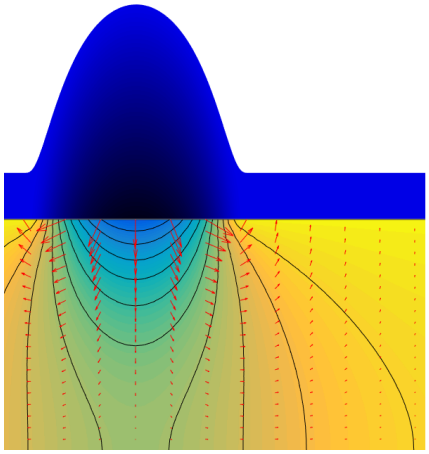
Universität Stuttgart

SimTech
Cluster of Excellence

Bachelor Thesis

Shallow Water Flow on Porous Media



Florian List

Supervisors:

Prof. Dr. Christian Rohde

Institute of Applied Analysis and
Numerical Simulation
University of Stuttgart

Prof. Dr. Florin Radu

Department of Mathematics

University of Bergen

September 2015

Contents

Introduction

The origins of ambitious efforts to understanding the mechanics of fluids date back to Ancient Greece when Archimedes discovered the famous principle of buoyancy bearing his name more than 2000 years ago. Ever since, outstanding scientists have devoted their research to the formulation of principles elucidating the physics of fluid flow in manifold phenomena. The introduction of infinitesimal calculus into fluid mechanics in the 18th century by Leonhard Euler amongst others constituted a landmark in the development of a sound mathematical framework and laid the foundation for the Navier–Stokes equations which are widely used for the modelling of fluid flow nowadays. However, technical innovations in the course of the industrial revolution necessitated the derivation of rules of thumb in order to make statements about the behaviour of mechanical systems since partial differential equations providing no analytical solutions were of little use when it came to engineering challenges in those days.
In 1856, Henry Darcy presented a formula describing flow through porous media, which he found phenomenologically in the context of drafts for a water system in the French city of Dijon [11]. For the description of fluid flow in open channels, a helpful simplification of the Navier–Stokes equations was suggested in 1871 by Adhémar Jean Claude Barré de Saint-Venant, known as the shallow water equations [12]. As to the flow in porous media, Lorenzo Adolph Richards published an equation representing flow of water in unsaturated domains in 1931, based upon an extension of Darcy's law to multiphase flow by Edgar Buckingham [48].
Since the computer age began various numerical algorithms for the approximate solution of differential equations have been established and today theorems considering the existence and uniqueness of analytical solutions of both the shallow water equations and Richards' equation are at hand, as well as numerous results on the convergence of numerical methods for solving these equations.
In the recent decades, emphasis has been placed on the investigation of coupled models incorporating different flow domains, motivated by environmental and technical issues. In this spirit, we will deal with the coupling of surface and subsurface flow in this thesis, which is of great interest in many hydrological applications such as the interaction between surface runoff and groundwater.
The thesis is structured as follows: in Chapter 1, fundamentals of fluid flow are introduced and the derivations of the partial differential equations for either flow domain are presented. Then, a coupled surface-subsurface model is formulated. Chapter 2 is dedicated to the conservation of mass and estimates for the energy of the coupled system. In Chapter 3, numerical methods for solving the governing equations are considered and an algorithm for the solution of the coupled system is derived. The performance of the coupling algorithm is investigated by means of numerical experiments in Chapter 4. Finally, we end with conclusions and an outlook to further research topics.

Chapter 1

Mathematical modelling

In this chapter, we derive a mathematical formulation for flow in a coupled system consisting of a surface $\Omega_{\text{ff}} \cong \mathbb{R}$ and a subsurface $\Omega_{\text{pm}} \subset \mathbb{R}^2$. To this end, we point out necessary assumptions allowing for the one-dimensional description of the surface flow and from which the shallow water equations are then deduced. Furthermore, we consider the derivation of Richards' equation for the subsurface flow from the balance of mass and a generalized form of Darcy's law. Finally, the governing partial differential equations are supplemented with appropriate coupling conditions in order to arrive at a sound mathematical description of the coupled model.

1.1 Surface flow: The shallow water equations

Mathematical modelling of fluid flow at the surface is often based on hyperbolic systems of partial differential equations expressing the conservation of physical quantities such as mass, momentum or energy. In one space dimension, these conservation laws take the form

$$\partial_t u(x,t) + \partial_x \Phi(u(x,t)) = f(x,t), \qquad (x,t) \in \mathbb{R} \times (0,T), \tag{1.1}$$

where the solution vector $u\colon \mathbb{R} \times [0,T) \to \mathbb{R}^m$ is gathering $m \in \mathbb{N}$ conserved properties and $T > 0$ is a period of time. The flux $\Phi\colon \mathbb{R}^m \to \mathbb{R}^m$ is a (usually non-linear) function of the solution vector and $f\colon \mathbb{R} \times [0,T) \to \mathbb{R}^m$ is a source term, independent of the solution in our considerations. Since equation (1.1) only expresses the evolution of quantities, initial conditions have to be imposed in addition, as well as boundary conditions in case of a bounded flow domain $\Omega_{\text{ff}} \subsetneq \mathbb{R}$. However, the correct choice of boundary conditions for hyperbolic systems is intricate and in what follows, we simply assume periodic boundary conditions for the surface flow. The shallow water equations are widely used for the modelling of surface flow and represent a prototype of a system of conservation laws. Despite a couple of assumptions underlying, they are often found to provide a sufficiently exact description of surface flow problems.

1.1.1 Derivation of the shallow water equations

Consider fluid flow in applications where the horizontal length scale is much larger than the vertical scale. This applies not only to undisturbed flows in flumes, but also to tidal waves in the sea or flood waves in rivers

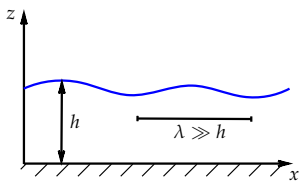


Figure 1.1: Tidal wave

featuring a wavelength λ [m] much greater than the water height h [m] (Figure 1.1). Then, the assumption of a hydrostatic pressure distribution for the pressure p [kg m^{-1} s^{-2}] and the neglect of the vertical velocity v_z [m s^{-1}] are justifiable. Furthermore, we assume that the horizontal velocity v [m s^{-1}] is independent of the vertical coordinate z. In addition, the compressibility of the fluid is neglected. We thus obtain the assumptions

$$\begin{aligned} p(x,z,t) &= \rho g(h(x,t)-z), \\ v_z(x,z,t) &= 0, \\ v(x,z,t) &= v(x,t), \\ \rho(x,z,t) &= \rho, \end{aligned} \tag{1.2}$$

in which ρ [kg m^{-3}] denotes the fluid density, g [m s^{-2}] is the gravitational constant and t [s] is the time. The variable z [m] stands for the vertical position and the horizontal dimension is described by the spatial variable x [m]. In addition, the Coriolis force may be included as in [39, chap. 2]. One could now introduce assumptions (1.2) into the Navier–Stokes equations as in [61, sec. 2.5] and integrate them over the water depth or pursue the approach in [57, sec. 1.3] incorporating a virtual fluid. As we settle for the one-dimensional case, we proceed by directly computing balances of mass and momentum of a one-dimensional flux at an arbitrary time $t \in [0,T)$ in a volume element

$$\Delta V = \Delta V(\xi,\Delta x,t) = \{(x,y,z) \in \mathbb{R}^3 \colon x \in [\xi,\xi+\Delta x],\ y \in [0,1],\ z = z(x,t) \in [0,h(x,t)]\},$$

for $\xi \in \Omega_{\mathrm{ff}}$ and $\Delta x > 0$, so that $\xi + \Delta x \in \Omega_{\mathrm{ff}}$. The width of ΔV is taken to be 1 [m]. The horizontal boundaries of ΔV are thus fixed and the height at each point is given by the water height h depending on the coordinate x and on the time t. Sufficient smoothness of h and v is presumed in the following.

Balance of mass

The balance of the mass m [kg] in ΔV accounts for fluxes inside the flow domain, fluxes through the boundaries and sources and sinks in the interior of ΔV (Figure 1.2). The Reynolds transport theorem yields after division by the constant width 1 [m]

$$1\mathrm{m}^{-1} \cdot \partial_t m(t) = \partial_t \int_\xi^{\xi+\Delta x} (\rho h)(x,t)\,dx + (\rho h v)(\xi+\Delta x,t) - (\rho h v)(\xi,t) = \int_\xi^{\xi+\Delta x} q_M(x,t)\,dx, \tag{1.3}$$

where q_M [kg m^{-2} s^{-1}] is an internal mass source. Note that fluxes over the boundaries of ΔV only occur at the two sides exhibiting normal vectors parallel to the x-axis since we consider flows in one horizontal direction, due to assumption $(1.2)_2$ and because the upper boundary of ΔV follows the flow of the fluid. Besides, these fluxes are independent of the vertical coordinate z in light of assumptions $(1.2)_3$ and $(1.2)_4$, and perpendicular to the boundaries since the vertical velocities equal zero owing to assumption $(1.2)_2$. Assumption $(1.2)_4$ permits to divide equation (1.3) by $\rho > 0$ and we get in the limit $\Delta x \to 0$ after division by Δx

$$\partial_t h(\xi,t) + \partial_x (hv)(\xi,t) = \frac{q_M(\xi,t)}{\rho}. \tag{1.4}$$

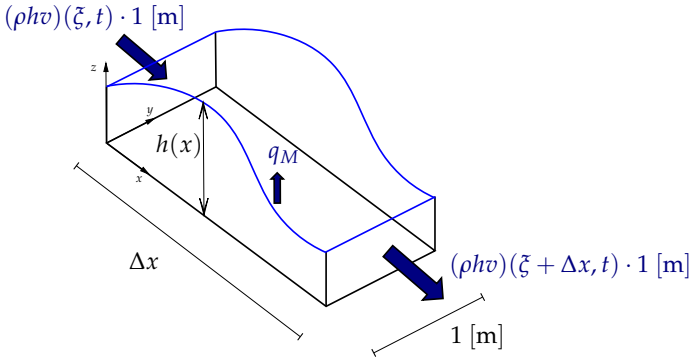


Figure 1.2: Balance of mass in the volume element ΔV

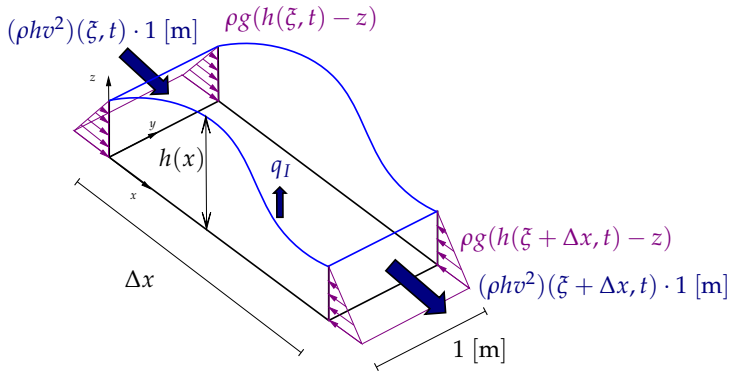


Figure 1.3: Balance of momentum in the volume element ΔV

Balance of momentum

Likewise, the conservation of the linear momentum I [kg m s^{-1}] in ΔV reads in consideration of the pressure forces at the boundaries (Figure 1.3)

$$\begin{aligned} 1\mathrm{m}^{-1}\cdot\partial_t I(t) &= \partial_t\int_\xi^{\xi+\Delta x}(\rho hv)(x,t)\,dx+(\rho hv^2)(\xi+\Delta x,t)-(\rho hv^2)(\xi,t)\\ &=\int_0^{h(\xi,t)}p(\xi,z,t)\,dz-\int_0^{h(\xi+\Delta x,t)}p(\xi+\Delta x,z,t)\,dz+\int_\xi^{\xi+\Delta x}q_I(x,t)\,dx, \end{aligned} \tag{1.5}$$

denominating the source term of linear momentum q_I [kg m^{-1} s^{-2}].
Inserting the assumption of a hydrostatic pressure distribution $(1.2)_1$ into equation (1.5) and dividing by $\rho>0$ yields further

$$\begin{aligned} &\partial_t\int_\xi^{\xi+\Delta x}(hv)(x,t)\,dx+(hv^2)(\xi+\Delta x,t)-(hv^2)(\xi,t)\\ &=g\int_0^{h(\xi,t)}(h(\xi,t)-z)\,dz-g\int_0^{h(\xi+\Delta x,t)}(h(\xi+\Delta x,t)-z)\,dz\\ &\quad+\frac{1}{\rho}\int_\xi^{\xi+\Delta x}q_I(x,t)\,dx\\ &=g\int_0^{h(\xi,t)}(h(\xi,t)-h(\xi+\Delta x,t))\,dz\\ &\quad-g\int_{h(\xi,t)}^{h(\xi+\Delta x,t)}h(\xi+\Delta x,t)\,dz+g\int_{h(\xi,t)}^{h(\xi+\Delta x,t)}z\,dz\\ &\quad+\frac{1}{\rho}\int_\xi^{\xi+\Delta x}q_I(x,t)\,dx. \end{aligned} \tag{1.6}$$

Dividing equation (1.6) by Δx and considering $\Delta x\to 0$, we obtain

$$\begin{aligned} \partial_t(hv)(\xi,t)+\partial_x(hv^2)(\xi,t) &= -g\,(\partial_x h)(\xi,t)\,h(\xi,t)-g\,(\partial_x h)(\xi,t)h(\xi,t)+g\,(\partial_x h)(\xi,t)h(\xi,t)+\frac{q_I(\xi,t)}{\rho}\\ &=-g\,(\partial_x h)(\xi,t)\,h(\xi,t)+\frac{q_I(\xi,t)}{\rho}\\ &=-\partial_x\left(\frac{1}{2}gh(\xi,t)^2\right)+\frac{q_I(\xi,t)}{\rho}. \end{aligned} \tag{1.7}$$

The momentum source term q_I is furnished by an empirical model taking the bottom stress, wind stresses, internal friction and small slopes into account.

System of shallow water equations

Altogether, we summarize the shallow water equations (also called Saint-Venant equations) (1.4) and (1.7) as a system of coupled non-linear equations:

$$\partial_t \begin{pmatrix} h \\ hv \end{pmatrix} + \partial_x \begin{pmatrix} hv \\ hv^2 + \frac{1}{2}gh^2 \end{pmatrix} = \frac{1}{\rho} \begin{pmatrix} q_M \\ q_I \end{pmatrix}. \tag{1.8}$$

Writing $u = (u_1, u_2)^T := (h, hv)^T$, the system (1.8) takes the form (1.1) with flux function

$$\Phi(u_1, u_2) = \begin{pmatrix} u_2 \\ \frac{u_2^2}{u_1} + \frac{1}{2}gu_1^2 \end{pmatrix}, \tag{1.9}$$

well-defined for $h > 0$, and source term $f = (f_1, f_2)^T := \rho^{-1}(q_M, q_I)^T$.
Applying some algebraic conversions, the system (1.8) can also be expressed in a reduced form for the variables h and v:

$$\partial_t \begin{pmatrix} h \\ v \end{pmatrix} + \partial_x \begin{pmatrix} hv \\ \frac{1}{2}v^2 + gh \end{pmatrix} = \frac{1}{\rho} \begin{pmatrix} q_M \\ \frac{q_I}{h} - \frac{q_M v}{h} \end{pmatrix}. \tag{1.10}$$

1.1.2 Simplifications of the shallow water equations

Further simplifications of the shallow water equations are widely used and have been deployed in the framework of coupled surface-subsurface models e.g. in [30, 37, 54]. They are obtained by neglecting several terms in the shallow water equations and by employing empirical models for the friction. For the sake of convenience, we assume that no mass sources occur, i.e. $q_M = 0$.

Kinematic wave equation

For the derivation of the kinematic wave equation, the Gauckler–Manning–Strickler formula (see [15, 38]) is used to describe the friction and moreover, the local inertia term $\partial_t v$, the convective inertia term $\partial_x(v^2/2)$ and the pressure term $\partial_x(gh)$ are assumed to be small and hence left out in system (1.10). We deduce

$$q_I = 0, \tag{1.11}$$

and split up the source of momentum in a part resulting from the bottom slope I_0 $[-]$ and a part acknowledging the friction expressed by the energy gradient I_E $[-]$, which gives

$$q_I = \rho g h(I_0 - I_E) = 0, \tag{1.12}$$

and consequently

$$I_0 = I_E. \tag{1.13}$$

The energy gradient is determined by Manning's formula

$$I_E = \frac{n^2 v^2}{r_{\text{hy}}^{4/3}}, \tag{1.14}$$

where n $[\text{m}^{-1/3}\,\text{s}]$ denotes the Gauckler–Manning coefficient accounting for the roughness and the sinuosity of the river bed and r_{hy} $[\text{m}]$ stands for the hydraulic radius. Furthermore, we assume a proportionality between the hydraulic radius and the water height, that is $r_{\text{hy}} = \alpha h$.
Rewriting Manning's formula (1.14) and making use of condition (1.13) yields

$$v = \frac{I_E^{1/2} r_{\text{hy}}^{2/3}}{n} = \frac{I_0^{1/2} r_{\text{hy}}^{2/3}}{n} = \frac{I_0^{1/2} \alpha^{2/3}}{n} h^{2/3} =: V h^{2/3}, \tag{1.15}$$

where the constant velocity is given by $V = (I_0^{1/2}\alpha^{2/3})/n$. Plugging equation (1.15) into the mass conservation in system (1.10) leaves the kinematic wave equation

$$\partial_t h + \partial_x \left(V h^{5/3} \right) = 0. \tag{1.16}$$

A characteristic feature of solutions of the kinematic wave equation (1.16), so-called kinematic waves, is that they steepen over time. Though the neglect of the velocity and pressure terms is a rigorous assumption, helpful results e.g. for the runoff of precipitation are achieved based on kinematic wave approximations (see, for instance, [7, sec. 5.5]).
As a compromise between the shallow water equations and the kinematic wave equation, retaining the pressure term and omitting the velocity terms gives rise to a scalar equation whose solution is known as diffusive wave [22, chap. 2].

Linear transport equation

The linear transport equation requires the same assumptions as the kinematic wave equation and emanates from a linear relation between the friction and the velocity employed instead of the Gauckler–Manning–Strickler formula, that is $I_E = \beta v$. Mimicking the derivation of the kinematic wave equation leads to the linear transport equation

$$\partial_t h + \partial_x (Vh) = 0, \tag{1.17}$$

where $V = I_0/\beta$.

1.2 Subsurface flow: Richards' equation

Unlike surface flow, fluid flow in the subsurface is characterized by the presence of several entities, such as a liquid phase, a gas phase and a solid matrix interacting with each other. As the exact structure in the subsurface is often unknown and computation resolving the detailed micro-structure would be expensive anyway, averaged equations are typically employed. Some of them have been found empirically, one of the most prominent being Darcy's law which is presented subsequently. Later, formulas describing the macroscopic flow behaviour have been derived mathematically from the Navier–Stokes equation by means of homogenization as shown in [23, chap. 3] and in [37, chap. 2], through volume averaging as in [62, chaps. 4–5] or through thermodynamically constrained averaging theory in [18, chap. 9]. First of all, we introduce some fundamental properties of subsurface flow.

1.2.1 Fundamental properties of subsurface flow

We assume the subsurface to consist of a material containing pores which enables the flow of fluids, for example sand or clay, termed a porous medium. In order to identify attributes specifying the physical behaviour on a macroscale, a representative elementary volume (REV) is defined which affords the averaging of properties while still permitting the distinction of inhomogeneities (Figure 1.4). A detailed introduction into the application of a REV in fluid mechanics can be found in [5, chap. 1] and [21, chap. 2].

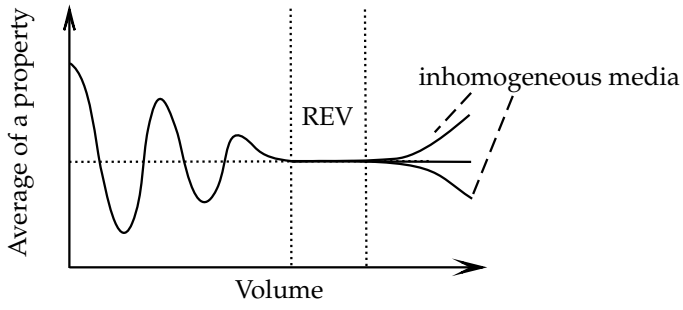


Figure 1.4: Definition of the REV [5, 21]

Porosity

We harness the REV to define the porosity ϕ $[-]$ of a porous medium as the ratio of the volume accessible to fluids within the REV, V_{pore} $[\text{m}^3]$, to the total volume of a REV, V_{total} $[\text{m}^3]$, namely

$$\phi := \frac{V_{\text{pore}}}{V_{\text{total}}}. \tag{1.18}$$

Discharge velocity

The quotient of the total discharge Q $[\text{m}^3\ \text{s}^{-1}]$ through a cross-sectional area A $[\text{m}^2]$ is referred to as the discharge velocity v $[\text{m}\ \text{s}^{-1}]$ (also called specific discharge or Darcy velocity)

$$v := \frac{Q}{A}. \tag{1.19}$$

In case of multidimensional flow, both Q and v are vectorial. The absolute value of the discharge velocity is less than the absolute value of the actual velocity of a fluid particle, as the definition of the former employs the area A parts of which are not accessible to the fluid (Figure 1.5).

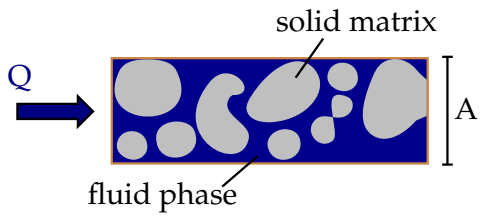


Figure 1.5: Schematic representation of flow through a porous medium

Saturation and water content

When it comes to multiphase flow, some more definitions are required. Let $\mathcal{A}$ be the set of fluid phases present in a porous medium, V_α $[\text{m}^3]$ the volume occupied by a phase $\alpha \in \mathcal{A}$ and let the fluids be immiscible. We define the saturation S_α $[-]$ of a phase $\alpha \in \mathcal{A}$ as

$$S_\alpha := \frac{V_\alpha}{V_{\text{pore}}}, \tag{1.20}$$

and its volume content θ_α as

$$\theta_\alpha := S_\alpha \phi = \frac{V_\alpha}{V_{\text{total}}}. \tag{1.21}$$

In case of the phase α being water, the property θ_α in definition (1.21) is called water content. We assume that the entire pore space is filled with fluids in $\mathcal{A}$, that is $\sum_{\alpha\in\mathcal{A}} V_\alpha = V_{\text{pore}}$. One thus asserts that

$$\begin{aligned} \sum_{\alpha\in\mathcal{A}} S_\alpha &= 1, \\ \sum_{\alpha\in\mathcal{A}} \theta_\alpha &= \phi. \end{aligned} \tag{1.22}$$

Capillary pressure

In what follows, we consider two fluid phases, i.e. $|\mathcal{A}| = 2$, the extension to more than two phases is straightforward and can be found in [21, sec. 2.4], as well as an elaborate explanation of capillarity in general. When

two fluid phases are present in a system, e.g. in case of the pore space in the subsurface being filled with both water and air, one observes that one phase exhibits a higher tendency to fill the narrow void spaces (named wetting phase, denoted by the index w) than the other phase which draws back from the small pores (named non-wetting phase, denoted by the index n). This is due to the complex interaction of cohesive and adhesive molecular forces.
This behaviour can be expressed by introducing the capillary pressure p_c [kg m^{-1} s^{-2}] as the difference between the pressure of the non-wetting phase and the wetting phase, namely

$$p_c := p_n - p_w. \tag{1.23}$$

An explicit formulation of the capillary pressure as a function of the interfacial tension and of the wetting angle is deduced from an equilibrium of forces at the interface of the phases as in [21, sec. 2.4], revealing that the capillary pressure is proportional to the reciprocal of the pore size. Turning away from microscopical considerations, a macroscopic quantity has to be determined which the capillary pressure depends on, so as to link the capillary pressure to the properties defined so far. The volume content (alternatively the saturation just as well) serves this purpose: When the content of the wetting phases decreases, the wetting fluid withdraws to narrow pores where capillary action is high which causes a rise of the capillary pressure. On the other hand, an increase of the content of the wetting phase provokes a decrease of the capillary action as the wetting fluid occupies large pores. Hence, we find the capillary pressure to be a function of the volume contents for a macro-scale description and since they can be expressed by each other via equation $(1.22)_2$, we write

$$p_c = p_c(\theta_w). \tag{1.24}$$

However, equation (1.23) is only valid for creeping flow and thus, under the influence of transient flow, dynamic capillary effects should be taken into account, as e.g. in [20, 27, 58]. Then, one gets

$$p_c(\theta_w, \partial_t \theta_w) = p_n - p_w + \tau(\theta_w)\, \partial_t \theta_w, \tag{1.25}$$

in which $\tau(\theta_w) \geq 0$ is a non-equilibrium coefficient. In addition, consideration of the hysteresis expressing a difference between the drainage and the imbibition in view of the capillary pressure might be necessary in some cases as carried out e.g. in [9].
In this thesis, we content ourselves with the capillary pressure expressed by relations (1.23) and (1.24).

Hydraulic conductivity

A macroscopic parameter describing the ease for a fluid phase to move through a porous medium whose pores are occupied by several fluid phases is the hydraulic conductivity K_α [m s^{-1}]. For a phase $\alpha \in \mathcal{A}$, it is defined as

$$K_\theta = K_\alpha(\theta_\alpha) = k\, k_{r\alpha}(\theta_\alpha)\, \frac{\rho_\alpha g}{\mu_\alpha}, \tag{1.26}$$

in which k [m^2] denotes the intrinsic permeability characterizing the porous medium (matrix-valued in the most general case), μ [kg m^{-1} s^{-1}] the dynamic viscosity of phase α, ρ_α [kg m^{-3}] the density of phase α and g [m s^{-2}] the gravity constant. The relative permeability $k_{r\alpha} = k_{r\alpha}(\theta_\alpha)$ [$-$] is taken to be a function of the volume content θ_α and epitomizes the obstructive influence of the presence of a phase on the flow behaviour of another one and vice versa. Its values lie within the range from 0 to 1, in particular $k_{r\alpha}(\theta_\alpha = \phi) = 1$ and $k_{r\alpha}(\theta_\alpha = 0) = 0$ hold true.
Since we intend to investigate the modelling of water flow in a possibly unsaturated porous medium close to the surface whose pores contain air beside water, it stands to reason to resort to Richards' equation instead of employing a full two-phase model.

1.2.2 Derivation of Richards' equation

Richards' equation, first formulated in 1931 by Lorenzo A. Richards [48], is employed for two-phase flow problems including a liquid phase w (wetting) and a gas phase n (non-wetting), in geological applications commonly water and air. For this reason, we name the property θ_w water content hereafter. It rests upon the prerequisite that the gas phase is connected to the exterior atmosphere implying that the assumption of

constant pressure for the gas phase is justified. We omit the evident dependence of the properties on the space and time variables in order to simplify the notation.
The balance of mass for the liquid phase $w \in \mathcal{A} = \{w, n\}$ in an arbitrary control volume $\Omega \subset \mathbb{R}^d$ for $d \in \{1, 2, 3\}$ with boundary $\partial\Omega$ and outward pointing normal vector n for phase w in a porous medium with porosity ϕ $[-]$ considering outflows or inflows over the boundary $\partial\Omega$ and internal sources is written as

$$\partial_t m_w = \partial_t \int_\Omega \rho_w S_w \phi \, dV + \int_{\partial\Omega} (\rho_w v_w) \cdot \vec{n} \, dA = \int_\Omega q_w \, dV. \tag{1.27}$$

We assume continuity of the integrands and change the order of integration and differentiation in the first integral of equation (1.27). Additionally, we apply the divergence theorem. Consequently

$$\int_\Omega \partial_t \left(\rho_w S_w \phi\right) \, dV + \int_\Omega \nabla \cdot \left(\rho_w v_w\right) \, dV = \int_\Omega q_w \, dV. \tag{1.28}$$

Since equation (1.28) holds for an arbitrary control volume Ω, we claim equality of the integrands in equation (1.28) to obtain

$$\partial_t \left(\rho_w S_w \phi\right) + \nabla \cdot \left(\rho_w v_w\right) = q_w. \tag{1.29}$$

Acknowledging the insignificance of the compressibility of the liquid phase w (e.g. for water at 25 °C: $4.6 \cdot 10^{-10}$ $[\mathrm{m\,s^2\,kg^{-1}}]$ [14]), we set $\rho_w(\vec{x}, t) \equiv \rho_w > 0$ in equation (1.29) and furthermore, we recall definition (1.21) of the water content θ_w. It follows that

$$\partial_t \theta_w + \nabla \cdot v_w = \frac{q_w}{\rho_w}. \tag{1.30}$$

In order to attain to an expression for the discharge velocity of the liquid phase v_w, an extended form of Darcy's law is introduced. Originally, Darcy's law was presented for the description of saturated flow in [11]. Buckingham was first to carry it over to the application on unsaturated flow. It then involves the hydraulic conductivity K_w $[\mathrm{m\,s^{-1}}]$ and states the relation

$$v_w = -K_w(\theta_w) \nabla(\psi_w + z), \tag{1.31}$$

between the discharge velocity v_w $[\mathrm{m\,s^{-1}}]$ and the negative gradient of the pressure head ψ_w $[\mathrm{m}]$ which is related to the pressure p_w by the relation $\psi_w := \frac{p_w}{\rho_w g}$. The extended Darcy law (1.31) is only valid for creeping flow characterized by small Reynolds numbers which we assume in the subsurface. It hence provides a simple equation expressing the balance of linear momentum for subsurface flow problems. A rigorous mathematical derivation of Darcy's law from the Navier–Stokes equation can be found in the references at the beginning of this section. Substituting equation (1.31) into equation (1.30) leads to

$$\partial_t \theta_w - \nabla \cdot \left[K_w(\theta_w) \nabla(\psi_w + z)\right] = \frac{q_w}{\rho_w}. \tag{1.32}$$

Finally, a constitutive relationship between the water content θ_w and the pressure height ψ_w is needed. We obtain it by recalling equations (1.23) and (1.24) and making the assumption of constant gas pressure in the domain, that is $p_n \equiv 0$. This is reasonable whenever the flow domain is interconnected and connected to the atmosphere [41, chap. 3]. We thus get

$$p_c(\theta_w) = -p_w. \tag{1.33}$$

If the parametrization $p_c(\theta_w)$ is strictly monotonically decreasing, which is usually the case, equation (1.33) can be inverted to obtain an explicit relationship

$$\theta_w = \theta_w(\psi_w). \tag{1.34}$$

By inserting equation (1.34) into equation (1.32), we arrive at Richards' equation in the mixed pressure-saturation form which writes

$$\partial_t \theta_w(\psi_w) - \nabla \cdot \left[K_w(\theta_w(\psi_w)) \, \nabla(\psi_w + z)\right] = \frac{q_w}{\rho_w}. \tag{1.35}$$

Richards' equation (1.35) is a non-linear parabolic partial differential equation for the pressure height ψ_w, which degenerates to an elliptic equation in saturated zones where $\partial_t \theta_w = 0$.

1.2.3 Simplifications of Richards' equation

In some special cases, it turns out that simplifications respectively slight modifications of Richards' equation (1.35) are feasible. In case of a fully saturated porous medium, one gets $\partial_t \theta(\psi_w) \equiv 0$, in addition $k_r \equiv 1$ and one ends up with an elliptic linear steady-state equation

$$-\nabla \cdot \left[\frac{k\rho_w g}{\mu_w} \nabla(\psi_w + z) \right] = \frac{q_w}{\rho_w}. \tag{1.36}$$

Expansion of Darcy's law by a viscous term for the saturated case was first proposed by Brinkman. Brinkman's approach is used to model highly porous media (e.g. in [4, 29, 36]) and leads to the system

$$\begin{aligned} v_w &= -\frac{k\rho_w g}{\mu_w} \nabla(\psi_w + z) + \varepsilon \Delta v_w, \\ \nabla \cdot v_w &= \frac{q_w}{\rho_w}, \end{aligned} \tag{1.37}$$

with a viscosity parameter $\varepsilon > 0$ $[\mathrm{m}^2]$. Anyhow, as both Darcy-type equation (1.36) and Brinkman's equation (1.37) are only applicable for one-phase flow problems, we employ Richards' equation (1.35) hereafter to examine the impact of the shallow water flow at the surface on the saturation in unsaturated regimes.

1.2.4 Parametrizations of hydraulic relationships

There exist several parametrizations of the water content $\theta_w(\psi_w)$ and the hydraulic conductivity $K_w(\psi_w)$ based on experiments, the Brooks–Corey model and the van Genuchten–Mualem model being prominent examples among them. For all numerical experiments in this thesis, we apply the van Genuchten–Mualem model [59] given by

$$\begin{aligned} \theta_w(\psi_w) &= \begin{cases} \theta_R + (\theta_S - \theta_R) \left[\frac{1}{1+(-\alpha\psi_w)^n} \right]^{\frac{n-1}{n}}, & \psi_w \leq 0, \\ \theta_S, & \psi_w > 0, \end{cases} \\ K_w(\psi_w) = K_w(\theta_w(\psi_w)) &= \begin{cases} K_S \theta_w(\psi_w)^{\frac{1}{2}} \left[1 - \left(1 - \theta_w(\psi_w)^{\frac{n}{n-1}} \right)^{\frac{n-1}{n}} \right]^2, & \psi_w \leq 0, \\ K_S, & \psi_w > 0, \end{cases} \end{aligned} \tag{1.38}$$

in which θ_S and K_S denote the water content respectively the hydraulic conductivity of the fully saturated porous medium, θ_R is the residual water content and α and n are curve fitting parameters associated with the soil properties. Typical curves are depicted in Figure 1.6.

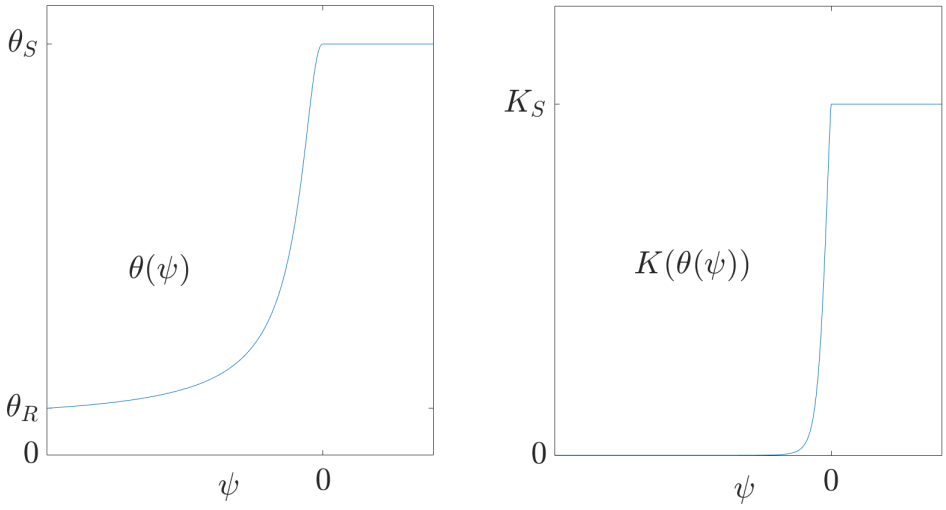


Figure 1.6: Typical profiles of $\theta(\psi)$ and $K(\theta(\psi))$ given by the van Genuchten–Mualem model

1.3 Coupled model

In this section, we bring the results of the preceding sections together and formulate a mathematical model for shallow water flow on a porous medium. Initially, we define the geometry of the coupled model. Then, coupling conditions ensuring the satisfaction of the conservation of physical quantities in the entire system are deduced which complement the governing partial differential equations.
We indicate surface and surface properties by the subscripts pm and ff and omit the qualifier for the phase of subsurface properties since all properties refer to the wetting phase which is present in both flow domains. The scalar surface velocity is denominated v and is distinguished by the vectorial velocity in the subsurface $v_{\text{pm}} := v_w$.

1.3.1 Geometry

Consider one-dimensional shallow water flow on the surface $\Omega_{\text{ff}} \cong \mathbb{R}$ coupled with two-dimensional subsurface flow (i.e. $d = 2$) in $\Omega_{\text{pm}} = (x_l, x_r) \times (z_b, z_t) \subset \mathbb{R}^2$ with $x_l < x_r$ and $z_b < z_t$. The boundary of the subsurface domain $\partial\Omega_{\text{pm}}$ and the surface Ω_{ff} coincide only at the interface $\Gamma \subset \mathbb{R}$ which is given by the non-empty set $\Gamma = \overline{\Omega_{\text{pm}}} \cap \Omega_{\text{ff}} = [x_l, x_r] \times \{z_t\}$ (Figure 1.7). We denote the spatial variable on the surface by x and in the subsurface by $\vec{x} = (x, z)$.

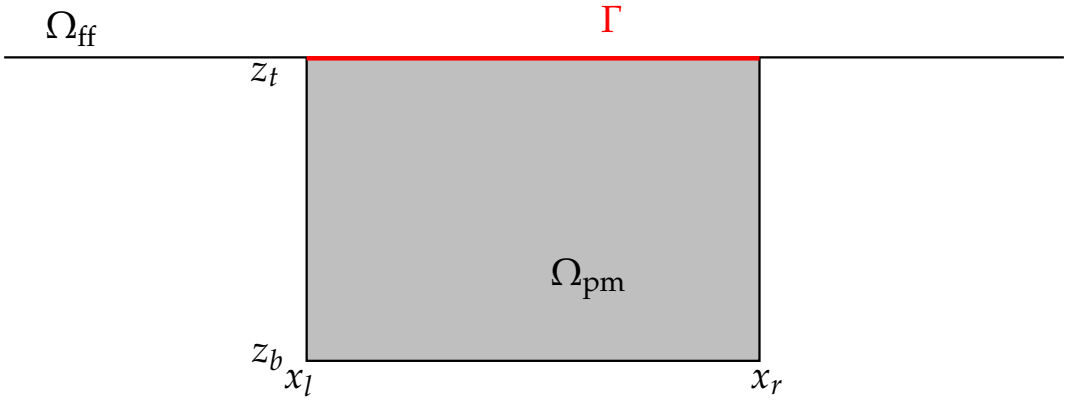


Figure 1.7: Scheme of the geometry

1.3.2 Coupling conditions

From subsurface to surface

To guarantee that the mass of the fluid is conserved in the coupled system, we claim the continuity of the flux across the interface by utilizing the vertical component of the subsurface flux at the interface as a source term for the surface flow, to be more specific

$$\frac{q_M}{\rho} = v_{\text{pm}} \cdot \vec{n} \qquad \text{on } \Gamma, \tag{1.39}$$

where q_M [kg m^{-2} s^{-1}] denotes the mass source in equation (1.4), $\vec{n}$ the outward pointing normal vector on Γ, ρ [kg m^{-3}] the density of the wetting phase. The discharge velocity v_{pm} [m s^{-1}] is given by relation (1.31) and tacitly extended by zero on $\Omega_{\text{ff}} \setminus \Gamma$. Thus, the influence of the subsurface flow on the surface flow is acknowledged.

From surface to subsurface

The converse coupling is undertaken by taking the pressure in the subsurface into account which is exerted by the weight of the mass of the fluid. This yields the Dirichlet coupling condition

$$\psi = h \qquad \text{on } \Gamma, \tag{1.40}$$

for the pressure height ψ [m], which shall be given by the water height h [m] at the interface Γ.
A slightly altered approach is presented in [37, sec. 4.2] based upon the assumption of an equilibrium of forces at the interface. It includes the definition of a normalized water height and implies that the interaction of forces from the two flow domains on each other only takes place at the interface.

1.3.3 Model formulation

We formulate the model of the coupled flow by endowing the shallow water equations (1.8) and Richards' equation (1.35) with initial data at $t = 0$ (distinguished by the subscript 0) and boundary conditions. Furthermore, we incorporate the coupling conditions (1.39) and (1.40). In the subsurface, we impose no-flow conditions on $\partial\Omega \setminus \Gamma$ and we do not consider any mass sources in neither flow domain except for the mutual mass exchange. Consequently, the model reads

$$\left.\begin{aligned}
\partial_t \begin{pmatrix} h \\ hv \end{pmatrix} + \partial_x \begin{pmatrix} hv \\ hv^2 + \frac{1}{2}gh^2 \end{pmatrix} &= \begin{pmatrix} v_{\mathrm{pm}} \cdot \vec{n} \\ q_I(h,v)/\rho \end{pmatrix}, && (x,t) \in \Omega_{\mathrm{ff}} \times (0,T), \\
h(x,0) &= h_0(x), && x \in \Omega_{\mathrm{ff}}, \\
v(x,0) &= v_0(x), && x \in \Omega_{\mathrm{ff}}, \\
\partial_t \theta(\psi) + \nabla \cdot v_{\mathrm{pm}} &= 0, && (\vec{x},t) \in \Omega_{\mathrm{pm}} \times (0,T), \\
v_{\mathrm{pm}} &= -K(\theta(\psi))\nabla(\psi + z), && (\vec{x},t) \in \Omega_{\mathrm{pm}} \times (0,T), \\
\psi(\vec{x},0) &= \psi_0(\vec{x}), && \vec{x} \in \Omega_{\mathrm{pm}}, \\
\psi &= h, && (\vec{x},t) \in \Gamma \times [0,T), \\
v_{\mathrm{pm}} \cdot \vec{n} &= 0, && (\vec{x},t) \in \partial\Omega_{\mathrm{pm}} \setminus \Gamma \times [0,T).
\end{aligned}\right\} \tag{1.41}$$

Chapter 2

Mass conservation and energy estimates

This chapter is devoted to the conservation of mass of the coupled surface-subsurface model (1.41) and to energy estimates for the uncoupled models and for a simplified coupled surface-subsurface model. We assume $\Omega_{\text{ff}} = \mathbb{R}$ in what follows which allows to neglect issues at the boundaries of the surface.

2.1 Mass conservation

The solutions of the shallow water equations as well as of Richards' equation are mass conservative since the conservation of mass is a crucial part in the derivations of these equations. The coupled model also features this physically relevant property as the following theorem shows. We set $\rho = 1$ [kg m^{-3}] and assume the subsurface flow domain Ω_{pm} in the x-z-plane as well as the surface flow domain Ω_{ff} in x-direction to have an extent of 1 [m] in y-direction so that the mass defined in the following corresponds to the physical understanding. We prove

Theorem 2.1.1 (Mass conservation for the coupled model) Let $\psi \in C^2(\Omega_{\text{pm}} \times [0,T))$, $h, v \in C^1(\mathbb{R} \times [0,T))$ be a classical solution of the coupled system (1.41). Assume that $h(\cdot,t)$ and $v(\cdot,t)$ have compact support for all $t \in [0,T)$. Then, the total mass

$$\mathcal{M} = \int_{\Omega_{\text{pm}}} \theta(\psi)\, d\vec{x} + \int_{\mathbb{R}} h\, dx \tag{2.1}$$

is conserved, that is

$$\frac{d}{dt}\mathcal{M} = 0. \tag{2.2}$$

Proof. We calculate for the subsurface flow making use of the divergence theorem

$$\frac{d}{dt}\int_{\Omega_{\text{pm}}} \theta(\psi)\, d\vec{x} = \int_{\Omega_{\text{pm}}} (\theta(\psi))_t\, d\vec{x} = -\int_{\Omega_{\text{pm}}} \nabla \cdot v_{\text{pm}}\, d\vec{x} = -\int_{\partial\Omega_{\text{pm}}} v_{\text{pm}} \cdot \vec{n}\, d\vec{\xi} = -\int_{\Gamma} v_{\text{pm}} \cdot \vec{n}\, d\vec{\xi},$$

since $v_{\text{pm}} \cdot \vec{n} = 0$ on $\partial\Omega_{\text{pm}} \setminus \Gamma$.
For the surface flow, we obtain

$$\frac{d}{dt}\int_{\mathbb{R}} h\, dx = \int_{\mathbb{R}} h_t\, dx = -\int_{\mathbb{R}} (hv)_x\, dx + \int_{\mathbb{R}} v_{\text{pm}} \cdot \vec{n}\, dx = hv|_{-\infty}^{\infty} + \int_{\Gamma} v_{\text{pm}} \cdot \vec{n}\, d\vec{\xi},$$

and thus

$$\frac{d}{dt}\mathcal{M} = -\int_{\Gamma} v_{\text{pm}} \cdot \vec{n}\, d\vec{\xi} + \int_{\Gamma} v_{\text{pm}} \cdot \vec{n}\, d\vec{\xi} = 0.$$

□

2.2 Energy estimates

Unlike the conservation of mass for the respective models for surface and subsurface flows, it might be not obvious that the solutions of the separate models conserve energies of a certain form or at least satisfy energy estimates. Therefore, we first examine the models individually and then consider estimates for coupled surface-subsurface models.

2.2.1 Surface flow

For surface flow, we prove the following theorems:

Theorem 2.2.1 (Conservation of energy for the linear transport equation and the kinematic wave equation) Let $h \in C^1(\mathbb{R} \times [0,T))$ be a classical solution of the linear transport equation (1.17) or of the kinematic wave equation (1.16) for $T > 0$. Assume further that $h(\cdot,t)$ has compact support for all $t \in [0,T)$. Then, the energy $W = W(h)$ of the system, given by

$$W(h) = \frac{h^2}{2}, \tag{2.3}$$

is conserved, that is

$$\frac{d}{dt}\int_{\mathbb{R}} W(h)\,dx = 0. \tag{2.4}$$

Proof. We compute in case of the linear transport equation

$$\frac{d}{dt}\int_{\mathbb{R}} W(h)\,dx = \int_{\mathbb{R}} h_t h\,dx = -V\int_{\mathbb{R}} h_x h\,dx = -V\int_{\mathbb{R}}\left(\frac{h^2}{2}\right)_x dx = -\,V\,\frac{h^2}{2}\Big|_{-\infty}^{\infty} = 0.$$

Similarly, we obtain for a solution of the kinematic wave equation

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{R}} W(h)\,dx = \int_{\mathbb{R}} h_t h\,dx = -V\int_{\mathbb{R}}\left(h^{5/3}\right)_x h\,dx = V\int_{\mathbb{R}} h^{5/3}h_x\,dx - Vh^{8/3}\Big|_{-\infty}^{\infty}\\ = \frac{3}{8}V\int_{\mathbb{R}}\left(h^{8/3}\right)_x dx = \frac{3}{8}\,Vh^{8/3}\Big|_{-\infty}^{\infty} = 0.\end{aligned}$$

□

Theorem 2.2.2 (Conservation of energy for the shallow water equations) Let $h, v \in C^1(\mathbb{R} \times [0,T))$ be a classical solution of the shallow water equations (1.8) for $T > 0$ and let be $q_M = q_I = 0$. Assume further that $h(\cdot,t)$ and $v(\cdot,t)$ have compact support for all $t \in [0,T)$. Then, the energy $W = W(h,v)$ of the system, given by

$$W(h,v) = \frac{gh^2}{2} + \frac{hv^2}{2}, \tag{2.5}$$

is conserved, that is

$$\frac{d}{dt}\int_{\mathbb{R}} W(h,v)\,dx = 0. \tag{2.6}$$

Proof. We start by introducing conservative variables h and $q = hv$, so that W takes the form

$$W(h,q) = \frac{gh^2}{2} + \frac{q^2}{2h}.$$

Recall that W is well-defined since $h > 0$.
We calculate the scalar product of system (1.8) with $\nabla_{(h,q)}W(h,q)$ which yields

$$h_t gh - \frac{1}{2}h_t\frac{q^2}{h^2} + q_t\frac{q}{h} = -q_x gh + \frac{1}{2}q_x\frac{q^2}{h^2} - \left(\frac{q^2}{h}\right)_x\frac{q}{h} - \left(\frac{gh^2}{2}\right)_x\frac{q}{h}. \qquad (\star)$$

We compute making use of equation $(\star)$

$$\begin{aligned}
\frac{d}{dt}\int_{\mathbb{R}} W(h,q)\,dx &= \int_{\mathbb{R}} h_t g h - \frac{1}{2}h_t\frac{q^2}{h^2} + q_t\frac{q}{h}\,dx \\
&= \int_{\mathbb{R}} -q_x g h + \frac{1}{2}q_x\frac{q^2}{h^2} - \left(\frac{q^2}{h}\right)_x\frac{q}{h} - \left(\frac{gh^2}{2}\right)_x\frac{q}{h}\,dx \\
&= \underbrace{\int_{\mathbb{R}} -q_x g h - q g h_x\,dx}_{=:I_1} + \underbrace{\int_{\mathbb{R}} \frac{1}{2}q_x\frac{q^2}{h^2} - \left(\frac{q^2}{h}\right)_x\frac{q}{h}\,dx}_{=:I_2} \\
&= I_1 + I_2.
\end{aligned}$$

Integration by parts and exploiting the compactness of the supports of h and v leads to

$$I_1 = \int_{\mathbb{R}} -q_x g h - q g h_x\,dx = \int_{\mathbb{R}} q g h_x - q g h_x\,dx - q g h|_{-\infty}^{\infty} = 0.$$

The second integral I_2 is treated similarly after transformation to the original variables h and v:

$$\begin{aligned}
I_2 &= \int_{\mathbb{R}} \frac{1}{2}q_x\frac{q^2}{h^2} - \left(\frac{q^2}{h}\right)_x\frac{q}{h}\,dx \\
&= \int_{\mathbb{R}} \frac{1}{2}(hv)_x v^2 - (hv^2)_x v\,dx \\
&= \int_{\mathbb{R}} -\frac{1}{2}h_x v^3 - \frac{3}{2}v_x h v^2\,dx \\
&= \int_{\mathbb{R}} \frac{3}{2}v_x h v^2 - \frac{3}{2}v_x h v^2\,dx - \frac{1}{2}h v^3\Big|_{-\infty}^{\infty} \\
&= 0.
\end{aligned}$$

Altogether, we conclude

$$\frac{d}{dt}\int_{\mathbb{R}} W(h,q)\,dx = 0.$$

□

2.2.2 Subsurface flow

Consider Richards' equation supplemented with no-flow boundary conditions

$$\begin{aligned}
\partial_t\theta(\psi) - \nabla\cdot[K(\theta(\psi))\nabla(\psi+z)] &= 0 \qquad && \text{on } \Omega_{\text{pm}}\times(0,T), \\
v_{\text{pm}}\cdot\vec{n} &= 0 && \text{on } \partial\Omega_{\text{pm}}\times[0,T),
\end{aligned} \tag{2.7}$$

for a time $T > 0$. In what follows, we assume that $\theta(\psi)$ and $K(\theta(\psi))$ are monotonically increasing, continuously differentiable and bounded.
It proves beneficial to convert Richards' equation (2.7) into a semi-linear partial differential equation, for which reason we perform the Kirchhoff transformation

$$\mathcal{K}:\mathbb{R}\to\mathbb{R},\quad \psi\mapsto u = \mathcal{K}(\psi) = \int_0^{\psi} K(\theta(\varphi))\,d\varphi.$$

Under the assumptions on K and θ, $\mathcal{K}$ is invertible and we can define

$$\begin{aligned}
b(u) &:= \theta(\mathcal{K}^{-1}(u)), \\
k(b(u)) &:= K\left(\theta(\mathcal{K}^{-1}(u))\right).
\end{aligned}$$

Applying the chain rule of differentiation, one finds

$$\nabla u = K(\theta(\psi))\nabla\psi,$$

and we arrive at the transformed Richards' equation for the unknown variable u

$$\partial_t b(u) - \nabla \cdot [\nabla u + k(b(u))\vec{e}_z] = 0, \tag{2.8}$$

where we defined $\vec{e}_z = \nabla z$. The boundary conditions take the form

$$v_{\text{pm}} \cdot \vec{n} = -\left(K(\theta(\psi))\nabla(\psi + z)\right) \cdot \vec{n} = (-\nabla u - k(b(u))\vec{e}_z) \cdot \vec{n} = 0 \qquad \text{on } \partial\Omega_{\text{pm}} \times [0, T). \tag{2.9}$$

In terms of the transformed variables we define a suitable energy for Richards' equation and prove an energy estimate.

Theorem 2.2.3 (Energy estimate for Richards' equation) Let $u \in C^2(\Omega_{\text{pm}} \times [0, T))$ be a classical solution of the transformed Richards' equation (2.8) endowed with no-flow boundary conditions (2.9). Assume that $u(\cdot, t)$ has compact support for all $t \in [0, T)$. Then, the following estimate for the energy

$$W(u) = \int_0^u b'(v)\mathcal{K}^{-1}(v)\, dv \tag{2.10}$$

holds:

$$\frac{d}{dt}\int_{\Omega_{\text{pm}}} W(u)\, d\vec{x} \leq -\int_{\partial\Omega_{\text{pm}}} u\vec{e}_z \cdot \vec{n}\, d\vec{\xi} =: \mathcal{G}(u). \tag{2.11}$$

The right hand side $\mathcal{G}(u)$ represents an energy source term due to gravitation.

Proof. We compute, using the divergence theorem

$$\begin{aligned}
\frac{d}{dt}\int_{\Omega_{\text{pm}}} W(u)\, d\vec{x} &= \int_{\Omega_{\text{pm}}} \partial_t W(u)\, d\vec{x} = \int_{\Omega_{\text{pm}}} W'(u)\partial_t u\, d\vec{x} = \int_{\Omega_{\text{pm}}} b'(u)\mathcal{K}^{-1}(u)\partial_t u\, d\vec{x} = \int_{\Omega_{\text{pm}}} \mathcal{K}^{-1}(u)\partial_t b(u)\, d\vec{x} \\
&= -\int_{\Omega_{\text{pm}}} \mathcal{K}^{-1}(u)\nabla \cdot v_{\text{pm}}\, d\vec{x} = \int_{\Omega_{\text{pm}}} \nabla\mathcal{K}^{-1}(u) \cdot v_{\text{pm}}\, d\vec{x} - \int_{\partial\Omega_{\text{pm}}} \mathcal{K}^{-1}(u) \underbrace{v_{\text{pm}} \cdot \vec{n}}_{=0}\, d\vec{\xi} \\
&= \int_{\Omega_{\text{pm}}} \frac{1}{K(\theta(\psi))} K(\theta(\psi))\nabla\psi \cdot v_{\text{pm}}\, d\vec{x} = \underbrace{-\int_{\Omega_{\text{pm}}} K(\theta(\psi))|\nabla\psi|^2\, d\vec{x}}_{\leq 0} - \int_{\Omega_{\text{pm}}} K(\theta(\psi))\nabla\psi \cdot \vec{e}_z\, d\vec{x} \\
&\leq \int_{\Omega_{\text{pm}}} -\nabla u \cdot \vec{e}_z\, d\vec{x} = \int_{\Omega_{\text{pm}}} u \underbrace{\nabla \cdot \vec{e}_z}_{=0}\, d\vec{x} - \int_{\partial\Omega_{\text{pm}}} u\vec{e}_z \cdot \vec{n}\, d\vec{x} = \mathcal{G}(u).
\end{aligned}$$

□

2.2.3 Coupled flow

We start with a simplified surface-subsurface model, given by the linear transport equation ($\gamma = 1$) or the kinematic wave equation ($\gamma = 5/3$) for the surface flow and by the Kirchhoff transformed Richards' equation in the subsurface

$$\left.\begin{aligned}
\partial_t h + \partial_x\left(Vh^\gamma\right) &= v_{\text{pm}} \cdot \vec{n}, && (x, t) \in \mathbb{R} \times (0, T), \\
h(x, 0) &= h_0(x), && x \in \mathbb{R}, \\
\partial_t b(u) + \nabla \cdot v_{\text{pm}} &= 0, && (\vec{x}, t) \in \Omega_{\text{pm}} \times (0, T), \\
v_{\text{pm}} &= -(\nabla u + k(b(u))\vec{e}_z), && (\vec{x}, t) \in \Omega_{\text{pm}} \times (0, T), \\
\mathcal{K}^{-1}(u(\vec{x}, 0)) &= \psi_0(\vec{x}), && \vec{x} \in \Omega_{\text{pm}}, \\
\mathcal{K}^{-1}(u) &= h, && (\vec{x}, t) \in \Gamma \times [0, T], \\
v_{\text{pm}} \cdot \vec{n} &= 0, && (\vec{x}, t) \in \partial\Omega_{\text{pm}} \setminus \Gamma \times [0, T].
\end{aligned}\right\} \tag{2.12}$$

We prove the following theorem:

Theorem 2.2.4 (Energy estimate for the coupled model) Let $u \in C^2(\Omega_{\text{pm}} \times [0,T))$, $h \in C^1(\mathbb{R} \times [0,T))$ be a classical solution of the coupled model (2.12). Assume that $u(\cdot,t)$ and $h(\cdot,t)$ have compact support for all $t \in [0,T)$. Then, the following estimate for the total energy

$$\mathcal{W}(h,u) = \int_{\mathbb{R}} \frac{h^2}{2}\, dx + \int_{\Omega_{\text{pm}}} \int_0^u b'(v)\mathcal{K}^{-1}(v)\, dv\, d\vec{x} \tag{2.13}$$

holds:

$$\frac{d}{dt}\mathcal{W}(h,u) \leq -\int_{\partial\Omega_{\text{pm}}} u\vec{e}_z \cdot \vec{n}\, d\vec{\xi} =: \mathcal{G}(u). \tag{2.14}$$

Proof. We mimic the proofs of Theorems 2.2.1 and 2.2.3 paying attention to the coupling.
Calculating the time derivative of the surface energy similarly to the uncoupled case, all terms cancel save a source term due to the flux from the subsurface and we obtain for both surface models

$$\frac{d}{dt}\int_{\mathbb{R}} \frac{h^2}{2}\, dx = \int_{\mathbb{R}} hh_t\, dx = \int_{\mathbb{R}} hv_{\text{pm}} \cdot \vec{n}\, dx = \int_{\Gamma} hv_{\text{pm}} \cdot \vec{n}\, d\vec{\xi}.$$

The subsurface model differs from the uncoupled one only by the coupling condition $(2.12)_6$ instead of the no-flow condition on Γ. Therefore, the integral over the boundary in the second line of the proof of Theorem 2.2.3 does not vanish completely, but we are left with

$$\frac{d}{dt}\int_{\Omega_{\text{pm}}}\int_0^u b'(v)\mathcal{K}^{-1}(v)\, dv\, d\vec{x} \leq -\int_{\partial\Omega_{\text{pm}}} u\vec{e}_z \cdot \vec{n}\, d\vec{\xi} - \int_{\Gamma} \mathcal{K}^{-1}(u)v_{\text{pm}} \cdot \vec{n}\, d\vec{\xi} = \mathcal{G}(u) - \int_{\Gamma} \psi v_{\text{pm}} \cdot \vec{n}\, d\vec{\xi}.$$

Now, we sum up the energies and exploit the boundary condition $\psi = h$ on Γ to get

$$\frac{d}{dt}\,\mathcal{W}(h,u) \leq \int_{\Gamma} hv_{\text{pm}} \cdot \vec{n}\, d\vec{\xi} + \mathcal{G}(u) - \int_{\Gamma} hv_{\text{pm}} \cdot \vec{n}\, d\vec{\xi} = \mathcal{G}(u),$$

which completes the proof. □

Remark 2.2.1 Pursuing the foregoing approach in order to achieve a similar energy estimate for the energy of the coupled model (1.41) involving the shallow water equations for the surface flow, given by

$$\mathcal{W}(h,u) = \int_{\mathbb{R}} \frac{h^2}{2} + \frac{hv^2}{2g}\, dx + \int_{\Omega_{\text{pm}}} \int_0^u b'(v)\mathcal{K}^{-1}(v)\, dv\, d\vec{x}, \tag{2.15}$$

leaves an additional integral over the boundary and one arrives at

$$\frac{d}{dt}\mathcal{W}(h,u) \leq -\int_{\partial\Omega_{\text{pm}}} u\vec{e}_z \cdot \vec{n}\, d\vec{\xi} + \int_{\Gamma} \left(\frac{-v^2}{2g}\, v_{\text{pm}} \cdot \vec{n} + v\frac{q_I}{\rho} \right) d\vec{\xi}. \tag{2.16}$$

Even though one could choose a friction term $\tau := q_I/\rho$ due to the coupling, eliminating the latter integral would require the choice of $\tau = (v/2g)\, v_{\text{pm}} \cdot \vec{n}$, whose physical justification as a friction term caused by the coupling is lacking.
In Section 4.3, the impact of this energy source term due to the coupling will be investigated numerically. ◁

Although we have shown the energy estimate only for simplified models, we will employ the shallow water equations as governing equations for the surface flow in our numerical simulations.

Chapter 3

Numerical methods

In order to obtain solutions of the coupled model (1.41), numerical methods have to be employed since the derivation of analytical solutions is not possible in general. Therefore, we consider numerical methods for (non-linear) hyperbolic systems which are to be applied to the shallow water equations (1.8). Then, numerical methods are used for Richards' equation (1.35). Finally, an algorithm implementing the coupled mathematical model is developed.

3.1 Surface flow

In this section, we provide an insight into the basics of numerical methods for the solution of hyperbolic systems of conservation laws, based upon the in-depth introduction by LeVeque [35].
Firstly, due to the non-linearity of system (1.1), smooth solutions cannot be expected for all times. This necessitates the extension of the concept of solutions and thus the definition of weak solutions of the system of conservation laws (1.1). We arrive at the weak formulation by multiplying system (1.1) with a test function $\phi \in C_0^1(\mathbb{R} \times [0,T), \mathbb{R}^m)$, in which $C_0^1(\mathbb{R} \times [0,T), \mathbb{R}^m)$ stands for the space of continuously differentiable functions with compact support mapping $\mathbb{R} \times [0,T)$ to $\mathbb{R}^m$. Then, integration by parts is applied to obtain

$$\int_0^T \int_{-\infty}^{\infty} [(\partial_t \phi) \cdot u + (\partial_x \phi) \cdot \Phi(u)] \, dx \, dt = - \int_{-\infty}^{\infty} \phi(x,0) \cdot u(x,0) \, dx - \int_0^T \int_{-\infty}^{\infty} \phi \cdot f \, dx \, dt. \tag{3.1}$$

The initial data go into the right hand side of the weak formulation (3.1). Note that the analogous integral evaluated at $t = T$ vanishes by virtue of the compact support of ϕ.

Definition 3.1.1 (Weak solution) The function $u : \mathbb{R} \times [0,T) \to \mathbb{R}^m$ is called a weak solution of the system of conservation laws (1.1) endowed with initial data $u(x,0) = u_0(x)$ if equation (3.1) holds for all $\phi \in C_0^1(\mathbb{R} \times [0,T), \mathbb{R}^m)$.

The prevalent discretization approach for hyperbolic systems is based on an explicit time discretization combined with a suitable finite difference / finite volume scheme in space. However, there exist other methods such as finite elements for hyperbolic equations, analysed e.g. in [24, 26].
When applying numerical schemes to non-linear equations with discontinuous solutions, several difficulties may occur, summarized in [35, chap. 12]:

- A method exhibiting stability in the linear case might compute oscillatory solutions for non-linear problems.
- A method might converge to a function which is not a weak solution of the system of conservation laws.
- A method might converge to a weak solution which is not the "physical" solution in terms of the so-called entropy condition being violated.

Fortunately, at least convergence towards a wrong solution can be prevented by claiming a particular form for the numerical method which is derived in what follows.

3.1.1 Conservative methods

As to the numerical methods, we consider a bounded surface flow domain $\Omega_{\text{ff}} = (\hat{x}_l, \hat{x}_r) \in \mathbb{R}$ and circumvent the difficulties coming with appropriate choices of Dirichlet or Neumann conditions for hyperbolic problems by imposing periodic boundary conditions, that is $u(\hat{x}_l, t) = u(\hat{x}_r, t)$ for all $t \in [0, T]$.
We furnish the flow domain Ω_{ff} with a uniform grid $\{x_j\}_{j=0}^{M}$ given by $x_j = \hat{x}_l + j\Delta x$ for $M \in \mathbb{N}$ and $\Delta x = (\hat{x}_r - \hat{x}_l)/M$ (Figure 3.1). The numerical solution of equation (1.8) at time $t^n = t^{n-1} + \Delta t$ at grid point x_j is denoted by U_j^n, where we define $t^0 = 0$. The time increment Δt might vary over the time to ensure a stability condition as will be explained later. To implement periodic boundary conditions, let us write $x_{-1} = x_{M-1}$ and $x_{M+1} = x_1$, so that no attention has to be turned on a special treatment at the boundaries.
Substituting the time derivative in (1.1) by a forward difference quotient and employing a central difference quotient for the spatial derivative evaluated at the cell centres $x_{j+1/2} := x_j + \Delta x/2$ for $j \in \{0, \dots, M-1\}$, $x_{-1/2} := x_{M-1/2}$ and $x_{M+1/2} := x_{1/2}$, yields for $j \in \{0, \dots, M\}$

$$\frac{1}{\Delta t}(U_j^{n+1} - U_j^n) + \frac{1}{\Delta x}[\Phi(U_{j+1/2}^n) - \Phi(U_{j-1/2}^n)] = f_j^n. \tag{3.2}$$

In the context of a finite volume scheme, U_j^n should be interpreted as an approximation to the average of u in cell $(x_{j-1/2}, x_{j+1/2})$, i.e.

$$U_j^n \approx \frac{1}{\Delta x} \int_{x-1/2}^{x+1/2} u(x, t^n)\, dx. \tag{3.3}$$

Solving equation (3.2) for the unknown value U_j^{n+1} leads to

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x}[\Phi(U_{j+1/2}^n) - \Phi(U_{j-1/2}^n)] + \Delta t f_j^n. \tag{3.4}$$

Approximating $\Phi(U_{j+1/2}^n)$ by a numerical flux function $F(U_j^n, U_{j+1}^n)$, the method takes the final form

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x}[F(U_j^n, U_{j+1}^n) - F(U_{j-1}^n, U_j^n)] + \Delta t f_j^n, \tag{3.5}$$

in case of the numerical flux function F only depending on two variables. The extension to more general numerical flux functions is plain and can be found in [35, chap. 12].
A numerical method of the form (3.5) is conservative, i.e. in case of $f \equiv 0$

$$\sum_{j=0}^{M} U_j^{n+1} = \sum_{j=0}^{M} U_j^n \tag{3.6}$$

holds, since

$$\sum_{j=0}^{M} \Big(F(U_j^n, U_{j+1}^n) - F(U_{j-1}^n, U_j^n) \Big) = 0. \tag{3.7}$$

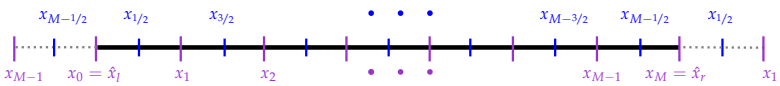


Figure 3.1: Uniform grid on Ω_{ff} for periodic boundary conditions

3.1.2 Convergence

For linear partial differential equations, a powerful outcome in the analysis of numerical methods is given by the Lax equivalence theorem stating that consistency and stability imply convergence of a linear method. This theorem does not carry over to the non-linear case, but still, a few helpful results concerning the convergence of numerical methods at least for scalar hyperbolic equations exist.
We proceed with the definition of fundamental attributes of numerical methods. For the time being, we assume

a fixed ratio $\Delta x/\Delta t = \text{const.}$, which will be motivated later and allows to drop the index Δx in the subsequent definitions.
Our emphasis is placed on attaining results for the actual error the numerical method produces as compared to the analytical solution of the conservation laws (1.1), which we assume to exist and to be unique here. This measure is referred to as the global error E_j^n, given by

$$E_j^n = U_j^n - u(x_j, t^n). \tag{3.8}$$

We intend to compare the numerical and the analytical solutions in integral norms, to be specific in the L^1-norm, since it is the natural norm when it comes to conservation laws. Hence, we define a piecewise constant function $U_{\Delta t}(x,t)$ by setting

$$U_{\Delta t}(x,t) = U_j^n, \qquad (x,t) \in [x_{j-1/2}, x_{j+1/2}) \times [t^n, t^{n+1}). \tag{3.9}$$

Then, the error function $E_{\Delta t}(x,t)$, defined as

$$E_{\Delta t}(x,t) = U_{\Delta t}(x,t) - u(x,t), \tag{3.10}$$

may be employed for the definition of convergence in integral norms.

Definition 3.1.2 (Convergence) A method is called convergent with respect to $\|\cdot\|_1$ if

$$\|E_{\Delta t}(\cdot,t)\|_1 \to 0 \qquad \text{for } \Delta t \to 0, \tag{3.11}$$

for all $t \in [0,T)$.

It turns out that in many cases, it is useful to split the global error into a local truncation error and a cumulative error. This dodge leads to the fundamental terms of consistency and stability on which we focus in what follows.
We require the definition of a finite difference operator $\mathcal{H}_{t,\Delta t}$, operating on functions of a spatial variable at a fixed time t.

Definition 3.1.3 (Finite difference operator) The finite difference operator $\mathcal{H}_{t,\Delta t}$ of a conservative method at time t for a time step Δt is defined as

$$\mathcal{H}_{t,\Delta t}[v](x) = v(x) - \frac{\Delta t}{\Delta x}\left[F(v(x), v(x+\Delta x)) - F(v(x-\Delta x), v(x))\right] + \Delta t f(x,t). \tag{3.12}$$

We will also employ the discrete finite difference operator $\tilde{\mathcal{H}}_{t,\Delta t}$ which maps discrete functions merely defined on the grid $\{x_j\}_{j=0}^M$ analogously, that is

$$\tilde{\mathcal{H}}_{t,\Delta t}[V](x_j) = V_j - \frac{\Delta t}{\Delta x}\left[F(V_j, V_{j+1}) - F(V_{j-1}, V_j))\right] + \Delta t f(x_j,t). \tag{3.13}$$

Next, we introduce the local error occurring when we execute the numerical method employing the pointwise evaluation of the analytical solution of system (1.1).

Definition 3.1.4 (Local truncation error) The local truncation error of a conservative method (3.5) is given by

$$L_{\Delta t}(x,t) = \frac{1}{\Delta t}\left[u(x,t+\Delta t) - u(x,t) + \frac{\Delta t}{\Delta x}\left[F(u(x,t), u(x+\Delta x,t)) - F(u(x-\Delta x,t), u(x,t))\right] - \Delta t f(x,t)\right]. \tag{3.14}$$

Exploiting definitions (3.12) and (3.14), we make the splitting of the error E concise by asserting that

$$u(x,t+\Delta t) = \mathcal{H}_{t,\Delta t}[u(\cdot,t)](x) + \Delta t L_{\Delta t}(x,t), \tag{3.15}$$

and hence

$$\begin{aligned} E_{\Delta t}(x,t+\Delta t) &= U_{\Delta t}(x,t+\Delta t) - u(x,t+\Delta t) \\ &= \underbrace{\left(\mathcal{H}_{t,\Delta t}[U_{\Delta t}(\cdot,t)](x) - \mathcal{H}_{t,\Delta t}[u(\cdot,t)](x)\right)}_{\text{cumulative error}} - \underbrace{\Delta t L_{\Delta t}(x,t)}_{\text{local error}}. \end{aligned} \tag{3.16}$$

The local truncation error provides a measure for the error that is locally generated by utilizing the very numerical method, whereas the cumulative part accounts for the error made in the previous time steps. It is worth mentioning that in the linear case the finite difference operator $\mathcal{H}_{t,\Delta t}$ is linear, too, which gives rise to a much simpler analysis of the cumulative error, heavily exploited in the proof of the Lax equivalence theorem.

3.1.3 Consistency

As $\Delta t \to 0$, convergence of the numerical solution to the analytical one can only happen if the local truncation error converges to 0. This requirement is referred to as consistency.

Definition 3.1.5 (Consistency) The conservative method (3.5) is consistent, if

$$\|L_{\Delta t}(\cdot, t)\|_1 \to 0, \tag{3.17}$$

for $\Delta t \to 0$.

A more detailed distinction in terms of the local order of convergence is often made. To obtain criteria for the numerical flux function F ensuring the consistency of the numerical scheme (3.5), we consider the case of constant flux, i.e. $u(x,t) \equiv u^*$. We then expect the numerical solution to be constant as well. This requires that

$$F(u^*, u^*) = \Phi(u^*), \tag{3.18}$$

for all $u^* \in \mathbb{R}$. Furthermore, a requirement of smoothness has to be imposed on F. Since proofs of consistency rely on the expansion in Taylor series, the following result for the consistency of conservative methods holds only in case that the numerical flux F is continuously differentiable.
We prove

Theorem 3.1.1 (Consistency in the smooth case) Let the numerical flux function F in method (3.5) be continuously differentiable and let equation (3.18) be satisfied for all $u^* \in \mathbb{R}$. Then, method (3.5) is consistent.

Proof. As consistency is a property of the numerical method employed and not of the solution, we can assume the solution u to have compact support in $[\hat{x}_l, \hat{x}_r] \subset \mathbb{R}$ and to be sufficiently smooth for our purposes.
We calculate, employing Taylor's formula

$$\begin{aligned}
|L_{\Delta t}(x,t)| &= \left| \frac{1}{\Delta t} \left[u(x,t+\Delta t) - u(x,t) + \frac{\Delta t}{\Delta x} \left[F(u(x,t), u(x+\Delta x,t)) - F(u(x-\Delta x,t), u(x,t)) \right] - \Delta t f(x,t) \right] \right| \\
&\leq \left| \frac{u(x,t+\Delta t) - u(x,t)}{\Delta t} - \partial_t u(x,t) \right| + |\partial_t u(x,t) + \partial_x \Phi(u(x,t)) - \Delta t f(x,t)| \\
&\quad + \left| \frac{F(u(x,t), u(x+\Delta x,t)) - F(u(x-\Delta x,t), u(x,t))}{\Delta x} - \partial_x \Phi(u(x,t)) \right| \\
&\leq \frac{\Delta t}{2} \max_{\tau \in [t, \Delta t]} \{ |\partial_{tt}^2 u(x,\tau)| \} + 0 + \left| \frac{F(u(x,t), u(x+\Delta x,t)) - F(u(x,t), u(x,t))}{\Delta x} \right. \\
&\quad + \frac{F(u(x,t), u(x,t)) - F(u(x-\Delta x,t), u(x,t))}{\Delta x} \\
&\quad \left. - \left[\partial_{u_1} F(u(x,t), u(x,t))\, \partial_x u(x,t) + \partial_{u_2} F(u(x,t), u(x,t))\, \partial_x u(x,t) \right] \right| \\
&\leq \frac{\Delta t}{2} \max_{\tau \in [t, \Delta t]} \{ |\partial_{tt}^2 u(x,\tau)| \} + \left| \frac{F(u(x,t), u(x,t)) - F(u(x-\Delta x,t), u(x,t))}{\Delta x} - \partial_{u_1} F(u(x,t), u(x,t))\, \partial_x u(x,t) \right| \\
&\quad + \left| \frac{F(u(x,t), u(x+\Delta x,t)) - F(u(x,t), u(x,t))}{\Delta x} - \partial_{u_2} F(u(x,t), u(x,t))\, \partial_x u(x,t) \right| \\
&\leq \frac{\Delta t}{2} \max_{\tau \in [t, \Delta t]} \{ |\partial_{tt}^2 u(x,\tau)| \} + \Delta x \max_{\xi \in [x-\Delta x, x+\Delta x]} \{ |\partial_{xx}^2 \Phi(u(\xi,t))| \}.
\end{aligned}$$

The hindmost inequality stems from the continuous differentiability of F rendering the differential quotients of F with respect to u well-defined in a neighbourhood of $u(x,t)$. We recall $\Delta x / \Delta t = \text{const.}$, from which follows

that for a constant $c > 0$

$$|L_{\Delta t}(x,t)| \leq c\Delta t \left(\max_{\tau\in[t,\Delta t]} \{|\partial_{tt}^2 u(x,\tau)|\} + \max_{\xi\in[x-\Delta x,x+\Delta x]} \{|\partial_{xx}^2 \Phi(u(\xi,t)|\} \right)$$

and thus, provided that the support of the solution is contained in a compact interval $[\hat{x}_l, \hat{x}_r]$,

$$\begin{aligned}
\|L_{\Delta t}(\cdot,t)\|_1 &= \int_{\hat{x}_l}^{\hat{x}_r} |L_{\Delta t}(x,t)|\, dx \\
&\leq c\Delta t \left(\max_{\substack{\tau\in[t,\Delta t]\\ \xi\in[\hat{x}_l,\hat{x}_r]}} \{|\partial_{tt}^2 u(\xi,\tau)|\} + \max_{\substack{\tau\in[t,\Delta t]\\ \xi\in[\hat{x}_l,\hat{x}_r]}} \{|\partial_{xx}^2 u(\xi,\tau)|\} \right) (\hat{x}_r - \hat{x}_l) \\
&= \mathcal{O}(\Delta t).
\end{aligned}$$

□

Note that this result strongly requires the continuous differentiability of the numerical flux F and sufficient smoothness of the solution u and thus does not carry over to settings only admitting weak solutions that lack of regularity, e.g. in case of discontinuous initial data or in the non-linear case as in the latter one smooth solutions may gradually develop discontinuities. However, condition (3.18) is of vital importance for convergence results also then.

3.1.4 The Lax–Wendroff theorem

We focus again on the numerical issues arising when solving non-linear hyperbolic systems. It turns out that the choice of a numerical scheme in conservative form (3.5) prevents the undesirable behaviour of convergence to a non-solution, which the following theorem by Lax and Wendroff [33] here presented in the version by Kröner [31, chap. 2] makes more precise.
Note that the Lax–Wendroff theorem does not make any statement if a method converges at all. This is because no stability condition is presumed and even more due to the fact that multiple weak solutions may exist in general, for which reason it is no wonder that sequences of $\Delta t, \Delta x$ exist leading to sequences of solutions which jump among approximations to different solutions. It suffices to assume Lipschitz continuity of the numerical flux here, together with condition (3.18).

Theorem 3.1.2 (Lax–Wendroff) Let $(\Delta x)_i$ and $(\Delta t)_i$ be sequences of grid parameters satisfying $(\Delta x)_i, (\Delta t)_i \to 0$ as $i \to 0$, where $(\Delta x)_i = c(\Delta t)_i$ for a fixed c. Let (U_i) denote the sequence of numerical solutions of system (1.1) computed with a conservative method (3.5) on the i-th grid. Assume that Φ is Lipschitz continuous and equation (3.18) holds. Let U_i^0 be given as

$$U_i^0 := \frac{1}{(\Delta x)_i} \int_{x_{i-1/2}}^{x_{i+1/2}} u_0(x)\, dx, \tag{3.19}$$

and let $U_i(x,t) := U_{(\Delta t)_i}(x,t)$ be the piecewise constant function as defined in equation (3.9).
Assume further that

1. There exists a constant $K \geq 0$ such that

$$\sup_i \sup_{\mathbb{R}\times[0,T)} |U_i(x,t)| \leq K. \tag{3.20}$$

2.

$$U_i \longrightarrow u \quad \text{a.e. in } \mathbb{R}\times[0,T) \text{ for } i \to \infty. \tag{3.21}$$

Then, u is a weak solution of system (1.1) in the weak sense (3.1).

Proof. Kröner [31, Theorem 2.3.1]. □

3.1.5 Stability

Even though the local error of a consistent method vanishes as $\Delta t \to 0$, this criterion is not sufficient to ensure convergence of the numerical solution towards the analytical solution as equation (3.16) shows.
We proceed with the definition of a stability criterion referred to as Lax–Richtmyer stability which is a necessary assumption in the Lax equivalence theorem for the linear case and adherence to which gives a limit to the choice of the time step Δt depending on the spatial grid with Δx.

Definition 3.1.6 (Lax-Richtmyer stability) A method is stable if for each time T there exist constants $C > 0$ and $\tau > 0$ such that

$$\|\tilde{\mathcal{H}}^n_{t,\Delta t}\|_1 \leq C \qquad \text{for all} \quad n\Delta t \leq T, \Delta t \leq \tau. \tag{3.22}$$

The superscript n denotes the n-fold application of the operator $\tilde{\mathcal{H}}_{t,\Delta t}$.

Using the foregoing definitions, we can put the Lax equivalence theorem to record:

Theorem 3.1.3 (Lax equivalence theorem) Let the initial value problem consisting of the conservation laws (1.1) supplemented with initial conditions at $t = 0$ be linear and well-posed. Then for a consistent method, stability is necessary and sufficient for the convergence of the method.

Proof. Strikwerda [56, Theorem 10.5.1]. □

For the analysis of convergence in the non-linear case, more general concepts of stability related to the total variation and to contraction properties are required. Moreover, results concerning the convergence towards the physically correct solution satisfying an entropy condition may demand further assumptions on the methods such as monotonicity. A survey on stability concepts for non-linear problems can be found in [35, chap. 15].

3.1.6 The Lax–Friedrichs method and the local Lax–Friedrichs method

We now shift focus to some examples of conservative methods to be applied for the discretization of the surface model in the coupled system (1.41). It turns out that the numerical scheme resulting from the intuitive choice $F(U_j, U_{j+1}) = \frac{1}{2}(\Phi(U_{j+1}) + \Phi(U_j))$ in the conservative scheme (3.5) behaves unstably and is therefore futile in practice. However, this approach can be mended by replacing U^n_j by the centred difference $\frac{1}{2}(U^n_{j-1} + U^n_{j+1})$ yielding the Lax–Friedrichs (LxF) method

$$U^{n+1}_j = \frac{1}{2}(U^n_{j-1} + U^n_{j+1}) - \frac{\Delta t}{2\Delta x}[\Phi(U^n_{j+1}) - \Phi(U^n_{j-1})] + \Delta t f^n_j. \tag{3.23}$$

It is written as a conservative scheme defining the numerical flux

$$F(U_j, U_{j+1}) = \frac{1}{2}(\Phi(U_{j+1}) + \Phi(U_j)) + \frac{\Delta x}{2\Delta t}\left(U_j - U_{j+1}\right), \tag{3.24}$$

which obviously satisfies condition (3.18).

Example 3.1.1 (Convergence of the Lax–Friedrichs method for the linear transport equation) Consider the application of the LxF method on the linear transport equation (1.17). The numerical flux is then given by

$$F(U_j, U_{j+1}) = \frac{1}{2}(\Phi(U_{j+1}) + \Phi(U_j)) + \frac{\Delta x}{2\Delta t}\left(U_j - U_{j+1}\right) = \left(\frac{V}{2} + \frac{\Delta x}{2\Delta t}\right) U_j + \left(\frac{V}{2} - \frac{\Delta x}{2\Delta t}\right) U_{j+1},$$

which is apparently continuously differentiable and hence consistency of the method is ensured by Theorem (3.1.1). To ascertain the Lax–Richtmyer stability, we assert that it suffices to show that $\|U^{n+1}\|_1 \leq \|U^n\|_1$ because then when have for the operator norm $\|\mathcal{H}^{n+1}_{t,\Delta t}\|_1 \leq \|\mathcal{H}_{t,\Delta t}\|^{n+1}_1 \leq 1$.
We calculate

$$\begin{aligned}\|U^{n+1}\|_1 &= \Delta x \sum_{j=0}^{M} |U^{n+1}_j| \\ &\leq \frac{\Delta x}{2}\left(\sum_{j=0}^{M}\left|\left(1 - \frac{V\Delta t}{\Delta x}\right) U^n_{j+1}\right| + \left|\left(1 + \frac{V\Delta t}{\Delta x}\right) U^n_{j-1}\right|\right).\end{aligned}$$

If we select $\Delta t \leq \Delta x/|V|$, the prefactors of U_{j-1}^n and U_{j+1}^n are non-negative and we infer that

$$\begin{aligned}\|U^{n+1}\|_1 &\leq \frac{\Delta x}{2}\left(\sum_{j=0}^{M}\left(1-\frac{V\Delta t}{\Delta x}\right)\left|U_{j+1}^n\right|+\left(1+\frac{V\Delta t}{\Delta x}\right)\left|U_{j-1}^n\right|\right)\\ &= \|U^n\|_1.\end{aligned}$$

Thus, the Lax equivalence theorem 3.1.3 ensures the convergence of the LxF method under the assumption that the time step satisfies $\Delta t \leq \Delta x/|V|$. ◁

When explicit time-marching schemes are employed for the solution of hyperbolic conservation laws, it is typical that stability of the method requires the time step to be sufficiently small depending on the spatial grid size. This phenomenon can be explained considering the bounded domains of dependence in the x-t space due to the finite propagation speed of solutions of hyperbolic equations. The resulting condition is named Courant–Friedrichs–Lewy condition (CFL-condition), first described in [10], and writes for explicit schemes in general

$$\frac{\lambda^{\max}\Delta t}{\Delta x} < 1, \tag{3.25}$$

in which $\lambda^{\max}$ denotes the maximum magnitude of the characteristic speed of the problem, given by the maximum absolute value among the eigenvalues of the Jacobian of the flux vector.
Choosing $(\lambda^{\max}\Delta t)/\Delta x = 1/2$ in order to fulfil the CFL-condition (3.25), one can rewrite the numerical flux of the LxF method as

$$F(U_j, U_{j+1}) = \frac{1}{2}(\Phi(U_{j+1}) + \Phi(U_j)) + \tilde{\lambda}^{\max}\left(U_j - U_{j+1}\right), \tag{3.26}$$

where $\tilde{\lambda}^{\max}$ denotes the maximum eigenvalue of the Jacobian of the flux vector Φ taken over $U_1, \ldots, U_M$.
However, the term $\lambda^{\max}\left(U_j - U_{j+1}\right)$ in the numerical flux of the LxF method inducing the stability of the method entails a numerical viscosity which causes smeared numerical solutions in vicinity to discontinuities. To alleviate this spurious behaviour, one could employ the local Lax–Friedrichs (LLxF) method instead, arising from the choice of the numerical flux function

$$F(U_j, U_{j+1}) = \frac{1}{2}(\Phi(U_{j+1}) + \Phi(U_j)) + \tilde{\lambda}^{\max}_{j+1/2}\left(U_j - U_{j+1}\right), \tag{3.27}$$

in which the subscript $j + 1/2$ indicates that the maximum is now only taken over U_j, U_{j+1}.

3.1.7 Godunov's method and Roe's approximate Riemann solver

The LxF method employs symmetric evaluations of the flux functions, regardless of the propagation of the solution. It turns out that making use of the knowledge about the domain of dependence significantly enhances the quality of the numerical approximation.
A widely used conservative method exploiting this information is Godunov's method, first proposed in [17]. It stems from the idea of solving a Riemann problem at each time level, this is an initial value problem with piecewise constant initial data exhibiting one discontinuity which occurs at $x_{j+1/2}$ for $j \in \{0, \ldots, M-1\}$ in Godunov's method, as the initial data between $x_{j-1/2}$ and $x_{j+1/2}$ is taken to be U_j^n at time level t^n. This approach yields a numerical flux

$$F(U_j^n, U_{j+1}^n) = \Phi(u^*(U_j^n, U_{j+1}^n)), \tag{3.28}$$

in which u^* denotes the solution of the Riemann problem at $x_{j+1/2}$ which can be shown to be constant over (t^n, t^{n+1}), provided that the CFL-condition is fulfilled. To see this, consider also the numerical example in Chapter 4 for the shallow water equations where one finds an expression for the constant water height at the initial shock location. The CFL-condition ensures that waves from neighbouring Riemann problems do not affect the cell.
The Riemann problems, although analytically solvable in theory, may be approximated as analytical solving is expensive. An attractive solver has been proposed by Roe [49]. It relies on the solution of the Riemann problem for a linearized equation on $(x_j, x_{j+1}) \times (t^n, t^{n+1})$ which writes

$$\hat{u}_t + \hat{A}(\hat{u}_j, \hat{u}_{j+1})\hat{u}_x = 0, \tag{3.29}$$

where we assume that no sources or sinks are present in the system for the sake of convenience. The Roe matrix $\hat{A}(\hat{u}_j, \hat{u}_{j+1})$ is subject to the following assumptions whose necessity is discussed in [35, sec. 14.2]:

1. Correct behaviour across discontinuities: $\hat{A}(\hat{u}_j, \hat{u}_{j+1})(\hat{u}_{j+1} - \hat{u}_j) = \Phi(\hat{u}_{j+1}) - \Phi(\hat{u}_j)$
2. Hyperbolicity: $\hat{A}(\hat{u}_j, \hat{u}_{j+1})$ is diagonizable with real eigenvalues $\hat{\lambda}_1, \dots, \hat{\lambda}_m$
3. Consistency with Jacobian: $\hat{A}(\hat{u}_j, \hat{u}_{j+1}) \to D\Phi(\bar{u})$ smoothly for $\hat{u}_j, \hat{u}_{j+1} \to \bar{u}$

Then, a calculation reveals that the numerical flux for Roe's solver can be written as

$$F(U_j^n, U_{j+1}^n) = \Phi(U_j^n) + \sum_{p=1}^{m} \hat{\lambda}_p^- \alpha_p \hat{r}_p, \tag{3.30}$$

where $\hat{\lambda}_p^- = \min\{\hat{\lambda}_p, 0\}$, $\hat{r}_p$ defined as the p-th eigenvector of $\hat{A}$ and α_p the corresponding coefficient in a eigenvector expansion of the difference $U_{j+1}^n - U_j^n$.
It is worth noting that the computed solution of Roe's solver might not fulfil the CFL-condition, but there exist modifications overcoming this defect as described e.g. in [19]. Besides, the idea of replacing the piecewise constant initial states as the input for the Riemann problems by more complex approximations gives rise to sophisticated methods providing highly accurate solutions such as MUSCL-schemes [60].

3.1.8 Numerical methods for the shallow water equations

We recall that for the shallow water equations the vector u contains $u = (u_1, u_2)^T = (h, hv)^T$ and the flux Φ is given by equation (1.9). Adjusting the time step requires the computation of the eigenvalues of the Jacobian $D\Phi$ which depend on the solution in the non-linear case. In case of the shallow water equations, one easily finds that

$$\lambda_{1/2} = v \pm \sqrt{gh}. \tag{3.31}$$

For the application of the (local) LxF method, the numerical flux function (3.24) or (3.27) is plugged into equation (3.5) and U_j^{n+1} can be computed successively.
In order to implement Roe's approximate Riemann solver, a Roe matrix $\hat{A}$ has to be determined. In terms of h and v, it is found as

$$\hat{A}(h_j, v_j, h_{j+1}, v_{j+1}) = \begin{pmatrix} 0 & 1 \\ -\hat{v}^2 + g\hat{h} & 2\hat{v} \end{pmatrix}, \tag{3.32}$$

where $\hat{h} = (h_j + h_{j+1})/2$ and $\hat{v} = \left(\sqrt{h_j} v_j + \sqrt{h_{j+1}} v_{j+1}\right) \Big/ \left(\sqrt{h_j} + \sqrt{h_{j+1}}\right)$.
This yields the eigenvalues and eigenvectors

$$\begin{aligned} \hat{\lambda}_{1/2}(h_j, v_j, h_{j+1}, v_{j+1}) &= \hat{v} \pm \sqrt{g\hat{h}}, \\ \hat{r}_{1/2}(h_j, v_j, h_{j+1}, v_{j+1}) &= (1, \hat{v} \pm \sqrt{g\hat{h}})^T. \end{aligned} \tag{3.33}$$

It follows for the coefficients $\alpha_{1/2}$ that

$$\alpha_{1/2}(h_j, v_j, h_{j+1}, v_{j+1}) = \frac{1}{2\sqrt{g\hat{h}}} \left[\pm \left((hv)_{j+1} - (hv)_j \right) \mp (h_{j+1} - h_j)(\hat{v} \mp \sqrt{g\hat{h}}) \right]. \tag{3.34}$$

so that all members of the Roe flux (3.30) in case of the shallow water equations are identified.

3.2 Subsurface flow

This section is devoted to the numerical treatment of Richards' equation. For the sake of convenience, we will write $K(\psi) = K(\theta(\psi))$ in this section. We impose no-flow Neumann boundary conditions for the time being, which will be replaced at the interface Γ by the coupling condition (1.40) in Section 3.3.

Thus, the model investigated in this section has the form

$$\left.\begin{aligned} \partial_t \theta(\psi) - \nabla \cdot [K(\psi)\nabla(\psi + z)] &= 0, \quad (\vec{x}, t) \in \Omega_{\text{pm}} \times (0, T), \\ -[K(\psi)\nabla(\psi + z)] \cdot \vec{n} &= 0, \quad (\vec{x}, t) \in \partial\Omega_{\text{pm}} \times [0, T), \end{aligned}\right\} \tag{3.35}$$

endowed with initial conditions $\psi(\vec{x}, 0) = \psi_0(\vec{x})$ for $\vec{x} \in \Omega_{\text{pm}}$. Problem (3.35) is degenerate inasmuch as it becomes elliptic wherever the flow domain becomes saturated, that is $\partial_t \theta(\psi) = 0$. It hence turns out that also for the subsurface model it is necessary to consider a weak formulation of the problem.
Let $H^1(\Omega_{\text{pm}})$ denote the standard Sobolev space of weakly differentiable functions on $L^2(\Omega_{\text{pm}})$ with derivative in $L^2(\Omega_{\text{pm}})$. In order to derive a weak formulation, we multiply Richards' equation $(3.35)_1$ by an arbitrary test function $\phi \in H^1(\Omega_{\text{pm}})$ and integrate over Ω_{pm} which yields for every $t \in (0, T)$

$$\int_{\Omega_{\text{pm}}} \partial_t \theta(\psi)\phi \, d\vec{x} - \int_{\Omega_{\text{pm}}} \nabla \cdot [K(\psi)\nabla(\psi + z)]\, \phi \, d\vec{x} = \int_{\Omega_{\text{pm}}} f\phi \, d\vec{x}. \tag{3.36}$$

Integration by parts applied to the second integral in (3.36) results in

$$\begin{aligned} &\int_{\Omega_{\text{pm}}} \partial_t \theta(\psi)\phi \, d\vec{x} + \int_{\Omega_{\text{pm}}} K(\psi)\nabla(\psi + z) \cdot \nabla\phi \, d\vec{x} - \int_{\partial\Omega_{\text{pm}}} \underbrace{K(\psi)\nabla(\psi + z) \cdot \vec{n}}_{=0} \, \phi \, d\vec{x} \\ &= \int_{\Omega_{\text{pm}}} f(\vec{x}, t)\phi \, d\vec{x}, \end{aligned} \tag{3.37}$$

which can be reformulated in a more compact way as

$$\langle \partial_t \theta(\psi), \phi \rangle + \langle K(\psi)\nabla(\psi + z), \nabla\phi \rangle = \langle f, \phi \rangle, \tag{3.38}$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product of scalar- or vector-valued functions on $L^2(\Omega_{\text{pm}})$.

Definition 3.2.1 (Weak solution) We call $\psi \in H^1(\Omega_{\text{pm}})$ a weak solution of problem (3.35) if equation (3.38) holds for all $\phi \in H^1(\Omega_{\text{pm}})$.

Results concerning the existence and uniqueness of weak solution of the Richards' equation can be found e.g. in [1]. The Galerkin method that we use for the discretization in space relies heavily on the concept of weak solutions. Prior to that, we turn towards the discretization in time and the linearization.

3.2.1 Discretization in time

We approximate the time derivative by the implicit Euler method featuring high stability as compared to explicit methods. Given the solution $\psi^n(\vec{x}) := \psi(\vec{x}, t^n)$ at a time level $t^n \in [0, T)$ for $n \in \mathbb{N}_0$, we select $\Delta t > 0$ respecting $t^{n+1} := t^n + \Delta t \leq T$ and compute the solution $\psi^{n+1}(\vec{x})$ at the next time level t^{n+1} substituting the time derivative $\partial_t \theta(\psi(\vec{x}, t))|_{t=t^{n+1}}$ in equation (3.38) by backward differences. Similarly, we define $f^n(\vec{x}) = f(\vec{x}, t^n)$. After multiplying with Δt we obtain the semi-discrete form of equation (3.38)

$$\left\langle \theta(\psi^{n+1}) - \theta(\psi^n), \phi \right\rangle + \Delta t \left\langle K(\psi^{n+1})\nabla(\psi^{n+1} + z), \nabla\phi \right\rangle = \Delta t \left\langle f^{n+1}, \phi \right\rangle. \tag{3.39}$$

3.2.2 Linearization

To linearize Richards' equation iterative methods are employed almost exclusively. To get rid of the nonlinearity caused by $K(\psi)$, one could employ methods based upon the Kirchhoff transformed Richards' equation as considered e.g. in [6]. We juxtapose several iterative methods in what follows. The superscript $j \in \mathbb{N}_0$ denotes the iteration step.
A classical approach is given by Newton's method, used for the solution of the Richards' equation e.g. in [34, 46]. It stands out due to quadratic convergence, however, convergence only takes place if the initial guess is close enough to the solution. Application of Newton's method yields:

Find $\psi^{n+1,j+1} \in H^1(\Omega_{\text{pm}})$, so that

$$\begin{aligned}
&\left\langle \theta(\psi^{n+1,j}), \phi \right\rangle + \left\langle \theta'(\psi^{n+1,j})(\psi^{n+1,j+1} - \psi^{n+1,j}), \phi \right\rangle - \langle \theta(\psi^n), \phi \rangle \\
&+ \Delta t \left\langle K(\psi^{n+1,j}) \left(\nabla \psi^{n+1,j+1} + \nabla z \right), \nabla \phi \right\rangle + \Delta t \left\langle K'(\psi^{n+1,j}) \left(\nabla \psi^{n+1,j} + \nabla z \right) (\psi^{n+1,j+1} - \psi^{n+1,j}), \nabla \phi \right\rangle \\
&= \Delta t \left\langle f^{n+1}, \phi \right\rangle,
\end{aligned} \tag{3.40}$$

for all $\phi \in H^1(\Omega_{\text{pm}})$. It is derived by expanding the non-linearities K and θ in truncated Taylor series. If this approach is only applied on θ, one arrives at a modified Picard scheme, analysed in [8, 34]. It reads:
Find $\psi^{n+1,j+1} \in H^1(\Omega_{\text{pm}})$, so that

$$\begin{aligned}
&\left\langle \theta(\psi^{n+1,j}), \phi \right\rangle + \left\langle \theta'(\psi^{n+1,j})(\psi^{n+1,j+1} - \psi^{n+1,j}), \phi \right\rangle - \langle \theta(\psi^n), \phi \rangle + \Delta t \left\langle K(\psi^{n+1,j}) \left(\nabla \psi^{n+1,j+1} + \nabla z \right), \nabla \phi \right\rangle \\
&= \Delta t \left\langle f^{n+1}, \phi \right\rangle,
\end{aligned} \tag{3.41}$$

for all $\phi \in H^1(\Omega_{\text{pm}})$. The simpler implementation of the modified Picard scheme as compared to Newton's method comes at the cost of the rate of convergence, which is only linear.
Another scheme, referred to as L-scheme hereafter, has been presented e.g. in [42, 43, 44, 53]. It ensures linear convergence for $L_\psi \geq \sup_\psi \frac{d}{d\psi}\theta(\psi)$ and is given by:
Find $\psi^{n+1,j+1} \in H^1(\Omega_{\text{pm}})$, so that

$$\begin{aligned}
&\left\langle \theta(\psi^{n+1,j}), \phi \right\rangle + L_\psi \left\langle \psi^{n+1,j+1} - \psi^{n+1,j}, \phi \right\rangle - \langle \theta(\psi^n), \phi \rangle + \Delta t \left\langle K(\psi^{n+1,j}) \left(\nabla \psi^{n+1,j+1} + \nabla z \right), \nabla \phi \right\rangle \\
&= \Delta t \left\langle f^{n+1}, \phi \right\rangle,
\end{aligned} \tag{3.42}$$

for all $\phi \in H^1(\Omega_{\text{pm}})$. The L-scheme does not require the evaluation of derivatives.
Besides, there exist schemes arising from the pressure-based form of Richards' equation which are not considered in this thesis. A combination of several schemes has been found effective in [34] and will be applied in the numerical example in Section 4.4.

3.2.3 Discretization in space

For the spatial discretization of the linearized equations (3.40), (3.41) respectively (3.42), we employ the finite element method (FEM). To this end, we restrict the solution space $H^1(\Omega_{\text{pm}})$ to a finite-dimensional subspace $X_h(\Omega_{\text{pm}}) \subset H^1(\Omega_{\text{pm}})$ to obtain the Galerkin formulation. The choice of the subspace and in particular of its basis leads us to the FEM.
Other methods widely used for the spatial discretization of the Richards' equation are the finite volume method (FVM), studied e.g. in [13], and the mixed finite element method (MFEM), presented e.g. in [42]. As to the linearization, we only pursue Newton's method henceforth since both the modified Picard scheme and the L-scheme can actually be considered as simplifications of Newton's method.

The Galerkin formulation

Let $X_h \subset H^1(\Omega_{\text{pm}})$ be a k-dimensional subspace with basis $\{\phi_1, \dots, \phi_k\}$ that shall be fixed in the following. We seek the Galerkin solution $\psi_h^{n+1,j+1} \in X_h$ which can be written as

$$\psi_h^{n+1,j+1}(\vec{x}) = \sum_{i=1}^{k} d_i^{n+1,j+1} \phi_i(\vec{x}), \tag{3.43}$$

with some coefficients $d_i^{n+1} \in \mathbb{R}$. Due to the bilinearity of the scalar product, equation (3.40) holds for all $\phi \in X_h$ if and only if it holds for all basis functions $\{\phi_1, \dots, \phi_k\}$ of X_h, that is

$$\begin{aligned}
&\left\langle \theta(\psi^{n+1,j}), \phi_s \right\rangle + \left\langle \theta'(\psi_h^{n+1,j})(\psi_h^{n+1,j+1} - \psi_h^{n+1,j}), \phi_s \right\rangle - \langle \theta(\psi_h^n), \phi_s \rangle \\
&+ \Delta t \left\langle K(\psi_h^{n+1,j}) \left(\nabla \psi_h^{n+1,j+1} + \nabla z \right), \nabla \phi_s \right\rangle + \Delta t \left\langle K'(\psi_h^{n+1,j}) \left(\nabla \psi_h^{n+1,j} + \nabla z \right) (\psi_h^{n+1,j+1} - \psi_h^{n+1,j}), \nabla \phi_s \right\rangle \\
&= \Delta t \left\langle f^{n+1}, \phi_s \right\rangle,
\end{aligned} \tag{3.44}$$

for all $s \in \{1, \dots, k\}$. Plugging representation (3.43) into equation (3.44) yields for $s \in \{1, \dots, k\}$

$$\begin{aligned}
&\left\langle \theta(\psi_h^{n+1,j}), \phi_s \right\rangle + \sum_{i=1}^{k} \left(d_i^{n+1,j+1} - d_i^{n+1,j} \right) \left\langle \theta'(\psi_h^{n+1,j}) \phi_i, \phi_s \right\rangle - \langle \theta(\psi^n), \phi_s \rangle \\
&+ \Delta t \sum_{i=1}^{k} d_i^{n+1,j+1} \left\langle K(\psi_h^{n+1,j}) \nabla \phi_i, \nabla \phi_s \right\rangle + \Delta t \left\langle K(\psi_h^{n+1,j}) \nabla z, \nabla \phi_s \right\rangle \\
&+ \Delta t \sum_{i=1}^{k} \left(d_i^{n+1,j+1} - d_i^{n+1,j} \right) \left\langle K'(\psi_h^{n+1,j}) \nabla \psi_h^{n+1,j}, \nabla \phi_s \right\rangle + \Delta t \sum_{i=1}^{k} \left(d_i^{n+1,j+1} - d_i^{n+1,j} \right) \left\langle K'(\psi_h^{n+1,j}) \nabla z, \nabla \phi_s \right\rangle \\
&= \Delta t \left\langle f^{n+1}, \phi_s \right\rangle.
\end{aligned} \tag{3.45}$$

In order to interpret equation (3.45) as row $s \in \{1, \dots, k\}$ of a system of linear equations, we require some definitions:

$$\begin{array}{llll}
\boldsymbol{W} & = (w_{lm}) \in \mathbb{R}^{k \times k}, & w_{lm} & := \langle \phi_m, \phi_l \rangle, \\
\boldsymbol{A}^{n,j} = \boldsymbol{A}(\psi_h^{n,j}) & = (a_{lm}^{n,j}) \in \mathbb{R}^{k \times k}, & a_{lm}^{n,j} & := \left\langle K(\psi_h^{n,j}) \nabla \phi_m, \nabla \phi_l \right\rangle, \\
\boldsymbol{\vartheta}^{n,j} = \boldsymbol{\vartheta}(\psi_h^{n,j}) & = (\vartheta_l^{n,j}) \in \mathbb{R}^{k}, & \vartheta_l^{n,j} & := \left\langle \theta(\psi_h^{n,j}), \phi_l \right\rangle, \\
\boldsymbol{b}^{n,j} = \boldsymbol{b}(\psi_h^{n,j}) & = (b_l^{n,j}) \in \mathbb{R}^{k}, & b_l^{n,j} & := \langle f^n, \phi_l \rangle - \left\langle K(\psi_h^{n,j}) \nabla z, \nabla \phi_l \right\rangle, \\
\boldsymbol{K}^{n,j} = \boldsymbol{K}(\psi_h^{n,j}) & = (\kappa_{lm}^{n,j}) \in \mathbb{R}^{k \times k}, & \kappa_{lm}^{n,j} & := \left\langle K'(\psi_h^{n,j}) \nabla \psi_h^{n,j} \phi_m \nabla \phi_l \right\rangle, \\
\boldsymbol{H}^{n,j} = \boldsymbol{H}(\psi_h^{n,j}) & = (\eta_{lm}^{n,j}) \in \mathbb{R}^{k \times k}, & \eta_{lm}^{n,j} & := \left\langle \theta'(\psi_h^{n,j}) \phi_m, \phi_l \right\rangle, \\
\boldsymbol{B}^{n,j} = \boldsymbol{B}(\psi_h^{n,j}) & = (\beta_{lm}^{n,j}) \in \mathbb{R}^{k \times k}, & \beta_{lm}^{n,j} & := - \left\langle K'(\psi_h^{n,j}) \phi_m \nabla z, \nabla \phi_l \right\rangle.
\end{array} \tag{3.46}$$

Equation (3.45) arises now as a row of the linear equation system for the difference between the unknown vector of coefficients $\boldsymbol{d}^{n+1,j+1} := (d_1^{n+1,j+1}, \dots, d_k^{n+1,j+1})^T \in \mathbb{R}^k$ and the recently computed vector $\boldsymbol{d}^{n+1,j}$,

$$\begin{aligned}
&\left[\boldsymbol{H}^{n+1,j} + \Delta t \boldsymbol{K}^{n+1,j} \cdot \boldsymbol{d}^{n+1,j} + \Delta t \boldsymbol{A}^{n+1,j} - \Delta t \boldsymbol{B}^{n+1,j} \right] (\boldsymbol{d}^{n+1,j+1} - \boldsymbol{d}^{n+1,j}) \\
&= - \left[\boldsymbol{\vartheta}^{n+1,j} - \boldsymbol{\vartheta}^n + \Delta t \boldsymbol{A}^{n+1,j} \boldsymbol{d}^{n+1} - \Delta t \boldsymbol{b}^{n+1,j} \right].
\end{aligned} \tag{3.47}$$

For completeness, we remark that linearization by the modified Picard scheme leads to

$$\left[\boldsymbol{H}^{n+1,j} + \Delta t \boldsymbol{A}^{n+1,j} \right] \boldsymbol{d}^{n+1,j+1} = \boldsymbol{\vartheta}^n - \boldsymbol{\vartheta}^{n+1,j} + \Delta t \boldsymbol{b}^{n+1,j} + \boldsymbol{H}^{n+1,j} \cdot \boldsymbol{d}^{n+1,j}, \tag{3.48}$$

and applying the L-scheme, one obtains

$$\left[L_\psi \boldsymbol{W} + \Delta t \boldsymbol{A}^{n+1,j} \right] \boldsymbol{d}^{n+1,j+1} = \boldsymbol{\vartheta}^n - \boldsymbol{\vartheta}^{n+1,j} + \Delta t \boldsymbol{b}^{n+1,j} + L_\psi \boldsymbol{W} \cdot \boldsymbol{d}^{n+1,j}. \tag{3.49}$$

The usual choice for the initial coefficients at time t^{n+1} is given by

$$\boldsymbol{d}^{n+1,0} = \boldsymbol{d}^n. \tag{3.50}$$

The finite element method

We opt for a finite-dimensional subspace $X_h \subset H^1(\Omega_{\text{pm}})$ admitting basis functions with small support to keep the matrices in the system of linear equations (3.47) sparse. Thus, we furnish the flow domain Ω_{pm} with a triangulation $\mathcal{T}_h$ containing triangles $T \in \mathcal{T}_h$ whose k corners are $\{\vec{x}_i \in \overline{\Omega_{\text{pm}}} : i \in \{1,\dots,k\}\}$ and select the finite-dimensional subspace

$$\hat{X}_h := \{\phi \in C^0(\Omega_{\text{pm}}) : \phi|_T \in \mathcal{P}_1(T), T \in \mathcal{T}_h\} \subset H^1(\Omega_{\text{pm}}), \tag{3.51}$$

where $\mathcal{P}_1(T)$ is defined as the space of linear polynomials on T. A basis of $\hat{X}_h$ is given by the set

$$\{\phi_i \in \hat{X}_h : \phi_i(\vec{x}_l) = \delta_{il}, i, l \in \{1,\dots,k\}\}. \tag{3.52}$$

The method of confining to a finite-dimensional solution space $\hat{X}_h$ whose basis is associated with nodes of a triangulation $\mathcal{T}_h$ to solve a system of linear equations arising from the weak formulation is what is known as the most elementary FEM. An elaborate introduction in FEM is given e.g. in [25]. The convergence of FEM for Richards' equation has been widely studied, e.g. in [2, 3, 40] for Galerkin FEM and in [45, 47, 52] for Mixed FEM.
We obtain a FEM-solution $\psi_h^n \in \hat{X}_h$ of problem (3.35) at each time level n by solving the system of linear equations (3.47) iteratively until

$$\|\psi_h^{n,j+1} - \psi_h^{n,j}\|_2 < \varepsilon_{\text{abs}} + \varepsilon_{\text{rel}}\|\psi^{n,j+1}\|_2, \tag{3.53}$$

for $\varepsilon_{\text{abs}}, \varepsilon_{\text{rel}} > 0$.

Implementation notes

For the evaluation of the occurring integrals, numerical quadrature is applied. Therefore, we define the standard triangle

$$T^* := \{\vec{\xi} = (\xi,\zeta)^T \in \mathbb{R}^2 : \xi + \zeta \le 1, \xi, \zeta \ge 0\}, \tag{3.54}$$

and compute the integrals on the triangles $T \in \mathcal{T}_h$ using the transformation theorem.
The unique affine transformation F from T^* to an arbitrary triangle $T \subset \mathbb{R}^2$ with corners $\vec{a}_1, \vec{a}_2, \vec{a}_3 \in \mathbb{R}^2$ satisfying $\vec{F}\left((0,0)^T\right) = \vec{a}_1, \vec{F}\left((1,0)^T\right) = \vec{a}_2, \vec{F}\left((0,1)^T\right) = \vec{a}_3$ is given by

$$\begin{cases} \vec{F} : T^* \to T, \\ \vec{F} : \vec{\xi} \longmapsto \vec{F}(\vec{\xi}), \end{cases} \tag{3.55}$$

where

$$\vec{F}(\vec{\xi}) = \underbrace{\begin{pmatrix} | & | \\ \vec{a}_2 - \vec{a}_1 & \vec{a}_3 - \vec{a}_1 \\ | & | \end{pmatrix}}_{:=C} \vec{\xi} + \begin{pmatrix} | \\ \vec{a}_1 \\ | \end{pmatrix}. \tag{3.56}$$

We illustrate the assembling of the matrices in system (3.47) using the example of an entry $a_{lm}^{n+1,j}$ of matrix $A^{n+1,j}$: Since the support of a basis function $\phi_s \in \hat{X}_h$ for $\hat{X}_h$ as defined in definition (3.52) is included in the triangles $\mathcal{T}_h^s := \{T \in \mathcal{T}_h : \vec{x}_s \text{ is corner of } T\}$, one only needs to evaluate the integrals on $\mathcal{T}_h^{lm} := \mathcal{T}_h^l \cap \mathcal{T}_h^m$ and sum up the results, i.e. $a_{lm}^{n+1,j} = \sum_{T \in \mathcal{T}_h^{lm}} a_{lm,T}^{n+1,j}$ with $a_{lm,T}^{n+1,j} := \left\langle K(\psi_h^{n+1,j})\nabla\phi_m, \nabla\phi_l \right\rangle_{L^2(T)}$.

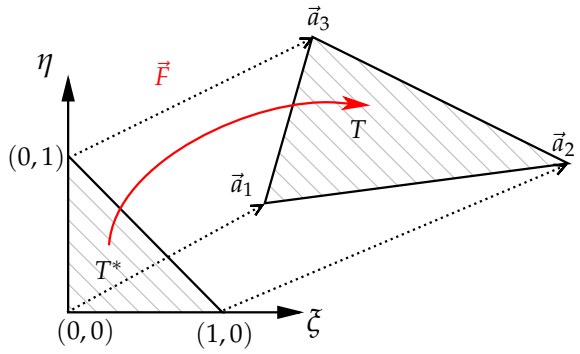


Figure 3.2: Affine transformation $\vec{F}$

On a triangle $T \in \mathcal{T}_h^{lm}$ with corners $\vec{x}_{T_1}, \vec{x}_{T_2}, \vec{x}_{T_3} \in \mathbb{R}^2$, we compute

$$\begin{aligned} a_{lm,T}^{n+1,j} &= \int_{T=\vec{F}(T^*)} K(\psi_h^{n+1,j}(\vec{x}))\, \nabla_{\vec{x}}\phi_m(\vec{x}) \cdot \nabla_{\vec{x}}\phi_l(\vec{x})\, d\vec{x} \\ &= |\det \boldsymbol{C}| \int_{T^*} K\left(\sum_{i=1}^{3} d_{T_i}^{n+1,j}\phi_{T_i}(\vec{F}(\vec{\xi}))\right) \nabla_{\vec{x}}\phi_m(\vec{F}(\vec{\xi})) \cdot \nabla_{\vec{x}}\phi_l(\vec{F}(\vec{\xi}))\, d\vec{\xi} \\ &= |\det \boldsymbol{C}| \int_{T^*} K\left(\sum_{i=1}^{3} d_{T_i}^{n+1,j}\phi_i^*(\vec{\xi})\right) \left((\boldsymbol{C}^{-1})^T \nabla_{\vec{\xi}}\phi_m^*(\vec{\xi})\right) \cdot \left((\boldsymbol{C}^{-1})^T \nabla_{\vec{\xi}}\phi_l^*(\vec{\xi})\right) d\vec{\xi}, \end{aligned} \tag{3.57}$$

$\{\phi_i^*\}_{i=1}^3$ being defined as the Lagrange basis on the standard triangle, for $\vec{\xi} = (\xi, \zeta)^T \in T^*$ given by

$$\begin{aligned} \phi_1^*(\vec{\xi}) &= 1 - \xi - \zeta, \\ \phi_2^*(\vec{\xi}) &= \xi, \\ \phi_3^*(\vec{\xi}) &= \zeta. \end{aligned} \tag{3.58}$$

In calculation (3.57), we employed the transformation theorem for integrals and the identities $\phi_{T_i} \circ \vec{F} = \phi_i^*$ and $\nabla_{\vec{x}} = (\boldsymbol{C}^{-1})^T \nabla_{\vec{\xi}}$. The latter one is due to the chain rule of differentiation. Furthermore, note that $\psi_h^{n+1,j}$ is merely composed of the three basis functions $\phi_{T_1}, \phi_{T_2}, \phi_{T_3}$ inside T since all the other basis functions vanish in T, to be more exact $\psi_h^{n+1,j}|_T \in \operatorname{span}\{\phi_{T_1}, \phi_{T_2}, \phi_{T_3}\}$.
Recall that for this setting, the no-flow boundary conditions are implicitly satisfied as they are included in the weak formulation (3.38) which the FEM is derived from.
The final expression in equation (3.57) is approximated by a quadrature of the form

$$\int_{T^*} v(\vec{\xi})\, d\vec{\xi} \approx \sum_{i=1}^{r} \gamma_i v(\vec{\eta}_i), \tag{3.59}$$

in which $\gamma_i > 0$ are weights and $\vec{\eta}_i \in T^*$ are evaluation points for $i \in \{1, \ldots, r\}$ under the assumption that the integrand $v : T^* \to \mathbb{R}$ is sufficiently regular to be evaluated pointwise.

3.3 Coupled flow

We direct our focus now on the implementation of the coupling conditions (1.39) and (1.40). Particularly condition (1.39) is slightly intricate since it requires post-processing of the pressure heights in order to determine the flux in the subsurface.

3.3.1 Coupling from subsurface to surface

The surface flow employs the vertical flux from the subsurface as a source or sink term, mathematically expressed by coupling condition (1.39).
In the framework of a finite volume scheme, the source term f_j^n should be considered as

$$f_j^n = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} f^n(x)\, dx. \tag{3.60}$$

Applying the midpoint rule for this integral, we find $f_j^n \approx f^n(x_j)$. For the coupled problem, we have $f_j^n = v_{\mathrm{pm}}(\psi_j^n) \cdot \vec{n}$. As the flux v_{pm} is not a primary variable of the Richards' equation, it has to be computed from the pressure heights. Recall that v_{pm} is given by

$$v_{\mathrm{pm}} = -K(\theta(\psi))\nabla(\psi + z), \tag{3.61}$$

which can be reconstructed numerically by finite differences at the nodes. The resulting values of $v_{\mathrm{pm}} \cdot \vec{n}$ are interpolated on the surface grid, e.g. using piecewise cubic Hermite polynomials as depicted in Figure 3.3.

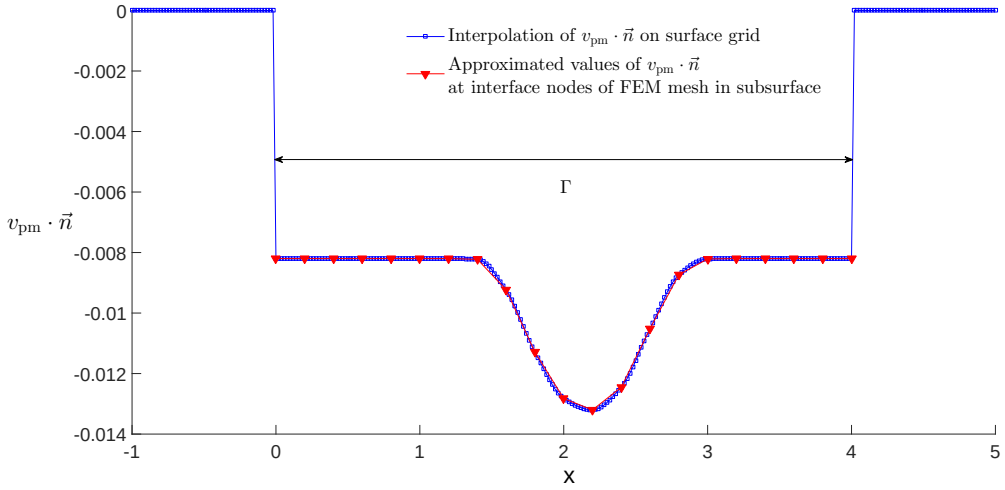


Figure 3.3: Interpolated values of $v_{\mathrm{pm}} \cdot \vec{n}$ on surface grid

3.3.2 Coupling from surface to subsurface

The converse coupling condition (1.40) states that the pressure head at the interface is given by the water head. Let ψ_m^{n+1} be the imposed pressure height at a particular node $\vec{x}_m$ at time $t = t^{n+1}$ and let d_m^n be the recently computed solution at this node (which is given by the corresponding coefficient for the m-th basis function because of the choice of the basis (3.52)).
The system of linear equation (3.47) describes no-flow conditions at the entire boundary up to now, in order to impose a Dirichlet condition for the pressure height at node $\vec{x}_m$, row m of the system of linear equations is replaced by $(\delta_{ml})_{l=1}^k$ and the m-th entry of the load vector is set to the difference of the pressure heights for the first iteration, i.e. $\psi_m^{n+1} - d_m^n$. The m-th row of the system writes consequently

$$1 \cdot (d_m^{n+1,1} - d_m^{n+1,0}) = \psi_m^{n+1} - d_m^n. \tag{3.62}$$

Using $d_m^{n+1,0} = d_m^n$, we see that $\psi_h^{n+1,1}(\vec{x}_m) = \sum_{i=1}^k d_i^{n+1,1}\phi_i(\vec{x}_m) = d_m^{n+1,1} = \psi_m^{n+1}$, indeed. From the second iteration on, the m-th row of the left hand side matrix remains $(\delta_{ml})_{l=1}^k$ and the m-th entry of the load vector is now substituted with zero as the Dirichlet condition is already taken into account.
For the modified Picard scheme and the L-scheme, the implementation is carried out similarly. In this way, we prescribe the pressure height at the interface via the water height of the current time level which is interpolated to the subsurface nodes being part of the interface.

3.3.3 Coupling algorithm

Since the velocities in the subsurface flow are usually considerably smaller than the ones of the surface flow and moreover because of the much higher costs to solve the non-linear two-dimensional subsurface system as compared to the surface flow, we expect to increase the efficiency of our algorithm by computing multiple steps at the surface before updating the subsurface flow, as proposed in the context of coupled surface-subsurface flows e.g. in [50]. Therefore, we introduce an algorithmic coupling constant

$$c_{\text{coupling}} = \frac{\text{Surface steps}}{\text{Subsurface steps}} \geq 1. \tag{3.63}$$

The resulting algorithm for the numerical solution of the coupled model (1.41) reads schematically:

Data: Time $T > 0$, coupling ratio c_{coupling}
Result: surface solution U consisting of h and v, subsurface solution $\boldsymbol{d}$

$U^0 \leftarrow u_0$;
$\boldsymbol{d}^0 \leftarrow \psi_0$;
determine Δt^0 satisfying the CFL-condition;
post-processing of $\boldsymbol{d}^0$: compute v_{pm}^0;
$t \leftarrow 0$;
$n \leftarrow 1$;
while $t < T$ **do**
  for $i = n : (n + c_{\text{coupling}} - 1)$ **do**
    compute U^i using v_{pm}^{n-1};
    determine Δt^i satisfying the CFL-condition;
    $t \leftarrow t + \Delta t^i$;
  end
  $n \leftarrow n + c_{\text{coupling}} - 1$;
  compute $\boldsymbol{d}^n$ using h^n solving system (3.40) iteratively with adapted entries at interface nodes;
  post-processing of $\boldsymbol{d}^n$: compute v_{pm}^n;
end

Algorithm 1: Coupling algorithm

Notice that for the subsurface problem a coarse time step results, given by the sum of all fine grid time steps lying in between. After the computation, the solution may be interpolated to a uniform time grid to simplify the visualization.

Chapter 4

Numerical simulations

In this chapter, we first check the correctness of our implementations of the individual systems using recognized examples. Then, we investigate Algorithm 1 with regard to conservation of mass and we examine the influence of several parameters on the numerical solution, such as the surface-subsurface coupling constant c_{coupling}, the surface grid size and the subsurface mesh size. Furthermore, we analyse numerically if the energy of the coupled problem as defined in Chapter 2 fulfils the second law of thermodynamics, despite a source term caused by the coupling whose sign is unknown a-priori. Finally, we consider a dam break scenario with realistic parameters; particular emphasis is placed on the partially dry bed on the surface.
The time step in all simulations for problems including a surface is chosen as $\Delta t = (1/2)\Delta x/\lambda_{\max}$. The time unit is 1 [s] in each simulation and the length scale is 1 [m].

4.1 Validation of uncoupled models

This section contains a numerical example for either flow domain for which an analytical solution is at hand in order to validate our implementation.

4.1.1 Riemann problem for the shallow water equations

This example represents the situation of an instantaneous dam break at $t = 0$ on a wet domain and is widely used to test the performance of numerical algorithms. Consider the Riemann problem for the shallow water equations (1.8) with the initial data

$$h_0(x) = \begin{cases} h_l, & x \le 0, \\ h_r, & x > 0, \end{cases}, \tag{4.1}$$

with $h_l > h_r > 0$ and $v_0 \equiv 0$. Furthermore, let be $q_M = q_I = 0$, thus no mass sources are present and no friction occurs.
As derived in [55], the analytical solution of this dam break problem consists of a right-going shock and a rarefaction wave propagating to the left, to be specific

$$h(x,t) = \begin{cases} h_l, & x \le x_A(t), \\ \frac{4}{9g}\left(\sqrt{gh_l} - \frac{x}{2t}\right)^2, & x_A(t) < x \le x_B(t), \\ \frac{c_m^2}{g}, & x_B(t) < x \le x_C(t), \\ h_r, & x_C(t) < x \end{cases} \qquad v(x,t) = \begin{cases} 0, & x \le x_A(t), \\ \frac{2}{3}\left(\frac{x}{t} + \sqrt{gh_l}\right), & x_A(t) < x \le x_B(t), \\ 2(\sqrt{gh_l} - c_m), & x_B(t) < x \le x_C(t), \\ 0, & x_C(t) < x, \end{cases} \tag{4.2}$$

in which $x_A(t) = -t\sqrt{gh_l}$, $x_B(t) = t(2\sqrt{gh_l} - 3c_m)$, $x_C(t) = t(2c_m^2(\sqrt{gh_l} - c_m))/(c_m^2 - \sqrt{gh_r})$ with the solution c_m of the equation $-8gh_rc_m^2\left(\sqrt{gh_l} - c_m\right)^2 + (c_m^2 - gh_r)^2(c_m^2 + gh_r) = 0$. The resulting water height at the initial shock location $x = 0$ is computed as $h_m = c_m^2/g$.
For the simulation, we consider the interval $\Omega_{\text{ff}} = (-3.5, 3.5)$ large enough so that the boundaries are not affected by the waves until the simulation time of $T = 0.25$. Furthermore we choose $h_l = 2$, $h_r = 1$, $g = 9.81$ and a grid size of $\Delta x = 10^{-3}$.

The results are depicted in Figure 4.1 for the LxF method, the LLxF method and Roe's approximate Godunov solver.

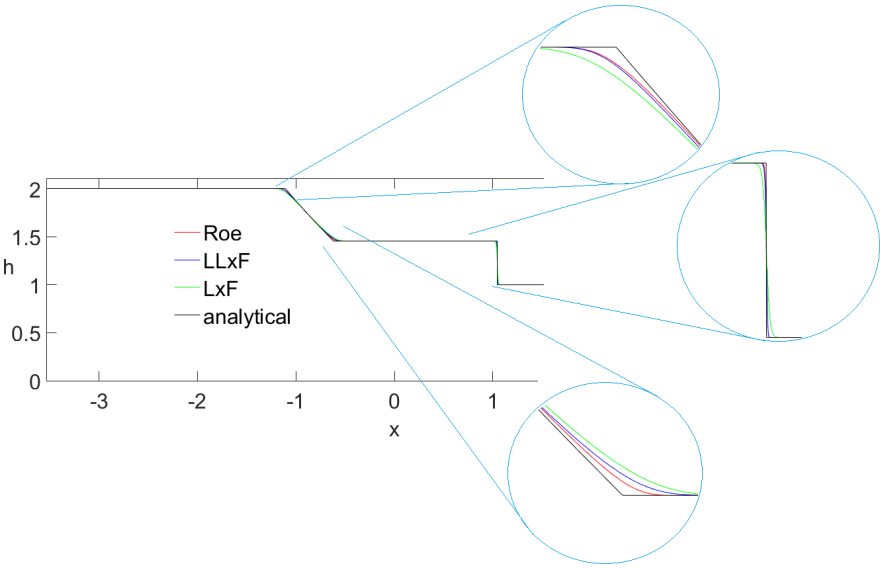


Figure 4.1: Computed water height of the Riemann problem for several solvers at $t = 0.25$

Figure 4.1 demonstrates that all methods investigated located the shock and the rarefaction correctly, Roe's solver yielding the least smearing, followed by the LLxF method and the LxF method.
The discrete $L^1(\Omega_{\text{ff}})$-error $\|h_{\text{numerical}}(\cdot, T) - h_{\text{analytical}}(\cdot, T)\|_1$ is 0.0043 for Roe's method, 0.0062 for the LLxF method and 0.0126 for the LxF method.

4.1.2 Hornung–Messing problem for Richards' equation

For Richards' equation, analytical solutions are known only in a few cases requiring particular parametrizations of $\theta(\psi)$ and $K(\psi)$. A popular example is furnished by the Hornung–Messing problem (see, for instance, [13, 51]), for which the (non-physical) hydraulic relationships are given by

$$\theta(\psi) = \begin{cases} \frac{\pi^2}{2} - 2\arctan^2(\psi), & \psi < 0, \\ \frac{\pi^2}{2}, & \psi \geq 0, \end{cases} \qquad K(\psi) = \begin{cases} \frac{2}{(1+\psi)^2}, & \psi < 0, \\ 2, & \psi \geq 0. \end{cases} \tag{4.3}$$

The flow domain $\Omega_{\text{pm}} = (0,1) \times (0,1)$ is aligned horizontally (the two spatial coordinates are denoted by x and y) and hence no gravitational convection term occurs. Therefore, Richards' equation takes the form

$$\partial_t \theta(\psi) - \nabla \cdot (K(\psi) \nabla \psi) = 0. \tag{4.4}$$

The exact solution is then given by

$$\psi(x, y, t) = \begin{cases} -\frac{1}{2}s, & s < 0, \\ -\tan\left(\frac{\exp(s)-1}{\exp(s)+1}\right), & s \geq 0, \end{cases} \tag{4.5}$$

where $s = x - y - t$. Thus, the domain is saturated wherever $s < 0$, in particular completely saturated for $t \geq 1$. Figure 4.2 shows the water content θ at $t = 0.2$ of the numerical solution computed on a uniform mesh consisting of 1250 triangles with a time step of $\Delta t = 0.01$. The L^2-error between the numerical solution and the analytical one is presented in Figure 4.3. The error curve exhibits the typical shape as in [13], i.e. the error vanished the more the domain became saturated, and attests the correctness of the implementation.

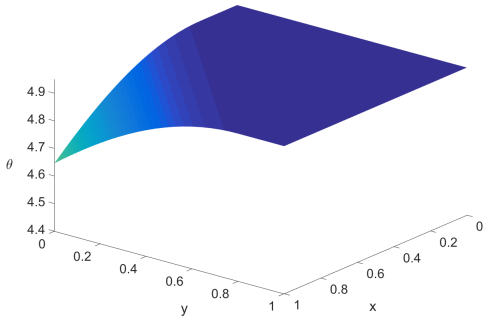


Figure 4.2: Water content θ at $t = 0.2$

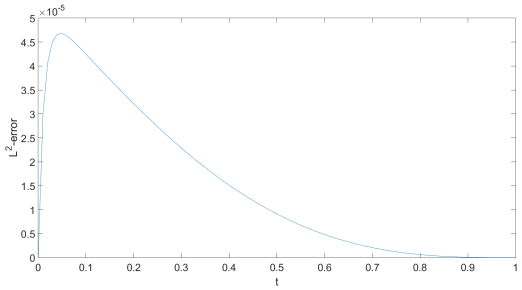


Figure 4.3: L^2-error between analytical and numerical solution

4.2 Numerical example for the coupled problem 1

Theorem 2.1.1 shows the conservation of mass for the coupled model (1.41) in theory. Therefore, it is interesting to analyse the loss or gain of mass of the numerical solution. The Galerkin FEM scheme used for the spatial discretization of the subsurface is not mass conservative in itself, which is why the coupled model is not expected to be mass conservative in opposition to discretizations utilizing FVM or MFEM. Besides, mass errors occur when interpolating the flux computed in the subsurface to the usually denser surface grid. Nonetheless, it is worth to check if the discrepancy between the initial mass and the mass at the end of the simulation is admissible.

For this purpose, we consider the setting presented in Table 4.1. The initial water height is smooth and its support lies in $(4/3, 8/3)$. Initially, the upper half of the subsurface is influenced by the water height (see Figure 4.4) and of course, the initial conditions satisfy the boundary condition and the Dirichlet coupling condition. At the surface boundaries, periodic boundary conditions are imposed.

The total mass M_{total} in the system is evaluated as

$$M_{\text{total}} = \frac{\Delta x}{2}(h_1 + h_M) + \Delta x \sum_{i=2}^{M-1} h_i + \frac{|T|}{3}\left(\sum_{T \in \mathcal{T}_h} \theta(\psi_{T_1}) + \theta(\psi_{T_2}) + \theta(\psi_{T_3})\right), \tag{4.6}$$

in which $|T|$ denotes the area of a triangle which is constant since we use a regular mesh, and the subscript T_i refers to the i-th corner of triangle T for $i \in \{1, 2, 3\}$.

We run a simulation on a dense grid ($\Delta x = 2.5 \cdot 10^{-4}$, 28800 triangles in the subsurface mesh, $c_{\text{coupling}} = 10$), and investigate the errors of the numerical solutions for several surface grid sizes, subsurface mesh sizes and coupling constants with respect to the dense grid solution $(h_{\text{dense}}, \psi_{\text{dense}})$, since no analytical solution of this problem is available. The most interesting moment in this numerical example is at $t \approx 0.45$ when the subsurface becomes fully saturated. The development of the absolute value of the subsurface velocity beforehand is displayed in Figure 4.5 and the water height at $t = 0.45$ is depicted in Figure 4.6.

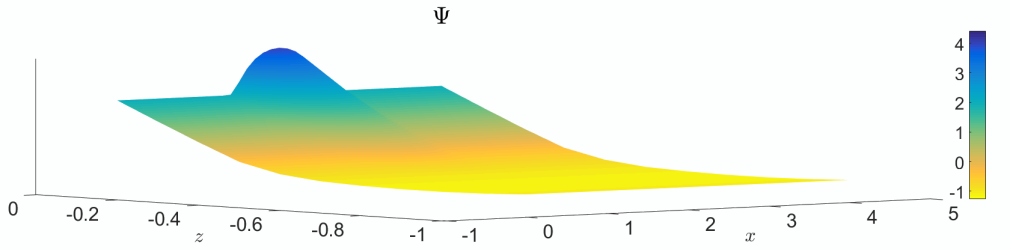


Figure 4.4: Initial pressure height

Geometry		
FF	$(\hat{x}_l, \hat{x}_r)$	$(-1, 5)$
PM	(x_l, x_r)	$(0, 4)$
|	(z_b, z_t)	$(-1, 0)$
Initial conditions		
FF	$h_0(x)$	$\begin{cases} 0.2 + 0.2\exp(1)\exp\left\{\left[\left((2-x)\big/(2/3)\right)^2 - 1\right]^{-1}\right\}, & x \in (4/3, 8/3), \\ 0.2, & \text{else} \end{cases}$
|	v_0	0
PM	$\psi_0(x, z)$	$\begin{cases} 2z^2 + 3z + h_0(x)(4z^2 + 4z + 1), & z \geq -1/2, \\ 2z^2 + 3z, & \text{else} \end{cases}$
Van Genuchten parameters		
PM	α	0.30
|	θ_S	0.40
|	θ_R	0.01
|	n	4.0
|	K_S	$1.5 \cdot 10^{-3}$
Physical parameters		
	g	9.81
FF	q_I	0
Simulation time		
	T	0.5
Numerical parameters		
FF	Numerical flux	Roe
PM	Solver	Newton
|	$\varepsilon_{\text{abs}}, \varepsilon_{\text{rel}}$	10^{-5}

Table 4.1: Simulation parameters

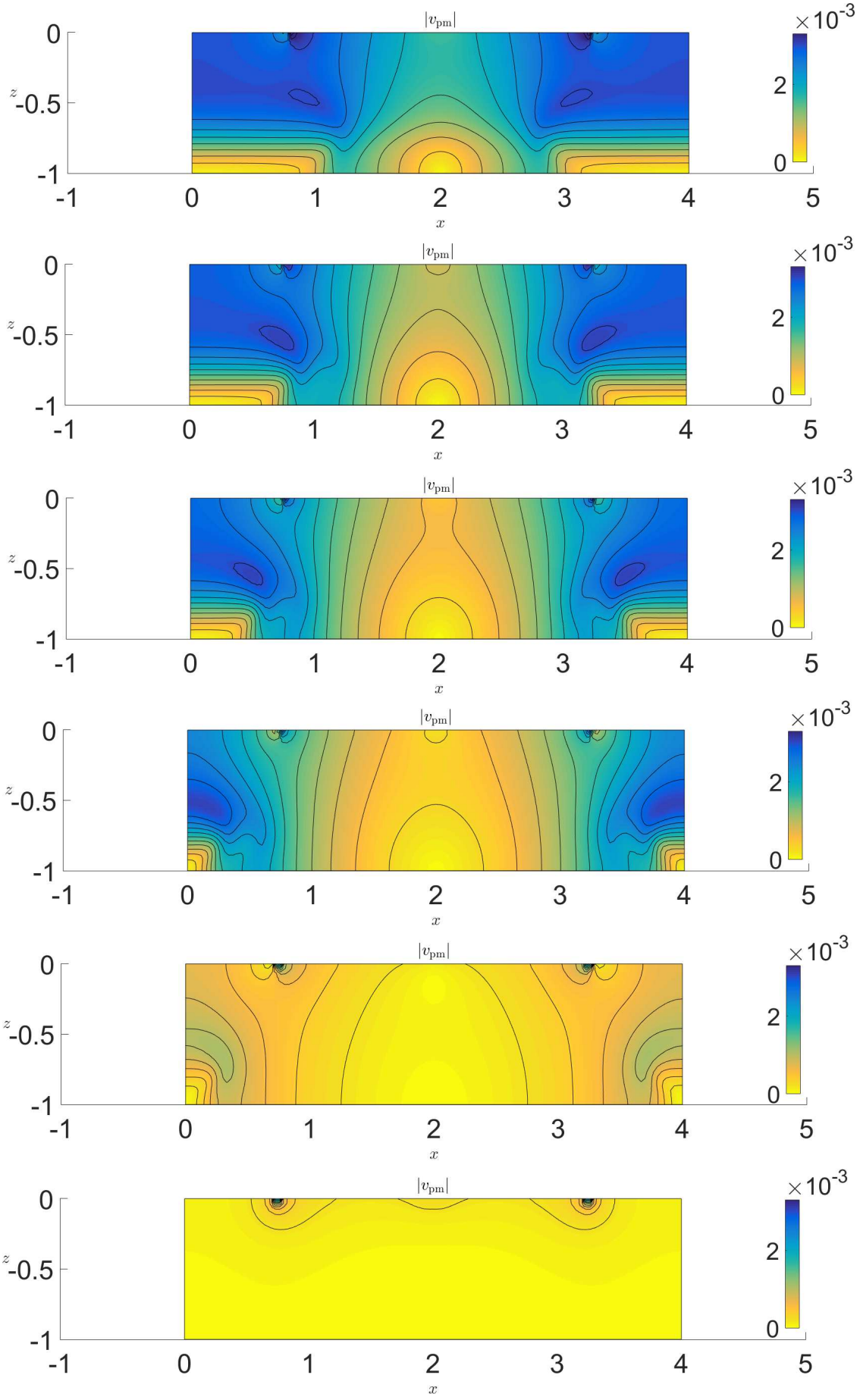


Figure 4.5: Absolute value of the velocity in the subsurface at $t = 0.438, 0.441, 0.443, 0.446, 0.449, 0.452$ (from top to bottom)

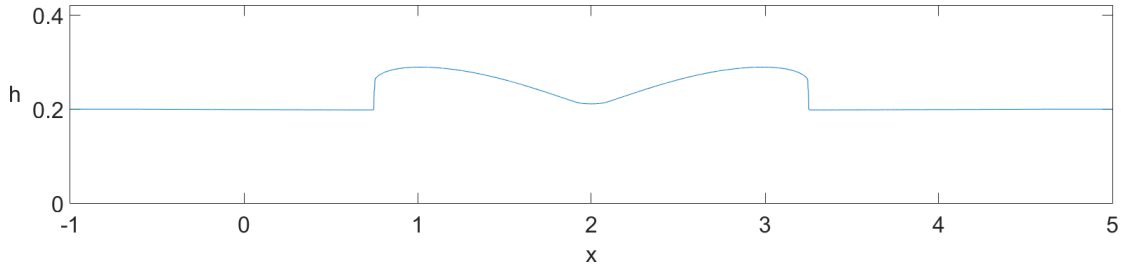


Figure 4.6: Water height when the subsurface becomes fully saturated

Figure 4.6 exhibits that two shocks have formed from the initially smooth surface water level – the breaking of the waves cannot be modelled by the one-dimensional shallow water equations. Below these shocks one observes high velocities, and as soon as the entire domain is saturated flow almost only takes place close to the interface (Figure 4.5).

c_{coupling}

First, we analyse the impact of the coupling parameter c_{coupling} on the error and on the conservation of mass. The surface grid width is chosen as $\Delta x = 10^{-3}$, the subsurface mesh consists of 200 triangles. The water height and the saturation for different values of c_{coupling} at $t = 0.264$ [s] are depicted in Figure 4.7.
Table 4.2 shows the balance of mass, the CPU times and the speed-up as compared to the computation for $c_{\text{coupling}} = 1$ for several coupling constants, namely

$$\text{speed-up}_i := \frac{\text{CPU time for } c_{\text{coupling}} = 1}{\text{CPU time for } c_{\text{coupling}} = i}. \tag{4.7}$$

It demonstrates that the conservation of mass is admissible for all c_{coupling} examined, with a maximum discrepancy of less than 0.01%. Larger coupling constants even yielded slightly better conservation of mass, possibly due to the mass error generated by the FEM discretization accumulating in each subsurface computation.

c_{coupling}	mass at $t = T$ / initial mass [%]	CPU time [s]	speed-up
1	100.00960	7406	1.00
2	100.00940	3819	1.94
5	100.00879	1562	4.74
10	100.00778	896	8.27
20	100.00576	479	15.46
30	100.00373	327	22.65
50	99.99969	208	35.61

Table 4.2: Conservation of mass and errors at $t = T$ for several c_{coupling}

Mass conservation alone is clearly not a sufficient criterion for the quality of the numerical solution. However, also the errors in comparison to the dense grid solution plotted in Figure 4.8 suggest that the quality of the solutions did not decrease substantially when the coupling constant c_{coupling} became larger to a certain extent, neither in the subsurface nor on the surface. For $c_{\text{coupling}} \leq 20$, no significant improvement of the L^1-error of the water height was observable since the errors due to the spatial and temporal discretization were apparently predominant as compared to the error caused by the surface-subsurface mass exchange. The L^2-error of the pressure height in the subsurface was slightly reduced when it came to smaller coupling constants.
Employing a time step for the subsurface much greater than the one for the surface leaded to a considerable speed-up of the CPU time.

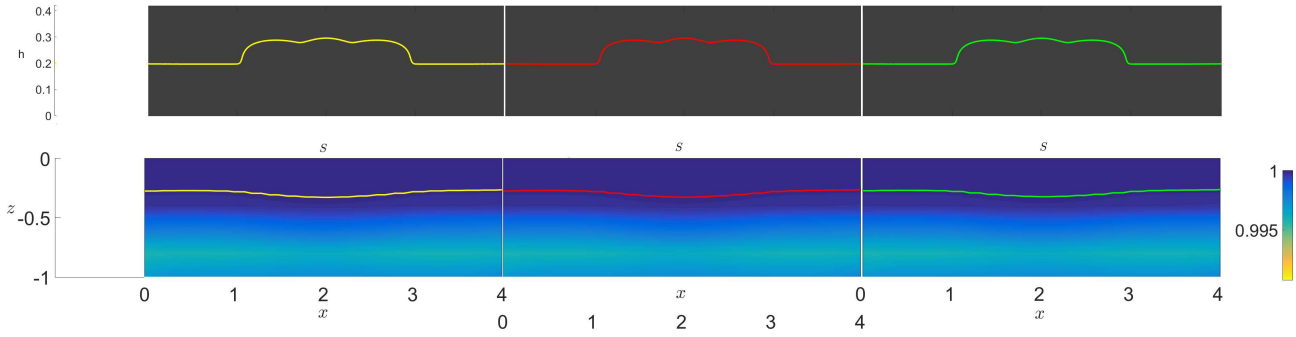


Figure 4.7: Water heights and saturations at $t = 0.264$ for $c_{\text{coupling}} = 1, 20, 50$ (from left to right), the coloured lines in the subsurface mark the front where $S = 99.99\%$

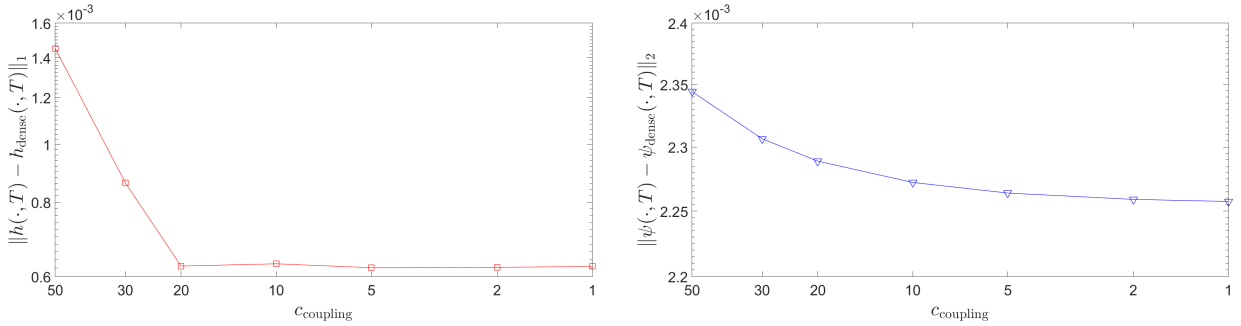


Figure 4.8: Errors for several c_{coupling} at $t = T$

Surface grid size

The choice of the surface grid size Δx has a crucial impact on the simulation since it limits the time step Δt via the CFL-condition additionally. The subsurface mesh is given as in the previous section. The coupling constant c_{coupling} is taken as 20.

Δx	mass at $t = T$ / initial mass [%]
10^{-2}	99.96911
$5 \cdot 10^{-3}$	99.98931
$2.5 \cdot 10^{-3}$	99.99959
$1.25 \cdot 10^{-3}$	100.00473
$6.25 \cdot 10^{-4}$	100.00730

Table 4.3: Conservation of mass at $t = T$ for several Δx

Table 4.3 shows that the mass at $t = T$ differed less than 0.01% from the initial mass for $\Delta x \leq 2.5 \cdot 10^{-3}$. When the surface grid size became smaller than $2.5 \cdot 10^{-3}$, the mass error at $t = T$ did slightly increase, but since the mass error was not monotonically increasing over the time it is possible that this result is only a snapshot and that the computations with a finer surface grid would lead to better results in the long term.
The errors of h and ψ for several values of Δx are plotted in Figure 4.9. Expectedly, the error of h decreased roughly linearly when Δx was reduced. The error of ψ remained approximately constant for $\Delta x \leq 5 \cdot 10^{-3}$.

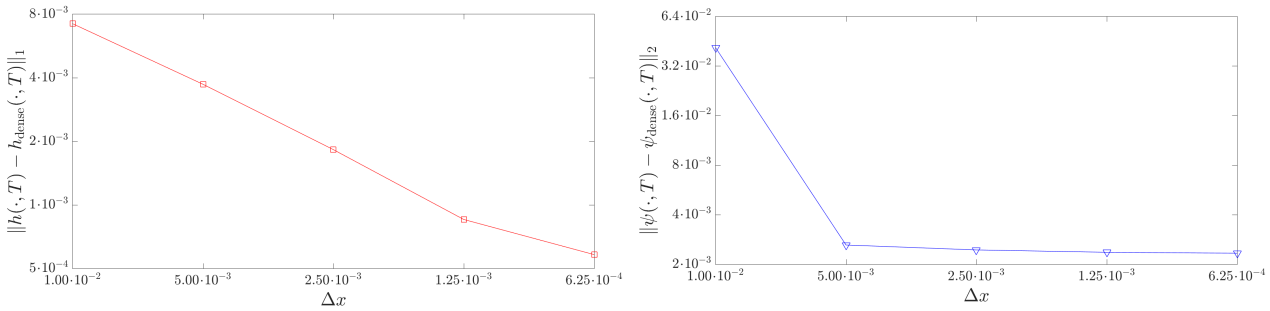


Figure 4.9: Errors for several Δx at $t = T$

Subsurface mesh size

We examine the mass error and the errors of h and ψ when the FEM mesh size is reduced. Table 4.4 shows the ratio of mass at $t = T$ to the initial mass. For all FEM meshes, the mass discrepancy was less than 0.006%.

No. of triangles	mass at $t = T$ / initial mass [%]
800	99.99815
3200	100.00576
7200	99.99567
12800	99.99552
20000	99.99545

Table 4.4: Discrete conservation of mass and errors at $t = T$ for several subsurface mesh sizes

Figure 4.10 demonstrates that the choice of the FEM mesh size did barely influence the error of the water height; refinement of the subsurface mesh without refining the surface grid even yielded a slight increase of the error. For more than 3200 triangles, the error of the pressure heights decreased when the subsurface mesh was further refined.

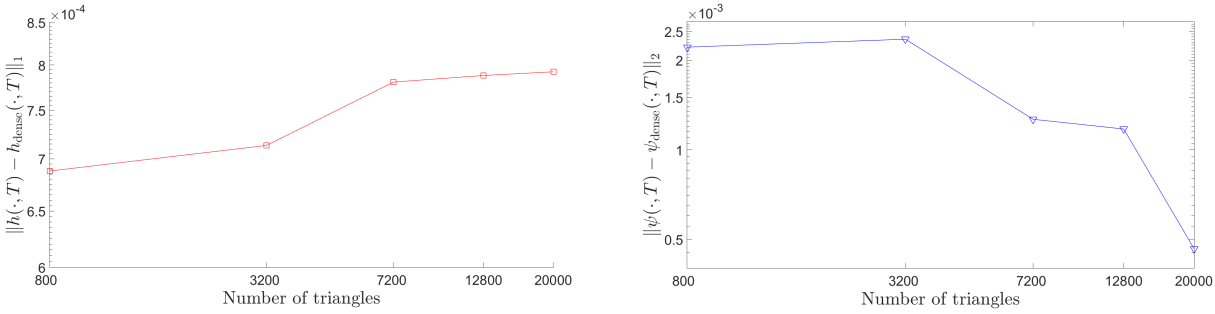


Figure 4.10: Errors for several subsurface mesh sizes at $t = T$

4.3 Numerical example for the coupled problem 2

This numerical example is to analyse the energy of the coupling problem introduced in Chapter 2 numerically, which reads

$$\mathcal{W}_{\text{total}} = \mathcal{W}_{\text{ff}} + \mathcal{W}_{\text{pm}}, \tag{4.8}$$

where

$$\begin{aligned} \mathcal{W}_{\text{ff}} &= \int_{\mathbb{R}} \frac{h^2}{2} + \frac{hv^2}{2g}\, dx, \\ \mathcal{W}_{\text{pm}} &= \int_{\Omega_{\text{pm}}} \int_0^u b'(v)\mathcal{K}^{-1}(v)\, dv\, d\vec{x}. \end{aligned} \tag{4.9}$$

Mathematically, an energy function should only increase over time if energy sources are present in the system. An energy source term in the coupled surface/subsurface system (1.41) is due to the gravitational convection, and in contrast to the coupled model consisting of the linear transport equation or the kinematic wave equation, a further term $\mathcal{G}_{\text{coupling}}$ caused by the coupling arises, so that one finds (see Remark 2.2.1)

$$\frac{d}{dt}\mathcal{W}_{\text{total}} \leq \mathcal{G}_{\text{grav}} + \mathcal{G}_{\text{coupling}}, \tag{4.10}$$

in which

$$\begin{aligned} \mathcal{G}_{\text{grav}} &= -\int_{\partial\Omega_{\text{pm}}} u\vec{e}_z \cdot \vec{n}\, d\vec{\xi}, \\ \mathcal{G}_{\text{coupling}} &= \int_{\Gamma} \frac{-v^2}{2g}\, v_{\text{pm}} \cdot \vec{n}\, d\vec{\xi}, \end{aligned} \tag{4.11}$$

in case of $q_I = 0$.
It is worth mentioning that in a steady state of the uncoupled Richards' equation, i.e. $\psi = c - z$ for $c \in \mathbb{R}$,

$$\mathcal{G}_{\text{grav}} \leq 0 \tag{4.12}$$

holds true for the rectangular geometry considered in this thesis, because

$$\vec{e}_z \cdot \vec{n} = \begin{cases} 0, & \text{at the lateral boundaries,} \\ +1, & \text{at the upper boundary,} \\ -1, & \text{at the lower boundary,} \end{cases} \tag{4.13}$$

since ψ is greater at the lower boundary than at the upper boundary of Ω_{pm} and due to $K(\theta(\psi)) > 0$.
As the physical justification for the term $\mathcal{G}_{\text{coupling}}$ is not clear, it would be desirable if this term had little effect on the energy.
We consider a similar setup as in Section 4.3, but the initial water height is piecewise constant, and the subsurface contains a zone D_{pm} of much lower permeability as compared to the remaining domain. The simulation parameters are listed in Table 4.5. As in the previous example, the subsurface is influenced by the initial water height for $z \geq -1/2$, and periodic surface boundary conditions are imposed.
Figures 4.11 and 4.12 show the water height, the pressure height, the saturation and the hydraulic conductivity at $t = 0.075$ and $t = 0.150$, respectively. At $t = 0.075$, large parts of the subsurface domain have been infiltrated by water, whereas the zone below D_{pm} has remained comparatively dry. The water height above the subsurface domain has reduced visibly owing to the seepage into the subsurface, this more on the left hand side of Ω_{pm} since the seepage is impeded on the right hand side due to the low permeability in D_{pm}. At $t = 0.150$, the entire domain to the left of D_{pm} as well as parts below D_{pm} are saturated.
As to the energy, Figure 4.13 demonstrates that both the surface energy and the subsurface energy are monotonically decreasing throughout the simulation, hence, the total energy as defined in equation (4.8) satisfies the second law of thermodynamics. This is due to the dissipative term $-\int_{\Omega_{\text{pm}}} K(\theta(\psi))|\nabla\psi|^2\, d\vec{x}$ (see the proof of Theorem 2.2.3) and the gravitational source term $\mathcal{G}_{\text{grav}}$ which is negative at all times, though it increases the more the domain becomes saturated. In contrast, the coupling term $\mathcal{G}_{\text{coupling}}$ is positive, but it is one order of magnitude smaller in comparison to $\mathcal{G}_{\text{grav}}$.
Altogether, we assert that in this numerical example, the coupling energy source term did not yield a violation of $(d/dt)\mathcal{W}_{\text{total}} \leq 0$.

Geometry		
FF	$(\hat{x}_l, \hat{x}_r)$	$(-1, 5)$
PM	(x_l, x_r)	$(0, 4)$
|	(z_b, z_t)	$(-1, 0)$
|	D_{pm}	$[3, 3.5] \times [-0.5, -0.25]$
Initial conditions		
FF	$h_0(x)$	$\begin{cases} 2, & x \in (4/3, 8/3), \\ 1, & \text{else} \end{cases}$
|	v_0	0
PM	$\psi_0(x, z)$	$\begin{cases} 2z^2 + 3z + h_0(x)(4z^2 + 4z + 1), & z \geq -1/2, \\ 2z^2 + 3z, & \text{else} \end{cases}$
Van Genuchten parameters		
PM	α	0.60
|	θ_S	0.50
|	θ_R	0.00
|	n	3.0
|	$K_S(x, z)$	$\begin{cases} 10^{-4} & (x, z) \in D_{\text{pm}}, \\ 0.1, & \text{else} \end{cases}$
Physical parameters		
	g	9.81
FF	q_I	0
Time evolution		
	T	0.15
	c_{coupling}	20
Numerical parameters		
FF	Numerical flux	Roe
|	Δx	10^{-3}
PM	# triangles	20000
|	Solver	Modified Picard
|	$\varepsilon_{\text{abs}}, \varepsilon_{\text{rel}}$	10^{-5}

Table 4.5: Simulation parameters

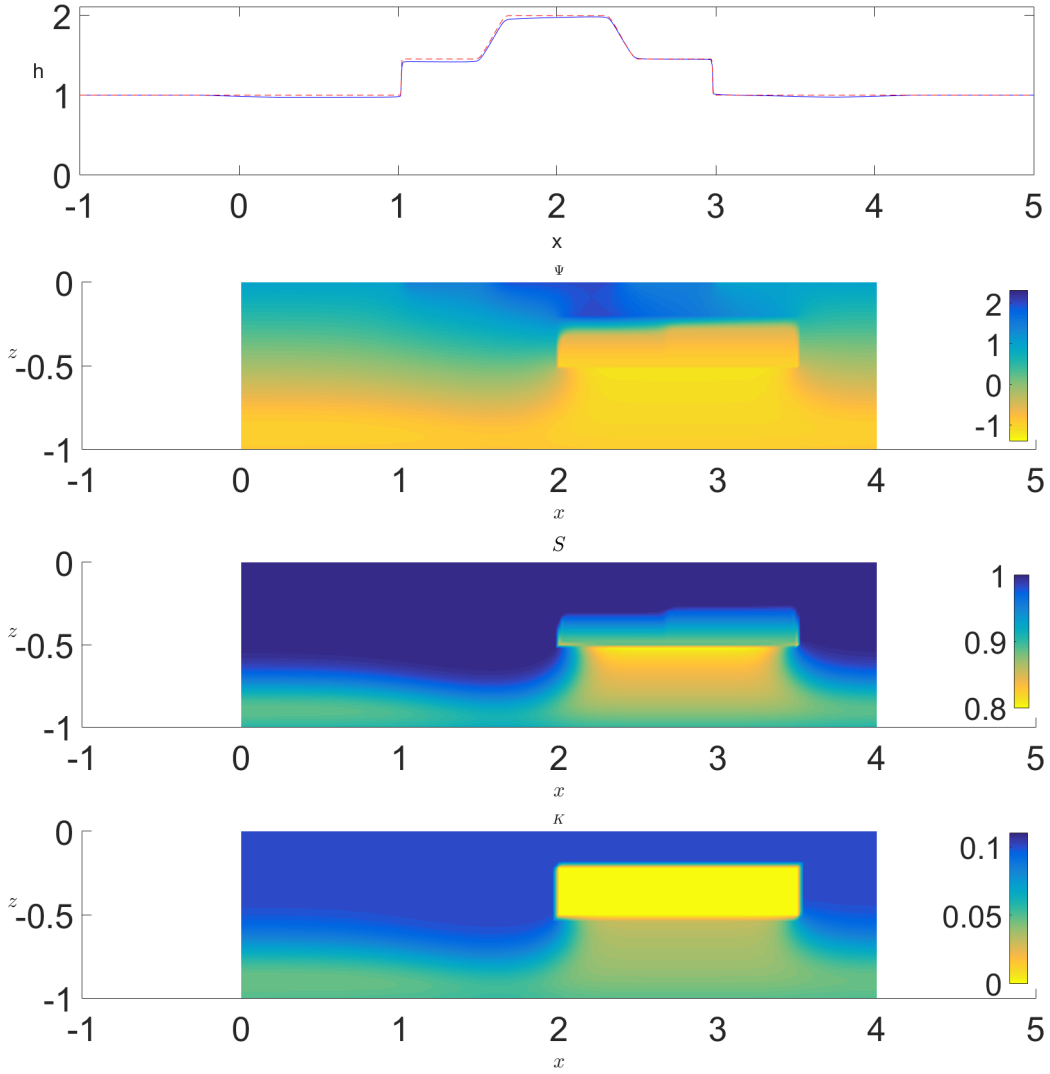


Figure 4.11: $h, \psi, S(\psi), K(\psi)$ at $t = 0.075$ (from top to bottom), the dashed red line marks the water height in the uncoupled case

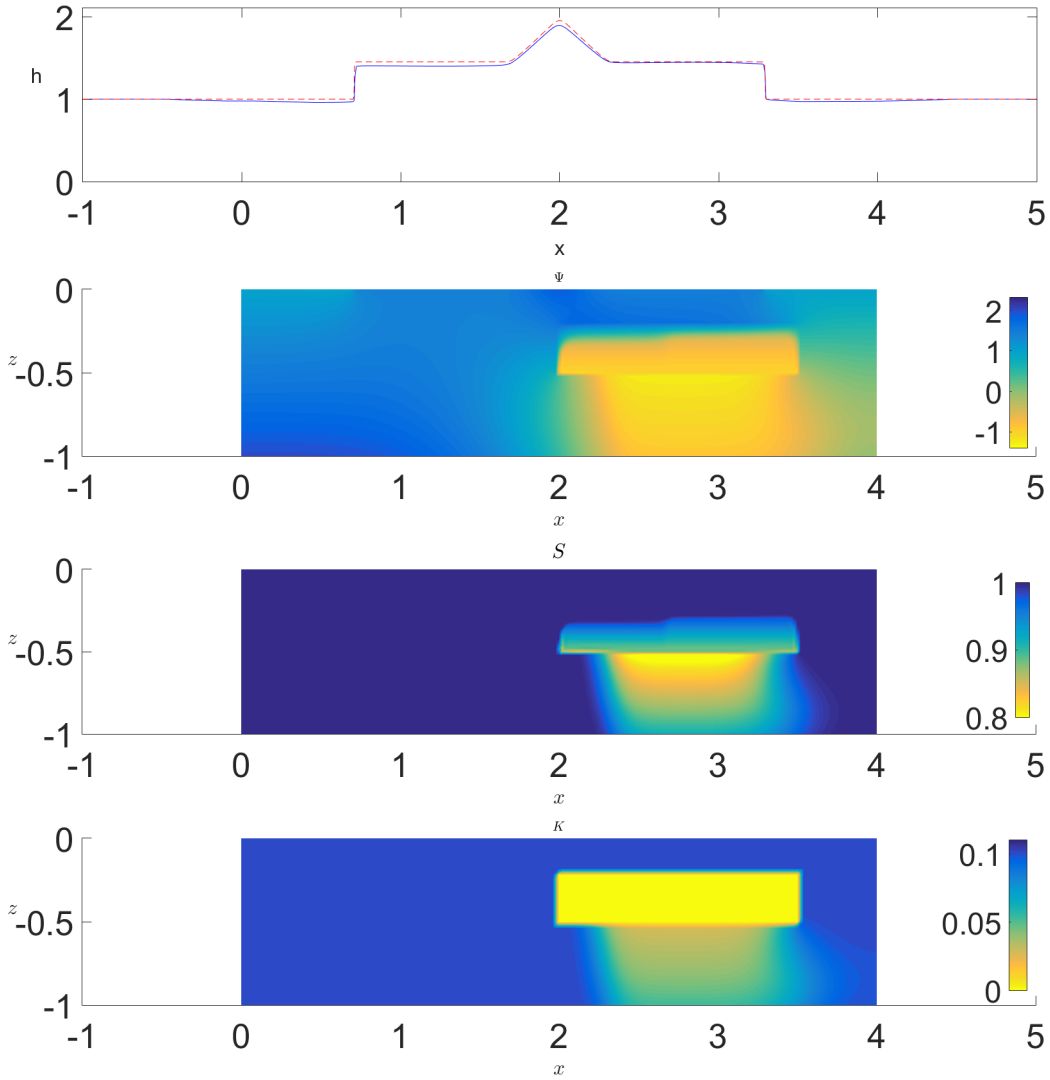


Figure 4.12: $h, \psi, S(\psi), K(\psi)$ at $t = 0.150$ (from top to bottom), the dashed red line marks the water height in the uncoupled case

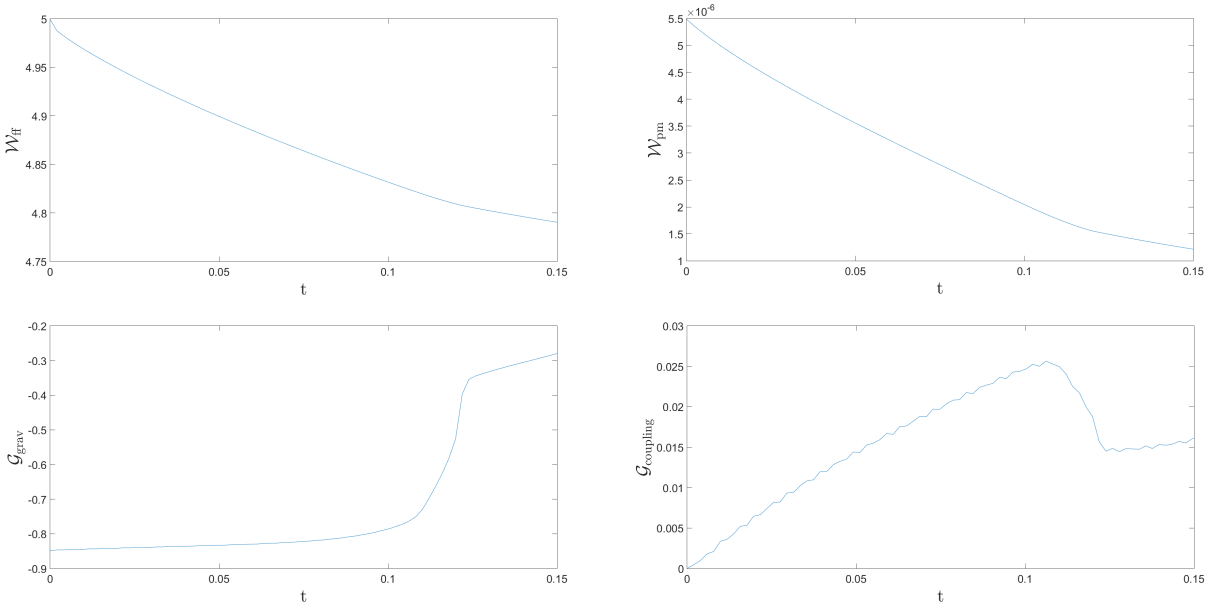


Figure 4.13: Energies $\mathcal{W}_{\text{ff}}$, $\mathcal{W}_{\text{pm}}$ and energy source terms $\mathcal{G}_{\text{grav}}$, $\mathcal{G}_{\text{coupling}}$

4.4 Numerical example for the coupled problem 3 (Realistic example)

This example models a dam break and the ensuing runoff and seepage into the subsurface. The simulation parameters are presented in Table 4.6. Initially, the water is impounded at a height of 10 for $x \in [0, 10]$, the subsurface beneath is fully saturated. On the right hand side of the dam, the bed is dry in the beginning, the groundwater level is at $z = -3$ and above, the subsurface is unsaturated. At $t = 0$ the dam is removed instantaneously on the surface as well as in the subsurface. At both surface boundaries no-flow conditions are imposed representing rock walls. The van Genuchten parameters are taken from [28] (Sample no. 2-3084) and correspond to sandy gravel in the Ringold Formation.
A numerical challenge in this example is the partially dry bed, i.e. $h = 0$. As reported e.g. in [16], the performance of Roe's solver is poor when the water height is taken as 0 since Roe's solver may produce negative water heights, which is why the LxF flux is employed in this simulation.
When the water height becomes small, inaccuracies arise at the computation of the velocity v by

$$v = \frac{hv}{h} \tag{4.14}$$

because of the small denominator. For this reason, a desingularization is applied as suggested in [32] and v is computed as

$$v = \frac{\sqrt{2}h(hv)}{\sqrt{h^4 + \max(h^4, \varepsilon)}}, \tag{4.15}$$

where $\varepsilon > 0$ is a small positive parameter which we choose as $\varepsilon = (\Delta x)^4$.
As regards the subsurface, a mixed linearization scheme is used: we start with 30 modified Picard iterations due to the higher robustness in comparison to Newton's method. If criterion (3.53) is not satisfied yet, 10 Newton iterations are carried out thereafter to exploit the quadratic order of convergence of Newton's method, and finally, L-scheme iterations are executed until convergence is achieved.
Wherever the water height becomes less than $\delta = 10^{-3}$, the Dirichlet coupling condition for the pressure in the subsurface (1.40) is replaced by a no-flow condition, i.e. $v_{\text{pm}} \cdot \vec{n} = 0$. This serves two purposes: it prevents negative water heights which would cause failure of the computation owing to imaginary values when calculating $\sqrt{gh}$, and furthermore, it permits negative pressure heights at the upper boundary of the subsurface without the boundary condition being violated.
The resulting water height and saturation at several time levels are depicted in Figures 4.14 and 4.15.

Geometry		
FF	$(\hat{x}_l, \hat{x}_r)$	$(0, 100)$
PM	(x_l, x_r)	$(0, 50)$
|	(z_b, z_t)	$(-5, 0)$
Initial conditions		
FF	$h_0(x)$	$\begin{cases} 10, & x \in [0, 10], \\ 0, & \text{else} \end{cases}$
|	v_0	0
PM	$\psi_0(x, z)$	$\begin{cases} 10 - z, & x \in [0, 10], \\ -3 - z, & \text{else} \end{cases}$
Van Genuchten parameters		
PM	α	0.97
|	θ_S	0.0579
|	θ_R	0.0125
|	n	1.57
|	K_S	$1.30 \cdot 10^{-3}$
Physical parameters		
	g	9.81
FF	q_I	0
Simulation time		
	T	15
Numerical parameters		
	c_{coupling}	1000
FF	Numerical flux	LxF
|	Δx	10^{-2}
PM	Solver	Mixed
|	No. of triangles	18000
|	$\varepsilon_{\text{abs}}, \varepsilon_{\text{rel}}$	10^{-5}

Table 4.6: Simulation parameters

One observes that the shock is gradually smoothed out and the dry/wet transition is characterized by a rarefaction wave. When the runoff reaches the right boundary, the water is reflected and a shock wave propagating in the opposite direction is forming.
The water is seeping into the subsurface by and by, and at the end of the simulation almost the entire subsurface is saturated.

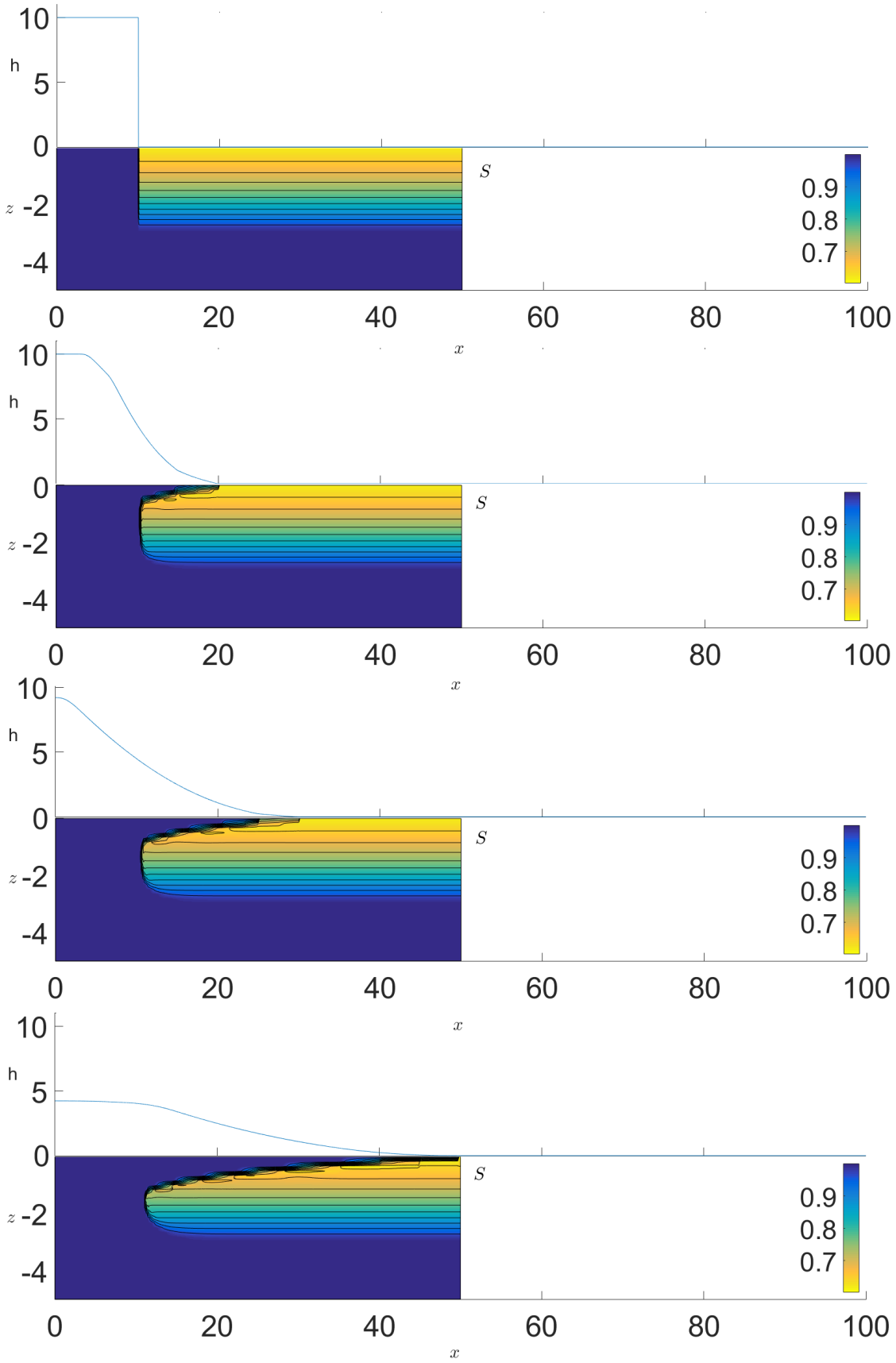


Figure 4.14: Water height and saturation at $t = 0, 0.5, 1, 2$ (from top to bottom)

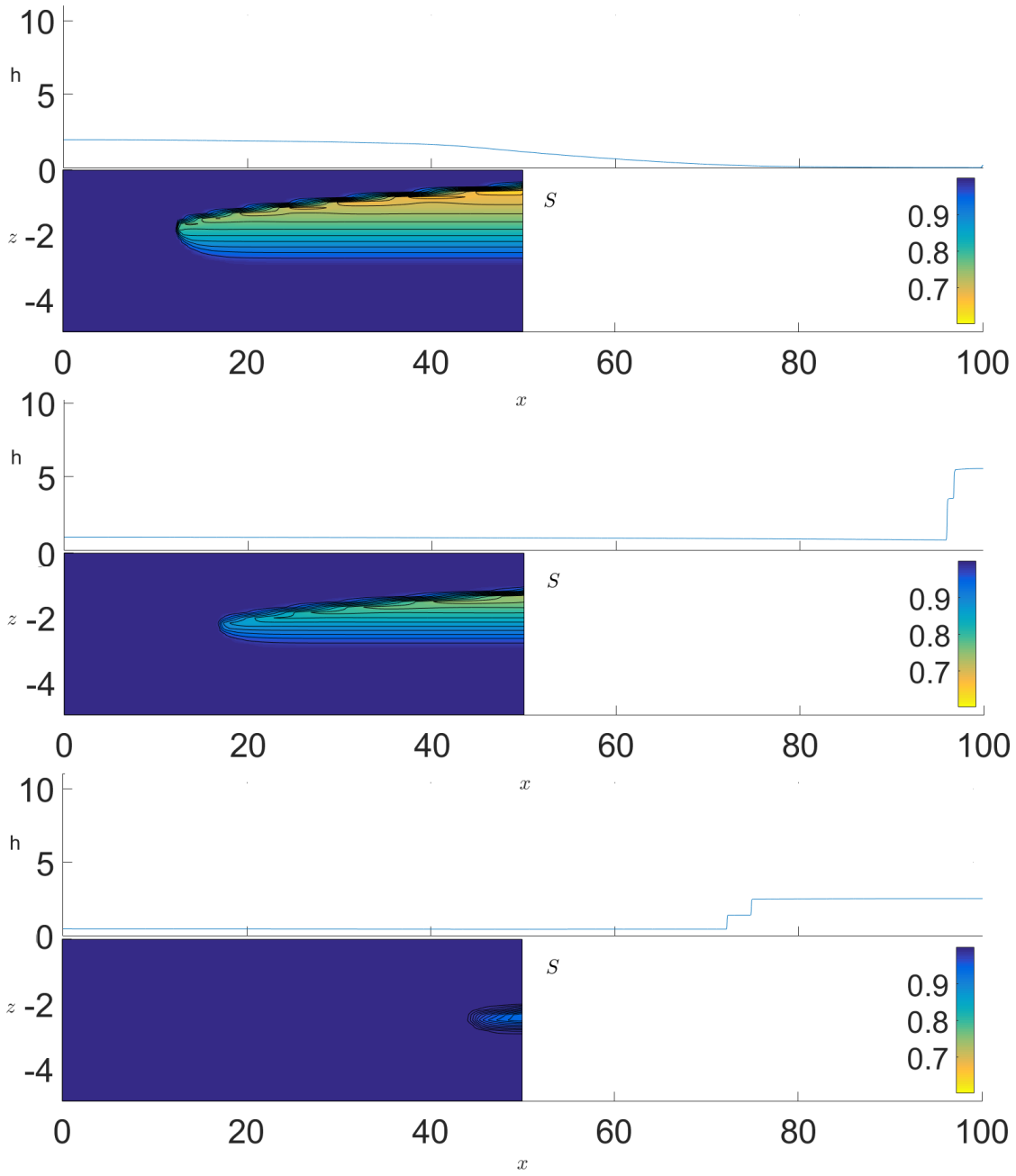


Figure 4.15: Water height and saturation at $t = 4, 8, 15$ (from top to bottom)

Chapter 5

Conclusions and outlook

In this thesis, a coupled mass-conservative surface-subsurface model has been formulated, based upon physical coupling conditions. For a simplified coupled model, an energy estimate has been proved. As regards numerical methods, the application of conservative finite volume schemes on the shallow water equations has been considered, in conjunction with an explicit Euler time discretization. For Richards' equation, linearization methods have been introduced and Galerkin FEM has been employed, along with an implicit Euler time discretization. We have developed an algorithm for solving the coupled surface-subsurface problem, which provides for the choice of a larger time step for the subsurface than for the surface. This is reasonable in consideration of the typically smaller velocities in the subsurface and allows to reduce the number of subsurface computations, which are considerably more expensive as compared to the surface computations owing to the non-linearity and the higher dimensionality of the subsurface. The analysis of the stability and convergence of the coupled algorithm is beyond the scope of this thesis, however, the numerical examples give very good results, also for an example with realistic simulation parameters.
Many extensions of this work are possible, one of the most interesting among those being a rigorous derivation of the coupled model via volume averaging from the Navier–Stokes equations, which could reveal further momentum coupling terms between surface and subsurface. Furthermore, analysis for the analytical solution of the coupled model as well as of the numerical solution computed with the coupling algorithm could be provided. For applications with highly permeable porous media, the coupled model can be extended to include dynamic capillary effects. Besides, complex bottom geometries can be considered, and simulations for two-dimensional surface flow and three-dimensional subsurface flow can be carried out. As to numerical issues, a multitude of spatial discretization methods can be applied on Richards' equation, high resolution schemes can be employed for the shallow water equations, and criteria for the choice of the ratio between subsurface and surface time step c_{coupling} can be determined.

List of Figures

List of Tables

Bibliography

[1] H. Alt and S. Luckhaus. Quasilinear elliptic-parabolic differential equations. *Math. Z.*, 183:311–341, 1983.

[2] T. Arbogast. An error analysis for Galerkin approximations to an equation of mixed elliptic-parabolic type. Technical Report TR90-33, Rice University, Houston, 1990.

[3] T. Arbogast, M. Obeyesekere, and M. Wheeler. Numerical methods for the simulation of flow in root-soil systems. *SIAM J. Numer. Anal.*, 30(6), 1993.

[4] J.-L. Auriault. On the domain of validity of Brinkman's equation. *Transport Porous Med.*, 79(2):215–223, 2009.

[5] J. Bear. *Dynamics of Fluids in Porous Media*. Dover Publ., 1988.

[6] H. Berninger, R. Kornhuber, and O. Sander. Fast and robust numerical solution of the Richards equation in homogeneous soil. *SIAM J. Numer. Anal.*, 49(6):2576–2597, 2011.

[7] K. Beven. *Rainfall-Runoff Modelling: The Primer*. John Wiley & Sons, 2012.

[8] M. Celia, E. Bouloutas, and R. Zarba. A general mass-conservative numerical solution for the unsaturated flow equation. *Water Resour. Res.*, 26(7):1483–1496, 1990.

[9] A. Cihan, J. Birkholzer, T. H. Illangasekare, and Q. Zhou. A modeling approach to represent hysteresis in capillary pressure-saturation relationship based on fluid connectivity in void space. *Water Resour. Res.*, 50(1):119–131, 2014.

[10] R. Courant, K. Friedrichs, and H. Lewy. Über die partiellen Differenzengleichungen der mathematischen Physik. *Math. Ann.*, 100(1):32–74, 1928.

[11] H. Darcy. *Les fontaines publiques de la ville de Dijon*. Dalmont, 1856.

[12] A. J. de Saint-Venant. Théorie du mouvement non-permanent des eaux avec application aux crues des rivières et à l'introduction des marées dans leur lit. *Comptes Rendus des Séances de l'Académie des Sciences*, 73:147–154, 1871.

[13] R. Eymard, M. Gutnic, and D. Hilhorst. The finite volume method for Richards equation. *Comput. Geosci.*, 3(3–4):259–294, 1999.

[14] R. A. Fine and F. J. Millero. Compressibility of water as a function of temperature and pressure. *J. Chem. Phys.*, 59(10):5529–5536, 1973.

[15] P. Gauckler. Etudes théoriques et pratiques sur l'écoulement et le mouvement des eaux. *Comptes Rendues de l'Académie des Sciences*, 64:818–822, 1867.

[16] D. L. George. Numerical Approximation of the Nonlinear Shallow Water Equations with Topography and Dry Beds: A Godunov-Type Scheme. Master's thesis, University of Washington, 2004.

[17] S. K. Godunov. A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. *Mat. Sb. (N.S.)*, 47(89)(3):271–306, 1959.

[18] W. G. Gray and C. T. Miller. *Introduction to the Thermodynamically Constrained Averaging Theory for Porous Medium Systems*. Springer, 2014.

[19] A. Harten and J. M. Hyman. Self-adjusting grid methods for one-dimensional hyperbolic conservation laws. *J. Comput. Phys.*, 50(2):235–269, 1983.

[20] S. Hassanizadeh, M. A. Celia, and H. Dahle. Dynamic effect in the capillary pressure-saturation relationship and its impacts on unsaturated flow. *Vadose Zone J.*, 1(1):37–57, 2002.

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

[22] R. Helmig and H. Class. *Fluidmechanik II*. Skript, Institut für Wasser- und Umweltsystemmodellierung, Universität Stuttgart, 2014.

[23] U. Hornung. *Homogenization and Porous Media*. Springer, 1997.

[24] P. Houston, C. Schwab, and E. Süli. Stabilized hp-finite element methods for first-order hyperbolic problems. *SIAM J. Numer. Anal.*, 37(5):1618–1643, 2000.

[25] C. Johnson. *Numerical Solution of Partial Differential Equations by the Finite Element Method*. Courier Corporation, 2012.

[26] C. Johnson, U. Nävert, and J. Pitkäranta. Finite element methods for linear hyperbolic problems. *Comput. Methods Appl. Math.*, 45(1–3):285–312, 1984.

[27] F.-M. Kalaydjian. Dynamic capillary pressure curve for water/oil displacement in porous media: Theory vs. experiment. In *SPE Annual Technical Converence and Exhibition*. Society of Petroleum Engineers, 1992.

[28] R. Khaleel and E. J. Freeman. Variability and scaling of hydraulic properties for 200 area soils, Hanford site. *U.S. Department of Energy*, 1995.

[29] S. Kim and W. B. Russel. Modelling of porous media by renormalization of the Stokes equations. *J. Fluid Mech.*, 154:269–286, 1985.

[30] S. J. Kollet and R. M. Maxwell. Integrated surface–groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model. *Adv. Water Res.*, 29:945–958, 2006.

[31] D. Kröner. *Numerical Schemes for Conservation Laws*. Wiley-Teubner, 1997.

[32] A. Kurganov and G. Petrova. A second-order well-balanced positivity preserving central-upwind scheme for the Saint-Venant system. *Commun. Math. Sci.*, 5(1), 2007.

[33] P. Lax and B. Wendroff. Systems of conservation laws. *Comm. Pure Appl. Math.*, 13:217–237, 1960.

[34] F. Lehmann and P. Ackerer. Comparison of iterative methods for improved solutions of the fluid flow equation in partially saturated porous media. *Transp. Porous Media*, 31:275–292, 1998.

[35] R. J. LeVeque. *Numerical Methods for Conservation Laws*. Birkhäuser, 1990.

[36] T. Lundgren. Slow flow through stationary random beds and suspensions of spheres. *J. Fluid Mech.*, 51(2):273–299, 1972.

[37] J. Magiera. *Coupling of Surface and Subsurface Flow*. Master's thesis, University of Stuttgart, 2014.

[38] R. Manning. On the flow of water in open channels and pipes. *Transactions of the Institution of Civil Engineers of Ireland*, 20:161–207, 1891.

[39] S. A. Moore. *Lateral Boundary Conditions in Numerical Ocean Models*. PhD thesis, University of Reading, 2004.

[40] R. H. Nochetto and C. Verdi. Approximation of degenerate parabolic problems using numerical integration. *SIAM J. Numer. Anal.*, 25(4):784–814, 1988.

[41] J. M. Nordbotten and M. A. Celia. *Geological Storage of CO2 - Modeling Approaches for Large-Scale Simulation*. John Wiley & Sons, 2011.

[42] I. Pop, F. Radu, and P. Knabner. Mixed finite elements for the Richards' equation: linearization procedure. *J. Comput. Appl. Math.*, 168(1):365–373, 2004.

[43] F. Radu. *Mixed finite element discretization of Richards' equation: error analysis and application to realistic infiltration problems*. PhD thesis, University of Erlangen-Nürnberg, Heidelberg, 2004.

[44] F. Radu, J. Nordbotten, I. Pop, and K. Kumar. A robust linearization scheme for finite volume based discretizations for simulation of two-phase flow in porous media. *J. Comput. Appl. Math.*, doi:10.1016/j.cam.2015.02.051, 2015.

[45] F. Radu, I. Pop, and P. Knabner. Order of convergence estimates for an Euler implicit, mixed finite element discretization of Richards' equation. *SIAM J. Numer. Anal.*, 42(4):1452–1478, 2004.

[46] F. Radu, I. Pop, and P. Knabner. Newton-type methods for the mixed finite element discretization of some degenerate parabolic equations. *Numerical Mathematics and Advanced Applications*, pages 1192–1200, 2006.

[47] F. Radu, I. Pop, and P. Knabner. Error estimates for a mixed finite element discretization of some degenerate parabolic equations. *Numer. Math.*, 109(2):285–311, 2008.

[48] L. A. Richards. Capillary conduction of liquids through porous mediums. *J. Appl. Phys.*, 1(5):318–333, 1931.

[49] P. Roe. Approximate Riemann solvers, parameter vectors, and difference schemes. *J. Comput. Phys.*, 43:357–372, 1981.

[50] I. Rybak and J. Magiera. A multiple-time-step technique for coupled free flow and porous medium systems. *J. Comput. Phys.*, 272:327–342, 2014.

[51] E. Schneid. *Hybrid-Gemischte Finite-Elemente-Diskretisierung der Richards-Gleichung*. PhD thesis, University of Erlangen-Nürnberg, 2000.

[52] E. Schneid, P. Knabner, and F. Radu. A priori error estimates for a mixed finite element discretization of the Richards' equation. *Numer. Math*, 98(2):353–370, 2004.

[53] M. Slodička. A robust and efficient linearization scheme for doubly nonlinear and degenerate parabolic problems arising in flow in porous media. *SIAM J. Sci. Comput.*, 23(5):1593–1614, 2002.

[54] P. Sochala, A. Ern, and S. Piperno. Mass conservative BDF-discontinuous Galerkin/explicit finite volume schemes for coupling subsurface and overland flows. *Comput. Methods Appl. Mech. Engrg.*, 198:2122–2136, 2009.

[55] J. J. Stoker. *Water Waves: The Mathematical Theory with Applications*. Interscience Publishers, 1957.

[56] J. Strikwerda. *Finite Difference Schemes and Partial Differential Equations*. Society for Industrial and Applied Mathematics, 2004.

[57] W. Tan. *Shallow Water Hydrodynamics - Mathematical Theory and Numerical Solution for a Two-dimensional System of Shallow-water Equations*. Elsevier, 1992.

[58] S. K. Teveldal. *Dynamic Capillary Effects in the Simulation of Flow and Transport in Porous Media: A New Linearisation Method*. Master's thesis, University of Bergen, 2014.

[59] M. T. van Genuchten. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.*, 44(5), 1980.

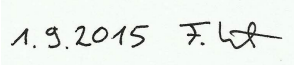
[60] B. van Leer. Towards the ultimate conservative difference scheme V. A second-order sequel to Godunov's method. *Comp. Phys.*, 32:101–136, 1978.

[61] C. Vreugdenhil. *Numerical Methods for Shallow-Water Flow*. Springer, 1994.

[62] S. Whitaker. *The Method of Volume Averaging*. Springer, 1999.

Declaration

I hereby declare that the work submitted is my own and that all passages and ideas that are not mine have been fully and properly acknowledged. Furthermore, I declare that this work has not been in parts or wholly published as a submission for another examination procedure and that all copies, both printed and electronic, are the same.



Date, F. List