UNIVERSIDAD DE CONCEPCION
DIRECCION DE POSTGRADO
CONCEPCION–CHILE

LEYES DE CONSERVACION Y ECUACIONES AFINES CON FLUJOS NO LOCALES E INVOLUCIONES

Tesis para optar al grado de
Doctor en Ciencias Aplicadas con mención en Ingeniería Matemática

Fernando Elías Betancourt Cerda

FACULTAD DE CIENCIAS FISICAS Y MATEMATICAS
DEPARTAMENTO DE INGENIERIA MATEMATICA

2011

LEYES DE CONSERVACION Y ECUACIONES AFINES CON FLUJOS NO LOCALES E INVOLUCIONES

Fernando Elías Betancourt Cerda

Prof. Guía de Tesis: Dr. Raimund Bürger (Universidad de Concepción).

Prof. Co-Guía de Tesis: Dr. Christian Rohde (Universidad de Stuttgart).

Director de Programa: Dr. Raimund Bürger (Universidad de Concepción).

COMISION EVALUADORA

Dr. Christophe Chalons, Universidad de París VII, Francia.

Dr. Pep Mulet, Universidad de Valencia, España.

Dr. John Towers, MiraCosta College, Estados Unidos de Norteamérica.

COMISION EXAMINADORA

Firma: _____
Dr. Raimund Bürger, Universidad de Concepción, Chile.

Firma: _____
Dr. Freddy Paiva, Universidad de Concepción, Chile.

Firma: _____
Dr. Christian Rohde, Universidad de Stuttgart, Alemania.

Firma: _____
Dr. Mauricio Sepúlveda, Universidad de Concepción, Chile.

Fecha Examen de Grado: _____

Calificación: _____

Concepción–Enero 2011

AGRADECIMIENTOS

Deseo expresar mi agradecimiento al profesor Dr. Raimund Bürger por el apoyo, motivación y entusiasmo sostenido durante todo el tiempo en que hemos trabajado en este proyecto. Valoro en sobremanera su gran calidad humana y su disposición a compartir su conocimiento y experiencia para ayudar en mi formación científica. Agradezco también al profesor Dr. Christian Rohde, mi otro director de tesis, por su gran profesionalismo y destacada calidad humana, además de la hospitalidad manifestada en mi estadía en la Universidad de Stuttgart. También quisiera agradecer en forma especial a los profesores Dr. Gabriel Barrenechea, Dr. Gabriel Gatica y Dr. Freddy Paiva que han dejado una huella importante en mi formación matemática, esforzándose por enseñarme las técnicas y conocimientos básicos necesarios para poder desarrollar este trabajo de tesis.

Agradezco también a la Comisión Nacional de Investigación en Ciencia y Tecnología (CONICYT), al Servicio Alemán de Intercambio Académico (DAAD), al proyecto FONDAP BASAL y al proyecto AMIRA- CORFO 08CM01-17 por el apoyo financiero que ha sido fundamental para llevar a buen término esta tesis.

Al Departamento de Ingeniería Matemática y en particular al Doctorado en Ciencias Aplicadas por su apoyo en la participación en Congresos. A mis amigos del CI^{2}MA en general, pero destacando en particular a Ricardo Ruiz y Frank Sanhueza, compañeros desde el primer año que hicieron mucho más amena mi estadía durante el doctorado. También a mis amigos de IANS (Stuttgart), en especial a Jan Kelkel y Adriana Lalegname por su simpatía y enriquecedora conversación.

A mi esposa, Pamela, por su paciencia al embarcarse conmigo en este proyecto que requirió tanto en lo personal como familiar muchos sacrificios.

Y finalmente a Dios, quien hace en mí todo lo bueno.

A mis padres, a mi esposa Pamela e hija Amanda.
Con mucho Amor...

ABSTRACT

This thesis has three aims. The first aim of the thesis is to study the well-posedness and to develop numerical methods for scalar conservation laws with nonlocal flux function modeling the phenomenon of aggregation in mathematical biology. The existence of weak solutions to a nonlocal strongly degenerate parabolic aggregation equation is proved using a finite difference method and compactness arguments. For uniqueness, we employ an entropy concept and prove the equivalence between weak and entropy solutions. The finite difference method is utilized to generate numerical examples that illustrate the aggregation process.

The second goal of the thesis is to study the well-posedness of a nonlocal conservation but now modeling sedimentation in process industry. We prove existence and uniqueness of entropy solutions for a nonlocal sedimentation equation, again using a finite difference method and standard compactness results. Depending on parameter values, a Lipschitz regularity result or a maximum principle independent by the time variable is found. By the finite difference scheme we obtain numerical examples and compare it with local model results. The layered sedimentation phenomenon is observed.

Finally, the Generalized Lagrange Multiplier Finite Volume Method, which was originally developed for the Maxwell equations, is extended to any hyperbolic Friedrichs system of conservation laws with involutions. We prove the convergence of the method. Moreover, the fulfillment of the involution in the weak sense when the mesh parameter goes to zero is shown. Numerical examples illustrate the properties of the method in the Maxwell equations and in the induction equation in the MHD system.

RESUMEN

La presente tesis tiene tres objetivos. El primero de ellos es el estudio de buen planteamiento y el desarrollo de métodos numéricos para una ley de conservación escalar con flujos no-locales, que modela el fenómeno de agregación en biología matemática. Se demuestra la existencia de solución débil de la ecuación no-local de agregación usando el método de las aproximaciones sucesivas y argumentos de compacidad. Para la unicidad se utiliza el concepcto de entropía y se prueba un resultado de equivalencia entre soluciones débiles y de entropía. Con el método de aproximación se desarrollan ejemplos numéricos que ilustran el fenómeno de agregación.

El segundo objetivo de la tesis es el estudio de buen planteamiento de una ley de consevación no-local, esta vez modelando el proceso de sedimentación. Para esta ecuación, se prueba la existencia de soluciones débiles de entropía por un método de diferencias finitas y argumentos de compacidad. La unicidad se obtiene por la técnica de doblamiento de variables. Dependiendo de ciertos valores de parámetros, se obtiene una regularidad Lipschitz o un Principio del Máximo independiente del tiempo. Con el método de aproximación se generan resultados numéricos que se comparan con los modelos clásicos locales. Se aprecia el fenómeno de sedimentación por capas.

Finalmente, se extiende el método de volúmenes finitos con multiplicadores de Lagrange generalizados, que originalmente fue desarrollado para las ecuaciones de Maxwell, a cualquier sistema hiperbólico de Friedrichs con restricciones de tipo involuciones. Se demuestra la convergencia del método a la solución deseada. Además se prueba el cumplimiento de la involución en el sentido débil. Ejemplos numéricos ilustran las propiedades del método en las ecuaciones de Maxwell y en la ecuación de inducción en magneto-hidrodinámica.

Contents

Introduction

The well-posedness study and the development of numerical methods for conservation laws with nonlocal flux functions have gained importance over the last years in mathematical modeling. Several authors have contributed in developing these equations in mathematical biology [28, 86, 105]. In particular, the aggregation phenomenon has been studied since 1980's by Nagai [90] and Nagai and Mimura [91, 92, 93]. More recent contributions on this topic have been made by Bertozzi et al. [15, 16, 17, 18, 19]. The study of numerical methods to deal with these equations has not received the same attention. In practice, one utilizes classical methods for conservation laws [77]. In the first chapter of this thesis, we show the well-posedness of a strongly degenerate parabolic equation modeling aggregation using a finite difference method. This tool, also provides a reliable numerical method that converges to the exact solution. Numerical examples illustrate the aggregation phenomenon.

The solid-liquid suspensions is a classical area of application of nonlinear conservation laws. The main single contribution was the kinematic sedimentation theory by Kynch [70], which describes the sedimentation of an ideal suspension of small rigid spheres dispersed in a viscous fluid. It is based on the postulate that the settling velocity of a particle is a function of the local solids concentration (or volume fraction). In the second chapter of the thesis, we develop a complementary theory to Kynch's, by assuming that the settling velocity of a particle does not depend only on the local concentration but on the concentration in a contiguous region of finite width. To incorporate this nonlocal behavior, we introduce a convolution with a kernel in the flux function. Again, using a finite difference method, we prove the well-posedness of the nonlocal equation. The main motivation for this nonlocal model is the layered sedimentation phenomenon, reported by Siano [109]. Several numerical examples compare the local and nonlocal model and illustrate the layered sedimentation.

In a wider context, conservation laws are related in several cases to hyperbolic partial differential equations. In the hyperbolic equations, the symmetric linear systems, also called Friedrichs systems, appear in several physical models like the Maxwell equations in electrodynamics. In the Maxwell equations, in addition to the system of PDEs, the

solution must satisfy a differential constraint in the magnetic and electrical fields, called an involution (cf. [39]). Involutions also appear in other physical problems like magneto-hydrodynamic (MHD) and in thermo-elastic problems [39]. The well-posedness of this problem is established in [39]. On the other hand, the development of stable and reliable numerical methods to deal with the involution is still a work in progress. Munz et al. [88, 89, 41] introduced the so-called Generalized Lagrangian Multiplier Finite Volume Method (GLMFVM) to compute approximate solutions for the Maxwell's system [88] and for the MHD equations [41]. This method works well and preserves the involution at the discrete level. In the third chapter of this thesis, we propose a general method that considers the ideas of Munz [88, 89] to deal with involutions and the finite volume method for Friedrichs systems developed by Vila and Villedieu [112], to face the problem of Friedrichs systems with involutions. The convergence of the method and the satisfaction of the involution in the weak sense in the limit when the mesh parameter goes to zero is proved. Numerical examples show the characteristics of the proposed method.

A nonlocal aggregation equation

In the first chapter we study the strongly degenerate parabolic equation

$$u_t + \left(\Phi' \left(\int_{-\infty}^{x} u(y,t)\,\mathrm{d}y \right) u(x,t) \right)_x = A(u)_{xx}, \quad x \in \mathbb{R}, \quad 0 < t \leq T,$$
$$u(x,0) = u_0(x) \geq 0, \quad x \in \mathbb{R}, \quad u_0 \in (L^1 \cap L^\infty)(\mathbb{R})$$

for the density $u = u(x,t) \geq 0$, where $A(u)$ is a diffusion function given by

$$A(u) := \int_0^u a(s)\,\mathrm{d}s,$$

where $a(u) \geq 0$ for $u \in \mathbb{R}$. This equation has been studied as a model of aggregation by a series of authors including Alt [3], Diaz, Nagai, and Shmarev [42], Nagai [90] and Nagai and Mimura [91, 92, 93], all of which assumed that $a(u) = 0$ at most at isolated values of u. We assume that $a(u) = 0$ on u-intervals of positive measure. For instance, if we consider

$$A(u) = \begin{cases} 0 & \text{si } u \leq u_c, \\ a_0(u - u_c) & \text{si } u > u_c, \end{cases}$$

the nonlocal equation model an aggregation-dispersion "threshold" process, i.e, the dispersion stars when u exceeds the critical value $u_c > 0$. We note that for $u < u_c$ the equation degenerates into a hyperbolic scalar conservation law. As an structural assumption, it is supposed that $A(s) \to +\infty$ as $s \to +\infty$.

The aggregation phenomenon is given by the nonlocal flux. For example, if we take

$$u_t + \left(-k \left[\int_{-\infty}^{x} u(y,t)\,\mathrm{d}y - \int_{x}^{\infty} u(y,t)\,\mathrm{d}y \right] u \right)_x = A(u)_{xx}, \quad k > 0,$$

here the convective term provides a mechanism that moves $u(x,t)$ to the right (respectively, to the left) if

$$\int_{-\infty}^{x} u(y,t)\,\mathrm{d}y < \int_{x}^{\infty} u(y,t)\,\mathrm{d}y \quad \text{(respectively, } \ldots > \ldots\text{)}.$$

In other words, an individual will move to the right (respectively, left) if the total population to its right is larger (respectively, smaller) than to its left. Now assume that the initial population is finite and define

$$C_0 := \int_{\mathbb{R}} u_0(x)\,\mathrm{d}x,$$

then we have

$$\Phi(v) = -kv(v - C_0) + \text{const.} \tag{1}$$

We need that $\Phi \in C^2(\mathbb{R})$, and that Φ has exactly one maximum. This assumption is introduced to facilitate some of the steps of the analysis; it is, however, not essential. Employing a function Φ with several separate extrema, the results remain valid.

The key observation made in previous work [3, 90, 91, 92, 93] is that if all coefficients are sufficiently smooth, and $u(x,t)$ is an L^1 solution of the problem, then its primitive is a solution of the local initial value problem

$$v_t + \Phi(v)_x = A(v_x)_x, \quad x \in \mathbb{R}, \quad t \in (0,T],$$
$$v(x,0) = v_0(x), \quad x \in \mathbb{R}, \quad v_0(x) := \int_{-\infty}^{x} u_0(\xi)\,\mathrm{d}\xi.$$

In **Chapter 1**, we use this idea to define a finite difference scheme for u based on a monotonic scheme for the primitive v. This scheme is an explicit version of the scheme developed by Evje and Karlsen [47]. The scheme for u can be obtained taking the discrete derivative of the values of the scheme for v. Using Lax-Wendroff arguments, we prove that the numerical solution generated by the scheme converges to a weak solution of the problem. For uniqueness, following the ideas of Carrillo [30] and Kobayasi [66], it is proved that any weak solution is also an entropy solution. Slight modifications of a result in [62], gives us the uniqueness of entropy solutions. To end the chapter, numerical experiments show the properties of the scheme and the aggregation phenomenon. This study has given rise the following paper:

- F. Betancourt, R. Bürger and K.H. Karlsen. "A strongly degenerate parabolic aggregation equation", accepted for publication in *Communications in Mathematical Sciences.*

A nonlocal sedimentation equation

In **Chapter 2** we study a family of nonlocal conservation laws

$$u_t + \big(u(1-u)^\alpha V(K_a * u)\big)_x = 0, \quad x \in \mathbb{R}, \quad t \in (0,T],$$
$$u(0,x) = u_0(x), \quad 0 \le u_0(x) \le 1, \quad x \in \mathbb{R}.$$

Here the solid fraction $u(x,t)$, depends only on the deep x and the time t. The parameter α satisfies either $\alpha = 0$ or $\alpha \ge 1$. The function V is a hindered settling factor that can be chosen, for example, as

$$V(w) = (1-w)^n, \quad n \ge 1,$$

according to Richardson and Zaki [100], and which is herein supposed to depend on

$$(K_a * u)(x,t) = \int_{-2a}^{2a} K_a(y) u(x-y,t)\,\mathrm{d}y,$$

where K_a is a symmetric, non-negative piecewise smooth kernel function with support on $[-2a,2a]$ with $a > 0$ and

$$\int_{\mathbb{R}} K_a(x)\,\mathrm{d}x = 1.$$

Usually, one defines $K = K(x)$ with support on $[-2,2]$ and sets $K_a(x) := a^{-1}K(a^{-1}x)$. This model can be motivated as follows. When diffusion is negligible, the kinematic Kynch theory [70] models the sedimentation process through

$$u_t(x,t) + \big(u(x,t)v_{\mathrm{s}}(x,t)\big)_x = 0,$$

where $v_{\mathrm{s}}(x,t)$ is the solids phase velocity, or settling velocity, at position x at time t. Considering the Richarson and Zaki formula for the settling velocity we have

$$v_{\mathrm{s}}(x,t) = v_{\mathrm{St}}(1-u(x,t))^n,$$

where v_{St} is the Stokes velocity. Under the assumption that V depends on $K_a * u$ instead of u (a justified explanation is detailed in 2.2.1), the Kynch equation turns now

$$u_t(x,t) + v_{\mathrm{St}}\Big(u(x,t)\big(1-(K_a*u)(x,t)\big)^n\Big)_x = 0.$$

A different approach consists in considering also the fluid mass conservation $-u_t + ((1-u)v_{\mathrm{f}})_x = 0$, where v_{f} is the fluid phase velocity. For batch settling we have the relation $v_{\mathrm{s}} = (1-u)v_{\mathrm{r}}$, where $v_{\mathrm{r}} := v_{\mathrm{s}} - v_{\mathrm{f}}$ is the solid-fluid relative velocity or slip velocity. This leads to the governing equation

$$u_t + \big(u(1-u)v_{\mathrm{r}}\big)_x = 0.$$

Assuming now that v_r (instead of v_s) has a nonlocal behavior and requiring that the local versions based on constitutive assumptions for either v_s or v_r should coincide, we state the constitutive assumption for v_r as $v_r = V(K_a * u)/(1-u)$. For instance, if we employ the Richardson-Zaki equation, then the exponent n should be reduced by one, so using the properly adapted Richardson-Zaki equation leads us to

$$v_s(x_0,t)/v_{St} = \big(1-u(x_0,t)\big)\big(1-(K_a*u)(x_0,t)\big)^{n-1},$$

from which the following conservation law is obtained

$$u_t + v_{St}\big(u(1-u)(1-K*u)^{n-1}\big)_x = 0.$$

We study the more general form of the last equation replacing $(1-u)$ by $(1-u)^\alpha$. On the other hand, the qualitative properties of these nonlocal equations are also interesting. The "effective" equations are dispersive. Dispersive equations usually present oscillations. These oscillations are interpreted as layers of different concentration.

Similarly to the work developed in **Chapter 1**, we get uniqueness for entropy solutions using a slight modification of a result in Karlsen and Risebro [62]. Existence is obtained using a difference-quadrature method, which is based on the classical Lax-Friedrichs scheme. We remark that for $\alpha = 0$ the solution is Lipschitz if the initial data do so. The Lipschitz regularity make possible to get uniqueness without the entropy concept. However, even though the solution is bounded for $T < +\infty$, it does not remain in the interval $[0,1]$ although $u_0 \in [0,1]$. On the other side for $\alpha \geq 1$, the solution is in general discontinuous but it remains in the interval $[0,1]$ provided the initial data do so. Numerical examples illustrate the behavior of the solution and the layered sedimentation. This chapter gave rise to the article:

- F. Betancourt, R. Bürger, K. H. Karlsen and E. M. Tory. "On nonlocal conservation laws modeling sedimentation", accepted for publication in *Nonlinearity*.

Finite Volume Schemes for Friedrichs systems with involutions

In **Chapter 3**, we study linear systems of conservation laws of Friedrichs type. In addition, we impose differential side conditions in the form of involutions [39]. We consider the spatially d-dimensional case with $d \geq 2$, $x = (x_1,\ldots,x_d)^T$ and time $t \geq 0$. For $t \leq T$, we define the functions $G^1,\ldots,G^d, D : \mathbb{R}^d \times [0,T] \to \mathbb{R}^{m\times m}$ with $m \in \mathbb{N}$, and $f : \mathbb{R}^d \times [0,T] \to \mathbb{R}^m$. Since the system is of Friedrichs type, we have that $G^1(x,t),\ldots,G^d(x,t)$ are symmetric matrices for all $(x,t) \in \mathbb{R}^d \times [0,T]$.

The initial value problem is given by:

$$\frac{\partial}{\partial t}u(x,t)+\sum_{i=1}^{d}\frac{\partial}{\partial x_i}\big(G^i(x,t)u(x,t)\big)+D(x,t)u(x,t)=f(x,t),$$
$$u(x,0)=u_0(x).$$

Moreover, we require the solution u to satisfy the linear differential side condition

$$\sum_{i=1}^{d}M_i\frac{\partial}{\partial x_i}\left(u(x,t)\right)=0,\qquad \Big((x,t)\in\mathbb{R}^d\times[0,T)\Big),$$

where M_i, $i=1,\dots,d$, are constant matrices. According to Dafermos [39], the differential constraint is called an involution for the system if and only if any solution of the system satisfies the involution, whenever the initial data do so.
Involutions appear frequently in applications. We mention the classical Maxwell system to describe electrodynamical processes (cf.[75]). The divergence of the electrical and magnetical field is constrained in this case. The induction equations in the (in)compressible electro- and magnetohydrodynamical equations provide similar examples but with (x,t)-dependence in the flux. Solutions of the equations of linear elasticity have to satisfy compatibility conditions on the deformation gradient, which result in an involutionary condition (cf. Chapter 5 of [39]). Yet, another example is the linear piezoelectrical system (see [84]). Let us mention that involutions of course appear also in the more challenging case of nonlinear conservation laws. Again, magnetohydrodynamics [38], electrohydrodynamics, nonlinear elasticity systems, but also Einstein's equations of general relativity are prominent examples.
On the analytical level an involutionary side condition is not problematic. The well-posedness is well known from [39]. By definition the involution is satisfied. Also standard numerical schemes are known to converge. However, without consideration of the involution in the numerical scheme the residuum in the side condition usually grows with increasing time [88]. In coupled processes this is a typical source of instabilities (cf.[88] and cites therein). Therefore, a wide range of stabilization methods has been suggested (e.g. [4, 20, 35, 58, 89]).
The motivation for this contribution is the work of Munz et al. [89]. They introduced in particular the so-called hyperbolic Generalized Lagrangian Multiplier Finite Volume Method (GLMFVM) to compute approximate solutions for Maxwell's system of linear electrodynamics. We formulate this approach for a general Friedrichs systems with involutions. The method is based on solving an extended system of relaxation-type. Let $a,\varepsilon>0$ and $u_0,\psi_0^\varepsilon:\mathbb{R}^d\to\mathbb{R}^m$ be given. Consider the following initial value problem for the unknown function: $w^\varepsilon:\mathbb{R}^d\times[0,T]\to\mathbb{R}^{2m}$, with $w^\varepsilon:=(u_1^\varepsilon,\dots,u_m^\varepsilon,\psi_1^\varepsilon,\dots,\psi_m^\varepsilon)^T$,

given by

$$\frac{\partial}{\partial t}u^{\varepsilon}+\sum_{i=1}^{d}\frac{\partial}{\partial x_i}\left(G^i(x,t)u^{\varepsilon}\right)+M_i^T\frac{\partial}{\partial x_i}\psi^{\varepsilon}+D(x,t)u^{\varepsilon}=f(x,t),$$
$$\frac{\partial}{\partial t}\psi^{\varepsilon}+\sum_{i=1}^{d}\frac{M_i}{\varepsilon}\frac{\partial}{\partial x_i}u^{\varepsilon}+a\psi^{\varepsilon}=0,$$
$$u^{\varepsilon}(x,0)=u_0^{\varepsilon}(x),\qquad \psi^{\varepsilon}(x,0)=\psi_0^{\varepsilon}(x)=0.$$

Since the last formulation is not symmetric we introduce the variable $\varphi^{\varepsilon}:=\psi^{\varepsilon}\sqrt{\varepsilon}$ and the system turns

$$\frac{\partial}{\partial t}U^{\varepsilon}+\sum_{i=1}^{d}\frac{\partial}{\partial x_i}(A^{\varepsilon,i}U^{\varepsilon})+BU^{\varepsilon}=F,\qquad U^{\varepsilon}(x,0)=U_0^{\varepsilon}(x):=\begin{pmatrix}u_0^{\varepsilon}(x)\\0\end{pmatrix},$$

where

$$U^{\varepsilon}:=\begin{pmatrix}u^{\varepsilon}\\\varphi^{\varepsilon}\end{pmatrix};\quad A^{\varepsilon,i}:=\begin{pmatrix}G^i & \frac{M_i^T}{\sqrt{\varepsilon}}\\ \frac{M_i}{\sqrt{\varepsilon}} & 0\end{pmatrix};\quad B:=\begin{pmatrix}D & 0\\0 & aI\end{pmatrix};\quad F:=\begin{pmatrix}f\\0\end{pmatrix}.$$

We prove that the symmetric system is well posed. It is also proved, under mild assumptions, that the solution of the extended systems equals the solution of the original system a.e. The main result is the convergence of the GLMFVM. This is done in Section 3.4. Following the theory developed by Vila and Villedieu [112] and Jovanovic and Rohde [59] we get

$$\|u_h^{\varepsilon}-u^{\varepsilon}\|_{L^2(\mathbb{R}^d\times[0,T];\mathbb{R}^m)}=\mathscr{O}\left(\varepsilon^{-1/4}h^{1/2}\right)$$

where h is the mesh parameter, $u_h^{\varepsilon}:\mathbb{R}^d\times[0,T]\to\mathbb{R}^m$ is the solution generated by the GLMFVM and u^{ε} is the solution of the extended system. A crucial fact is that coupling ε and h the estimate does not depend critically on ε. As a Corollary, it is found that the involution is satisfied in the weak sense when the mesh parameter goes to zero . Numerical examples realize the characteristics of the GLMFVM. As a result of this research we have:

- F. Betancourt and C. Rohde. "Finite-Volume Schemes for Friedrichs Systems with Involutions" (in preparation).

Introducción

El estudio de buen planteamiento y el desarrollo de métodos numéricos para leyes de conservación escalares con flujos no-locales ha tomado gran importancia en el último tiempo dentro del área de la modelación matemática. Diversos autores han hecho numerosos avances en el desarrollo de estas ecuaciones dentro de la modelación de procesos biológicos [28, 86, 105]. En particular, el fenómeno de agregación ha sido estudiado desde la década de los 80's por Nagai [90] y, Nagai y Mimura [91, 92, 93]. Contribuciones más recientes han sido hechas por Bertozzi y colaboradores [15, 16, 17, 18, 19]. El estudio de métodos numéricos sencillos para estos problemas no ha sido objeto de un mayor estudio. En la práctica se recurre a los métodos clásicos conocidos en leyes de conservación [77]. En el primer capítulo de esta tesis, usando el método de las aproximaciones sucesivas, se presenta el análisis de buen planteamiento de una ley de conservación escalar parabólica fuertemente degenerada que modela el fenómeno de agregación tipo "enjambre". Con dicha técnica, además de probarse el buen planteamiento del problema, se obtiene un método numérico confiable que aproxima y converge a la solución exacta deseada. Ejemplos numéricos ilustran el fenómeno de agregación.

Las suspensiones sólido-líquido son un área clásica de aplicación en leyes de conservación no-lineales. La más importante contribución en el estudio de este proceso fue hecha por Kynch [70], quien desarrolló una teoría de tipo cinemática. El supuesto clave de la teoría de Kynch es que la velocidad de sedimentación de una partícula depende sólo de la concentración en el mismo punto donde se encuentra la partícula. En el segundo capítulo de la tesis se desarrolla una teoría complementaria a la desarrollada por Kynch donde se supone que la velocidad de sedimentación de una partícula depende no sólo de la concentración en el lugar donde está la partícula, sino que de la concentración en una vecindad de ésta. Dicha dependencia se produce a través de la incorporación de una convolución con un kernel en la función flujo. Nuevamente, a través del método de aproximaciones sucesivas, se demuestra el buen planteamiento de la ecuación no-local. La motivación principal del modelo no-local es tratar de interpretar y reproducir el fenómeno de sedimentación por capas reportado por Siano [109]. Variados ejemplos numéricos muestran los efectos en el proceso de sedimentación de la incorporación del término no-local.

En un contexo más amplio, las leyes de conservación están relacionadas en muchos casos con las ecuaciones en derivadas parciales de tipo hiperbólico. Dentro de las ecuaciones hiperbólicas, los sistemas hiperbólicos de leyes de conservación lineales simétricos, llamados también de Friedrichs, aparecen en algunos modelos físicos como por ejemplo las ecuaciones del electromagnetismo. Adicionalmente, en este ejemplo, se requiere que la solución del sistema satisfaga una restricción de tipo diferencial llamada involución [39], que corresponde a las restricciones sobre el campo eléctrico y magnético. Involuciones aparecen en otros problemas físicos como lo son la magnetohidrodinámica y los procesos termoelásticos. El buen planteamiento de este problema ya ha sido abordado [39]. Sin embargo, el desarrollo de métodos numéricos estables que consideren la involución es aún un problema en estudio. Munz y colaboradores [88, 89, 41] desarrollaron el llamado GLMFVM (Generalized Lagrange Multiplier Finite Volume Method) por sus siglas en inglés, para incluir las involuciones en el caso del electromagnetismo (lineal) y magnetohidrodinámica (no-lineal). Dicho método funciona de buena manera y preserva la involución a nivel discreto. En el tercer capítulo de la tesis se propone un método de volúmenes finitos, el cual mezcla las ideas de Munz para incorporar las involuciones en el sistema hiperbólico, con el método de volúmenes finitos diseñado por Vila y Villedieu [112] para sistemas de Friedrichs. Se demuestra la convergencia a la solución deseada y la satisfacción de la involución en el sentido débil. Ejemplos numéricos dan cuenta de las características de la solución numérica generada así como de las propiedades del método.

Una ecuación de agregación no-local

En el capítulo 1 se estudia el problema de valores iniciales para la ecuación parabólica fuertemente degenerada

$$u_t + \left(\Phi' \left(\int_{-\infty}^{x} u(y,t)\,\mathrm{d}y \right) u(x,t) \right)_x = A(u)_{xx}, \quad x \in \mathbb{R}, \quad 0 < t \leq T,$$
$$u(x,0) = u_0(x) \geq 0, \quad x \in \mathbb{R}, \quad u_0 \in (L^1 \cap L^\infty)(\mathbb{R})$$

donde $u = u(x,t) \geq 0$ es un tipo de densidad o concentración, $A(u)$ es el coeficiente de difusión dado por

$$A(u) := \int_0^u a(s)\,\mathrm{d}s,$$

donde $0 \leq a(u) < +\infty$. Esta ecuación ha sido estudiada por varios autores, entre ellos Alt [3], Diaz, Nagai, y Shmarev [42], Nagai [90] y Nagai y Mimura [91, 92, 93]. Todos los autores nombrados han asumido que $a(u) = 0$ sólo en valores aislados de u. En el caso tratado se supone que $a(u) = 0$ en u-intervalos de medida finita pero positiva. Por ejemplo

si se define

$$A(u) = \begin{cases} 0 & \text{si } u \leq u_{\mathrm{c}}, \\ a_0(u-u_{\mathrm{c}}) & \text{si } u > u_{\mathrm{c}}, \end{cases}$$

la ecuación representa un fenómeno de agregación-dispersión con umbral, es decir, la dispersión dada por el término parabólico, se activa cuando u excede el valor crítico $u_{\mathrm{c}} > 0$. Este hecho le da sentido a la denominación de ecuación parabólica fuertemente degenerada dado que si $u \leq u_c$ la ecuación es hiperbólica, en cambio, si $u > u_c$ la ecuación es del tipo parabólico. Como supuesto estructural en el término parabólico se asume que $A(s) \to \infty$ cuando $s \to \infty$.

El fenómeno de agregación está dado por la parte no-local en la función de flujo convectiva. Si se considera la ecuación

$$u_t + \left(-k \left[\int_{-\infty}^{x} u(y,t)\,\mathrm{d}y - \int_{x}^{\infty} u(y,t)\,\mathrm{d}y \right] u \right)_x = A(u)_{xx}, \quad k > 0.$$

El término convectivo entrega un mecanismo que mueve $u(x,t)$ a la derecha (respectivamente a la izquierda) si

$$\int_{-\infty}^{x} u(y,t)\,\mathrm{d}y < \int_{x}^{\infty} u(y,t)\,\mathrm{d}y \quad \text{(respectivamente, } \ldots > \ldots\text{)},$$

dicho de otro modo, un individuo se mueve a la derecha (respectivamente a la izquierda) si la cantidad de individuos es mayor a su derecha (respectivamente a su izquierda). Si se define

$$C_0 := \int_{\mathbb{R}} u_0(x)\,\mathrm{d}x,$$

entonces

$$\Phi(v) = -kv(v - C_0) + \text{const.}$$

Sobre la parte convectiva, se asume que la función Φ tiene un sólo máximo. Esta hipótesis es introducida sólo para facilitar algunas pasos dentro del análisis, sin embargo, no es esencial. Si se considera una función Φ con extremos separados, los resultados siguen siendo válidos.

La observación clave hecha en trabajos previos [3, 90, 91, 92, 93] , es que si todos los coeficientes son suaves y $u(x,t)$ es una solución en L^1 del problema, su primitiva v es una solución de

$$\begin{aligned} v_t + \Phi(v)_x &= A(v_x)_x, \quad x \in \mathbb{R}, \quad t \in (0,T], \\ v(x,0) &= v_0(x), \quad x \in \mathbb{R}, \quad v_0(x) := \int_{-\infty}^{x} u_0(\xi)\,\mathrm{d}\xi. \end{aligned}$$

En el capítulo 1, se utiliza esta idea, presentada aquí formalmente, para definir un esquema de diferencias finitas para la ecuación de agregación basándose en un esquema (monótono) para su primitiva v. El esquema usado es la versión explícita del esquema desarrollado por Evje y Karlsen [47]. El esquema para u se obtiene tomando la derivada discreta de los valores del esquema para v. A través de modificaciones standard de argumentos de tipo Lax-Wendroff, se prueba que la solución numérica generada por el esquema para u converge a una solución débil del problema.
Para probar unicidad de la solución se utiliza el marco de soluciones de entropía. Siguiendo la línea de los trabajos de Carillo [30] y Kobayasi [66] se prueba que toda solución débil de la ecuación de agregación es también solución de entropía. Con pequeñas modificaciones respecto de un resultado de unicidad en [62], se demuestra que las soluciones de entropía son únicas. Para finalizar, ejemplos numéricos ilustran el fenómeno de agregación y las propiedades de convergencia del esquema.
El estudio de esta ecuación dio origen al artículo:

- F. Betancourt, R. Bürger y K.H. Karlsen. "A strongly degenerate parabolic aggregation equation", aceptado para publicación en *Communications in Mathematical Sciences.*

Ecuaciones no-locales en sedimentación

En el capítulo 2 de esta tesis se estudia una familia de leyes de conservación con flujo no-local definidas por

$$\begin{aligned} u_t + \big(u(1-u)^\alpha V(K_a * u)\big)_x = 0, \quad x \in \mathbb{R}, \quad t \in (0,T], \\ u(0,x) = u_0(x), \quad 0 \le u_0(x) \le 1, \quad x \in \mathbb{R}. \end{aligned}$$

Donde la fracción de sólidos $u(x,t)$, se considera una función sólo de la profundidad x y del tiempo t. El parámetro α satisface $\alpha = 0$ o bien $\alpha \ge 1$. La función V, conocida como factor de obstaculización, está dada por

$$V(w) = (1-w)^n, \quad n \ge 1,$$

de acuerdo a lo propuesto por Richardson y Zaki [100], y que se supone depende de

$$(K_a * u)(x,t) = \int_{-2a}^{2a} K_a(y) u(x-y,t)\,\mathrm{d}y,$$

donde K_a es un kernel simétrico, no negativo y suave a trozos con soporte en el intervalo $[-2a, 2a]$ con $a > 0$ y

$$\int_{\mathbb{R}} K_a(x)\,\mathrm{d}x = 1.$$

Es común definir $K = K(x)$ con soporte en $[-2,2]$ y considerar $K_a(x) := a^{-1}K(a^{-1}x)$.
El modelo puede ser motivado como sigue. Si la difusión es despreciable, la teoría cinemática de Kynch [70] modela el proceso de sedimentación a través de la ecuación

$$u_t(x,t) + \big(u(x,t)v_{\mathrm{s}}(x,t)\big)_x = 0,$$

donde $v_{\mathrm{s}}(x,t)$ es la velocidad de fase sólida, también llamada velocidad de sedimentación, en la posición x en el instante t. Considerando la fórmula de Richardson y Zaki para la velocidad de sedimentación, se tiene que

$$v_{\mathrm{s}}(x,t) = v_{\mathrm{St}}(1-u(x,t))^n,$$

donde v_{St} es la velocidad de Stokes. Bajo el supuesto que V depende de $K_a * u$ en vez de u (una justificación detallada se encuentra en la subsección 2.2.1), la ecuación del modelo de Kynch toma la forma

$$u_t(x,t) + v_{\mathrm{St}}\Big(u(x,t)\big(1-(K_a*u)(x,t)\big)^n\Big)_x = 0$$

Una ecuación diferente se obtiene considerando también la ecuación de conservación del fluido $-u_t + \big(v_{\mathrm{f}}(1-u)\big)_x = 0$, donde v_{f} es la velocidad de la fase líquida. Para sedimentación batch se tiene que $v_{\mathrm{s}} = (1-u)v_{\mathrm{r}}$, donde $v_{\mathrm{r}} := v_{\mathrm{s}} - v_{\mathrm{f}}$ es la velocidad relativa entre fases. Esto lleva a la ecuación para la concentración de sólidos

$$u_t + \big(u(1-u)v_{\mathrm{r}}\big)_x = 0.$$

Si se supone ahora que v_{r} (en vez de v_{s}) tiene un comportamiento no-local y que las ecuaciones constitutivas para v_{r} y v_{s} coinciden, se obtiene que $v_{\mathrm{r}} = V(K_a*u)/(1-u)$. Por ejemplo, si se emplea la ecuación de Richardson y Zaki esto lleva a

$$v_{\mathrm{s}}(x,t)/v_{\mathrm{St}} = \big(1-u(x,t)\big)\big(1-(K_a*u)(x,t)\big)^{n-1},$$

con la cual se llega a la ley de conservación

$$u_t + v_{\mathrm{St}}\big(u(1-u)(1-K_a*u)^{n-1}\big)_x = 0.$$

Se estudia la forma más general de la última ecuación reemplazando el término $(1-u)$ por $(1-u)^\alpha$. Dentro de las propiedades cualitativas de la ecuación no-local, se destaca que su ecuación "efectiva" [114], es de carácter dispersivo. Las ecuaciones dispersivas se caracterizan por presentar oscilaciones. Se interpretan dichas oscilaciones como "capas" de sedimento de distinta concentración.
De forma análoga a lo desarrollado en el capítulo 1, se establece la unicidad de soluciones

utilizando el concepto de entropía de Kruzkov y usando una variación respecto de un resultado en [62]. La existencia de solución de entropía para la ecuación de sedimentación no-local se logra a través de un esquema de diferencias finitas y cuadratura basado en el clásico esquema de Lax-Friedrichs. Se destaca el hecho de que para $\alpha = 0$ la solución es Lipschitz continua si el dato inicial lo es. Esta regularidad hace posible prescindir del concepto de entropía para obtener la unicidad. Sin embargo, aunque la solución permanece acotada para todo tiempo $T < +\infty$, ésta escapa del intervalo $[0,1]$ aún cuando $u_0 \in [0,1]$. Por otro lado, para $\alpha \geq 1$, la solución es en general discontinua aunque el dato inicial sea suave, pero ésta se mantiene en el intervalo $[0,1]$. Ejemplos numéricos ilustran el comportamiento de la solución de entropía de la ecuación no-local. Los resultados anteriores constituyen el artículo:

- F. Betancourt, R. Bürger, K. H. Karlsen y E. M. Tory. "On nonlocal conservation laws modeling sedimentation", aceptado para publicación en *Nonlinearity*.

Esquemas de Volúmenes Finitos para Sistemas de Friedrichs con Involuciones

En el capítulo 3 de esta tesis se estudian sistemas lineales de leyes de conservación del tipo Friedrichs, que además deben satisfacer restricciones de tipo diferencial denominadas involuciones [39]. Se considera el caso con d dimensiones espaciales, $d \geq 2$, $x = (x_1, \ldots, x_d)^T$, y tiempo $t \geq 0$. Para $t \leq T$, se definen las funciones $G^1, \ldots, G^d, D : \mathbb{R}^d \times [0,T] \to \mathbb{R}^{m \times m}$ con $m \in \mathbb{N}$, y $f : \mathbb{R}^d \times [0,T] \to \mathbb{R}^m$. Puesto que el sistema es de tipo Friedrichs, se tiene que las funciones matriciales $G^1(x,t), \ldots, G^d(x,t)$ son simétricas para todo $(x,t) \in \mathbb{R}^d \times [0,T]$. El problema de valores iniciales a resolver está dado por:

$$\frac{\partial}{\partial t} u(x,t) + \sum_{i=1}^{d} \frac{\partial}{\partial x_i} \left(G^i(x,t) u(x,t) \right) + D(x,t) u(x,t) = f(x,t),$$
$$u(x,0) = u_0(x).$$

La solución del sistema anterior debe además satisfacer la restricción diferencial

$$\sum_{i=1}^{d} M_i \frac{\partial}{\partial x_i} \left(u(x,t) \right) = 0, \qquad \Big((x,t) \in \mathbb{R}^d \times [0,T) \Big),$$

donde M_i, $i = 1, \ldots, d$, son matrices constantes. De acuerdo a la definición dada por Dafermos [39], la restricción diferencial es una involución si y sólo si toda solución del sistema satisface la restricción siempre que el dato inicial lo haga.
Las restricciones de tipo involución aparecen con frecuencia en modelos físicos. Dentro de ellos, destaca el sistema de Maxwell que describe los procesos electrodinámicos. La

divergencia del campo eléctrico y magnético son restricciones en ese caso. La ecuación de inducción en electro y magneto hidrodinámica son ejemplos similares pero con una dependencia de (x,t) en la función de flujo. Las soluciones de las ecuaciones de la elasticidad lineal tienen que satisfacer condiciones de compatibilidad en el gradiente de la deformación, lo que se traduce en una condición de involución (ver [39] Cap. 5). Otro ejemplo son los sistemas piezo-eléctricos. Obviamente, las involuciones aparecen en leyes de conservación no-lineales, nuevamente, electro y magneto hidrodinámica, elasticidad no-lineal así como también las ecuaciones de Einstein de relatividad general son ejemplos importantes.

Desde el punto de vista analítico las involuciones no son problematicas. El buen planteamiento es conocido [39]. Por definición, la involución se satisface. Para métodos numéricos standard se conoce la convergencia. Sin embargo, sino se considera la involución en el esquema numérico, el residuo en la involución puede crecer con el tiempo [88]. En métodos numéricos acoplados, ésta es una típica fuente de inestabilidades (ver [88] y las referencias ahí citadas). Varios métodos de estabilización han sido reportados [4, 20, 35, 58, 89]. La motivación de este estudio es el trabajo de Munz y colaboradores [88, 89]. Ellos introducen el llamado Método de Volúmenes Finitos con Multiplicador de Lagrange Generalizado, GLMFVM por sus siglas en inglés, para el cálculo del sistema de Maxwell en electrodinámica. En el último capítulo de esta tesis, se reformula este método para cualquier sistema de Friedrichs con involuciones. Este método se basa en la resolución de un sistema extendido de tipo relajación. Sean a, $\varepsilon > 0$ y u_0, $\psi_0^\varepsilon : \mathbb{R}^d \to \mathbb{R}^m$ dados. Se considera el problema de Cauchy para la incógnita $w^\varepsilon : \mathbb{R}^d \times [0,T] \to \mathbb{R}^{2m}$, con $w^\varepsilon := (u_1^\varepsilon, \ldots, u_m^\varepsilon, \psi_1^\varepsilon, \ldots, \psi_m^\varepsilon)^T$, dado por

$$\begin{aligned} \frac{\partial}{\partial t}u^\varepsilon + \sum_{i=1}^{d} \frac{\partial}{\partial x_i}\left(G^i(x,t)u^\varepsilon\right) + M_i^T \frac{\partial}{\partial x_i}\psi^\varepsilon + D(x,t)u^\varepsilon &= f(x,t), \\ \frac{\partial}{\partial t}\psi^\varepsilon + \sum_{i=1}^{d} \frac{M_i}{\varepsilon}\frac{\partial}{\partial x_i}u^\varepsilon + a\psi^\varepsilon &= 0, \\ u^\varepsilon(x,0) = u_0^\varepsilon(x), \qquad \psi^\varepsilon(x,0) &= \psi_0^\varepsilon(x) = 0. \end{aligned}$$

Como la formulación anterior no es simétrica se introduce la variable $\varphi^\varepsilon := \psi^\varepsilon\sqrt{\varepsilon}$ con lo que el sistema a resolver se reduce a

$$\frac{\partial}{\partial t}U^\varepsilon + \sum_{i=1}^{d} \frac{\partial}{\partial x_i}(A^{\varepsilon,i}U^\varepsilon) + BU^\varepsilon = F, \qquad U^\varepsilon(x,0) = U_0^\varepsilon(x) := \begin{pmatrix} u_0^\varepsilon(x) \\ 0 \end{pmatrix},$$

con

$$U^\varepsilon := \begin{pmatrix} u^\varepsilon \\ \varphi^\varepsilon \end{pmatrix}; \quad A^{\varepsilon,i} := \begin{pmatrix} G^i & \frac{M_i^T}{\sqrt{\varepsilon}} \\ \frac{M_i}{\sqrt{\varepsilon}} & 0 \end{pmatrix}; \quad B := \begin{pmatrix} D & 0 \\ 0 & aI \end{pmatrix}; \quad F := \begin{pmatrix} f \\ 0 \end{pmatrix}.$$

Se prueba que el problema simétrico extendido está bien planteado. También se demuestra, bajo supuestos no restrictivos, que la solución del sistema extendido es igual a la solución del sistema original en casi todo punto. Luego se introduce el método de volúmenes finitos GLMFVM. El resultado principal es la convergencia del método GLMFVM. Esta se realiza en la sección 3.4. Siguiendo la teoría desarrollada por Vila y Villedieu [112] y Jovanovic y Rohde [59] se obtiene que

$$\|u_h^\varepsilon - u^\varepsilon\|_{L^2(\mathbb{R}^d \times [0,T];\mathbb{R}^m)} = \mathscr{O}\left(\varepsilon^{-1/4} h^{1/2}\right)$$

donde h es el parámetro de malla, $u_h^\varepsilon : \mathbb{R}^d \times [0,T] \to \mathbb{R}^m$ es la solución generada por el método GLMFVM y u^ε la solución del sistema extendido. Un hecho importante es que acoplando ε con h la estimación no depende críticamente del parámetro ε. Como Corolario se demuestra que la involución se satisface en el límite cuando h y ε van a 0. Ejemplos numéricos ilustran la importancia de considerar la involución en el método de aproximación. Como resultado de esta investigación se tiene en preparación el artículo:

- F. Betancourt y C. Rohde. “Finite-Volume Schemes for Friedrichs Systems with Involutions”.

Chapter 1

Strongly degenerate parabolic aggregation equation

This chapter is concerned with a strongly degenerate convection-diffusion equation in one space dimension whose convective flux involves a nonlinear function of the total mass to one side of the given position. This equation can be understood as a model of aggregation of the individuals of a population with the solution representing their local density. The aggregation mechanism is balanced by a degenerate diffusion term describing the effect of dispersal. In the strongly degenerate case, solutions of the nonlocal problem are usually discontinuous and need to be defined as weak solutions. A finite difference scheme for the nonlocal problem is formulated and its convergence to the unique weak solution is proved. This scheme emerges from taking divided differences of a monotone scheme for the local PDE for the primitive. Some numerical examples illustrate the behaviour of solutions of the nonlocal problem, in particular the aggregation phenomenon.

1.1 Introduction

1.1.1 Scope

This chapter is related to the initial value problem for a strongly degenerate convection-diffusion equation of the form

$$u_t + \left(\Phi' \left(\int_{-\infty}^{x} u(y,t)\,\mathrm{d}y \right) u(x,t) \right)_x = A(u)_{xx}, \quad x \in \mathbb{R}, \quad 0 < t \leq T, \tag{1.1}$$

$$u(x,0) = u_0(x) \geq 0, \quad x \in \mathbb{R}, \quad u_0 \in (L^1 \cap L^\infty)(\mathbb{R}) \tag{1.2}$$

for the density $u = u(x,t) \geq 0$, where $A(u)$ is a diffusion function given by $A(u) := \int_0^u a(s)\,\mathrm{d}s$, where $a(u) \geq 0$ for $u \in \mathbb{R}$. The model (1.1), (1.2) was studied as a model of aggregation

by a series of authors including Alt [3], Diaz, Nagai, and Shmarev [42], Nagai [90] and Nagai and Mimura [91, 92, 93], all of which assumed that $a(u) = 0$ at most at isolated values of u. It is the purpose of this work to study (1.1), (1.2) under the more general assumption that $a(u) = 0$ on bounded u-intervals on which (1.1) reduces to a first-order conservation law with nonlocal flux. We assume that $A(s) \to \infty$ as $s \to \infty$.

The key observation made in previous work [3, 90, 91, 92, 93] is that if all coefficients are sufficiently smooth, and $u(x,t)$ is an L^1 solution of the problem (1.1), (1.2), then the primitive defined by

$$v(x,t) := \int_{-\infty}^{x} u(\xi,t)\,\mathrm{d}\xi, \quad t \in (0,T], \tag{1.3}$$

is a solution of the local initial value problem

$$v_t + \Phi(v)_x = A(v_x)_x, \quad x \in \mathbb{R}, \quad t \in (0,T], \tag{1.4}$$

$$v(x,0) = v_0(x), \quad x \in \mathbb{R}, \quad v_0(x) := \int_{-\infty}^{x} u_0(\xi)\,\mathrm{d}\xi. \tag{1.5}$$

As a nonlinear but local PDE, (1.4) is more amenable to well-posedness and numerical analysis. In this work we use that the transformation to the local equation (1.4) is also possible in the strongly degenerate case, in which solutions of (1.1) are usually discontinuous and need to be defined as weak solutions. We prove that any weak solution is also an entropy solution. This property allows us to use available L^1 stability and uniqueness results in the framework of entropy solutions.

The core, and essential novelty, of this contribution is the formulation and convergence proof of a finite difference scheme for (1.1), (1.2) (in short, "u-scheme"). The scheme is based on a monotone difference scheme for the initial value problem (1.4), (1.5) (in short, "v-scheme") in the strongly degenerate case, which in turn is a special case of the schemes formulated and analyzed by Evje and Karlsen [47] for the more general doubly degenerate equation $v_t + \Phi(v)_x = B(A(v_x))_x$. The u-scheme is obtained by taking finite differences of the numerical solution values generated by the v-scheme. The v-scheme is, in particular, monotonicity preserving, so the discrete approximations for v are always monotonically increasing when the initial datum v_0 is, and therefore the u-scheme produces nonnegative solutions. Moreover, by modifications of standard compactness and Lax-Wendroff-type arguments it is proved that the numerical approximations generated by the u-scheme converge to the unique weak solution of (1.1), (1.2). An appealing feature is that the primitive (1.3) never needs to be calculated explicitly (except for the computation of v_0). Numerical examples illustrate the behaviour of solutions of (1.1), (1.2), and recorded error histories demonstrate the convergence of the v- and u-schemes.

1.1.2 Assumptions

We assume that u_0 has compact support, and that there exists a constant $\mathscr{M}$ such that

$$\mathrm{TV}(u_0) < \mathscr{M}. \tag{1.6}$$

We also need that $\Phi \in C^2(\mathbb{R})$, and that Φ has exactly one maximum:

$$\exists v^* > 0: \quad \Phi'(v^*) = 0, \quad \Phi'(v) > 0 \text{ for } v < v^*, \quad \Phi'(v) < 0 \text{ for } v > v^*. \tag{1.7}$$

This assumption is introduced to facilitate some of the steps of our analysis; it is, however, not essential. In fact, in our convergence analysis of Section 1.4 we need to discuss the local behaviour of the numerical solution for v close to where it includes the value v^* since that value is critical in the definition of the numerical flux. If we employ a function Φ that has several separate extrema, then the locations of solution values including extrema are spatially well separated since the discrete analogue of v_x is bounded, and the techniques of Section 1.4 can be extended to that case in a straightforward manner. We recall that the function A is defined via

$$A(u) := \int_0^u a(s)\,\mathrm{d}s, \quad \text{where } a(u) \geq 0 \text{ for } u \in \mathbb{R}.$$

The assumptions on A are the following:

$$A(s) \to \infty \quad \text{as } s \to \infty; \quad \exists M_a > 0: \quad a(s) < M_a \quad \text{for all } s \in \mathbb{R}. \tag{1.8}$$

Our analysis is restricted to a finite final time T, since some of the constants appearing in the convergence analysis, which serves here as an existence proof, actually depend on T. The L^∞ bound on u is, however, independent of T.

1.1.3 Motivation

Equation (1.1), or some specific cases of it, were studied in a series of papers [3, 42, 90, 91, 92, 93], in all of which it is assumed that $a(u) = 0$ at most at isolated values of u, so that it is always ensured that $A'(u) > 0$ for $u \geq 0$. The interpretation of (1.1) as a model of the aggregation of populations (e.g., of animals) can be illustrated as follows. Assume that $u(x,t)$ is the density of the population under study, and consider the equation

$$u_t + \left(-k \left[\int_{-\infty}^{x} u(y,t)\,\mathrm{d}y - \int_{x}^{\infty} u(y,t)\,\mathrm{d}y \right] u \right)_x = A(u)_{xx}, \quad k > 0. \tag{1.9}$$

Here, the convective term provides a mechanism that moves $u(x,t)$ to the right (respectively, to the left) if

$$\int_{-\infty}^{x} u(y,t)\,\mathrm{d}y < \int_{x}^{\infty} u(y,t)\,\mathrm{d}y \quad \text{(respectively, } \ldots > \ldots\text{)}.$$

In other words, an animal will move to the right (respectively, left) if the total population to its right is larger (respectively, smaller) than to its left. Now assume that the initial population is finite and define

$$C_0 := \int_{\mathbb{R}} u_0(x)\,\mathrm{d}x. \tag{1.10}$$

It is then clear that (1.9) is an example of (1.1) if $\Phi'(v) = -k(2v - C_0)$, i.e.,

$$\Phi(v) = -kv(v - C_0) + \text{const.} \tag{1.11}$$

The aggregation mechanism is balanced by nonlinear diffusion described by the term $A(u)_{xx}$, termed density-dependent dispersal in mathematical ecology. A novel feature addressed by the present analysis is a "threshold effect", i.e. dispersal only sets on when the density u exceeds a critical value $u_c > 0$. The underlying idea is that the individuals, animals or humans, would react to variations of the local density only if that density exceeds a critical value. A similar "behavioristic" motivation of degenerate diffusion was advanced in the context of a traffic model, see [27, 101]. This effect is considered in the present model since A may degenerate on intervals. For example, for a constant $a_0 > 0$ we may consider

$$a(u) = \begin{cases} 0 & \text{for } u \le u_c, \\ a_0 & \text{for } u > u_c, \end{cases} \quad \text{i.e.,} \quad A(u) = \begin{cases} 0 & \text{for } u \le u_c, \\ a_0(u - u_c) & \text{for } u > u_c. \end{cases} \tag{1.12}$$

To illustrate some of the consequences of the presence of a strongly degenerating diffusion term, and to compare our findings with the most recent results obtained for multi-dimensional aggregation equations, let us consider a strongly degenerating integrated diffusion coefficient $A(u)$ and the local degenerate parabolic PDE

$$u_t + f(x,t,u)_x = A(u)_{xx}, \quad (x,t) \in \Pi_T; \quad u(x,0) = u_0(x), \quad x \in \mathbb{R}, \tag{1.13}$$

where f should depend smoothly on x and u. It is well known that even in the absence of a convective term ($f \equiv 0$), i.e., for the problem

$$u_t = A(u)_{xx}, \quad (x,t) \in \Pi_T; \quad u(x,0) = u_0(x), \quad x \in \mathbb{R}, \tag{1.14}$$

solutions of (1.13) may form discontinuities from smooth initial data in finite time due to the strong degeneracy of $A(u)$. The appearance of discontinuities motivates why solutions of strongly degenerate parabolic PDEs are studied as weak solutions. However, the appearance of discontinuities solely due to degenerate diffusion does not necessarily require the introduction of an entropy solution concept to ensure uniqueness. In fact, the uniqueness in L^1 of weak solutions of (1.14) is a classical result [23]. This result carries

over to such cases of (1.13) that can be transformed to (1.14), for example the linear case $f(x,u) = \alpha u$, where $\alpha \in \mathbb{R}$ is a constant, or may possibly depend on x and t (in the latter case, restrictions on the choice of $\alpha(x)$ may apply).

This discussion motivates why we expect solutions of the present problem (1.1), (1.2) to form discontinuities even from smooth initial data, so this problem should be studied in a suitably defined space of weak solutions. We may write (1.1) as

$$u_t + \big(\Phi'\big(v(x,t)\big)u\big)_x = A(u)_{xx}. \tag{1.15}$$

In this work we demonstrate that for the present equation (1.15) weak solutions are entropy solutions. The main importance of identifying weak solutions as entropy solutions lies in the easy access to stability and uniqueness results for entropy solutions (see [32, 62]) which can be applied to (1.1), (1.2), as will be done in Section 1.3.2.

1.1.4 Related work

More recently, aggregation equations of the form

$$u_t + \nabla \cdot (u \nabla K * u) = \Delta A(u) \tag{1.16}$$

have seen an enormous amount of interest, where the typical case is $A \equiv 0$. Here, K denotes an interaction potential, and $K * u$ denotes spatial convolution. The nonlocal and diffusive terms account for long-range and short-range interactions, respectively, as is emphasized in [28]. The derivation of (1.16) from microscopic interacting particle systems and related models, and for particular choices of K and A, is presented in [16, 22, 28, 85, 86]. Related models also include equations with fractional dissipation that cannot be cast in the form (1.16), see e.g. [78, 79].

The essential research problem associated with (1.16) (or variants of this equation) is the well-posedness of this equation together with bounded initial data $u(x,0) = u_0(x)$ for $x \in \mathbb{R}^d$, where d denotes the number of space dimensions. While the short-time existence of a unique smooth solution for smooth initial data is known in most situations, one wishes to determine criteria in terms of the functions K and A (or related diffusion terms), and possibly of u_0, that either ensure that smooth solutions exist globally in time, or that compel that solutions of (1.16) will blow up in finite time. This problem is analyzed in [8, 15, 16, 17, 18, 19, 22, 28, 31, 76, 78, 79, 82] (this list is far from being complete).

Here and it what follows, "blow-up" of a solution refers to L^∞ norm blow-up (as opposed to the finite time loss of classical regularity generic to problems with degenerate diffusion). The occurrence of blow-up was analyzed in terms of the properties of K for $A \equiv 0$ in [16, 17]; if K is radial, i.e., $K = K(|x|)$, then blow-up occurs if the Osgood condition for the characteristic ODEs is violated, as occurs e.g. for $K(x) = \exp(-|x|)$, while for a

C^2 kernel this does not occur [16]. Li and Rodrigo [78, 79] consider this particular kernel and describe the circumstances under which blow-up occurs if the aggregation equation is equipped with fractional diffusion. Special cases of (1.16) have also been studied in the context of Patlak-Keller-Segel models, where K is the fundamental solution to an elliptic PDE (see e.g. [8, 21]).

We can write (1.1) as a one-dimensional version of (1.16) only in very special cases. However, and as was already pointed out in [91], (1.9) can be written as

$$u_t + (u\tilde{K} * u)_x = A(u)_{xx} \tag{1.17}$$

with the odd kernel $\tilde{K}(x) = -k\,\mathrm{sgn}(x)$. Equation (1.17), or equivalently, (1.1) with Φ given by (1.11), becomes a one-dimensional example of (1.16) if we observe that $\tilde{K} * u = K' * u$, where K' denotes the derivative of K, if we choose the even kernel

$$K(x) = -k|x| + C, \tag{1.18}$$

where C is a constant. We can write this as $K(x) = -\kappa(|x|)$ for $\kappa(r) = r - C$. Suppose that one uses this kernel in the multi-dimensional equation (1.16). It is then straightforward to verify that in absence of dispersal ($A \equiv 0$), the kernel (1.18) satisfies the integral condition for blow-up in finite time, see [16]. One result of our analysis is then that the condition (1.8) is sufficient to ensure that L^∞ blow-up of solutions of (1.1) does not occur.

In fact, in the context of aggregation models that are based either on (1.1) or on the more recently studied equation (1.16), the present work is the first that incorporates a strongly degenerate diffusion term, i.e., involves a function $A(u)$ that is flat on a u-interval of positive length. So far, diffusion terms that have been considered in (1.1) degenerate at most at isolated u-values. Nagai and Mimura [91] studied the Cauchy problem for equation (1.1) under the assumptions $A(0) = 0$, $A'(u) > 0$ being an odd function. The initial function for the Cauchy problem in [91] is assumed to be bounded, nonnegative and integrable. Nagai and Mimura [91] prove existence and uniqueness of a bounded and continuous solution to the initial value problem. In [92] the asymptotic behaviour of solutions to the same problem was studied for the specific choice

$$A(u) = u^m, \quad m > 1. \tag{1.19}$$

It seems that the analysis of (1.16) with degenerate diffusion has just started. Li and Zhang [82] study this equation in one space dimension for the diffusion function $A(u) = u^3/3$, which degenerates at $u = 0$ only. On the other hand, the numerical simulations presented herein show that under strongly degenerate diffusion, typical features of the aggregation phenomenon such as "clumped" solutions with very sharp edges [105] appear.

Let us briefly mention some of the recent results concerning (1.16). If diffusion is absent ($A \equiv 0$), (1.16) becomes an inviscid nonlocal transport law, which are well known to

be have better regularity properties than general quasi-linear conservation laws. In particular, Laurent [76] and Bertozzi and Laurent [17] show that if the initial condition is smooth, then solutions of (1.16) remain smooth for as long as the L^p norms remains bounded. In particular, discontinuities can only occur if they were present in the initial data. Moreover, according to [82], the addition of nonlinear diffusion will cause higher regularity of weak solutions to be lost in finite time, i.e., the spatial gradient of the solution will experience L^∞ blow-up in finite time. This contrasts with the expected solution behaviour of (1.1), (1.2) described in Section 1.1.3, namely that strong discontinuities form from smooth data.

Regarding uniqueness of weak solutions, it has been shown that in dimensions two and higher, entropy conditions are not required to ensure that weak solutions to (1.16) are unique. For the inviscid case, see Bertozzi and Brandman [15] or Bertozzi et al. [18]. For the case with diffusion, uniqueness is shown by Bertozzi and Slepčev in [19] and in more generality by Bedrossian et al. [8]. These results are consistent with ours.

1.1.5 Outline of the chapter

The remainder of this chapter is organized as follows. In Section 1.2 we state the definition of weak and entropy solutions of (1.1), (1.2). While it is standard to verify that any entropy solution is a weak solution, we able to prove that for the present equation, any weak solution is an entropy solution. In Section 1.3.1 we state jump conditions that can be derived from the definition of weak solutions, and in Section 1.3.2 we prove the uniqueness of a weak solution, using that any weak solution is, in fact, an entropy solution. Section 1.4 presents a convergence analysis for the u-scheme. In Section 1.4.1, the schemes are described. Section 1.4.2 contains a series of lemmas stating uniform estimates on the numerical approximations generated by the v- and the u-schemes, which allow to employ standard compactness arguments to deduce that both schemes converge to the unique weak solution. The final convergence result (Theorem 1.4.1) and its proof are presented in Section 1.4.3. This proof follows a standard Lax-Wendroff argument. A finite speed of propagation property is proven in 1.4.4. Some numerical examples are presented in Section 1.5.

1.2 Definition of a weak solution

Definition 1.2.1 *A measurable function u is said to be a* weak solution *of the initial value problem* (1.1)*,* (1.2) *if it satisfies the following conditions:*

1. *We have* $u \in L^\infty(\Pi_T) \cap L^\infty(0,T;BV(\mathbb{R}))$*, and* $A(u) \in L^2(0,T;H^1(\mathbb{R}))$*, where* $\Pi_T := \mathbb{R} \times (0,T)$.

2. *The initial condition* (1.2) *is satisfied in the following sense:*

$$\lim_{t\downarrow 0}\int_{\mathbb{R}}\big|u(x,t)-u_0(x)\big|\,\mathrm{d}x=0. \tag{1.20}$$

3. *If $v(x,t)$ is defined by* (1.3)*, then the following equality is satisfied for all test functions $\phi\in C_0^\infty(\Pi_T)$:*

$$\iint_{\Pi_T}\Big\{u\big(\phi_t+\Phi'(v)\phi_x\big)+A(u)\phi_{xx}\Big\}\,\mathrm{d}x\,\mathrm{d}t=0. \tag{1.21}$$

Definition 1.2.2 *A measurable, nonnegative function u is an* entropy solution *of* (1.1), (1.2) *if it satisfies items (1) and (2) of Definition 1.2.1 and if for all nonnegative test functions $\varphi\in C_0^\infty(\Pi_T)$, the following entropy inequality is satisfied:*

$$\begin{aligned}\forall k\in\mathbb{R}:\iint_{\Pi_T}\Big\{&|u-k|\big(\varphi_t+\Phi'(v)\varphi_x\big)-\operatorname{sgn}(u-k)uk\Phi''(v)\varphi\\&+\big|A(u)-A(k)\big|\varphi_{xx}\Big\}\,\mathrm{d}x\,\mathrm{d}t\geq 0.\end{aligned} \tag{1.22}$$

It is straightforward to check that an entropy solution of the initial value problem (1.1), (1.2) is a weak solution.

Lemma 1.2.1 *Assume that u is an entropy solution of the initial value problem* (1.1), (1.2) *(cf. Definition 2.4.1). Then u is a weak solution (cf. Definition 1.2.1).*

Proof. Choosing $k\geq\|u\|_{L^\infty(\Pi_T)}$ in (1.22) we obtain

$$\iint_{\Pi_T}\Big\{-(u-k)\big(\phi_t+\Phi'(v)\phi_x\big)-A(u)\phi_{xx}\Big\}\,\mathrm{d}x\,\mathrm{d}t\geq -k\iint_{\Pi_T}u\Phi''(v)\phi\,\mathrm{d}x\,\mathrm{d}t$$

or equivalently,

$$\begin{aligned}&\iint_{\Pi_T}\Big\{u\big(\phi_t+\Phi'(v)\phi_x\big)+A(u)\phi_{xx}\Big\}\,\mathrm{d}x\,\mathrm{d}t\\&\leq k\iint_{\Pi_T}\Big\{\phi_t+\big(\Phi'(v)\phi\big)_x\Big\}\,\mathrm{d}x\,\mathrm{d}t=0.\end{aligned} \tag{1.23}$$

On the other hand, since we look for nonnegative solutions, it suffices to set $k=0$ in (1.22) to deduce that we always have

$$\iint_{\Pi_T}\Big\{u\big(\phi_t+\Phi'(v)\phi_x\big)+A(u)\phi_{xx}\Big\}\,\mathrm{d}x\,\mathrm{d}t\geq 0.$$

Combining this with (1.23) we see that u satisfies (1.21). □

The following lemma states that conversely, any weak solution of the initial value problem (1.1), (1.2) is an entropy solution. Lemma 1.2.2 is inspired by Carrillo [30] and Kobayasi [66, Lemmas 3.1 and 3.3].

Lemma 1.2.2 *Let u be a weak solution of problem* (1.1)*,* (1.2)*, then u is also an entropy solution.*

Proof. Let us define $\alpha(x,t) := \Phi'(v(x,t))$. Then we recall that u is a weak solution of (1.1) if for all test functions $\phi \in C_0^\infty(\Pi_T)$,

$$\iint_{\Pi_T} \Big\{ u\big(\phi_t + \alpha(x,t)\phi_x\big) + A(u)\phi_{xx} \Big\} \,\mathrm{d}x\,\mathrm{d}t = 0$$

or equivalently,

$$\iint_{\Pi_T} \Big\{ u\big(\phi_t + \alpha(x,t)\phi_x\big) - A(u)_x\phi_x \Big\} \,\mathrm{d}x\,\mathrm{d}t = 0. \tag{1.24}$$

In what follows we will utilize the functions defined by

$$H_0(x) := \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x \le 0, \end{cases} \quad H_1(x) := \begin{cases} 1 & \text{if } x \ge 0, \\ 0 & \text{if } x < 0, \end{cases} \quad H_\varepsilon(x) := \begin{cases} 1 & \text{if } x > \varepsilon, \\ x/\varepsilon & \text{if } x \in [0,\varepsilon], \\ 0 & \text{if } x < 0. \end{cases}$$

and the multi-valued function (see [30, 66])

$$H(x) := \begin{cases} 1 & \text{if } x > 0, \\ [0,1] & \text{if } x = 0, \\ 0 & \text{if } x < 0. \end{cases}$$

To simplify the argument, let us concentrate on the case of a single u-interval $[m,M]$ of degeneracy, assuming that $A'(s) = 0$ for $s \in [m,M]$ and $A'(s) > 0$ for $s \notin [m,M]$, where $0 \le m, M < \infty$. Now let us use as a test function $\phi(x,t) = H_\varepsilon(A(u) - A(k))\varphi(x,t)$ with $k \notin [m,M]$, where φ is an admissible test function. Following the proof of Lemma 2.4 in [62] we find

$$\begin{aligned} \iint_{\Pi_T} \Big\{ |u-k|^+ \varphi_t + H_0(u-k)\big(\alpha(x,t)(u-k) - A(u)_x\big)\varphi_x \\ - H_0(u-k)\alpha_x(x,t)k\varphi \Big\} \,\mathrm{d}x\,\mathrm{d}t \ge 0 \quad \text{for } k \notin [m,M], \end{aligned} \tag{1.25}$$

where $|z|^+ := H_0(z)z$. Since $M < \infty$ we can construct a sequence $\{s_n\}_{n\in\mathbb{N}}$ such that $s_n > M$, $s_n \to M$ and $H_0(u - s_n) \to H_0(u - M)$ as $n \to \infty$. Setting $k = s_n$ in (1.25) and sending $n \to \infty$, we get

$$\begin{aligned} \iint_{\Pi_T} \Big\{ |u-M|^+ \varphi_t + \alpha(x,t)|u-M|^+ \varphi_x - H_0(u-M)A(u)_x\varphi_x \\ - H_0(u-M)\alpha_x(x,t)M\varphi \Big\} \,\mathrm{d}x\,\mathrm{d}t \ge 0. \end{aligned} \tag{1.26}$$

Similarly, we may construct a sequence $\{s_n\}_{n\in\mathbb{N}}$ such that $s_n < m$, $s_n \to m$ and $H_0(u-s_n) \to H_1(u-m)$ as $n\to\infty$. Setting $k = s_n$ in (1.25) and sending $n\to\infty$ yields

$$\begin{aligned}\iint_{\Pi_T}\Big\{&|u-m|^+\varphi_t + \alpha(x,t)|u-m|^+\varphi_x - H_1(u-m)A(u)_x\varphi_x\\ &- H_1(u-m)\alpha_x(x,t)m\varphi\Big\}\,\mathrm{d}x\,\mathrm{d}t \geq 0.\end{aligned} \tag{1.27}$$

Now, we take in the entropy inequality (1.27) a test function $\varphi(x,t) = \xi(x,t)\zeta(x,t)$, where ζ is a smooth function such that $0\leq\zeta\leq 1$ and ξ is an admissible test function, and in (1.26) we use $\varphi(x,t) = \xi(x,t)(1-\zeta(x,t))$. Adding both resulting expressions we obtain the inequality $I_1+I_2+I_3+I_4\geq 0$ with the following terms, where we drop the argument (x,t) wherever convenient:

$$\begin{aligned}
I_1 &:= \iint_{\Pi_T}\Big\{\big(|u-m|^+ - |u-M|^+\big)(\xi\zeta)_t + |u-M|^+\xi_t\Big\}\,\mathrm{d}x\,\mathrm{d}t,\\
I_2 &:= \iint_{\Pi_T}\Big\{\alpha\big(|u-m|^+ - |u-M|^+\big)(\xi\zeta)_x + \alpha|u-M|^+\xi_x\Big\}\,\mathrm{d}x\,\mathrm{d}t,\\
I_3 &:= \iint_{\Pi_T}\Big\{\alpha_x\big(H_0(u-M)M - H_1(u-m)m\big)\xi\zeta - \alpha_x H_0(u-M)M\xi\Big\}\,\mathrm{d}x\,\mathrm{d}t,\\
I_4 &:= -\iint_{\Pi_T}\big(H_1(u-m) - H_0(u-M)\big)A(u)_x(\xi\zeta)_x\,\mathrm{d}x\,\mathrm{d}t\\
&\quad - \iint_{\Pi_T} H_0(u-M)A(u)_x\xi_x\,\mathrm{d}x\,\mathrm{d}t.
\end{aligned} \tag{1.28}$$

Assume now that $\rho_n = \rho_n(x)$ is a standard sequence of mollifier functions in $\mathbb{R}$, and let us define $|u-m|_n^+ := |u-m|^+ * \rho_n$ and $|u-M|_n^+ := |u-M|^+ * \rho_n$ for $n\in\mathbb{N}$. Now we select the function $\zeta = \zeta(x,t)$ defined by

$$\zeta = \zeta_{n,\varepsilon} := H_\varepsilon\big(|u-m|_n^+ + m - s - |u-M|_n^+\big), \quad s\in[m,M].$$

Let us denote the versions of I_p obtained by replacing $|\cdot|^+$ by $|\cdot|_n^+$ and $\zeta = \zeta_{n,\varepsilon}$ by $I_p(n,\varepsilon)$, $p = 1,\dots,4$. Since m and s are constant and $\xi\zeta_{n,\varepsilon}$ has compact support, we get after an integration by parts

$$\begin{aligned}
I_1(n,\varepsilon) &= \iint_{\Pi_T}\Big\{\big(|u-m|_n^+ + m - s - |u-M|_n^+\big)(\xi\zeta_{n,\varepsilon})_t + |u-M|_n^+\xi_t\Big\}\,\mathrm{d}x\,\mathrm{d}t\\
&= \iint_{\Pi_T}\Big\{-H_\varepsilon\big(|u-m|_n^+ + m - s - |u-M|_n^+\big)\\
&\qquad\qquad \times\big(|u-m|_n^+ + m - s - |u-M|_n^+\big)_t\xi + |u-M|_n^+\xi_t\Big\}\,\mathrm{d}x\,\mathrm{d}t.
\end{aligned}$$

Taking $\varepsilon\downarrow 0$ and again integrating by parts yields

$$\begin{aligned}
I_1(n,0) &= \iint_{\Pi_T}\Big\{-\Big(\big||u-m|_n^+ + m - s - |u-M|_n^+\big|^+\Big)_t\xi + |u-M|_n^+\xi_t\Big\}\,\mathrm{d}x\,\mathrm{d}t\\
&= \iint_{\Pi_T}\Big\{\big||u-m|_n^+ + m - s - |u-M|_n^+\big|^+\xi_t + |u-M|_n^+\xi_t\Big\}\,\mathrm{d}x\,\mathrm{d}t,
\end{aligned}$$

and letting $n \to \infty$ we find that $|u-m|_n^+ + m - s - |u-M|_n^+$ converges to $|u-m|^+ + m - s - |u-M|^+$ in $L^1(\mathbb{R})$ and $\zeta_{n,0} = H_0(|u-m|_n^+ + m - s - |u-M|_n^+)$, or at least a subsequence, converges weak-$*$ to some $\tilde{H}$ in $L^\infty(\Pi_T)$. Since H is maximal monotone, it follows that $\tilde{H} \in H(|u-m|^+ + m - s - |u-M|^+)$. Noting that $H(w)w = H_0(w)w$ for any function w, we arrive at

$$\begin{aligned} I_1 &= \iint_{\Pi_T} \Big\{ \big| |u-m|^+ + m - s - |u-M|^+ \big|^+ + |u-M|^+ \Big\} \xi_t \,\mathrm{d}x\,\mathrm{d}t \\ &= \iint_{\Pi_T} |u-s|^+ \xi_t \,\mathrm{d}x\,\mathrm{d}t. \end{aligned} \tag{1.29}$$

Next, we deal with I_4. Since $A'(u) = 0$ for $u \in [m, M]$, we have

$$H_0(u-M)A(u)_x = H_0(u-s)A(u)_x = H_1(u-m)A(u)_x$$

for all $s \in [m, M]$, which gives

$$I_4 = \lim_{n\to\infty} \lim_{\varepsilon \downarrow 0} I_4(n,\varepsilon) = -\int_{\Pi_T} H_0(u-s)A(u)_x \xi_x \,\mathrm{d}x\,\mathrm{d}t. \tag{1.30}$$

To deal with I_2, we proceed in a similar way as for I_1. We get

$$\begin{aligned} I_2(n,\varepsilon) = \iint_{\Pi_T} \Big\{ & \alpha\big(|u-m|_n^+ + m - s - |u-M|_n^+\big)(\xi\,\zeta_{n,\varepsilon})_x \\ & + (s-m)\alpha(\xi\,\zeta_{n,\varepsilon})_x + \alpha|u-M|_n^+ \xi_x \Big\} \,\mathrm{d}x\,\mathrm{d}t \\ = -\iint_{\Pi_T} & \alpha_x H_\varepsilon\big(|u-m|_n^+ + m - s - |u-M|_n^+\big) \\ & \times \big(|u-m|_n^+ + m - s - |u-M|_n^+\big)\xi \,\mathrm{d}x\,\mathrm{d}t \\ - \iint_{\Pi_T} & \alpha\big(|u-m|_n^+ + m - s - |u-M|_n^+\big)_x \\ & \times H_\varepsilon\big(|u-m|_n^+ + m - s - |u-M|_n^+\big)\xi \,\mathrm{d}x\,\mathrm{d}t \\ + \iint_{\Pi_T} & \alpha|u-M|_n^+ \xi_x \,\mathrm{d}x\,\mathrm{d}t + \iint_{\Pi_T} \alpha(s-m)(\xi\,\zeta_{n,\varepsilon})_x \,\mathrm{d}x\,\mathrm{d}t. \end{aligned}$$

Taking $\varepsilon \downarrow 0$ we get, after integration by parts,

$$\begin{aligned}
I_2(n,0) &= -\iint_{\Pi_T} \alpha_x \big| |u-m|_n^+ + m - s - |u-M|_n^+ \big|^+ \xi \,\mathrm{d}x\,\mathrm{d}t \\
&\quad - \iint_{\Pi_T} \alpha \Big(\big| |u-m|_n^+ + m - s - |u-M|_n^+ \big|^+ \Big)_x \xi \,\mathrm{d}x\,\mathrm{d}t \\
&\quad + \iint_{\Pi_T} \alpha |u-M|_n^+ \xi_x \,\mathrm{d}x\,\mathrm{d}t + \lim_{\varepsilon\downarrow 0} \iint_{\Pi_T} \alpha (s-m)(\xi\,\zeta_{n,\varepsilon})_x \,\mathrm{d}x\,\mathrm{d}t \\
&= -\iint_{\Pi_T} \alpha_x \big| |u-m|_n^+ + m - s - |u-M|_n^+ \big|^+ \xi \,\mathrm{d}x\,\mathrm{d}t \\
&\quad + \iint_{\Pi_T} \alpha_x \big| |u-m|_n^+ + m - s - |u-M|_n^+ \big|^+ \xi \,\mathrm{d}x\,\mathrm{d}t \\
&\quad + \iint_{\Pi_T} \alpha \big| |u-m|_n^+ + m - s - |u-M|_n^+ \big|^+ \xi_x \,\mathrm{d}x\,\mathrm{d}t \\
&\quad + \iint_{\Pi_T} \alpha |u-M|_n^+ \xi_x \,\mathrm{d}x\,\mathrm{d}t + \lim_{\varepsilon\downarrow 0} \iint_{\Pi_T} \alpha (s-m)(\xi\,\zeta_{n,\varepsilon})_x \,\mathrm{d}x\,\mathrm{d}t,
\end{aligned}$$

where the two first terms obviously cancel. Sending $n \to \infty$ and proceeding like in the term I_1 we arrive at

$$\begin{aligned}
I_2 &= \iint_{\Pi_T} \alpha \Big(|u-M|^+ + \big| |u-m|^+ + m - s - |u-M|^+ \big|^+ \Big) \xi_x \,\mathrm{d}x\,\mathrm{d}t \\
&\quad + \lim_{\substack{\varepsilon\downarrow 0 \\ n\to\infty}} \iint_{\Pi_T} (s-m)\alpha (\xi\,\zeta_{n,\varepsilon})_x \,\mathrm{d}x\,\mathrm{d}t \\
&= \iint_{\Pi_T} \alpha |u-s|^+ \xi_x \,\mathrm{d}x\,\mathrm{d}t + \lim_{\substack{\varepsilon\downarrow 0 \\ n\to\infty}} \iint_{\Pi_T} (s-m)\alpha (\xi\,\zeta_{n,\varepsilon})_x \,\mathrm{d}x\,\mathrm{d}t. \qquad (1.31)
\end{aligned}$$

The last term of the last expression will be incorporated into the analysis of I_3. In fact, taking into account that

$$\begin{aligned}
I_3(n,\varepsilon) = \iint_{\Pi_T} \alpha_x \Big\{ &\big(H_0(u-M)M - H_1(u-m)m \big) \\
&\times H_\varepsilon \big(|u-m|_n^+ + m - s - |u-M|_n^+ \big) - H_0(u-M)M \Big\} \xi \,\mathrm{d}x\,\mathrm{d}t
\end{aligned}$$

and that

$$\lim_{\substack{\varepsilon\downarrow 0 \\ n\to\infty}} \iint_{\Pi_T} (s-m)\alpha (\xi\,\zeta_{n,\varepsilon})_x \,\mathrm{d}x\,\mathrm{d}t = -\lim_{\substack{\varepsilon\downarrow 0 \\ n\to\infty}} \iint_{\Pi_T} (s-m)\alpha_x \xi\,\zeta_{n,\varepsilon} \,\mathrm{d}x\,\mathrm{d}t,$$

we obtain

$$
\begin{aligned}
&I_3 + \lim_{\substack{\varepsilon\downarrow 0\\ n\to\infty}} \iint_{\Pi_T} (s-m)\alpha(\xi\,\zeta_{n,\varepsilon})_x \,\mathrm{d}x\,\mathrm{d}t \\
&= \lim_{\substack{\varepsilon\downarrow 0\\ n\to\infty}} \left(I_3(n,\varepsilon) - \iint_{\Pi_T} (s-m)\alpha_x \xi\,\zeta_{n,\varepsilon}\,\mathrm{d}x\,\mathrm{d}t \right) \\
&= \lim_{\substack{\varepsilon\downarrow 0\\ n\to\infty}} \iint_{\Pi_T} \Big\{ \alpha_x \big(H_0(u-M)M - s + m - H_1(u-m)m \big)\,\xi\,\zeta_{n,\varepsilon} - \alpha_x H_0(u-M)M\xi \Big\}\,\mathrm{d}x\,\mathrm{d}t.
\end{aligned}
$$

Proceeding as in the cases I_1 and I_2 we find

$$
\begin{aligned}
&I_3 + \lim_{\substack{\varepsilon\downarrow 0\\ n\to\infty}} \iint_{\Pi_T} (s-m)\alpha(\xi\,\zeta_{n,\varepsilon})_x \,\mathrm{d}x\,\mathrm{d}t \\
&= \iint_{\Pi_T} \alpha_x \Big\{ \big(H_0(u-M)M + m - s - H_1(u-m)m \big) \\
&\qquad\qquad \times \tilde{H}\big(|u-m|^+ + m - s - |u-M|^+ \big) - H_0(u-M)M \Big\} \xi\,\mathrm{d}x\,\mathrm{d}t.
\end{aligned}
$$

If $u > M$, then the expression in curled brackets in the last integrand equals

$$
\{\dots\} = (M-s)\tilde{H}(M-s) - M = -s = -\tilde{H}(u-s)s.
$$

Likewise, for each of the cases $M \geq u > s > m$, $M > s \geq u > m$, and $M > s > u \geq m$ we verify that $\{\dots\} = -\tilde{H}(u-s)s$, and also for $m > u$ we obtain

$$
\{\dots\} = (m-s)\tilde{H}(m-s) = 0 = -\tilde{H}(u-s)s,
$$

and finally, the result is also valid if $s = m$ or $s = M$. We therefore conclude that

$$
I_3 + \lim_{\substack{\varepsilon\downarrow 0\\ n\to\infty}} \iint_{\Pi_T} (s-m)\alpha(\xi\,\zeta_{n,\varepsilon})_x \,\mathrm{d}x\,\mathrm{d}t = -\int_{\Pi_T} \alpha_x \tilde{H}(u-s)s\xi\,\mathrm{d}x\,\mathrm{d}t. \tag{1.32}
$$

Now, combining (1.29)–(1.32), we obtain the inequality

$$
\begin{aligned}
\iint_{\Pi_T} \Big\{ &|u-s|^+ \big(\xi_t + \alpha(x,t)\xi_x \big) - \alpha_x(x,t)\tilde{H}(u-s)s\xi \\
&- H_0(u-s)A(u)_x\xi_x \Big\}\,\mathrm{d}x\,\mathrm{d}t \geq 0 \quad \text{for all } s \in [m,M].
\end{aligned}
$$

Now, for any $s \in [m,M)$ there exists a sequence $\{s_n\}_{n\in\mathbb{N}}$ such that $s_n < s < M$ and $s_n \to s$. Then $\tilde{H}(u-s_n) \to H_0(u-s)$ and $H_0(u-s_n) \to H_0(u-s)$ almost everywhere. Hence we get

$$
\begin{aligned}
\iint_{\Pi_T} \Big\{ &|u-s|^+ \big(\xi_t + \alpha(x,t)\xi_x \big) - \alpha_x(x,t)H_0(u-s)s\xi \\
&- H_0(u-s)A(u)_x\xi_x \Big\}\,\mathrm{d}x\,\mathrm{d}t \geq 0 \quad \text{for all } s \in [m,M].
\end{aligned} \tag{1.33}
$$

This proof uses only that α_x is bounded, which is indeed the case for our choice $\alpha(x,t) = \Phi'(v(x,t))$, since Φ is assumed to be smooth and $v_x(x,t) = u(x,t)$ is bounded. On the other hand, considering in the weak formulation a test function $\phi(x,t) = H_\varepsilon(A(k)-A(u))\varphi(x,t)$ with $k \notin [m,M]$ and following essentially the same steps as before we find

$$\begin{aligned}\iint_{\Pi_T} \Big\{ |s-u|^+ \big(\xi_t + \alpha(x,t)\xi_x\big) + \alpha_x(x,t) H_0(s-u) s\xi \\ + H_0(s-u)A(u)_x\xi_x \Big\} \,\mathrm{d}x\,\mathrm{d}t \geq 0 \quad \text{for all } s \in [m,M].\end{aligned} \tag{1.34}$$

Adding (1.33) and (1.34) we get

$$\begin{aligned}\iint_{\Pi_T} \Big\{ |u-s| \big(\xi_t + \alpha(x,t)\xi_x\big) - \alpha_x(x,t)\,\mathrm{sgn}(u-s) s\xi \\ - \mathrm{sgn}(u-s)A(u)_x\xi_x \Big\} \,\mathrm{d}x\,\mathrm{d}t \geq 0 \quad \text{for all } s \in [m,M].\end{aligned} \tag{1.35}$$

Moreover (1.35) is valid for all $s \notin [m,M]$ (cf. [62]). This implies that any weak solution is an entropy solution. □

1.3 Jump conditions and uniqueness

1.3.1 Rankine-Hugoniot condition

Assume that u is a weak solution having a discontinuity at a point $(x_0,t_0) \in \Pi_T$ between the approximate limits u^+ and u^- of u taken with respect to $x > x_0$ and $x < x_0$, respectively. Standard results from the theory of strongly degenerate parabolic equations imply that such a discontinuity is possible only if $A(u)$ is flat for $u \in \mathscr{I}(u^-,u^+) := [\min\{u^-,u^+\}, \max\{u^-,u^+\}]$. In that case, the propagation velocity of the jump is given by the Rankine-Hugoniot condition, which is derived by standard arguments from the weak formulation (1.21):

$$s = \frac{1}{u^+ - u^-}\Big(\Phi'(v^+)u^+ - \Phi'(v^-)u^- - \big(A(u)_x\big)^+ + \big(A(u)_x\big)^-\Big). \tag{1.36}$$

Here, $(A(u)_x)^+$ and $(A(u)_x)^-$ denote the approximate limits of $A(u)_x$ taken with respect to $x > x_0$ and $x < x_0$, respectively, and v^+ and v^- denote the corresponding limits of v. Since v is continuous, we actually have $v^+ = v^-$, and (2.29) reduces to

$$s = \Phi'\big(v(x_0,t_0)\big) - \frac{(A(u)_x)^+ - (A(u)_x)^-}{u^+ - u_-}.$$

1.3.2 Uniqueness of weak solutions

The uniqueness of weak solutions is an immediate consequence of Lemma 1.2.2 and a result proved in [62] (cf. also [32]) regarding continuous dependence of entropy solutions with respect to the flux function. More precisely, we have the following theorem.

Theorem 1.3.1 *Let u and $\bar{u}$ be two weak solutions of* (1.1), (1.2) *(in the sense of Definition 1.2.1) with initial data u_0 and $\bar{u}_0$, respectively. Then there exists a constant $C = C(\max|\Phi'|)$ such that*

$$\big\|u(\cdot,t)-\bar{u}(\cdot,t)\big\|_{L^1(\mathbb{R})} \le C\|u_0-\bar{u}_0\|_{L^1(\mathbb{R})}, \qquad \forall t\in(0,T].$$

In particular, weak solutions of (1.1), (1.2) *are unique.*

Proof. According to Lemma 1.2.2, u and $\bar{u}$ are entropy solutions (in the sense of Definition 2.4.1) with initial data u_0 and $\bar{u}_0$, respectively. To be able to apply the L^1 stability and uniqueness results from [32, 62], we rewrite the equations satisfied by u and $\bar{u}$ as

$$u_t + \big(V(x,t)u\big)_x = A(u)_{xx}, \qquad V(x,t) := \Phi'\left(\int_{-\infty}^{x} u(y,t)\,\mathrm{d}y\right),$$

with initial data $u(0,x)=u_0(x)$ and

$$\bar{u}_t + (\bar{V}(x,t)\bar{u})_x = A(\bar{u})_{xx}, \qquad \bar{V}(x,t) := \Phi'\left(\int_{-\infty}^{x} \bar{u}(y,t)\,\mathrm{d}y\right).$$

with initial data $\bar{u}(0,x)=\bar{u}_0(x)$, respectively. Keeping in mind that u and $\bar{u}$ are of bounded variation, i.e., $u,\bar{u}\in L^\infty(0,T;BV(\mathbb{R}))$, we now may apply Theorem 1.3 in [62] to conclude that there exists a constant C such that

$$\begin{aligned}\big\|u(\cdot,t)-\bar{u}(\cdot,t)\big\|_{L^1(\mathbb{R})} &\le \|u_0-\bar{u}_0\|_{L^1(\mathbb{R})} + \int_0^t \big|V_x(x,s)-\bar{V}_x(x,s)\big|\,\mathrm{d}s \\ &\quad + \int_0^t \big|V(x,s)-\bar{V}(x,s)\big|\,\mathrm{TV}(u(\cdot,s))\,\mathrm{d}s \\ &\le \|u_0-\bar{u}_0\|_{L^1(\mathbb{R})} + C\int_0^t \big|V_x(x,s)-\bar{V}_x(x,s)\big|\,\mathrm{d}s.\end{aligned}$$

Observe that

$$\int_0^t \big|V_x(x,s)-\bar{V}_x(x,s)\big|\,\mathrm{d}s \le \max|\Phi'|\int_0^t \big|u(x,s)-\bar{u}(x,s)\big|\,\mathrm{d}s,$$

so that by the Gronwall inequality we arrive at

$$\big\|u(\cdot,t)-\bar{u}(\cdot,t)\big\|_{L^1(\mathbb{R})} \le \exp\big(\max|\Phi'|\,t\big)\|u_0-\bar{u}_0\|_{L^1(\mathbb{R})}.$$

□

1.4 Convergence analysis of numerical schemes

1.4.1 Preliminaries

We define the vectors $U^n := \{u^n_{j+1/2}\}_{j\in\mathbb{Z}}$ and $V^n := \{v^n_j\}_{j\in\mathbb{Z}}$, and discretize $\mathbb{R}$ by $x_j := j\Delta x$, $j \in \mathbb{Z}$, and the time interval $[0,T]$ by $t_n = n\Delta t$, $n = 0,\dots,N$, $\Delta t := T/N$, $N \in \mathbb{N}$. We denote by $u^n_{j+1/2}$ the cell average over $I_j := [x_j, x_{j+1}]$ at time t_n and $j \in \mathbb{Z}$. We also define $\lambda := \Delta t/\Delta x$ and $\mu := \Delta t/\Delta x^2 = \lambda/\Delta x$ and wherever convenient use the spatial difference operators $\Delta_+\phi_j := \phi_{j+1} - \phi_j$, $\Delta_-\phi_j := \phi_j - \phi_{j-1}$, and

$$\Delta^2\phi_j := \Delta_+\Delta_-\phi_j = \phi_{j+1} - 2\phi_j + \phi_{j-1}.$$

We assume that the initial datum u_0 is discretized via

$$u^0_{j+1/2} := \frac{1}{\Delta x}\int_{I_j} u_0(\xi)\,d\xi, \quad j \in \mathbb{Z}.$$

Moreover, we define the operator $\mathscr{S}_{\Delta x}$ and its inverse $\mathscr{S}^{-1}_{\Delta x}$ via

$$\mathscr{S}_{\Delta x}(U^n; j) := \Delta x \sum_{l=-\infty}^{j-1} u^n_{l+1/2}, \quad \mathscr{S}^{-1}_{\Delta x}(V^n; j) := \frac{v^n_{j+1} - v^n_j}{\Delta x}. \tag{1.37}$$

Clearly, $\mathscr{S}_{\Delta x}$ and $\mathscr{S}^{-1}_{\Delta x}$ are the discrete analogues of the integral and differential operators that convert $u(\cdot,t_n)$ into $v(\cdot,t_n)$ and vice versa, respectively. Since we assume that u_0 is compactly supported, the sum in (1.37) is actually finite.

The numerical scheme for the initial value problem (1.1), (1.2) can be compactly written as follows:

$$U^{n+1} = \left[\mathscr{S}^{-1}_{\Delta x} \circ \mathscr{H} \circ \mathscr{S}_{\Delta x}\right] U^n, \quad n = 0,\dots,N-1, \tag{1.38}$$

where the basic idea is to utilize a standard scheme of the form

$$V^{n+1} = \mathscr{H}(V^n), \quad n = 0,\dots,N-1 \tag{1.39}$$

for approximate solutions of the local PDE (1.4), starting from the initial data

$$v^0_j := \Delta x \sum_{l=-\infty}^{j-1} u^0_{l+1/2} = \int_{-\infty}^{x_j} u_0(\xi)\,\mathrm{d}\xi, \quad j \in \mathbb{Z}.$$

Clearly, if C_0 is the total mass defined in (1.10), then we have that

$$0 \le v^0_j \le C_0, \quad v^0_j \le v^0_{j+1} \quad \text{for all } j \in \mathbb{Z}. \tag{1.40}$$

Let us emphasize here that (1.38) implies that

$$U^n = \big[\mathscr{S}_{\Delta x}^{-1} \circ \mathscr{H} \circ \mathscr{S}_{\Delta x}\big]^n U^0 = \big[\mathscr{S}_{\Delta x}^{-1} \circ \mathscr{H}^n \circ \mathscr{S}_{\Delta x}\big] U^0.$$

This means that for the actual computation of U^n from U^0, the operators $\mathscr{S}_{\Delta x}$ and $\mathscr{S}_{\Delta x}^{-1}$ need to be applied only once, and not for every time step.

To derive properties of the scheme (1.38), we first analyze the scheme (1.39), which is here given by the marching formula

$$v_j^{n+1} = v_j^n - \lambda \Delta_+ \big[h\big(v_{j-1}^n, v_j^n\big) - A\big(\Delta_- v_j^n/\Delta x\big)\big], \quad j \in \mathbb{Z}, \quad n = 0,1,2,\ldots, \tag{1.41}$$

where λ is subject to the CFL condition stated below, and

$$h(w,z) := \Phi(0) + \Phi_+(w) + \Phi_-(z) \tag{1.42}$$

is the Engquist-Osher flux [44], where we define the functions

$$\Phi_+(v) := \int_0^v \max\big\{0, \Phi'(s)\big\}\,\mathrm{d}s, \quad \Phi_-(v) := \int_0^v \min\big\{0, \Phi'(s)\big\}\,\mathrm{d}s. \tag{1.43}$$

We assume that Δt and Δx satisfy the CFL stability condition

$$2\lambda \max_{v\in[0,C_0]} \big|\Phi'(v)\big| + 2\mu \max_{u\in\mathbb{R}} \big|a(u)\big| \le 1. \tag{1.44}$$

The scheme for u can be written as

$$u_{j+1/2}^{n+1} = u_{j+1/2}^n - \lambda \Delta_+ G_j^n + \mu \Delta^2 A\big(u_{j+1/2}^n\big), \quad j \in \mathbb{Z}, \quad n = 0,1,2,\ldots, \tag{1.45}$$

where we define

$$G_j^n := \frac{1}{\Delta x}\Delta_+ h\big(v_{j-1}^n, v_j^n\big) = \frac{1}{\Delta x}\left(\int_{v_{j-1}^n}^{v_j^n} \Phi'_+(s)\,\mathrm{d}s + \int_{v_j^n}^{v_{j+1}^n} \Phi'_-(s)\,\mathrm{d}s\right). \tag{1.46}$$

For the ease of reference, we will refer to (1.41)–(1.43) and (1.42), (1.43), (1.45), (1.46) as "v-scheme" and "u-scheme", respectively. Both schemes are, in particular, conservative, so the total mass C_0 is preserved.

The v-scheme (1.41)–(1.43) is a special case of the scheme studied by Evje and Karlsen [47] for the more general doubly degenerate parabolic equation $v_t + \Phi(v)_x = B(A(v_x))_x$. While Evje and Karlsen prove that their scheme converges to an entropy solution of that equation, we are here only interested in the property that the scheme is monotone, therefore TVD and monotonicity preserving, and produces solutions for which the discrete analogue of v_x is uniformly bounded. This makes it possible here to take finite differences of that scheme to generate the u-scheme for the nonlocal equation (1.1) satisfied by $u = v_x$. The convergence of the u-scheme will be analyzed separately.

1.4.2 Uniform estimates on $\{v_j^n\}$ and $\{u_{j+1/2}^n\}$

We establish the compactness and regularity estimates on the discrete solutions $\{v_j^n\}$ and $\{u_{j+1/2}^n\}$ in a series of lemmas. In Lemma 1.4.1 we prove that the v-scheme is monotone, and derive from this that the numerical solution $\{v_j^n\}$ satisfies an L^1 Lipschitz continuity in time property (Lemma 1.4.2). This result, in combination with the unboundedness of $A(u)$ for $u \to \infty$, allows us to prove (in Lemma 1.4.3) a uniform L^∞ bound for $\{u_{j+1/2}^n\}$. Then, in Lemma 1.4.4, we prove that the spatial total variation of $A(U^n)$ is uniformly bounded. With the help of Lemma 1.4.5, which states that the cell that includes v^* can move at most one position to the left or the right in one time step, we are then able to show (Lemma 1.4.6) that the spatial total variation of U^n is bounded uniformly with respect to the discretization parameters; the bound depends, however, on the final time T. Then, in Lemma 1.4.7, we prove that the solution $\{u_{j+1/2}^n\}$ is L^1 Hölder continuous in time. Finally, we establish in Lemmas 1.4.8 and 1.4.9 L^2 inequalities related to spatial and temporal translates of $\{A(u_{j+1/2}^n)\}$. The series of lemmas then permits us to prove the main convergence result, Theorem 1.4.1, which states that the numerical solutions $\{u_{j+1/2}^n\}$ produced by the u-scheme indeed converge to the unique weak solution (under conditions, and in a sense made precise in the theorem).

While Lemmas 1.4.1 to 1.4.7 are based on the original CFL condition (1.44), we need to employ a strengthened condition ((1.58), stated in Lemma 1.4.8) to prove Lemmas 1.4.8 and 1.4.9, and eventually Theorem 1.4.1. Finally, we mention that the proofs of Lemmas 1.4.1 to 1.4.3 follow the treatment in [47].

Lemma 1.4.1 *Under the CFL condition* (1.44)*, the v-scheme defined by* (1.41)–(1.43) *is monotone.*

Proof. We rewrite the scheme (1.41) as

$$v_j^{n+1} = \mathscr{H}\big(v_{j-1}^n, v_j^n, v_{j+1}^n\big) =: \mathscr{H}_j^n, \quad j \in \mathbb{Z}, \quad n = 0, 1, \ldots, N-1.$$

Since $a \geq 0$, we then have

$$\frac{\partial \mathscr{H}_j^n}{\partial v_{j\pm1}^n} = \mp\lambda \min\big\{0, \Phi'\big(v_{j\pm1}^n\big)\big\} + \mu a\big(\Delta_\pm v_j^n/\Delta x\big) \geq 0,$$

while the CFL condition (1.44) implies that

$$\begin{aligned}\frac{\partial \mathscr{H}_j^n}{\partial v_j^n} &= 1 - \lambda\big(\max\big\{0, \Phi'\big(v_j^n\big)\big\} - \min\big\{0, \Phi'\big(v_j^n\big)\big\}\big) - \mu\Delta_+ a\big(\Delta_- v_j^n/\Delta x\big) \\ &= 1 - \lambda\big|\Phi'\big(v_j^n\big)\big| - \mu\Delta_+ a\big(\Delta_- v_j^n/\Delta x\big) \geq 0.\end{aligned}$$

□

As a monotone scheme, the scheme (1.41) is total variation diminishing (TVD) and monotonicity preserving. Since (1.41) represents an explicit three-point scheme, for a fixed discretization $(\Delta x, \Delta t)$ we will always have

$$v_j^n = 0 \quad \text{for } j < -\mathscr{K}, \quad v_j^n = C_0 \quad \text{for } j > \mathscr{K} \tag{1.47}$$

for a sufficiently large constant $\mathscr{K} > 0$. Thus, we can state the following corollary.

Corollary 1.4.1 *If* (1.40) *and the CFL condition* (1.44) *hold, then the numerical solution* $\{v_j^n\}$ *produced by the v-scheme* (1.41)–(1.43) *satisfies*

$$0 \le v_j^n \le C_0, \quad v_j^n \le v_{j+1}^n \quad \text{for all } j \in \mathbb{Z},\, n = 1, \ldots, N. \tag{1.48}$$

As a direct consequence, the numerical solution values $V^n = \{v_j^n\}_{j\in\mathbb{Z}}$ *satisfy the (trivial) uniform total variation bound*

$$\mathrm{TV}(V^n) = \sum_{j\in\mathbb{Z}} \left| v_{j+1}^n - v_j^n \right| = C_0.$$

Lemma 1.4.2 *The numerical solution* $\{v_j^n\}$ *produced by the v-scheme* (1.41)–(1.43) *satisfies the* L^1 *Lipschitz continuity in time property, i.e., there exists a constant* C_1*, which is independent of* $\Delta := (\Delta x, \Delta t)$*, such that*

$$\sum_{j\in\mathbb{Z}} \left| v_j^{n+1} - v_j^n \right| \le C_1 \lambda. \tag{1.49}$$

Proof. For $j \in \mathbb{Z}$, the quantity $w_j^{n+1/2} := v_j^{n+1} - v_j^n$ satisfies

$$\begin{aligned} w_j^{n+3/2} - w_j^{n+1/2} = &-\lambda \Delta_+ \left[h\left(v_{j-1}^{n+1}, v_j^{n+1}\right) - h\left(v_{j-1}^n, v_j^n\right) \right] \\ &+ \lambda \Delta_+ \left[A\left(\Delta_- v_j^{n+1}/\Delta x\right) - A\left(\Delta_- v_j^n/\Delta x\right) \right]. \end{aligned} \tag{1.50}$$

We define

$$\theta(s) := \begin{cases} 1/s & \text{if } s \neq 0, \\ 0 & \text{otherwise,} \end{cases}$$

and the quantities

$$\begin{aligned} B_j^{n+1/2} &:= \left[h\left(v_{j-1}^n, v_j^{n+1}\right) - h\left(v_{j-1}^n, v_j^n\right) \right] \theta\left(v_j^{n+1} - v_j^n\right), \\ C_j^{n+1/2} &:= \left[h\left(v_j^{n+1}, v_{j+1}^{n+1}\right) - h\left(v_j^n, v_{j+1}^{n+1}\right) \right] \theta\left(v_j^{n+1} - v_j^n\right), \\ D_j^{n+1/2} &:= \left[A\left(\Delta_+ v_j^{n+1}/\Delta x\right) - A\left(\Delta_+ v_j^n/\Delta x\right) \right] \theta\left(\Delta_+ v_j^{n+1} - \Delta_+ v_j^n\right). \end{aligned} \tag{1.51}$$

Since h is a monotonically non-decreasing function of its first argument and a monotonically non-increasing function of its second argument, and A is a monotonically non-decreasing function, we have

$$C_j^{n+1/2} \geq 0, \quad D_j^{n+1/2} \geq 0, \quad B_j^{n+1/2} \leq 0. \tag{1.52}$$

After some manipulations and using (1.48) we obtain from (1.50)

$$\begin{aligned} w_j^{n+3/2} = w_j^{n+1/2}\big[1 - \lambda C_j^{n+1/2} + \lambda B_j^{n+1/2} - \lambda\big(D_{j-1}^{n+1/2} + D_j^{n+1/2}\big)\big] \\ + w_{j-1}^{n+1/2}\lambda\big(C_{j-1}^{n+1/2} + D_{j-1}^{n+1/2}\big) + w_{j+1}^{n+1/2}\lambda\big(-B_{j+1}^{n+1/2} + D_j^{n+1/2}\big). \end{aligned}$$

Using the CFL condition we find

$$\begin{aligned} \big|w_j^{n+3/2}\big| \leq \big|w_j^{n+1/2}\big|\big[1 - \lambda\big(C_j^{n+1/2} - B_j^{n+1/2} + D_{j-1}^{n+1/2} + D_j^{n+1/2}\big)\big] \\ + \big|w_{j-1}^{n+1/2}\big|\lambda\big(C_{j-1}^{n+1/2} + D_{j-1}^{n+1/2}\big) + \big|w_{j+1}^{n+1/2}\big|\lambda\big(-B_{j+1}^{n+1/2} + D_j^{n+1/2}\big). \end{aligned}$$

Summing this over $j \in \mathbb{Z}$, using (1.52) and (1.48) we obtain

$$\sum_{j\in\mathbb{Z}} \big|w_j^{n+3/2}\big| \leq \sum_{j\in\mathbb{Z}} \big|w_j^{n+1/2}\big|,$$

which implies that

$$\sum_{j\in\mathbb{Z}} \big|w_j^{n+3/2}\big| \leq \sum_{j\in\mathbb{Z}} \big|w_j^{1/2}\big|.$$

From (1.41) with $n = 0$ we get

$$\sum_{j\in\mathbb{Z}} \big|w_j^{1/2}\big| = \sum_{j\in\mathbb{Z}} \big|v_j^1 - v_j^0\big| = \sum_{j\in\mathbb{Z}} \lambda\big|\Delta_+\big(h\big(v_{j-1}^0, v_j^0\big) - A\big(\Delta_- v_j^0/\Delta x\big)\big)\big|.$$

Using (1.6) we arrive at (1.49). □

Lemma 1.4.3 *The numerical solution $\{v_j^n\}$ produced by the v-scheme* (1.41)–(1.43) *satisfies the inequality $|\Delta_+ v_j^n/\Delta x| \leq C_3$ with a constant C_3, which is independent of Δ. Equivalently, the solution $\{u_{j+1/2}^n\}$ generated by the u-scheme* (1.42), (1.43), (1.45), (1.46) *satisfies the uniform L^∞ bound*

$$\big|u_{j+1/2}^n\big| \leq C_3 \quad \text{for all } j \in \mathbb{Z},\ n = 0, \ldots, N. \tag{1.53}$$

Proof. It is sufficient to show that $A(\Delta_+ v_j^n/\Delta x) \le C_2$ for a constant C_2 that is independent of Δ. Taking into account (1.47) we get

$$\begin{aligned}
&\big|A\big(\Delta_+ v_j^n/\Delta x\big)\big| - \big|h\big(v_j^n, v_{j+1}^n\big)\big| \\
&\le \big|A\big(\Delta_+ v_j^n/\Delta x\big) - h\big(v_j^n, v_{j+1}^n\big)\big| \\
&= \left|\Phi(0) + \sum_{k=-\infty}^{j} \Delta_-\big(A\big(\Delta_+ v_k^n/\Delta x\big) - h\big(v_k^n, v_{k+1}^n\big)\big)\right| \\
&= \left|\sum_{k=-\infty}^{j} \frac{v_k^{n+1} - v_k^n}{\lambda} + \Phi(0)\right| \le \frac{1}{\lambda}\sum_{k\in\mathbb{Z}} \big|v_k^{n+1} - v_k^n\big| + \big|\Phi(0)\big|.
\end{aligned}$$

Due to Lemma 1.4.2, we see that $|A(\Delta_+ v_j^n/\Delta x)| \le C_2$ if we choose $C_2 = C_1 + |\Phi(0)|$. Considering (1.8), it concludes the proof. □

Lemma 1.4.4 *The solution* $\{u_{j+1/2}^n\}$ *generated by the u-scheme* (1.42), (1.43), (1.45), (1.46) *satisfies the following inequality, where the constant* C_4 *is independent of* Δ*:*

$$\mathrm{TV}\big(A(U^n)\big) = \sum_{j\in\mathbb{Z}} \big|\Delta_+ A\big(u_{j-1/2}^n\big)\big| \le C_4.$$

Proof. Using the marching formula (1.41) we can write

$$\begin{aligned}
\big|\Delta_+ A\big(u_{j-1/2}^n\big)\big| &\le \frac{1}{\lambda}\big|v_j^{n+1} - v_j^n\big| + \big|\Delta_+ h\big(v_{j-1}^n, v_j^n\big)\big| \\
&\le \frac{1}{\lambda}\big|v_j^{n+1} - v_j^n\big| + \big|\big[h\big(v_j^n, v_{j+1}^n\big) - h\big(v_j^n, v_j^n\big)\big]\theta\big(v_{j+1}^n - v_j^n\big)\big|\big|\Delta_+ v_j^n\big| \\
&\quad + \big|\big[h\big(v_j^n, v_j^n\big) - h\big(v_{j-1}^n, v_j^n\big)\big]\theta\big(v_j^n - v_{j-1}^n\big)\big|\big|\Delta_- v_j^n\big|.
\end{aligned}$$

Summing over $j \in \mathbb{Z}$ yields

$$\sum_{j\in\mathbb{Z}} \big|\Delta_+ A\big(u_{j-1/2}^n\big)\big| \le \frac{1}{\lambda}\sum_{j\in\mathbb{Z}} \big|v_j^{n+1} - v_j^n\big| + 2\|\Phi'\|_\infty \sum_{j\in\mathbb{Z}} \big|\Delta_+ v_j^n\big|.$$

The right-hand side is uniformly bounded due to Lemma 1.4.2 and Corollary 1.4.1. □

Lemma 1.4.4 does, in general, not permit to establish a uniform bound on the spatial total variation $\mathrm{TV}(U^n)$ of the solution values $\{u_{j+1/2}^n\}$ generated by the u-scheme.

We now prove that $\mathrm{TV}(U^n)$ is nevertheless uniformly bounded, but by a bound that depends on the final time T. Our analysis will appeal to assumption (1.7). From (1.47) and (1.48) we deduce that if $v^* < C_0$, where we recall that C_0 is defined in (1.10), and $\{v_j^n\}$ is the numerical solution produced by the v-scheme (1.41)–(1.43), then at each time level there exists a unique index k such that $v_k^n < v^* \le v_{k+1}^n$. The following lemma informs about the behavior of this index with each time iteration. (In light of the discussion of Section 1.1.3, the case $v^* < C_0$ is the most relevant for the phenomenon of aggregation.)

Lemma 1.4.5 *Assume that $v^* < C_0$, and that the data $\{v_j^n\}_{j\in\mathbb{Z}}$ and $\{v_j^{n+1}\}_{j\in\mathbb{Z}}$ have been produced by the v-scheme* (1.41)–(1.43) *starting from the monotone data $\{v_j^0\}_{j\in\mathbb{Z}}$ under the CFL condition* (1.44)*. Let $k,\bar{k}\in\mathbb{Z}$ be the uniquely defined indices that satisfy $v_k^n < v^* \le v_{k+1}^n$ and $v_{\bar{k}}^{n+1} < v^* \le v_{\bar{k}+1}^{n+1}$, respectively. Then $\bar{k}\in\{k-1,k,k+1\}$.*

Proof. Since $v_k^n < v^* \le v_{k+1}^n$ we analyze two cases: $v_k^n < v^* < v_{k+1}^n$ and $v_k^n < v^* = v_{k+1}^n$. In the first, the monotonicity of the v-scheme and (1.48) imply that

$$v_{k-1}^{n+1} \le v_k^n < v^* < v_{k+1}^n \le v_{k+2}^{n+1},$$

such that either $v_{k-1}^{n+1} < v^* \le v_k^{n+1}$, or $v_k^{n+1} < v^* \le v_{k+1}^{n+1}$, or $v_{k+1}^{n+1} < v^* < v_{k+2}^{n+1}$, which means that $\bar{k} = \{k-1,k,k+1\}$. In the second, we find that

$$v_{k-1}^{n+1} \le v_k^n < v^* = v_{k+1}^n \le v_{k+2}^{n+1},$$

so either $v_{k-1}^{n+1} < v^* = v_k^{n+1}$, or $v_k^{n+1} < v^* \le v_{k+1}^{n+1}$, or $v_{k+1}^{n+1} < v^* \le v_{k+2}^{n+1}$. We conclude the proof by noting that $v_{k+2}^{n+1} < v^*$ is impossible due to the monotonicity of the v-scheme and (1.48). □

The next lemma states the announced bound on $\mathrm{TV}(U^n)$.

Lemma 1.4.6 *Assume that the CFL condition* (1.44) *is satisfied. Then there exist constants C_5 and C_6, which are independent of Δ, such that the solution values $U^n = \{u_{j+1/2}^n\}_{j\in\mathbb{Z}}$ satisfy the uniform total variation bound*

$$\mathrm{TV}(U^n) = \sum_{j\in\mathbb{Z}} \left|u_{j+1/2}^n - u_{j-1/2}^n\right| \le \left(C_5 + \mathrm{TV}(U^0)\right)\exp(C_6 T), \quad n = 1,\dots,N. \tag{1.54}$$

Proof. From (1.45) we obtain

$$\Delta_+ u_{j-1/2}^{n+1} = \Delta_+ u_{j-1/2}^n - \mu\Delta_+\Delta^2 h\left(v_{j-1}^n, v_j^n\right) + \mu\Delta_+\Delta^2 A\left(u_{j-1/2}^n\right).$$

Let us assume that $v^* < C_0$, so that there exists an index k such that $v_k^n < v^* \le v_{k+1}^n$ (cf. Lemma 1.4.5), and let us split $\mathbb{Z}$ into the subsets

$$\begin{aligned} \mathscr{A} &:= \mathscr{A}^n := \{j\in\mathbb{Z}\,|\,j \le k-2\},\\ \mathscr{B} &:= \mathscr{B}^n := \{j\in\mathbb{Z}\,|\,k-2 < j \le k+2\},\\ \mathscr{C} &:= \mathscr{C}^n := \{j\in\mathbb{Z}\,|\,k+2 < j\}. \end{aligned} \tag{1.55}$$

(In case that $v^* \ge C_0$, the following arguments for $j\in\mathscr{A}$ apply to all $j\in\mathbb{Z}$, i.e. we may choose $\mathscr{A} = \mathbb{Z}$, and formally $\mathscr{B} = \mathscr{C} = \varnothing$.)

Let $w_j^n := \Delta_+ u_{j-1/2}^n$ and $a_j^n := \Delta_+ A(u_{j-1/2}^n)\theta(\Delta_+ u_{j-1/2}^n)$. For $j\in\mathscr{A}$, we obtain

$$w_j^{n+1} = w_j^n - \mu\Delta_-\Delta^2\Phi\left(v_j^n\right) + \mu\Delta^2\left(a_j^n w_j^n\right). \tag{1.56}$$

Using a Taylor expansion about v_j^n we find that there exist numbers $\alpha_j^n \in [v_j^n, v_{j+1}^n]$ and $\beta_j^n \in [v_{j-1}^n, v_j^n]$ such that

$$\Delta^2\Phi\big(v_j^n\big) = \Phi'(v_j^n)w_j^n\Delta x + \frac{1}{2}\Phi''\big(\alpha_j^n\big)\big(\Delta_+ v_j^n\big)^2 + \frac{1}{2}\Phi''\big(\beta_j^n\big)\big(\Delta_- v_j^n\big)^2.$$

Substituting this into (1.56) we obtain

$$\begin{aligned}
w_j^{n+1} &= w_j^n - \lambda\Delta_-\big(\Phi'\big(v_j^n\big)w_j^n\big) + \mu\Delta^2\big(a_j^n w_j^n\big) - \frac{\mu}{2}\Delta_-\big(\Phi''\big(\alpha_j^n\big)\big(\Delta_+ v_j^n\big)^2\big) \\
&\quad - \frac{\mu}{2}\Delta_-\big(\Phi''\big(\beta_j^n\big)\big(\Delta_- v_j^n\big)^2\big) \\
&= w_j^n - \lambda\Delta_-\big(\Phi'\big(v_j^n\big)w_j^n\big) + \mu\Delta^2\big(a_j^n w_j^n\big) \\
&\quad - \frac{\mu}{2}\Big(\Delta_-\Phi''\big(\alpha_j^n\big)\big(\Delta_+ v_j^n\big)^2 + \Phi''\big(\alpha_{j-1}^n\big)\big(v_{j+1}^n - v_{j-1}^n\big)w_j^n\Delta x \\
&\qquad + \Delta_-\Phi''\big(\beta_j^n\big)\big(\Delta_- v_j^n\big)^2 + \Phi''\big(\beta_{j-1}^n\big)\big(v_j^n - v_{j-2}^n\big)w_{j-1}^n\Delta x\Big) \\
&= w_j^n\big[1 - \lambda\Phi'\big(v_j^n\big) - 2\mu a_j^n\big] + w_{j-1}^n\big[\mu a_{j-1}^n + \lambda\Phi'\big(v_{j-1}^n\big)\big] + \mu w_{j+1}^n a_{j+1}^n \\
&\quad + \mathscr{O}(\Delta t)\big(w_{j-1}^n + w_j^n + \Delta_+ v_j^n + \Delta_- v_j^n\big).
\end{aligned}$$

In an analogous way, we find for $j \in \mathscr{C}$

$$\begin{aligned}
w_j^{n+1} &= w_j^n\big[1 + \lambda\Phi'\big(v_j^n\big) - 2\mu a_j^n\big] + w_{j+1}^n\big[\mu a_{j+1}^n - \lambda\Phi'\big(v_{j+1}^n\big)\big] + \mu w_{j-1}^n a_{j-1}^n \\
&\quad + \mathscr{O}(\Delta t)\big(w_j^n + w_{j+1}^n + \Delta_+ v_j^n + \Delta_- v_j^n\big).
\end{aligned}$$

Now we deal with $j \in \mathscr{B}$. For $j = k-1$, using that v^* is a maximum of Φ and following analogous steps as before, we get

$$\begin{aligned}
w_{k-1}^{n+1} &= w_{k-1}^n - \mu\big(\Phi(v_{k+1}^n) - \Phi(v^*) + \Delta_-\Delta^2\Phi\big(v_{k-1}^n\big)\big) + \mu\Delta^2\big(a_{k-1}^n w_{k-1}^n\big) \\
&= w_{k-1}^n - \mu\big(\Phi'(\xi)\big(v_{k+1}^n - v^*\big) + \Delta_-\Delta^2\Phi\big(v_{k-1}^n\big)\big) + \mu\Delta^2\big(a_{k-1}^n w_{k-1}^n\big) \\
&= w_{k-1}^n - \mu\big(\big(\Phi'(\xi) - \Phi'(v^*)\big)\big(v_{k+1}^n - v^*\big) + \Delta_-\Delta^2\Phi\big(v_{k-1}^n\big)\big) \\
&\quad + \mu\Delta^2\big(a_{k-1}^n w_{k-1}^n\big) \\
&= w_{k-1}^n\big[1 - \lambda\Phi'\big(v_{k-1}^n\big) - 2\mu a_{k-1}^n\big] + w_{k-2}^n\big[\mu a_{k-2}^n + \lambda\Phi'\big(v_{k-2}^n\big)\big] + \mu w_k^n a_k^n \\
&\quad + \mathscr{O}(\Delta t)\big(1 + w_{k-2}^n + w_{k-1}^n + \Delta_+ v_{k-1}^n + \Delta_- v_{k-1}^n\big).
\end{aligned}$$

For $j=k$, using that $\Phi'(v^*)=0$ we compute

$$\begin{aligned}
w_k^{n+1} &= w_k^n - \mu\big[\Phi\big(v_{k+2}^n\big) - 2\Phi(v_{k+1}^n) + \Phi(v_k^n) - \big\{\Phi(v_k^n) - 2\Phi(v_{k-1}^n) + \Phi\big(v_{k-2}^n\big)\big\}\big] \\
&\quad - \mu\big[\Phi(v_{k-1}^n) - \Phi(v_k^n) + 2\big(\Phi(v^*) - \Phi(v_k^n)\big) + \Phi(v^*) - \Phi(v_{k+1}^n)\big] \\
&\quad + \mu\Delta^2\big(a_k^n w_k^n\big) \\
&= w_k^n - \mu\big[\Delta_+\Delta^2\Phi(v_k^n) + \Delta_-\Delta^2\Phi(v_k^n)\big] + \mu\Delta^2\big(a_k^n w_k^n\big) \\
&\quad - \mu\big[\Phi\big(v_{k-1}^n\big) - \Phi(v_k^n) + 2\big(\Phi(v^*) - \Phi(v_k^n)\big) + \Phi(v^*) - \Phi\big(v_{k+1}^n\big)\big] \\
&= w_k^n\big(1-2\mu a_k^n\big) + w_{k-1}^n\big[\mu a_{k-1}^n + \lambda\Phi'\big(v_{k-1}^n\big)\big] + w_{k+1}^n\big[\mu a_{k+1}^n - \lambda\Phi'\big(v_{k+1}^n\big)\big] \\
&\quad + \mathscr{O}(\Delta t)\big(1 + w_{k-1}^n + w_k^n + w_{k+1}^n + \Delta_+ v_k^n + \Delta_- v_k^n\big).
\end{aligned}$$

For $j=k+1$ and $j=k+2$, the following steps are analogous to the previous cases. Using that $\Phi'(v^*)=0$ we obtain

$$\begin{aligned}
w_{k+1}^{n+1} &= w_{k+1}^n - \mu\big[\Delta_+\Delta^2\Phi(v_{k+1}^n) + 3(\Phi(v_k^n) - \Phi(v^*)) + \Phi(v_k^n) - \Phi(v_{k-1}^n)\big] \\
&\quad + \mu\Delta^2\big(a_{k+1}^n w_{k+1}^n\big) \\
&= w_{k+1}^n\big[1 + \lambda\Phi'\big(v_{k+1}^n\big) - 2\mu a_{k+1}^n\big] + w_k^n\mu a_k^n + w_{k+2}^n\big[\mu a_{k+2}^n - \lambda\Phi'\big(v_{k+2}^n\big)\big] \\
&\quad + \mathscr{O}(\Delta t)\big(1 + w_{k+1}^n + w_{k+2}^n + \Delta_+ v_{k+1}^n + \Delta_- v_{k+1}^n\big), \\
w_{k+2}^{n+1} &= w_{k+2}^n - \mu\big[\Delta_+\Delta^2\Phi(v_{k+2}^n) + \Phi(v_k^n) - \Phi(v^*)\big] + \mu\Delta^2\big(a_{k+2}^n w_{k+2}^n\big) \\
&= w_{k+2}^n\big[1 + \lambda\Phi'\big(v_{k+2}^n\big) - 2\mu a_{k+2}^n\big] + w_{k+3}^n\big[\mu a_{k+3}^n - \lambda\Phi'\big(v_{k+3}^n\big)\big] + \mu w_{k+1}^n a_{k+1}^n \\
&\quad + \mathscr{O}(\Delta t)\big(1 + w_{k+2}^n + w_{k+3}^n + \Delta_+ v_{k+2}^n + \Delta_- v_{k+2}^n\big).
\end{aligned}$$

Finally, summing over j we find that there exist constants C_6 and C_7 such that

$$\sum_{j\in\mathbb{Z}} \big|w_j^{n+1}\big| \le \sum_{j\in\mathbb{Z}} \big|w_j^n\big|\,(1 + C_6\Delta t) + C_7\Delta t,$$

which implies that

$$\sum_{j\in\mathbb{Z}} \Big|w_j^{n+1}\Big| \le \sum_{j\in\mathbb{Z}} \big|w_j^0\big| \exp(C_6 T) + \frac{C_7}{C_6}\exp(C_6 T),$$

which proves (1.54). □

The next lemma states L^1 Hölder continuity with respect to the variable t of the solution generated by (1.45).

Lemma 1.4.7 *The solution $\{u_{j+1/2}^n\}$ generated by the u-scheme* (1.42), (1.43), (1.45), (1.46) *satisfies the following inequality, where the constant C_8 is independent of Δ:*

$$\sum_{j\in\mathbb{Z}} \big|u_{j+1/2}^m - u_{j+1/2}^n\big|\Delta x \le C_8\sqrt{\Delta t(m-n)} \quad \text{for } m>n,\ m,n\in\mathbb{N}_0. \tag{1.57}$$

Proof. We first establish weak Lipschitz continuity in the time variable. To this end, let $\phi(x)$ be a test function and $\phi_j := \phi(j\Delta x)$. Multiplying equation (1.45) by $\phi_j \Delta x$, summing over n and j and applying a summation by parts, we get

$$\left|\Delta x \sum_{j\in\mathbb{Z}} \phi_j\big(u^{n+1}_{j+1/2} - u^n_{j+1/2}\big)\right| \le \left|\Delta t \sum_{j\in\mathbb{Z}} G^n_j\big(\phi_j - \phi_{j-1}\big)\right| + \left|\lambda \sum_{j\in\mathbb{Z}} \big(\phi_j - \phi_{j-1}\big)\big(A\big(u^n_{j+1/2}\big) - A\big(u^n_{j-1/2}\big)\big)\right|.$$

Using Lemma 1.4.4 and the fact that ϕ is smooth we obtain

$$\left|\Delta x \sum_{j\in\mathbb{Z}} \phi_j\big(u^{n+1}_{j+1/2} - u^n_{j+1/2}\big)\right| \le C\|\phi'\|\Delta t,$$

where C is independent of Δ and ϕ. Consequently, for $m > n$ the following weak continuity result holds:

$$\left|\Delta x \sum_{j\in\mathbb{Z}} \phi_j\big(u^m_{j+1/2} - u^n_{j+1/2}\big)\right| \le C\|\phi'\|\Delta t(m-n).$$

Since $E_j := u^m_{j+1/2} - u^n_{j+1/2}$ has bounded variation on $\mathbb{R}$, we arrive at the inequality (1.57) by proceeding as in [46, Lemma 3.6]. □

Now, following the treatment in [61] we prove an L^2 estimate for the discrete version of $A(u)_x$.

Lemma 1.4.8 *Assume that the following strengthened CFL condition is satisfied for a constant $\varepsilon > 0$:*

$$\mathrm{CFL}_\varepsilon := 2\lambda \max_{u\in\mathbb{R}}\big|\Phi'(u)\big| + 4\mu \max_{u\in\mathbb{R}} a(u) \le 1 - \varepsilon. \tag{1.58}$$

Then the solution $\{u^n_{j+1/2}\}$ generated by the u-scheme (1.45), (1.46) *satisfies the following inequality, where the constant C_9 depends on ε, but is independent of Δ:*

$$\sum_{n=1}^{N}\sum_{j\in\mathbb{Z}} \left(\frac{\Delta_- A(u^n_{j+1/2})}{\Delta x}\right)^2 \Delta t\Delta x \le C_9. \tag{1.59}$$

Proof. Multiplying (1.45), by $u^n_{j+1/2}\Delta x$, summing the result over $n = 0,\dots,N-1$ and $j \in$

$\mathbb{Z}$, and using summations by parts we get

$$\begin{aligned}
&\lambda \sum_{n=0}^{N-1} \sum_{j\in\mathbb{Z}} \big(\Delta_- A\big(u^n_{j+1/2}\big)\big)\big(\Delta_- u^n_{j+1/2}\big) \\
&= \Delta t \sum_{n=0}^{N-1} \sum_{j\in\mathbb{Z}} G^n_j\big(\Delta_- u^n_{j+1/2}\big) - \frac{\Delta x}{2} \sum_{n=0}^{N-1} \sum_{j\in\mathbb{Z}} \Big(\big(u^{n+1}_{j+1/2}\big)^2 - \big(u^n_{j+1/2}\big)^2\Big) \\
&\quad + \frac{\Delta x}{2} \sum_{n=0}^{N-1} \sum_{j\in\mathbb{Z}} \big(u^{n+1}_{j+1/2} - u^n_{j+1/2}\big)^2,
\end{aligned}$$

where we used that

$$\big(u^{n+1}_{j+1/2} - u^n_{j+1/2}\big)u^n_{j+1/2} = \frac{1}{2}\Big[\big(u^{n+1}_{j+1/2}\big)^2 - \big(u^n_{j+1/2}\big)^2 - \big(u^{n+1}_{j+1/2} - u^n_{j+1/2}\big)^2\Big].$$

In light of Lemma 1.4.3, we can also write

$$\big(\Delta_- A\big(u^n_{j+1/2}\big)\big)\big(\Delta_- u^n_{j+1/2}\big) \geq \frac{1}{a^*}\big(\Delta_- A\big(u^n_{j+1/2}\big)\big)^2, \quad a^* := \max_u a(u),$$

since $a(u) \geq 0$. Using this observation, we find that

$$\begin{aligned}
\frac{\lambda}{a^*} \sum_{n=0}^{N-1} \sum_{j\in\mathbb{Z}} \big(\Delta_- A\big(u^n_{j+1/2}\big)\big)^2 &\leq \Delta t \sum_{n=0}^{N-1} \sum_{j\in\mathbb{Z}} G^n_j\big(\Delta_- u^n_{j+1/2}\big) + \frac{\Delta x}{2} \sum_{j\in\mathbb{Z}} \big(u^0_{j+1/2}\big)^2 \\
&\quad + \frac{\Delta x}{2} \sum_{n=0}^{N-1} \sum_{j\in\mathbb{Z}} \big(u^{n+1}_{j+1/2} - u^n_{j+1/2}\big)^2.
\end{aligned} \tag{1.60}$$

On the other hand, from (1.45) and the inequality $(a+b)^2 \leq 2a^2 + 2b^2$ we obtain

$$\frac{1}{2}\big(u^{n+1}_{j+1/2} - u^n_{j+1/2}\big)^2 \leq \lambda^2\big(\Delta_+ G^n_j\big)^2 + 2\mu^2\Big(\big(\Delta_+ A\big(u^n_{j+1/2}\big)\big)^2 + \big(\Delta_- A\big(u^n_{j+1/2}\big)\big)^2\Big).$$

Multiplying the last inequality by Δx and summing the result over n and j yields

$$\begin{aligned}
\frac{\Delta x}{2} \sum_{n=0}^{N-1} \sum_{j\in\mathbb{Z}} \big(u^{n+1}_{j+1/2} - u^n_{j+1/2}\big)^2 &\leq \frac{\Delta t^2}{\Delta x} \sum_{n=0}^{N-1} \sum_{j\in\mathbb{Z}} \big(\Delta_+ G^n_j\big)^2 \\
&\quad + 4\mu^2 \Delta x \sum_{n=0}^{N-1} \sum_{j\in\mathbb{Z}} \big(\Delta_- A\big(u^n_{j+1/2}\big)\big)^2.
\end{aligned}$$

The new CFL condition (1.58) now implies that

$$4\mu^2\Delta x = 4\mu\frac{\Delta t}{\Delta x} \leq \frac{\Delta t(1-\varepsilon)}{\Delta x\, a^*},$$

and therefore

$$\begin{aligned}&\frac{\Delta x}{2}\sum_{n=0}^{N-1}\sum_{j\in\mathbb{Z}}\big(u_{j+1/2}^{n+1}-u_{j+1/2}^{n}\big)^2\\&\leq\frac{\Delta t^2}{\Delta x}\sum_{n=0}^{N-1}\sum_{j\in\mathbb{Z}}\big(\Delta_+G_j^n\big)^2+\frac{\Delta t(1-\varepsilon)}{\Delta x\,a^*}\sum_{n=0}^{N-1}\sum_{j\in\mathbb{Z}}\big(\Delta_-A\big(u_{j+1/2}^n\big)\big)^2.\end{aligned}\tag{1.61}$$

Summing (1.60) and (1.61) yields

$$\begin{aligned}&\frac{\varepsilon\lambda}{a^*}\sum_{n=0}^{N-1}\sum_{j\in\mathbb{Z}}\big(\Delta_-A\big(u_{j+1/2}^n\big)\big)^2\\&\leq\Delta t\sum_{n=0}^{N-1}\sum_{j\in\mathbb{Z}}G_j^n\big(\Delta_-u_{j+1/2}^n\big)+\frac{\Delta x}{2}\sum_{j\in\mathbb{Z}}\big(u_{j+1/2}^0\big)^2+\frac{\Delta t^2}{\Delta x}\sum_{n=0}^{N-1}\sum_{j\in\mathbb{Z}}\big(\Delta_-G_{j+1}^n\big)^2\leq C,\end{aligned}$$

where we used Lemma 1.4.6, the bound on G_j^n and the fact that $\Delta t=\mathscr{O}(\Delta x^2)$. □

With the help of Lemma 1.4.8 we can prove

Lemma 1.4.9 *Under the assumptions of Lemma 1.4.8 there exists a constant C_{10} which is independent of Δ such that*

$$\sum_{j\in\mathbb{Z}}\big|A\big(u_{j+1/2}^m\big)-A\big(u_{j+1/2}^n\big)\big|^2\Delta x\leq C_{10}(m-n)\Delta t\quad\text{for } m>n.\tag{1.62}$$

Proof. Using Lemma 1.4.3, the fact that $A'(u)\geq 0$ and (1.45) we get

$$\begin{aligned}&\sum_{j\in\mathbb{Z}}\big(A\big(u_{j+1/2}^m\big)-A\big(u_{j+1/2}^n\big)\big)^2\Delta x\\&\leq a^*\sum_{j\in\mathbb{Z}}\big(A\big(u_{j+1/2}^m\big)-A\big(u_{j+1/2}^n\big)\big)\big(u_{j+1/2}^m-u_{j+1/2}^n\big)\Delta x=:\mathscr{A}+\mathscr{B},\end{aligned}\tag{1.63}$$

where we define

$$\begin{aligned}\mathscr{A}&:=-\Delta t\,a^*\sum_{j\in\mathbb{Z}}\big(A\big(u_{j+1/2}^m\big)-A\big(u_{j+1/2}^n\big)\big)\sum_{l=n}^{m-1}\Delta_+G_j^l,\\\mathscr{B}&:=\lambda\,a^*\sum_{j\in\mathbb{Z}}\big(A\big(u_{j+1/2}^m\big)-A\big(u_{j+1/2}^n\big)\big)\sum_{l=n}^{m-1}\Delta^2A\big(u_{j+1/2}^l\big).\end{aligned}$$

Summing by parts we get

$$\mathscr{A}=\Delta t\,a^*\sum_{j\in\mathbb{Z}}\sum_{l=n}^{m-1}G_j^l\big(\Delta_-A\big(u_{j+1/2}^m\big)-\Delta_-A\big(u_{j+1/2}^n\big)\big).$$

We can write

$$\mathscr{A} = \Delta t \Delta x\, a^* \sum_{j\in\mathbb{Z}} \sum_{l=n}^{m-1} G_j^l \left(\frac{\Delta_- A\big(u_{j+1/2}^m\big)}{\Delta x} - \frac{\Delta_- A\big(u_{j+1/2}^n\big)}{\Delta x} \right).$$

Using that $ab \le a^2 + b^2$, we find

$$\begin{aligned}\mathscr{A} \le \frac{\Delta t}{2} a^* \sum_{j\in\mathbb{Z}} \sum_{l=n}^{m-1} \left|G_j^l\right| \left[\left(\frac{\Delta_- A(u_{j+1/2}^m)}{\Delta x} \right)^2 + \left(\frac{\Delta_- A(u_{j+1/2}^n)}{\Delta x} \right)^2 \right] \Delta x \\ + \Delta t a^* \sum_{j\in\mathbb{Z}} \sum_{l=n}^{m-1} \left|G_j^l\right| \Delta x = \mathscr{O}\big((m-n)\Delta t\big),\end{aligned}$$

where we have used Lemma 1.4.8 and the bound on G_j^n.

Proceeding in the same way for $\mathscr{B}$ yields

$$\begin{aligned}\mathscr{B} &= -\lambda a^* \sum_{j\in\mathbb{Z}} \Bigg\{ \big[A\big(u_{j+1/2}^m\big) - A\big(u_{j+1/2}^n\big) - \big(A\big(u_{j-1/2}^m\big) - A\big(u_{j-1/2}^n\big)\big)\big] \\ &\quad \times \sum_{l=n}^{m-1} \Delta_- A\big(u_{j+1/2}^l\big) \Bigg\} \\ &= -\lambda a^* \sum_{j\in\mathbb{Z}} \Bigg\{ \big(\Delta_- A\big(u_{j+1/2}^m\big) - \Delta_- A\big(u_{j+1/2}^n\big)\big) \sum_{l=n}^{m-1} \Delta_- A\big(u_{j+1/2}^l\big) \Bigg\} \\ &= -\lambda a^* \sum_{j\in\mathbb{Z}} \sum_{l=n}^{m-1} \Big(\Delta_- A\big(u_{j+1/2}^m\big) \cdot \Delta_- A\big(u_{j+1/2}^l\big) - \Delta_- A\big(u_{j+1/2}^n\big) \cdot \Delta_- A\big(u_{j+1/2}^l\big) \Big) \\ &\le 2(m-n)\Delta t\, a^* \sum_{j\in\mathbb{Z}} \left(\frac{\Delta_- A(u_{j+1/2}^n)}{\Delta x} \right)^2 \Delta x = \mathscr{O}\big((m-n)\Delta t\big).\end{aligned}$$

Inserting into (1.63) that $\mathscr{A}, \mathscr{B} = \mathscr{O}((m-n)\Delta t)$ concludes the proof. □

Let us now denote by u^Δ the piecewise constant function

$$u^\Delta(x,t) := \sum_{n=0}^{N-1} \sum_{j\in\mathbb{Z}} \chi_{jn}(x,t) u_{j+1/2}^n,$$

where χ_{jn} denotes the characteristic function of $I_j \times [t_n, t_{n+1})$, and let us denote by v^Δ its primitive. From the L^∞ bound (Lemma 1.4.3), the uniform bound on the total variation in space (Lemma 1.4.6) and the L^1 Hölder continuity in time result (Lemma 1.4.7) we infer that there is a constant C such that

$$\|u^\Delta\|_{L^\infty(\Pi_T)} + \|u^\Delta\|_{L^1(\Pi_T)} \le C; \quad \left|u^\Delta(\cdot,t)\right|_{BV(\mathbb{R})} \le C \quad \text{for all } t \in (0,T] \tag{1.64}$$

uniformly as $\Delta x, \Delta t \downarrow 0$, while Lemmas 1.4.8 and 1.4.9 imply (cf. [48, 55]) that there are constants C_{11} and C_{12} independent of Δ such that

$$\begin{aligned} \Big\| A(u^\Delta(\cdot+y,\cdot)) - A(u^\Delta(\cdot,\cdot)) \Big\|_{L^2(\Pi_T)} &\le C_{11}\sqrt{|y|(|y|+\Delta x)}, \\ \Big\| A(u^\Delta(\cdot,\cdot+\tau)) - A(u^\Delta(\cdot,\cdot)) \Big\|_{L^2(\Pi_{T-\tau})} &\le C_{12}\sqrt{\tau}. \end{aligned} \tag{1.65}$$

1.4.3 Convergence to the weak solution

Theorem 1.4.1 *Assume that Δx and Δt satisfy the CFL$_\varepsilon$ condition* (1.58)*, and that u_0 is compactly supported and satisfies* (1.6)*. Then the piecewise constant solutions u^Δ generated by the u-scheme* (1.42)*,* (1.43)*,* (1.45)*,* (1.46) *converge in the strong topology of $L^1(\Pi_T)$ to the unique weak solution of* (1.1)*,* (1.2) *(in the sense of Definition 1.2.1).*

Proof. Since $u^\Delta \in L^\infty(\Pi_T) \cap L^\infty(0,T;BV(\mathbb{R})) \cap C^{1/2}(0,T;L^1(\mathbb{R}))$, we deduce from (1.64) that there exists a sequence $\{\Delta_i\}_{i\in\mathbb{N}}$ with $\Delta_i \downarrow 0$ for $i \to \infty$ and a function $u \in L^\infty(\Pi_T) \cap L^1(\Pi_T) \cap L^\infty(0,T;BV(\mathbb{R}))$ such that $u^\Delta \to u$ a.e. on Π_T. Moreover, in light of (1.65) we have $A(u^\Delta) \to A(u)$ strongly on $L^2_{\mathrm{loc}}(\Pi_T)$, and we have that $A(u) \in L^2(0,T;H^1(\mathbb{R}))$. Lemma 1.4.7 ensures that u satisfies the initial condition (2.26). It remains to prove that u satisfies the weak formulation (1.2.1). To this end, we apply a standard Lax-Wendroff-type argument. Now, multiplying (1.46) by $\int_{I_j} \varphi(x,t_n)\,\mathrm{d}x$, where $I_j := [x_j, x_{j+1}]$ and φ is a suitable smooth test function, and summing the results over $j \in \mathbb{Z}$, we obtain $W_1 + W_2 + W_3 = 0$, where we define

$$\begin{aligned} W_1 &:= \sum_{n=0}^{N-1} \sum_{j\in\mathbb{Z}} \big(u^{n+1}_{j+1/2} - u^n_{j+1/2}\big) \int_{I_j} \varphi(x,t_n)\,\mathrm{d}x, \\ W_2 &:= \lambda \sum_{n=0}^{N-1} \sum_{j\in\mathbb{Z}} \Delta_+ G^n_j \int_{I_j} \varphi(x,t_n)\,\mathrm{d}x, \\ W_3 &:= -\mu \sum_{n=0}^{N-1} \sum_{j\in\mathbb{Z}} \Delta^2 A(u^n_{j+1/2}) \int_{I_j} \varphi(x,t_n)\,\mathrm{d}x. \end{aligned}$$

By standard summation by parts and using that φ has compact support, we get

$$\begin{aligned} W_1 &= -\Delta t \sum_{n=0}^{N-1} \sum_{j\in\mathbb{Z}} u^{n+1}_{j+1/2} \int_{I_j} \frac{\varphi(x,t_{n+1}) - \varphi(x,t_n)}{\Delta t}\,\mathrm{d}x, \\ W_2 &:= -\Delta t \sum_{n=0}^{N-1} \sum_{j\in\mathbb{Z}} G^n_{j+1} \int_{I_j} \frac{\varphi(x+\Delta x,t_n) - \varphi(x,t_n)}{\Delta x}\,\mathrm{d}x, \\ W_3 &:= -\Delta t \sum_{n=0}^{N-1} \sum_{j\in\mathbb{Z}} A(u^n_{j+1/2}) \int_{I_j} \frac{\varphi(x+\Delta x,t_n) - 2\varphi(x,t_n) + \varphi(x-\Delta x,t_n)}{\Delta x^2}\,\mathrm{d}x. \end{aligned}$$

A direct application of the convergence of u^Δ gives us the desired result for W_1 and W_3. It remains to analyze W_2. Since for each fixed $n \in \{0,\dots,N-1\}$, the data $\{v_j^n\}_{j\in\mathbb{Z}}$ are monotone, there exists an index k such that $v_k^n < v^* \le v_{k+1}^n$. Thus, if $\mathscr{A}$, $\mathscr{B}$ and $\mathscr{C}$ are the sets defined in (1.55), we may write $W_2 = W_{2,\mathscr{A}} + W_{2,\mathscr{B}} + W_{2,\mathscr{C}}$, where the subindex denotes the summation over j from the sets $\mathscr{A}$, $\mathscr{B}$ and $\mathscr{C}$, respectively. For $j \in \mathscr{A}$ we have

$$G_{j+1}^n = \frac{\Delta_+\Phi(v_j^n)}{\Delta x} = \frac{\Delta_+\Phi(v_j^n)}{\Delta_+ v_j^n} u_{j+1/2}^n \quad \text{if } u_{j+1/2}^n \neq 0 \text{ and } G_{j+1}^n = 0 \text{ otherwise,}$$

since $\Phi' > 0$ for the values of v_j^n with $j \in \mathscr{A}$. For $j \in \mathscr{C}$ we have

$$G_{j+1}^n = \frac{\Delta_+\Phi(v_j^n)}{\Delta_+ v_j^n} u_{j+1/2}^n + \frac{\Delta^2\Phi(v_{j+1}^n)}{\Delta x} \quad \text{if } u_{j+1/2}^n \neq 0 \text{ and } G_{j+1}^n = 0 \text{ otherwise.}$$

We can write

$$\begin{aligned}\Delta^2\Phi\big(v_{j+1}^n\big) = \Phi'\big(v_{j+1}^n\big)\Delta_+ u_{j+1/2}^n \Delta x \\ + \frac{1}{2}\left(\Phi''\big(\xi_{j+3/2}^n\big)\big(\Delta_+ v_{j+1}^n\big)^2 + \Phi''\big(\xi_{j+1/2}\big)\big(\Delta_+ v_j^n\big)\right),\end{aligned}$$

where $\xi_{j+1/2}^n \in [v_j^n, v_{j+1}^n]$. Using Lemmas 1.4.3 and 1.4.6 we get

$$W_{2,\mathscr{C}} = -\Delta t \sum_{n=0}^{N-1} \sum_{j\in\mathscr{C}} \frac{\Delta_+\Phi(v_j^n)}{\Delta_+ v_j^n} u_{j+1/2}^n \int_{I_j} \frac{\varphi(x+\Delta x, t_n) - \varphi(x,t_n)}{\Delta x}\,\mathrm{d}x + \mathscr{O}(\Delta x).$$

The case $j \in \mathscr{B}$ can be treated in the same way as $j \in \mathscr{C}$ using that $\Phi'(v^*) = 0$ and that $\mathscr{B}$ is a finite index set. Taking the limit $\Delta \downarrow 0$ we finally get the result. □

1.4.4 Finite Speed of Propagation

In this subsection we prove that the solution u of (1.1)-(1.2) presents finite speed of propagation. To get this result we need the additional assumptions

$$\Phi'' < 0; \quad a(u) = 0 \quad \text{for } u \le u_c,\ u_c \ge 0; \quad a(u) > 0 \quad \text{for } u > u_c. \tag{1.66}$$

From an intuitive point of view, for an initial "mass distribution", the left portion (for which $v < v^*$) of the total "population" (mass) moves to the right and the right portion moves to the left. One then expects that one "herd" of bounded spatial extension forms which moves to the left, right, or is stationary depending on whether the initially left-moving or right-moving individuals outnumber, or are equal to, those moving initially in opposite direction. This is indeed the case, as will be shown in the following characterization of the process of aggregation-dispersion by a travelling wave analysis.

Lemma 1.4.10 *Assume that* (1.66) *holds, and let* $u(x,t)$ *be a solution of* (1.1), (1.2). *Then there exist constants* α *and* β *such that for* $t \le T$, $u(x,t) = 0$ *outside of* $\alpha \le x - st \le \beta$, *where*

$$s = \sigma(0, C_0) := \frac{\Phi(C_0) - \Phi(0)}{C_0}.$$

Proof. We consider the following regularized equation

$$\partial_t v_\varepsilon + \partial_x \Phi(v_\varepsilon) = \partial_x A_\varepsilon(\partial_x v_\varepsilon), \quad x \in \mathbb{R}, \quad t \in (0,T], \tag{1.67}$$

$$v_\varepsilon(-\infty) = 0 \quad \text{and} \quad v_\varepsilon(+\infty) = C_0 \tag{1.68}$$

(see (1.10)), where $A_\varepsilon(u) = A(u) + \varepsilon u$. We seek a travelling wave solution $v_\varepsilon(x,t) = w_\varepsilon(x - st)$ of (1.67), (1.68). Hence, $w_\varepsilon(\xi)$ (with $\xi := x - st$) is a solution of the following problem, where $' = \mathrm{d}/\mathrm{d}\xi$:

$$-s w_\varepsilon' + \Phi(w_\varepsilon)' = \left(A_\varepsilon(w_\varepsilon')\right)', \tag{1.69}$$

$$w_\varepsilon(-\infty) = 0 \quad \text{and} \quad w_\varepsilon(+\infty) = C_0. \tag{1.70}$$

Integrating (1.69) we find

$$-s w_\varepsilon + \Phi(w_\varepsilon) - A_\varepsilon(w_\varepsilon') \equiv \overline{w} (= \text{const.}).$$

Applying (1.70) we get $\overline{w} = \Phi(0)$ and $s = \sigma(0, C_0)$. By using $\Phi'' < 0$ we observe that

$$A_\varepsilon(w_\varepsilon') = \Phi(w_\varepsilon) - [\Phi(0) + \sigma(0, C_0) w_\varepsilon] \ge 0 \text{ for } 0 \le w_\varepsilon \le C_0.$$

Since $A_\varepsilon(\cdot)$ is strictly increasing we can write

$$w_\varepsilon' = A_\varepsilon^{-1}\left(\Phi(w_\varepsilon) - \left[\Phi(0) + \sigma(0, C_0) w_\varepsilon\right]\right) \ge 0.$$

Hence, w_ε is determined by the relation

$$\xi_2 - \xi_1 = \int_{w_\varepsilon(\xi_1)}^{w_\varepsilon(\xi_2)} \frac{\mathrm{d}\tilde{w}_\varepsilon}{A_\varepsilon^{-1}(\gamma(\tilde{w}_\varepsilon))}, \quad \text{for } \xi_1, \xi_2 \in \mathbb{R},$$

where

$$\gamma(\tilde{w}_\varepsilon) := \Phi(\tilde{w}_\varepsilon) - [\Phi(0) + \sigma(0, C_0)\tilde{w}_\varepsilon].$$

Thus we have constructed a solution w_ε for the problem (1.69), (1.70). The above integral gives a well-defined function w_ε since $A_\varepsilon^{-1}(\gamma(w_\varepsilon)) > 0$ for $w_\varepsilon \in (0, C_0)$ and $w_\varepsilon' > 0$. Now, we study the limit $\varepsilon \downarrow 0$. We note that

$$A_\varepsilon : [0, u_c) \to [0, \varepsilon u_c) \Rightarrow A_\varepsilon^{-1} : [0, \varepsilon u_c) \to [0, u_c) \quad \text{for } w_\varepsilon < u_c,$$
$$A_\varepsilon : [u_c, \infty) \to [\varepsilon u_c, \infty) \Rightarrow A_\varepsilon^{-1} : [\varepsilon u_c, \infty) \to [u_c, \infty) \quad \text{for } u_c \le w_\varepsilon.$$

Let ξ_1, ξ_2 be fixed. We may assume that $\varepsilon u_c < w_\varepsilon(\xi_1) < w_\varepsilon(\xi_2)$ (otherwise we could choose a smaller ε using that $w'_\varepsilon > 0$ in $(0, C_0)$). Thus

$$\int_{w_\varepsilon(\xi_1)}^{w_\varepsilon(\xi_2)} \frac{\mathrm{d}\tilde{w}_\varepsilon}{A_\varepsilon^{-1}(\gamma(\tilde{w}_\varepsilon))} \leq \frac{C_0}{u_c}. \tag{1.71}$$

Moreover, $A_\varepsilon^{-1} \xrightarrow{\varepsilon \to 0} A_0^{-1}$ pointwise, where $A_0^{-1} : [0, +\infty) \to [u_c, +\infty)$ is the inverse function of A restricted to $[u_c, +\infty)$. Using the Lebesgue theorem we obtain passing to the limit $\varepsilon \downarrow 0$

$$\xi_2 - \xi_1 = \int_{w(\xi_1)}^{w(\xi_2)} \frac{\mathrm{d}\tilde{w}}{A_0^{-1}(\gamma(\tilde{w}))}, \quad \text{for } \xi_1, \xi_2 \in \mathbb{R}. \tag{1.72}$$

Since the right-hand side of (1.72) is bounded, there exist two constants α y β such that $w(\xi) = 0$ for $\xi \leq \alpha$ and $w(\xi) = C_0$ for $\xi \geq \beta$ and $w'(\xi) > 0$ for $\alpha < \xi < \beta$. Let $W(\xi) := w'(\xi)$, which is the desired function and satisfies $W(\xi) = 0$ outside (α, β) and $W(\xi) > 0$ on (α, β).

Let u be the solution of (1.1), (1.2), and let $v(x,t)$ and $v_0(x)$ be defined by (1.3) and (1.5), respectively. Since u_0 is assumed to have compact support, there exist constants α_1 and β_1 such that $v_0(x) = 0$ for $x \leq \alpha_1$ and $v_0(x) = C_0$ for $x \geq \beta_1$. On the other hand, we have that v is a solution of (1.4), (1.5). From the above discussion we can construct two travelling wave solutions $w_1(x - st)$ and $w_2(x - st)$ of (1.4), (1.5) with the same speed $s = \sigma(0, C_0)$ which satisfy $w_1(x) \leq v_0(x) \leq w_2(x)$, $w_2(x - st) = 0$ for $x - st \leq \alpha$, and $w_1(x - st) = C_0$ for $x - st \geq \beta$, where we used that (1.1) and (1.4) are invariant under translation of x. We also know that the solution v of (1.4), (1.5) is monotone (as a function of x, for each fixed t), and we find

$$w_1(x - st) \leq v(x,t) \leq w_2(x - st) \text{ on } \Pi_T.$$

Therefore $v(x,t) = 0$ if $x - st \leq \alpha$ and $v(x,t) = C_0$ if $x - st \geq \beta$, which implies $u(x,t) = 0$ outside $\alpha \leq x - st \leq \beta$ for $t \leq T$. □

1.5 Numerical examples

The examples presented here illustrate the qualitative behavior of weak solutions of the initial value problem (1.1), (1.2) and the convergence properties of the numerical scheme. For the first purpose, we select a relatively fine discretization and present the corresponding numerical solutions as three-dimensional sequences of profiles at selected times or contour plots that should almost be free of numerical artefacts, while the convergence properties of the scheme are demonstrated by error histories in some examples. For all numerical examples we specify Δx and use $\mu = \Delta t / \Delta x^2 = 0.1$, i.e., $\Delta t = 0.1 \Delta x^2$.

1.5.1 Example 1

In Example 1 we calculate the numerical solution of (1.1), (1.2) for $\Phi(v) = -(1-v)^2$ and the degenerating integrated diffusion coefficient (1.12) with $u_c = 10$ and $a_0 = 0.1$. The initial datum is given by

$$u_0(x) = \begin{cases} 5 & \text{for } 0.1 \le x \le 0.2, \quad 7 \quad \text{for } 0.8 \le x \le 0.9, \\ 8 & \text{for } 0.6 \le x \le 0.7, \quad 0 \quad \text{otherwise.} \end{cases}$$

Note that $C_0 = 2$ in Example 1, where C_0 is defined in (1.10), and $v^* = C_0/2 = 1$, so that the function Φ corresponds to (1.11), where the constant of integration is -1, and that u_0 is chosen such that (1.6) is satisfied. Moreover, in our case $\Phi''(v) = -2 < 0$, and $\Phi(0) = \Phi(C_0) = -1$. Lemma 1.4.10 shows that under these conditions, and for the integrated diffusion coefficient given by (1.19), the solution converges in time to a compactly supported, stationary travelling-wave solution, which represents the aggregated group of individuals and is defined by the time-independent version of (1.1).

In Figure 1.1 we show the numerical approximations for v and u for $0 \le t \le 0.5$ and for $\Delta x = 0.001$. As predicted, for each fixed time the data $\{v_j^n\}$ are monotonically increasing, and the numerical solution for u indeed displays the aggregation phenomenon, and terminates in a stationary profile.

In Table 1.1 we show the error at $t_1 = 0.1$ and $t_2 = 0.25$ in the L^1 norm for u (denoted as $e_u^{t_i}$, $i = 1,2$) and in the L^∞ norm for v (denoted as $e_v^{t_i}$, $i = 1,2$), where we take as a reference the solution calculated with $\Delta x = 0.0002$. We find an experimental rate of convergence in both cases greater than one. For small Δx this behavior is possibly related to the proximity of the reference solution. One should expect a real order of convergence at most one since the v-scheme is monotone. Similar convergence rates have been observed for the other examples.

In Figure 1.2 we compare the numerical approximations for v and a for different mesh sizes at the simulated time $t = 0.25$.

1.5.2 Example 2

This example represents a slight modification of Example 1, namely we choose Φ and A as in Example 1, but we consider a smooth initial datum u_0 given by $u_0(x) = 2 - 2\cos(4\pi x)$ for $x \in [0,1]$ and $u_0(x) = 0$ otherwise. In Figure 1.3 we show the numerical approximation of u for $0 \le t \le 0.5$ and $\Delta x = 0.001$. We observe that strong discontinuities form after finite time from the smooth initial datum. This behavior contrasts with the regularity of other related problems [17, 76], where discontinuities can occur only if they are present in the initial data.

Δx	$e_v^{t_1}$	conv. rate	$e_v^{t_2}$	conv. rate	$e_u^{t_1}$	conv. rate	$e_u^{t_2}$	conv. rate
0.020	0.239	-	0.317	-	0.915	-	0.695	-
0.010	0.133	0.845	0.146	1.122	0.513	0.834	0.442	0.655
0.005	0.061	1.135	0.069	1.070	0.246	1.062	0.200	1.144
0.004	0.048	1.018	0.054	1.090	0.181	1.369	0.164	0.891
0.002	0.021	1.168	0.024	1.161	0.082	1.150	0.073	1.163
0.001	0.008	1.360	0.009	1.399	0.036	1.167	0.032	1.200

Table 1.1: Example 1: Numerical error at $t_1 = 0.1$ and $t_2 = 0.25$.

1.5.3 Example 3

We now choose Φ and u_0 as in Example 1, but define A by

$$A(u) = \begin{cases} 0.05u & \text{for } 0 \le u \le 5, \\ 0.25 & \text{for } 5 < u \le 10, \\ 0.05u - 0.25 & \text{for } u > 10. \end{cases}$$

Figure 1.4 shows the results for $\Delta x = 0.001$ and $t \in [0, 0.5]$. Again, a stationary single-peak solution is forming, including a jump between $u = 5$ and $u = 10$, in agreement with the flatness of $A(u)$ for $u \in [5, 10]$.

1.5.4 Example 4

We now utilize the function $\Phi(v) = -0.5(\cos(v\pi) + 1)$ combined with the degenerating integrated diffusion coefficient (1.12) with $u_c = 10$ and $a_0 = 0.1$ from Examples 1 and 2 and the initial datum

$$u_0(x) = \begin{cases} 10 & \text{for } x \in [0.05, 0.15], \\ 14 & \text{for } x \in [0.3, 0.5], \\ & \end{cases} \quad \begin{cases} 9 & \text{for } x \in [0.6, 0.7], \\ 8 & \text{for } x \in [0.9, 1], \\ 0 & \text{otherwise.} \end{cases}$$

The result is shown in Figure 1.5 for $\Delta x = 0.001$. We observe the formation of three groups, but the third moves to the right "looking for more" mass since it is not a full state, in the sense of Lemma 1.4.10 for the formation of stationary travelling waves. In addition to Figure 1.5 we show in Figure 1.6 a contour plot of the numerical approximation of v for this example. The contour lines of v correspond to trajectories of "individuals". This example has been included to illustrate the solution behaviour when Φ has several extrema and inflection points.

1.5.5 Example 5

Here we calculate the numerical approximation of u for A as in Examples 1, 2 and 4, but with Φ and u_0 given by the respective equations

$$\Phi(v) = \begin{cases} -0.5(\cos(v\pi)+1) & \text{for } 0 \le v \le 2, \\ (v-2)^2 - 1 & \text{for } v > 2, \end{cases}$$

$$u_0(x) = \begin{cases} 14 & \text{for } x \in [0.15, 0.3], \quad 18 \quad \text{for } x \in [0.8, 0.95], \\ 17 & \text{for } x \in [0.6, 0.7], \quad 0 \quad \text{otherwise.} \end{cases}$$

In Figure 1.7 we show the result for $\Delta x = 0.001$. We see that the spare mass (i.e. the mass that can not get in the first group) "dilutes" to the right. This dissipation of the right-moving mass is driven by the choice of Φ, and not by that of A. Clearly, as in the previous example, Φ does not satisfy the assumption stated in (1.7). This example illustrates that solutions of (1.1), (1.2) will not always evolve into a finite number of stationary or moving, aggregated "herds".

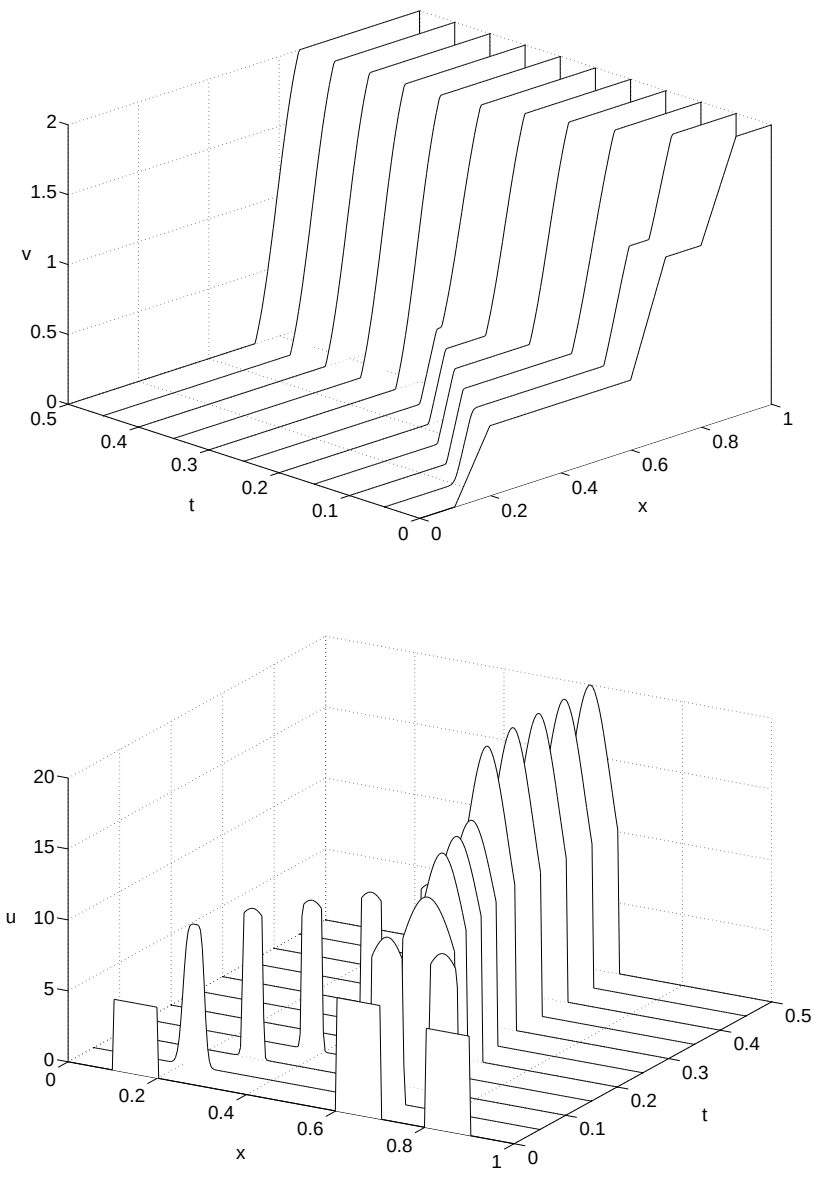

Figure 1.1: Example 1: Numerical approximation of v (top) and corresponding approximation of u (bottom), obtained via (1.38) with $\Delta x = 0.001$.

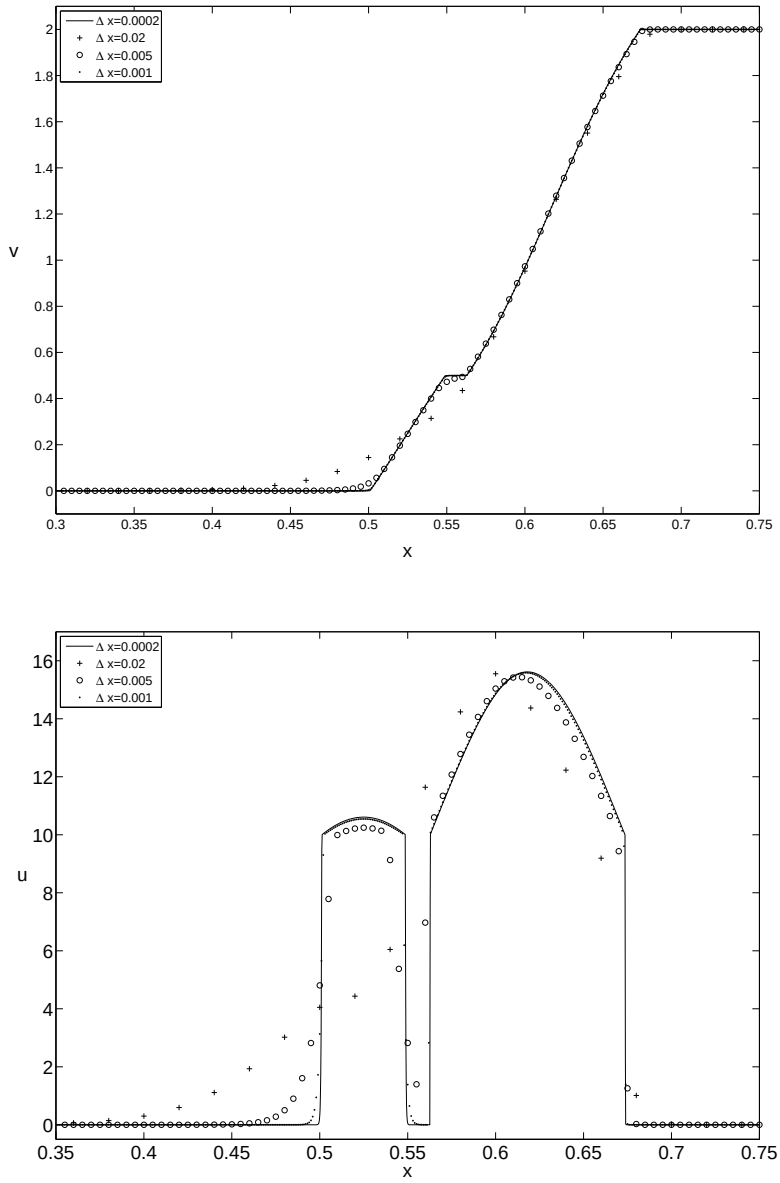

Figure 1.2: Example 1: Numerical approximation of v (top) and u (bottom) for several mesh sizes at $t = 0.25$.

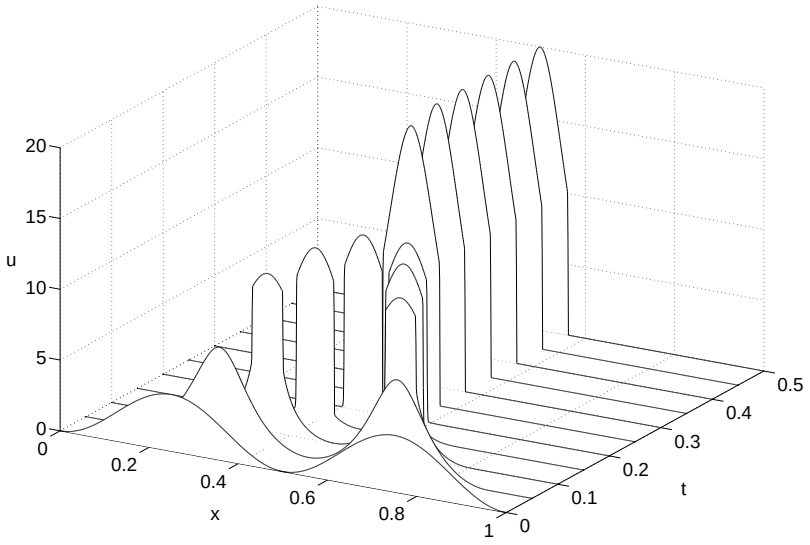

Figure 1.3: Example 2: Numerical approximation of u, obtained via (1.38) with $\Delta x = 0.001$.

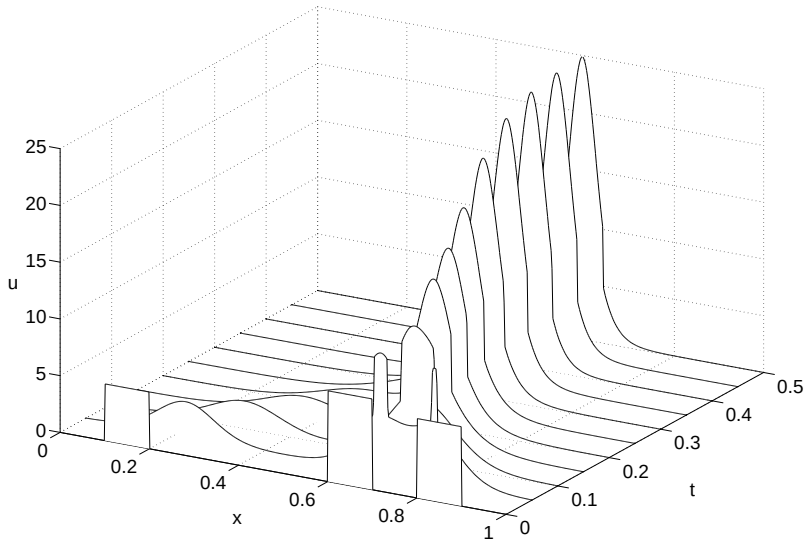

Figure 1.4: Example 3: Numerical approximation of u, obtained via (1.38) for $\Delta x = 0.001$.

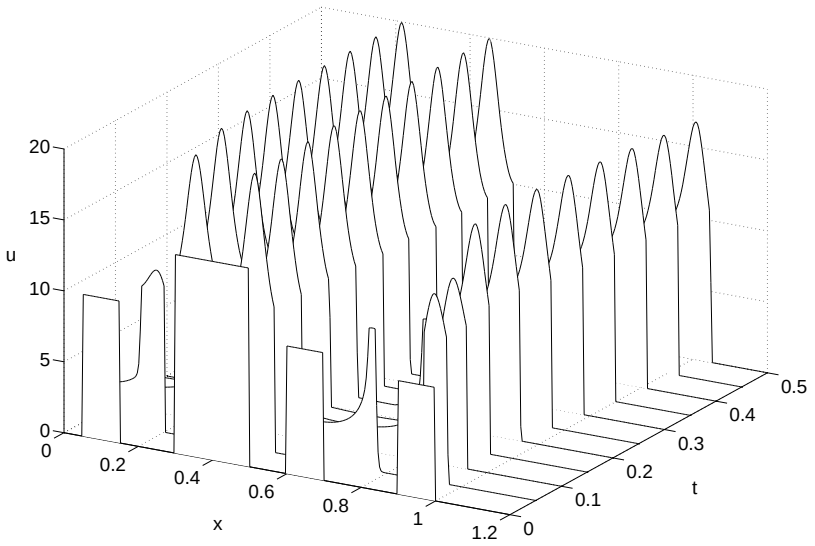

Figure 1.5: Example 4: Numerical approximation of u, obtained via (1.38) for $\Delta x = 0.001$.

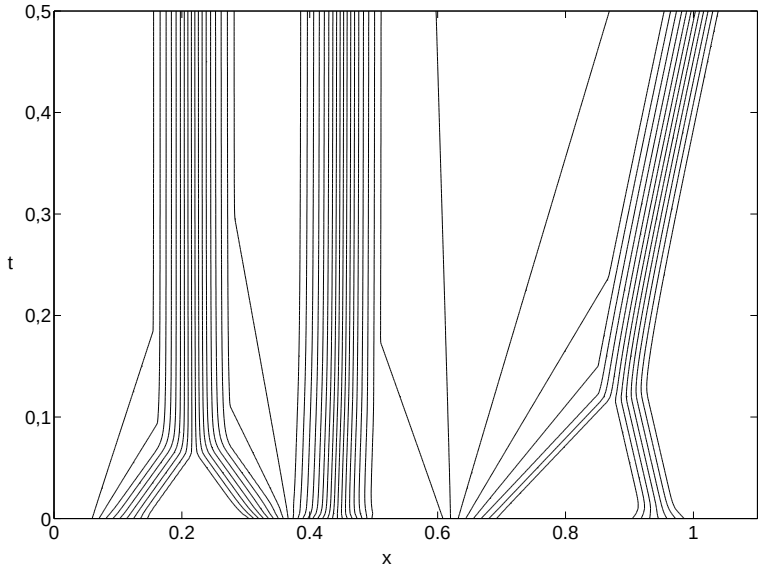

Figure 1.6: Example 4: Contour lines of the numerical approximation of v for $\Delta x = 0.001$.

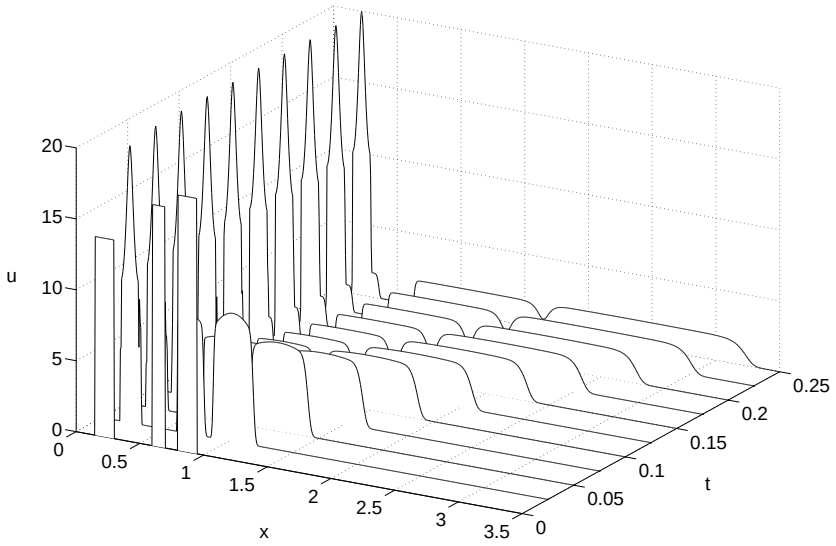

Figure 1.7: Example 5: Numerical approximation of u, obtained via (1.38) for $\Delta x = 0.001$.

Chapter 2

On nonlocal conservation laws modeling sedimentation

The well-known kinematic sedimentation model by Kynch states that the settling velocity of small equal-sized particles in a viscous fluid is a function of the local solids volume fraction. This assumption converts the one-dimensional solids continuity equation into a scalar, nonlinear conservation law with a non-convex and local flux. The present work deals with a modification of this model, and is based on the assumption that either the solids phase velocity or the solid-fluid relative velocity at a given position and time depends on the concentration in a neighborhood via convolution with a symmetric kernel function with finite support. This assumption is justified by theoretical arguments arising from stochastic sedimentation models, and leads to a conservation law with a nonlocal flux. The alternatives of velocities for which the nonlocality assumption can be stated lead to different algebraic expressions for the factor that multiplies the nonlocal flux term. In all cases, solutions are in general discontinuous and need to be defined as entropy solutions. An entropy solution concept is introduced, jump conditions are derived and uniqueness of entropy solutions in shown. Existence of entropy solutions is established by proving convergence of a difference-quadrature scheme. It turns out that only for the assumption of nonlocality for the relative velocity it is ensured that solutions of the nonlocal equation assume physically relevant solution values between zero and one. Numerical examples illustrate the behaviour of entropy solutions of the nonlocal equation.

2.1 Introduction

2.1.1 Scope

We study a family of conservation laws with nonlocal flux defined by

$$u_t + \big(u(1-u)^\alpha V(K_a * u)\big)_x = 0, \quad x \in \mathbb{R}, \quad t \in (0,T], \tag{2.1}$$

together with the initial datum

$$u(0,x) = u_0(x), \quad 0 \le u_0(x) \le 1, \quad x \in \mathbb{R}. \tag{2.2}$$

Under idealizing assumptions, (2.1) represents a one-dimensional model for the sedimentation of small equal-sized spherical solid particles dispersed in a viscous fluid, where the local solids volume fraction $u = u(x,t)$ as a function of depth x and time t is sought. The parameter α satisfies either $\alpha = 0$ or $\alpha \ge 1$; for both choices there is justification from literature, and we study both in parallel. The function V is a hindered settling factor that can be chosen, for example, as

$$V(w) = (1-w)^n, \quad n \ge 1, \tag{2.3}$$

according to Richardson and Zaki [100], and which is herein supposed to depend on

$$(K_a * u)(x,t) = \int_{-2a}^{2a} K_a(y) u(x+y,t)\,\mathrm{d}y,$$

where K_a is a symmetric, non-negative piecewise smooth kernel function with support on $[-2a,2a]$ for a parameter $a > 0$ and $\int_{\mathbb{R}} K_a(x)\,\mathrm{d}x = 1$. Usually, one defines a kernel $K = K(x)$ with support on $[-2,2]$ and sets $K_a(x) := a^{-1}K(a^{-1}x)$. Clearly, (2.1) can be considered as a nonlocal version of the kinematic sedimentation model due to Kynch [70], which gives rise to the local scalar conservation law

$$u_t + \big(uV(u)\big)_x = 0, \quad x \in \mathbb{R}, \quad t \in (0,T]. \tag{2.4}$$

In this chapter we study the well-posedness of (2.1), (2.2). We establish uniqueness of solutions by an entropy solution concept, and existence by proving convergence of a difference-quadrature scheme based on the standard Lax-Friedrichs scheme. It turns out that for $\alpha = 0$, solutions are bounded by a constant that depends on the final time T, and are Lipschitz continuous if u_0 is Lipschitz continuous. In contrast, for $\alpha \ge 1$ solutions are in general discontinuous even if u_0 is smooth, but assume values within the interval $[0,1]$ for all times. Some numerical examples illustrate the solution behaviour, in particular the so-called effect of layering in sedimenting suspensions and the differences between the cases $\alpha = 0$ and $\alpha \ge 1$.

2.1.2 Motivation of the nonlocal flux

Kynch [70] carried out an analysis of sedimentation in which the suspension was approximated by a continuum. When diffusion is negligible, the one-dimensional continuity equation is [29]

$$u_t(x,t) + \big(u(x,t)v_{\rm s}(x,t)\big)_x = 0, \tag{2.5}$$

where $v_{\rm s}(x,t)$ is the solids phase velocity, or settling velocity, at position x at time t, and (2.4) corresponds to the assumption that $v_{\rm s}$ is an explicit function of u, $v_{\rm s} = v_{\rm St}V(u)$, where $v_{\rm St}$ is the Stokes velocity, i.e., the settling velocity of a single sphere in an unbounded fluid. If V is given by (2.3), that is, we employ

$$v_{\rm s}(x_0,t) = v_{\rm St}\big(1-u(x_0,t)\big)^n, \tag{2.6}$$

and assume that V depends on $K_a * u$ instead of u (detailed justification of this assumption will be provided in Section 2.2), then (2.5) takes the form

$$u_t + v_{\rm St}\big(u(1-K_a*u)^n\big)_x = 0. \tag{2.7}$$

A different approach consists in considering the solid and fluid mass conservation equations (2.5) and $-u_t + ((1-u)v_{\rm f})_x = 0$, where $v_{\rm f}$ is the fluid phase velocity. For batch settling we have the relation $v_{\rm s} = (1-u)v_{\rm r}$, where $v_{\rm r} := v_{\rm s} - v_{\rm f}$ is the solid-fluid relative velocity or slip velocity. This leads to the governing equation

$$u_t + \big(u(1-u)v_{\rm r}\big)_x = 0. \tag{2.8}$$

Assuming now that $v_{\rm r}$ (instead of $v_{\rm s}$) has a nonlocal behaviour and requiring that the local versions based on constitutive assumptions for either $v_{\rm s}$ or $v_{\rm r}$ should coincide, we state the constitutive assumption for $v_{\rm r}$ as $v_{\rm r} = V(K_a*u)/(1-u)$. For instance, if we employ (2.3), then the exponent n should be reduced by one, so using the properly adapted Richardson-Zaki equation leads us to

$$v_{\rm s}(x_0,t)/v_{\rm St} = \big(1-u(x_0,t)\big)\big(1-(K_a*u)(x_0,t)\big)^{n-1},$$

from which we obtain the conservation law

$$u_t + v_{\rm St}\big(u(1-u)(1-K*u)^{n-1}\big)_x = 0. \tag{2.9}$$

Equations (2.7) and (2.9) represent the respective cases $\alpha = 0$ and $\alpha = 1$. It is relevant to write the exponent in (2.9) as "$n-1$" only if predictions made by the two versions are to be compared; since n can be chosen arbitrarily, for the mathematical analysis it is sufficient to consider the generic model (2.1)–(2.3). As we prove in this work, the basic

difference in solution behaviour between (2.8) and (2.9) is that solutions of (2.8) may assume values larger than one, while those of (2.9) are strictly limited to $[0,1]$. It seems to us that only (2.9) is suitable for the simulation of the complete sedimentation process from the dilute limit to the densely packed bed. Moreover, formulating a constitutive assumption for v_r rather than for v_s is consistent with one consequence of the principle of material objectivity (see e.g. [81]) stating that a constitutive relation should only be formulated for an objective quantity: not a single velocity (such as v_s), but only the difference between two velocities (such as v_r) is objective. In fact, already Richardson and Zaki [100] recognized that the functional relationship was between v_r and $1-u$ (V_s and ε in their notation). Equation (2.9) is the nonlocal approach that is analogous to theirs.

2.1.3 Approximate dispersive local PDE and invariant region

Insight into qualitative properties of the nonlocal PDE (2.1) can be gained by analyzing an approximate local PDE (the "effective" local PDE [114]) obtained by Taylor expansion of $K_a * u$. In a formal calculation, since K_a is even, we have $K_a * u = u + M_2 a^2 u_{xx} + \mathscr{O}(a^4)$, where $2M_2$ is the second moment of K_a, i.e.

$$2M_2 = \frac{1}{a^2}\int_{-2a}^{2a} K_a(x)x^2\,\mathrm{d}x.$$

Thus, we can write

$$\begin{aligned} V(K_a * u) &= V\big(u + M_2 a^2 u_{xx} + \mathscr{O}(a^4)\big) \approx V(u) + a^2 V'(u)\big(M_2 u_{xx} + \mathscr{O}(a^2)\big) \\ &\approx V(u) + a^2 M_2 V'(u) u_{xx}. \end{aligned}$$

Assuming that the length scale of the solution is much larger than a, we replace $V(K_a * u)$ in (2.1) by $V(u) + a^2 M_2 V'(u) u_{xx}$ and obtain the approximate diffusive-dispersive local PDE

$$u_t + \big(u(1-u)^\alpha V(u)\big)_x = -a^2 M_2 \big(V'(u)u(1-u)^\alpha u_{xx}\big)_x. \tag{2.10}$$

Note that (2.10) depends on the choices of α and V independently; one cannot simply "absorb" $(1-u)^\alpha$ into the choice of V. Thus, for example, we expect qualitatively different solutions in the respective cases $\alpha = 0$ and $\alpha = 1$ with V given by (2.3) with exponents n and $n-1$, although both assumptions lead to the same PDE if V depends locally on u. Specifically, (2.10) reveals why we should expect bounded solutions for $\alpha \geq 1$. In fact, dispersive equations do, in general, not have invariant regions, i.e., one cannot guarantee that the solution takes values in a bounded u-interval for all times. However, for $\alpha \geq 1$ the term sitting inside the derivative on the right-hand side of (2.10) is multiplied by $u(1-u)$, regardless of the algebraic form of V, so for $u = 0$ and $u = 1$ (2.10) degenerates to the

first-order conservation law $u_t + (u(1-u)^\alpha V(u))_x = 0$. The factor $u(1-u)$ has a "saturating" effect; it prevents solution values from leaving the interval $[0,1]$. Thus, we should expect that also the nonlocal PDE (2.1) satisfies an invariant region principle for $\alpha \geq 1$. This is indeed the case, as will be proved in Lemma 2.5.2 of Section 2.5.

2.1.4 Related work

Zumbrun [114] studied an equation equivalent to (2.1) in the case $\alpha = 0$ and $V(w) = v_{\mathrm{St}}(1-\beta w)$. This model for the sedimentation of a dilute particle in a viscous fluid was advanced by Rubinstein [102], and arises as the limiting case for $d \to 0$ from the more general equation

$$u_t + v_{\mathrm{st}}\big(u(1-\beta K_a * u)\big)_x = d u_{xx}, \quad v_{\mathrm{St}}, \beta, d > 0, \tag{2.11}$$

derived from a kinetic theory by Rubinstein and Keller [103, 104] (see also our Section 2.2). For $\beta = 6.55$ [7], this model had been proposed earlier by Caflisch and Papanicolaou [33]. In coordinates $x' = x - v_{\mathrm{St}} t$ and for $\beta = 1$ (equivalent to rescaling u) and $d = 0$, (2.11) reduces to the equation actually studied in [114], namely

$$u_t + \big(u K_a * u\big)_x = 0, \tag{2.12}$$

where $K_a(x) := a^{-1} K(a^{-1}x)$ and K is the truncated parabola given by

$$K(x) = \frac{3}{8}\left(1 - \frac{x^2}{4}\right) \text{ for } |x| < 2; \quad K(x) = 0 \text{ otherwise.} \tag{2.13}$$

Zumbrun [114] showed global existence of weak solutions for the initial value problem (2.2), (2.12) in L^∞ and uniqueness in the class BV. Furthermore, he derived the effective local, dispersive, KdV-like PDE

$$u_t + (u^2)_x = -M_2 a^2 (u u_{xx})_x, \tag{2.14}$$

and showed by analyzing (2.14) that (2.12) supports travelling waves, but not viscous shocks. This result is based on the symmetry of K, which makes (2.12) completely dispersive. Moreover, an L^2 stability argument is invoked to conclude that smooth solutions of the Burgers-like first-order conservation law $u_t + (u^2)_x = 0$ arise from smooth solutions of (2.12) as $a \to 0$. Zumbrun [114] (see also [65]) also studied the effect of artificial diffusion added to (2.12), corresponding to $d > 0$, and showed that for the corresponding effective local PDE, i.e. (2.14) with $d u_{xx}$ added to the right-hand side, solutions of shock initial data converge to a stable, oscillatory travelling wave. He then discussed whether the resulting model is possibly sufficient to explain the phenomenon of layering in sedimentation. Much of his analysis is for a more general, but symmetric kernel K. Whatever

the exact form of $K(x)$, it is clear that the interval over which it applies scales with the sphere radius a. We will compare our findings with those of Zumbrun in Section 2.5.4, see also Section 2.7.

Another spatially one-dimensional, nonlocal sedimentation model was studied by Sjögreen et al. [110]. Starting from a more involved model, they consider a hyperbolic-elliptic model problem given by (2.5) coupled with $-\eta(v_s)_{xx} + v_s = u$, where $\eta > 0$ is a viscosity parameter. Clearly, at any fixed position x_0, $v_s(x_0,t)$ will depend on $u(\cdot,t)$ as a whole; the nonlocal dependence is not limited to a neighborhood, as in [114] and herein. They prove that their model has a smooth solution, and present numerical solutions obtained by a high-order difference scheme.

The (local) kinematic model of sedimentation (2.4) is similar to the well-known Lighthill-Whitham-Richards (LWR) model of vehicular traffic. Sopasakis and Katsoulakis [111] extended the LWR model to a nonlocal version by a "look-ahead" rule, i.e. drivers choose their velocity taking account the density on a stretch of road ahead of them. Kurganov and Polizzi [67] showed that an extension of the well-known Nesshayu-Tadmor (NT) central nonoscillatory scheme [94] is suitable for the nonlocal model of [111], which can be written as (2.1) for $\alpha = 1$ and $V(w) = \exp(-w)$, and if we replace K_a by $K(y) = \mathscr{H}(y)\gamma^{-1}\varphi(y/\gamma)$, where $\mathscr{H}$ is the Heaviside function, $\gamma > 0$ is a constant proportional to the look-ahead distance, and either $\varphi = 1$ (according to [111]) or $\varphi(z) = 2 - 2z$ (as proposed in [67]) for $0 \le z \le 1$, and $\varphi = 0$ elsewhere. As pointed out in [67], the basic methods for conservation laws with local flux that should be adapted for (2.1) are central rather than upwind schemes for conservation laws, since the latter usually involve the (approximate) solution of Riemann problems, and no Riemann solver is available for (2.1). Though the second-order NT scheme produces better resolution, we herein rely on the Lax-Friedrichs scheme to be consistent with the entropy analysis.

Related models with a nonlocal convective flux that have been analyzed within an entropy solution framework include the continuum model for the flow of pedestrians by Hughes [56], which gives rise to a multi-dimensional conservation law with a nonlocal flux; see also [36, 37]. However, in contrast to (2.1) the nonlocality in that model is not introduced by explicit convolution but via the solution of an eikonal equation. An entropy solution framework is employed in [43] to establish well-posedness for a hyperbolic-elliptic approximation of the original model of [56].

Another equation that can formally be expressed in the form (2.1), namely for $\alpha = 0$, $V(w) = w$ and with K_a replaced by the Cauchy kernel so that $K_a * u$ becomes the Hilbert transform Hu, is studied in [34]. This equation arises from several applications, including a one-dimensional model of the two-dimensional vortex sheet problem [5], and is analyzed in [34] with respect to existence of smooth solutions for smooth initial data. Equations that can formally be written as a first-order conservation law with nonlocal flux

also arise from models of opinion formation [2].

2.1.5 Outline of the chapter

The remainder of this chapter is organized as follows. In Section 2.2 we motivate the assumption of nonlocal dependence of settling velocities and argue that it may describe the layering phenomenon in sedimentation. In Section 2.3 we describe the numerical scheme, which involves the approximate computation of $K_a * u$ by a quadrature formula. We state some assumptions on the functions u_0 and V and on the mesh for the numerical scheme, and derive some estimates on differences of the discrete convolution. In Section 2.4 we state the definition of entropy solutions of (2.1), (2.2), the jump conditions, and prove that entropy solutions are L^1 contractive with respect to initial data, and in particular unique. Section 2.5 is devoted to the proof of convergence of approximate solutions generated by the numerical scheme to entropy solutions, which is achieved by standard compactness bounds (Sect. 2.5.1), a cell entropy inequality, and Lax-Wendroff-type arguments (Sect. 2.5.2). In Section 2.5.3 we prove that for the case $\alpha = 0$, solutions are actually Lipschitz continuous provided that u_0 is Lipschitz continuous. In Section 2.6 we present numerical examples, paying particular attention to the layering phenomenon. Conclusions, limitations and possible extensions are addressed in Section 2.7.

2.2 Motivation of the nonlocal sedimentation model

2.2.1 Nonlocal dependence of settling velocities

The solution of the one-dimensional continuity equation (2.5) requires an initial condition, possibly boundary conditions, and an equation relating v_s to $u = u(x,t)$. Theoretical studies of this relationship in dilute, uniformly mixed suspensions of identical spheres are numerous. Those by Kermack et al. [64] and Batchelor [7] are especially notable. When $u(x,t)$ is not constant, the relationship between v_s and u is no longer obvious. The key assumption of Kynch's theory [70] is that v_s is determined by the "local solids concentration", which is the concentration $u(x,t)$ at a specified height and time. This relationship is expressed as $v_s = v_{St}V(u)$. This treatment is analyzed in detail by Bustos et al. [29]. In the three-dimensional reality approximated by the one-dimensional theory, $u(x,t)$ is the solids concentration at a horizontal plane [25, 97]. Though this is an excellent approximation (for the dependence of v_s on u), we may still improve it by a nonlocal dependence, as will be argued in this section.

The locality of the dependence in Kynch's theory contrasts sharply with the theoretical result that the velocity of each particle is determined by the size, position and orientation

of all particles and the nature of the boundaries, if any [52]. (Of course, the orientation is irrelevant for spheres.) As a compromise between this result and the assumption by Kynch, Pickard and Tory [96] postulated that the settling velocity, $v_s(x_0,t)$, of a test particle at x_0 is governed by a parameter

$$c(x_0,t) = \int_I w(x)u(x_0+x,t)\,\mathrm{d}x, \qquad I \subseteq \mathbb{R}, \tag{2.15}$$

that is the convolution of local solids concentration with a weighting function. This parameter was introduced in the context of a stochastic model for which the smoothing effect was important [98] and later generalized to polydisperse suspensions [107]. The function $w(x)$ was specified to be positive in a neighborhood of zero, unimodal, and uniformly bounded with uniformly bounded mean, mode, and variance. When u is constant, we require that $c(x_0,t) = u(x_0,t)$. This implies that $\int_I w(x)\,\mathrm{d}x = 1$, see [53]. This means that the velocity of a sphere at x_0 is governed by the concentration in a contiguous region of finite width. In the limiting case, $w(x)$ is replaced by $\delta(x)$ and the sifting property of the Dirac delta function equates the parametric and local solids concentrations [97].

Beenakker and Mazur [12, 13] calculated the mean velocity of a test sphere in a dilute suspension of identical spheres settling toward an infinite horizontal flat plate. Assuming that all the spheres were placed according to a uniform distribution subject only to the condition that they do not overlap the test sphere or the solid boundary [108], they obtained an explicit expression for the mean velocity of a sphere at a given height. Neglecting terms of $\mathscr{O}(a/h)$, where a is the radius of the test sphere and h is its distance from the boundary, this can be written as [102]

$$v_s(x_0,t) = v_{St}\left(1 + \int_{-2a}^{2a} H_a(x)u(x_0+x,t)\,\mathrm{d}x\right), \tag{2.16}$$

where

$$H_a(x) = \frac{15}{8a}\left(\frac{1}{4}\left(\frac{x}{a}\right)^2 - 1\right). \tag{2.17}$$

Note that only spheres in the interval $[x_0-2a, x_0+2a]$ affect the mean velocity of the test sphere [12, 108]. This results from an exact cancellation, before taking the limit, of large terms in the regions above and below this interval [108]. Taking the limits first yields a divergent sedimentation velocity upon integration [12]. Equation (2.16) does not contradict the result that the velocity of the test sphere is affected by all the spheres in a suspension [11, 52, 83] because the variance of velocity is determined by the positions of all the spheres [108]. Velocity fluctuations in sedimentation, which account for hydrodynamic diffusion, are still being studied intensively [87].

When u is constant, insertion of (2.17) into (2.16) yields

$$v_s(x_0,t) = v_{St}\big(1-5u(x_0,t)\big), \quad \text{i.e., } V(u) = 1-5u. \tag{2.18}$$

This result also holds in a linear concentration gradient because the additional term in the integrand is an odd function (since $H_a(x)$ is even), and the integral is between symmetric limits. Apart from discontinuities, concentration normally varies smoothly over distances much greater than $4a$. Hence, this equation is a good approximation in nonlinear gradients, but only in very dilute suspensions. Higher-order two-sphere interactions can be added [12] to yield Batchelor's result for identical spheres [7], which is

$$v_s(x_0,t) = v_{St}\big(1-6.55u(x_0,t)\big), \quad \text{i.e., } V(u) = 1-6.55u. \tag{2.19}$$

This equation works well for colloidal dispersions in which Brownian motion maintains an essentially uniform distribution of sphere centers. However, experiments with non-Brownian spheres suggest that (2.18) is more accurate. Though the velocity of the spheres relative to the fluid is independent of the shape of the container [13], equation (2.17) applies only to dilute suspensions settling towards an infinite flat plate. Nevertheless, it seems likely, given the success of Kynch's theory, that a generalization of (2.17) should be a reasonable approximation at higher concentrations and for suspensions in finite containers.

Three-sphere and higher interactions are important at higher concentrations [9, 10, 60]. Special treatments involving intensive computation are necessary for concentrated suspensions [54, 71, 72, 73, 74, 95]. At higher concentrations, the dependence of v_s on u is nonlinear. The Richardson-Zaki [100] equation (2.6), corresponding to $V(u)$ given by (2.3), is widely used to predict the position of the interface and the propagation of concentration changes.

In the Pickard-Tory model, the dependence of the settling velocity, $v_s(x_0,t)$, on $c(x_0,t)$ rather than $u(x_0,t)$ is similar to the dependence in [11], but not as specific. If we combine their model with the Richardson-Zaki equation, we obtain

$$\begin{aligned} \frac{v_s(x_0,t)}{v_{St}} &= \big(1-c(x_0,t)\big)^n \\ &= 1 - n\int_I w(x)u(x_0+x,t)\,dx + \frac{n(n-1)}{2}\left[\int_I w(x)u(x_0+x,t)\,dx\right]^2 - \dots. \end{aligned} \tag{2.20}$$

We can choose w to be an even function. Then (2.20) implies that $c(x_0,t) = u(x_0,t)$ in a linear concentration gradient. When u is small and constant or linear, we obtain the approximation $v_s(x_0,t) \approx v_{St}(1-nu(x_0,t))$, which agrees with (2.18) and (2.19).

Again, we choose w to be an even function and require that $\int_I w(x)\,dx = 1$. Then (2.20) can be written as $v_s(x_0,t) = v_{St}(1-K*u)^n$, which yields (2.7), where $K*u$ is the convolution of K with u and $\int_I K(x)\,dx = 1$.

2.2.2 Layered sedimentation in suspensions

Initially homogeneous suspensions of hydrophobic colloidal particles do not always sediment in smooth continuous fashion. Instead, layers of different concentrations are often observed after settling has proceeded for a time [109]. This phenomenon is accentuated when a very dilute suspension has an initial concentration gradient [109]. The upward propagation of a concentration gradient from the bottom of the container will eventually obliterate the layered form if we study this phenomenon in a closed vessel rather than just focussing on the zone slightly below the suspension-supernate interface.

The weighting functions $H_a(x), K(x), w(x)$, and $W(x)$ have an important influence near discontinuities in concentration. When the interval I overlaps the packed bed, these weighting functions introduce a concentration gradient [6]. Where Kynch's theory predicts a jump in u from u_0 to $u_{\max}$, which correponds to a so-called mode of sedimentation MS-1 [26, 29], weighting functions produce the same increase over a finite distance [6]. However, this gradient does not expand.

Our assumption of nonlocal dependence of the settling velocity provides an explanation of the layering phenomenon. In fact, when I overlaps the suspension-supernatant interface, the spheres near that interface settle faster than those below. For example, according to (2.16), a particle at the interface of a uniform suspension has an initial velocity of $v(x_0,t) = v_{\mathrm{St}}(1-2.5u(x_0,t))$ compared to that given by (2.18) for a sphere that is $2a$ or more below the interface. This causes an increase in concentration from u_0 to $u_0+\Delta u$ in a small region just below the interface. However, spheres near the bottom of this region settle faster than those in its middle because I includes a sub-region with $u = u_0$ as well as one with $u = u_0 + \Delta u$. This concentration disturbance should propagate down the settling column. If equation (2.5) applies, the result would seem to be a gradual increase in concentration and perhaps some instability if the concentration near the top remains higher than that near the bottom. Since concentrated suspensions settle much more slowly than dilute ones, it would seem that layering would occur only in very dilute suspensions where slight increases in concentration cause only slight changes in settling velocity.

2.3 Preliminaries

2.3.1 Assumptions and numerical scheme

We discretize (2.1) on a fixed grid given by $x_j = j\Delta x$ for $j \in \mathbb{Z}$ and $t_n = n\Delta t$ for $n \le N := T/\Delta t$, where T is the finite final time. As usual, u_j^n approximates the cell average

$$u_j^n \approx \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} u(y,t_n)\,\mathrm{d}y, \tag{2.21}$$

and we define $U^n := (\dots, u^n_{j-1}, u^n_j, u^n_{j+1}, \dots)^{\mathrm{T}}$. The initial datum u_0 is discretized accordingly. We use the standard spatial difference operators $\Delta_+ u^n_j := u^n_{j+1} - u^n_j$, $\Delta_- u^n_j := u^n_j - u^n_{j-1}$, and $\Delta^2 u^n_j := \Delta_+\Delta_- u^n_j = u^n_{j+1} - 2u^n_j + u^n_{j-1}$. The obvious difficulty in defining a numerical scheme for (2.1) arises from the discretization of the integral. We approximate it by a quadrature formula given by

$$(K_a * u)^n_j \approx \tilde{u}^n_{a,j} := \sum_{i=-l}^{l} \gamma_i u^n_{j-i}, \quad \text{where } \gamma_i = \int_{x_{i-1/2}}^{x_{i+1/2}} K_a(y)\,\mathrm{d}y \text{ and } l = \left\lceil \frac{2a}{\Delta x} \right\rceil + 1,$$

i.e., l is the smallest integer larger or equal to $(2a/\Delta x) + 1$.

Due to the properties of K (Eq. (2.13)), $\gamma_{-l} + \dots + \gamma_l = 1$. The computations and the numerical analysis are based on the Lax-Friedrichs scheme for a standard non-linear scalar conservation law. We summarize all assumptions on the initial datum u_0, the velocity function V and the mesh.

Assumption 2.3.1 *We assume that u_0 has compact support, $u_0(x) \geq 0$ for $x \in \mathbb{R}$ and $u_0 \in BV(\mathbb{R})$. The function $u \mapsto V(u)$ and its derivatives are locally Lipschitz continuous for $u \geq 0$ (which occurs, for example, if $V(\cdot)$ is a polynomial). When we send $\Delta x, \Delta t \downarrow 0$ then it is understood that $\lambda := \Delta t/\Delta x$ is kept constant.*

In addition to Assumption 2.3.1 for the case $\alpha \geq 1$ we suppose the following.

Assumption 2.3.2 *The inital datum satisfies $u_0(x) \leq 1$ for all $x \in \mathbb{R}$.*

Remark 2.3.1 *The same analysis remains valid for any smooth, positive and not necessarily compactly supported kernel with $\|K\|_1 = 1$ and $\|\partial_x K_a\|_1 < \infty$.*

From now on we let the function u^Δ be defined by

$$u^\Delta(x,t) = U^n_j \quad \text{for } (x,t) \in [j\Delta x, (j+1)\Delta x) \times [n\Delta t, (n+1)\Delta t).$$

We now prove two lemmas that will be used for the convergence analysis.

Although K_a (Eq. (2.13)) is just Lipschitz continuous on $\mathbb{R}$, on its support it is a smooth function. Having this in mind we can prove the following lemma.

Lemma 2.3.1 *Suppose that $u^\Delta(\cdot, t_n) \in L^1_{\mathrm{loc}}(\mathbb{R})$. Then*

$$\left|\Delta_+ \tilde{u}^n_{a,j}\right| \leq \|\partial_x K_a\|_\infty \left\|u^\Delta(\cdot, t_n)\right\|_1 \Delta x \quad \text{for } j \in \mathbb{Z}. \tag{2.22}$$

Proof. We compute

$$\Delta_- \tilde{u}^n_{a,j} = \sum_{i=-l}^{l} \gamma_i \Delta_+ u^n_{j-i} = \sum_{i=-l}^{l-1} u^n_{j-i}(\gamma_{i+1} - \gamma_i) + \gamma_l\left(u^n_{j+1+l} - u^n_{j-l}\right). \tag{2.23}$$

Since $K_a(2a) = 0$, we have

$$\gamma_l = \int_{2a-\Delta x/2}^{2a} K_a(x)\,\mathrm{d}x = \int_{2a-\Delta x/2}^{2a} \big|K_a(x) - K_a(2a)\big|\,\mathrm{d}x \le \|\partial_x K_a\|_\infty \frac{\Delta x^2}{4}.$$

For $-l \le i \le l-1$ we find

$$\begin{aligned}\gamma_{i+1} - \gamma_i &= \int_{x_{i+1/2}}^{x_{i+3/2}} \big(K_a(x) - K_a(x-\Delta x)\big)\,\mathrm{d}x = \int_{x_{i+1/2}}^{x_{i+3/2}} \partial_x K_a(\xi_{i+1})\Delta x\,\mathrm{d}x \\ &\le \|\partial_x K_a\|_\infty \Delta x^2,\end{aligned}$$

where $\xi_{i+1} \in [x_{i-1/2}, x_{i+3/2}]$. Applying the last two inequalities to the right-hand side of (2.23) and using that $u^\Delta(\cdot,t_n) \in L^1_{\mathrm{loc}}(\mathbb{R})$, we obtain

$$\big|\Delta_+ \tilde{u}^n_{a,j}\big| \le \|\partial_x K_a\|_\infty \left(\sum_{i=-l}^{l-1} |u^n_{j-i}| + \frac{|u^n_{j+1+l}| + |u^n_{j-l}|}{4} \right) \Delta x^2,$$

which implies (2.22). □

In what follows, $\mathscr{C}_a$ always denotes a constant that is independent of $\Delta := (\Delta x, \Delta t)$, but depends on a, and that may change from one line to the next.

Lemma 2.3.2 *Suppose that $u^\Delta(\cdot,t_n) \in L^1_{\mathrm{loc}}(\mathbb{R}) \cap L^\infty(\mathbb{R})$. Then*

$$\big|\Delta^2 \tilde{u}^n_{a,j}\big| \le \mathscr{C}_a \Delta x^2 \quad \text{for } j \in \mathbb{Z}. \tag{2.24}$$

Proof. We calculate

$$\begin{aligned}\Delta^2 \tilde{u}^n_{a,j} &= \sum_{i=-l}^{l} \big(\gamma_i u^n_{j+1-i} - 2\gamma_i u^n_{j-i} + \gamma_i u^n_{j-1-i}\big) \\ &= \sum_{i=-l+1}^{l-1} u^n_{j-i}\Delta^2\gamma_i + u^n_{j+l}\Delta_+\gamma_{-l} - u^n_{j-l}\Delta_-\gamma_l + \gamma_l\big(\Delta_+ u^n_{j+l} - \Delta_- u^n_{j-l}\big). \end{aligned} \tag{2.25}$$

Lemma 2.3.1 implies that there exists a constant $\mathscr{C}_a$ such that $\gamma_l \le \mathscr{C}_a \Delta x^2$, and therefore $\Delta_+\gamma_{-l} \le \mathscr{C}_a\Delta x^2$ and $\Delta_-\gamma_l \le \mathscr{C}_a\Delta x^2$. Using the Taylor Theorem we get for $i \in \{-l+1, \dots, l-1\}$

$$\begin{aligned}\big|\Delta^2\gamma_i\big| &= \left| \int_{x_{i-1/2}}^{x_{i+1/2}} \big(K_a(x+\Delta x) - 2K_a(x) + K_a(x-\Delta x)\big)\,\mathrm{d}x \right| \\ &= \left| \int_{x_{i-1/2}}^{x_{i+1/2}} \left(\partial_x^2 K_a(\xi_i^+)\frac{\Delta x^2}{2} + \partial_x^2 K_a(\xi_i^-)\frac{\Delta x^2}{2} \right) \mathrm{d}x \right| \le \|\partial_x^2 K_a\|_\infty \Delta x^3,\end{aligned}$$

where $\xi_i^+ \in [x_{i-1/2}, x_{i+3/2}]$ and $\xi_i^- \in [x_{i-3/2}, x_{i+1/2}]$. Consequently, using that $u^\Delta(\cdot,t_n) \in L^1_{\mathrm{loc}}(\mathbb{R}) \cap L^\infty(\mathbb{R})$, we obtain from (2.25) the desired estimate (2.24). □

2.4 Definition and uniquenss of entropy solutions

2.4.1 Definition of an entropy solution and jump conditions

It is well known that solutions to a standard nonlinear conservation law like (2.4) are in general discontinuous even if the initial datum u_0 is smooth. The same will occur with the nonlocal equation (2.1), so we need to define solutions as weak solutions. Since weak solutions of conservation laws are, in general, not unique, a selection criterion must be imposed in order to single out the physically relevant solution. We select the solution through an entropy criterion, and the sought solutions are entropy solutions defined as follows. To facilitate notation we define $f(u) := u(1-u)^\alpha$.

Definition 2.4.1 *A measurable, non-negative function u is an* entropy solution *of the initial value problem* (2.1)*,* (2.2) *if it satisfies the following conditions:*

1. *We have* $u \in L^\infty(\Pi_T) \cap L^1(\Pi_T) \cap BV(\Pi_T)$.
2. *The initial condition* (2.2) *is satisfied in the following sense:*
$$\lim_{t\downarrow 0} \int_{\mathbb{R}} \big|u(x,t) - u_0(x)\big| \,\mathrm{d}x = 0. \tag{2.26}$$
3. *For all non-negative test functions* $\varphi \in C_0^\infty(\Pi_T)$*, the following entropy inequality is satisfied:*
$$\begin{aligned} \forall k \in \mathbb{R}: \quad \iint_{\Pi_T} \Big\{ &|u-k|\varphi_t + \operatorname{sgn}(u-k)\big(f(u)-f(k)\big)V(K_a * u)\varphi_x \\ &- \operatorname{sgn}(u-k) f(k) V'(K_a * u)(\partial_x K_a * u)\varphi \Big\} \,\mathrm{d}x\,\mathrm{d}t \geq 0. \end{aligned} \tag{2.27}$$

The Kružkov-type [69] entropy inequality (2.27) follows from a standard vanishing viscosity argument. It is also standard to deduce that an entropy solution is, in particular, a weak solution of (2.1), (2.2), which is defined by (1) and (2) of Definition 2.4.1, and the following equality, which must hold for all $\varphi \in C_0^\infty(\Pi_T)$:

$$\iint_{\Pi_T} \Big\{ u\,\varphi_t + f(u)V(K_a * u)\varphi_x - f(u)V'(K_a * u)(\partial_x K_a * u)\varphi \Big\} \,\mathrm{d}x\,\mathrm{d}t = 0. \tag{2.28}$$

Assume that u is an entropy solution having a discontinuity at a point $(x_0,t_0) \in \Pi_T$ between the approximate limits u^+ and u^- of u taken with respect to $x > x_0$ and $x < x_0$, respectively. The propagation velocity s of the jump is given by the Rankine-Hugoniot condition, which is derived in a standard way from (2.28):

$$s = \sigma(u^+,u^-)V(K_a * u), \quad \sigma(u,v) := \frac{f(u)-f(v)}{u-v}, \tag{2.29}$$

where we utilize that $(K_a * u)(\cdot,t)$ is a Lipschitz continuous function of x. In addition, a discontinuity between two solution values needs to satisfy the following jump entropy condition, which is a consequence of (2.27):

$$\forall k \in \big(\min\{u^-,u^+\},\max\{u^-,u^+\}\big) : \sigma(u^+,k)V(K_a*u) \le s \le \sigma(u^-,k)V(K_a*u).$$

2.4.2 Uniqueness of entropy solutions

The uniqueness of entropy solutions is a consequence of a result proved in [62] regarding continuous dependence of entropy solutions with respect to the flux function. Precisely, we have the following theorem.

Theorem 2.4.1 *Assume that u and v are entropy solutions of* (2.1), (2.2) *with initial data* u_0 *and* v_0*, respectively. Then there exists a constant* C_1 *such that*

$$\big\|u(\cdot,t)-v(\cdot,t)\big\|_{L^1(\mathbb{R})} \le C_1 \|u_0-v_0\|_{L^1(\mathbb{R})} \quad \forall t \in (0,T].$$

In particular, an entropy solution of (2.1), (2.2) *is unique.*

Proof. Let u and v be entropy solutions of the respective initial value problems

$$\begin{aligned} u_t + \big(V(x,t)f(u)\big)_x = 0, \quad V(x,t) := V\big((K_a*u)(x,t)\big); \quad u(x,0)=u_0(x),\\ v_t + \big(\tilde V(x,t)f(v)\big)_x = 0, \quad \tilde V(x,t) := V\big((K_a*v)(x,t)\big); \quad v(x,0)=v_0(x). \end{aligned}$$

Following the proof of Theorem 1.3 of [62] and keeping in mind that u and v are of bounded variation, we obtain the following inequality, where $J := [0,\|u\|_\infty]$:

$$\begin{aligned} \big\|u(\cdot,t)-v(\cdot,t)\big\|_{L^1(\mathbb{R})} \le{}& \|u_0-v_0\|_{L^1(\mathbb{R})} \\ &+ \|f\|_{L^\infty(J)} \int_0^t \int_{\mathbb{R}} \big|V_x(x,s)-\tilde V_x(x,s)\big|\,\mathrm{d}x\,\mathrm{d}s \\ &+ \|f\|_{\mathrm{Lip}(J)} \int_0^t \int_{\mathbb{R}} \big|V(x,s)-\tilde V(x,s)\big|\big|v_x(x,t)\big|\,\mathrm{d}x\,\mathrm{d}s, \end{aligned} \tag{2.30}$$

where v_x must be understood in the sense of measures. Now we observe that

$$\begin{aligned} \big|V(x,s)-\tilde V(x,s)\big| &= \big|V\big((K_a*u)(x,s)\big)-V\big((K_a*v)(x,s)\big)\big| \\ &\le \big\|V'\big\|_\infty \big|\big(K_a*(u-v)\big)(x,s)\big| \\ &\le \big\|V'\big\|_\infty \|K_a\|_\infty \big\|u(\cdot,s)-v(\cdot,s)\big\|_{L^1(\mathbb{R})}, \\ \big|V_x(x,s)-\tilde V_x(x,s)\big| &= \big|V'\big(K_a*u(x,s)\big)(\partial_x K_a*u)(x,s) \\ &\qquad - V'\big(K_a*v(x,s)\big)(\partial_x K_a*v)(x,s)\big| \\ &\le \|V'\|_\infty \big|\big(\partial_x K_a*(u-v)\big)(x,s)\big| \end{aligned}$$

$$+ \|\partial_x K_a * v\|_\infty \|V''\|_\infty \big| \big(K_a * (u-v)\big)(x,s)\big|.$$

Inserting the last expressions into the integrands in (2.30), using the properties of the kernel K_a and the fact that v has bounded variation we arrive at

$$\big\|u(\cdot,t) - v(\cdot,t)\big\|_{L^1(\mathbb{R})} \le \|u_0 - v_0\|_{L^1(\mathbb{R})} + C_2 \int_0^t \big\|u(\cdot,s) - v(\cdot,s)\big\|_{L^1(\mathbb{R})} \,\mathrm{d}s.$$

Applying the integral form of the Gronwall inequality we finally obtain

$$\big\|u(\cdot,t) - v(\cdot,t)\big\|_{L^1(\mathbb{R})} \le \|u_0 - v_0\|_{L^1(\mathbb{R})} \big(1 + C_2 t \exp(C_2 t)\big).$$

The second statement of the theorem follows by taking $u_0 = v_0$. □

2.5 Convergence analysis and existence of entropy solutions

2.5.1 Compactness estimates

We define $V_j^n := V(\tilde{u}_{a,j}^n)$. Then the marching formula for the approximation of solutions of (2.1), (2.2) reads

$$u_j^{n+1} = \frac{u_{j-1}^n + u_{j+1}^n}{2} - \frac{\lambda}{2} u_{j+1}^n \big(1 - u_{j+1}^n\big)^\alpha V_{j+1}^n + \frac{\lambda}{2} u_{j-1}^n \big(1 - u_{j-1}^n\big)^\alpha V_{j-1}^n. \tag{2.31}$$

We assume that $\lambda = \Delta t/\Delta x$ satisfies the following CFL condition:

$$\lambda \max_{u \le u^*} \big|V(u)\big| < 1 \text{ for } \alpha = 0,\ u^* := \|K_a\|_\infty \|u_0\|_1; \tag{2.32}$$

$$\lambda \max_{0 \le u \le 1} \big|V(u)\big| < 1 \text{ for } \alpha \ge 1. \tag{2.33}$$

By the conservativity of the scheme (2.31) and the CFL condition, we immediately obtain the following lemma.

Lemma 2.5.1 *Under Assumption 2.3.1, the numerical approximation generated by* (2.31) *in the case* $\alpha = 0$ *satisfies*

$$\|U^n\|_1 \le \|U_0\|_1 \quad \text{for } 0 \le n \le N.$$

The next step in the numerical analysis is to prove the L^∞ stability. Appealing to Lemma 2.3.1, we are in a position to prove the following lemma.

Lemma 2.5.2 *The numerical approximation generated by* (2.31) *satisfies*

$$0 \le u_j^n \le \begin{cases} C_3 & \text{if } \alpha = 0, \\ 1 & \text{if } \alpha \ge 1, \end{cases} \quad \text{for } j \in \mathbb{Z} \text{ and } 0 \le n \le N, \tag{2.34}$$

where the constant C_3 is independent of Δ but depends on T.

Proof. We can rewrite (2.31) as

$$u_j^{n+1} = \frac{u_{j-1}^n}{2}\left(1+\lambda\left(1-u_{j-1}^n\right)^{\alpha} V_{j-1}^n\right) + \frac{u_{j+1}^n}{2}\left(1-\lambda\left(1-u_{j+1}^n\right)^{\alpha} V_{j+1}^n\right). \tag{2.35}$$

We consider first the case $\alpha = 0$. Using Assumption 2.3.1 we have

$$\tilde{u}_{a,j}^n = \sum_{i=-l}^{l}\left(\int_{x_{i-1/2}}^{x_{i+1/2}} K_a(y)\,\mathrm{d}y\right) u_{j-i}^n \le \|K_a\|_\infty \sum_{i=-l}^{l} u_{j-i}^n \Delta x \le \|K_a\|_\infty \|u_0\|_1,$$

and thanks to the local Lipschitz continuity of V we can bound $|V(\tilde{u}_{a,j}^n)|$ as a function of $\|K_a\|_\infty$ and $\|u_0\|_1$. Moreover, $|V'(\tilde{u}_{a,j}^n)|$ and $|V''(\tilde{u}_{a,j}^n)|$ can be bounded thanks to the assumptions on V and its derivatives. We can write

$$u_j^{n+1} = u_{j+1}^n\left(\frac{1}{2} - \frac{\lambda}{2}V_{j+1}^n\right) + u_{j-1}^n\left(\frac{1}{2} + \frac{\lambda}{2}V_{j+1}^n\right) - \frac{\lambda}{2}u_{j-1}^n\left(\Delta_+ V_j^n + \Delta_- V_j^n\right).$$

With Lemma 2.3.1 and the CFL condition we get

$$\begin{aligned} \left|u_j^{n+1}\right| &\le \left|u_{j+1}^n\right|\left(\frac{1}{2} - \frac{\lambda}{2}V_{j+1}^n\right) + \left|u_{j-1}^n\right|\left(\frac{1}{2} + \frac{\lambda}{2}V_{j+1}^n\right) \\ &\quad + \lambda\left|u_{j-1}^n\right|\|V'\|_\infty \|\partial_x K_a\|_\infty \|u_0\|_1 \Delta x \\ &\le \|U^n\|_\infty (1 + C_4 \Delta t), \end{aligned}$$

which means that

$$\left|u_j^{n+1}\right| \le \|U^0\|_\infty (1+C_4\Delta t)^n = \|U^0\|_\infty \left(1 + C_4\frac{T}{n}\right)^n \le \|u_0\|_\infty \exp(C_4 T). \tag{2.36}$$

To handle the case $\alpha \ge 1$, we assume that $u_j^n \le 1$ for all $j \in \mathbb{Z}$ (Assumption 2.3.2) and rewrite (2.31) as

$$\begin{aligned} u_j^{n+1} &= \frac{u_{j+1}^n}{2}\left(1+\lambda u_{j+1}^n\left(1-u_{j+1}^n\right)^{\alpha-1} V_{j+1}^n\right) - \frac{\lambda}{2}u_{j+1}^n\left(1-u_{j+1}^n\right)^{\alpha-1} V_{j+1}^n \\ &\quad + \frac{u_{j-1}^n}{2}\left(1-\lambda u_{j-1}^n (1-u_{j-1}^n)^{\alpha-1} V_{j-1}^n\right) + \frac{\lambda}{2}u_{j-1}^n (1-u_{j-1}^n)^{\alpha-1} V_{j-1}^n \\ &\le \frac{u_{j+1}^n}{2}\left(1+\lambda u_{j+1}^n\left(1-u_{j+1}^n\right)^{\alpha-1} V_{j+1}^n\right) - \frac{\lambda}{2}\left(u_{j+1}^n\right)^2\left(1-u_{j+1}^n\right)^{\alpha-1} V_{j+1}^n \end{aligned}$$

$$+\frac{u_{j-1}^n}{2}\left(1-\lambda u_{j-1}^n(1-u_{j-1}^n)^{\alpha-1}V_{j-1}^n\right)+\frac{\lambda}{2}u_{j-1}^n(1-u_{j-1}^n)^{\alpha-1}V_{j-1}^n$$
$$=\frac{u_{j+1}^n}{2}+u_{j-1}^n\left(\frac{1}{2}-\frac{\lambda}{2}u_{j-1}^n(1-u_{j-1}^n)^{\alpha-1}V_{j-1}^n\right)+\frac{\lambda}{2}u_{j-1}^n(1-u_{j-1}^n)^{\alpha-1}V_{j-1}^n.$$

Because of the CFL condition, the last right-hand side is a convex combination of u_{j+1}^n, u_{j-1}^n and one. We therefore conclude that $u_j^{n+1}\leq 1$. The other inequality, $u_j^{n+1}\geq 0$ provided that $u_j^n\geq 0$ for all $j\in\mathbb{Z}$, follows in both cases $\alpha=0$ and $\alpha\geq 1$ from the CFL condition. □

Remark 2.5.1 *Lemma 2.5.2 represents the most important estimate of this work. Based on the discussion of the (local) effective PDE* (2.10) *we argued in Section 2.1.3 that one should expect an "invariant region" principle, namely that solutions assume values in* $[0,1]$*, to hold for* (2.1)*,* (2.2) *with* $\alpha\geq 1$*. The estimate* (2.34) *shows that this property indeed holds. This is an exceptional feature, since an invariant region principle does not hold for dispersive equations in general, and is not valid for* (2.1) *with* $\alpha=0$*. In fact, from* (2.36) *we deduce that for* $\alpha=0$*, one can guarantee that the model* (2.1)*,* (2.2) *produces physically relevant results only if* $\|u_0\|_\infty$ *and the final time* T *are sufficiently small. The requirement of smallness for* $\|u_0\|_\infty$ *is consistent with the observation that the model development in Section 2.2.1 for* $\alpha=0$ *is rigorously valid for dilute suspensions only.*

Since $u_j\geq 0$, we readily obtain the following corollary.

Corollary 2.5.1 *Under Assumption 2.3.1, the numerical solution generated by* (2.31) *in the case* $\alpha\geq 1$ *satisfies*

$$\|U^n\|_1\leq\|U_0\|_1\quad\text{for } 0\leq n\leq N.$$

With the help of Lemma 2.3.2 we may prove the following uniform bound of total variation of the numerical approximation generated by (2.31).

Lemma 2.5.3 *The numerical approximation generated by* (2.31) *satisfies the following total variation bound, where* C_5 *does not depend on* Δ*:*

$$\sum_{j\in\mathbb{Z}}\left|u_j^n-u_{j-1}^n\right|\leq C_5\quad\text{for } 0\leq n\leq N.$$

Proof. Defining $w_{j-1/2}^n:=u_j^n-u_{j-1}^n$ we get from the marching formula (2.31)

$$w_{j-1/2}^{n+1}=w_{j+1/2}^n\left(\frac{1}{2}-\frac{\lambda}{2}f'\left(\xi_{j+1/2}^n\right)V_{j+1}^n\right)+w_{j-3/2}^n\left(\frac{1}{2}+\frac{\lambda}{2}f'\left(\xi_{j-3/2}^n\right)V_{j-1}^n\right)$$

$$
\begin{aligned}
&-\frac{\lambda}{2}\left(\Delta_+ V_j^n\right)\left(f'\left(\xi_{j-1/2}^n\right)w_{j-1/2}^n + f'\left(\xi_{j-3/2}^n\right)w_{j-3/2}^n\right)\\
&+\frac{\lambda}{2}f\left(u_{j-2}^n\right)\left(-\Delta^2 V_j^n - \Delta^2 V_{j-1}^n\right),
\end{aligned}
$$

where $\xi_{j-1/2}^n \in [u_j^n, u_{j-1}^n]$. Using the Taylor theorem we obtain

$$
\Delta^2 V_j^n = V'\left(\tilde{u}_{a,j}^n\right)\Delta^2 \tilde{u}_{a,j}^n + \frac{1}{2}V''\left(\alpha_{j+1/2}^n\right)\left(\Delta_+ \tilde{u}_{a,j}^n\right)^2 + \frac{1}{2}V''\left(\alpha_{j-1/2}^n\right)\left(\Delta_- \tilde{u}_{a,j}^n\right)^2,
$$

where

$$
\alpha_{j+1/2}^n \in [\tilde{u}_{a,j}^n \wedge \tilde{u}_{a,j+1}^n, \tilde{u}_{a,j}^n \vee \tilde{u}_{a,j+1}^n]
$$

(where we define, as usual, $a \wedge b = \min\{a,b\}$ and $a \vee b = \max\{a,b\}$). Thus, Lemmas 2.3.1 and 2.3.2 imply that

$$
\Delta^2 V_j^n = \mathscr{O}(\Delta x^2).
$$

Due to the CFL condition and using that $f(0) = 0$, we obtain that there exists a constant $\mathscr{C}_a$ such that

$$
\begin{aligned}
\left|w_{j-1/2}^{n+1}\right| \leq \left|w_{j+1/2}^n\right|\left(\frac{1}{2} - \frac{\lambda}{2}f'\left(\xi_{j+1/2}^n\right)V_{j+1}^n\right) + \left|w_{j-3/2}^n\right|\left(\frac{1}{2} + \frac{\lambda}{2}f'(\xi_{j-3/2}^n)V_{j-1}^n\right)\\
+ \mathscr{C}_a \Delta t\left(\left|w_{j-1/2}^n\right| + \left|w_{j-3/2}^n\right| + \left|u_{j-2}^n\right|\Delta x\right).
\end{aligned}
$$

Summing over j and using Lemma 2.5.1 we find that there exist constants C_6 and C_7, which depend on a but not on Δ, such that

$$
\mathrm{TV}(U^{n+1}) \leq \mathrm{TV}(U^n)(1 + C_6\Delta t) + C_7\Delta t.
$$

Finally, summing over n we obtain

$$
\begin{aligned}
\mathrm{TV}(U^{n+1}) &\leq \mathrm{TV}(U^0)(1 + C_6\Delta t)^{n+1} + C_7\Delta t\sum_{p=0}^{n}(1 + C_6\Delta t)^p\\
&\leq \mathrm{TV}(u_0)\exp(C_6 T)\left(1 + \frac{C_7}{C_6}\right).
\end{aligned}
$$

□

We also need that u^Δ satisfies the uniform L^1-Lipschitz continuity property with respect to time. This follows directly from the previous results.

Lemma 2.5.4 *The numerical approximation generated by* (2.31) *satisfies the following inequality, where C_8 depends on a, but not on Δ:*

$$
\sum_{j\in\mathbb{Z}}\left|u_j^{n+1} - u_j^n\right| \leq C_8\lambda \quad \text{for } 0 \leq n < N.
$$

Proof. Using the marching formula (2.31) we write

$$\begin{aligned} u_j^{n+1} - u_j^n &= \frac{1}{2}\Delta_+ u_j^n - \frac{1}{2}\Delta_- u_j^n - \frac{\lambda}{2}\left(f\left(u_{j+1}^n\right) - f\left(u_{j-1}^n\right)\right) V_{j+1}^n \\ &\quad - \frac{\lambda}{2} f\left(u_{j-1}^n\right)\left(V_{j+1}^n - V_{j-1}^n\right) \\ &= \frac{1}{2}\Delta_+ u_j^n - \frac{1}{2}\Delta_- u_j^n - \frac{\lambda}{2}\left(f'\left(\xi_{j+1/2}^n\right)\left(u_{j+1}^n - u_j^n\right)\right. \\ &\quad \left. + f'\left(\xi_{j-1/2}^n\right)\left(u_j^n - u_{j-1}^n\right)V_{j+1}^n\right) - f'(\xi) u_{j-1}^n \Delta_+ V_j^n. \end{aligned}$$

In the last expression we used that $f(0) = 0$. We conclude the proof by appealing to Lemma 2.5.3 and the fact that $\Delta t = \mathscr{O}(\Delta x)$. □

2.5.2 Satisfaction of the entropy condition and existence result

From Helly's theorem we have that u^Δ converges to a function $u \in L^\infty(\Pi_T) \cap L^1(\Pi_T) \cap BV(\Pi_T)$ as $\Delta \to 0$. It remains to prove that u satisfies the entropy inequality (2.27).

Theorem 2.5.1 *Assume that Assumptions 2.3.1 and 2.3.2 hold. Then the numerical solution generated by* (2.31) *converges to the unique entropy solution of* (2.1)*,* (2.2)*.*

Proof. We define the function

$$G_j^n(u, v, U^n) := \frac{1}{2}\left(u - \lambda f(u) V_{j+1}^n + v + \lambda f(v) V_{j-1}^n\right).$$

We can rewrite the scheme (2.31) as $u_j^{n+1} = G_j^n(u_{j+1}^n, u_{j-1}^n, U^n)$. Under the CFL condition, G_j^n is monotone its first two arguments for all $j \in \mathbb{Z}$, $0 \le n < N$. Using this property and omitting the third argument, which is always U^n, we obtain

$$\begin{aligned} \left|u_j^{n+1} - G_j^n(k,k)\right| &= \left|G_j^n\left(u_{j+1}^n, u_{j-1}^n\right) - G_j^n(k,k)\right| \\ &\le \left|G_j^n\left(u_{j+1}^n \vee k, u_{j-1}^n \vee k\right) - G_j^n\left(u_{j+1}^n \wedge k, u_{j-1}^n \wedge k\right)\right| \\ &= \left|u_j^n - k\right| - \left(\mathscr{G}_{j+}^n - \mathscr{G}_{j-}^n\right), \end{aligned} \tag{2.37}$$

where we define

$$\mathscr{G}_{j\pm}^n := \frac{\lambda}{2}\left[\left(f\left(u_{j\pm1}^n \vee k\right) - f\left(u_{j\pm1}^n \wedge k\right)\right) V_{j\pm1}^n - \frac{1}{\lambda}\Delta_\pm\left(\left|u_j^n - k\right|\right)\right].$$

On the other hand,

$$\begin{aligned} &\left|u_j^{n+1} - k + \frac{\lambda}{2} f(k)\left(V_{j+1}^n - V_{j-1}^n\right)\right| \\ &\ge \left|u_j^{n+1} - k\right| + \operatorname{sgn}\left(u_j^{n+1} - k\right)\frac{\lambda}{2} f(k)\left(V_{j+1}^n - V_{j-1}^n\right). \end{aligned} \tag{2.38}$$

Combining (2.37) and (2.38) we arrive at the "cell entropy inequality"

$$\left|u_j^{n+1}-k\right|-\left|u_j^n-k\right|+\mathscr{G}_{j+}^n-\mathscr{G}_{j-}^n+\operatorname{sgn}\left(u_j^{n+1}-k\right)\frac{\lambda}{2}f(k)\left(V_{j+1}^n-V_{j-1}^n\right)\leq 0. \tag{2.39}$$

We now establish convergence to a solution that satisfies (2.27) by a Lax-Wendroff-type argument. Multiplying the j-th inequality in (2.39) by $\int_{I_j}\varphi(x,t_n)\,\mathrm{d}x$, where φ is a non-negative test function, and summing the results over $j\in\mathbb{Z}$ and $0\leq n\leq N-1$ we obtain the inequality $E_1+E_2+E_3\leq 0$, where we define

$$\begin{aligned}
E_1 &:= \sum_{n=0}^{N-1}\sum_{j\in\mathbb{Z}}\left(\left|u_j^{n+1}-k\right|-\left|u_j^n-k\right|\right)\int_{I_j}\varphi(x,t_n)\,\mathrm{d}x,\\
E_2 &:= \frac{\lambda}{2}f(k)\sum_{n=0}^{N-1}\sum_{j\in\mathbb{Z}}\operatorname{sgn}(u_j^{n+1}-k)\left(V_{j+1}^n-V_{j-1}^n\right)\int_{I_j}\varphi(x,t_n)\,\mathrm{d}x,\\
E_3 &:= \sum_{n=0}^{N-1}\sum_{j\in\mathbb{Z}}\left(\mathscr{G}_{j+}^n-\mathscr{G}_{j-}^n\right)\int_{I_j}\varphi(x,t_n)\,\mathrm{d}x.
\end{aligned}$$

By a standard summation by parts and using that φ has compact support, we get

$$E_1=-\Delta t\sum_{n=0}^{N-1}\sum_{j\in\mathbb{Z}}\left|u_j^{n+1}-k\right|\int_{I_j}\frac{\varphi(x,t_{n+1})-\varphi(x,t_n)}{\Delta t}\,\mathrm{d}x.$$

For E_2, we write $E_2=E_2^a+E_2^b$ where

$$\begin{aligned}
E_2^a &:= \frac{\lambda}{2}f(k)\sum_{n=0}^{N-1}\sum_{j\in\mathbb{Z}}\left(\operatorname{sgn}\left(u_j^{n+1}-k\right)-\operatorname{sgn}\left(u_j^n-k\right)\right)\left(V_{j+1}^n-V_{j-1}^n\right)\int_{I_j}\varphi(x,t_n)\,\mathrm{d}x,\\
E_2^b &:= \frac{\lambda}{2}f(k)\sum_{n=0}^{N-1}\sum_{j\in\mathbb{Z}}\operatorname{sgn}\left(u_j^n-k\right)\left(V_{j+1}^n-V_{j-1}^n\right)\int_{I_j}\varphi(x,t_n)\,\mathrm{d}x.
\end{aligned}$$

Again summing by parts yields

$$\begin{aligned}
E_2^a = &-\frac{\lambda}{2}f(k)\sum_{n=0}^{N-1}\sum_{j\in\mathbb{Z}}\operatorname{sgn}\left(u_j^{n+1}-k\right)\int_{I_j}\varphi(x,t_n)\,\mathrm{d}x\\
&\qquad\times\left[V_{j+1}^{n+1}-V_{j-1}^{n+1}-\left\{V_{j+1}^n-V_{j-1}^n\right\}\right]\\
&-\frac{\lambda}{2}f(k)\sum_{n=0}^{N-1}\sum_{j\in\mathbb{Z}}\operatorname{sgn}\left(u_j^{n+1}-k\right)\left(V_{j+1}^n-V_{j-1}^n\right)\\
&\qquad\times\int_{I_j}\left(\varphi(x,t_{n+1})-\varphi(x,t_n)\right)\mathrm{d}x.
\end{aligned}$$

Lemmas 2.5.4 and 2.3.1 and the fact that $\gamma_{i+1}-\gamma_i=\mathscr{O}(\Delta x^2)$, $\gamma_l=\mathscr{O}(\Delta x^2)$ yield

$$V^n_{j+1}-V^n_{j-1}=V'\big(\tilde{U}^n_{a,j}\big)\big(\tilde{u}^n_{a,j+1}-\tilde{u}^n_{a,j-1}\big)+\mathscr{O}(\Delta x^2)=\mathscr{O}(\Delta x),$$

and

$$\begin{aligned}
&\tilde{u}^{n+1}_{a,j+1}-\tilde{u}^{n+1}_{a,j-1}-\big\{\tilde{u}^n_{a,j+1}-\tilde{u}^n_{a,j-1}\big\}\\
&=\sum_{i=-l}^{l-1}\big(u^{n+1}_{j-i}+u^{n+1}_{j-i-1}\big)(\gamma_{i+1}-\gamma_i)+\gamma_l\big(u^{n+1}_{j+1+l}+u^{n+1}_{j+l}-u^{n+1}_{j-l}-u^{n+1}_{j-l-1}\big)\\
&\quad-\left\{\sum_{i=-l}^{l-1}\big(u^{n}_{j-i}+u^{n}_{j-i-1}\big)(\gamma_{i+1}-\gamma_i)+\gamma_l\big(u^{n}_{j+1+l}+u^{n}_{j+l}-u^{n}_{j-l}-u^{n}_{j-l-1}\big)\right\}\\
&=\sum_{i=-l}^{l-1}\big(u^{n+1}_{j-i}-u^{n}_{j-i}+u^{n+1}_{j-i-1}-u^{n}_{j-i-1}\big)(\gamma_{i+1}-\gamma_i)\\
&\quad+\gamma_l\big(u^{n+1}_{j+1+l}+u^{n+1}_{j+l}-u^{n+1}_{j-l}-u^{n+1}_{j-l-1}-\Big[u^{n}_{j+1+l}+u^{n}_{j+l}-u^{n}_{j-l}-u^{n}_{j-l-1}\Big]\big)\\
&=\mathscr{O}(\Delta x^2).
\end{aligned}$$

Then, we can write

$$\begin{aligned}
E^a_2=&-\frac{\lambda}{2}f(k)\sum_{n=0}^{N-1}\sum_{j\in\mathbb{Z}}\operatorname{sgn}(u^{n+1}_j-k)\int_{I_j}\varphi(x,t_n)\,\mathrm{d}x\times\\
&\times\Big[\big(V'\big(\tilde{u}^{n+1}_{a,j}\big)-V'\big(\tilde{u}^n_{a,j}\big)\big)\big(\tilde{u}^{n+1}_{a,j+1}-\tilde{u}^{n+1}_{a,j-1}\big)\\
&\quad+V'\big(\tilde{u}^n_{a,j}\big)\big(\tilde{u}^{n+1}_{a,j+1}-\tilde{u}^{n+1}_{a,j-1}-\tilde{u}^n_{a,j+1}+\tilde{u}^n_{a,j-1}\big)\Big]+\mathscr{O}(\Delta x).
\end{aligned}$$

Noting that $\gamma_i=\mathscr{O}(\Delta x)$ we have

$$\tilde{u}^{n+1}_{a,j}-\tilde{u}^n_{a,j}=\sum_{i=-l}^{l}\gamma_i\big(u^{n+1}_{j-i}-u^n_{j-i}\big)=\mathscr{O}(\Delta x),$$

and we conclude that $E^a_2=\mathscr{O}(\Delta x)$. Analogously, we obtain

$$E^b_2=\Delta t f(k)\sum_{n=0}^{N-1}\sum_{j\in\mathbb{Z}}\operatorname{sgn}\big(u^n_j-k\big)V'\big(\tilde{u}^n_{a,j}\big)\frac{\tilde{u}^n_{a,j+1}-\tilde{u}^n_{a,j-1}}{2\Delta x}\int_{I_j}\varphi(x,t_n)\,\mathrm{d}x+\mathscr{O}(\Delta x).$$

It remains to analyze E_3. Another summation by parts gives us

$$\begin{aligned}
E_3=-\frac{\lambda}{2}\sum_{n=0}^{N-1}\sum_{j\in\mathbb{Z}}\Big\{&\big(f\big(u^n_j\vee k\big)-f\big(u^n_j\wedge k\big)\big)V^n_j\\
&\times\int_{I_j}\big(\varphi(x+\Delta x,t_n)-\varphi(x-\Delta x,t_n)\big)\,\mathrm{d}x\Big\}
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{2}\sum_{n=0}^{N-1}\sum_{j\in\mathbb{Z}}\left|u_j^n-k\right|\int_{I_j}\left(\varphi(x+\Delta x,t_n)-2\varphi(x,t_n)+\varphi(x-\Delta x,t_n)\right)\mathrm{d}x\\
&=-\Delta t\sum_{n=0}^{N-1}\sum_{j\in\mathbb{Z}}\Bigg\{\operatorname{sgn}\left(u_j^n-k\right)\left(f\left(u_j^n\right)-f(k)\right)V_j^n\\
&\qquad\times\int_{I_j}\frac{\varphi(x+\Delta x,t_n)-\varphi(x-\Delta x,t_n)}{2\Delta x}\,\mathrm{d}x\Bigg\}+\mathscr{O}(\Delta x).
\end{aligned}
$$

To conclude we must show that

$$
\begin{aligned}
\mathscr{A}&:=\Delta t\sum_{n=1}^{N-1}\sum_{j\in\mathbb{Z}}\left|\tilde{u}_{a,j}^n\Delta x-\int_{I_j}K_a*u(x,t_n)\,\mathrm{d}x\right|\to 0\quad\text{as }\Delta\to 0,\\
\mathscr{B}&:=\Delta t\sum_{n=1}^{N-1}\sum_{j\in\mathbb{Z}}\left|\tilde{u}_{a,j+1}^n-\tilde{u}_{a,j}^n-\int_{I_j}(\partial_x K_a*u)(x,t_n)\,\mathrm{d}x\right|\to 0\quad\text{as }\Delta\to 0.
\end{aligned}
$$

First, we proceed for $\mathscr{A}$. Using the definitions of $\tilde{u}_{a,j}^n$ and γ, we find that

$$
\begin{aligned}
\mathscr{A}&=\Delta t\sum_{n=1}^{N-1}\sum_{j\in\mathbb{Z}}\left|\sum_{i=-l}^{l}\int_{I_i}K_a(y)u_{j-i}^n\Delta x\,\mathrm{d}y-\int_{I_j}\sum_{i=-l}^{l}\int_{I_i}K_a(y)u(x-y,t_n)\,\mathrm{d}y\,\mathrm{d}x\right|\\
&=\Delta t\sum_{n=1}^{N-1}\sum_{j\in\mathbb{Z}}\left|\sum_{i=-l}^{l}\int_{I_i}\int_{I_j}K_a(y)u_{j-i}^n\,\mathrm{d}x\,\mathrm{d}y-\sum_{i=-l}^{l}\int_{I_i}\int_{I_j}K_a(y)u(x-y,t_n)\,\mathrm{d}x\,\mathrm{d}y\right|\\
&=\Delta t\sum_{n=1}^{N-1}\sum_{j\in\mathbb{Z}}\left|\sum_{i=-l}^{l}\int_{I_i}\int_{I_j}K_a(y)\left(u_{j-i}^n-u(x-y,t_n)\right)\mathrm{d}x\,\mathrm{d}y\right|\\
&\le\Delta t\sum_{n=1}^{N-1}\sum_{i=-l}^{l}\int_{I_i}K_a(y)\sum_{j\in\mathbb{Z}}\int_{I_j}\left|u_{j-i}^n-u(x-y,t_n)\right|\mathrm{d}x\,\mathrm{d}y.
\end{aligned}
$$

Using the convergence of u^Δ and the bound of K_a we get the result. Now, we continue with $\mathscr{B}$. Proceeding as above, we find that

$$
\begin{aligned}
\mathscr{B}&=\Delta t\sum_{n=1}^{N-1}\sum_{j\in\mathbb{Z}}\Bigg|\sum_{i=-l}^{l}\int_{I_i}K_a(y)\left(u_{j+1-i}^n-u_{j-i}^n\right)\mathrm{d}y\\
&\qquad\qquad-\int_{I_j}\sum_{i=-l}^{l}\int_{I_i}\partial_y K_a(y)u(x-y)\,\mathrm{d}y\,\mathrm{d}x\Bigg|\\
&\le\Delta t\sum_{n=1}^{N-1}\sum_{j\in\mathbb{Z}}\Bigg|\sum_{i=-l+1}^{l-1}\int_{I_i}\left(K_a(y+\Delta x)-K_a(y)\right)u_{j-i}^n\,\mathrm{d}y\\
&\qquad\qquad-\int_{I_j}\sum_{i=-l+1}^{l-1}\int_{I_i}\partial_y K_a(y)u(x-y)\,\mathrm{d}y\,\mathrm{d}x\Bigg|
\end{aligned}
$$

$$+\Delta t\sum_{n=1}^{N-1}\sum_{j\in\mathbb{Z}}\left|\int_{I_{-l}}K_a(y)u_{j+l+1}^n\,\mathrm{d}y-\int_{I_j}\int_{I_{-l}}\partial_y K_a(y)u(x-y)\,\mathrm{d}y\,\mathrm{d}x\right|$$
$$+\Delta t\sum_{n=1}^{N-1}\sum_{j\in\mathbb{Z}}\left|-\int_{I_l}K_a(y)u_{j-l}^n\,\mathrm{d}y-\int_{I_j}\int_{I_l}\partial_y K_a(y)u(x-y)\,\mathrm{d}y\,\mathrm{d}x\right|.$$

The last two terms of the last inequality are $\mathscr{O}(\Delta x)$ since $\partial_x K_a$ is bounded and $K_a(2a)=0$. Finally, we use a Taylor expansion and the convergence of u^Δ to get the result for almost all $k\in\mathbb{R}$. Proceeding as in Lemmas 4.3 and 4.4 of [63] we may extend the analysis to all $k\in\mathbb{R}$. □

2.5.3 An additional regularity result for $\alpha=0$

Lemma 2.5.5 *Assume that $\alpha=0$. Then the numerical solution generated by* (2.31) *converges to a Lipschitz continuous function u provided u_0 is also Lipschitz continuous.*

Proof. Defining $w_{j+1/2}^n:=(u_{j+1}^n-u_j^n)/\Delta x$ we obtain from (2.31)

$$w_{j-1/2}^{n+1}=w_{j+1/2}^n\left(\frac{1}{2}-\frac{\lambda}{2}V_{j+1}^n\right)+w_{j-3/2}^n\left(\frac{1}{2}+\frac{\lambda}{2}V_{j+1}^n\right)-w_{j-1/2}^n\frac{\lambda}{2}\Delta_+V_j^n$$
$$-w_{j-3/2}^n\frac{\lambda}{2}\left(V_{j+1}^n-V_{j-2}^n\right)-u_{j-1}^n\frac{\lambda}{2\Delta x}\left(\Delta^2V_j^n+\Delta^2V_{j-1}^n\right).$$

Using the CFL condition we have

$$\left|w_{j-1/2}^{n+1}\right|\le\left|w_{j+1/2}^n\right|\left(\frac{1}{2}-\frac{\lambda}{2}V_{j+1}^n\right)+\left|w_{j-3/2}^n\right|\left(\frac{1}{2}+\frac{\lambda}{2}V_{j+1}^n\right)+\frac{\lambda}{2}\left|w_{j-1/2}^n\right|\left|\Delta_+V_j^n\right|$$
$$+\frac{\lambda}{2}\left|w_{j-3/2}^n\right|\left|V_{j+1}^n-V_{j-2}^n\right|+\left|u_{j-1}^n\right|\frac{\lambda}{2\Delta x}\left|\Delta^2V_j^n+\Delta^2V_{j-1}^n\right|.$$

Lemmas 2.3.1, 2.3.2 and 2.5.2 imply that there exist constants C_9 and C_{10} such that

$$\left|w_{j-1/2}^{n+1}\right|\le\|W^n\|_\infty(1+C_9\Delta t)+C_{10}\Delta t.$$

Following the same steps as in the proof of Lemma 2.5.3 we obtain

$$\left|w_{j-1/2}^{n+1}\right|\le\|W^0\|_\infty\exp(C_9T)\left(1+\frac{C_{10}}{C_9}\right).$$

To conclude we notice that

$$w_{j+1/2}^0=\frac{u_{j+1}^0-u_j^0}{\Delta x}=\frac{1}{\Delta x^2}\left(\int_{x_{j+1/2}}^{x_{j+3/2}}u_0(y)\,\mathrm{d}y-\int_{x_{j-1/2}}^{x_{j+1/2}}u_0(y)\,\mathrm{d}y\right)$$

$$= \frac{1}{\Delta x^2}\left(\int_{x_{j-1/2}}^{x_{j+1/2}} \big(u_0(y+\Delta x) - u_0(y)\big)\,\mathrm{d}y\right) \leq \|u_0\|_{\mathrm{Lip}}.$$

The next step is to prove an analogous estimate for the discrete time derivative. Using (2.31) we can write

$$\begin{aligned} u_j^{n+1} - u_j^n =& \frac{u_{j+1}^n - u_j^n}{2} - \frac{u_j^n - u_{j-1}^n}{2} - \frac{\lambda}{2} V_{j+1}^n \left(u_{j+1}^n - u_{j-1}^n\right) \\ & - \frac{\lambda}{2} u_{j-1}^n \left(V_{j+1}^n - V_{j-1}^n\right). \end{aligned}$$

Multiplying this by Δt^{-1} and using that $\Delta t = \mathcal{O}(\Delta x)$ we find that there exists a constant C_{11}, which is independent of Δ, such that

$$\begin{aligned} \frac{u_j^{n+1} - u_j^n}{\Delta t} &= \frac{u_{j+1}^n - u_j^n}{2\Delta t} - \frac{u_j^n - u_{j-1}^n}{2\Delta t} - \frac{V_{j+1}^n}{2\Delta x}\left(u_{j+1}^n - u_{j-1}^n\right) - \frac{u_{j-1}^n}{2\Delta x}\left(V_{j+1}^n - V_{j-1}^n\right) \\ &\leq C_{11}\left(\frac{\left|u_{j+1}^n - u_j^n\right|}{2\Delta x} + \frac{\left|u_j^n - u_{j-1}^n\right|}{2\Delta x}\right) - \frac{V_{j+1}^n}{2\Delta x}\left(u_{j+1}^n - u_{j-1}^n\right) \\ &\quad - \frac{u_{j-1}^n}{2\Delta x}\left(V_{j+1}^n - V_{j-1}^n\right) \end{aligned}$$

Then, using that u^Δ is Lipschitz continuous respect to the space variable and Lemma 2.3.1, we get that the solution generated by the numerical method converges to a Lipschitz continuous function. □

Remark 2.5.2 *Lemma 2.5.5 is not a surprise since in the simplest case, V constant, the conservation law becomes a linear advection equation, whose solution has a regularity that is the same as that of the initial data. Moreover, the limit function u will be a Lipschitz continuous weak solution of* (2.1)*,* (2.2) *will automatically be an entropy solution, and stability and uniqueness are immediate from Theorem 2.4.1.*

2.5.4 Comparison with the analysis by Zumbrun [114]

The equation studied by Zumbrun, (2.12), is equivalent (up to a coordinate transformation) to (2.1) with $\alpha = 0$ and $V(w) = w$. The local existence of a bounded solution u with bounded spatial derivative u_x (provided that u_0 has corresponding properties) is proved in [114] by a fixed-point argument applied to the transport equation $u_t + (uK_a * v)_x = 0$ with given v. In general, global solutions inherit their regularity from u_0; in particular, if $\mathrm{TV}(u_0)$ is bounded, then (2.12) will have a BV solution u, which is unique following an L^1 argument with a discussion of entropy production terms at isolated discontinuities. In the

present work, existence of a solution of (2.1), (2.2) is shown by the convergence of a difference scheme, covering a wider range of cases of α and V. Moreover, our Lemma 2.5.5 is a rough equivalent of Zumbrun's result concerning the regularity of u in terms of that of u_0. Both the analysis of [114] and ours rely on estimates on u or u^Δ that blow up when $a \to 0$. This holds, in particular, for the L^∞-stability estimates of [114, Sect. 2]. However, as is shown in [114, Sect. 4], an L^2-stability argument can be invoked to prove that smooth solutions of (2.12) converge in L^∞ at an $\mathcal{O}(a^2)$ rate to smooth solutions of $u_t + (u^2)_x = 0$. (Of course, this result holds for smooth u_0 and a sufficiently small final time T.) The proof of this result in [114] depends on the linearity of V, and does not carry over to more general functions V or to $\alpha = 1$.

A detailed discussion is devoted in [114] to the existence of travelling wave solutions to (2.12) and (2.14), that is, of solutions of the form $u(x,t) = \varphi(x - st)$ with $\varphi(\xi) \to \varphi(\pm\infty)$ as $\xi \to \infty$ with either $\varphi(\infty) \neq \varphi(-\infty)$, as for a "viscous shock", or $\varphi(\infty) = \varphi(-\infty)$ corresponding to a solitary wave solution. (The "viscous shock"-type solution is of particular interest in the context of the sedimentation model, since it corresponds to the evolution of the suspension-supernate interface.) Roughly speaking, the result of [114] is that neither (2.12) nor (2.14) admit viscous shocks, but that both equations do admit solitary wave solutions. However, as mentioned in Section 2.1.4, solutions of (2.14) with an additional diffusion term du_{xx}, $d > 0$ with Riemann-like initial data do converge to a stable oscillatory travelling wave of "viscous shock" type. The analysis of travelling waves for (2.1) is outside the scope of this contribution, but the numerical results presented in Section 2.6 suggest that (2.1) for all values of α with a nonlinear function V equally supports oscillatory travelling waves of "viscous shock" type.

2.6 Numerical Examples

The numerical examples illustrate the qualitative behaviour of the solutions of (2.1), (2.2), with $\alpha = 0$ and $\alpha \geq 1$, and demostrate the convergence properties of the numerical scheme. For the first purpose, we select a relatively fine discretization and present the corresponding numerical solution as profiles at selected times, while the convergence properties of the scheme are illustrated by partly including error histories in some examples.

2.6.1 Example 1

We calculate the numerical solution of (2.1), (2.2) with $\alpha = 0$ for the hindered settling factor (2.3) with $n = 5$, and the kernel K given by (2.13) with $a = 0.2$. We are especially interested in phenomena produced at the suspension-supernate interface of a sedimenting

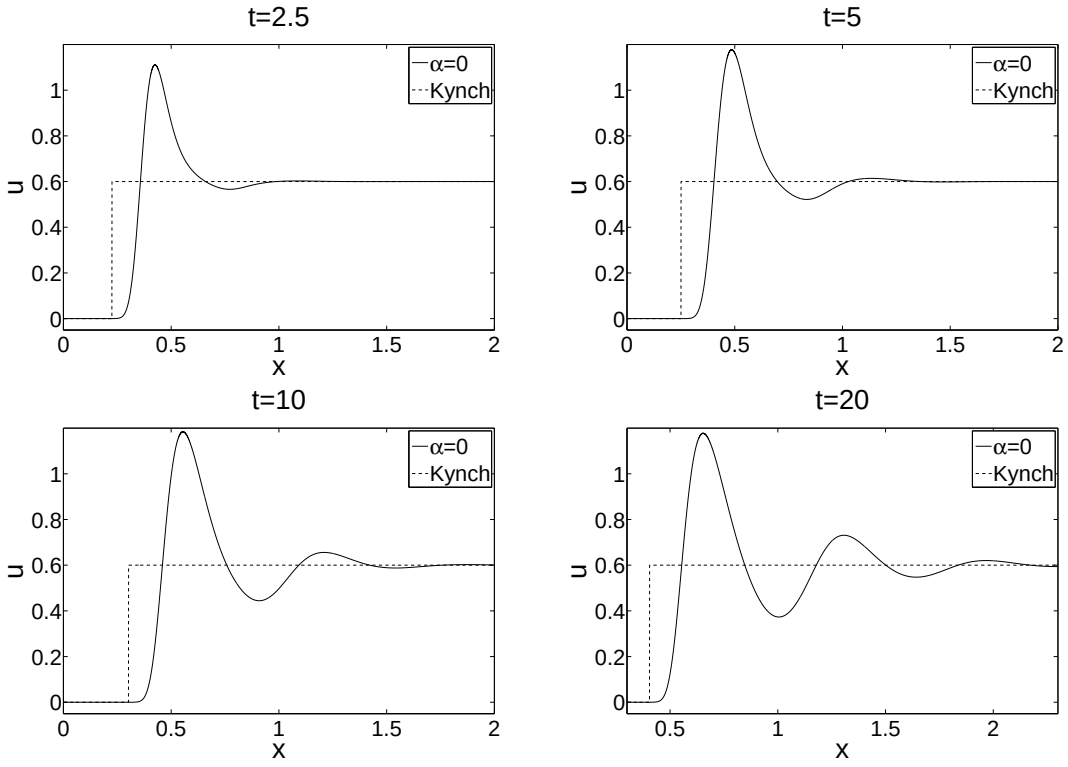

Figure 2.1: Example 1: Numerical solution of (2.1), (2.2) with $\alpha = 0$ and $a = 0.2$ for the hindered settling factor (2.3) with $n = 5$, for an initially concentrated suspension at $t = 2.5$, 5, 10 and 20.

suspension, and therefore employ the following Riemann initial data corresponding to the initial state of this interface for a concentrated and a dilute suspension, respectively:

$$u_0(x) = \begin{cases} 0.0 & \text{for } x \leq 0.2, \\ 0.6 & \text{for } x > 0.2, \end{cases} \quad \text{and} \quad u_0(x) = \begin{cases} 0.0 & \text{for } x \leq 0.2, \\ 0.01 & \text{for } x > 0.2. \end{cases} \tag{2.40}$$

In both cases we use $\Delta x = 0.0005$ and $\lambda = 0.2$. The results are shown in Figures 2.1 and 2.2 for the respective cases of an initially concentrated and dilute suspension. As predicted in Section 2.2.2, we obtain the formation of layers of mass due to the non-constancy of the initial data. We also plot the corresponding solution for the local equation (2.4), which we call the "Kynch solution."

We can conjecture from these simulations, that even though u_0 is not smooth, the presence of the kernel has a regularizating effect since we do not observe the formation of discontinuities. Moreover, we see that the numerical solution is not in $[0,1]$ for the concentrated suspension accordingly with Lemma 2.5.2 even though u_0 assumes values

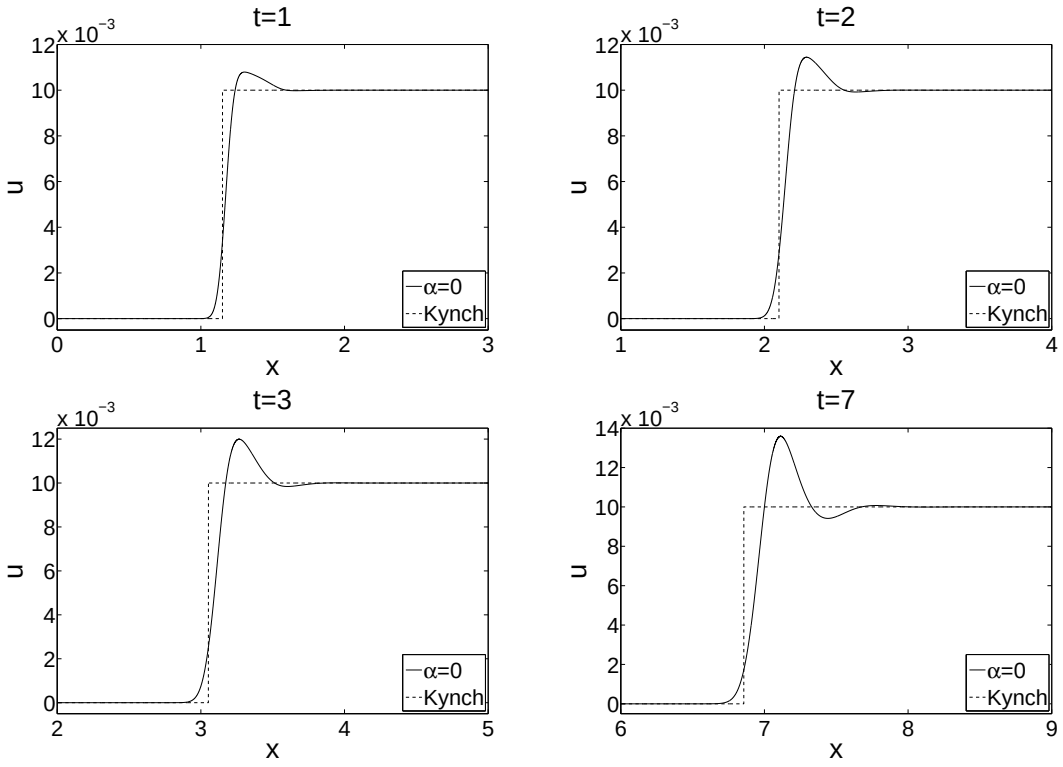

Figure 2.2: Example 1: Numerical solution of (2.1), (2.2) with $\alpha = 0$ and $a = 0.2$ for the hindered settling factor (2.3) with $n = 5$, for an initially dilute suspension at $t = 1, 2, 3$ and 7.

from that interval. In Table 2.1 we show the error at $t_1 = 1$ and $t_2 = 3$ in the L^1 norm for u (denoted by $e^{t_i}_{c/d}$, $i = 1,2$) where we take as a reference the solution calculated with $\Delta x = 0.0005$. As expected for the Lax-Friedrichs method, we obtain an experimental order of convergence one. In addition to Table 2.1 we show in Figure 2.3 the "graphical" approximation.

2.6.2 Example 2

We study now the behaviour of the numerical solution to (2.1), (2.2) with $\alpha = 1$. We use V as given by (2.3) with $n = 4$ and K given by (2.13) with $a = 0.2$. We again utilize the initial datum (2.40) with $\Delta x = 0.0005$ and $\lambda = 0.2$. The results are plotted in Figures 2.4 and 2.5.

We observe the presence of layers but of smaller amplitude than those observed in Example 3.5.1. We explain this by the different flux function. We also observe more pro-

Δx	$e_c^{t_1}$	conv. rate	$e_c^{t_2}$	conv. rate	$e_d^{t_1}$	conv. rate	$e_d^{t_2}$	conv. rate
1.00E-2	8.62E-3	-	1.33E-2	-	1.29E-4	-	2.25E-4	-
5.00E-3	6.24E-3	0.47	1.07E-2	0.31	8.34E-5	0.63	1.53E-4	0.56
4.00E-3	5.45E-3	0.60	9.07E-3	0.46	7.10E-5	0.72	1.33E-4	0.61
2.00E-3	3.26E-3	0.74	6.37E-3	0.61	3.91E-5	0.86	7.89E-5	0.75
1.25E-3	1.99E-4	1.05	4.11E-3	0.93	2.27E-5	1.16	4.73E-5	1.09

Table 2.1: Example 1: Numerical error for u at $t_1 = 1$ and $t_2 = 3$.

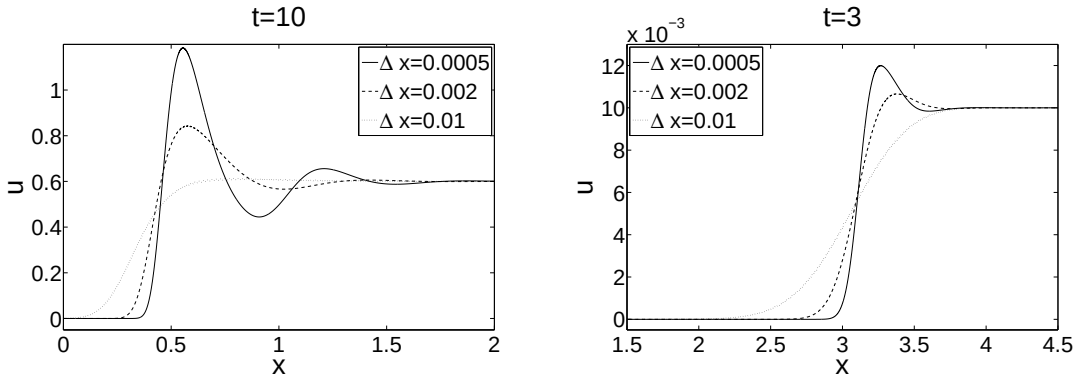

Figure 2.3: Example 1: Numerical solution of (2.1), (2.2) with $\alpha = 0$ and $a = 0.2$ for the hindered settling factor (2.3) with $n = 5$ for $\Delta x = 0.01$, $\Delta x = 0.002$ and $\Delta x = 0.0005$.

nounced gradients in the solution, which is in agreement with results proved in Section 2.5. In Table 2.2 we show the error at $t_1 = 1$ and $t_2 = 3$ in the L^1 norm for u where we take as a reference the solution calculated with $\Delta x = 0.0005$ as in Example 3.5.1. We again get an experimental order of convergence one. Figure 2.6 shows the graphical approximation.

2.6.3 Example 3

We now examine how changes in the parameter a affect qualitatively the numerical solution of (2.1), (2.2) for $\alpha = 0$ and $\alpha = 1$. We use (2.3) with $n = 5$ for $\alpha = 0$ and correspondingly, (2.3) with $n = 4$ for $\alpha = 1$. In both cases, K is given by (2.13) with the parameter $a = 0.4$, 0.2, 0.1 and 0.01. The initial datum is (2.40) for the two cases of a concentrated and a dilute suspension with $\Delta x = 0.0005$ and $\lambda = 0.2$. Figure 2.7 shows the results at $t = 10$ and $t = 7$ in the concentrated and dilute case, respectively.

The case $a = 0.01$ was calculated with $\Delta x = 0.0002$ since if we consider the parameter

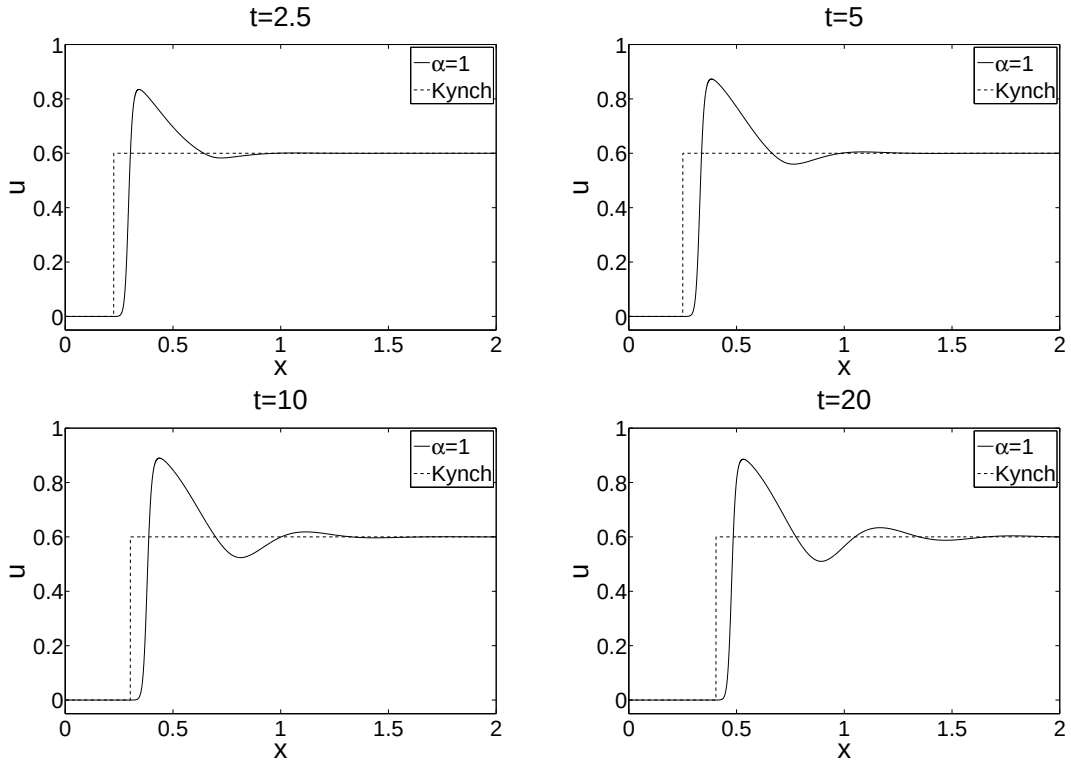

Figure 2.4: Example 2: Numerical solution of (2.1), (2.2) with $\alpha = 1$ and $a = 0.2$ for the hindered settling factor (2.3) with $n = 4$ for an initially concentrated suspension at $t = 2.5$, 5, 10 and 20.

a "close" to Δx we get the Kynch result because the stencil of the convolution includes just a few points, and the numerical scheme can be viewed as a mollification scheme [1]. We observe a more strongly oscillatory behaviour with $a = 0.2$ and $a = 0.1$, and that the period of the oscillation is proportional to the value of a for both cases. The peak in the case $\alpha = 0$ occurs for $a = 0.4$ and in the case $\alpha = 1$ there is no difference between the peak with $a = 0.2$ and $a = 0.4$. We explain this by the dispersive behaviour of the formulation.

2.6.4 Example 4

The idea of the present example is try to reproduce the layered sedimentation observed by Siano [109] in a batch process. The obvious difficulty appears when we are "close" to the boundary since in a batch process we have a zero flux condition and for the numerical computations we have to extrapolate values in order to compute the numerical fluxes. To solve this problem, we assume that outside the volume control we have initial concentra-

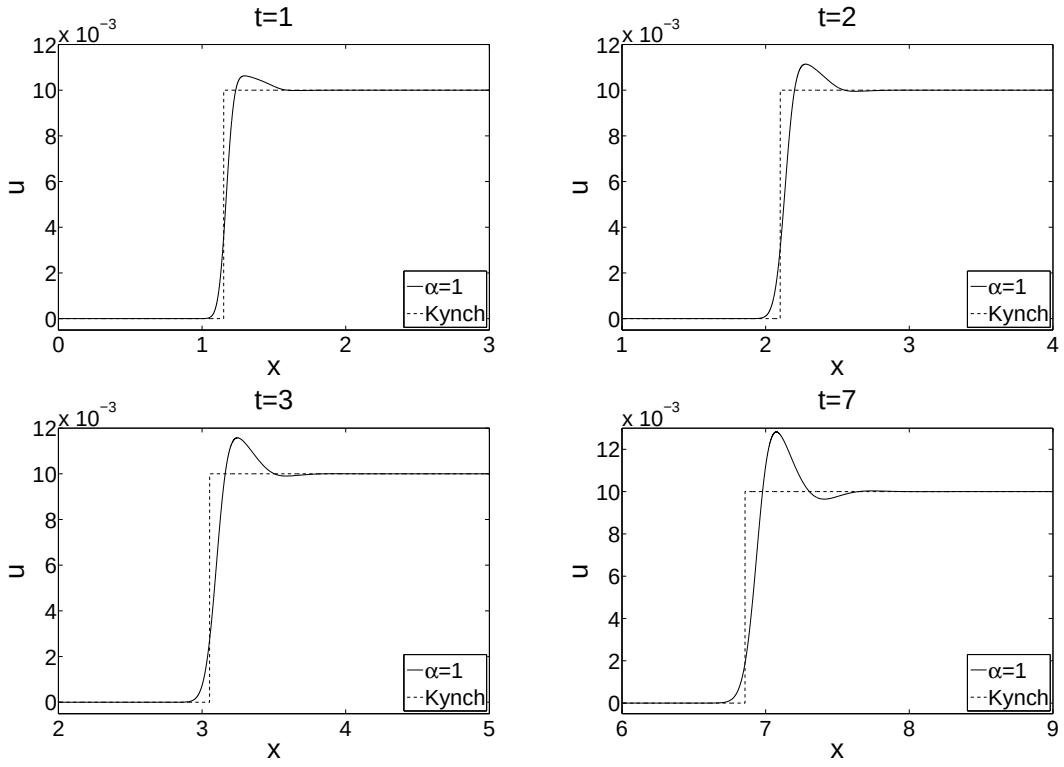

Figure 2.5: Example 2: Numerical solution of (2.1), (2.2) with $\alpha = 1$ and $a = 0.2$ for the hindered settling factor (2.3) with $n = 4$ for an initially dilute suspension at $t = 1$, 2, 3 and 7.

tion values, 0 to the left and 1 to the right. In Figures 2.8 and 2.9 we show the numerical results for $\alpha = 1$, with $V(u) = (1-u)^4$, K as in (2.13), $a = 0.025$, $\Delta x = 0.00025$, $\lambda = 0.5$ and the respective initial datum for concentrated and dilute suspensions given by

$$u_0(x) = \begin{cases} 0 & \text{for } x < 0, \\ 0.5 & \text{for } 0 \le x < 1, \\ 1 & \text{for } x \ge 1 \end{cases} \quad \text{and} \quad u_0(x) = \begin{cases} 0 & \text{for } x < 0, \\ 0.05 & \text{for } 0 \le x < 1, \\ 1 & \text{for } x \ge 1. \end{cases} \tag{2.41}$$

In each figure we also plot the solution obtained by the local model (Kynch solution). We observe that the layers smooth after a while.

2.6.5 Example 5

In Figures 2.10–2.12 we plot the solution for u^Δ for $\alpha = 1$, with $V(u) = (1-u)^4$, K as in (2.13), $a = 0.025$ and $a = 0.05$ and we consider two different initial data. For the first

Δx	$e_c^{t_1}$	conv. rate	$e_c^{t_2}$	conv. rate	$e_d^{t_1}$	conv. rate	$e_d^{t_2}$	conv. rate
1.00E-2	7.06E-3	-	9.66E-3	-	1.28E-4	-	2.22E-4	-
5.00E-3	4.67E-3	0.59	6.93E-3	0.48	8.18E-5	0.64	1.46E-4	0.60
4.00E-3	3.95E-3	0.76	5.97E-3	0.67	6.96E-5	0.73	1.27E-4	0.63
2.00E-3	2.08E-3	0.92	3.30E-3	0.86	3.83E-5	0.86	7.43E-5	0.77
1.25E-3	1.15E-3	1.26	1.84E-4	1.24	2.22E-5	1.16	4.43E-5	1.10

Table 2.2: Example 2: Numerical error for u at $t_1 = 1$ and $t_2 = 3$.

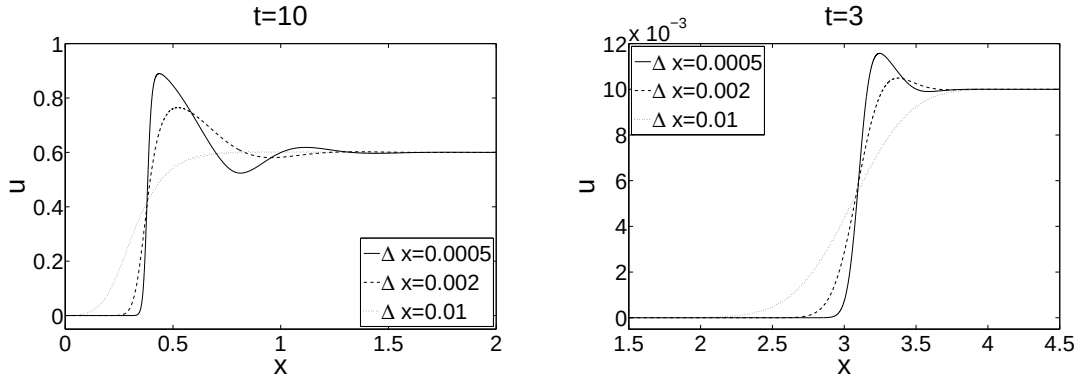

Figure 2.6: Example 2: Numerical solution of (2.1), (2.2) with $\alpha = 1$ and $a = 0.2$ for the hindered settling factor (2.3) with $n = 4$ for an initially concentrated (left) and dilute (right) suspension for $\Delta x = 0.01$, $\Delta x = 0.002$ and $\Delta x = 0.0005$.

one we take u_0 as in Example 4 and u_0 given by

$$u_0(x) = \begin{cases} 0 & \text{for } x < 0, \\ 0.4 + 0.2x & \text{for } 0 \le x < 1, \\ 1 & \text{for } x \ge 1 \end{cases} \tag{2.42}$$

for the concentrated case and

$$u_0(x) = \begin{cases} 0 & \text{for } x < 0, \\ 0.04 + 0.02x & \text{for } 0 \le x < 1, \\ 1 & \text{for } x \ge 1, \end{cases} \tag{2.43}$$

for the dilute case. We also use a nonlinear scale in color in order to highlight the layering phenomenon, which is supposed to appear in the range of concentrations close to the initial concentration. We observe the presence of layers in the case with u_0 given by the

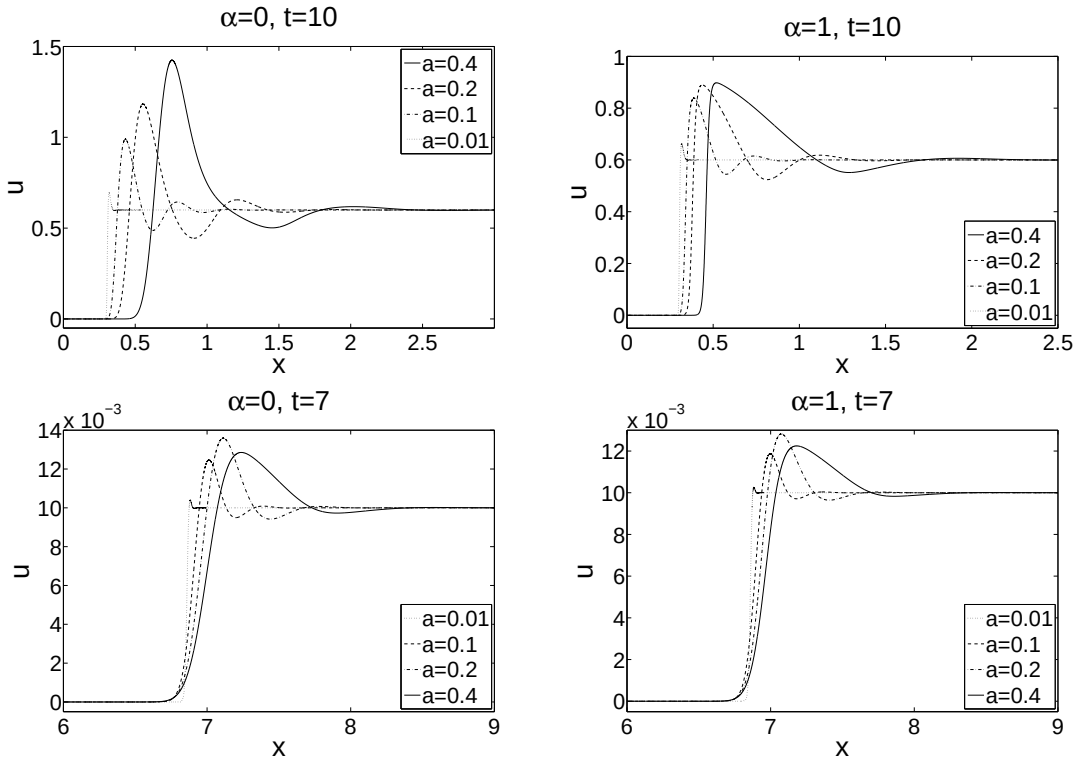

Figure 2.7: Example 3: Numerical solution of (2.1), (2.2) for the indicated values of α with $a = 0.4$, 0.2, 0.1 and 0.01 (top) for an initially concentrated suspension, at $t = 10$ and (bottom) for an initially dilute suspension, at $t = 7$.

Riemann data (2.41) in a more pronounced form than for the linear initial data (2.42) and (2.43). As we explain in Section 2.2.2, the presence of layers occurs only if the initial concentration exhibits strong variation, e.g. a jump between zero and a positive constant. We also see, comparing Figures 2.10 and 2.12, that the "width" of the layer is proportional to the parameter a.

2.7 Conclusions

We study a greater variety of models than the one proposed in [114], which corresponds to $\alpha = 0$ and a linear function V. The model corresponding to $\alpha = 1$ is consistent with (2.18) and (2.19) in the dilute limit $\varphi \to 0$, but assumes values in $[0,1]$ only and therefore can be applied to the whole range of concentrations. The treatment of the boundary conditions can possibly be improved. Our analysis shows that a reasonably simple difference-quadrature schemes converges to the entropy solution. However, since it is

based on the Lax-Friedrichs scheme, high-order versions should be used for practical computations.

We have conducted numerical experiments aiming at assessing whether (2.1) can possibly explain the phenomenon of layering in sedimentation. The numerical experiments, and especially the plots of Figures 2.10–2.12, illustrate that (2.1) indeed produces patterns that are similar to layering, namely vertical fluctuations of concentration of $\mathscr{O}(a)$ with beneath the suspension-supernate interface. These oscillatory travelling waves of "viscous shock" type disappear when they start to interfere with solution information propagating upwards (in the direction of decreasing x). One should mention, however, that this phenomenon differs from "layering" as observed by Siano [109] in that the solution exhibits oscillations rather than staircasing. As mentioned in [114] it would be interesting to explore further whether (2.1) produces solutions more similar to the staircasing phenomenon if this equation were equipped with additional standard or nonstandard diffusion terms.

Finally, a systematic travelling wave analysis of (2.1), which would extend the results of [114], is still lacking. Such an analysis could explain whether new phenomena, e.g. nonclassical shocks, should be expected when one considers the formal limit $a \to 0$ of entropy solutions of (2.1), especially in the case $\alpha \geq 1$. Unfortunately, most of the constants appearing in the compactness estimates of Section 2.5.1 are not uniform with respect to a, i.e. they blow up when $a \to 0$. It is therefore not clear at the moment whether a sequence of entropy solutions converges to a meaningful limit as $a \to 0$.

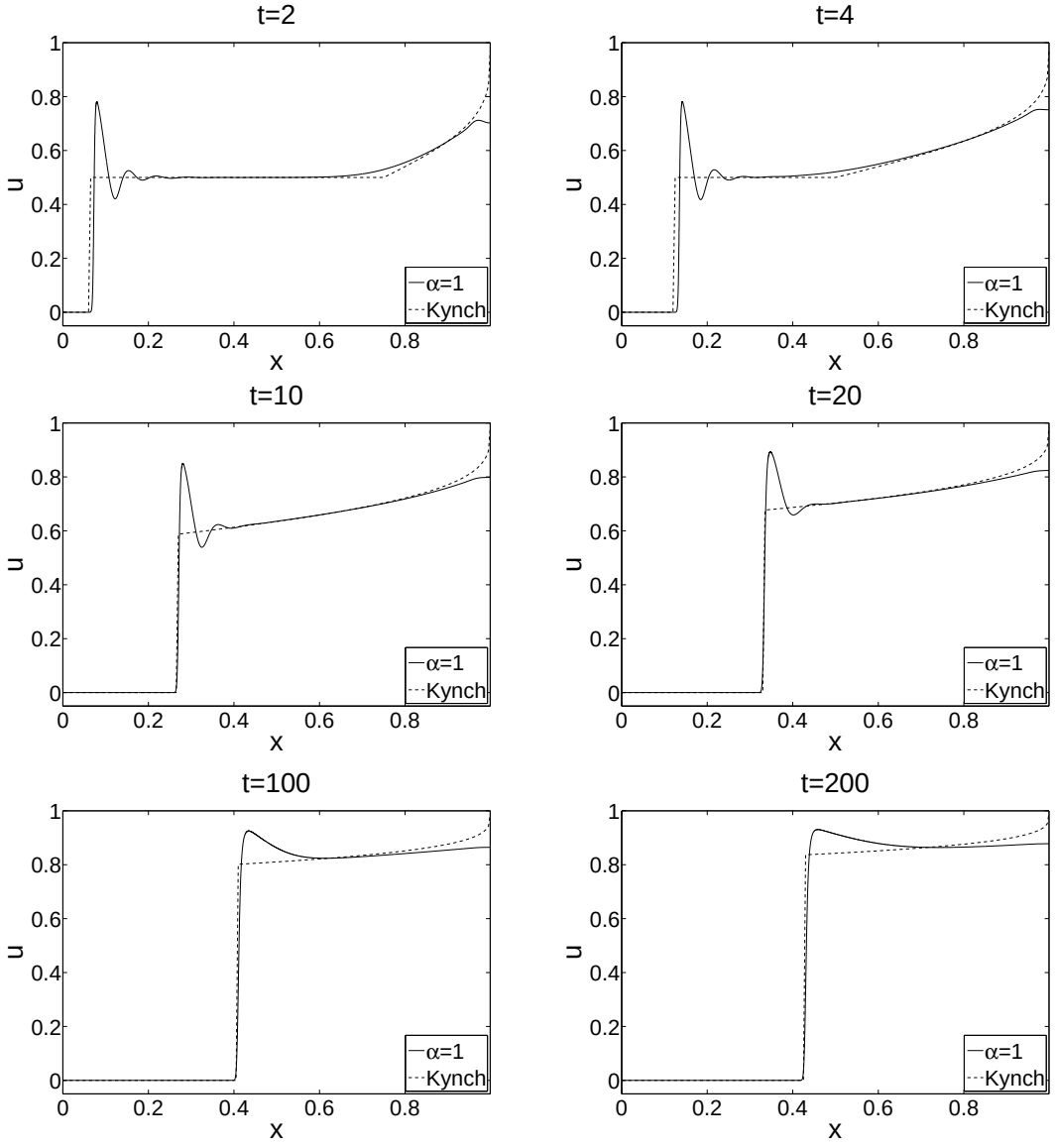

Figure 2.8: Example 4: Numerical solution of (2.1), (2.2) with $\alpha = 1$ for the hindered settling factor (2.3) with $n = 4$ and $a = 0.025$ for an initially concentrated suspension.

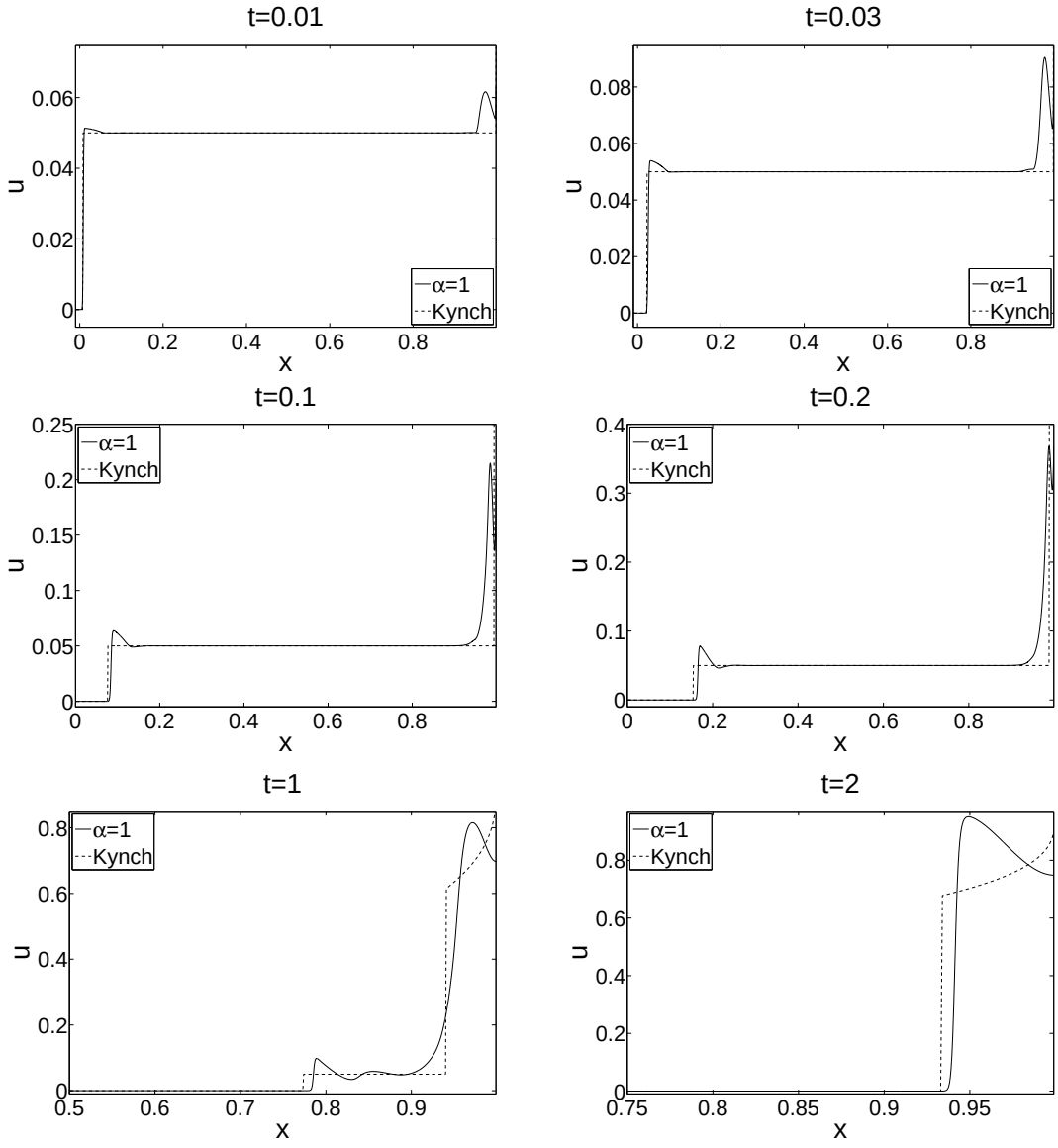

Figure 2.9: Example 4: Numerical solution of (2.1), (2.2) with $\alpha = 1$ for the hindered settling factor (2.3) with $n = 4$ and $a = 0.025$ for an initially dilute suspension.

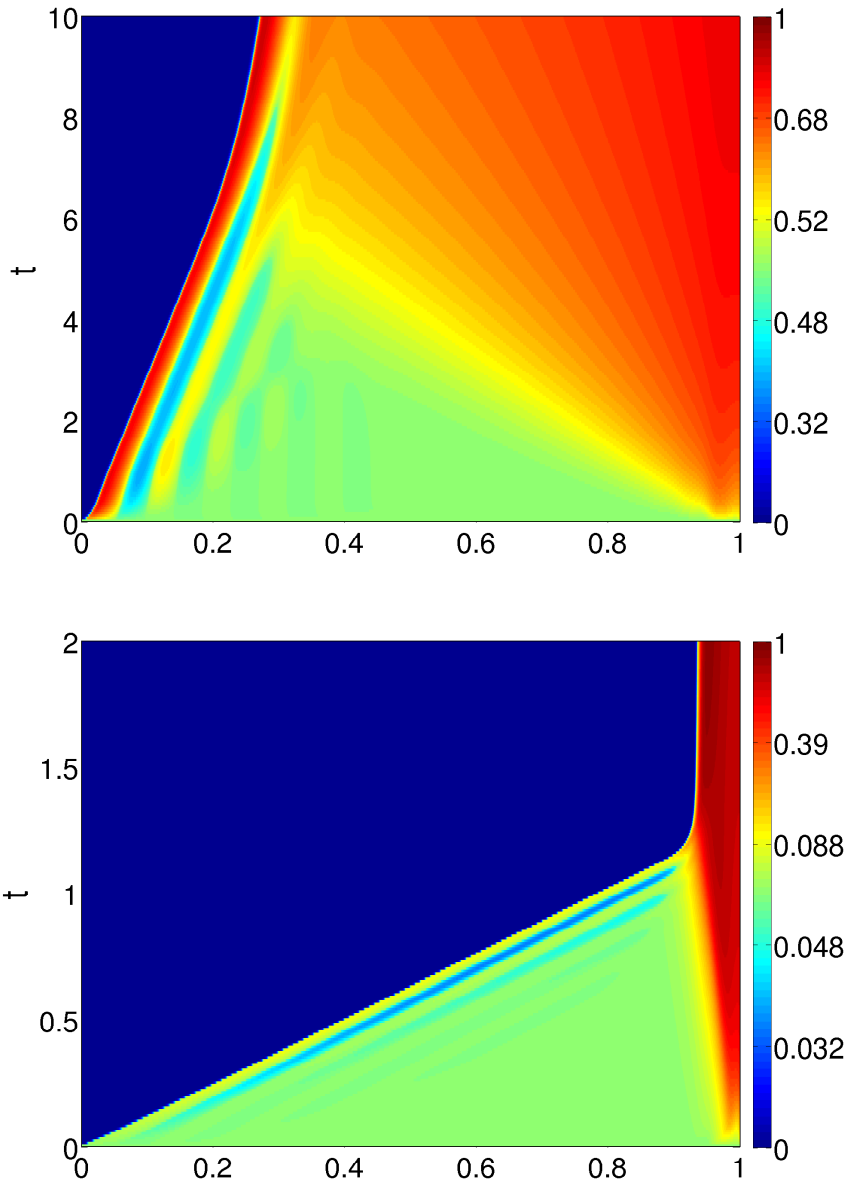

Figure 2.10: Example 5: Numerical solution of (2.1), (2.2) with $\alpha = 1$ for the hindered settling factor (2.3) with $n = 4$ and $a = 0.025$ for an initially concentrated (above) and dilute (below) suspension with u_0 constant.

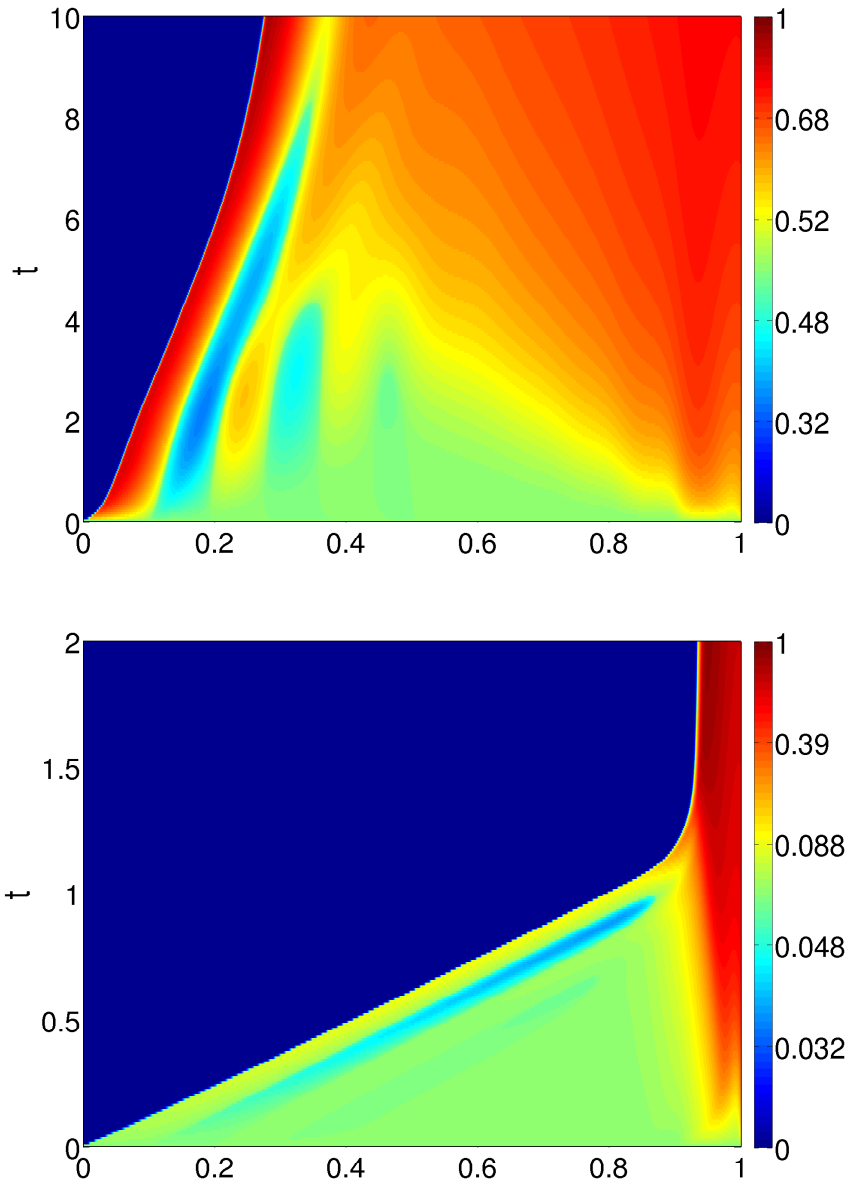

Figure 2.11: Example 5: Numerical solution of (2.1), (2.2) with $\alpha = 1$ for the hindered settling factor (2.3) with $n = 4$ and $a = 0.05$ for an initially concentrated (above) and dilute (below) suspension with u_0 constant.

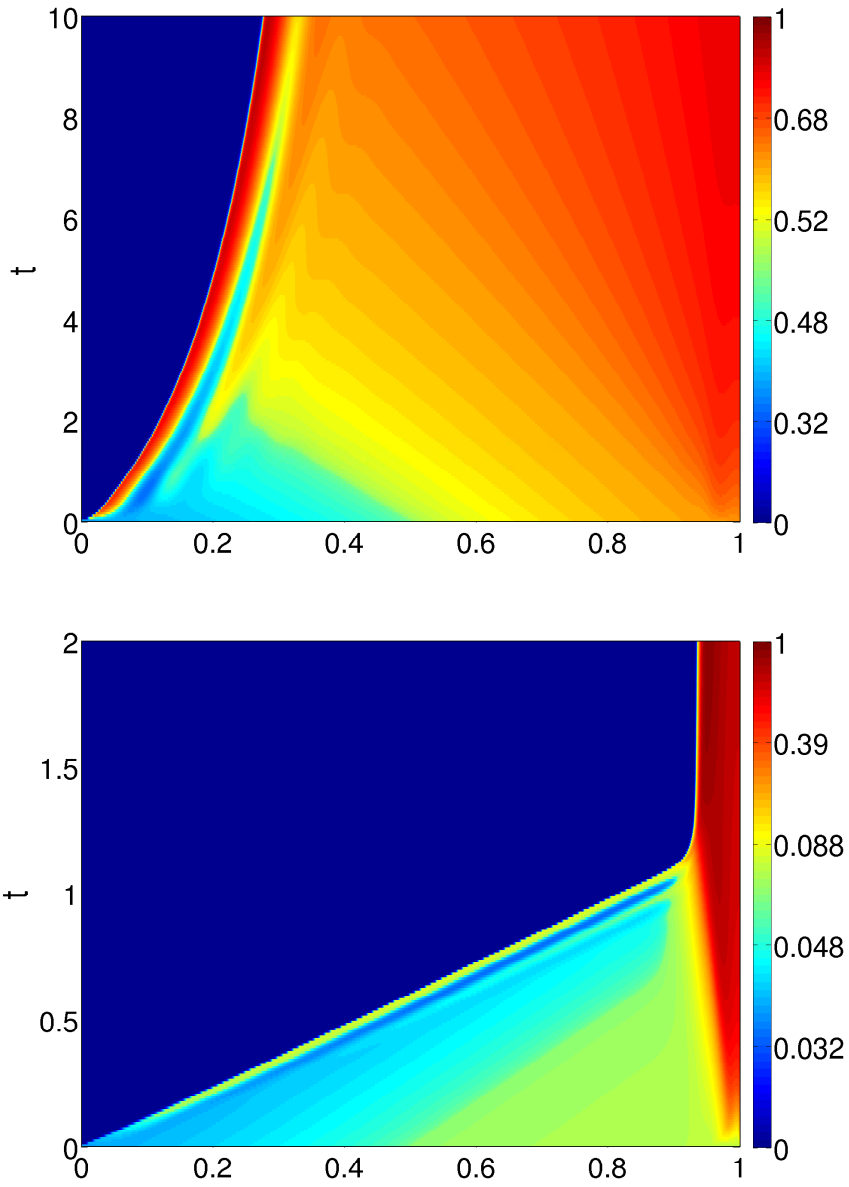

Figure 2.12: Example 5: Numerical solution of (2.1), (2.2) with $\alpha = 1$ for the hindered settling factor (2.3) with $n = 4$ and $a = 0.025$ for an initially concentrated (above) and dilute (below) suspension with a linear initial concentration u_0.

Chapter 3

Finite-Volume Schemes for Friedrichs Systems with Involutions

In applications solutions of systems of hyperbolic balance laws often have to satisfy additional side conditions. We consider initial value problems for the general class of Friedrichs systems where the solutions are constrained by differential conditions given in the form of involutions. These occur in particular in electrodynamics, electro- and magnetohydrodynamics as well as in elastodynamics. Neglecting the involution on the discrete level typically leads to instabilities.
To overcome this problem in electrodynamical applications it has been suggested in Munz et al. (2000) to solve an extended system. Here we suggest an extended formulation to the general class of constrained Friedrichs systems. It is proven for Finite-Volume schemes that the discrete solution of the extended system converges to the weak solution of the original system for vanishing discretization and extension parameter under appropriate scalings. Moreover we show that the involution is weakly satisfied in the limit. The proofs rely on a reformulation of the extension as a relaxation-type approximation and careful use of the convergence theory for Finite-Volume methods for systems of Friedrichs type. Numerical experiments illustrate our analytical results.

3.1 Introduction

In this chapter, we study linear systems of balance laws, namely $(m \times m)$-systems of Friedrichs [50] type with $m \in \mathbb{N}$. We consider the spatially d-dimensional case with $d \geq 2$, space coordinates $x = (x_1, \ldots, x_d)^T$, and time $t \geq 0$. For $T > 0$, let $G^1, \ldots, G^d, D : \mathbb{R}^d \times [0,T] \to \mathbb{R}^{m \times m}$ and $f : \mathbb{R}^d \times [0,T] \to \mathbb{R}^m$ be given (matrix-valued) functions. We suppose that the matrices $G^1(x,t), \ldots, G^d(x,t)$ are symmetric for all $(x,t) \in \mathbb{R}^d$. Then the initial value problem for the unknown vector-valued function $u : \mathbb{R}^d \times [0,T] \to \mathbb{R}^m$ takes

the form

$$\frac{\partial}{\partial t}u(x,t)+\sum_{i=1}^{d}\frac{\partial}{\partial x_i}\left(G^i(x,t)u(x,t)\right)+D(x,t)u(x,t)=f(x,t), \tag{3.1}$$

$$u(x,0)=u_0(x). \tag{3.2}$$

Here $u_0:\mathbb{R}^d\to\mathbb{R}^m$ denotes the initial function. Moreover we require the solution u to satisfy a linear differential side condition of the form

$$\sum_{i=1}^{d}M_i\frac{\partial}{\partial x_i}\left(u(x,t)\right)=0,\qquad \Big((x,t)\in\mathbb{R}^d\times[0,T)\Big). \tag{3.3}$$

Here M_i, $i=1,\dots,d$, are constant $(m\times m)$-matrices. Following the notion of Dafermos [39, 40] for the side condition (3.3), we restrict ourselves to involutions.

Definition 3.1.1 *The differential constraint* (3.3) *is called an* ***involution*** *for the system* (3.1) *if and only if any (weak) solution of* (3.1)-(3.2) *(weakly) satisfies* (3.3)*, whenever the initial data do so.*

Involutions appear frequently in applications. We mention the classical Maxwell system to describe electrodynamical processes (cf.[75]). The divergence of the electrical and magnetical field is constrained in this case. The induction equations in the (in)compressible electro- and magnetohydrodynamical equations provide similar examples but with (x,t)-dependence in the flux (Sect. 3.5 below). Solutions of the equations of linear elasticity have to satisfy compatibility conditions on the deformation gradient, which result in an involutionary condition (cf. Chapter 5 of [39]) Yet another example is the linear piezoelectrical system (see [84]). In Sect. 3.5 we present some of these examples in more detail. Let us mention that involutions of course appear also in the more challenging case of nonlinear conservation laws. Again magnetohydrodynamics [38], electrohydrodynamics, nonlinear elasticity systems, but also Einstein's equations of general relativity are prominent examples.
On the analytical level an involutionary side condition is not problematic. The well-posedness for (3.1)-(3.3) is well known from [39]. By definition the involution (3.3) is satisfied. Also standard numerical schemes are known to converge. However, without consideration of (3.3) in the numerical scheme the residuum in the side condition usually grows with increasing time. In coupled processes this is a typical source of instabilities (cf.[88] and cites therein). Therefore a wide range of stabilization methods has been suggested (e.g. [4, 20, 35, 58, 89]).
The motivation for this contribution is the work of Munz *et al.* [89]. They introduced in particular the so-called hyperbolic Generalized Lagrangian Multiplier Finite Volume

method (GLM-FV) to compute approximate solutions for Maxwell's system of linear electrodynamics. We formulate this approach for the general problem (3.1)-(3.2) with involution (3.3). While the original approach is motivated by a generalization of a Finite-Element type method [4] for a constrained wave equation we consider the approach as the approximation of (3.1)-(3.3) by an extended *relaxation-type system*. Relaxation approximations of systems of conservation laws have been intensively studied in the last decade (see [113] for an overview).
To be precise let $a, \varepsilon > 0$ and $u_0, \psi_0^\varepsilon : \mathbb{R}^d \to \mathbb{R}^m$ be given. Consider the following initial value problem for the unknown function:
$w^\varepsilon : \mathbb{R}^d \times [0,T] \to \mathbb{R}^{2m}$, $w^\varepsilon := (u_1^\varepsilon, \ldots, u_m^\varepsilon, \psi_1^\varepsilon, \ldots, \psi_m^\varepsilon)^T$
satisfying

$$\frac{\partial}{\partial t} u^\varepsilon + \sum_{i=1}^{d} \frac{\partial}{\partial x_i} \left(G^i(x,t) u^\varepsilon \right) + M_i^T \frac{\partial}{\partial x_i} \psi^\varepsilon + D(x,t) u^\varepsilon = f(x,t), \tag{3.4}$$

$$\frac{\partial}{\partial t} \psi^\varepsilon + \sum_{i=1}^{d} \frac{M_i}{\varepsilon} \frac{\partial}{\partial x_i} u^\varepsilon + a \psi^\varepsilon = 0, \tag{3.5}$$

and

$$u^\varepsilon(x,0) = u_0^\varepsilon(x), \qquad \psi^\varepsilon(x,0) = \psi_0^\varepsilon(x). \tag{3.6}$$

We will show in Section 3.2 that the initial value problem for the extended system (3.4)-(3.6) is well-posed. For vanishing parameter ε we prove under mild assumptions on the coefficients (see Proposition 3.2.1) that $u^\varepsilon = u$, a.e., where u is the solution of (3.1)-(3.2). In Section 3.3 we present the Generalized Lagrangian Multiplier Finite Volume Method (GLM-FV) for the general system (3.1)-(3.3). For mesh parameter $h > 0$ this gives us the mesh function $u_h^\varepsilon : \mathbb{R}^d \times [0,T] \to \mathbb{R}^m$. The method will be analyzed in Sect. 3.4. By careful investigation of the convergence theory from Vila and Villedieu [112] and Jovanovic and Rohde [59] we obtain (see Theorem 3.4.1)

$$\| u_h^\varepsilon - u^\varepsilon \|_{L^2(\mathbb{R}^d \times [0,T]; \mathbb{R}^m)} = \mathscr{O}\left(\varepsilon^{-1/4} h^{1/2} \right) \tag{3.7}$$

The crucial fact is that the estimate does not depend critically on the parameter ε. This expresses the dissipative character of the approximation (3.4)-(3.6). Moreover we will show that the weak constraint error goes to zero if h and ε vanishes (Corollary 3.4.1).

Up to our knowledge convergence statements as Corollary 3.4.1 have not been derived for any of the existing methods to handle involutionary systems ([4, 20, 35, 58, 89]) .
The assumptions, definitions, general results on Friedrichs systems and some notation are summarized in Section 3.2, while Section 3.3 is devoted to the numerical scheme. Section 3.4 contains the analysis of the scheme and in particular the proofs of the main convergence theorems (Theorem 3.4.1 and Corollary 3.4.1). In the last section we present

applications and numerical examples.
Finally we comment on related work for nonlinear systems of hyperbolic balance laws. The approach of Munz *et al.* [89] has been transferred to the system of compressible magnetohydrodynamics in [41]. Discontinuous-Galerkin methods with locally divergent-free ansatz functions have been introduced in [80] and studied for the MHD equations in [20]. Even much earlier Powell [99] suggested an (non-relaxation) extension of the magnetohydrodynamical system, see also [45]. A general approach can be found in Torrilhon [106], which has been applied to nonlinear systems in [49] and in [57].

3.2 Preliminaries

For $d,m \in \mathbb{N}$ we denote by $L^2(\mathbb{R}^d;\mathbb{R}^m)$, $H^1(\mathbb{R}^d;\mathbb{R}^m)$ the usual Lebesgue and Sobolev spaces equipped with the norms $\|\cdot\|_{L^2(\mathbb{R}^d;\mathbb{R}^m)}$, $\|\cdot\|_{H^1(\mathbb{R}^d;\mathbb{R}^m)}$, respectively. $C_b^{0,1}(\mathbb{R}^d\times[0,T])$ is the set of bounded, Lipschitz continuous functions on $\mathbb{R}^d\times[0,T]$. Furthermore, for $l\in\mathbb{N}$ we need the Bochner spaces $C^l([0,T];X)$ and $L^2(0,T;X)$ where X is an arbitrary function space. The corresponding norms are denoted by $\|\cdot\|_{C^l([0,T];X)}$ and $\|\cdot\|_{L^2(0,T;X)}$. For $M:\mathbb{R}^d\times[0,T]\to\mathbb{R}^{m\times m}$ we define

$$\|M\| = \sup_{(x,t)\in\mathbb{R}^d\times[0,T]} \|M(x,t)\|_2,$$

where $\|\cdot\|_2$ denotes the spectral norm. By $\mathscr{C}>0$ we denote a generic constant (that can change from a line to the next!) independent on h and ε.

Definition 3.2.1 *We say $u\in L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^m))$ is called a **weak solution** of* (3.1)-(3.3) *if*

$$\int_{\mathbb{R}^d}\int_0^T \left(u\cdot\frac{\partial\rho}{\partial t} + \sum_{i=1}^d G^i(x,t)u\cdot\frac{\partial\rho}{\partial x_i} - D(x,t)u\cdot\rho + f\cdot\rho \right) dxdt = \int_{\mathbb{R}^d} u_0(x)\cdot\rho(x,0)\,dx \tag{3.8}$$

holds for all $\rho\in C_0^\infty(\mathbb{R}^d\times[0,T);\mathbb{R}^m)$.

Recall that, since (3.3) is an involution (see Definition 3.1.1), it is satisfied automatically in the weak sense for any weak solution. We specify all assumptions on the coefficients in Assumption 3.2.1 below. We note that in particular the regularity statement (i) can be relaxed, however, it does not lead to a better result in terms of the order of convergence.

Assumption 3.2.1 *Consider the initial value problem* (3.1)-(3.3).

(i) The mappings $D, G^1, \dots, G^d \in C^\infty(\mathbb{R}^d \times [0,T], \mathbb{R}^{m\times m})$ *satisfy*

$$G^i(x,t)^T = G^i(x,t) \quad \forall (x,t) \in \mathbb{R}^d \times [0,T] \ (i = 1,\dots,d),$$
$$\sum_{i=1}^{d} \left(\|\partial_t^j \partial_x^\alpha G^i\| + \|\partial_t^j \partial_x^\alpha D\| \right) < +\infty \qquad \forall \alpha \in \mathbb{N}_0^d, j \in \mathbb{N}_0.$$

(ii) The functions u_0, f *satisfy* $u_0 \in H^1(\mathbb{R}^d;\mathbb{R}^m)$, $f \in L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^m))$ *and* u_0 *also satisfies* (3.3).

(iii) The $(m\times m)$*-matrices* M_i *are constant for* $i = 1,\dots,d$.

We proceed with the presentation of the extended GLM formulation (3.4)-(3.6).

Definition 3.2.2 *For* $a, \varepsilon > 0$ *the function* $(u^\varepsilon, \psi^\varepsilon)^T \in L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))$ *is called a* ***weak solution*** *of the extended problem* (3.4)-(3.6) *if*

$$\int_{\mathbb{R}^d}\int_0^T \left(u^\varepsilon \cdot \frac{\partial \rho}{\partial t} + \sum_{i=1}^{d} \left(G^i(x,t)u^\varepsilon + M_i^T \psi^\varepsilon \right) \cdot \frac{\partial \rho}{\partial x_i} - D(x,t)u^\varepsilon \cdot \rho + f \cdot \rho \right) dxdt$$
$$= \int_{\mathbb{R}^d} u_0(x) \cdot \rho(x,0)\, dx,$$
$$\int_{\mathbb{R}^d}\int_0^T \left(\psi^\varepsilon \cdot \frac{\partial \omega}{\partial t} + \sum_{i=1}^{d} \frac{M_i}{\varepsilon} u^\varepsilon \cdot \frac{\partial \omega}{\partial x_i} - a\psi^\varepsilon \cdot \omega \right) dxdt = \int_{\mathbb{R}^d} \psi_0^\varepsilon(x) \cdot \omega(x,0)\, dx,$$

holds for all $\rho, \omega \in C_0^\infty(\mathbb{R}^d \times [0,T);\mathbb{R}^m)$.

This approach generalies the idea of Munz et al. [89] to arbitrary Friedrichs systems. The small parameter ε has to be identified with the ratio $\mathscr{C}_h/\mathscr{C}_p$ in Dedner et al. [41]. Note that (at least formally) we recover the original formulation (3.1)-(3.3) by letting $\varepsilon \to 0$ in (3.4)-(3.6).
Regarding to the system (3.4)-(3.6) we use additional assumptions for the Lagrange multiplier ψ^ε.

Assumption 3.2.2 *For* $\varepsilon > 0$ *consider the initial value problem* (3.4)-(3.6).

(i) Assumption 3.2.1 holds.

(ii) $\psi_0^\varepsilon \equiv 0$.

To analyze (3.1)-(3.3) one can use pseudo-differential calculus. We have the following well-posedness result ([14], Chapter 2).

Theorem 3.2.1 *Suppose that Assumption 3.2.1 holds. Then there exists a unique weak solution u of* (3.1)-(3.2) *and we have* $u \in C([0,T];H^1(\mathbb{R}^d;\mathbb{R}^m))$. *In addition there exists a constant* $\mathscr{C} > 0$ *such that we have for* $t \in [0,T]$ *the estimate*

$$\|u(\cdot,t)\|_{H^1(\mathbb{R}^d;\mathbb{R}^m)} + \|u_t(\cdot,t)\|_{L^2(\mathbb{R}^d;\mathbb{R}^m)} \leq \mathscr{C}\left(\|u_0\|_{H^1(\mathbb{R}^d;\mathbb{R}^m)} + \int_0^t \|f(\cdot,s)\|_{H^1(\mathbb{R}^d;\mathbb{R}^m)} ds\right).$$

Moreover, if $u_0 \in C_0^\infty(\mathbb{R}^d;\mathbb{R}^m)$ *and* $f \in C_0^\infty([0,T]\times\mathbb{R}^d;\mathbb{R}^m)$ *then u is a classical solution and belongs to the space* $C_0^\infty([0,T]\times\mathbb{R}^d;\mathbb{R}^m)$.

The extended GLM formulation (3.4)-(3.6) is not symmetric. Therefore, we consider the change of variables $\varphi^\varepsilon := \psi^\varepsilon\sqrt{\varepsilon}$. In a blockmatrix structure (3.4)-(3.6) read

$$\frac{\partial}{\partial t}U^\varepsilon + \sum_{i=1}^d \frac{\partial}{\partial x_i}(A^{\varepsilon,i}U^\varepsilon) + BU^\varepsilon = F, \qquad U^\varepsilon(x,0) = U_0^\varepsilon(x) := \begin{pmatrix} u_0^\varepsilon(x) \\ 0 \end{pmatrix}, \tag{3.9}$$

with

$$U^\varepsilon := \begin{pmatrix} u^\varepsilon \\ \varphi^\varepsilon \end{pmatrix}; \quad A^{\varepsilon,i} := \begin{pmatrix} G^i & \frac{M_i^T}{\sqrt{\varepsilon}} \\ \frac{M_i}{\sqrt{\varepsilon}} & 0 \end{pmatrix}; \quad B := \begin{pmatrix} D & 0 \\ 0 & aI \end{pmatrix}; \quad F := \begin{pmatrix} f \\ 0 \end{pmatrix}.$$

In particular it is clear that we are again in the framework of symmetric systems and the extended formulation leads to a hyperbolic system.

Remark 3.2.1 *We note that from Assumption 3.2.2 we have* $\|A^{\varepsilon,i}\| = \mathscr{O}(\varepsilon^{-1/2})$, $\|B\| = \mathscr{O}(a)$ *and* $\|\mathrm{div}A\| = \mathscr{O}(1)$ *with* $\mathrm{div}A := \sum_{i=1}^d (A^{\varepsilon,i})_{x_i}$.

We want a generic estimate (like in Theorem 3.2.1) for the system (3.9), but it is not clear how the constant will depend on ε. However, an analogous estimate holds, i.e, we have the following result.

Lemma 3.2.1 *Let Assumption 3.2.2 be satisfied. There exists a constant* $\mathscr{C} > 0$, *independent of* ε, *such that for* $t \in [0,T]$ *we have*

$$\begin{aligned}\|u^\varepsilon(\cdot,t)\|^2_{H^1(\mathbb{R}^d;\mathbb{R}^m)} + \|\varphi^\varepsilon(\cdot,t)\|^2_{H^1(\mathbb{R}^d;\mathbb{R}^m)} + \|u_t^\varepsilon(\cdot,t)\|^2_{L^2(\mathbb{R}^d;\mathbb{R}^m)} + \|\varphi_t^\varepsilon(\cdot,t)\|^2_{L^2(\mathbb{R}^d;\mathbb{R}^m)} \\ \leq \mathscr{C}\left(\|u_0\|^2_{H^1(\mathbb{R}^d;\mathbb{R}^m)} + \int_0^t \|f(\cdot,s)\|^2_{H^1(\mathbb{R}^d;\mathbb{R}^m)} ds\right).\end{aligned}$$

Proof. We compute the following energy estimates

$$\frac{d}{dt}\left(\int_{\mathbb{R}^d} \frac{|u^\varepsilon|^2}{2} dx\right) + \int_{\mathbb{R}^d}\left(\sum_{i=1}^d u^\varepsilon \cdot G^i \frac{\partial u^\varepsilon}{\partial x_i} + u^\varepsilon \cdot \frac{M_i^T}{\sqrt{\varepsilon}} \frac{\partial \varphi^\varepsilon}{\partial x_i}\right) dx$$

$$+\int_{\mathbb{R}^d}(u^\varepsilon\cdot(D+\mathrm{div}G)u^\varepsilon-u^\varepsilon\cdot f)\,dx=0,$$
$$\frac{d}{dt}\left(\int_{\mathbb{R}^d}\frac{|\varphi^\varepsilon|^2}{2}\right)+\int_{\mathbb{R}^d}\left(\sum_{i=1}^d\varphi^\varepsilon\cdot\frac{M_i}{\sqrt{\varepsilon}}\frac{\partial u^\varepsilon}{\partial x_i}+a|\varphi^\varepsilon|^2\right)dx=0.$$

We see that thanks to Assumption 3.2.2 (symmetry of G^i) we have

$$\frac{\partial}{\partial x_i}\left((G^iu^\varepsilon)^Tu^\varepsilon\right)=2(u^\varepsilon)^TG^i\frac{\partial u^\varepsilon}{\partial x_i}+(u^\varepsilon)^T\frac{\partial G^i}{\partial x_i}u^\varepsilon.$$

Adding the above energy estimates and using the last expression we obtain

$$\frac{d}{dt}\int_{\mathbb{R}^d}\left(\frac{|u^\varepsilon|^2}{2}+\frac{|\varphi^\varepsilon|^2}{2}\right)dx+\int_{\mathbb{R}^d}\left(u^\varepsilon\cdot(D+\frac{1}{2}\mathrm{div}G)u^\varepsilon+a|\varphi^\varepsilon|^2\right)dx=\int_{\mathbb{R}^d}f\cdot u^\varepsilon\,dx.$$

Using Assumption 3.2.2 (properties of D,G) and applying Gronwall inequality we get finally

$$\int_{\mathbb{R}^d}\left(\frac{|u^\varepsilon|^2}{2}+\frac{|\varphi^\varepsilon|^2}{2}\right)dx\le e^{\mathscr{C}t}\left(\int_{\mathbb{R}^d}|u_0|^2\,dx+\int_0^t\int_{\mathbb{R}^d}|f|^2dxdt\right).$$

Reasoning as above we find analogous estimates for $\|U^\varepsilon_{x_i}\|^2_{L^2(\mathbb{R}^d;\mathbb{R}^{2m})}$ with $i=1,\dots,d$. In order to get estimates for $\|u^\varepsilon_t\|$ and $\|\varphi^\varepsilon_t\|$ we use (3.9), the bound for $\|U^\varepsilon\|^2_{L^2(\mathbb{R}^d;\mathbb{R}^{2m})}$ and $\|U^\varepsilon_{x_i}\|^2_{L^2(\mathbb{R}^d;\mathbb{R}^{2m})}$, Remark 3.2.1 and Assumption 3.2.1(ii). □

With Lemma 3.2.1 we can conclude as in Theorem 3.2.1 that the following theorem holds.

Theorem 3.2.2 *Suppose that Assumption 3.2.2 holds. Then there exists a unique weak solution U^ε of* (3.9) *with $U^\varepsilon\in C([0,T];H^1(\mathbb{R}^d;\mathbb{R}^{2m}))$. In addition there exists a constant $\mathscr{C}>0$ independent of ε such that*

$$\|u^\varepsilon(\cdot,t)\|_{H^1(\mathbb{R}^d;\mathbb{R}^m)}+\|\varphi^\varepsilon(\cdot,t)\|_{H^1(\mathbb{R}^d;\mathbb{R}^m)}+\|u^\varepsilon_t(\cdot,t)\|_{L^2(\mathbb{R}^d;\mathbb{R}^m)}+\|\varphi^\varepsilon_t(\cdot,t)\|_{L^2(\mathbb{R}^d;\mathbb{R}^m)}$$
$$\le\mathscr{C}\left(\|u^\varepsilon_0\|_{H^1(\mathbb{R}^d;\mathbb{R}^m)}+\int_0^t\|f(\cdot,s)\|_{H^1(\mathbb{R}^d;\mathbb{R}^m)}\,ds\right).$$

Moreover, if $U^\varepsilon_0\in C^\infty_0(\mathbb{R}^d;\mathbb{R}^{2m})$ and $F\in C^\infty_0([0,T]\times\mathbb{R}^d;\mathbb{R}^{2m})$ then U^ε is a classical solution and lies in the space $C^\infty_0([0,T]\times\mathbb{R}^d;\mathbb{R}^{2m})$.

Now we are in a position to estimate the error $\|u^\varepsilon-u\|$. The corresponding result in the special case of electrodynamics can be found in [88].

Proposition 3.2.1 *Let u be the weak solution of* (3.1)-(3.3) *and $U^\varepsilon=(u^\varepsilon,\varphi^\varepsilon)^T$ the weak solution of* (3.9). *Suposse that $u^\varepsilon_0=u_0$. Under Assumption 3.2.2, we have for all $t\in[0,T]$*

$$u^\varepsilon(\cdot,t)=u(\cdot,t),\quad a.e.,$$
$$\psi^\varepsilon(\cdot,t)=0,\quad a.e.$$

Proof. Defining $\overline{u} := u^\varepsilon - u$, a direct computation yields

$$\frac{d}{dt}\left(\int_{\mathbb{R}^d}\frac{|\overline{u}|^2}{2}dx\right)+\int_{\mathbb{R}^d}\left(\sum_{i=1}^d \overline{u}\cdot G^i\frac{\partial\overline{u}}{\partial x_i}+\overline{u}\cdot\frac{M_i^T}{\sqrt{\varepsilon}}\frac{\partial\varphi^\varepsilon}{\partial x_i}\right)dx+\int_{\mathbb{R}^d}(\overline{u}\cdot(D+\mathrm{div}G)\overline{u})\,dx=0,$$
$$\frac{d}{dt}\left(\int_{\mathbb{R}^d}\frac{|\varphi^\varepsilon|^2}{2}dx\right)+\int_{\mathbb{R}^d}\left(\sum_{i=1}^d \varphi^\varepsilon\frac{M_i}{\sqrt{\varepsilon}}\frac{\partial\overline{u}}{\partial x_i}+a|\varphi^\varepsilon|^2\right)dx=0.$$

Adding the last two equations and using Assumption 3.2.2 (symmetry of G^i) we get

$$\frac{d}{dt}\int_{\mathbb{R}^d}\left(\frac{|\overline{u}|^2}{2}+\frac{|\varphi^\varepsilon|^2}{2}\right)dx+\int_{\mathbb{R}^d}\left(\overline{u}^T(D+\frac{1}{2}\mathrm{div}G)\overline{u}+a|\varphi^\varepsilon|^2\right)dx=0,$$

which implies after applying Gronwall inequality and using that $\|\overline{u}_0\|_{L^2(\mathbb{R}^d;\mathbb{R}^m)}=0$ the estimate

$$\int_{\mathbb{R}^d}\left(|\overline{u}(\cdot,t)|^2+|\varphi^\varepsilon(\cdot,t)|^2\right)dx\le\mathscr{C}\int_{\mathbb{R}^d}|\varphi_0^\varepsilon|^2dx=0.$$

□

Proposition 3.2.1 shows the equivalence of solutions of the extended formulation (3.9) and the solution of the original problem (3.1)-(3.3) for $\varepsilon>0$. However it is not clear how ε can be chosen (asymptotically) in computations and whether the constraint (3.3) is satisfied in the limit.

3.3 Finite-Volume Discretization

We approximate the solution of (3.9) by a Finite-Volume scheme on unstructured meshes. This construction follows [59]. We begin with some standard generalities on Finite-Volume schemes.

Definition 3.3.1 *For some index set $I\subset\mathbb{N}$ let a family $\left\{K^i\right\}_{i\in I}$ of open non-empty sets be given. This family is called a triangulation if each element is a convex polyhedron and*

$$\cup_{i\in I}K^i=\mathbb{R}^d,\ K^i\cap K^j=\emptyset\ \forall i,j\in I\ i\neq j\ \text{and}\ h:=\sup_{i\in I}\left\{\mathrm{diam(K^i)}\right\}<\infty.$$

We denote the family $\left\{K^i\right\}_{i\in I}$ by $\mathscr{T}_h$ and introduce the following notations for $K\in\mathscr{T}_h$

$$\begin{aligned}|K|&:\text{ area of }K,\\ e\in\partial K&:\text{ an edge of }K\text{ with length }|e|,\\ n_{e,K}=(n_{e,K}^1,\dots,n_{e,K}^d)^T&:\text{ unit outward normal to the edge }e\text{ of }K,\\ K_e&:\text{ neighboring cell of }K\text{ with }\overline{K}\cap\overline{K}_e=e.\end{aligned}$$

For $N \in \mathbb{N}$, let $0 = t^1 < t^2 \ldots < t^N = T$ be a partition of the interval $[0,T]$. We denote $\Delta t^n = t^{n+1} - t^n$ for $n \in \mathscr{N} \cup \{N\}, \mathscr{N} = \{0,\ldots,N-1\}$.
For each $n \in \{0,\ldots,N\}, K \in \mathscr{T}_h$, and $e \in \partial K$ we define for $S : \mathbb{R}^d \times [0,T] \to \mathbb{R}^{2m}$

$$S_K^n := \frac{1}{\Delta t^n |K|} \int_{t^n}^{t^{n+1}} \int_K S(x,t)dxdt, \qquad S_e^n := \frac{1}{\Delta t^n |e|} \int_{t^n}^{t^{n+1}} \int_e S(\zeta,t)d\zeta dt,$$
$$S_K(t) := \frac{1}{|K|} \int_K S(x,t)dx, \qquad S_e(t) := \frac{1}{|e|} \int_e S(\zeta,t)\,d\zeta.$$

For the sake of clarity we summarize all the assumptions on the mesh.

Assumption 3.3.1 *Let $\mathscr{T}_h$ be a triangulation (Def. 3.3.1) of $\mathbb{R}^d$. There exist constants $\eta > 0$ and $\nu > 0$ such that*

$$|K| \geq \eta h^d, \quad |\partial K| \leq \nu h^{d-1} \qquad (\forall K \in \mathscr{T}_h, \forall e \in \mathscr{E}(K)).$$

Moreover we assume that the time step Δt is constant, i.e., $\Delta t^n = \Delta t$.

Definition 3.3.2 *The GLM-FV approximation $U_h^\varepsilon := (u_h^\varepsilon, \varphi_h^\varepsilon)^T : \mathbb{R}^d \times [0,T) \to \mathbb{R}^{2m}$ of* (3.1)-(3.3) *with initial data $U_0^\varepsilon = (u_0^\varepsilon, \varphi_0^\varepsilon)^T$ is given by*

$$U_h^\varepsilon(x,t) = V_k^{\varepsilon,n} \qquad \text{for } (x,t) \in K \times [t^n, t^{n+1}).$$

The vectors $V_K^{\varepsilon,n} \in \mathbb{R}^{2m}$ are given for $n = 0$ and $K \in \mathscr{T}_h$ by

$$V_K^{\varepsilon,0} = \frac{1}{|K|} \int_K U_0^\varepsilon(x)dx,$$

and iteratively for $n \in \mathscr{N}$ by

$$V_K^{\varepsilon,n+1} = V_K^{\varepsilon,n} - \frac{\Delta t}{|K|} \sum_{e \in \partial K} |e| g_{e,K}^n (V_K^{\varepsilon,n}, V_{K_e}^{\varepsilon,n}) - \Delta t B_K^n V_K^{\varepsilon,n} + \Delta t F_K^n. \tag{3.10}$$

The numerical flux $g_{e,K}^n : \mathbb{R}^{2m} \times \mathbb{R}^{2m} \to \mathbb{R}^{2m}$ is defined for $K \in \mathscr{T}_h$ and $e \in \partial K$ by

$$g_{e,K}^n(U,V) = -C_{e,K}^{\varepsilon,n} V + D_{e,K}^{\varepsilon,n} U, \tag{3.11}$$

with

$$A_{e,K}^{\varepsilon,n} := \sum_{i=1}^d n_{e,K}^i (A^{\varepsilon,i})_e^n, \quad C_{e,K}^{\varepsilon,n} := -O^T \Lambda^- O, \quad D_{e,K}^{\varepsilon,n} := O^T \Lambda^+ O, \tag{3.12}$$

and

$$A_{e,K}^{\varepsilon,n} = O^T \Lambda^+ O + O^T \Lambda^- O, \tag{3.13}$$

where $\Lambda^+(\Lambda^-)$ is a diagonal matrix which entries are the positive (negative) eigenvalues of $A_{e,K}^{\varepsilon,n}$.

As long as a, ε are fixed in Definition 3.3.2 the GLM-FV method gives an approximate for the weak solution of (3.1)-(3.3). However, we will choose $\varepsilon = \varepsilon(h)$ with $\varepsilon(h) \to 0$ as $h \to 0$, so that the GLM-FV method is supposed to fulfill the constraint (3.3) whenever the classical FVM do not. This is the problem to solve. The crucial question here is how to determine $\varepsilon(h)$ to get (an optimal order of) convergence and to fulfill the side condition (3.3).

Remark 3.3.1 *(i) Note that thanks to the hyperbolicity of the formulation the decomposition* (3.13) *makes sense.*

(ii) The Definition 3.3.2 leads to a consistent upwind numerical scheme. Note that we have the symmetric relation $C_{e,K}^{\varepsilon,n} = D_{e,K_e}^{\varepsilon,n}$. This leads to $g_{e,K}^n(U,V) = -g_{e,K_e}^n(V,U)$ for $U,V \in \mathbb{R}^{2m}$ and ensures that the scheme is conservative.

(iii) Using (ii) the iteration (3.10) *can be written as*

$$V_K^{\varepsilon,n+1} = V_K^{\varepsilon,n} - \frac{\Delta t}{|K|} \sum_{e \in \partial K} |e| C_{e,K}^{\varepsilon,n} (V_K^{\varepsilon,n} - V_{K_e}^{\varepsilon,n}) - \Delta t \left(B_K^n V_K^{\varepsilon,n} + \mathrm{div} A_K^n V_K^{\varepsilon,n} - F_K^n \right). \tag{3.14}$$

(iv) We note that since $\|C_{e,K}^{\varepsilon,n}\|$ is a function of $A^{\varepsilon,i}$ we have that $\|C_{e,K}^{\varepsilon,n}\| = \mathcal{O}\left(\varepsilon^{-1/2}\right)$ and $\|B_K^n\| = \mathcal{O}(a)$.

3.4 Convergence of the GLM-FV scheme

Our main goal in this section is to determine how the incorporation of $\varepsilon = \varepsilon(h)$ affects the rate of convergence of the component u_h^ε of U_h^ε given by Definition 3.3.2 to the solution u of (3.1)-(3.3) as $h \to 0$. To do this we carefully track the parameter ε in the constants that appear in the finite volume error analysis in [112] (see [59]). We assume throughout this section that Assumptions 3.2.1-3.3.1 hold.

3.4.1 Stability results

We start with a result that can be seen as a local stability lemma.

Lemma 3.4.1 *Under the CFL condition*

$$\sup_{K \in \mathcal{T}_h, e \in \partial K} \frac{\Delta t |\partial K| \|C_{e,K}^{\varepsilon,n}\|}{|K|} < 1 - \delta, \qquad \delta \in (0,1) \tag{3.15}$$

the solution U_h^ε generated by the GLM-FV method satisfies

$$
\begin{aligned}
&|V_K^{\varepsilon,n+1}|^2 - |V_K^{\varepsilon,n}|^2 + 2\Delta t (V_K^{\varepsilon,n+1})^T [B_K^n V_K^{\varepsilon,n} + \mathrm{div} A_K^n V_K^{\varepsilon,n} - F_K^n] \\
&+ \frac{\Delta t}{|K|} \sum_{e\in\partial K} |e| \big((V_K^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n} V_K^{\varepsilon,n} - (V_{K_e}^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n} V_{K_e}^{\varepsilon,n} \big) \\
&\qquad \le -\delta \frac{\Delta t}{|K|} \sum_{e\in\partial K} (V_{K_e}^{\varepsilon,n} - V_K^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n} (V_{K_e}^{\varepsilon,n} - V_K^{\varepsilon,n}) |e|. \qquad (3.16)
\end{aligned}
$$

Proof. We represent $V_K^{\varepsilon,n+1}$ as the convex decomposition $V_K^{\varepsilon,n+1} = (\sum_{e\in\partial K} |e| V_{e,K}^{\varepsilon,n+1})/|\partial K|$. Thereby we used (3.14) and

$$
V_{e,K}^{\varepsilon,n+1} := V_K^{\varepsilon,n} - \frac{\Delta t |\partial K|}{|K|} C_{e,K}^{\varepsilon,n} (V_K^{\varepsilon,n} - V_{K_e}^{\varepsilon,n}) - \Delta t ((B_K^n V_K^{\varepsilon,n} + \mathrm{div} A_K^n V_K^{\varepsilon,n} - F_K^n).
$$

We define also

$$
W_{e,K}^{\varepsilon,n+1} := V_K^{\varepsilon,n} - \frac{\Delta t |\partial K|}{|K|} C_{e,K}^{\varepsilon,n} (V_K^{\varepsilon,n} - V_{K_e}^{\varepsilon,n}). \qquad (3.17)
$$

Scalar multiplication of $W_{e,K}^{\varepsilon,n+1}$ with $V_K^{\varepsilon,n}$ and the symmetry of $C_{e,K}^{\varepsilon,n}$ gives

$$
\frac{1}{2}|W_{e,K}^{\varepsilon,n+1}|^2 - \frac{1}{2}|V_K^{\varepsilon,n}|^2 = -\frac{\Delta t |\partial K|}{2|K|} \big((V_K^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n} V_K^{\varepsilon,n} - (V_{K_e}^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n} V_{K_e}^{\varepsilon,n} \big) + Q, \qquad (3.18)
$$

with

$$
Q = \frac{1}{2}|W_{e,K}^{\varepsilon,n+1} - V_K^{\varepsilon,n}|^2 - \frac{\Delta t |\partial K|}{2|K|} \big((V_{K_e}^{\varepsilon,n} - V_K^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n} (V_{K_e}^{\varepsilon,n} - V_K^{\varepsilon,n}) \big).
$$

A straightforward calculation shows that

$$
Q = -\frac{\Delta t |\partial K|}{2|K|} \left((V_{K_e}^{\varepsilon,n} - V_K^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n} \left[I - \frac{\Delta t |\partial K|}{|K|} C_{e,K}^{\varepsilon,n} \right] (V_{K_e}^{\varepsilon,n} - V_K^{\varepsilon,n}) \right).
$$

Here I is the unit matrix in $\mathbb{R}^{2m\times 2m}$. Using the CFL condition (3.15) we get

$$
\begin{aligned}
Q &\le \frac{\Delta t |\partial K|}{2|K|} (V_{K_e}^{\varepsilon,n} - V_K^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n} (V_{K_e}^{\varepsilon,n} - V_K^{\varepsilon,n}) \\
&\quad + (1-\delta) \frac{\Delta t |\partial K|}{2|K|} (V_{K_e}^{\varepsilon,n} - V_K^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n} (V_{K_e}^{\varepsilon,n} - V_K^{\varepsilon,n}) \\
&= -\delta \frac{\Delta t |\partial K|}{2|K|} \big((V_{K_e}^{\varepsilon,n} - V_K^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n} (V_{K_e}^{\varepsilon,n} - V_K^{\varepsilon,n}) \big). \qquad (3.19)
\end{aligned}
$$

From the convex decomposition of $V_K^{e,n+1}$ and (3.17) we also see

$$
V_K^{\varepsilon,n+1} = \sum_{e\in\partial K} \frac{|e|}{|\partial K|} \left(W_{e,K}^{\varepsilon,n+1} - \Delta t (B_K^n V_K^{\varepsilon,n} + \mathrm{div} A_K^n V_K^{\varepsilon,n} + \Delta t F_K^n \right).
$$

Scalar multiplication of the last expression with $V_K^{\varepsilon,n+1}$ yields

$$\begin{aligned}\frac{1}{2}|V_K^{\varepsilon,n+1}|^2 &= -\Delta t (V_K^{\varepsilon,n+1})^T \left(B_K^n V_K^{\varepsilon,n} + \mathrm{div} A_K^n V_K^{\varepsilon,n}\right) + \Delta t (V_K^{\varepsilon,n+1})^T F_K^n \\ &\quad + \frac{1}{2}\sum_{e\in\partial K}\frac{|e|}{|\partial K|}\left(|W_{e,K}^{\varepsilon,n+1}|^2 - |W_{e,K}^{\varepsilon,n+1} - V_K^{\varepsilon,n+1}|^2\right) \\ &\leq -\Delta t (V_K^{\varepsilon,n+1})^T \left(B_K^n V_K^{\varepsilon,n} + \mathrm{div} A_K^n V_K^{\varepsilon,n} - F_K^n\right) + \frac{1}{2}\sum_{e\in\partial K}\frac{|e|}{|\partial K|}|W_{e,K}^{\varepsilon,n+1}|^2.\end{aligned}$$

Replacing the expression for $|W_{e,K}^{\varepsilon,n+1}|^2$ in (3.18) and using (3.19) we finally obtain

$$\begin{aligned}\frac{1}{2}|V_K^{\varepsilon,n+1}|^2 &\leq \frac{1}{2}|V_K^{\varepsilon,n}|^2 - \Delta t (V_K^{\varepsilon,n+1})^T \left(B_K^n V_K^{\varepsilon,n} + \mathrm{div} A_K^n V_K^{\varepsilon,n}\right) + \Delta t (V_K^{\varepsilon,n+1})^T F_K^n \\ &\quad - \frac{\Delta t}{2|K|}\sum_{e\in\partial K}|e|\left((V_K^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n} V_K^{\varepsilon,n} - (V_{K_e}^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n} V_{K_e}^{\varepsilon,n}\right) \\ &\quad - \delta\frac{\Delta t}{2|K|}\sum_{e\in\partial K}|e|(V_{K_e}^{\varepsilon,n} - V_K^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n}(V_{K_e}^{\varepsilon,n} - V_K^{\varepsilon,n}).\end{aligned}$$

□

With Lemma 3.4.1 we can now prove a global L^2-stability result as discrete counterpart to Lemma 3.2.1.

Proposition 3.4.1 *Assume that the CFL condition* (3.15) *is satisfied for a given* $\delta \in (0,1)$. *Then, for* $h \leq \mu/2\gamma$ *and* $\varepsilon \leq 1$, *the GLM-FV approximation* U_h^ε *satisfies for all* $0 \leq t \leq T$

$$\|U_h^\varepsilon(\cdot,t)\|_{L^2(\mathbb{R}^d;\mathbb{R}^{2m})} \leq \mathscr{C}\left(\|U_0^\varepsilon\|_{L^2(\mathbb{R}^d;\mathbb{R}^{2m})} + \|F\|_{L^2(0,T;L^2(\mathbb{R}^d;\mathbb{R}^{2m}))}\right). \tag{3.20}$$

Here $\gamma := 1 + \|B\| + \|\mathrm{div} A\|$ *and* $\mu := \sqrt{\varepsilon}\,\sup_{K\in\mathscr{T}_h, e\in\partial K}\|C_{e,K}^{\varepsilon,n}\|$. *The constant* $\mathscr{C}$ *depend on the data but not on* ε. *Moreover, the discrete space derivatives of* U_h^ε *satisfy the following weak estimate:*

$$\begin{aligned}&\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\sum_{e\in\partial K}(V_{K_e}^{\varepsilon,n} - V_K^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n}(V_{K_e}^{\varepsilon,n} - V_K^{\varepsilon,n})|e|\Delta t \\ &\qquad \leq \frac{\mathscr{C}}{\delta}\left(\|U_0^\varepsilon\|_{L^2(\mathbb{R}^d;\mathbb{R}^{2m})} + \|F\|_{L^2(0,T;L^2(\mathbb{R}^d;\mathbb{R}^{2m}))}\right)^2.\end{aligned} \tag{3.21}$$

Again, the constant $\mathscr{C} > 0$ *depends only on the data of the problem.*

Proof. First, by adding the nil quantity

$$-\Delta t\sum_{e\in\partial K}(V_K^{\varepsilon,n})^T A_{e,K}^{\varepsilon,n} V_K^{\varepsilon,n}|e| + |K|\Delta t (V_K^{\varepsilon,n})^T \mathrm{div} A_K^n V_K^{\varepsilon,n}$$

to the R.H.S of (3.16), we obtain the following form of the local energy estimate:

$$\begin{aligned}\frac{|K||V_K^{\varepsilon,n+1}|^2}{2} \leq \frac{|K||V_K^{\varepsilon,n}|^2}{2} &- \Delta t|K|(V_K^{\varepsilon,n+1})^T\left(\left(B_K^n V_K^{\varepsilon,n} + \mathrm{div}A_K^n V_K^{\varepsilon,n}\right) - F_K^n\right)\\ &- \frac{\Delta t}{2}\sum_{e\in\partial K}|e|\left((V_K^{\varepsilon,n})^T D_{e,K}^n V_K^{\varepsilon,n} - (V_{K_e}^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n} V_{K_e}^{\varepsilon,n}\right)\\ &+ \frac{\Delta t|K|}{2}(V_K^{\varepsilon,n})^T\mathrm{div}A_K^n V_K^{\varepsilon,n} - \frac{\delta}{2}\Delta t\sum_{e\in\partial K}|e|(V_{K_e}^{\varepsilon,n}-V_K^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n}(V_{K_e}^{\varepsilon,n}-V_K^{\varepsilon,n}).\end{aligned} \tag{3.22}$$

Summing (3.22) over all volumes, we obtain

$$\begin{aligned}&\sum_{K\in\mathscr{T}_h}|K||V_K^{\varepsilon,n+1}|^2 + \delta\Delta t\sum_{K\in\mathscr{T}_h}\sum_{e\in\partial K}|e|(V_{K_e}^{\varepsilon,n}-V_K^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n}(V_{K_e}^{\varepsilon,n}-V_K^{\varepsilon,n})\\ &\leq \sum_{K\in\mathscr{T}_h}|K|\left(|V_K^{\varepsilon,n}|^2 + \Delta t(V_K^{\varepsilon,n})^T\mathrm{div}A_K^n V_K^{\varepsilon,n}\right)\\ &\quad - 2\Delta t\sum_{K\in\mathscr{T}_h}|K|(V_K^{\varepsilon,n+1})^T\left(\left(B_K^n V_K^{\varepsilon,n} + \mathrm{div}A_K^n V_K^{\varepsilon,n}\right) - F_K^n\right)\\ &\leq (1+\kappa\Delta t)\|U_h^\varepsilon(t^n,\cdot)\|^2_{L^2(\mathbb{R}^d;\mathbb{R}^{2m})} + 2\Delta t\|U_h^\varepsilon(t^{n+1},\cdot)\|_{L^2(\mathbb{R}^d;\mathbb{R}^{2m})}\Big(\sum_{K\in\mathscr{T}_h}|K||F_K^n|^2\Big)^{1/2}\\ &\quad + 2\Delta t(\kappa+\beta)\|U_h^\varepsilon(t^{n+1},\cdot)\|_{L^2(\mathbb{R}^d;\mathbb{R}^{2m})}\|U_h^\varepsilon(t^n,\cdot)\|_{L^2(\mathbb{R}^d;\mathbb{R}^{2m})}\end{aligned}$$

Using that $2ab \leq a^2+b^2$, we get

$$\begin{aligned}&\|U_h^\varepsilon(t^{n+1},\cdot)\|^2_{L^2(\mathbb{R}^d;\mathbb{R}^{2m})}(1-\Delta t\gamma) + \delta\Delta t\sum_{K\in\mathscr{T}_h}\sum_{e\in\partial K}|e|(V_{K_e}^{\varepsilon,n}-V_K^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n}(V_{K_e}^{\varepsilon,n}-V_K^{\varepsilon,n})\\ &\quad\leq (1+2\gamma\Delta t)\,\|U_h^\varepsilon(t^n,\cdot)\|^2_{L^2(\mathbb{R}^d;\mathbb{R}^{2m})} + \Delta t\sum_{K\in\mathscr{T}_h}|K||F_K^n|^2.\end{aligned} \tag{3.23}$$

Using the hypothesis over h and ε we find $\gamma\Delta t \leq 1/2$ (we remark that γ is independent of ε). Hence we have $1 \leq (1-\Delta t\gamma)^{-1} \leq 1+2\gamma\Delta t$ and we deduce from (3.23) that:

$$\begin{aligned}&\|U_h^\varepsilon(t^{n+1},\cdot)\|^2_{L^2(\mathbb{R}^d;\mathbb{R}^{2m})} + \delta\Delta t\sum_{K\in\mathscr{T}_h}\sum_{e\in\partial K}|e|(V_{K_e}^{\varepsilon,n}-V_K^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n}(V_{K_e}^{\varepsilon,n}-V_K^{\varepsilon,n})\\ &\quad\leq (1+2\gamma\Delta t)^2\|U_h^\varepsilon(t^n,\cdot)\|^2_{L^2(\mathbb{R}^d;\mathbb{R}^{2m})} + \Delta t(1+2\gamma\Delta t)\sum_{K\in\mathscr{T}_h}|K||F_K^n|^2\\ &\quad\leq (1+6\gamma\Delta t)\|U_h^\varepsilon(t^n,\cdot)\|^2_{L^2(\mathbb{R}^d;\mathbb{R}^{2m})} + 2\Delta t\sum_{K\in\mathscr{T}_h}|K||F_K^n|^2.\end{aligned}$$

Iterating the last inequality, we get for any $t\in[0,T)$:

$$\|U_h^\varepsilon(t,\cdot)\|^2_{L^2(\mathbb{R}^d;\mathbb{R}^{2m})} + \gamma\sum_{n=0}^{N}\sum_{K\in\mathscr{T}_h}\Delta t(1+\mathscr{C}\Delta t)^{N-n}\sum_{e\in\partial K}|e|(V_{K_e}^{\varepsilon,n}-V_K^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n}(V_{K_e}^{\varepsilon,n}-V_K^{\varepsilon,n})$$

$$\leq \left(1+\mathscr{C}\frac{T}{N}\right)^N \|U_0^\varepsilon\|^2_{L^2(\mathbb{R}^d;\mathbb{R}^{2m})} + 2\sum_{n=0}^{N}\Delta t(1+\mathscr{C}\Delta t)^{N-n}\sum_{K\in\mathscr{T}_h}|K||F_K^n|^2,$$

with N being the integer part of $T/\Delta t$ and $\mathscr{C} := 6\gamma$. We finally obtain that $\forall t \in [0,T)$:

$$\|U_h^\varepsilon(t,\cdot)\|^2_{L^2(\mathbb{R}^d;\mathbb{R}^{2m})} + \gamma\sum_{n=0}^{N}\sum_{K\in\mathscr{T}_h}\sum_{e\in\partial K}\Delta t|e|(V_{K_e}^{\varepsilon,n}-V_K^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n}(V_{K_e}^{\varepsilon,n}-V_K^{\varepsilon,n})$$
$$\leq \exp(\mathscr{C}T)\left(\|U_0^\varepsilon\|^2_{L^2(\mathbb{R}^d;\mathbb{R}^{2m})} + 2\|F\|^2_{L^2(0,T;L^2(\mathbb{R}^d;\mathbb{R}^{2m}))}\right).$$

□

3.4.2 A Comparison Result

In the next step we consider the difference between the exact solution U^ε of (3.9) and some function $V \in L^2(\mathbb{R}^d\times(0,T);\mathbb{R}^{2m})$ in terms of residual errors. To achieve the main result we will put $V = U_h^\varepsilon$. Following the book of Kröner [68] we introduce two useful measures.

Definition 3.4.1 *Let $V \in L^2(\mathbb{R}^d\times(0,T);\mathbb{R}^{2m})$ be given. The consistency measure $\mu_V : C^1([0,T];L^2(\mathbb{R}^d;\mathbb{R}^{2m}))\cap C([0,T];H^1(\mathbb{R}^d;\mathbb{R}^{2m})) \to \mathbb{R}$ and the dissipation measure $\nu_V : C_b^{0,1}(\mathbb{R}^d\times[0,T]) \to \mathbb{R}$ are defined for $\pi \in C^1([0,T];L^2(\mathbb{R}^d;\mathbb{R}^{2m}))\cap C([0,T];H^1(\mathbb{R}^d;\mathbb{R}^{2m}))$ and $\omega \in C_b^{0,1}(\mathbb{R}^d\times[0,T])$ by*

$$<\mu_V,\pi> := -\int_0^T\int_{\mathbb{R}^d}\left(V^T\partial_t\pi + \sum_{i=1}^d V^T A^{\varepsilon,i}\partial_i\pi\right)dxdt$$
$$+\int_0^T\int_{\mathbb{R}^d}(V^TB^T - F^T)\pi\,dxdt - \int_{\mathbb{R}^d}(U_0^\varepsilon)^T\pi(x,0)\,dx,$$
$$<\nu_V,\omega> := -\int_0^T\int_{\mathbb{R}^d}\left(|V|^2\partial_t\omega + \sum_{i=1}^d V^TA^{\varepsilon,i}V\partial_i\omega\right)dxdt$$
$$+\int_0^T\int_{\mathbb{R}^d}\left(V^T(\mathrm{div}A + B + B^T)V - 2F^TV\right)\omega\,dxdt - \int_{\mathbb{R}^d}|U_0^\varepsilon|^2\omega(x,0)\,dx.$$

We cite Proposition 2.4 of [59]

Proposition 3.4.2 *Let $U_0^\varepsilon \in C_0^\infty(\mathbb{R}^d;\mathbb{R}^{2m})$, $F \in C_0^\infty([0,T]\times\mathbb{R}^d;\mathbb{R}^{2m})$, $V \in L^2((0,T)\times\mathbb{R}^d;\mathbb{R}^{2m})$ be and define $\alpha := \|B+B^T+\mathrm{div}A\|$. Then we have*

$$\int_0^T\int_{\mathbb{R}^d}\exp(-\alpha t)|U^\varepsilon - V|^2dxdt \leq <\nu_V,\theta> -2<\mu_V,\theta U^\varepsilon> \tag{3.24}$$

where U^ε is the exact solution of (3.9) *and $\theta : [0,T]\to\mathbb{R}$ is defined by:*
$\theta(t) = \exp(-\alpha t)(T-t)$ with $t\in[0,T]$.

3.4.3 The Error Estimate

To prove Theorem 3.4.1 we will apply the estimate of Proposition 3.4.2 to the approximate solution U_h^ε generated by the GLM-FV method.

Theorem 3.4.1 *Under the CFL-condition* (3.15)*, the GLM-FV approximation U_h^ε converges towards the solution U^ε of* (3.9) *in $L^2([0,T]\times\mathbb{R}^d;\mathbb{R}^{2m})$. Moreover U_h^ε satisfies the following error estimate*

$$\|U^\varepsilon - U_h^\varepsilon\|_{L^2([0,T]\times\mathbb{R}^d;\mathbb{R}^{2m})} \leq \mathscr{C}\varepsilon^{-1/4}h^{1/2}. \tag{3.25}$$

In (3.25) *$\mathscr{C}$ is a positive constant that depends only on δ, U_0, F and T but not on ε.*

In order to prove Theorem 3.4.1 we use a proposition of [112] and a lemma of [59]. We recall that $U_h^\varepsilon \in L^2([0,T]\times\mathbb{R}^d;\mathbb{R}^{2m})$ thanks to Proposition 3.4.1.

Lemma 3.4.2 (Proposition 5.1 in Vila and Villedieu (2003)) *If we choose $V = U_h^\varepsilon$ in Definition 3.4.1 we get*

$$< \mu_{U_h^\varepsilon}, \pi >= \sum_{l=1}^{7} \mathscr{R}_h^l(\pi), \qquad < \nu_{U_h^\varepsilon}, \omega > \leq \sum_{l=1}^{7} \mathscr{E}_h^l(\omega) - \delta Q_h^\varepsilon(\omega),$$

where $\omega : \mathbb{R}^d \times [0,T] \to [0,\infty)$ and $\pi : \mathbb{R}^d \times [0,T] \to \mathbb{R}^{2m}$ are smooth functions with compact support in x. Here we used

$$\mathscr{R}_h^1(\pi) = \sum_{n\in\mathscr{N}} \sum_{K\in\mathscr{T}_h} |K|(V_K^{\varepsilon,n+1} - V_K^{\varepsilon,n})^T(\pi_K(t^{n+1}) - \pi_K^n),$$

$$\mathscr{R}_h^2(\pi) = \sum_{n\in\mathscr{N}} \sum_{K\in\mathscr{T}_h} \sum_{e\in\partial K} \Delta t|e|(V_K^{\varepsilon,n} - V_{K_e}^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n}(\pi_e^n - \pi_K^n),$$

$$\mathscr{R}_h^3(\pi) = \int_{\mathbb{R}^d} (U_h^\varepsilon(x,0) - U_0^\varepsilon(x))^T \pi(x,0)dx,$$

$$\mathscr{R}_h^4(\pi) = \sum_{n\in\mathscr{N}} \sum_{K\in\mathscr{T}_h} (V_K^{\varepsilon,n})^T \left[\sum_{e\in\partial K} \Delta t|e|A_{e,K}^n \pi_e^n - \int_{t^n}^{t^{n+1}} \int_K \sum_{i=1}^{d} \frac{\partial}{\partial x_i}(A^{\varepsilon,i}\pi)dxdt \right],$$

$$\mathscr{R}_h^5(\pi) = \sum_{n\in\mathscr{N}} \sum_{K\in\mathscr{T}_h} \Delta t|K|(V_K^{\varepsilon,n})^T \left[\frac{1}{\Delta t|K|} \int_{t^n}^{t^{n+1}} \int_K (\mathrm{div}A)\pi dtdx - (\mathrm{div}A)_K^n \pi_K^n \right],$$

$$\mathscr{R}_h^6(\pi) = -\sum_{n\in\mathscr{N}} \sum_{K\in\mathscr{T}_h} \Delta t|K|(V_K^{\varepsilon,n})^T \left[\frac{1}{\Delta t|K|} \int_{t^n}^{t^{n+1}} \int_K B^T \pi dtdx - (B_K^n)^T \pi_K^n \right],$$

$$\mathscr{R}_h^7(\pi) = -\sum_{n\in\mathscr{N}} \sum_{K\in\mathscr{T}_h} \Delta t|K| \left[\frac{1}{\Delta t|K|} \int_{t^n}^{t^{n+1}} \int_K F^T \pi dtdx - (F_K^n)^T \pi_K^n \right],$$

$$\mathscr{E}_h^1(\omega) = \sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h} |K|(|V_K^{\varepsilon,n+1}|^2 - |V_K^{\varepsilon,n}|^2)(\omega_K(t^{n+1}) - \omega_K^n),$$

$$\mathscr{E}_h^2(\omega) = \sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\sum_{e\in\partial K} \Delta t|e|(V_K^{\varepsilon,n} - V_{K_e}^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n}(V_K^{\varepsilon,n} + V_{K_e}^{\varepsilon,n})(\omega_e^n - \omega_K^n),$$

$$\mathscr{E}_h^3(\omega) = \int_{\mathbb{R}^d}(|U_h^\varepsilon(x,0)|^2 - |U_0^\varepsilon(x)|^2)\omega(x,0)dx,$$

$$\mathscr{E}_h^4(\omega) = \sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}(V_K^{\varepsilon,n})^T\left[\sum_{e\in\partial K}\Delta t|e|A_{e,K}^n\omega_e^n - \int_{t^n}^{t^{n+1}}\int_K\sum_{i=1}^d\frac{\partial}{\partial x_i}(A^{\varepsilon,i}\omega)dxdt\right]V_K^{\varepsilon,n},$$

$$\begin{aligned}\mathscr{E}_h^5(\omega) = \sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}(V_K^{\varepsilon,n})^T\Big[&\int_{t^n}^{t^{n+1}}\int_K\left(B + B^T + 2(\mathrm{div}A)\right)\omega\,dxdt\\ &- |K|\Delta t(B_K^n + (B_K^n)^T + 2(\mathrm{div}A)_K^n)\omega_K^n\Big]V_K^{\varepsilon,n},\end{aligned}$$

$$\mathscr{E}_h^6(\omega) = \sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h} 2\Delta t|K|(V_K^{\varepsilon,n})^T\left[\frac{1}{\Delta t|K|}\int_{t^n}^{t^{n+1}}\int_K F\,\omega\,dxdt - (F_K^n)\omega_K^n\right],$$

$$\mathscr{E}_h^7(\omega) = -\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h} 2\Delta t|K|(V_K^{\varepsilon,n+1} - V_K^{\varepsilon,n})^T\left((B_K^n + (\mathrm{divA})_{\mathrm{K}}^{\mathrm{n}})V_K^{\varepsilon,n} - F_K^n\right)\omega_K^n,$$

and

$$Q_h^\varepsilon(\omega) = \sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\sum_{e\in\partial K}\Delta t|e|(V_{K_e}^{\varepsilon,n} - V_K^n)^T C_{e,K}^{\varepsilon,n}(V_{K_e}^{\varepsilon,n} - V_K^{\varepsilon,n})\omega_K^n.$$

Lemma 3.4.3 (Lemma 4.3 in Jovanovic and Rohde (2004)) *Let $z \in H^1(K;\mathbb{R}^{2m})$, then there exists a constant $\mathscr{C} > 0$ such that for $K \in \mathscr{T}_h$*

$$\int_K |z - z_k|^2 dx \le \mathscr{C}h^2\int_K |Dz|^2 dx, \quad \int_e |z - z_k|^2 d\zeta \le \mathscr{C}h\int_K |Dz|^2 dx \qquad (e \in \partial K).$$

Here $\mathscr{C}$ depends only on d and m.

We conclude with the proof of Theorem 3.4.1.

Proof. We suppose first that $U_0^\varepsilon \in C_0^\infty(\mathbb{R}^d;\mathbb{R}^{2m})$, $F \in C_0^\infty([0,T]\times\mathbb{R}^d;\mathbb{R}^{2m})$. Accordingly to Theorem 3.2.2 we have $U^\varepsilon \in C_0^\infty([0,T]\times\mathbb{R}^d;\mathbb{R}^{2m})$. Applying Proposition 3.4.2 and Lemma 3.4.2 with $\omega = \theta$ and $\pi = \theta U^\varepsilon$ we just have to estimate

$$\sum_{l=1}^7\left[\mathscr{E}_h^l(\theta) - 2\mathscr{R}_h^l(\theta U^\varepsilon)\right] - \delta Q_h^\varepsilon(\theta).$$

We first consider two terms that will appear many times in our calculation, namely

$$\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\theta^n|K||V_K^{\varepsilon,n+1} - V_K^{\varepsilon,n}|^2, \qquad \sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}|K||V_K^{\varepsilon,n+1} - V_K^{\varepsilon,n}|^2.$$

From (3.14) we obtain

$$|V_K^{\varepsilon,n+1}-V_K^{\varepsilon,n}|^2 \leq \mathscr{C}\Delta t^2 |B_K^n V_K^{\varepsilon,n} + \mathrm{div} A_K^n V_K^{\varepsilon,n} - F_K^n|^2 + \mathscr{C}\frac{\Delta t^2}{|K|^2}\sum_{e\in\partial K}|e|^2 |C_{e,K}^{\varepsilon,n}(V_K^{\varepsilon,n}-V_{K_e}^{\varepsilon,n})|^2. \tag{3.26}$$

Multiplying (3.26) by $|K|\theta^n$, summing over all the elements, using the CFL condition (3.15) and the stability result (Proposition 3.4.1) we get that

$$\begin{aligned}
&\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\theta^n|K||V_K^{\varepsilon,n+1}-V_K^{\varepsilon,n}|^2\\
\leq&\mathscr{C}\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\theta^n(\Delta t)^2|K|\left(\|B_K^n\|^2|V_K^{\varepsilon,n}|^2+\|\mathrm{div}A_K^n\|^2|V_K^{\varepsilon,n}|^2+|F_K^n|^2\right)\\
&+\mathscr{C}\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\theta^n\frac{(\Delta t)^2}{|K|}\sum_{e\in\partial K}|e|^2\|C_{e,K}^{\varepsilon,n}\|(V_K^{\varepsilon,n}-V_{K_e}^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n}(V_K^{\varepsilon,n}-V_{K_e}^{\varepsilon,n})\\
\leq&\mathscr{C}Q_h^\varepsilon+\mathscr{C}\Delta t\left(\|U_0^\varepsilon\|_{L^2(\mathbb{R}^d;\mathbb{R}^{2m})}+\|F\|_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))}\right)^2.
\end{aligned}\tag{3.27}$$

In an analogous way we find that

$$\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}|K||V_K^{\varepsilon,n+1}-V_K^{\varepsilon,n}|^2\leq\mathscr{C}\left(\|U_0^\varepsilon\|_{L^2(\mathbb{R}^d;\mathbb{R}^{2m})}+\|F\|_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))}\right)^2. \tag{3.28}$$

In both cases $\mathscr{C}$ is independent of ε.

Term $\mathscr{E}_h^1-2\mathscr{R}_h^1$: We define

$$\begin{aligned}
\mathscr{R}_h^{1,a}(\theta U^\varepsilon)&:=\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}|K|(V_K^{\varepsilon,n+1}-V_K^{\varepsilon,n})^T U_K^\varepsilon(t^{n+1})(\theta(t^{n+1})-\theta^n),\\
\mathscr{R}_h^{1,b}(\theta U^\varepsilon)&:=\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}|K|(V_K^{\varepsilon,n+1}-V_K^{\varepsilon,n})^T(\theta^n U_K^\varepsilon(t^{n+1})-(\theta U^\varepsilon)_K^n).
\end{aligned}$$

Since U^ε is a C^1-function in time (Theorem 3.2.2) and applying the Cauchy-Schwarz inequality we obtain

$$\begin{aligned}
|\theta^n U_K^\varepsilon(t^{n+1})-(\theta U^\varepsilon)_K^n|&\leq\frac{1}{|K|\Delta t}\left|\int_{t^n}^{t^{n+1}}\int_K(U^\varepsilon(x,t^{n+1})-U^\varepsilon(x,t))dx\theta(t)\,dt\right|\\
&\leq\mathscr{C}\theta^n\left(\frac{\Delta t}{|K|}\right)^{1/2}\left(\int_{t^n}^{t^{n+1}}\int_K\left|\frac{\partial U^\varepsilon}{\partial t}\right|^2dxdt\right)^{1/2},
\end{aligned}$$

where $\mathscr{C}$ is independent of ε.

Using the Cauchy-Schwarz inequality we get

$$|\mathscr{R}_h^{1,b}(\theta U^\varepsilon)|$$
$$\leq \mathscr{C}\Delta t^{1/2}\left[\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\int_{t^n}^{t^{n+1}}\int_K\left|\frac{\partial U^\varepsilon}{\partial t}\right|^2 dxdt\right]^{1/2}\left[\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\theta^n|K||V_K^{\varepsilon,n+1}-V_K^{\varepsilon,n}|^2\right]^{1/2},$$

which gives using Theorem 3.2.2, (3.27) and the CFL condition (3.15)

$$|\mathscr{R}_h^{1,b}(\theta U^\varepsilon)| \leq \mathscr{C}\left(\Delta t Q_h^\varepsilon\right)^{1/2}\left(\|U_0^\varepsilon\|_{H^1(\mathbb{R}^d;\mathbb{R}^{2m})}+\|F\|_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))}\right)$$
$$+\mathscr{C}\Delta t\left(\|U_0^\varepsilon\|_{H^1(\mathbb{R}^d;\mathbb{R}^{2m})}+\|F\|_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))}\right)^2. \tag{3.29}$$

On the other hand

$$|\mathscr{E}_h^1-2\mathscr{R}_h^{1,a}|\leq\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}|K||\theta(t^{n+1})-\theta^n||V_K^{\varepsilon,n+1}-V_K^{\varepsilon,n}||V_K^{\varepsilon,n+1}+V_K^{\varepsilon,n}-2U_K^\varepsilon(t^{n+1})|.$$

Since $|\theta(t^{n+1})-\theta^n|\leq\mathscr{C}\Delta t$, we get by the Cauchy-Schwarz inequality

$$|\mathscr{E}_h^1-2\mathscr{R}_h^{1,a}|\leq\mathscr{C}\left[\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\Delta t|K||V_K^{\varepsilon,n+1}-V_K^{\varepsilon,n}|^2\right]^{1/2}$$
$$\times\left[\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\Delta t|K||V_K^{\varepsilon,n+1}+V_K^{\varepsilon,n}-2U_K^\varepsilon(t^{n+1})|^2\right]^{1/2}.$$

We consider now the following splitting

$$V_K^{\varepsilon,n+1}+V_K^{\varepsilon,n}-2U_K^\varepsilon(t^{n+1})=V_K^{\varepsilon,n+1}-V_K^{\varepsilon,n}+2(V_K^{\varepsilon,n}-(U^\varepsilon)_K^n)+2((U^\varepsilon)_K^n-U_K^\varepsilon(t^{n+1})),$$

It is easy to check that

$$|V_K^{\varepsilon,n}-(U^\varepsilon)_K^n|\leq\frac{1}{(\Delta t|K|)^{1/2}}\left(\int_{t^n}^{t^{n+1}}\int_K|U^\varepsilon-U_h^\varepsilon|^2dxdt\right)^{1/2},$$

$$|U_K^\varepsilon(t^{n+1})-(U^\varepsilon)_K^n|\leq\left(\frac{\Delta t}{|K|}\right)^{1/2}\left(\int_{t^n}^{t^{n+1}}\int_K\left|\frac{\partial U^\varepsilon}{\partial t}\right|^2dxdt\right)^{1/2}.$$

With these inequalities, using (3.28), Theorem 3.2.2 and the CFL condition (3.15) we obtain

$$\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\Delta t|K||V_K^{\varepsilon,n+1}+V_K^{\varepsilon,n}-2U_K^\varepsilon(t^{n+1})|^2$$

$$\leq \mathscr{C}\|U^\varepsilon - U_h^\varepsilon\|^2_{L^2([0,T]\times\mathbb{R}^d;\mathbb{R}^{2m})} + \mathscr{C}\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\Delta t|K||V_K^{\varepsilon,n+1} - V_K^{\varepsilon,n}|^2$$
$$+ \mathscr{C}(\Delta t)^2\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\int_{t^n}^{t^{n+1}}\int_K\left|\frac{\partial U^\varepsilon}{\partial t}\right|^2 dxdt$$
$$\leq \mathscr{C}\|U^\varepsilon - U_h^\varepsilon\|^2_{L^2([0,T]\times\mathbb{R}^d;\mathbb{R}^{2m})} + \mathscr{C}\Delta t\left(\|U_0^\varepsilon\|_{H^1(\mathbb{R}^d;\mathbb{R}^{2m})} + \|F\|_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))}\right)^2,$$

again using (3.28) we get

$$|\mathscr{E}_h^1 - 2\mathscr{R}_h^{1,a}| \leq \mathscr{C}\Delta t\left(\|U_0^\varepsilon\|_{H^1(\mathbb{R}^d;\mathbb{R}^{2m})} + \|F\|_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))}\right)^2. \tag{3.30}$$
$$+ \mathscr{C}\Delta t^{1/2}\|U^\varepsilon - U_h^\varepsilon\|_{L^2([0,T]\times\mathbb{R}^d;\mathbb{R}^{2m})}\left(\|U_0^\varepsilon\|_{H^1(\mathbb{R}^d;\mathbb{R}^{2m})} + \|F\|_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))}\right).$$

Term $\mathscr{E}_h^2 - 2\mathscr{R}_h^2$: First we note that since θ depends only on t we have

$$\mathscr{E}_h^2 = 0. \tag{3.31}$$

Using the Cauchy-Schwarz inequality and the bound of $\|C_{e,K}^{\varepsilon,n}\|$ (Remark 3.3.1) we obtain

$$|\mathscr{R}_h^2| \leq \sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\sum_{e\in\partial K}\theta^n\Delta t|e|\left|\left(C_{e,K}^{\varepsilon,n}(V_K^{\varepsilon,n} - V_{K_e}^{\varepsilon,n})\right)^T((U^\varepsilon)_e^n - (U^\varepsilon)_K^n)\right|$$
$$\leq \mathscr{C}\left(\frac{1}{\sqrt{\varepsilon}}\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\sum_{e\in\partial K}\theta^n\Delta t|e|\ (V_K^{\varepsilon,n} - V_{K_e}^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n}(V_K^{\varepsilon,n} - V_{K_e}^{\varepsilon,n})\right)^{1/2}$$
$$\times\left(\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\sum_{e\in\partial K}\theta^n\Delta t|e||(U^\varepsilon)_e^n - (U^\varepsilon)_K^n|^2\right)^{1/2}$$
$$\leq \mathscr{C}\left(\frac{Q_h^\varepsilon}{\sqrt{\varepsilon}}\right)^{1/2}\left(\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\sum_{e\in\partial K}\theta^n\Delta t|e||(U^\varepsilon)_e^n - (U^\varepsilon)_K^n|^2\right)^{1/2}.$$

Thanks to the Cauchy-Schwarz inequality and Lemma 3.4.3 we find

$$|(U^\varepsilon)_e^n - (U^\varepsilon)_K^n| \leq \frac{1}{|e|^{1/2}}\left(\int_e|(U^\varepsilon)^n(x) - (U^\varepsilon)_K^n|^2dx\right)^{1/2} \leq \mathscr{C}\frac{h^{1/2}}{|e|^{1/2}}\left(\int_K|DU^\varepsilon|^2dx\right)^{1/2}.$$

Moreover, from Lemma 3.4.3 and since $U^\varepsilon(\cdot,t)\in H^1(\mathbb{R}^d;\mathbb{R}^{2m})$ (Theorem 3.2.2) we get

$$\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\sum_{e\in\partial K}\theta^n\Delta t|e||(U^\varepsilon)_e^n - (U^\varepsilon)_K^n|^2 \leq \mathscr{C}h\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\theta^n\int_{t^n}^{t^{n+1}}\int_K|DU^\varepsilon(x,t)|^2dxdt$$
$$\leq \mathscr{C}h\left(\|U_0^\varepsilon\|^2_{H^1(\mathbb{R}^d;\mathbb{R}^{2m})} + \|F\|^2_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))}\right)^2,$$

and finally

$$|\mathscr{R}_h^2| \le \mathscr{C}(Q_h^\varepsilon)^{1/2}\left(\frac{h}{\sqrt{\varepsilon}}\right)^{1/2}\left(\|U_0^\varepsilon\|_{H^1(\mathbb{R}^d;\mathbb{R}^{2m})} + \|F\|_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))}\right). \tag{3.32}$$

Term $\mathscr{E}_h^3 - 2\mathscr{R}_h^3$: From a direct calculation and Lemma 3.4.3, we obtain

$$|\mathscr{E}_h^3 - 2\mathscr{R}_h^3| \le T\int_{\mathbb{R}^d} |U_h^\varepsilon(0,x) - U_0^\varepsilon(x)|^2\,dx \le \mathscr{C}h^2. \tag{3.33}$$

Term $\mathscr{E}_h^4 - 2\mathscr{R}_h^4$: We first note that

$$\int_{t^n}^{t^{n+1}}\int_K \sum_{i=1}^d \frac{\partial}{\partial x_i}\big(A^{\varepsilon,i}U^\varepsilon\theta\big)\,dxdt = (T_1)_K^n + (T_2)_K^n + (T_3)_K^n + (T_4)_K^n + (T_5)_K^n + (T_6)_K^n, \tag{3.34}$$

where

$$\begin{aligned}
(T_1)_K^n &= \int_{t^n}^{t^{n+1}}\int_K \theta(t)\mathrm{div}A(x,t)(U^\varepsilon(x,t) - U_K^\varepsilon(t))\,dxdt,\\
(T_2)_K^n &= \int_{t^n}^{t^{n+1}}\int_K \theta(t)(\mathrm{div}A(x,t) - \mathrm{div}A^n(x))U_K^\varepsilon(t)\,dxdt,\\
(T_3)_K^n &= \int_{t^n}^{t^{n+1}}\int_K \theta(t)\mathrm{div}A^n(x)U_K^\varepsilon(t)\,dxdt,\\
(T_4)_K^n &= \int_{t^n}^{t^{n+1}}\int_K \theta(t)\sum_{i=1}^d (A^{\varepsilon,i}(x,t) - (A^{\varepsilon,i})_K(t))\frac{\partial}{\partial x_i}U^\varepsilon(x,t)\,dxdt,\\
(T_5)_K^n &= \int_{t^n}^{t^{n+1}}\int_K \theta(t)\sum_{i=1}^d ((A^{\varepsilon,i})_K(t) - (A^{\varepsilon,i})_K^n)\frac{\partial}{\partial x_i}U^\varepsilon(x,t)\,dxdt,\\
(T_6)_K^n &= \int_{t^n}^{t^{n+1}}\int_K \theta(t)\sum_{i=1}^d (A^{\varepsilon,i})_K^n\frac{\partial}{\partial x_i}U^\varepsilon(x,t)\,dxdt.
\end{aligned}$$

We get

$$\begin{aligned}
|\mathscr{R}_h^4| \le{}& \sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\left|(V_K^{\varepsilon,n})^T\int_{t^n}^{t^{n+1}}\left(\sum_{e\in\partial K}A_{e,K}^{\varepsilon,n}U_e^\varepsilon(t)\theta(t)|e|\right)dt - (T_3)_K^n - (T_6)_K^n\right|\\
&+ \sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\big(|(V_K^{\varepsilon,n})^T(T_1)_K^n| + |(V_K^{\varepsilon,n})^T(T_2)_K^n| + |(V_K^{\varepsilon,n})^T(T_4)_K^n| + |(V_K^{\varepsilon,n})^T(T_5)_K^n|\big).
\end{aligned} \tag{3.35}$$

It is easy to check that

$$\sum_{n\in\mathscr{N}}\sum_{K\in\mathscr{T}_h}\big(|(V_K^{\varepsilon,n})^T(T_1)_K^n| + |(V_K^{\varepsilon,n})^T(T_2)_K^n| + |(V_K^{\varepsilon,n})^T(T_4)_K^n| + |(V_K^{\varepsilon,n})^T(T_5)_K^n|\big)$$

$$\leq \mathscr{C} h \left(\|U_0^\varepsilon\|_{H^1(\mathbb{R}^d;\mathbb{R}^{2m})} + \|F\|_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))} \right)^2$$

holds where $\mathscr{C}$ does not depend on ε.
For example using the Cauchy-Schwarz inequality, Lemma 3.4.3, the regularity of U^ε ($U^\varepsilon(\cdot,t) \in H^1(\mathbb{R}^d;\mathbb{R}^{2m})$ by Theorem 3.2.2) and Proposition 3.4.1 we obtain

$$\begin{aligned}
&\sum_{n\in\mathscr{N}} \sum_{K\in\mathscr{T}_h} |(V_K^{\varepsilon,n})^T (T_1)_K^n| \\
&\quad \leq \mathscr{C} \sum_{n\in\mathscr{N}} \sum_{K\in\mathscr{T}_h} |V_K^{\varepsilon,n}| \left(\int_{t^n}^{t^{n+1}} \int_K 1 \cdot |U^\varepsilon(x,t) - U_K^\varepsilon(t)| dx dt \right) \\
&\quad \leq \mathscr{C} \sum_{n\in\mathscr{N}} \sum_{K\in\mathscr{T}_h} \Delta t^{1/2} |K|^{1/2} |V_K^{\varepsilon,n}| \left(\int_{t^n}^{t^{n+1}} \int_K |U^\varepsilon(x,t) - U_K^\varepsilon(t)|^2 dx dt \right)^{1/2} \\
&\quad \leq \mathscr{C} h \left(\|U_0^\varepsilon\|_{H^1(\mathbb{R}^d;\mathbb{R}^{2m})} + \|F\|_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))} \right)^2.
\end{aligned}$$

Returning to the first term on the R.H.S of (3.35) we note that thanks to Green's formula we find $\sum_{e\in\partial K} \sum_{i=1}^d n_{e,K}^i (A^{\varepsilon,i})_K^n U_K^\varepsilon(t) \theta(t) |e| = 0$. Therefore using the last expression and the Green formula we have that

$$(T_6)_K^n = \int_{t^n}^{t^{n+1}} \Big(\sum_{e\in\partial K} \sum_{i=1}^d n_{e,K}^i (A^{\varepsilon,i})_K^n U_e^\varepsilon(t) \theta(t) |e| \Big) dt.$$

This leads to

$$\begin{aligned}
&\sum_{n\in\mathscr{N}} \sum_{K\in\mathscr{T}_h} \left| (V_K^{\varepsilon,n})^T \int_{t^n}^{t^{n+1}} \left(\sum_{e\in\partial K} A_{e,K}^{\varepsilon,n} U_e^\varepsilon(t) \theta(t) |e| \right) dt - (T_3)_K^n - (T_6)_K^n \right| \\
&\quad = \sum_{n\in\mathscr{N}} \sum_{K\in\mathscr{T}_h} \left| (V_K^{\varepsilon,n})^T \int_{t^n}^{t^{n+1}} \left[\sum_{e\in\partial K} \left(A_{e,K}^{\varepsilon,n} - \sum_{i=1}^d n_{e,K}^i (A^{\varepsilon,i})_K^n \right) (U_e^\varepsilon(t) - U_K^\varepsilon(t)) \, \theta(t) |e| \right] dt \right| \\
&\quad \leq \sum_{n\in\mathscr{N}} \int_{t^n}^{t^{n+1}} \left(\sum_{K\in\mathscr{T}_h} \sum_{e\in\partial K} |V_K^{\varepsilon,n}|^2 \left| A_{e,K}^{\varepsilon,n} - \sum_{i=1}^d n_{e,K}^i (A^{\varepsilon,i})_K^n \right|^2 |e| \right)^{1/2} \\
&\qquad \times \left(\sum_{K\in\mathscr{T}_h} \sum_{e\in\partial K} |e| |U_e^\varepsilon(t) - U_K^\varepsilon(t)|^2 \right)^{1/2} \theta(t) dt.
\end{aligned}$$

We note that thanks to Lemma 3.4.3 and the regularity of A (Assumption 3.2.2)

$$|e| |U_e^\varepsilon(t) - U_K^\varepsilon(t)|^2 \leq \mathscr{C} h \int_{K\in\mathscr{T}_h} |DU^\varepsilon|^2 dx, \qquad |A_{e,K}^{\varepsilon,n} - \sum_{i=1}^d n_{e,K}^i (A^{\varepsilon,i})_K^n|^2 \frac{|e|}{|K|} \leq \mathscr{C} h.$$

Finally with the help of the stability result (Proposition 3.4.1) we obtain

$$|\mathscr{R}_h^4| \leq \mathscr{C} h \left(\|U_0^\varepsilon\|_{H^1(\mathbb{R}^d;\mathbb{R}^{2m})} + \|F\|_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))} \right)^2. \tag{3.36}$$

$\mathscr{E}_h^4$ can be treated in the same way as $\mathscr{R}_h^4$.

Terms $\mathscr{R}_h^5, \mathscr{R}_h^6, \mathscr{R}_h^7, \mathscr{E}_h^5 \mathscr{E}_h^6$: It is easy to check that

$$\begin{aligned} |\mathscr{R}_h^l| &\leq \mathscr{C} h \left(\|U_0^\varepsilon\|_{H^1(\mathbb{R}^d;\mathbb{R}^{2m})} + \|F\|_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))} \right)^2 \quad l = 5,6,7. \\ |\mathscr{E}_h^l| &\leq \mathscr{C} h \left(\|U_0^\varepsilon\|_{H^1(\mathbb{R}^d;\mathbb{R}^{2m})} + \|F\|_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))} \right)^2 \quad l = 5,6. \end{aligned} \tag{3.37}$$

For example, using the Cauchy-Schwarz inequality, Lemma 3.4.3, the regularity of F and θ, Theorem 3.2.2 and the CFL condition we obtain

$$\begin{aligned} \left|\mathscr{R}_h^7\right| &\leq \sum_{n\in\mathscr{N}} \sum_{K\in\mathscr{T}_h} \int_{t^n}^{t^{n+1}} \int_K |F(x,t)\cdot(\pi(x,t) - \pi_K(t) + \pi_K(t) - \pi_K^n)|\, dxdt \\ &\quad + \sum_{n\in\mathscr{N}} \sum_{K\in\mathscr{T}_h} \left(\int_{t^n}^{t^{n+1}} \int_K F^2\, dxdt \right)^{1/2} \left(\int_{t^n}^{t^{n+1}} \int_K |\pi_K(t) - \pi_K^n|^2\, dxdt \right)^{1/2} \\ &\leq \mathscr{C} h \sum_{n\in\mathscr{N}} \sum_{K\in\mathscr{T}_h} \left(\int_{t^n}^{t^{n+1}} \int_K F^2\, dxdt \right)^{1/2} \left(\int_{t^n}^{t^{n+1}} \int_K |D\pi|^2\, dxdt \right)^{1/2} \\ &\quad + \mathscr{C} \Delta t \sum_{n\in\mathscr{N}} \sum_{K\in\mathscr{T}_h} \left(\int_{t^n}^{t^{n+1}} \int_K F^2\, dxdt \right)^{1/2} \left(\int_{t^n}^{t^{n+1}} \int_K \left|\frac{\partial \pi}{\partial t}\right|^2 dxdt \right)^{1/2} \\ &\leq \mathscr{C}\,(h+\Delta t) \left(\|U_0^\varepsilon\|_{H^1(\mathbb{R}^d;\mathbb{R}^{2m})} + \|F\|_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))} \right)^2. \end{aligned}$$

Term $\mathscr{E}_h^7$: Using Cauchy-Schwarz inequality and (3.27) we obtain

$$\begin{aligned} |\mathscr{E}_h^7| &\leq \sum_{n\in\mathscr{N}} \sum_{K\in\mathscr{T}_h} 2\Delta t |K| |(V_K^{\varepsilon,n+1} - V_K^{\varepsilon,n})^T ((B_K^n + \mathrm{div} A_K^n) V_K^{\varepsilon,n} - F_K^n)\theta^n| \\ &\leq \mathscr{C} \left(\sum_{n\in\mathscr{N}} \sum_{K\in\mathscr{T}_h} \Delta t |K| |V_K^{\varepsilon,n+1} - V_K^{\varepsilon,n}|^2 \theta^n \right)^{1/2} \\ &\quad \times \left(\sum_{n\in\mathscr{N}} \sum_{K\in\mathscr{T}_h} \Delta t |K| |(B_K^n + \mathrm{div} A_K^n) V_K^{\varepsilon,n} - F_K^n|^2 \theta^n \right)^{1/2} \\ &\leq \mathscr{C} \left(\Delta t Q_h^\varepsilon + \Delta t^2 \left(\|U_0^\varepsilon\|_{H^1(\mathbb{R}^d;\mathbb{R}^{2m})} + \|F\|_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))} \right)^2 \right)^{1/2} \\ &\quad \times \left(\sum_{n\in\mathscr{N}} \sum_{K\in\mathscr{T}_h} \Delta t |K| |(B_K^n + \mathrm{div} A_K^n) V_K^{\varepsilon,n} - F_K^n|^2 \theta^n \right)^{1/2}. \end{aligned}$$

Using the stability result (Proposition 3.4.1), the bound on B (Assumption 3.2.2) and θ, and the regularity of F (Assumption 3.2.2) we find

$$\sum_{n\in\mathcal{N}}\sum_{K\in\mathcal{T}_h}\Delta t|K|(|(B_K^n+\mathrm{div}A_K^n)V_K^{\varepsilon,n}|^2+|F_K^n|^2\theta^n \leq \mathscr{C}\left(\|U_0^\varepsilon\|_{H^1(\mathbb{R}^d;\mathbb{R}^{2m})}+\|F\|_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))}\right)^2,$$

and finally we get

$$|\mathscr{E}_h^7| \leq \mathscr{C}(\Delta t)^{1/2}(Q_h^\varepsilon)^{1/2}\left(\|U_0^\varepsilon\|_{H^1(\mathbb{R}^d;\mathbb{R}^{2m})}+\|F\|_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))}\right) + \mathscr{C}\Delta t\left(\|U_0^\varepsilon\|_{H^1(\mathbb{R}^d;\mathbb{R}^{2m})}+\|F\|_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))}\right)^2. \tag{3.38}$$

Now, applying Proposition 3.4.2 and combining the estimates (3.29), (3.30), (3.31), (3.32), (3.33), (3.36), (3.37) and (3.38) we get

$$\begin{aligned}&\int_0^T\int_{\mathbb{R}^d}\exp(-\alpha t)|U^\varepsilon-U_h^\varepsilon|^2dxdt+\delta Q_h^\varepsilon\\ &\leq \mathscr{C}(h+\Delta t)\left(\|U_0^\varepsilon\|_{H^1(\mathbb{R}^d;\mathbb{R}^{2m})}+\|F\|_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))}\right)^2\\ &\quad+\mathscr{C}\left(\Delta t^{1/2}\|U^\varepsilon-U_h^\varepsilon\|_{L^2([0,T]\times\mathbb{R}^d;\mathbb{R}^{2m})}+\left(\frac{h}{\sqrt{\varepsilon}}Q_h^\varepsilon\right)^{1/2}\right)\\ &\quad\times\left(\|U_0^\varepsilon\|_{H^1(\mathbb{R}^d;\mathbb{R}^{2m})}+\|F\|_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))}\right).\end{aligned}$$

Appropriate application of Young's inequality yields to

$$\frac{\exp(-\alpha T)}{2}\|U^\varepsilon-U_h^\varepsilon\|^2_{L^2([0,T]\times\mathbb{R}^d;\mathbb{R}^{2m})}+\frac{\delta}{2}Q_h^\varepsilon \tag{3.39}$$

$$\leq \mathscr{C}\frac{h}{\sqrt{\varepsilon}}\left(\|U_0^\varepsilon\|_{H^1(\mathbb{R}^d;\mathbb{R}^{2m})}+\|F\|_{L^2(0,T;H^1(\mathbb{R}^d;\mathbb{R}^{2m}))}\right)^2. \tag{3.40}$$

Using the fact that $Q_h^\varepsilon\geq 0$ we conclude the theorem for smooth data. To derive the statement for the non-smooth case a standard mollification argument for the coefficients and the initial data can be used. □

Remark 3.4.1 *From Theorem 3.25 it is not clear how we can get benefit from the new formulation since choosing ε constant we recover the result of [112]. Numerical results will show the role of ε in the conservation of the constraint at the numerical level.*

For a first order method we can not expect to get a convergence rate directly for the expression $\sum_{i=1}^d M^i(u_h^\varepsilon)_{x_i}$. However it is possible to obtain a weak convergence estimate.

Corollary 3.4.1 *Suppose that the CFL-condition* (3.15) *holds. Let* $U_h^\varepsilon = (u_h^\varepsilon, \varphi_h^\varepsilon)$ *be the GLM-FV approximation from Definition 3.3.2. Then we have*

$$\left| \int_0^T \int_{\mathbb{R}^d} \sum_{i=1}^d M^i u_h^\varepsilon \cdot \frac{\partial \rho}{\partial x_i} dx dt \right| \leq \mathscr{C} \left(h\sqrt{\varepsilon} \right)^{1/2}, \forall \rho \in C_0^\infty([0,T) \times \mathbb{R}^d, \mathbb{R}^m).$$

Proof. We consider $\pi = (0,\rho)^T$ and $V = U_h^\varepsilon$ in Definition 3.4.1 with $\rho \in C_0^\infty([0,T) \times \mathbb{R}^d, \mathbb{R}^m)$. Then

$$\left| \int_0^T \int_{\mathbb{R}^d} \sum_{i=1}^d (\frac{M^i}{\sqrt{\varepsilon}} u_h^\varepsilon)^T \frac{\partial \rho}{\partial x_i} dx dt \right| = \left| - < \mu_{V_h}, \pi > - \int_0^T \int_{\mathbb{R}^d} (\varphi_h^\varepsilon)^T \partial_t \rho \, dx dt + a \int_0^T \int_{\mathbb{R}^d} (\varphi_h^\varepsilon)^T \rho \, dx dt \right|.$$

From Lemma 3.4.2 we know $< \mu_{V_h}, \pi > = \sum_{l=1}^7 \mathscr{R}_h^l(\pi)$. Moreover from Theorem 3.4.1 it is clear that

$$\left| \mathscr{R}_h^l(\pi) \right| \leq \mathscr{C} h \qquad \text{for} \quad l = 3,4,5,6,7.$$

It remains to focus just on $\mathscr{R}_h^1$ and $\mathscr{R}_h^2$. We have

$$\mathscr{R}_h^1(\pi) = \sum_{n \in \mathscr{N}} \sum_{K \in \mathscr{T}_h} |K| \big((\varphi_h^\varepsilon)_K^{n+1} - (\varphi_h^\varepsilon)_K^n \big)^T \big(\rho_K(t^{n+1}) - \rho_K^n \big).$$

Since ρ is a smooth function we find

$$|\rho_K(t^{n+1}) - \rho_K^n| \leq \left(\frac{\Delta t}{|K|} \right)^{1/2} \left(\int_{t^n}^{t^{n+1}} \int_K \left| \frac{\partial \rho}{\partial t} \right|^2 dx dt \right)^{1/2}.$$

Applying now the Cauchy-Schwarz inequality we get

$$\left| \mathscr{R}_h^1(\pi) \right| \leq \mathscr{C} \Delta t^{1/2} \left(\sum_{n \in \mathscr{N}} \sum_{K \in \mathscr{T}_h} |K| |(\varphi_h^\varepsilon)_K^{n+1} - (\varphi_h^\varepsilon)_K^n|^2 \right)^{1/2}.$$

Using (3.28) we find $|\mathscr{R}_h^1(\pi)| \leq \mathscr{C} \Delta t^{1/2}$.
For $\mathscr{R}_h^2$ using again the Cauchy-Schwarz inequality we have

$$\begin{aligned} \left| \mathscr{R}_h^2(\pi) \right| \leq & \mathscr{C} \varepsilon^{-1/4} \left(\sum_{n \in \mathscr{N}} \sum_{K \in \mathscr{T}_h} \sum_{e \in \partial K} \Delta t |e| (V_K^{\varepsilon,n} - V_{K_e}^{\varepsilon,n})^T C_{e,K}^{\varepsilon,n} (V_K^{\varepsilon,n} - V_{K_e}^{\varepsilon,n}) \right)^{1/2} \\ & \times \left(\sum_{n \in \mathscr{N}} \sum_{K \in \mathscr{T}_h} \sum_{e \in \partial K} \Delta t |e| (\pi_e^n - \pi_K^n)^2 \right)^{1/2} \end{aligned}$$

Using Lemma 3.4.3 and Proposition 3.4.1 we find

$$\left| \mathscr{R}_h^2(\pi) \right| \leq \mathscr{C} \left(\frac{h}{\sqrt{\varepsilon}} \right)^{1/2}.$$

Now we have altogether

$$\left|\int_0^T\int_{\mathbb{R}^d}\sum_{i=1}^d\left(\frac{M^i}{\sqrt{\varepsilon}}u_h^\varepsilon\right)^T\frac{\partial\rho}{\partial x_i}dxdt\right|\leq\left|\int_0^T\int_{\mathbb{R}^d}\left((\varphi_h^\varepsilon)^T\partial_t\rho-a(\varphi_h^\varepsilon)^T\rho\right)dxdt\right|+\mathscr{C}\left(\frac{h}{\sqrt{\varepsilon}}\right)^{1/2}.$$

Multiplying by $\sqrt{\varepsilon}$, using the Cauchy-Schwarz inequality and Theorem 3.25, we finally get

$$\begin{aligned}\left|\int_0^T\int_{\mathbb{R}^d}\sum_{i=1}^d(M^iu_h^\varepsilon)^T\frac{\partial\rho}{\partial x_i}dxdt\right|&\leq\mathscr{C}\left(\|\varphi_h^\varepsilon\|_{L^2([0,T]\times\mathbb{R}^d;\mathbb{R}^m)}\sqrt{\varepsilon}+(h\sqrt{\varepsilon})^{1/2}\right)\\&\leq\mathscr{C}\left(h\sqrt{\varepsilon}\right)^{1/2}.\end{aligned}$$

□

Corollary 3.4.1 showed the role of ε in the formulation. This parameter may possible to improve the rate of convergence of the involution, at least in the weak sense. Example 3.5.2 will show this behavior. In Example 3.5.3 we will see that the absence of ε yields to serious problems in the involution. Example 3.5.4 will give us some light on the function of a.

3.5 Numerical Examples

The purpose of this section is to illustrate the qualitative and quantitative behavior of the numerical solution generated by the GLM-FVM. We present the L^2-error respect to the exact solution when it is known and the discrete error of the involution also in the L^2-norm. In all the examples the time step is calculated according to $\Delta t = h\sqrt{\varepsilon}/8$. The simulations were performed using an Intel Proccesor Celeron of 2.26 GHz and 1024 MB of RAM memory.

3.5.1 Example 1

One example of Friedrichs systems with involutions are provided by the Maxwell equations given in $\mathbb{R}^3\times(0,\infty)$ by the system

$$\begin{aligned}\partial_t\mathbf{E}-\nabla\times\mathbf{B}=-\mathbf{j},\quad\partial_t\mathbf{B}+\nabla\times\mathbf{E}=0,\\\nabla\cdot\mathbf{E}=\rho,\quad\nabla\cdot\mathbf{B}=0.\end{aligned}\tag{3.41}$$

Here $\mathbf{E}=\mathbf{E}(x,t)\in\mathbb{R}^3$, $\mathbf{B}=\mathbf{B}(x,t)\in\mathbb{R}^3$, $\mathbf{j}=\mathbf{j}(x,t)\in\mathbb{R}^3$ and $\rho=\rho(x,t)\in\mathbb{R}$ denote the electric field, the magnetic induction, the current density and the charge density respectively.

	$\varepsilon = h^0$			$\varepsilon = h^{1/3}$		
h	L^2[error]	EOC	CPU	L^2[error]	EOC	CPU
0.04	4.98E-1	-	1.04E-1	5.72E-1	-	1.72E-1
0.02	3.11E-1	0.68	8.48E-1	3.87E-1	0.56	1.60E 0
0.01	1.75E-1	0.83	6.84E 0	2.37E-1	0.71	1.46E+1
0.005	9.34E-2	0.91	5.83E+1	1.36E-1	0.80	1.38E+2
0.004	7.57E-2	0.94	1.14E+2	1.13E-1	0.83	2.74E+2
0.0025	4.82E-2	0.96	4.87E+2	7.57E-2	0.85	1.07E+3

Table 3.1: Ex. 1: Numerical results for the GLM-FV($\varepsilon = h^{2/3}$) and FV method applied to Maxwell equations for different mesh size.

For the computations we consider the homogeneous Maxwell equations in two space dimensions (i.e. $\mathbf{j} = 0$, $\rho = 0$, B_1, B_2 and E_3 are constants), periodic boundary conditions on the computational domain $(0,1)\times(0,1)$, $t \in [0,1]$, and we use a Cartesian mesh with mesh parameter $h > 0$. We use $\varepsilon = \varepsilon(h)$ with $\varepsilon = h^\alpha$, $\alpha = \{0, 1/3, 2/3\}$. We set $a = 1$. In the case of the finite volume method directly applied to (3.1), (3.2) (referred to as FVM) we use $\Delta t = h/8$. It is easy to check that an exact solution of (3.41) (which is periodic) is given by

$$\begin{aligned} E_1(x_1,x_2,t) &= -\frac{k_\perp}{k_\parallel}\sin(k_\perp x_2)\cos(k_\parallel x_1 - \omega t), \\ E_2(x_1,x_2,t) &= \cos(k_\perp x_2)\sin(k_\parallel x_1 - \omega t), \\ B_3(x_1,x_2,t) &= \frac{\omega}{k_\parallel c^2}\cos(k_\perp x_2)\sin(k_\parallel x_1 - \omega t). \end{aligned} \tag{3.42}$$

Here $c > 0$ is the light speed, and the longitudinal and transverse wave numbers $k_\parallel$ and $k_\perp$, respectively, are related to the frequency ω according to $k_\parallel^2 + k_\perp^2 = \frac{\omega^2}{c^2}$. For this simulation we suppose $c = 1$ and $k_\parallel = k_\perp = 2\pi$. Initial condition for E_1, E_2 and B_3 are chosen according to (3.42) together with the results for the application of the GLM-FV method. The idea in this numerical example is to illustrate the rate of convergence predicted by Theorem 3.25. The results of the GLM-FVM are presented in Tables 3.1-3.2.

The rate of convergence takes a bigger value than we predicted. This is not a surprise because the rate predicted in the case without including the involution (FVM) is $1/2$ [112] but numerical simulations show an experimental convergence order of 1. Our case follows the same rule, the rate observed (2/3 and 5/6) is two times the rate predicted (1/3 and 5/12). The case $\varepsilon = 1$ follows the same behavior than the FVM. We do not show the error in the discrete computation of $\nabla\cdot\mathbf{E}$ because $\nabla\cdot\mathbf{E} = 0$ is preserved to machine precision.

	$\varepsilon = h^{2/3}$			FVM		
h	L^2[error]	EOC	CPU	L^2[error]	EOC	CPU
0.04	6.50E-1	-	2.96E-1	4.07E-1	-	8.40E-2
0.02	4.85E-1	0.42	2.98E 0	2.43E-1	0.74	6.36E-1
0.01	3.31E-1	0.55	3.12E+1	1.35E-1	0.85	5.31E 0
0.005	2.14E-1	0.63	3.38E+2	7.25E-2	0.90	4.46E+1
0.004	1.85E-1	0.65	6.91E+2	5.93E-2	0.90	9.04E+1
0.0025	1.34E-1	0.69	2.42E+3	3.91E-2	0.89	3.64E+2

Table 3.2: Ex. 1: Numerical results for the GLM-FV($\varepsilon = h^{2/3}$) and FV method applied to Maxwell equations for different mesh size.

Example 3.5.1 suggests that the use of the GLM-FV method does not pay off since the higher computational cost comes with an even worse convergence rate (compared to the original FV method). The next three examples show the benefits of the approach if the divergence error becomes more important.

3.5.2 Example 2

Another physical example of, in fact non-linear, conservation laws with involutions are the MHD equations in $\mathbb{R}^3 \times (0,\infty)$ given by

$$\begin{aligned} \partial_t \rho + \nabla \cdot (\rho \mathbf{u}) &= 0, \\ \partial_t (\rho \mathbf{u}) + \nabla \cdot \big(\rho \mathbf{u} \otimes \mathbf{u} + (p + \frac{1}{2}|\mathbf{B}|^2) I - \mathbf{B} \otimes \mathbf{B}\big) &= 0, \\ \partial_t \mathbf{B} + \nabla \cdot (\mathbf{u} \otimes \mathbf{B} - \mathbf{B} \otimes \mathbf{u}) &= 0, \\ \nabla \cdot \mathbf{B} &= 0. \end{aligned}$$

Here $\rho = \rho(x,t) \in \mathbb{R}$ is the density, $\mathbf{u} = \mathbf{u}(x,t) \in \mathbb{R}^3$ is the velocity field, $\mathbf{B} = \mathbf{B}(x,t) \in \mathbb{R}^3$ the magnetic field, $p = p(\rho) \in \mathbb{R}$ the pressure and I the identity matrix. For simplicity we write down the isentropic version. We consider the system of the MHD equations in two space dimensions and we suppose that the velocity field $\mathbf{u} = (u^1, u^2)^T$ is given. If we add the "source" term (which is zero due to $\nabla \cdot \mathbf{B} = 0$) $-\mathbf{u}\nabla \cdot \mathbf{B}$ to the induction equation we get (after some manipulations) the induction system

$$\partial_t \mathbf{B} + \partial_{x_1}(A^1 \mathbf{B}) + \partial_{x_2}(A^2 \mathbf{B}) + C\mathbf{B} = 0,$$

where

$$A^i = \begin{pmatrix} u^i & 0 \\ 0 & u^i \end{pmatrix}, \quad C = -\begin{pmatrix} \partial_{x_1} u^1 & \partial_{x_2} u^1 \\ \partial_{x_1} u^2 & \partial_{x_2} u^2 \end{pmatrix} \quad (i = 1,2).$$

	$\varepsilon = h^0$		$\varepsilon = h^{1/3}$		$\varepsilon = h^{2/3}$	
h	div-err.	CPU	div-err.	CPU	div-err.	CPU
0.04	5.84E-1	1.52E-1	2.89E-1	1.72E-1	1.66E-1	4.40E-1
0.02	4.87E-1	1.21E 0	2.37E-1	1.44E 0	1.18E-1	4.39E 0
0.01	3.55E-1	9.39E 0	1.67E-1	1.29E+1	6.81E-2	4.89E+1
0.005	2.30E-1	7.79E+1	9.72E-2	1.19E+2	3.63E-2	2.91E+2
0.004	1.96E-1	1.52E+2	7.99E-2	3.75E+2	3.02E-2	6.02E+2
0.0025	1.35E-1	6.22E+2	5.17E-2	1.30E+3	1.79E-2	3.95E+3

Table 3.3: Ex. 2: Divergence error for the GLM-FVM applied to a Induction Equation for defined mesh size and different choice of ε at $t = 0.5$.

This linear system fits exactly to our setting. We use again periodic boundary conditions in the domain $(0,1)^2$, $t \in [0,1)$ and Cartesian mesh. We also take $a = 1$. The idea of this numerical example is to study the behavior of the discrete version of $\nabla \cdot \mathbf{B}$ generated by the GLM-FV approximation. The velocity and initial condition are taken from [51]. These values are

$$B_0^1(x_1,x_2) = \partial_{x_2} A(x_1,x_2), B_0^2 = -\partial_{x_1} A(x_1,x_2),$$

where

$$A(x_1,x_2) = \frac{1}{2\pi}\sin(2\pi x_1)\cos(2\pi x_2) + x_2 - x_1,$$

and

$$\mathbf{u}(x_1,x_2) = (1,1) + 0.25(\cos(2\pi x_1) + 2\sin(2\pi x_2), \sin(2\pi x_1) + 2\cos(2\pi x_2)).$$

We present the results of the error $\nabla \cdot \mathbf{B}$ in the L^2-norm at $t = 0.5$ in Tables 3.3 and 3.4. The L^2-norm of the discrete $\nabla \cdot \mathbf{B}$, denoted by $\mathrm{div}_h \mathbf{B}_h^n$, is calculated as follows:

$$\mathrm{div}_h \mathbf{B}_h^n = \sqrt{\sum_{K \in \mathscr{T}_h} \left(\sum_{e \in \partial K} \mathbf{B}_{K_e}^n \cdot n_{e,K} \right)^2}.$$

We see that for both methods (FV and GLM-FV), the error in $\nabla \cdot \mathbf{B}$ converges to zero. However we observe also that the error in the GLM-FV method remains considerably lower than in the FV method and moreover we get better results for smaller values of $\varepsilon = \varepsilon(h)$.

3.5.3 Example 3

Even though the GLM-FV method damps the divergence error in Example 3.5.2 much better than the FVM, one might conclude that also the FVM leads to stable computations

	$\varepsilon = h^{5/6}$		FVM	
h	div-err.	CPU	div-err.	CPU
0.04	1.31E-1	3.60E-1	9.99E-1	3.60E-2
0.02	8.08E-2	3.83E 0	8.24E-1	2.88E-1
0.01	4.33E-2	4.08E+1	5.99E-1	2.27E 0
0.005	2.26E-2	4.49E+2	3.86E-1	1.89E+1
0.004	1.69E-2	9.60E+2	3.28E-1	3.70E+1
0.0025	1.05E-2	5.10E+3	2.26E-1	1.54E+2

Table 3.4: Ex. 2: Divergence error for the GLM-FVM applied to a Induction Equation for defined mesh size and different choice of ε at $t = 0.5$.

in the homogeneous case. This is not true for the inhomogeneous case with $\mathbf{j} \neq 0$ in the Maxwell equations as we will demonstrate below. In fact in almost all practical computations, the Maxwell equations are coupled to other equations via source terms. An uncontrollable increase in the divergence error can stop the computation. This problem was the motivation of Munz et al. [88] to develop the GLM. We use the same two-dimensional system as in Example 3.5.1 but with $\mathbf{j} \neq 0$. Taking the divergence in the Maxwell system we get

$$\partial_t(\nabla \cdot \mathbf{E}) = \partial_t \rho = -\nabla \cdot \mathbf{j}. \tag{3.43}$$

If we consider $\rho = 0$ we find the compatibility condition $\nabla \cdot \mathbf{j} = 0$. We want to study now the behavior of the GLM-FV method for the Maxwell equations under a small perturbation on the condition $\nabla \cdot \mathbf{j} = 0$. To do so we consider the following current density (for which the divergence is not zero)

$$j_1(x,y) = -1.001\frac{k_\perp}{k_\parallel}\sin(k_\perp y)\cos(k_\parallel x), \quad j_2(x,y) = \cos(k_\perp y)\sin(k_\parallel x).$$

with $k_\parallel$ and $k_\perp$ as in Example 1. Moreover we consider an initial electrical field $\mathbf{E}_0$ such that $\nabla \cdot \mathbf{E}_0 = 0$. We again set $a = 1$.
Accordingly to (3.43) we can expect a linear growth for $\mathrm{div}_h\mathbf{E}_h$ with respect to time in the case without correction. In Figure 3.1 we find the expected linear growth of $\mathrm{div}_h\mathbf{E}_h$ (calculated as in Example 3.5.2 for each time t) in the FV method and the bounded error in the GLM-FVM.

3.5.4 Example 4

In this example we want to study the influence of the relaxation parameter a considering $a = a(h)$. We again work with the homogeneous Maxwell equations as in Example

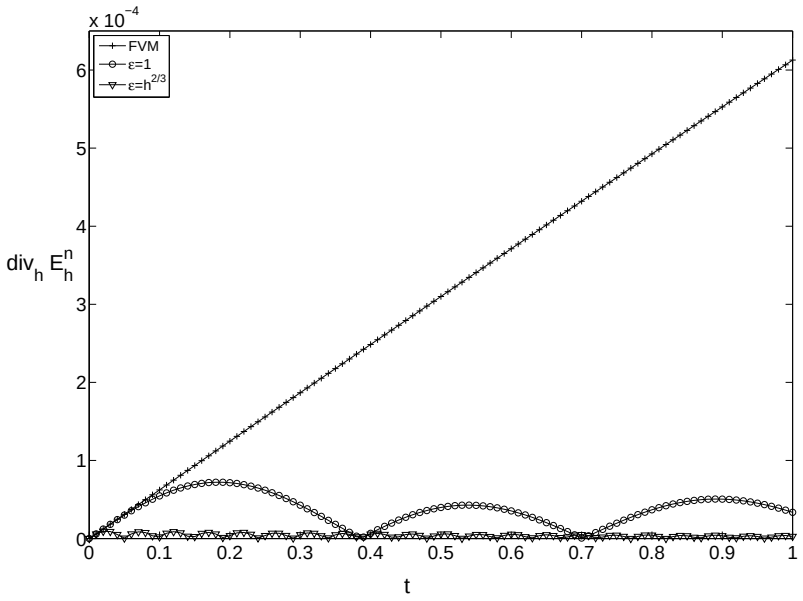

Figure 3.1: Ex. 3: Comparison of the divergence error between FVM and GLM-FV method for $\varepsilon = 1$ and $\varepsilon = h^{2/3}$ with $h = 0.0025$

3.5.1, but with the difference that we take an initial data u_0 that is not divergence free, i.e., div $\mathbf{E}_0 \neq 0$. For the first part we use $\varepsilon = h^{2/3}$ and $a(h) = h^{-q}$, with $q = 0, \frac{1}{3}, \frac{2}{3}, \frac{4}{3}$. We also simulated the FVM case. The results are showed in Tables 3.5 and 3.6.

	$a = 1$		$a = h^{-1/3}$		$a = h^{-2/3}$	
h	div-err.	L^2[error]	div-err.	L^2[error]	div-err.	L^2[error]
0.02	1.98E-1	4.74E-1	1.62E-1	4.74E-1	1.18E-1	4.74E-1
0.01	2.35E-1	3.18E-1	1.68E-1	3.18E-1	1.04E-1	3.18E-1
0.005	2.68E-1	1.99E-1	1.65E-1	1.99E-1	8.97E-2	1.99E-1
0.0025	2.94E-1	1.19E-1	1.55E-1	1.19E-1	7.80E-2	1.18E-1
0.00125	3.14E-1	7.10E-2	1.42E-1	6.87E-2	6.91E-2	6.84E-2

Table 3.5: Ex. 4: Errors for the GLM-FVM applied to a homogeneous Maxwell equations for defined $\varepsilon = \varepsilon(h)$ and different choice of $a = a(h)$.

First of all we note that the absence of the parameter $a = a(h)$ yields to a poor result in order to fulfill the divergence constraint. In the classical FVM, div$\mathbf{E}$ increase with we decrease the mesh. The case when $a = 1$ does not damp the divergence error showing the importance of coupling a with h. On the other side, in the case $a = h^{-4/3}$ the schemes also

	$a = h^{-4/3}$		FV	
h	div-err.	L^2[error]	div-err.	L^2[error]
0.02	1.64E-1	4.74E-1	5.20E-1	2.30E-1
0.01	2.04E-1	3.18E-1	5.70E-1	1.24E-1
0.005	2.57E-1	1.99E-1	5.98E-1	6.93E-2
0.0025	3.21E-1	1.20E-1	6.13E-1	4.96E-2
0.00125	3.91E-1	7.17E-2	6.21E-1	4.65E-2

Table 3.6: Ex. 4: Errors for the GLM-FVM applied to a homogeneous Maxwell equations for defined $\varepsilon = \varepsilon(h)$ and different choice of $a = a(h)$.

does not damp the divergence constraint.

We now isolate the behavior of a respect to ε. We computed the same simulations but with $\varepsilon = 1$ fixed and $a = h^{-1/3}, h^{-2/3}$ (see Table 3.7). When $a = h^{-1/3}$ we have a correct

	$a = h^{-1/3}$		$a = h^{-2/3}$	
h	div-err.	L^2[error]	div-err.	L^2[error]
0.02	2.18E-1	2.98E-1	2.07E-1	2.97E-1
0.01	2.20E-1	1.60E-1	2.40E-1	1.60E-1
0.005	2.15E-1	7.93E-2	2.90E-1	8.01E-2
0.0025	2.11E-1	4.05E-2	3.54E-1	4.35E-2
0.00125	2.10E-1	3.01E-2	4.21E-1	3.64E-2

Table 3.7: Ex. 4: Errors for the GLM-FVM applied to a homogeneous Maxwell equations for $\varepsilon = 1$ fixed and $a = h^{-1/3}, h^{-2/3}$.

damping of the divergence error but with $a = h^{-2/3}$ we find that the divergence error does not decrease as long as $h \to 0$.

Even tough in the framework of explicit schemes we can not get a theoretical result of convergence with $a = a(h)$, we speculate that an analogous result is possible if we consider a semi-implicit scheme. In that case, the rate predicted by Theorem 3.4.1 should include a term of the form $\mathcal{O}(a^{\overline{q}} h^q)$.

3.6 Conclusions

In this contribution we suggested a numerical method for Friedrichs systems with constraints in the form of an involution. It relies on an extended reformulation. This approach is a generalization of the approach for the equations of electrodynamics due to

Munz *et al*. [89] We have proven that the extended method gives convergence to a weak solution of the original initial value problem if a classical finite-volume discretization is applied. Moreover we have shown that the extended system allows to control the error in the primary unknowns and the constraint error. Several numerical examples which underline the analytical findings are added. The most important observation is that the GLM-FV method is stable under small perturbations on the side condition while the original finite volume method is not.

Future analytical investigations should address initial boundary value problems for Friedrichs systems. More important is the case of nonlinear conservation laws with (linear) involutions, e.g. Lundquist's equation of ideal magnetohydrodynamics subject to solenoidal magnetic fields. For this kind of nonlinear problems disregard of the constraint can lead to negative pressure and/or density values and thus to the simulation's abort. However the GLM-FV scheme still works convincingly (cf.[41]) but any rigorous argument is missing.

Chapter 4

General Conclusions

Here we present a summary with the main contributions and conclusions of the thesis.

- We propose a strongly degenerate parabolic equation modelling aggregation of "swarm" type. It is proved the existence of weak solutions for the initial value problem using a finite difference approach based on the Engquist-Osher method for the primitive of the function and using compactness arguments and Lax-Wendroff techniques. We prove the equivalence between weak and entropy solutions. Uniqueness is a corollary of a result in [62], which uses the doubling of variables technique. Numerical computations are performed finding the aggregation phenomenon. Solutions are discontinuous even though the initial data are not, in agreement with classical results of strongly degenerates parabolic equations. We prove the existence of travelling-wave solutions and their finite speed of propagation. Extension to higher space dimensions is not clear from the model. The equivalence of entropy and weak solutions for related equations is an open problem.

- We study a family of scalar conservation laws with nonlocal flux function modelling sedimentation. The existence and uniqueness of entropy solutions is proved using a difference-quadrature scheme and slight modifications of a result in [62]. For $\alpha = 0$, we find a Lipschitz regularity provided the data do so. In this case, the solution remains bounded for all time $T < +\infty$, however, the bound grows with the time. In the other case, $\alpha \geq 1$, we speculate discontinuities even though the data is smooth, as a counterpart, a Maximum Principle is valid for all time. From the physical point of view, the case $\alpha \geq 1$ is the relevant model. Numerical experiments are computed finding oscillations since the efective equation is dispersive. We interpret these oscillations as the formation of layer of different concentration. The formation of traveling waves is observed, being their analysis a future work to do. Other open problem corresponds to the asymptotic limit when the support of the kernel goes

to zero. Zumbrun [114] studied this limit but using a linear framework. This is not possible in the case studied in this thesis. Other alternatives will be considered.

- We study a Finite Volume method for Friedrichs systems with Involutions. The proposed method generalies the method developed by Munz et al. [88]. We prove the convergence of the numerical approximation to the unique solution of the problem. Moreover, it is shown that the involution is satisfied in the limit when the mesh parameter goes to zero. Numerical examples illustrated the performance of the method in the Maxwell equations and the induction equation in MHD. An open problem is the study of the initial-boundary value problem. Another open problem, much more challenging, is to develop a reliable numerical method that includes the involution in non-linear hyperbolic systems, like the Lundquist equations of MHD.

Chapter 5

Conclusiones Generales (en español)

A continuación, se presenta un resumen con los principales aportes y conclusiones generadas en esta tesis.

- Se propuso una ecuación parabólica fuertemente degenerada que modela el fenómeno de agregación tipo "enjambre". Se probó la existencia de solución para el problema de valores iniciales, en el marco de soluciones débiles, usando una aproximación basada en el esquema de Engquist-Osher para la primitiva de la función y utilizando argumentos de compacidad y de tipo Lax-Wendroff. Se demostró la equivalencia entre soluciones débiles y de entropía. La unicidad se obtuvo con técnicas de doblamiento de variables y del hecho de que las soluciones débiles son también de entropía. Se demostró la existencia de soluciones del tipo onda viajera y además se probó que la solución posee velocidad finita de propagación bajo supuestos adicionales. Se realizaron experiencias numéricas encontrándose el fenómeno de agregación. Las soluciones son discontinuas aunque el dato inicial no lo sea, encontrándose concordancia con los resultados clásicos para este tipo de ecuaciones. Extensiones a un mayor número de dimensiones espaciales no es clara del modelo. La equivalencia entre soluciones débiles y de entropía para ecuaciones relacionadas es un problema abierto.

- Se estudió una familia de leyes de conservación escalares con flujo no-local que modelan el proceso de sedimentación. Se demostró la existencia de solución para las ecuaciones usando un método de diferencias finitas con cuadratura y argumentos de compacidad, y para la unicidad, se usaron argumentos del tipo Kružkov. Para el caso $\alpha = 0$ se probó que la solución es continua si el dato también lo es. La solución permanece acotada para todo tiempo T, sin embargo la cota encontrada crece con el valor de T. Para $\alpha \geq 1$, se especula que la función presenta discontinuidades independiente de la regularidad del dato, pero por otra parte, se halló un Principio

del Máximo independiente del tiempo. Luego, desde el punto de vista físico, el caso $\alpha \geq 1$ corresponde a un modelo válido. Se realizaron experiencias numéricas, con el objetivo de reproducir el fenómeno de sedimentación por capas, interpretándose las oscilaciones que presenta la solución como un proceso de formación de capas de sedimento de distinta concentración. En los ejemplos numéricos se aprecia la formación de estructuras tipo ondas viajeras, quedando abierto su estudio. Otro problema por resolver corresponde al límite cuando el soporte del kernel tiende a cero. Zumbrun [114] estudió este límite en un caso particular haciendo uso de una estructura de tipo lineal. Para el caso que se estudió en el capítulo 2, esto no es posible. Otras alternativas serán estudiadas a futuro.

- En el tercer capítulo se estudió un método numérico de volúmenes finitos para sistemas de Friedrichs con restricciones en la forma de involuciones. El método corresponde a una generalización del método desarrollado por Munz y colaboradores [88]. Se probó la convergencia del método a la solución débil del problema. Además se demostró la satisfacción de la involución cuando el parámetro de malla tiende a cero. Ejemplos numéricos ilustran el desempeño del método en la caso de las ecuaciones de Maxwell y de la ecuación de inducción en magneto-hidrodinámica. Un problemas abierto corresponde al estudio del problema de valores iniciales y de contorno. Otro problema abierto corresponde al desarrollo de métodos numéricos convergentes que consideren restricciones tipo involuciones en el caso más general de sistemas hiperbólicos no-lineales, como lo son las ecuaciones de Lundquist en magneto-hidrodinámica.

Bibliography

[1] C.D. Acosta and C.E. Mejía. Approximate solution of hyperbolic conservation laws by discrete mollification. *Appl. Numer. Math.*, 59:2256–2265, 2009.

[2] G. Aletti, G. Naldi, and G. Toscani. First-order continuous models of opinion formation. *SIAM J. Appl. Math.*, 67:837–853, 2007.

[3] W. Alt. Degenerate diffusion equations with drift functionals modelling aggregation. *Nonlin. Anal. TMA*, 9:811–836, 1985.

[4] F. Assous, P. Degond, E. Heintze, P.A. Raviart, and J. Segré. On a finite element method for solving the three-dimensional Maxwell equations. *J. Comput. Phys.* , 109: 222–237, 1993.

[5] G.R. Baker, X. Li, and A.C. Morlet. Analystic structure of two 1D-transport equations with nonlocal fluxes. *Phys. D*, 91:349–375, 1996.

[6] M. Bargiel, R.A. Ford, and E.M. Tory. Simulation of sedimentation of polydisperse suspensions: a particle-based approach. *AIChE J.*, 51:2457–2468, 2005.

[7] G.K. Batchelor. Sedimentation in a dilute dispersion of spheres. *J. Fluid Mech.*, 52:245–268, 1972.

[8] J. Bedrossian, N. Rodríguez, and A. Bertozzi. Local and global well-posedness for aggregation equations and Patlak-Keller-Segel models with degenerate diffusion. Preprint; submitted.

[9] C.W.J. Beenakker and P. Mazur. Diffusion of spheres in suspension: Three-body hydrodynamic interaction effects. *Phys. Lett.*, 91A:290–291, 1982.

[10] C.W.J. Beenakker and P. Mazur. Diffusion of spheres in a concentrated suspension II. *Physica A*, 126:349–370, 1984.

[11] C.W.J. Beenakker, W. van Saarloos, and P. Mazur. Many-sphere hydrodynamic interactions III. The influence of a plane wall. *Physica A*, 127:451–472, 1984.

[12] C.W.J. Beenakker and P. Mazur. On the Smoluchowski paradox in a sedimenting suspension. *Phys. Fluids*, 28:767–769, 1985.

[13] C.W.J. Beenakker and P. Mazur. Is sedimentation container-shape dependent?. *Phys. Fluids*, 28:3203–3206, 1985.

[14] S. Benzoni-Gavage and D. Serre. *Multidimensional Hyperbolic Partial Differential Equations. First-order Systems and Applications.* Oxford, Oxford Science Publications, 2007.

[15] A.L. Bertozzi and J. Brandman. Finite-time blow-up of L^∞-weak solutions of an aggregation equation. *Commun. Math. Sci.*, 8:45–65, 2010.

[16] A.L. Bertozzi, J.A. Carrillo, and T. Laurent. Blow-up in multidimensional aggregation equations with mildly singular interaction kernels. *Nonlinearity*, 22:683–710, 2009.

[17] A.L. Bertozzi and T. Laurent. Finite-time blow-up of solutions of an aggregation equation in $\mathbb{R}^n$. *Comm. Math. Phys.*, 274:717–735, 2007.

[18] A.L. Bertozzi, T. Laurent, and J. Rosado. L^p theory for the multidimensional aggregation equation. *Comm. Pure Appl. Math.*, 64:45–83, 2011.

[19] A.L. Bertozzi and D. Slepčev. Existence and uniqueness of solutions to an aggregation equation with degenerate diffusion. *Comm. Pure Appl. Anal.*, 9:1617–1637, 2010.

[20] N. Besse and D. Kröner. Convergence of locally divergence-free discontinuous Galerkin methods for the induction equations of the 2D-MHD equations. *M2AN Math. Mod. Num. Anal.*, 39:1177-1202, 2005.

[21] A. Blanchet, J.A. Carrillo, and P. Laurençot. Critical mass for a Patlak-Keller-Segel model with degenerate diffusion in higher dimensions. *Calc. Var.*, 35:133–168, 2009.

[22] M. Bodnar and J.J.L. Velazquez. An integro-differential equation arising as a limit of individual cell-based models. *J. Differential Equations*, 222:341–380, 2006.

[23] H. Brézis and M.G. Crandall. Uniqueness of solutions of the initial-value problem for $u_t - \Delta\varphi(u) = 0$. *J. Math. Pures et Appl.*, 58:1979, 153–163.

[24] J.U. Brackbill and D.C. Barnes. The effect of nonzero $\nabla \cdot \mathbf{B}$ on the numerical solution of the magnetohydrodynamic equations. *J. Comput. Phys.* , 35:426-430, 1980.

[25] R. Bürger, K.H. Karlsen, E.M. Tory, and W.L. Wendland. Model equations and instability regions for the sedimentation of polydisperse suspensions of spheres. *ZAMM Z. Angew. Math. Mech.*, 82:699–722, 2002.

[26] R. Bürger and E.M. Tory. On upper rarefaction waves in batch settling. *Powder Technol.*, 108:74–87, 2000.

[27] R. Bürger and K.H. Karlsen. On a diffusively corrected kinematic-wave traffic model with changing road surface conditions. *Math. Models Methods Appl. Sci.,* 13:1767–1799, 2003.

[28] M. Burger, V. Capasso, and D. Morale. On an aggregation model with long and short range interactions. *Nonlin. Anal. Real World Appl.*, 8:939–958, 2007.

[29] M.C. Bustos, F. Concha, R. Bürger, and E.M. Tory. *Sedimentation and Thickening. Phenomenological Foundation and Mathematical Theory.* Dordrecht, Kluwer Academic Publishers, 1999.

[30] J. Carrillo. Entropy solutions for nonlinear degenerate problems. *Arch. Rational Mech. Anal.*, 147:269–361, 1999.

[31] J.A. Carrillo, M. Di Francesco, A. Figalli, T. Laurent, and D. Slepčev. Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations. *Duke Math. J.*, to appear.

[32] G.-Q. Chen and K.H. Karlsen. Quasilinear anisotropic degenerate parabolic equations with time-space dependent diffusion coefficients. *Commun. Pure Appl. Anal.*, 4:241–266, 2005.

[33] R. Caflisch and G.C. Papanicolaou. Dynamic theory of suspensions with Brownian effects. *SIAM J. Appl. Math.*, 43:885–906, 1983.

[34] A. Castro and D. Córdoba. Global existence, singularities and ill-posedness for a nonlocal flux. *Adv. Math.*, 219:1916–1936, 2008.

[35] B. Cockburn, F. Li, and C.W. Shu. Locally divergence-free discontinuous Galerkin methods for the Maxwell equations. *J. Comput. Phys.* , 194:588-610, 2004.

[36] R.M. Colombo, G. Facchi, G. Maternini, and M.D. Rosini. On the continuum modelling of crowds. In: *Hyperbolic Problems: Theory, Numerics and Applications.* Proc. Sympos. Appl. Math., 67, Part 2, Eds.: E. Tadmor, J.-G. Liu, and A. Tzavaras. pp:517–526, Providence, Amer. Math. Soc., 2009.

[37] R.M. Colombo, M. Herty, and M. Mercier. Control of the continuity equation with a non local flow. *ESAIM: Contr. Opt. Calc. Var.*, to appear.

[38] P.A. Davidson. *An Introduction to Magnetohydrodynamics.* Cambridge, Cambridge University Press, 2001.

[39] C. Dafermos. *Hyperbolic Conservation Laws in Continuum Physics.* Berlin, Springer, 1991.

[40] C. Dafermos. Hyperbolic conservation laws with involutions and contingent entropies. In: *Recent advances in nonlinear partial differential equations and applications.* Providence, Amer. Math. Soc., 2007.

[41] A. Dedner, F. Kemm, D. Kröner, C.D. Munz, T. Schnitzer, and M. Wesenberg. Hyperbolic divergence cleaning for the MHD equations. *J. Comput. Phys.*, 175:645-673, 2003.

[42] J.I. Diaz, T. Nagai, and S.I. Shmarev. On the interfaces in a nonlocal quasilinear degenerate equation arising in population dynamics. *Japan J. Indust. Appl. Math.*, 13:385–415, 1996.

[43] M. Di Francesco, P.A. Markowich, J.-F. Pietschmann, M.-T. Wolfram. On the Hughes' model for pedestrian flow: The one-dimensional case. Preprint; submitted.

[44] B. Engquist and S. Osher. One-sided difference approximations for nonlinear conservation laws. *Math. Comp.*, 36:321–351, 1981.

[45] C.R. Evans and J.F. Hawley. Simulation of general relativistic magnetohydrodynamic flows: A constrained transport method. *Astrophys. J.* , 332:659, 1998.

[46] S. Evje and K.H. Karlsen. Monotone difference approximations of BV solutions to degenerate convection-diffusion equations. *SIAM J. Numer. Anal.*, 37:1838–1860, 2000.

[47] S. Evje and K.H. Karlsen. Discrete approximations of BV solutions to doubly degenerate parabolic equations. *Numer. Math.*, 86:377–417, 2000.

[48] R. Eymard, T. Gallouët, R. Herbin, and A. Michel. Convergence of a finite volume scheme for nonlinear degenerate parabolic equations. *Numer. Math.*, 92:41–82, 2002.

[49] M. Fey and M. Torrilhon. A constrained transport upwind scheme for divergence-free advection. In: *Hyperbolic Problems: Theory, Numerics and Applications.* Eds.: T. Y. Hou and E. Tadmor. New York, Springer, 2003.

[50] K.O. Friedrichs. Symmetric hyperbolic linear differential equations.*Commun. Pure Appl. Math.* , 11:345–392, 1958.

[51] F.G. Fuchs, K.H. Karlsen, S. Mishra, and N.H. Risebro. Stable upwind schemes for the magnetic induction equation. *M2AN Math. Mod. Num. Anal.*, 43:825–852, 2009.

[52] J. Happel and H. Brenner. *Low Reynolds Number Hydrodynamics.* Dordrecht, Martinus Nijhoff, 1986.

[53] C.H. Hesse and E.M. Tory. The stochastics of sedimentation. In: *Sedimentation of Small Particles in a Viscous Fluid.* Ed.: E.M. Tory. pp: 199–239. Southampton, Computational Mechanics Publications, 1996.

[54] K. Höfler. *Räumliche Simulation von Zweiphasenflüssen*. Diplomarbeit, Institut für Computeranwendungen, U. Stuttgart, Germany, 1997.

[55] H. Holden, K.H. Karlsen, and N.H. Risebro. On uniqueness and existence of entropy solutions of weakly coupled systems of nonlinear degenerate parabolic equations. *Electron. J. Differential Equations*, 46:1–31, 2003.

[56] R.L. Hughes. A continuum theory for the flow of pedestrians. *Transp. Res. B*, 36:507–535, 2002.

[57] R. Jeltsch and M. Torrilhon Solenoidal initial conditions for locally divergence-free MHD simulations. In: *Modeling, Simulation and Optimization of Complex Processes.* Eds.: H.G. Bock, E. Kostina, H.X. Phu, and R. Rannacher. Berlin, Springer, 2005.

[58] B. Jiang, J. Wu and L.A. Povinelli. The origin of spurious solutions in computational electromagnetics, *J. Comput. Phys.* , 125:104–123, 1996.

[59] V. Jovanovic and C. Rohde. Finite-volume schemes for Friedrichs systems in multiple space dimensions: *A Priori* and *A Posteriori* error estimates. *Numer. Methods Partial Differential Eq.* , 21:104–131, 2004.

[60] M.T. Kamel and E.M. Tory. Sedimentation of clusters of identical spheres I. Comparison of methods for computing velocities. *Powder Technol.*, 59:227–248, 1989. Erratum, ibid. 94:266, 1997.

[61] K.H. Karlsen and N.H. Risebro. Convergence of finite difference schemes for viscous and inviscid conservation laws with rough coefficients. *M2AN Math. Model. Numer. Anal.*, 35:239–269, 2001.

[62] K.H. Karlsen and N.H. Risebro. On the uniqueness and stability of entropy solutions for nonlinear degenerate parabolic equations with rough coefficients. *Discr. Contin. Dyn. Syst.*, 9:1081–1104, 2003.

[63] K.H. Karlsen, N.H. Risebro, and J.D. Towers. L^1 stability for entropy solutions of nonlinear degenerate parabolic convection-diffusion equations with discontinuous coefficients. *Skr. K. Nor. Vid. Selsk.*, 3:1–49, 2003

[64] W.O. Kermack, A.G. M'Kendrick, and E. Ponder. The stability of suspensions. III. The velocities of sedimentation and of cataphoresis of suspensions in a viscous fluid. *Proc. Roy. Soc. Edinburgh*, 49:170–197, 1929.

[65] M. Khodja and K. Zumbrun. Stabilité des profils de choc pour une équation dispersive. *C. R. Acad. Sci. Paris Sér. I*, 325:163–166, 1997.

[66] K. Kobayasi. The equivalence of weak solutions and entropy solutions of nonlinear degenerate second-order equations. *J. Differential Equations*, 189:383–395, 2003.

[67] A. Kurganov and A. Polizzi. Non-oscillatory central schemes for traffic flow models with Arrhenius look-ahead dynamics. *Netw. Heterog. Media*, 4:431–451, 2009.

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

[69] S.N. Kružkov. First order quasilinear equations in several independent variables. *Math. USSR Sbornik*, 10:217–243, 1970.

[70] G.J. Kynch. A theory of sedimentation. *Trans. Faraday Soc.*, 48:166–176, 1952.

[71] A.J.C. Ladd. Dynamical simulations of sedimenting spheres. *Phys. Fluids A*, 5:299–310, 1993.

[72] A.J.C. Ladd. Numerical simulation of particulate suspensions via a discretized Boltzmann equation, Part I. Theoretical foundation. *J. Fluid Mech.*, 271:285–309, 1994.

[73] A.J.C. Ladd. Numerical simulation of particulate suspensions via a discretized Boltzmann equation, Part II. Numerical results. *J. Fluid Mech.*, 271:311–339, 1994.

[74] A.J.C. Ladd. Sedimentation of homogeneous suspensions of non-Brownian spheres. *Phys. Fluids*, 9:491–499, 1997.

[75] Landau and Lifshitz. *Electrodynamics of Continuous Media.* Burlington, Elsevier Butterworth-Heinemann, 1984.

[76] T. Laurent. Local and global existence for an aggregation equation. *Commun. Part. Diff. Eqs.*, 32:1941–1964, 2007.

[77] R.J. Le Veque. *Numerical Methods for Conservation Laws.* Basel, Birkhäuser, 1992.

[78] D. Li and J. Rodrigo. Finite-time singularities of an aggregation equation in $\mathbb{R}^n$ with fractional dissipation. *Commun. Math. Phys.*, 287:687–703, 2009.

[79] D. Li and J. Rodrigo. Refined blowup criteria and nonsymmetric blowup of an aggregation equation. *Adv. in Math.*, 220:1717–1738, 2009.

[80] F. Li and C.W. Shu. Locally divergence-free discontinuous Galerkin methods for MHD equations, *J. Sci. Comput.* , 22:413–442, 2005.

[81] I-S. Liu. *Continuum Mechanics.* Berlin, Springer, 2002.

[82] D. Li and X. Zhang. On a nonlocal aggregation model with nonlinear diffusion. *Discr. Cont. Dyn. Syst.*, 27:301–323, 2010.

[83] P. Mazur and W. van Saarloos. Many-sphere hydrodynamic interactions and mobilities in a suspension. *Physica A*, 115:21–57, 1982.

[84] D. Mercier and S. Nicaise. Existence, uniqueness and regularity results for piezo-electrical systems. *SIAM J. Math. Anal.*, 37:651–672, 2005.

[85] A. Mogilner, L. Edelstein-Keshet, L. Bent, and A. Spiros. Mutual interactions, potentials, and individual distance in a social aggregation. *J. Math. Biol.*, 47:353–389, 2003.

[86] D. Morale, V. Capasso, and K. Oelschläger. An interacting particle system modelling aggregation behavior: from individuals to populations. *J. Math. Biol.*, 50:49–66, 2005.

[87] P.J. Mucha, S.-Y. Tee, D.A. Weitz, B.I. Shraiman, and M.P. Brenner. A model for velocity fluctuations in sedimentation. *J. Fluid Mech.*, 501:71–104, 2004.

[88] C.D. Munz, P. Ommes, R. Schneider, E. Sonnendrücker, and U. Voss. Maxwell's equations when the charge conservation is not satisfied. *C.R. Acad. Sci. Paris Sér I Math.*, 328:431–436, 1999.

[89] C.D. Munz, P. Ommes, R. Schneider, E. Sonnendrücker, and U. Voss. Divergence correction techniques for Maxwell solvers based on a hyperbolic model. *J. Comput. Phys.*, 161:484–511, 2000.

[90] T. Nagai. Some nonlinear degenerate diffusion equations with a nonlocally convective term in ecology. *Hiroshima Math. J.*, 13:165–202, 1983.

[91] T. Nagai and M. Mimura. Some nonlinear degenerate diffusion equations related to population dynamics. *J. Math. Soc. Japan*, 35:539–562, 1983.

[92] T. Nagai and M. Mimura. Asymptotic behavior for a nonlinear degenerate diffusion equation in population dynamics. *SIAM J. Appl. Math.*, 43:449–464, 1983.

[93] T. Nagai and M. Mimura. Asymptotic behavior of the interfaces to a nonlinear degenerate diffusion equation in population dynamics. *Japan J. Appl. Math.*, 3:129–161, 1986.

[94] H. Nessyahu and E. Tadmor. Non-oscillatory central differencing for hyperbolic conservation laws. *J. Comput. Phys.*, 87:408–463, 1990.

[95] N.-Q. Nguyen and A.J.C. Ladd. Sedimentation of hard-sphere suspensions at low Reynolds number. *J. Fluid Mech.*, 525:73–104, 2005.

[96] D.K. Pickard and E.M. Tory. A Markov model for sedimentation. *J. Math. Anal. Appl.*, 60:349–369, 1977.

[97] D.K. Pickard and E.M. Tory. Experimental implications of a Markov model for sedimentation. *J. Math. Anal. Appl.*, 72:150–176, 1979.

[98] D.K. Pickard and E.M. Tory. A Markov model for sedimentation: Fundamental issues and insights. In: *Advances in the Statistical Sciences.* Eds.: I.B. MacNeill and G.J. Umphrey, Vol. IV, Stochastic Hydrology. pp:1–25, Dordrecht, D. Reidell, 1987.

[99] K.G. Powell. An approximate Riemann solver for magnetohydrodynamics (That works in more than one dimension). *ICASE-Report 94-24*, 1994.

[100] J.F. Richardson and W.N. Zaki. Sedimentation and fluidization: Part I. *Trans. Instn Chem. Engrs (London)* 32:35–53, 1954.

[101] E. Rouvre and G. Gagneux. Solution forte entropique de lois scalaires hyperboliques-paraboliques dégénérées. *C. R. Acad. Sci. Paris Sér. I*, 329:599–602, 1999.

[102] J. Rubinstein. Evolution equations for stratified dilute suspensions. *Phys. Fluids A*, 2:3–6, 1990.

[103] J. Rubinstein and J.B. Keller. Particle distribution functions in suspensions. *Phys. Fluids A*, 1:1632–1641, 1989.

[104] J. Rubinstein and J.B. Keller. Sedimentation of a dilute suspension. *Phys. Fluids A*, 1:637–643, 1989.

[105] C.M. Topaz, A.L. Bertozzi, and M.A. Lewis. A nonlocal continuum model for biological aggregation. *Bull. Math. Biol.*, 68:1601–1623, 2006.

[106] M. Torrilhon. Numerical pseudo-convergence for a MHD model system. In: *Hyperbolic Problems: Theory, Numerics and Applications*. Proc. 10th Int. Conf. Hyperbolic Problems in Osaka, Japan 2004. Eds.: F. Asakura et al., Yokohama, Yokohama Publishers, 2006.

[107] E.M. Tory and D.K. Pickard. Extensions and refinements of a Markov model for sedimentation. *J. Math. Anal. Appl.*, 86:442–470, 1982.

[108] E.M. Tory and M.T. Kamel. On the divergence problem in calculating particle velocities in dilute dispersions of identical spheres II. Effect of a plane wall, *Powder Technol.*, 55:51–59, 1988. Erratum, *ibid.* 94:265, 1997.

[109] D.B. Siano. Layered sedimentation in suspensions of monodisperse spherical colloidal particles. *J. Colloid Interface Sci.* 68:111–127, 1979.

[110] B. Sjögreen, K. Gustavsson, and R. Gudmundsson. A model for peak formation in the two-phase equations. *Math. Comp.*, 76:1925–1940, 2007.

[111] A. Sopasakis and M.A. Katsoulakis. Stochastic modelling and simulation of traffic flow: asymmetric single exclusion process with Arrhenius look-ahead dynamics. *SIAM J. Appl. Math.*, 66:921–944, 2006.

[112] J.P. Vila and P. Villedieu. Convergence of an explicit finite volume scheme for first order symmetric systems. *Numer. Math.*, 94:573–602, 2003.

[113] W.A. Yong. Basic aspects of hyperbolic relaxation systems. In: *Advances in the Theory of Shocks Waves*, Progress in Nonlinear Differential Equations and their Applications 47, Eds.: Freistühler et al., pp:259–305, Boston, Birkhaüser, 2001.

[114] K. Zumbrun. On a nonlocal dispersive equation modeling particle suspensions. *Quart. Appl. Math.*, 57:573–60, 1999.