transitions*, Arch. Rational Mech. Anal. **138** (1997), 105-136.

[2] N. Bedjaoui, P.G. LeFloch, *Diffusive-dispersive traveling waves and kinetic relations. I. Nonconvex hyperbolic conservation laws*, J. Differential Equations **178** (2002), 574-607.

[3] A.L. Bertozzi, A. Münch and M. Shearer, *Undercompressive shocks in thin film flows*, Phys. D **134** (1999), 431-464.

[4] C. Cuesta, *Pseudo-parabolic equations with driving convection term*, Ph.D. thesis, VU Amsterdam, The Netherlands, 2003.

[5] C. Cancès, C. Choquet, Y. Fan and I.S. Pop, *Existence of weak solutions to a degenerate pseudo-parabolic equation modeling two-phase flow in porous media*, CASA Report **10-75**, Eindhoven University of Technology, 2010.

[6] C. Cuesta, C. J. van Duijn, and J. Hulshof, *Infiltration in porous media with dynamic capillary pressure: Travelling waves*, European J. Appl. Math. **11** (2000), 381-397.

[7] C. Cuesta and J. Hulshof, *A model problem for groundwater flow with dynamic capillary pressure: stability of travelling waves*, Nonlinear Anal. **52** (2003), 1199-1218.

[8] C.M. Cuesta, I.S. Pop, *Numerical schemes for a pseudo-parabolic Burgers equation: discontinuous data and long-time behaviour*, J. Comput. Appl. Math. **224** (2009), 269-283.

[9] L. Cueto-Felgueroso, R. Juanes, *Stability analysis of a phase-field model of gravity-driven unsaturated flow through porous media*, Phys. Rev. E **79** (2009), 036301.

[10] D.A. DiCarlo, *Experimental measurements of saturation overshoot on infiltration*, Water Resour. Res. **40** (2004), W04215.1-W04215.

[11] C.J. van Duijn, L.A. Peletier and I.S. Pop, *A new class of entropy solutions of the Buckley-Leverett equation*, SIAM J. Math. Anal. **39** (2007), 507-536.

[12] W. Düll, *Some qualitative properties of solutions to a pseudoparabolic equation modeling solvent uptake in polymeric solids*, Comm. Partial Differential Equations **31** (2006), 1117-1138.

[13] A.G. Egorov, R.Z. Dautov, J.L. Nieber and A.Y. Sheshukov, *Stability analysis of gravity-driven infiltrating flow*, Water Resour. Res. **39** (2003), 1266.

[14] Y. Fan and I.S. Pop, *A class of degenerate pseudo-parabolic equation": existence, uniqueness of weak solutions, and error estimates for the Euler-implicit discretization*, Mathematical Methods in the Applied Sciences **34** (2011), 2329-2339.

[15] S.M. Hassanizadeh M. Celia and H. Dahle, *Dynamic effect in the capillary pressure-saturation relationship and its impact on unsaturated flow*, Vadose Zone J. **1** (2002), 38-57.

[16] S.M. Hassanizadeh and W.G. Gray, *Thermodynamic basis of capillary pressure in porous media*, Water Resour. Res. **29** (1993), 3389-3405.

[17] B.T. Hayes, P.G. LeFloch, *Nonclassical shocks and kinetic relations: scalar conservation laws*, Arch. Ration. Mech. Anal. **139** (1997), 1-56.

[18] R. Helmig, A. Weiss, B. Wohlmuth, *Dynamic capillary effects in heterogeneous porous media*, Comput. Geosci. **11** (2007), 261-274.

[19] D. Jacobs, B. McKinney, M. Shearer, *Travelling wave solutions of the modified Korteweg de Vries-Burgers equation*, J. Differential Equations **116** (1995), 448-467.

[20] J.R. King, C.M. Cuesta, *Small and waiting-time behavior of a Darcy flow model with a dynamic pressure saturation relation*, SIAM J. Appl. Math. **66** (2006), 1482-1511.

[21] F. Kissling, P.G. LeFloch, C. Rohde, *A kinetic decomposition for singular limits of non-local conservation laws*, J. Differential Equations (2009), doi:10.1016/j.jde.2009.05.006

[22] P.G. LeFloch, *Hyperbolic systems of conservation laws–the theory of classical and nonclassical shock waves*, Lectures Math. ETH Zürich, Birkhäu ser Verlag, Basel, 2002.

[23] L. Malaguti, C. Marcelli, *Travelling wavefronts in reaction-diffusion equations with convection effects and non-regular terms*, Math. Nachr. **242** (2002), 148-164.

[24] L. Malaguti, C. Marcelli, *Sharp profiles in degenerate and doubly degenerate Fisher-KPP equations*, J. Differential Equations **195** (2003), 471-496.

[25] S. Manthey, S.M. Hassanizadeh, R. Helmig, R. Hilfer, *Dimensional analysis of two-phase flow including a rate-dependent capillary pressure-saturation relationship*, Adv. Water Res. **31** (2008), 1137-1150.

[26] A. Mikelić, *A global existence result for the equations describing unsatured flow in porous media with dynamic capillary pressure*, submitted.

[27] A. Mikelić, H. Bruining, *Analysis of model equations for stress-enhanced diffusion in coal layers. I. Existence of a weak solution*, SIAM J. Math. Anal. **40** (2008), 1671-1691.

[28] J.L. Nieber, R.Z. Dautov, A.G. Egorov and A.Y. Sheshukov, *Dynamic capillary pressure mechanism for instability in gravity-driven flows; review and extension to very dry conditions*, in Transp. Porous Med. **58** (2005), 147-172.

[29] M. Peszyńska, S. Yi, *Numerical methods for unsaturated flow with dynamic capillary pressure in heterogeneous porous media*. Int. J. Numer. Anal. Model. **5** (2008), 126-149.

[30] M. Ptashnyk, *Degenerate quaslinear pseudoparabolic equations with memory terms and variational inequalities*, Nonlinear Anal. **66** (2007), 2653-2675.

[31] C. Rohde, *Scalar conservation laws with mixed local and non-local diffusion-dispersion terms*, SIAM J. Math. Anal. **37** (2005), 103-129.

[32] F. Sánchez-Garduño, P.K. Maini, *Existence and uniqueness of a sharp travelling wave in degenerate non-linear diffusion Fisher-KPP equation*, J. Math. Biol. **33** (1994), 163-192.

Chapter 3

Existence, uniqueness of weak solutions to a simplified model

In this chapter, we focus on the following pseudo-parabolic equation:

$$u_t + \nabla \cdot \mathbf{F}(u) = \nabla \cdot \Big(H(u)\nabla u\Big) + \tau \Delta u_t. \tag{3.0.1}$$

This equation is motivated by the model of a two-phase porous media flow, where dynamic effects are taken into account in the phase pressure difference. The corresponding equation is proposed in [12],

$$u_t + \nabla \cdot \mathbf{F}(u) = \nabla \cdot \Big(H(u)\nabla p\Big), \tag{3.0.2}$$

with $p = p_c(u) + \tau\partial_t u$. Here u stands for the water saturation, $\mathbf{F}$ and H are the water fractional flow function and the capillary induced diffusion function. The difference between (3.0.1) and (3.0.2) is that H only appears in the second order term in (3.0.1).

This chapter has appeared as a paper at Mathematical Methods in the Applied Sciences (Fan and Pop, 2011)

In this case no sign restriction need to be imposed on H. We study the existence and uniqueness of weak solutions to (3.0.1) in this chapter, complemented with initial and boundary conditions. We do so by applying a discretization in time, for which we also give error estimates.

Pseudo-parabolic equations arise in many real life applications such as radiation with time delay [16], seepage in fissured rocks [3], heat conduction models [26] and models for lightning propagation [2], etc. Existence and uniqueness of weak solutions to nonlinear pseudo-parabolic equations are proved in [20], while the existence of weak solutions for degenerate cases is studied in [4, 17, 18]. A nonlinear parabolic-Sobolev-type equation is studied in [28], and the homogenization of a closely related pseudoparabolic system is considered in [21]. Travelling wave solutions and their relation to non-standard shock solutions to hyperbolic conservation laws are investigated in [5, 8] for linear higher order terms. This analysis is pursued in [7] for degenerate situations. Numerical schemes for dynamic capillary effects in heterogeneous porous media are given in [13]. The case of discontinuous initial data is analyzed in [6]. The super-convergence of a finite element approximation to similar equation is investigated in [1] and time-stepping Galerkin methods are analyzed in [10] and [11], where two difference-approximation schemes are considered. In [25], Fourier spectral methods for pesudo-parabolic equations are analyzed.

To investigate the equation (3.0.1), we make the following assumptions:

- (A1) Ω is an open, bounded and connected domain in $\mathbb{R}^d$, with Lipschitz continuous boundary. With $T > 0$ given, we denote $Q = \Omega \times (0, T]$.
- (A2) $\tau > 0$ is a given number.
- (A3) The functions $\mathbf{F}$ and H are Lipschitz continuous and bounded, $|\mathbf{F}| \leq M$, $|H| \leq M$ for some $M > 0$. Moreover, L is an upper bound for the Lipschitz constants of $\mathbf{F}$ and H.

Remark 3.0.2 *Observe that since the term involving the mixed derivative is linear, no assumption on the sign of H is needed, enlarging the context that is usually encountered in physical models, where $H \geq 0$. We also extend the context for $\boldsymbol{F}$ appeared in engineer* *literature (where $\boldsymbol{F} = \boldsymbol{v} f(u)$ with $\boldsymbol{v}$ a divergence free vector field) to more general cases: bounded Lipschitz continuous functions.*

In this chapter, $L^2(\Omega)$ stands for the square Lebesgue integrable functions on Ω , $W^{1,2}(\Omega)$ requests the same also for the derivatives of first order. $W_0^{1,2}(\Omega)$ is a subset of $W^{1,2}(\Omega)$ whose elements have zero boundary values (in the trace sense), and $W^{-1,2}(\Omega)$ is the dual space of $W_0^{1,2}(\Omega)$. Besides, C denotes a generic positive number.

The initial and boundary conditions of (3.0.1) are given as follows:

$$u(\cdot, 0) = u^0, \quad \text{and} \quad u|_{\partial\Omega} = 0, \tag{3.0.3}$$

where $u^0 \in W_0^{1,2}(\Omega)$. We seek a solution to the following:

Problem P Find $u \in W^{1,2}(0, T; W_0^{1,2}(\Omega))$ such that

$$\begin{aligned} &\int_0^T \int_\Omega u_t \phi dxdt - \int_0^T \int_\Omega \mathbf{F}(u) \cdot \nabla\phi dxdt \\ &+ \int_0^T \int_\Omega H(u) \nabla u \cdot \nabla\phi dxdt + \tau \int_0^T \int_\Omega \nabla u_t \cdot \nabla\phi dxdt = 0, \end{aligned} \tag{3.0.4}$$

for any $\phi \in L^2(0, T; W_0^{1,2}(\Omega))$.

The results of this chapter are summarized in the theorem below:

Theorem: Assuming (A1)-(A3), Problem P has a unique solution.

Another result concerns the Euler implicit time discretization of Problem P, for which we prove the convergence by obtaining optimal error estimates.

This chapter is organized as follows: Section 3.1 provides the existence of solution to Problem P. The uniqueness of the solution is proved in Section 3.2. In Section 3.3, some error estimates for an Euler implicit time discretization scheme are obtained, and in Section 3.4, an iterative approach for solving the time discretization nonlinear problems is discussed and some numerical computations are given to verify the theoretical results.

3.1 Existence

We show the existence of a solution to Problem P by the method of Rothe (see [14]), based on the Euler implicit time stepping. Before defining the time discretization we mention the following elementary inequality, which will be used several times later:

$$ab \leq \frac{1}{2\delta}a^2 + \frac{\delta}{2}b^2, \quad \text{for any} \quad a, b \in \mathbb{R} \quad \text{and} \quad \delta > 0. \tag{3.1.1}$$

3.1.1 The time discretization

With $N \in \mathbb{N}$, let $h = T/N$ and consider the following:

Problem P^{n+1} Given $u^n \in W_0^{1,2}(\Omega), n \in \{0, 1, 2, ..., N-1\}$, find $u^{n+1} \in W_0^{1,2}(\Omega)$ such that

$$\begin{aligned}&(u^{n+1} - u^n, \phi) + h\Big(\nabla \cdot \mathbf{F}(u^{n+1}), \phi\Big) + h\Big(H(u^{n+1})\nabla u^{n+1}, \nabla\phi\Big)\\&+\tau\Big(\nabla(u^{n+1} - u^n), \nabla\phi\Big) = 0,\end{aligned} \tag{3.1.2}$$

for any $\phi \in W_0^{1,2}(\Omega)$, here $(\cdot, \cdot)$ means L^2 inner product.

Our final goal is to prove the existence of a solution to Problem P. To do so, we consider a series of time-discrete solutions, and then pass the time step to 0. Therefore we are only interested in small time step. Particularly, in this section we assume

$$h \leq \frac{\tau}{4M}, \tag{3.1.3}$$

for which the results concerning Problem P^{n+1} are obtained.

Lemma 3.1.1 *If h satisfies (3.1.3), Problem P^{n+1} has a unique solution.*

Proof . To prove Lemma 3.1.1, we first define

$$\mathcal{G}(y) := \int_0^y \Big(H(v) + \frac{\tau}{h}\Big)dv, \tag{3.1.4}$$

and it satisfies

$$\Big(u_1 - u_2, \mathcal{G}(u_1) - \mathcal{G}(u_2)\Big) \geq \int_\Omega min\Big(\mathcal{G}'(\cdot)\Big) \cdot (u_1 - u_2)^2 dx \geq 3M\|u_1 - u_2\|^2_{L^2(\Omega)}. \tag{3.1.5}$$

Further, we have for the function $\mathcal{G}^{-1}$

$$\frac{h}{\tau + hM} \leq (\mathcal{G}^{-1})' = \frac{1}{\mathcal{G}'} \leq \frac{h}{\tau - hM}.$$

Define

$$\begin{aligned}
&a : W_0^{1,2}(\Omega) \times W_0^{1,2}(\Omega) \to \mathbb{R}, \quad a(v, \phi) = \Big(\mathcal{G}^{-1}(v), \phi\Big) + h\Big(\nabla \cdot \mathbf{F}(\mathcal{G}^{-1}(v)), \phi\Big) + h(\nabla v, \nabla \phi),\\
&b : W_0^{1,2}(\Omega) \to \mathbb{R}, \quad b(\phi) = (u^n, \phi) + \tau(\nabla u^n, \nabla \phi).
\end{aligned}$$

Clearly, b is a linear bounded functional and for each $v \in W_0^{1,2}(\Omega)$, $\phi \mapsto a(v, \phi)$ is a linear bounded functional. Furthermore, for small enough h,

$$\begin{aligned}
a(v_1, v_1 - v_2) - a(v_2, v_1 - v_2) &\geq \frac{h}{\tau + hM}\|v_1 - v_2\|^2_{L^2(\Omega)} + h\|\nabla(v_1 - v_2)\|^2_{L^2(\Omega)}\\
&\quad - \frac{Lh^2}{\tau - hM}\|v_1 - v_2\|_{L^2(\Omega)} \cdot \|\nabla(v_1 - v_2)\|_{L^2(\Omega)}\\
&\geq \left(h - \frac{Lh^2}{2(\tau - hM)}\right)\|\nabla(v_1 - v_2)\|^2_{L^2(\Omega)}\\
&\quad + \left(\frac{h}{\tau + hM} - \frac{Lh^2}{2(\tau - hM)}\right)\|v_1 - v_2\|^2_{L^2(\Omega)}\\
&\geq C\|v_1 - v_2\|^2_{W^{1,2}(\Omega)},
\end{aligned}$$

and it is easy to check that

$$|a(v_1, \phi) - a(v_2, \phi)| \leq C\|v_1 - v_2\|_{W^{1,2}(\Omega)} \cdot \|\phi\|_{W^{1,2}(\Omega)}.$$

Therefore, the nonlinear Lax-Milgram theorem ([30], p.174-175) provides the existence and uniqueness of a $v \in W_0^{1,2}(\Omega)$ such that

$$\Big(\mathcal{G}^{-1}(v), \phi\Big) + h\Big(\nabla \cdot \mathbf{F}(\mathcal{G}^{-1}(v)), \phi\Big) + h(\nabla v, \nabla \phi) = (u^n, \phi) + \tau(\nabla u^n, \nabla \phi), \tag{3.1.6}$$

for any $\phi \in W_0^{1,2}(\Omega)$. By the properties of $\mathcal{G}$, $u = \mathcal{G}^{-1}(v)$ solves Problem P^{n+1}, and this solution is unique. □

3.1.2 A priori estimates

Having established the existence for the time discretization problems, we proceed with investigating Problem P. To this end, we start with some a priori estimates. Before giving the estimates, we mention the discrete Gronwall inequality which will be used later.

Lemma 3.1.2 ***Discrete Gronwall inequality***: *If $\{y_n\}, \{f_n\}$ and $\{g_n\}$ are nonnegative sequences and*

$$y_n \le f_n + \sum_{0\le k<n} g_k y_k \text{ for } n \ge 0,$$

then

$$y_n \le f_n + \sum_{0\le k<n} f_k g_k \exp\Big(\sum_{k<j<n} g_j\Big), \text{ for } n \ge 0.$$

Lemma 3.1.3 *If h is small enough, for any $n \in \{0, 1, 2, ..., N-1\}$ we have:*

$$\|u^{n+1}\|^2_{L^2(\Omega)} + \tau\|\nabla u^{n+1}\|^2_{L^2(\Omega)} \le C, \tag{3.1.7}$$

$$\|u^{n+1} - u^n\|^2_{L^2(\Omega)} + \frac{\tau}{2}\|\nabla(u^{n+1} - u^n)\|^2_{L^2(\Omega)} \le Ch^2, \tag{3.1.8}$$

here C is independent of n.

Proof . **1**. Taking $\phi = u^{n+1}$ in (3.1.2) gives

$$\begin{aligned} &\|u^{n+1}\|^2_{L^2(\Omega)} + h\Big(\nabla\cdot\mathbf{F}(u^{n+1}), u^{n+1}\Big) + \tau\|\nabla u^{n+1}\|^2_{L^2(\Omega)} \\ &+h\int_\Omega H(u^{n+1})|\nabla u^{n+1}|^2 dx = (u^n, u^{n+1}) + \tau(\nabla u^n, \nabla u^{n+1}). \end{aligned} \tag{3.1.9}$$

Since u^{n+1} vanishes on $\partial\Omega$, with $\mathcal{F}(u^{n+1}) = \int_0^{u^{n+1}} \mathbf{F}(v)dv$ and Gauss' theorem we have

$$\Big(\nabla\cdot\mathbf{F}(u^{n+1}), u^{n+1}\Big) = -\int_\Omega \mathbf{F}(u^{n+1})\cdot\nabla u^{n+1} dx = -\int_\Omega \nabla\cdot[\mathcal{F}(u^{n+1})]dx = \int_{\partial\Omega} \nu\cdot\mathcal{F}(0)dx = 0.$$

By (3.1.1) and $|H| \le M$, this yields

$$\|u^{n+1}\|^2_{L^2(\Omega)} + \tau\|\nabla u^{n+1}\|^2_{L^2(\Omega)} \le \|u^n\|^2_{L^2(\Omega)} + \tau\|\nabla u^n\|^2_{L^2(\Omega)} + 2hM\|\nabla u^{n+1}\|^2_{L^2(\Omega)}, \tag{3.1.10}$$

for all n. Because $u^0 \in W_0^{1,2}(\Omega)$, we have

$$\|u^n\|^2_{L^2(\Omega)} + (\tau - 2hM)\|\nabla u^n\|^2_{L^2(\Omega)} \le C + 2hM \sum_{j=1}^{n-1} \|\nabla u^j\|^2_{L^2(\Omega)}. \tag{3.1.11}$$

Recalling (3.1.3) we have

$$\frac{\tau}{2}\|\nabla u^n\|^2_{L^2(\Omega)} \le C + 2hM \sum_{j=1}^{n-1} \|\nabla u^j\|^2_{L^2(\Omega)}. \tag{3.1.12}$$

Using Lemma 3.1.2, we obtain

$$\|\nabla u^n\|^2_{L^2(\Omega)} \le C, \quad \text{for any } n \in \mathbb{N}. \tag{3.1.13}$$

Further, using Poincaré inequality we have

$$\|u^n\|^2_{L^2(\Omega)} \le C, \quad \text{for any } n \in \mathbb{N}. \tag{3.1.14}$$

2. Taking $\phi = u^{n+1} - u^n$ in (3.1.2) gives

$$\begin{aligned} &\|u^{n+1} - u^n\|^2_{L^2(\Omega)} - h\Big(\mathbf{F}(u^{n+1}), \nabla(u^{n+1} - u^n)\Big) \\ &+h\Big(H(u^{n+1})\nabla u^{n+1}, \nabla(u^{n+1} - u^n)\Big) + \tau\|\nabla(u^{n+1} - u^n)\|^2_{L^2(\Omega)} = 0. \end{aligned} \tag{3.1.15}$$

Using (3.1.1) and the boundedness of $\mathbf{F}$ and H, we have

$$\begin{aligned} 0 \ge \|u^{n+1} - u^n\|^2_{L^2(\Omega)} + \tau\|\nabla(u^{n+1} - u^n)\|^2_{L^2(\Omega)} - \frac{\tau}{4}\|\nabla(u^{n+1} - u^n)\|^2_{L^2(\Omega)} \\ -\frac{Ch^2}{\tau} - \frac{\tau}{4}\|\nabla(u^{n+1} - u^n)\|^2_{L^2(\Omega)} - \frac{h^2M^2}{\tau}\|\nabla u^{n+1}\|^2_{L^2(\Omega)}, \end{aligned} \tag{3.1.16}$$

leads to

$$\|u^{n+1} - u^n\|^2_{L^2(\Omega)} + \frac{\tau}{2}\|\nabla(u^{n+1} - u^n)\|^2_{L^2(\Omega)} \le Ch^2, \text{ for any } n \in \mathbb{N}. \quad \square \tag{3.1.17}$$

Remark 3.1.1 *From (3.1.17) one immediately obtains*

$$\sum_{k=1}^{N} \|u^k - u^{k-1}\|^2_{L^2(\Omega)} \le Ch, \tag{3.1.18}$$

$$\sum_{k=1}^{N} \|\nabla(u^k - u^{k-1})\|^2_{L^2(\Omega)} \le Ch. \tag{3.1.19}$$

3.1.3 Existence

To show the existence of a solution to Problem P, we start by defining

$$U_N(t) = u^{k-1} + \frac{t - t^{k-1}}{h}(u^k - u^{k-1}), \quad \text{and} \quad \overline{U}_N(t) = u^k, \tag{3.1.20}$$

for $t^{k-1} = (k-1)h \le t < t^k = kh, k = 1, 2...N$. We have the following result:

Theorem 3.1.1 *Assuming (A1)-(A3), Problem P has a solution.*

Proof . According to the a priori estimates in Lemma 3.1.3,

$$\begin{aligned}\int_0^T \|U_N(t)\|^2_{L^2(\Omega)} dt &= \sum_{k=1}^N \int_{t^{k-1}}^{t^k} \|u^{k-1} + \frac{t - t^{k-1}}{h}(u^k - u^{k-1})\|^2_{L^2(\Omega)} dt \\ &\le 2\sum_{k=1}^N \int_{t^{k-1}}^{t^k} \left(\|u^{k-1}\|^2_{L^2(\Omega)} + \|u^k - u^{k-1}\|^2_{L^2(\Omega)}\right) dt \\ &\le C.\end{aligned} \tag{3.1.21}$$

Similarly,

$$\int_0^T \|\nabla U_N(t)\|^2_{L^2(\Omega)} dt \le C, \tag{3.1.22}$$

$$\int_0^T \|\partial_t U_N\|^2_{L^2(\Omega)} dt = \frac{1}{h}\sum_{k=1}^N \|u^k - u^{k-1}\|^2_{L^2(\Omega)} \le C, \tag{3.1.23}$$

and

$$\begin{aligned}\int_0^T \|\partial_t \nabla U_N\|^2_{L^2(\Omega)} dt &= \sum_{k=1}^N \int_{t^{k-1}}^{t^k} \|\frac{1}{h}\nabla(u^k - u^{k-1})\|^2_{L^2(\Omega)} dt \\ &= \frac{1}{h}\sum_{k=1}^N \|\nabla(u^k - u^{k-1})\|^2_{L^2(\Omega)} \le C.\end{aligned} \tag{3.1.24}$$

Therefore $\{U_N\}_{N\in\mathbb{N}}$ is uniformly bounded in $W^{1,2}(0,T;W_0^{1,2}(\Omega))$, so it has a subsequence (still denoted as $\{U_N\}$) that converges weakly to some $U \in W^{1,2}(0,T;W_0^{1,2}(\Omega))$. Again, using the compact imbedding of $W^{1,2}(Q)$ into $L^2(Q)$, we have U_N converges strongly to U in $L^2(Q)$.

We now exploit a general principle that relates the piecewise linear and the piecewise constant interpolation (see e.g. [15] for a proof of the corresponding lemma): if one interpolation converges strongly in $L^2(Q)$, then the other interpolation also converges strongly in $L^2(Q)$. From the convergence of U_N, we conclude that $\overline{U}_N$ also converges strongly in $L^2(Q)$.

Further, observe that $H(\overline{U}_N)\nabla\overline{U}_N$ is bounded in $(L^2(\Omega))^d$, therefore it has a weak limit χ. To identify this limit, we take $\phi \in C_0^\infty(\Omega)$ as test function. Since $H(\overline{U}_N) \to H(U)$ strongly in $L^2(\Omega)$ and $\nabla\overline{U}_N \rightharpoonup \nabla U$ weakly in $(L^2(\Omega))^d$, we have

$$\Big(H(\overline{U}_N)\nabla\overline{U}_N, \nabla\phi\Big) \to \Big(H(U)\nabla U, \nabla\phi\Big).$$

This implies that $H(\overline{U}_N)\nabla\overline{U}_N \rightharpoonup H(U)\nabla U$ in distributional sense. By the uniqueness of the limit, we have $\chi = H(U)\nabla U$.

From (3.1.2), we know

$$\begin{aligned}&\int_0^T\int_\Omega \partial_t U_N(t)\phi dxdt - \int_0^T\int_\Omega \mathbf{F}\Big(\overline{U}_N(t)\Big)\cdot\nabla\phi dxdt\\&+\int_0^T\int_\Omega H\Big(\overline{U}_N(t)\Big)\nabla\Big(\overline{U}_N(t)\Big)\cdot\nabla\phi dxdt + \tau\int_0^T\int_\Omega \partial_t\nabla U_N(t)\cdot\nabla\phi dxdt = 0,\end{aligned} \tag{3.1.25}$$

for any $\phi \in L^2(0,T;W_0^{1,2}(\Omega))$,

Using the weak convergence of U_N and $\overline{U}_N$, we consider a sequence $h \to 0$ and pass to the limit in (3.1.25). This shows that U is a solution to Problem P. □

Remark 3.1.2 *As will be proved in the following section, the solution of Problem P is unique. Therefore the convergence holds along any sequence* $\{h_n\}_{n\in\mathbb{N}} \searrow 0$.

3.2 Uniqueness

Here we show that the solution to Problem P is unique. To do so, we use the function G_g, defined as the weak solution to the Possion equation

$$-\Delta G_g = g \in L^2(\Omega) \quad \text{in} \quad \Omega, \tag{3.2.1}$$

with boundary condition $G_g|_{\partial\Omega} = 0$. It is easy to show that

$$G_g \in W_0^{1,2}(\Omega) \text{ and } \|G_g\|_{W^{1,2}(\Omega)} = \|g\|_{W^{-1,2}(\Omega)} \leq C\|g\|_{L^2(\Omega)}. \tag{3.2.2}$$

For the ease of writing, we define

$$\mathcal{H}(y) := \int_0^y H(v)dv. \tag{3.2.3}$$

We have the following result:

Theorem 3.2.1 *Assuming (A1)-(A3), the solution of Problem P is unique.*

Proof . Assume u and v are two solutions, we have $(u-v)(\cdot,0) = 0$ and for any $\tilde{t} > 0$,

$$\begin{aligned} &\int_0^{\tilde{t}}\int_\Omega (u-v)_t \phi dxdt - \int_0^{\tilde{t}}\int_\Omega \Big(\mathbf{F}(u) - \mathbf{F}(v)\Big)\cdot\nabla\phi dxdt \\ &+ \int_0^{\tilde{t}}\int_\Omega \nabla\Big(\mathcal{H}(u) - \mathcal{H}(v)\Big)\cdot\nabla\phi dxdt + \tau\int_0^{\tilde{t}}\int_\Omega \nabla(u-v)_t\cdot\nabla\phi dxdt = 0, \end{aligned} \tag{3.2.4}$$

for any $\phi \in L^2(0,T;W_0^{1,2}(\Omega))$.

According to (3.2.2), we know that $G_{u-v} \in W_0^{1,2}(\Omega)$ and

$$(\nabla G_{u-v}, \nabla\psi) = (u-v,\psi), \text{ for any } \psi \in W_0^{1,2}(\Omega). \tag{3.2.5}$$

Besides,

$$\|G_{u-v}\|_{W^{1,2}(\Omega)} \leq C\|u-v\|_{L^2(\Omega)}. \tag{3.2.6}$$

Note that G_{u-v} also depends on t implicitly, through u and v. For any $\tilde{t} > 0$, using (3.2.5) we have

$$
\begin{aligned}
&\int_0^{\tilde{t}} \int_\Omega (u-v)_t G_{u-v} dxdt \\
&= \int_\Omega \left[(u-v)G_{u-v}\right]_0^{\tilde{t}} dx - \int_\Omega \int_0^{\tilde{t}} (u-v)\partial_t G_{u-v} dtdx \\
&= \int_\Omega \left[|\nabla G_{u-v}|^2\right]_0^{\tilde{t}} dx - \int_0^{\tilde{t}} \int_\Omega \nabla G_{u-v} \cdot \nabla \partial_t G_{u-v} dxdt \\
&= \frac{1}{2} \int_\Omega |\nabla G_{u-v}(\cdot, \tilde{t})|^2 dx,
\end{aligned}
\tag{3.2.7}
$$

as $G_{u-v}(\cdot, 0) = 0$. Further, since $\mathbf{F}$ is Lipschitz, we have

$$
\begin{aligned}
&\int_0^{\tilde{t}} \int_\Omega \Big(\mathbf{F}(u) - \mathbf{F}(v)\Big) \cdot \nabla G_{u-v} dxdt \\
&\leq C \int_0^{\tilde{t}} \int_\Omega |u-v| \cdot |\nabla G_{u-v}| dxdt \leq C \int_0^{\tilde{t}} \int_\Omega |u-v|^2 dxdt.
\end{aligned}
\tag{3.2.8}
$$

Next, the boundedness of H implies

$$
\begin{aligned}
&\int_0^{\tilde{t}} \int_\Omega \nabla\Big(\mathcal{H}(u) - \mathcal{H}(v)\Big) \cdot \nabla G_{u-v} dxdt \\
&= \int_0^{\tilde{t}} \int_\Omega \Big(\mathcal{H}(u) - \mathcal{H}(v)\Big)(u-v) dxdt \geq -M \int_0^{\tilde{t}} \int_\Omega |u-v|^2 dxdt.
\end{aligned}
\tag{3.2.9}
$$

Finally,

$$
\begin{aligned}
&\tau \int_0^{\tilde{t}} \int_\Omega \partial_t \nabla(u-v) \cdot \nabla G_{u-v} dxdt \\
&= \tau \int_0^{\tilde{t}} \int_\Omega \partial_t (u-v)(u-v) dxdt = \frac{\tau}{2} \int_\Omega (u-v)(\cdot, \tilde{t})^2 dx.
\end{aligned}
\tag{3.2.10}
$$

Therefore taking $\phi = G_{u-v}$ in (3.2.4) gives

$$
\frac{1}{2} \|\nabla G_{u-v}(\cdot, \tilde{t})\|^2_{L^2(\Omega)} + \frac{\tau}{2} \|(u-v)(\cdot, \tilde{t})\|^2_{L^2(\Omega)} \leq (C+M) \int_0^{\tilde{t}} \int_\Omega |u-v|^2 dxdt. \tag{3.2.11}
$$

By Gronwall's inequality, $\|(u-v)(\cdot, \tilde{t})\|_{L^2(\Omega)} = 0$. Since $\tilde{t}$ is arbitrary, this gives uniqueness.

3.3 Error estimates

From the above we see that the approximating sequence U_N converges strongly to U in $L^2(Q)$. In this section, we will estimate the error $U_N - U$. Recalling (3.2.4), for any $\tilde{t} > 0$ we have

$$\begin{aligned}&\int_0^{\tilde{t}}\int_\Omega \partial_t U_N(t)\phi dxdt - \int_0^{\tilde{t}}\int_\Omega \mathbf{F}\Big(\overline{U}_N(t)\Big)\cdot\nabla\phi dxdt\\&+\int_0^{\tilde{t}}\int_\Omega \nabla\mathcal{H}\Big(\overline{U}_N(t)\Big)\cdot\nabla\phi dxdt+\tau\int_0^{\tilde{t}}\int_\Omega \partial_t\nabla U_N(t)\cdot\nabla\phi dxdt=0.\end{aligned}\tag{3.3.1}$$

for any $\phi \in L^2(0,T;W_0^{1,2}(\Omega))$. For simple writing, we denote

$$e_u(t) = u(t) - U_N(t), \quad \text{and} \quad e_{\mathcal{H}}(t) = \mathcal{H}\Big(u(t)\Big) - \mathcal{H}\Big(U_N(t)\Big),\tag{3.3.2}$$

here $\mathcal{H}$ is defined in (3.2.3). Obviously, $e_u, e_{\mathcal{H}} \in W_0^{1,2}(\Omega)$ and $e_u(\cdot,0) = e_{\mathcal{H}}(\cdot,0) = 0$.

Theorem 3.3.1 *The following estimate holds:*

$$\|e_u\|_{L^\infty\left(0,T;L^2(\Omega)\right)} \le Ch.\tag{3.3.3}$$

Proof . Subtracting (3.3.1) from (3.0.4) gives

$$\begin{aligned}&\int_0^{\tilde{t}}\int_\Omega \partial_t e_u\phi dxdt - \int_0^{\tilde{t}}\int_\Omega \Big(\mathbf{F}(u(t)) - \mathbf{F}(\overline{U}_N(t))\Big)\cdot\nabla\phi dxdt\\&+\int_0^{\tilde{t}}\int_\Omega \nabla\Big(\mathcal{H}(u(t)) - \mathcal{H}(\overline{U}_N(t))\Big)\cdot\nabla\phi dxdt+\tau\int_0^{\tilde{t}}\int_\Omega \partial_t\nabla e_u\cdot\nabla\phi dxdt=0.\end{aligned}\tag{3.3.4}$$

To estimate the discretization error we proceed as in [19, 22, 23, 27] (where degenerate parabolic equations are considered) and test with the function G_{e_u} satisfying

$$(\nabla G_{e_u}, \nabla\psi) = (e_u, \psi), \text{ for any } \psi \in W_0^{1,2}(\Omega).\tag{3.3.5}$$

By (3.2.2), G_{e_u} belongs to $W_0^{1,2}(\Omega)$. Moreover,

$$\|G_{e_u}\|_{W^{1,2}(\Omega)} \le C\|e_u\|_{L^2(\Omega)}.\tag{3.3.6}$$

As in Section 3 we have for any $\tilde{t} > 0$

$$\int_0^{\tilde{t}} \int_\Omega \partial_t e_u G_{e_u} dxdt = \frac{1}{2} \int_\Omega \left(\nabla G_{e_u}(\cdot, \tilde{t})\right)^2 dx. \tag{3.3.7}$$

Further,

$$\begin{aligned}
& \int_0^{\tilde{t}} \int_\Omega \left(\mathbf{F}(u(t)) - \mathbf{F}(\overline{U}_N(t))\right) \cdot \nabla G_{e_u} dxdt \\
\leq\ & C_1 \int_0^{\tilde{t}} \|e_u\|^2_{L^2(\Omega)} dxdt + \int_0^{\tilde{t}} \int_\Omega \left(\mathbf{F}(U_N(t)) - \mathbf{F}(\overline{U}_N(t))\right) \cdot \nabla G_{e_u} dxdt \\
\leq\ & C_1 \int_0^{\tilde{t}} \|e_u\|^2_{L^2(\Omega)} dxdt + C_2 \int_0^{\tilde{t}} \|U_N - \overline{U}_N\|_{L^2(\Omega)} \|\nabla G_{e_u}\|_{L^2(\Omega)} dt,
\end{aligned} \tag{3.3.8}$$

Here C_1, C_2 are two positive numbers. Since $U_N - \overline{U}_N = \frac{t_k - t}{h}(u^k - u^{k-1})$, for $t \in (t^{k-1}, t^k)$. By (3.1.18), we get $\|U_N - \overline{U}_N\|_{L^2(\Omega)} \leq Ch$, therefore

$$\int_0^{\tilde{t}} \int_\Omega \left(\mathbf{F}(u(t)) - \mathbf{F}(\overline{U}_N(t))\right) \cdot \nabla G_{e_u} dxdt \leq \left(C_1 + \frac{1}{2}\right) \int_0^{\tilde{t}} \int_\Omega |e_u|^2 dxdt + \frac{C_2^2 h^2}{2}. \tag{3.3.9}$$

Similarly,

$$\begin{aligned}
& - \int_0^{\tilde{t}} \int_\Omega \nabla\left(\mathcal{H}(u(t)) - \mathcal{H}(\overline{U}_N(t))\right) \cdot \nabla G_{e_u} dxdt \\
&= - \int_0^{\tilde{t}} \int_\Omega \nabla e_{\mathcal{H}} \cdot \nabla G_{e_u} dxdt - \int_0^{\tilde{t}} \int_\Omega \nabla\left(\mathcal{H}(U_N(t)) - \mathcal{H}(\overline{U}_N(t))\right) \cdot \nabla G_{e_u} dxdt \\
&= - \int_0^{\tilde{t}} \int_\Omega e_{\mathcal{H}} e_u dxdt - \int_0^{\tilde{t}} \int_\Omega \left(\mathcal{H}(U_N(t)) - \mathcal{H}(\overline{U}_N(t))\right) \cdot e_u dxdt \\
&\leq M \int_0^{\tilde{t}} \int_\Omega |e_u|^2 dxdt + \frac{M}{2} \int_0^{\tilde{t}} \int_\Omega |U_N(t) - \overline{U}_N(t)|^2 dxdt + \frac{1}{2} \int_0^{\tilde{t}} \int_\Omega |e_u|^2 dxdt \\
&\leq \left(M + \frac{1}{2}\right) \int_0^{\tilde{t}} \int_\Omega |e_u|^2 dxdt + C_3 h^2,
\end{aligned} \tag{3.3.10}$$

with C_3 a positive number. Further,

$$\tau \int_0^{\tilde{t}} \int_\Omega \partial_t \nabla e_u \nabla G_{e_u} dxdt = \tau \int_\Omega e_u \partial_t e_u dxdt = \frac{\tau}{2} \int_\Omega e_u(\cdot, \tilde{t})^2 dx. \tag{3.3.11}$$

Using (3.3.7-3.3.11) and taking $\psi = G_{e_u}$ in (3.3.4) give

$$
\begin{aligned}
&\frac{1}{2}\int_{\Omega}\left(\nabla G_{e_u}(\cdot,\tilde{t})\right)^2 dx + \frac{\tau}{2}\int_{\Omega} e_u(\cdot,\tilde{t})^2 dx \\
\leq &\left(\frac{C_2^2}{2} + C_3\right)h^2 + (M + C_1 + 1)\int_0^{\tilde{t}}\int_{\Omega}|e_u|^2 dxdt.
\end{aligned}
\tag{3.3.12}
$$

Using Gronwall's inequality, we obtain the estimate

$$
\|e_u\|_{L^\infty\left(0,T;L^2(\Omega)\right)} \leq Ch. \qquad \square \tag{3.3.13}
$$

Remark 3.3.1 *From (3.3.13), since H is Lipschitz continuous, we immediately obtain*

$$
\|e_{\mathcal{H}}(\cdot,t)\|_{L^\infty\left(0,T;L^2(\Omega)\right)} \leq Ch. \tag{3.3.14}
$$

3.4 Numerical examples

In this section, we give two numerical examples to verify the theoretical findings.

3.4.1 Example 1

We solve the following equation in $Q = (0,1) \times (0,1]$

$$
\frac{\partial u}{\partial t} = \frac{1}{6}\frac{\partial}{\partial x}\left([u]_+\frac{\partial u}{\partial x}\right) + \frac{1}{6}\frac{\partial^3 u}{\partial^2 x\partial t} - \frac{1}{2(1+t)^2}, \tag{3.4.1}
$$

with initial and boundary conditions

$$
u(x,0) = x(1-x), \quad u(0,t) = u(1,t) = 0. \tag{3.4.2}
$$

Here

$$
[u]_+ = \begin{cases} u & \text{if} \quad u > 0, \\ 0 & \text{if} \quad u \leq 0. \end{cases} \tag{3.4.3}
$$

For the equation (3.4.1), the analytical solution is

$$
u(x,t) = \frac{x(1-x)}{1+t}. \tag{3.4.4}
$$

In the following, we use this solution to test the numerical scheme. Before giving the numerical results, we present an iterative scheme to solve the time discretization problems. To do so, taking $h = 1/N (N \in \mathbb{N})$ and denoting $f(t) = \frac{1}{2(1+t)^2}$, formally we get

$$\frac{u^n - u^{n-1}}{h} = \frac{1}{6}\partial_x\left([u^n]_+\partial_x u^n\right) + \frac{1}{6}\partial_{xx}\left(\frac{u^n - u^{n-1}}{h}\right) - f(t^n). \tag{3.4.5}$$

Define the Kirchhoff transform

$$v = \beta(u) := \frac{1}{6}\int_0^u (h[s]_+ + 1)ds = \begin{cases} \frac{h}{12}u^2 + \frac{1}{6}u, & \text{if} \quad u > 0, \\ \frac{1}{6}u, & \text{if} \quad u \le 0, \end{cases} \tag{3.4.6}$$

instead of solving (3.4.5), we seek $v^n = \beta(u^n)$ such that

$$\beta^{-1}(v^n) - \partial_{xx}v^n = u^{n-1} - \frac{1}{6}\partial_{xx}u^{n-1} - hf(t^n). \tag{3.4.7}$$

with $v^n = 0$ at $x = 0$ and $x = 1$. To solve (3.4.7), we observe that $(\beta^{-1})' \le 6$, and define an iteration method inspired from [29], pp. 90-100 (also see e.g. [9], [24]):

$$6v^{n,i} - \partial_{xx}v^{n,i} = 6v^{n,i-1} - \beta^{-1}(v^{n,i-1}) + \alpha(u^{n-1}, t^n), \tag{3.4.8}$$

where $i = 1, 2...$ and

$$\alpha(u^{n-1}, t^n) = u^{n-1} - \frac{1}{6}\partial_{xx}u^{n-1} - hf(t^n). \tag{3.4.9}$$

This iteration requires a starting point $v^{n,0}$. As will be proved below, the iteration is convergent for any $v^{n,0}$. However, for the practical reasons, we choose $v^{n,0} = v^{n-1} = \beta(u^{n-1})$.

Lemma 3.4.1 *The iteration method (3.4.8) is convergent in the $W^{1,2}(0,1)$ norm.*

Proof . We write (3.4.9) in weak form, find $v^{n,i} \in W_0^{1,2}(0,1)$ such that

$$\left(6v^{n,i}, \phi\right) + \left(\partial_x v^{n,i}, \partial_x\phi\right) = \left(6v^{n,i-1} - \beta^{-1}(v^{n,i-1}), \phi\right) + \left(\alpha(u^{n-1}, t^n), \phi\right). \tag{3.4.10}$$

for any $\phi \in W_0^{1,2}(0,1)$. Similarly,

$$\left(6v^{n,i-1}, \phi\right) + \left(\partial_x v^{n,i-1}, \partial_x\phi\right) = \left(6v^{n,i-2} - \beta^{-1}(v^{n,i-2}), \phi\right) + \left(\alpha(u^{n-1}, t^n), \phi\right). \tag{3.4.11}$$

Subtracting (3.4.10) from (3.4.11),

$$\begin{aligned} & 6\left(v^{n,i} - v^{n,i-1}, \phi\right) + \left(\partial_x(v^{n,i} - v^{n,i-1}), \partial_x\phi\right) \\ & = 6\left(v^{n,i-1} - v^{n,i-2}, \phi\right) - \left(\beta^{-1}(v^{n,i-1}) - \beta^{-1}(v^{n,i-2}), \phi\right). \end{aligned} \tag{3.4.12}$$

Taking $\phi = v^{n,i} - v^{n,i-1}$ gives,

$$\begin{aligned} & 6\|v^{n,i} - v^{n,i-1}\|^2_{L^2(\Omega)} + \|\partial_x(v^{n,i} - v^{n,i-1})\|^2_{L^2(\Omega)} \\ \leq\ & \|v^{n,i} - v^{n,i-1}\|_{L^2(\Omega)} \cdot \|6\left(v^{n,i-1} - v^{n,i-2}\right) - \left(\beta^{-1}(v^{n,i-1}) - \beta^{-1}(v^{n,i-2})\right)\|_{L^2(\Omega)}. \end{aligned} \tag{3.4.13}$$

From the definition of β, we have

$$\beta'(u) = \begin{cases} \frac{1}{6}(hu+1), & \text{if} \quad u \geq 0 \\ \frac{1}{6}, & \text{otherwise.} \end{cases} \tag{3.4.14}$$

Therefore

$$(\beta^{-1})'(v) = \frac{1}{\beta'(u)} \in (0, 6]. \tag{3.4.15}$$

From (3.4.13), we obtain

$$6\left\|v^{n,i} - v^{n,i-1}\right\|^2_{L^2(\Omega)} + \|\partial_x(v^{n,i} - v^{n,i-1})\|^2_{L^2(\Omega)} \leq 6\|v^{n,i} - v^{n,i-1}\|^2_{L^2(\Omega)} \cdot \|v^{n,i-1} - v^{n,i-2}\|^2_{L^2(\Omega)}. \tag{3.4.16}$$

Using Poincaré inequality, $\|u\|_{L^2(0,1)} \leq \|\partial_x u\|_{L^2(0,1)}$ for any $u \in W_0^{1,2}(0,1)$. Therefore

$$\begin{aligned} & \|v^{n,i} - v^{n,i-1}\|^2_{L^2(\Omega)} + \frac{1}{6}\|\partial_x(v^{n,i} - v^{n,i-1})\|^2_{L^2(\Omega)} \\ \leq\ & \frac{1}{2}\|v^{n,i} - v^{n,i-1}\|^2_{L^2(\Omega)} + \frac{1}{2}\|v^{n,i-1} - v^{n,i-2}\|^2_{L^2(\Omega)} \\ \leq\ & \frac{1}{2}\|v^{n,i} - v^{n,i-1}\|^2_{L^2(\Omega)} + \frac{3}{8}\|v^{n,i-1} - v^{n,i-2}\|^2_{L^2(\Omega)} + \frac{1}{8}\|\partial_x(v^{n,i-1} - v^{n,i-2})\|^2_{L^2(\Omega)} \end{aligned} \tag{3.4.17}$$

Define $\|v^{n,i}\|^2 = \|v^{n,i} - v^{n,i-1}\|^2_{L^2(\Omega)} + \frac{1}{3}\|\partial_x(v^{n,i} - v^{n,i-1})\|^2_{L^2(\Omega)}$ (equivalent to the $W^{1,2}$ norm), we obtain

$$\|v^{n,i}\|^2 \leq \frac{3}{4}\|v^{n,i-1}\|^2, \tag{3.4.18}$$

using Banach fixed point theorem, we obtain the convergence of the iteration method (3.4.9). □

We compute the numerical solution u^N of (3.4.1) and estimate the error $e_u = u - u^N$, with u the exact solution of (3.4.1). For simplicity, we only compute e_u at $t = 1$. To this aim, finite difference scheme on uniform mesh with $\Delta x = 10^{-5}$ is coupled with different time stepping $h = 10^{-1}, 10^{-2}, 10^{-3}$ and 10^{-4}. To solve the nonlinear problem at any two steps, we perform 3 to 4 iterations. This is sufficient to achieve $\|v^{n,i} - v^{n,i-1}\|_{L^2(\Omega)} \leq 10^{-5}$. The numerical results are presented in Table 3.1. As follows from Theorem 3.3.1, the error satisfies

$$\|e_u(\cdot, 1)\|_{L^2(\Omega)} \leq Ch. \tag{3.4.19}$$

This is confirmed by the Table 3.1. In particular, we estimate C to 0.066.

h	$\|e_u(\cdot, 1)\|_{L^2(\Omega))}$	ratio($\|e_u\|/h$)
10^{-1}	6.1997×10^{-3}	6.1997×10^{-2}
10^{-2}	6.447×10^{-4}	6.447×10^{-2}
10^{-3}	6.4632×10^{-5}	6.4632×10^{-2}
10^{-4}	6.5842×10^{-6}	6.5842×10^{-2}

Table 3.1: Errors $e_u(\cdot, 1)$ for different h

Figure 3.1 (left) displays the numerical solutions for $u(\cdot, 1)$ at $t = 1$, compared to the analytical solution. Clearly, the numerical solution converges to the analytical solution when h is small enough.

3.4.2 Example 2

Here we solve the following equation in $Q = (0, 2\pi) \times (0, 1]$ with negative function 'H'

$$\frac{\partial u}{\partial t} = -\frac{\partial}{\partial x}\left(u\frac{\partial u}{\partial x}\right) + \frac{\partial^3 u}{\partial x^2 \partial t} + 2\sin x + (1+t)^2 \cos 2x, \tag{3.4.20}$$

with initial and boundary conditions

$$u(x, 0) = \sin x, \quad u(0, t) = u(2\pi, t) = 0, \tag{3.4.21}$$

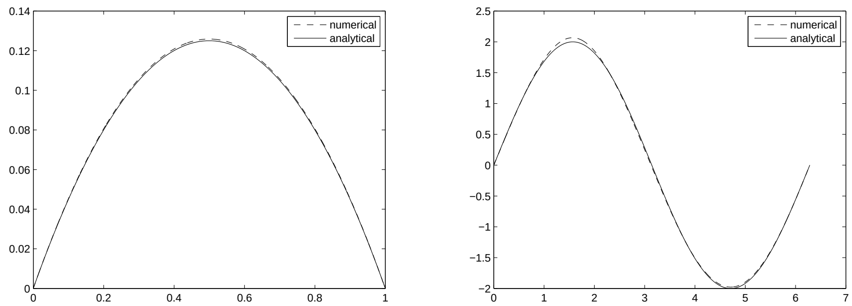


FIGURE 3.1: NUMERICAL SOLUTION AND ANALYTICAL SOLUTION FOR EXAMPLE 1 WITH $\Delta x = 10^{-5}, h = 10^{-2}$(LEFT) AND EXAMPLE 2 WITH $\Delta x = 2\pi \times 10^{-5}, h = 10^{-2}$(RIGHT)

and the analytical solution is

$$u(x,t) = (1+t)\sin x. \tag{3.4.22}$$

Using the similar numerical method in Example 1, here we only give the numerical results. Again for simplicity, we only compute e_u at $t = 1$, and we take $\Delta x = 2\pi \times 10^{-5}$ coupled with different time stepping $h = 0.1, 0.05, 0.01$ and 0.005. The numerical results are presented in Table 3.2. As follows from Theorem 3.3.1, the error satisfies

$$\|e_u(\cdot,1)\|_{L^2(\Omega)} \le Ch. \tag{3.4.23}$$

This is confirmed by the Table 3.2. In particular, we estimate C to 1.45.

h	$\|e_u(\cdot,1)\|_{L^2(\Omega))}$	ratio($\|e_u\|/h$)
0.1	0.14386	1.4386
0.05	0.06878	1.3756
0.01	0.013383	1.3383
0.005	0.0067537	1.3507

TABLE 3.2: ERRORS $e_u(\cdot,1)$ FOR DIFFERENT h

Figure 3.1 (right) also displays the numerical solutions for $u(\cdot,1)$ at $t = 1$, compared to the analytical solution. Clearly, the numerical solution converges to the analytical solution when h is small enough.

References

[1] D.N. Arnold, Jr.J. Douglas and V. Thomée, *Superconvergence of a finite element approximation to the solution of a Sobolev equation in a single space variable*, Math. Comp. **36** (1981), 53–63.

[2] B.C. Aslan, W.W. Hager and S. Moskow, *A generalized eigenproblem for the Laplacian which arises in lightning*, J. Math. Anal. Appl. **341** (2008), 1028–1041.

[3] G.I. Barenblatt, I.P. Zheltov and I.N. Kochina, *Basic concepts in the theory of seepage of homogeneous liquids in fissured rocks (strata)*, J. Appl. Math. Mech. **24** (1960), 1286–1303.

[4] C. Cancès, C. Choquet, Y. Fan and I.S. Pop, *Existence of weak solutions to a degenerate pseudo-parabolic equation modeling two-phase flow in porous media*, CASA Report **10-75**, Eindhoven University of Technology, 2010.

[5] C. Cuesta, J. Hulshof, *A model problem for groundwater flow with dynamic capillary pressure: stability of travelling waves*, Nonlinear Anal. **52** (2003), 1199–1218.

[6] C.M. Cuesta, I.S. Pop, *Numerical schemes for a pseudo-parabolic Burgers equation: discontinuous data and long-time behaviour*, J. Comput. Appl. Math. **224** (2009), 269–283.

[7] C.J. van Duijn, Y. Fan, L.A. Peletier and I.S. Pop, *Travelling wave solutions for degenerate pseudo-parabolic equation modelling two-phase flow in porous media*, CASA Report **10-01**, TU Eindhoven, 2010.

[8] C.J. van Duijn, L.A. Peletier and I.S. Pop, *A new class of entropy solutions of the Buckley-Leverett equation*, SIAM J. Math. Anal. **39** (2007), 507–536.

[9] V.M. Devigne, I.S. Pop, C.J. van Duijn, T. Clopeau, *A numerical scheme for the pore-scale simulation of crystal dissolution and precipitation in porous media*, SIAM J. Numer. Anal. **46** (2008), 895–919.

[10] R.E. Ewing, *Time-stepping Galerkin methods for nonlinear Sobolev partial differential equations*, SIAM J. Numer. Anal. **15** (1978), 1125–1150.

[11] W.H. Ford and T.W. Ting, *Unifrom error estimates for difference approximations to nonlinear pseudo-parabolic partial differential equations*, SIAM J. Numer. Anal. **11** (1974), 155–169.

[12] S.M. Hassanizadeh and W.G. Gray, *Thermodynamic basis of capillary pressure in porous media*, Water Resour. Res. **29** (1993), 3389–3405.

[13] R. Helmig, A. Weiss, B. Wohlmuth, *Dynamic capillary effects in heterogeneous porous media*, Comput. Geosci. **11** (2007), 261–274.

[14] J. Kačur, *Method of Rothe in evolution equations*, Teubner Texts Math. **80**, Teubner Verlagsgesellschaft, Leipzig, 1985.

[15] M. Lenzinger and B. Schweizer, *Two-phase flow equations with outflow boundary conditions in the hydrophobic-hydrophilic case*, Nonlinear Anal. **73** (2010), 840–853.

[16] E. Milne, *The diffusion of imprisoned radiation through a gas*, J. London Math. Soc. **1** (1926), 40–51.

[17] A. Mikelić, H. Bruining, *Analysis of model equations for stress-enhanced diffusion in coal layers. I. Existence of a weak solution*, SIAM J. Math. Anal. **40** (2008), 1671–1691.

[18] A. Mikelić, *A global existence result for the equations describing unsatured flow in porous media with dynamic capillary pressure*, J. Differential Equations **248** (2010), 1561–1577.

[19] R.H. Nochetto and C. Verdi, *Approximation of degenerate parabolic problems using numerical integration*, SIAM J. Numer. Anal. 25 (1988), pp. 784-814.

[20] M. Ptashnyk, *Nonlinear pseudoparabolic equations as sigular limit of reaction-diffusion equations*, Appl. Anal. **85** (2006), 1285–1299.

[21] M. Peszyńska, R. Showalter, S. Yi, *Homogenization of a pseudoparabolic system*, Appl. Anal. **88** (2009), 1265–1282.

[22] I.S. Pop, *Error estimates for a time discretization method for the Richards' equation*, Comput. Geosci. 6 (2002), pp. 141-160.

[23] I.S. Pop and W.A. Yong, *A numerical approach to degenerate parabolic equations*, Numer. Math. 92 (2002), pp. 357-381.

[24] I.S. Pop, F.A. Radu and P. Knabner, *Mixed finite elements for the Richards' equation: linearization procedure*, J. Comput. and Appl. Math. 168 (2004), pp. 365-373.

[25] A. Quarteroni, *Fourier spectral methods for pseudo-parabolic equations*, SIAM J. Numer. Anal. **24** (1987), 323–335.

[26] L.I. Rubinstein, *On the problem of the process of propagation of heat in heterogeneous media*, IZV. Akad. Nauk SSSR, Ser. Geogr. **1** (1948).

[27] F.A. Radu, I.S. Pop and P. Knabner, *Order of convergence estimates for an Euler implicit, mixed finite element discretization of Richards' equation*, SIAM J. Numer. Anal. 42 (2004), pp. 1452-1478.

[28] R.E. Showalter, *A nonlinear parabolic-Sobolev equation*, J. Math. Anal. Appl. **50** (1975), 183–190.

[29] J. Smoller, *Shock Waves and Reaction-Diffusion Equations*, Springer-Verlag, New York, 1994.

[30] E. Zeidler, *Applied Functional Analysis*, Springer-Verlag, New York, 1999.

Chapter 4

Existence of weak solutions to the original degenerate model

Pseudo-parabolic equations appear as models for many real life applications, such as lightning [2], seepage in fissured rocks [4], radiation with time delay [26] and heat conduction models [35]. We consider a pseudo-parabolic equation modeling two-phase flow in porous media, where dynamic effects are complementing the capillary pressure-saturation relationship. With a given maximal time $T > 0$ and for all $x \in \Omega$ a bounded domain in $\mathbb{R}^d$ ($d = 1, 2,$ or 3) having a Lipschitz continuous boundary $\partial\Omega$, we investigate the equation

$$\partial_t u + \nabla \cdot \mathbf{F}(u, x, t) = \nabla \cdot \Big(H(u)\nabla p_c\Big), \quad (x, t) \in Q := \Omega \times (0, T]. \tag{4.0.1}$$

This equation is obtained by including Darcy's law for both phases in the mass conservation laws. Here u stands for water saturation, $\mathbf{F}$ and H are the water fraction flow

This chapter has been submitted to Mathematical Models and Methods in Applied Sciences, and it is a collaborative work with C. Cancès, C. Choquet, I.S. Pop function and the capillary induced diffusion function, while p_c is the capillary pressure term. Such models are proposed in [19, 31]. For recent works providing experimental evidence for the dynamic effects in the phase pressure difference we refer to [5, 12, 21]. Similar models, but considering an "apparent saturation" are discussed in [3]. Here we consider a simplified situation, where

$$p_c = u + \tau\partial_t u. \tag{4.0.2}$$

Then (4.0.1) becomes

$$\partial_t u + \nabla \cdot \mathbf{F}(u, x, t) = \nabla \cdot \Big(H(u)\nabla(u + \tau\partial_t u)\Big). \tag{4.0.3}$$

The functions H and $\mathbf{F}$ depend on the specific model, in particular on the relative permeabilities. Commonly encountered in the engineering literature are relative permeabilities of power-like types, u^{p+1} and $(1-u)^{q+1}$, where p and q are positive real numbers. This leads to

$$H(u) = \frac{K}{\mu}\frac{u^{p+1}(1-u)^{q+1}}{u^{p+1} + M(1-u)^{q+1}}, \mathbf{F}(u, x, t) = \mathbf{Q}(x, t)\frac{u^{p+1}}{u^{p+1} + M(1-u)^{q+1}} + H(u)\rho\mathbf{g}, \tag{4.0.4}$$

where K is the permeability of the porous medium. For the sake of simplicity, we assume the porous medium to be isotropic. Next, μ and $\tilde{\mu}$ are the viscosities of the two phases, whereas $M = \frac{\tilde{\mu}}{\mu} > 0$ is the viscosity ratio of the two fluids, and τ is a positive constant standing for the damping coefficient. Further, $\mathbf{Q}$ is the total flow in the porous medium, satisfying $\nabla\cdot\mathbf{Q} = 0$, and $\mathbf{g}$ is the gravity vector. Finally, ρ denotes the difference between the phase densities.

With the given function H, (4.0.3) becomes degenerate whenever $u = 0$ or $u = 1$. Note that the expression (4.0.4) makes sense only for $u \in [0, 1]$. For completeness we extend H continuously by 0 outside this interval. Therefore the function H is nonnegative, bounded and Lipschitz continuous on the entire $\mathbb{R}$. Similarly, the vector-valued function $\mathbf{F}$ is extended by $\mathbf{Q}$ when $u > 1$ and by $\mathbf{0}$ when $u < 0$, leading to a bounded Lipschitz continuous function. Note that the functions in (4.0.4) are just typical examples.

Throughout this paper we assume

$$
\begin{aligned}
&H : \mathbb{R} \to \mathbb{R} \text{ is } \textit{Lipschitz continuous}, \text{ nonnegative, and satisfies} \\
&\qquad H(u) > 0 \text{ if } 0 < u < 1, \text{ and } H(u) = 0 \text{ otherwise;} \\
&\mathbf{F} : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}^d \text{ is } \textit{Lispchitz continuous}, \text{ satisfying} \\
&\qquad \nabla_x \cdot \mathbf{F}(v, x, t) = 0 \quad \text{for all } v \in \mathbb{R}, (x, t) \in Q.
\end{aligned}
\tag{A1}
$$

Pseudo-parabolic equations like (4.0.3) have been investigated in the mathematical literature for decades. Short time existence of solutions with constant, compact support is obtained in [15], whereas a nonlinear Sobolev type equation is studied in [36]. The existence and uniqueness of weak solutions for some nonlinear pseudo-parabolic equations, where the degeneracy may appear in only one term, are proved in [17] and [33]. Long time existence of weak solutions to a closely related model is proved in [27, 28]. We further refer to [24] for the analysis of a non-degenerate pseudo-parabolic model that includes hysteresis.

The connection between pseudo-parabolic equations and shock solutions to hyperbolic conservation laws is investigated in [14] for the case of a constant function H. The analysis there, based on travelling waves, is continued in [13]. In both cases, undercompressive shocks are obtained for values of τ exceeding a threshold value. Nonclassical shocks are also obtained in [6], but in a heterogeneous medium, and in [23], but based on a different regularization. Travelling wave solutions for a pseudo-parabolic equation involving a convex flux function are analyzed in [9, 10, 30].

Concerning numerical methods for pseudo-parabolic equations, the superconvergence of a finite element approximation of a similar equation is investigated in [1] and time-stepping Galerkin methods are analyzed in [16] and [18], where two finite difference approximation schemes are considered. Further, Fourier spectral methods are analyzed in [34]. For homogeneous media, discontinuous initial data and corresponding numerical schemes for pseudo-parabolic equations are considered in [11], whereas for heterogeneous media we refer to [20]. We also mention [32] for a review of different numerical methods for pseudo-parabolic equations.

In this chapter we prove the existence of weak solutions to the degenerate pseudo-parabolic equation (4.0.3), which are introduced below. Existence results for closely related models are proved in [27] and particularly in [28]. Despite the model we consider in this paper is slightly different of the one studied in [28] (our problem degenerates both in 0 and 1 while a single degeneracy in 0 is considered in [28]), both in [28] and in the present work, a main difficulty is the degenerate nonlinearity characterizing the viscous term in equation (4.0.1). The existence proof given here applies to a more general framework, and (thus) involves different mathematical tools. The existence result obtained in [28] requires stronger assumptions on this degenerate nonlinearity than in the present work. First, Theorem 5 in [28] applies iff H is some regular function in the form (4.0.4) with sufficiently large exponents p and q, while our proof works for any nonnegative function H satisfying (A1). In particular we may consider (4.0.4) with $p = q = 0$, while a similar analysis to the one proposed in [28] would require that $p, q > 4$. Next hypothesis (H6) in [28] means that the initial saturation cannot take any degeneracy value (0 or 1) on a non-zero measure subset of Ω. This assumption is weakened or even removed in the present work (see assumption (A2) below and the subsequent comments). Finally we obtain a saturation satisfying the physical property of belonging to $[0, 1]$ while [28] only ensures that the saturation is nonnegative.

To obtain the existence result we employ regularization and compactness arguments. The main difficulty appears in dealing with the nonlinear and degenerate term involving the third order derivative, for which we combine the Div-Curl Lemma (see e.g. [29, 37]) with equi-integrability properties. A simplified approach is possible whenever the degeneracy H can be controlled by the convective term $\mathbf{F}$, specifically if the product $H^{-1/2}\,\mathbf{F}$ is a bounded function. This is obtained, e.g., if $\mathbf{Q} \equiv \mathbf{0}$, as considered in [28]. In this case one can use the structure of the equation as in [8] to obtain uniform L^6 estimates for $\partial_t u$, and then apply the Div-Curl Lemma directly. Here we consider a rather general convective flux $\mathbf{F}$ that makes this latter strategy fail.

Below we use standard notation in the theory of partial differential equations, such as $L^2(\Omega), W^{1,2}(\Omega)$ and $W_0^{1,2}(\Omega)$. $W^{-1,2}(\Omega)$ denotes the dual space of $W_0^{1,2}(\Omega)$, while $L^2(0, T; W_0^{1,2}(\Omega))$ denotes the Bochner space of $W_0^{1,2}(\Omega)$-valued functions. By $(\cdot, \cdot)$ we mean the inner product in either $L^2(\Omega)$, or $(L^2(\Omega))^d$, and $\|\cdot\|$ stands for the corresponding norm. Furthermore, C denotes a generic positive real number and we define the set

$$V := C_D + W_0^{1,2}(\Omega).$$

The equation (4.0.3) is complemented by the following initial and boundary conditions

$$u(\cdot,0) = u^0 \in V, \quad \text{and} \quad u|_{\partial\Omega} = C_D. \tag{4.0.5}$$

The initial data is assumed to be in $W^{1,2}(\Omega)$. Furthermore, it satisfies $0 \le u^0 \le 1$ almost everywhere in Ω, while $C_D \in (0,1)$ is a constant. The extension to non-constant boundary data is possible, but requires more technical steps, given in [28]. We eliminate these here for the sake of presentation. An important requirement here is that C_D is neither 0 nor 1. The reason for this will become clear in the proof of the main result.

In this paper, a weak solution satisfies

Problem P Find $u \in L^\infty(0,T;V)$, $\partial_t u \in L^2(Q)$ such that $u(\cdot,0) = u^0$, $\sqrt{H(u)}\nabla\partial_t u \in L^2(0,T;L^2(\Omega)^d)$, and such that

$$\begin{aligned} &\int_0^T \int_\Omega \partial_t u\phi dxdt - \int_0^T \int_\Omega \mathbf{F}(u,x,t)\cdot\nabla\phi dxdt \\ &+ \int_0^T \int_\Omega H(u)\nabla u\cdot\nabla\phi dxdt + \tau\int_0^T \int_\Omega H(u)\nabla\partial_t u\cdot\nabla\phi dxdt = 0, \end{aligned} \tag{4.0.6}$$

for any $\phi \in L^2(0,T;W_0^{1,2}(\Omega))$.

As $H(u)$ vanishes at $u = 0$ and 1, (4.0.3) becomes degenerate. We define the functions

$$G,\Gamma : \mathbb{R} \to \mathbb{R}\cup\{\pm\infty\}, \qquad G(u) = \int_{C_D}^u \frac{1}{H(v)}dv, \text{ and } \Gamma(u) = \int_{C_D}^u G(v)dv. \tag{4.0.7}$$

Clearly, Γ is a convex function satisfying $\Gamma(C_D) = \Gamma'(C_D) = 0$, implying

$$\Gamma(u) \ge 0, \quad \text{for all} \quad u \in \mathbb{R}. \tag{4.0.8}$$

The existence results in the following sections are obtained under the assumption

$$\int_\Omega \Gamma(u^0)dx < \infty. \tag{A2}$$

For the particular function H in (4.0.4), this assumption is fulfilled if, for example, $0 < p,q < 1$. Whenever $p \ge 1$, (A2) requires that meas$\{u^0 = 0\} = 0$. Similarly, $q \ge 1$

requires meas$\{u^0 = 1\} = 0$. The construction of Γ is inspired by [27, 28], where a generalized Kullback entropy is defined. As will be proved below, (A2) implies

$$\int_\Omega \Gamma(u(t))dx < C,$$

uniformly for $t \in (0, T]$.

The main result of this paper is Theorem 4.2.1, providing the existence of weak solutions to Problem P. Further, the resulting solution is essentially bounded by 0 and 1. As mentioned before, a similar existence result is proved in [28], but in a more restrictive framework. Specifically, the assumptions in Theorem 5 of [28] imply an asymptotic behavior of the nonlinear function H that can be related to sufficiently large exponents p, q in (4.0.4). The proof here works for any C^1-function H that is non negative. For example, one can take $p, q > 0$ in (4.0.4). Further, by Assumption (H6) in [28] the initial data can not take any degeneracy value (0 or 1) on a non-zero measure subset of Ω, which is allowed here if H behaves sub-quadratically close to 0 or 1 ($p < 1$, respectively $q < 1$).

Remark 4.0.1 *Equation* (4.0.3) *is a simplified model for two-phase flow in porous media, where dynamic effects are taken into account in the capillary pressure. However, this model contains the main mathematical difficulties related to such models: a degenerate nonlinearity in the terms involving the higher order derivatives. More realistic models are proposed in [19, 31]. With minor modifications, the present analysis can be extended for dealing with the cases considered e.g. in [9, 10, 30]. For instance, a capillary pressure of the form*

$$p_c = p(u) + \tau\partial_t u$$

may be treated following the ideas presented below, provided that p is increasing, with $\sqrt{p'} \in L^1(0, 1)$ *and* $H(\cdot)p'(\cdot) \in L^\infty(0, 1)$. *In particular the degeneracy* $p'(u) = 0$ *for some u is allowed, as well as* $\lim_{s\to\{0,1\}} p'(s) = +\infty$. *Note that under these finer assumptions, the definition of the solution to the Problem P has to be modified slightly (see [7]).*

We start by studying a regularized problem in Section 4.1, where we replace H by the strictly positive function $H_\delta = H + \delta$. Some a priori estimates are provided in Section 4.1 and the existence of weak solutions for H_δ is proved. In Section 4.2, the existence of weak solutions to equation (4.0.3) is proved by compactness arguments.

4.1 The regularized problem

To overcome the problems that are due to the degeneracy, we regularize Problem P by perturbing $H(u)$:

$$H_\delta(u) = H(u) + \delta, \tag{4.1.1}$$

where δ is a small positive number. Then we consider the equation:

$$\partial_t u + \nabla \cdot \mathbf{F}(u, x, t) = \nabla \cdot \Big(H_\delta(u) \nabla(u + \tau \partial_t u)\Big), \tag{4.1.2}$$

and investigate the limit case as $\delta \to 0$. In particular, we seek a solution to the following **Problem $\mathbf{P}_\delta$** Find $u \in W^{1,2}(0, T; V)$ such that $u(\cdot, 0) = u^0$, $\nabla\partial_t u \in L^2(0, T; L^2(\Omega)^d)$ and

$$\begin{aligned} &\int_0^T \int_\Omega \partial_t u\phi dxdt - \int_0^T \int_\Omega \mathbf{F}(u, x, t) \cdot \nabla\phi dxdt \\ &+ \int_0^T \int_\Omega H_\delta(u)\nabla u \cdot \nabla\phi dxdt + \tau \int_0^T \int_\Omega H_\delta(u)\nabla\partial_t u \cdot \nabla\phi dxdt = 0, \end{aligned} \tag{4.1.3}$$

for any $\phi \in L^2(0, T; W_0^{1,2}(\Omega))$.

Clearly, any solution to Problem P_δ depends on δ. However within Section 4.1, δ will be fixed. For the ease of reading, the δ-subscript will be omitted. We start by showing that P_δ has a solution. To do so, we use the Rothe method [22] and investigate firstly a sequence of time discrete problems.

4.1.1 The time discretization

Setting $\Delta t = T/N (N \in \mathbb{N})$, we consider the Euler-implicit discretization of Problem P_δ which leads to a sequence of time discretized problems. Specifically, we consider
Problem $\mathbf{P}_\delta^{n+1}$ Let $n \in \{0, 1, 2, ..., N-1\}$, and $u^n \in V$ given. Find $u^{n+1} \in V$ such that

$$\begin{aligned} &(u^{n+1} - u^n, \phi) + \Delta t\Big(\nabla \cdot \mathbf{F}(u^{n+1}, x, t), \phi\Big) + \Delta t\Big(H_\delta(u^{n+1})\nabla u^{n+1}, \nabla\phi\Big) \\ &+\tau\Big(H_\delta(u^{n+1})\nabla(u^{n+1} - u^n), \nabla\phi\Big) = 0, \end{aligned} \tag{4.1.4}$$

for any $\phi \in W_0^{1,2}(\Omega)$.

For obtaining estimates we will use the elementary Young inequality

$$ab \le \frac{1}{2\delta}a^2 + \frac{\delta}{2}b^2, \quad \text{for any} \quad a, b \in \mathbb{R} \quad \text{and} \quad \delta > 0. \tag{4.1.5}$$

We have the following result:

Proposition 4.1.1 *Problem P_δ^{n+1} has a solution.*

For proving Proposition 4.1.1, we introduce the auxiliary problem:

Problem $\mathbf{P}_{aux}$ Given $v_1, v_2 \in V$, Find $w \in V$ such that

$$\begin{aligned}&(\Delta t + \tau)\Big(H_\delta(w)\nabla w, \nabla\phi\Big) + (w, \phi) + \Delta t\Big(\nabla \cdot \mathbf{F}(w, x, t), \phi\Big)\\&= (v_2, \phi) + \tau\Big(H_\delta(v_1)\nabla v_2, \nabla\phi\Big),\end{aligned} \tag{4.1.6}$$

for any $\phi \in W_0^{1,2}(\Omega)$.
Then we have the following:

Lemma 4.1.1 *Problem P_{aux} has a solution.*

Proof . Define

$$\mathcal{G}(y) := \int_{C_D}^{y} H_\delta(s)ds, \tag{4.1.7}$$

then we have

$$\frac{1}{1+\delta} \le (\mathcal{G}^{-1})' = \frac{1}{\mathcal{G}'} \le \frac{1}{\delta}. \tag{4.1.8}$$

and

$$(u_1 - u_2)(\mathcal{G}(u_1) - \mathcal{G}(u_2)) \ge \delta|u_1 - u_2|^2, \tag{4.1.9}$$

Further, define

$$\begin{aligned}&a : W_0^{1,2}(\Omega) \times W_0^{1,2}(\Omega) \to \mathbb{R},\\&a(v, \phi) = \Big(\mathcal{G}^{-1}(v), \phi\Big) + \Delta t\Big(\nabla \cdot \mathbf{F}(\mathcal{G}^{-1}(v), x, t), \phi\Big) + (\Delta t + \tau)(\nabla v, \nabla\phi),\\&b : W_0^{1,2}(\Omega) \to \mathbb{R}, \quad b(\phi) = (v_2, \phi) + \tau\Big(H_\delta(v_1)\nabla v_2, \nabla\phi\Big).\end{aligned}$$

Clearly, b is a linear bounded functional and for each $v \in W_0^{1,2}(\Omega)$, $\phi \mapsto a(v,\phi)$ is a linear bounded functional. Furthermore, for small enough Δt

$$\begin{aligned} a(v_1, v_1 - v_2) - a(v_2, v_1 - v_2) \geq &(\Delta t + \tau)\|\nabla(v_1 - v_2)\|^2_{L^2(\Omega)} + \frac{1}{\delta}\|v_1 - v_2\|^2_{L^2(\Omega)} \\ &- \frac{L\Delta t}{\delta}\|\nabla(v_1 - v_2)\|_{L^2(\Omega)} \cdot \|v_1 - v_2\|_{L^2(\Omega)} \\ \geq &C\|v_1 - v_2\|^2_{W^{1,2}(\Omega)}, \end{aligned}$$

and it is easy to check that

$$|a(v_1,\phi) - a(v_2,\phi)| \leq C\|v_1 - v_2\|_{W^{1,2}(\Omega)} \cdot \|\phi\|_{W^{1,2}(\Omega)}.$$

Therefore, the nonlinear Lax-Milgram theorem ([38], pp.174-175) provides the existence and uniqueness of a solution $v \in W_0^{1,2}(\Omega)$ such that

$$\begin{aligned} &\Big(\mathcal{G}^{-1}(v),\phi\Big) + \Delta t\Big(\nabla\cdot\mathbf{F}(\mathcal{G}^{-1}(v),x,t),\phi\Big) + (\Delta t + \tau)(\nabla v, \nabla\phi) \\ &= (v_2,\phi) + \tau\Big(H_\delta(v_1)\nabla v_2, \nabla\phi\Big), \end{aligned}$$

for any $\phi \in W_0^{1,2}(\Omega)$. Taking $w = \mathcal{G}^{-1}(v) \in V$ gives a solution to Problem P_{aux}. □

Proof of Proposition 4.1.1. Using Lemma 4.1.1, for given $u^n \in V$ and $v_i \in V$ $(i \in \mathbb{N})$, there exists a $v_{i+1} \in V$ such that

$$\begin{aligned} &(\Delta t + \tau)\Big(H_\delta(v_{i+1})\nabla v_{i+1}, \nabla\phi\Big) + (v_{i+1},\phi) + \Delta t\Big(\nabla\cdot\mathbf{F}(v_{i+1},x,t),\phi\Big) \\ &= (u^n,\phi) + \tau\Big(H_\delta(v_i)\nabla u^n, \nabla\phi\Big), \end{aligned} \tag{4.1.10}$$

for any $\phi \in W_0^{1,2}(\Omega)$. Therefore we construct a sequence $\{v_i\}_{i\in\mathbb{N}}$ by iteration (for simplicity, here we take $v_1 = u^n$).
Next, taking $\phi = v_{i+1} - C_D$ in (4.1.10) leads to

$$\begin{aligned} &(\Delta t + \tau)\|\sqrt{H_\delta(v_{i+1})}\nabla v_{i+1}\|^2_{L^2(\Omega)} + \|v_{i+1}\|^2_{L^2(\Omega)} \\ &= (C_D, v_{i+1}) + (u^n, v_{i+1}) + \tau\Big(H_\delta(v_i)\nabla u^n, \nabla v_{i+1}\Big), \end{aligned} \tag{4.1.11}$$

where we used

$$\Big(\nabla\cdot\mathbf{F}(v_{i+1},x,t), v_{i+1} - C_D\Big) = -\int_\Omega \mathbf{F}(v_{i+1},x,t)\cdot\nabla v_{i+1}dx = -\int_{\partial\Omega}\gamma\cdot\mathcal{F}(C_D)dx = 0,$$

with $\mathcal{F}(w) := \int_{C_D}^{w} \mathbf{F}(v,x,t)dv$ and γ denoting the outer normal vector to $\partial\Omega$. Therefore,

$$\begin{aligned}
&(\Delta t+\tau)\|\sqrt{H_\delta(v_{i+1})}\nabla v_{i+1}\|^2_{L^2(\Omega)} + \|v_{i+1}\|^2_{L^2(\Omega)}\\
&\le \frac{1}{4}\|v_{i+1}\|^2_{L^2(\Omega)} + C_D^2\text{meas}(\Omega) + \frac{1}{2}\|v_{i+1}\|^2_{L^2(\Omega)} + \frac{1}{2}\|u^n\|^2_{L^2(\Omega)}\\
&+\frac{(\Delta t+\tau)\delta}{2}\|\nabla v_{i+1}\|^2_{L^2(\Omega)} + \frac{\tau^2(1+\delta)^2}{2(\Delta t+\tau)\delta}\|\nabla u^n\|^2_{L^2(\Omega)}.
\end{aligned}$$

Using $u^n \in V$ and Ω is bounded domain, we obtain the following:

$$\frac{(\Delta t+\tau)\delta}{2}\|\nabla v_{i+1}\|^2_{L^2(\Omega)} + \frac{1}{4}\|v_{i+1}\|^2_{L^2(\Omega)} \le C. \tag{4.1.12}$$

Here C is a positive constant, which does not depend on i. Hence, there exists a $v \in V$ such that $v_i \rightharpoonup v$ weakly. Consequently, $v_i \to v$ strongly in $L^2(\Omega)$ and for any $\phi \in W_0^{1,2}(\Omega)$

$$(v_i,\phi) \to (v,\phi),$$
$$\Big(\nabla\cdot\mathbf{F}(v_i,x,t),\phi\Big) \to \Big(\nabla\cdot\mathbf{F}(v,x,t),\phi\Big).$$

Next, we prove

$$\Big(H_\delta(v_i)\nabla v_i,\nabla\phi\Big) \to \Big(H_\delta(v)\nabla v,\nabla\phi\Big).$$

The idea in proving the above limit will be used again later. We start by observing that $H_\delta(v_i)\nabla v_i$ is bounded in $(L^2(\Omega))^d$, therefore it has a weak limit χ. To identify this in $L^2(\Omega)$, we take any $\phi \in C_0^\infty(\Omega)$ as test function. Since $H_\delta(v_i) \to H_\delta(v)$ strongly in $L^2(\Omega)$ and $\nabla v_i \rightharpoonup \nabla v$ weakly in $(L^2(\Omega))^d$, we have

$$\Big(H_\delta(v_i)\nabla v_i,\nabla\phi\Big) \to \Big(H_\delta(v)\nabla v,\nabla\phi\Big).$$

This implies that $H_\delta(v_i)\nabla v_i \rightharpoonup H_\delta(v)\nabla v$ in distributional sense. By the uniqueness of the limit, we have $\chi = H_\delta(v)\nabla v$. Using the same idea, we prove

$$\Big(H_\delta(v_i)\nabla u^n,\nabla\phi\Big) \to \Big(H_\delta(v)\nabla u^n,\nabla\phi\Big).$$

Passing to the limit $i \to +\infty$ in (4.1.10), we have

$$(\Delta t+\tau)\Big(H_\delta(v)\nabla v,\nabla\phi\Big)+(v,\phi)+\Delta t\Big(\nabla\cdot\mathbf{F}(v,x,t),\phi\Big) = (u^n,\phi)+\tau\Big(H_\delta(v)\nabla u^n,\nabla\phi\Big). \tag{4.1.13}$$

Taking $u^{n+1} = v$ gives a solution to Problem P_δ^{n+1}. □

In proving the existence of a solution to Problem P_δ, we use the following elementary results

Proposition 4.1.2 *Let $k \in \{0, 1, ..., N\}, m \geq 1$. For any set of m-dimensional real vectors $\boldsymbol{a}^k, \boldsymbol{b}^k \in \mathbb{R}^m$, we have the following identities:*

$$\sum_{k=1}^{N} < \boldsymbol{a}^k - \boldsymbol{a}^{k-1}, \sum_{n=k}^{N} \boldsymbol{b}^n >= \sum_{k=1}^{N} < \boldsymbol{a}^k, \boldsymbol{b}^k > - < \boldsymbol{a}^0, \sum_{k=1}^{N} \boldsymbol{b}^k >, \tag{4.1.14}$$

$$\sum_{k=1}^{N} < \boldsymbol{a}^k - \boldsymbol{a}^{k-1}, \boldsymbol{a}^k >= \frac{1}{2}(|\boldsymbol{a}^N|^2 - |\boldsymbol{a}^0|^2 + \sum_{k=1}^{N} |\boldsymbol{a}^k - \boldsymbol{a}^{k-1}|^2), \tag{4.1.15}$$

$$\sum_{k=1}^{N} < \sum_{k=n}^{N} \boldsymbol{a}^k, \boldsymbol{a}^n >= \frac{1}{2}|\sum_{k=1}^{N} \boldsymbol{a}^k|^2 + \frac{1}{2}\sum_{k=1}^{N} |\boldsymbol{a}^k|^2. \tag{4.1.16}$$

4.1.2 A priori estimates

For the existence of a solution to Problem P_δ, we apply compactness arguments based on the following a priori estimates.

Proposition 4.1.3 *For any $n \geq 1$, we have the following:*

$$\|\nabla u^n\|_{L^2(\Omega)} \leq C, \tag{4.1.17}$$

$$\int_\Omega \Gamma_\delta(u^n)dx \leq C, \tag{4.1.18}$$

$$\|u^n - u^{n-1}\|^2_{L^2(\Omega)} + \tau\| \sqrt{H_\delta(u^n)}\nabla(u^n - u^{n-1})\|^2_{L^2(\Omega)} \leq C(\Delta t)^2, \tag{4.1.19}$$

$$\|u^n\|_{L^2(\Omega)} \leq C. \tag{4.1.20}$$

Here C does not depend on δ.

Proof . **1.** Taking $\phi = G_\delta(u^{n+1}) = \int_{C_D}^{u^{n+1}} \frac{1}{H_\delta(v)}dv \in W_0^{1,2}(\Omega)$ in (4.1.4) gives

$$\begin{aligned} &\Big(u^{n+1} - u^n, G_\delta(u^{n+1})\Big) + (\tau + \Delta t)\|\nabla u^{n+1}\|^2_{L^2(\Omega)} \\ -&\tau(\nabla u^n, \nabla u^{n+1}) + \Delta t\Big(\nabla \cdot \mathbf{F}(u^{n+1}, x, t), G_\delta(u^{n+1})\Big) = 0. \end{aligned} \tag{4.1.21}$$

Define $\mathcal{G}(u^{n+1},x,t) := \int_{C_D}^{u^{n+1}} G_\delta(v)\partial_v \mathbf{F}(v,x,t)dv$. By (A1) we have

$$\left(\nabla\cdot\mathbf{F}(u^{n+1},x,t), G_\delta(u^{n+1})\right) = \int_\Omega \nabla\cdot\mathcal{G}(u^{n+1},x,t)dx = \int_{\partial\Omega}\gamma\cdot\mathcal{G}(C_D)dx = 0.$$

Here γ denotes the outer normal vector to $\partial\Omega$. Further, as in (4.0.7) we define $\Gamma_\delta(u) := \int_{C_D}^u G_\delta(v)dv$ and note that $\Gamma_\delta''(u) = \frac{1}{H_\delta(u)} > 0$, thus

$$(u^{n+1}-u^n)G_\delta(u^{n+1}) \geq \Gamma_\delta(u^{n+1}) - \Gamma_\delta(u^n). \tag{4.1.22}$$

Summing (4.1.22) in (4.1.21) up from 0 to $n-1$ gives

$$0 \geq \int_\Omega \Gamma_\delta(u^n)dx - \int_\Omega \Gamma_\delta(u^0)dx + (\Delta t+\tau)\sum_{k=1}^n \|\nabla u^k\|^2_{L^2(\Omega)} - \tau\sum_{k=1}^n(\nabla u^k,\nabla u^{k-1}). \tag{4.1.23}$$

By (4.1.15) we have

$$\begin{aligned} 0 \geq &\int_\Omega \Gamma_\delta(u^n)dx - \int_\Omega \Gamma_\delta(u^0)dx + \Delta t\sum_{k=1}^n \|\nabla u^k\|^2_{L^2(\Omega)} \\ &+\frac{\tau}{2}\|\nabla u^n\|^2_{L^2(\Omega)} - \frac{\tau}{2}\|\nabla u^0\|^2_{L^2(\Omega)} + \frac{\tau}{2}\sum_{k=1}^n\|\nabla(u^k-u^{k-1})\|^2_{L^2(\Omega)}, \end{aligned} \tag{4.1.24}$$

implying

$$\begin{aligned} &\int_\Omega \Gamma_\delta(u^n)dx + \Delta t\sum_{k=1}^n \|\nabla u^k\|^2_{L^2(\Omega)} + \frac{\tau}{2}\|\nabla u^n\|^2_{L^2(\Omega)} + \frac{\tau}{2}\sum_{k=1}^n\|\nabla(u^k-u^{k-1})\|^2_{L^2(\Omega)} \\ &\leq \int_\Omega \Gamma_\delta(u^0)dx + \frac{\tau}{2}\|\nabla u^0\|^2_{L^2(\Omega)}. \end{aligned} \tag{4.1.25}$$

Recalling (4.0.8) and (A2), as H_δ is bounded and $u^0 \in W^{1,2}(\Omega)$, we have

$$\int_\Omega \Gamma_\delta(u^0)dx = \int_\Omega\int_{C_D}^{u^0}\int_{C_D}^u \frac{1}{H_\delta(v)}dvdudx \leq \int_\Omega\int_{C_D}^{u^0}\int_{C_D}^u\frac{1}{H(v)}dvdudx \leq C,$$

where C does not depend on δ. Therefore,

$$\int_\Omega \Gamma_\delta(u^n)dx \leq C, \tag{4.1.26}$$

$$\|\nabla u^n\|_{L^2(\Omega)} \leq C, \quad \sum_{k=1}^n \|\nabla(u^k-u^{k-1})\|^2_{L^2(\Omega)} \leq C. \tag{4.1.27}$$

2. Taking $\phi = u^n - u^{n-1} \in W_0^{1,2}(\Omega)$ in (4.1.4) written at time $t_n = n\Delta t$, we have

$$\begin{aligned}&\|u^n - u^{n-1}\|^2_{L^2(\Omega)} + \Delta t\Big(\nabla \cdot \mathbf{F}(u^n, x, t), u^n - u^{n-1}\Big) + \\ &\Delta t\Big(H_\delta(u^n)\nabla u^n, \nabla(u^n - u^{n-1})\Big) + \tau\|\sqrt{H_\delta(u^n)}\nabla(u^n - u^{n-1})\|^2_{L^2(\Omega)} = 0.\end{aligned} \tag{4.1.28}$$

By (4.1.5) and (A1),

$$\begin{aligned}&\|u^n - u^{n-1}\|^2_{L^2(\Omega)} - \frac{1}{2}\|u^n - u^{n-1}\|^2_{L^2(\Omega)} - \frac{(C\Delta t)^2}{2}\|\nabla u^n\|^2_{L^2(\Omega)} \\ &-\frac{(\Delta t)^2}{2\tau}\|\sqrt{H_\delta(u^n)}\nabla u^n\|^2_{L^2(\Omega)} - \frac{\tau}{2}\|\sqrt{H_\delta(u^n)}\nabla(u^n - u^{n-1})\|^2_{L^2(\Omega)} \\ &+\tau\|\sqrt{H_\delta(u^n)}\nabla(u^n - u^{n-1})\|^2_{L^2(\Omega)} \le 0.\end{aligned} \tag{4.1.29}$$

According to (4.1.27), since H_δ is bounded, we obtain

$$\|u^n - u^{n-1}\|^2_{L^2(\Omega)} + \tau\|\sqrt{H_\delta(u^n)}\nabla(u^n - u^{n-1})\|^2_{L^2(\Omega)} \le C(\Delta t)^2. \tag{4.1.30}$$

As $H_\delta \ge \delta$, we also derive

$$\|u^n - u^{n-1}\|_{L^2(\Omega)} \le C\Delta t \quad \text{and} \quad \|\nabla(u^n - u^{n-1})\|_{L^2(\Omega)} \le \frac{C\Delta t}{\sqrt{\delta}}. \tag{4.1.31}$$

3. Finally, since $u^n - C_D \in W_0^{1,2}(\Omega)$,

$$\|u^n\|_{L^2(\Omega)} \le \|u^n - C_D\|_{L^2(\Omega)} + \|C_D\|_{L^2(\Omega)} \le C(\Omega)\|\nabla(u^n - C_D)\|_{L^2(\Omega)} + C \le C. \quad \square \tag{4.1.32}$$

4.1.3 Existence for Problem P_δ

Using Proposition 4.1.3, we now prove the existence of a solution to the regularized Problem P_δ.

Theorem 4.1.1 *Problem P_δ has a solution.*

Proof . We start by defining

$$U_N(t) = u^{k-1} + \frac{t - t^{k-1}}{\Delta t}(u^k - u^{k-1}), \tag{4.1.33}$$

for $t^{k-1} = (k-1)\Delta t \le t < t^k = k\Delta t, k = 1, 2...N$. Clearly, $U_N|_{\partial\Omega} = C_D$. Then we have

$$\begin{aligned}
\int_0^T \|U_N(t)\|^2_{L^2(\Omega)} dt &= \sum_{k=1}^N \int_{t^{k-1}}^{t^k} \|u^{k-1} + \frac{t-t^{k-1}}{\Delta t}(u^k - u^{k-1})\|^2_{L^2(\Omega)} dt \\
&\le 2 \sum_{k=1}^N \int_{t^{k-1}}^{t^k} \left(\|u^{k-1}\|^2_{L^2(\Omega)} + \|u^k - u^{k-1}\|^2_{L^2(\Omega)}\right) dt \\
&= 2\Delta t \sum_{k=1}^N \left(\|u^{k-1}\|^2_{L^2(\Omega)} + \|u^k - u^{k-1}\|^2_{L^2(\Omega)}\right) \\
&\le C,
\end{aligned} \tag{4.1.34}$$

and

$$\begin{aligned}
\int_0^T \|\nabla U_N(t)\|^2_{L^2(\Omega)} dt &= \sum_{k=1}^N \int_{t^{k-1}}^{t^k} \|\nabla u^{k-1} + \frac{t-t^{k-1}}{\Delta t}\nabla(u^k - u^{k-1})\|^2_{L^2(\Omega)} dt \\
&\le 2 \sum_{k=1}^N \int_{t^{k-1}}^{t^k} \left(\|\nabla u^{k-1}\|^2_{L^2(\Omega)} + \|\nabla(u^k - u^{k-1})\|^2_{L^2(\Omega)}\right) dt \\
&= 2\Delta t \sum_{k=1}^N \left(\|\nabla u^{k-1}\|^2_{L^2(\Omega)} + \|\nabla(u^k - u^{k-1})\|^2_{L^2(\Omega)}\right) \\
&\le C.
\end{aligned} \tag{4.1.35}$$

Additionally,

$$\begin{aligned}
\int_0^T \|\partial_t U_N\|^2_{L^2(\Omega)} dt &= \sum_{k=1}^N \int_{t^{k-1}}^{t^k} \|\frac{1}{\Delta t}(u^k - u^{k-1})\|^2_{L^2(\Omega)} dt \\
&= \frac{1}{\Delta t} \sum_{k=1}^N \|u^k - u^{k-1}\|^2_{L^2(\Omega)} \le C
\end{aligned} \tag{4.1.36}$$

and, by (4.1.31),

$$\begin{aligned}
\int_0^T \|\partial_t \nabla U_N\|^2_{L^2(\Omega)} dt &= \sum_{k=1}^N \int_{t^{k-1}}^{t^k} \|\frac{1}{\Delta t}\nabla(u^k - u^{k-1})\|^2_{L^2(\Omega)} dt, \\
&= \frac{1}{\Delta t} \sum_{k=1}^N \|\nabla(u^k - u^{k-1})\|^2_{L^2(\Omega)} \le \frac{C}{\delta}.
\end{aligned} \tag{4.1.37}$$

By (4.1.34), (4.1.35), (4.1.36), (4.1.37), there exists a subsequence of $\{U_N\}$ (still denoted as $\{U_N\}$) such that, as $N \to \infty$,

$$U_N \to U \quad \text{strongly in} \quad L^2(Q), \tag{4.1.38}$$

$$\partial_t U_N \rightharpoonup \partial_t U \quad \text{weakly in} \quad L^2(Q), \tag{4.1.39}$$

$$\nabla U_N \rightharpoonup \nabla U \quad \text{weakly in} \quad L^2(0,T;L^2(\Omega)^d), \tag{4.1.40}$$

$$\nabla \partial_t U_N \rightharpoonup \nabla \partial_t U \quad \text{weakly in} \quad L^2(0,T;L^2(\Omega)^d). \tag{4.1.41}$$

Now we prove that U solves Problem P_δ. Firstly, for any $\phi \in L^2(0,T;W_0^{1,2}(\Omega))$, (4.1.4) implies

$$\begin{aligned}&\left(\frac{u^k-u^{k-1}}{\Delta t}, \int_{t^{k-1}}^{t^k}\phi dt\right) + \left(\nabla\cdot\mathbf{F}(u^k,x,t), \int_{t^{k-1}}^{t^k}\phi dt\right)\\ &+\left(H_\delta(u^k)\nabla u^k, \int_{t^{k-1}}^{t^k}\nabla\phi dt\right) + \tau\left(H_\delta(u^k)\nabla\frac{u^k-u^{k-1}}{\Delta t}, \int_{t^{k-1}}^{t^k}\nabla\phi dt\right) = 0,\end{aligned} \tag{4.1.42}$$

for $k = 1, 2, \ldots, N$. Define

$$\overline{U}_N(t) = u^k, \tag{4.1.43}$$

for $\quad t^{k-1} = (k-1)\Delta t \le t < t^k = k\Delta t, k = 1, 2...N$. Then $\overline{U}_N|_{\partial\Omega} = C_D$ and

$$\begin{aligned}&\int_0^T\int_\Omega \partial_t U_N \phi dxdt - \int_0^T\int_\Omega \mathbf{F}(\overline{U}_N,x,t)\cdot\nabla\phi dxdt\\ &+\int_0^T\int_\Omega H_\delta(\overline{U}_N)\nabla\overline{U}_N\cdot\nabla\phi dxdt + \tau\int_0^T\int_\Omega H_\delta(\overline{U}_N)\nabla\partial_t U_N\cdot\nabla\phi dxdt = 0.\end{aligned} \tag{4.1.44}$$

We now exploit a general principle that relates the piecewise linear and the piecewise constant interpolation (see e.g. [25] for a proof of the corresponding lemma): if one interpolation converges strongly in $L^2(Q)$, then the other interpolation also converges strongly in $L^2(Q)$. From the convergence of U_N, we conclude that $\overline{U}_N$ also converges strongly in $L^2(Q)$. Then we obtain $F(\overline{U}_N) \to F(U)$ strongly in $(L^2(Q))^d$ and $H_\delta(\overline{U}_N) \to H_\delta(U)$ strongly in $L^2(Q)$. Employing the same idea as in the proof of Proposition 4.1.1, we have

$$H_\delta(\overline{U}_N)\nabla\overline{U}_N \rightharpoonup H_\delta(U)\nabla U \quad \text{weakly in } L^2(0,T;L^2(\Omega)^d), \tag{4.1.45}$$

$$H_\delta(\overline{U}_N)\nabla\partial_t U_N \rightharpoonup H_\delta(U)\nabla\partial_t U \quad \text{weakly in } L^2(0,T;L^2(\Omega)^d). \tag{4.1.46}$$

Combining the latter results with (4.1.44), we obtain that U is a solution to Problem P_δ.
□

4.2 Existence for Problem P

For any $\delta > 0$, Section 4.1 provides a solution u_δ to the regularized Problem P_δ. In this section, we identify a sequence $\{\delta_n\}_{n\in\mathbb{N}}$ tending to 0, providing the limit u of the sequence $\{u_{\delta_n}\}_{n\in\mathbb{N}}$, which solves Problem P. This involves a compactness argument, and therefore convergence should always be understood along a subsequence. From Assumption (A.2), Proposition 4.1.3 and Theorem 4.1.1, we have the following

Proposition 4.2.1 *We have the following estimates:*

$$\|u_\delta\|_{L^2(Q)} \le C, \tag{4.2.1}$$
$$\|\partial_t u_\delta\|_{L^2(Q)} \le C, \tag{4.2.2}$$
$$\|\sqrt{H_\delta(u_\delta)}\nabla\partial_t u_\delta\|_{L^2(0,T;(L^2(\Omega))^d)} \le C, \tag{4.2.3}$$
$$\|\nabla u_\delta\|_{L^\infty\left(0,T;(L^2(\Omega))^d\right)} \le C, \tag{4.2.4}$$
$$\int_\Omega \Gamma_\delta(u_\delta(t))dx \le C, \quad \textit{for a.e.} \quad t > 0, \tag{4.2.5}$$

where C does not depend on δ.

By Proposition 4.2.1, there exists a $u \in H^1(Q)$ such that,

$$u_{\delta_n} \to u \quad \text{strongly in } L^2(Q), \quad \text{and a.e. on Q}, \tag{4.2.6}$$
$$\partial_t u_{\delta_n} \rightharpoonup \partial_t u \quad \text{weakly in } L^2(Q), \tag{4.2.7}$$

as well as

$$\nabla u_{\delta_n} \rightharpoonup \nabla u \quad \text{in } L^\infty(0,T;(L^2(\Omega))^d) \quad \text{in the weak-}\star \text{ sense}. \tag{4.2.8}$$

Further, from (4.2.3) there exists a $\zeta = (\zeta_1, ..., \zeta_d) \in L^2(0,T;L^2(\Omega)^d)$ such that,

$$\sqrt{H_{\delta_n}(u_{\delta_n})}\partial_t \nabla u_{\delta_n} \rightharpoonup \zeta \quad \text{weakly in } L^2(0,T;L^2(\Omega)^d). \tag{4.2.9}$$

Let $\psi \in C_0^\infty(Q)$, then for all n, u_{δ_n} satisfies

$$A_n + B_n + C_n + D_n = 0, \tag{4.2.10}$$

where

$$\begin{aligned}
A_n &= \iint_Q \partial_t u_{\delta_n} \psi dxdt,\\
B_n &= -\iint_Q \mathbf{F}(u_{\delta_n}, x, t) \cdot \nabla\psi dxdt,\\
C_n &= \iint_Q H_{\delta_n}(u_{\delta_n})\nabla u_{\delta_n} \cdot \nabla\psi dxdt,\\
D_n &= \iint_Q H_{\delta_n}(u_{\delta_n})\partial_t \nabla u_{\delta_n} \cdot \nabla\psi dxdt.
\end{aligned}$$

In view of the above, A_n, B_n and C_n converge to the desired limit as $n \to \infty$. We thus focus on the limit of D_n. To this end, let $j \in \{1, \dots, d\}$ be fixed and decompose the variable $x \in \mathbf{R}^d$ into $(x_j, \tilde{x}_j) \in \mathbf{R} \times \mathbf{R}^{d-1}$. Define

$$\Omega_j(\tilde{x}_j) := \{x_j \in \mathbf{R} \mid (x_j, \tilde{x}_j) \in \Omega\}, \quad \text{and} \quad Q_j(\tilde{x}_j) := \Omega_j(\tilde{x}_j) \times (0, T).$$

We note that

$$D_n = \sum_{j=1}^{d} \int_{\mathbf{R}^{d-1}} D_{j,n}(\tilde{x}_j) d\tilde{x}_j, \tag{4.2.11}$$

where, for a.e. $\tilde{x}_j \in \mathbf{R}^{d-1}$,

$$D_{j,n}(\tilde{x}_j) = \iint_{Q_j(\tilde{x}_j)} H_{\delta_n}(u_{\delta_n})\partial_t \partial_{x_j} u_{\delta_n} \partial_{x_j} \psi dx_j dt.$$

We have the following

Lemma 4.2.1 *For almost every $\tilde{x}_j \in \mathbf{R}^{d-1}$,*

$$\lim_{n\to\infty} D_{j,n}(\tilde{x}_j) = D_j(\tilde{x}_j) := \iint_{Q_j(\tilde{x}_j)} H(u)\partial_t \partial_{x_j} u \partial_{x_j} \psi dx_j dt.$$

Proof We deduce from Proposition 4.2.1 that, for almost every $\tilde{x}_j$,

$$\|\partial_{x_j} u_{\delta_n}(\cdot, \tilde{x}_j)\|_{L^2(Q_j(\tilde{x}_j))} \le C(\tilde{x}_j), \tag{4.2.12}$$
$$\|\sqrt{H(u_{\delta_n}(\cdot, \tilde{x}_j))}\partial_t \partial_{x_j} u_{\delta_n}(\cdot, \tilde{x}_j)\|_{L^2(Q_j(\tilde{x}_j))} \le C(\tilde{x}_j), \tag{4.2.13}$$
$$\|\partial_t u_{\delta_n}(\cdot, \tilde{x}_j)\|_{L^2(Q_j(\tilde{x}_j))} \le C(\tilde{x}_j), \tag{4.2.14}$$

where $C(\tilde{x}_j) \in L^2(\mathbf{R}^{d-1})$.

From (4.2.13) and in view of (4.2.9), we deduce

$$\sqrt{H(u_{\delta_n}(\cdot, \tilde{x}_j))}\partial_t \partial_{x_j} u_{\delta_n}(\cdot, \tilde{x}_j) \rightharpoonup \zeta_j(\tilde{x}_j) \quad \text{weakly in } L^2(Q_j(\tilde{x}_j)). \tag{4.2.15}$$

We define an auxiliary C^2 function $A : \mathbb{R} \to \mathbb{R}$ such that

$$\begin{aligned} &\frac{A}{\sqrt{H}} \in L^\infty(0,1), \quad \frac{A'}{\sqrt{H}} \in L^\infty(0,1), \quad A'' \in L^\infty(0,1), \\ &\qquad\qquad \text{and} \quad A(s) > 0 \text{ if } s \in (0,1). \end{aligned} \tag{4.2.16}$$

For instance, if $H(u) \sim u^{p+1}$ in the vicinity of 0 (as encountered e.g. in (4.0.4)), one can consider $A(u) \sim u^{max(1,(p+3)/2)}$. The construction in the vicinity of 1 is similar. Note that (4.2.16) implies that $A(\cdot)$ is 0 outside $(0,1)$. Furthermore, the fractions in (4.2.16) are extended by 0 outside $(0,1)$.

Define the differential operator $\tilde{\nabla} := (\partial_{x_j}, \partial_t)^T$, and, for fixed $\tilde{x}_j$ in a full measure subset of $\mathbf{R}^{d-1}$, the two vector-valued functions

$$V_n(\tilde{x}_j) = \Big(A'(u_{\delta_n}(\cdot, \tilde{x}_j))\partial_t u_{\delta_n}(\cdot, \tilde{x}_j), 0\Big), \qquad W_n(\tilde{x}_j) = \Big(\partial_{x_j} u_{\delta_n}(\cdot, \tilde{x}_j), \partial_t u_{\delta_n}(\cdot, \tilde{x}_j)\Big). \tag{4.2.17}$$

For reader's convenience, we remove the parameter $\tilde{x}_j$ in the sequel. By (4.2.12)–(4.2.14) and the properties of A, we obtain that V_n and W_n are uniformly bounded in $(L^2(Q_j))^2$. Since $\tilde{\nabla} \times W_n = \tilde{\nabla} \times (\tilde{\nabla} u_{\delta_n}) = 0$, so $\{\tilde{\nabla} \times W_n, n \in \mathbb{N}\}$ is a compact subset of $W^{-1,2}(Q_j)$.

Moreover, the sequence $\{\nabla \cdot V_n, n \in \mathbb{N}\}$ is uniformly bounded in $L^2(0,T; L^1(\Omega_j))$, as

$$\partial_{x_j}(A'(u_{\delta_n})\partial_t u_{\delta_n}) = A''(u_{\delta_n})\partial_t u_{\delta_n} \partial_{x_j} u_{\delta_n} + \frac{A'(u_{\delta_n})}{\sqrt{H(u_{\delta_n})}} \sqrt{H(u_{\delta_n})}\partial_t \partial_{x_j} u_{\delta_n}, \tag{4.2.18}$$

a.e. in $\omega_j(\tilde{x}_j) = \{(x_j, t) \in Q_j \mid \quad u(x_j, t, \tilde{x}_j) \in (0,1)\}$ and in fact in the entire Q_j in view of the extension of the fractions in (4.2.16). The embedding $L^2(0,T; L^1(\Omega_j)) \hookrightarrow W^{-1,2}(Q_j)$ being compact (note that $\Omega_j \subset \mathbf{R}$), then, applying the Div-Curl Lemma [29, 37], we get

$$V_n \cdot W_n = A'(u_{\delta_n})\partial_t u_{\delta_n}\partial_{x_j} u_{\delta_n} \rightharpoonup A'(u)\partial_t u \partial_{x_j} u \quad \text{weakly in} \quad \mathcal{D}'(Q_j). \tag{4.2.19}$$

Finally, let $\mathcal{A}$ be a primitive form of A. As before, the equality

$$\partial_t\partial_{x_j}\mathcal{A}(u_{\delta_n}) = A'(u_{\delta_n})\partial_t u_{\delta_n}\partial_{x_j} u_{\delta_n} + \frac{A(u_{\delta_n})}{\sqrt{H(u_{\delta_n})}}\sqrt{H(u_{\delta_n})}\partial_t\partial_{x_j} u_{\delta_n}, \tag{4.2.20}$$

holding a.e. in ω_j can be extended to Q_j. Since $\frac{A(u_{\delta_n})}{\sqrt{H(u_{\delta_n})}}$ converges a.e. in Q_j to $\frac{A(u)}{\sqrt{H(u)}}$ and is essentially bounded uniformly w.r.t. n, we obtain the strong convergence in $L^2(Q_j)$. Together with the weak convergence in (4.2.9), we pass to the limit ($n \to \infty$) in (4.2.20) and obtain

$$\partial_t\partial_{x_j}\mathcal{A}(u) = A'(u)\partial_t u\partial_{x_j} u + \frac{A(u)}{\sqrt{H(u)}}\zeta_j. \tag{4.2.21}$$

In the distributional sense, this implies

$$A'(u)\partial_t u\partial_{x_j} u + A(u)\partial_t\partial_{x_j} u = A'(u)\partial_t u\partial_{x_j} u + \frac{A(u)}{\sqrt{H(u)}}\zeta_j. \tag{4.2.22}$$

As a consequence, for almost every $\tilde{x}_j \in \mathbf{R}^{d-1}$,

$$\zeta_j(\tilde{x}_j) = \sqrt{H(u(\cdot, \tilde{x}_j))}\partial_t\partial_{x_j} u(\cdot, \tilde{x}_j). \tag{4.2.23}$$

Because of (4.2.15) and the strong $L^2(Q_j)$ convergence of $\sqrt{H_{\delta_n}(u_\delta(\cdot, \tilde{x}_j))}$ to $\sqrt{H(u(\cdot, \tilde{x}_j))}$, one has for almost every $\tilde{x}_j$ in $\mathbf{R}^{d-1}$,

$$\lim_{n\to\infty} D_{j,n}(\tilde{x}_j) = \iint_{Q_j(\tilde{x}_j)} \sqrt{H(u(\cdot, \tilde{x}_j))}\zeta_j(\tilde{x}_j)\partial_{x_j}\psi dx_j dt = D_j(\tilde{x}_j).$$

Proposition 4.2.2 *Let u be the limit in* (4.2.6)–(4.2.8). *Then, for all $\psi \in C_0^\infty(Q)$,*

$$\lim_{n\to\infty} \iint_Q H_{\delta_n}(u_{\delta_n})\partial_t \nabla u_{\delta_n} \cdot \nabla\psi dx dt = \iint_Q H(u)\partial_t \nabla u \cdot \nabla\psi dx dt. \tag{4.2.24}$$

Proof Note that, thanks to (4.2.11), for proving Proposition 4.2.2, it is sufficient to show that, for any $j \in \{1, \dots, d\}$,

$$\lim_{n\to\infty} \int_{\mathbf{R}^{d-1}} D_{j,n}(\tilde{x}_j) d\tilde{x}_j = \int_{\mathbf{R}^{d-1}} D_j(\tilde{x}_j) d\tilde{x}_j.$$

Since Ω is bounded, the functions $D_{j,n}$ are compactly supported. Further, the Cauchy-Schwarz inequality gives

$$\begin{aligned}\left(D_{j,n}(\tilde{x}_j)\right)^2 &\leq C\left(\iint_{Q_j(\tilde{x}_j)} H_{\delta_n}(u_{\delta_n})\left(\partial_t\partial_{x_j}u_{\delta_n}\right)^2 dx_j dt + \iint_{Q_j(\tilde{x}_j)} H_{\delta_n}(u_{\delta_n})\left(\partial_{x_j}\psi\right)^2 dx_j dt\right)\\ &\leq C\left(\iint_{Q_j(\tilde{x}_j)} H_{\delta_n}(u_{\delta_n})\left(\partial_t\partial_{x_j}u_{\delta_n}\right)^2 dx_j dt + \iint_{Q_j(\tilde{x}_j)} (\partial_{x_j}\psi)^2 dx_j dt\right),\end{aligned}$$

and therefore

$$\begin{aligned}\int_{\mathbf{R}^{d-1}} \left(D_{j,n}(\tilde{x}_j)\right)^2 d\tilde{x}_j &\leq C\left(\iint_Q H_{\delta_n}(u_{\delta_n})\left(\partial_t\partial_{x_j}u_{\delta_n}\right)^2 dxdt + \iint_Q (\partial_{x_j}\psi)^2 dxdt\right),\\ &\leq C\iint_Q H_{\delta_n}(u_{\delta_n})\left(\partial_t\partial_{x_j}u_{\delta_n}\right)^2 dxdt + C.\end{aligned}$$

By (4.2.3), $D_{j,n}$ is uniformly bounded in $L^2(\mathbf{R}^{d-1})$. Hence the sequence $\{D_{j,n}\}_n$ is equi-integrable. Now (4.2.24) follows by Lemma 4.2.1 and Vitali's theorem.

Theorem 4.2.1 *Problem P has a solution u. Furthermore, this solution is essentially bounded by 0 and 1 in Q.*

Proof Let u be the limit in (4.2.6)–(4.2.8). To show that u is a weak solution of Problem P, it is sufficient to show that the terms A_n, B_n, C_n, D_n in (4.2.10) have the desired limit. Identifying the limits of A_n, B_n, C_n is straightforward due to (4.2.6)–(4.2.8) and the strong L^2 convergence of $H_{\delta_n}(u_{\delta_n})$ to $H(u)$. For D_n we recall Proposition 4.2.2.

It remains to prove that $0 \leq u \leq 1$ a.e. in Q. To this end we consider $\epsilon > 0$ arbitrary, take $t \in (0, T)$, and define $\Omega^-_{\epsilon,n}(t) := \{x \in \Omega \mid u_{\delta_n}(x, t) < -\epsilon\}$. Then

$$\Gamma_{\delta_n}(u_{\delta_n}) = \int_{C_D}^{u_{\delta_n}} \int_{C_D}^{w} \frac{1}{H_{\delta_n}(v)} dv dw = \frac{(C_D - u_{\delta_n})^2}{2\delta_n}, \tag{4.2.25}$$

a.e. in $\Omega^-_{\epsilon,n}(t)$. Recalling (4.2.5), for all $\delta_n > 0$ and a.e. t, we write

$$C \geq \int_\Omega \Gamma_{\delta_n}(u_{\delta_n}(x,t))dx \geq \int_{\Omega^-_{\epsilon,n}(t)} \Gamma_{\delta_n}(u_{\delta_n}(x,t))dx \geq \frac{(C_D+\epsilon)^2}{2\delta_n}\ \mathrm{meas}(\Omega^-_{\epsilon,n}(t)). \tag{4.2.26}$$

Letting $\delta_n \to 0$, we obtain

$$\lim_{n\to\infty} \mathrm{meas}(\Omega^-_{\epsilon,n}(t)) = 0, \tag{4.2.27}$$

for a.e. $t \in (0,T]$. However, by (4.2.13) and (4.2.14), $u_{\delta_n} \to u$ in $C([0,T];L^2(\Omega))$, thus $u_{\delta_n}(\cdot,t) \to u(\cdot,t)$ a.e in Ω, for all t. Passing to the limit $\epsilon \to 0$ gives the lower bound for u. Similarly, we have $u \leq 1$ a.e., and the theorem is proved. □

References

[1] D.N. Arnold, J. Douglas Jr. and V. Thomée, *Superconvergence of a Finite Element Approximation to the Solution of a Sobolev Equation in a Single Space Variable*, Math. Comp. **36** (1981), 53–63.

[2] B.C. Aslan, W.W. Hager and S. Moskow, *A generalized eigenproblem for the Laplacian which arises in lightning*, J. Math. Anal. Appl. **341** (2008), 1028–1041.

[3] G.I. Barenblatt, T.W. Patzek, *The mathematical model of nonequilibrium effects in water-oil displacement*, SPE Journal **8**(2003), No.4, 409–416.

[4] G.I. Barenblatt, I.P. Zheltov and I.N. Kochina, *Basic concepts in the theory of seepage of homogeneous liquids in fissured rocks (strata)*, J. Appl. Math. Mech. **24** (1960), 1286–1303.

[5] S. Bottero, *Advances in theories of capillarity in porous media*, Ph'D Thesis, Utrecht University (2009).

[6] C. Cancès, *Asymptotic behavior of two-phase flows in heterogeneous porous media for capillarity depending only on space. II. Nonclassical shocks to model oil-trapping*, SIAM J. Math. Anal. **42**, 971–995.

[7] C. Cancès, C. Choquet, Y. Fan and I.S. Pop, *An existence result related to two-phase flows with dynamical capillary pressure*, MAMERN11 proceeding (2011).

[8] C. Choquet, *Parabolic and degenerate parabolic models for pressure-driven transport problems*, Math. Models Methods Appl. Sci. **20** (2010), 543–566.

[9] C. Cuesta, C.J. van Duijn, and J. Hulshof, *Infiltration in porous media with dynamic capillary pressure: Travelling waves*, European J. Appl. Math. **11** (2000), 381-397.

[10] C. Cuesta, J. Hulshof, *A model problem for groundwater flow with dynamic capillary pressure: stability of travelling waves*, Nonlinear Anal. **52** (2003), 1199–1218.

[11] C.M. Cuesta, I.S. Pop, *Numerical schemes for a pseudo-parabolic Burgers equation: Discontinuous data and long-time behaviour*, J. Comp. Appl. Math. **224** (2009), 269–283.

[12] D.A. DiCarlo, *Experimental measurements of saturation overshoot on infiltration*, Water Resour. Res. **40** (2004), W04215.1-W04215.

[13] C.J. van Duijn, Y. Fan, L.A. Peletier and I.S. Pop, *Travelling wave solutions for degenerate pseudo-parabolic equation modelling two-phase flow in porous media*, CASA Report **10-01**, Eindhoven University of Technology, 2010.

[14] C.J. van Duijn, L.A. Peletier and I.S. Pop, *A new class of entropy solutions of the Buckley-Leverett equation*, SIAM J. Math. Anal. **39** (2007), 507–536.

[15] W. Düll, *Some qualitative properties of solutions to a pseudoparabolic equation modeling solvent uptake in polymeric solids*, Comm. Partial Differential Equations **31** (2006), 1117-1138.

[16] R.E. Ewing, *Time-stepping Galerkin methods for nonlinear Sobolev partial differential equations*, SIAM J. Numer. Anal. **15** (1978), 1125–1150.

[17] Y. Fan and I.S. Pop, *A class of degenerate pseudo-parabolic equation: existence, uniqueness of weak solutions, and error estimates for the Euler-implicit discretization*, Mathematical Methods in the Applied Sciences **34** (2011), 2329-2339.

[18] W.H. Ford and T.W. Ting, *Unifrom error estimates for difference approximations to nonlinear pseudo-parabolic partial differential equations*, SIAM J. Numer. Anal. **11** (1974), 155–169.

[19] S.M. Hassanizadeh and W.G. Gray, *Thermodynamic basis of capillary pressure in porous media*, Water Resour. Res. **29** (1993), 3389–3405.

[20] R. Helmig, A. Weiss and B. Wohlmuth, *Dynamic capillary effects in heterogeneous porous media*, Comput. Geosci. **11** (2007), 261–274.

[21] V. Joekar-Niasar and S.M. Hassanizadeh, *Effect of fluids properties on non-equilibrium capillarity effects; dynamic pore-network modelling*, International Journal of Multiphase Flow, submitted.

[22] J. Kačur, *Method of Rothe in evolution equations*, Teubner Texts Math. **80**, Teubner Verlagsgesellschaft, Leipzig, 1985.

[23] F. Kissling, P.G. LeFloch and C. Rohde, *A kinetic decomposition for singular limits of non-local conservation laws*, J. Differential Equations **247**(2009), 3338–356.

[24] A. Lamacz, A. Rätz and B. Schweizer, *A well-posed hysteresis model for flows in porous media and applications to fingering effects*, Report **12** (2010), Fakultät für Mathematik, Universität Dortmund, 2010.

[25] M. Lenzinger and B. Schweizer, *Two-phase flow equations with outflow boundary conditions in the hydrophobic-hydrophilic case*, Nonlinear Anal. **73** (2010), 840–853.

[26] E. Milne, *The diffusion of imprisoned radiation through a gas*, J. London Math. Soc. **1** (1926), 40–51.

[27] A. Mikelić, H. Bruining, *Analysis of model equations for stress-enhanced diffusion in coal layers. I. Existence of a weak solution*, SIAM J. Math. Anal. **40** (2008), 1671–1691.

[28] A. Mikelić, *A global existence result for the equations describing unsatured flow in porous media with dynamic capillary pressure*, J. Differential Equations **248** (2010), 1561–1577.

[29] F. Murat, *Compacité par compensation*, Ann. Sc. Norm. Sup. Pisa **5** (1978), 489–507.

[30] J.L. Nieber, R.Z. Dautov, A.G. Egorov, A.Y. Sheshukov, *Dynamic capillary pressure mechanism for instability in gravity-driven flows; review and extension to very dry conditions*, Transp. Porous Med. **58** (2005), 147-172.

[31] D. Pavone, *Macroscopic equations derived from space averaging for immiscible two-Phase flow in porous media*, Oil & Gas Science and Technology - Rev. IFP **44**(1989), No. 1, 29–41.

[32] M. Peszyńska, S. Yi, *Numerical methods for unsaturated flow with dynamic capillary pressure in heterogeneous porous media*. Int. J. Numer. Anal. Model. **5** (2008), 126-149.

[33] M. Ptashnyk, *Nonlinear pseudoparabolic equations as sigular limit of reaction-diffusion equations*, Appl. Anal. **85** (2006), 1285–1299.

[34] A. Quarteroni, *Fourier spectral methods for pseudo-parabolic equations*, SIAM J. Numer. Anal. **24** (1987), 323–335.

[35] L.I. Rubinstein, *On the problem of the process of propagation of heat in heterogeneous media*, IZV. Akad. Nauk SSSR, Ser. Geogr. **1** (1948).

[36] R.E. Showalter, *A nonlinear parabolic-Sobolev equation*, Journal of Mathematical Analysis And Applications 50 (1975), pp. 183-190.

[37] Tartar, L. *Compensated compactness and applications to partial differential equations*, Nonlinear analysis and mechanics: Heriot-Watt Symposium, Vol. IV, Pitman, 1979, **39**, 136-212.

[38] E. Zeidler, *Applied Functional Analysis*, Springer-Verlag, 1999.

Chapter 5

Equivalent formulations and numerical schemes

In this chapter we study a mathematical model for two-phase flow in porous media, where dynamic effects are included in the difference of the phase pressure (see [12]). With a given maximal time $T > 0$ and for all $x \in \Omega$ (a bounded domain in $\mathbb{R}^d$), we consider the following equation

$$\partial_t u + \nabla \cdot \mathbf{F}(u) = \nabla \cdot \Big(H(u)\nabla p\Big), \quad (x,t) \in Q := \Omega \times (0,T]. \tag{5.0.1}$$

The scalar equation above is a simplified model, resulting from inserting the Darcy law for both phases into the mass conservation laws (see [2]), with u denoting the water saturation. In physical terms, the functions $\mathbf{F}$ and H represent the fraction flow, respectively the capillary induced diffusion function. Finally, p is the term expressing the pressure difference between the two phases.

This chapter has been submitted to ACOMEN (Advanced COmputational Methods in ENgineering) 2011, and it is a collaborative work with I.S. Pop

Based on experimental measurements, a monotonically increasing relationship between p and u can be determined (see [2]). In most of the cases, the measurements are carried out under equilibrium conditions. The models proposed in [12] take the dynamics into account by letting p depend on the time derivative of the saturation. Here we consider

$$p = p_c(u) + \tau\partial_t u, \tag{5.0.2}$$

with τ being a positive parameter. Then (5.0.1) becomes

$$\partial_t u + \nabla\cdot\mathbf{F}(u) = \nabla\cdot\Big(H(u)\nabla(p_c(u) + \tau\partial_t u)\Big), \quad (x,t)\in Q. \tag{5.0.3}$$

The model is completed by the initial and boundary conditions:

$$u(\cdot,0) = u^0 \text{ in } \Omega, \text{ and } u(\cdot,t) = 0 \text{ on } \partial\Omega, \text{ for } t\in(0,T]. \tag{5.0.4}$$

Formally, (5.0.3) can be transformed into a system form. First, one can use the particular form of the phase pressure difference to obtain

$$\begin{cases} \partial_t u + \nabla\cdot\mathbf{F}(u) = \nabla\cdot\Big(H(u)\nabla p\Big), \\ \qquad\qquad\quad p = p_c(u) + \tau\partial_t u, \end{cases} \quad \text{for all } (x,t)\in Q. \tag{5.0.5}$$

Next, since $\partial_t u = \big(p - p_c(u)\big)/\tau$, the first equation in (5.0.5) gives an elliptic equation for p, leading to

$$\begin{cases} \partial_t u + \nabla\cdot\mathbf{F}(u) = \nabla\cdot\Big(H(u)\nabla p\Big), \\ p - \tau\nabla\cdot\Big(H(u)\nabla p\Big) = p_c(u) - \tau\nabla\cdot\mathbf{F}(u), \end{cases} \quad \text{for all } (x,t)\in Q. \tag{5.0.6}$$

Note that whereas (5.0.3) is standard (conformal), (5.0.5) and (5.0.6) can be associated with a mixed formulation: in both cases the first equation is the mass balance equation, while the second one identifies the capillary pressure. The formulations (5.0.3), (5.0.5) and (5.0.6) are equivalent from formal point of view. However, these may lead to different numerical schemes, combining conservative approaches for (5.0.3) and the first equations in (5.0.5) or (5.0.6), and ordinary differential equations methods, respectively elliptic approaches. Such schemes are discussed in [6]. The aim of this chapter is to give a rigorous proof for the equivalence of the three formulations, and for the corresponding numerical discretizations.

In what follows standard notations in the functional analysis and the theory of partial differential equations are used: $L^2(\Omega)$, $W^{1,2}(\Omega)$ and $W_0^{1,2}(\Omega)$. Further, $W^{-1,2}(\Omega)$ denotes the dual of $W_0^{1,2}(\Omega)$, and for any Banach space X we let $L^2(0,T;X)$ be the Bochner space of X-valued functions. By $(\cdot,\cdot)$ we mean the inner product in either $L^2(\Omega)$, or $L^2(0,T;X)$ (if X is a Hilbert space), and $\|\cdot\|$ stands for the corresponding norm. Furthermore, C denotes a generic positive number.

The analysis below is carried out under the following assumptions:

- (A1). Ω is an open, bounded and connected domain in $\mathbb{R}^d$ $(d = 1,2,3)$, with Lipschitz continuous boundary $\partial\Omega$.

- (A2). τ is a given positive number.

- (A3). $\mathbf{F} : \mathbb{R} \to \mathbb{R}^d$, $H : \mathbb{R} \to \mathbb{R}$ and $p_c : \mathbb{R} \to \mathbb{R}$ are Lipschitz and C^1. p_c is monotonically increasing and $p_c(0) = 0$. For H, we assume $m_1 \leq H \leq M_1$, where $0 < m_1, M_1 < \infty$ are given constants. Further, we denote by L a Lipschitz constant for the functions F, H and p_c.

- (A4). The function u^0 lies in $W_0^{1,2}(\Omega)$.

Remark 5.0.1 *Taking $p_c(0) = 0$ in (A3) is not necessary needed, as only the pressure gradients are involved in the model. If $p_c(0) \neq 0$, one can eventually take $p = p_c(u) - p_c(0) + \tau\partial_t u$.*

5.1 Equivalent formulations

In this section we introduce three different formulations of the model in weak forms, consider their time discretization, and study the equivalence in both continuous and time discrete cases. In doing so we follow the ideas in [17]. Once this is achieved, existence, uniqueness and convergence results available for the original scalar model (see e. g. [3–6,9,14–16,18] and the references therein) can be directly transferred to the alternative system formulations.

5.1.1 The continuous case

In a weak form the problem (5.0.3)-(5.0.4) reads

Problem 1: Find $u \in W^{1,2}(0,T;W_0^{1,2}(\Omega))$ such that $u(\cdot,0) = u^0$ and

$$(\partial_t u, \phi) - \Big(\mathbf{F}(u), \nabla\phi\Big) + \Big(H(u)\nabla(p_c(u) + \tau\partial_t u), \nabla\phi\Big) = 0, \tag{5.1.1}$$

for any $\phi \in L^2(0,T;W_0^{1,2}(\Omega))$.

To avoid the confusion between u given by the conformal formulation, and the solution pair obtained in either of the following two system formulations we denote the saturation in the latter by v. Then a weak form of (5.0.5) and (5.0.4) is

Problem 2: Find $v \in W^{1,2}(0,T;L^2(\Omega))$ and $p \in L^2(0,T;W_0^{1,2}(\Omega))$ such that $v(\cdot,0) = u^0$ and

$$(\partial_t v, \phi) - \Big(\mathbf{F}(v), \nabla\phi\Big) + \Big(H(v)\nabla p, \nabla\phi\Big) = 0, \tag{5.1.2}$$

$$(p, \psi) = \Big(p_c(v), \psi\Big) + \tau(\partial_t v, \psi), \tag{5.1.3}$$

for any $\phi \in L^2(0,T;W_0^{1,2}(\Omega))$, $\psi \in L^2(0,T;L^2(\Omega))$.

Finally, a weak counterpart of (5.0.6) and (5.0.4) is

Problem 3: Find $v \in W^{1,2}(0,T;L^2(\Omega))$ and $p \in L^2(0,T;W_0^{1,2}(\Omega))$ such that $v(\cdot,0) = u^0$ and

$$(\partial_t v, \phi) - \Big(\mathbf{F}(v), \nabla\phi\Big) + \Big(H(v)\nabla p, \nabla\phi\Big) = 0, \tag{5.1.4}$$

$$(p, \psi) + \tau\Big(H(v)\nabla p, \nabla\psi\Big) = \Big(p_c(v), \psi\Big) + \tau\Big(\mathbf{F}(v), \nabla\psi\Big), \tag{5.1.5}$$

for any $\phi \in L^2(0,T;W_0^{1,2}(\Omega))$, $\psi \in L^2(0,T;W_0^{1,2}(\Omega))$.

Note that in both Problems 2 and 3, v and $\partial_t v$ are only assumed in $L^2(0,T;W^{-1,2}(\Omega))$. However, the equivalence results below provide $v = u$ (and thus $\partial_t v = \partial_t u$), implying that $v, \partial_t v \in L^2(0,T;W_0^{1,2}(\Omega))$.

Proposition 5.1.1 *If (v, p) solves Problem 2, then $v \in W^{1,2}(0,T; W_0^{1,2}(\Omega))$.*

Proof . Taking any $V \subset\subset \Omega$ (meaning that its support is strictly included in Ω), with $i \in \{1, \ldots, d\}$ and e_i being the unit vector in the i^{th} direction, if $h \in \mathbb{R}$ is s. t. $|h| < \frac{1}{2}\,\mathrm{dist}(V, \partial\Omega)$ one can consider the he_i spatial translation for (5.1.3), where the inner products are restricted to V

$$\Big(p(\cdot + he_i, \cdot), \psi\Big) = \Big(p_c(v(\cdot + he_i, \cdot)), \psi\Big) + \tau\Big(\partial_t v(\cdot + he_i, \cdot), \psi\Big), \tag{5.1.6}$$

for any $\psi \in L^2(0,T; L^2(V))$. Given a function $g : \Omega \to \mathbb{R}$ we denote by

$$\delta_{h,i} g(x) = \frac{g(x + he_i) - g(x)}{h} \quad \text{for} \quad x \in V,$$

the extension to time dependent functions being straightforward. Subtracting (5.1.3) from (5.1.6) and dividing by h, taking $\psi = \delta_{h,i} v$ and integrating the result in time over the interval $(0, t)$ with t arbitrary in $(0, T]$ one gets

$$\begin{aligned}\frac{\tau}{2}\|\delta_{h,i} v(\cdot, t)\|^2_{L^2(V)} - \frac{\tau}{2}\left\|\delta_{h,i} u^0\right\|^2_{L^2(V)} \\ + \int_0^t \Big(\delta_{h,i} p_c(v(\cdot, s)), \delta_{h,i} v(\cdot, s)\Big)\, ds = \int_0^t \Big(\delta_{h,i} p(\cdot, s), \delta_{h,i} v(\cdot, s)\Big)\, ds.\end{aligned}$$

Since p_c is monotonically increasing and Lipschitz, the third term of the left side is positive. Further, since $u^0 \in W_0^{1,2}(\Omega)$ and $p \in L^2(0,T; W_0^{1,2}(\Omega))$, using the Cauchy-Schwarz inequality we have

$$\tau\|\delta_{h,i} v(\cdot, t)\|^2_{L^2(V)} \leq C + \int_0^t \|\delta_{h,i} v(\cdot, s)\|^2_{L^2(V)} ds. \tag{5.1.7}$$

By Gronwall's inequality we obtain

$$\|\delta_{h,i} v(\cdot, t)\|^2_{L^2(V)} \leq C, \tag{5.1.8}$$

for any $t \in (0, T]$. Moreover, the constant C above does not depend on h, the direction i, or the subset V. According to Lemma 7.24, pp. 169 in [11], $v \in W^{1,2}(\Omega)$ for any $t \in (0, T]$, while the uniform estimate in (5.1.8) shows that $v \in L^2(0,T; W^{1,2}(\Omega))$.

Furthermore, on $\partial\Omega$ we have $p = 0$ (in the trace sense). Furthermore, u^0 vanishes on $\partial\Omega$ as well. Note that (5.1.3) can be interpreted as an ordinary differential equation with unknown v and this equation has a unique solution for any given initial data since p_c is assumed C^1. Then uniqueness results for ordinary differential equations guarantee that the trace of v vanishes on $\partial\Omega$ for any $t \in (0, T]$.

In this way we have showed that $v \in L^2(0, T; W_0^{1,2}(\Omega))$ and $p_c(v) \in L^2(0, T; W_0^{1,2}(\Omega))$. Moreover, from (5.1.3) we have $\partial_t v \in L^2(0, T; W_0^{1,2}(\Omega))$, implying $v \in W^{1,2}(0, T; W_0^{1,2}(\Omega))$. □

Proposition 5.1.2 *Let $u \in W^{1,2}(0, T; W_0^{1,2}(\Omega))$ solve Problem 1. Then $(v, p) \in W^{1,2}(0, T; L^2(\Omega)) \times L^2(0, T; W^{1,2}(\Omega))$ defined as*

$$(v, p) = \left(u, p_c(u) + \tau\partial_t u\right) \tag{5.1.9}$$

solves Problem 2. Conversely, if $(v, p) \in W^{1,2}(0, T; L^2(\Omega)) \times L^2(0, T; W^{1,2}(\Omega))$ solves Problem 2, then $u = v$ solves Problem 1.

Proof . "⇒" Let $u \in W^{1,2}(0, T; W_0^{1,2}(\Omega))$ be a solution of Problem 1 and (v, p) defined in (5.1.9). Then $(v, p) \in W^{1,2}(0, T; L^2(\Omega)) \times L^2(0, T; W_0^{1,2}(\Omega))$, and (5.1.2), (5.1.3) are satisfied straightforwardly for test functions in $L^2(0, T; W_0^{1,2}(\Omega))$. Moreover, density arguments show that (5.1.3) also holds for test functions in $L^2(0, T; L^2(\Omega))$.

"⇐" Let $(v, p) \in W^{1,2}(0, T; L^2(\Omega)) \times L^2(0, T; W^{1,2}(\Omega))$ solve Problem 2. By Proposition 5.1.1, $v \in W^{1,2}(0, T; W_0^{1,2}(\Omega))$. Hence $u = v$ has the desired regularity. Therefore,

$$\nabla p = \nabla p_c(v) + \tau\nabla\partial_t v,$$

in L^2 sense, and for almost every x and t. Inserting the above into (5.1.2) gives (5.1.1). □

In a similar fashion we have

Proposition 5.1.3 *Let $u \in W^{1,2}(0, T; W_0^{1,2}(\Omega))$ solve Problem 1, then $(v, p) \in W^{1,2}(0, T; L^2(\Omega)) \times L^2(0, T; W_0^{1,2}(\Omega))$ defined as*

$$(v, p) = \left(u, p_c(u) + \tau\partial_t u\right) \tag{5.1.10}$$

solves Problem 3. Conversely, if $(v, p) \in W^{1,2}(0, T; L^2(\Omega)) \times L^2(0, T; W_0^{1,2}(\Omega))$ solves Problem 3, then $u = v$ solves Problem 1.

Proof . "⇒" As for Proposition 5.1.2, if $u \in W^{1,2}(0,T;W_0^{1,2}(\Omega))$ solves Problem 1, the (v,p) defined in (5.1.10) satisfies $(v,p) \in W^{1,2}(0,T;L^2(\Omega)) \times L^2(0,T;W_0^{1,2}(\Omega))$, implying (5.1.4). Further, $\partial_t v = \big(p - p_c(v)\big)/\tau$ in L^2 sense. Then, replacing u by v and $\partial_t u$ by $\big(p - p_c(v)\big)/\tau$ in (5.1.1) gives (5.1.5).

"⇐" Let $(v,p) \in W^{1,2}(0,T;L^2(\Omega)) \times L^2(0,T;W_0^{1,2}(\Omega))$ solving Problem 3 and $u = v$, then $u \in W^{1,2}(0,T;L^2(\Omega))$ and $u(\cdot,0) = u^0$. Further, since by definition

$$\tau\partial_t u + p_c(u) = p \in L^2(0,T;W_0^{1,2}(\Omega)),$$

Again using Proposition 5.1.1, we obtain $u \in W^{1,2}(0,T;W_0^{1,2}(\Omega))$, and the validity of (5.1.1). □

A direct consequence of Propositions 5.1.1 and 5.1.2 is

Proposition 5.1.4 *A pair* $(v_1,p_1) \in W^{1,2}(0,T;L^2(\Omega)) \times L^2(0,T;W_0^{1,2}(\Omega))$ *solves Problem 2 iff the pair* $(v_2,p_2) \in W^{1,2}(0,T;L^2(\Omega)) \times L^2(0,T;W_0^{1,2}(\Omega))$ *defined as*

$$(v_2,p_2) = (v_1,p_1)$$

solves Problem 3.

According to Propositions 5.1.2, 5.1.3 and 5.1.4, we have the following theorem

Theorem 5.1.1 *Problems 1, 2 and 3 are equivalent.*

We conclude this section with the following observation. The formulations considered here are equivalent only if the solution u of Problem 1 (and the corresponding v's) exists in the space $W^{1,2}(0,T;W_0^{1,2}(\Omega))$, whereas $p \in L^2(0,T;W_0^{1,2}(\Omega))$. This regularity holds in the nondegenerate case, when H is bounded away from 0 and ∞. Whenever H vanishes for some values of u (the degenerate case), this being a characteristic of many models used in the subsurface, the weak solution has less regularity (see e.g. [5, 14]). Specifically, $\partial_t \nabla u$ (and further ∇p) may fail belonging to L^2 in the regions where u takes the degeneracy values. Nevertheless, this lacking regularity is compensated by the vanishing of H so that the term $H(u)\partial_t \nabla u$ still remains L^2. As mentioned before, in this case the equivalence may not hold. An indication in this sense can be found in [7], where the existence of travelling waves for (5.0.3) is proved for one spatial dimension. The degenerate case does not allow for travelling waves in the classical sense, and requires an extended concept (the sharp waves) where u is continuous, but its derivative can have a discontinuity at points where u takes the degeneracy value. Consequently, p becomes discontinuous there. In this case, difficulties appear in the second equation of (5.0.6), which suggests that $p \in W^{1,2}$. In the one dimensional situation this also implies its continuity.

5.1.2 The semidiscrete case

In this section we extend the results obtained previously for the three formulations of the original problem, and show that the equivalence holds for the time discretization as well. In this sense we only consider the Euler implicit method. For simplicity, the time step is taken fixed, but the results can be extended straightforwardly to discretizations based on a variable/adaptive stepping.

Having shown the equivalence of the different time discrete formulations, the convergence results available for the conformal discretization (see e. g. [1, 8–10, 13]) provide rigorous convergence results for the other two formulations as well.

To define the time discretization we let $N \in \mathbb{N}$, take $\Delta t = T/N$, and $t_n = n\Delta t$. For the original formulation in (5.0.3) we consider the following sequence of problems:

Problem 4: Given $u^{n-1} \in W_0^{1,2}(\Omega)$, find $u^n \in W_0^{1,2}(\Omega)$ such that

$$\left(\frac{u^n - u^{n-1}}{\Delta t}, \phi\right) - \left(\mathbf{F}(u^n), \nabla\phi\right) + \left(H(u^n)\nabla\left(p_c(u^n) + \tau\frac{u^n - u^{n-1}}{\Delta t}\right), \nabla\phi\right) = 0, \quad (5.1.11)$$

for any $\phi \in W_0^{1,2}(\Omega)$.

Clearly, u^n approximates the solution u of Problem 1, taken at the time t_n. At $n = 1$ we start with the initial data, u^0. In a similar way, for the system in (5.0.5) we have

Problem 5: Given $v^{n-1} \in W_0^{1,2}(\Omega)$, find $(v^n, p^n) \in W_0^{1,2}(\Omega) \times W_0^{1,2}(\Omega)$ such that

$$\left(\frac{v^n - v^{n-1}}{\Delta t}, \phi\right) - \Big(\mathbf{F}(v^n), \nabla\phi\Big) + \Big(H(v^n)\nabla p^n, \nabla\phi\Big) = 0, \tag{5.1.12}$$

$$(p^n, \psi) = \Big(p_c(v^n), \psi\Big) + \tau\left(\frac{v^n - v^{n-1}}{\Delta t}, \psi\right), \tag{5.1.13}$$

for any $\phi \in W_0^{1,2}(\Omega)$, $\psi \in L^2(\Omega)$.

Finally, the discrete counterpart of (5.0.6) is

Problem 6: Given $v^{n-1} \in W_0^{1,2}(\Omega)$, find $(v^n, p^n) \in W_0^{1,2}(\Omega) \times W_0^{1,2}(\Omega)$ such that

$$\left(\frac{v^n - v^{n-1}}{\Delta t}, \phi\right) - \Big(\mathbf{F}(v^n), \nabla\phi\Big) + \Big(H(v^n)\nabla p^n, \nabla\phi\Big) = 0, \tag{5.1.14}$$

$$(p^n, \psi) + \tau\Big(H(v^n)\nabla p^n, \nabla\psi\Big) = \Big(p_c(v^n), \psi\Big) + \tau\Big(\mathbf{F}(v^n), \nabla\psi\Big). \tag{5.1.15}$$

for any $\phi \in W_0^{1,2}(\Omega)$, $\psi \in W_0^{1,2}(\Omega)$.

As before, when discretizing the system forms the sequence of problems are started with the initial value, $v^0 = u^0$. Note that in Problem 5 it is sufficient to assume that $v^n \in L^2(\Omega)$. Then, (5.1.13) immediately implies that $v^n \in W_0^{1,2}(\Omega)$. Further, in Problem 6 it is sufficient to assume that $v^n \in L^2(\Omega)$ as well.

Similar to the continuous case we have the equivalence of the three formulations in the time discrete case. To prove these results we only focus on one time step t_n and assume that $v^{n-1} = u^{n-1}$, i. e. the saturations at t_{n-1} appearing in Problems 4, 5, and 6 are the same.

Proposition 5.1.5 *Let $u^n \in W_0^{1,2}(\Omega)$ solve Problem 4. Then $(v^n, p^n) \in W_0^{1,2}(\Omega) \times W_0^{1,2}(\Omega)$ defined as*

$$(v^n, p^n) = \left(u^n, p_c(u^n) + \tau\frac{u^n - u^{n-1}}{\Delta t}\right) \tag{5.1.16}$$

solves Problem 5. Conversely, if $(v^n, p^n) \in W_0^{1,2}(\Omega) \times W_0^{1,2}(\Omega)$ solves Problem 5, then $u^n = v^n$ solves Problem 4.

Proof . "⇒" Let $u^n \in W_0^{1,2}(\Omega)$ solve Problem 4 and (v^n, p^n) be defined in (5.1.16). Since $v^{n-1} \in W_0^{1,2}(\Omega)$, $(v^n, p^n) \in W_0^{1,2}(\Omega) \times W_0^{1,2}(\Omega)$. From (5.1.11) and (5.1.16), we have (5.1.12). Further, (5.1.13) holds for test functions in $W_0^{1,2}(\Omega)$, but this can be extended to L^2 functions by density arguments.

"⇐" Let $(v^n, p^n) \in W_0^{1,2}(\Omega) \times W_0^{1,2}(\Omega)$ solve Problem 5 and $u^n = v^n$. Note that all terms in (5.1.13) are in $W^{1,2}$, therefore the equality holds (in L^2 sense and thus in a. e. sense) for the gradients as well. Since $H(u^n)$ is essentially bounded, we have

$$H(u^n)\nabla p^n = H(u^n)\left(\nabla p_c(v^n) + \tau\nabla\left(\frac{v^n - v^{n-1}}{\Delta t}\right)\right),$$

in L^2 sense. Now substituting the above into (5.1.12) gives (5.1.11). □

In the same spirit we have

Proposition 5.1.6 *Let $u^n \in W_0^{1,2}(\Omega)$ solve Problem 4. Then $(v^n, p^n) \in W_0^{1,2}(\Omega) \times W_0^{1,2}(\Omega)$ defined as*

$$(v^n, q^n) = \left(u^n, p_c(u^n) + \tau\frac{u^n - u^{n-1}}{\Delta t}\right) \tag{5.1.17}$$

solves Problem 6. Conversely, if $(v^n, p^n) \in W_0^{1,2}(\Omega) \times W_0^{1,2}(\Omega)$ solves Problem 6, then $u^n = v^n$ solves Problem 4.

Proof . "⇒" Let $u^n \in W_0^{1,2}(\Omega)$ solve (5.1.11) and (v^n, p^n) be defined by (5.1.17). As for Proposition 5.1.4, $(v^n, p^n) \in W_0^{1,2}(\Omega) \times W_0^{1,2}(\Omega)$ and (5.1.14) holds true. For (5.1.15) we use the definition of p^n into (5.1.11) to eliminate the terms containing the difference $u^n - u^{n-1}$.

"⇐" Let $(v^n, p^n) \in W_0^{1,2}(\Omega) \times W_0^{1,2}(\Omega)$ solve Problem 6. Then $u^n = v^n \in W_0^{1,2}(\Omega)$. Subtracting (5.1.15) from (5.1.14) multiplied by τ gives

$$(p^n, \phi) = \left(p_c(u^n) + \tau\frac{u^n - u^{n-1}}{\Delta t}, \phi\right)$$

for all $\phi \in W_0^{1,2}(\Omega)$. Then clearly, $p^n = p_c(u^n) + \tau\frac{u^n - u^{n-1}}{\Delta t}$, and the proof is continued as in Proposition 5.1.4. □

By Proposition 5.1.4 and Proposition 5.1.5, we have the following

Proposition 5.1.7 $(v_1^n, p_1^n) \in W_0^{1,2}(\Omega)\times W_0^{1,2}(\Omega)$ *solves Problem 5 iff* $(v_2^n, p_2^n) \in W_0^{1,2}(\Omega)\times W_0^{1,2}(\Omega)$ *defined as*

$$(v_1^n, p_1^n) = (v_2^n, p_2^n) \tag{5.1.18}$$

solves Problem 6.

Therefore, we have the following theorem

Theorem 5.1.2 *Problems 4, 5 and 6 are equivalent.*

References

[1] D.N. Arnold, J. Douglas Jr. and V. Thomée, *Superconvergence of a Finite Element Approximation to the Solution of a Sobolev Equation in a Single Space Variable*, Math. Comp. **36** (1981), 53–63.

[2] J. Bear and Y. Bachmat, *ntroduction to modelling of transport phenomena in porous media*, Kluwer Academic, Dordrecht, 1991.

[3] A. Lamacz, A. Rätz and B. Schweizer, *A well-posed hysteresis model for flows in porous media and applications to fingering effects*, Report **12** (2010), Fakultät für Mathematik, Universität Dortmund, 2010.

[4] B. Schweizer, *The Richards equation with hysteresis and degenerate capillary pressure*, submitted.

[5] C. Cancès, C. Choquet, Y. Fan and I.S. Pop, *Existence of weak solutions to a degenerate pseudo-parabolic equation modeling two-phase flow in porous media*, CASA Report **10-75**, Eindhoven University of Technology, 2010.

[6] C.M. Cuesta, I.S. Pop, *Numerical schemes for a pseudo-parabolic Burgers equation: Discontinuous data and long-time behaviour*, J. Comp. Appl. Math. **224** (2009), 269–283.

[7] C.J. van Duijn, Y. Fan, L.A. Peletier and I.S. Pop, *Travelling wave solutions for degenerate pseudo-parabolic equation modelling two-phase flow in porous media*, CASA Report **10-01**, Eindhoven University of Technology, 2010.

[8] R.E. Ewing, *Time-stepping Galerkin methods for nonlinear Sobolev partial differential equations*, SIAM J. Numer. Anal. **15** (1978), 1125–1150.

[9] Y. Fan and I.S. Pop, *A class of degenerate pseudo-parabolic equation: existence, uniqueness of weak solutions, and error estimates for the Euler-implicit discretization*, Mathematical Methods in the Applied Sciences **34** (2011), 2329-2339.

[10] W.H. Ford and T.W. Ting, *Unifrom error estimates for difference approximations to nonlinear pseudo-parabolic partial differential equations*, SIAM J. Numer. Anal. **11** (1974), 155–169.

[11] D. Gilbarg, Neil S. Trudinger, *Elliptic partial differential equations of second order*, Springer-Verlag, Berlin, 2001.

[12] S.M. Hassanizadeh and W.G. Gray, *Thermodynamic basis of capillary pressure in porous media*, Water Resour. Res. **29** (1993), 3389–3405.

[13] R. Helmig, A. Weiss and B. Wohlmuth, *Dynamic capillary effects in heterogeneous porous media*, Comput. Geosci. **11** (2007), 261–274.

[14] A. Mikelić, *A global existence result for the equations describing unsatured flow in porous media with dynamic capillary pressure*, J. Differential Equations **248** (2010), 1561–1577.

[15] M. Ptashnyk, *Nonlinear pseudoparabolic equations as sigular limit of reaction-diffusion equations*, Appl. Anal. **85** (2006), 1285–1299.

[16] M. Peszyńska, S. Yi, *Numerical methods for unsaturated flow with dynamic capillary pressure in heterogeneous porous media*. Int. J. Numer. Anal. Model. **5** (2008), 126-149.

[17] F.A. Radu, I.S. Pop and P. Knabner, *Order of convergence estimates for an Euler implicit, mixed finite element discretization of Richards' equation*, SIAM J. Numer. Anal. 42 (2004), pp. 1452-1478.

[18] R.E. Showalter, *A nonlinear parabolic-Sobolev equation*, Journal of Mathematical Analysis And Applications 50 (1975), pp. 183-190.

Chapter 6

A fully discrete scheme

In Chapter 2, we have solved the pseudo-parabolic equation (2.4.1) by using a semi-implicit Euler finite volume scheme, where a direct approximation of the term involving the mixed derivative is included. Here we introduce another numerical scheme, which is inspired by the equivalent formulations given in Chapter 5. We introduce the pressure p as an additional unknown and solve the equation (2.4.1) as a system, avoiding approximating the third order mixed derivative directly. In the following, we will describe the scheme for one-dimensional spatial case in details.

6.1 The numerical schemes

Recalling the scaling (2.4.3) in Chapter 2, we consider the equation

$$\partial_t u + \partial_x f(u) = \partial_x\Big(H(u)\partial_x(u + \tau\partial_t u)\Big). \tag{6.1.1}$$

In order to make the context complete, we repeat the scheme introduced in Chapter 2 and name it Scheme 1 here. With $\Delta x, \Delta t$ denoting the spatial and time step, the Scheme 1 is described in the following.

Scheme 1.

$$\frac{u_i^n - u_i^{n-1}}{\Delta t} + \frac{F^{n-1}(u_i, u_{i+1}) - F^{n-1}(u_{i-1}, u_i)}{\Delta x} = 0, \tag{6.1.2}$$

where the numerical flux $F^{n-1}(u_i, u_{i+1})$ is defined by

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

For the coefficient $H_{i+\frac{1}{2}}^{n-1}$, we use the arithmetic average value:

$$H_{i+\frac{1}{2}}^{n-1} = \frac{1}{2}\Big(H(u_i^{n-1}) + H(u_{i+1}^{n-1})\Big).$$

Let $p = u + \tau\partial_t u$, we transform (6.1.1) into a system

$$\begin{cases} p - u + \tau\partial_x f(u) = \tau\partial_x\Big(H(u)\partial_x p\Big), \\ p = u + \tau\partial_t u. \end{cases} \tag{6.1.3}$$

We introduce the following

Scheme 2.

$$\begin{cases} -\tau\dfrac{H_{i+\frac{1}{2}}^{n-1}\frac{p_{i+1}^n - p_i^n}{\Delta x} - H_{i-\frac{1}{2}}^{n-1}\frac{p_i^n - p_{i-1}^n}{\Delta x}}{\Delta x} + p_i^n - \theta u_i^n = (1-\theta)u_i^{n-1} - \tau\dfrac{f(u_i^{n-1}) - f(u_{i-1}^{n-1})}{\Delta x}, \\ p_i^n = u_i^n + \tau\dfrac{u_i^n - u_i^{n-1}}{\Delta t}. \end{cases} \tag{6.1.4}$$

In this way, we only need to approximate the derivatives up to order two. Note here we use the upwind discretization of the first order term $\partial_x f(u)$. Further, p_i^n, u_i^n are the estimated values of $p(i\Delta x, n\Delta t), u(i\Delta x, n\Delta t)$ with $i = 1, \cdots, N$. θ is a positive number satisfying $0 < \theta < 1$. For $H_{i+\frac{1}{2}}^{n-1}$ we use the same as Scheme 1. Rewriting (6.1.4) gives

$$\begin{cases} -\Delta t p_i^n + (\Delta t + \tau)u_i^n = \tau u_i^{n-1}, \\ -\tau H_{i+\frac{1}{2}}^{n-1} p_{i+1}^n + \Big(\tau(H_{i+\frac{1}{2}}^{n-1} + H_{i-\frac{1}{2}}^{n-1}) + (\Delta x)^2\Big)p_i^n - \tau H_{i-\frac{1}{2}}^{n-1} p_{i-1}^n \\ -\theta(\Delta x)^2 u_i^n = (1-\theta)(\Delta x)^2 u_i^{n-1} - \tau\Delta x\Big(f(u_i^{n-1}) - f(u_{i-1}^{n-1})\Big). \end{cases} \tag{6.1.5}$$

In the vector form, this becomes

$$MU^n = \mathcal{F}(U^{n-1}), \tag{6.1.6}$$

where $U^n = (u_1^n, \cdots, u_N^n, p_1^n, \cdots, p_N^n)^T$, and M is a $2N \times 2N$ matrix defined as

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

with $A = (\Delta t + \tau)I, B = -\Delta t I, C = -\theta(\Delta x)^2 I$ (I is the $N \times N$ identity matrix), whereas the $N \times N$ matrix D is defined by

$$\begin{gathered} D_{1,1} = -\tau H(u_1^{n-1}), \quad D_{1,2} = \tau H(u_1^{n-1}) + (\Delta x)^2, \\ D_{i,i-1} = -\tau H_{i-1/2}^{n-1}, \quad D_{i,i} = \tau(H_{i-1/2}^{n-1} + H_{i+1/2}^{n-1}) + (\Delta x)^2, \quad D_{i,i+1} = -\tau H_{i+1/2}^{n-1}, \\ D_{N,N-1} = -\tau H(u_N^{n-1}), \quad D_{N,N} = \tau H(u_N^{n-1}) + (\Delta x)^2, \end{gathered}$$

for $i = 2, \cdots, N-1$ and all the other elements of D are zero. Whenever $0 < \theta < 1$, the matrix M is strictly diagonal dominant, hence it is invertible. As in Chapter 5, one can prove the equivalence between Schemes 1 and 2 under the assumption H is essentially bounded. To test Scheme 2, in the following we will give two numerical results.

6.2 The test cases

The tests are carried out for two situations. In the first case, H is identically 1, so the higher order terms are linear; whereas the second case is nonlinear and possibly degenerate.

6.2.1 The linear case

We take here $H(u) = 1$ for all u. For the function f, we take $f(u) = \frac{u^2}{u^2+2(1-u)^2}$ when $u \in [0, 1]$, $f(u) = 0$ if $u < 0$ and $f(u) = 1$ if $u > 1$. To check the numerical results, we use the computations in [2] as benchmarks. As in [2], we show the graphs by scaling the time t here. We compare the results; for $\tau = 5$, $u_r = 0$, with two different inflow values: $u_B = 0.9$ and $u_B = 0.55$ (see Figure 6.1). The solid line is computed by solving the equation directly, whereas the dashed line is computed by solving the equation as a system. The computation time is $t = 250$. As we can see in Fig. 6.1, the results computed by Scheme 2 agree well with the benchmarks.

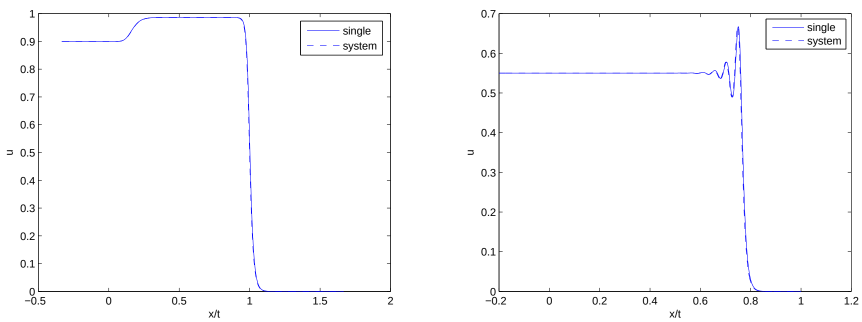


FIGURE 6.1: COMPARISON OF THE NUMERICAL RESULTS PROVIDED BY SCHEMES 1 AND 2, WITH $\tau = 5$, $u_r = 0$ AND FOR TWO DIFFERENT INFLOW VALUES: $u_B = 0.9$ (LEFT) AND $u_B = 0.55$ (RIGHT)

6.2.2 The nonlinear case

Here we take $H(u) = \frac{u^{1.5}(1-u)^{1.5}}{u^{1.5}+2.5(1-u)^{1.5}}$ and $f(u) = \frac{u^{1.5}}{u^{1.5}+2.5(1-u)^{1.5}}$ for $0 \leq u \leq 1$. Further $H(u) = f(u) = 0$ if $u < 1$, and $H(u) = 0$, $f(u) = 1$ if $u > 1$. Here we use the numerical results in [1] (also see Chapter 2) as benchmarks. The two numerical results shown in Fig 6.2 are computed by taking $\tau = 0.1$, $u_r = 0.1$, with inflow value $u_B = 1$ (left), respectively $\tau = 0.1$, $u_r = 0$ with inflow value $u_B = 0.9$ (right). The solid and dashed lines are computed by using Schemes 1 and 2 respectively. Here we compute the solutions on the spatial interval $(-5, 25)$. Again, a good agreement can be observed.

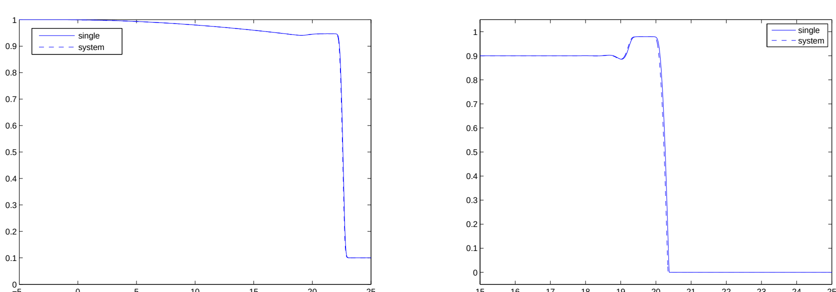


FIGURE 6.2: COMPARISON OF THE NUMERICAL RESULTS PROVIDED BY SCHEMES 1 AND 2: $\tau = 0.1$, $u_r = 0.1$ WITH INFLOW VALUE $u_B = 1$ (LEFT), $\tau = 0.1$, $u_r = 0$ WITH INFLOW VALUE $u_B = 0.9$ (RIGHT)

Based on the tests we compare the two schemes, which are equivalent. However, the advantages of Scheme 2 are:

1. There is no need to approximate the mixed third order derivative, we only deal with first and second derivatives being involved.
2. Scheme 2 is easier to implement than Scheme 1.
The disadvantages of Scheme 2 are:
1. It requires solving a system instead of one equation, which requires more computation resources.
2. It is not clear whether the equivalence of the schemes still holds in the degenerate case.

References

[1] C.J. van Duijn, Y. Fan, L.A. Peletier and I.S. Pop, *Travelling wave solutions for degenerate pseudo-parabolic equation modelling two-phase flow in porous media*, CASA Report **10-01**, TU Eindhoven, 2010.

[2] C.J. van Duijn, L.A. Peletier and I.S. Pop, *A new class of entropy solutions of the Buckley-Leverett equation*, SIAM J. Math. Anal. **39** (2007), 507–536.

Appendix A

The proof of Remark 2.1.3

In Remark 2.1.3 of Chapter 2, we mentioned that f is 'convex-concave' when $0 < p < 1, 0 < q < 1$. Here we prove the result for $p > 0, q > 0$:

Lemma: Assume $p > 0, q > 0, M > 0$ and $f(u) = \frac{u^{1+p}}{u^{1+p}+M(1-u)^{1+q}}$. Consider $u \in [0, 1]$, then there exists a unique $u^* \in (0, 1)$ such that $f''(u^*) = 0$, $f''(u) > 0$ for $0 < u < u^*$ and $f''(u) < 0$ for $u^* < u < 1$.

Proof . Set $f_1 = u^{1+p}$ and $f_2 = M(1 - u)^{1+q}$, we have

$$f_1 > 0, f_1' = (p + 1)u^p > 0, f_1'' = (p + 1)pu^{p-1} > 0,$$

and

$$f_2 > 0, f_2' = -M(q + 1)(1 - u)^q < 0, f_2'' = Mq(q + 1)(1 - u)^{q-1} > 0,$$

whenever $0 < u < 1$. Therefore for f, we have

$$f'(u) = \frac{f_1'(u)f_2(u) - f_1(u)f_2'(u)}{\left(f_1(u) + f_2(u)\right)^2} > 0. \tag{A.0.1}$$

Further,

$$f''(u) = \frac{f_1''(u)f_2(u) - f_1(u)f_2''(u) - 2\Big(f_1'(u)f_2(u) - f_1(u)f_2'(u)\Big)\Big(f_1'(u) + f_2'(u)\Big)}{\Big(f_1(u) + f_2(u)\Big)^3}. \tag{A.0.2}$$

The above lemma means there is only one zero point for f'' when $0 < u < 1$. We prove this by an equivalent statement: for any $\lambda > 0$, there exists at most two solutions for equation $f'(u) = \lambda$ in $(0, 1)$.

To do so, define $a(u) = f_1'(u)f_2(u) - f_1(u)f_2'(u)$, $b(u) = \Big(f_1(u) + f_2(u)\Big)^2$. Then

$$f'(u) = C \text{ iff } a(u) = \lambda b(u). \tag{A.0.3}$$

Consider $g(u) = a(u) - \lambda b(u)$, then

$$g(u) = M(p+1)u^p(1-u)^{q+1} + M(q+1)u^{p+1}(1-u)^q - \lambda\Big(u^{2p+2} + 2Mu^{p+1}(1-u)^{q+1} + M^2(1-u)^{2q+2}\Big).$$

Dividing by $u^p(1-u)^q$, we have

$$g(u) = 0 \text{ iff } M(p+1)(1-u) + (q+1)u - \lambda\Big(u^{p+2}(1-u)^{-q} + 2Mu(1-u) + M^2u^{-p}(1-u)^{q+2}\Big) = 0.$$

Denote the expression on the left by h, we have

$$\begin{aligned}
h''(u)/(-\lambda) = {} & (p+2)(p+1)u^p(1-u)^{-q} + (p+2)qu^{p+1}(1-u)^{-q-1} \\
& + (p+2)qu^{p+1}(1-u)^{-q-1} + q(q+1)u^{p+2}(1-u)^{-q-2} \\
& + M^2(p+1)pu^{-p-2}(1-u)^{-q-2} + M^2p(q+2)u^{-p-1}(1-u)^{q+1} \\
& + M^2p(q+2)u^{-p-1}(1-u)^{-q-1} + M^2(q+2)(q+1)u^{-p}(1-u)^q - 4M \\
& > (p+2)(p+1)u^p(1-u)^{-q} + M^2(q+2)(q+1)u^{-p}(1-u)^q - 4M \\
& \geq 2M\sqrt{(p+2)(p+1)(q+2)(q+1)} - 4M > 0
\end{aligned}$$

Therefore $h(u) = 0$ has at most two solutions. Therefore the same holds for g, implying that $f'(u) = \lambda$ for at most 2 values of u . □

Appendix B

The Div-Curl Lemma

.

In Chapter 4, we have used an important compensated-compactness argument, the Div-Curl Lemma. It was introduced by F. Murat[1] 1978, and has various forms. Here we give the classical form and the proof, for more details we refer to L.C. Evans[2].

Notation. If $w \in L^2(U;\mathbb{R}^n)$, $w = (w^1, ..., w^n)$, we define curl $w \in W^{-1,2}(U; M^{n\times n})$ by

$$(\mathrm{curl} w)_{ij} := w^i_{x_j} - w^j_{x_i}, \text{ for } 1 \leq i, j \leq n. \tag{B.0.1}$$

Here $M^{n\times n}$ denotes the set of $n \times n$ matrices.

Div-Crul Lemma: Let $\{v_k\}_{k\in\mathbb{N}}$, $\{w_k\}_{k\in\mathbb{N}}$ be two bounded sequences in $L^2(U;\mathbb{R}^n)$ such that

1. $\{\mathrm{div} v_k\}_{k\in\mathbb{N}}$ lies in a compact subset of $W^{-1,2}(U)$,
2. $\{\mathrm{curl} w_k\}_{k\in\mathbb{N}}$ lies in a compact subset of $W^{-1,2}(U; M^{n\times n})$.

Suppose further $v_k \rightharpoonup v$, $w_k \rightharpoonup w$ weakly in $L^2(U;\mathbb{R}^n)$. Then

$$v_k \cdot w_k \to v \cdot w, \tag{B.0.2}$$

in the sense of distribution.

Proof . Consider for each $k = 1$ ($k \in \mathbb{N}$), the vector field $u_k \in W^{2,2}(U;\mathbb{R}^n)$ solving

$$\begin{cases} -\Delta u_k = w_k, & \text{in } U, \\ u_k = 0, & \text{on } \partial U. \end{cases} \tag{B.0.3}$$

in the weak sense. Since $\{w_k\}_{k\in\mathbb{N}}$ is bounded in $L^2(U;\mathbb{R}^n)$, $\{u_k\}_{k\in\mathbb{N}}$ is bounded in $W^{2,2}(U;\mathbb{R}^n)$.

Now set $z_k := -\mathrm{div} u_k$, $y_k := w_k - Dz_k \quad (k \in \mathbb{N})$. Then $\{z_k\}_{k\in\mathbb{N}}$ is bounded in $W^{1,2}(U)$. Additionally, if $1 \le i \le n$,

$$\begin{aligned} y_k^i &= w_k^i - z_{k,x_i} \\ &= \sum_{j=1}^n -u^i_{k,x_jx_j} + \sum_{j=1}^n u^j_{k,x_ix_j} \\ &= \sum_{j=1}^n (u^j_{k,x_i} - u^i_{k,x_j})_{x_j}. \end{aligned} \tag{B.0.4}$$

In view of hypothesis 2, we infer from (B.0.3) that $\{\mathrm{curl} u_k\}_{k\in\mathbb{N}}$ lies in a compact subset of $W^{1,2}_{\mathrm{loc}}(U; M \times M)$. Thus from (B.0.4) it follows that $\{y_k\}_{k\in\mathbb{N}}$ is contained in a compact subset of $L^2_{\mathrm{loc}}(U;\mathbb{R}^n)$.

Hence we may suppose, upon passing to subsequences as necessary, that

$$z_k \rightharpoonup z \text{ weakly in } W^{1,2}(U), \quad y_k \to y \text{ strongly in } L^2_{\mathrm{loc}}(U;\mathbb{R}^n). \tag{B.0.5}$$

where $z = -\mathrm{div} u$, $y = w - Dz$, for $u \in W^{2,2}(U;\mathbb{R}^n)$ solving

$$\begin{cases} -\Delta u = w, \text{ in } U, \\ u = 0 \text{ on } \partial U. \end{cases} \tag{B.0.6}$$

Now, if $\phi \in C_0^\infty(U)$, we have

$$\int_U v_k \cdot w_k \phi dx = \int_U v_k \cdot (y_k + Dz_k)\phi dx. \tag{B.0.7}$$

According to (B.0.5),

$$\int_U v_k \cdot y_k \phi dx = \int_U v \cdot y \phi dx. \tag{B.0.8}$$

In addition, hypothesis 1 and (B.0.5) allow us to compute

$$\begin{aligned} &\int_U v_k \cdot D z_k \phi dx = -\int_U v_k \cdot D\phi z_k dx - < \mathrm{div} v_k, z_k \phi > \\ &\to -\int_U v \cdot D\phi z dx - < \mathrm{div} v, z\phi >= \int_U v \cdot D z \phi dx. \end{aligned} \tag{B.0.9}$$

Here $<,>$ is the pairing of $W^{-1,2}(U)$ and $W_0^{1,2}(U)$. Thus

$$\int_U v_k \cdot w_k \phi dx \to \int_U v \cdot (y + Dz)\phi dx = \int_U v \cdot w \phi dx. \tag{B.0.10}$$

References

[1]. F. Murat, *Compacité par compensation*, Ann. Sc. Norm. Sup. Pisa **5**(1978), 489-507.

[2]. L.C. Evans, *Weak convergence methods for nolinear partial differential equations*, CBMS No.74, AMS.

Appendix C

The Vitali Convergence Theorem

The Vitali Convergence Theorem is a generalization of the Dominated Convergence Theorem of Henri Lebesgue. It is named after Italian mathematician Giuseppe Vitali. It is useful whenever a dominating function cannot be found, so the the Dominated Convergence Theorem cannot be applied directly.

Vitali Convergence Theorem: Let (X, Σ, μ) be a measure space, $p \geq 1$ and let $f_n : X \to \mathbb{R}$ belong to $L^p(X, \Sigma, \mu)$ for each $n \in \mathbb{N}$. Then $f_n \to f$ ($f \in L^p(X, \Sigma, \mu)$) if and only if

- f_n converges in measure to f.
- $|f_n|^p$ are uniformly integrable in the sense that,

  for every $\varepsilon > 0$, there exists some $t \geq 0$ such that, for all $n \in \mathbb{N}$,

$$\int_{|f_n|>t} |f_n(x)|^p d\mu(x) < \varepsilon, \tag{C.0.1}$$

and for every $\varepsilon > 0$, there exists some measurable set $E \subseteq X$ with finite μ-measure such that, for all $n \in \mathbb{N}$,

$$\int_{X \setminus E} |f_n(x)|^p d\mu(x) < \varepsilon. \tag{C.0.2}$$

Proof . "$\Rightarrow$" Fix $t > 0$, define $f_{mn} = |f_m - f_n|$ and $E_{mn} := \{f_{mn} \geq t\}$. Then

$$\mu(E_{mn})^{\frac{1}{p}} = \frac{1}{t}\|t\mathbf{1}_{E_{mn}}\| \leq \frac{1}{t}\|f_{mn}\| \to 0, \text{ as } m, n \to +\infty. \tag{C.0.3}$$

Here $\mathbf{1}$ is the characteristic function. Choose N such that $\|f_n - f_N\| < \varepsilon$ for all $n \geq N$, then the family $\{|f_1|^p, ..., |f_{N-1}|^p, |f_N|^p\}$ is uniformly integrable because it consists of only finitely many integrable functions. So for every $\varepsilon > 0$, there is $\delta > 0$ such that $\mu(E) < \delta$ implies $\|f_n\mathbf{1}(E)\| < \varepsilon$ for $n \leq N$. On the other hand, for $n > N$,

$$\|f_n\mathbf{1}_E\| \leq \|(f_n - f_N)\mathbf{1}_E\| + \|f_N\mathbf{1}_E\| < 2\varepsilon, \tag{C.0.4}$$

for the same set E, and thus the entire infinite sequence $\{|f_n|^p\}$ is uniformly integrable too.

Select N such that $\|f_n - f_N\| < \varepsilon$ for all $n \geq N$. Let ϕ be a simple function such that $\|f_N - f\| < \varepsilon$. Then $\|f_n - \phi\| < 2\varepsilon$ for all $n \geq N$. Let A_N be the support of ϕ, which must have finite measure. It follows that

$$\begin{aligned} \|f_n\mathbf{1}_{X\setminus A_N}\| = \|f_n - f_n\mathbf{1}_{A_N}\| &\leq \|f_n - \phi\| + \|\phi - f_n\mathbf{1}_{A_N}\| \\ &= \|f_n - \phi\| + \|(\phi - f_n)\mathbf{1}_{A_N}\| \\ &< 2\varepsilon + 2\varepsilon. \end{aligned} \tag{C.0.5}$$

For each $n < N$, we can similarly construct sets A_n of finite measure, such that $\|f_n\mathbf{1}_{X\setminus A_n}\| < \varepsilon$. If we set $A = \cup_{j=1}^N A_i$, a finite union, then A has finite measure, and clearly $\||f_n\mathbf{1}_{X\setminus A}| < 4\varepsilon$ for any n.

"$\Leftarrow$". For any set B,

$$\|f_{mn}\| = \|f_{mn}\mathbf{1}_{B\setminus E_{mn}}\| + \|f_{mn}\mathbf{1}_{E_{mn}}\| + \|f_{mn}\mathbf{1}_{X\setminus B}\|. \tag{C.0.6}$$

Choose B of finite measure such that $\|f_n\mathbf{1}_{X\setminus B}\| < \varepsilon$ for every n. Then $\|f_{mn}\mathbf{1}_{X\setminus B}\| < 2\varepsilon$.

Let $t = \frac{\varepsilon}{\mu(B)^{1/p}} > 0$, choose $\delta > 0$ so that $\|f_n\mathbf{1}_E\| < \varepsilon$ whenever $\mu(E) < \delta$. Then take N such that if $m, n > N$, we have $\mu(E_{mn}) < \delta$. It follows immediately $\|f_{mn}\mathbf{1}_{E_{mn}}\| < 2\varepsilon$.

Finally, $\|f_{mn}\mathbf{1}_{A\setminus E_{mn}}\| \leq t\mu(A)^{1/p} = \varepsilon$, since $f_{mn} < t$ on the complement of E_{mn}. Hence $\|f_{mn}\| < 5\varepsilon$ for $m, n \geq N$. □

Summary

Dynamic Capillarity in Porous Media - Mathematical Analysis

In this thesis we investigate pseudo-parabolic equations modelling the two-phase flow in porous media, where dynamic effects in the difference of the phase pressures are included. Specifically, the time derivative of saturation is taken into account.

The first chapter investigates the travelling wave (TW) solutions, the existence and uniqueness of smooth TW solutions are proved by ordinary differential equation techniques. The existence depends on the parameters involved. There is a threshold value for one of the parameters, the damping coefficient. When the damping coefficient is beyond the threshold value, smooth TW solutions do not exist any more. Instead, non-smooth (sharp) TW solutions are introduced and their existence is shown. The theoretical results for both smooth and non-smooth TW solutions are confirmed by numerical computations.

In the next chapter, the analysis is extended to weak solutions. In a simplified case where the mixed (time and space) order-three derivative term is linear, the existence and uniqueness of a weak solution are obtained. Next, the complex model involving nonlinear and possibly degenerate capillary induced diffusion function is considered. Then the existence is obtained by employing regularization techniques, compensated compactness and equi-integrability arguments.

Further, inspired by the nature of the capillary pressure, different formulations of the equation are introduced and their equivalence is proved.

Finally, two fully discrete numerical schemes are implemented, and compared to each other by means of two specific examples.

Samenvatting

Dynamische capillariteit in poreuze media - wiskundige analyse

In dit proefschrift beschouwen we pseudo-parabolische vergelijkingen als model voor twee-fasestromingen door poreuze media, inclusief de dynamische effecten als gevolg van drukverschillen tussen de fasen. In het bijzonder modelleren we deze effecten door de tijdsafgeleide van de verzadiging mee te nemen.

Het eerste hoofdstuk gaat over lopende-golf-oplossingen (LG) van de vergelijking. Het bestaan en de uniciteit van gladde LG-oplossingen wordt bewezen met technieken uit de gewone differentiaalvergelijkingen. Het bestaan van zulke oplossingen hangt af van de parameters in kwestie. De dempingscoëfficiënt heeft een kritieke waarde; als deze waarde wordt overschreden bestaan er geen gladde LG-oplossingen meer. Voor dat geval hebben we niet-gladde (scherpe) LG-oplossingen geïntroduceerd, en hebben we het bestaan ervan aangetoond. De theoretische resultaten voor zowel gladde en niet-gladde LG-oplossingen worden bevestigd door numerieke berekeningen.

In het volgende hoofdstuk breiden we de analyse uit naar zwakke oplossingen. In het gesimplificeerde geval waarin de samengestelde derde-orde (ruimte- en tijds-) afgeleide lineair is, wordt het bestaan en de uniciteit van een zwakke afgeleide verkregen. Daarnaast hebben we het uitgebreidere model bekeken waarin een niet-lineaire en mogelijk gedegenereerde functie voorkomt als gevolg van de capillariteit. In dat geval wordt het bestaan van zwakke oplossingen verkregen door middel van regularisatietechnieken, gecompenseerde compactheidstechnieken en equi-integreerbaarheidsargumenten.

Verder introduceren we verschillende formuleringen van de vergelijking, geïnspireerd op de aard van capillaire druk, en hebben we de equivalentie van deze formuleringen bewezen.

Tenslotte zijn er twee volledig discrete numerieke methoden geïmplementeerd, en met elkaar vergeleken aan de hand van twee specifieke voorbeelden.

Acknowledgement

This thesis presents the outcome of my PhD study at Eindhoven University of Technology from January 2008 to November 2011. It has been a very pleasant and unique experience to work in the CASA (Center for Analysis, Scientific computing and Applications) group over the past years. Therefore, I would like to take this opportunity to thank all the people who have supported and helped me.

First of all, I would like to thank my supervisor and copromotor Sorin Pop for his constant guidance, advice and suggestions during the past years. Without him, this thesis would not be possible. Second of all, I am very grateful to my promotor prof.dr.ir. C.J. van Duijn, who encouraged and inspired me through several valuable discussions during his very busy work as the rector.

I would also like to express my sincere gratitude to the members of my defense committee including prof.dr.ir. L.A. Peletier, prof.dr. J. Hulshof, prof.dr. M.A. Peletier, prof.dr.ir. S.M. Hassanizadeh, prof.dr. Ben Schweizer, prof.dr. Christian Rohde.

Next, I would like to express my profound respect and gratitude to prof.dr.ir. L.A. Peletier, I learned a lot from his dedication and attitude to mathematics when we work together on a paper. I am also grateful to him about his valuable suggestions and advice on my thesis. I would also like to thank prof.dr. M.A. Peletier for his mentoring now and then (Wednesday morning sessions, CASA-days and so on), he even taught me how to play squash better :). Furthermore, many thanks to prof.dr. J. Hulshof for his time reading the draft of my thesis and for his useful comments.

I would also like to thank prof.dr.ir. S.M. Hassanizadeh and prof.dr.-ing. Rainer Helmig for giving me opportunity to visit Institute of Hydraulic Engineering, Dept. of Hydromechanics and Modelling of Hydrosystems at University of Stuttgart for two months. I am also indebted to prof.dr. Christian Rohde and prof. dr. Barbara Wohlmuth for their useful suggestion on my research during my stay. I would also thank all the members of Rainer's group for their hospitality and help during my stay: Alex, Andreas, Anozie, Benjamin, David, Karin, Katherina, Klaus, Maria, Markus, Melanie, Onur, Philipp, Prudence, Yufei. Special thanks go to NUPUS (Non-linearities and upscaling in porous media) for financially supporting the visit and other conferences.

I am also thankful to all the CASA members, who create a great working atmosphere. Many thanks go to our two secretaries Marèse and Enna for taking care of all administrative details. In particular, I would like to thank: Agnieszka, Ali, Andriy, Bas, Badr, Carlo, David, Eric, Evgeniya, Erwin, Hans, Kakuba, Laura, Lucia, Maria, Maria, Mark, Marko, Maxim, Miguel, Mirela, Oleg, Patricio, Peter, Roxana, Steffen, Sudhir, Tasnim, Valeriu, Xiulei, Yves, Zoran. Special thanks go to Michiel Renger, who helped me with the writing of 'Samenvatting'. Furthermore, I would like to thank my officemate Kundan Kumar, with whom I shared a lot of happy time together: nice scientific and non-scientific discussions, tennis playing, squash playing and so on.

Many thanks go to Liu Lei for being a wonderful roommate, for the fun time (dinner, tennis and basketball games, wii fighting and so on) together. I also would like to thank all my friends for enriching my personal life in Holland.

Last but not the least, I would like to thank my wife Qin Lu for her unconditional support in the last four years.

Curriculum Vitae

Yabin Fan was born on March 12, 1983 in Henan, China. He graduated from The Affiliated Middle School of Henan Normal University, and then started his study in the School of Mathematical Sciences at Fudan University in 2000. After four years of Bachelor and one year of Master studies, he came to Europe and followed the Erasmus Mundus Master Program "Industrial Mathematics". He graduated from both the Eindhoven University of Technology and the University of Kaiserslautern.

In January 2008, he started a PhD in the Department of Mathematics and Computer Science, Eindhoven University of Technology, under the supervision of prof.dr.ir. C.J. van Duijn and dr. I.S. Pop. The results of the PhD study are presented in this thesis.