Two-phase flow in porous media

F. Wessels

28th June 2013

Supervised by:
prof. dr. I. S. Pop
Abstract

Step by step a parabolic partial differential equation for two-phase flow in porous media is derived. Especially Darcy’s law and the mass balance law are being used in the derivation. A solution has been constructed for the limit case that the diffusion term is zero. This unique solution to this hyperbolic equation is constructed using a combination of shocks, rarefaction waves and constant solutions. This solution is compared with numerical approximations using different finite difference schemes like central, upwind and Lax-Friedrichs. Properties of the schemes are investigated, like whether they respect the mass balance law and are consistent and stable. Conclusion is that the forward in time, backward in space (upwind) scheme works very well. Finally, a finite difference scheme for the original partial differential equation with the diffusion term was made using iteration with only linear equations.
## Contents

1. **Introduction** .................................................. 1

2. **Derivation of the model** .................................. 2
   2.1 Darcy’s law ................................................. 2
   2.2 The rock permeability with a two-phase flow .......... 3
   2.3 Mass balance ............................................. 3
   2.4 The dimensionless model ................................. 5

3. **Solutions to** $u_t + (f(u))_x = 0$ .................... 7
   3.1 Introduction ............................................. 7
   3.2 Admissible shocks (travelling waves) ................ 7
   3.3 Rarefaction ............................................. 9
   3.4 Solution as combination ................................ 9

4. **Numerical approximations, Finite Difference Methods** 12
   4.1 Boundary conditions ................................... 12
   4.2 Discretization scheme tests and possible restrictions 13
     4.2.1 Stability check example with Fourier analysis for the case \( M = p = d = 1 \) 13
     4.2.2 Mass balance check example ..................... 14
   4.3 Explicit schemes for $u_t + (f(u))_x = 0$ ............ 15
     4.3.1 Forward in time and derivative of \( f \) times central difference 15
     4.3.2 Forward in time and space ...................... 16
     4.3.3 Lax-Friedrichs scheme ......................... 17
     4.3.4 Forward in time backward in space (upwind) .... 17
   4.4 Implicit on diffusion, explicit on convection ....... 20
     4.4.1 Integrate \( D(u) \) .................................. 20
     4.4.2 Chord method .................................... 21

Bibliography .................................................. 23
1 Introduction

The behavior of two non-solids under pressure in a porous medium is investigated. As an example water and oil are chosen. A porous medium is a material containing pores which can be filled with a liquid or gas. A typical example is sand, see Figure 1. We will make a simple model in one dimension, where at $t = 0$ the two non-solids, for example water and oil, are separated. There is a straight surface between the two liquids as seen in Figure 2. The effect of a liquid traveling through for example paper without external forces is called the capillary pressure. In our model we expect both oil and water to travel by this capillary pressure throughout the porous medium. Besides that, we assume a mixture of the two liquids is injected from the left. We first derive a partial differential equation. Next we try to solve this equation both analytically and numerically. An example can be an oil company which is extracting oil from the underground by injecting water in it [6]. A question could be how long it will take until it is not profitable anymore, or to find the optimized pressure which the water is being injected with. An example with gases is CO2 storage [5].
2 Derivation of the model

In this section we follow the procedure in [3], pp. 127-139. There are a lot of variables we will use while deriving the differential equation. Some will be explained first.

- **Porosity**: \( \phi \in (0, 1) := \) volume fraction of empty space of the porous soil.

- **Saturation**: \( S \in [0, 1] := \) fraction of the empty space filled with a liquid, where \( S_w \) is the water saturation and \( S_o \) is the oil saturation. In our model we assume there will be only water or oil in the gaps. This means \( S_w + S_o = 1 \). Furthermore, it is important to realise that \( S \) is dependent of the location \( x \). In every point \( x \), which can be considered as an infinitely small interval, the fraction will always be 1 or 0. So instead, we see \( S \) as a fraction of the particular liquid in a small interval, or in higher dimension, the fraction in a small area around \( x \).

- **Connate water saturation**: \( S_{wc} := \) maximum water saturation capillary trapped.

- **Residual oil saturation**: \( S_{or} := \) maximum oil saturation capillary trapped.

Because of the assumption of a maximum water and oil saturation, we rescale our model, defining

\[
S_{wd} = \frac{S_w - S_{wc}}{1 - S_{wc} - S_{or}},
\]

and

\[
S_{od} = \frac{S_o - S_{or}}{1 - S_{wc} - S_{or}},
\]

such that \( S_{wd}, S_{od} \in [0, 1] \).

2.1 Darcy’s law

When for example a liquid flows trough porous media due to a pressure, then the fluid will lose some of its pressure along the way. Think of it as a sort of friction of the fluid with the porous medium. Darcy’s law, published by Henry Darcy [1], says that the difference in pressure is directly proportional to the flux of the liquid. It is analogous to Ohm’s law in electrical networks, where the pressure difference can be seen as the potential difference and the flux is now the electric current. Due to the resistance of a wire in an electric network, the electric current will get smaller when taking a longer wire (of the same material and thickness). As the resistance of a wire you can imagine that Darcy’s law should be dependent on which liquid and which porous medium we have, i.e.,

\[
Q = -\frac{kA}{\mu} \frac{(P_b - P_a)}{L}.
\]

Here \( k \) denotes the rock permeability, \( \mu \) the fluid viscosity, \( L \) is the length, \( A \) is the cross section of the medium, \( Q \) denotes the total flux and \( (P_b - P_a) \) is the pressure difference.
where \( P_b < P_a \), see Figure 3. We would like to calculate the flow at every point \( x \). To do so we divide both sides in (3) by the area \( A \). The pressure difference between point \( a \) and \( b \) over the length \( L \) now becomes the gradient of \( P \),

\[
q = -\frac{k}{\mu} \nabla P.
\]  

(4)

2.2 The rock permeability with a two-phase flow

With homogenisation techniques it can be shown that Darcy’s law holds for a constant \( k \) which only depends on the type of soil. In this case there are two non-solids, water and oil. Water tends to draw into the corners which results in a slight reduction in the oil permeability. Oil forms bubbles in the middle of the pores which results in a more significant reduction in the relative water permeability. Thus \( k \) should also be dependent on the water and oil concentrations. In our model we will assume

\[
q_\alpha = -\frac{k_\alpha}{\mu_\alpha} k \text{ grad} P_\alpha, \ \alpha \in \{w, o\}, \quad (5)
\]

where \( k_w = k_w(S_{wd}) \) and \( k_o = k_o(S_{od}) \), functions of the rescaled water and oil saturations, see equation (1) and (2). We expect that if the porous medium is fully saturated with water, equation (4) should be equal to (5), thus \( k \) does not depend on \( S_{wd} \). The same holds for oil, such that \( k_w(S_{wc}) = k_o(S_{or}) = 1 \). Examples for the functions \( k_w \) and \( k_o \) are:

\[
k_w(S_{wd}) = S_{wd}^p, \\
k_o(S_{od}) = S_{od}^d = (1 - S_{wd})^d.
\]  

(6)

Here \( p \) and \( d \) usually varies between 1.5 and 2.5. With experiments a more accurate and more complicated function can be fit, also for other non-solids.

2.3 Mass balance

Next a conservation equation is derived, the mass balance equation. Let \( V \) be an arbitrary volume in an open space, enclosed by surface \( S \). The mass balance equation is nothing else then using the fact that the change of mass per unit time in \( V \) is equal to the total mass

Figure 3: Graphical representation of Darcy’s law
that crosses surface $S$ per unit time. This is an assumption that can be made, because the only way to lose or gain mass besides when mass crosses $S$, is by forming new atoms which is assumed not to happen. We get

$$\frac{d}{dt} \int_V \rho'_\alpha \phi S_{\alpha d} \, dV + \int_S \rho'_\alpha q'_\alpha \cdot \mathbf{n} \, dS = 0, \quad \alpha \in \{w,o\},$$

(7)

where $\rho'_\alpha$ is the density, mass per unit rock volume, and $q'_\alpha$ the flux of fluid per unit rock surface. To make life easier we can write $\rho'_\alpha = \rho_\alpha \phi S_{\alpha d}$ and $q'_\alpha = q_\alpha / \phi S_{\alpha d}$. Here $\rho_\alpha$ is the density of the fluid, and $q_\alpha$ is the discharge of the fluid. Next, the divergence theorem can be used on the second integral in (7).

$$\frac{d}{dt} \int_V \rho_\alpha \phi S_{\alpha d} \, dV + \int_V \text{div}(\rho_\alpha q_\alpha) \, dV = 0, \quad \alpha \in \{w,o\}. \tag{8}$$

Also is assumed that volume $V$ does not depend on time $t$, so the time derivative can be placed inside the integral. Since the mass balance equation would hold for any arbitrary volume $V$, we can write

$$\frac{\partial}{\partial t} (\rho_\alpha \phi S_{\alpha d}) + \text{div}(\rho_\alpha q_\alpha) = 0, \quad \alpha \in \{w,o\}. \tag{9}$$

We assume the porosity is constant and the fluids are incompressible, thus a constant density, which leads to

$$\phi \frac{\partial S_{\alpha d}}{\partial t} + \text{div} q_\alpha = 0, \quad \alpha \in \{w,o\}. \tag{10}$$

Because we look only in $1D$, $\text{div} q_\alpha = \frac{\partial}{\partial x} q_\alpha$, thus we get

$$\phi \frac{\partial S_{\alpha d}}{\partial t} + \frac{\partial}{\partial x} q_\alpha = 0, \quad \alpha \in \{w,o\}. \tag{11}$$
Known is that $S_{wd} + S_{od} = 1$, thus $\phi \frac{\partial S_{wd}}{\partial t} + S_{od} = 0$. This means $\frac{\partial}{\partial x}(q_w + q_o) = 0$. Let $q = q_w + q_o$, then $q = q(t)$, not depending on $x$. If fluid is injected from the left with constant rate $q_{inj}$ we get

$$q_w + q_o = q_{inj}. \quad (12)$$

Finally, we need the capillary pressure $P_c$, the average pressure difference between oil and water, thus

$$P_c = P_o - P_w. \quad (13)$$

We will make the assumption that

$$P_c = P_c(S_{wd}) = \sigma \sqrt{\frac{\phi}{k} J(S_{wd})}, \quad (14)$$

which is gained from experiments. Here $\sigma$ is a constant, which denotes the interfacial tension between the two liquids in the pores. $J(S_{wd})$ is a strictly decreasing function. We assume $J$ to be a function of the form

$$J(S_{wd}) = -S_{wd}^\lambda, \quad (15)$$

where $0 < \lambda$. Other functions for $J$ are possible, but generally $J$ is a strictly decreasing function of $S_{wd}$.

We have $\phi \frac{\partial S_{wd}}{\partial t} + \text{div} q_w = 0$. Now $q_w$ will be expressed in terms of $S_{wd}$. We know

$$q_o = -\frac{k_o}{\mu_o} k_o \frac{\partial P_o}{\partial x} = -\frac{k_o}{\mu_o} k_o \frac{\partial P_c}{\partial x} - \frac{k_o}{\mu_o} k \frac{\partial P_w}{\partial x} = -\frac{k_o}{\mu_o} k_o \frac{\partial P_c}{\partial x} + \frac{k_o}{\mu_o} k w \frac{\partial P_w}{\partial x} \quad (16)$$

using first (5), then (13) and finally (5) again to write the last term in terms of $q_w$. After substitution into (12) we get

$$q_w = \frac{k_w}{k_w + k_o} q_{inj} + \frac{k_w}{k_o} \frac{k_o}{\mu_o} k_o \frac{\partial P_c}{\partial x}. \quad (17)$$

Now (17) will be substituted into (11) where we choose $\alpha = w$, which results in

$$\phi \frac{\partial S_{wd}}{\partial t} + \frac{\partial}{\partial x} \left( \frac{k_w}{k_w + k_o} q_{inj} + \frac{k_w}{k_o} \frac{k_o}{\mu_o} k w \frac{\partial P_c}{\partial x} \right) = 0. \quad (18)$$

### 2.4 The dimensionless model

We now define the constants $M$ and $\varepsilon$ and make (17) dimensionless by dividing $x$ with a reference length $L$. If we have for example an oil field with a length of 20 kilometer, then $L$ is 10 kilometer such that the interval length that we used is 2, $[-1, 1]$. To make time
dimensionless, we multiply it with the flux divided by the length $L$ (of the pores). Thus we have the following dimensionless variables:

$$
\hat{x} = \frac{x}{L},
$$

$$
\hat{t} = \frac{q_{\text{inj}}}{\Phi L} t,
$$

$$
M := \frac{\mu_o}{\mu_w},
$$

$$
\varepsilon := \frac{\sigma \sqrt{k \phi}}{q_{\text{inj}} \mu_o L}.
$$

(19)

For notational convenience we set $u = S_{wd}$, $x = \hat{x}$ and $t = \hat{t}$ to get the final differential equation

$$
\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) - \varepsilon \frac{\partial}{\partial x} \left( D(u) \frac{\partial u}{\partial x} \right) = 0.
$$

(20)

Choosing $k_w(u) = u^p$ and $k_o = (1 - u)^d$ gives

$$
f(u) = \frac{M u^p}{(1 - u)^d + M u^p},
$$

(21)

$$
D(u) = f(u)(1 - u)^d \frac{dJ}{du},
$$

(22)

$$
J(u) = u^\lambda.
$$

(23)

Of course, other choices for $k_w$ and $k_o$ are possible, but for the calculations later on, $u^p$ and $(1 - u)^d$ are chosen. See Figure 7 for the function $f(u)$ with the chosen functions for $k_w$ and $k_o$.

![Figure 5: $f(u)$ plotted for different values of $p = d$ and $M$](image-url)
3 Solutions to $u_t + (f(u))_x = 0$

3.1 Introduction

In many practical situations, the capillary pressure term is very small. The main reason for this is due to equation (19), where we divide by the length. Usually the length $L$ is very large. In this section we first consider $\varepsilon$ to be 0, so there is no capillary pressure. We then get the differential equation

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

where

$$f(u) = \frac{Mu^p}{(1-u)^d + Mu^p}. \quad (25)$$

How will the solutions look like, are there solutions for every $M$, $d$ and $p$ and will a solution to this problem look like the solution of the differential equation with the capillary pressure term?

3.2 Admissible shocks (travelling waves)

Shocks are discontinuous weak solutions to the problem, where $u$ jumps from one value to another, and this jump travels in time with a certain speed. In nature or physics, there is often also a diffusion term present, very small, just like in the original problem. By adding such a diffusion term some constraints of a possible shock can be derived. Let us define $u_l$ and $u_r$ as the left and right limit at the shock. We start off with

$$u_t + (f(u))_x = \nu u_{xx}, \quad (26)$$

where $\nu > 0$, $u(-\infty, t) = u_l$ and $u(+\infty, t) = u_r$. An assumption that $u(x, t) = v(\eta)$, where $\eta = \frac{x-ct}{\nu}$ will be made. Substituting $v$ gives

$$-\frac{c}{\nu} v' + \frac{1}{\nu} (f(v))' = \frac{\nu}{\nu^2} v'' \Rightarrow -cv' + (f(v))' = v'',$$

where $\nu$ disappeared. Integrations gives

$$-cv + f(v) = v' + A. \quad (28)$$

Here $A \in \mathbb{R}$ is a constant. In the limit we get for $v$: $u(-\infty, t) = v(-\infty) = u_l$ and $u(+\infty, t) = v(+\infty) = u_r$. Next the following theorem will be applied.

**Theorem 1.** Assume that the function $v : \mathbb{R} \rightarrow \mathbb{R}$ is differentiable and that both $v(x)$ and $v'(x)$ have a limit as $x$ goes to infinity. Then $\lim_{x \rightarrow \infty} v'(x) = 0.$
Proof. Assume $\lim_{x \to \infty} v'(x) = \tilde{C} \neq 0$, let us choose $\tilde{C} > 0$ without loss of generality. Then there exists an $M_1 > 0$ s.t. for all $x > M_1$, $v'(x) > \frac{1}{2}\tilde{C} > 0$.

Also given is that $\lim_{x \to \infty} v(x) = \hat{C}$ for a $\hat{C} \in \mathbb{R}$. Thus for all $\varepsilon > 0$ there exists an $M_2 > 0$ s.t. for all $x > M_2$,

\[
|v(x + 1) - v(x)| = |v(x + 1) - \hat{C} + \hat{C} - v(x)| \leq |v(x + 1) - \hat{C}| + |\hat{C} - v(x)| < 2\varepsilon = \frac{2}{5}\tilde{C} \text{ and } |v(x + 1) - v(x)| = |v'(z)| > \frac{1}{2}\tilde{C}.
\]

This can not both be true, thus $\tilde{C} = 0$.

Thus for smooth enough $v$, $v'(\pm \infty) = 0$ which leads to

\[
-cu_t + f(u_t) = A \text{ and } -cu_r + f(u_r) = A. \tag{29}
\]

From this we can see that $A = 0$. For the traveling wave speed $c$ we get

\[
c = \frac{f(u_r) - f(u_l)}{u_r - u_l}, \tag{30}
\]

where the traveling wave speed does not depend on $\nu$. What remains is the first order equation,

\[
v' = f(v) - f(u_l) - c(v - u_l). \tag{31}
\]

For this initial value problem it is proven, in [2], pp. 24-31, that there exists a unique solution to this problem, which is in our case strictly decreasing. It implies that $|\frac{f(v) - f(u_l)}{v - u_l}| > c$.

![Figure 6: Possible shocks plotted in red for $u_r = 0$ plotted against $f(u)$ for $p = d = 2$ and $M = 1$](image)

Thus if we assume $u_r = 0$, there exists a shock solution connecting $u = \alpha$ to $u = 0$ where $\alpha$ is the solution of $uf'(u) = f(u)$. If for example $d = p = 2$ and $M = 1$ we get $\alpha = \frac{1}{2} \sqrt{2} \approx 0.707$, see Figure 6.
3.3 Rarefaction

In this section we will try to solve (24) differently using rarefaction waves. Let \( \hat{u}(x,t) \) be a solution to (24). Because for all \( k \in \mathbb{N} \), \( \frac{\delta u(kx,kt)}{\delta (kx)} = k \frac{\delta u(kx,kt)}{\delta (kt)} = k \frac{\delta u(kx,kt)}{\delta t} \), \( \hat{u}(kx,kt) \) is also a solution. Thus for all \( k \in \mathbb{N} \), \( \hat{u}(x,\frac{1}{k}) = \hat{u}(kx,1) \). Consequently \( u \) must be of the form \( u(x,t) = r(\eta) \) with \( \eta = \frac{x}{t} \). Substituting \( r \) for \( u \) in (24) gives

\[
(-\eta + f'(r))r' = 0,
\]

with boundary conditions \( r(-\infty) = 1 \) and \( r(\infty) = 0 \). This means that either \( r' = 0 \) or \( f'(r) = \eta \). From the first case can be obtained that \( r(\eta) \) is constant and thus from this constant solutions are obtained. From the second case we get that

\[
\eta = f'(\eta) = \frac{M(1-\eta)^{d-1}\eta^{p-1}(p(1-u)+du)}{((1-\eta)^d+M\eta^p)^2}.
\]

Figure 7: \( r(\eta) \) plotted for different values of \( d = p \) and \( M \)

3.4 Solution as combination

In the previous sections three different solutions has been discovered. Shocks, rarefaction waves and constant solutions. The speed of a rarefaction wave can be calculated in different ways. The speed is by definition equal to \( \frac{dx}{dt} \). If \( \hat{u}(x(t),t) \) is a solution of the partial differential equation, then

\[
\frac{d}{dt}\hat{u}(x(t),t) = \hat{u}_t + \hat{u}_x \frac{dx}{dt},
\]

(34)
from which can be seen that the speed $\frac{dx}{dt}$ is equal to $f'(\hat{u})$. Since for a rarefraction wave it holds that $\hat{u} = r(\eta) = (f')^{-1}(\eta)$ it follows that the speed is equal to

$$f'(f'(\eta)^{-1}) = \eta = \frac{x}{t}. \quad (35)$$

Another way to calculate the speed is by realizing that $x(t) = \eta t$ and thus that $\frac{dx}{dt} = \eta$.

The speed of the shock is somewhat harder to calculate. A shock happens just in a point, thus it is not possible to just substitute the shock solution into $f'(u)$. Fortunately the shock speed can be derived from the mass balance law. Let us look at an interval $(a, b)$ in which a shock travels in a period of time. Besides the shock the solution is just a constant solution, with value $u_l$ left of the shock and $u_r$ right of the shock. Then the mass balance law can be applied to the interval $(a, b)$. In words, if the shock is not fast enough, mass piles up before the shock. This can not happen because of the assumption that the porous medium is always fully saturated. Same argument for when the shock is too fast. In formula, the mass balance law can be written as

$$\frac{d}{dt} \int_a^b u(x,t) \, dx = f(a) - f(b). \quad (36)$$

The first part can be rewritten into

$$\frac{d}{dt} \int_a^b u(x,t) \, dx = \frac{d}{dt} \int_a^{x(t)} u(x,t) \, dx + \frac{d}{dt} \int_{x(t)}^b u(x,t) \, dx, \quad (37)$$

where $x(t)$ is the location of the shock. Next the $\frac{d}{dt}$ can be moved inside the integral,

$$\frac{d}{dt} \int_a^{x(t)} u(x,t) \, dx = \frac{d}{dt} \int_{x(t)}^b u(x,t) \, dx = \int_a^{x(t)} \frac{\partial}{\partial t} u(x,t) \, dx + \int_{x(t)}^b \frac{\partial}{\partial x} u(x,t) \, dx + x'(t)(u(x(t) - 0, t) - u(u(x(t) + 0, t)). \quad (38)$$

The term $x(t) - 0$ means the shock from the left en $x(t) + 0$ means the shock from the right, we get $x'(t)(u(x(t) - 0, t) - u(u(x(t) + 0, t)) = x'(t)(u_l, t) - u(u_r, t)).$ The integral $\int_a^{x(t)} \frac{\partial}{\partial x} u(x,t) \, dx$ can be changed into

$$\int_a^{x(t)} \frac{\partial}{\partial x} f(u) \, dx = -(f(u_l) - f(a)), \quad (39)$$

using $u_t = -(f(u))_x$. Doing this also for the right part, and substitute this into (36) results in

$$f(u_r) - f(u_l) + f(a) - f(b) + x'(t)(u_l - u_r) = f(a) - f(b). \quad (40)$$
From this the shock speed can be derived,

\[ \frac{f(u_l) - f(u_r)}{u_l - u_r} = x'(t). \]  \hspace{1cm} (41)

In our situation \( u_r = 0 \) and \( u_l \) can be at most \( \alpha \).

The speed of the rarefraction wave is given by \( f'(u) \) and the speed of the shock is given by \( \frac{f(u)}{u} \) and we have already seen that at exactly one point the shock speed and the rarefraction speed coincide, at \( \alpha \)! In Figure 8 a solution for \( M = 1, \ d = p = 2 \) is plotted. Here \( \alpha = \frac{1}{2} \sqrt{2} \approx 0.707 \).

![Figure 8: Combination of rarefraction waves with shocks plotted for \( d = p = 2, \ M = 1 \) and \( i \in \{1, 2, ..., 10\} \)](image)
4 Numerical approximations, Finite Difference Methods

Some notations when using discretization schemes. For time the parameter $l$ is used as a super index, and for $x$, $k$ is used as a sub index. All discretizations use $\Delta t = \frac{1}{m-1}$ and $\Delta x = \frac{1}{n-1}$, where $m$ is the amount of grid point on the time-axes, and $n$ the amount of time steps on the $x$-axe. The discretized area is denote by $[0, T] \times [a, b]$ where $a, b, T \in \mathbb{R}$, $a < b$ and $T > 0$. So the partial differential equation $\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0$, using the explicit upwind scheme will look like

$$\frac{u_{k}^{l+1} - u_{k}^{l}}{\Delta t} + \frac{u_{k}^{l} - u_{k-1}^{l}}{\Delta x} = 0, \quad (42)$$

which is a simple case of the derived partial differential equation section 3, when $\varepsilon = 0$ and $p = d = M = 1$.

4.1 Boundary conditions

What to take for the boundary conditions and until which $T$ and which $x$ will we go? We already assumed that $u(x, t) = 1$ for $x < 0$ and $u(x, t) = 0$ for $x > 0$. It is $1 - H(x)$ where $H(x)$ is the Heaviside-function. Thus in our discretization scheme, $u_{k}^{0} = 1$ for $k < \frac{n}{2}$ and $u_{k}^{0} = 0$ for $k > \frac{n}{2}$. If $n$ is even, then at 0 we define, $u_{\frac{n}{2}}^{0} := \frac{1}{2}$.

We choose $t$ to go until 0.5 and $x \in [-1, 1]$, because then in all schemes that where tried so far, the solution in time did not yet change at the boundary $(-1, t)$ and $(1, t)$. This is also the reason we choose Dirichlet boundary conditions, because we expect no changes for all $t \in [0, 0.5]$. Thus for all $l \in \{0, 1, ..., m\}$, $u_{0}^{l} = 1$ and $u_{n}^{l} = 0$, see Figure 9.

![Figure 9: Discretization, $m = 3$ and $n = 8)](image-url)
4.2 Discretization scheme tests and possible restrictions

To be sure a discretization scheme works, consistency, stability and convergence have to be checked. Consistency means that the discretization scheme converges to the actual partial differential equation when taking the limit \((\Delta x, \Delta t) \to (0, 0)\). Often Taylor approximations are used in order to prove this. Stability means the differential operator does not amplify noise, roundoff-errors do not grow. This can be checked using Fourier analysis. Convergence means that the numerical solution approximates the real solution for small \(\Delta x\) and \(\Delta t\). If a scheme is consistent and stable, then under conditions it converges. Unfortunately one of those conditions is that the solution is differentiable. In Section 3 we have seen that the solution to (24) is not differentiable. Also we have a non-differentiable initial value at \(t = 0\), namely \(1 - H(x)\), where \(H(x)\) is the Heaviside function.

Fortunately from a physical point of view, there are more restrictions to add to the discretization schemes. An example is that already is known that the solution \(u\) can not be negative, while the discretized solution is negative. To overcome this, it is sometimes relevant to use lower order schemes. Also it is possible to test if the scheme is consistent with a balance law. The two liquids can not just disappear, so the discretization scheme should be consistent with the mass balance law. In this case \(u\) is representing the water concentration. Earlier the assumption was made that the liquids are incompressible. Thus to calculate the mass from concentration done by multiplying it with the density, \(\Phi\) and a scaling factor. All three are constants, thus if the balance law does not hold for the concentration, it also does not odd for the mass.

4.2.1 Stability check example with Fourier analysis for the case \(M = p = d = 1\)

If \(M = p = d = 1\), \(f(u)\) simply reduces to \(f(u) = u\). In this case we have a linear partial differential equation \(u_t + u_x = 0\) which is a model equation for, for instance, the 1D Euler equations describing gas flow in a shock tube. A forward in time and central difference in space discretization scheme is used as an example.

\[
\frac{u_{k+1}^l - u_k^l}{\Delta t} + \frac{u_{k+1}^l - u_k^{l-1}}{2\Delta x} = 0. \tag{43}
\]

The error is defined as \(\epsilon_k^l = u(x_k, t_l) - u_k^l\), where \(u_k^l\) is the discrete solution and \(u(x_k, t_l)\) is the exact solution in point \((x_k, t_l)\). This results in

\[
\epsilon_{k+1}^l - \epsilon_k^l + \frac{\Delta t}{2\Delta x} (\epsilon_{k+1}^l - \epsilon_{k-1}^l) = d_k^l, \tag{44}
\]

where \(d_k^l\) is the local discretization error that can be found using Taylor approximations. The scheme is consistent when \(d_k^l\) goes to zero for small \(\Delta x\) and \(\Delta t\). Now all we have to show is that the left part is bounded. Next the assumption is made that

\[
\epsilon(x, t) = \sum_{j=1}^{J} c^j e^{i\frac{2\pi j}{L} x}, \tag{45}
\]
where \( c \) is a constant and \( L \) is the interval length, which can be assumed to be 1. Since the differential equation is linear, it is allowed to only check for one random \( j \) what happens with the round-off error. Substituting \( \epsilon(x,t) = e^{ct} e^{i\pi j x} \) into the left part of (44) gives

\[
e^{c\Delta t} = 1 - \frac{\Delta t}{2\Delta x} (e^{i\pi j \Delta x} - e^{-i\pi j \Delta x}) = 1 - \frac{i\Delta t}{\Delta x} \sin(\pi j \Delta x),
\]

where it is used that \( \epsilon_{k+1}^{l+1} = e^{c(t+\Delta t)} e^{i\pi j x} \), \( \epsilon_{k+1}^{l} = e^{ct} e^{i\pi j (x+\Delta x)} \) and \( \epsilon_{k+1}^{l} = e^{ct} e^{i\pi j (x-\Delta x)} \).

Defining the amplification factor as

\[
G := \frac{\epsilon_{k+1}^{l+1}}{\epsilon_{k}^{l}} = e^{c\Delta t},
\]

we should have that \( |G| \leq 1 \). Since \( |e^{c\Delta t}| = |1 - \frac{i\Delta t}{\Delta x} \sin(\pi j \Delta x)| \) we get that for any \( \Delta t \), there is a \( j \in \mathbb{N} \) such that \( |G| > 1 \), thus the finite difference scheme from (43) will always amplify the errors.

### 4.2.2 Mass balance check example

If a numerical scheme respects the mass balance, then the scheme should not add mass. Let us consider the example where in an interval \((a,b)\) there is some mass, but at the boundary, point \(a\) and \(b\), there is an equal flux for all \( t \in (0,T) \), where \( t \) represents time. Then the total mass should be the same for each \( t \). Thus we get

\[
u_t + (f(u))_x = 0, f(a) - f(b) = 0.
\]

Integrating over \((a,b)\) gives

\[
\int_{a}^{b} \frac{\partial}{\partial t} u \, dx + \int_{a}^{b} (f(u))_x \, dx = 0.
\]

Since the interval \((a,b)\) is not dependent on time, the time derivative can be moved outside the integral,

\[
\frac{d}{dt} \int_{a}^{b} u \, dx + f(b) - f(a) = 0.
\]

Since \( f(a) - f(b) = 0 \) we get

\[
\int_{a}^{b} u \, dx = c \in \mathbb{R}.
\]

Now looking at a forward in time explicit numerical scheme, we get

\[
u_{k+1}^{l} = u_{k}^{l} - \Delta t L(u_{1}^{l}, \ldots, u_{n}^{l}, \Delta x) = 0,
\]

where \( L \) represents a random difference scheme for \((f(u))_x\). Since \( \sum_{k=1}^{n} u_{k}^{l+1} = \sum_{k=1}^{n} u_{k}^{l} \) it follows that

\[
\sum_{k=1}^{n} L(u_{1}^{l}, \ldots, u_{n}^{l}, \Delta x) = 0.
\]
As example the central difference scheme is considered, \( L = \frac{f(u_{k+1}^l) - f(u_{k-1}^l)}{2\Delta x} \). It not hard to see that \( \sum_{k=1}^{n} L = f(u_n^l) - f(u_1^l) = f(a) = f(b) = 0 \). The conclusion is that the central difference scheme for \( u_x \) in this case respects the mass balance law.

4.3 Explicit schemes for \( u_t + (f(u))_x = 0 \)

First equation 20 is considered only now without the diffusion part. The reason for this is that a solution to this partial differential equation is found in section 3 and thus that it is possible to compare the real solution with a numerical solution. Consider

\[
\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = 0,
\]

where \( f(u) = \frac{M u^p}{(1-u)^d + M u^p} \). We notice that \( f(u) \) is a positive function from \([0,1]\) to \([0,1]\), see Figure 7. We will now test some finite difference schemes, [7] Ch. 13, and compare them with the real solution.

4.3.1 Forward in time and derivative of \( f \) times central difference

Because \( f \) is differentiable for all \( M, d \) and \( p \), the differential equation \( u_t + (f(u))_x = 0 \) can be rewritten as

\[
u_t + f'(u)u_x = 0.
\]

A simple numerical approach for this equation is

\[
\frac{u_{k+1}^l - u_k^l}{\Delta t} + f'(u_k^l) \frac{u_{k+1}^l - u_{k-1}^l}{2\Delta x} = 0.
\]

In this explicit scheme central difference is used for \( u_x \). So to calculate \( u_{k+1}^l \) we move everything else to the right. The results can be seen in Figure 10. Clearly this is not the

Figure 10: Results of a forward in time non-linear in space scheme, choosing \( n = m = 513 \) for \( p = d = 2 \) and \( M = 1 \)
result wanted. To check if this scheme is consistent we use Taylor approximations. We have 
\[ \frac{u_{k+1}^l - u_k^l}{\Delta x} = (u_k^l)_x + \mathcal{O}((\Delta x)^2) \] 
and 
\[ \frac{u_{k+1}^l - u_k^l}{\Delta t} = (u_k^l)_t + \mathcal{O}(\Delta t). \]
Because \( f' \) is a nice smooth function, this means this scheme is order 1 in time and order 2 in space consistent.

To check stability is in this case hard because for \( p = d = 2 \), \( f(u) \) is not linear. The case \( p = d = M = 1 \) is already calculated in section 4.2.1 where the conclusion was that central difference is not reducing the round-off error. Still the result in Figure 10 is so bad that it should have another cause.

To check whether this discretization scheme is consistent with the mass balance law, summing the scheme over \( k \) over every \( u_k^l \) is enough. Since also \( f'(u) \) is non-linear, it can be seen immediately that 
\[ \sum_{k=1}^n f'(u_k^l)(u_{k+1}^l - u_k^l) \neq L(u_1^l, u_n^l). \] Thus the scheme is really not consistent with the mass balance law!

Looking close to what happens, it can be seen that the solution stays exactly the same over time. The reason for this is that 
\[ f'(1) = f'(0) = 0, \] 
thus 
\[ f'(u_k^l)(u_{k+1}^l - u_k^l) = 0 \forall k \in \{1, 2, ..., n\}. \]

### 4.3.2 Forward in time and space

We now consider a forward scheme in both space and time. The result can be seen in Figure 11.

\[
\frac{u_{k+1}^l - u_k^l}{\Delta t} + \frac{f(u_{k+1}^l) - f(u_k^l)}{\Delta x} = 0
\]

\[ (57) \]

Figure 11: Results of the forward in time and space scheme, choosing \( n = m = 33 \) for \( p = d = 2 \) and \( M = 1 \)

This is also a result far from the expected solution, calculated in section 3. The scheme
is consistent and even respects the mass balance. What happens is that the mass coming from the left piles up at \( x = 0 \). Let us check on stability following the same steps as in section 4.2.1, for the case \( d = p = M = 1 \), only now with a forward in space scheme.

\[
\epsilon_{k+1}^l = \epsilon_k^l - \frac{\Delta t}{\Delta x} (\epsilon_{k+1}^l - \epsilon_k^l). \tag{58}
\]

Substituting \( \epsilon(x,t) = e^{ct}e^{i\pi jx} \) into (58) gives

\[
e^{c\Delta t} = 1 + \frac{\Delta t}{\Delta x} (1 - e^{i\pi j\Delta x}). \tag{59}
\]

Rewriting this as

\[
e^{c\Delta t} = 1 + 2\frac{\Delta t}{\Delta x} \sin^2\left(\frac{\pi j\Delta x}{2}\right) - 2i \frac{\Delta t}{\Delta x} \sin\left(\frac{\pi j\Delta x}{2}\right) \cos\left(\frac{\pi j\Delta x}{2}\right), \tag{60}
\]

leads to

\[
e^{c\Delta t} = \sqrt{1 + 4\frac{\Delta t}{\Delta x} (1 + \frac{\Delta t}{\Delta x}) \sin^2\left(\frac{\pi j\Delta x}{2}\right)}. \tag{61}
\]

This means \(|e^{c\Delta t}| > 1\) and thus this scheme is not stable at all!

### 4.3.3 Lax-Friedrichs scheme

The Lax-Friedrich scheme is given by

\[
\frac{u_{k+1}^l - \frac{1}{2}(u_{k+1}^l - u_{k-1}^l)}{\Delta t} + \frac{f(u_{k+1}^l) - f(u_{k}^l)}{\Delta x} = 0. \tag{62}
\]

Using Taylor approximations for the consistency check we get that this system converges with \( O((\Delta t)^2, (\Delta x)^4) \). Doing stability analysis following the same steps as in section 4.2.1, we get that it is stable for \(-1 \leq \frac{\Delta t}{\Delta x} \leq 1\). In Figure 12 the results are plotted. With this low resolution it is easy to see the diffusion error, both at the shock as at \( x = 0 \). The next scheme will even have better results.

### 4.3.4 Forward in time backward in space (upwind)

This time a forward in time and backward in space scheme is used. The speed, \( f'(u) \), is positive for all \( u \in (0,1) \). This means the flow is coming from the left. This scheme uses \( f'(u_{k-1}^l) \), information from where the flow is coming from, thus this scheme is called upwind,

\[
\frac{u_{k+1}^l - u_{k}^l}{\Delta t} + \frac{f(u_{k}^l) - f(u_{k-1}^l)}{\Delta x} = 0. \tag{63}
\]

In Figure 13, the results are plotted. Clearly there is still a huge error exactly at the location of where the shock should be. This has two reasons. One is that with the slightest error in the \( x \) direction, the shock moves in the \( x \) direction which results immediately in an
error with the size of the shock. This is the reason for using the 1-norm divided by \( n \), s.t. it is clear that the average error goes down with higher values of \( n, m \). The second reason is explained in Figure 14. The black line stands for the exact location of the shock. The red line represents the border of where the solution is zero, interpreted in the discrete real solution. The blue line represents the border of zero’s calculated by the upwind scheme. This suggests a shift in the shock of the real solution of \( s \in [0, \Delta x] \) to the right depending on the slope of the shock speed \( f'(\alpha) \) and on the ratio \( \frac{\Delta t}{\Delta x} \). In (64) a formula can be seen which was made to calculate the shift \( s \). The results with this formula have improved.

\[
s = \frac{f'(\alpha) \Delta x}{\Delta t} - 1 - f'(\alpha) \frac{\Delta t}{\Delta x}.
\]  

(64)

When doing stability analysis on this scheme for the case \( d = p = M = 1 \) following the same steps as in section 4.2.1 we get an amplification factor of

\[
e^{c \Delta t} = 1 - \frac{\Delta t}{\Delta x}(1 - e^{-i\pi j \Delta x}) = 1 - 2 \frac{\Delta t}{\Delta x} \sin^2\left(\frac{\pi j \Delta x}{2}\right) - 2i \frac{\Delta t}{\Delta x} \sin\left(\frac{\pi j \Delta x}{2}\right) \cos\left(\frac{\pi j \Delta x}{2}\right).
\]

(65)

\[
|e^{c \Delta t}| = \sqrt{1 - 4 \frac{\Delta t}{\Delta x}(1 - \frac{\Delta t}{\Delta x}) \sin^2\left(\frac{\pi j \Delta x}{2}\right)}.
\]

So for \( \frac{\Delta t}{\Delta x} \leq 1 \) the scheme is stable. Since we have \( t \in [0, 0.5] \) and \( x \in [-1, 1] \), we have that

\[
\frac{1}{2 \frac{m-1}{n-1}} = \frac{1}{4 m - 1} \leq 1.
\]

(66)

Equation (65) also suggests that choosing \( \frac{\Delta t}{\Delta x} = 1 \) the discretization is exact. Nevertheless in other cases for \( p, d \) and \( M \) there is a positive but non-linear convection term. It is possible to calculate a bound for \( \frac{\Delta t}{\Delta x} \), but we have not done that. Still (66) suggests that the results could be better choosing \( \Delta x \) smaller than \( \Delta t \). After some simulations the optimal fraction \( \frac{n}{m} \) has been found for \( p = d = 2, M = 1 \). In Figure 15 the result is plotted using shift calculated with (64) and with \( \frac{n}{m} \approx 2.27 \). Comparing this result with the results in Figure 13, this result is almost a factor 3 better with using lower gridsize.
Figure 13: Plots of the result, the difference with the real solution and the 1-norm divided by $n$ for $n, m \in \{129, 513, 2049\}$

Figure 14: Illustration of the cause of the shift in the shock
4.4 Implicit on diffusion, explicit on convection

In this section the term $\frac{\partial}{\partial x} D(u) \frac{\partial u}{\partial x}$ is used as implicit. The term $(f(u))_x$ is explicitly calculated using the upwind scheme, described in section 4.3.4. Since implicit schemes are used, there has to be an iteration at every time step which will be denoted with a subindex next to the space index.

4.4.1 Integrate $D(u)$

Suppose $v = \beta(u) = \int_0^u D(u) \, du$, [8]. Substituting gives

$$\frac{\partial \beta^{-1}(v)}{\partial t} + \frac{\partial}{\partial x} f(\beta^{-1}(v)) - \varepsilon \frac{\partial^2}{\partial x^2} v = 0.$$  \hspace{1cm} (67)

This looks very nice, because the explicit diffusion term is now linear! The discretization scheme looks like

$$\frac{\beta^{-1}(v_{k}^{l+1}) - \beta^{-1}(v_{k}^{l})}{\Delta t} + \frac{f(\beta^{-1}(v_{k}^{l})) - f(\beta^{-1}(v_{k-1}^{l}))}{\Delta x} = \varepsilon \frac{v_{k+1}^{l+1} - 2v_{k}^{l+1} + v_{k-1}^{l+1}}{(\Delta x)^2},$$  \hspace{1cm} (68)

using a nice second order scheme for the diffusive part. Important to notice is that the domain has changed. Since $v$ is now the variable to work with, the domain changes from $[0, 1]$ to $[0, \rho]$ where $\rho = \int_0^1 D(u)$. Later on $u$ can be calculated by taking $\beta^{-1}(v)$. Unfortunately $\beta$ can not be calculated very easily. Instead, if the interval $(0, 1)$ is divided into $n - 1$ intervals, it is easier to use interpolation, to calculate the surface under $\beta$ of each interval. From that data, the inverse function $\beta^{-1}(v)$ can also be calculated. Let us define

$$A(v^{l+1}) = \frac{\varepsilon \Delta t}{(\Delta x)^2} \begin{bmatrix} 2 & -1 & \ldots & \ldots & 0 \\ -1 & 2 & -1 & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & \ldots & \ldots & -1 & 2 & -1 \\ \end{bmatrix} v^{l+1} + \beta^{-1}(v^{l+1}),$$  \hspace{1cm} (69)

Figure 15: Result with the shift for $n = 2000$, $m = 880$, $d = p = 2$, $M = 1$ using upwind scheme.
and

\[
g = \begin{bmatrix}
\frac{\varepsilon \Delta t}{(\Delta x)^2} \rho \\
0 \\
\vdots \\
0
\end{bmatrix} + \beta^{-1}(v^l) - \frac{\Delta t}{\Delta x} \begin{bmatrix}
(f(\beta^{-1}(v^l_1)) - f(\beta^{-1}(\rho))) \\
(f(\beta^{-1}(v^l_2)) - f(\beta^{-1}(v^l_2))) \\
\vdots \\
(f(\beta^{-1}(v^l_{n-1})) - f(\beta^{-1}(v^l_{n-2})))
\end{bmatrix},
\]

such that (68) implies \( A(u) = g \), where at the boundary \( \forall l \in \mathbb{N} \), Dirichlet boundary conditions are assumed. Instead of \( u_1^l = 1 \) and \( u_n^l = 0 \), we now have \( v_1^l = \rho \) and \( v_n^l = 0 \). For every step \( l \), \( v^{l+1} \) can be calculated using Newton’s method, where \( F(u) = A(u) - g \),

\[
F'(v_{(i)})(v) = F(v_{(i)}),
\]

\[
v_{(i+1)} = v_{(i)} - v.
\]

A good choice for \( v_{(i)} \) is \( v^l \), because \( v^{l+1} \) will not differ much from \( v^l \), especially with small \( \Delta t \), thus a better choice than the zero vector and a higher chance of converging. A result is plotted in Figure 16.

### 4.4.2 Chord method

There is even a faster way to iterate, see [9], pp. 365-373 and [10], pp. 895-919. It is called the Chord method. We start with rewriting (68) into

\[
\frac{\beta^{-1}(v^{l+1}_k) - \beta^{-1}(v^l_k) + v^l_k - v^{l+1}_k}{\Delta t} = \frac{f(\beta^{-1}(v^l_k)) - f(\beta^{-1}(v^{l+1}_k))}{\Delta x} + \varepsilon \frac{v^{l+1}_k - 2v^l_k + v^{l+1}_k}{(\Delta x)^2}.
\]

We make use of the fact that of \( v^l_k \) comes close to \( v^{l+1}_k \), then \( \beta^{-1}(v^{l+1}_k) - \beta^{-1}(v^l_k) \) is close to zero and the equation almost linear! Let us define \( \tau_{(i)} \) as the iteration and \( \tau \) will be defined as

\[
\tau_{k,(i)} = \frac{\beta^{-1}(v_{(i)}) - \beta^{-1}(v^l_k)}{v_{(i)} - v^l_k},
\]

if \( v_{k,(i)} \neq v^l_k \) and

\[
\tau_{k,(i)} = (\beta^{-1}(v_{k,(i)}))',
\]

if \( v_{k,(i)} = v^l_k \) for all \( k \in \{1, 2, ..., n\} \). Now a linear equation can be solved with implicit part;

\[
A := \begin{bmatrix}
\frac{2\varepsilon \Delta t}{(\Delta x)^2} + \tau_{2,(i)} & -\frac{\varepsilon \Delta t}{(\Delta x)^2} & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
-\frac{\varepsilon \Delta t}{(\Delta x)^2} & \frac{2\varepsilon \Delta t}{(\Delta x)^2} + \tau_{3,(i)} & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \cdots \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \frac{2\varepsilon \Delta t}{(\Delta x)^2} + \tau_{n-2,(i)} & -\frac{\varepsilon \Delta t}{(\Delta x)^2} \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \frac{2\varepsilon \Delta t}{(\Delta x)^2} + \tau_{n-1,(i)}
\end{bmatrix} v_{(i+1)},
\]
\[
\begin{bmatrix}
\frac{\varepsilon \Delta t}{(\Delta x)^2} \rho \\
0 \\
\vdots \\
0
\end{bmatrix}
+ \frac{\Delta t}{\Delta x}
\begin{bmatrix}
(f(\beta^{-1}(v_1^1)) - f(\beta^{-1}(\rho))) \\
(f(\beta^{-1}(v_3^1)) - f(\beta^{-1}(v_2^1))) \\
\vdots \\
(f(\beta^{-1}(v_{n-1}^1)) - f(\beta^{-1}(v_{n-2}^1)))
\end{bmatrix}
+ \begin{bmatrix}
\tau_{1,(i)v_1,(i)} \\
\tau_{2,(i)v_2,(i)} \\
\vdots \\
\tau_{n,(i)v_n,(i)}
\end{bmatrix}.
\]

From this \( v_{k,(i+1)} \) can be obtained. This method is linear and very fast. A first impression of the speed of this Chord method is given in the next table. Here the parameters are: \( p = d = 2, M = 1, \lambda = .5, \varepsilon = 1, TOL = 10^{-4} \) using the \( \sqrt{\Delta x} \) times the 2-norm and max iterations is set to 25.

<table>
<thead>
<tr>
<th>n=m</th>
<th>9</th>
<th>17</th>
<th>33</th>
<th>65</th>
</tr>
</thead>
<tbody>
<tr>
<td>Newton (s)</td>
<td>1.8408</td>
<td>13.7905</td>
<td>107.0167</td>
<td>734.0315</td>
</tr>
<tr>
<td>Koord (s)</td>
<td>0.1092</td>
<td>0.2496</td>
<td>0.7962</td>
<td>3.9624</td>
</tr>
</tbody>
</table>

Unfortunately implementing this method is not that easy. The problems lays in the derivative of the \( \beta^{-1} \) function. On the boundaries \( \beta^{-1} \) and \( (\beta^{-1})' \) goes to infinity, which leads to an almost singular matrix \( A \) when solving the linear equations. There are ways to overcome this problem. Instead of using \( 1 - H(x) \) one could take \( 0.9 - 0.8H(x) \) as a boundary condition at \( t = 0 \). This method is so fast that it probably pays of the effort, but it will not be included in this report. It is something to do for in the future. In Figure 16 the result of the Newton and the Chord method is plotted. It is hard to say something about how good the results are, but it seems oké.

Figure 16: Results of left Newton method and right Chord method, choosing \( n = m = 64 \) for \( p = d = 2, M = 1 \) and \( \varepsilon = 1 \)
References


