

The Finite Volume Particle Method
– **Recent Developments and Applications**

Dissertation
zur Erlangung des Doktorgrades
der Fakultät für Mathematik, Informatik
und Naturwissenschaften
der Universität Hamburg

vorgelegt
im Fachbereich Mathematik

von
Christine Kaland
aus Hamburg

Hamburg
2013

Als Dissertation angenommen vom Fachbereich
Mathematik der Universität Hamburg

Auf Grund der Gutachten von Prof. Dr. Jens Struckmeier
und Prof. Dr. Thomas Götz

Hamburg, den 17.12.2013

Prof. Dr. Armin Iske
Leiter des Fachbereichs Mathematik

Acknowledgements

I would like to thank my supervisor Prof. Dr. Jens Struckmeier for the interesting topic, his help and his great encouragement. I could always rely on his support during the crucial phases of the work. His continual confidence and patience during the last years were a fundamental help in finishing this thesis.

Moreover, I would like to thank Dr. Stefan Heitmann for many fruitful discussions. I also would like to express my gratitude to Prof. Dr. Thomas Götz for agreeing to act as second examiner.

Special thanks go to my friends Philipp Wruck and Felix Salzinger, who always supported me with their close and honest friendship. Philipp, your comments on this thesis helped a lot. Last but not least, I want to express my deepest gratitude to my companion and best friend Morten, my parents Gisela and Joachim, my sister Lena and my aunt Stephanie for their support and love. Especially Morten and Lena deserve my most profound thanks for all those discussions on the topic and a very great time. Thank you very much for being there for me during the last years.

Contents

Introduction	1
1 Preliminaries	3
1.1 Hyperbolic conservation laws	3
1.2 Finite Volume Methods	7
1.3 Kinetic schemes	11
2 The Finite Volume Particle Method	15
2.1 Derivation of the FVPM	16
2.2 The geometrical coefficients and properties of the scheme	21
2.3 Stability and entropy results	32
3 The FVPM with B-splines	45
3.1 Introduction	45
3.2 The FVPM with B-splines in one spatial dimension	49
3.3 Special cases	56
4 The FVPM for granular flows	61
4.1 The Savage-Hutter equations	61
4.2 A kinetic FVM for the SH equations	62
4.3 The FVPM with kinetic fluxes	71
5 Numerical results	83
5.1 The 1D Euler equations	84
5.2 A linearized piston problem with moving boundary	88
5.3 The SH equations	96
5.4 Solid body rotation	101
Conclusions	108
Appendix	109
Nomenclature	113
Bibliography	115

Introduction

Many dynamical processes can be described by partial differential equations, and convection dominated phenomena as occurring in fluid dynamics are often modeled by hyperbolic conservation laws. As an example, the Euler equations of gas dynamics, the shallow water equations or the Savage-Hutter equations are used to characterize different forms of inviscid fluid flow. Since the structure of hyperbolic conservation laws is different from that of elliptic or parabolic equations, there is a need for specially designed numerical methods that take into account characteristics of the solutions, like discontinuities or the transport direction. Classically, partial differential equations are handled using Finite Element Methods (FEM), Finite Difference Methods (FDM) or Finite Volume Methods (FVM) on structured or unstructured meshes [BS02], [CL91], [GR96], [EGH00], [Krö97], [LeV92], [LeV02]. In these methods, a mesh is a covering of the computational domain consisting of pairwise disjoint polygons – the cells – which have to satisfy several geometrical and conformity constraints in order to ensure a good approximation of the solution. A main drawback of these methods when applied to problems with complicated or even time-dependent computational domains is the time consuming construction of such a mesh.

Therefore, meshless or particle schemes like the Smoothed Particle Hydrodynamics [GM77], [Mon94], [MK94], [Mon05], the Finite Pointset Method [Kuh99] or the Finite Volume Particle Method [HSS00], [Tel00], [Tel05], [TS08] have been designed to overcome these difficulties. These methods have in common that they use particles for the discretization of the equations instead of a mesh. These particles carry the information of interest and are usually allowed to move, i.e. they transport the information throughout the computational domain. Especially for time-dependent domains [TS08], fluid-structure interaction problems [TAH⁺07] or crack-propagation [MTM05], these methods seem to be promising. In this thesis, we will study the Finite Volume Particle Method (FVPM), a meshless method, that was constructed to combine the advantages of Finite Volume Methods and meshless particle schemes.

In FVPM, the particles are sufficiently smooth, compactly supported functions that cover the computational domain. Moreover, they have to build a partition of unity such that their supports have to overlap. The exchange of information between the particles is managed via this overlap: if the supports of two particles overlap, the particles are interpreted as neighbours and may interact. Moreover, the interaction is modelled using a numerical flux function. Because of the meshless character, the FVPM is very flexible and thus can be applied for example to problems on moving domains where conventional mesh generation

would be too complicated or not efficient. The general structure of a FVM provides an interpretation of the FVPM as a generalization of these classical mesh-based schemes and prevents the need of artificial viscosity due to the use of numerical flux functions. As will be shown, the coefficients of the FVPM can be seen as generalized surface areas and normal vectors. These geometrical coefficients have to satisfy certain conditions in order to get some of the classical results of FVM like stability, monotonicity, discrete entropy estimates or a Lax-Wendroff type consistency result.

The thesis is organized as follows: In Chapter 1, we start with some basics on conservation laws and their numerical treatment by FVM. In addition, we will give a brief review on the convergence analysis for FVM. In preparation for the numerical examples concerning the Savage-Hutter equations and because of their close relation to meshless methods, the idea of kinetic schemes will be mentioned at the end of the first chapter.

In Chapter 2, we will derive the FVPM and mention the main properties of the scheme in general and of the geometrical coefficients in particular. We will describe an adjustment of the scheme such that constant states will be preserved in the case of moving particles. Furthermore, we will give L^∞ - and L^1 -stability results as well as a weak BV-stability result and show a discrete entropy inequality. In Chapter 3, we provide a modification of the FVPM in one spatial dimension using B-splines. This new scheme has the advantage that geometrical coefficients can be computed exactly in a very short time. Especially for problems on moving domains, the coupling of the FVPM with B-splines represents an efficient alternative to the original FVPM. Chapter 4 deals with a kinetic FVM for the Savage-Hutter equations. In Section 4.3, the applicability of the FVPM to the kinetic framework will be shown. Numerical results will be presented in Chapter 5. Beyond the classical Riemann problem for the Euler equations, a linearized piston problem on a moving domain and the Savage-Hutter equations are considered. The thesis closes with a short conclusion, reviewing the main results and sketching some ideas how to further improve the method.

Chapter 1

Preliminaries

The following chapter gives an overview on hyperbolic conservation laws. In Section 1.1, we give a short introduction to the notion of adequate solutions of hyperbolic conservation laws, including the concept of weak solutions. The main results concerning existence and uniqueness of solutions will be mentioned together with a short overview on the relevant literature. Especially for scalar conservation laws, we will define entropy and entropy process solutions and give a short insight into the existing theory.

In Section 1.2, Finite Volume Methods (FVM) for hyperbolic conservation laws will be derived. As the existence of an entropy solution is often shown with the help of a convergent numerical scheme, the important result concerning existence, uniqueness and the convergence of a FVM will be presented in this section. Since we will adopt some of the results that ensure convergence of the FVM to the Finite Volume Particle Method (FVPM) in Chapter 2, the main steps of the proof of the convergence result are formulated at the end of Section 1.2.

In Section 1.3, our focus will be on kinetic schemes for hyperbolic conservation laws. This paragraph serves as an introduction to Chapter 4, where the FVPM will be coupled with a kinetic scheme that was developed to solve a system of balance equations for granular flow, namely the Savage-Hutter equations.

1.1 Hyperbolic conservation laws

In this section, we describe the main properties of hyperbolic conservation laws and their solutions. We begin with the definition of conservation laws. More details can be found in [Kr 97], [MNRR96], [EGH00], [GR96], [Bre05], [LeV92] or [LeV02].

Definition 1.1.1 (Conservation laws). *A system of partial differential equations of the form*

$$\begin{aligned}\partial_t \mathbf{u} + \nabla \cdot \mathbf{F}(\mathbf{u}) &= \mathbf{0} & \forall \mathbf{x} \in \mathbb{R}^d, t > 0 \\ \mathbf{u}(\mathbf{x}, 0) &= \mathbf{u}_0(\mathbf{x}) & \forall \mathbf{x} \in \mathbb{R}^d,\end{aligned}\tag{1.1}$$

where $\mathbf{u} : \mathbb{R}^d \times \mathbb{R}_+ \rightarrow \mathbb{R}^m$, $\mathbf{F} = (\mathbf{f}_1, \dots, \mathbf{f}_d) : \mathbb{R}^m \rightarrow \mathbb{R}^{m \times d}$, $\mathbf{u}_0 \in L^\infty(\mathbb{R}^d)$ and $\nabla \cdot \mathbf{F} := \sum_{j=1}^d \partial_{x_j} \mathbf{f}_j(\mathbf{u})$ is called a system of conservation laws in d space dimensions.

Many equations of practical interest are of hyperbolic type, e.g. the Euler equations of gas dynamics or the equations of magneto-hydrodynamics. Those systems are defined as follows.

Definition 1.1.2 (Hyperbolic systems of conservation laws). *A system of conservation laws (1.1) is called (strictly) hyperbolic if for any $\mathbf{u} \in \mathbb{R}^m$ and any $\mathbf{n} = (n_1, \dots, n_d)^T \in \mathbb{R}^d$, $\mathbf{n} \neq \mathbf{0}$, the matrix*

$$\mathbf{A}(\mathbf{u}, \mathbf{n}) = \sum_{j=1}^d n_j D\mathbf{f}_j(\mathbf{u}) \quad (1.2)$$

has only real (and distinct) eigenvalues $\lambda_1, \dots, \lambda_m$ and m linearly independent right eigenvectors $\mathbf{r}_1, \dots, \mathbf{r}_m$.

Hyperbolic systems can be characterized by the eigenvalues and corresponding eigenvectors of the Jacobian flux matrices $\mathbf{A}(\mathbf{u}, \mathbf{n})$ as information travels with a speed given by the eigenvalues. As an example of a hyperbolic system we consider the Euler equations of gas dynamics for compressible flow.

Example 1.1.3 (Euler equations of gas dynamics). *The Euler equations of gas dynamics are a system of hyperbolic conservation laws that describe the motion of compressible fluids. In two dimensions, the Euler equations read*

$$\partial_t \mathbf{u} + \nabla \cdot \mathbf{F}(\mathbf{u}) = \mathbf{0}, \quad (1.3)$$

where the vector of conserved quantities \mathbf{u} and the flux \mathbf{F} are given by

$$\mathbf{u} = \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ \rho E \end{pmatrix}, \quad \mathbf{f}_1 = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ u(\rho E + p) \end{pmatrix}, \quad \mathbf{f}_2 = \begin{pmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ v(\rho E + p) \end{pmatrix}, \quad \mathbf{F} = (\mathbf{f}_1, \mathbf{f}_2),$$

and ρ , u , v , E and p denote the density, the velocity in x - and y -direction, the specific total energy and the pressure of the fluid, respectively. To obtain a closed system, the Euler equations (1.3) are endowed with an equation of state and, of course, suitable initial and boundary conditions. For an ideal polytropic gas the equation of state reads

$$p = (\gamma - 1) \rho \left(E - \frac{u^2 + v^2}{2} \right)$$

where γ is the ratio of specific heat capacities at constant pressure c_p and at constant volume c_v . The eigenvalues of the matrix $\mathbf{A}(\mathbf{u}, \mathbf{n})$ for the Euler equations are given by

$$\lambda_1 = \mathbf{u} \cdot \mathbf{n} - c, \quad \lambda_2 = \mathbf{u} \cdot \mathbf{n}, \quad \lambda_3 = \mathbf{u} \cdot \mathbf{n} + c$$

and c denotes the speed of sound defined by

$$c = \sqrt{\frac{\gamma p}{\rho}}.$$

In contrast to solutions of elliptic or parabolic problems, solutions of hyperbolic equations may become discontinuous in finite time, even for arbitrarily smooth initial data. For that reason, the notion of a weak solution is introduced, which allows the solution to be discontinuous.

Definition 1.1.4 (Weak solution). *A function $\mathbf{u} \in L^\infty(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R}^m)$ is called a weak solution of the hyperbolic system (1.1) if*

$$\int_{\mathbb{R}^d \times \mathbb{R}_+} (\mathbf{u} \varphi_t + \mathbf{F}(\mathbf{u}) \nabla \varphi) dt d\mathbf{x} + \int_{\mathbb{R}^d} \mathbf{u}_0(\mathbf{x}) \varphi(\mathbf{x}, 0) d\mathbf{x} = 0 \quad (1.4)$$

holds for all test functions $\varphi \in C_0^1(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R})$.

Of course, smooth weak solutions $\mathbf{u} \in (C^1 \cap L^\infty)(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R}^m)$ are classical solutions.

The existence and uniqueness of weak solutions of general two- or multi-dimensional systems of hyperbolic conservation laws is still an open problem.

In the one-dimensional case, the existence of a weak solution of strictly hyperbolic systems of conservation laws with each characteristic field being either genuinely nonlinear or linearly degenerate was shown by Glimm in 1965 [Gli65] under the assumption that $\mathbf{u}_0 \in (L_{loc}^\infty \cap BV)(\mathbb{R}, \mathbb{R}^m)$. Here, the space $BV(\mathbb{R}^d, \mathbb{R}^m)$ denotes the space of all functions in L_{loc}^1 with bounded total variation:

$$BV(\mathbb{R}^d, \mathbb{R}^m) = BV(\mathbb{R}^d, \mathbb{R})^m = \{u \in L_{loc}^1(\mathbb{R}^d, \mathbb{R}) \mid TV(u) < \infty\}^m$$

with

$$TV(u) = \sup \left\{ \int_{\mathbb{R}^d} u(\mathbf{x}) \operatorname{div} \boldsymbol{\varphi}(\mathbf{x}) d\mathbf{x} \mid \boldsymbol{\varphi} \in C_0^\infty(\mathbb{R}^d, \mathbb{R}^d), \|\boldsymbol{\varphi}\|_\infty \leq 1 \right\}.$$

In 1995, Bressan was able to show uniqueness of that weak solution for one-dimensional systems in the class of BV solutions using the semigroup technique, compare [Bre05].

In the nonlinear scalar case, one can show that there exists at least one weak solution. Unfortunately, this weak solution is not unique in general. For that reason, so-called entropy solutions are considered to ensure the weak solution to be the physically correct one. This can be done with the help of the concept of vanishing viscosity, which was introduced by Lax in 1954 [Lax54]. The idea of this approach is to introduce a small diffusive term $\varepsilon \Delta \mathbf{u}$ in the conservation law. The resulting parabolic equation is known to admit a unique smooth solution. Then, the physically correct solution to the conservation law should coincide with the limit of the parabolic solution for $\varepsilon \rightarrow 0$. Although the concept of entropies exists also for systems of conservation laws [GR96], we will concentrate on scalar conservation laws in the following. We begin with the notion of entropy solutions.

Definition 1.1.5 (Entropy weak solution). *A function $u \in L^\infty(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R})$ is called an entropy weak solution of the hyperbolic conservation law (1.1) for $m = 1$ if*

$$\int_{\mathbb{R}^d \times \mathbb{R}_+} (\eta(u) \varphi_t + \boldsymbol{\Phi}(u) \nabla \varphi) dt d\mathbf{x} + \int_{\mathbb{R}^d} \eta(u_0(\mathbf{x})) \varphi(\mathbf{x}, 0) d\mathbf{x} \geq 0, \quad (1.5)$$

holds for all test functions $\varphi \in C_0^1(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R})$, all convex functions $\eta \in C^1(\mathbb{R}, \mathbb{R})$ and for all $\Phi \in C^1(\mathbb{R}, \mathbb{R}^d)$ such that $\nabla \cdot \Phi = \eta' \nabla \cdot \mathbf{F}$. The functions η and Φ are called entropy and entropy flux, respectively.

By a suitable choice of the entropy function η in Definition 1.1.5, it can be shown that every entropy solution is a weak solution.

The existence of a unique entropy solution in the multi-dimensional nonlinear scalar case was first shown by Vol'pert in [Vol67] assuming $u_0 \in BV(\mathbb{R}^d, \mathbb{R})$ and $\mathbf{F} \in C^1(\mathbb{R}, \mathbb{R}^d)$. This was generalized by Kruzkov [Kru70] to the case $\mathbf{F} = \mathbf{F}(\mathbf{x}, t, u(\mathbf{x}, t))$ and $u_0 \in L^\infty(\mathbb{R}^d, \mathbb{R})$ under the stronger assumption $\mathbf{F} \in C^3(\mathbb{R}^d \times \mathbb{R}_+ \times \mathbb{R}, \mathbb{R}^d)$. In fact, this result extends to scalar balance equations with a source term. Moreover, it was shown that the class of entropies and corresponding entropy fluxes can be restricted to $|u - \kappa|$ and $\text{sgn}(u - \kappa) [\mathbf{F}(\mathbf{x}, t, u) - \mathbf{F}(\mathbf{x}, t, \kappa)] \forall \kappa \in \mathbb{R}$, respectively. Therefore, we give the following equivalent definition of an entropy solution.

Definition 1.1.6 (Entropy weak solution due to Kruzkov). *A function $u \in L^\infty(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R})$ is called an entropy weak solution of the hyperbolic conservation law (1.1) for $m = 1$ if*

$$\begin{aligned} \int_{\mathbb{R}^d \times \mathbb{R}_+} |u(\mathbf{x}, t) - \kappa| \varphi_t + (\mathbf{F}(u(\mathbf{x}, t) \top \kappa) - \mathbf{F}(u(\mathbf{x}, t) \perp \kappa)) \nabla \varphi \, dt \, d\mathbf{x} \\ + \int_{\mathbb{R}^d} |u_0(\mathbf{x}) - \kappa| \varphi(\mathbf{x}, 0) \, d\mathbf{x} \geq 0 \end{aligned} \quad (1.6)$$

holds for all test functions $\varphi \in C_0^1(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R}_+)$ and for all $\kappa \in \mathbb{R}$. Here, for $a, b \in \mathbb{R}$ $a \top b = \max(a, b)$ and $a \perp b = \min(a, b)$.

Remark 1.1.7. *Although there are many results concerning the existence and uniqueness of solutions and the convergence of suitable numerical schemes for more general fluxes $\mathbf{F}(\mathbf{x}, t, u(\mathbf{x}, t))$, we will give definitions like the one above only for fluxes $\mathbf{F}(u(\mathbf{x}, t))$ as we are interested in solutions to (1.1).*

Remark 1.1.8. *Note that for $u \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^d$ and $t \in \mathbb{R}_+$, there holds*

$$\text{sgn}(u(\mathbf{x}, t) - \kappa) [\mathbf{F}(u(\mathbf{x}, t)) - \mathbf{F}(\kappa)] = \mathbf{F}(u(\mathbf{x}, t) \top \kappa) - \mathbf{F}(u(\mathbf{x}, t) \perp \kappa) \quad \forall \kappa \in \mathbb{R}.$$

In [EGH95], Eymard, Gallouët and Herbin showed existence and uniqueness of the entropy solution for $\mathbf{F}(\mathbf{x}, t, u(\mathbf{x}, t)) = \mathbf{v}(\mathbf{x}, t) f(u(\mathbf{x}, t))$ with $\mathbf{v} \in (L^\infty \cap C^1)(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R}^d)$ and $f \in C^1(\mathbb{R}, \mathbb{R})$ under the additional condition $\text{div}_x \mathbf{v}(\mathbf{x}, t) = 0$. This result was generalized by Chainais-Hillairet in [Cha99] for fluxes $\mathbf{F} = \mathbf{F}(\mathbf{x}, t, u(\mathbf{x}, t))$ with $\mathbf{F} \in C^1(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R}, \mathbb{R}^d)$ with $\text{div}_x \mathbf{F}(\mathbf{x}, t, u) = \sum_{i=1}^d \partial F_i(\mathbf{x}, t, u) / \partial x_i = 0$. For the proof, the authors in [EGH95] introduced the notion of entropy process solutions, which generalizes the concept of entropy solutions. The idea of entropy process solutions is closely related to the concept of measure valued solutions introduced by DiPerna [DiP85].

Definition 1.1.9 (Entropy process solution). *A function $u \in L^\infty(\mathbb{R}^d \times \mathbb{R}_+ \times (0, 1), \mathbb{R})$ is called an entropy process solution of the hyperbolic conservation law (1.1) for $m = 1$ if*

$$\begin{aligned} & \int_{\mathbb{R}^d \times \mathbb{R}_+ \times (0, 1)} |u(\mathbf{x}, t, \alpha) - \kappa| \varphi_t + (\mathbf{F}(u(\mathbf{x}, t, \alpha) \top \kappa) - \mathbf{F}(u(\mathbf{x}, t, \alpha) \perp \kappa)) \nabla \varphi \, d\alpha \, dt \, d\mathbf{x} \\ & + \int_{\mathbb{R}^d} |u_0(\mathbf{x}) - \kappa| \varphi(\mathbf{x}, 0) \, d\mathbf{x} \geq 0, \end{aligned} \quad (1.7)$$

holds for all test functions $\varphi \in C_0^1(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R}_+)$ and for all $\kappa \in \mathbb{R}$.

After the existence of such an entropy process solution could be shown, the authors in [EGH95] were able to proof that the solution does not depend on the additional parameter α and thus is an entropy solution. Since the proof of existence and uniqueness of such a solution relies on the convergence proof of suitable numerical schemes, we will give the main idea of the proof in the next section, where we will discuss the concept of Finite Volume Methods.

1.2 Finite Volume Methods

In this section, we will introduce Finite Volume Methods as it is done in [Kr97]. We consider the conservation law

$$\begin{aligned} \partial_t \mathbf{u} + \nabla \cdot \mathbf{F}(\mathbf{u}) &= 0 & \forall \mathbf{x} \in \mathbb{R}^d, \, t > 0 \\ \mathbf{u}(\mathbf{x}, 0) &= \mathbf{u}_0(\mathbf{x}) & \forall \mathbf{x} \in \mathbb{R}^d, \end{aligned} \quad (1.8)$$

where $\mathbf{u} : \mathbb{R}^d \times \mathbb{R}_+ \rightarrow \mathbb{R}^m$, $\mathbf{F} : \mathbb{R}^m \rightarrow \mathbb{R}^{m \times d}$ and $\mathbf{u}_0 \in L^\infty(\mathbb{R}^d)$. For simplicity, we assume $d = 2$ for the moment.

Definition 1.2.1 (Unstructured grid). *Let $\Omega \subset \mathbb{R}^2$, $I \subset \mathbb{N}$ an index set and $k \in \mathbb{N}$. Moreover, let a k -polygon be a closed polygon with k vertices. The set*

$$\mathcal{T}_h := \{T_i \mid T_i \subseteq \Omega, T_i \text{ is a } k\text{-polygon for } i \in I\}$$

is called an unstructured grid of Ω if

1. $\Omega = \cup_{i \in I} T_i$
2. for $i \neq j$, we have $T_i \cap T_j = \emptyset$ or $T_i \cap T_j$ is a common vertex of T_i and T_j or $T_i \cap T_j$ is a common edge of T_i and T_j .

The elements $T_i \in \mathcal{T}_h$ are called cells, the common edge between the cells T_i and T_j will be denoted by σ_{ij} . If the cells T_i and T_j have a common edge, T_j is called a neighbouring cell of T_i . The index set of all neighbouring cells of T_i will be denoted by $N(i)$, i.e.

$$N(i) := \{j \in I \mid T_j \text{ is a neighbouring cell of } T_i\}.$$

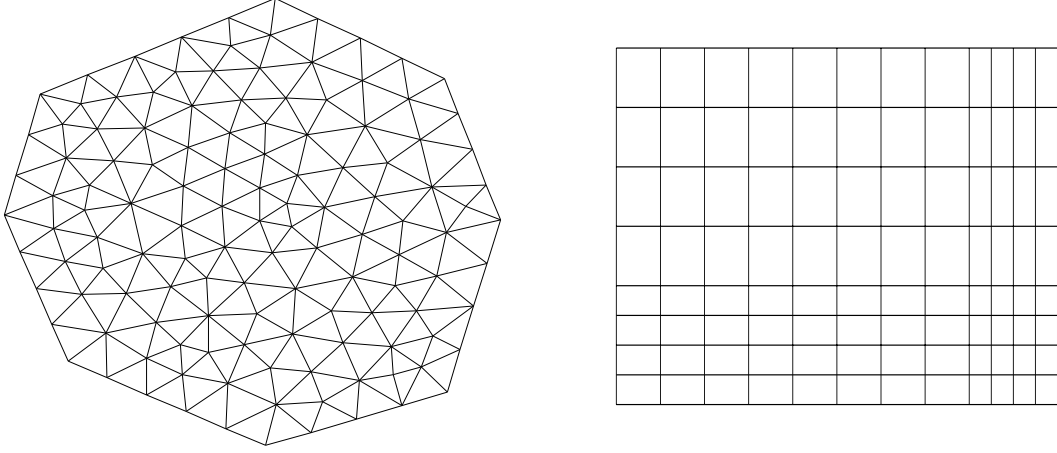


Figure 1.1: Typical grids using triangles (left) and rectangles (right) as cells.

The parameter h is defined as

$$h := \sup_{i \in I} \text{diam}(T_i)$$

and corresponds to the resolution of the grid.

Usually, a suitable grid has to satisfy some regularity conditions. For details see [Krö97], [Cha99] or [EGH00]. Typical grids for $k = 3$ and $k = 4$, respectively, are shown in Figure 1.1. The idea of FVM consists in approximating the solution $\mathbf{u}(\cdot, t)$ of (1.8) on each cell $T_i \in \mathcal{T}_h$ by the mean value of $\mathbf{u}(\cdot, t)$ over that cell:

$$\mathbf{u}(\mathbf{x}, t) \approx \mathbf{u}_i(t) := \frac{1}{|T_i|} \int_{T_i} \mathbf{u}(\mathbf{x}, t) \, d\mathbf{x} \quad \text{for } \mathbf{x} \in T_i,$$

where $|T_i|$ denotes the volume of cell T_i . If we now integrate Equation (1.8) over cell T_i , we get

$$\begin{aligned} 0 &= \frac{d}{dt} \int_{T_i} \mathbf{u}(\mathbf{x}, t) \, d\mathbf{x} + \int_{\partial T_i} \mathbf{F}(\mathbf{u}(s, t)) \mathbf{n}_i \, ds \\ &\approx |T_i| \frac{d\mathbf{u}_i(t)}{dt} + \sum_{j \in \mathcal{N}(i)} \int_{\sigma_{ij}} \mathbf{F}(\mathbf{u}(s, t)) \mathbf{n}_{ij} \, ds \\ &\approx |T_i| \frac{d\mathbf{u}_i(t)}{dt} + \sum_{j \in \mathcal{N}(i)} |\sigma_{ij}| \mathbf{g}(\mathbf{u}_i(t), \mathbf{u}_j(t), \mathbf{n}_{ij}) \end{aligned}$$

where σ_{ij} denotes the edge between the i -th and j -th k -polygon, $|\sigma_{ij}|$ describes its length and \mathbf{n}_{ij} denotes the outer normal vector to σ_{ij} . The function \mathbf{g} is a numerical flux function approximating the flux on the boundary σ_{ij} . Thus, we have the following system of ordinary differential equations for the time-dependent functions $\mathbf{u}_i(t)$:

$$\frac{d\mathbf{u}_i(t)}{dt} = -\frac{1}{|T_i|} \sum_{j \in \mathcal{N}(i)} |\sigma_{ij}| \mathbf{g}(\mathbf{u}_i(t), \mathbf{u}_j(t), \mathbf{n}_{ij}) \quad (1.9)$$

together with the initial condition

$$\mathbf{u}_i(0) = \frac{1}{|T_i|} \int_{T_i} \mathbf{u}(\mathbf{x}, 0) \, d\mathbf{x}. \quad (1.10)$$

Equations (1.9)–(1.10) can be discretized in time using a simple forward Euler discretization with $t^n := n \Delta t$ and time steps Δt leading to the fully discretized explicit FVM

$$\begin{aligned} \mathbf{u}_i^{n+1} &= \mathbf{u}_i^n - \frac{\Delta t}{|T_i|} \sum_{j \in \mathcal{N}(i)} |\sigma_{ij}| \mathbf{g}(\mathbf{u}_i, \mathbf{u}_j, \mathbf{n}_{ij}), \\ \mathbf{u}_i^0 &= \frac{1}{|T_i|} \int_{T_i} \mathbf{u}(\mathbf{x}, 0) \, d\mathbf{x}. \end{aligned} \quad (1.11)$$

The numerical solution $\mathbf{u}_h(\mathbf{x}, t)$ is then given by a suitable reconstruction formula. Here, we assumed a piecewise constant solution, so we simply set

$$\mathbf{u}_h(\mathbf{x}, t) := \mathbf{u}_i^n \text{ for } (\mathbf{x}, t) \in T_i \times [t^n, t^{n+1}), \quad (1.12)$$

but of course other reconstruction techniques are possible, e.g. [Son97].

In order to ensure that the numerical solution \mathbf{u}_h satisfies properties that hold for the exact solution, several conditions on the numerical flux function \mathbf{g} , the time step Δt and the data are to be set. The most important conditions are the following.

- **Conservativity:** Equations (1.8) describe the evolution of conserved quantities, i.e. mass is neither produced nor destroyed. To ensure the conservation numerically, the numerical flux function has to guarantee that the amount of material moving from cell T_i to cell T_j is exactly the same as the amount arriving in cell T_j coming from cell T_i , thus it should hold

$$\mathbf{g}(\mathbf{u}_i, \mathbf{u}_j, \mathbf{n}_{ij}) = -\mathbf{g}(\mathbf{u}_j, \mathbf{u}_i, -\mathbf{n}_{ij}).$$

- **Consistency:** To ensure the preservation of constant states, the numerical flux should satisfy

$$\mathbf{g}(\mathbf{u}, \mathbf{u}, \mathbf{n}) = \mathbf{F}(\mathbf{u}) \mathbf{n}$$

as can be seen immediately from the approximation

$$\int_{\sigma_{ij}} \mathbf{F}(\mathbf{u}(s, t)) \mathbf{n}_{ij} \, ds \approx |\sigma_{ij}| \mathbf{g}(\mathbf{u}_i(t), \mathbf{u}_j(t), \mathbf{n}_{ij}).$$

- **CFL condition:** Usually, the scheme (1.11) is only stable in some suitable sense, if the time step size Δt is sufficiently small and the upper bound depends on the mesh parameter h . Such a condition is called CFL condition, named after Courant, Friedrichs and Lewy.

As stated in the last section, existence and uniqueness of the entropy solution and a convergence result for a suitable Finite Volume Method are in general closely related. We will now give the main ideas for the proof of existence and uniqueness of the entropy solution as well as the convergence of a Finite Volume Method towards this entropy solution in the multi-dimensional scalar case as it was done in [EGH95] and [Cha99]. We begin with the main result in [Cha99].

Theorem 1.2.2. *Consider the scalar case $u : \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R}$ and assume that $u_0 \in L^\infty(\mathbb{R}^d, \mathbb{R})$, $\mathbf{F} \in C^1(\mathbb{R}^d \times \mathbb{R}_+ \times \mathbb{R}, \mathbb{R}^d)$, $\partial_u \mathbf{F}$ being locally Lipschitz continuous and $\sum_{i=1}^d \frac{\partial f_i}{\partial x_i}(\mathbf{x}, t, u) = 0$. Assume further that for all compact sets $K \subset \mathbb{R}$, there exists a constant $C_K < \infty$ such that $|\partial_u \mathbf{F}(\mathbf{x}, t, u)| \leq C_K$ for almost every $(\mathbf{x}, t, u) \in \mathbb{R}^d \times \mathbb{R}_+ \times K$. Then the scalar equation*

$$\begin{aligned} \partial_t u + \nabla \cdot \mathbf{F}(\mathbf{x}, t, u) &= 0 & \forall \mathbf{x} \in \mathbb{R}^d, t > 0 \\ u(\mathbf{x}, 0) &= u_0(\mathbf{x}) & \forall \mathbf{x} \in \mathbb{R}^d \end{aligned}$$

admits a unique entropy solution. Consider moreover the numerical scheme (1.11) with a consistent, conservative, Lipschitz continuous and monotone numerical flux function and a suitable CFL condition on a sufficiently regular grid, see [Cha99]. Then the approximate solution constructed by (1.11)-(1.12) converges towards the entropy solution in $L_{loc}^p(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R})$ for $1 \leq p < \infty$.

An important tool in the proof is the notion of the nonlinear weak-* convergence introduced in [EGH95].

Definition 1.2.3 (Nonlinear weak-* convergence). *Let Ω be an open set in \mathbb{R}^d and $(u_n)_{n \in \mathbb{N}}$ be a sequence in $L^\infty(\Omega)$. The sequence $(u_n)_{n \in \mathbb{N}}$ is said to converge in a nonlinear weak-* sense, if there exists a function $u \in L^\infty(\Omega \times (0, 1))$ such that*

$$\int_{\Omega} h(u_n(\mathbf{x})) \phi(\mathbf{x}) d\mathbf{x} \rightarrow \int_{\Omega} \int_0^1 h(u(\mathbf{x}, \alpha)) \phi(\mathbf{x}) d\alpha d\mathbf{x} \quad (1.13)$$

for all $\phi \in L^1(\Omega)$, and $h \in C(\mathbb{R}, \mathbb{R})$. We will use the abbreviation $u_n \xrightarrow{n.w} u$.*

Note that the nonlinear weak-* convergence is equivalent to the weak-* convergence of $h(u_n)$ towards a function U_h with

$$U_h(\mathbf{x}) := \int_0^1 h(u(\mathbf{x}, \alpha)) d\alpha.$$

The following result for bounded sequences in $L^\infty(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R})$ of Eymard, Gallouët and Herbin in [EGH95] gives rise to the notion of the entropy process solution.

Proposition 1.2.4. *Let $(u_n)_{n \in \mathbb{N}}$ be a bounded sequence in $L^\infty(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R})$. Then there exists a subsequence $(u_{n_k})_{k \in \mathbb{N}}$ and a function $u \in L^\infty(\mathbb{R}^d \times \mathbb{R}_+ \times (0, 1), \mathbb{R})$ such that $u_{n_k} \xrightarrow{n.w*} u$. If additionally, u does not depend on the parameter α , then $(u_{n_k})_{k \in \mathbb{N}}$ converges strongly in $L_{loc}^p(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R})$ for all $1 \leq p < \infty$.*

The main steps in the proof of Theorem 1.2.2 are the following: First, it is shown that the numerical solution is stable in L^∞ and in some weak BV sense, meaning that estimates on the discrete time and space derivatives of the approximate solution are of order $h^{-1/2}$. Subsequently, one can derive a discrete entropy inequality for the coefficients u_i^n and an entropy inequality for the reconstructed solution. Then, the nonlinear weak-* convergence, which follows from the L^∞ -stability, allows to pass to the limit in the entropy inequality for the approximate solution. The occurring error terms in the entropy inequality for the reconstruction can be controlled with the help of the weak BV-stability result. Thus, the nonlinear weak-* limit of the subsequence $(u_{n_k})_{k \in \mathbb{N}}$ is an entropy process solution. Afterwards, the authors in [EGH95] and [Cha99] proved the uniqueness and the independence of α of the entropy process solution, i.e. the entropy process solution coincides with the unique entropy solution. Applying the above proposition, strong convergence in $L^p_{loc}(\mathbb{R}^d \times \mathbb{R}_+)$ follows immediately.

At the end of Chapter 2, we will adopt some of the results used in the above mentioned convergence theory in [EGH95] and [Cha99]. In fact, we will show stability in L^∞ and in a weak BV-sense for the FVPM as well as in L^1 at least for special cases. Moreover, we will proof a discrete entropy inequality that can be seen as a generalization of the well-known discrete entropy inequality for FVM.

In the next section, we give an introduction to kinetic schemes. Shortly speaking, kinetic schemes are FVM or FDM whose numerical flux function is the moment of a numerical flux function for a microscopic version of the equation under consideration. The idea and the principle of the construction of kinetic schemes are important for Chapter 4, where we combine the FVPM with a kinetic scheme for the Savage-Hutter equations.

1.3 Kinetic schemes

When considering a dilute gas, the mean free path can no longer be neglected. Instead of the equations for continuum mechanics, one has to regard the kinetic theory of gases. In this section, we will give a very brief review on the fundamental principles of the kinetic theory of dilute gases and the resulting numerical schemes. Since we want to apply the FVPM to an existing kinetic scheme for the Savage-Hutter equations (see Chapter 4), these basics will be important for our further proceeding. For details on the kinetic theory of dilute gases and kinetic numerical schemes see for example [Cer90], [Cer94], [Per02], [LPT94], [KNS96] or [Jun00].

We begin with a system that contains N atoms or particles of a monatomic gas. The atoms are assumed to be hard balls that collide elastically. For the description of the whole system, we consider the phase space that consists of the positions $\mathbf{x}_i \in \Omega \subset \mathbb{R}^3$ and the velocities $\boldsymbol{\xi}_i \in \mathbb{R}^3$, $i = 1, \dots, N$, of all N particles. Thus, the phase space \mathfrak{D} has dimension $6N$ and N usually will be of order 10^{20} [Cer90]. Since we have to assume that the dynamics of the

atoms are coupled, we will end up with a very large system of differential equations for the appropriate description of the dynamics of the original system. Therefore, the introduction of a probability density $f : \mathfrak{D} \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ for the description of the whole system seems to be reasonable. The corresponding equation one ends up with is the Liouville equation

$$\frac{\partial f}{\partial t} + \sum_{i=1}^N \boldsymbol{\xi}_i \frac{\partial f}{\partial \mathbf{x}_i} + \sum_{i=1}^N \mathbf{X}_i \frac{\partial f}{\partial \boldsymbol{\xi}_i} = 0 \quad (1.14)$$

where \mathbf{X}_i denotes the force acting on particle i . If we consider an ideal gas, we can neglect the intermolecular forces between the atoms unless their distance is smaller than a characteristic distance, the so-called molecular diameter σ . Thus, outside this region, Equation (1.14) holds with $\mathbf{X}_i = \mathbf{0}$, $i = 1, \dots, N$, and the collision terms are treated as suitable boundary terms. If we consider the density $f_1(\mathbf{x}, \boldsymbol{\xi}, t)$ for just one particle in a set of N particles, we will end up with the Boltzmann equation

$$\frac{\partial f_1^\varepsilon}{\partial t} + \boldsymbol{\xi} \nabla_{\mathbf{x}} f_1^\varepsilon = \frac{Q(f_1^\varepsilon, f_1^\varepsilon)}{\varepsilon}, \quad (1.15)$$

where Q is the so-called collision operator and ε denotes the Knudsen number which measures the mean free path of the particles between their collision with other particles. The collision operator takes the form

$$Q(f, g) := \alpha \int_{\mathbb{R}^3} \int_{S_+} (f' g'_* + g' f'_* - f g_* - g f_*) |(\boldsymbol{\xi} - \boldsymbol{\xi}_*) \cdot \mathbf{n}| d\boldsymbol{\xi}_* d\mathbf{n},$$

where

$$f' := f(\mathbf{x}, \boldsymbol{\xi}', t), \quad f_* := f(\mathbf{x}, \boldsymbol{\xi}_*, t), \quad f'_* := f(\mathbf{x}, \boldsymbol{\xi}'_*, t),$$

$\boldsymbol{\xi}'$ denotes the velocity of the particle after the collision, the star is used to distinguish the two colliding particles and α is a constant. If we consider the collision operator, we will automatically ask for functions f for which $Q(f, f) = 0$ holds. It can be shown that

$$Q(f, f) = 0 \quad (1.16)$$

$$\Leftrightarrow f(\mathbf{x}, \boldsymbol{\xi}, t) = \exp(a(\mathbf{x}, t) + \mathbf{b}(\mathbf{x}, t) \cdot \boldsymbol{\xi} + c(\mathbf{x}, t) |\boldsymbol{\xi}|^2) \quad (1.17)$$

with $a, c : \mathbb{R}^3 \times \mathbb{R}_+ \rightarrow \mathbb{R}$ and $\mathbf{b} : \mathbb{R}^3 \times \mathbb{R}_+ \rightarrow \mathbb{R}^3$. Density functions of the form (1.17) are called Maxwell distributions or Maxwellians and because of (1.16) equilibrium distributions. Usually, the Maxwell distribution is given in the form

$$f_{\text{Max}}(\mathbf{x}, \boldsymbol{\xi}, t) = A \exp(-\alpha (\boldsymbol{\xi} - \mathbf{v})^2)$$

where A, α and \mathbf{v} are constant in $\boldsymbol{\xi}$. This form is much more descriptive, because the constant \mathbf{v} is in fact the macroscopic velocity in the Euler equations. In the limit $\varepsilon \rightarrow 0$, it can be shown that the collision operator satisfies $Q(f, f) = 0$ in the lowest order. Thus, f can be assumed to be a Maxwellian.

To link the microscopic approach to the macroscopic one, we define the macroscopic values as the moments of the kinetic distribution function f :

$$\rho := \int_{\mathbb{R}^3} f d\boldsymbol{\xi}, \quad \rho \mathbf{v} := \int_{\mathbb{R}^3} \boldsymbol{\xi} f d\boldsymbol{\xi}, \quad \rho E := \int_{\mathbb{R}^3} \frac{|\boldsymbol{\xi}|^2}{2} f d\boldsymbol{\xi}. \quad (1.18)$$

Let us now assume that the kinetic density f is the Maxwellian

$$\mathcal{M}(\rho, \mathbf{v}, T; \mathbf{x}, \boldsymbol{\xi}, t) := \frac{\rho}{(2\pi T)^{\frac{d}{2}}} \exp\left(-\frac{|\boldsymbol{\xi} - \mathbf{v}|^2}{2T}\right)$$

where the temperature T is related to the energy by $T = 2/3(E - |\mathbf{v}|^2/2)$. Multiplying the Boltzmann equation with the collision invariants 1, $\boldsymbol{\xi}$ and $|\boldsymbol{\xi}|^2/2$ and integrating over $\boldsymbol{\xi}$, we will end up with the Euler equations (1.3). Thus, we have found a new formulation for the Euler equations:

$$\frac{\partial f}{\partial t} + \boldsymbol{\xi} \nabla_{\mathbf{x}} f = \bar{Q}, \quad f = \mathcal{M}, \quad (1.19)$$

with $Q(f, f)/\varepsilon \rightarrow \bar{Q}$ for $\varepsilon \rightarrow 0$. In the following, we will call (1.19) the kinetic formulation of the Euler equations. Note that the advantage of the kinetic formulation lies in the linearity of the scalar operator $(\partial/\partial t + \boldsymbol{\xi} \nabla_{\mathbf{x}})$. The nonlinearity of Equation (1.3) is completely shifted to the right hand side and the additional constraint $f = \mathcal{M}$ to ensure that the resulting f is always a Maxwellian. Note also that spatial homogeneous solutions of the Boltzmann equation imply stationary solutions of the corresponding macroscopic equations, because the divergence terms vanish.

Let us now consider Equation (1.1). Following [Jun00], we will call the following system a kinetic formulation of system (1.1):

$$\left(\frac{\partial}{\partial t} + \boldsymbol{\xi} \nabla_{\mathbf{x}}\right) \mathbf{f} = \mathbf{Q}, \quad \mathbf{f}(\mathbf{x}, t, \boldsymbol{\xi}) = \boldsymbol{\mu}(\mathbf{u}(\mathbf{x}, t), \boldsymbol{\xi}), \quad \int_{\mathbb{R}^d} \mathbf{Q} \, d\boldsymbol{\xi} = 0 \quad (1.20)$$

for a constraint function $\boldsymbol{\mu} : \mathbb{R}^m \times \mathbb{R}^d \rightarrow \mathbb{R}^m$ that satisfies

$$\int_{\mathbb{R}^d} \boldsymbol{\mu}(\mathbf{u}, \boldsymbol{\xi}) \, d\boldsymbol{\xi} = \mathbf{u}, \quad \int_{\mathbb{R}^d} \xi_j \boldsymbol{\mu}(\mathbf{u}, \boldsymbol{\xi}) \, d\boldsymbol{\xi} = \mathbf{f}_j(\mathbf{u}), \quad j = 1, \dots, d. \quad (1.21)$$

In this formulation, $\boldsymbol{\mu}$ with its constraints (1.21) describes a class of so-called equilibrium densities. For the two-dimensional Euler equations, we have

$$\mathbf{u} = \begin{pmatrix} \rho \\ \rho \mathbf{v} \\ \rho E \end{pmatrix}, \quad \boldsymbol{\mu}(\mathbf{u}, \boldsymbol{\xi}) = \begin{pmatrix} 1 \\ \boldsymbol{\xi} \\ \frac{1}{2}|\boldsymbol{\xi}|^2 \end{pmatrix} \mathcal{M}(\rho, \mathbf{u}, T; \boldsymbol{\xi}), \quad \mathbf{Q} = \begin{pmatrix} 1 \\ \boldsymbol{\xi} \\ \frac{1}{2}|\boldsymbol{\xi}|^2 \end{pmatrix} \bar{Q}.$$

Note that in the kinetic formulation (1.20), (1.21), the unknown \mathbf{f} and the collision operator are no longer scalar functions. Of course, for a given system of conservation laws, we can bring (1.20), (1.21) in a scalar form. For example, for a 2×2 system in one spatial dimension, we have

$$\begin{aligned} \frac{\partial f}{\partial t} + \xi \partial_x f &= \bar{Q}, \quad \int_{\mathbb{R}} \bar{Q} \, d\xi = \int_{\mathbb{R}} \xi \bar{Q} \, d\xi = 0, \\ u_1 &= \int_{\mathbb{R}} f \, d\xi, \quad u_2 = \int_{\mathbb{R}} \xi f \, d\xi, \quad \mathbf{u} = (u_1, u_2)^T. \end{aligned}$$

To get a numerical method for the original equations, we can now formulate a numerical scheme for the kinetic formulation, multiply it with 1, $\boldsymbol{\xi}$ and $|\boldsymbol{\xi}|^2/2$ and integrate over $\boldsymbol{\xi}$. A

numerical scheme that originates in a numerical method of the kinetic formulation will be referred to as a kinetic scheme in the following.

We will now sketch the construction of a semi-discrete kinetic scheme. Note that if we neglect the collision operator in (1.20), we have a simple linear advection equation for \mathbf{f} :

$$\frac{\partial \mathbf{f}}{\partial t} + \boldsymbol{\xi} \nabla_{\mathbf{x}} \mathbf{f} = 0. \quad (1.22)$$

A fundamental idea in the construction of kinetic schemes is to enforce the constraint $\mathbf{f}(\mathbf{x}, t, \boldsymbol{\xi}) = \boldsymbol{\mu}(\mathbf{u}(\mathbf{x}, t), \boldsymbol{\xi})$ only at the discrete points $t^n = n \Delta t$, where Δt denotes the time step size and $n = 1, \dots, N_T$ for some $N_T \in \mathbb{N}$. Given $\mathbf{u}^n(\mathbf{x})$, we start with $\mathbf{f}(\mathbf{x}, t^n, \boldsymbol{\xi}) = \boldsymbol{\mu}(\mathbf{u}^n(\mathbf{x}), \boldsymbol{\xi})$ and compute an update of \mathbf{f} according to the advection Equation (1.22) for $t^n \leq t \leq t^{n+1}$:

$$\mathbf{f}(\mathbf{x}, t, \boldsymbol{\xi}) = \boldsymbol{\mu}(\mathbf{u}^n(\mathbf{x} - (t - t^n)\boldsymbol{\xi}), \boldsymbol{\xi})$$

But, by neglecting the collision operator, we will leave the class of equilibrium densities in general. Therefore, we set

$$\mathbf{u}^{n+1}(\mathbf{x}) := \int_{\mathbb{R}^d} \boldsymbol{\mu}(\mathbf{u}^n(\mathbf{x} - \Delta t \boldsymbol{\xi}), \boldsymbol{\xi}) \, \mathrm{d}\boldsymbol{\xi}$$

and get automatically an equilibrium density $\mathbf{f}(\mathbf{x}, t^{n+1}, \boldsymbol{\xi}) := \boldsymbol{\mu}(\mathbf{u}^{n+1}(\mathbf{x}), \boldsymbol{\xi})$.

In practice, the solution of the advection equation can be realized by a suitable discretization, e.g. by a FDM, a FVM or a FEM. Then, the discrete evolution for the linear advection equation is integrated over $\boldsymbol{\xi}$ to get a scheme for the original nonlinear equation. In the case of a FVM, the discretization of the advection part is managed by a numerical flux function. Thus, after integration, the kinetic Ansatz yields a FVM for the nonlinear macroscopic equations where the numerical flux function is nothing but the integral of the microscopic flux function.

In Chapter 4, we will describe a kinetic numerical scheme for the Savage-Hutter (SH) equations in one space dimension. The SH equations are a system of two equations for the height and the velocity of a granular mass and contain a source term. Concerning the SH equations, a main difficulty relies in the preservation of nontrivial constant states that exists due to the source term. In [KS08], a kinetic scheme was developed that is able to describe the complete dynamics of a granular mass and that preserves granular masses at rest. In Chapter 4, we will combine this scheme with the FVPM.

Chapter 2

The Finite Volume Particle Method

Classical methods like the FEM or the FVM discussed in the last section are not always applicable to complex problems like extrusion and molding processes, fluid structure interaction problems or propagation of cracks, where the computational domain and thus the mesh deforms significantly. The need for expensive mesh generation or remeshing procedures makes conventional schemes time-consuming and thus ineffective.

To overcome these difficulties, so-called meshless methods or particle methods grew up. Instead of using a mesh, these methods employ positive and compactly supported functions, whose supports cover the computational domain. These functions are called particles and carry all the information that is usually provided by a mesh.

Well-known representatives in the class of meshless methods are the Smoothed Particle Hydrodynamics (SPH), the Finite Pointset Method (FPM) and the Finite Volume Particle Method (FVPM). Before we go into detail about the FVPM, we will give a short survey on SPH and FPM. A general overview on meshless methods as well as further examples can be found in [DO95], [BKO⁺96], [KNS96], [Str97], [FM04], [GS05], [LL07] and [GS07].

SPH was originally developed for applications to astrophysical problems and was first mentioned in 1977 by Lucy [Luc77] and independently by Gingold and Monaghan [GM77]. The scheme is a Lagrangian particle method in which a material, e.g. a fluid, is supposed to consist of small particles which move with the fluid velocity. These particles build a set of interpolation points in reconstruction formulas. Several improvements and extensions of the method followed and turned the method into a powerful and inspiring instrument for the CFD community, e.g. [Mon94] and [Mon05].

In FPM, approximations of spatial derivatives of a smooth and scalar valued function f are based on the least squares method. For this purpose, Taylor series expansions of f are made in the positions of the particles and interpreted as linear systems for the derivatives. For details see for example [Kuh99], [TAH⁺07] or [GS07]. An extension of the FPM for the incompressible Navier-Stokes equations can be found in [TAH⁺07].

In this chapter, we begin with the derivation of the FVPM, as it has been done for example in [JS01] or [Tel05]. Section 2.2 deals with general and mostly well-known properties of the scheme. At the beginning of that section, we are concerned with the approximation quality of the particle cell averages. Next, we describe some properties of the geometrical coefficients,

which are of prime importance for the FVPM, as they ensure for example conservativity. As a new result, we show how to implement the velocity of the particles into the numerical flux function such that the scheme preserves constant states. At the end of Section 2.2, we describe two well-established correction procedures and show how conservativity of the scheme, which will be destroyed by one of the correction methods, can be regained using a special particle movement.

In Section 2.3, we show some of our central results, namely L^∞ -stability and the positivity of the scheme. For non-moving particles, we give a result which we denote as weak BV-estimates and show L^1 -stability under suitable assumptions. At the end, we are able to prove a discrete in-cell entropy inequality. All results in Section 2.3 are essential tools in the convergence analysis for FVM. To our knowledge, they have not been shown yet in this form and thus can be seen as first steps of a convergence proof of the FVPM.

2.1 Derivation of the FVPM

We consider the problem

$$\begin{aligned} \partial_t \mathbf{u} + \nabla \cdot \mathbf{F}(\mathbf{u}) &= \mathbf{0} & \forall \mathbf{x} \in \Omega(t) \subset \mathbb{R}^d, \ t > 0 \\ \mathbf{u}(\mathbf{x}, 0) &= \mathbf{u}_0(\mathbf{x}) & \forall \mathbf{x} \in \Omega(0) \end{aligned} \quad (2.1)$$

together with some appropriate boundary conditions. Here, $\mathbf{u}(\mathbf{x}, t) \in \mathbb{R}^m$ denotes the vector of unknowns of the conservation law (2.1). The computational domain $\Omega(t)$ is bounded and varies in time, i.e. the temporal variation of $\Omega(t)$ can be described with the help of a function denoting the velocity of the boundary $\Gamma(t)$ of $\Omega(t)$:

$$\Gamma(t) = \{x(t) : x(t) = x_0 + \int_0^t \mathbf{v}(x(s), s) \, ds, \ x_0 \in \Gamma_0\}$$

with a given continuous velocity field \mathbf{v} .

Instead of introducing a grid structure on the spatial domain $\Omega(t)$, we consider N functions $\psi_i : \mathbb{R}^d \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$, $i = 1, \dots, N$, associated with positions $\mathbf{x}_i(t) \in \mathbb{R}^d$. The functions ψ_i will be called particles, the positions \mathbf{x}_i the corresponding particle positions. The particles have to satisfy $0 \leq \psi_i(\mathbf{x}, t) \leq 1$ and build a partition of unity, that is

$$\sum_{i=1}^N \psi_i(\mathbf{x}, t) = 1, \quad \forall \mathbf{x} \in \Omega(t), \ t \geq 0.$$

For the construction of such a partition of unity, we start with a Lipschitz continuous and compactly supported function $W : \mathbb{R}^d \rightarrow \mathbb{R}_+$, e.g.

$$W(\mathbf{x}, h) = \begin{cases} (-\|\mathbf{x}\|^2 + \frac{h^2}{2}), & \|\mathbf{x}\| < \frac{h}{2} \\ (\|\mathbf{x}\| - h)^2, & \|\mathbf{x}\| \in [\frac{h}{2}, h) \\ 0, & \text{otherwise.} \end{cases} \quad (2.2)$$

The value h is the so-called smoothing length and $\|\cdot\|$ denotes the Euclidean norm in \mathbb{R}^d .

According to Shepard's renormalization method [She68], the family $\{\psi_i, i = 1, \dots, N\}$, where

$$\psi_i(\mathbf{x}, t) := \frac{W_i(\mathbf{x}, t)}{\sigma(\mathbf{x}, t)},$$

$$W_i(\mathbf{x}, t) := W(\mathbf{x} - \mathbf{x}_i(t), h)$$

and

$$\sigma(\mathbf{x}, t) := \sum_{i=1}^N W_i(\mathbf{x}, t),$$

constitutes a partition of unity.

Now, we want to test Equation (2.1) against the test functions ψ_i . With

$$\int_{\Omega(t)} (\partial_t \mathbf{u} + \nabla \cdot \mathbf{F}(\mathbf{u})) \psi_i \, d\mathbf{x} = \mathbf{0}, \quad i = 1, \dots, N$$

and integration by parts, we obtain

$$\begin{aligned} \frac{d}{dt} \int_{\Omega(t)} \mathbf{u} \psi_i \, d\mathbf{x} &= \int_{\Omega(t)} \partial_t (\mathbf{u} \psi_i) \, d\mathbf{x} + \int_{\partial\Omega(t)} ((\mathbf{u} \psi_i) \cdot \mathbf{v}) \, \mathbf{n} \, ds \\ &= \int_{\Omega(t)} (\partial_t \mathbf{u}) \psi_i \, d\mathbf{x} + \int_{\Omega(t)} \mathbf{u} (\partial_t \psi_i) \, d\mathbf{x} + \int_{\partial\Omega(t)} ((\mathbf{u} \psi_i) \cdot \mathbf{v}) \, \mathbf{n} \, ds \\ &= - \int_{\Omega(t)} (\nabla \cdot \mathbf{F}(\mathbf{u})) \psi_i \, d\mathbf{x} + \int_{\Omega(t)} \mathbf{u} (\partial_t \psi_i) \, d\mathbf{x} + \int_{\partial\Omega(t)} ((\mathbf{u} \psi_i) \cdot \mathbf{v}) \, \mathbf{n} \, ds \\ &= \int_{\Omega(t)} (\mathbf{F}(\mathbf{u}) \nabla \psi_i + \mathbf{u} \partial_t \psi_i) \, d\mathbf{x} - \int_{\partial\Omega(t)} \psi_i (\mathbf{F}(\mathbf{u}) - \mathbf{u} \cdot \mathbf{v}) \, \mathbf{n} \, ds. \end{aligned}$$

Note that, here and in the following, the product \cdot denotes the outer product that maps from $\mathbb{R}^m \times \mathbb{R}^d$ to $\mathbb{R}^{m \times d}$.

Some easy calculations (for the proof see [JS01]) verify the following

Proposition 2.1.1.

$$\partial_t \psi_i(\mathbf{x}, t) = \sum_{j=1}^N \left(\dot{\mathbf{x}}_j \psi_j(\mathbf{x}, t) \frac{\nabla W_j(\mathbf{x}, t)}{\sigma(\mathbf{x}, t)} - \dot{\mathbf{x}}_i \psi_j(\mathbf{x}, t) \frac{\nabla W_i(\mathbf{x}, t)}{\sigma(\mathbf{x}, t)} \right), \quad (2.3)$$

$$\nabla \psi_i(\mathbf{x}, t) = \sum_{j=1}^N \left(\psi_j(\mathbf{x}, t) \frac{\nabla W_i(\mathbf{x}, t)}{\sigma(\mathbf{x}, t)} - \psi_i(\mathbf{x}, t) \frac{\nabla W_j(\mathbf{x}, t)}{\sigma(\mathbf{x}, t)} \right). \quad (2.4)$$

In the next step, we define local averages $\mathbf{u}_i(t)$ of the function $\mathbf{u}(\mathbf{x}, t)$ as

$$\mathbf{u}_i(t) := \frac{1}{V_i(t)} \int_{\Omega(t)} \mathbf{u}(\mathbf{x}, t) \psi_i(\mathbf{x}, t) \, d\mathbf{x}, \quad (2.5)$$

where

$$V_i(t) := \int_{\Omega(t)} \psi_i(\mathbf{x}, t) \, d\mathbf{x}. \quad (2.6)$$

With

$$\mathbf{\Gamma}_{ij}(\mathbf{x}, t) := \frac{\psi_i(\mathbf{x}, t)}{\sigma(\mathbf{x}, t)} \nabla W_j(\mathbf{x}, t), \quad (2.7)$$

we can now write

$$\partial_t \psi_i(\mathbf{x}, t) = - \sum_{j=1}^N \left(\dot{\mathbf{x}}_i \mathbf{\Gamma}_{ji}(\mathbf{x}, t) - \dot{\mathbf{x}}_j \mathbf{\Gamma}_{ij}(\mathbf{x}, t) \right), \quad (2.8)$$

$$\nabla \psi_i(\mathbf{x}, t) = \sum_{j=1}^N (\mathbf{\Gamma}_{ji}(\mathbf{x}, t) - \mathbf{\Gamma}_{ij}(\mathbf{x}, t)). \quad (2.9)$$

Remark 2.1.2. *Because of*

$$\text{supp}(\mathbf{\Gamma}_{ij}) = \text{supp} \left(\frac{\psi_i(\mathbf{x}, t)}{\sigma(\mathbf{x}, t)} \nabla W_j(\mathbf{x}, t) \right),$$

the functions $\mathbf{\Gamma}_{ij}$ are located in the intersections of the supports of ψ_i and ψ_j .

Remark 2.1.3. *The particles ψ_i satisfy the transport equation*

$$\partial_t \psi_i + \dot{\mathbf{x}}_i \nabla \psi_i = \sum_{j=1}^N (\dot{\mathbf{x}}_j - \dot{\mathbf{x}}_i) \mathbf{\Gamma}_{ij},$$

where the left hand side describes the movement and the right hand side the deformations of the particles caused by the relative movement: whenever the relative particle positions change, the particles itself have to change, too, in order to satisfy the partition of unity property.

With the above definitions and calculations, we get the following system of ODEs

$$\begin{aligned} \frac{d}{dt}(V_i \mathbf{u}_i) &= \sum_{j=1}^N \int_{\Omega(t)} \mathbf{F}(\mathbf{u})(\mathbf{\Gamma}_{ji} - \mathbf{\Gamma}_{ij}) - \mathbf{u}(\dot{\mathbf{x}}_i \mathbf{\Gamma}_{ji} - \dot{\mathbf{x}}_j \mathbf{\Gamma}_{ij}) d\mathbf{x} - \int_{\partial\Omega(t)} \psi_i(\mathbf{F}(\mathbf{u}) - \mathbf{u} \cdot \mathbf{v}) \mathbf{n} ds \\ &= \sum_{j=1}^N \int_{\Omega(t)} \left[(\mathbf{F}(\mathbf{u}) - \mathbf{u} \cdot \dot{\mathbf{x}}_i) \mathbf{\Gamma}_{ji} - (\mathbf{F}(\mathbf{u}) - \mathbf{u} \cdot \dot{\mathbf{x}}_j) \mathbf{\Gamma}_{ij} \right] d\mathbf{x} - \mathbf{B}_i, \end{aligned} \quad (2.10)$$

where

$$\mathbf{B}_i = \int_{\partial\Omega(t)} \psi_i(\mathbf{F}(\mathbf{u}) - \mathbf{u} \cdot \mathbf{v}) \mathbf{n} ds. \quad (2.11)$$

Under the assumption that the smoothing length h is not too large in comparison to the distance of the particles, i.e. the particles do not overlap too much, and using the assumption that the unknown \mathbf{u} can be approximated by some $\bar{\mathbf{u}}_{ij}$ on $\text{supp}(\psi_i) \cap \text{supp}(\psi_j)$, we conclude

$$\frac{d}{dt}(V_i \mathbf{u}_i) \approx - \sum_{j=1}^N (\mathbf{F}(\bar{\mathbf{u}}_{ij}) - \bar{\mathbf{u}}_{ij} \cdot \bar{\mathbf{x}}_{ij}) \int_{\Omega(t)} (\mathbf{\Gamma}_{ij} - \mathbf{\Gamma}_{ji}) d\mathbf{x} - \mathbf{B}_i,$$

where $\bar{\mathbf{x}}_{ij}$ is some constant value being close to both $\dot{\mathbf{x}}_i$ and $\dot{\mathbf{x}}_j$. We will show in Theorem 2.2.10 how to choose an appropriate approximation, such that constant states are preserved even in the case when particles move.

Next, we define geometrical coefficients

$$\gamma_{ij}(t) := \int_{\Omega(t)} \mathbf{\Gamma}_{ij}(\mathbf{x}, t) d\mathbf{x}, \quad (2.12)$$

$$\beta_{ij}(t) := \gamma_{ij}(t) - \gamma_{ji}(t), \quad (2.13)$$

and $\mathbf{n}_{ij} = \beta_{ij}/\|\beta_{ij}\|$. We introduce a numerical flux function $\mathbf{g}_{ij} := \mathbf{g}(\mathbf{u}_i, \dot{\mathbf{x}}_i, \mathbf{u}_j, \dot{\mathbf{x}}_j, \mathbf{n}_{ij})$ that is consistent not with \mathbf{F} but with the modified flux $(\mathbf{F}(\mathbf{u}) - \mathbf{u} \cdot \dot{\mathbf{x}})$, meaning that

$$\mathbf{g}(\mathbf{u}, \dot{\mathbf{x}}_i, \mathbf{u}, \dot{\mathbf{x}}_j, \mathbf{n}_{ij}) = (\mathbf{F}(\mathbf{u}) - \mathbf{u} \cdot \dot{\mathbf{x}}_{ij}) \mathbf{n}_{ij}. \quad (2.14)$$

Thus, we can formulate a semi-discrete form of the FVPM

$$\frac{d}{dt}(V_i \mathbf{u}_i) = - \sum_{j=1}^N \|\beta_{ij}\| \mathbf{g}_{ij} - \mathbf{B}_i \quad (2.15)$$

with initial condition

$$\mathbf{u}_i(0) = \frac{1}{V_i(0)} \int_{\Omega(0)} \mathbf{u}_0(\mathbf{x}) \psi_i(\mathbf{x}) d\mathbf{x}. \quad (2.16)$$

Note that, because we are interested in the unknowns \mathbf{u}_i , we need an additional equation for the volumes V_i . The most simple choice seems to be Definition (2.6) for the volumes

$$V_i(t) = \int_{\Omega(t)} \psi_i(\mathbf{x}, t) d\mathbf{x}. \quad (2.17)$$

Alternatively, we can differentiate Equation (2.6). Using (2.8), we end up with

$$\dot{V}_i(t) = \sum_{j=1}^N (\gamma_{ij} \dot{\mathbf{x}}_j - \gamma_{ji} \dot{\mathbf{x}}_i) + \int_{\partial\Omega(t)} \psi_i \mathbf{v} \cdot \mathbf{n} ds. \quad (2.18)$$

All in all, the semi-discrete form of the FVPM reads

$$\frac{d}{dt}(V_i \mathbf{u}_i) = - \sum_{j=1}^N \|\beta_{ij}\| \mathbf{g}_{ij} - \mathbf{B}_i \quad (2.19a)$$

$$\dot{V}_i = \sum_{j=1}^N (\gamma_{ij} \dot{\mathbf{x}}_j - \gamma_{ji} \dot{\mathbf{x}}_i) + \int_{\partial\Omega(t)} \psi_i \mathbf{v} \cdot \mathbf{n} ds \quad (2.19b)$$

$$\mathbf{u}_i(0) = \frac{1}{V_i(0)} \int_{\Omega(0)} \mathbf{u}_0(\mathbf{x}) \psi_i(\mathbf{x}, 0) d\mathbf{x} \quad (2.19c)$$

$$V_i(0) = \int_{\Omega(0)} \psi_i(\mathbf{x}, 0) d\mathbf{x} \quad (2.19d)$$

together with a prescribed velocity field for the motion of the particles. As a reconstruction, the following formula is reasonable for the semi-discrete scheme:

$$\mathbf{u}_h(\mathbf{x}, t) = \sum_{i=1}^N \mathbf{u}_i(t) \psi_i(\mathbf{x}, t). \quad (2.20)$$

Originally, in [HSS00], the terms related with the particle movement in (2.10) have been discretized separately instead of treating them as part of a numerical flux function. The semi-discrete version for this approach then reads

$$\frac{d}{dt}(V_i \mathbf{u}_i) = - \sum_{j=1}^N [||\beta_{ij}|| \mathbf{f}_{ij} - (\gamma_{ij} \dot{\mathbf{x}}_j \mathbf{u}_i - \gamma_{ji} \dot{\mathbf{x}}_i \mathbf{u}_j)] - \mathbf{B}_i \quad (2.21a)$$

$$\dot{V}_i = \sum_{j=1}^N (\gamma_{ij} \dot{\mathbf{x}}_j - \gamma_{ji} \dot{\mathbf{x}}_i) + \int_{\partial\Omega(t)} \psi_i \mathbf{v} \cdot \mathbf{n} \, ds \quad (2.21b)$$

$$\mathbf{u}_i(0) = \frac{1}{V_i(0)} \int_{\Omega(0)} \mathbf{u}_0(\mathbf{x}) \psi_i(\mathbf{x}, 0) \, d\mathbf{x} \quad (2.21c)$$

$$V_i(0) = \int_{\Omega(0)} \psi_i(\mathbf{x}, 0) \, d\mathbf{x}, \quad (2.21d)$$

again with given particle movement and the reconstruction formula (2.20). Here, $\mathbf{f}_{ij} := \mathbf{f}(\mathbf{u}_i, \mathbf{u}_j, \mathbf{n}_{ij})$ denotes a numerical flux function that is consistent with the original flux \mathbf{F} from (2.1).

Remark 2.1.4. *The approach (2.21a)–(2.21d) has the drawback of suffering from instabilities in some situations and for special particle movement, compare [Kec02] or [Tel05]. Although we will use mainly scheme (2.19a)–(2.19d) and (2.20) in what follows, we mention (2.21a)–(2.21d) because it preserves constant states even for moving particles by construction (see Theorem 2.2.8). In order to construct a stable scheme that preserves constant states, this property will be adopted for scheme (2.19a)–(2.19d) in Theorem 2.2.10 by a special way of embedding the particle movement in the numerical flux function.*

In the following chapters, we will use a simple Euler discretization of (2.19a)–(2.19d) to get a fully discretized scheme, although higher order methods are possible. The Euler discretized

scheme with appropriate initial conditions reads

$$V_i^{n+1} \mathbf{u}_i^{n+1} = V_i^n \mathbf{u}_i^n - \Delta t \left(\sum_{j \in \mathcal{N}(i)} \|\beta_{ij}^n\| \mathbf{g}_{ij}^n + \mathbf{B}_i^n \right) \quad (2.22a)$$

$$V_i^{n+1} = V_i^n + \Delta t \sum_{j=1}^N (\gamma_{ij}^n \dot{\mathbf{x}}_j^n - \gamma_{ji}^n \dot{\mathbf{x}}_i^n) + C_i^n \quad (2.22b)$$

$$\mathbf{u}_i^0 = \frac{1}{V_i^0} \int_{\Omega} \mathbf{u}_0(\mathbf{x}) \psi_i(\mathbf{x}, 0) d\mathbf{x} \quad (2.22c)$$

$$V_i^0 = \int_{\Omega} \psi_i(\mathbf{x}, 0) d\mathbf{x}, \quad (2.22d)$$

where \mathbf{B}_i^n and C_i^n are suitable discretizations of the term \mathbf{B}_i in (2.11) and $C_i := \int_{\partial\Omega} \psi_i \mathbf{v} \mathbf{n} ds$, respectively, $\mathbf{g}_{ij}^n := \mathbf{g}(\mathbf{u}_i^n, \dot{\mathbf{x}}_i^n, \mathbf{u}_j^n, \dot{\mathbf{x}}_j^n, \mathbf{n}_{ij}^n)$, $\gamma_{ij}^n := \gamma_{ij}(t^n)$ and $\dot{\mathbf{x}}_i^n := \dot{\mathbf{x}}_i(t^n)$. In what follows, we will often use the abbreviations

$$\mathbf{g}_{ij}(\mathbf{u}, \mathbf{v}) := \mathbf{g}(\mathbf{u}, \dot{\mathbf{x}}_i, \mathbf{v}, \dot{\mathbf{x}}_j, \mathbf{n}_{ij})$$

and

$$\mathbf{g}_{ij}^n := \mathbf{g}(\mathbf{u}_i^n, \dot{\mathbf{x}}_i^n, \mathbf{u}_j^n, \dot{\mathbf{x}}_j^n, \mathbf{n}_{ij}^n).$$

As a reconstruction formula for the fully discretized scheme one may choose for example

$$\mathbf{u}_h(\mathbf{x}, t) = \sum_{n=0}^{N_T} \sum_{i=1}^N \mathbf{u}_i^n \psi_i(\mathbf{x}, t) \mathbb{1}_{[t^n, t^{n+1})}(t). \quad (2.23)$$

We will see in the next section that this choice of reconstruction is conservative (see Theorem 2.2.6 and Remark 2.2.7).

2.2 The geometrical coefficients and properties of the scheme

In this section, we describe the geometrical coefficients and some important properties of the FVPM that are closely related to those coefficients. We start with well-known approximation results for the generalized cell averages and the reconstruction formula (see Lemma 2.2.1, Theorem 2.2.2), followed by some crucial properties of the coefficients (see Proposition 2.2.4). These properties imply for example the conservativity of the scheme (see Theorem 2.2.6). Subsequently, we show that the FVPM preserves constant states. This has already been shown for non-moving particles, but was still an open question for the moving case. In Theorem 2.2.10, we propose an average particle velocity in order to implement the particle movement in the numerical flux function. With this average particle velocity, the FVPM preserves constant states even when the particles are moving. Some important properties of this particle velocity are figured out in Proposition 2.2.13.

During this section, we always assume that the computational domain Ω is an open subset of \mathbb{R}^d and that the family of functions ψ_i form a partition of unity on Ω as long as we do not

make other assumptions.

We begin this part by showing that the order of the approximation with the help of cell averages is best in the barycenters. For the FVM and the corresponding cell averages, this result can be found in [Son95] and [AS09]. In [Kec02], the result was adopted for the generalized cell averages (2.5).

Lemma 2.2.1 (Approximation property). *Let $\mathbf{u} \in C^2(\Omega \times \mathbb{R}_+)$ and the barycenters of the cells $\Omega_i = \text{supp}(\psi_i)$ be defined as*

$$\mathbf{b}_i(t) := \frac{1}{V_i(t)} \int_{\Omega} \mathbf{x} \psi_i(\mathbf{x}, t) d\mathbf{x}.$$

Then the approximation \mathbf{u}_i of \mathbf{u} is of second order, i.e.

$$\mathbf{u}_i(t) = \mathbf{u}(\mathbf{b}_i, t) + \mathcal{O}(h^2),$$

where h is the smoothing length.

Proof. The proof can be found in [Son95] and [AS09] and is given here for completeness. For simplicity, we will drop the second argument t in the following. We use the Taylor series expansion

$$\mathbf{u}(\mathbf{x}) = \mathbf{u}(\mathbf{b}_i) + \mathbf{u}'(\mathbf{b}_i)(\mathbf{x} - \mathbf{b}_i) + \mathcal{O}(h^2)$$

and get

$$\begin{aligned} \mathbf{u}_i - \mathbf{u}(\mathbf{b}_i) &= \frac{1}{V_i} \int_{\Omega} (\mathbf{u} - \mathbf{u}(\mathbf{b}_i)) \psi_i d\mathbf{x} \\ &= \frac{1}{V_i} \int_{\Omega} \mathbf{u}'(\mathbf{b}_i) (\mathbf{x} - \mathbf{b}_i) \psi_i d\mathbf{x} + \mathcal{O}(h^2) \\ &= \mathbf{u}'(\mathbf{b}_i) \left(\frac{1}{V_i} \int_{\Omega} \mathbf{x} \psi_i d\mathbf{x} - \mathbf{b}_i \right) + \mathcal{O}(h^2) \\ &= \mathcal{O}(h^2). \end{aligned}$$

□

With the help of Lemma 2.2.1, we can directly deduce the following approximation property for the reconstruction formula (2.20) as stated in [Kec02].

Theorem 2.2.2 (Approximation property). *For the reconstruction (2.20) we have for all $(\mathbf{x}, t) \in \Omega \times [0, T]$*

$$\mathbf{u}_h(\mathbf{x}, t) = \mathbf{u}(\mathbf{x}, t) + \mathcal{O}(h). \quad (2.24)$$

Moreover, it was shown in [Tel05] that for $u(\cdot, t) \in H^1(\Omega, \mathbb{R})$ and Ω being an open subset of \mathbb{R}^2 , the reconstruction (2.20) satisfies the estimate

$$\|u_h(\cdot, t) - u(\cdot, t)\|_{L^2(\Omega, \mathbb{R})} \leq C h \|\nabla u(\cdot, t)\|_{L^2(\Omega, \mathbb{R})}$$

where C is a constant independent of h . The proof uses the condition of a maximal overlap of the partition of unity according to (2.41) and $\|\psi_i(\cdot, t)\|_{L^\infty(\Omega, \mathbb{R})} \leq C_\infty$ for all particles.

We state now that the partition of unity satisfies the following properties. For the proof see [Kec02].

Proposition 2.2.3. *For a multiindex α and $W_i \in C^{|\alpha|}$, with $|\alpha| := \sum_{j=1}^d \alpha_j$, the functions ψ_i satisfy*

$$\begin{aligned} |D^\alpha \psi_i(\mathbf{x}, t)| &\leq \frac{C_\alpha}{h^{|\alpha|}} \\ \left| \int_{\Omega} g(\mathbf{x}) \psi_i(\mathbf{x}, t) d\mathbf{x} \right| &\leq C h^{d+q} \quad \text{if } g(\mathbf{x}) = \mathcal{O}(h^q), \end{aligned}$$

where C_α and C are h -independent constants.

Besides these results, it is easy to show some crucial characteristics of the geometrical coefficients and the generalized volumes.

Proposition 2.2.4 (Properties of the geometrical coefficients). *The geometrical coefficients defined in (2.13) satisfy*

$$\beta_{ij} = -\beta_{ji} \quad \forall i, j = 1, \dots, N, \quad (2.25a)$$

$$\beta_{ii} = 0 \quad \forall i = 1, \dots, N, \quad (2.25b)$$

$$\beta_{ij} = 0, \text{ if } \text{supp}(\psi_i) \cap \text{supp}(\psi_j) = \emptyset, \quad (2.25c)$$

$$\beta_{ij} = 2 \int_{\Omega} \psi_i \nabla \psi_j d\mathbf{x} - \int_{\partial\Omega} \psi_i \psi_j \mathbf{n} ds, \quad (2.25d)$$

$$\sum_j \beta_{ij} = - \int_{\partial\Omega} \psi_i \mathbf{n} ds \quad \forall i = 1, \dots, N, \quad (2.25e)$$

$$\|\beta_{ij}\| = \mathcal{O}(h^{d-1}), \quad (2.25f)$$

$$V_i = \mathcal{O}(h^d). \quad (2.25g)$$

In the one-dimensional case, we have for every $\bar{x} \in \mathbb{R}$

$$\sum_{x_i \geq \bar{x}}^N \sum_{x_j \geq \bar{x}}^N \beta_{ij} = 1. \quad (2.26)$$

Proof. The proofs of the above properties can be found in [JS01], [Tel05] or [Kec02] and are given here for completeness. (2.25a), (2.25b) and (2.25c) follow directly from Definition (2.13). For (2.25d), we use the construction of the partition of unity and get

$$\nabla \psi_i = \frac{\nabla W_i}{\sigma} - \psi_i \frac{\nabla \sigma}{\sigma}$$

and

$$\nabla W_i = \sigma \nabla \psi_i + \psi_i \nabla \sigma$$

and therefore

$$\begin{aligned} \beta_{ij} &= \int_{\Omega} (\mathbf{\Gamma}_{ij} - \mathbf{\Gamma}_{ji}) \, d\mathbf{x} = \int_{\Omega} \left(\frac{\psi_i}{\sigma} \nabla W_j - \frac{\psi_j}{\sigma} \nabla W_i \right) \, d\mathbf{x} \\ &= \int_{\Omega} (\psi_i \nabla \psi_j - \psi_j \nabla \psi_i) \, d\mathbf{x} = 2 \int_{\Omega} (\psi_i \nabla \psi_j) \, d\mathbf{x} - \int_{\partial\Omega} \psi_i \psi_j \mathbf{n} \, ds. \end{aligned}$$

Now this and the partition of unity property can be used to show (2.25e):

$$\begin{aligned} \sum_j \beta_{ij} &= 2 \sum_j \int_{\Omega} \psi_i \nabla \psi_j \, d\mathbf{x} - \sum_j \int_{\partial\Omega} \psi_i \psi_j \mathbf{n} \, ds \\ &= 2 \int_{\Omega} \psi_i \nabla \underbrace{\sum_j \psi_j}_{=0} \, d\mathbf{x} - \int_{\partial\Omega} \psi_i \sum_j \psi_j \mathbf{n} \, ds = - \int_{\partial\Omega} \psi_i \mathbf{n} \, ds. \end{aligned}$$

The properties (2.25f) and (2.25g) can be deduced directly from Proposition (2.2.3). For (2.26), we define the function $\tilde{\sigma}_{\bar{x}}(x, t) := \sum_{x_i \geq \bar{x}} \psi_i(x, t)$ and obtain

$$\begin{aligned} \sum_{x_i \geq \bar{x}} \sum_{x_j \geq \bar{x}} \beta_{ij} &= 2 \int_{\Omega} \tilde{\sigma}_{\bar{x}} \partial_x \tilde{\sigma}_{\bar{x}} \, dx = 2 \int_{\Omega} (\partial_x \tilde{\sigma}_{\bar{x}})^2 \, dx \\ &= \lim_{x \rightarrow \infty} \tilde{\sigma}_{\bar{x}}^2 - \lim_{x \rightarrow -\infty} \tilde{\sigma}_{\bar{x}}^2 = 1. \end{aligned}$$

□

Remark 2.2.5. Note that in the one-dimensional case and for $\Omega = [a, b] \subset \mathbb{R}$, property (2.25e) reads

$$\sum_j \beta_{ij} = - \int_{\partial\Omega} \psi_i \mathbf{n} \, ds = \psi_i(a) - \psi_i(b).$$

Theorem 2.2.6 (Conservativity). *Under the condition that the numerical flux function is conservative, i.e.*

$$\mathbf{g}_{ij}(\mathbf{u}, \mathbf{v}) = -\mathbf{g}_{ji}(\mathbf{v}, \mathbf{u}), \quad (2.27)$$

the FVPM (2.19a)–(2.19d) is conservative in the sense that

$$\frac{d}{dt} \left(\sum_{i=1}^N V_i \mathbf{u}_i \right) = - \int_{\partial\Omega} \left(\mathbf{F}(\mathbf{u}) - \mathbf{u} \cdot \mathbf{v} \right) \mathbf{n} \, ds. \quad (2.28)$$

Proof. The proof can be found in [HSS00], [JS01] or [Tel05] and is given here for completeness. Because of the conservation property of the numerical flux function (2.27) and the skew-symmetry condition of the geometrical coefficients (2.25a), we can write

$$\begin{aligned} \frac{d}{dt} \left(\sum_{i=1}^N V_i \mathbf{u}_i \right) &= - \sum_{i,j=1}^N \|\beta_{ij}\| \mathbf{g}_{ij} - \sum_{i=1}^N \mathbf{B}_i \\ &= - \sum_{i=1}^N \int_{\partial\Omega} \psi_i \left(\mathbf{F}(\mathbf{u}) - \mathbf{u} \cdot \mathbf{v} \right) \mathbf{n} \, ds \\ &= - \int_{\partial\Omega} \left(\mathbf{F}(\mathbf{u}) - \mathbf{u} \cdot \mathbf{v} \right) \mathbf{n} \, ds. \end{aligned}$$

□

Remark 2.2.7. *Note that this result ensures a conservation property of the reconstruction (2.20), because we have*

$$\begin{aligned} \frac{d}{dt} \left(\int_{\Omega} \mathbf{u}_h \, d\mathbf{x} \right) &= \frac{d}{dt} \left(\sum_i \mathbf{u}_i \int_{\Omega} \psi_i \, d\mathbf{x} \right) = \frac{d}{dt} \left(\sum_i \mathbf{u}_i V_i \right) \\ &= - \int_{\partial\Omega} \left(\mathbf{F}(\mathbf{u}) - \mathbf{u} \cdot \mathbf{v} \right) \mathbf{n} \, ds. \end{aligned} \quad (2.29)$$

Next, we will formulate in which way the different discretizations of the FVPM preserve constant states under the assumption of a non-moving computational domain.

Theorem 2.2.8 (Preservation of constant states). *Consider the semi-discrete scheme (2.21a)–(2.21d) and a non-moving spatial domain Ω . Let us further assume a conservative and consistent numerical flux function*

$$\mathbf{f}(\mathbf{u}_i, \mathbf{u}_j, \mathbf{n}_{ij}) = -\mathbf{f}(\mathbf{u}_j, \mathbf{u}_i, -\mathbf{n}_{ij})$$

and an arbitrary particle motion $\hat{\mathbf{x}}_i$, $i = 0, \dots, N$. If for $t \in \mathbb{R}_+$

$$\mathbf{u}_j(t) = \mathbf{u}_i(t) \quad \forall j \in \mathcal{N}(i),$$

scheme (2.21a)–(2.21d) preserves constant states in the sense that

$$\frac{d}{dt} \mathbf{u}_i(t) = 0.$$

Proof. For non-moving particles, this has already been done and can be found for example in [Tel05]. For the case of moving particles, we refer to [Kad]. □

Remark 2.2.9. *Note that the fully discretized scheme, obtained by an explicit forward Euler method in time in the semi-discrete version (2.21a)–(2.21d), guarantees the preservation of constant states in the sense that the coefficients satisfy*

$$\mathbf{u}_i^{n+1} = \mathbf{u}_i^n \quad \forall n \in \mathbb{N}, \forall i = 0, \dots, N,$$

if $\mathbf{u}_i^0 = \mathbf{u} = \text{const}$ for all $i = 0, \dots, N$.

We will now show that for a special approximation of the velocities $\dot{\mathbf{x}}_i$ and $\dot{\mathbf{x}}_j$ of neighbouring particles ψ_i and ψ_j in the numerical flux function, scheme (2.19a)–(2.19d) preserves constant states, too. In Proposition 2.2.13, we will show why (2.30) should be considered as a good approximation of the particle velocities $\dot{\mathbf{x}}_i$ and $\dot{\mathbf{x}}_j$.

Theorem 2.2.10 (Preservation of constant states). *Assume again an arbitrary particle motion and a conservative numerical flux function \mathbf{g}_{ij} that satisfies the consistency condition (2.14) with the following approximation of the movement of the particles*

$$(\bar{\mathbf{x}}_{ij})_k := \frac{(\gamma_{ij})_k (\dot{\mathbf{x}}_j)_k - (\gamma_{ji})_k (\dot{\mathbf{x}}_i)_k}{(\beta_{ij})_k}, \quad \text{for } k = 1, \dots, d, \quad (2.30)$$

where the index k denotes the k -th component. Provided that the spatial domain Ω is non-moving and for $t \in \mathbb{R}_+$, we have

$$\mathbf{u}_j(t) = \mathbf{u}_i(t) \quad \forall j \in \mathcal{N}(i),$$

the semi-discrete scheme (2.19a)–(2.19d) preserves constant states in the sense that

$$\frac{d}{dt} \mathbf{u}_i(t) = 0.$$

Proof. The proof in case of fixed particles and the fully discretized scheme may be found in [Tel05]. Let us now assume that the particles are moving according to the prescribed velocity field $\dot{\mathbf{x}}_i$, $i = 0, \dots, N$. Then, we have

$$\begin{aligned} \frac{d}{dt}(V_i \mathbf{u}_i) &= - \sum_{j=1}^N \|\beta_{ij}\| \mathbf{g}_{ij} - \mathbf{B}_i = - \sum_{j=1}^N (\mathbf{F}(\mathbf{u}_i) - \mathbf{u}_i \cdot \bar{\mathbf{x}}_{ij}) \beta_{ij} - \mathbf{B}_i \\ &= \sum_{j=1}^N (\mathbf{u}_i \cdot \bar{\mathbf{x}}_{ij}) \beta_{ij} = \mathbf{u}_i \sum_{j=1}^N \bar{\mathbf{x}}_{ij} \beta_{ij} \\ &= \mathbf{u}_i \sum_{j=1}^N \sum_{k=1}^d \frac{(\gamma_{ij})_k (\dot{\mathbf{x}}_j)_k - (\gamma_{ji})_k (\dot{\mathbf{x}}_i)_k}{(\beta_{ij})_k} (\beta_{ij})_k \\ &= \mathbf{u}_i \sum_{j=1}^N (\gamma_{ij} \dot{\mathbf{x}}_j - \gamma_{ji} \dot{\mathbf{x}}_i) = \mathbf{u}_i \dot{V}_i. \end{aligned}$$

The proposition follows directly. □

Remark 2.2.11. *Again, the property of preserving constant states holds for the fully discretized FVPM (2.22a)–(2.22d), too.*

Remark 2.2.12. *Note that Theorems 2.2.8 and 2.2.10 ensure the preservation of constant states in the reconstruction (2.20), since $\mathbf{u}_i(t) \equiv \bar{\mathbf{u}} = \text{const}$ implies*

$$\frac{d}{dt} \mathbf{u}_h(\mathbf{x}, t) = \frac{d}{dt} \sum_i \mathbf{u}_i(t) \psi_i(\mathbf{x}, t) = \bar{\mathbf{u}} \underbrace{\frac{d}{dt} \sum_i \psi_i(\mathbf{x}, t)}_{\equiv 1} = 0.$$

The following Proposition states some important properties of the approximation (2.30).

Proposition 2.2.13. *The approximate particle velocity (2.30) satisfies the following.*

i) *For neighbouring particles ψ_i and ψ_j , (2.30) fulfills the symmetry condition*

$$\bar{\mathbf{x}}_{ij} = \bar{\mathbf{x}}_{ji}. \quad (2.31)$$

In particular, (2.30) ensures $\bar{\mathbf{x}}_{ij} = \dot{\mathbf{x}}$ if $\dot{\mathbf{x}}_i = \dot{\mathbf{x}}_j = \dot{\mathbf{x}}$.

ii) *In the one-dimensional case, (2.30) can be written as*

$$\bar{x}_{ij} = \lambda \dot{x}_j + (1 - \lambda) \dot{x}_i. \quad (2.32)$$

If the function $\sigma = \sum_{i=1}^N W_i(x, t)$ varies not too much spatially, meaning that $\partial_x \sigma$ satisfies the estimate

$$\left| \frac{\int_{\Omega} \psi_i \psi_j (\sigma^{-1}) \partial_x \sigma \, dx}{\int_{\Omega} \psi_i \partial_x \psi_j \, dx} \right| \leq 1, \quad (2.33)$$

\bar{x}_{ij} is a convex combination of \dot{x}_i and \dot{x}_j .

Proof. The symmetry (2.31) can be computed directly for each component. In the one-dimensional case, we can write $\bar{x}_{ij} = \lambda \dot{x}_j + \mu \dot{x}_i$ with $\lambda = \frac{\gamma_{ij}}{\beta_{ij}}$ and $\mu = -\frac{\gamma_{ji}}{\beta_{ij}}$ which satisfy

$$\lambda + \mu = \frac{\gamma_{ij}}{\beta_{ij}} - \frac{\gamma_{ji}}{\beta_{ij}} = 1$$

and therefore (2.32) holds. Under condition (2.33), we obtain

$$\begin{aligned} \lambda &= \frac{\gamma_{ij}}{\beta_{ij}} = \frac{\int \psi_i \sigma^{-1} \partial_x W_j \, dx}{2 \int \psi_i \partial_x \psi_j \, dx} \\ &= \frac{\int (\psi_i \partial_x \psi_j + \psi_i \psi_j \sigma^{-1} \partial_x \sigma) \, dx}{2 \int \psi_i \partial_x \psi_j \, dx} \\ &= \frac{1}{2} + \frac{1}{2} \frac{\int \psi_i \psi_j \sigma^{-1} \partial_x \sigma \, dx}{\int \psi_i \partial_x \psi_j \, dx}, \end{aligned}$$

and therefore $0 \leq \lambda \leq 1$. □

Remark 2.2.14. Note that the symmetry (2.31) is important for the conservativity of the numerical flux \mathbf{g}_{ij} in case of moving particles, because

$$\mathbf{g}_{ij}(\mathbf{u}, \mathbf{u}) = \mathbf{g}(\mathbf{u}, \dot{\mathbf{x}}_i, \mathbf{u}, \dot{\mathbf{x}}_j, \mathbf{n}_{ij}) = (\mathbf{F}(\mathbf{u}) - \mathbf{u} \cdot \bar{\mathbf{x}}_{ij}) \mathbf{n}_{ij} = -(\mathbf{F}(\mathbf{u}) - \mathbf{u} \cdot \bar{\mathbf{x}}_{ji}) \mathbf{n}_{ji} = -\mathbf{g}_{ji}(\mathbf{u}, \mathbf{u}).$$

Remark 2.2.15. In higher dimensions, (2.30) builds a convex combination in each coordinate, as long as (2.33) is valid in each coordinate. A convex combination in general cannot be guaranteed.

We will now give some examples where $\nabla\sigma \equiv 0$ and thus, (2.33) is satisfied.

Example 2.2.16.

- (one-dimensional hat functions) Consider a one-dimensional setting where each particle has only two neighbours: $\frac{\Delta x_i}{2} \leq h \leq \Delta x_i$, where $\Delta x_i := x_{i+1} - x_i$. For the partition of unity, we use the piecewise linear polynomials

$$W_i(x, t) = \begin{cases} \frac{x - x_i + h}{h}, & x \in [x_i - h, x_i) \\ \frac{x - x_i + h}{h}, & x \in [x_i, x_i + h) \\ 0, & \text{else.} \end{cases}$$

Then, some easy computations lead to

$$\gamma_{i,j} = \begin{cases} \frac{1}{2}, & j = i + 1 \\ -\frac{1}{2}, & j = i - 1 \\ 0, & \text{else} \end{cases}$$

and

$$\beta_{i,j} = \begin{cases} 1, & j = i + 1 \\ -1, & j = i - 1 \\ 0, & \text{else.} \end{cases}$$

In that case, we end up with

$$\bar{\mathbf{x}}_{ij} = \frac{\dot{\mathbf{x}}_i + \dot{\mathbf{x}}_j}{2}.$$

- (one-dimensional quadratic kernel) Consider again a one-dimensional setting. The partition of unity is now computed using piecewise quadratic polynomials

$$W_i(x, t) = \begin{cases} (x - x_i + h)^2, & x \in [x_i - h, x_i - \frac{h}{2}) \\ -(x - x_i)^2 + \frac{h^2}{2}, & x \in [x_i - \frac{h}{2}, x_i + \frac{h}{2}) \\ (x - x_i - h)^2, & x \in [x_i + \frac{h}{2}, x_i + h) \\ 0, & \text{else.} \end{cases}$$

The particles are assumed to be distributed equidistantly with six neighbours each and a smoothing length of $h = 2\Delta x$. In this case, the function σ satisfies $\sigma \equiv h^2 = \text{const}$ [Tel05], such that the integral in (2.33) vanishes. Thus, we end up as before with

$$\bar{x}_{ij} = \frac{\dot{x}_i + \dot{x}_j}{2}.$$

- (two-dimensional quadratic kernel) Condition (2.33) is naturally satisfied if the functions W_i already form a partition of unity, because then $\sigma = \sum_i W_i \equiv 1$ and $\nabla\sigma \equiv 0$. This is, for example, the case for the two-dimensional hat functions based on

$$W(x, y) = \begin{cases} \frac{-x+h}{h}, & x \in [0, h), y \in [x-h, 0) \\ \frac{-x+y+h}{h}, & x \in [0, h), y \in [0, x) \\ \frac{x-y+h}{h}, & x \in [0, h), y \in [x, h) \\ \frac{x+h}{h}, & x \in [-h, 0), y \in [0, x+h) \\ \frac{-y+h}{h}, & x \in [-h, 0), y \in [x, 0) \\ \frac{y+h}{h}, & x \in [-h, 0), y \in [-h, x) \\ 0, & \text{else} \end{cases}$$

and equidistantly distributed particles with a smoothing length of $h = \Delta x = \Delta y$.

- (B-splines) B-splines naturally build a partition of unity by construction and can therefore be used in the FVPM as can be seen in Chapter 3.

In [Tel05] and [Tel00], it was shown for a scalar conservation law in one space dimension that the FVPM (2.22a)–(2.22d) is monotone under suitable assumptions. In particular, the numerical flux function $g_{ij}^n(u, v)$ is assumed to be in C^1 and monotone nondecreasing in u and nonincreasing in v . Furthermore, the time step Δt has to satisfy the CFL-like condition

$$\Delta t \leq \frac{\min_i V_i^n}{L \max_i \sum_j |\beta_{ij}|}, \quad (2.34)$$

where L is the Lipschitz constant of g_{ij} . In Section 2.3, we will show monotonicity of the FVPM (2.22a)–(2.22d) in several space dimensions under a slightly different CFL condition and weaker conditions on the smoothness of g_{ij} , see Lemma 2.3.6. We will need the monotonicity to show a discrete entropy inequality, compare Lemma 2.3.8.

In [JS01], the authors showed a Lax-Wendroff result for the semi-discrete formulation of the FVPM in the case of a scalar conservation law in one spatial dimension: Under suitable conditions on the partition of unity and under the condition that the coefficients $u_i(t)$ are bounded and converge in some sense towards u , then u is a weak solution of the Cauchy problem. In addition, it can be shown that the reconstruction $\sum_i u_i \psi_i$ converges towards u in $L_{loc}^1(\mathbb{R}^+, L_{loc}^1(\mathbb{R}, \mathbb{R}))$. The proof makes use of property (2.26).

We end this section with an overview on procedures that ensure the important properties of the coefficients formulated in Proposition 2.2.4 by correcting them numerically.

Correction procedures

We have seen that the properties of the geometrical coefficients β_{ij} are important for the resulting method: they ensure, for example, conservativity and the preservation of constant states. Therefore, it is of great interest to ensure these properties numerically. Since in the computation of the coefficients β_{ij} we have to integrate numerically in each time step, and motivated by the fact that numerical quadrature is cost expensive, it is best to compute the coefficients by a coarse integration technique and then to correct them, so that the desired properties are fulfilled. Obviously, it is easy to guarantee conditions (2.25a)–(2.25c) in contrast to condition (2.25e).

The following fast and efficient correction procedure, where the error with respect to condition (2.25e) is shifted from one particle to the “next” neighbour, was given by Keck in [Kec02] and [HK03]. If the boundary integrals in (2.25e) can be computed exactly, it is easy to verify that (2.25a)–(2.25c) and (2.25e) are satisfied.

- Compute approximations to the coefficients $\tilde{\beta}_{ij} =: \tilde{\beta}_{ij}^{(0)}$ satisfying (2.25a)–(2.25c) by using a coarse and fast integration technique.
- Define the errors

$$\tilde{\beta}_{ii}^{(0)} = \sum_j \beta_{ij} - \sum_j \tilde{\beta}_{ij}^{(0)} = - \int_{\partial\Omega} \psi_i \mathbf{n} \, ds - \sum_j \tilde{\beta}_{ij}^{(0)}$$

and save them on the main diagonal of the matrix $(\tilde{\beta})_{ij} = \tilde{\beta}_{ij}$.

- For $i = 1, \dots, N-1$: choose an index j such that $\Omega_i \cap \Omega_j \neq \emptyset$ and $j > i$ and set

$$\begin{aligned} \tilde{\beta}_{ij}^{(i)} &= \tilde{\beta}_{ij}^{(i-1)} + \tilde{\beta}_{ii}^{(i-1)} \\ \tilde{\beta}_{ji}^{(i)} &= \tilde{\beta}_{ji}^{(i-1)} - \tilde{\beta}_{ii}^{(i-1)} \\ \tilde{\beta}_{ii}^{(i)} &= \tilde{\beta}_{ii}^{(i-1)} - \tilde{\beta}_{ii}^{(i-1)} = \mathbf{0} \\ \tilde{\beta}_{jj}^{(i)} &= \tilde{\beta}_{jj}^{(i-1)} + \tilde{\beta}_{ii}^{(i-1)}. \end{aligned}$$

For all other elements set $\tilde{\beta}_{kl}^{(i)} = \tilde{\beta}_{kl}^{(i-1)}$.

Then, $\beta_{ij} := \tilde{\beta}_{ij}^{(N-1)}$ satisfies (2.25a)–(2.25c) and (2.25e) by construction. For the exact computation of the integrals in (2.25e) in the one-dimensional case, see Remark (2.2.5).

The next idea of correcting the coefficients can be found in [Tel05] and [TS08]. This procedure is very simple to handle, but has the fundamental drawback that it destroys the conservativity of the scheme. As before, in each time step, we compute the coefficients $\tilde{\beta}_{ij}^n$ by a fast and

coarse numerical integration technique. Then, for all interior particles i , we add the error term $-\Delta t \mathbf{G}(\mathbf{u}_i^n) \tilde{\boldsymbol{\beta}}_{ii}^n$, where $\tilde{\boldsymbol{\beta}}_{ii}^n = \sum_{j \in \mathcal{N}(i)} \tilde{\boldsymbol{\beta}}_{ij}^n$, so the scheme writes

$$V_i^{n+1} \mathbf{u}_i^{n+1} = V_i^n \mathbf{u}_i^n - \Delta t \sum_{j \in \mathcal{N}(i)} \|\tilde{\boldsymbol{\beta}}_{ij}^n\| \mathbf{g}_{ij}^n - \Delta t \mathbf{G}(\mathbf{u}_i^n) \tilde{\boldsymbol{\beta}}_{ii}^n \quad (2.35)$$

with $\mathbf{G}(\mathbf{u}) := \mathbf{F}(\mathbf{u}) - \mathbf{u} \cdot \dot{\mathbf{x}}_i$. In general, this scheme is not conservative any more, although the error is only of order $\mathcal{O}(h^{d-1+q})$ if q is the order of the numerical integration. Nevertheless, we will construct a possible movement of the particles such that conservativity can be guaranteed at least in some special cases and under appropriate conditions for this correction method. To maintain the conservativity of the scheme, we are interested in getting

$$\begin{aligned} \sum_i \mathbf{G}(\mathbf{u}_i^n) \tilde{\boldsymbol{\beta}}_{ii}^n &= 0 \\ \Leftrightarrow \sum_i (\mathbf{F}(\mathbf{u}_i^n) - \mathbf{u}_i^n \cdot \mathbf{a}_i^n) \tilde{\boldsymbol{\beta}}_{ii}^n &= 0 \\ \Leftrightarrow \sum_i (\mathbf{u}_i^n \cdot \tilde{\boldsymbol{\beta}}_{ii}^n) \mathbf{a}_i^n &= \sum_i \mathbf{F}(\mathbf{u}_i^n) \tilde{\boldsymbol{\beta}}_{ii}^n, \end{aligned} \quad (2.36)$$

where $\mathbf{u}_i^n = (u_{i,1}^n, \dots, u_{i,m}^n)^T \in \mathbb{R}^m$, $\mathbf{a}_i^n = (a_{i,1}^n, \dots, a_{i,d}^n)^T \in \mathbb{R}^d$ and $\tilde{\boldsymbol{\beta}}_{ii}^n \in \mathbb{R}^d$ are the cell averages, the velocity of the particles in the particle positions x_i at time t^n and the accumulated error term of the geometrical coefficients, respectively, and the product \cdot is the outer product that maps from $\mathbb{R}^m \times \mathbb{R}^d$ into $\mathbb{R}^{m \times d}$, i.e.

$$\mathbf{u}_i^n \cdot \tilde{\boldsymbol{\beta}}_{ii}^n = (\tilde{\beta}_{ii_1}^n \mathbf{u}_i^n, \dots, \tilde{\beta}_{ii_d}^n \mathbf{u}_i^n) = \begin{pmatrix} \tilde{\beta}_{ii_1}^n u_{i,1}^n & \tilde{\beta}_{ii_2}^n u_{i,1}^n & \dots & \tilde{\beta}_{ii_d}^n u_{i,1}^n \\ \tilde{\beta}_{ii_1}^n u_{i,2}^n & \tilde{\beta}_{ii_2}^n u_{i,2}^n & \dots & \tilde{\beta}_{ii_d}^n u_{i,2}^n \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\beta}_{ii_1}^n u_{i,m}^n & \tilde{\beta}_{ii_2}^n u_{i,m}^n & \dots & \tilde{\beta}_{ii_d}^n u_{i,m}^n \end{pmatrix}.$$

Conclusively, (2.36) reads

$$\sum_i \mathbf{F}(\mathbf{u}_i^n) \tilde{\boldsymbol{\beta}}_{ii}^n = \sum_i (\mathbf{u}_i^n \cdot \tilde{\boldsymbol{\beta}}_{ii}^n) \mathbf{a}_i^n = \underbrace{\left(\mathbf{u}_1^n \cdot \tilde{\boldsymbol{\beta}}_{11}^n, \dots, \mathbf{u}_N^n \cdot \tilde{\boldsymbol{\beta}}_{NN}^n \right)}_{=: \mathbf{M}} \mathbf{A}^n \quad (2.37)$$

where $\mathbf{A}^n := \left(\underbrace{a_{1,1}^n, \dots, a_{1,d}^n}_{\mathbf{a}_1^n}, a_{2,1}^n, \dots, a_{N,d}^n \right)^T \in \mathbb{R}^{dN}$. Usually, the number of particles N will

be much greater than m , so we have to deal with a highly underdetermined linear system, which can only be solved if

$$\sum_i \mathbf{F}(\mathbf{u}_i^n) \tilde{\boldsymbol{\beta}}_{ii}^n \in \text{span}\{\mathbf{M}_j, j = 1, \dots, dN\},$$

where \mathbf{M}_j denotes the j th column of \mathbf{M} . So, if we have

$$\text{rank}(\mathbf{M}) = \text{rank} \left(\mathbf{M} \left| \sum_i \mathbf{F}(\mathbf{u}_i^n) \tilde{\boldsymbol{\beta}}_{ii}^n \right. \right),$$

Equation (2.37) can be solved using the pseudoinverse, for example. The solution of the above system in combination with the correction procedure in (2.35) leads to a conservative scheme in which the geometric coefficients have all the desired properties.

Unfortunately, solving a pseudoinverse in every time step can be very time consuming.

As an alternative, we show in Chapter 3 how to compute the geometrical coefficients exactly in an efficient way using B-splines. In our numerical computations, we mainly use the correction procedure due to Keck or B-splines.

2.3 Stability and entropy results

In this section, we will adopt some of the results for FVM stated in [Cha99] and [EGH95] for the FVPM, namely a L^∞ -stability result (see Lemma 2.3.1), a monotonicity and a discrete entropy inequality (see Lemmas 2.3.6 and 2.3.8) and, under suitable assumptions on the partition of unity, a weak discrete BV-stability result and L^1 -stability (see Lemmas 2.3.5 and 2.3.3) for the scalar multi-dimensional case. Crucial for the proof of the L^∞ -stability is the fact that u_i^{n+1} is a convex combination of u_j^n for all $j \in \mathcal{N}(i)$. This has the direct consequence that the FVPM preserves positive states (see Lemma 2.3.2).

The results mentioned above are essential elements in a Finite Volume convergence analysis for scalar equations. Therefore, they are first steps in a convergence analysis for the FVPM. Note that a L^∞ -stability result and monotonicity have already been shown in [Tel00], but only for the one-dimensional scalar case. In this section, all results apply to scalar conservation laws in arbitrary space dimensions. Moreover, the upper bound in Lemma 2.3.1 is clearly optimal.

Let us consider the scalar conservation law in several spatial dimensions

$$\begin{aligned} \partial_t u(\mathbf{x}, t) + \nabla \cdot \mathbf{F}(u(\mathbf{x}, t)) &= 0, & (\mathbf{x}, t) \in \mathbb{R}^d \times \mathbb{R}_+ \\ u(\mathbf{x}, 0) &= u_0(\mathbf{x}), & \mathbf{x} \in \mathbb{R}^d \end{aligned} \quad (2.38)$$

and the following scheme:

$$\begin{aligned} V_i^{n+1} u_i^{n+1} &= V_i^n u_i^n - \Delta t \sum_{j \in \mathcal{N}(i)} \|\beta_{ij}^n\| g_{ij}^n(u_i^n, u_j^n) \\ V_i^{n+1} &= V_i^n + \Delta t \sum_{j \in \mathcal{N}(i)} (\gamma_{ij}^n \dot{\mathbf{x}}_j^n - \gamma_{ji}^n \dot{\mathbf{x}}_i^n) \\ u_i^0 &= \frac{1}{V_i^0} \int_{\Omega} u_0(\mathbf{x}) \psi_i(\mathbf{x}, 0) d\mathbf{x} \\ V_i^0 &= \int_{\Omega} \psi(\mathbf{x}, 0) d\mathbf{x}, \end{aligned} \quad (2.39)$$

with the associated reconstruction formula

$$u_h(\mathbf{x}, t) = \sum_{n \in \mathbb{N}} \sum_{i \in \mathcal{T}} u_i^n \psi_i(\mathbf{x}, t) \mathbf{1}_{[t^n, t^{n+1})}(t), \quad (2.40)$$

where we denote the index set of all particles by \mathcal{T} . As before, we define

$$\Omega_i(t) := \text{supp}(\psi_i(\cdot, t))$$

and its discrete variant

$$\Omega_i^n := \text{supp}(\psi_i(\cdot, t^n)).$$

For simplicity, we will write for the numerical flux function $g_{ij}^n(u, v)$ or simply g_{ij}^n instead of $g(u, \mathbf{x}_i, v, \mathbf{x}_j, n_{ij}^n)$. In the case of non-moving particles, we will omit the index n in the numerical flux function, too. Furthermore, we make the following assumptions on the partition of unity, the numerical flux function, the time step size as well as the data.

1. There exists a maximal and a minimal number of overlapping particles

$$\exists \mu, M \in \mathbb{R} : \forall (\mathbf{x}, t) \in \mathbb{R}^d \times \mathbb{R}_+ \quad 1 \leq \mu \leq |\{i \in \mathcal{T} : \mathbf{x} \in \Omega_i(t)\}| \leq M. \quad (2.41)$$

The geometrical coefficients and the volumes are of order $d - 1$ and d , respectively. Moreover, the diameters of the patches are bounded. In particular, we assume

$$\|\mathcal{B}_{ij}^n\| \leq \frac{h^{d-1}}{\alpha}, \quad V_i^n \geq \alpha h^d, \quad (2.42)$$

$$\delta(\Omega_i^n) \leq 2h. \quad (2.43)$$

Further on, (2.25a)–(2.25e) have to be satisfied.

2. The numerical flux function $g_{ij}^n(u, v)$ satisfies:

$$\left. \begin{aligned} &\bullet g_{ij}^n(u, v) \text{ is Lipschitz continuous in both arguments } u \text{ and } v \\ &\quad \text{with Lipschitz constant } L, \\ &\bullet g_{ij}^n(u, v) \text{ has to be monotone nondecreasing in } u \text{ and monotone} \\ &\quad \text{nonincreasing in } v, \\ &\bullet g_{ij}^n(u, v) = -g_{ji}^n(v, u), \\ &\bullet \text{ In addition, } g_{ij}^n(u, v) \text{ has to satisfy the consistency condition (2.14).} \end{aligned} \right\} \quad (2.44)$$

3. The time step size Δt has to satisfy the CFL condition

$$\Delta t \leq \frac{(1 - \xi) \alpha^2 h}{2 L C_N} \quad (2.45)$$

with some $\xi \in (0, 1)$.

4. The assumptions on the initial data u_0 and on the flux function \mathbf{F} are:

$$\begin{aligned} &u_0 \in L^\infty(\mathbb{R}^d, \mathbb{R}) \text{ with } \|u_0\|_{L^\infty(\mathbb{R}^d, \mathbb{R})} \leq B \text{ for some } B \in \mathbb{R}, \\ &\mathbf{F} \in C^1(\mathbb{R}, \mathbb{R}^d) \text{ and } \frac{d\mathbf{F}}{ds} \text{ is locally Lipschitz continuous.} \end{aligned} \quad (2.46)$$

Note that under conditions (2.41)–(2.43), it is easy to show that every particle has only a finite number of neighbours. Conversely, it can be shown that a finite number of neighbours per particle implies that every point in the computational domain is covered by at most finitely many particles. The proof of this can be found in the Appendix.

L^∞ -stability

In the one-dimensional case and for scalar conservation laws, an L^∞ -stability analysis was made in [Tel00] and [Tel05] with a result of the form

$$\|u_h(\cdot, t)\|_{L^\infty(\mathbb{R}, \mathbb{R})} \leq e^{CT} \|u_0\|_{L^\infty(\mathbb{R}, \mathbb{R})} \quad \forall t < T$$

with $C, T \in \mathbb{R}$. In this paragraph, we will give an improved L^∞ -stability result for the scalar multi-dimensional case.

Lemma 2.3.1 (L^∞ -stability). *Under the assumptions (2.41)–(2.46) and (2.25a)–(2.25e), the scheme (2.39) with the reconstruction (2.40) is stable in the sense that*

$$\|u_h(\cdot, t)\|_{L^\infty(\mathbb{R}^d, \mathbb{R})} \leq \|u_0\|_{L^\infty(\mathbb{R}^d, \mathbb{R})} \quad \forall t > 0. \quad (2.47)$$

Proof. The first part of the proof is very close to the proof in [Cha99]. The second part relies on the partition of unity property of the family of functions $\psi_i, i \in \mathcal{T}$.

First, we will show that u_i^{n+1} is a convex combination of u_i^n and $u_j^n, j \in \mathcal{N}(i)$, i.e.

$$u_i^{n+1} = \sum_{j \in \mathcal{N}(i) \cup \{i\}} \lambda_j u_j^n,$$

where $0 \leq \lambda_j \leq 1, j \in \mathcal{N}(i) \cup \{i\}$ and $\sum_j \lambda_j = 1$. For this purpose, we write the scheme as

$$\begin{aligned} u_i^{n+1} &= \frac{V_i^n}{V_i^{n+1}} u_i^n - \frac{\Delta t}{V_i^{n+1}} \sum_{j \in \mathcal{N}(i)} \|\beta_{ij}^n\| g_{ij}^n(u_i^n, u_j^n) \\ &= u_i^n - \frac{V_i^{n+1} - V_i^n}{V_i^{n+1}} u_i^n - \frac{\Delta t}{V_i^{n+1}} \sum_{j \in \mathcal{N}(i)} \|\beta_{ij}^n\| g_{ij}^n(u_i^n, u_j^n) \\ &= u_i^n - \frac{\Delta t}{V_i^{n+1}} \left[\sum_{j \in \mathcal{N}(i)} \|\beta_{ij}^n\| g_{ij}^n(u_i^n, u_j^n) + u_i^n \sum_{j \in \mathcal{N}(i)} (\gamma_{ij}^n \bar{\mathbf{x}}_j^n - \gamma_{ji}^n \bar{\mathbf{x}}_i^n) \right] \\ &= u_i^n - \frac{\Delta t}{V_i^{n+1}} \left[\sum_{j \in \mathcal{N}(i)} \|\beta_{ij}^n\| g_{ij}^n(u_i^n, u_j^n) + u_i^n \sum_{j \in \mathcal{N}(i)} \beta_{ij}^n \bar{\mathbf{x}}_{ij}^n \right] \\ &= u_i^n - \frac{\Delta t}{V_i^{n+1}} \left[\sum_{j \in \mathcal{N}(i)} \|\beta_{ij}^n\| g_{ij}^n(u_i^n, u_j^n) - \sum_{j \in \mathcal{N}(i)} \|\beta_{ij}^n\| (\mathbf{F}(u_i^n) - u_i^n \bar{\mathbf{x}}_{ij}^n) \mathbf{n}_{ij}^n \right] \\ &= u_i^n - \frac{\Delta t}{V_i^{n+1}} \sum_{j \in \mathcal{N}(i)} \|\beta_{ij}^n\| (g_{ij}^n(u_i^n, u_j^n) - g_{ij}^n(u_i^n, u_i^n)) \\ &= u_i^n - \frac{\Delta t}{V_i^{n+1}} \sum_{\substack{j \in \mathcal{N}(i), \\ u_j^n \neq u_i^n}} \|\beta_{ij}^n\| \left(\frac{g_{ij}^n(u_i^n, u_j^n) - g_{ij}^n(u_i^n, u_i^n)}{u_i^n - u_j^n} \right) (u_i^n - u_j^n). \end{aligned}$$

Obviously, with

$$\lambda_i = 1 - \frac{\Delta t}{V_i^{n+1}} \sum_{\substack{j \in \mathcal{N}(i), \\ u_j^n \neq u_i^n}} \|\beta_{ij}^n\| \left(\frac{g_{ij}^n(u_i^n, u_j^n) - g_{ij}^n(u_i^n, u_i^n)}{u_i^n - u_j^n} \right),$$

$$\lambda_j = \frac{\Delta t}{V_i^{n+1}} \|\beta_{ij}^n\| \left(\frac{g_{ij}^n(u_i^n, u_j^n) - g_{ij}^n(u_i^n, u_i^n)}{u_i^n - u_j^n} \right), \quad j \in \mathcal{N}(i), u_j^n \neq u_i^n$$

we have $\sum_j \lambda_j = 1$.

Because of the monotonicity properties and the Lipschitz continuity of g_{ij}^n , we have

$$0 \leq \frac{g_{ij}^n(u_i^n, u_j^n) - g_{ij}^n(u_i^n, u_i^n)}{u_i^n - u_j^n} \leq L,$$

and this implies, thanks to the CFL condition and the assumptions made on V_i and β_{ij} , that

$$\frac{\Delta t}{V_i^{n+1}} \sum_{\substack{j \in \mathcal{N}(i), \\ u_j^n \neq u_i^n}} \|\beta_{ij}^n\| \left(\frac{g_{ij}^n(u_i^n, u_j^n) - g_{ij}^n(u_i^n, u_i^n)}{u_i^n - u_j^n} \right) \leq \frac{(1-\xi) \alpha^2 h}{2 L C_N} \frac{1}{\alpha h^d} L C_N \frac{h^{d-1}}{\alpha} = \frac{(1-\xi)}{2}.$$

Thus, we have

$$0 \leq \lambda_j \leq 1,$$

meaning that u_i^{n+1} is a convex combination of $u_j^n, j \in \mathcal{N}(i)$. This leads us to

$$\sup_{i \in \mathcal{T}} |u_i^{n+1}| \leq \sup_{i \in \mathcal{T}} |u_i^n|.$$

Now, using the partition of unity property of the family $(\psi_i)_{i \in \mathcal{T}}$, we obtain for $t \in [t^n, t^{n+1})$

$$\begin{aligned} \|u_h(\cdot, t)\|_{L^\infty} &= \text{ess sup}_{\mathbf{x} \in \mathbb{R}^d} \left| \sum_{i \in \mathcal{T}} u_i^n \psi_i(\mathbf{x}, t) \right| \leq \sup_{i \in \mathcal{T}} |u_i^n| \\ &\leq \sup_{i \in \mathcal{T}} |u_i^{n-1}| \leq \dots \leq \sup_{i \in \mathcal{T}} |u_i^0| \leq \|u_0\|_{L^\infty}. \end{aligned}$$

The last inequality follows from the fact that for $A, B \in \mathbb{R}$, we have

$$A \leq \frac{1}{V_i^0} \int_{\mathbb{R}^d} u_0(\mathbf{x}) \psi_i(\mathbf{x}, 0) d\mathbf{x} \leq B \quad (2.48)$$

if $A \leq u_0(\mathbf{x}) \leq B$ for all $\mathbf{x} \in \mathbb{R}^d$. □

Note that we can set $\xi = 0$ in the proof above, but not for the weak BV-stability.

Positivity and L^1 -estimates

The first lemma in this paragraph concerns the preservation of positive states, as this is often desired in the numerical treatment of problems which are known to have a positive solution. This lemma follows directly from the results of the last paragraph.

Lemma 2.3.2 (Positivity). *Under the assumptions of Lemma 2.3.1, the FVPM (2.39)–(2.40) preserves positivity, i.e. $u_i^n \geq 0$ for all $i \in \mathbb{Z}$ implies $u_i^{n+1} \geq 0$ for all $i \in \mathbb{Z}$. If, in particular, $u_0(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathbb{R}^d$, the same is true for the coefficients u_i^n for all $i \in \mathbb{Z}$, $n \in \mathbb{N}$.*

Proof. In the proof of Lemma 2.3.1, we have shown that u_i^{n+1} is a convex combination of $u_j^n, j \in \mathcal{N}(i)$ and the assertion follows immediately from (2.48). \square

Now we can show that the FVPM is L^1 -diminishing under the condition that $u_0(\mathbf{x})$ has the same sign for all $\mathbf{x} \in \mathbb{R}^d$.

Lemma 2.3.3 (L^1 -estimates). *Consider the scheme (2.39)–(2.40) and assume $u_0 \in L^1(\mathbb{R}^d, \mathbb{R})$ and $u_0(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathbb{R}^d$ or $u_0(\mathbf{x}) \leq 0$ for all $\mathbf{x} \in \mathbb{R}^d$. Suppose further that the assumptions of Lemma 2.3.2 hold. Then we have the following L^1 -estimate:*

$$\|u_h(\cdot, t^n)\|_{L^1(\mathbb{R}^d, \mathbb{R})} \leq \|u_0\|_{L^1(\mathbb{R}^d, \mathbb{R})}$$

for all $n \in \mathbb{N}$.

Proof. Because of Lemma 2.3.2, the coefficients u_i^n will have the same sign for all $i \in \mathbb{Z}$ and all $n \in \mathbb{N}$. Thus, due to the conservativity we have

$$\|u_h(\cdot, t^n)\|_{L^1(\mathbb{R}^d, \mathbb{R})} = \|u_h(\cdot, t^{n-1})\|_{L^1(\mathbb{R}^d, \mathbb{R})} = \cdots = \|u_h(\cdot, 0)\|_{L^1(\mathbb{R}^d, \mathbb{R})}.$$

It remains to show $\|u_h(\cdot, 0)\|_{L^1(\mathbb{R}^d, \mathbb{R})} \leq \|u_0\|_{L^1(\mathbb{R}^d, \mathbb{R})}$. We have

$$\begin{aligned} \|u_h(\cdot, 0)\|_{L^1(\mathbb{R}^d, \mathbb{R})} &= \int_{\mathbb{R}^d} \left| \sum_{i \in \mathbb{Z}} u_i^0 \psi_i(\mathbf{x}) \right| d\mathbf{x} \leq \sum_{i \in \mathbb{Z}} |u_i^0| \int_{\mathbb{R}^d} \psi_i(\mathbf{x}) d\mathbf{x} \\ &= \sum_{i \in \mathbb{Z}} \left| \frac{\int_{\mathbb{R}^d} u_0(\mathbf{y}) \psi_i(\mathbf{y}) d\mathbf{y}}{V_i} \right| \int_{\mathbb{R}^d} \psi_i(\mathbf{x}) d\mathbf{x} = \sum_{i \in \mathbb{Z}} \left| \int_{\mathbb{R}^d} u_0(\mathbf{y}) \psi_i(\mathbf{y}) d\mathbf{y} \right| \\ &\leq \sum_{i \in \mathbb{Z}} \int_{\mathbb{R}^d} |u_0(\mathbf{y})| \psi_i(\mathbf{y}) d\mathbf{y} = \|u_0\|_{L^1(\mathbb{R}^d, \mathbb{R})}. \end{aligned}$$

\square

In the next part, we will show so-called weak BV-estimates for non-moving particles. The proof follows [Cha99] and [EGH00] closely, where the authors formulated the weak BV-stability for the Finite Volume setting on unstructured grids.

Weak BV-estimates

Before we show the weak BV-stability, we state a helpful lemma that can be found in [EGGH98].

Lemma 2.3.4. *Let $g : \mathbb{R} \rightarrow \mathbb{R}$ be a monotone, Lipschitz continuous function with Lipschitz constant $G > 0$. Then g satisfies*

$$\left| \int_a^b (g(x) - g(a)) dx \right| \geq \frac{1}{2G} (g(b) - g(a))^2 \quad \forall a, b \in \mathbb{R}.$$

Proof. See [EGGH98]. □

Furthermore, we define

$$\begin{aligned} N_T &:= \max\{n \in \mathbb{N} : n \leq \frac{T}{\Delta t} - 1\} \\ T_R &:= \{i \in \mathcal{T} : \Omega_i \subset B(0, R)\} \\ E_R^n &:= \{(i, j) \in \mathcal{T}^2 : i \in T_R \text{ or } j \in T_R, j \in \mathcal{N}(i), u_i^n > u_j^n\} \\ \sigma_{ij} &:= \Omega_i \cap \Omega_j \\ \mathfrak{S} &:= \{\sigma_{ij} : \sigma_{ij} \subset B(0, R) \setminus B(0, R - 2h)\}. \end{aligned}$$

where $T > 0$ denotes a fixed time and $R > 0$. With the assumptions (2.41)–(2.43) made above, one can ensure that there exist constants C_1, C_2, C_3 , such that

$$\begin{aligned} |E_R^n| &\leq C_1 h^{-d}, \\ |T_R| &\leq C_2 h^{-d}, \\ |\mathfrak{S}| &\leq C_3 h^{1-d}. \end{aligned}$$

Details on these conclusions can be found in the Appendix in Proposition 5.4.4.

Now we will show an estimate that is often called weak BV-stability, see for example [Cha99], [EGH95] and [EGGH98], because it provides estimates of order $h^{-1/2}$ on the discrete derivatives of the numerical solution in space and time.

Lemma 2.3.5 (Weak BV-stability). *Assume non-moving particles. Under the conditions (2.41)–(2.46), (2.25a)–(2.25e) and for $R > 0$, $T > 0$, there exists a constant $C_{bv} \in \mathbb{R}$ depending only on $F, u_0, L, C_N, \alpha, \xi, R$ and T such that*

$$\begin{aligned} \sum_{n=0}^{N_T} \Delta t \sum_{(i,j) \in E_R^n} \|\beta_{ij}\| &\left[\max_{u_j^n \leq c \leq d \leq u_i^n} (g_{ij}(d, c) - g_{ij}(d, d)) \right. \\ &\left. + \max_{u_j^n \leq c \leq d \leq u_i^n} (g_{ij}(d, c) - g_{ij}(c, c)) \right] \leq \frac{C_{bv}}{\sqrt{h}} \end{aligned} \quad (2.49)$$

and

$$\sum_{n=0}^{N_T} \sum_{i \in T_R} V_i |u_i^{n+1} - u_i^n| \leq \frac{C_{bv}}{\sqrt{h}} \quad \forall h < R. \quad (2.50)$$

Proof. We write the scheme as

$$V_i (u_i^{n+1} - u_i^n) + \Delta t \sum_{j \in \mathcal{N}(i)} \|\beta_{ij}\| (g_{ij}(u_i^n, u_j^n) - g_{ij}(u_i^n, u_i^n)) = 0,$$

multiply it by u_i^n and sum over n and $i \in T_R$. Defining

$$B_1 := \sum_{n=0}^{N_T} \sum_{i \in T_R} V_i u_i^n (u_i^{n+1} - u_i^n)$$

and

$$B_2 := \sum_{n=0}^{N_T} \sum_{i \in T_R} \Delta t \sum_{j \in \mathcal{N}(i)} u_i^n \|\beta_{ij}\| (g_{ij}(u_i^n, u_j^n) - g_{ij}(u_i^n, u_i^n))$$

yields $B_1 + B_2 = 0$. Now we rewrite the term B_2 and get

$$\begin{aligned} B_2 = & \sum_{n=0}^{N_T} \sum_{(i,j) \in E_R^n} \Delta t \left[u_i^n \|\beta_{ij}\| (g_{ij}(u_i^n, u_j^n) - g_{ij}(u_i^n, u_i^n)) \right. \\ & \left. + u_j^n \|\beta_{ji}\| (g_{ji}(u_j^n, u_i^n) - g_{ji}(u_j^n, u_j^n)) \right] \\ & - \sum_{n=0}^{N_T} \sum_{i \in T_R} \sum_{\substack{j \in \mathcal{N}(i) \\ j \notin T_R}} \Delta t u_j^n \|\beta_{ji}\| (g_{ji}(u_j^n, u_i^n) - g_{ji}(u_j^n, u_j^n)). \end{aligned}$$

By defining

$$\begin{aligned} B_3 := & \sum_{n=0}^{N_T} \sum_{(i,j) \in E_R^n} \Delta t \left[u_i^n \|\beta_{ij}\| (g_{ij}(u_i^n, u_j^n) - g_{ij}(u_i^n, u_i^n)) \right. \\ & \left. + u_j^n \|\beta_{ji}\| (g_{ji}(u_j^n, u_i^n) - g_{ji}(u_j^n, u_j^n)) \right], \end{aligned}$$

we can easily estimate

$$\begin{aligned} |B_3 - B_2| &= \left| \sum_{n=0}^{N_T} \sum_{i \in T_R} \sum_{\substack{j \in \mathcal{N}(i) \\ j \notin T_R}} \Delta t u_j^n \|\beta_{ij}\| (g_{ij}(u_i^n, u_j^n) - g_{ij}(u_j^n, u_j^n)) \right| \\ &\leq \sum_{n=0}^{N_T} \Delta t \sum_{i \in T_R} \sum_{\substack{j \in \mathcal{N}(i) \\ j \notin T_R}} \underbrace{|u_j^n|}_{\leq B} \underbrace{\|\beta_{ij}\|}_{\leq h^{d-1} \alpha^{-1}} \underbrace{|(g_{ij}(u_i^n, u_j^n) - g_{ij}(u_j^n, u_j^n))|}_{\leq 2LB} \\ &\leq 2 \sum_{n=0}^{N_T} \Delta t \sum_{i \in T_R} \sum_{\substack{j \in \mathcal{N}(i) \\ j \notin T_R}} B^2 L \frac{h^{d-1}}{\alpha} \\ &\leq \frac{2TB^2L}{\alpha} h^{d-1} \underbrace{|\mathfrak{S}|}_{\leq C_3 h^{1-d}} \leq \tilde{C}_2, \end{aligned}$$

where

$$\tilde{C}_2 := \frac{2TB^2L}{\alpha} C_3.$$

By introducing the function

$$\begin{aligned} z_{ij}(x) &:= \int_0^x \|\beta_{ij}\| s \left(\frac{\partial g_{ij}}{\partial u}(s, s) + \frac{\partial g_{ij}}{\partial v}(s, s) \right) ds \\ &= \int_0^x \|\beta_{ij}\| s \frac{d}{ds} g_{ij}(s, s) ds \\ &= \|\beta_{ij}\| \left[x g_{ij}(x, x) - \int_0^x g_{ij}(s, s) ds \right] \end{aligned}$$

and using

$$\begin{aligned} z_{ij}(b) - z_{ij}(a) &= \|\beta_{ij}\| \left[b(g_{ij}(b, b) - g_{ij}(a, b)) - a(g_{ij}(a, a) - g_{ij}(a, b)) \right. \\ &\quad \left. - \int_a^b (g_{ij}(s, s) - g_{ij}(a, b)) \, ds \right], \end{aligned}$$

we can write B_3 as

$$B_3 = B_4 + B_5,$$

where

$$B_4 := \sum_{n=0}^{N_T} \sum_{(i,j) \in E_R^n} \Delta t [z_{ij}(u_j^n) - z_{ij}(u_i^n)]$$

and

$$B_5 := \sum_{n=0}^{N_T} \sum_{(i,j) \in E_R^n} \Delta t \|\beta_{ij}\| \int_{u_i^n}^{u_j^n} (g_{ij}(s, s) - g_{ij}(u_i^n, u_j^n)) \, ds.$$

Because of the consistency of the numerical flux g_{ij} and the assumptions made on the geometrical coefficients β_{ij} , we have

$$\sum_{j \in \mathcal{N}(i)} z_{ij}(x) = \sum_{j \in \mathcal{N}(i)} \|\beta_{ij}\| \left(x g_{ij}(x, x) - \int_0^x g_{ij}(s, s) \, ds \right) = 0,$$

and therefore

$$\begin{aligned} B_4 &= - \sum_{n=0}^{N_T} \sum_{(i,j) \in E_R^n} \Delta t \left[z_{ij}(u_i^n) - \underbrace{z_{ij}(u_j^n)}_{=-z_{ji}(u_j^n)} \right] \\ &= - \sum_{n=0}^{N_T} \sum_{(i,j) \in E_R^n} \Delta t [z_{ij}(u_i^n) + z_{ji}(u_j^n)] \\ &= - \sum_{n=0}^{N_T} \Delta t \left[\sum_{i \in T_R} \underbrace{\sum_{j \in \mathcal{N}(i)} z_{ij}(u_i^n)}_{=0} + \sum_{i \in T_R} \sum_{\substack{j \in \mathcal{N}(i) \\ j \notin T_R}} z_{ji}(u_j^n) \right] \\ &= \sum_{n=0}^{N_T} \Delta t \sum_{i \in T_R} \sum_{\substack{j \in \mathcal{N}(i) \\ j \notin T_R}} z_{ij}(u_j^n) \\ &= \sum_{n=0}^{N_T} \Delta t \sum_{i \in T_R} \sum_{\substack{j \in \mathcal{N}(i) \\ j \notin T_R}} \int_0^{u_j^n} \|\beta_{ij}\| s \frac{d}{ds} g_{ij}(s, s) \, ds. \end{aligned}$$

This implies

$$|B_4| \leq T L B^2 \frac{h^{d-1}}{\alpha} \underbrace{|\mathfrak{G}|}_{\leq C_3 h^{1-d}} \leq T L \frac{B^2}{\alpha} C_3 =: \tilde{C}_4.$$

For the estimate of B_5 , we use the monotonicity of the numerical flux and Lemma 2.3.4 to get:

$$B_5 \geq \frac{1}{4L} \sum_{n=0}^{N_T} \Delta t \sum_{(i,j) \in E_R^n} \left[\|\beta_{ij}\| \max_{u_j^n \leq c \leq d \leq u_i^n} (g_{ij}(d, c) - g_{ij}(d, d))^2 + \|\beta_{ij}\| \max_{u_i^n \leq c \leq d \leq u_j^n} (g_{ij}(d, c) - g_{ij}(c, c))^2 \right].$$

At the end, we get an estimate of B_1 using the technique of summation by parts:

$$\begin{aligned} B_1 &= \sum_{n=0}^{N_T} \sum_{i \in T_R} V_i u_i^n (u_i^{n+1} - u_i^n) \\ &\geq \alpha h^d \left(\frac{1}{2} \sum_{n=0}^{N_T} \sum_{i \in T_R} u_i^n (u_i^{n+1} - u_i^n) + \frac{1}{2} \sum_{n=0}^{N_T} \sum_{i \in T_R} u_i^n (u_i^{n+1} - u_i^n) \right) \\ &= \alpha h^d \left(\frac{1}{2} \sum_{n=0}^{N_T} \sum_{i \in T_R} u_i^n (u_i^{n+1} - u_i^n) - \frac{1}{2} \sum_{n=0}^{N_T} \sum_{i \in T_R} u_i^{n+1} (u_i^{n+1} - u_i^n) \right. \\ &\quad \left. - \frac{1}{2} \underbrace{\sum_{i \in T_R} (u_i^0)^2}_{\leq \frac{B^2}{2} C_2 h^{-d}} + \frac{1}{2} \underbrace{\sum_{i \in T_R} (u_i^{N_T+1})^2}_{\geq 0} \right) \\ &\geq -\alpha h^d \frac{1}{2} \sum_{n=0}^{N_T} \sum_{i \in T_R} (u_i^{n+1} - u_i^n)^2 - \tilde{C}_5, \end{aligned}$$

where

$$\tilde{C}_5 := \frac{\alpha B^2 C_2}{2}.$$

Applying the Cauchy-Schwarz inequality to the scheme (2.39) yields

$$\begin{aligned} (u_i^{n+1} - u_i^n)^2 &= \frac{\Delta t^2}{V_i^2} \left(\sum_{j \in \mathcal{N}(i)} \|\beta_{ij}\| (g_{ij}(u_i^n, u_j^n) - g_{ij}(u_i^n, u_i^n)) \right)^2 \\ &\leq \frac{h^{-2d}}{\alpha^2} \frac{(1-\xi) \alpha^2 h}{2 L C_N} \frac{h^{d-1}}{\alpha} \sum_{j \in \mathcal{N}(i)} \Delta t \|\beta_{ij}\| (g_{ij}(u_i^n, u_j^n) - g_{ij}(u_i^n, u_i^n))^2 \underbrace{\sum_{l \in \mathcal{N}(i)} 1}_{\leq C_N} \\ &\leq \frac{(1-\xi)}{2 \alpha L} h^{-d} \sum_{j \in \mathcal{N}(i)} \Delta t \|\beta_{ij}\| (g_{ij}(u_i^n, u_j^n) - g_{ij}(u_i^n, u_i^n))^2, \end{aligned}$$

and therefore

$$\begin{aligned}
B_1 &\geq -\alpha h^d \frac{1}{2} \frac{(1-\xi)}{2\alpha L} h^{-d} \sum_{n=0}^{N_T} \sum_{i \in T_R} \sum_{j \in \mathcal{N}(i)} \Delta t \|\beta_{ij}\| (g_{ij}(u_i^n, u_j^n) - g_{ij}(u_i^n, u_i^n))^2 - \tilde{C}_5 \\
&= -\frac{(1-\xi)}{4L} \sum_{n=0}^{N_T} \sum_{i \in T_R} \sum_{j \in \mathcal{N}(i)} \Delta t \|\beta_{ij}\| (g_{ij}(u_i^n, u_j^n) - g_{ij}(u_i^n, u_i^n))^2 - \tilde{C}_5 \\
&= -\frac{(1-\xi)}{4L} \sum_{n=0}^{N_T} \left[\sum_{(i,j) \in E_R^n} \Delta t \|\beta_{ij}\| \left((g_{ij}(u_i^n, u_j^n) - g_{ij}(u_i^n, u_i^n))^2 \right. \right. \\
&\quad \left. \left. + (g_{ji}(u_j^n, u_i^n) - g_{ji}(u_j^n, u_j^n))^2 \right) - \sum_{i \in T_R} \sum_{\substack{j \in \mathcal{N}(i) \\ j \notin T_R}} \Delta t \|\beta_{ij}\| (g_{ji}(u_j^n, u_i^n) - g_{ji}(u_j^n, u_j^n))^2 \right] \\
&\quad - \tilde{C}_5 \\
&\geq -\frac{(1-\xi)}{4L} \sum_{n=0}^{N_T} \sum_{(i,j) \in E_R^n} \Delta t \|\beta_{ij}\| \left((g_{ij}(u_i^n, u_j^n) - g_{ij}(u_i^n, u_i^n))^2 \right. \\
&\quad \left. + (g_{ij}(u_i^n, u_j^n) - g_{ij}(u_j^n, u_j^n))^2 \right) - \tilde{C}_5 \\
&\geq -\frac{(1-\xi)}{4L} \sum_{n=0}^{N_T} \sum_{(i,j) \in E_R^n} \Delta t \|\beta_{ij}\| \left(\max_{u_j^n \leq c \leq d \leq u_i^n} (g_{ij}(d, c) - g_{ij}(d, d))^2 \right. \\
&\quad \left. + \max_{u_j^n \leq c \leq d \leq u_i^n} (g_{ij}(d, c) - g_{ij}(c, c))^2 \right) - \tilde{C}_5.
\end{aligned}$$

Defining

$$\begin{aligned}
A_1 &:= \sum_{n=0}^{N_T} \sum_{(i,j) \in E_R^n} \Delta t \|\beta_{ij}\| \left(\max_{u_j^n \leq c \leq d \leq u_i^n} (g_{ij}(d, c) - g_{ij}(d, d))^2 \right. \\
&\quad \left. + \max_{u_j^n \leq c \leq d \leq u_i^n} (g_{ij}(d, c) - g_{ij}(c, c))^2 \right),
\end{aligned}$$

we now have $B_1 + B_2 = 0$, $|B_3 - B_2| \leq \tilde{C}_2$, $B_3 = B_4 + B_5$, $|B_4| \leq \tilde{C}_4$, $B_5 \geq \frac{A_1}{4L}$ and $B_1 \geq -\frac{(1-\xi)}{4L} A_1 - \tilde{C}_5$ and thus

$$\begin{aligned}
A_1 &\leq 4L B_5 = 4L (B_3 - B_4) = 4L (B_3 - B_2 - B_1 - B_4) \\
&\leq 4L \left(|B_3 - B_2| + |B_4| + \tilde{C}_5 + \frac{(1-\xi)}{4L} A_1 \right) \\
&\leq 4L \left(\tilde{C}_2 + \tilde{C}_4 + \tilde{C}_5 + \frac{(1-\xi)}{4L} A_1 \right).
\end{aligned}$$

It follows

$$A_1 \leq \frac{4L}{\xi} \left(\tilde{C}_2 + \tilde{C}_4 + \tilde{C}_5 \right) =: \tilde{C}_6.$$

Next, we apply the Cauchy-Schwarz inequality two times to each term of the left hand side of (2.49) and get

$$\begin{aligned} & \sum_{n=0}^{N_T} \Delta t \sum_{(i,j) \in E_R^n} \|\beta_{ij}\| \max_{u_j^n \leq c \leq d \leq u_i^n} \left(g_{ij}(d, c) - g_{ij}(d, d) \right) \\ & \leq \sum_{n=0}^{N_T} \Delta t \left(\sum_{(i,j) \in E_R^n} \|\beta_{ij}\|^2 \max_{u_j^n \leq c \leq d \leq u_i^n} \left(g_{ij}(d, c) - g_{ij}(d, d) \right)^2 \right)^{1/2} \underbrace{\left(\sum_{(i,j) \in E_R^n} 1 \right)^{1/2}}_{\leq (C_1 h^{-d})^{1/2}} \\ & \leq C_1^{1/2} h^{-d/2} \left(\sum_{n=0}^{N_T} \Delta t^2 \right)^{1/2} \left(\sum_{n=0}^{N_T} \sum_{(i,j) \in E_R^n} \|\beta_{ij}\|^2 \max_{u_j^n \leq c \leq d \leq u_i^n} \left(g_{ij}(d, c) - g_{ij}(d, d) \right)^2 \right)^{1/2} \\ & = C_1^{1/2} h^{-d/2} \left(\sum_{n=0}^{N_T} \Delta t \right)^{1/2} \left(\sum_{n=0}^{N_T} \sum_{(i,j) \in E_R^n} \Delta t \|\beta_{ij}\|^2 \max_{u_j^n \leq c \leq d \leq u_i^n} \left(g_{ij}(d, c) - g_{ij}(d, d) \right)^2 \right)^{1/2} \\ & \leq C_1^{1/2} h^{-d/2} T^{1/2} \frac{h^{(d-1)/2}}{\sqrt{\alpha}} \tilde{C}_6^{1/2} \\ & \leq \tilde{C}_7 h^{-1/2}, \end{aligned}$$

where

$$\tilde{C}_7 := \left(\frac{T C_1 \tilde{C}_6}{\alpha} \right)^{1/2}.$$

To get the estimate (2.50), we state that

$$V_i |u_i^{n+1} - u_i^n| \leq \Delta t \sum_{j \in \mathcal{N}(i)} \|\beta_{ij}\| |g_{ij}(u_i^n, u_j^n) - g_{ij}(u_i^n, u_i^n)|$$

and thus

$$\begin{aligned} & \sum_{n=0}^{N_T} \sum_{i \in T_R} V_i |u_i^{n+1} - u_i^n| \leq \sum_{n=0}^{N_T} \sum_{i \in T_R} \sum_{j \in \mathcal{N}(i)} \Delta t \|\beta_{ij}\| |g_{ij}(u_i^n, u_j^n) - g_{ij}(u_i^n, u_i^n)| \\ & = \sum_{n=0}^{N_T} \sum_{(i,j) \in E_R^n} \Delta t \left[\|\beta_{ij}\| |g_{ij}(u_i^n, u_j^n) - g_{ij}(u_i^n, u_i^n)| + \|\beta_{ji}\| |g_{ji}(u_j^n, u_i^n) - g_{ji}(u_j^n, u_j^n)| \right] \\ & \quad - \sum_{n=0}^{N_T} \sum_{i \in T_R} \sum_{\substack{j \in \mathcal{N}(i) \\ j \notin T_R}} \Delta t \|\beta_{ji}\| |g_{ji}(u_j^n, u_i^n) - g_{ji}(u_j^n, u_j^n)| \\ & \leq 2 \left(\tilde{C}_7 h^{-1/2} + \frac{T L B C_3}{\alpha} \right) \\ & \leq \tilde{C}_8 h^{-1/2}, \end{aligned}$$

where

$$\tilde{C}_8 := 2 \left(\tilde{C}_7 + \frac{T L B C_3}{\alpha} R^{1/2} \right).$$

The assertion follows by setting $C_{bv} := \tilde{C}_8$. \square

Discrete entropy inequality

Next, we will show a discrete entropy inequality for scheme (2.39) using Kruzkov's entropy from Definition 1.1.6. In order to do this, we need that the scheme is monotone in the classical sense, meaning that the function

$$G(u_i^n, u_j, j \in \mathcal{N}(i)) := \frac{V_i^n}{V_i^{n+1}} u_i^n - \frac{\Delta t}{V_i^{n+1}} \sum_{j \in \mathcal{N}(i)} \|\beta_{ij}^n\| g_{ij}^n(u_i^n, u_j^n) \quad (2.51)$$

is monotonically nondecreasing in each argument. In [Tel05] and [Tel00], the FVPM was shown to be monotone for scalar conservation laws in one spatial dimension for a differentiable numerical flux function and under the assumption of a slightly different CFL condition (see (2.34)). With the techniques used in the proof of the L^∞ -stability, it is easy to show the monotonicity without these restrictions.

Lemma 2.3.6 (Monotonicity). *Under the assumptions (2.41)–(2.46), the FVPM (2.39) is monotone.*

Proof. Since the numerical flux function $g_{ij}(u, v)$ is monotonically nonincreasing in v , the function G is monotonically nondecreasing in u_j^n for all $j \in \mathcal{N}(i)$ and $j \neq i$. To show the monotonicity of G in the first argument, let us assume $\tilde{u}_i^n < u_i^n$. With the same arguments as in the proof of Lemma 2.3.1, we conclude that

$$\begin{aligned} & \frac{V_i^{n+1}}{V_i^n} \left(G(\tilde{u}_i^n, u_j, j \in \mathcal{N}(i)) - G(u_i^n, u_j, j \in \mathcal{N}(i)) \right) \\ &= \left[1 - \frac{\Delta t}{V_i^n} \sum_{j \in \mathcal{N}(i)} \|\beta_{ij}^n\| \frac{g_{ij}(\tilde{u}_i^n, u_j^n) - g_{ij}(u_i^n, u_j^n)}{\tilde{u}_i^n - u_i^n} \right] (\tilde{u}_i^n - u_i^n) \leq 0 \end{aligned}$$

under the CFL condition

$$\Delta t \leq \frac{\alpha^2 h}{L C_N}. \quad (2.52)$$

Thus, a time step size Δt satisfying (2.52) is enough to ensure monotonicity of the scheme. Note that condition (2.45) is more restrictive and hence monotonicity follows. \square

Remark 2.3.7. *Note that we need the more restrictive CFL condition (2.45) only for the weak BV-stability result. For the monotonicity as well as L^∞ -stability, it is sufficient to assume (2.52).*

Now, we can formulate a discrete entropy inequality. It is the discrete analogue to Definition 1.1.6.

Lemma 2.3.8 (Discrete entropy inequality). *Under the assumptions (2.41)–(2.46), for all $c \in \mathbb{R}$, $i \in \mathcal{T}$ and $n \in \mathbb{N}$ we have the following discrete entropy inequality for the scheme (2.39).*

$$\frac{|u_i^{n+1} - c|V_i^{n+1} - |u_i^n - c|V_i^n}{\Delta t} + \sum_{j \in \mathcal{N}(i)} \|\beta_{ij}^n\| \left[g_{ij}^n(u_i^n \top c, u_j^n \top c) - g_{ij}^n(u_i^n \perp c, u_j^n \perp c) \right] \leq 0. \quad (2.53)$$

Proof. If we write the scheme as

$$u_i^{n+1} = G_{ij}^n(u_i^n, u_j^n, j \in \mathcal{N}(i))$$

with G from (2.51), we have a monotonically nondecreasing function (under the CFL condition) which satisfies $G(c, \dots, c) = c \forall c \in \mathbb{R}$ and

$$\begin{aligned} G_{ij}^n(u_i^n \top c, u_j^n \top c, j \in \mathcal{N}(i)) &\geq u_i^{n+1} \top c, \\ G_{ij}^n(u_i^n \perp c, u_j^n \perp c, j \in \mathcal{N}(i)) &\leq u_i^{n+1} \perp c. \end{aligned}$$

Subtracting the second inequality from the first one, we get (2.53). \square

Remark 2.3.9. *Note that in the case of non-moving particles, (2.53) is equivalent to the well-known discrete entropy inequality for Kruzkov's entropy-entropy-flux pair for a standard FVM*

$$\frac{|u_i^{n+1} - c| - |u_i^n - c|}{\Delta t} + \frac{1}{V_i} \sum_{j \in \mathcal{N}(i)} \|\Gamma_{ij}\| \left[g_{ij}(u_i^n \top c, u_j^n \top c) - g_{ij}(u_i^n \perp c, u_j^n \perp c) \right] \leq 0,$$

where Γ_{ij} denotes the area of the surface between two neighbouring cells i and j , see [Kru70], [Cha99] and [EGH00].

In this section, we have shown L^∞ -stability, positivity, a generalized monotonicity result and a discrete entropy inequality for the FVPM applied to scalar conservation laws in arbitrary space dimensions. Thus, according to Proposition 1.2.4, we have shown the convergence of a subsequence in the nonlinear weak-* sense. Moreover, we proved a weak BV-stability and L^1 -stability under suitable assumptions. All these results are important tools in a convergence analysis for the FVPM.

It remains to show an entropy inequality for the reconstructed solution that naturally includes some error terms. These error terms can be controlled with the help of the weak BV-stability, see [Cha99] and [EGH00]. Passing to the limit in this entropy inequality together with the notion of nonlinear weak-* convergence yields that the nonlinear weak-* limit of u_h is an entropy process solution (see Definition 1.1.9). It is well known that the entropy process solution of problem (2.38) coincides with the unique entropy solution. In particular, it does not depend on the additional parameter α , thus strong convergence in $L_{loc}^p(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R})$ follows, compare e.g. [Cha99] and [EGH00].

Chapter 3

The FVPM with B-splines

Up to now, we have computed the coefficients numerically and corrected them afterwards or added a correction term directly to the right hand side of the scheme to force the scheme to be conservative without losing other important properties like the preservation of constant states. In the last section, we have seen that this can lead to the disadvantage that we cannot choose an arbitrary particle motion any more. Of course, the best way to ensure the desired properties of the scheme is still an exact evaluation of the respective integrals.

Using Shepard's method of constructing a partition of unity, as described in Section (2.1), leads to piecewise rational functions whose exact integration can be done efficiently in some cases, but is too expensive in general. A possibility to compute the geometrical coefficients exactly is shown in [QN11], where the authors take piecewise constant and overlapping shape functions. In this way, the computational effort is reduced considerably, but the drawback is clearly the restriction to piecewise constant particles.

In this section, we will use B-splines, which are given as piecewise polynomials and already form a partition of unity, to compute the geometrical coefficients exactly.

B-Splines are a standard approximation tool and are therefore used in many numerical methods such as FEM or meshless methods, see for example [Höl03], [ZP13] or [WA10]. But, to the author's knowledge, there exists no meshless method where moving B-splines are used as particles without any renormalization technique. In our scheme, B-splines are used as test and Ansatz functions. Since B-splines already build a partition of unity, the integrands in (2.12) are piecewise polynomials and thus, the geometrical coefficients can be computed exactly. Because the B-splines are allowed to move, we have to take into account additional terms. But, as we will see later, these terms can be computed exactly, too.

3.1 Introduction

We begin with the definition of B-splines and their main properties. For details as well as for further applications of B-splines, we refer to [DeB01] or [SW05].

Definition 3.1.1 (B-splines in one spatial dimension over \mathbb{R}). *Consider a knot sequence $\{x_i\}_{i \in \mathbb{Z}}$ with $x_i \leq x_{i+1}$ which builds a partition of \mathbb{R} , i.e. $\bigcup_{i \in \mathbb{Z}} [x_i, x_{i+1}) = \mathbb{R}$, and the*

functions

$$w_i^m(x) = \begin{cases} \frac{x - x_i}{x_{i+m} - x_i}, & \text{if } x_i < x_{i+m} \\ 0, & \text{else} \end{cases} \quad (3.1)$$

with $i \in \mathbb{Z}$ and $m \in \mathbb{N}$. The functions B_i^m defined by

$$B_i^0(x) = \begin{cases} 1, & \text{if } x \in [x_i, x_{i+1}) \\ 0, & \text{else} \end{cases} \quad (3.2)$$

and

$$B_i^m(x) = w_i^m(x) B_i^{m-1}(x) + (1 - w_{i+1}^m(x)) B_{i+1}^{m-1}(x) \quad (3.3)$$

with $i \in \mathbb{Z}$ and $m \in \mathbb{N}$ are called B-splines of degree m associated with the knot sequence $\{x_i\}_{i \in \mathbb{Z}}$.

Thus, B_i^0 are the characteristic functions over the intervals $[x_i, x_{i+1})$ and B_i^1 are the piecewise linear hat functions.

Originally, Curry and Schoenberg introduced B-splines using divided differences [CS66] and a different normalization ending up with

$$M_i^m(x) := (m+1) [x_i, \dots, x_{i+m+1}] (\cdot - x)_+^m, \quad x \in \mathbb{R}, \quad (3.4)$$

where

$$[x_i, \dots, x_{i+n}] f$$

denotes the n -th divided difference of a function $f : \mathbb{R} \rightarrow \mathbb{R}$ which is defined as the leading coefficient of the interpolation polynomial of degree n that interpolates the data $(x_i, f(x_i)), \dots, (x_{i+n}, f(x_{i+n}))$. Further on, $(\cdot)_+^n : \mathbb{R} \rightarrow \mathbb{R}_+$ is defined as

$$x_+^0 := \begin{cases} 1, & x > 0 \\ 0, & \text{else} \end{cases}$$

for $n = 0$ and

$$x_+^n := (\max(0, x))^n \quad \text{for } n \geq 1.$$

One can show that

$$M_i^m(x) = \frac{m+1}{x_{i+m+1} - x_i} B_i^m(x). \quad (3.5)$$

See [DeB01] for details.

Remark 3.1.2. Note that Definition 3.1.1 does not exclude that knots may occur more than once in the knot sequence.

Proposition 3.1.3 (Properties of B-splines). Assume $x_i < x_{i+m+1}$ for all $i \in \mathbb{Z}$. Then, the function B_i^m defined in 3.1.1 consists of at most $m+1$ nontrivial polynomial pieces of degree at most m , has local support and is positive on (x_i, x_{i+m+1}) . Especially, the B-spline B_i^m satisfies

$$\begin{aligned}\text{supp}(B_i^m) &= \begin{cases} (x_i, x_{i+m+1}), & \text{if } x_i < x_{i+m} \\ [x_i, x_{i+m+1}), & \text{if } x_i = x_{i+1} = \dots = x_{i+m}, \end{cases} \\ B_i^m(x_i) &= \begin{cases} 0, & \text{if } x_i < x_{i+m} \\ 1, & \text{if } x_i = x_{i+1} = \dots = x_{i+m}. \end{cases}\end{aligned}$$

Moreover, the functions B_i^m always build a partition of unity if the knot sequence forms a partition of \mathbb{R} :

$$\sum_{i \in \mathbb{Z}} B_i^m(x) = 1 \quad \forall x \in \mathbb{R}. \quad (3.6)$$

The first statements can be verified directly by Definition 3.1.1. The proof of (3.6) can be found for example in [DeB01] or [SW05].

Definition 3.1.4 (B-splines on intervals). *Consider an interval $[a, b] \subset \mathbb{R}$ and a knot vector $X = (x_0, \dots, x_n)$ with $a = x_0 < x_1 \leq \dots \leq x_{n-1} < x_n = b$ and the extended vector $X_m = (x_{-m} = \dots = x_0 < x_1 \leq x_2 \leq \dots \leq x_{n-1} < x_n = \dots = x_{n+m})$. Let the functions $B_0^m, \dots, B_{n+m-1}^m$ be as defined in 3.1.1, but with $B_{n+m-1}^m(b) := 1$. Then the functions $B_0^m, \dots, B_{n+m-1}^m$ are called B-splines of degree m on the interval $[a, b]$.*

Remark 3.1.5. *Note that the change in the last B-spline $B_{n+m-1}^m(b) = 1$ is strictly required, when the B-splines shall build a partition of unity on the closed interval $[a, b]$. Without this change, Definition 3.1.1 directly yields $B_i^m(b) = 0$ for all $i = 0, 1, \dots, n + m - 1$ and consequently $\sum_{i=0}^{n+m-1} B_i^m(b) = 0$.*

Figure 3.1 shows the B-splines B_i^m for $m = 1$, $m = 2$ and $m = 5$ on the interval $[a, b] = [0, 7]$ for the knotvectors $X_1 = (0, 0, 1, 2.5, 3, 5, 6.2, 7, 7)$, $X_2 = (0, 0, 0, 1, 2.5, 3, 5, 6.2, 7, 7, 7)$ and $X_5 = (0, 0, 0, 0, 0, 0, 1, 2.5, 3, 5, 6.2, 7, 7, 7, 7, 7)$.

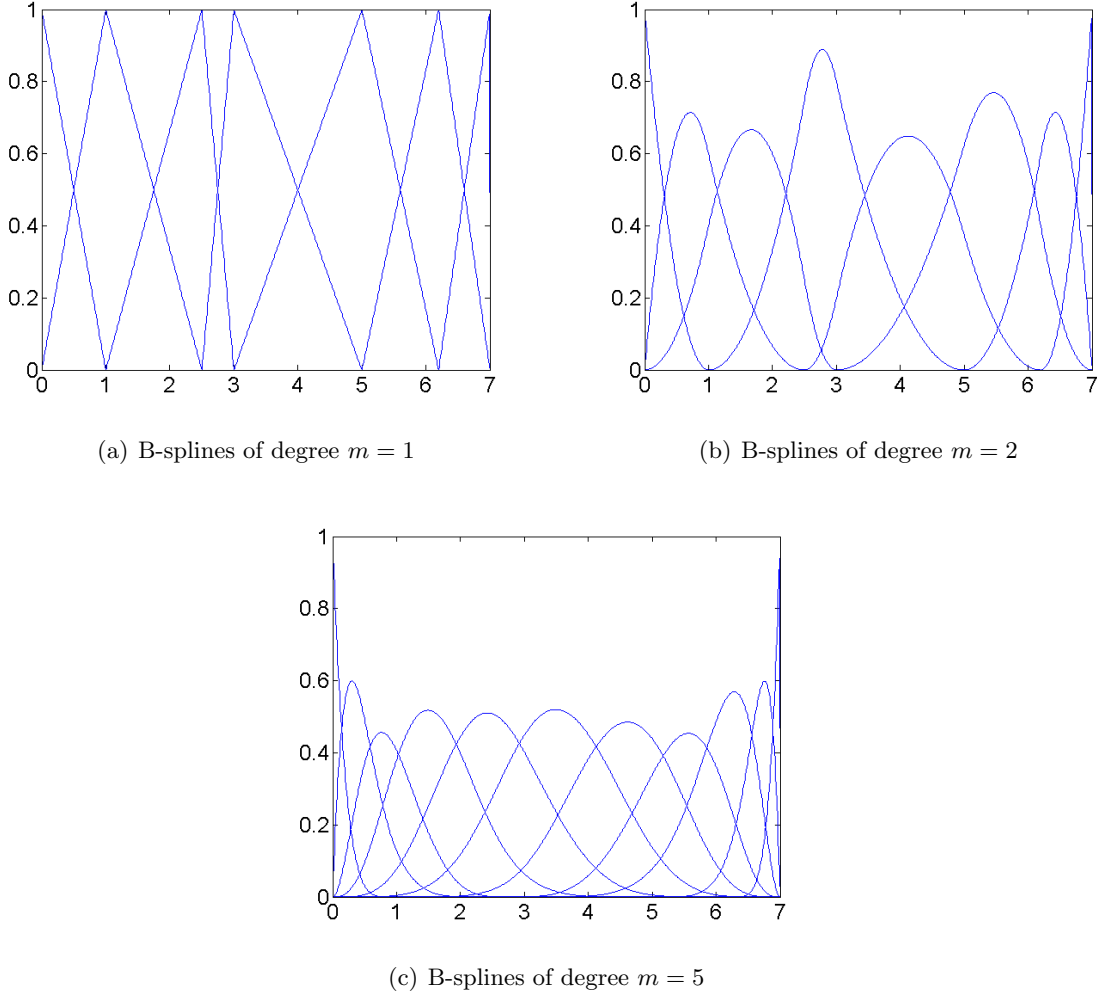
The following lemma will allow to simplify the formulation of the FVPM with B-splines considerably. The proof can be found in [DeB01] and is given here for completeness.

Lemma 3.1.6. *For $m \geq 0$, we have*

$$\int_{\mathbb{R}} B_i^m(x) dx = \frac{x_{i+m+1} - x_i}{(m+1)}. \quad (3.7)$$

Proof. We consider the Taylor expansion for $f \in C^{m+1}$ and assume $x, x_i, \dots, x_{i+m+1} \in [a, b]$ for some $a, b \in \mathbb{R}$:

$$\begin{aligned}f(x) &= \sum_{r \leq m} D^r f(a) \frac{(x-a)^r}{r!} + \int_a^x (x-s)^m \frac{D^{m+1} f(s)}{m!} ds \\ &= \sum_{r \leq m} D^r f(a) \frac{(x-a)^r}{r!} + \int_a^b (x-s)_+^m \frac{D^{m+1} f(s)}{m!} ds.\end{aligned}$$

Figure 3.1: B-splines on the interval $[0, 7]$.

If we now apply the divided difference $[x_i, \dots, x_{i+m+1}]$ to both sides, we get

$$[x_i, \dots, x_{i+m+1}]f = \int_a^b [x_i, \dots, x_{i+m+1}](\cdot - s)_+^m \frac{D^{m+1}f(s)}{m!} ds.$$

Choosing $f(x) = x^{m+1}$ leads to

$$1 = \int_a^b M_i^m(s) ds = \int_{\mathbb{R}} M_i^m(s) ds$$

with M_i^m defined in (3.4). Using (3.5), we get

$$\int_{\mathbb{R}} B_i^m(x) dx = \frac{x_{i+m+1} - x_i}{m+1}.$$

□

3.2 The FVPM with B-splines in one spatial dimension

As we have seen in the last paragraph, B-splines are piecewise polynomial and already build a partition of unity. Thus, we are able to evaluate the integrals in (2.12) and (2.13) exactly without much effort instead of using expensive numerical integration techniques. In return, we have to take into account additional terms appearing in the scheme for B-splines of degree $m > 1$. Further on, we will see that one loses the freedom of choosing arbitrary many neighbours for each particle.

In the following, we will derive the FVPM for the problem (2.1) in one spatial dimension using B-splines as particles. We begin with the case of B-splines of degree $m = 1$, because in this case the scheme can be written in a very simple way by using a slightly different discretization than will be used in the general case.

First Ansatz: B-splines of degree $m = 1$

Consider an interval $[a, b] \subset \mathbb{R}$, points $a = x_{-1} = x_0 \leq x_1 \leq \dots \leq x_N = x_{N+1} = b$ and B-splines of degree $m = 1$ that form a partition of unity over the interval $[a, b]$ as defined in 3.1.4. The points x_0, \dots, x_N will be denoted as particle positions which are allowed to move, thus having $x_i = x_i(t)$ for $i = 0, \dots, N$.

The corresponding B-splines B_i^1 are denoted as particles. Of course, the auxiliary points x_{-1} and x_{N+1} should move with the same velocities as the points x_0 and x_N , respectively, such that we have $x_{-1}(t) = x_0(t)$ and $x_N(t) = x_{N+1}(t) \forall t \geq 0$ and a partition of unity according to Definition 3.1.4 is always guaranteed. Further on, we restrict the movement of the points such that we have $x_i(t) \leq x_{i+1}(t) \forall t \geq 0$, meaning that particles are not allowed to overtake each other. Obviously, the B-spline B_i^m defined in 3.1.1 only depends on the points $x_i(t), \dots, x_{i+m+1}(t)$. After renumbering and introducing quantities

$$h_i(t) = x_{i+1}(t) - x_i(t),$$

the B-splines of degree $m = 1$ can be written at each time $t \geq 0$ as

$$B_i^1(x, t) = W(x - x_i(t), h_{i-1}(t), h_i(t)), \quad i = 0, \dots, N$$

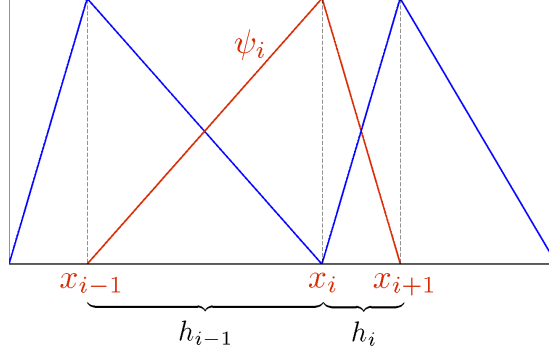
with the function

$$W(x, h_1, h_2) = \begin{cases} 1 + \frac{x}{h_1} & \text{for } -h_1 \leq x \leq 0 \\ 1 - \frac{x}{h_2} & \text{for } 0 \leq x \leq h_2 \\ 0 & \text{else.} \end{cases}$$

These B-splines will now be taken as particles, meaning that we set

$$\psi_i(x, t) = B_i^1(x, t),$$

compare Figure 3.2. Note that in general, the position $x_i(t)$ of particle ψ_i at time t is no longer the center of the support of that particle.

Figure 3.2: Sketch of a particle ψ_i .

Using

$$\begin{aligned} \partial_t \psi_i(x, t) &= -\dot{x}_i(t) \nabla_x \psi_i(x, t) + \dot{h}_{i-1}(t) \partial_{h_1} W(x - x_i(t), h_{i-1}(t), h_i(t)) \\ &\quad + \dot{h}_i(t) \partial_{h_2} W(x - x_i(t), h_{i-1}(t), h_i(t)), \end{aligned}$$

we can write for a non-moving domain $\Omega = [a, b] \subset \mathbb{R}$ with the outer normal vector n

$$\begin{aligned} \frac{d}{dt} \int_{\Omega} \mathbf{u} \psi_i dx &= \int_{\Omega} \left[(-\dot{x}_i \nabla_x \psi_i + \dot{h}_{i-1} \partial_{h_1} W(x - x_i, h_{i-1}, h_i) \right. \\ &\quad \left. + \dot{h}_i \partial_{h_2} W(x - x_i, h_{i-1}, h_i)) \mathbf{u} + \mathbf{F}(\mathbf{u}) \nabla_x \psi_i \right] dx - \int_{\partial\Omega} \mathbf{F}(\mathbf{u}) \psi_i n ds \\ &= \int_{\Omega} (\mathbf{F}(\mathbf{u}) - \dot{x}_i \cdot \mathbf{u}) \nabla_x \psi_i dx + \dot{h}_{i-1} \int_{\Omega} \partial_{h_1} W(x - x_i, h_{i-1}, h_i) \mathbf{u} dx \\ &\quad + \dot{h}_i \int_{\Omega} \partial_{h_2} W(x - x_i, h_{i-1}, h_i) \mathbf{u} dx - \int_{\partial\Omega} \mathbf{F}(\mathbf{u}) \psi_i n ds. \end{aligned}$$

As in Chapter 2, we define local averages as

$$\mathbf{u}_i(t) = \frac{1}{V_i(t)} \int_{\Omega} \mathbf{u}(x, t) \psi_i(x, t) dx,$$

where

$$V_i(t) = \int_{\Omega} \psi_i(x, t) dx = \frac{1}{2} (x_{i+1}(t) - x_{i-1}(t)) = \frac{1}{2} (h_i(t) + h_{i-1}(t)).$$

Assuming now that \mathbf{u} can be approximated by $\bar{\mathbf{u}}_{i,i-1}$ and $\bar{\mathbf{u}}_{i,i+1}$ on (x_{i-1}, x_i) and (x_i, x_{i+1}) , respectively, and using

$$\begin{aligned} \partial_{h_1} W(x - x_i, h_{i-1}, h_i)|_{(x_i, x_{i+1})} &= 0 \\ \partial_{h_2} W(x - x_i, h_{i-1}, h_i)|_{(x_{i-1}, x_i)} &= 0 \end{aligned}$$

as well as

$$\int_{x_{i-1}}^{x_i} \partial_{h_1} W(x - x_i, h_{i-1}, h_i) dx = \int_{x_i}^{x_{i+1}} \partial_{h_2} W(x - x_i, h_{i-1}, h_i) dx = \frac{1}{2}$$

and

$$\int_{x_{i-1}}^{x_i} \nabla_x \psi_i dx = - \int_{x_i}^{x_{i+1}} \nabla_x \psi_i dx = 1,$$

we can write

$$\begin{aligned} \frac{d}{dt} (V_i \mathbf{u}_i) &\approx (\mathbf{F}(\bar{\mathbf{u}}_{i,i-1}) - \bar{\mathbf{u}}_{i,i-1} \dot{x}_i) - (\mathbf{F}(\bar{\mathbf{u}}_{i,i+1}) - \bar{\mathbf{u}}_{i,i+1} \dot{x}_i) \\ &\quad + \frac{1}{2} \dot{h}_{i-1} \bar{\mathbf{u}}_{i,i-1} + \frac{1}{2} \dot{h}_i \bar{\mathbf{u}}_{i,i+1} - \tilde{\mathbf{B}}_i \\ &= \left(\mathbf{F}(\bar{\mathbf{u}}_{i,i-1}) - \bar{\mathbf{u}}_{i,i-1} \frac{\dot{x}_i + \dot{x}_{i-1}}{2} \right) - \left(\mathbf{F}(\bar{\mathbf{u}}_{i,i+1}) - \bar{\mathbf{u}}_{i,i+1} \frac{\dot{x}_i + \dot{x}_{i+1}}{2} \right) - \tilde{\mathbf{B}}_i \end{aligned}$$

where

$$\tilde{\mathbf{B}}_i = \int_{\partial\Omega} \mathbf{F}(\mathbf{u}) \psi_i n ds.$$

Introducing a numerical flux function $\mathbf{g}_{ij} := \mathbf{g}(\mathbf{u}_i, \dot{x}_i, \mathbf{u}_j, \dot{x}_j)$ consistent with $(\mathbf{F}(\mathbf{u}) - \mathbf{u} \dot{x})$ in the sense that

$$g(\mathbf{u}, \dot{x}_i, \mathbf{u}, \dot{x}_j) = (\mathbf{F}(\mathbf{u}) - \mathbf{u} \frac{\dot{x}_i + \dot{x}_j}{2}),$$

we can approximately write

$$\frac{d}{dt} (V_i \mathbf{u}_i) = -(\mathbf{g}_{i,i+1} + \mathbf{g}_{i,i-1}) - \tilde{\mathbf{B}}_i. \quad (3.8)$$

For the volumes it follows immediately from the shape of the B-splines of degree $m = 1$ or by Lemma 3.1.6 that

$$V_i(t) = \int_{\Omega} \psi_i(x, t) dx = \frac{x_{i+1}(t) - x_{i-1}(t)}{2}$$

and therefore

$$\dot{V}_i = \frac{\dot{x}_{i+1} - \dot{x}_{i-1}}{2}. \quad (3.9)$$

Together with the initial values

$$V_i(0) = \int_{\Omega} \psi_i(x, 0) dx = \frac{x_{i+1}(0) - x_{i-1}(0)}{2}, \quad (3.10)$$

$$\mathbf{u}_i(0) = \int_{\Omega} \mathbf{u}(x, 0) \psi_i(x, 0) dx \quad (3.11)$$

and the reconstruction formula

$$\mathbf{u}_h(x, t) = \sum_{i=0}^N \mathbf{u}_i(t) \psi_i(x, t),$$

we have a semi-discrete version of the FVPM with B-splines of degree $m = 1$ in one spatial dimension. It follows immediately that this scheme is conservative and preserves constant states. In fact, the only difference to scheme (2.19a)–(2.19d) for the case of two neighbours per particle and piecewise linear shape functions described in Chapter 2 lies in the initial values and the reconstruction formula.

FVPM with B-splines of higher degree

We will now derive the FVPM with B-splines of higher degree in one spatial dimension. Unfortunately, in the general case, the scheme will contain additional terms resulting from the more complicated dependence of the B-splines on the particle positions. As an advantage, these terms and all other occuring integrals can be computed exactly.

By construction, for every $t \geq 0$, the B-spline B_i^m defined in 3.1.1 can be written as a function W^m depending only on $x - x_i$ and the distances between the knots x_{i+1} and x_i , x_{i+2} and $x_{i+1}, \dots, x_{i+m+1}$ and x_{i+m} , i.e. we set

$$\psi_i^m(x, t) = W^m(x - x_i(t), h_i(t), \dots, h_{i+m}(t)) \quad (3.12)$$

for $W^m : \mathbb{R} \times \mathbb{R}_{>0}^{m+1} \rightarrow \mathbb{R}$ and $h_i(t) := x_{i+1}(t) - x_i(t)$. As before, the particles will be moved according to the movement of the knots $x_i(t)$. Note that, in constrast to our first Ansatz, $x_i(t)$ cannot be assigned to the unique maximum of the particle ψ_i^m any longer, except for degree $m = 1$ and after renumbering, see Figure 3.3. For this reason, we will get two different numerical schemes in the case $m = 1$: scheme (3.8)–(3.11) based on our first Ansatz and scheme (3.26)–(3.29) based on the general Ansatz to be introduced now.

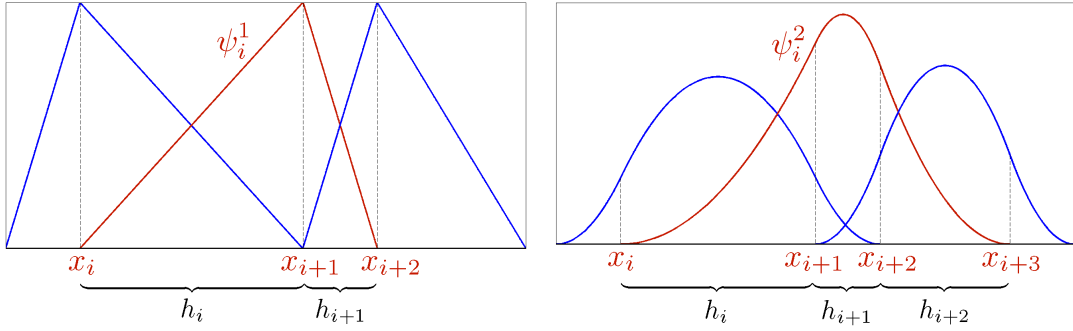


Figure 3.3: Sketch of particles ψ_i^1 (left) and ψ_i^2 (right).

To simplify notation, we set

$$W_i^m(x, t, h_i(t), \dots, h_{i+m}(t)) := W^m(x - x_i(t), h_i(t), \dots, h_{i+m}(t)).$$

Moreover, we will often neglect the arguments and write W_i^m instead of $W_i^m(x, t, h_i(t), \dots, h_{i+m}(t))$.

Proposition 3.2.1. *The time derivative of particle i can be written as*

$$\begin{aligned} \partial_t \psi_i^m(x, t) &= \sum_{j=1}^N (\dot{x}_j \psi_i^m \nabla_x \psi_j^m - \dot{x}_i \psi_j^m \nabla_x \psi_i^m) \\ &\quad + \sum_{j=1}^N \sum_{k=0}^m (\dot{h}_{i+k} (\partial_{h_{i+k}} W_i^m) \psi_j^m - \dot{h}_{j+k} (\partial_{h_{j+k}} W_j^m) \psi_i^m). \end{aligned} \quad (3.13)$$

Proof. First, we state

$$\partial_t \psi_i^m(x, t) = -\dot{x}_i \nabla_x \psi_i^m + \sum_{k=0}^m \dot{h}_{i+k} \partial_{h_{i+k}} W_i^m. \quad (3.14)$$

Using $\sum_j \psi_j^m \equiv 1$, we get

$$\begin{aligned} \partial_t \psi_i^m(x, t) &= \sum_{j=1}^N \psi_j^m \partial_t \psi_i^m - \psi_i^m \frac{d}{dt} \sum_{j=1}^N \psi_j^m \\ &= \sum_{j=1}^N (\dot{x}_j \psi_i^m \nabla_x \psi_j^m - \dot{x}_i \psi_j^m \nabla_x \psi_i^m) \\ &+ \sum_{j=1}^N \sum_{k=0}^m (\dot{h}_{i+k} (\partial_{h_{i+k}} W_i^m) \psi_j^m - \dot{h}_{j+k} (\partial_{h_{j+k}} W_j^m) \psi_i^m). \end{aligned}$$

□

Remark 3.2.2. Using (3.14), we see that the particles satisfy the transport equation

$$\partial_t \psi_i^m + \dot{x}_i \nabla_x \psi_i^m = \sum_{k=0}^m \dot{h}_{i+k} \partial_{h_{i+k}} W_i^m.$$

As before, the spatial derivative can be represented in the symmetric form

$$\nabla_x \psi_i^m = \sum_{j=1}^N (\psi_j^m \nabla_x \psi_i^m - \psi_i^m \nabla_x \psi_j^m). \quad (3.15)$$

In analogy to Section 2.1, we define geometrical coefficients

$$\begin{aligned} \gamma_{ij} &= \int_{\Omega} \Gamma_{ij} dx, \\ \beta_{ij} &= \gamma_{ij} - \gamma_{ji}, \\ \delta_{ji}^k &= \int_{\Omega} (\partial_{h_{i+k}} W_i^m) \psi_j^m dx \end{aligned}$$

and the boundary term

$$\tilde{B}_i = \int_{\partial\Omega} \mathbf{F}(\mathbf{u}) \psi_i^m n ds$$

to get

$$\begin{aligned} \frac{d}{dt} (V_i \mathbf{u}_i) &= \frac{d}{dt} \int_{\Omega} \mathbf{u} \psi_i^m dx = \sum_{j=1}^N \int_{\Omega} (\dot{x}_j \psi_i^m \nabla_x \psi_j^m - \dot{x}_i \psi_j^m \nabla_x \psi_i^m) \mathbf{u} dx \\ &+ \sum_{j=1}^N \sum_{k=0}^m \int_{\Omega} (\dot{h}_{i+k} (\partial_{h_{i+k}} W_i^m) \psi_j^m - \dot{h}_{j+k} (\partial_{h_{j+k}} W_j^m) \psi_i^m) \mathbf{u} dx \\ &+ \sum_{j=1}^N \int_{\Omega} \mathbf{F}(\mathbf{u}) (\psi_j^m \nabla_x \psi_i^m - \psi_i^m \nabla_x \psi_j^m) dx - \tilde{B}_i. \end{aligned}$$

As before, we take a numerical flux function $\mathbf{g}_{ij} := \mathbf{g}(\mathbf{u}_i, \dot{x}_i, \mathbf{u}_j, \dot{x}_j, n_{ij})$ that is conservative and consistent with $(\mathbf{F}(\mathbf{u}) - \mathbf{u} \cdot \bar{\dot{x}}_{ij})$ with the approximated particle velocity (2.30) and end up with the following ODE

$$\frac{d}{dt}(V_i \mathbf{u}_i) = - \sum_{j=1}^N |\beta_{ij}| \mathbf{g}_{ij} + \sum_{j=1}^N \sum_{k=0}^m (\dot{h}_{i+k} \delta_{ji}^k \mathbf{u}_j - \dot{h}_{j+k} \delta_{ij}^k \mathbf{u}_i) - \tilde{\mathbf{B}}_i. \quad (3.16)$$

Note that the approximated particle velocity (2.30) forms a convex combination of the particle velocities \dot{x}_i and \dot{x}_j as condition (2.33) is always satisfied.

To complete the modified FVPM, we need an additional equation for the volumes V_i . From Lemma 3.1.6 we can conclude directly

$$\dot{V}_i = \frac{d}{dt} \int_{\Omega} \psi_i^m dx = \frac{\dot{x}_{i+m+1} - \dot{x}_i}{m+1}. \quad (3.17)$$

Thus, we can now formulate a semi-discrete version of the FVPM with B-splines consisting of (3.16), (3.17), the initial conditions

$$\mathbf{u}_i(0) = \frac{1}{V_i(0)} \int_{\Omega} \mathbf{u}_0(x) \psi_i^m(x, 0) dx \quad (3.18)$$

$$V_i(0) = \int_{\Omega} \psi_i^m(x, 0) dx = \frac{x_{i+m+1}(0) - x_i(0)}{m+1} \quad (3.19)$$

and the reconstruction formula

$$\mathbf{u}_h(x, t) = \sum_{i=1}^N \mathbf{u}_i(t) \psi_i^m(x, t). \quad (3.20)$$

Remark 3.2.3. *For the evolution of the volumes V_i , we can write equivalently*

$$\dot{V}_i = \int_{\Omega} \partial_t \psi_i^m dx = \sum_{j=1}^N \left[(\gamma_{ij} \dot{x}_j - \gamma_{ji} \dot{x}_i) + \sum_{k=0}^m (\dot{h}_{i+k} \delta_{ji}^k - \dot{h}_{j+k} \delta_{ij}^k) \right]. \quad (3.21)$$

Again, we can show that the FVPM with B-splines is conservative and preserves constant states if a conservative numerical flux function that is consistent with the averaged particle velocity (2.30) is used.

Theorem 3.2.4 (Conservativity). *Consider a non-moving domain $\Omega \subset \mathbb{R}$ and assume that the numerical flux function g_{ij} is conservative and consistent with the modified flux function $(\mathbf{F}(\mathbf{u}) - \mathbf{u} \cdot \bar{\dot{x}}_{ij})$ with the averaged particle velocity (2.30). Then, the method (3.16)–(3.19) is conservative in the sense that*

$$\frac{d}{dt} \sum_{i=1}^N (V_i \mathbf{u}_i) = - \int_{\partial\Omega} \mathbf{F}(\mathbf{u}) n ds. \quad (3.22)$$

The proof is similar to that in Chapter 2.

Theorem 3.2.5 (Preservation of constant states). *Under the assumptions of Theorem 3.2.4 scheme (3.16)–(3.19) preserves constant states in the sense that*

$$\mathbf{u}_j(t) = \mathbf{u}_i(t) \quad \forall j \in \mathcal{N}(i)$$

implies

$$\frac{d}{dt} \mathbf{u}_i(t) = 0.$$

Proof. We use (3.21) to see that

$$\begin{aligned} \frac{d}{dt}(V_i \mathbf{u}_i) &= - \sum_{j=1}^N (\mathbf{F}(\mathbf{u}_i) - \mathbf{u}_i \cdot \bar{\mathbf{x}}_{ij}) \beta_{ij} + \sum_{j=1}^N \sum_{k=0}^m (\dot{h}_{i+k} \delta_{ji}^k - \dot{h}_{j+k} \delta_{ij}^k) \mathbf{u}_i - \tilde{\mathbf{B}}_i \\ &= \sum_{j=1}^N \left[(\gamma_{ij} \dot{x}_j - \gamma_{ji} \dot{x}_i) + \sum_{k=0}^m (\dot{h}_{i+k} \delta_{ji}^k - \dot{h}_{j+k} \delta_{ij}^k) \right] \mathbf{u}_i \\ &= \dot{V}_i \mathbf{u}_i. \end{aligned}$$

□

Remark 3.2.6. *Again, the conservativity and the property of preservation of constant states can be naturally transferred to a fully discretized version of the FVPM using for example an explicit forward Euler method in time which we will use in our numerical computations in Chapter 5.*

We end this section with the derivation of some helpful properties of the coefficients and the boundary term, which can be deduced directly from the properties of B-splines.

Proposition 3.2.7. *Consider the domain $\Omega = [a, b] \subset \mathbb{R}$ and N particles $\psi_i^m(x, t)$ over the extended knot vector $(x_1, x_2, \dots, x_{N+1}, \dots, x_{N+m+1})$ with $x_1 = \dots = x_{m+1}$ and $x_{N+1} = \dots = x_{N+m+1}$ (see Definition 3.1.4). Under the assumption $x_{m+1} < x_{m+2} < \dots < x_{N+1}$, the coefficients and the boundary term of the FVPM with B-splines satisfy*

$$\beta_{ij} = 2 \gamma_{ij}, \quad (3.23)$$

$$\bar{x}_{ij} = \frac{\dot{x}_i + \dot{x}_j}{2}, \quad (3.24)$$

$$\tilde{\mathbf{B}}_i(t) = \begin{cases} -\mathbf{F}(\mathbf{u}(a, t)), & \text{for } i = 1 \\ \mathbf{F}(\mathbf{u}(b, t)), & \text{for } i = N \\ \mathbf{0}, & \text{else.} \end{cases} \quad (3.25)$$

Proof. For the coefficients γ_{ij} , we have

$$\gamma_{ij} = \int_{\Omega} \psi_i^m \partial_x \psi_j^m dx = \int_a^b \psi_i^m \partial_x \psi_j^m dx = [\psi_i^m \psi_j^m]_a^b - \int_a^b \psi_j^m \partial_x \psi_i^m dx.$$

By Definition 3.1.4 and Proposition 3.1.3, the particles satisfy

$$\psi_i(a) = \begin{cases} 1, & \text{for } i = 1 \\ 0, & \text{else,} \end{cases} \quad \psi_i(b) = \begin{cases} 1, & \text{for } i = N \\ 0, & \text{else.} \end{cases}$$

Thus

$$\gamma_{ij} = -\gamma_{ji}$$

and

$$\beta_{ij} = \gamma_{ij} - \gamma_{ji} = 2\gamma_{ij}.$$

This yields

$$\bar{x}_{ij} = \frac{\gamma_{ij} \dot{x}_i - \gamma_{ji} \dot{x}_j}{\beta_{ij}} = \frac{\dot{x}_i + \dot{x}_j}{2}.$$

Alternatively, (3.24) follows by Proposition 2.2.13 (ii) in Chapter 2.

For the boundary term, we get directly

$$\tilde{\mathbf{B}}_i(t) = \int_{\partial\Omega} \mathbf{F}(\mathbf{u}) \psi_i^m n \, ds = [\mathbf{F}(\mathbf{u}) \psi_i^m]_a^b = \begin{cases} -\mathbf{F}(\mathbf{u}(a, t)), & \text{for } i = 1 \\ \mathbf{F}(\mathbf{u}(b, t)), & \text{for } i = N \\ \mathbf{0}, & \text{else.} \end{cases}$$

□

In the next section, we will present two special cases of the FVPM with B-splines, namely the cases for B-splines of degree $m = 1$ and $m = 2$ in one spatial dimension.

3.3 Special cases

We begin with B-splines of degree $m = 1$ for the scheme (3.16)–(3.19).

For $m = 1$, we get for the coefficients

$$\begin{aligned} \beta_{ij} &= \begin{cases} 1 & \text{for } j = i + 1 \\ -1 & \text{for } j = i - 1 \\ 0 & \text{else,} \end{cases} & \gamma_{ij} &= \begin{cases} \frac{1}{2} & \text{for } j = i + 1 \\ -\frac{1}{2} & \text{for } j = i - 1 \\ 0 & \text{else,} \end{cases} \\ \delta_{ji}^0 &= \begin{cases} \frac{1}{2} & \text{for } j = i + 1 \\ -\frac{1}{6} & \text{for } j = i - 1 \\ 0 & \text{else,} \end{cases} & \delta_{ji}^1 &= \begin{cases} \frac{1}{3} & \text{for } j = i + 1 \\ 0 & \text{else.} \end{cases} \end{aligned}$$

Using a forward Euler method in time for Equation (3.16), we obtain

$$\begin{aligned}
V_i^{n+1} \mathbf{u}_i^{n+1} &= V_i^n \mathbf{u}_i^n - \Delta t [g_{i,i+1}^n(\mathbf{u}_i^n, \mathbf{u}_{i+1}^n) + g_{i,i-1}^n(\mathbf{u}_i^n, \mathbf{u}_{i-1}^n)] \\
&\quad + \Delta t (\dot{h}_i^n \delta_{i+1,i}^0 \mathbf{u}_{i+1}^n - \dot{h}_{i+1}^n \delta_{i,i+1}^0 \mathbf{u}_i^n + \dot{h}_{i+1}^n \delta_{i+1,i}^1 \mathbf{u}_{i+1}^n - \dot{h}_{i+2}^n \delta_{i,i+1}^1 \mathbf{u}_i^n \\
&\quad + \dot{h}_i^n \delta_{i-1,i}^0 \mathbf{u}_{i-1}^n - \dot{h}_{i-1}^n \delta_{i,i-1}^0 \mathbf{u}_i^n + \dot{h}_{i-1}^n \delta_{i-1,i}^1 \mathbf{u}_{i-1}^n - \dot{h}_i^n \delta_{i,i-1}^1 \mathbf{u}_i^n) - \Delta t \tilde{\mathbf{B}}_i^n \\
&= V_i^n \mathbf{u}_i^n - \Delta t [g_{i,i+1}^n(\mathbf{u}_i^n, \mathbf{u}_{i+1}^n) + g_{i,i-1}^n(\mathbf{u}_i^n, \mathbf{u}_{i-1}^n)] \\
&\quad + \Delta t \left(\frac{1}{2} \dot{h}_i^n \mathbf{u}_{i+1}^n + \frac{1}{6} \dot{h}_{i+1}^n \mathbf{u}_i^n + \frac{1}{3} \dot{h}_{i+1}^n \mathbf{u}_{i+1}^n \right. \\
&\quad \left. - \frac{1}{6} \dot{h}_i^n \mathbf{u}_{i-1}^n - \frac{1}{2} \dot{h}_{i-1}^n \mathbf{u}_i^n - \frac{1}{3} \dot{h}_i^n \mathbf{u}_i^n \right) - \Delta t \tilde{\mathbf{B}}_i^n,
\end{aligned} \tag{3.26}$$

where $\tilde{\mathbf{B}}_i^n$ is a suitable discretization of the boundary term $\tilde{\mathbf{B}}_i$ in (3.16) and $V_i^n = V_i(n \Delta t)$, $\mathbf{u}_i^n = \mathbf{u}_i(n \Delta t)$, $\dot{h}_i^n = \dot{h}_i(n \Delta t)$ and $g_{ij}^n(\mathbf{u}, \mathbf{v}) = g_{ij}(\mathbf{u}, \dot{x}_i^n, \mathbf{v}, \dot{x}_j^n, n_{ij}^n)$ with $\dot{x}_i^n = \dot{x}_i(n \Delta t)$ and $n_{ij}^n = n_{ij}(n \Delta t)$. The initial values and the volume updates are given by

$$\mathbf{u}_i^0 = \frac{2}{x_{i+2}(0) - x_i(0)} \int_{x_i(0)}^{x_{i+2}(0)} \mathbf{u}_0(x) \psi_i^1(x, 0) dx, \tag{3.27}$$

$$V_i^0 = \frac{x_{i+2}(0) - x_i(0)}{2}, \tag{3.28}$$

$$V_i^{n+1} = V_i^n + \frac{\dot{x}_{i+2}^n - \dot{x}_i^n}{2}, \tag{3.29}$$

where

$$\psi_i^1(x, t) = \begin{cases} \frac{x - x_i(t)}{h_i(t)}, & x \in [x_i(t), x_i(t) + h_i(t)] \\ 1 - \frac{x - x_i(t) - h_i(t)}{h_{i+1}(t)}, & x \in [x_i(t) + h_i(t), x_i(t) + h_i(t) + h_{i+1}(t)] \\ 0, & \text{else.} \end{cases}$$

Obviously, scheme (3.26)–(3.29) is more complicated than scheme (3.8)–(3.11). In Section 5.2, we implemented both schemes for the problem of a moving piston in a one-dimensional domain. Both schemes behave well and provide similar convergence rates. Therefore, scheme (3.8)–(3.11) is the preferred one in the case $m = 1$.

Next, we consider the scheme (3.16)–(3.19) with B-splines of degree $m = 2$. In this case, we have four neighbours for every particle located sufficiently far away from the boundaries:

$$\mathcal{N}(i) = \{i - 2, i - 1, i + 1, i + 2\}, \quad i = 3, 4, \dots, N - 2. \tag{3.30}$$

An Euler discretization of (3.16)–(3.19) yields

$$\begin{aligned} V_i^{n+1} \mathbf{u}_i^{n+1} &= V_i^n \mathbf{u}_i^n - \Delta t \sum_{j \in \mathcal{N}(i)} |\beta_{ij}^n| \mathbf{g}_{ij}^n(\mathbf{u}_i^n, \mathbf{u}_j^n) \\ &\quad + \sum_{j \in \mathcal{N}(i)} \sum_{k=0}^2 \left(\dot{h}_{i+k}^n \left(\delta_{ji}^k \right)^n \mathbf{u}_j^n - \dot{h}_{j+k}^n \left(\delta_{ij}^k \right)^n \mathbf{u}_i^n \right) - \Delta t \tilde{\mathbf{B}}_i \end{aligned} \quad (3.31)$$

$$V_i^{n+1} = V_i^n + \frac{\dot{x}_{i+3}^n - \dot{x}_i^n}{3} \quad (3.32)$$

$$\mathbf{u}_i^0 = \frac{1}{V_i^0} \int_{\Omega} \mathbf{u}_0(x) \psi_i^2(x, 0) dx \quad (3.33)$$

$$V_i^0 = \frac{x_{i+3}(0) - x_i(0)}{3} \quad (3.34)$$

with

$$\begin{aligned} \psi_i^2(x, t) &= W^2(x - x_i(t), h_i(t), h_{i+1}(t), h_{i+2}(t)) \\ &= \begin{cases} \frac{(x - x_i)^2}{h_i(h_i + h_{i+1})}, & x \in [x_i, x_i + h_i) \\ \frac{(x - x_i)(h_i + h_{i+1} - (x - x_i))}{h_{i+1}(h_i + h_{i+1})} \\ \quad + \frac{((x - x_i) - h_i)(h_i + h_{i+1} + h_{i+2} - (x - x_i))}{h_{i+1}(h_{i+1} + h_{i+2})}, & x \in [x_i + h_i, \\ & \quad x_i + h_i + h_{i+1}) \\ \frac{(h_i + h_{i+1} + h_{i+2} - (x - x_i))^2}{h_{i+2}(h_{i+1} + h_{i+2})}, & x \in [x_i + h_i + h_{i+1}, \\ & \quad x_i + h_i + h_{i+1} + h_{i+2}) \\ 0, & \text{else,} \end{cases} \end{aligned}$$

where we omitted the argument t in the last term to simplify notation. Note that the upper index in the function ψ_i^2 denotes the degree of the B-spline. Note further that we used x_i, h_i, h_{i+1} and h_{i+2} as the only arguments of ψ_i^2 in the above representation. Since we have to compute the partial derivatives $\partial_{h_i} W_i$, $\partial_{h_{i+1}} W_i$ and $\partial_{h_{i+2}} W_i$, this notation is much more advisable.

The geometrical coefficients β_{ij} can be computed exactly and are given by

$$\begin{aligned} \beta_{i,i+1} &= \frac{1}{3} \left[\frac{h_{i+1}(2h_i + h_{i+1})}{(h_i + h_{i+1})(h_{i+1} + h_{i+2})} - \frac{h_{i+1}^2}{(h_{i+1} + h_{i+2})(h_{i+2} + h_{i+3})} + \frac{h_{i+1} + 3h_{i+2}}{(h_{i+1} + h_{i+2})} \right] \\ &= -\beta_{i+1,i}, \quad i = 3, \dots, N-2 \end{aligned}$$

and

$$\beta_{i,i+2} = \frac{h_{i+2}^2}{3(h_{i+1} + h_{i+2})(h_{i+2} + h_{i+3})} = -\beta_{i+2,i}, \quad i = 2, \dots, N-2.$$

For particles near the boundary, the coefficients β_{ij} have to be computed separately. For the computation of the additional coefficients δ_{ji}^k , the functions $\partial_{h_i} W_i$, $\partial_{h_{i+1}} W_i$ and $\partial_{h_{i+2}} W_i$ are needed. These are given by (3.35)–(3.37). To compute the coefficients δ_{ji}^k , we have to determine the integrals

$$\delta_{ji}^k = \int_{\Omega} (\partial_{h_{i+k}} W_i^m) \psi_j^m dx, \quad j \in \mathcal{N}(i), \quad k = 0, 1, 2.$$

If we write

$$\begin{aligned} \delta_{i-2,i}^k &= \int_{x_i}^{x_{i+1}} (\partial_{h_{i+k}} W_i) \psi_{i-2} dx, \\ \delta_{i-1,i}^k &= \int_{x_i}^{x_{i+1}} (\partial_{h_{i+k}} W_i) \psi_{i-1} dx + \int_{x_{i+1}}^{x_{i+2}} (\partial_{h_{i+k}} W_i) \psi_{i-1} dx, \\ \delta_{i+1,i}^k &= \int_{x_{i+1}}^{x_{i+2}} (\partial_{h_{i+k}} W_i) \psi_{i+1} dx + \int_{x_{i+2}}^{x_{i+3}} (\partial_{h_{i+k}} W_i) \psi_{i+1} dx, \\ \delta_{i+2,i}^k &= \int_{x_{i+2}}^{x_{i+3}} (\partial_{h_{i+k}} W_i) \psi_{i+2} dx, \end{aligned}$$

for $k = 0, 1, 2$, every integrand is a polynomial of degree 4 at most and can be computed exactly with Gaussian integration using 3 integration points. Alternatively, the above formulas can be computed symbolically. In our numerical problem tested in Section 5.2, we decided to compute the integrals symbolically using the MATLAB Symbolic Math Toolbox and evaluate the results in every time step, because this method is much faster in MATLAB.

Remark 3.3.1. *A first idea to construct a scheme like the modified FVPM with B-splines in the last sections in a two-dimensional setting $\Omega \subseteq \mathbb{R}^2$, is to build B-splines by a tensor product structure*

$$\psi_{(ij)}^m(x, y, t) = \psi_i^m(x, t) \psi_j^m(y, t),$$

where $\psi_i^m(x, t)$ and $\psi_j^m(y, t)$ are one-dimensional B-splines constructed as in Section 3.1 and 3.2 [DeB01].

In the next chapter, we show how the FVPM can be applied to balance equations like the Savage-Hutter equations. For this purpose, we couple the FVPM with an existing kinetic scheme for the Savage-Hutter equations.

$$\begin{aligned}
\partial_{h_i} W_i &= \begin{cases} -\frac{(x-x_i)^2(2h_i+h_{i+1})}{(h_i(h_i+h_{i+1}))^2}, & x \in (x_i, x_i+h_i) \\ \frac{(x-x_i)^2}{h_{i+1}(h_i+h_{i+1})^2} + \frac{2(x-x_i)-2h_i-h_{i+1}-h_{i+2}}{h_{i+1}(h_{i+1}+h_{i+2})}, & x \in (x_i+h_i, x_i+h_i+h_{i+1}) \\ 2\frac{h_i+h_{i+1}+h_{i+2}-(x-x_i)}{h_{i+2}(h_{i+1}+h_{i+2})}, & x \in (x_i+h_i+h_{i+1}, x_i+h_i+h_{i+1}+h_{i+2}) \\ 0, & \text{else,} \end{cases} \quad (3.35) \\
\partial_{h_{i+1}} W_i &= \begin{cases} -\frac{(x-x_i)^2}{h_i(h_i+h_{i+1})^2}, & x \in (x_i, x_i+h_i) \\ \frac{(x-x_i)^2(h_i+2h_{i+1})}{(h_{i+1}(h_i+h_{i+1}))^2} - \frac{(x-x_i)}{h_{i+1}^2} + \frac{(x-x_i-h_i)^2(h_{i+2}+2h_{i+1})}{(h_{i+1}(h_{i+1}+h_{i+2}))^2} - \frac{(x-x_i-h_i)}{h_{i+1}^2}, & x \in (x_i+h_i, x_i+h_i+h_{i+1}) \\ 2\frac{(h_i+h_{i+1}+h_{i+2}-(x-x_i))}{h_{i+2}(h_{i+1}+h_{i+2})} - \frac{(h_i+h_{i+1}+h_{i+2}-(x-x_i))^2}{h_{i+2}(h_{i+1}+h_{i+2})^2}, & x \in (x_i+h_i+h_{i+1}, x_i+h_i+h_{i+1}+h_{i+2}) \\ 0, & \text{else,} \end{cases} \quad (3.36) \\
\partial_{h_{i+2}} W_i &= \begin{cases} \frac{((x-x_i)-h_i)^2}{h_{i+1}(h_{i+1}+h_{i+2})^2}, & x \in (x_i+h_i, x_i+h_i+h_{i+1}) \\ 2\frac{(h_i+h_{i+1}+h_{i+2}-(x-x_i))}{h_{i+1}(h_{i+1}+h_{i+2})} - \frac{(h_i+h_{i+1}+h_{i+2}-(x-x_i))^2(h_{i+1}+2h_{i+2})}{(h_{i+1}(h_{i+1}+h_{i+2}))^2}, & x \in (x_i+h_i+h_{i+1}, x_i+h_i+h_{i+1}+h_{i+2}) \\ 0, & \text{else} \end{cases} \quad (3.37)
\end{aligned}$$

Chapter 4

The FVPM for granular flows

Dense snow avalanches and landslides regularly entail enormous damage. Therefore it would be extremely helpful if one could predict the dynamic behaviour of such granular masses, particularly to predict the place where the mass comes to rest. Furthermore, industrial problems dealing with the dynamic behaviour of granular material like the modeling of grain in silos are of great interest.

This chapter deals with the dynamics of granular flow problems that can be described by the Savage-Hutter (SH) equations. In Section 4.1, we will give a short introduction to the SH equations and describe the main difficulties that arise when dealing with the dynamics of granular masses.

Section 4.2 is concerned with a (semi-)kinetic representation of the SH equations and a resulting Finite Volume Method. This so-called kinetic scheme is able to describe the whole dynamic of a granular mass from starting to stopping. Moreover, it can be shown that it preserves the steady states of granular masses of rest. For the construction of the kinetic formulation, we use an Ansatz of Perthame and Simeoni presented in [PS01]. The contents and results of Sections 4.1 and 4.2 can be found in [KS08] and are given here only for the sake of completeness. In Section 4.3, we apply the FVPM to the SH equations by adopting the kinetic fluxes. This provides an example where the FVPM can be used successfully to solve hyperbolic conservation laws with a source term.

4.1 The Savage-Hutter equations

In [SH89], Savage and Hutter deduced a model based on the incompressible Euler equations to describe the dynamics of granular masses moving on a bottom or ground topography, which is constant in time and spatially only slowly varying. Under the additional assumption that the internal friction angle equals the dynamic one, the SH equations in one spatial dimension read

$$\begin{aligned}\partial_t h + \partial_x(hu) &= 0, \\ \partial_t(hu) + \partial_x(hu^2 + b\frac{h^2}{2}) &= gh,\end{aligned}\tag{4.1}$$

where

$$b = \varepsilon k \cos \zeta$$

and

$$g = g(u) = \sin \zeta - \operatorname{sgn}(u) \cos \zeta \tan \delta \quad \text{for } u \neq 0,$$

in the unknowns $h, hu : \mathbb{R} \times (0, \infty) \rightarrow \mathbb{R}$. Here, h and u denote the height of the granular material and its average velocity, respectively. Moreover ζ is referred to as the inclination angle of the bottom topography against the horizontal, δ denotes the dynamic friction angle, ε the ratio of characteristic height to characteristic length of the mass under consideration and k is an earth pressure coefficient. The system is hyperbolic as long as the height stays nonnegative. In [SH91], this model is extended to smoothly varying bottom topographies, which results in an additional x -dependence in the flux function. The effect of friction is modelled in both equations using a simple Coulomb model. Other friction models that are not subject of this paper are discussed in [MVB⁺03].

The source term gh on the right hand side of the second equation of (4.1) causes non-constant steady states, which actually do make physical sense. But in contrast to other models for fluid dynamical processes like the shallow water equations, there are infinitely many steady states, which are described by

$$\partial_t h = 0, \quad u = 0. \tag{4.2}$$

If we insert (4.2) into the SH equations (4.1), we get the following relation between the spatial derivative of h and the source term

$$b \partial_x h = g.$$

Since we are interested in granular masses at rest, we have to consider static friction, which is bounded by the dynamic friction due to the Coulomb model. Therefore, the source term and hence the steady states are not uniquely defined for $u = 0$. But we can give a simple criterion on $\partial_x h$ to decide whether the mass is in equilibrium or not. Height profiles h which satisfy that condition will be named admissible profiles in the following. An exact definition of admissible profiles is given in Section 4.2.

4.2 A kinetic FVM for the SH equations

In this section, we present a (semi-)kinetic representation of the SH equations with constant inclination angle such that $b = \text{const}$ and a numerical scheme based on that formulation. A (semi-)kinetic representation of the general SH equations for smoothly varying chutes can be found easily and is given in [KS08]. The construction of the kinetic representation relies on an Ansatz of Perthame and Simeoni presented in [PS01]. The numerical scheme will be shown to be conservative and consistent and has the desired property of preserving nontrivial steady states. Moreover, we can apply a stability result from [PS01] which ensures nonnegativity of the height under some weak conditions.

Other numerical approaches to the SH equations can be found in [TNGH02], where the authors present a combination of a shock-capturing non-oscillatory central difference

scheme (NOC) and a front-tracking method. This scheme produces very satisfactory results, particularly for an avalanche moving down an incline and coming to rest at a flat runout zone.

Consider a constant inclination angle $\zeta = \text{const}$ and thus $b = \text{const}$. Similar to the kinetic formulation of the Euler equations or the shallow water equations in [Cer94], [Str95], [PS01] and [Per02], we are looking for a density of particles $M(x, \xi, t) = M(h, \xi - u)$ satisfying

$$\begin{pmatrix} h \\ hu \\ hu^2 + \frac{1}{2}bh^2 \end{pmatrix} = \int_{\mathbb{R}} \begin{pmatrix} 1 \\ \xi \\ \xi^2 \end{pmatrix} M(h, \xi - u) d\xi. \quad (4.3)$$

In the following, we proceed as Perthame and Simeoni in [PS01] for the shallow water equations and define M by making the Ansatz

$$M(h, \xi - u) = \sqrt{h} \chi\left(\frac{\xi - u}{\sqrt{h}}\right) \quad (4.4)$$

with a nonnegative function $\chi : \mathbb{R} \rightarrow \mathbb{R}$ satisfying

$$\begin{aligned} \chi(\omega) &= \chi(-\omega), \\ \int_{\mathbb{R}} \chi(\omega) d\omega &= 1, \\ \int_{\mathbb{R}} \omega^2 \chi(\omega) d\omega &= \frac{b}{2}. \end{aligned} \quad (4.5)$$

It is easy to see that every function M of the form (4.4) satisfies (4.3): because of the properties (4.5), we obtain for the moments of M

$$\begin{aligned} \int_{\mathbb{R}} M(h, \xi - u) d\xi &= \int_{\mathbb{R}} \sqrt{h} \chi\left(\frac{\xi - u}{\sqrt{h}}\right) d\xi \\ &= \int_{\mathbb{R}} \sqrt{h} \chi(\omega) \sqrt{h} d\omega \\ &= h, \\ \int_{\mathbb{R}} \xi M(h, \xi - u) d\xi &= \int_{\mathbb{R}} \xi \sqrt{h} \chi\left(\frac{\xi - u}{\sqrt{h}}\right) d\xi \\ &= \int_{\mathbb{R}} (\sqrt{h}\omega + u) \sqrt{h} \chi(\omega) \sqrt{h} d\omega \\ &= h^{3/2} \int_{\mathbb{R}} \omega \chi(\omega) d\omega + hu \int_{\mathbb{R}} \chi(\omega) d\omega \\ &= hu, \end{aligned}$$

$$\begin{aligned}
\int_{\mathbb{R}} \xi^2 M(h, \xi - u) d\xi &= \int_{\mathbb{R}} \xi^2 \sqrt{h} \chi\left(\frac{\xi - u}{\sqrt{h}}\right) d\xi \\
&= \int_{\mathbb{R}} (\sqrt{h} \omega + u)^2 \sqrt{h} \chi(\omega) \sqrt{h} d\omega \\
&= h^2 \int_{\mathbb{R}} \omega^2 \chi(\omega) d\omega + 2 h^{3/2} u \int_{\mathbb{R}} \omega \chi(\omega) d\omega + h u^2 \int_{\mathbb{R}} \chi(\omega) d\omega \\
&= h u^2 + \frac{1}{2} b h^2.
\end{aligned}$$

These relations lead to the following theorem, the same theorem as Perthame and Simeoni formulate in [PS01] for the shallow water equations. For the proof see [KS08].

Theorem 4.2.1. *The pair (h, hu) is a strong solution of the Equation (4.1) if and only if $M(h, \xi - u)$ of the form (4.4) solves the (semi-)kinetic equation*

$$\partial_t M(h, \xi - u) + \xi \partial_x M(h, \xi - u) + g(u) \partial_\xi M(h, \xi - u) = Q(t, x, \xi) \quad (4.6)$$

with $Q(t, x, \xi)$ satisfying

$$\int_{\mathbb{R}} Q d\xi = 0 \quad \text{and} \quad \int_{\mathbb{R}} \xi Q d\xi = 0. \quad (4.7)$$

Remark 4.2.2. *Note that there are still macroscopic values in the kinetic Equation (4.6). Therefore, we call the Vlasov type Equation (4.6) (semi-)kinetic.*

Remark 4.2.3. *The only nonlinearity of Equation (4.6) is caused by the collision operator Q . Therefore, (4.6) is a lot easier to handle than the original system.*

Now we have to choose the function χ . In contrast to the shallow water equations, where friction is neglected and the only physical steady states are those of a "lake at rest", the SH equations admit infinitely many steady states of granular masses at rest. It is clear that the height profile of such a granular mass is not uniquely defined. Moreover, the gradient of the profile has to be somehow bounded to avoid the moving of the material. If one inserts the criterion $\partial_t h = 0$ and $u = 0$ for a mass at rest into the SH equations, one finds the following relation between the height h and the function g

$$b \partial_x h = g. \quad (4.8)$$

Note that g is not yet defined for $u = 0$, as described in Section 4.1. Nevertheless, for $u = 0$ the values of g do not exceed certain bounds, because g is defined as the sum of a gravitational force term

$$g_1 := \sin \zeta$$

and a second term, which expresses the Coulomb friction:

$$g_2 := -\frac{u}{|u|} \cos \zeta \tan \delta \quad \text{for } u \neq 0.$$

In a simple Coulomb friction model, static friction does not exceed dynamic friction and therefore

$$-\cos \zeta \tan \delta \leq g_2 \leq \cos \zeta \tan \delta \quad \text{for } u = 0.$$

These considerations lead us directly to the definition of an admissible profile.

Definition 4.2.4 (Admissible profile). *A mass profile h is said to be admissible if its spatial derivative $\partial_x h$ satisfies*

$$\min[\sin \zeta + R, \max(b \partial_x h, \sin \zeta - R)] = b \partial_x h, \quad (4.9)$$

where $\pm R$ denotes the force term resulting from friction:

$$R = \cos \zeta \tan \delta.$$

Moreover, we demand the collision operator Q to vanish in equilibrium. This condition does make sense in analogy to classical gas dynamics. Up to now, the only constraint on Q is the vanishing of the first two moments.

Therefore, for steady states of the form $\partial_t h = 0$ and $u = 0$, Equation (4.6) becomes

$$\xi \partial_x M(h, \xi) + g \partial_\xi M(h, \xi) = 0. \quad (4.10)$$

As described in [PS01], one can convert this equation into an ordinary differential equation for χ using (4.8) and the substitution $\omega = \xi/\sqrt{h}$. In particular, for admissible profiles $\partial_x h \neq 0$, we get

$$\begin{aligned} & \frac{\partial_x h}{2} \left[\frac{\xi}{\sqrt{h}} \chi \left(\frac{\xi}{\sqrt{h}} \right) - \frac{\xi^2}{h} \chi' \left(\frac{\xi}{\sqrt{h}} \right) \right] + g \chi' \left(\frac{\xi}{\sqrt{h}} \right) = 0 \\ \Leftrightarrow & \quad \frac{\partial_x h}{2} \left[\omega \chi(\omega) - \omega^2 \chi'(\omega) + 2b \chi'(\omega) \right] = 0 \\ \Leftrightarrow & \quad \left[\omega \chi(\omega) + (2b - \omega^2) \chi'(\omega) \right] = 0. \end{aligned} \quad (4.11)$$

Under the conditions (4.5), Equation (4.11) admits the unique solution

$$\chi(\omega) = C \sqrt{(2b - \omega^2)_+}, \quad \text{where } C = \frac{1}{\pi b}. \quad (4.12)$$

Finally, the microscopic equilibrium density for a constant inclination angle is specified by

$$M(h, \xi - u) = \sqrt{h} \chi \left(\frac{\xi - u}{\sqrt{h}} \right) = \frac{\sqrt{2h}}{\pi \sqrt{b}} \sqrt{\left(1 - \frac{(\xi - u)^2}{2bh}\right)_+}. \quad (4.13)$$

To preserve the steady states numerically, the discretization of the kinetic equation and, in particular, that of the source term is extremely important. In [PS01] the authors construct a scheme with reflections following the concept of upwinding the sources at interfaces as described in [KPS03] and [PS03]. Here, we do not follow this concept since internal friction is a volumic force rather than an interface force and the interpretation as a surface effect is not evident. But away from physical meaning, there are ideas to treat friction as a bottom topography [MVB⁺02].

We will now present a Finite Volume scheme in one spatial dimension for the SH equations (4.1) based on the (semi-)kinetic representation, i.e. a numerical method that preserves the steady states of a granular mass at rest. Furthermore, it can be shown easily that the scheme

preserves the nonnegativity of the materials' height at least in regions of large deformations. The numerical scheme is constructed by integrating a discretization of the kinetic Equation (4.6). The details of this approach can be found in [KS08]. Here, we will only present the main ideas and the resulting Finite Volume scheme for the SH equations.

For the construction of the scheme we proceed as follows: First, we discretize the kinetic Equation (4.6) directly in the density M . Integration of the method leads to a consistent macroscopic FVM, which is conservative and in some sense stable in the first component. Since this method is not able to preserve steady states, we introduce a modified microscopic numerical flux and a particular discretization of the source term. At the end, we choose a combination of the two approaches and use the the second one for the starting and stopping process and the first one otherwise.

We consider an equidistant mesh of $\mathbb{R} \times (0, \infty)$ with points $(x_i, t_n) = (i \Delta x, n \Delta t)$, $i \in \mathbb{Z}, n \in \mathbb{N}$ and cells $C_i = [x_{i-1/2}, x_{i+1/2}]$ with $x_{i+1/2} = (x_i + x_{i+1})/2$. Based on that mesh, we define cell averages

$$\begin{aligned} h_i^n &:= \frac{1}{|C_i|} \int_{C_i} h(x, t) \, dx, \\ q_i^n &:= (hu)_i^n := \frac{1}{|C_i|} \int_{C_i} h(x, t) u(x, t) \, dx, \end{aligned}$$

and $u_i^n := (hu)_i^n / h_i^n$.

With these quantities, we define the discrete density of particles $M_i^n(\xi)$ by

$$M_i^n(\xi) := M(h_i^n, \xi - u_i^n).$$

In order to discretize Equation (4.6), we firstly neglect the collision operator Q and get

$$f_i^{n+1}(\xi) - M_i^n(\xi) + \lambda \xi (M_{i+1/2}^n(\xi) - M_{i-1/2}^n(\xi)) + \Delta t g_i^n(M_\xi)_i^n(\xi) = 0, \quad (4.14)$$

where $\lambda := \Delta t / \Delta x$ and the microscopic fluxes $M_{i+1/2}^n$ are defined by the Upwind formula

$$M_{i+1/2}^n = \begin{cases} M_i^n(\xi), & \text{if } \xi \geq 0, \\ M_{i+1}^n(\xi), & \text{if } \xi < 0. \end{cases} \quad (4.15)$$

Since we neglect the collision operator, the density $f_i^{n+1}(\xi)$ is no longer an equilibrium density, i.e. it no longer satisfies the conditions (4.3). Therefore, the resulting density is denoted by $f_i^{n+1}(\xi)$ instead of M_i^{n+1} , but can be projected back to the class of equilibrium densities via the moments

$$\mathbf{U}_i^{n+1} := \begin{pmatrix} h_i^{n+1} \\ q_i^{n+1} \end{pmatrix} := \int_{\mathbb{R}} \begin{pmatrix} 1 \\ \xi \end{pmatrix} f_i^{n+1}(\xi) \, d\xi. \quad (4.16)$$

As described in Section 1.3, this procedure is common practice in constructing kinetic schemes [Str95], [PS01], [Per02]. The microscopic density $M_i^{n+1}(\xi)$ is obtained from the data at time level t_n as follows:

1. receive $f_i^{n+1}(\xi)$ from Equation (4.14),
2. compute the macroscopic quantity $\mathbf{U}_i^{n+1} = (h_i^{n+1}, q_i^{n+1})^T$ via Equation (4.16),
3. the density $M_i^{n+1}(\xi) = M(h_i^{n+1}, \xi - u_i^{n+1})$ is an equilibrium density again, i.e. it satisfies (4.3).

To discretize the source term, we simply choose $g_i^n = g(u_i^n)$ for $u_i^n \neq 0$ and a discretized form of (4.9) for admissible profiles. If the profile is not in equilibrium, we know that static friction is bounded by dynamic friction. These facts motivate the following definition of g_i^n :

$$g_i^n := \begin{cases} \tilde{g}_i^n, & \text{if } u_i^n = 0 \\ g(u_i^n), & \text{otherwise,} \end{cases}$$

where

$$\tilde{g}_i^n = \min \left[\sin \zeta + R, \max \left(\beta ((h_{i+1}^n - h_{i-1}^n)/2\Delta x), \sin \zeta - R \right) \right].$$

A macroscopic scheme can be obtained by building the first two moments of the microscopic scheme.

Instead of Equation (4.14), we will use the microscopic scheme

$$f_i^{n+1}(\xi + \Delta t g_i^n) - M_i^n(\xi) + \lambda \xi (M_{i+1/2}^n(\xi) - M_{i-1/2}^n(\xi)) = 0 \quad (4.17)$$

in further considerations (see Remark (4.2.5)).

Computing the first two moments of Equation (4.17), leads us to

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \lambda \left(\mathbf{F}_{i+1/2}^n - \mathbf{F}_{i-1/2}^n \right) + \Delta t \mathbf{S}_i^n \quad (4.18)$$

with

$$\mathbf{S}_i^n := \begin{pmatrix} 0 \\ g_i^n h_i^{n+1} \end{pmatrix}$$

and the macroscopic numerical fluxes

$$\mathbf{F}_{i+1/2}^n = \mathbf{F}(\mathbf{U}_i^n, \mathbf{U}_{i+1}^n) := \int_{\xi \geq 0} \xi \begin{pmatrix} 1 \\ \xi \end{pmatrix} M_i^n(\xi) d\xi + \int_{\xi < 0} \xi \begin{pmatrix} 1 \\ \xi \end{pmatrix} M_{i+1}^n(\xi) d\xi.$$

For the components of $\mathbf{F}_{i+1/2}^n = ((F_{i+1/2}^n)^1, (F_{i+1/2}^n)^2)^T$, we get

$$(F_{i+1/2}^n)^1 = \underbrace{\int_{\xi \geq 0} \xi M_i^n(\xi) d\xi}_{=: I_{1,i}^n} + \underbrace{\int_{\xi < 0} \xi M_{i+1}^n(\xi) d\xi}_{=: I_{2,i+1}^n},$$

and

$$(F_{i+1/2}^n)^2 = \underbrace{\int_{\xi \geq 0} \xi^2 M_i^n(\xi) d\xi}_{=: I_{3,i}^n} + \underbrace{\int_{\xi < 0} \xi^2 M_{i+1}^n(\xi) d\xi}_{=: I_{4,i+1}^n},$$

where $I_{1,i}^n = I_1(h_i^n, u_i^n)$, $I_{2,i}^n = I_2(h_i^n, u_i^n)$, $I_{3,i}^n = I_3(h_i^n, u_i^n)$ and $I_{4,i}^n = I_4(h_i^n, u_i^n)$ and the functions I_1, \dots, I_4 are given by

$$I_1(h, u) := \begin{cases} hu, & \text{if } a \leq -1 \\ \frac{2}{3\pi} \sqrt{2b} h^{3/2} \cos^3(\arcsin(a)) \\ + \frac{2}{\pi} hu \left(\frac{\pi}{4} - \frac{1}{2} \arcsin(a) - \frac{1}{4} \sin(2 \arcsin(a)) \right), & \text{if } -1 < a < 1 \\ 0, & \text{if } a \geq 1, \end{cases} \quad (4.19)$$

$$I_2(h, u) := \begin{cases} 0, & \text{if } a \leq -1 \\ -\frac{2}{3\pi} \sqrt{2b} h^{3/2} \cos^3(\arcsin(a)) \\ + \frac{2}{\pi} hu \left(\frac{\pi}{4} + \frac{1}{2} \arcsin(a) + \frac{1}{4} \sin(2 \arcsin(a)) \right), & \text{if } -1 < a < 1 \\ hu, & \text{if } a \geq 1, \end{cases} \quad (4.20)$$

$$I_3(h, u) := \begin{cases} hu^2 + \frac{1}{2}bh^2, & \text{if } a \leq -1 \\ \frac{4}{\pi} bh^2 \left(\frac{\pi}{16} - \frac{1}{8} \arcsin(a) + \frac{1}{32} \sin(a) \right) \\ - \frac{4}{3\pi} \sqrt{2b} h^{3/2} u \cos^3(\arcsin(a)) \\ + \frac{2}{\pi} hu^2 \left(\frac{\pi}{4} - \frac{1}{2} \arcsin(a) - \frac{1}{4} \sin(2 \arcsin(a)) \right), & \text{if } -1 < a < 1 \\ 0, & \text{if } a \geq 1, \end{cases} \quad (4.21)$$

$$I_4(h, u) := \begin{cases} 0, & \text{if } a \leq -1 \\ \frac{4}{\pi} bh^2 \left(\frac{1}{8} \arcsin(a) + \frac{\pi}{16} - \frac{1}{32} \sin(4 \arcsin(a)) \right) \\ - \frac{4}{3\pi} \sqrt{2b} h^{3/2} u \cos^3(\arcsin(a)) \\ + \frac{2}{\pi} hu^2 \left(\frac{\pi}{4} + \frac{1}{2} \arcsin(a) + \frac{1}{4} \sin(2 \arcsin(a)) \right), & \text{if } -1 < a < 1 \\ hu^2 + \frac{1}{2}bh^2, & \text{if } a \geq 1, \end{cases} \quad (4.22)$$

with $a := a(h, u) := -u/\sqrt{2bh}$.

Remark 4.2.5. *The only difference between the resulting macroscopic schemes relying on (4.14) and (4.17) is the term $\Delta t g_i^n h_i^n$. Using (4.17), one ends up with $\Delta t g_i^n h_i^{n+1}$ instead.*

In [KS08] it is figured out why this scheme can not preserve admissible profiles in general.

Therefore, we propose the following modifications in the microscopic numerical flux function and the discretized source term.

As new microscopic fluxes, we choose

$$\tilde{M}_{i+1/2}^n = \begin{cases} M\left(\frac{h_i^n + h_{i+1}^n}{2}, \xi - u_i^n\right), & \text{if } \xi \geq 0 \\ M\left(\frac{h_i^n + h_{i+1}^n}{2}, \xi - u_{i+1}^n\right), & \text{if } \xi < 0. \end{cases} \quad (4.23)$$

This means that the actual upwinding takes place only in the velocity u and no longer in the height h . With (4.23), the new macroscopic fluxes $\tilde{\mathbf{F}}_{i+1/2}^n = ((\tilde{F}_{i+1/2}^n)^1, (\tilde{F}_{i+1/2}^n)^2)^T$ read

$$(\tilde{F}_{i+1/2}^n)^1 = \underbrace{\int_{\xi \geq 0} \xi M\left(\frac{h_i^n + h_{i+1}^n}{2}, \xi - u_i^n\right) d\xi}_{=: \tilde{I}_{1,i,i+1}^n} + \underbrace{\int_{\xi < 0} \xi M\left(\frac{h_i^n + h_{i+1}^n}{2}, \xi - u_{i+1}^n\right) d\xi}_{=: \tilde{I}_{2,i,i+1}^n}$$

and

$$(\tilde{F}_{i+1/2}^n)^2 = \underbrace{\int_{\xi \geq 0} \xi^2 M\left(\frac{h_i^n + h_{i+1}^n}{2}, \xi - u_i^n\right) d\xi}_{=: \tilde{I}_{3,i,i+1}^n} + \underbrace{\int_{\xi < 0} \xi^2 M\left(\frac{h_i^n + h_{i+1}^n}{2}, \xi - u_{i+1}^n\right) d\xi}_{=: \tilde{I}_{4,i,i+1}^n}$$

with

$$\begin{aligned} \tilde{I}_{1,i,j}^n &= I_1\left(\frac{h_i^n + h_j^n}{2}, u_i^n\right), & \tilde{I}_{2,i,j}^n &= I_2\left(\frac{h_i^n + h_j^n}{2}, u_j^n\right), \\ \tilde{I}_{3,i,j}^n &= I_3\left(\frac{h_i^n + h_j^n}{2}, u_i^n\right), & \tilde{I}_{4,i,j}^n &= I_4\left(\frac{h_i^n + h_j^n}{2}, u_j^n\right) \end{aligned}$$

and I_1, \dots, I_4 given by (4.19)–(4.22).

Additionally, we change the source term for slowly moving or non-moving masses, i.e. masses that satisfy $0 < |q_i^n| \leq |\Delta t g(u_i^n) h_i^{n+1}|$ or $u_i^n = 0$: for admissible profiles, we replace h_i^{n+1} in the source term by $((h_{i+1}^n + 2h_i^n + h_{i-1}^n)/4)$ resulting in the modified source term

$$\tilde{\mathbf{S}}_i^n := \begin{pmatrix} 0 \\ \tilde{g}_i^n \left(\frac{h_{i+1}^n + 2h_i^n + h_{i-1}^n}{4} \right) \end{pmatrix}.$$

For fast moving masses, we still take \mathbf{S}_i^n and for slowly moving masses, we use a combination of \mathbf{S}_i^n and $\tilde{\mathbf{S}}_i^n$ with the coefficients

$$\mu := \begin{cases} -q_i^n / \Delta t g(q_i^n) h_i^{n+1}, & \text{if } q_i^n \neq 0 \\ 0, & \text{else} \end{cases}$$

and

$$(1 - \mu) = \begin{cases} (1 + q_i^n / \Delta t g(q_i^n) h_i^{n+1}), & \text{if } q_i^n \neq 0 \\ 1, & \text{else.} \end{cases}$$

Here, we assumed that the height h_i^n is always positive such that we have $\text{sgn}(q_i^n) = \text{sgn}(u_i^n)$ and $q_i^n = 0$ iff $u_i^n = 0$. In Theorem 4.2.8, we give a justification of that assumption.

All together, we end up with the macroscopic scheme

$$U_i^{n+1} = \begin{cases} U_i^n - \lambda (\tilde{F}_{i+1/2}^n - \tilde{F}_{i-1/2}^n) + \Delta t (\mu S_i^n + (1 - \mu) \tilde{S}_i^n), & \text{if } |\mu| \leq 1 \\ U_i^n - \lambda (\tilde{F}_{i+1/2}^n - \tilde{F}_{i-1/2}^n) + \Delta t S_i^n, & \text{else.} \end{cases} \quad (4.24)$$

Remark 4.2.6. Note that the coefficient μ does not always result in a convex combination of the different discretizations of the source terms S_i^n and \tilde{S}_i^n . Assume for example $q_i^n, h_i^n > 0$ for some $i \in \mathbb{Z}$ and $n \in \mathbb{N}$. Moreover, let $\zeta, \delta \in (0, \pi/2)$ with $\zeta > \delta$. Then, we get

$$g(q_i^n) = \sin \zeta - \cos \zeta \tan \delta > 0$$

and thus $\mu < 0$ in case of $h_i^{n+1} > 0$. But, under the assumption that the heights h_i^n and h_i^{n+1} are positive and $q_i^n < 0$ or $q_i^n > 0$ and $\zeta < \delta$, we always get $\mu > 0$.

It is easy to see that scheme (4.24) is able to preserve admissible states and can handle the starting and stopping of inadmissible profiles. For details see [KS08]. But, in the case of strongly inadmissible profiles, scheme (4.24) causes small oscillations. To get rid of this drawback, we decided to couple the two schemes (4.18) and (4.24) by using the first one for rapid flows and the second one for the simulation of starting and stopping processes as well as for masses in equilibrium.

The resulting coupled scheme thus reads

$$U_i^{n+1} = \begin{cases} U_i^n - \lambda (\tilde{F}_{i+1/2}^n - \tilde{F}_{i-1/2}^n) + \Delta t (\mu S_i^n + (1 - \mu) \tilde{S}_i^n), & \text{if } |\mu| \leq 1 \\ U_i^n - \lambda (F_{i+1/2}^n - F_{i-1/2}^n) + \Delta t S_i^n, & \text{else.} \end{cases} \quad (4.25)$$

Written in the components h_i^n and q_i^n , the method (4.25) can be written as

Scheme 4.2.7.

$$\left. \begin{aligned} h_i^{n+1} &= h_i^n - \lambda ((\tilde{F}_{i+1/2}^n)^1 - (\tilde{F}_{i-1/2}^n)^1) \\ q_i^{n+1} &= -\lambda ((\tilde{F}_{i+1/2}^n)^2 - (\tilde{F}_{i-1/2}^n)^2) \\ &\quad + (1 - \mu) \Delta t \tilde{g}_i^n ((h_{i+1}^n + 2h_i^n + h_{i-1}^n)/4) \end{aligned} \right\} \text{if } |\mu| \leq 1$$

$$\left. \begin{aligned} h_i^{n+1} &= h_i^n - \lambda ((F_{i+1/2}^n)^1 - (F_{i-1/2}^n)^1) \\ q_i^{n+1} &= q_i^n - \lambda ((F_{i+1/2}^n)^2 - (F_{i-1/2}^n)^2) + \Delta t g(q_i^n) h_i^{n+1} \end{aligned} \right\} \text{else}$$

where

$$\begin{aligned} F_{i+1/2}^n &= ((F_{i+1/2}^n)^1, (F_{i+1/2}^n)^2)^T = (I_{1,i}^n + I_{2,i+1}^n, I_{3,i}^n + I_{4,i+1}^n)^T, \\ \tilde{F}_{i+1/2}^n &= ((\tilde{F}_{i+1/2}^n)^1, (\tilde{F}_{i+1/2}^n)^2)^T = (\tilde{I}_{1,i,i+1}^n + \tilde{I}_{2,i,i+1}^n, \tilde{I}_{3,i,i+1}^n + \tilde{I}_{4,i,i+1}^n)^T \end{aligned}$$

and \tilde{g}_i^n and μ are given above.

In [PS01], the authors formulate a theorem for their kinetic method solving the shallow water equations with a source which ensures nonnegativity of the height h . Because part of the method was constructed following the Ansatz made in [PS01], we can easily apply that theorem to the scheme (4.2.7) in the case of rapidly moving granular masses.

Theorem 4.2.8. *Under the CFL condition*

$$\max_{i \in \mathbb{Z}} (|u_i^n| + \sqrt{2bh_i^n}) \leq \frac{\Delta x}{\Delta t} = \frac{1}{\lambda} \quad \text{and} \quad |q_i^n| \geq |\Delta t g(q_i^n) h_i^{n+1}| \quad \forall i \in \mathbb{Z},$$

the height h in scheme 4.2.7 remains nonnegative, i.e. we have $h_i^n \geq 0$ for all $i \in \mathbb{Z}$ and $n \in \mathbb{N}$, if $h_i^0 \geq 0$ for all $i \in \mathbb{Z}$.

Proof. The proof of Theorem 4.2.8 is similar to that in [PS01] and uses that $f_i^{n+1}(\xi + \Delta t g(u_i^n)) \geq 0$ for all $\xi \in \mathbb{R}$ and $i \in \mathbb{Z}$ as it is a convex combination of nonnegative quantities. The proposition follows directly from

$$h_i^{n+1} = \int_{\mathbb{R}} f_i^{n+1}(\xi + \Delta t g(u_i^n)) d\xi.$$

□

In the case $|q_i^n| < |\Delta t g(q_i^n) h_i^{n+1}|$, the considered mass is only moving very slowly or not at all, so we can expect that the height is kept nonnegative, too.

It is easy to see that the method is conservative in the first component and consistent in the sense that

$$\mathbf{F}(\mathbf{U}, \mathbf{U}) = \tilde{\mathbf{F}}(\mathbf{U}, \mathbf{U}) = \begin{pmatrix} hu \\ hu^2 + \frac{1}{2}bh^2 \end{pmatrix}.$$

Several numerical tests of the method can be found in [KS08].

In the next section, we will show how this scheme can be managed within the particle framework of the FVPM. In Chapter 5, we will repeat some of the numerical examples in [KS08] computed with the FVPM with kinetic fluxes.

4.3 The FVPM with kinetic fluxes

As in Section 4.2, we begin with the (semi-)kinetic representation of the SH equations:

$$\partial_t M + \xi \partial_x M + G(u) \partial_\xi M = Q \tag{4.26}$$

where

$$G(u) := \begin{cases} g(u), & \text{if } u \neq 0 \\ \min((\sin \zeta + R), \max(b \partial_x h, (\sin \zeta - R))), & \text{if } u = 0 \end{cases} \quad (4.27)$$

and

$$\int_{\mathbb{R}} \begin{pmatrix} 1 \\ \xi \end{pmatrix} Q \, d\xi = \mathbf{0}. \quad (4.28)$$

Now we define the generalized cell averages according to the FVPM:

$$\begin{aligned} h_i(t) &:= \frac{1}{V_i(t)} \int_{\Omega} h(x, t) \psi_i(x, t) \, dx, \\ (hu)_i(t) = q_i(t) &:= \frac{1}{V_i(t)} \int_{\Omega} h(x, t) u(x, t) \psi_i(x, t) \, dx, \\ u_i(t) &:= \frac{q_i(t)}{h_i(t)} \end{aligned}$$

and

$$M_i(t, \xi) := \frac{1}{V_i(t)} \int_{\Omega} M(h(x, t), \xi - u(x, t)) \psi_i(x, t) \, dx.$$

Remember from Chapter 2 that we have

$$\begin{aligned} \partial_t \psi_i(x, t) &= - \sum_{j \in \mathcal{N}(i)} (\dot{x}_i(t) \Gamma_{ji}(x, t) - \dot{x}_j(t) \Gamma_{ij}(x, t)), \\ \partial_x \psi_i(x, t) &= \sum_{j \in \mathcal{N}(i)} (\Gamma_{ji}(x, t) - \Gamma_{ij}(x, t)). \end{aligned}$$

By defining

$$\begin{aligned} Y &:= G(u) \partial_{\xi} M(h, \xi - u), \\ Y_i &:= \frac{1}{V_i(t)} \int_{\Omega} Y \psi_i \, dx, \\ Q_i &:= \frac{1}{V_i(t)} \int_{\Omega} Q \psi_i \, dx \end{aligned}$$

and integrating by parts, we thus get

$$\begin{aligned} \frac{d}{dt} (V_i M_i) &= \frac{d}{dt} \int_{\Omega} M \psi_i \, dx = \int_{\Omega} (\partial_t M \psi_i + M \partial_t \psi_i) \, dx \\ &= -\xi \int_{\Omega} \partial_x M \psi_i \, dx + \int_{\Omega} M \partial_t \psi_i \, dx - \int_{\Omega} G(u) \partial_{\xi} M \psi_i \, dx + \int_{\Omega} Q \psi_i \, dx \\ &= \xi \sum_{j \in \mathcal{N}(i)} \int_{\Omega} M (\Gamma_{ji} - \Gamma_{ij}) \, dx - \sum_{j \in \mathcal{N}(i)} \int_{\Omega} M (\dot{x}_i \Gamma_{ji} - \dot{x}_j \Gamma_{ij}) \, dx \\ &\quad - V_i Y_i + V_i Q_i. \end{aligned}$$

In analogy to Section 4.2, we approximate M in the first integral on the right hand side by the upwind formula

$$M_{ij} = \begin{cases} M(h_i, \xi - u_i) n_{ij}, & \text{if } \xi n_{ij} \geq 0 \\ M(h_j, \xi - u_j) n_{ij}, & \text{if } \xi n_{ij} < 0 \end{cases} \quad (4.29)$$

and get

$$\frac{d}{dt} (V_i M_i) = -\xi \sum_{j \in \mathcal{N}(i)} |\beta_{ij}| M_{ij} - \sum_{j \in \mathcal{N}(i)} \int_{\Omega} M (\dot{x}_i \Gamma_{ji} - \dot{x}_j \Gamma_{ij}) dx - V_i Y_i + V_i Q_i. \quad (4.30)$$

To get a macroscopic scheme, we build the moments of Equation (4.30). With $\mathbf{U}_i := (h_i, (hu)_i)^T = (h_i, q_i)^T$ and because of

$$\begin{aligned} \int_{\mathbb{R}} \left(\frac{1}{\xi} \right) \frac{d}{dt} (V_i M_i) d\xi &= \frac{d}{dt} \int_{\mathbb{R}} \left(\frac{1}{\xi} \right) \int_{\Omega} M(h, \xi - u) \psi_i dx d\xi \\ &= \frac{d}{dt} \int_{\Omega} \int_{\mathbb{R}} \left(\frac{1}{\xi} \right) M(h, \xi - u) d\xi \psi_i dx \\ &= \frac{d}{dt} \int_{\Omega} \begin{pmatrix} h \\ hu \end{pmatrix} \psi_i = \frac{d}{dt} (V_i \mathbf{U}_i), \end{aligned}$$

we obtain

$$\begin{aligned} \frac{d}{dt} (V_i \mathbf{U}_i) &= - \sum_{j \in \mathcal{N}(i)} |\beta_{ij}| \int_{\mathbb{R}} \left(\frac{1}{\xi} \right) \xi M_{ij} d\xi - \sum_{j \in \mathcal{N}(i)} \int_{\Omega} \int_{\mathbb{R}} \left(\frac{1}{\xi} \right) M d\xi (\dot{x}_i \Gamma_{ji} - \dot{x}_j \Gamma_{ij}) dx \\ &\quad - V_i \int_{\mathbb{R}} \left(\frac{1}{\xi} \right) Y_i d\xi + V_i \int_{\mathbb{R}} \left(\frac{1}{\xi} \right) Q_i d\xi. \end{aligned} \quad (4.31)$$

To simplify the comparability to Scheme 4.2.7, let us assume fixed particles for the moment.

In analogy to Section 4.2, the macroscopic numerical fluxes \mathbf{F}_{ij} are defined as the moments of the microscopic numerical fluxes M_{ij} :

$$\begin{aligned} \mathbf{F}_{ij} &= \mathbf{F}(\mathbf{U}_i, \mathbf{U}_j, n_{ij}) = ((F_{ij})^1, (F_{ij})^2)^T := \int_{\mathbb{R}} \left(\frac{1}{\xi} \right) \xi M_{ij} d\xi \\ &= \left(\int_{\xi n_{ij} \geq 0} \left(\frac{1}{\xi} \right) \xi M(h_i, \xi - u_i) d\xi + \int_{\xi n_{ij} < 0} \left(\frac{1}{\xi} \right) \xi M(h_j, \xi - u_j) d\xi \right) n_{ij}. \end{aligned}$$

For the first component, we compute

$$\begin{aligned}
 (F_{ij})^1 &= \begin{cases} \int_{\xi \geq 0} \xi M(h_i, \xi - u_i) d\xi + \int_{\xi < 0} \xi M(h_j, \xi - u_j) d\xi, & \text{if } n_{ij} = 1 \\ - \left(\int_{\xi \geq 0} \xi M(h_j, \xi - u_j) d\xi + \int_{\xi < 0} \xi M(h_i, \xi - u_i) d\xi \right), & \text{if } n_{ij} = -1 \end{cases} \\
 &= \begin{cases} I_{1,i} + I_{2,j}, & \text{if } n_{ij} = 1 \\ -(I_{1,j} + I_{2,i}), & \text{if } n_{ij} = -1 \end{cases}
 \end{aligned}$$

where $I_{1,i} = I_1(h_i, u_i)$ and $I_{2,i} = I_2(h_i, u_i)$ with I_1 and I_2 given by (4.19) and (4.20), respectively.

In the same manner, we obtain for the second component

$$\begin{aligned}
 (F_{ij})^2 &= \begin{cases} \int_{\xi \geq 0} \xi^2 M(h_i, \xi - u_i) d\xi + \int_{\xi < 0} \xi^2 M(h_j, \xi - u_j) d\xi, & \text{if } n_{ij} = 1 \\ - \left(\int_{\xi \geq 0} \xi^2 M(h_j, \xi - u_j) d\xi + \int_{\xi < 0} \xi^2 M(h_i, \xi - u_i) d\xi \right), & \text{if } n_{ij} = -1 \end{cases} \\
 &= \begin{cases} I_{3,i} + I_{4,j}, & \text{if } n_{ij} = 1 \\ -(I_{3,j} + I_{4,i}), & \text{if } n_{ij} = -1. \end{cases}
 \end{aligned}$$

Again, we have $I_{3,i} = I_3(h_i, u_i)$ and $I_{4,i} = I_4(h_i, u_i)$ with I_3 and I_4 given in (4.21) and (4.22), respectively.

With

$$\int_{\mathbb{R}} \begin{pmatrix} 1 \\ \xi \end{pmatrix} M d\xi = \begin{pmatrix} h \\ hu \end{pmatrix}, \quad \int_{\mathbb{R}} \begin{pmatrix} 1 \\ \xi \end{pmatrix} \partial_{\xi} M d\xi = - \begin{pmatrix} 0 \\ h \end{pmatrix}$$

and

$$\int_{\mathbb{R}} \begin{pmatrix} 1 \\ \xi \end{pmatrix} Q d\xi = \mathbf{0},$$

we get

$$\begin{aligned}
 \frac{d}{dt} \mathbf{U}_i &= -\frac{1}{V_i} \sum_{j \in \mathcal{N}(i)} |\beta_{ij}| \mathbf{F}_{ij} - \frac{1}{V_i} \int_{\Omega} G(u) \int_{\mathbb{R}} \begin{pmatrix} 1 \\ \xi \end{pmatrix} \partial_{\xi} M d\xi \psi_i dx + \frac{1}{V_i} \int_{\Omega} \int_{\mathbb{R}} \begin{pmatrix} 1 \\ \xi \end{pmatrix} Q d\xi \psi_i dx \\
 &= -\frac{1}{V_i} \sum_{j \in \mathcal{N}(i)} |\beta_{ij}| \mathbf{F}_{ij} + \frac{1}{V_i} \int_{\Omega} G(u) \begin{pmatrix} 0 \\ h \end{pmatrix} \psi_i dx.
 \end{aligned}$$

For the second component of the last term on the right hand side, we write

$$\frac{1}{V_i} \int_{\Omega} G(u) h \psi_i dx \approx \begin{cases} \frac{1}{V_i} \int_{\Omega} g(u) h \psi_i dx, & \text{if } u_i \neq 0 \\ \frac{1}{V_i} \int_{\Omega} \min((\sin \zeta + R), \max(b \partial_x h, (\sin \zeta - R))) h \psi_i dx, & \text{if } u_i = 0 \end{cases}$$

and approximate

$$\frac{1}{V_i} \int_{\Omega} g(u) h \psi_i dx = (g(u) h)_i \approx g(u_i) h_i$$

in the case $u_i \neq 0$. For $u_i = 0$, we write

$$\begin{aligned} \frac{1}{V_i} \int_{\Omega} b \partial_x h h \psi_i dx &= \frac{b}{2 V_i} \int_{\Omega} \partial_x (h^2) \psi_i dx \\ &= -\frac{b}{2 V_i} \int_{\Omega} h^2 \partial_x \psi_i dx + \frac{b}{2 V_i} \int_{\partial \Omega} h^2 \psi_i n ds \\ &= -\frac{b}{2 V_i} \sum_{j \in \mathcal{N}(i)} \int_{\Omega} h^2 (\Gamma_{ji} - \Gamma_{ij}) dx + \frac{b}{2 V_i} \int_{\partial \Omega} h^2 \psi_i n ds \\ &\approx \frac{b}{2 V_i} \sum_{j \in \mathcal{N}(i)} \beta_{ij} \left(\frac{h_i + h_j}{2} \right)^2 + D_i \end{aligned}$$

with a suitable approximation of the boundary term

$$D_i \approx \frac{b}{2 V_i} \int_{\partial \Omega} h^2 \psi_i n ds.$$

All in all, we get the following discretization of the source term:

$$\frac{1}{V_i} \int_{\Omega} G(u) h \psi_i dx \approx \begin{cases} g(u_i) h_i, & \text{if } u_i \neq 0 \\ \min \left((\sin \zeta + R) h_i, \max \left(\frac{b}{2 V_i} \sum_{j \in \mathcal{N}(i)} \beta_{ij} \left(\frac{h_i + h_j}{2} \right)^2, (\sin \zeta - R) h_i \right) \right), & \text{if } u_i = 0 \end{cases} \quad (4.32)$$

Similar to (4.29), we define the modified numerical fluxes where the upwinding takes place only in the variable u_i^n :

$$\tilde{M}_{ij} := \begin{cases} M \left(\frac{h_i + h_j}{2}, \xi - u_i \right) n_{ij}, & \text{if } \xi n_{ij} \geq 0 \\ M \left(\frac{h_i + h_j}{2}, \xi - u_j \right) n_{ij}, & \text{if } \xi n_{ij} < 0 \end{cases}$$

and

$$\begin{aligned}
\tilde{\mathbf{F}}_{ij} &= \tilde{\mathbf{F}}(\mathbf{U}_i, \mathbf{U}_j, n_{ij}) = \left((\tilde{F}_{ij})^1, (\tilde{F}_{ij})^2 \right)^T := \int_{\mathbb{R}} \begin{pmatrix} 1 \\ \xi \end{pmatrix} \xi \tilde{M}_{ij} d\xi \\
&= \left(\int_{\xi n_{ij} \geq 0} \begin{pmatrix} 1 \\ \xi \end{pmatrix} \xi M \left(\frac{h_i + h_j}{2}, \xi - u_i \right) d\xi \right. \\
&\quad \left. + \int_{\xi n_{ij} < 0} \begin{pmatrix} 1 \\ \xi \end{pmatrix} \xi M \left(\frac{h_i + h_j}{2}, \xi - u_j \right) d\xi \right) n_{ij} .
\end{aligned}$$

For the components $(\tilde{F}_{ij})^1$ and $(\tilde{F}_{ij})^2$, we compute

$$\begin{aligned}
(\tilde{F}_{ij})^1 &= \begin{cases} \int_{\xi \geq 0} \xi M \left(\frac{h_i + h_j}{2}, \xi - u_i \right) d\xi \\ \quad + \int_{\xi < 0} \xi M \left(\frac{h_i + h_j}{2}, \xi - u_j \right) d\xi , & \text{if } n_{ij} = 1 \\ \\ - \left(\int_{\xi \geq 0} \xi M \left(\frac{h_i + h_j}{2}, \xi - u_j \right) d\xi \right. \\ \quad \left. + \int_{\xi < 0} \xi M \left(\frac{h_i + h_j}{2}, \xi - u_i \right) d\xi \right) , & \text{if } n_{ij} = -1 \end{cases} \\
&= \begin{cases} \tilde{I}_{1,i,j} + \tilde{I}_{2,i,j} , & \text{if } n_{ij} = 1 \\ -(\tilde{I}_{1,j,i} + \tilde{I}_{2,j,i}) , & \text{if } n_{ij} = -1 \end{cases}
\end{aligned}$$

and

$$\begin{aligned}
(\tilde{F}_{ij})^2 &= \begin{cases} \int_{\xi \geq 0} \xi^2 M \left(\frac{h_i + h_j}{2}, \xi - u_i \right) d\xi \\ \quad + \int_{\xi < 0} \xi^2 M \left(\frac{h_i + h_j}{2}, \xi - u_j \right) d\xi , & \text{if } n_{ij} = 1 \\ \\ - \left(\int_{\xi \geq 0} \xi^2 M \left(\frac{h_i + h_j}{2}, \xi - u_j \right) d\xi \right. \\ \quad \left. + \int_{\xi < 0} \xi^2 M \left(\frac{h_i + h_j}{2}, \xi - u_i \right) d\xi \right) , & \text{if } n_{ij} = -1 \end{cases} \\
&= \begin{cases} \tilde{I}_{3,i,j} + \tilde{I}_{4,i,j} , & \text{if } n_{ij} = 1 \\ -(\tilde{I}_{3,j,i} + \tilde{I}_{4,j,i}) , & \text{if } n_{ij} = -1 \end{cases}
\end{aligned}$$

where

$$\begin{aligned}\tilde{I}_{1,i,j} &= I_1 \left(\frac{h_i + h_j}{2}, u_i \right), & \tilde{I}_{2,i,j} &= I_2 \left(\frac{h_i + h_j}{2}, u_j \right), \\ \tilde{I}_{3,i,j} &= I_3 \left(\frac{h_i + h_j}{2}, u_i \right), & \tilde{I}_{4,i,j} &= I_4 \left(\frac{h_i + h_j}{2}, u_j \right)\end{aligned}$$

and I_1, \dots, I_4 given by (4.19)–(4.22).

Now, we proceed as in Section 4.2 and use a combination of the discretizations of the source term given in (4.32) for slowly moving masses. In analogy to Scheme 4.2.7, we define the fully discretized FVPM with fixed particles and kinetic fluxes for the SH equations:

Scheme 4.3.1.

$$\left. \begin{aligned} h_i^{n+1} &= h_i^n - \frac{\Delta t}{V_i} \sum_{j \in \mathcal{N}(i)} |\beta_{ij}^n| (\tilde{F}_{ij}^n)^1 \\ q_i^{n+1} &= -\frac{\Delta t}{V_i} \sum_{j \in \mathcal{N}(i)} |\beta_{ij}^n| (\tilde{F}_{ij}^n)^2 + (1 - \mu) \Delta t \tilde{g}_i^n \end{aligned} \right\} \text{ if } |\mu| \leq 1$$

$$\left. \begin{aligned} h_i^{n+1} &= h_i^n - \frac{\Delta t}{V_i} \sum_{j \in \mathcal{N}(i)} |\beta_{ij}^n| (F_{ij}^n)^1 \\ q_i^{n+1} &= q_i^n - \frac{\Delta t}{V_i} \sum_{j \in \mathcal{N}(i)} |\beta_{ij}^n| (F_{ij}^n)^2 + \Delta t g(q_i^n) h_i^{n+1} \end{aligned} \right\} \text{ else}$$

with

$$\tilde{g}_i^n = \min \left((\sin \zeta + R) h_i^n, \max \left(\frac{b}{2V_i} \sum_{j \in \mathcal{N}(i)} \beta_{ij}^n \left(\frac{h_i^n + h_j^n}{2} \right)^2, (\sin \zeta - R) h_i^n \right) \right),$$

$$\mu = \begin{cases} -\frac{q_i^n}{\Delta t g(q_i^n) h_i^{n+1}}, & \text{if } u_i^n \neq 0 \\ 0, & \text{else} \end{cases}$$

and

$$\mathbf{F}_{ij}^n = ((F_{ij}^n)^1, (F_{ij}^n)^2)^T, \quad \tilde{\mathbf{F}}_{ij}^n = ((\tilde{F}_{ij}^n)^1, (\tilde{F}_{ij}^n)^2)^T.$$

It can be seen easily that Scheme 4.3.1 preserves the desired steady states of admissible granular profiles at rest.

Theorem 4.3.2 (Preservation of granular masses at rest). *Scheme 4.3.1 preserves admissible profiles of granular masses at rest which are characterized by*

$$u_i^n = 0 \quad \forall i \in \mathbb{Z}, \tag{4.33}$$

$$\begin{aligned} &\min \left((\sin \zeta + R) h_i^n, \max \left(\frac{b}{2V_i} \sum_{j \in \mathcal{N}(i)} \beta_{ij}^n \left(\frac{h_i^n + h_j^n}{2} \right)^2, (\sin \zeta - R) h_i^n \right) \right) \\ &= \frac{b}{2V_i} \sum_{j \in \mathcal{N}(i)} \beta_{ij}^n \left(\frac{h_i^n + h_j^n}{2} \right)^2 \quad \forall i \in \mathbb{Z}. \end{aligned} \tag{4.34}$$

Proof. Inserting (4.33) and (4.34) in the equation for q_i^{n+1} of Scheme 4.3.1 yields

$$q_i^{n+1} = -\frac{\Delta t}{V_i} \sum_{j \in \mathcal{N}(i)} |\beta_{ij}^n| (\tilde{F}_{ij}^n)^2 + \Delta t \frac{b}{2V_i} \sum_{j \in \mathcal{N}(i)} \beta_{ij}^n \left(\frac{h_i^n + h_j^n}{2} \right)^2 = 0,$$

as can be seen easily by evaluating the terms $\tilde{I}_{3,i,j}^n$ and $\tilde{I}_{4,i,j}^n$. Thus, we have $u_i^{n+1} = 0$ for all $i \in \mathbb{Z}$. Moreover, we get

$$h_i^{n+1} = h_i^n \quad \forall i \in \mathbb{Z}$$

by inserting $u_i^n = 0$ into $\tilde{I}_{1,i,j}^n$ and $\tilde{I}_{2,i,j}^n$. □

In Chapter 5, we apply Scheme 4.3.1 to some of the examples given in [KS08]. As we will see, Scheme 4.3.1 can manage the starting and stopping of inadmissible profiles and preserves the desired steady states of granular masses at rest.

Although we use only fixed particles in our numerical examples for the SH equations, we will now give an idea how a particle movement can be incorporated in the FVPM for the SH equations. We start by noting that non-constant steady states will not be preserved by the particle averaging with a general particle movement. Assume some steady state $u(x, t) = u(x)$ of a general conservation law. Then, moving particles will cause

$$\frac{d}{dt} u_i(t) = -\frac{\dot{V}_i(t)}{V_i^2(t)} \int_{\Omega} u(x) \psi_i(x, t) dx + \frac{1}{V_i(t)} \int_{\Omega} u(x) \partial_t \psi_i(x, t) dx.$$

For the reconstruction

$$u_h(x, t) = \sum_{i \in \mathbb{Z}} u_i(t) \psi_i(x, t),$$

we thus get

$$\begin{aligned} \frac{d}{dt} u_h(x, t) = \sum_{i \in \mathbb{Z}} \left[\left(-\frac{\dot{V}_i(t)}{V_i^2(t)} \int_{\Omega} u(x) \psi_i(x, t) dx + \frac{1}{V_i(t)} \int_{\Omega} u(x) \partial_t \psi_i(x, t) dx \right) \psi_i(x, t) \right. \\ \left. + u_i(t) \partial_t \psi_i(x, t) \right] \neq 0 \end{aligned} \tag{4.35}$$

in general. Even if we had a scheme that ensures

$$\frac{d}{dt} u_i(t) = 0 \quad \forall i \in \mathbb{Z},$$

we still get for the time derivative of the reconstruction

$$\frac{d}{dt} u_h(x, t) = \sum_{i \in \mathbb{Z}} u_i(t) \partial_t \psi_i(x, t) = \sum_{i, j \in \mathbb{Z}} u_i (\dot{x}_j \Gamma_{ij} - \dot{x}_i \Gamma_{ji}) \neq 0$$

usually.

Therefore, particles should be non-moving if the granular mass is in equilibrium. Although a Lagrangian movement of the particles is not recommended in cases of discontinuous velocities

[Tel05], it could be an adequate choice for the SH equations.

To incorporate the particle velocity in the numerical scheme, we start with Equation (4.31) again:

$$\begin{aligned} \frac{d}{dt} (V_i \mathbf{U}_i) &= - \sum_{j \in \mathcal{N}(i)} |\beta_{ij}| \int_{\mathbb{R}} \left(\frac{1}{\xi} \right) \xi M_{ij} d\xi - \sum_{j \in \mathcal{N}(i)} \int_{\Omega} \int_{\mathbb{R}} \left(\frac{1}{\xi} \right) M d\xi (\dot{x}_i \Gamma_{ji} - \dot{x}_j \Gamma_{ij}) dx \\ &\quad - V_i \int_{\mathbb{R}} \left(\frac{1}{\xi} \right) Y_i d\xi + V_i \int_{\mathbb{R}} \left(\frac{1}{\xi} \right) Q_i d\xi. \end{aligned}$$

For the second term on the right hand side, we write

$$\begin{aligned} \sum_{j \in \mathcal{N}(i)} \int_{\Omega} \int_{\mathbb{R}} \left(\frac{1}{\xi} \right) M d\xi (\dot{x}_i \Gamma_{ji} - \dot{x}_j \Gamma_{ij}) dx &\approx \sum_{j \in \mathcal{N}(i)} \int_{\Omega} \int_{\mathbb{R}} \left(\frac{1}{\xi} \right) M d\xi (\Gamma_{ji} - \Gamma_{ij}) dx \bar{x}_{ij} \\ &\approx \sum_{j \in \mathcal{N}(i)} |\beta_{ij}| \int_{\mathbb{R}} \left(\frac{1}{\xi} \right) M_{ij} d\xi \bar{x}_{ij} \end{aligned}$$

and define

$$\begin{aligned} \mathbf{K}_{ij} &= \mathbf{K}(\mathbf{U}_i, \mathbf{U}_j, n_{ij}) = ((K_{ij})^1, (K_{ij})^2)^T := \int_{\mathbb{R}} \left(\frac{1}{\xi} \right) M_{ij} d\xi \\ &= \left(\int_{\xi n_{ij} \geq 0} \left(\frac{1}{\xi} \right) M(h_i, \xi - u_i) d\xi + \int_{\xi n_{ij} < 0} \left(\frac{1}{\xi} \right) M(h_j, \xi - u_j) d\xi \right) n_{ij}. \end{aligned}$$

Note that because of $(K_{ij})^2 = (F_{ij})^1$, we just have to compute the first component of \mathbf{K}_{ij} :

$$\begin{aligned} K_{ij}^1 &= \int_{\mathbb{R}} M_{ij}(\xi) d\xi \\ &= \begin{cases} \underbrace{\int_{\xi \geq 0} M(h_i, \xi - u_i) d\xi}_{=: J_{1,i}} + \underbrace{\int_{\xi < 0} M(h_j, \xi - u_j) d\xi}_{=: J_{2,j}}, & \text{if } n_{ij} = 1 \\ - \left(\int_{\xi \leq 0} M(h_i, \xi - u_i) d\xi + \int_{\xi > 0} M(h_j, \xi - u_j) d\xi \right), & \text{if } n_{ij} = -1 \end{cases} \\ &= \begin{cases} J_{1,i} + J_{2,j}, & \text{if } n_{ij} = 1 \\ - (J_{2,i} + J_{1,j}), & \text{if } n_{ij} = -1. \end{cases} \end{aligned}$$

For the integrals, we get $J_{1,i} = J_1(h_i, u_i)$ and $J_{2,i} = J_2(h_i, u_i)$ with J_1 and J_2 given by

$$J_1(h, u) = \begin{cases} h, & \text{if } a \leq -1 \\ h \left[\frac{1}{2} - \frac{1}{\pi} (a \cos(\arcsin(a)) + \arcsin(a)) \right], & \text{if } -1 < a < -1 \\ 0, & \text{if } a \geq 1, \end{cases} \quad (4.36)$$

$$J_2(h, u) = \begin{cases} 0, & \text{if } a \leq -1 \\ \frac{h}{\pi} \left[a \cos(\arcsin(a)) + \arcsin(a) + \frac{\pi}{2} \right], & \text{if } -1 < a < -1 \\ h, & \text{if } a \geq 1 \end{cases} \quad (4.37)$$

with $a := a(h, u) := -u/\sqrt{2bh}$.

In the same manner, we proceed with the modified flux \tilde{M}_{ij} and get

$$\begin{aligned} \tilde{\mathbf{K}}_{ij} &= \tilde{\mathbf{K}}(\mathbf{U}_i, \mathbf{U}_j, n_{ij}) = \left((\tilde{K}_{ij})^1, (\tilde{K}_{ij})^2 \right)^T := \int_{\mathbb{R}} \begin{pmatrix} 1 \\ \xi \end{pmatrix} \tilde{M}_{ij} d\xi \\ &= \left(\int_{\xi n_{ij} \geq 0} \begin{pmatrix} 1 \\ \xi \end{pmatrix} M(h_i, \xi - u_i) d\xi + \int_{\xi n_{ij} < 0} \begin{pmatrix} 1 \\ \xi \end{pmatrix} M(h_j, \xi - u_j) d\xi \right) n_{ij}. \end{aligned}$$

Again, the second component of $\tilde{\mathbf{K}}_{ij}$ is already known as $\tilde{K}_{ij}^2 = \tilde{F}_{ij}^1$ and we only have to compute \tilde{K}_{ij}^1 :

$$\begin{aligned} \tilde{K}_{ij}^1 &= \int_{\mathbb{R}} M_{ij}(\xi) d\xi \\ &= \begin{cases} \underbrace{\int_{\xi \geq 0} M\left(\frac{h_i + h_j}{2}, \xi - u_i\right) d\xi}_{=: \tilde{J}_{1,i,j}} + \underbrace{\int_{\xi < 0} M\left(\frac{h_i + h_j}{2}, \xi - u_j\right) d\xi}_{=: \tilde{J}_{2,i,j}}, & \text{if } n_{ij} = 1 \\ - \left(\int_{\xi \leq 0} M\left(\frac{h_i + h_j}{2}, \xi - u_i\right) d\xi + \int_{\xi > 0} M\left(\frac{h_i + h_j}{2}, \xi - u_j\right) d\xi \right), & \text{if } n_{ij} = -1 \end{cases} \\ &= \begin{cases} \tilde{J}_{1,i,j} + \tilde{J}_{2,i,j}, & \text{if } n_{ij} = 1 \\ - (\tilde{J}_{2,j,i} + \tilde{J}_{1,j,i}), & \text{if } n_{ij} = -1. \end{cases} \end{aligned}$$

The integrals can be computed as

$$\tilde{J}_{1,i,j} = J_1\left(\frac{h_i + h_j}{2}, u_i\right), \quad \tilde{J}_{2,i,j} = J_2\left(\frac{h_i + h_j}{2}, u_i\right),$$

with J_1 and J_2 defined in (4.36) and (4.37), respectively.

Thus, we end up with the macroscopic numerical flux functions

$$\mathbf{G}_{ij}(\mathbf{U}_i, \mathbf{U}_j) := \mathbf{G}(\mathbf{U}_i, \dot{x}_i, \mathbf{U}_j, \dot{x}_j, n_{ij}) := \mathbf{F}_{ij} - \mathbf{K}_{ij} \cdot \bar{\mathbf{x}}_{ij}$$

and

$$\tilde{G}_{ij}(\mathbf{U}_i, \mathbf{U}_j) := \tilde{G}(\mathbf{U}_i, \dot{x}_i, \mathbf{U}_j, \dot{x}_j, n_{ij}) := \tilde{\mathbf{F}}_{ij} - \tilde{\mathbf{K}}_{ij} \cdot \bar{x}_{ij}$$

where \bar{x}_{ij} is the averaged particle velocity given in (2.30).

It is easy to see that both flux functions are consistent in the sense that

$$\mathbf{G}_{ij}(\mathbf{U}, \mathbf{U}) = \tilde{\mathbf{G}}_{ij}(\mathbf{U}, \mathbf{U}) = \left[\begin{pmatrix} hu \\ hu^2 + \frac{1}{2}bh^2 \end{pmatrix} - \mathbf{U} \cdot \bar{x}_{ij} \right] n_{ij}.$$

Altogether, the fully discretized FVPM for the SH equations and moving particles reads

Scheme 4.3.3.

$$\left. \begin{aligned} V_i^{n+1} h_i^{n+1} &= V_i^n h_i^n - \Delta t \sum_{j \in \mathcal{N}(i)} |\beta_{ij}^n| \left(\tilde{G}_{ij}^n \right)^1 \\ V_i^{n+1} q_i^{n+1} &= -\Delta t \sum_{j \in \mathcal{N}(i)} |\beta_{ij}^n| \left(\tilde{G}_{ij}^n \right)^2 + (1 - \mu) \Delta t \tilde{g}_i^n \end{aligned} \right\} \text{ if } |\mu| < 1$$

$$\left. \begin{aligned} V_i^{n+1} h_i^{n+1} &= V_i^n h_i^n - \Delta t \sum_{j \in \mathcal{N}(i)} |\beta_{ij}^n| \left(G_{ij}^n \right)^1 \\ V_i^{n+1} q_i^{n+1} &= V_i^n q_i^n - \Delta t \sum_{j \in \mathcal{N}(i)} |\beta_{ij}^n| \left(G_{ij}^n \right)^2 + \Delta t g(q_i^n) V_i^{n+1} h_i^{n+1} \end{aligned} \right\} \text{ else}$$

with

$$\tilde{g}_i^n = \min \left((\sin \zeta + R) V_i^n h_i^n, \max \left(\frac{b}{2} \sum_{j \in \mathcal{N}(i)} \beta_{ij}^n \left(\frac{h_i^n + h_j^n}{2} \right)^2, (\sin \zeta - R) V_i^n h_i^n \right) \right),$$

$$\mu = \begin{cases} -\frac{V_i^n q_i^n}{\Delta t g(q_i^n) V_i^{n+1} h_i^{n+1}}, & \text{if } u_i^n \neq 0 \\ 0, & \text{else} \end{cases}$$

and

$$\mathbf{G}_{ij}^n = ((G_{ij}^n)^1, (G_{ij}^n)^2)^T, \quad \tilde{\mathbf{G}}_{ij}^n = ((\tilde{G}_{ij}^n)^1, (\tilde{G}_{ij}^n)^2)^T.$$

As noted above, the velocity of the particles should be zero for admissible profiles. Under that assumption, it is easy to see that Scheme 4.3.3 preserves the steady states of granular masses at rest because in that case Scheme 4.3.3 coincides with Scheme 4.3.1.

Chapter 5

Numerical results

In this chapter, we will test the FVPM on several examples. We begin with a simple Riemann Problem for the one-dimensional Euler equations in Section 5.1. For this problem, we use fixed particles and give a numerical convergence analysis for different kernels and smoothing lengths, corresponding to a different number of neighbours per particle. The results differ only slightly from the results presented in [Tel05] for a similar problem.

In Section 5.2, we apply the FVPM to a problem with time-varying computational domain, namely the linearized piston problem in one spatial dimension. For this problem, we use moving particles. On the boundary, the particles move exactly with the boundary and in the interior of the computational domain, the velocities of the particles are linearly interpolated. In our numerical computations, we use both the FVPM described in Chapter 2 and the FVPM with B-splines developed in Chapter 3. For the original FVPM with a simple piecewise linear kernel and different boundary conditions, this problem has already been studied in [Lam01].

The SH equations, as introduced in Chapter 4, are considered in Section 5.3. Our numerical results obtained with the FVPM agree very well with the results obtained with the kinetic scheme described in [KS08] and Section 4.2. In Section 5.4, we apply the FVPM to a two-dimensional solid body rotation problem.

Since we are operating on bounded computational domains, we need boundary conditions, i.e. a suitable discretization of the boundary term (2.11)

$$\mathbf{B}_i = \int_{\partial\Omega} \psi_i (\mathbf{F}(\mathbf{u}) - \mathbf{u} \cdot \mathbf{v}) \mathbf{n} \, ds. \quad (5.1)$$

As described in [Tel05], we simply use (2.25e) for inflow or outflow boundary conditions to get

$$\mathbf{B}_i \approx (\mathbf{F}(\mathbf{u}_i^*) - \mathbf{u}_i^* \cdot \mathbf{v}_i^*) \int_{\partial\Omega} \psi_i \mathbf{n} \, ds = (\mathbf{u}_i^* \cdot \mathbf{v}_i^* - \mathbf{F}(\mathbf{u}_i^*)) \sum_{j \in \mathcal{N}(i)} \beta_{ij} \quad (5.2)$$

where \mathbf{u}_i^* and \mathbf{v}_i^* denote approximations of \mathbf{u} and \mathbf{v} on the boundary $\Omega_i \cap \partial\Omega$, respectively. If we assume, that particle i with $\Omega_i \cap \partial\Omega \neq \emptyset$ moves exactly with the same velocity as the

domain, we get

$$\mathbf{B}_i \approx (\mathbf{u}_i^* \cdot \dot{\mathbf{x}}_i - \mathbf{F}(\mathbf{u}_i^*)) \sum_{j \in \mathcal{N}(i)} \beta_{ij}.$$

In the case of inflow boundary conditions, we will set \mathbf{u}_i^* on the prescribed inflow value for \mathbf{u} . For outflow boundary conditions, we will use the reconstruction and set $\mathbf{u}_i^*(\mathbf{x}, t) = \mathbf{u}_h(\mathbf{x}_i, t)$. The treatment of solid wall boundary conditions is also described in [Tel05] and figured out for the two-dimensional Euler equations. As the normal velocity of the flux through the boundary vanishes, we will end up with

$$\mathbf{F}(\mathbf{u}) \mathbf{n} = \begin{pmatrix} 0 \\ p n_1 \\ p n_2 \\ 0 \end{pmatrix}$$

and thus

$$\int_{\partial\Omega(t)} \psi_i \mathbf{F}(\mathbf{u}) \, ds \approx \begin{pmatrix} 0 \\ -p_i \sum_{j \in \mathcal{N}(i)} \beta_{ij} \\ 0 \end{pmatrix}.$$

Alternatively, one may handle inflow or outflow boundary conditions directly, at least for non-moving domains Ω . For inflow boundary conditions, this can be done by prescribing the values for \mathbf{u}_i^n if $\Omega_i \cap \partial\Omega^{in} \neq \emptyset$, where $\partial\Omega^{in}$ denotes the part of the boundary $\partial\Omega$ where inflow conditions are considered. In this case, we set $\mathbf{u}_i^n = \frac{1}{V_i^n} \int_{\Omega} \mathbf{u}(\mathbf{x}, t^n) \psi_i(\mathbf{x}, t^n) \, d\mathbf{x}$ for all $n = 0, \dots, N_T$. Outflow boundary conditions can be implemented by simply extrapolating the values from inner particles. Then, the update (2.22a) is made only for inner particles. In this approach, the boundary particles are also called ghost particles.

In some of the following examples, we compute the experimental order of convergence (EOC) according to

$$\text{EOC}_1 := \frac{\log(|\mathbf{u}^1 - \mathbf{u}_h^1|_1) - \log(|\mathbf{u}^2 - \mathbf{u}_h^2|_1)}{\log(N_2) - \log(N_1)} \quad (5.3)$$

where $|\cdot|_1$ denotes the discrete L^1 -norm, \mathbf{u}_h^i denotes the numerical solution computed with N_i particles and \mathbf{u}^i denotes the exact solution, both evaluated at the particle positions $\mathbf{x}_1, \dots, \mathbf{x}_{N_i}$, $i = 1, 2$. The EOC for the discrete L^2 -norm is constructed analogously.

5.1 The 1D Euler equations

We begin this chapter with some results for a standard Riemann problem for the one-dimensional Euler equations

$$\partial_t \mathbf{u} + \nabla \cdot \mathbf{F}(\mathbf{u}) = \mathbf{0} \quad \text{for } (x, t) \in [0, 1] \times \mathbb{R}_+$$

with

$$\mathbf{u} = \begin{pmatrix} \rho \\ \rho u \\ \rho E \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ u(\rho E + p) \end{pmatrix}.$$

As already mentioned in Section 1.1, the system is completed with the equation of state

$$p = (\gamma - 1) \rho \left(E - \frac{u^2}{2} \right)$$

where $\gamma = 1.4$. The initial data are given by

$$\rho(x, 0) = \begin{cases} 4, & x \leq 0.5 \\ 1, & x > 0.5, \end{cases} \quad u(x, 0) = 0, \quad p(x, 0) = \begin{cases} 1.6, & x \leq 0.5 \\ 0.4, & x > 0.5. \end{cases}$$

The exact solution consists of a rarefaction wave, a contact discontinuity and a shock wave. As a first example, we compute approximate solutions with fixed equidistant particles. As boundary conditions, we use inflow boundary conditions on the left boundary $\mathbf{u}(0, t) = \mathbf{u}(0, 0) = (\rho(x, 0), u(x, 0), p(x, 0))^T \forall t \leq T^*$, where T^* denotes a sufficiently small time such that the evolving waves do not meet the boundary. On the right boundary, we use outflow boundary conditions. The boundary conditions are implemented using formula (5.2).

In the first example, we choose the smoothing length to be $h = 5/6 \Delta x$ meaning that every inner particle has exactly two neighbours. Note that this implies that the geometrical coefficients always satisfy $\beta_{ij} \in \{-1, 1\}$, see Chapter 2. In Figure 5.1, we can see the convergence for $N = 100, 300$ and 500 particles. The exact and the approximate solutions are evaluated at time $t = 0.3$.

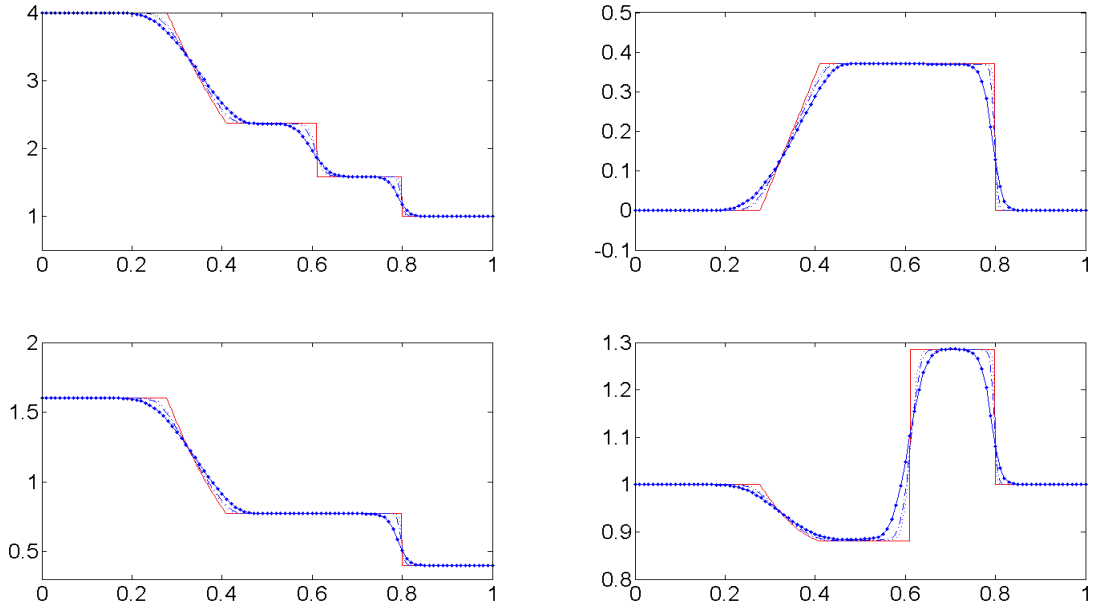


Figure 5.1: Euler equations: density (upper left), velocity (upper right), pressure (lower left) and energy (lower right) computed with equidistant particles and a smoothing length $h = 5/6 \Delta x$ at time $t = 0.3$.

As a second example, we use irregular distributed particles. For that purpose, we take equidistant particles and disturb the particle positions x_i by adding a small random number. The smoothing length is $h = 2\Delta x$. For equidistant particles, this would result in six neighbours for each inner particle. We compute the geometrical coefficients using a precise numerical integration method. Again, the particles do not move, so we have to compute the coefficients only once at the beginning of the computation. For the computation, we use again $N = 100, 300$ and 500 particles and evaluate the exact and the approximate solutions at time $t = 0.3$. In Figure 5.2, we can clearly see that the solution is smeared out in nonsmooth regions compared to the example above.

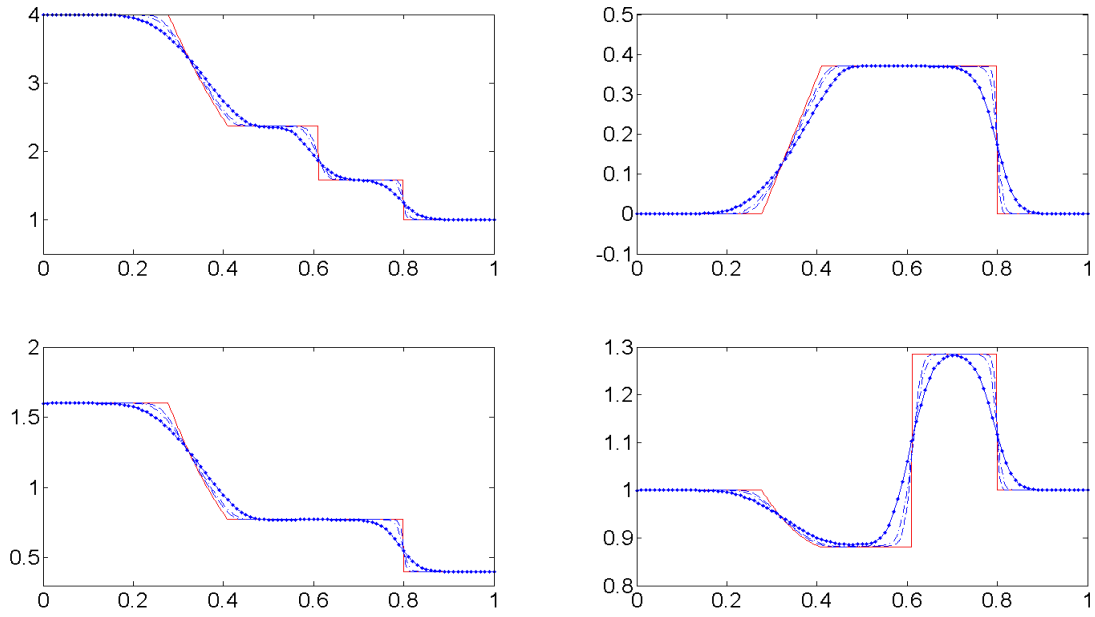


Figure 5.2: Euler equations: density (upper left), velocity (upper right), pressure (lower left) and energy (lower right) computed with irregular particles at time $t = 0.3$.

At the end of this section, we compute the numerical convergence rates in the discrete L^1 - and L^2 -norm for the density ρ , the velocity u , the pressure p and the energy e at time $t = 0.3$. We use equidistant particles and a smoothing length $h = 2\Delta x$, meaning that each particle being sufficiently in the interior of the computational domain has exactly six neighbours. In this case, the geometrical coefficients β_{ij} can be computed exactly as

$$\beta_{ij} = \begin{cases} \pm \frac{7}{16}, & j = i \pm 1 \\ \pm \frac{1}{4}, & j = i \pm 2 \\ \pm \frac{1}{48}, & j = i \pm 3 \\ 0, & \text{else,} \end{cases}$$

see [Tel05] for details. The results for $N = 25, 50, 100, 200, 400$ and 800 particles can be

seen in Table 5.1. where we used Formula (5.3) for the computation of EOC_{L^1} and EOC_{L^2} . The observed convergence rate for the density in L^1 resembles that in [Tel05] for a similar problem. The worse convergence rates for the density and the energy may be caused by the additional discontinuities occurring in these solutions compared with velocity and pressure. In [Krö97] and for the same problem in two space dimensions, the author gets similar L^1 convergence rates of $EOC_{L^1} \approx 0.62$ and $EOC_{L^1} \approx 0.59$ on different grids for the density. As a numerical scheme, a first order FVM with a numerical flux function of Steger and Warming has been used. For more details see [Krö97].

N	$ \rho - \rho_h _1$	EOC_{L^1}	$ \rho - \rho_h _2$	EOC_{L^2}
25	$1.836237e - 01$		$2.291921e - 01$	
50	$1.304578e - 01$	0.4932	$1.752139e - 01$	0.3874
100	$8.966017e - 02$	0.5410	$1.359276e - 01$	0.3663
200	$5.910732e - 02$	0.6011	$1.029453e - 01$	0.4010
400	$3.811848e - 02$	0.6328	$7.784011e - 02$	0.4033
800	$2.435570e - 02$	0.6462	$5.984039e - 02$	0.3794
N	$ u - u_h _1$	EOC_{L^1}	$ u - u_h _2$	EOC_{L^2}
25	$5.317025e - 02$		$6.919373e - 02$	
50	$3.565628e - 02$	0.5765	$5.608921e - 02$	0.3029
100	$2.222067e - 02$	0.6823	$4.232139e - 02$	0.4063
200	$1.345759e - 02$	0.7235	$3.130398e - 02$	0.4350
400	$7.899747e - 03$	0.7685	$2.263215e - 02$	0.4680
800	$4.543966e - 03$	0.7979	$1.627234e - 02$	0.4760
N	$ p - p_h _1$	EOC_{L^1}	$ p - p_h _2$	EOC_{L^2}
25	$8.427131e - 02$		$1.079944e - 01$	
50	$5.601806e - 02$	0.5891	$8.278192e - 02$	0.3836
100	$3.545578e - 02$	0.6599	$6.079546e - 02$	0.4454
200	$2.196382e - 02$	0.6909	$4.371940e - 02$	0.4757
400	$1.326389e - 02$	0.7276	$3.066004e - 02$	0.5119
800	$7.850071e - 03$	0.7567	$2.114738e - 02$	0.5359
N	$ e - e_h _1$	EOC_{L^1}	$ e - e_h _2$	EOC_{L^2}
25	$5.669022e - 02$		$7.747364e - 02$	
50	$3.805264e - 02$	0.5751	$5.936258e - 02$	0.3842
100	$2.522890e - 02$	0.5929	$4.753148e - 02$	0.3207
200	$1.676008e - 02$	0.5900	$3.903484e - 02$	0.2841
400	$1.100486e - 02$	0.6069	$3.210244e - 02$	0.2821
800	$7.134643e - 03$	0.6252	$2.605069e - 02$	0.3014

Table 5.1: Error and convergence rates for the Euler equations with equidistant particles and $h = 2 \Delta x$ at time $t = 0.3$.

5.2 A linearized piston problem with moving boundary

In this section, we consider a linear one-dimensional system of conservation laws that can be used to model the behaviour of a gas in a tube that is compressed by a piston moving back and forth. The linear system is derived from the isentropic Euler equations of gas dynamics. As a detailed description of this example can be found in [Lam01], we only sketch the main steps in the governing of the equations. Also in [Lam01], the FVPM was successfully applied to this problem, but only for a piecewise linear kernel and with different boundary conditions. We use this example to show that the FVPM with B-splines, as described in Chapter 3, works well. For this purpose, we use B-splines of different order and give a numerical convergence analysis. Moreover, the computation times of the different methods are compared. The algorithms were implemented in MATLAB and all computations were done on a 2.3 GHz Intel Core i5 machine.

Let us consider a one-dimensional tube of length L given by the interval $I := [0, L]$ and filled with a compressible gas. Assume that the right boundary is fixed and the left one is moving back and forth with a given velocity. Moreover, we consider the compressible and isentropic Euler equations

$$\partial_t \begin{pmatrix} \rho \\ \rho v \end{pmatrix} + \partial_x \begin{pmatrix} \rho v \\ \rho v^2 + p \end{pmatrix} = 0, \quad (5.4)$$

where p denotes the pressure and v the velocity of the gas. The system is endowed with the additional state equation $p = c \rho^\gamma$. In the following, we consider small perturbations around the equilibrium state

$$\begin{pmatrix} p_0 \\ v_0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

If we now linearize system (5.4) around this equilibrium state, we end up with two equations for the linearized pressure \tilde{p} and the linearized velocity \tilde{v} of the gas:

$$\partial_t \tilde{\mathbf{u}} + \tilde{\mathbf{A}} \partial_x \tilde{\mathbf{u}} = 0,$$

where

$$\tilde{\mathbf{u}} = \begin{pmatrix} \tilde{p} \\ \tilde{v} \end{pmatrix}$$

and

$$\tilde{\mathbf{A}} = \begin{pmatrix} 0 & \gamma p_0 \\ \frac{p_0^{-1/\gamma}}{c} & 0 \end{pmatrix}.$$

In order to get dimensionless variables, we introduce $\bar{t} = \tilde{t}/T$, $\bar{x} = \tilde{x}/L$, $p = \tilde{p}/(\gamma p_0)$ and $v = T \tilde{v}/L$ where T and L denote a characteristic time and the length of the tube, respectively. Then, the dimensionless system reads

$$\partial_{\bar{t}} \mathbf{u} + \mathbf{A} \partial_{\bar{x}} \mathbf{u} = 0 \quad (5.5)$$

with $\mathbf{u} = (p, v)^T$ and

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

System (5.5) is completed with the initial and boundary values

$$\begin{aligned} p(x, 0) &= 1 \\ v(x, 0) &= 0 \\ v(1, t) &= 0 \\ v(a(t), t) &= v_a(t). \end{aligned}$$

Again, the boundary conditions are implemented using Formula (5.2).

The eigenvalues and eigenvectors of system (5.5) are given by

$$\lambda_1 = -1, \lambda_2 = 1, \mathbf{w}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \mathbf{w}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

and the exact solution can be easily computed using the method of characteristics. For a piston moving back and forth with piecewise constant velocity

$$v_a(t) = \begin{cases} \tilde{v}_a, & 0 \leq t \leq t_1 \\ -\tilde{v}_a, & t_1 < t \leq 2t_1, \end{cases}$$

$\tilde{v}_a \leq \lambda_2$, the exact solution for $t \leq t_1$ is given by

$$\begin{aligned} p(x, t) &= \begin{cases} 1 + \tilde{v}_a, & \tilde{v}_a t \leq x \leq \lambda_2 t \\ 1, & t < x \leq 1, \end{cases} \\ v(x, t) &= \begin{cases} \tilde{v}_a, & \tilde{v}_a t \leq x \leq \lambda_2 t \\ 0, & t < x \leq 1. \end{cases} \end{aligned}$$

For $t_1 < t \leq 2t_1$, the exact solution reads

$$\begin{aligned} p(x, t) &= \begin{cases} 1 - \tilde{v}_a, & \tilde{v}_a(1 - t) \leq x \leq \lambda_2 t + (\tilde{v}_a - 1)t_1 \\ 1 + \tilde{v}_a, & \lambda_2 t + (\tilde{v}_a - 1)t_1 < x \leq \lambda_2 t \\ 1, & \lambda_2 t < x \leq 1, \end{cases} \\ v(x, t) &= \begin{cases} -\tilde{v}_a, & \tilde{v}_a(1 - t) \leq x \leq \lambda_2 t + (\tilde{v}_a - 1)t_1 \\ \tilde{v}_a, & \lambda_2 t + (\tilde{v}_a - 1)t_1 < x \leq \lambda_2 t \\ 0, & \lambda_2 t < x \leq 1. \end{cases} \end{aligned}$$

In this problem, particles which intersect the boundary are moved exactly with the velocity of the boundary. The velocity of inner particles is computed as an interpolation of the velocities of the two boundaries.

For our computations, we set $v_a = 0.25$. The piston is assumed to compress the gas until $t = 0.5$, such that it will be back in its initial position at $t = 1$.

In the following, we compare the FVPM using piecewise linear shape functions and piecewise quadratic shape functions with the FVPM with B-splines of degree $m = 1$ and $m = 2$ regarding the accuracy, the rate of convergence and the computation time. In all examples, we consider particles which are equidistant initially and then moved according to the movement of the piston.

We first use piecewise linear functions $W_i(x, t)$ and a smoothing length $h = 2/3 \Delta x^0$, such that we have only two neighbours for each inner particle and thus $\beta_{ij} \in \{-1, 1\}$ for all $i, j \in \mathcal{N}(i)$. Here, Δx^0 denotes the distance between the initial particle positions. The smoothing length of the particles and the movement of the piston are sufficiently small such that the configuration of neighbours does not change during computation. Thus, the coefficients β_{ij} stay constant. In contrast to this, the volumes V_i will change in time and have to be computed in every time step. However, as stated before, it is very easy to realize an exact integration in the case of piecewise linear functions $W_i(x, t)$. The errors in discrete L^1 - and L^2 -norms and the corresponding convergence rates for the FVPM with piecewise linear shape functions $W_i(x, t)$ can be seen in Table 5.2.

Next, we compare these results with the errors and convergence rates for the FVPM with B-splines of degree $m = 1$. For that purpose, we implemented both methods (3.8)–(3.11), where we used a forward Euler discretization in time, and (3.26)–(3.29). In Tables 5.3 and 5.4, it can be seen that both methods provide adequate convergence rates compared with the results presented in Table 5.2. The computation times for the three methods are presented in the upper part of Table 5.7. Since the problem can be solved by each of these three methods without much effort, the computation times do not differ significantly. In the case of the FVPM with piecewise linear shape functions and a more complex overlapping of the particles, the computation time would surely increase. But even in that situation, there exists an efficient way to compute the geometrical coefficients [Lam01].

After that, we have a look at the FVPM with piecewise quadratic shape functions $W_i(x, t)$ and smoothing lengths of $h = 1.1 \Delta x^0$, $h = 1.25 \Delta x^0$ and $h = 1.5 \Delta x^0$ resulting in four neighbours for each inner particle initially. In this example, we have to compute the coefficients β_{ij} and V_i in every time step. For that purpose, we use Gaussian numerical integration and Keck's correction method. For the volumes

$$V_i = \int_{x_i-h}^{x_i+h} \psi_i dx = \int_{x_i-h}^{x_i+h} \frac{W_i}{\sigma} dx,$$

we get best results if we divide the domain $[x_i-h, x_i+h]$ into three parts, W_i being polynomial of degree 2 on each part:

$$V_i = \int_{x_i-h}^{x_i-h/2} \frac{W_i}{\sigma} dx + \int_{x_i-h/2}^{x_i+h/2} \frac{W_i}{\sigma} dx + \int_{x_i+h/2}^{x_i+h} \frac{W_i}{\sigma} dx.$$

For each subinterval, we use 3 integration points. For the computation of the coefficients β_{ij} ,

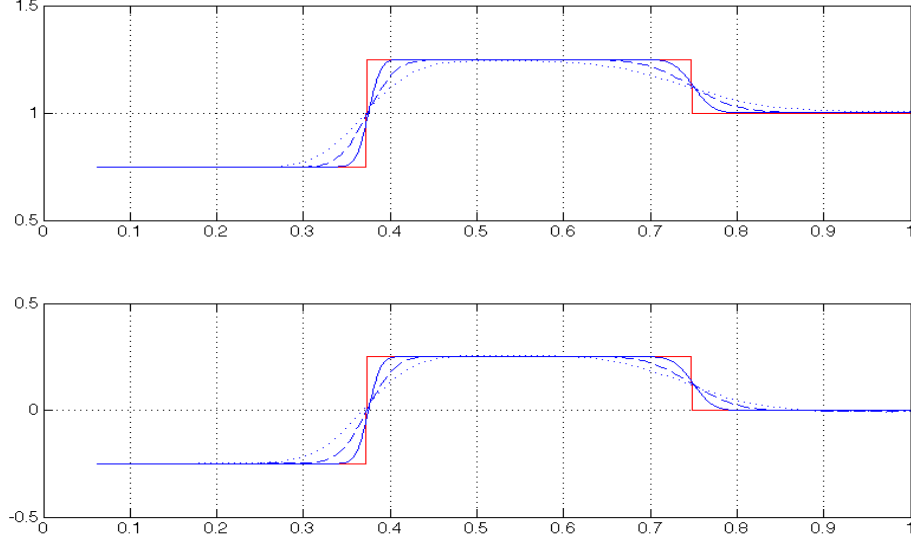


Figure 5.3: The piston problem: pressure and velocity of the gas. Red: exact solution, blue: numerical solution computed with the FVPM with B-splines of degree $m = 2$ and 100 (dotted), 300 (dashed) and 1600 particles.

we split the integral into two integrals:

$$\beta_{ij} = 2 \int_{\max(x_i, x_j) - h}^{(x_i + x_j)/2} \psi_i \partial_x \psi_j \, dx + 2 \int_{(x_i + x_j)/2}^{\min(x_i, x_j) - h} \psi_i \partial_x \psi_j \, dx$$

and use 3 integration points on each domain. The errors and the convergence of the method can be seen in Table 5.5. These results are compared with that from the FVPM with B-splines of degree $m = 2$, whose errors and rates of convergence are shown in Table 5.6. Both methods show similar convergence rates in the discrete L^1 - and L^2 -norm.

The computation times in seconds of both methods can be found in the lower part of Table 5.7. Obviously, the modified FVPM with B-splines is much faster than the original FVPM. This is primarily due to the computation of the coefficients by numerical integration. Note that the computation times and the discrepancies between them may vary significantly in other languages than MATLAB. Therefore, the comparison of absolute computation times makes only limited sense.

N	$ p - p_h _1$	EOC_{L^1}	$ p - p_h _2$	EOC_{L^2}
25	$5.586147e - 02$		$8.613473e - 02$	
50	$3.970762e - 02$	0.4924	$6.976309e - 02$	0.3041
100	$2.818071e - 02$	0.4947	$5.638197e - 02$	0.3072
200	$2.030610e - 02$	0.4728	$4.846928e - 02$	0.2182
400	$1.429303e - 02$	0.5066	$4.006131e - 02$	0.2749
800	$1.013196e - 02$	0.4964	$3.387245e - 02$	0.2421
1600	$7.144707e - 03$	0.5040	$2.835968e - 02$	0.2563
N	$ v - v_h _1$	EOC_{L^1}	$ v - v_h _2$	EOC_{L^2}
25	$5.529669e - 02$		$8.606280e - 02$	
50	$3.945640e - 02$	0.4869	$6.997673e - 02$	0.2985
100	$2.786748e - 02$	0.5017	$5.638907e - 02$	0.3115
200	$2.008645e - 02$	0.4724	$4.849070e - 02$	0.2177
400	$1.415388e - 02$	0.5050	$4.006193e - 02$	0.2755
800	$1.005208e - 02$	0.4937	$3.387435e - 02$	0.2420
1600	$7.100593e - 03$	0.5015	$2.835974e - 02$	0.2563

Table 5.2: Error and convergence rates for the FVPM with piecewise linear shape functions at time $t = 0.75$ and with $h = 2/3 \Delta x^0$.

N	$ p - p_h _1$	EOC_{L^1}	$ p - p_h _2$	EOC_{L^2}
25	$5.385102e - 02$		$7.805576e - 02$	
50	$3.934070e - 02$	0.4530	$6.692482e - 02$	0.2220
100	$2.805524e - 02$	0.4878	$5.643740e - 02$	0.2459
200	$1.998551e - 02$	0.4893	$4.760950e - 02$	0.2454
400	$1.416659e - 02$	0.4965	$4.006953e - 02$	0.2487
800	$1.003384e - 02$	0.4976	$3.370726e - 02$	0.2494
1600	$7.100074e - 03$	0.4990	$2.834803e - 02$	0.2498
N	$ v - v_h _1$	EOC_{L^1}	$ v - v_h _2$	EOC_{L^2}
25	$5.324860e - 02$		$7.791734e - 02$	
50	$3.929769e - 02$	0.4383	$6.692305e - 02$	0.2194
100	$2.805477e - 02$	0.4862	$5.643740e - 02$	0.2459
200	$1.998551e - 02$	0.4893	$4.760950e - 02$	0.2454
400	$1.416659e - 02$	0.4965	$4.006953e - 02$	0.2487
800	$1.003384e - 02$	0.4976	$3.370726e - 02$	0.2494
1600	$7.100074e - 03$	0.4990	$2.834803e - 02$	0.2498

Table 5.3: Error and convergence rates for the FVPM with B-splines of degree $m = 1$ (scheme (3.8)–(3.11)) at time $t = 0.75$.

N	$ p - p_h _1$	EOC_{L^1}	$ p - p_h _2$	EOC_{L^2}
25	$5.378349e - 02$		$7.888601e - 02$	
50	$3.885186e - 02$	0.4692	$6.584225e - 02$	0.2608
100	$2.798613e - 02$	0.4733	$5.646677e - 02$	0.2216
200	$1.991220e - 02$	0.4911	$4.743282e - 02$	0.2515
400	$1.415722e - 02$	0.4921	$4.007760e - 02$	0.2431
800	$1.002499e - 02$	0.4979	$3.368940e - 02$	0.2505
1600	$7.098213e - 03$	0.4981	$2.836257e - 02$	0.2483
N	$ v - v_h _1$	EOC_{L^1}	$ v - v_h _2$	EOC_{L^2}
25	$5.297022e - 02$		$7.873064e - 02$	
50	$3.879630e - 02$	0.4493	$6.584033e - 02$	0.2580
100	$2.798556e - 02$	0.4712	$5.646677e - 02$	0.2216
200	$1.991220e - 02$	0.4910	$4.743282e - 02$	0.2515
400	$1.415722e - 02$	0.4921	$4.007760e - 02$	0.2431
800	$1.002499e - 02$	0.4979	$3.368940e - 02$	0.2505
1600	$7.098213e - 03$	0.4981	$2.836257e - 02$	0.2483

Table 5.4: Error and convergence rates for the FVPM with B-splines of degree $m = 1$ (scheme (3.26)–(3.29)) at time $t = 0.75$.

$h = 1.1\Delta x^0$	N	$ p - p_h _1$	EOC_{L^1}	$ p - p_h _2$	EOC_{L^2}
	25	$5.717630e - 02$		$8.783970e - 02$	
	50	$4.055706e - 02$	0.4955	$7.134051e - 02$	0.3002
	100	$2.876460e - 02$	0.4957	$5.763901e - 02$	0.3077
	200	$2.118084e - 02$	0.4415	$4.995715e - 02$	0.2064
	400	$1.538061e - 02$	0.4616	$4.120906e - 02$	0.2777
	800	$1.144423e - 02$	0.4265	$3.529777e - 02$	0.2234
	1600	$8.575832e - 03$	0.4163	$2.977336e - 02$	0.2456
$h = 1.1\Delta x^0$	N	$ v - v_h _1$	EOC_{L^1}	$ v - v_h _2$	EOC_{L^2}
	25	$5.637238e - 02$		$8.771490e - 02$	
	50	$4.043996e - 02$	0.4792	$7.144753e - 02$	0.2959
	100	$2.867055e - 02$	0.4962	$5.763104e - 02$	0.3100
	200	$2.076068e - 02$	0.4657	$4.990703e - 02$	0.2076
	400	$1.465274e - 02$	0.5027	$4.118900e - 02$	0.2770
	800	$1.049790e - 02$	0.4811	$3.525405e - 02$	0.2245
	1600	$7.492939e - 03$	0.4865	$2.974303e - 02$	0.2452

to be continued on next page

$h = 1.25\Delta x^0$	N	$ p - p_h _1$	EOC_{L^1}	$ p - p_h _2$	EOC_{L^2}
	25	$6.067745e - 02$		$9.037867e - 02$	
	50	$4.339368e - 02$	0.4837	$7.350599e - 02$	0.2981
	100	$3.114745e - 02$	0.4784	$5.984847e - 02$	0.2965
	200	$2.267578e - 02$	0.4580	$5.176039e - 02$	0.2095
	400	$1.614304e - 02$	0.4902	$4.287385e - 02$	0.2718
	800	$1.166733e - 02$	0.4684	$3.674861e - 02$	0.2224
	1600	$8.436395e - 03$	0.4678	$3.111323e - 02$	0.2402
$h = 1.25\Delta x^0$	N	$ v - v_h _1$	EOC_{L^1}	$ v - v_h _2$	EOC_{L^2}
	25	$5.991149e - 02$		$9.026237e - 02$	
	50	$4.325466e - 02$	0.4700	$7.373853e - 02$	0.2917
	100	$3.078546e - 02$	0.4906	$5.986074e - 02$	0.3008
	200	$2.228819e - 02$	0.4660	$5.179544e - 02$	0.2088
	400	$1.575288e - 02$	0.5007	$4.287258e - 02$	0.2728
	800	$1.129054e - 02$	0.4805	$3.675292e - 02$	0.2222
	1600	$8.064306e - 03$	0.4855	$3.110896e - 02$	0.2405
$h = 1.5\Delta x^0$	N	$ p - p_h _1$	EOC_{L^1}	$ p - p_h _2$	EOC_{L^2}
	25	$6.957848e - 02$		$9.633679e - 02$	
	50	$4.955719e - 02$	0.4895	$7.821947e - 02$	0.3006
	100	$3.543613e - 02$	0.4839	$6.395492e - 02$	0.2905
	200	$2.596215e - 02$	0.4488	$5.503163e - 02$	0.2168
	400	$1.893909e - 02$	0.4550	$4.566847e - 02$	0.2691
	800	$1.412734e - 02$	0.4229	$3.895068e - 02$	0.2296
	1600	$1.065109e - 02$	0.4075	$3.287157e - 02$	0.2448
$h = 1.5\Delta x^0$	N	$ v - v_h _1$	EOC_{L^1}	$ v - v_h _2$	EOC_{L^2}
	25	$6.797417e - 02$		$9.578145e - 02$	
	50	$4.920334e - 02$	0.4662	$7.819990e - 02$	0.2926
	100	$3.517455e - 02$	0.4842	$6.391164e - 02$	0.2911
	200	$2.541279e - 02$	0.4690	$5.494849e - 02$	0.2180
	400	$1.801065e - 02$	0.4967	$4.563017e - 02$	0.2681
	800	$1.291017e - 02$	0.4803	$3.888240e - 02$	0.2309
	1600	$9.247077e - 03$	0.4814	$3.281835e - 02$	0.2446

Table 5.5: Error and convergence rates for the FVPM with piecewise quadratic shape functions at time $t = 0.75$ and with smoothing lengths $h = 1.1 \Delta x^0$, $h = 1.25 \Delta x^0$ and $h = 1.5 \Delta x^0$. The coefficients are computed using Gaussian numerical integration.

N	$ p - p_h _1$	EOC_{L^1}	$ p - p_h _2$	EOC_{L^2}
25	$7.802649e - 02$		$9.764169e - 02$	
50	$5.297692e - 02$	0.5586	$7.676638e - 02$	0.3470
100	$3.598215e - 02$	0.5581	$6.214926e - 02$	0.3047
200	$2.477685e - 02$	0.5383	$5.147634e - 02$	0.2718
400	$1.715379e - 02$	0.5305	$4.297345e - 02$	0.2605
800	$1.193744e - 02$	0.5230	$3.601821e - 02$	0.2547
1600	$8.334053e - 03$	0.5184	$3.024013e - 02$	0.2523
N	$ v - v_h _1$	EOC_{L^1}	$ v - v_h _2$	EOC_{L^2}
25	$5.542434e - 02$		$8.194314e - 02$	
50	$4.204249e - 02$	0.3987	$6.999622e - 02$	0.2273
100	$3.165262e - 02$	0.4095	$5.959725e - 02$	0.2320
200	$2.312872e - 02$	0.4526	$5.050249e - 02$	0.2389
400	$1.652668e - 02$	0.4849	$4.258699e - 02$	0.2459
800	$1.169947e - 02$	0.4984	$3.587279e - 02$	0.2475
1600	$8.243379e - 03$	0.5051	$3.017864e - 02$	0.2494

Table 5.6: Error and convergence rates for the FVPM with B-splines of degree $m = 2$ at time $t = 0.75$.

N	B-splines, $m = 1$ (3.8)–(3.11)	B-splines, $m = 1$ (3.26)–(3.29)	piecew. lin. shape function
25	0.0861	0.0851	0.0862
50	0.1864	0.1795	0.1911
100	0.7048	0.6936	0.7147
200	2.7572	2.8064	2.8248
400	10.9403	11.0227	12.3276
N	B-Splines, $m = 2$	piecew. quadr. shape function, $h = 1.1\Delta x$	factor
25	0.2333	2.9697	≈ 12.7
50	0.4062	11.8122	≈ 29.1
100	1.5611	49.6814	≈ 31.8
200	6.2568	223.4927	≈ 35.7
400	26.2550	2066.5929	≈ 78.7

Table 5.7: Computation times in seconds for solving the piston problem up to time $t = 0.75$.

5.3 The SH equations

In this section, we consider the SH equations in one spatial dimension with Riemann initial data in a first example and parabolic initial data in a second one. For all examples, we use a piecewise quadratic kernel and the time step size is $\Delta t = \Delta x/3$.

We start with a classical Riemann problem for the SH equations on the computational domain $\Omega = [0, 1]$ without a source, i.e. $\zeta = \delta = 0$, and the initial data

$$\begin{aligned} h(x, 0) &= \begin{cases} h_l, & \text{if } x \leq x_0 \\ h_r, & \text{if } x > x_0 \end{cases} \\ u(x, 0) &= 0, \end{aligned}$$

where $h_l = 1$, $h_r = 0.5$ and $x_0 = 0.5$. The exact solution for this problem is given by

$$\begin{aligned} h(x, t) &= \begin{cases} h_l, & \text{if } x \leq \lambda_l^1 t + x_0 \\ \frac{1}{9b} \left[\left(u_l + 2\sqrt{b h_l} \right) - \left(\frac{x - x_0}{t} \right) \right]^2, & \text{if } \lambda_l^1 t + x_0 < x \leq \lambda_m^1 t + x_0 \\ h_m, & \text{if } \lambda_m^1 t + x_0 < x \leq s t + x_0 \\ h_r, & \text{if } x > s t + x_0, \end{cases} \\ u(x, t) &= \begin{cases} u_l, & \text{if } x \leq \lambda_l^1 t + x_0 \\ \frac{2}{3} \left[\left(\frac{x - x_0}{t} \right) + \sqrt{b h_l} - \frac{1}{2} u_l \right], & \text{if } \lambda_l^1 t + x_0 < x \leq \lambda_m^1 t + x_0 \\ u_m, & \text{if } \lambda_m^1 t + x_0 < x \leq s t + x_0 \\ u_r, & \text{if } x > s t + x_0 \end{cases} \end{aligned}$$

and λ_l^1 , λ_m^1 and λ_r^1 are defined as

$$\lambda_l^1 = \lambda^1(h_l, u_l), \quad \lambda_m^1 = \lambda^1(h_m, u_m), \quad \lambda_r^1 = \lambda^1(h_r, u_r),$$

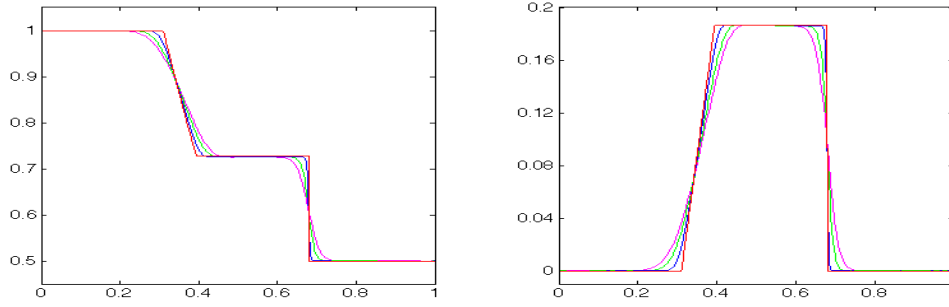
where $\lambda^1(h, u) = u - \sqrt{b h}$ denotes the first eigenvalue of the SH equations and $s = (h_m u_m - h_r u_r)/(h_m - h_r)$ denotes the shock speed. The value h_m is given by

$$(h_r - h_m) \sqrt{\frac{b}{2} \left(\frac{1}{h_m} + \frac{1}{h_r} \right)} + 2 \left(\sqrt{b h_l} - \sqrt{b h_m} \right) = u_r - u_l.$$

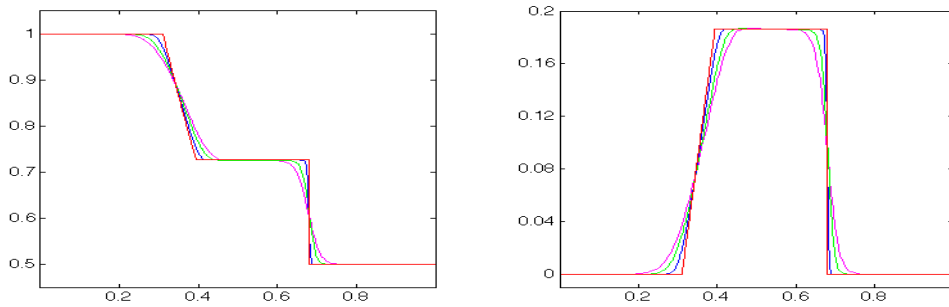
For $t = 0.3$, the results computed with Scheme 4.3.1 and a smoothing length $h = 2 \Delta x$ are compared with the exact solution in Figure 5.4. We used both equidistantly distributed as well as randomly disturbed particle positions. For the latter, we first constructed equidistant particles and disturbed the particle positions x_i by adding a random number $r \in (-0.25 h, 0.25 h)$ afterwards. No significant difference can be seen between the results in Figure 5.4. But we have to remark that the solution for the randomly disturbed particle positions sometimes becomes false, in case of $r \in (-0.5 h, 0.5 h)$ for example. For the equidistantly distributed particles, the smoothing length $h = 2 \Delta x$ results in 6 neighbours for each inner particle. For the randomly disturbed particles, the minimal and maximal number of neighbours for inner

particles is 4 and 8, respectively.

The EOCs for the classical Riemann problem for the SH equations can be found in Table 5.8. We used equidistant particles with smoothing lengths $h = 1.25 \Delta x$ and $h = 2 \Delta x$ resulting in 4 and 6 neighbours for inner particles, respectively. Moreover, the EOCs for the randomly disturbed particles and $h = 2 \Delta x$ are shown. Although the error for disturbed particles is slightly larger than for equidistant particles, the convergence rates seem to be a little better.



(a) Equidistant particle positions.



(b) Randomly disturbed particle positions.

Figure 5.4: The SH equations: height (left) and velocity (right) of the granular mass. Red: exact solution, magenta, green and blue: computational solution with $N = 200, 400$ and 1600 particles, a smoothing length $h = 2 \Delta x$ and equidistant particles (a) as well as randomly disturbed particle positions (b) .

$h = 1.25\Delta x$	N	$ h - h_h _1$	EOC_{L^1}	$ h - h_h _2$	EOC_{L^2}
	25	$2.688724e - 02$		$4.331827e - 02$	
	50	$1.775274e - 02$	0.5989	$3.179321e - 02$	0.4463
	100	$1.154324e - 02$	0.6210	$2.421813e - 02$	0.3926
	200	$7.241052e - 03$	0.6728	$1.798139e - 02$	0.4296
	400	$4.342156e - 03$	0.7378	$1.260675e - 02$	0.5123
	800	$2.577831e - 03$	0.7523	$9.064102e - 03$	0.4760
	1600	$1.524313e - 03$	0.7580	$6.553677e - 03$	0.4679
$h = 1.25\Delta x$	N	$ u - u_h _1$	EOC_{L^1}	$ u - u_h _2$	EOC_{L^2}
	25	$2.009522e - 02$		$3.231550e - 02$	
	50	$1.292351e - 02$	0.6369	$2.370580e - 02$	0.4470
	100	$8.364980e - 03$	0.6276	$1.822500e - 02$	0.3793
	200	$5.224095e - 03$	0.6792	$1.358851e - 02$	0.4235
	400	$3.148072e - 03$	0.7307	$9.750044e - 03$	0.4789
	800	$1.855980e - 03$	0.7623	$6.897379e - 03$	0.4994
	1600	$1.094709e - 03$	0.7616	$4.972079e - 03$	0.4722
$h = 2\Delta x$	N	$ h - h_h _1$	EOC_{L^1}	$ h - h_h _2$	EOC_{L^2}
	25	$3.536548e - 02$		$5.056982e - 02$	
	50	$2.406173e - 02$	0.5556	$3.814586e - 02$	0.4068
	100	$1.587797e - 02$	0.5997	$2.936683e - 02$	0.3773
	200	$1.017705e - 02$	0.6417	$2.208192e - 02$	0.4113
	400	$6.299573e - 03$	0.6920	$1.618504e - 02$	0.4482
	800	$3.755617e - 03$	0.7462	$1.148840e - 02$	0.4945
	1600	$2.184898e - 03$	0.7815	$8.008097e - 03$	0.5206
$h = 2\Delta x$	N	$ u - u_h _1$	EOC_{L^1}	$ u - u_h _2$	EOC_{L^2}
	25	$2.731579e - 02$		$3.822565e - 02$	
	50	$1.771638e - 02$	0.6247	$2.850676e - 02$	0.4232
	100	$1.159073e - 02$	0.6121	$2.217433e - 02$	0.3624
	200	$7.390988e - 03$	0.6491	$1.679123e - 02$	0.4012
	400	$4.586912e - 03$	0.6882	$1.259393e - 02$	0.4150
	800	$2.720151e - 03$	0.7538	$8.904937e - 03$	0.5001
	1600	$1.581580e - 03$	0.7823	$6.263809e - 03$	0.5076
$h = 2\Delta x$, randomly disturbed	N	$ h - h_h _1$	EOC_{L^1}	$ h - h_h _2$	EOC_{L^2}
	25	$3.917843e - 02$		$5.321011e - 02$	
	50	$2.490328e - 02$	0.6537	$3.946125e - 02$	0.4313
	100	$1.651274e - 02$	0.5928	$3.081731e - 02$	0.3567
	200	$1.058071e - 02$	0.6421	$2.249888e - 02$	0.4539
	400	$6.536702e - 03$	0.6948	$1.663628e - 02$	0.4355
	800	$3.888203e - 03$	0.7495	$1.179245e - 02$	0.4965
	1600	$2.246328e - 03$	0.7915	$8.047857e - 03$	0.5512

to be continued on next page

	N	$ u - u_h _1$	EOC_{L^1}	$ u - u_h _2$	EOC_{L^2}
$h = 2\Delta x$, randomly disturbed	25	$3.035758e - 02$		$4.201710e - 02$	
	50	$1.833132e - 02$	0.7277	$2.929060e - 02$	0.5205
	100	$1.206759e - 02$	0.6032	$2.333436e - 02$	0.3280
	200	$7.759506e - 03$	0.6371	$1.744616e - 02$	0.4195
	400	$4.766688e - 03$	0.7030	$1.301106e - 02$	0.4232
	800	$2.815193e - 03$	0.7598	$9.139008e - 03$	0.5096
	1600	$1.623429e - 03$	0.7942	$6.287387e - 03$	0.5396

Table 5.8: Error and convergence rates for Scheme 4.3.1 at time $t = 0.75$ and with smoothing lengths $h = 1.25 \Delta x$ and $h = 2 \Delta x$. The coefficients are computed using Gaussian numerical integration.

In the second example, which can be seen as a one-dimensional breaking grain silo over an incline, we assume an inclination angle $\zeta = \frac{\pi}{15}$ and a friction angle $\delta = \frac{\pi}{5}$. The computational domain is $\Omega = [0, 1]$ and the initial data are given by

$$\begin{aligned} h(x, 0) &= \begin{cases} h_l, & \text{if } x \in [0, x_1] \\ h_m, & \text{if } x \in (x_1, x_2] \\ h_r, & \text{if } x \in (x_2, 1] \end{cases} \\ u(x, 0) &= 0, \end{aligned}$$

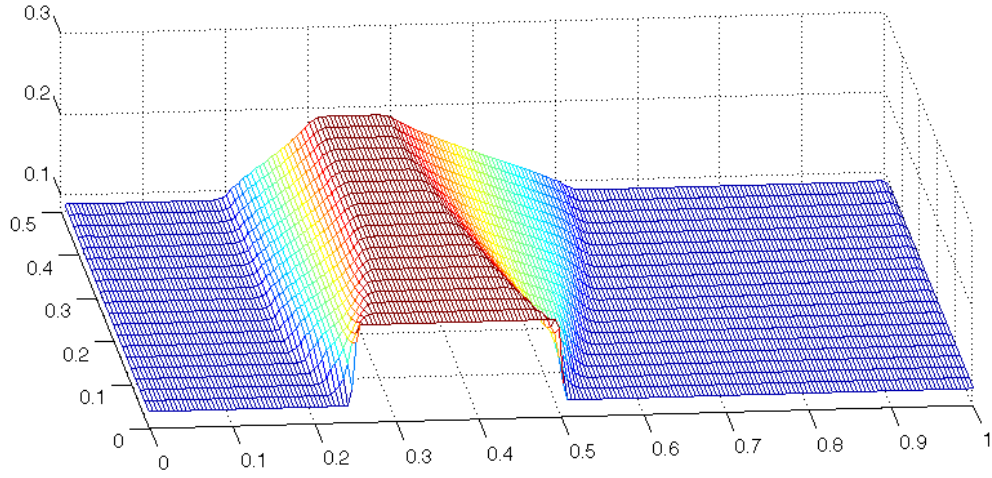
where $h_l = h_r = 0.1$, $h_m = 0.2$, $x_1 = 0.25$ and $x_2 = 0.5$. Because of the inclination, we expect the mass to move unsymmetrically. We first use equidistant particles with a smoothing length $h = 2 \Delta x$ resulting in six neighbours for each inner particle and afterwards randomly disturbed particle positions with the same smoothing length. The construction of the randomly disturbed particle positions is as in the previous example: we start with equidistant particle positions and change them by adding a random number $r \in (-0.25 h, 0.25 h)$.

The time evolution of the initial profile and the velocity for $N = 200$ randomly disturbed particles is shown in Figure 5.5. It can be seen clearly that the mass moves more to the right than to the left and stops after a while. To make the unsymmetric movement of the granular mass more evident, the top view of the time evolution of the velocity is shown in Figure 5.5 (b). The results agree very well with the results for a similar problem treated in [KS08].

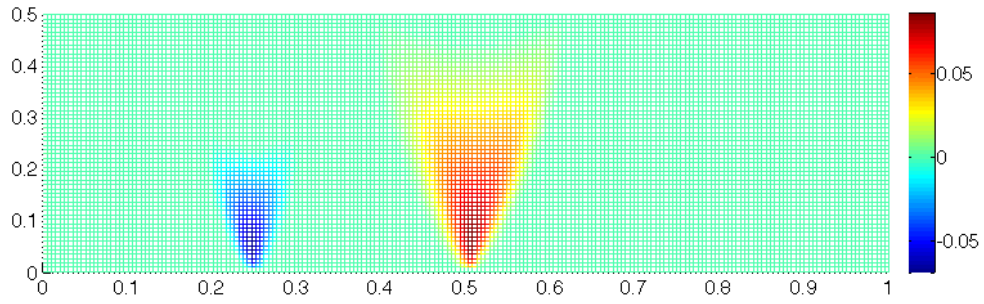
Our last example concerning the SH equations is the spreading of a parabolic initial profile, where we set $\zeta = 0$ and $\delta = \pi/15$. The initial profile is given by

$$h(x, 0) = \max \left(\frac{K}{10}, K - l(x - x_0)^2 \right),$$

where $K = 1$, $l = 4$ and $x_0 = 2.5$. The time evolution of the computational solution for $N = 200$ and randomly disturbed particle positions with a smoothing length $h = 1.25 \Delta x$ is shown in Figure 5.6. We can see that the mass starts to move symmetrically and finally stops



(a) Height for randomly disturbed particle positions.



(b) Velocity for randomly disturbed particle positions.

Figure 5.5: The SH equations: Time evolution of height (a) and velocity (b) for a Riemannian initial profile over an incline with $\zeta = \pi/15$, $\delta = \pi/5$ and $N = 200$ randomly disturbed particle positions up to $t = 0.5$.

due to friction. Again, the results for equidistant and randomly disturbed particle positions differ only slightly and agree very well with the results in [KS08].

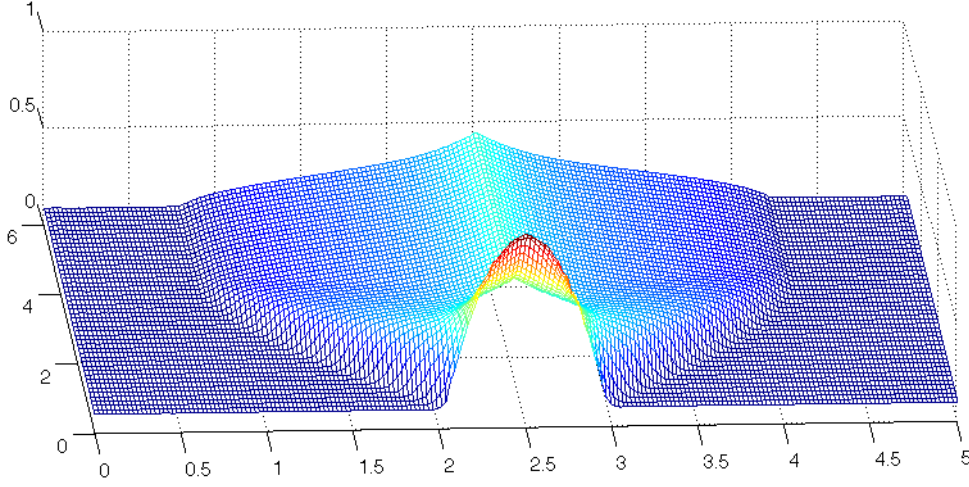


Figure 5.6: The SH equations: Time evolution of the height for a parabolic initial profile with $\zeta = 0$, $\delta = \pi/15$ and $N = 200$ randomly disturbed particle positions up to $t = 6$.

5.4 Solid body rotation

In this section, we consider the so-called solid-body rotation. The corresponding equation and the initial condition in two space dimensions reads

$$\begin{aligned} \partial_t u + 2y \partial_x u - 2x \partial_y u &= 0 \quad \text{for } (x, y, t) \in [-1, 1]^2 \times \mathbb{R}_+, \\ u(x, y, 0) &= u_0(x, y). \end{aligned} \quad (5.6)$$

Note that Equation (5.6) is of the form

$$\partial_t u + \partial_x(v^1(x, y)u) + \partial_y(v^2(x, y)u) = 0,$$

with a divergence-free velocity field $\mathbf{v} = (v^1, v^2) = (2y, -2x)$. The streamlines of the flow are level sets of the function $f(x, y) = x^2 + y^2$ and hence circles. Therefore, the exact solution of problem (5.6) describes the clockwise rotation of the initial condition. With the method of characteristics, it can be easily computed as

$$u(x, y, t) = u_0(x \cos(2t) - y \sin(2t), x \sin(2t) + y \cos(2t)).$$

Thus, a complete rotation of the initial condition will be achieved for $t = k\pi$ and $k = 1, 2, \dots$

For the numerical computation, we use an equidistant particle configuration suggested in [Tel05]. Suppose that we have N^2 particles Ψ_I with $I = (i, j)$, $i, j = 1, \dots, N$, thus N particles in each direction. We use non-moving particles with a piecewise quadratic kernel in two dimensions that is constructed from a piecewise quadratic kernel in one dimension by

$$\mathcal{W}_I(x, y, h) := W_i(x, h) W_j(y, h)$$

and thus $\Psi_I(x, y, h) = \psi_i(x, h) \psi_j(y, h)$. Here, \mathcal{W}_I and Ψ_I denote the kernel and the particle functions in two dimensions and W_i and ψ_i the one-dimensional kernel and particles, respectively. This construction naturally leads to a square-shaped support of the particles Ψ_I . We distribute the particles equidistantly in the domain $\Omega = [-1, 1]^2$ with a smoothing length $h = \Delta x = \Delta y$, where Δx and Δy denote the distance between two particles in x- and y-direction. According to (2.2), we set $W_i(x, h) = W(x - x_i, h)$ with

$$W(x, h) = \begin{cases} (x + h)^2, & x \in [-h, -\frac{h}{2}) \\ -x^2 + \frac{h^2}{2}, & x \in [-\frac{h}{2}, \frac{h}{2}) \\ (x - h)^2, & x \in [\frac{h}{2}, h) \\ 0, & \text{else} \end{cases}$$

and get

$$\psi(x - x_i, h) = \psi_i(x, h) = \begin{cases} 2 \left(1 + \frac{x - x_i}{h}\right)^2, & x \in [x_i - h, x_i - \frac{h}{2}) \\ 1 - 2 \left(\frac{x - x_i}{h}\right)^2, & x \in [x_i - \frac{h}{2}, x_i + \frac{h}{2}) \\ 2 \left(1 - \frac{x - x_i}{h}\right)^2, & x \in [x_i + \frac{h}{2}, x_i + h) \\ 0, & \text{else.} \end{cases}$$

This configuration leads to $V_I = h^2$ and eight neighbours for all inner particles Ψ_I . The geometrical coefficients β_{IJ} can be computed exactly to

$$\beta_{IJ} = \begin{cases} (\pm \frac{7}{60}h, \pm \frac{7}{60}h)^T, & J = (i \pm 1, j \pm 1) \\ (\pm \frac{23}{30}h, 0)^T, & J = (i \pm 1, j) \\ (0, \pm \frac{23}{30}h)^T, & J = (i, j \pm 1) \\ (0, 0)^T, & \text{else,} \end{cases} \quad (5.7)$$

see [Tel05] for details.

The characteristics for this problem are circles and thus dictate an inflow-outflow pattern on the boundary. In the numerical computation for this example, we use ghost particles as described at the beginning of this chapter. On the inflow part of the boundary, we set $u = u_0$.

As rotating objects, we choose first a cone and a rectangular solid as an example of a discontinuous initial condition:

$$u_0(x, y) = \begin{cases} 1, & 0.1 < x < 0.6, -0.25 < y < 0.25 \\ 1 - \frac{r}{0.35}, & r < 0.35 \\ 0, & \text{else} \end{cases} \quad (5.8)$$

with $r := \sqrt{(x + 0.45)^2 + y^2}$.

As a second example, we choose the smooth initial condition

$$u_0(x, y) = \Psi_{(0.35, 0)}(x, y, \tilde{h}) := \psi(x - 0.35, \tilde{h}) \psi(y, \tilde{h}) \quad (5.9)$$

with a smoothing length of $\tilde{h} = 0.5$.

Moreover, we use the Enquist-Osher flux function that is defined in the following way [Kr97]:

$$g(u, v, \mathbf{n}_{ij}) := c_{ij}^+(u) + c_{ij}^-(v),$$

where

$$c_{ij}^+(u) := c_{ij}(0) + \int_0^u \max(c'_{ij}(s), 0) ds, \quad (5.10)$$

$$c_{ij}^-(u) := \int_0^u \min(c'_{ij}(s), 0) ds \quad (5.11)$$

and

$$c_{ij}(u) := \mathbf{f}(u) \cdot \mathbf{n}_{ij}.$$

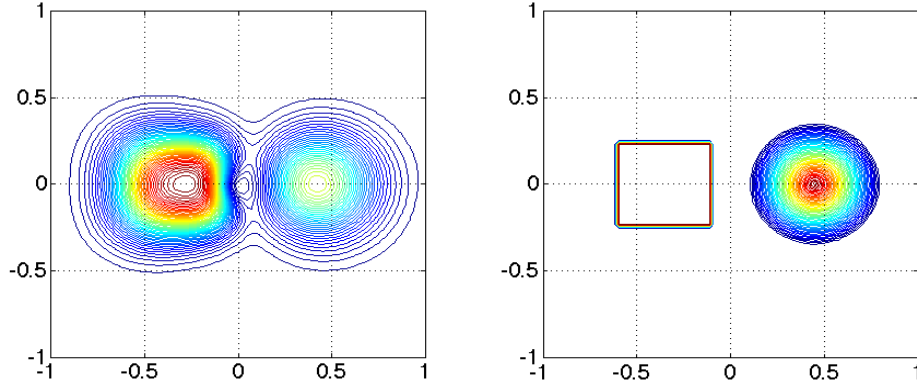
For the solid body rotation problem, we have $\mathbf{f}(u) = 2(yu, -xu)$.

In Figure 5.7, the isolines of the numerical and the exact solution for $N \times N = 100 \times 100$ particles at time $t = \pi/2$ are shown. The initial conditions (5.8) and (5.9) and the corresponding numerical results at time $t = \pi/8$, $t = \pi/4$ and $t = \pi/2$ can be found in Figure 5.8 and 5.9, respectively. In both cases, it can be seen clearly that the solution is smeared out as time proceeds.

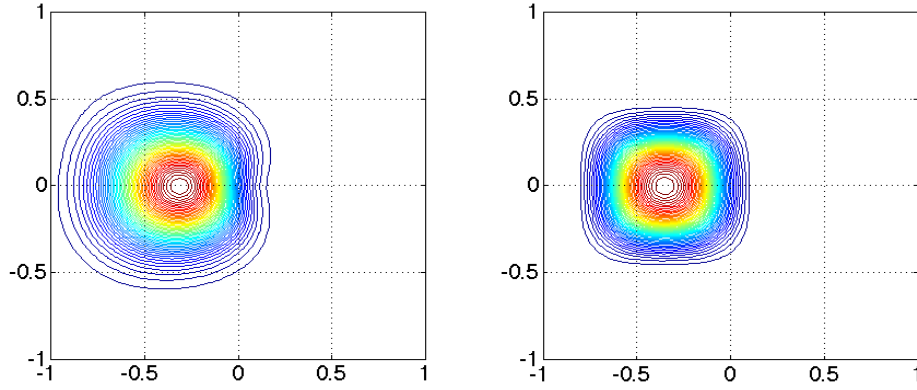
The numerical convergence rates for the numerical solutions in $t = \pi/2$ with the discontinuous initial condition (5.8) and the smooth initial condition (5.9) can be found in Table 5.9 and 5.10, respectively. The results seem to be a little bit better than those in [Tel05] for a much simpler linear advection problem. Note that for smooth initial data, convergence rates of almost one are achieved.

$N \times N$	$ u - u_h _1$	EOC_{L^1}	$ u - u_h _2$	EOC_{L^2}
60×60	$2.913056e - 01$		$2.959874e - 001$	
80×80	$2.559090e - 01$	0.4503	$2.666400e - 01$	0.3630
100×100	$2.256024e - 01$	0.5649	$2.498392e - 01$	0.2917
120×120	$2.067484e - 01$	0.4787	$2.351285e - 01$	0.3328
140×140	$1.886838e - 01$	0.5931	$2.248879e - 01$	0.2889

Table 5.9: Convergence rates for the solid body rotation problem with discontinuous initial data (5.8) and $t = \pi/2$.



(a) discontinuous initial condition



(b) smooth initial condition

Figure 5.7: Solid body rotation: isolines of numerical and exact solution with discontinuous and smooth initial condition at time $t = \frac{\pi}{2}$. Computed with 100×100 particles.

$N \times N$	$ u - u_h _1$	EOC_{L^1}	$ u - u_h _2$	EOC_{L^2}
60×60	$9.893387e - 02$		$1.089388e - 001$	
80×80	$7.894127e - 02$	0.7847	$8.632665e - 02$	0.8087
100×100	$6.542729e - 02$	0.8415	$7.107600e - 02$	0.8711
120×120	$5.567490e - 02$	0.8853	$6.012179e - 02$	0.9180
140×140	$4.831997e - 02$	0.9191	$5.190007e - 02$	0.9540

Table 5.10: Convergence rates for the solid body rotation problem with smooth initial data (5.9) and $t = \pi/2$.

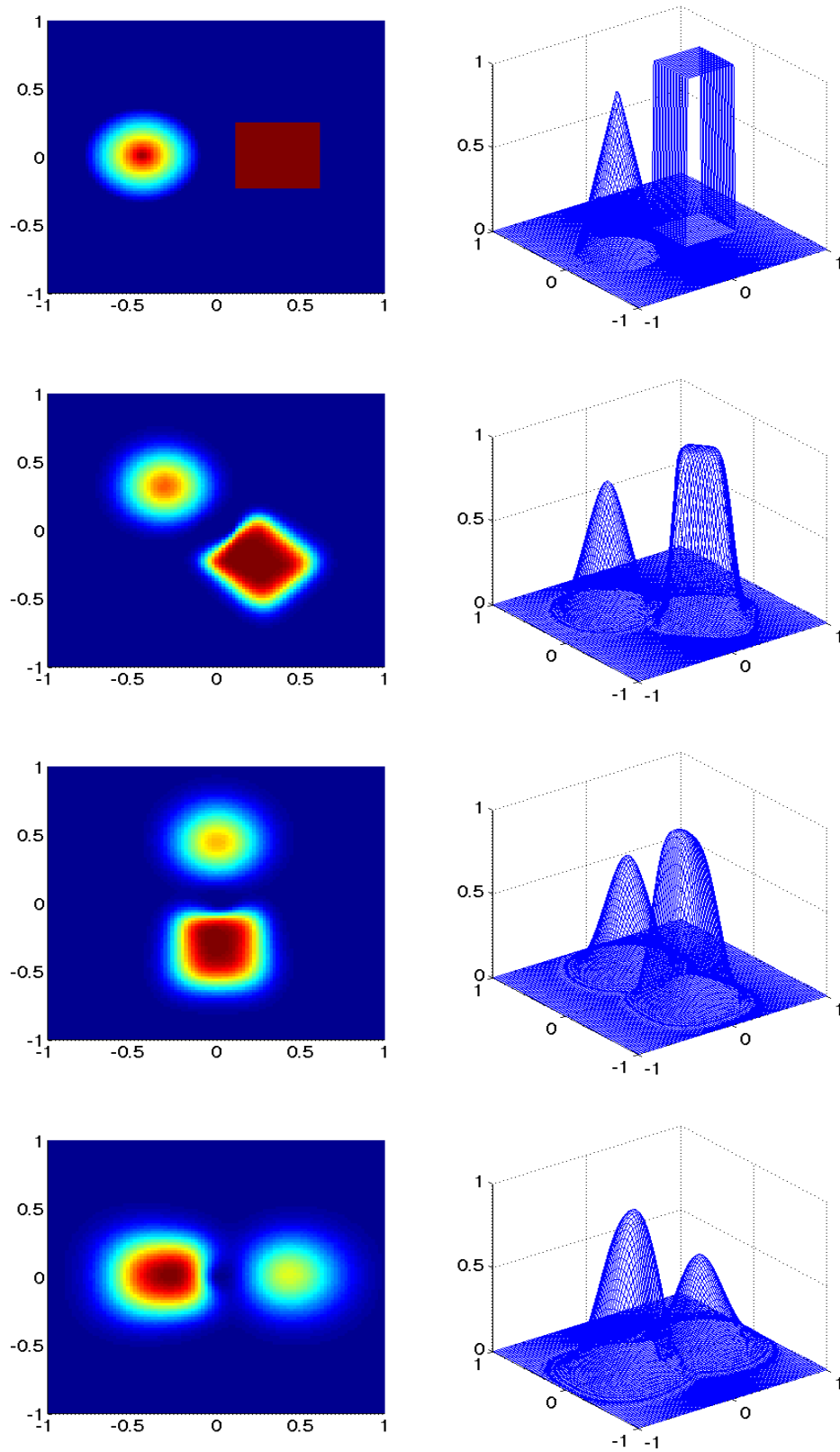


Figure 5.8: Solid body rotation: discontinuous initial condition and numerical solution for $t = \frac{\pi}{8}$, $t = \frac{\pi}{4}$ and $t = \frac{\pi}{2}$. Computed with 100×100 particles.

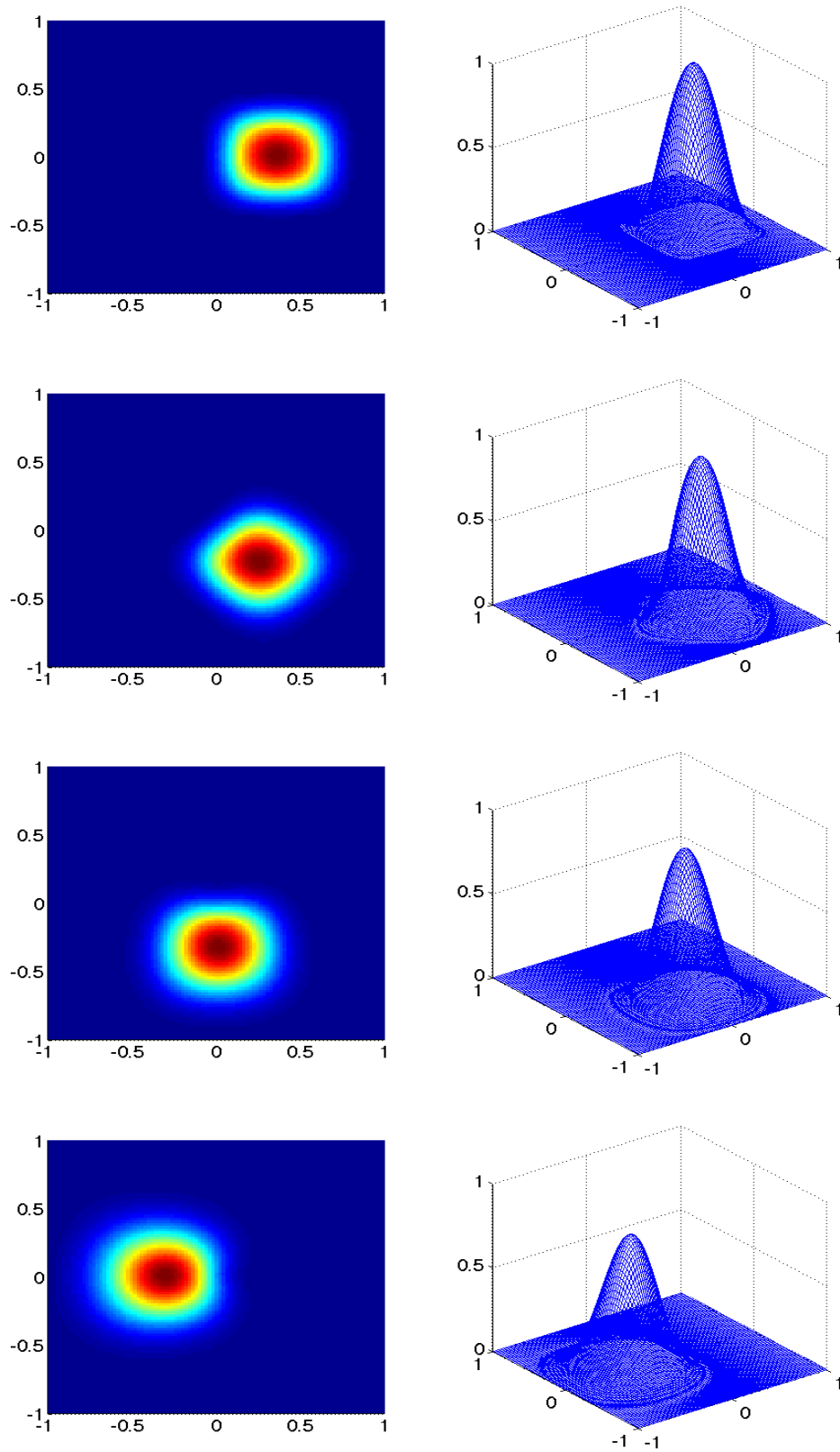


Figure 5.9: Solid body rotation: smooth initial condition and numerical solution for $t = \frac{\pi}{8}$, $t = \frac{\pi}{4}$ and $t = \frac{\pi}{2}$. Computed with 100×100 particles.

Conclusions

This work deals with the Finite Volume Particle Method (FVPM), a meshless method constructed for the numerical treatment of conservation laws. As this method can be interpreted as a generalization of the FVM, a first step in a convergence analysis of the FVPM was made by the proof of some important results that are crucial for the analysis of FVM. For scalar equations in several space dimensions, we could show an improved L^∞ -stability result. As a consequence, we get the convergence of a subsequence in a nonlinear weak-* sense. Under some assumptions, an L^1 - and a weak BV-stability result were proven. Moreover, we were able to prove monotonicity and thus a discrete entropy inequality for general scalar conservation laws in arbitrary space dimensions.

In another part of this thesis, we modified the FVPM in one spatial dimension by using B-splines as shape functions. In this approach, the support of the shape functions may vary in time so that additional terms appear in the scheme. On the other hand, these new terms and the geometrical coefficients can be computed exactly. Thus, the resulting scheme is much more efficient for problems where the particles have to move, e.g. because of a time-dependent computational domain.

After that, we coupled the FVPM with a kinetic Ansatz in order to get a meshfree method for the SH equations. A crucial requirement on numerical methods for the SH equations is the preservation of granular masses at rest, i.e. non-homogenous steady states that make physical sense. For that purpose, we discretized the kinetic representation of the SH equations with the help of the generalized cell averages ending up with a FVPM with kinetic fluxes and a special treatment of the source term. As we have already designed a kinetic FVM for the SH equations earlier, we were able to compare the results of both methods.

In the last part, we tested the FVPM on several numerical examples. First, we implemented the FVPM for the compressible Euler equations in one spatial dimension and for fixed particles. For a classical Riemann problem, we chose particles with different smoothing lengths and compared the EOCs in space with results in [Tel05] for similar test cases. In a second example, we applied the FVPM to a moving boundary problem, i.e. a one-dimensional piston problem. In this case, the particles have to move because of the movement of the boundary. We compared the EOCs and the computation time of the original FVPM with numerically computed geometrical coefficients with that of the modified FVPM with B-splines. Both piecewise linear and piecewise quadratic shape functions with different smoothing lengths were used for the FVPM. For the new scheme, we used first and second order B-splines. It was shown that the EOCs are similar, while the FVPM with B-splines is much faster

and thus considerably more effective. Next, we tested the kinetic FVPM developed for the SH equations on several examples. We started with a classical Riemann problem without source term. The observed EOCs are similar to that observed for the Riemann problem for the Euler equations. The second test case is a Riemann problem over an incline and with simultaneous consideration of friction forces. In this case, we chose the inclination angle to be smaller than the friction angle. It is seen that the granular mass clearly moves more in downward direction and stops after a while. In our last example for the SH equations, we start with a granular mass shaped like a parabolic cap. We take friction into account but no inclination. The mass moves symmetrically and finally stops. All the results for the SH equations agree very well with those from the kinetic scheme in [KS08]. At the end, we implemented the FVPM for a linear solid body rotation problem. We tested the method with smooth and discontinuous initial data. For smooth initial data, we achieved convergence rates of almost one.

The next step concerning the convergence analysis of the FVPM for scalar equations is clearly the construction of an entropy inequality for the reconstructed solution. With the notion of the entropy process solution and the nonlinear weak-* convergence, it is then possible to go to the limit in this entropy inequality and get estimates for the error terms with the help of the weak BV-stability. Thus, it should be possible to show convergence of the numerical solution obtained with the FVPM towards the entropy process solution. As it is well known that the entropy process solution is independent of the additional parameter α and coincides with the unique entropy solution, one gets strong convergence towards the entropy solution in L^p_{loc} according to Proposition 1.2.4.

Concerning the SH equations, an extension to two or three dimensions would be interesting as well as the implementation of moving particles. As figured out in Chapter 4, in such a particle movement the particles should stop when the granular mass tends towards an equilibrium. It would be interesting to see if the kinetic FVPM is able to preserve the granular masses at rest, although this requirement should not be too difficult to achieve.

Appendix

On partitions of unity

The first proposition provides a bound on the sum of all Lebesgue measures of particle patches contained in a set X with Lebesgue measure V . With the help of this proposition, we will show that under the assumptions (2.41), (2.42) and (2.43), each particle has a finite number of neighbours.

Proposition 5.4.1. *Let $X \subseteq \mathbb{R}^d$ be a measurable set with Lebesgue measure $m(X) =: V$. Moreover, let $\mathfrak{X} := \{X_i\}_{i \in I}$, where the sets X_i are measurable sets with equal Lebesgue measure $m(X_i) =: v$ and satisfy*

$$\bigcup_{i \in I} X_i \subseteq X$$

such that no point $\mathbf{x} \in X$ is contained in more than M elements of \mathfrak{X} , i.e.

$$\forall \mathbf{x} \in X : \quad |\{i \in I : \mathbf{x} \in X_i\}| \leq M. \quad (5.12)$$

Then

$$|I| \cdot v \leq M \cdot V.$$

Proof. We can assume that $X = \bigcup_{i \in I} X_i$. For $j = 1, \dots, M$, let $Y_j \subseteq X$ be the set of points in X that are contained in precisely j elements of \mathfrak{X} . Since these sets are clearly disjoint, we have

$$\sum_{j=1}^M m(Y_j) = V.$$

If the sets X_i are disjoint, we have $M = 1$ and

$$|I| \cdot v = V.$$

In the general case, we count points in Y_2 twice, points in Y_3 three times etc. Hence, we have

$$V = |I| \cdot v - 1 \cdot m(Y_2) - 2 \cdot m(Y_3) - \dots - (M-1) \cdot m(Y_M).$$

This yields

$$|I| \cdot v = V + \sum_{j=1}^M (j-1) \cdot m(Y_j)$$

$$\begin{aligned}
&\leq V + (M - 1) \sum_{j=1}^M m(Y_j) \\
&= V + (M - 1) \cdot V \\
&= M \cdot V.
\end{aligned}$$

□

With this, we can show that every particle has only a finite number of neighbours.

Proposition 5.4.2. *Let $\Omega \subseteq \mathbb{R}^d$, $\{\psi_i\}_{i \in \mathcal{T}}$ be a partition of unity on Ω and $\Omega_i := \text{supp}(\psi_i)$ for $i \in \mathcal{T}$. Under the assumptions (2.41), (2.42) and (2.43), each particle has only a finite number of neighbours.*

Proof. We denote by $m(\Omega_i)$ the Lebesgue measure of Ω_i . Because of (2.42), we have

$$m(\Omega_i) = \int_{\Omega_i} 1 \, d\mathbf{x} \geq \int_{\mathbb{R}^d} \psi_i \, d\mathbf{x} = V_i \geq \alpha h^d.$$

Moreover, let $\mathbf{y}_i \in \Omega_i$. Because of (2.43), the ball $B(\mathbf{y}_i, 4h)$ centered in \mathbf{y}_i with radius $4h$ contains Ω_i and Ω_j for all $j \in \mathcal{N}(i)$. Using Proposition 5.4.1, we have

$$M \cdot m(B(\mathbf{y}_i, 4h)) \geq \sum_{j \in \mathcal{N}(i) \cup \{i\}} m(\Omega_j) \geq (C_{N,i} + 1) \cdot \alpha h^d$$

where $C_{N,i} := |\mathcal{N}(i)|$. We denote by \mathcal{M}_d the Lebesgue measure of the unit ball in \mathbb{R}^d and obtain

$$C_{N,i} \leq \frac{M \cdot m(B(\mathbf{y}_i, 4h))}{\alpha \cdot h^d} = \frac{4^d \cdot M \cdot \mathcal{M}_d}{\alpha} = \text{const.}$$

□

For completeness, we will show that a finite number of neighbours for each particle implies that every point in the considered domain is covered by a finite number of particles. Thus, the assumptions that each point of a domain is covered by a finite number of particles and that each particle has a finite number of neighbours are equivalent.

Proposition 5.4.3. *We consider a partition of unity $\{\psi_i\}_{i \in \mathcal{T}}$ defined on a domain $\Omega \subseteq \mathbb{R}^d$ and assume that each particle has only a finite number of neighbours:*

$$\exists C_N \in \mathbb{R} : |\mathcal{N}(i)| \leq C_N \quad \forall i \in \mathcal{T}.$$

Then each point $\mathbf{x} \in \mathbb{R}^d$ is covered by at most $M = C_N + 1$ particles, i.e.

$$\forall \mathbf{x} \in \mathbb{R}^d : |\{i \in \mathcal{T} : \mathbf{x} \in \Omega_i\}| \leq C_N + 1.$$

Proof. Assume a particle ψ_i with $C_{N,i} := |\mathcal{N}(i)|$. We consider the set

$$A := \Omega_i \cap \left(\bigcap_{j \in \mathcal{N}(i)} \Omega_j \right),$$

which might be nonempty. In this case, we have

$$\forall \mathbf{x} \in A : |\{k \in \mathcal{T} : \mathbf{x} \in \Omega_k\}| \geq C_{N,i} + 1.$$

On the other hand, we have

$$\forall \mathbf{x} \in A : |\{k \in \mathcal{T} : \mathbf{x} \in \Omega_k\}| \leq C_{N,i} + .$$

Indeed, the assumption that there exists another particle ψ_l with $l \notin \mathcal{N}(i)$ and

$$A \cap \Omega_l \neq \emptyset$$

automatically implies that

$$\Omega_i \cap \Omega_l \neq \emptyset,$$

thus the particles ψ_i and ψ_l are neighbours, in contrast to the assumption. If the set A is empty, the set $\{k \in \mathcal{T} : \mathbf{x} \in \Omega_k\}$ is even smaller for every $\mathbf{x} \in \mathbb{R}^d$.

The Proposition follows by setting $C_N := \max_{i \in \mathcal{T}} C_{N,i}$. \square

With the above propositions, it is easy to show the bounds stated in Chapter 2.

Proposition 5.4.4. *Under the assumptions (2.41), (2.42) and (2.43), the sets*

$$\begin{aligned} T_R &:= \{i \in \mathcal{T} : \Omega_i \subset B(0, R)\} \\ E_R^n &:= \{(i, j) \in \mathcal{T}^2 : i \in T_R \vee j \in T_R, j \in \mathcal{N}(i), u_i^n > u_j^n\} \\ \mathfrak{S} &:= \{\sigma_{ij} : \sigma_{ij} \subset B(0, R) \setminus B(0, R - 2h)\} \end{aligned}$$

are bounded. More precisely, we have $|E_R^n| \leq C_1 h^{-d}$, $|T_R| \leq C_2 h^{-d}$ and $|\mathfrak{S}| \leq C_3 h^{1-d}$.

Proof. By Proposition 5.4.1, we know that

$$|T_R| \leq C_2 h^{-d}$$

with

$$C_2 := \frac{m(B(0, R)) \cdot M}{\alpha}.$$

Because of Proposition 5.4.2, each particle has only a finite number of neighbours C_N . This directly yields

$$|E_R^n| \leq C_N |T_R| = C_1 h^{-d}$$

with

$$C_1 := \frac{m(B(0, R)) \cdot M \cdot C_N}{\alpha}.$$

To estimate \mathfrak{S} , note that we just have to count the particles in the annulus $B(0, R) \setminus B(0, R - 4h)$, multiplied by the maximal number of neighbours per particle C_N . Thus, we have

$$\begin{aligned} |\mathfrak{S}| &\leq |\{i \in \mathcal{T} : \Omega_i \subset B(0, R) \setminus B(0, R - 4h)\}| \cdot C_N \\ &\leq \frac{\mathcal{M}_d \cdot (R^d - (R - 4h)^d)}{\alpha \cdot h^d} \cdot M \cdot C_N \end{aligned}$$

$$\leq C_3 h^{1-d}$$

with

$$C_3 := \frac{\mathcal{M}_d \cdot \sum_{k=1}^d \binom{d}{k} \cdot R^{d-k} \cdot 4^k \cdot M \cdot C_N}{\alpha} .$$

□

Nomenclature

FDM	Finite Difference Method
FEM	Finite Element Method
FVM	Finite Volume Method
FVPM	Finite Volume Particle Method
SPH	Smoothed Particle Hydrodynamics
CFL condition	Condition on the time step size, named after Courant, Friedrichs and Lewy
EOC	Experimental order of convergence
d	Spatial dimension
Ω	Computational domain
$\partial\Omega$	Boundary of the computational domain
\boldsymbol{v}	Velocity of the boundary of the computational domain
\boldsymbol{x}	Spatial coordinate
t	Time variable
\boldsymbol{n}	Outer unit normal
ρ	Density
u, v	x - and y -component of the velocity, resp.
p	Pressure
E	Energy
$\xrightarrow{n.w*}$	Nonlinear weak-* convergence
$m(\cdot)$	Lebesgue measure
\mathcal{M}_d	Lebesgue measure of the unit ball in \mathbb{R}^d
$\delta(\cdot)$	Diameter
$a \perp b, a \top b$	$\min(a, b), \max(a, b)$

L^p	Lebesgue space of p -integrable functions
L^∞	Lebesgue space of essentially bounded functions
L^p_{loc}	Locally p -integrable functions
BV	Space of functions in L^1_{loc} with bounded total variation
$\ \cdot\ _{L^p}$	L^p -norm
$ \cdot _{L^p}$	Discrete L^p -norm
L	Lipschitz constant
\mathcal{T}	Index set of particles
i, j	Indices of particles
N	Number of particles on bounded domains
$\mathcal{N}(i)$	Index set of all neighbours of particle ψ_i
$C_{N,i}$	Number of neighbours of particle $\psi_i : C_{N,i} := \mathcal{N}(i) $
C_N	Upper bound on the number of neighbours per particle: $ \mathcal{N}(i) \leq C_N \ \forall i \in \mathcal{T}$
n	Time step
Δt	Time step size
ψ_i, ψ_i^n	Particle
Ω_i, Ω_i^n	Support of particle ψ_i : $\Omega_i := \text{supp}(\psi_i)$
μ, M	Minimal and maximal number of overlapping particles: $1 \leq \mu \leq \{i \in \mathcal{T} : \mathbf{x} \in \Omega_i(t)\} \leq M \quad \forall (\mathbf{x}, t) \in \mathbb{R}^d \times \mathbb{R}_+$
$\mathbf{x}_i, \mathbf{x}_i^n$	Position of particle ψ_i
$\dot{\mathbf{x}}_i, \dot{\mathbf{x}}_i^n$	Velocity of particle ψ_i
$\bar{\mathbf{x}}_{ij}$	Average particle velocity
\mathbf{b}_i	Barycenter of particle ψ_i
V_i, V_i^n	Generalized volume of particle ψ_i
$\gamma_{ij}, \gamma_{ij}^n, \beta_{ij}, \beta_{ij}^n$	Geometrical coefficients
$\mathbf{B}_i, \tilde{\mathbf{B}}_i, C_i, D_i$	Boundary terms
h	Smoothing length
$\mathbf{g}_{ij}, \mathbf{g}_{ij}^n$	Numerical flux function
B_i^m	B-spline of degree m
h_i	Distance between particles in the context of B-splines: $h_i := x_{i+1} - x_i$

Bibliography

- [AS09] Ansorge R., Sonar T.: Mathematical Models of Fluid Dynamics. WILEY-VCH (2009)
- [BKO⁺96] Belytschko T., Krongauz Y., Organ D., Fleming M., Krysl P.: Meshless methods: An overview and recent developments. *Comput. Methods Appl. Mech. Engrg.* 139, 3-47 (1996)
- [Bre05] Bressan A.: Hyperbolic Systems of Conservation Laws. The One-Dimensional Cauchy Problem. Oxford University Press (2005)
- [BS02] Brenner S.C., Scott L.R.: The Mathematical Theory of Finite Element Methods. Springer-Verlag, New York (2002)
- [Cer90] Cercignani C.: Mathematical Methods in Kinetic Theory (2nd ed.). Plenum Press, New York (1990)
- [Cer94] Cercignani C., Illner R., Pulvirenti: Mathematical Theory of Dilute Gases. *Applied Mathematical Sciences* 106, Springer-Verlag, New York (1994)
- [Cha99] Chainais-Hillairet C.: Finite volume schemes for a nonlinear hyperbolic equation. Convergence towards the entropy solution and error estimate. *M2AN*, Vol. 33, No. 1, 129-156 (1999)
- [CL91] Ciarlet P.G., Lions J.L. (Eds.): Finite Element Methods (Part 1). *Handbook of Numerical Analysis*, Vol. 2, North-Holland, Amsterdam (1991)
- [CS66] Curry H.B., Schoenberg I.J.: On Polya Frequency Functions IV: The Fundamental Spline Functions and their Limits. *J. d'Analyse Math.*, Vol. 17, No. 1, 71-107 (1966)
- [DeB01] de Boor C.: A Practical Guide to Splines. Springer-Verlag (2001)
- [DiP85] DiPerna R.J.: Measure-valued solutions to conservation laws. *Arch. Rat. Mech. Anal.* 88, 223-270 (1985)
- [DO95] Duarte C.A., Oden J. T.: A Review of Some Meshless Methods to Solve Partial Differential Equations. TICAM Report 95-06 (1995)

- [EGGH98] Eymard R., Gallouët T., Ghilani M., Herbin R.: Error estimates for the approximate solutions of a nonlinear hyperbolic equation given by finite volume schemes. *IMA Journal of Numerical Analysis*, 18, 563-594 (1998)
- [EGH95] Eymard R., Gallouët T., Herbin R.: Existence and uniqueness of the entropy solution to a nonlinear hyperbolic equation. *Chin. Ann. of Math.* 16B: 1, 1-14 (1995)
- [EGH00] Eymard R., Gallouët T., Herbin R.: Finite Volume Methods. *Handbook of Numerical Analysis*, Vol. VII, 713-1018. Editors: P.G. Ciarlet and J.L. Lions (2000)
- [FM04] Fries T.-P., Matthies H.-G.: Classification and Overview of Meshfree Methods. *Informatikbericht Nr.: 2003-3*, Institute of Scientific Computing, TU Braunschweig (2004)
- [Gli65] Glimm J.: Solutions in the Large for Nonlinear Hyperbolic Systems of Equations. *Comm. on Pure and Appl. Math.*, Vol. 18, 697-715 (1965)
- [GM77] Gingold R.A., Monaghan J. J.: Smoothed Particle Hydrodynamics. *Mon. Not. Roy. Astr. Soc.* 181, 375-389 (1977)
- [GR96] Godlewski E., Raviart P.-A.: Numerical Approximation of Hyperbolic Systems of Conservation Laws. Springer-Verlag (1996)
- [GS05] Griebel M., Schweitzer M.A. (Editors): Meshfree Methods for Partial Differential Equations II. *Lecture Notes in Computational Science and Engineering*, Vol. 43, Springer-Verlag (2005)
- [GS07] Griebel M., Schweitzer M.A. (Editors): Meshfree Methods for Partial Differential Equations III. *Lecture Notes in Computational Science and Engineering*, Vol. 57, Springer-Verlag (2007)
- [HK03] Hietel D., Keck R.: Consistency by Coefficient-Correction in the Finite-Volume-Particle Method. in *Meshfree Methods for Partial Differential Equations*, Griebel M., Schweitzer M.A. (Editors), *Lecture Notes in Computational Science and Engineering*, Vol. 26, Springer, Berlin, 211-222 (2003)
- [Höl03] Hölbig K.: Finite Element Methods with B-Splines. SIAM (2003)
- [HSS00] Hietel D., Steiner K., Struckmeier J.: A finite volume particle method for compressible flows. *Math. Models Methods Appl. Sci.* 10, 1363-1382 (2000)
- [JS01] Junk M., Struckmeier J.: Consistency analysis of mesh-free methods for conservation laws. *GAMM-Mitteilungen* 2, 96-126 (2001)
- [Jun00] Junk M.: A new Perspective on Kinetic Schemes. *SIAM J. NUMER. ANAL.*, Vol. 38, No. 5, 1603-1625 (2000)

- [Kad] Kadrnka L.: PhD Thesis. Dep. of Math., Universität Hamburg, in progress
- [Kec02] Keck R.: The Finite Volume Particle Method - A Meshless Projection Method for Incompressible Flow. PhD Thesis, Dep. of Math., Universität Kaiserslautern, Shaker Verlag (2002)
- [KNS96] Klar A., Neunzert H., Struckmeier J.: Particle Methods: Theory and Applications. ICIAM 95: Proceedings of the Third International Congress on Industrial and Applied Mathematics held in Hamburg, Germany, 1995, ed. by K. Kirchgassner, O. Mahrenholtz, R. Mennicken, Berlin: Akad. Verl., Mathematical Research, Vol. 87 (1996)
- [KPS03] Katsaounis Th., Perthame B., Simeoni C.: Upwinding Sources at Interfaces in conservation laws. Appl.Math.Lett. (2003)
- [Krö97] Kröner D.: Numerical Schemes for Conservation Laws. Wiley-Teubner (1997)
- [Kru70] Kružkov S.N.: First order quasilinear equations in several independent variables. Math. USSR Sb., 10, 217-243 (1970)
- [KS08] Kaland C., Struckmeier J.: A kinetic scheme for the Savage-Hutter equations. Math. Methods in the Appl. Sc., Vol. 31, 1922-1945 (2008)
- [Kuh99] Kuhnert J.: General Smoothed Particle Hydrodynamics. Shaker Verlag (1999)
- [Lam01] Lamichhane B.P.: The Applications of Finite Volume Particle Method for Moving Boundary. Master Thesis, Dep. of Math., University of Kaiserslautern (2001)
- [Lax54] Lax P.D.: Weak Solutions of Nonlinear Hyperbolic Equations and Their Numerical Computation. Comm. Pure Apl. Math. VII, 159-193 (1954)
- [LeV92] LeVeque R.J.: Numerical Methods for Conservation Laws. Birkhäuser Verlag (1992)
- [LeV02] LeVeque R.J.: Finite Volume Methods for Hyperbolic Problems. Cambridge University Press (2002)
- [LL07] Li S., Liu W.K.: Meshfree Particle Methods. Springer-Verlag (2007)
- [LPT94] Lions P.L., Perthame B., Tadmor E.: A Kinetic Formulation of Multidimensional Scalar Conservation Laws and Related Equations. Journal of the AMS, Vol. 7, No. 1 (1994)
- [Luc77] Lucy L.: A numerical approach to testing the fission hypothesis. Astronomical J. 82, 1013-1024 (1977)
- [MK94] Monaghan J. J., Kocharyan A.: SPH simulation of multi-phase flow. Comp. Phys. Comm. 87, 225-235 (1994)

- [MNR96] J. Málek and J. Nečas and M. Rokyta and M. Růžička Weak and Measure-valued Solutions to Evolutionary PDEs. Chapman & Hall (1996)
- [Mon94] Monaghan J. J.: Simulating free surface flows with SPH. *J. Computat. Phys.* 110, 399-406 (1994)
- [Mon05] Monaghan J. J.: Smoothed Particle Hydrodynamics. *Rep. Prog. Phys.* 68, 1703-1759 (2005)
- [MTM05] Muravin B., Turkel E., Muravin G.: Solution of a Dynamic Main Crack Interaction with a System of Micro-Cracks by the Element Free Galerkin Method. In: *Meshfree Methods for Partial Differential Equations II*, eds. Griebel M., Schweitzer M.A., *Lecture Notes in Computational Science and Engineering*, Vol. 43, 149-167, Springer-Verlag (2005)
- [MVB⁺02] Mangeney-Castelnau A., Vilotte J.P., Bristeau M.O., Bouchut F., Perthame B., Simeoni C., Yernini S.: A new kinetic scheme for Saint-Venant equations applied to debris avalanches. INRIA: Rapport de recherche no. 4646 (2002)
- [MVB⁺03] Mangeney-Castelnau A., Vilotte J.P., Bristeau M.O., Bouchut F., Perthame B., Simeoni C., Yernini S.: Numerical modeling of avalanches based on Saint-Venant equations using a kinetic scheme. *Journal of Geophysical Research*, Vol. 108, No.B11, 2527 (2003)
- [Per02] Perthame B.: *Kinetic Formulation of Conservation Laws*. Oxford University Press (2002)
- [PS01] Perthame B., Simeoni C.: A kinetic scheme for the Saint-Venant system with a source term. *Calcolo*, 38, 201-231 (2001)
- [PS03] Perthame B., Simeoni C.: *Convergence of the Upwind Interface Source method for hyperbolic conservation laws*. Springer (2003)
- [QN11] Quinlan N.J., Nestor R.M.: Fast exact evaluation of particle interaction vectors in the finite volume particle method. *Meshfree Methods for Partial Differential Equations V*, eds. Griebel M., Schweitzer M.A., *Lecture Notes in Computational Science and Engineering*, Vol. 79, Springer, 219-234 (2011)
- [SH89] Savage S.B., Hutter K.: The motion of a finite mass of granular material down a rough incline. *J. Fluid Mech.*, Vol. 199, 177-215 (1989)
- [SH91] Savage S.B., Hutter K.: The dynamics of avalanches of granular materials from initiation to runout. Part 1: Analysis. *Acta Mechanica*, Vol. 86, 201-223 (1991)
- [She68] Shepard D.: A two-dimensional interpolation function for irregularly spaced data. *Proceedings of A.C.M. National Conference*, 517-524 (1968)

- [Son95] Sonar T.: Multivariate Rekonstruktionsverfahren zur numerischen Berechnung hyperbolischer Erhaltungsgleichungen. Technical Report 95-02, Deutsche Forschungsanstalt für Luft- und Raumfahrt e. V., Institut für Strömungsmechanik, Göttingen (1995)
- [Son97] Sonar T.: Mehrdimensionale ENO-Verfahren. Advances in Numerical Mathematics, B. G. Teubner, Stuttgart (1997)
- [Str95] Struckmeier J.: On a kinetic Model for Shallow Water Waves. Mathematical Methods in the Appl. Sciences, Vol.18, 709-722 (1995)
- [Str97] Struckmeier J.: Particle Methods in Fluid Dynamics. Proceedings of the 3rd Summer Conference on Numerical Modelling in Continuum Mechanics, Prague 1997, Part I, ed. by M Feistauer, R. Rannacher and K. Kozel, 47-61 (1997).
- [SW05] Schaback R., Wendland H.: Numerische Mathematik. Springer-Verlag (2005)
- [TAH⁺07] Tiwari S., Antonov S., Hietel D., Kuhnert J., Olawsky F., Wegener R.: A Mesh-free Method for Simulations of Interactions between Fluids and Flexible Structures. In: Meshfree Methods for Partial Differential Equations III, eds. Griebel M., Schweitzer M.A., Lecture Notes in Computational Science and Engineering, Springer-Verlag (2007)
- [Tel00] Teleaga D.: Numerical Studies of a Finite-Volume Particle Method for Conservation Laws. Master Thesis, Dep. of Math., Universität Kaiserslautern (2000)
- [Tel05] Teleaga D.: A Finite-Volume Particle Method for Conservation Laws. PhD Thesis, Dep. of Math., Universität Hamburg, Logos Verlag (2005)
- [TNGH02] Tai Y.C., Noelle S., Gray J.M.N.T., Hutter K.: Shock-Capturing and Front-Tracking Methods for Granular Avalanches. J. of Computational Physics 175, 269-301 (2002)
- [TS08] Teleaga D., Struckmeier J.: A finite-volume particle method for conservation laws on moving domains. Int. J. Numer. Meth. Fluids (2008)
- [Vol67] Vol’pert A. I.: The spaces BV and quasilinear equations. Math. USSR Sb. 2, 225-267 (1967)
- [WA10] Westover L.M., Adeeb S.M.: Cubic Spline Meshless Method for Numerical Analysis of the Two-Dimensional Navier-Stokes Equations. International Journal of Numerical Analysis and Modeling, Series B, Vol. 1, No. 2, 172-196 (2010)
- [ZP13] Zarebnia M. and Parvaz R.: Septic B-spline collocation method for numerical solution of the Kuramoto-Sivashinsky equation. International Journal of Mathematical Sciences 7(4) (2013)

Summary

This thesis is concerned with the Finite Volume Particle Method (FVPM), a meshless numerical method for solving hyperbolic conservation laws. After a general overview on hyperbolic conservation laws and the idea of the standard Finite Volume Method (FVM), the FVPM is deduced following the derivation of the FVM. We show L^∞ -stability, a weak BV-stability result, positivity, monotonicity and a discrete entropy inequality of the FVPM for scalar equations in several space dimensions. Subsequently, the FVPM is combined with B-splines in a one-dimensional setting and coupled with an existing kinetic scheme to handle the dynamical behaviour of granular masses. Numerical results are shown for the one-dimensional Euler equations, a one-dimensional linear piston problem with moving boundary, the Savage-Hutter equations and a two-dimensional solid body rotation problem.

Zusammenfassung

Diese Arbeit befasst sich mit der Analyse und Weiterentwicklung der Finite Volumen Partikel Methode (FVPM), einer gitterfreien Methode zur numerischen Berechnung von Lösungen hyperbolischer Erhaltungsgleichungen. Nach einem kurzen Überblick über hyperbolische Erhaltungsgleichungen und der Einführung in das Gebiet der Finite Volumen Methoden (FVM) leiten wir die FVPM her. Als zentrale Ergebnisse zeigen wir L^∞ -Stabilität, eine schwache BV-Stabilität, die Positivität des Verfahrens, Monotonie sowie eine diskrete Entropie-Ungleichung für skalare Erhaltungsgleichungen in mehreren Raumdimensionen. Anschließend formulieren wir die FVPM mit B-Splines als Ansatz- und Testfunktionen in einer Raumdimension und ermöglichen damit eine effiziente Berechnung der geometrischen Koeffizienten. Darüber hinaus nutzen wir eine bereits erfolgreich umgesetzte kinetische Methode zur Lösung der Savage-Hutter Gleichungen und erhalten daraus eine Formulierung der FVPM zur Beschreibung von Strömungen granularer Medien. Schließlich werden numerische Resultate für die eindimensionalen Eulergleichungen, ein eindimensionales Kolben-Problem mit bewegtem Rand, die Savage-Hutter Gleichungen sowie ein zweidimensionales Rotationsproblem angegeben.

Lebenslauf

22.10.1980	geboren in Hamburg
08/1991 - 07/2000	Theodor-Heuss-Schule Pinneberg
07/2000	Erhalt der Allgemeinen Hochschulreife (Abitur)
10/2000 - 09/2001	Studium der Allgemeinen Ingenieurwissenschaften an der Technischen Universität Hamburg-Harburg
10/2001 - 09/2004	Studium der Technomathematik an der Universität Hamburg
10/2004 - 03/2005	Teilnahme am Erasmus-Austauschprogramm: Studium der Mathematik an der Université Henri Poincaré in Nancy, Frankreich
04/2005 - 03/2007	Studium der Technomathematik an der Universität Hamburg
03/2007	Diplom im Fach Technomathematik
04/2007 - 03/2013	Wissenschaftliche Mitarbeiterin am Fachbereich Mathematik der Universität Hamburg
09/2013 - heute	Wissenschaftliche Mitarbeiterin bei HITEC e.V., Mitarbeit im Forschungs- und Entwicklungsprojekt zur Erstellung eines MATLAB-Tools zur Auslegung von Rohrreaktoren mit der Firma ENCOS GmbH