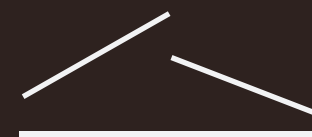




DISCONTINUOUS GALERKIN FINITE ELEMENT METHODS FOR (NON)CONSERVATIVE PARTIAL DIFFERENTIAL EQUATIONS

DISCONTINUOUS GALERKIN
FINITE ELEMENT METHODS
FOR (NON)CONSERVATIVE
PARTIAL DIFFERENTIAL
EQUATIONS



SANDER RHEBERGEN

ISBN 978-90-365-2964-8

SANDER RHEBERGEN

**DISCONTINUOUS GALERKIN FINITE ELEMENT
METHODS FOR (NON)CONSERVATIVE PARTIAL
DIFFERENTIAL EQUATIONS**

Sander Rhebergen

The research presented in this thesis was carried out at the group of Numerical Analysis and Computational Mechanics (NACM), Department of Applied Mathematics, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands.



UNIVERSITY OF TWENTE.

The research was supported by the Institute of Mechanics, Processes and Control Twente (IMPACT) and partly by ADIGMA, a European project on the development of adaptive higher order variational methods for aerospace applications.



Copyright © by S. Rhebergen, Enschede, The Netherlands, 2010.

Printed by Wöhrmann Printing Service, Zutphen, The Netherlands.

ISBN 978-90-365-2964-8

DOI 10.3990/1/9789036529648

**DISCONTINUOUS GALERKIN FINITE ELEMENT
METHODS FOR (NON)CONSERVATIVE PARTIAL
DIFFERENTIAL EQUATIONS**

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof. dr. H. Brinksma,
volgens het besluit van het College voor Promoties
in het openbaar te verdedigen
op vrijdag 5 februari 2010 om 15.00 uur

door

Sander Rhebergen
geboren op 19 oktober 1981
te Kagondo, Tanzania

Dit proefschrift is goedgekeurd door de promotor
prof. dr. ir. J.J.W van der Vegt

Contents

1	Introduction	1
2	DGFEM for hyperbolic nonconservative pdes: Theory	9
2.1	Nonconservative hyperbolic partial differential equations	9
2.2	Space-time DGFEM discretization	13
2.3	The NCP numerical flux	23
3	DGFEM for hyperbolic nonconservative pdes: Applications	29
3.1	Obtaining and solving the discrete system	29
3.2	One dimensional test cases	32
3.3	Effect of the path in phase space on the numerical solution	53
4	DGFEM for shallow two-phase flows	59
4.1	Depth-averaged two-phase flows	59
4.2	The DGFEM discretization	63
4.3	Verification	69
4.4	Validation	75
5	h-Multigrid optimization for higher order accurate ST-DG	81
5.1	Space-time DG for the 2D advection-diffusion equation	81
5.2	h -Multigrid algorithms for linear systems	86
5.3	Fourier analysis of discrete operators	88
5.4	Optimizing multigrid	98
5.5	Rescaling	99

5.6	Testing multigrid performance	106
6	Multigrid for higher order accurate ST-DG: Euler equations	113
6.1	Space-time DG for the Euler equations	113
6.2	Multigrid methods	116
6.3	Fourier analysis for p - and hp -multigrid	120
6.4	Multigrid algorithms applied to the Euler equations	123
7	Alternative derivation of the DG weak formulation	131
7.1	Borel measures in DG	133
7.2	Space-time DGFEM weak formulation	135
7.3	Derivation based on Borel measures	138
7.4	Test cases	147
8	Conclusions and recommendations	153
A	Space DGFEM for hyperbolic nonconservative pdes	159
B	The three-dimensional two-phase flow model	163
	Bibliography	165
	Summary	173
	Samenvatting	175
	Acknowledgments	177
	List of publications	179
	About the author	181

CHAPTER 1

Introduction

Debris flows are flows of water-saturated slurry mixtures [37, 52, 62]). Examples are mud slides initiated by heavy rainfall on eroded mountain sides consisting of mixtures of rock, sand and mud; and volcanic debris flows in which the flow may be a mixture of volcanic debris and water (see Fig. 1a). These flows often cause major destruction to buildings and infrastructure, with accompanying loss of human lives. In industrial applications, dense liquid-solid flows, such as slurry flows, are used in pipeline transportation (see Fig. 1b). This form of transportation has relatively low operation and maintenance costs, and is friendly to the environment [48]. Other applications occur for instance in liquid fluidized beds [38].

The first objective of the research in this thesis is to be able to solve hydrodynamic models of two-phase flows which describe the motion of the above mentioned debris flows. These models contain many interesting aspects, e.g., the presence of nonconservative products, stiff source terms and flows with free-surfaces. In this thesis we will provide some of the tools necessary for solving these models by space- and/or space-time discontinuous Galerkin (DG) finite element methods.

The second objective of this thesis is to develop fast multigrid methods for the solution of the algebraic system of equations originating from the space-time DG discretization. In particular for higher order accurate DG discretizations of practical problems, a significant improvement in computational performance is essential. In the next sections we will discuss the main topics in more detail.



(a) The lahar developed on the slopes of Santiaguito volcano [33]. Photograph courtesy of U.S. Geological Survey.



(b) Slurry and sediment transport in pipelines. Photograph courtesy of LI-Ceengineering [34].

Figure 1.1: *Examples of two-phase flows.*

Nonconservative products. We are interested in solving dispersed two-phase two-fluid models. The use of a discontinuous Galerkin (DG) method for these problems is of interest because it can deal efficiently with unstructured and deforming grids, local mesh refinement (h -adaptation), adjustment of the polynomial order in each element (p -refinement), and parallel computation. These benefits stem from the compact stencil used in DG methods, i.e., the solution on an element depends only on the data of its immediate neighboring elements. Furthermore, the DG finite element method easily deals with shocks and other discontinuities in the solution. Dispersed two-phase two-fluid models contain, however, nonconservative products which are introduced in the governing equations in the modeling procedure [22, 23]. This poses serious problems which motivated the research in Chapters 2 and 3 to present a way to genuinely deal with nonconservative products in a DG finite element context.

Systems of equations containing nonconservative products cannot be transformed into divergence form, i.e., equations of the form $\partial_t u + \partial_x f(u) + g(u) \partial_x u = 0$ cannot be written as $\partial_t u + \partial_x h(u) = 0$. This causes problems once the solution becomes discontinuous, since the weak solution in the classical sense of distributions then does not exist. Consequently, no classical Rankine-Hugoniot shock conditions can be defined. To overcome these problems we use the theory of Dal Maso, LeFloch and Murat (DLM) [54] for nonconservative products. In this theory a definition is given for nonconservative products of the type $g(u) \partial_x u$, where $g : \mathbb{R}^m \rightarrow \mathbb{R}^m$ is a smooth function, but $u :]a, b[\rightarrow \mathbb{R}^m$ may admit discontinuities. Using this theory, a notion of a weak solution can be given to the Riemann problem for nonconservative hyperbolic partial differential equa-

tions. In Chapter 2 we will use the DLM theory to propose a new discontinuous Galerkin (DG) finite element method suitable for hyperbolic partial differential equations in nonconservative form. A problem with the DLM theory is, however, the introduction of a path in phase space connecting the left and right state across a discontinuity. It is possible to derive an expression for this path by constructing entropy solutions to the nonconservative hyperbolic equations (see LeFloch [46]), but this construction can be a very difficult as well as costly job. In Chapter 3 we will therefore investigate the influence of this path in phase space on the numerical solution.

Over the years several authors have been developing numerical methods suitable for nonconservative hyperbolic partial differential equations with non-smooth solutions. Toumi [75] introduced a generalized Roe solver based on the DLM theory, which was later applied by Toumi and Kumbaro [76] to shock tube problems and two-fluid problems. The work by Toumi [75] was also used by Parés [57], Castro, Gallardo and Parés [14] and Parés and Castro [58] to develop numerical schemes in the finite volume context. Alternative approaches in which the DLM theory is not used are followed by Saurel and Abgrall [66] and Xing and Shu [87]. The latter work considers high order well-balanced finite volume WENO and Runge-Kutta discontinuous Galerkin methods for systems containing nonconservative products. The schemes of Xing and Shu are designed such that they exactly maintain the balance laws at the discrete level for certain steady state solutions. In Chapter 2 we use the DLM theory in a DG finite element context to give the nonconservative products a proper definition at locations where discontinuities are present. This work differs from the previously mentioned works in that we do not formulate a weak formulation based on generalized Roe solvers. Instead, we present and use a new numerical flux in the context of the DLM theory.

Depth-averaged two-phase flows. In many flows the height H of the flow is much smaller than the length L of the flow, $H/L \ll 1$. For these flows, depth-averaging techniques are commonly used to simplify the three dimensional equations. Examples include the shallow water equations derived from the incompressible Navier-Stokes equations or the Savage-Hutter equations for dry granular flow [67]. Recently, Pitman and Le [62] and Le [45] derived a depth-averaged model for two-phase flows based on a three dimensional continuum model for two-phase flows as derived by Jackson [38] (see Appendix B for the three dimensional continuum model). We remark, however, that with the assumptions made in the depth-averaging process by Le [45], the same depth-averaged model can be derived from the three dimensional model of Drew and Lahey [22]. In Chapter 4 we have slightly extended the depth-averaged model by also including extra friction terms to simulate turbulent friction and we

present a discontinuous Galerkin finite element method for the depth-averaged two-phase flow model.

Much of the research conducted with depth averaged models for liquid-solid flows focuses on correctly predicting the final depositions of debris avalanches and their behavior over natural terrains (Denlinger and Iverson [20], Patra et al. [60, 59], Pouliquen and Forterre [63], Tai et al. [70], Wang et al. [83]). In Chiou et al. [15] and Gray et al. [26] also the influence of obstacles on granular flows is investigated. We are, however, interested in the behavior of debris flows through contractions and in Chapter 4 we will perturb a steady-state two-phase flow with a low particle volume fraction by introducing an upstream avalanche of particles for a short period, thus temporarily increasing the particle volume fraction. This experiment was done by Akers and Bokhove [2] (see Fig. 1.2) and we use this experiment to qualitatively validate the depth-averaged two-phase flow model.

Multigrid. The space-time DG method is implicit in time and requires the solution of a system of algebraic equations at each time step. To solve this system we consider multigrid techniques, which are very efficient and versatile techniques for the solution of large systems of (non)linear algebraic equations. During the past decades many different multigrid algorithms have been developed and applied to a wide variety of problems. Furthermore, an extensive mathematical analysis has been conducted for many multigrid algorithms resulting in detailed knowledge about the design of optimal multigrid algorithms, their performance and efficient implementation.

In this thesis we consider the use of a pseudo-time multigrid technique originally developed by Jameson [39] and further extended in [55]. This method was applied in the early 2000's in a space-time DG context in [79] for second order accurate discretizations of the Euler equations and has been a preferred method to solve space-time DG discretizations since [4, 40, 41, 42, 61, 64]. Our objective is to improve this method for higher order space-time DG discretizations since it preserves the locality of the DG scheme and is very useful in a FAS multigrid scheme for nonlinear problems.

The main components in a multigrid algorithm are an iterative method and coarsened approximations of the algebraic system. In addition, restriction and prolongation operators are necessary to connect the various approximations of the algebraic system. In case of partial differential equations the coarsened algebraic systems can be obtained by either discretizing the equations on coarser meshes, resulting in h -multigrid algorithms [30, 42, 80, 79], or by using discretizations with different orders of accuracy, which give p -multigrid methods [9, 24, 50, 53]. Of course combinations of both techniques are possible, resulting in hp -multigrid methods [56, 68]. In Chapter 5 we will consider h -

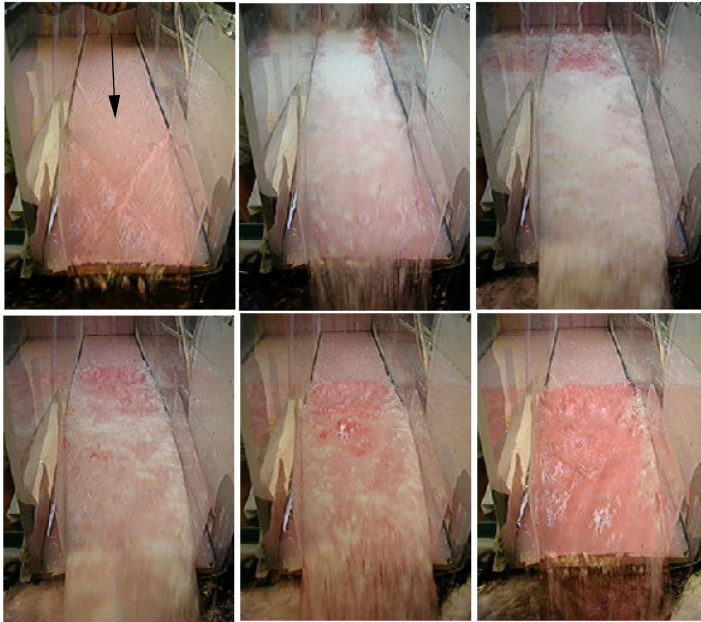


Figure 1.2: When water flow enters a contraction at a certain speed, a steady state in the contraction is reached with oblique hydraulic jumps (top left). This steady-state is perturbed by an upstream avalanche of polystyrene beads (just inserted in the top middle frame). There is a transition period in the top right, bottom left and bottom middle frames in which one second elapses between each frame. A second steady state, an upstream steady shock, is reached (bottom right) [2].

multigrid methods while in Chapter 6 we discuss some preliminary results using also p - and hp -multigrid.

The design of the iterative method (also known as a smoother) and the restriction and prolongation operators are crucial for multigrid performance. Also, the coarsening of the algebraic system can have a significant impact. For linear problems discrete Fourier analysis can provide detailed information on these aspects. This is achieved by analyzing the full two- or three-level multigrid algorithm. Due to its complexity, the analysis of multigrid algorithms is frequently restricted to two-level analysis, or even the simpler analysis of only the multigrid smoother. For many problems this results in a rather poor prediction of the actual multigrid performance. It is therefore important to consider realistic

model problems and extend the analysis to three grid levels. This can significantly enhance the accuracy of the analysis and is essential when optimizing the multigrid algorithm, see e.g. [86].

The h -multigrid algorithm using a pseudo-time integration method discussed in this thesis was originally developed in [42, 79] for second order accurate space-time DG discretizations of the compressible Euler and Navier-Stokes equations. The algorithm is easy to implement and parallelize, even on locally refined meshes, and is insensitive to initial conditions. For higher order accurate DG discretizations the multigrid performance was, however, not satisfactory. In Chapter 5 we therefore discuss the analysis of the h -multigrid algorithm and demonstrate how the multigrid performance for higher order DG discretizations can be improved.

In Chapter 5 we perform a two- and three-level Fourier analysis for space-time DG discretizations of a linear PDE. To better approximate “real” problems, such as the compressible Navier-Stokes equations, we consider the 2D advection-diffusion equation instead of the 1D model problem as was done in [42]. The Fourier analysis then serves as a tool to optimize our h -multigrid algorithms. In [79], an explicit Runge-Kutta time integrator was optimized for single-grid 1D computations. This Runge-Kutta scheme was used as a smoother in the h -multigrid algorithm without further modifications. In Chapter 5 we optimize explicit Runge-Kutta smoothers in combination with the three-level h -multigrid algorithm. This combined optimization process should result in a significantly improved algorithm. The optimization is performed for a three-level h -multigrid algorithm for the solution of the 2D advection-diffusion equation discretized with a second or third order accurate space-time DG discretization.

It is known that multigrid methods converge slowly on grids with high aspect ratio cells. High aspect ratio cells cause a strong coupling in one direction and a weak coupling in other directions, see e.g. [77]. Different techniques are proposed in the literature to tackle this problem, e.g. semi-coarsening and line implicit smoothing. Both, however, become impractical for three dimensional computations on unstructured grids. In order to address this problem, recently, Lucas et al. [49] proposed a Newton linearization in combination with a Krylov subspace technique for unsteady flow computations. To maintain the locality of the discontinuous Galerkin discretization, we prefer however the use of pseudo-time integration methods using explicit Runge-Kutta time integrators. In order to reduce the stiffness of the discretization introduced by the high aspect ratio cells, we have derived a rescaling in Chapter 5 to be added to the pseudo-time algorithm. By investigating how the scaling of the discretization changes on high and low aspect ratio cells, a rescaling can be derived to better balance the discretizations in all directions. A similar argument holds for the change from inviscid to viscous flows in a boundary layer.

In Chapter 6 we perform a three-level Fourier analysis of h -, p - and hp -multigrid techniques applied again to the 2D advection-diffusion equation. We test the multigrid schemes then through the computation of inviscid flow over a NACA0012 airfoil, solving the Euler equations of gas dynamics. We found that the choice of the basis functions in the numerical simulations has a large effect on the stability of the computations. In order to investigate this, we have calculated the spectrum of a third order accurate space-time DG discretization for the 2D Euler equations for the flow over a NACA0012 airfoil. This provided useful information on the best choice of basis functions.

An alternative derivation. Many partial differential equations describing fluid flow contain second (and higher) order derivatives. Obtaining a DG discretization for these higher order derivatives is non-trivial and many different DG methods exist to deal with these terms. Even for first order partial differential equations there are some issues regarding the derivation of the DG weak formulation. In Chapter 7 we aim at a mathematically more consistent derivation of DG discretizations. This derivation is different from the classical derivation in that we introduce generalized DG derivatives based on bounded Borel measures. On element boundaries the generalized DG derivative is well defined despite the discontinuity in the numerical approximation at the element faces. Using this alternative approach we recover the standard weak formulation for hyperbolic partial differential equations (PDE's). For parabolic and elliptic PDE's, two DG formulations can be obtained, a new weak formulation and the weak formulation proposed by Brezzi et al. [13]. Numerical simulations are conducted to compare both weak formulations.

Outline. In Chapter 2 we derive the space-time DG finite element formulation for nonconservative partial differential equations and state the space DG finite element formulation as a special case in Appendix A. In DG methods, the numerical flux plays an essential role. In Chapter 2 we therefore also derive a numerical flux for systems with nonconservative products (NCP-flux) which can also be applied to moving grids.

We apply the DG finite element method to two depth-averaged and dispersed multiphase systems in Chapter 3 and show numerical results using a linear path in phase space. We then investigate the effect of the different paths in phase space on the numerical solution.

In Chapter 4 we present a discontinuous Galerkin finite element method for the depth-averaged two-phase flow model as derived by Pitman and Le [62] and Le [45]. We numerically verify and validate the method and the model. The DG method does not guarantee monotone solutions around discontinuities and sharp gradients and thus numerical oscillations can develop. To prevent these

numerical oscillations we investigate and clarify the WENO slope limiter given in [51] in combination with Krivodonova's discontinuity detector [43].

Space-time DG methods of partial differential equations result in large systems of algebraic equations that need to be solved at each time-step. To efficiently solve these systems, we combine pseudo-time integration methods with multigrid techniques. In Chapter 5, a two- and three-level Fourier analysis of the multigrid algorithms is conducted. The Fourier analysis provides the spectral radius of the multigrid algorithm which gives a prediction of the asymptotic rate of convergence of the multigrid method. We optimize the h -multigrid method by minimizing the spectral radius. In Chapter 6 we further analyze p - and hp -multigrid using Fourier analysis and perform "real-life" simulations.

An alternative derivation of the discontinuous Galerkin (DG) finite element weak formulation is given in Chapter 7. We introduce generalized DG derivatives based on Borel measures which lead to a mathematically more consistent derivation. We compare numerical results of a new DG weak formulation for higher-order derivatives with the well known method of Brezzi et al. [13]. For this we consider the compressible Navier-Stokes equations in which we simulate viscous flow past a cylinder and a NACA0012 airfoil. Conclusions are drawn in Chapter 8.

Discontinuous Galerkin finite element methods for hyperbolic nonconservative partial differential equations: Theory

In this chapter we present a discontinuous Galerkin finite element (DGFEM) formulation for systems containing nonconservative products, such as occur in dispersed multiphase flow equations. The main criterium we pose on the weak formulation is that if the system of nonconservative partial differential equations can be transformed into conservative form, then the formulation must reduce to that for conservative systems. Standard DGFEM formulations cannot be applied to nonconservative systems of partial differential equations. We therefore introduce the theory of weak solutions for nonconservative products into the DGFEM formulation. We also introduce a new numerical flux that is able to deal with nonconservative products.

2.1 Nonconservative hyperbolic partial differential equations

The main topic of this chapter is the derivation of a formulation for DGFEM suitable for nonlinear hyperbolic partial differential equations in nonconservative form. We use the theory of Dal Maso, LeFloch and Murat (DLM) [54] to overcome the absence of a weak solution in the classical sense of distributions for these types of equations. In an article by Dal Maso, LeFloch and Murat [54],

10 Chapter 2: DGFEM for hyperbolic nonconservative pdes: Theory

a definition was given for nonconservative products of the form $g(u)\partial_x u$, where $g : \mathbb{R}^m \rightarrow \mathbb{R}^m$ is a smooth function, but $u :]a, b[\rightarrow \mathbb{R}^m$ may admit discontinuities. They assumed u to be a function of bounded variation (BV), viz. a Lebesgue integrable function whose first derivative is a bounded Borel measure, and the product $g(u)\partial_x u$ is defined as a Borel measure on $]a, b[$. Such a definition is necessary when g is not the differential of a smooth function q , i.e., there is no q such that $g(u)\partial_x u$ admits a conservative form $\partial_x q$. The following example, given by LeFloch [46], illustrates the DLM theory.

Consider the function $u(x)$ composed of two constant vectors u_L and u_R in \mathbb{R}^m with $u_L \neq u_R$:

$$u(x) = u_L + \mathcal{H}(x - x_d)(u_R - u_L), \quad x \in]a, b[, \quad (2.1)$$

where $x_d \in]a, b[$ and $\mathcal{H} : \mathbb{R} \rightarrow \mathbb{R}$ is the Heaviside function with $\mathcal{H}(x) = 0$ if $x < 0$ and $\mathcal{H}(x) = 1$ if $x > 0$. Consider any smooth function $g : \mathbb{R}^m \rightarrow \mathbb{R}^m$. We see immediately that $g(u)\partial_x u$ is not defined at $x = x_d$ since here $|\partial_x u| \rightarrow \infty$. Dal Maso, LeFloch and Murat [54] introduce therefore a smooth regularization u^ε of the discontinuous function u . They show that in this particular case, if the total variation of u^ε remains uniformly bounded with respect to ε :

$$g(u) \frac{du}{dx} \equiv \lim_{\varepsilon \rightarrow 0} g(u^\varepsilon) \frac{du^\varepsilon}{dx}$$

gives a sense to the nonconservative product as a bounded measure. This limit, however, depends on how we choose u^ε . Introduce a Lipschitz continuous path $\phi : [0, 1] \rightarrow \mathbb{R}^m$, satisfying $\phi(0) = u_L$ and $\phi(1) = u_R$, connecting u_L and u_R in \mathbb{R}^m . The following regularization u^ε for u then emerges:

$$u^\varepsilon(x) = \begin{cases} u_L, & \text{if } x \in]a, x_d - \varepsilon[\\ \phi\left(\frac{x - x_d + \varepsilon}{2\varepsilon}\right), & \text{if } x \in]x_d - \varepsilon, x_d + \varepsilon[\\ u_R, & \text{if } x \in]x_d + \varepsilon, b[\end{cases} \quad \varepsilon > 0. \quad (2.2)$$

Using this regularization, LeFloch [46] states that when ε tends to zero, then:

$$g(u^\varepsilon) \frac{du^\varepsilon}{dx} \rightarrow C \delta_{x_d}, \quad \text{with } C = \int_0^1 g(\phi(\tau)) \frac{d\phi}{d\tau}(\tau) d\tau,$$

vaguely in the sense of measures on $]a, b[$, where δ_{x_d} is the Dirac measure at x_d . We see that the limit of $g(u^\varepsilon)\partial_x u^\varepsilon$ depends on ϕ . There is one exception, namely if a $q : \mathbb{R}^m \rightarrow \mathbb{R}$ exists with $g = \partial_u q$. In this case $C = q(u_R) - q(u_L)$. We are, however, interested in the case when such a function q does not exist. We then see that the definition of the nonconservative product $g(u)\partial_x u$ must depend on the path ϕ chosen in the regularization. In Section 3.3, we will

investigate the effect of different paths ϕ on the numerical solution. For now, assume that the path ϕ is given. In Dal Maso, LeFloch and Murat [54] it is assumed that the path belongs to a fixed family of paths in \mathbb{R}^m . These paths are Lipschitz continuous maps $\phi : [0, 1] \times \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ which satisfy the following properties:

(H1) $\phi(0; u_L, u_R) = u_L, \phi(1; u_L, u_R) = u_R,$

(H2) $\phi(\tau; u_L, u_L) = u_L,$

(H3) $\left| \frac{\partial \phi}{\partial \tau}(\tau; u_L, u_R) \right| \leq K |u_L - u_R|, \text{ a.e. in } [0, 1].$

Dal Maso, LeFloch and Murat [54] consider functions $u :]a, b[\rightarrow \mathbb{R}^m$ of bounded variation, viz. $u \in BV(]a, b[, \mathbb{R}^m)$. These are functions of $L^1(]a, b[, \mathbb{R}^m)$ whose first order derivative is a bounded Borel measure on the interval $]a, b[$. Since u is BV, u admits a countable set of discontinuity points and at each such point x_d , a left trace $u_L = \lim_{\varepsilon \downarrow 0} u(x_d - \varepsilon)$ and a right trace $u_R = \lim_{\varepsilon \downarrow 0} u(x_d + \varepsilon)$ exist. For more on Borel measures, BV functions and related topics, see, e.g., [89].

Based on the family of paths satisfying (H1)-(H3), the following theorem is given by Dal Maso, LeFloch and Murat [54]:

Theorem 2.1.1. *Let $u :]a, b[\rightarrow \mathbb{R}^m$ be a function of bounded variation and $g : \mathbb{R}^m \rightarrow \mathbb{R}^m$ be a continuous function. Then, there exists a unique real-valued bounded Borel measure μ on $]a, b[$ characterized by the two following properties:*

1. *If u is continuous on a Borel set $B \subset]a, b[$, then:*

$$\mu(B) = \int_B g(u) \frac{du}{dx} d\lambda,$$

where λ is the Borel measure.

2. *If u is discontinuous at a point x_d of $]a, b[$, then:*

$$\mu(\{x_d\}) = \int_0^1 g(\phi(\tau; u_L, u_R)) \frac{\partial \phi}{\partial \tau}(\tau; u_L, u_R) d\tau.$$

By definition, this measure μ is the nonconservative product of $g(u)$ by $\partial_x u$ and is denoted by $\mu = [g(u) \frac{du}{dx}]_\phi$.

In this chapter we will derive a space-time DGFEM weak formulation for nonlinear hyperbolic systems of partial differential equations in nonconservative form in multi-dimensions:

$$U_{i,0} + F_{ik,k} + G_{ikr} U_{r,k} = 0, \quad \bar{x} \in \mathbb{R}^q, \quad t > 0, \tag{2.3}$$

12 Chapter 2: DGFEM for hyperbolic nonconservative pdes: Theory

with $U \in \mathbb{R}^m$, $F \in \mathbb{R}^m \times \mathbb{R}^q$, $G \in \mathbb{R}^m \times \mathbb{R}^q \times \mathbb{R}^m$; we use the comma notation to denote partial differentiation and the summation convention on repeated indices. Here, $(\cdot)_{,0}$ denotes partial differentiation with respect to time and $(\cdot)_{,k}$ ($k = 1, \dots, q$) partial differentiation with respect to the spatial coordinates. In a space-time context, space and time variables are, however, not explicitly distinguished. A point at time $t = x_0$ with position $\bar{x} = (x_1, x_2, \dots, x_q)$ has Cartesian coordinates $x = (x_0, \bar{x}) \in \mathbb{R}^{q+1}$. We can write (2.3) then as:

$$T_{ikr}U_{r,k} = 0, \quad x \in \mathbb{R}^{q+1}, \quad x_0 > 0, \quad k = 0, 1, 2, \dots, q, \quad (2.4)$$

with $U \in \mathbb{R}^m$ and $T \in \mathbb{R}^m \times \mathbb{R}^{q+1} \times \mathbb{R}^m$ given by:

$$T_{ikr} = \begin{cases} \delta_{ir}, & \text{if } k = 0, \\ D_{ikr}, & \text{otherwise,} \end{cases} \quad (2.5)$$

where δ represents the Kronecker delta symbol and where $D_{ikr} = \partial F_{ik} / \partial U_r + G_{ikr}$. Dal Maso, LeFloch and Murat [54] give a similar theorem to Theorem 2.1.1 for the nonconservative term $T_{ikr}U_{r,k}$ in multi-dimensions. As before, assume a given family of Lipschitz continuous paths $\phi : [0, 1] \times \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ that satisfy, for some $K > 0$ and for all $U^L, U^R \in \mathbb{R}^m$ and $\tau \in [0, 1]$, the properties:

$$(H1) \quad \phi_r(0; U^L, U^R) = U_r^L, \quad \phi_r(1; U^L, U^R) = U_r^R,$$

$$(H2) \quad \phi_r(\tau; U^L, U^L) = U_r^L,$$

$$(H3) \quad \left| \frac{\partial \phi_r}{\partial \tau}(\tau; U^L, U^R) \right| \leq K |U_r^L - U_r^R|, \text{ a.e. in } [0, 1],$$

$$(H4) \quad \phi_r(\tau; U^L, U^R) = \phi_r(1 - \tau; U^R, U^L).$$

Note that property H4 has been added, which does not have to be satisfied in the one dimensional case. Let $\Omega \subset \mathbb{R}^{q+1}$ with $\Omega = \Omega_u \cup S_u \cup I_u$ where Ω_u is the set of points of approximate continuity, S_u the set of points of approximate jump and I_u contains the irregular points. The DLM theorem then states:

Theorem 2.1.2. *Let $U : \Omega \rightarrow \mathbb{R}^m$ be a bounded function of bounded variation defined on an open subset Ω of \mathbb{R}^{q+1} and $T : \mathbb{R}^m \rightarrow \mathbb{R}^m$ be a locally bounded Borel function. Then there exists a unique family of real-valued bounded Borel measures μ_i on Ω , $i = 1, 2, \dots, m$ such that*

1. if B is a Borel subset of Ω_u , then:

$$\mu_i(B) = \int_B T_{ikr}U_{r,k} d\lambda, \quad (2.6)$$

where λ is the Borel measure;

2. if B is a Borel subset of S_u , then:

$$\mu_i(B) = \int_{B \cap S_u} \int_0^1 T_{ikr}(\phi(\tau; U^L, U^R)) \frac{\partial \phi_r}{\partial \tau}(\tau; U^L, U^R) d\tau n_k^L dH^q, \quad (2.7)$$

with U^L and U^R the left and right traces at the discontinuity, where H^q denotes the q -dimensional Hausdorff measure and where we choose n^L the outward normal with respect to the left state;

3. if B is a Borel subset of I_u , then $\mu_i(B) = 0$.

The measure μ_i is the nonconservative product of T_{ikr} by $U_{r,k}$, denoted by:

$$\mu_i = [T_{ikr} U_{r,k}]_{\phi}. \quad (2.8)$$

In particular, a piecewise \mathcal{C}^1 function U is a weak solution of (2.4) if and only if the following two conditions are satisfied [14]:

1. U is a classical solution in the domains where it is \mathcal{C}^1 .
2. At a discontinuity U satisfies the generalized Rankine-Hugoniot conditions:

$$-\sigma(U_i^R - U_i^L) + F_{ik}(U^R)\bar{n}_k^L - F_{ik}(U^L)\bar{n}_k^L + \int_0^1 G_{ikr}(\phi(\tau; U^L, U^R)) \frac{\partial \phi_r}{\partial \tau}(\tau; U^L, U^R) d\tau \bar{n}_k^L = 0, \quad (2.9)$$

where σ is the speed of propagation of the discontinuity, U^L and U^R are the left and right limits of the solution at the discontinuity and \bar{n}^L is the space component of the space-time normal n^L (see e.g. LeFloch [46]).

When $G(U)$ is the Jacobian of some flux function $Q(U)$, jump conditions (2.9) are independent of the path and reduce to the Rankine-Hugoniot condition:

$$H_{ik}(U^R)\bar{n}_k^L - H_{ik}(U^L)\bar{n}_k^L = \sigma(U_i^R - U_i^L), \quad (2.10)$$

where $H = F + Q$.

2.2 Space-time DGFEM discretization

In this section we will introduce the formulation for space-time DGFEM for systems of hyperbolic partial differential equations containing nonconservative products. We will start by introducing space-time elements, function spaces, trace operators and basis functions, after which we derive the space-time DG formulation. In Appendix A we also give the formulation for space DGFEM.

2.2.1 Space-time elements

In the space-time DGFEM method, the space and time variables are not distinguished. A point at time $t = x_0$ with position vector $\bar{x} = (x_1, x_2, \dots, x_q)$ has Cartesian coordinates (x_0, \bar{x}) in the open domain $\mathcal{E} \subset \mathbb{R}^{q+1}$. At time t , the flow domain $\Omega(t)$ is defined as:

$$\Omega(t) := \{\bar{x} \in \mathbb{R}^q : (t, \bar{x}) \in \mathcal{E}\}.$$

By taking t_0 and T as the initial and final time of the evolution of the space-time flow domain, the space-time domain boundary $\partial\mathcal{E}$ consists of the hyper-surfaces:

$$\begin{aligned} \Omega(t_0) &:= \{x \in \partial\mathcal{E} : x_0 = t_0\}, \\ \Omega(T) &:= \{x \in \partial\mathcal{E} : x_0 = T\}, \\ \mathcal{Q} &:= \{x \in \partial\mathcal{E} : t_0 < x_0 < T\}. \end{aligned}$$

The time interval $[t_0, T]$ is partitioned using the time levels $t_0 < t_1 < \dots < T$, where the n -th time interval is defined as $I_n = (t_n, t_{n+1})$ with length $\Delta t_n = t_{n+1} - t_n$. The space-time domain \mathcal{E} is then divided into N_t space-time slabs $\mathcal{E}^n = \mathcal{E} \cap I_n$. Each space-time slab \mathcal{E}^n is bounded by $\Omega(t_n)$, $\Omega(t_{n+1})$ and $\mathcal{Q}^n = \partial\mathcal{E}^n / (\Omega(t_n) \cup \Omega(t_{n+1}))$.

The flow domain $\Omega(t_n)$ is approximated by $\Omega_h(t_n)$, where $\Omega_h(t) \rightarrow \Omega(t)$ as $h \rightarrow 0$, with h the radius of the smallest sphere completely containing the largest space-time element. The domain $\Omega_h(t_n)$ is divided into N_n non-overlapping spatial elements $K_j(t_n)$. Similarly, $\Omega(t_{n+1})$ is approximated by $\Omega_h(t_{n+1})$. We can relate each element $K_j^n = K_j(t_n)$ to a master element $\hat{K} \subset \mathbb{R}^q$ through the mapping F_K^n :

$$F_K^n : \hat{K} \rightarrow K_j^n : \bar{\xi} \mapsto \bar{x} = \sum_i x_i(K_j^n) \chi_i(\bar{\xi})$$

with x_i the spatial coordinates of the vertices of the spatial element K_j^n and χ_i the standard Lagrangian shape functions defined on element \hat{K} . The space-time elements \mathcal{K}_j^n are constructed by connecting K_j^n with K_j^{n+1} using linear interpolation in time, resulting in the mapping $G_{\mathcal{K}}^n$ from the master element $\hat{\mathcal{K}} \subset \mathbb{R}^{q+1}$ to the space-time element \mathcal{K}^n :

$$\begin{aligned} G_{\mathcal{K}}^n : \hat{\mathcal{K}} \rightarrow \mathcal{K}^n : \xi \mapsto (t, \bar{x}) &= \left(\frac{1}{2}(t_{n+1} + t_n) + \frac{1}{2}(t_{n+1} - t_n)\xi_0, \right. \\ &\left. \frac{1}{2}(1 - \xi_0)F_K^n(\bar{\xi}) + \frac{1}{2}(1 + \xi_0)F_K^{n+1}(\bar{\xi})\right). \end{aligned}$$

The tessellation \mathcal{T}_h^n of the space-time slab \mathcal{E}_h^n consists of all space-time elements \mathcal{K}_j^n ; thus the tessellation \mathcal{T}_h of the discrete flow domain $\mathcal{E}_h := \cup_{n=0}^{N_t-1} \mathcal{E}_h^n$ then is defined as $\mathcal{T}_h := \cup_{n=0}^{N_t-1} \mathcal{T}_h^n$.

The element boundary $\partial\mathcal{K}_j^n$, which is the union of open faces of \mathcal{K}_j^n , consists of three parts: $K_j(t_n^+) = \lim_{\epsilon \downarrow 0} K_j(t_n + \epsilon)$, $K_j(t_{n+1}^-) = \lim_{\epsilon \downarrow 0} K_j(t_{n+1} - \epsilon)$ and $\mathcal{Q}_j^n = \partial\mathcal{K}_j^n / (K_j(t_n^+) \cup K_j(t_{n+1}^-))$. Define the grid velocity $v \in \mathbb{R}^q$ as $v = \Delta\bar{x} / \Delta t$. The outward space-time normal vector at an element boundary point on $\partial\mathcal{K}_j^n$ is given by:

$$n = \begin{cases} (1, \bar{0}) & \text{at } K_j(t_{n+1}^-), \\ (-1, \bar{0}) & \text{at } K_j(t_n^+), \\ (-v_k \bar{n}_k, \bar{n}) & \text{at } \mathcal{Q}_j^n, \end{cases} \quad (2.11)$$

where $\bar{0} \in \mathbb{R}^q$. Note that since the space-time normal vector n has length one, the space component \bar{n} of the space-time normal has a length $|\bar{n}| = 1 / \sqrt{1 + v \cdot v}$. It can be convenient to split the element boundaries into separate faces. In addition to the faces $K_j(t_n^+)$ and $K_j(t_{n+1}^-)$, we also define therefore interior and boundary faces. An interior face is shared by two neighboring elements \mathcal{K}_i^n and \mathcal{K}_j^n , such that $\mathcal{S}_{ij}^n = \mathcal{Q}_i^n \cap \mathcal{Q}_j^n$, and a boundary face is defined as $\mathcal{S}_{Bj}^n = \partial\mathcal{E}^n \cap \mathcal{Q}_j^n$. The set of interior faces in time slab I^n is denoted by \mathcal{S}_I^n and the set of all boundary faces by \mathcal{S}_B^n . The total set of faces is denoted by $\mathcal{S}_{I,B}^n = \mathcal{S}_I^n \cup \mathcal{S}_B^n$.

2.2.2 Function spaces and trace operators

We consider approximations of $U(x, t)$ and functions $V(x, t)$ in the finite element space V_h , which is defined as:

$$V_h = \{V \in (L^2(\mathcal{E}_h))^m : V|_{\mathcal{K}} \circ G_{\mathcal{K}} \in (P^p(\hat{\mathcal{K}}))^m, \forall \mathcal{K} \in \mathcal{T}_h\},$$

where $L^2(\mathcal{E}_h)$ is the space of square integrable functions on \mathcal{E}_h and $P^p(\hat{\mathcal{K}})$ denotes the space of polynomials of degree at most p on the reference element $\hat{\mathcal{K}}$. Here m denotes the dimension of U .

We now introduce some operators as defined in Klaij et al. [41]. The trace of a function $f \in V_h$ at the element boundary $\partial\mathcal{K}^L$ is defined as:

$$f^L = \lim_{\epsilon \downarrow 0} f(x - \epsilon n^L),$$

with n^L the unit outward space-time normal at $\partial\mathcal{K}^L$. When only the space components of the outward normal vector are considered we will use the notation \bar{n}^L . A function $f \in V_h$ has a double valued trace at element boundaries $\partial\mathcal{K}$. The traces of a function f at an internal face $\mathcal{S} = \mathcal{K}^L \cap \mathcal{K}^R$ are denoted by f^L and f^R . The jump of f at an internal face $\mathcal{S} \in \mathcal{S}_I^n$ in the direction k of a Cartesian coordinate system is defined as:

$$[[f]]_k = f^L \bar{n}_k^L + f^R \bar{n}_k^R,$$

16 Chapter 2: DGFEM for hyperbolic nonconservative pdes: Theory

with $\bar{n}_k^R = -\bar{n}_k^L$. The average of f at $\mathcal{S} \in \mathcal{S}_T^n$ is defined as:

$$\{\!\{f\}\!\} = \frac{1}{2}(f^L + f^R).$$

The jump operator satisfies the following product rule at $\mathcal{S} \in \mathcal{S}_T^n$ for $\forall g \in V_h$ and $\forall f \in V_h$, which can be proven by direct verification:

$$\llbracket g_i f_{ik} \rrbracket_k = \{\!\{g_i\}\!\} \llbracket f_{ik} \rrbracket_k + \llbracket g_i \rrbracket_k \{\!\{f_{ik}\}\!\}. \quad (2.12)$$

Consequently, we can relate element boundary integrals to face integrals:

$$\sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{Q}} g_i^L f_{ik}^L \bar{n}_k^L d\mathcal{Q} = \sum_{\mathcal{S} \in \mathcal{S}_T^n} \int_{\mathcal{S}} \llbracket g_i f_{ik} \rrbracket_k d\mathcal{S} + \sum_{\mathcal{S} \in \mathcal{S}_B^n} \int_{\mathcal{S}} g_i^L f_{ik}^L \bar{n}_k^L d\mathcal{S}. \quad (2.13)$$

2.2.3 Weak formulation

In this section we derive a space-time DGFEM weak formulation for equations containing nonconservative products. Before discussing the space-time DGFEM weak formulation for equations containing nonconservative products, we first introduce as a reference the space-time DGFEM weak formulation for equations in conservative form (see, e.g., van der Vegt and van der Ven [79]).

Consider partial differential equations in conservative form:

$$U_{i,0} + H_{ik,k} = 0, \quad \bar{x} \in \mathbb{R}^q, \quad x_0 > 0, \quad (2.14)$$

where $U \in \mathbb{R}^m$ and $H \in \mathbb{R}^m \times \mathbb{R}^q$. Using the approach discussed in van der Vegt and van der Ven [79], the space-time DG formulation for (2.14) can be stated as:

Find a $U \in V_h$ such that for all $V \in V_h$:

$$\begin{aligned} 0 = & - \sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} \left(V_{i,0} U_i + V_{i,k} H_{ik} \right) d\mathcal{K} \\ & + \sum_{\mathcal{K} \in \mathcal{T}_h^n} \left(\int_{K(t_{n+1}^-)} V_i^L U_i^L dK - \int_{K(t_n^+)} V_i^L U_i^L dK \right) \\ & + \sum_{\mathcal{S} \in \mathcal{S}_T^n} \int_{\mathcal{S}} (V_i^L - V_i^R) \{\!\{H_{ik} - v_k U_i\}\!\} \bar{n}_k^L d\mathcal{S} + \sum_{\mathcal{S} \in \mathcal{S}_B^n} \int_{\mathcal{S}} V_i^L (H_{ik}^L - v_k U_i^L) \bar{n}_k^L d\mathcal{S}. \end{aligned} \quad (2.15)$$

Note that at this point no numerical fluxes have been introduced yet into the DG formulation. We continue now with equations containing nonconservative products. Let $U \in V_h$. We know that the numerical solution is continuous on an

element and discontinuous across a face, so, using Theorem 2.1.2, U is a weak solution to (2.4) if:

$$0 = \int_{\mathcal{E}_h} V_i d\mu_i \quad (2.16)$$

$$\begin{aligned} &= \sum_{\mathcal{K} \in \mathcal{T}_h} \int_{\mathcal{K}} V_i (U_{i,0} + D_{ikr} U_{r,k}) dK \\ &+ \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\int_{K(t_{n+1}^-)} \widehat{V}_i \left(\int_0^1 \delta_{ir} \frac{\partial \phi_r}{\partial \tau} (\tau; U_L, U_R) d\tau n_0^L \right) dK \right. \\ &\quad \left. + \int_{K(t_n^+)} \widehat{V}_i \left(\int_0^1 \delta_{ir} \frac{\partial \phi_r}{\partial \tau} (\tau; U_L, U_R) d\tau n_0^L \right) dK \right) \\ &+ \sum_{S \in \mathcal{S}_I} \int_S \widehat{V}_i \left(\int_0^1 D_{ikr} (\phi(\tau; U^L, U^R)) \frac{\partial \phi_r}{\partial \tau} (\tau; U^L, U^R) d\tau \bar{n}_k^L \right. \\ &\quad \left. + \int_0^1 \frac{\partial \phi_i}{\partial \tau} (\tau; U^L, U^R) d\tau n_0^L \right) dS \quad (2.17) \end{aligned}$$

$$\begin{aligned} &= \sum_{\mathcal{K} \in \mathcal{T}_h} \int_{\mathcal{K}} V_i (U_{i,0} + D_{ikr} U_{r,k}) dK \\ &+ \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\int_{K(t_{n+1}^-)} \widehat{V}_i (U_i^R - U_i^L) n_0^L dK + \int_{K(t_n^+)} \widehat{V}_i (U_i^R - U_i^L) n_0^L dK \right) \\ &+ \sum_{S \in \mathcal{S}_I} \int_S \widehat{V}_i \left(\int_0^1 D_{ikr} (\phi(\tau; U^L, U^R)) \frac{\partial \phi_r}{\partial \tau} (\tau; U^L, U^R) d\tau \bar{n}_k^L \right. \\ &\quad \left. - v_k \delta_{ir} \int_0^1 \frac{\partial \phi_r}{\partial \tau} (\tau; U^L, U^R) d\tau \bar{n}_k^L \right) dS \quad (2.18) \end{aligned}$$

$$\begin{aligned} &= \sum_{\mathcal{K} \in \mathcal{T}_h} \int_{\mathcal{K}} V_i (U_{i,0} + D_{ikr} U_{r,k}) dK \\ &+ \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\int_{K(t_{n+1}^-)} \widehat{V}_i (U_i^R - U_i^L) dK - \int_{K(t_n^+)} \widehat{V}_i (U_i^R - U_i^L) dK \right) \\ &+ \sum_{S \in \mathcal{S}_I} \int_S \widehat{V}_i \left(\int_0^1 D_{ikr} (\phi(\tau; U^L, U^R)) \frac{\partial \phi_r}{\partial \tau} (\tau; U^L, U^R) d\tau \bar{n}_k^L \right) dS \\ &+ \sum_{S \in \mathcal{S}_I} \int_S \widehat{V}_i \llbracket v_k U_i \rrbracket_k dS, \quad (2.19) \end{aligned}$$

where $V \in V_h$ is an arbitrary test function. Furthermore, \widehat{V} is the value (numerical flux) of the test function V on a face S and δ represents the Kronecker

18 Chapter 2: DGFEM for hyperbolic nonconservative pdes: Theory

delta symbol. In (2.19) we used the definition of n_0^L as given in (2.11). The crucial point in obtaining the DG formulation is the choice of the numerical flux for the test function V . Using $D_{ikr} = \partial F_{ik}/\partial U_r + G_{ikr}$, (2.19) can be rewritten as:

$$\begin{aligned}
 0 = & \sum_{\mathcal{K} \in \mathcal{T}_h} \int_{\mathcal{K}} V_i (U_{i,0} + F_{ik,k} + G_{ikr} U_{r,k}) d\mathcal{K} \\
 & + \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\int_{K(t_{n+1}^-)} \widehat{V}_i (U_i^R - U_i^L) dK - \int_{K(t_n^+)} \widehat{V}_i (U_i^R - U_i^L) dK \right) \\
 & + \sum_{\mathcal{S} \in \mathcal{S}_I} \int_{\mathcal{S}} \widehat{V}_i \left(\int_0^1 G_{ikr}(\phi(\tau; U^L, U^R)) \frac{\partial \phi_r}{\partial \tau}(\tau; U^L, U^R) d\tau \bar{n}_k^L \right) d\mathcal{S} \\
 & - \sum_{\mathcal{S} \in \mathcal{S}_I} \int_{\mathcal{S}} \widehat{V}_i \llbracket F_{ik} - v_k U_i \rrbracket_k d\mathcal{S}.
 \end{aligned} \tag{2.20}$$

We choose the numerical flux for V such that if there exists a Q , with $G_{ikr} = \partial Q_{ik}/\partial U_r$, then the DG formulation for the system containing nonconservative products reduces to the conservative space-time DGFEM weak formulation given by (2.15) with $H_{ik} = F_{ik} + Q_{ik}$.

Theorem 2.2.1. *If the numerical flux \widehat{V} for the test function V in (2.20) is defined as:*

$$\widehat{V} = \begin{cases} \llbracket V \rrbracket & \text{at } \mathcal{S} \in \mathcal{S}_I, \\ 0 & \text{at } K(t_n) \subset \Omega_h(t_n) \quad \forall n, \end{cases} \tag{2.21}$$

then the DG formulation (2.20) will reduce to the conservative space-time DGFEM formulation (2.15) when there exists a Q such that $G_{ikr} = \partial Q_{ik}/\partial U_r$ so that $H_{ik} = F_{ik} + Q_{ik}$.

Proof Assume there is a Q , such that $G_{ikr} = \partial Q_{ik}/\partial U_r$. We immediately see that:

$$\int_0^1 G_{ikr}(\phi(\tau; U^L, U^R)) \frac{\partial \phi_r}{\partial \tau}(\tau; U^L, U^R) d\tau \bar{n}_k^L = -\llbracket Q_{ik} \rrbracket_k. \tag{2.22}$$

Integrating by parts the volume integral in (2.20) and using (2.22) we obtain:

$$\begin{aligned}
0 = & - \sum_{\mathcal{K} \in \mathcal{T}_h} \int_{\mathcal{K}} (V_{i,0} U_i + V_{i,k} (F_{ik} + Q_{ik})) d\mathcal{K} \\
& + \sum_{\mathcal{K} \in \mathcal{T}_h} \int_{\partial \mathcal{K}} V_i^L (U_i^L n_0^L + (F_{ik}^L + Q_{ik}^L) \bar{n}_k^L) d(\partial \mathcal{K}) \\
& + \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\int_{K(t_{n+1}^-)} \widehat{V}_i (U_i^R - U_i^L) dK - \int_{K(t_n^+)} \widehat{V}_i (U_i^R - U_i^L) dK \right) \\
& - \sum_{S \in \mathcal{S}_I} \int_S \widehat{V}_i [F_{ik} + Q_{ik} - v_k U_i]_k dS.
\end{aligned} \tag{2.23}$$

We write $H_{ik} = F_{ik} + Q_{ik}$. Using the definition of the normal vector (2.11), the element boundary integral in (2.23) becomes:

$$\begin{aligned}
\sum_{\mathcal{K} \in \mathcal{T}_h} \int_{\partial \mathcal{K}} V_i^L (U_i^L n_0^L + H_{ik}^L \bar{n}_k^L) d(\partial \mathcal{K}) &= \sum_{\mathcal{K} \in \mathcal{T}_h} \int_{\mathcal{Q}} V_i^L (H_{ik}^L - v_k U_i^L) \bar{n}_k^L d\mathcal{Q} \\
&+ \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\int_{K(t_{n+1}^-)} V_i^L U_i^L dK - \int_{K(t_n^+)} V_i^L U_i^L dK \right).
\end{aligned} \tag{2.24}$$

We will now use relations (2.12) and (2.13) to write the element boundary integrals as face integrals:

$$\begin{aligned}
& \sum_{\mathcal{K} \in \mathcal{T}_h} \int_{\mathcal{Q}} V_i^L (H_{ik}^L - v_k U_i^L) \bar{n}_k^L d\mathcal{Q} \\
= & \sum_{S \in \mathcal{S}_I} \int_S \llbracket V_i (H_{ik} - v_k U_i) \rrbracket_k dS + \sum_{S \in \mathcal{S}_B} \int_S V_i^L (H_{ik}^L - v_k U_i^L) \bar{n}_k^L dS \\
= & \sum_{S \in \mathcal{S}_I} \int_S \left(\{V_i\} [H_{ik} - v_k U_i]_k + (V_i^L - V_i^R) \{H_{ik} - v_k U_i\} \bar{n}_k^L \right) dS \\
& + \sum_{S \in \mathcal{S}_B} \int_S V_i^L (H_{ik}^L - v_k U_i^L) \bar{n}_k^L dS.
\end{aligned} \tag{2.25}$$

20 Chapter 2: DGFEM for hyperbolic nonconservative pdes: Theory

Combining (2.23), (2.24) and (2.25) we obtain:

$$\begin{aligned}
0 = & - \sum_{\mathcal{K} \in \mathcal{T}_h} \int_{\mathcal{K}} (V_{i,0} U_i + V_{i,k} H_{ik}) d\mathcal{K} \\
& + \sum_{K \in \mathcal{T}_h} \left(\int_{K(t_{n+1}^-)} V_i^L U_i^L dK - \int_{K(t_n^+)} V_i^L U_i^L dK \right) \\
& + \sum_{K \in \mathcal{T}_h} \left(\int_{K(t_{n+1}^-)} \widehat{V}_i (U_i^R - U_i^L) dK - \int_{K(t_n^+)} \widehat{V}_i (U_i^R - U_i^L) dK \right) \quad (2.26) \\
& + \sum_{S \in \mathcal{S}_I} \int_S (\{V_i\} [H_{ik} - v_k U_i]_k + (V_i^L - V_i^R) \{H_{ik} - v_k U_i\} \bar{n}_k^L) dS \\
& - \sum_{S \in \mathcal{S}_I} \int_S \widehat{V}_i [H_{ik} - v_k U_i]_k dS + \sum_{S \in \mathcal{S}_B} \int_S V_i^L (H_{ik}^L - v_k U_i^L) \bar{n}_k^L dS.
\end{aligned}$$

The term $\{V_i\} [H_{ik} - v_k U_i]_k$ is set to zero in the space-time DG formulation for conservative systems by arguing that the formulation must be conservative. For a general nonconservative system we can not use this argument. Instead, we note that by taking $\widehat{V} = \{V\}$ on the faces $S \in \mathcal{S}_I$, the contribution $\int_S \{V_i\} [H_{ik} - v_k U_i]_k dS$ cancels with $-\int_S \widehat{V}_i [H_{ik} - v_k U_i]_k dS$. Furthermore, taking $\widehat{V} = 0$ on the time faces $K(t_n) \subset \Omega_h(t_n) \forall n$, we obtain the space-time DGFEM weak formulation for conservative systems given by (2.15). \square

Theorem 2.2.1 allows us to finalize the derivation of the DGFEM formulation for hyperbolic nonconservative partial differential equations. First, we start with

the volume integral of (2.20) and integrate by parts, to obtain:

$$\begin{aligned}
0 &= \sum_{\mathcal{K} \in \mathcal{T}_h} \int_{\mathcal{K}} (-V_{i,0} U_i - V_{i,k} F_{ik} + V_i G_{ikr} U_{r,k}) d\mathcal{K} \\
&+ \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\int_{K(t_{n+1}^-)} V_i^L U_i^L dK - \int_{K(t_n^+)} V_i^L U_i^L dK \right) \\
&+ \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\int_{K(t_{n+1}^-)} \widehat{V}_i (U_i^R - U_i^L) dK - \int_{K(t_n^+)} \widehat{V}_i (U_i^R - U_i^L) dK \right) \\
&+ \sum_{S \in \mathcal{S}_I} \int_S (\{V_i\} [F_{ik} - v_k U_i]_k + (V_i^L - V_i^R) \{F_{ik} - v_k U_i\} \bar{n}_k^L) d\mathcal{S} \quad (2.27) \\
&+ \sum_{S \in \mathcal{S}_B} \int_S V_i^L (F_{ik}^L - v_k U_i^L) \bar{n}_k^L d\mathcal{S} \\
&+ \sum_{S \in \mathcal{S}_I} \int_S \widehat{V}_i \left(\int_0^1 G_{ikr}(\phi(\tau; U^L, U^R)) \frac{\partial \phi_r}{\partial \tau}(\tau; U^L, U^R) d\tau \bar{n}_k^L \right) d\mathcal{S} \\
&- \sum_{S \in \mathcal{S}_I} \int_S \widehat{V}_i [F_{ik} - v_k U_i]_k d\mathcal{S},
\end{aligned}$$

where we used relation (2.11) for the time component of the space-time normal vector and relations (2.12) and (2.13) to write the element boundary integrals as face integrals. For the numerical flux for the test function V in (2.27) we use (2.21), and thus obtain:

$$\begin{aligned}
0 &= \sum_{\mathcal{K} \in \mathcal{T}_h} \int_{\mathcal{K}} (-V_{i,0} U_i - V_{i,k} F_{ik} + V_i G_{ikr} U_{r,k}) d\mathcal{K} \\
&+ \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\int_{K(t_{n+1}^-)} V_i^L U_i^L dK - \int_{K(t_n^+)} V_i^L U_i^L dK \right) \\
&+ \sum_{S \in \mathcal{S}_I} \int_S (V_i^L - V_i^R) \{F_{ik} - v_k U_i\} \bar{n}_k^L d\mathcal{S} \\
&+ \sum_{S \in \mathcal{S}_B} \int_S V_i^L (F_{ik}^L - v_k U_i^L) \bar{n}_k^L d\mathcal{S} \\
&+ \sum_{S \in \mathcal{S}_I} \int_S \{V_i\} \left(\int_0^1 G_{ikr}(\phi(\tau; U^L, U^R)) \frac{\partial \phi_r}{\partial \tau}(\tau; U^L, U^R) d\tau \bar{n}_k^L \right) d\mathcal{S}. \quad (2.28)
\end{aligned}$$

Theorem 2.2.1 states that the weak formulation given by (2.28) can be reduced to the space-time DGFEM formulation (2.15), when a Q exists such

22 Chapter 2: DGFEM for hyperbolic nonconservative pdes: Theory

that $G_{ikr} = \partial Q_{ik} / \partial U_r$. However, this formulation is generally numerically unstable. Problematic in the conservative space-time DGFEM formulation are the interior $(V_i^L - V_i^R) \{ \{ H_{ik} - v_k U_i \} \bar{n}_k^L$ and boundary $V_i^L (H_{ik}^L - v_k U_i^L) \bar{n}_k^L$ flux terms, see (2.15). Generally, a stabilizing term is added to these flux terms, together forming an upwind numerical flux. Furthermore, the following upwind flux is introduced in the conservative space-time DGFEM formulation at the time faces, a formulation naturally ensuring causality in time:

$$\widehat{U} = \begin{cases} U^L & \text{at } K(t_{n+1}^-) \\ U^R & \text{at } K(t_n^+) \end{cases}. \quad (2.29)$$

It replaces the traces of U taken from the interior of $\mathcal{K} \in \mathcal{T}_h^n$. In (2.28), we also introduce the upwind flux (2.29) at the time faces. We also need a stabilizing term in (2.28). To understand how we add our stabilizing term, consider again the conservative space-time formulation. As mentioned above, a stabilizing term is added to $\{ \{ H_{ik} - v_k U_i \} \}$. Denote this stabilizing term as H^{stab} , then $(\{ \{ H_{ik} - v_k U_i \} \} + H_{ik}^{stab}) \bar{n}_k^L = \widehat{H}_i$, where \widehat{H}_i is the space-time numerical flux. In the nonconservative space-time formulation (2.28) we add a stabilizing term to the conservative part $\{ \{ F_{ik} - v_k U_i \} \}$, but we also need to add a stabilizing part due to the nonconservative product. For the nonconservative product there is no counterpart for $\{ \{ F_{ik} - v_k U_i \} \}$. This term is hidden in the volume integral and in the last term of (2.28). We add the stabilizing term for the nonconservative product P_{ik}^{nc} to the stabilizing term for the conservative product P_{ik}^c : $(\{ \{ F_{ik} - v_k U_i \} \} + P_{ik}^c + P_{ik}^{nc}) \bar{n}_k^L = \widehat{P}_i^{nc}$. By introducing a ghost value U^R at the boundary, we can use the same expressions also at a boundary face. An expression for $\widehat{P}_i^{nc}(U^L, U^R, v, \bar{n}^L)$ is derived in Section 2.3, such that it reduces to the numerical flux in the conservative case, \widehat{H}_i . Finally, the space-time DGFEM weak formulation for partial differential equations containing nonconservative products (2.3) is:

Find a $U \in V_h$ such that for all $V \in V_h$:

$$\begin{aligned} 0 &= \sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} (-V_{i,0} U_i - V_{i,k} F_{ik} + V_i G_{ikr} U_{r,k}) d\mathcal{K} \\ &+ \sum_{\mathcal{K} \in \mathcal{T}_h^n} \left(\int_{K(t_{n+1}^-)} V_i^L U_i^L dK - \int_{K(t_n^+)} V_i^L U_i^R dK \right) \\ &+ \sum_{S \in \mathcal{S}^n} \int_S (V_i^L - V_i^R) \widehat{P}_i^{nc} dS \\ &+ \sum_{S \in \mathcal{S}^n} \int_S \{ \{ V_i \} \} \left(\int_0^1 G_{ikr}(\phi(\tau; U^L, U^R)) \frac{\partial \phi_r}{\partial \tau}(\tau; U^L, U^R) d\tau \bar{n}_k^L \right) dS, \end{aligned} \quad (2.30)$$

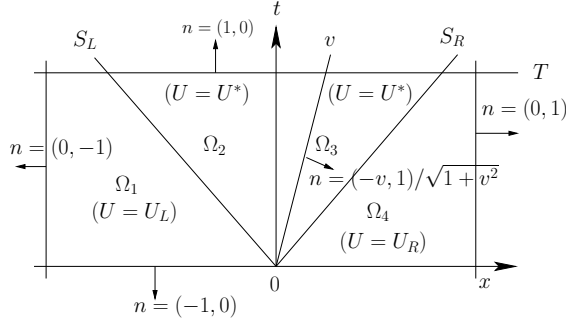


Figure 2.1: *Wave pattern of the solution for the Riemann problem. Here S_L and S_R are the fastest left and right moving signal velocities and v is the velocity of the element boundary point.*

Note that due to the introduction of the upwind flux at the time faces, each space-time slab only depends on the previous space-time slab so that the summation over all space-time slabs could be dropped.

2.3 The NCP numerical flux

In Section 2.2 we derived a weak formulation for space-time DGFEM for systems of equations containing a nonconservative product. To obtain an expression for the flux $\hat{P}_i^{nc}(U^L, U^R, v, \bar{n}^L)$ in (2.30), we first discuss the numerical flux \hat{U} , and then derive the numerical flux for NonConservative Products, or NCP-flux.

Consider the following nonconservative hyperbolic system:

$$\partial_t U + \partial_x F(U) + G(U) \partial_x U = 0, \quad (2.31)$$

where $U \in \mathbb{R}^m$, with m the number of components of U , similarly $F(U) \in \mathbb{R}^m$, $G(U) \in \mathbb{R}^{m \times m}$ and $x \in \mathbb{R}$ is along the normal of the face. To approximate the Riemann solution of (2.31) we consider only the fastest left and right moving waves of the system with velocities S_L and S_R and the grid velocity. In the star region (see Figure 2.1), which is the domain enclosed by the waves S_L and S_R , the averaged exact solution \bar{U}^* is defined as:

$$\bar{U}^* = \frac{1}{T(S_R - S_L)} \int_{TS_L}^{TS_R} U(x, T) dx. \quad (2.32)$$

In what follows we obtain a relation for \bar{U}^* from the weak formulation of (2.31).

24 Chapter 2: DGFEM for hyperbolic nonconservative pdes: Theory

Using Gauss' theorem we obtain over the control volume $\Omega_1 \cup \Omega_2$ the relation:

$$\begin{aligned} \int_{x_L}^{S_L T} U_L dx + \int_{S_L T}^{vT} U(x, T) dx &= \int_{x_L}^0 U_L dx + \\ \int_0^T F_L dt - \int_0^T (F(U(vt, t)) - vU(vt, t)) dt &- \int_{\Omega_2} G(U) \partial_x U dx dt \\ - \int_0^T \int_0^1 G(\phi_{LL^*}(\tau; U_L, U_L^*)) \frac{\partial \phi_{LL^*}}{\partial \tau}(\tau; U_L, U_L^*) d\tau dt, \end{aligned} \quad (2.33)$$

where $F_L = F(U_L)$ and $U_L^* = \lim_{s \downarrow S_L} U^*(st, t)$ is the trace of U^* taken from the interior of Ω_2 , which is constant along the wave S_L due to the self similarity of the solution in the star region. Replace the exact integrand in the second integral on the left hand side of (2.33) with the approximate solution \bar{U}^* . Furthermore, using the self similarity of the solution in the star region [54], we obtain:

$$\begin{aligned} \int_{\Omega_2} G(U) \partial_x U dx dt &= \int_{t=0}^T \int_{x=S_L t}^{vt} G(U) \partial_x U dx dt \\ &= \int_{t=0}^T \int_{S_L}^v G(U^*) \partial_s U^* \partial_{xs} |J| ds dt \quad (2.34) \\ &= T \int_{S_L}^v G(U^*) \partial_s U^* ds, \end{aligned}$$

where we used the coordinate transformation $x = st$, $t = t$, which has a Jacobian $|J| = t$. Introduce the trace of U^* taken from the interior of Ω_2 along the line $x = vt$ as: $U_{Lv}^* = \lim_{s \uparrow v} U^*(st, t)$ and the path $\phi_{L^*v} : [0, 1] \times \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ with:

$$\phi_{L^*v}(\tau; U_L^*, U_{Lv}^*) = U^*(s), \quad \text{if } S_L < s < v.$$

By connecting these two paths into the path $\phi_{Lv} : [0, 1] \times \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$, such that $\phi_{Lv}(\tau; U_L, U_{Lv}^*) = \phi_{LL^*} \cup \phi_{L^*v}$, redefining τ and using (2.34), the integral contributions due to the nonconservative product on the righthand side of (2.33) can be combined, resulting in:

$$S_L U_L + (v - S_L) \bar{U}^* = F_L - F^v - \int_0^1 G(\phi_{Lv}(\tau; U_L, U_{Lv}^*)) \frac{\partial \phi_{Lv}}{\partial \tau}(\tau; U_L, U_{Lv}^*) d\tau, \quad (2.35)$$

where $F^v = F(U(vt, t)) - vU(vt, t)$ which is constant along $x = vt$. Similarly,

using Gauss' theorem for the control volume $\Omega_3 \cup \Omega_4$ yields:

$$\begin{aligned} \int_{vT}^{S_RT} U(x, T) dx + \int_{S_RT}^{x_R} U_R dx &= \int_0^{x_R} U_R dx - \\ \int_0^T F_R dt + \int_0^T (F(U(vt, t)) - vU(vt, t)) dt &- \int_{\Omega_3} G(U) \partial_x U dx dt - \\ \int_0^T \int_0^1 G(\phi_{R^*R}(\tau; U_R^*, U_R)) \frac{\partial \phi_{R^*R}}{\partial \tau}(\tau; U_R^*, U_R) d\tau dt, &\quad (2.36) \end{aligned}$$

where $F_R = F(U_R)$ and $U_R^* = \lim_{s \uparrow S_R} U^*(st, t)$ is the trace of U^* taken from the interior of Ω_3 , which is constant along the wave S_R . Furthermore, denote the trace of U^* taken from the interior of Ω_3 along the line $x = vt$ as: $U_{Rv}^* = \lim_{s \downarrow v} U^*(st, t)$. Replace the exact integrand in the first integral on the left hand side of (2.36) with the average of the exact solution \bar{U}^* . Introduce the path $\phi_{vR^*} : [0, 1] \times \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ with:

$$\phi_{vR^*}(\tau; U_{Rv}^*, U_R^*) = U^*(s), \quad \text{if } v < s < S_R,$$

and the path $\phi_{vR} : [0, 1] \times \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ such that $\phi_{vR}(\tau; U_{Rv}^*, U_R) = \phi_{R^*R} \cup \phi_{vR^*}$ after redefining τ . Using the self similarity of the solution in the star region Ω_3 , similar to (2.34), the integral contributions on the righthand side of (2.36) can be combined, resulting in:

$$(S_R - v)\bar{U}^* - S_R U_R = F^v - F_R - \int_0^1 G(\phi_{vR}(\tau; U_{Rv}^*, U_R)) \frac{\partial \phi_{vR}}{\partial \tau}(\tau; U_{Rv}^*, U_R) d\tau. \quad (2.37)$$

Note that $U_{Lv}^* = U_{Rv}^*$ since the solution U is smooth across $\partial\bar{\Omega}_2 \cap \partial\bar{\Omega}_3$, where $\bar{\Omega}_2$ and $\bar{\Omega}_3$ are the closures of Ω_2 and Ω_3 . Now, introduce the path $\bar{\phi} : [0, 1] \times \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ (see Figure 2.2) and redefine τ such that $\bar{\phi}(\tau; U_L, U_R) = \phi_{Lv} \cup \phi_{vR}$ then, by adding (2.35) and (2.37) and rearranging terms, we obtain:

$$\begin{aligned} \bar{U}^* &= \frac{S_R U_R - S_L U_L + F_L - F_R}{S_R - S_L} - \\ &\quad \frac{1}{S_R - S_L} \int_0^1 G(\bar{\phi}(\tau; U_L, U_R)) \frac{\partial \bar{\phi}}{\partial \tau}(\tau; U_L, U_R) d\tau. \quad (2.38) \end{aligned}$$

This equation is still exact if we would know the path $\bar{\phi}$. Note from Figure 2.1 that outside the star region the solution is still at its initial values at $t = 0$, denoted by U_L and U_R . Within the star region bounded by the slowest and fastest signal speed S_L and S_R , respectively, an averaged star state solution \bar{U}^*

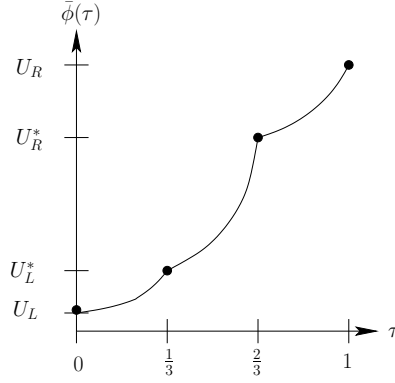


Figure 2.2: Combining the paths to form $\bar{\phi}_{LR}(\tau; U_L, U_R) = \bar{\phi}_{LL^*} \cup \bar{\phi}_{L^*v} \cup \bar{\phi}_{vR^*} \cup \bar{\phi}_{R^*R}$.

is assumed. We define the numerical flux for U as:

$$\hat{U} = \begin{cases} U_L, & \text{if } v \leq S_L, \\ \bar{U}^*, & \text{if } S_L < v < S_R, \\ U_R, & \text{if } v \geq S_R, \end{cases}$$

where the averaged star state solution \bar{U}^* is given by (2.38) and v is the velocity of the element boundary point.

We now continue to derive an expression for $\hat{P}^{nc}(U_L, U_R, v, \bar{n}^L)$. Define

$$\int_0^\tau G(\bar{\phi}(\tilde{\tau}; U_1, U_2)) \frac{\partial \bar{\phi}}{\partial \tilde{\tau}}(\tilde{\tau}; U_1, U_2) d\tilde{\tau} \equiv \int_0^\tau d\mathcal{G}(\bar{\phi}(\tilde{\tau}; U_1, U_2)),$$

so that:

$$\int_0^1 G(\bar{\phi}(\tilde{\tau}; U_1, U_2)) \frac{\partial \bar{\phi}}{\partial \tilde{\tau}}(\tilde{\tau}; U_1, U_2) d\tilde{\tau} = \mathcal{G}(U_2) - \mathcal{G}(U_1),$$

using conditions H1-H4. Denote $\mathcal{G}(U_k) = \mathcal{G}_k$ and introduce $\tilde{\mathcal{G}}_k = \mathcal{G}_k - \{\!\!\{ \mathcal{G} \}\!\!\}$, for $k = 1, 2$ with $\{\!\!\{ \mathcal{G} \}\!\!\} = (\mathcal{G}_1 + \mathcal{G}_2)/2$. Note that $\mathcal{G}_2 - \mathcal{G}_1 = \tilde{\mathcal{G}}_2 - \tilde{\mathcal{G}}_1$. From (2.35) and (2.37), the definition of the paths, conditions H1-H4 and assuming $U_{L^*v}^* = U_{R^*v}^* = \bar{U}^*$, we then obtain:

$$S_L U_L + (v - S_L) \bar{U}^* = F_L - F^v - \tilde{\mathcal{G}}^* + \tilde{\mathcal{G}}_L, \quad (2.39)$$

and:

$$(S_R - v) \bar{U}^* - S_R U_R = F^v - F_R - \tilde{\mathcal{G}}_R + \tilde{\mathcal{G}}^*, \quad (2.40)$$

where $\mathcal{G}_L = \mathcal{G}(U_L)$, $\mathcal{G}_R = \mathcal{G}(U_R)$ and $\mathcal{G}^* = \mathcal{G}(\bar{U}^*)$. Subtracting (2.40) from (2.39) and rearranging the terms, we obtain:

$$F^v + \tilde{\mathcal{G}}^* = \{\{\tilde{\mathcal{G}}\}\} + \{\{F\}\} + \frac{1}{2}((S_R - v)\bar{U}^* + (S_L - v)\bar{U}^* - S_L U_L - S_R U_R),$$

with $\{\{\tilde{\mathcal{G}}\}\} \equiv (\tilde{\mathcal{G}}_L + \tilde{\mathcal{G}}_R)/2 = 0$. Similarly, by adding (2.39) and (2.40) together and rearranging terms, we obtain:

$$F_L + \tilde{\mathcal{G}}_L = F_L - \frac{1}{2} \int_0^1 G(\bar{\phi}(\tau; U_L, U_R)) \frac{\partial \phi}{\partial \tau}(\tau; U_L, U_R) d\tau,$$

and:

$$F_R + \tilde{\mathcal{G}}_R = F_R + \frac{1}{2} \int_0^1 G(\bar{\phi}(\tau; U_L, U_R)) \frac{\partial \phi}{\partial \tau}(\tau; U_L, U_R) d\tau.$$

The NCP numerical flux $\hat{P}^{nc}(U_L, U_R, v, \bar{n}^L)$ is defined in Ω_1 as $F_L + \tilde{\mathcal{G}}_L$, in $\Omega_2 \cup \Omega_3$ as $F^v + \tilde{\mathcal{G}}^*$ and in Ω_4 as $F_R + \tilde{\mathcal{G}}_R$ (see also (2.30)). The NCP-flux is thus given by:

$$\hat{P}_i^{nc}(U_L, U_R, v, \bar{n}^L) = \begin{cases} F_{ik}^L \bar{n}_k^L - \frac{1}{2} \int_0^1 G_{ikr}(\bar{\phi}(\tau; U_L, U_R)) \frac{\partial \bar{\phi}_r}{\partial \tau}(\tau; U_L, U_R) d\tau \bar{n}_k^L & \text{if } S_L > v, \\ \{\{F_{ik}\}\} \bar{n}_k^L + \frac{1}{2}((S_R - v)\bar{U}_i^* + (S_L - v)\bar{U}_i^* - S_L U_i^L - S_R U_i^R) & \text{if } S_L < v < S_R, \\ F_{ik}^R \bar{n}_k^L + \frac{1}{2} \int_0^1 G_{ikr}(\bar{\phi}(\tau; U_L, U_R)) \frac{\partial \bar{\phi}_r}{\partial \tau}(\tau; U_L, U_R) d\tau \bar{n}_k^L & \text{if } S_R < v, \end{cases} \quad (2.41)$$

with \bar{U}^* given by (2.38). Note that if G is the Jacobian of some flux function Q , then $\hat{P}^{nc}(U_L, U_R, v, \bar{n}^L)$ is exactly the HLL flux derived for moving grids in van der Vegt and van der Ven [79].

Discontinuous Galerkin finite element methods for hyperbolic nonconservative partial differential equations: Applications

In the previous chapter and in Appendix A we presented space- and space-time discontinuous Galerkin finite element (DGFEM) formulations for systems containing nonconservative products, in which we introduced the theory of weak solutions for nonconservative products into the DGFEM formulation. This leads to the new question how to define the path connecting left and right states across a discontinuity. The effect of different paths on the numerical solution is investigated in this chapter and found to be small. We furthermore apply our scheme to two different systems of partial differential equations. We consider the shallow water equations, where topography leads to nonconservative products, in which the known, possibly discontinuous, topography is formally taken as an unknown in the system. We also consider a simplification of a depth-averaged two-phase flow model which contains more intrinsic nonconservative products.

3.1 Obtaining and solving the discrete system

By replacing the trial function U and the test function V in the DGFEM weak formulation by their polynomial approximations, a system of algebraic equations is obtained. Solving this system for the expansion coefficients of the trial function U results in the DGFEM solution of our problem. Depending on whether

we are solving the DGFEM weak formulation in space-time or in space, different basis functions and solving algorithms are used, which we discuss next.

3.1.1 Space-time DGFEM basis functions and solving method

Polynomial approximations for the trial function U and the test functions V in each element $\mathcal{K} \in \mathcal{T}_h^n$ are introduced as:

$$U(t, \bar{x})|_{\mathcal{K}} = \hat{U}_m \psi_m(t, \bar{x}) \text{ and } V(t, \bar{x})|_{\mathcal{K}} = \hat{V}_l \psi_l(t, \bar{x}), \quad (3.1)$$

with ψ_m the basis functions, $\bar{x} \in \mathbb{R}^q$, and expansion coefficients \hat{U}_m and \hat{V}_l , respectively, for $m, l = 0, 1, 2, \dots, N$, where N depends on the order of accuracy and the space dimension q . In this chapter the basis functions are defined such that the test and trial functions can be split into an element mean at time t_{n+1} and a fluctuating part. The basis functions ψ_m are given by:

$$\psi_m = \begin{cases} 1, & \text{for } m = 0 \\ \varphi_m(t, \bar{x}) - \frac{1}{|K_j(t_{n+1}^-)|} \int_{K_j(t_{n+1}^-)} \varphi_m(t, \bar{x}) dK & \text{for } m = 1, 2, \dots, N, \end{cases}$$

where the functions $\varphi_m(x)$ in element \mathcal{K} are related to the basis functions $\hat{\varphi}_m(\xi)$, with $\hat{\varphi}_m(\xi) \in P^p(\hat{\mathcal{K}})$ and ξ the local coordinates in the master element $\hat{\mathcal{K}}$, through the mapping $G_{\mathcal{K}}$:

$$\varphi_m = \hat{\varphi}_m \circ G_{\mathcal{K}}^{-1}.$$

By replacing U and V in the weak formulation (2.30) by their polynomial expansions (3.1), a system of algebraic equations for the expansion coefficients of U is obtained. For each physical time step, the system can be written as:

$$\mathcal{L}(\hat{U}^n; \hat{U}^{n-1}) = 0. \quad (3.2)$$

This system of coupled non-linear equations is solved in this chapter by adding a pseudo-time derivative:

$$|K^n| \frac{\partial \hat{U}^n}{\partial \tau} = -\frac{1}{\Delta t} \mathcal{L}(\hat{U}^n; \hat{U}^{n-1}), \quad (3.3)$$

which is integrated to steady-state in pseudo-time. Following van der Vegt and van der Ven [79] and Klaij et al. [40], we use the explicit Runge-Kutta method for inviscid flow with Melson correction which is given by:

Algorithm 1 Five-stage explicit Runge-Kutta scheme:

1. Initialize $\hat{V}^0 = \hat{U}$.

2. For all stages $s = 1$ to 5:

$$(I + \alpha_s \lambda I) \hat{V}^s = \hat{V}^0 + \alpha_s \lambda (\hat{V}^{s-1} - \mathcal{L}(\hat{V}^{s-1}, \hat{U}^{n-1}) / |K^n|).$$

3. Update $\hat{V} = \hat{V}^5$.

The coefficient λ is defined as $\lambda = \Delta\tau / \Delta t$, with $\Delta\tau$ the pseudo-time step and Δt the physical time step. The Runge-Kutta coefficients α_s are defined as: $\alpha_1 = 0.0797151$, $\alpha_2 = 0.163551$, $\alpha_3 = 0.283663$, $\alpha_4 = 0.5$ and $\alpha_5 = 1.0$.

3.1.2 Space DGFEM basis functions and solving method

Polynomial approximations for the trial function U and the test function V in each element K_j are introduced:

$$U(t, \bar{x})|_{K_j} = \hat{U}_m \psi_m(\bar{x}), \text{ and } V(t, \bar{x})|_{K_j} = \hat{V}_l \psi_l(\bar{x}) \quad (3.4)$$

for $m, l = 0, 1, 2, \dots, M$, where M depends on the order of accuracy and the space dimension, and where the basis functions ψ , in this chapter, are given by:

$$\psi_m = \begin{cases} 1 & \text{for } m = 0 \\ \varphi_m(\bar{x}) - \frac{1}{|K_j|} \int_{K_j} \varphi_m(\bar{x}) dK & \text{for } m = 1, 2, \dots, M. \end{cases}$$

The functions $\varphi_m(\bar{x})$ in element K_j are related to the basis functions $\hat{\varphi}_m(\xi)$ on the master element \hat{K} through the mapping F :

$$\varphi_m = \hat{\varphi}_m \circ F_K^{-1}$$

with $\hat{\varphi}_m(\xi) \in P^p(\hat{K})$ and ξ the local coordinates in the master element \hat{K} . By replacing U and V in the weak formulation (A.11) by their polynomial expansions (3.4), we arrive at the following system of ordinary differential equations for the expansion coefficients \hat{U} of the variables U :

$$M \frac{d\hat{U}}{dt} = \mathcal{L}^{DG}(\hat{U}), \quad (3.5)$$

with M the element mass matrices defined as $M_{ij} = \int_{K_j} \psi_i \psi_j dK$ and $\mathcal{L}^{DG}(\hat{U})$ the space part of the weak formulation with U and V replaced by their polynomial expansions. In this chapter, to solve this system of ordinary differential equations, we use an explicit TVD third order Runge-Kutta method (see e.g. Gottlieb and Shu [25]).

3.1.3 Slope limiters

In our space- and space-time DGFEM computations, when the solution may admit discontinuities, we use a slope limiter to deal with overshoots and undershoots. In this chapter we use a simple minmod function (see e.g. Cockburn and Shu [19]). Let \bar{U}_k represent the mean of U on element \mathcal{K}_k and let \hat{U}_k represent the slope, then the solution in an element is given by:

$$U_k = \bar{U}_k + \psi(x)m(\hat{U}_k, \bar{U}_{k+1} - \bar{U}_k, \bar{U}_k - \bar{U}_{k-1}),$$

where the minmod function m is defined as:

$$m(a_1, a_2, a_3) = \begin{cases} s \min_{1 \leq n \leq 3} |a_n| & \text{if } s = \text{sign}(a_1) = \text{sign}(a_2) = \text{sign}(a_3) \\ 0 & \text{otherwise.} \end{cases}$$

3.2 One dimensional test cases

3.2.1 The one dimensional shallow water equations with topography

We consider a non-dimensional form of the shallow water system with topography. The system reads:

$$U_{i,0} + F_{i,1} + G_{ij}U_{j,1} = 0, \text{ for } i, j = 1, 2, 3 \quad (3.6)$$

with:

$$U = \begin{bmatrix} b \\ h \\ hu \end{bmatrix}, \quad F = \begin{bmatrix} 0 \\ hu \\ hu^2 + \frac{1}{2}F^{-2}h^2 \end{bmatrix}, \quad G(U) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ F^{-2}h & 0 & 0 \end{bmatrix}. \quad (3.7)$$

Here b is the topography, h the water depth, u the flow velocity and F the Froude number defined as $F = u_0^*/\sqrt{g^*h_0^*}$, where the starred values denote reference values. The eigenvalues of $\partial F/\partial U + G(U)$ are given by:

$$\lambda_1 = u - \sqrt{F^{-2}h}, \quad \lambda_2 = 0, \quad \lambda_3 = u + \sqrt{F^{-2}h}. \quad (3.8)$$

When taking $\phi = U_L + \tau(U_R - U_L)$, the NCP-flux for (3.6) on a fixed grid becomes:

$$\hat{P}^{nc} = \begin{cases} F^L - \frac{1}{2}V^{nc}, & \text{if } S_L > 0, \\ F^{hll} - (S_R + S_L)V^{nc}/(2(S_R - S_L)), & \text{if } S_L < 0 < S_R, \\ F^R + \frac{1}{2}V^{nc}, & \text{if } S_R < 0, \end{cases}$$

in which F^{hll} is the HLL-flux [74]:

$$F^{hll} = \frac{S_R F_L - S_L F_R + S_L S_R (U_R - U_L)}{S_R - S_L}$$

and V^{nc} appears in the extra term due to the nonconservative product:

$$V^{nc} = [0, 0, -F^{-2} \{ \{ h \} \} [b]]^T.$$

In the numerical flux, as derived in Section 2.3, we take:

$$S_L = \min(u_L - \sqrt{F^{-2} h_L}, u_R - \sqrt{F^{-2} h_R}) \text{ and}$$

$$S_R = \max(u_L + \sqrt{F^{-2} h_L}, u_R + \sqrt{F^{-2} h_R}).$$

Test cases 1 and 2: rest flow

For test cases 1 and 2 we only consider the solution determined with space-time DGFEM calculations using linear basis functions and the linear path $\phi = U_L + \tau(U_R - U_L)$. Consider flow at rest over a discontinuous topography with initial and boundary conditions:

- Test case 1. Initial conditions: $b(x, 0) = 1$ if $x < 0$ and $b(x, 0) = 0$ if $x > 0$, $h(x, 0) + b(x, 0) = 2$, $hu(x, 0) = 0$. Boundary conditions: $b(-5, t) = 1$, $h(-5, t) = 1$, $u(-5, t) = 0$, $b(5, t) = 0$, $h(5, t) = 2$, $u(5, t) = 0$.
- Test case 2. Initial condition: $b(x, 0) = 0$ if $x < 0$ and $b(x, 0) = 1$ if $x > 0$, $h(x, 0) + b(x, 0) = 2$, $hu(x, 0) = 0$. Boundary conditions: $b(-5, t) = 0$, $h(-5, t) = 2$, $u(-5, t) = 0$, $b(5, t) = 1$, $h(5, t) = 1$, $u(5, t) = 0$.

In Figure 3.1 we show the steady state solution, calculated using a time step of $\Delta t = 10^{21}$ on a grid with 100 cells and a Froude number of $F = 0.2$. We solve the system of non-linear equations using a pseudo time stepping integration method (see van der Vegt and van der Ven [79]). As stopping criterium in the pseudo time-stepping calculation we take that the maximum residual must be smaller than 10^{-13} . A pseudo time stepping CFL number of $CFL^{pseudo} = 0.8$ is used.

For the space DGFEM weak formulation we prove theoretically, that when using linear basis functions and taking the path $\phi = U_L + \tau(U_R - U_L)$, rest flow remains at rest. Consider the one dimensional version of the space DGFEM

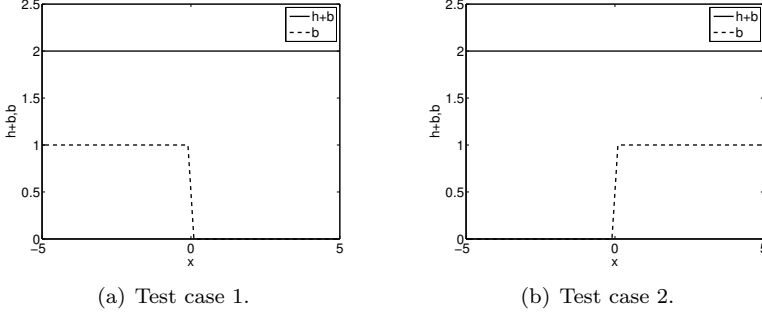


Figure 3.1: *Flow at rest over a discontinuous topography.* $F = 0.2$, 100 cells, $\Delta t = 10^{21}$.

weak formulation (A.11) for the shallow water equations:

$$\begin{aligned}
 0 &= \sum_k \int_{K_k} (V_i U_{i,0} - V_{i,1} F_i + V_i G_{ij} U_{j,1}) dK \\
 &+ \sum_{S \in \mathcal{S}_I} \int_S \{V_i\} \left(\int_0^1 G_{ij}(\phi(\tau; U_L, U_R)) \frac{\partial \phi_j}{\partial \tau}(\tau; U_L, U_R) d\tau \right) \bar{n}^L dS \\
 &+ \sum_{S \in \mathcal{S}_I} \int_S (V_i^L - V_i^R) \hat{P}_i^{nc} dS.
 \end{aligned}$$

We only consider cell K_k where the contributions satisfy:

$$\begin{aligned}
 0 &= \int_{K_k} (V_i U_{i,0} - V_{i,1} F_i + V_i G_{ij} U_{j,1}) dK \\
 &+ \int_{S_{k+1}} \frac{1}{2} V_i^L \left(\int_0^1 G_{ij}(\phi(\tau; U_L, U_R)) \frac{\partial \phi_j}{\partial \tau}(\tau; U_L, U_R) d\tau \right) \bar{n}^L + V_i^L \hat{P}_i^{nc} dS \\
 &+ \int_{S_k} \frac{1}{2} V_i^R \left(\int_0^1 G_{ij}(\phi(\tau; U_L, U_R)) \frac{\partial \phi_j}{\partial \tau}(\tau; U_L, U_R) d\tau \right) \bar{n}^L - V_i^R \hat{P}_i^{nc} dS.
 \end{aligned} \tag{3.9}$$

For the numerical flux we take the star-state solution given by (2.38). For rest flow, using $\phi = U_L + \tau(U_R - U_L)$ and $h_L + b_L = h_R + b_R$ the star-state solution is given by:

$$\bar{U}^* = \frac{1}{S_R - S_L} [S_R b_R - S_L b_L, \quad S_R h_R - S_L h_L, \quad 0]^T, \tag{3.10}$$

so that the numerical flux $\widehat{P}^{nc} = \{\{F\}\} + \frac{1}{2}(S_L(\bar{U}^* - U_L) + S_R(\bar{U}^* - U_R))$ is given by:

$$\widehat{P}^{nc} = \left[\frac{S_L S_R (b_R - b_L)}{S_R - S_L}, \quad \frac{S_L S_R (h_R - h_L)}{S_R - S_L}, \quad \frac{1}{4} \mathbf{F}^{-2} (h_L^2 + h_R^2) \right]^T. \quad (3.11)$$

Also, using $\phi = U_L + \tau(U_R - U_L)$ and $h_L + b_L = h_R + b_R$ we can show that:

$$\int_0^1 G_{ij}(\phi(\tau; U_L, U_R)) \frac{\partial \phi_j}{\partial \tau}(\tau; U_L, U_R) d\tau = [0, 0, -\mathbf{F}^{-2} \{\{b\}\} \{\{h\}\}]^T.$$

We can write (3.9) now as:

$$0 = \int_{K_k} (V_i U_{i,0} - V_{i,1} F_i + G_{ij} U_{j,1}) dK + \int_{S_{k+1}} V_i^L \mathcal{P}_i^p dS - \int_{S_k} V_i^R \mathcal{P}_i^m dS, \quad (3.12)$$

where \mathcal{P}^p and \mathcal{P}^m are given by:

$$\begin{aligned} \mathcal{P}^p &= \frac{1}{2} \int_0^1 G_{ij}(\phi(\tau; U_L, U_R)) \frac{\partial \phi_j}{\partial \tau}(\tau; U_L, U_R) d\tau + \widehat{P}^{nc} \\ &= \left[\frac{S_L S_R (b_R - b_L)}{S_R - S_L}, \quad \frac{S_L S_R (h_R - h_L)}{S_R - S_L}, \quad \frac{1}{2} \mathbf{F}^{-2} h_L^2 \right]^T \\ \mathcal{P}^m &= \frac{1}{2} \int_0^1 G_{ij}(\phi(\tau; U_L, U_R)) \frac{\partial \phi_j}{\partial \tau}(\tau; U_L, U_R) d\tau - \widehat{P}^{nc} \\ &= \left[-\frac{S_L S_R (b_R - b_L)}{S_R - S_L}, \quad -\frac{S_L S_R (h_R - h_L)}{S_R - S_L}, \quad \frac{1}{2} \mathbf{F}^{-2} h_R^2 \right]^T. \end{aligned}$$

Using linear basis functions we can evaluate the integrals as follows:

$$\int_{K_k} V_i U_{i,0} dK = \Delta x \bar{V}_i|_{K_k} \partial_t \bar{U}_i|_{K_k} + \frac{\Delta x}{3} \widehat{V}_i|_{K_k} \partial_t \widehat{U}_i|_{K_k}, \quad (3.13a)$$

$$\begin{aligned} - \int_{K_k} V_{i,1} F_i dK &= - \int_{-1}^1 \widehat{V}_i|_{K_k} F(\bar{U}_i|_{K_k} + \widehat{U}_i|_{K_k} \xi) d\xi \\ &= -\widehat{V}_i|_{K_k} [0, 0, \frac{1}{3} \mathbf{F}^{-2} \hat{h}_k^2 + \mathbf{F}^{-2} \bar{h}_k^2]^T \end{aligned} \quad (3.13b)$$

$$\begin{aligned} \int_{K_k} V_i G_{ij} U_{j,1} dK &= \int_{-1}^1 (\bar{V}_i|_{K_k} + \widehat{V}_i|_{K_k} \xi) G(\bar{U}_i|_{K_k} + \widehat{U}_i|_{K_k} \xi) \widehat{U}_i|_{K_k} d\xi \\ &= \bar{V}_i|_{K_k} [0, 0, 2\mathbf{F}^{-2} \bar{h}_k \hat{\theta}_k]^T + \widehat{V}_i|_{K_k} [0, 0, \frac{2}{3} \mathbf{F}^{-2} \hat{h}_k \hat{\theta}_k]^T \end{aligned} \quad (3.13c)$$

$$\int_{S_{k+1}} V_i^L \mathcal{P}_i^p dS = (\bar{V}|_{K_k} + \widehat{V}|_{K_k}) \begin{bmatrix} \frac{S_{k+1}^L S_{k+1}^R (b_{k+1}^R - b_{k+1}^L)}{S_{k+1}^R - S_{k+1}^L} \\ \frac{S_{k+1}^L S_{k+1}^R (h_{k+1}^R - h_{k+1}^L)}{S_{k+1}^R - S_{k+1}^L} \\ \frac{1}{2} F^{-2} (\bar{h}_k + \hat{h}_k)^2 \end{bmatrix}, \quad (3.13d)$$

$$- \int_{S_k} V_i^R \mathcal{P}_i^m dS = -(\bar{V}|_{K_k} - \widehat{V}|_{K_k}) \begin{bmatrix} \frac{S_k^L S_k^R (b_k^R - b_k^L)}{S_k^R - S_k^L} \\ \frac{S_k^L S_k^R (h_k^R - h_k^L)}{S_k^R - S_k^L} \\ \frac{1}{2} F^{-2} (\bar{h}_k - \hat{h}_k)^2 \end{bmatrix}, \quad (3.13e)$$

where $\bar{(\cdot)}$ and $\widehat{(\cdot)}$ are the means and slopes, respectively, of the approximation for U and V . Adding the vectors (3.13b)-(3.13e), we note that the third element of this sum is zero using $h_L + b_L = h_R + b_R$ and the fact that the slope of $h + b = 0$ (so $\widehat{U}|_{K_k} = (-\hat{h}_k, \hat{h}_k, 0)$). Note that in (3.13d) and (3.13e) we have $b_{k+1}^R - b_{k+1}^L + h_{k+1}^R - h_{k+1}^L = 0$ and $b_k^R - b_k^L + h_k^R - h_k^L = 0$, respectively so that:

$$\partial_t(\bar{h}_k + \bar{b}_k) = 0, \quad \partial_t(\hat{h}_k + \hat{b}_k) = 0, \quad \partial_t \bar{h} u_k = 0, \quad \partial_t \widehat{h} u_k = 0,$$

meaning that for rest flow $h + b$ remains constant.

Test case 3: Subcritical flow over a bump

We now consider subcritical flow with a Froude number of $F = 0.2$ over a bump. The topography reads:

$$b(x) = \begin{cases} a(b - (x - x_p))(b + (x - x_p))b^{-2} & \text{for } |x - x_p| \leq b, \\ 0 & \text{otherwise.} \end{cases} \quad (3.14)$$

We use $x_p = 10$, $a = 0.5$ and $b = 2$ as in [71]. The exact steady state solution for this test case is found by solving the following third order equation in u [31, 71]:

$$F^2 u^3 / 2 + (b - F^2 / 2 - 1)u + 1 = 0 \text{ with } hu = 1. \quad (3.15)$$

The domain $x \in [0, 20]$ is divided into 40, 80, 160 and 320 cells. We consider DGFEM and STDGFEM calculations using the linear path $\phi = U_L + \tau(U_R - U_L)$. For space DGFEM calculations, a CFL number of $CFL = 0.8$ is taken and when the residuals are smaller than 10^{-11} the calculation is stopped. For STDGFEM calculations we consider the solution after one physical time step of $\Delta t = 10^{21}$. We can do this because we want to consider the steady state solution. As stopping criterium in the pseudo time-stepping calculation we take that the maximum residual must be smaller than 10^{-11} . A pseudo time stepping CFL number of $CFL^{pseudo} = 0.8$ is used.

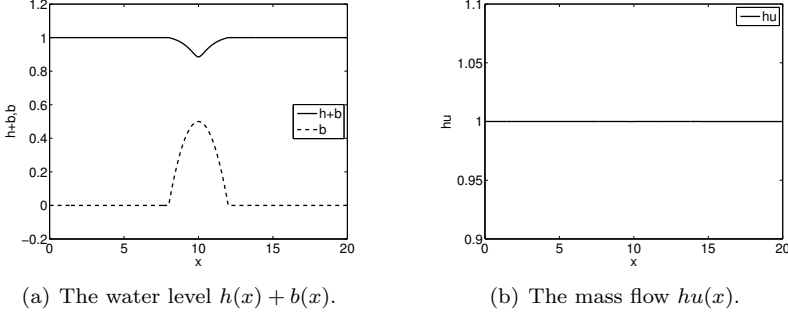


Figure 3.2: *Test case 3: steady-state solution calculated using space DGFEM, $F = 0.2$, 320 cells.*

The initial condition is $h + b = 1$ and $hu = 1$ and the boundary conditions are: $b(0, t) = 0$, $h(0, t) = 1$, $u(0, t) = 1$, $b(1, t) = 0$, $h(1, t) = 1$ and $u(1, t) = 1$. The steady state solution is given in Figure 3.2. The order of convergence is determined by looking at the L^2 and the L^{max} norm of the numerical error in $z = h + b$ and hu with respect to the exact solution:

$$\|z_{num} - z_{exact}\|_2 = \left(\sum_{k=1}^{N_{cells}} \int_{K_k} (z_{K_k}^{num} - z_{K_k}^{exact})^2 \right)^{1/2}, \quad (3.16)$$

and:

$$\|z_{num} - z_{exact}\|_{max} = \max\{|z_{num}^i - z_{exact}^i| : 1 \leq i \leq N_{cells}\}. \quad (3.17)$$

The order of convergence using DGFEM and STDGFEM is given in Table 3.1 using linear basis functions and in Table 3.2 using quadratic basis functions. Using linear basis functions we obtain second order convergence and using quadratic basis functions we obtain third order convergence for both space-DGFEM and space-time DGFEM calculations.

Test case 4: Supercritical flow over a bump

Next, we consider supercritical flow with a Froude number of $F = 1.9$ over a bump. We use the same topography (3.14) and the exact solution can be found by solving (3.15). The domain $x \in [0, 20]$ is again divided into 40, 80, 160 and 320 cells and we consider DGFEM and STDGFEM calculations using the linear path $\phi = U_L + \tau(U_R - U_L)$. For space DGFEM calculations, time steps of $\Delta t = 0.01$ are made. Using linear basis functions, a CFL number of $CFL = 0.3$

DGFEM								
N_{cells}	$h + b$				hu			
	L^2 error	p	L^{max} error	p	L^2 error	p	L^{max} error	p
40	$0.1133 \cdot 10^{-2}$	-	$0.6513 \cdot 10^{-2}$	-	$0.1265 \cdot 10^{-2}$	-	$0.3302 \cdot 10^{-2}$	-
80	$0.3193 \cdot 10^{-3}$	1.8	$0.2387 \cdot 10^{-2}$	1.4	$0.1944 \cdot 10^{-3}$	2.7	$0.8030 \cdot 10^{-3}$	2.0
160	$0.8364 \cdot 10^{-4}$	1.9	$0.6989 \cdot 10^{-3}$	1.8	$0.2764 \cdot 10^{-4}$	2.8	$0.1369 \cdot 10^{-3}$	2.6
320	$0.2119 \cdot 10^{-4}$	2.0	$0.1847 \cdot 10^{-3}$	1.9	$0.3798 \cdot 10^{-5}$	2.9	$0.2931 \cdot 10^{-4}$	2.2

STDGFEM								
N_{cells}	$h + b$				hu			
	L^2 error	p	L^{max} error	p	L^2 error	p	L^{max} error	p
40	$0.1141 \cdot 10^{-2}$	-	$0.6559 \cdot 10^{-2}$	-	$0.1262 \cdot 10^{-2}$	-	$0.3285 \cdot 10^{-2}$	-
80	$0.3194 \cdot 10^{-3}$	1.8	$0.2387 \cdot 10^{-2}$	1.5	$0.1943 \cdot 10^{-3}$	2.7	$0.8029 \cdot 10^{-3}$	2.0
160	$0.8365 \cdot 10^{-4}$	1.9	$0.6989 \cdot 10^{-3}$	1.8	$0.2763 \cdot 10^{-4}$	2.8	$0.1369 \cdot 10^{-3}$	2.6
320	$0.2119 \cdot 10^{-4}$	2.0	$0.1847 \cdot 10^{-3}$	1.9	$0.3797 \cdot 10^{-5}$	2.9	$0.2929 \cdot 10^{-4}$	2.2

Table 3.1: L^2 and L^{max} error for $h + b$ and hu using DGFEM and STDGFEM for test case 3. Second order convergence rates are shown for $F = 0.2$.

is taken and when the residuals are smaller than 10^{-11} the calculation is stopped. For the STDGFEM calculation we consider again the solution after one physical time step of $\Delta t = 10^{21}$. The same stopping criteria as in the subcritical flow case are used. Using linear basis functions, we use a pseudo time stepping CFL number of $CFL^{pseudo} = 0.8$. For quadratic basis functions, on the grids with 40 and 160 cells, a pseudo time stepping CFL number of $CFL^{pseudo} = 0.4$ is employed and on the grids with 80 and 320 cells a pseudo time stepping CFL number of $CFL^{pseudo} = 0.8$.

The initial condition is $h + b = 1$ and $hu = 1$ and transmissive boundary conditions are given at $x = 0$ and at $x = 20$, i.e., $U^b = U^L$, where U^b is the vector of the boundary data and U^L is the vector with the data calculated at the boundary from inside the domain. The steady-state solution is shown in Figure 3.3. The order of convergence is again determined by computing the L^2 and the L^{max} norm of the numerical error in $h + b$ and hu with respect to the exact solution as defined in (3.16) and (3.17). The order of convergence using DGFEM and STDGFEM is given in Table 3.3 using linear basis functions and in Table 3.4 using quadratic basis functions.

We see that the space- and space-time DGFEM calculations results in second order convergence for $h + b$ using linear basis functions and in third order

DGFEM								
N_{cells}	$h + b$				hu			
	L^2 error	p	L^{\max} error	p	L^2 error	p	L^{\max} error	p
40	$0.3210 \cdot 10^{-3}$	-	$0.1466 \cdot 10^{-2}$	-	$0.8352 \cdot 10^{-3}$	-	$0.3124 \cdot 10^{-2}$	-
80	$0.4622 \cdot 10^{-4}$	2.8	$0.2670 \cdot 10^{-3}$	2.5	$0.1269 \cdot 10^{-3}$	2.7	$0.5562 \cdot 10^{-3}$	2.5
160	$0.6303 \cdot 10^{-5}$	2.9	$0.3567 \cdot 10^{-4}$	2.9	$0.1689 \cdot 10^{-4}$	2.9	$0.7186 \cdot 10^{-4}$	3.0
320	$0.7931 \cdot 10^{-6}$	3.0	$0.4459 \cdot 10^{-5}$	3.0	$0.2144 \cdot 10^{-5}$	3.0	$0.8860 \cdot 10^{-5}$	3.0

STDGFEM								
N_{cells}	$h + b$				hu			
	L^2 error	p	L^{\max} error	p	L^2 error	p	L^{\max} error	p
40	$0.3278 \cdot 10^{-3}$	-	$0.1836 \cdot 10^{-2}$	-	$0.2339 \cdot 10^{-3}$	-	$0.1170 \cdot 10^{-2}$	-
80	$0.4433 \cdot 10^{-4}$	2.9	$0.3195 \cdot 10^{-3}$	2.5	$0.3721 \cdot 10^{-4}$	2.7	$0.2401 \cdot 10^{-3}$	2.3
160	$0.4556 \cdot 10^{-5}$	3.3	$0.3142 \cdot 10^{-4}$	3.3	$0.5513 \cdot 10^{-5}$	2.8	$0.3596 \cdot 10^{-4}$	2.7
320	$0.5522 \cdot 10^{-6}$	3.0	$0.4407 \cdot 10^{-5}$	2.8	$0.7489 \cdot 10^{-6}$	2.9	$0.5218 \cdot 10^{-5}$	2.8

Table 3.2: L^2 and L^{\max} error for $h + b$ and hu using DGFEM and STDGFEM for test case 3. Third order convergence rates are shown for $F = 0.2$.

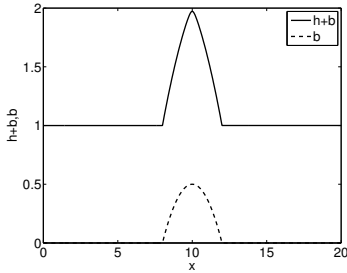
convergence for $h + b$ using quadratic basis functions. We do not show the order of convergence for hu because the error for hu is of the order of machine precision on all meshes for the space DGFEM calculations and stabilizes around 10^{-8} for the space-time DGFEM calculations.

Test case 5: Transcritical flow over a bump

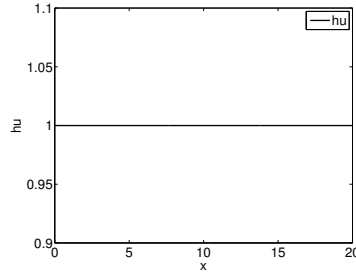
For this test case we consider the steady state solution of transcritical flow with a shock over a bump. The topography is given by:

$$b(x) = \begin{cases} 0.2 - 0.05(x - 10)^2 & \text{if } 8 \leq x \leq 12, \\ 0 & \text{otherwise,} \end{cases}$$

which is the same as that used by Xing and Shu [87]. The initial condition is $h + b = 0.5$ and $hu = 0$ and the boundary conditions are: $b(0, t) = 0$, $hu(0, t) = 0.18$, $b(25, t) = 0$, $h(25, t) = 0.33$, $hu(25, t) = 0.18$. The remaining boundary data are set equal to the data calculated at the boundary from inside the domain. In our computations, we take $F^{-2} = 9.812$. Simulations concern space-time DGFEM. We consider the solution after one physical time step of $\Delta t = 10^{21}$ on a grid with 200 cells using a pseudo time stepping CFL number



(a) The water level $h(x) + b(x)$.



(b) The mass flow $hu(x)$.

Figure 3.3: *Test case 4: steady-state solution calculated using space DGFEM, $F = 1.9$, 320 cells.*

N_{cells}	DGFEM $h + b$				STDGFEM $h + b$			
	L^2 error	p	L^{\max} error	p	L^2 error	p	L^{\max} error	p
40	$0.7543 \cdot 10^{-2}$	-	$0.4619 \cdot 10^{-1}$	-	$0.7543 \cdot 10^{-2}$	-	$0.4619 \cdot 10^{-1}$	-
80	$0.1281 \cdot 10^{-2}$	2.6	$0.9406 \cdot 10^{-2}$	2.3	$0.1281 \cdot 10^{-2}$	2.6	$0.9406 \cdot 10^{-2}$	2.3
160	$0.3188 \cdot 10^{-3}$	2.0	$0.2615 \cdot 10^{-2}$	1.8	$0.3188 \cdot 10^{-3}$	2.0	$0.2615 \cdot 10^{-2}$	1.8
320	$0.7914 \cdot 10^{-4}$	2.0	$0.6883 \cdot 10^{-3}$	1.9	$0.7914 \cdot 10^{-4}$	2.0	$0.6883 \cdot 10^{-3}$	1.9

Table 3.3: L^2 and L^{\max} error for $h + b$ using DGFEM and STDGFEM for test case 4. Second order convergence rates are shown for $F = 1.9$.

of $CF L^{pseudo} = 0.8$. To deal with the shock, we used the slope limiter as discussed in Section 3.1.3. The solution is given in Figure 3.4 and compares well with results in [31].

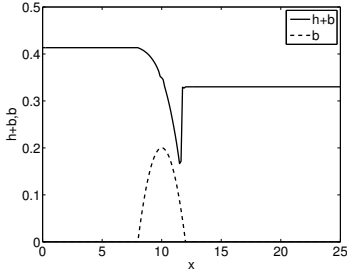
Test case 6: Perturbation of a steady state solution

We repeat a test case as was formulated in Xing and Shu [87] which was originally proposed by LeVeque [47]. Consider a topography given by:

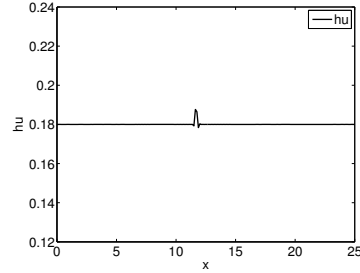
$$b(x) = \begin{cases} 0.25(\cos(10\pi(x - 1.5)) + 1) & \text{if } 1.4 \leq x \leq 1.6, \\ 0 & \text{otherwise.} \end{cases}$$

N_{cells}	DGFEM $h + b$				STDGFEM $h + b$			
	L^2 error	p	L^{\max} error	p	L^2 error	p	L^{\max} error	p
40	$0.1293 \cdot 10^{-2}$	-	$0.5034 \cdot 10^{-2}$	-	$0.9181 \cdot 10^{-3}$	-	$0.4946 \cdot 10^{-2}$	-
80	$0.1944 \cdot 10^{-3}$	2.7	$0.9383 \cdot 10^{-3}$	2.4	$0.1624 \cdot 10^{-3}$	2.5	$0.1127 \cdot 10^{-2}$	2.1
160	$0.2892 \cdot 10^{-4}$	2.7	$0.1545 \cdot 10^{-3}$	2.6	$0.1830 \cdot 10^{-4}$	3.1	$0.1382 \cdot 10^{-3}$	3.0
320	$0.3724 \cdot 10^{-5}$	3.0	$0.2111 \cdot 10^{-4}$	2.9	$0.2253 \cdot 10^{-5}$	3.0	$0.2002 \cdot 10^{-4}$	2.8

Table 3.4: L^2 and L^{\max} error for $h + b$ using DGFEM and STDGFEM for test case 4. Third order convergence rates are shown for $F = 1.9$.



(a) The water level $h(x) + b(x)$.



(b) The mass flow $hu(x)$.

Figure 3.4: Test case 5: steady-state transcritical flow with a shock, $\Delta t = 10^{21}$, $N_{cells} = 200$, $CFL^{\tau} = 0.8$, $F^{-2} = 9.812$.

The initial conditions are given by:

$$hu(x, 0) = 0, \quad h(x, 0) = \begin{cases} 1 - b(x) + \epsilon & \text{if } 1.1 \leq x \leq 1.2, \\ 1 - b(x) & \text{otherwise.} \end{cases}$$

At the boundaries, we use transmissive boundary conditions. We take $F^{-2} = 9.812$. The same two cases as in Xing and Shu [87] were run: $\epsilon = 0.2$ (big pulse) and $\epsilon = 0.001$ (small pulse). We used space-time DGFEM to compute the solution on a uniform grid with 200 cells and 3000 cells. On the grid with 200 cells, a physical time step of $\Delta t = 0.0002$ was used. On the grid with 3000 cells, we used a physical time step of $\Delta t = 0.00002$. A pseudo time stepping CFL number of $CFL^{pseudo} = 0.4$ was used. In Figures 3.5 and 3.6 we show the fine and coarse mesh solution, as in [87], for the water level $h(x) + b(x)$ and mass flow $hu(x)$ at time $t = 0.2$ for the big pulse test case and the small pulse

test case, respectively.

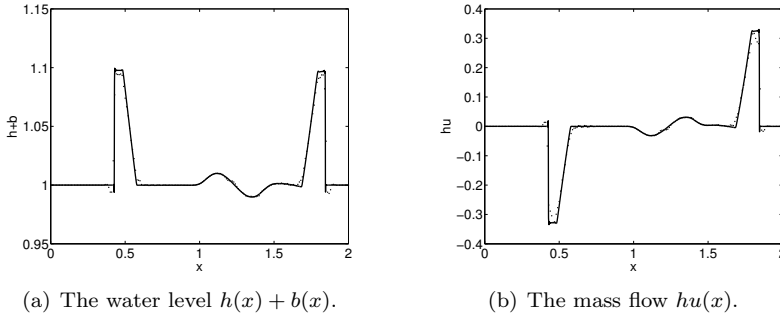


Figure 3.5: *Test case 6: perturbation of a steady state solution with a big pulse at time $t = 0.2$, $\epsilon = 0.2$. Line: $N_{cells} = 3000$. Dots: $N_{cells} = 200$.*

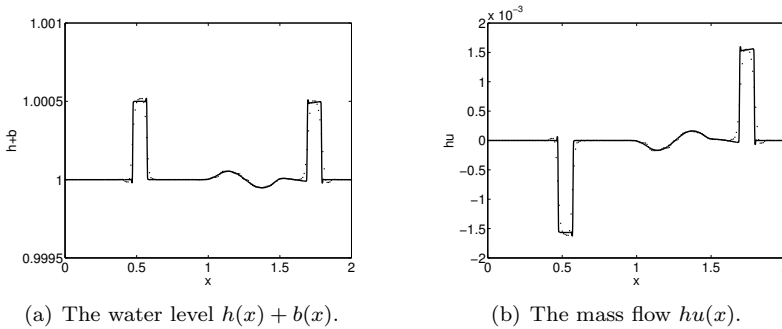


Figure 3.6: *Test case 6: perturbation of a steady state solution with a small pulse at time $t = 0.2$, $\epsilon = 0.001$. Line: $N_{cells} = 3000$. Dots: $N_{cells} = 200$.*

Test case 7: Dam break problem over a rectangular bump

A dam break problem is simulated over a rectangular hump, as in [87]. The topography is given by:

$$b(x) = \begin{cases} 8 & \text{if } |x - 750| \leq 1500/8, \\ 0 & \text{otherwise,} \end{cases}$$

for $x \in [0, 1500]$. The initial conditions are given by:

$$hu(x, 0) = 0, \quad h(x, 0) = \begin{cases} 20 - b(x) & \text{if } x \leq 750, \\ 15 - b(x) & \text{otherwise,} \end{cases}$$

and as boundary conditions we take: $b(0, t) = 0$, $h(0, t) = 20$, $hu(0, t) = 0$, $b(1500, t) = 0$, $h(1500, t) = 15$ and $hu(1500, t) = 0$. We take $F^{-2} = 9.812$. With space-time DGFEM the solution was computed on a uniform grid with 400 cells and 4000 cells. On the grid with 400 cells, a physical time step of $\Delta t = 0.02$ was used and on the grid with 4000 cells, the physical time step was $\Delta t = 0.002$. The pseudo time stepping CFL number was $CFL^{pseudo} = 0.8$. In Figures 3.7 and 3.8 we show the solution for the water level $h(x) + b(x)$ at time $t = 15$ and at time $t = 60$, respectively.

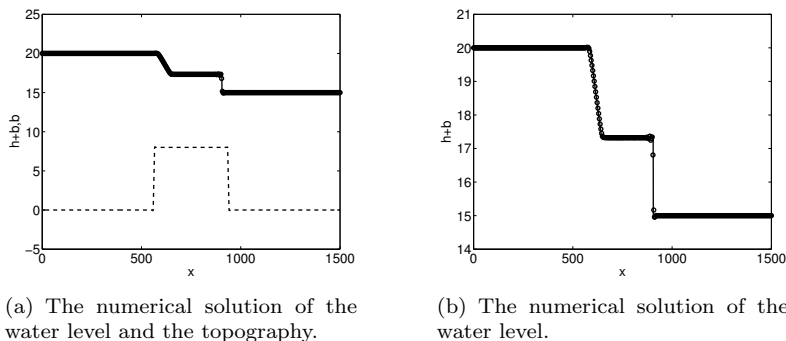


Figure 3.7: *Test case 7: the dam breaking problem at time $t = 15$. Line: 4000 cells. Dots: 400 cells.*

Conclusions

For the shallow water equations with topography we showed numerical results of seven test cases calculated using the space- and/or space-time DGFEM discretizations we developed for nonconservative hyperbolic partial differential equations. For all test cases we obtained good results. For test cases 1 and 2 we showed that rest flow remained unchanged despite having discontinuities in the topography. In test cases 3 and 4 we solved subcritical and supercritical flow over a bump demonstrating that the scheme is second order accurate for linear basis functions and third order accurate for quadratic basis functions. In test cases 5, 6 and 7 we showed that we resolved also more complex test cases with discontinuous solutions.

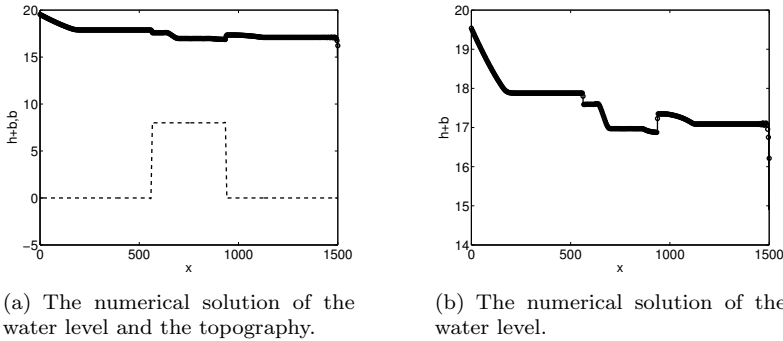


Figure 3.8: *Test case 7: the dam breaking problem at time $t = 60$. Line: 4000 cells. Dots: 400 cells.*

3.2.2 The depth averaged two-fluid model

In this section we consider two fluid models (also known as Eulerian models) in which the particle phase is treated as a continuum by averaging over individual particles. Two frequently used models for two-fluid equations, are those derived by Anderson and Jackson [5], and Drew and Lahey [22] and Enwald et al. [23]. Apart from their derivation, the difference between these systems of equations is how the fluid-phase shear stress (if included) is multiplied by the solid volume fraction in the momentum equations (see also van Wachem et al. [81]). In the limiting case that pressure is the only fluid stress, both formulations are equal.

We will consider a simplification of these equations, namely the depth-averaged two fluid model derived by Pitman and Le [62]. They start with the system of Anderson and Jackson [5] and use the shallow flow assumption, $H/L \ll 1$, where H is the characteristic length of the flow in the z -direction and L the characteristic length of the flow in the y -direction. The derivation is similar to the way the shallow water equations are derived from the Navier-Stokes equations. Since the pressure is the only fluid stress, the same depth-averaged two fluid model also follows from the system derived by Drew and Lahey [22] and Enwald et al. [23].

The dimensionless depth-averaged two fluid model of Pitman and Le [62], ignoring source terms for simplicity, can be written as:

$$U_{i,0} + F_{i,1} + G_{ij}U_{j,1} = 0, \text{ for } i, j = 1, 2, 3, 4, \quad (3.18)$$

where:

$$\begin{aligned}
 U &= \begin{bmatrix} h(1-\alpha) \\ h\alpha \\ h\alpha v \\ hu(1-\alpha) \\ b \end{bmatrix}, \quad F = \begin{bmatrix} h(1-\alpha)u \\ h\alpha v \\ h\alpha v^2 + \frac{1}{2}\varepsilon(1-\rho)\alpha_{xx}gh^2\alpha \\ hu^2 + \frac{1}{2}\varepsilon gh^2 \\ 0 \end{bmatrix} \\
 G(U) &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \varepsilon\rho\alpha gh & \varepsilon\rho\alpha gh & 0 & 0 & \varepsilon(1-\rho)\alpha_{xx}gh\alpha + \varepsilon\rho\alpha gh & \\ \frac{2u^2\alpha}{1-\alpha} - \alpha u^2 - \varepsilon gh\alpha & -\varepsilon gh\alpha - \alpha u^2 & u(\alpha-1) & u\alpha - \frac{2u\alpha}{1-\alpha} & (1-\alpha)\varepsilon gh & \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{3.19}
 \end{aligned}$$

Again we have taken the topography b as unknown. The meaning of the different symbols are: $h(x, t)$ is the depth of the flow, $v(x, t)$ the velocity of the solid phase, $u(x, t)$ the velocity of the fluid phase, $\alpha(x, t)$ the volume fraction of the solid phase, $b(x)$ the topography term, $\varepsilon = H/L$, ρ is the ratio between the fluid density and the solid density, $\alpha_{xx} = k_{ap}$, where k_{ap} is the Earth pressure coefficient and g is the z -component of the scaled gravity. Note that in the limit $\alpha \rightarrow 0$, this model reduces to the shallow water equations with εg akin to F^{-2} :

$$\begin{aligned}
 \partial_t h + \partial_x(hu) &= 0, \\
 \partial_t(hu) + \partial_x(hu^2 + \frac{1}{2}\varepsilon gh^2) &= -\varepsilon gh\partial_x b.
 \end{aligned} \tag{3.20}$$

In the limit $\alpha \rightarrow 1$, the depth-averaged two-fluid model model reduces to:

$$\begin{aligned}
 \partial_t h + \partial_x(hv) &= 0, \\
 \partial_t(hv) + \partial_x(hv^2 + \frac{1}{2}\varepsilon k_{ap}gh^2) &= -\varepsilon k_{ap}gh\partial_x b,
 \end{aligned} \tag{3.21}$$

which is the Savage-Hutter model without source terms, a model that simulates avalanches of dry granular matter [36].

In our simulations, we set the Earth pressure coefficient to be $\alpha_{xx} = 1$ and take $\varepsilon = 1$. To compute the eigenvalues of $\partial F/\partial U + G(U)$, we use the LAPACK package. The biggest eigenvalue is used for S_R and the smallest eigenvalue is used for S_L in the NCP numerical flux.

Test case 8: Two-phase subcritical flow

As in the case of the shallow water equations with topography, also for the two-phase flow model we consider the steady state solution for subcritical flow over a bump. We consider the same topography (3.14). The reference solution is found by solving:

$$\partial_x U = A^{-1}S, \tag{3.22}$$

where U , A and S are given by:

$$\begin{aligned}
 U &= [h(1 - \alpha), \quad h\alpha]^T, \quad S = \begin{bmatrix} -(1 - \alpha)hg\partial_x b \\ -gh\alpha\partial_x b \end{bmatrix} \\
 A &= \begin{bmatrix} u^2(1 - \alpha) - 2u^2 + gh(1 - \alpha) & u^2(1 - \alpha) + gh(1 - \alpha) \\ \frac{1}{2}(1 + \rho)gh\alpha & \frac{1}{2}(1 - \rho)g(1 + \alpha)h + gph\alpha - v^2 \end{bmatrix},
 \end{aligned} \tag{3.23}$$

with the topography derivative a known function and steady state discharges:

$$hu(1 - \alpha) = q_1, \quad hv\alpha = q_2, \tag{3.24}$$

with q_1 and q_2 integration constants. Here we take $q_1 = 0.2$, $q_2 = 0.1$, $g = 1$ and $\rho = 0.5$ and as initial condition $h(1 - \alpha) = 1$, $h\alpha = 0.6$, $hu(1 - \alpha) = 0.2$ and $hv\alpha = 0.1$. We use the STDGFEM formulation to calculate the solution. We consider one physical time step of $\Delta t = 10^{21}$ and use a pseudo time stepping integration method to solve the system of non-linear equations. We determine the solution on a domain $x \in [0, 20]$ divided into 40, 80, 160 and 320 cells. As stopping criterium in the pseudo time-stepping method we take that the maximum residual must be smaller than 10^{-8} . The pseudo time stepping CFL number is $CFL^{pseudo} = 0.1$. At the boundaries, we define the exterior trace to be the same as the initial condition. The numerical flux decides then what to do with this information. The steady state solution is given in Figure 3.9. The order of convergence is determined by computing the L^2 and L^{max} norm of the error, similar as to what is done in (3.16) and (3.17). The order of convergence is given in Table 3.5. Using linear basis functions, we obtain second order convergence as expected.

Test case 9: Two-phase supercritical flow

We will now consider the steady state solution of two-phase supercritical flow over a bump with (3.14) as topography. The exact solution is found by solving (3.22)-(3.24), now with $q_1 = 4$ and $q_2 = 2$. Other constants remain as in test case 9 and we use the same solution strategy. The steady state solution is given in Figure 3.10 and the order of convergence is given in Table 3.6. Again, using linear basis functions, we obtain second order convergence for the variables $h(1 - \alpha) + b$ and $h\alpha + b$. We do not see second order convergence for the variables $hu(1 - \alpha)$ and $hv\alpha$ because the error for these solutions stabilizes around 10^{-8} , the value of the maximum residual.

Test case 10: A two-phase dam break problem

For the depth-averaged two-phase flow model we consider a dam break type test case. Consider two mixtures separated by a membrane. The left mixture

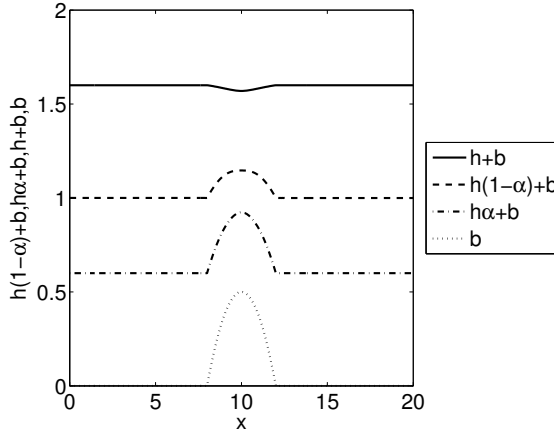


Figure 3.9: *Test case 8: steady-state solution for a subcritical two-phase flow calculated with STDGFEM using 320 cells. Shown are the total flow height $h+b$, the flow height due to the fluid phase $h(1-\alpha)$, the flow height due to the solids phase $h\alpha$ and the topography b .*

STDGFEM								
N_{cells}	$h(1-\alpha)+b$				$h\alpha+b$			
	L^2 error	p	L^{max} error	p	L^2 error	p	L^{max} error	p
40	$0.8171 \cdot 10^{-3}$	-	$0.2308 \cdot 10^{-2}$	-	$0.1404 \cdot 10^{-2}$	-	$0.4194 \cdot 10^{-2}$	-
80	$0.2025 \cdot 10^{-3}$	2.0	$0.5584 \cdot 10^{-3}$	2.0	$0.3537 \cdot 10^{-3}$	2.0	$0.9903 \cdot 10^{-3}$	2.1
160	$0.4871 \cdot 10^{-4}$	2.1	$0.1322 \cdot 10^{-3}$	2.1	$0.8511 \cdot 10^{-4}$	2.1	$0.2306 \cdot 10^{-3}$	2.1
320	$0.9789 \cdot 10^{-5}$	2.3	$0.2651 \cdot 10^{-4}$	2.3	$0.1712 \cdot 10^{-4}$	2.3	$0.4597 \cdot 10^{-4}$	2.3
N_{cells}	$hu(1-\alpha)$				$hv(\alpha)$			
	L^2 error	p	L^{max} error	p	L^2 error	p	L^{max} error	p
40	$0.3672 \cdot 10^{-4}$	-	$0.1442 \cdot 10^{-3}$	-	$0.1212 \cdot 10^{-4}$	-	$0.3409 \cdot 10^{-4}$	-
80	$0.5911 \cdot 10^{-5}$	2.6	$0.3448 \cdot 10^{-4}$	2.1	$0.1791 \cdot 10^{-5}$	2.8	$0.8054 \cdot 10^{-5}$	2.1
160	$0.1049 \cdot 10^{-5}$	2.5	$0.8471 \cdot 10^{-5}$	2.0	$0.3807 \cdot 10^{-6}$	2.2	$0.2048 \cdot 10^{-5}$	2.0
320	$0.1723 \cdot 10^{-6}$	2.6	$0.2078 \cdot 10^{-5}$	2.0	$0.5115 \cdot 10^{-7}$	2.9	$0.4861 \cdot 10^{-6}$	2.1

Table 3.5: L^2 and L^{max} error for $h(1-\alpha)+b$, $h\alpha+b$, $hu(1-\alpha)$ and $hv\alpha$ using STDGFEM for test case 8.

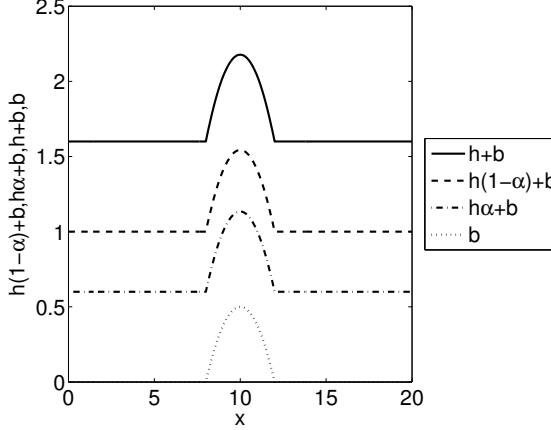


Figure 3.10: Test case 9: steady-state solution for a supercritical two-phase flow calculated using STDGFEM using 320 cells. Shown are the total flow height $h + b$, the flow height due to the fluid phase $h(1 - \alpha)$, the flow height due to the solids phase $h\alpha$ and the topography b .

STDGFEM								
N_{cells}	$h(1 - \alpha) + b$				$h\alpha + b$			
	L^2 error	p	L^{\max} error	p	L^2 error	p	L^{\max} error	p
40	$0.2400 \cdot 10^{-2}$	-	$0.5674 \cdot 10^{-2}$	-	$0.2359 \cdot 10^{-2}$	-	$0.5575 \cdot 10^{-2}$	-
80	$0.6060 \cdot 10^{-3}$	2.0	$0.1402 \cdot 10^{-2}$	2.0	$0.5958 \cdot 10^{-3}$	2.0	$0.1378 \cdot 10^{-2}$	2.0
160	$0.1459 \cdot 10^{-3}$	2.1	$0.3339 \cdot 10^{-3}$	2.1	$0.1434 \cdot 10^{-3}$	2.1	$0.3280 \cdot 10^{-3}$	2.1
320	$0.2933 \cdot 10^{-4}$	2.3	$0.6678 \cdot 10^{-4}$	2.3	$0.2884 \cdot 10^{-4}$	2.3	$0.6561 \cdot 10^{-4}$	2.3

Table 3.6: L^2 and L^{\max} error for $h(1 - \alpha) + b$, $h\alpha + b$, $hu(1 - \alpha)$ and $h\nu\alpha$ using STDGFEM for test case 9.

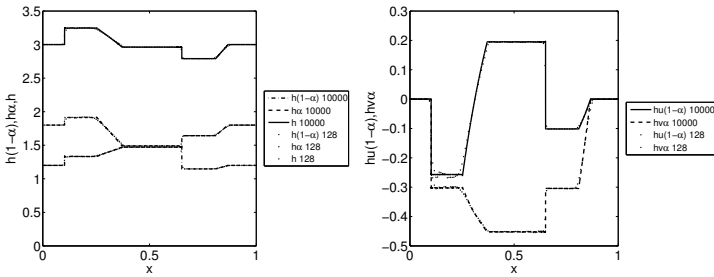
DGFEM								
N_{cells}	L^2 of $h(1-\alpha)$	p	L^2 of $h\alpha$	p	L^2 of $hu(1-\alpha)$	p	L^2 of $h\nu\alpha$	p
32	$0.1238 \cdot 10^{-1}$	-	$0.7030 \cdot 10^{-2}$	-	$0.1263 \cdot 10^{-1}$	-	$0.1384 \cdot 10^{-1}$	-
64	$0.1125 \cdot 10^{-1}$	0.1	$0.5780 \cdot 10^{-2}$	0.3	$0.1155 \cdot 10^{-1}$	0.1	$0.8164 \cdot 10^{-2}$	0.8
128	$0.6231 \cdot 10^{-2}$	0.9	$0.3391 \cdot 10^{-2}$	0.8	$0.7114 \cdot 10^{-2}$	0.7	$0.4465 \cdot 10^{-2}$	0.9
256	$0.4379 \cdot 10^{-2}$	0.5	$0.2751 \cdot 10^{-2}$	0.3	$0.4494 \cdot 10^{-2}$	0.7	$0.3828 \cdot 10^{-2}$	0.2
512	$0.3085 \cdot 10^{-2}$	0.5	$0.1875 \cdot 10^{-2}$	0.6	$0.3536 \cdot 10^{-2}$	0.3	$0.3275 \cdot 10^{-2}$	0.2

Table 3.7: L^2 error and convergence rate for $h(1-\alpha)$, $h\alpha$, $hu(1-\alpha)$ and $h\nu\alpha$ using DGFEM for test case 10. The convergence rates are shown for the solution at $t = 0.175$. With L^2 of U we mean $\|U_N - U_{2N}\|_2$.

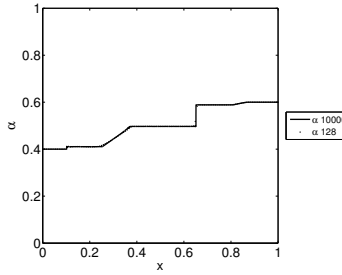
has a solid volume fraction of $\alpha = 0.4$ and the right mixture has a solid volume fraction of $\alpha = 0.6$. At time $t = 0$ we remove the membrane. We want to know how the mixtures behave. We consider the solution on the domain $[0, 1]$. As initial condition we take $U(x, 0) = U_L$ if $x < 0.5$ and $U(x, 0) = U_R$ if $x > 0.5$, where $U_L = [1.8, 1.2, 0, 0, 0]^T$ and $U_R = [1.2, 1.8, 0, 0, 0]^T$. The constants in the computation are taken as $g = 1$ and $\rho = 0.5$. We compute the solution on a domain with 16, 32, 64, 128, 256, 512 or 1024 elements. We consider DGFEM calculations using the linear path $\phi = U_L + \tau(U_R - U_L)$. The solution is determined at $t = 0.175$ using a time step of $\Delta t = 0.0001$. The solutions of $h(1-\alpha)$, $h\alpha$, b and h are depicted in Figure 3.11a, the solutions of $hu(1-\alpha)$ and $h\nu\alpha$ are depicted in Figure 3.11b and the solution of α is depicted in Figure 3.11c in which we compare the solutions on a grid with 128 elements to the solutions computed on a grid with 10000 elements. Apart from some small spurious oscillations obtained on the grid with 128 elements, the solutions compare very well with the solutions obtained on the grid with 10000 elements. Since we do not have an exact solution, we compute the order behavior using the following approach:

$$\frac{\|U_N - U_{2N}\|_2}{\|U_{2N} - U_{4N}\|_2} = 2^p, \quad (3.25)$$

where p is the order of convergence, U_N the solution on a mesh consisting of N cells, and $\|\cdot\|_2$ is the L^2 norm. The order behavior is shown in Table 3.7. Due to the presence of shocks we cannot obtain second order accuracy. Instead we obtain a convergence rate of approximately $\mathcal{O}(h^{1/2})$.



(a) Solution of $h(1 - \alpha)$, $h\alpha$, b and (b) Solution of $hu(1 - \alpha)$ and $hv\alpha$.
 h .



(c) Solution of α .

Figure 3.11: Test case 10. The solution computed on a mesh with 128 elements compared to the solution computed on a mesh with 10000 elements at time $t = 0.175$ using space DGFEM.

3.2.3 River bed evolution under shallow flows

Test case 11: hydraulic and sediment transport through a contraction

Consider the non-dimensional form of the shallow water equations and the bed evolution equation (for details see Tassi et al. [72, 73]):

$$A_{ir}U_{r,0} + F_{ik,k} + G_{ikr}U_{r,k} = 0, \quad (3.26)$$

where $U = [h, hu_1, hu_2, b]^T$ and:

$$A = \begin{bmatrix} \epsilon & 0 & 0 & 0 \\ 0 & \epsilon & 0 & 0 \\ 0 & 0 & \epsilon & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad F = \begin{bmatrix} hu_1 & hu_2 \\ hu_1^2 + F^{-2}h^2/2 & hu_1u_2 \\ hu_1u_2 & hu_2^2 + F^{-2}h^2/2 \\ |u|^{\beta-1}u_1 & |u|^{\beta-1}u_2 \end{bmatrix},$$

$$G_{k=1} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & F^{-2}h \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad G_{k=2} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & F^{-2}h \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where ϵ is the ratio between the sediment and hydrodynamic discharge and β is a constant. In most rivers far less sediment than water is transported so that $\epsilon \ll 1$. In our calculations we take $\epsilon = 0$, $\beta = 3$ and $F = 0.1$.

An extra complication in this test case is matrix A in (3.26) since it is a singular matrix when $\epsilon = 0$. This is a problem when deriving the numerical flux and the wave speeds S_L and S_R . However, since we solve the system of algebraic equations in pseudo-time, we need the numerical flux on the space faces only in the space-time normal direction. To obtain the numerical flux on a fixed grid, note that the normal in the time direction is 0, so that, after augmenting with a pseudo time derivative, (3.26) is changed to:

$$\partial_\tau U_r + F_{ik,k} + G_{ikr}U_{r,k} = 0. \quad (3.27)$$

The numerical flux is then determined in the space normal direction to a face (see Tassi et al. [72, 73]). For one dimensional numerical examples solving (3.26) including convergence rates with space and space-time DGFEM we refer to Tassi et al. [73].

In this test case we consider hydraulic and sediment transport through a contraction. The mesh considered is given in Figure 3.12. In Tassi et al. [73] we show results of this test case using space DGFEM and here we use space-time DGFEM. The physical time step is $\Delta t = 0.0001$. For the pseudo-time stepping, the pseudo-time CFL number is $CFL^{pseudo} = 0.8$. Furthermore, if residuals converged to a tolerance of 10^{-6} in the pseudo-time integration, we considered the system to be solved. In Figures 3.13, 3.14 and 3.15 we show the mass

flow hu , hv and the bed elevation b at time $t = 0.005$ which in physical time corresponds to a few months. As in Kubatko et al. [44], we observe that the bed experiences erosion in the converging part of the channel due to an increase in the flow velocity and the development of a mound in the diverging part of the channel. The results compare qualitatively well with those presented [44] and are the same as we obtained using space DGFEM in Tassi et al. [73].

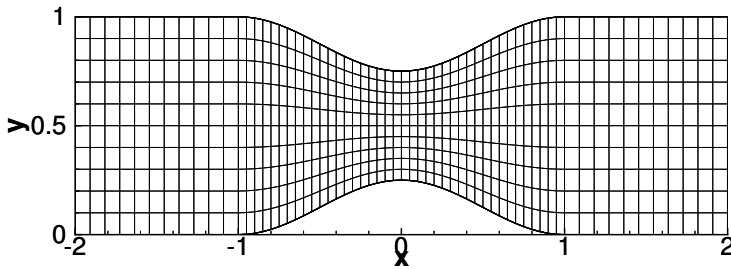


Figure 3.12: *Test case 11: the mesh.*

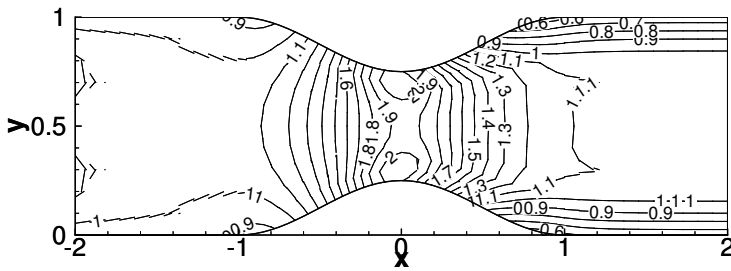


Figure 3.13: *Test case 11: flow and sediment transport in a contraction channel: mass flow $hu(x)$ at time $t = 0.005$.*

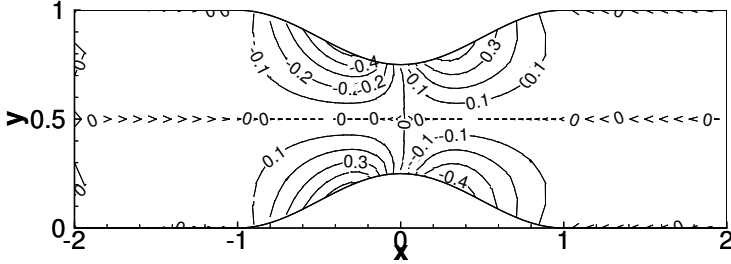


Figure 3.14: *Test case 11: flow and sediment transport in a contraction channel: mass flow $hv(x)$ at time $t = 0.005$.*

3.3 Effect of the path in phase space on the numerical solution

3.3.1 Polynomial paths

In the numerical test cases discussed in the previous sections a linear path was taken: $\phi = U^L + \tau(U^R - U^L)$. In this section, we will investigate the effect of different paths on our numerical results. To determine this effect we again consider test case 10 in Section 3.2.2 for which we expect to find the biggest effect of the path due to the shock waves in the solution. We use the following paths and note that in one dimension property (H4) can be neglected:

$$\begin{aligned}
 \phi_{2v1} &= U^L + \tau^2(U^R - U^L), & \phi_{2v2} &= U^R + (1 - \tau)^2(U^L - U^R), \\
 \phi_{5v1} &= U^L + \tau^5(U^R - U^L), & \phi_{5v2} &= U^R + (1 - \tau)^5(U^L - U^R), \\
 \phi_{20v1} &= U^L + \tau^{20}(U^R - U^L), & \phi_{20v2} &= U^R + (1 - \tau)^{20}(U^L - U^R).
 \end{aligned}
 \tag{3.28}$$

In Figure 3.16, $h(1 - \alpha)$, $h\alpha$, b and h are shown on the whole domain and also a zoom-in on the left shock wave. The deviations shown in these figures are approximately also seen in the mass flow variables and the void fraction.

In these computations it is important to have a good numerical integration scheme to approximate the path integral. Incorrectly approximating the path integral results in solutions having incorrect faster or slower shock speeds. A two-point Gauss integration scheme is sufficient when taking ϕ linear or when using ϕ_{2v1} and ϕ_{2v2} . For the other paths we split the domain $[0, 1]$ into 8 non-intersecting uniform intervals and within each interval we evaluate the integral in the two Gauss points corresponding to that particular interval. To conclude

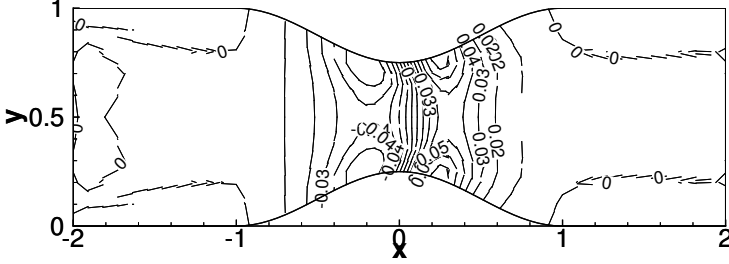


Figure 3.15: *Test case 11: flow and sediment transport in a contraction channel: bottom profile $b(x)$ at time $t = 0.005$.*

for this test case, when properly integrated any choice of paths in (3.28) leads to the same numerical solution with only minor differences.

3.3.2 Toumi paths

In this section we will consider paths similar to those chosen in Toumi [75]. These paths are different from those of the previous section in that these paths are C^0 . We will compare the solutions determined with the following five paths with the solution determined with a linear path:

$$\begin{aligned}
 \phi_{T1}(\tau; U^L, U^R) &= \begin{cases} (U_1^L + 2\tau(U_1^R - U_1^L), U_2^L, U_3^L, U_4^L, U_5^L), & \text{for } \tau \in [0, \frac{1}{2}] \\ (U_1^R, U_2^L + (2\tau - 1)(U_2^R - U_2^L), U_3^L + (2\tau - 1)(U_3^R - U_3^L), & \text{for } \tau \in [\frac{1}{2}, 1], \\ U_4^L + (2\tau - 1)(U_4^R - U_4^L), U_5^L + (2\tau - 1)(U_5^R - U_5^L)), & \end{cases} \\
 \phi_{T2}(\tau; U^L, U^R) &= \begin{cases} (U_1^L, U_2^L + 2\tau(U_2^R - U_2^L), U_3^L, U_4^L, U_5^L), & \text{for } \tau \in [0, \frac{1}{2}] \\ (U_1^L + (2\tau - 1)(U_1^R - U_1^L), U_2^R, U_3^L + (2\tau - 1)(U_3^R - U_3^L), & \text{for } \tau \in [\frac{1}{2}, 1], \\ U_4^L + (2\tau - 1)(U_4^R - U_4^L), U_5^L + (2\tau - 1)(U_5^R - U_5^L)), & \end{cases} \\
 \phi_{T3}(\tau; U^L, U^R) &= \begin{cases} (U_1^L, U_2^L, U_3^L + 2\tau(U_3^R - U_3^L), U_4^L, U_5^L), & \text{for } \tau \in [0, \frac{1}{2}] \\ (U_1^L + (2\tau - 1)(U_1^R - U_1^L), U_2^L + (2\tau - 1)(U_2^R - U_2^L), U_3^R, & \text{for } \tau \in [\frac{1}{2}, 1], \\ U_4^L + (2\tau - 1)(U_4^R - U_4^L), U_5^L + (2\tau - 1)(U_5^R - U_5^L)), & \end{cases} \\
 \phi_{T4}(\tau; U^L, U^R) &= \begin{cases} (U_1^L, U_2^L, U_3^L, U_4^L + 2\tau(U_4^R - U_4^L), U_5^L), & \text{for } \tau \in [0, \frac{1}{2}] \\ (U_1^L + (2\tau - 1)(U_1^R - U_1^L), U_2^L + (2\tau - 1)(U_2^R - U_2^L), & \text{for } \tau \in [\frac{1}{2}, 1], \\ U_3^L + (2\tau - 1)(U_3^R - U_3^L), U_4^R, U_5^L + (2\tau - 1)(U_5^R - U_5^L)), & \end{cases} \\
 \phi_{T5}(\tau; U^L, U^R) &= \begin{cases} (U_1^L, U_2^L, U_3^L, U_4^L, U_5^L + 2\tau(U_5^R - U_5^L)), & \text{for } \tau \in [0, \frac{1}{2}] \\ (U_1^L + (2\tau - 1)(U_1^R - U_1^L), U_2^L + (2\tau - 1)(U_2^R - U_2^L), & \text{for } \tau \in [\frac{1}{2}, 1]. \\ U_3^L + (2\tau - 1)(U_3^R - U_3^L), U_4^L + (2\tau - 1)(U_4^R - U_4^L), U_5^R), & \end{cases}
 \end{aligned} \tag{3.29}$$

In the implementation the integrals are computed using a two-point Gauss integration rule. In Figure 3.17, $h(1 - \alpha)$, $h\alpha$, b and h are shown on the whole

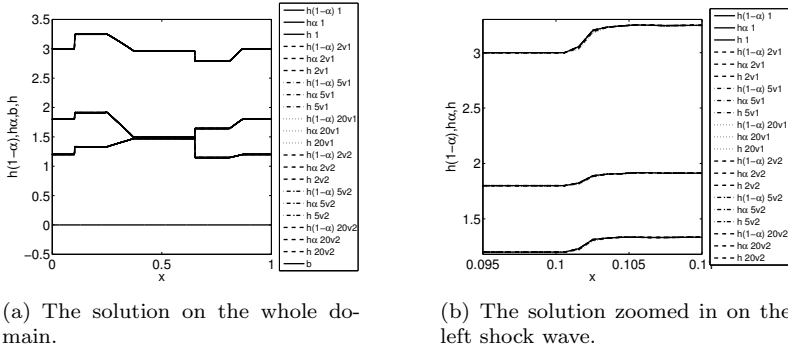
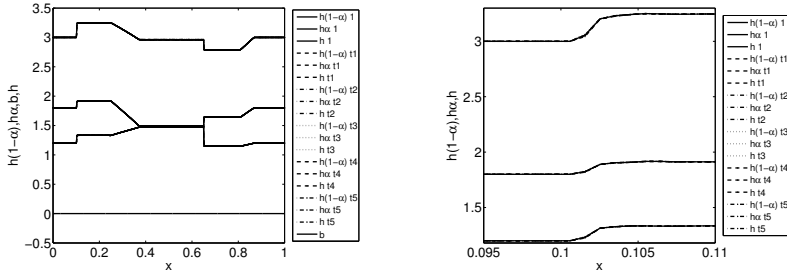


Figure 3.16: Solution of $h(1 - \alpha)$, $h\alpha$, b and h calculated on a mesh with 1024 elements at time $t = 0.175$ using the paths defined in (3.28).

domain and also zoomed in on the left shock wave. The deviations shown in these figures are approximately also seen in the mass flow variables and the void fraction. We see that the final solution determined with the paths given in (3.29) are all very similar. The choice of one of these paths does not have a big effect on the final solution compared to the linear path.

3.3.3 Refining the mesh

As a final check we further refine our mesh. We will calculate the solution on a mesh with 10000 elements. We only do this for the linear path, ϕ_{20v1} (see (3.28)) and ϕ_{T1} (see (3.29)) and compare these solutions with the numerical solution determined with the linear path on a mesh with 1024 elements. In Figure 3.18, $h(1 - \alpha)$, $h\alpha$, b and h are shown on the whole domain and also zoomed in on the left shock wave. The deviations shown in these figures are approximately also seen in the mass flow variables and the void fraction. To obtain these figures, the integral of the nonconservative product for each path was evaluated differently. For the linear path a two point Gauss integration scheme was used for the whole domain $[0, 1]$. For the path ϕ_{20v1} we divided the domain $[0, 1]$ into 16 nonintersecting uniform domains and within each domain we used again a two point Gauss integration scheme. For the path ϕ_{T1} the domain $[0, 1]$ was divided into 8 nonintersecting uniform domains and within each domain we used a two point Gauss integration scheme. As we see in these figures, the differences in the numerical solution for all the paths are minimal. The slight differences in the shock speed are more likely to be caused by the numerical integration scheme than the difference in the path. If we were to



(a) The solution on the whole domain.

(b) The solution zoomed in on the left shock wave.

Figure 3.17: Solution of $h(1 - \alpha)$, $h\alpha$, b and h calculated on a mesh with 1024 elements at time $t = 0.175$ using the paths defined in (3.29).

determine the numerical solution using the path ϕ_{20v1} by dividing the domain $[0, 1]$ into 8 nonintersecting uniform domains instead of 16, the differences in shock speed in comparison to the other paths will increase, so it is important to have a good approximation for the integral of the nonconservative product. We conclude that it is important to have a good numerical integration scheme to approximate the path integral. Using a linear path, a two points Gauss integration scheme, without refinement, suffices. We saw that it does not matter which path is chosen, but choosing the linear path, due to the simple integration scheme, is by far the cheapest and easiest choice.

3.3.4 Contact waves

In Parés and Castro [58] a test case is presented for the shallow water equations in which they state that the selection of the path is critical in order to satisfactorily capture stationary contact discontinuities related to bottom discontinuities (see [58]). We repeat this test case. Consider the shallow water equations given by (3.6) and (3.7). Following Parés and Castro [58], the initial condition is given by $U = U^L$ if $x \leq 0$ and $U = U^R$ if $x > 0$, where $U^L = [0, 1, \sqrt{2g}]^T$ and $U^R = [-1, 0.6527036446614, \sqrt{2g}]^T$ if $x > 0$, where $g = 9.81$. These initial conditions are such that the states U^L and U^R are connected by an entropic contact discontinuity (see [58]). The boundary conditions are given by: $b(-5, t) = 0$, $h(-5, t) = 1$, $hu(-5, t) = \sqrt{2g}$, $b(5, t) = -1$. The remaining boundary data are set equal to the data calculated at the boundary from inside the domain. The steady state solution is calculated using a physical time step of $\Delta t = 10^{21}$ and a pseudo CFL number of $CFL^{pseudo} = 0.8$. We consider the solution on a mesh

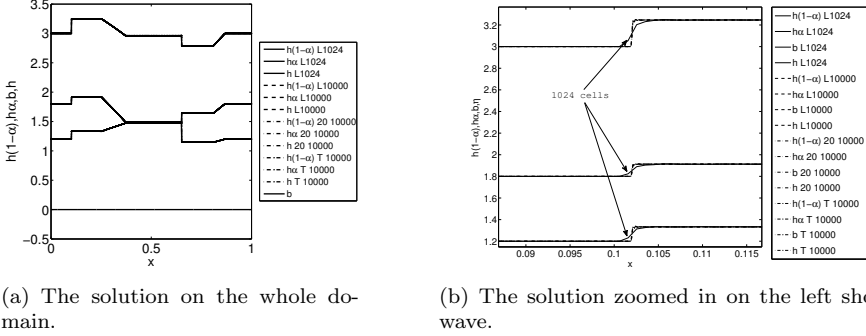


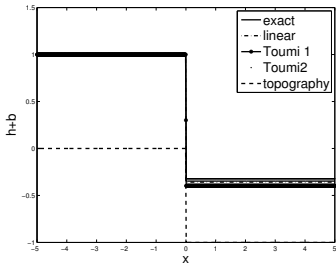
Figure 3.18: Solution of $h(1 - \alpha)$, $h\alpha$, b and h calculated on a mesh with 10000 elements at time $t = 0.175$ using the linear path, ϕ_{20v1} and ϕ_{T1} .

with 1000 elements on which the contact wave falls exactly on a face, and on a mesh with 999 elements so that the contact wave falls exactly in the middle of an element. The effect of three paths are considered, namely the linear path $\phi(\tau; U^L, U^R) = U^L + \tau(U^R - U^L)$ and two Toumi like paths:

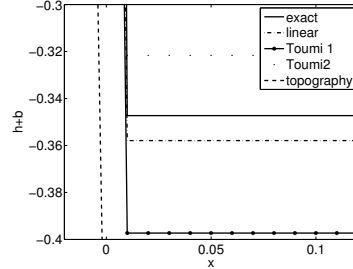
$$\phi_{T1}(\tau; U^L, U^R) = \begin{cases} (U_1^L + 2\tau(U_1^R - U_1^L), U_2^L), & \text{for } \tau \in [0, \frac{1}{2}] \\ (U_1^R, U_2^L + (2\tau - 1)(U_2^R - U_2^L)), & \text{for } \tau \in [\frac{1}{2}, 1] \end{cases}$$

$$\phi_{T2}(\tau; U^L, U^R) = \begin{cases} (U_1^L, U_2^L + 2\tau(U_2^R - U_2^L)), & \text{for } \tau \in [0, \frac{1}{2}] \\ (U_1^L + (2\tau - 1)(U_1^R - U_1^L), U_2^R), & \text{for } \tau \in [\frac{1}{2}, 1]. \end{cases}$$

Note that the path for $U_3 = hu$ is irrelevant since the nonconservative product for the shallow water equations only involve b and h . The solution on the mesh with 1000 elements is shown in Figure 3.19 and the solution on the mesh with 999 elements is shown in Figure 3.20. We see that the solution of a steady contact discontinuity experiences a similar dependence on the path as observed by Parés and Castro [58], also after refining the mesh to 10000 and 9999 elements, respectively. The numerical dissipation introduced when the contact discontinuity is not exactly at an element face has, however, a strong regularizing effect (compare Figures 3.19 and 3.20) and significantly reduces the dependence of the solution on the path. This effect will even be stronger in multi-dimensional problems since the discontinuities then rarely coincide with mesh lines, but it is always important to check the dependence of the solution on the chosen path.

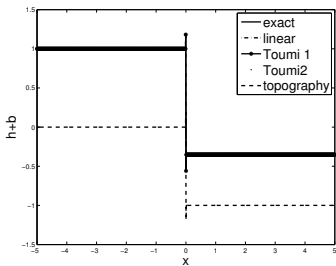


(a) Comparison of the computed water level at the stationary state with the exact solution

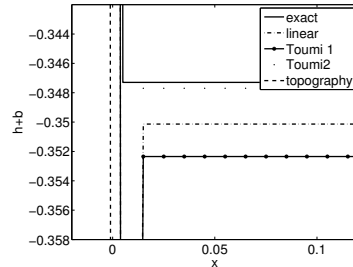


(b) Zoom of the water level right of the discontinuity in the topography.

Figure 3.19: A comparison of the computed solution of a contact discontinuity related to the discontinuous topography with the exact solution. The solution was computed on a mesh with 1000 elements using a physical time step of $\Delta t = 10^{21}$ and a pseudo CFL number $CFL^{pseudo} = 0.8$.



(a) Comparison of the computed water level at the stationary state with the exact solution



(b) Zoom of the water level right of the discontinuity in the topography.

Figure 3.20: A comparison of the computed solution of a contact discontinuity related to the discontinuous topography with the exact solution. The solution was computed on a mesh with 999 elements using a physical time step of $\Delta t = 10^{21}$ and a pseudo CFL number $CFL^{pseudo} = 0.8$.

Discontinuous Galerkin finite element method for shallow two-phase flows

In this chapter we present a discontinuous Galerkin finite element method for a depth-averaged two-phase flow model. This model contains nonconservative products for which we developed a discontinuous Galerkin finite element formulation in Chapter 2. We qualitatively validate the model against a laboratory experiment and show the abilities of the model to capture physical phenomena. To be able to perform these test cases, a WENO slope limiter is investigated in conjunction with a discontinuity detector to detect regions where spurious oscillations appear.

4.1 Depth-averaged two-phase flows

In shallow flows, the characteristic height H of the flow is typically much smaller than its characteristic length L , $H/L = \varepsilon \ll 1$. Variations in the vertical are small and we can simplify the governing equations by averaging the flow over the depth. In doing so, depth-averaged quantities are assumed to be independent of the vertical coordinate, at leading order in ε . In this section we introduce the depth-averaged two-phase flow equations derived by Le [45]. Note that the depth-averaged two-phase flow equations derived by Le [45] are slightly different from the depth-averaged two-phase flow equations derived by Pitman and Le [62]. The difference is that the momentum of the mixture of the Le model can be written in flux conservative form, while this is not the case for the

momentum of the mixture of the Pitman and Le model.

Le [45] derived a depth-averaged flow model by depth-averaging the three dimensional continuum model for two-phase flows as derived by Jackson [38] (see Appendix B for the three dimensional model). Using the summation convention on repeated indices and the comma notation to denote partial differentiation, the scaled non-dimensional depth-averaged flow model is:

$$U_{i,t} + F_{ik,k} + G_{ikr}V_{r,k} = S_i, \quad i, r = 1, \dots, 6, \quad k = 1, 2. \quad (4.1)$$

Note that $G_{ikr}V_{r,k}$ is a nonconservative product. In (4.1)

$U = [h(1 - \alpha), h\alpha, h\alpha v_i, h(1 - \alpha)u_i]^T$, $V = [h, \alpha, v_i, u_i]^T$ and

$$F_k = \begin{bmatrix} h(1 - \alpha)u_k \\ h\alpha v_k \\ h\alpha v_i v_k + \varepsilon(1 - \rho)\varphi_{ik}\alpha\frac{1}{2}g_3 h^2 \\ h(1 - \alpha)u_i u_k \end{bmatrix}, \quad G_k = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \varepsilon\rho\alpha g_3 h & 0 & 0 & 0 \\ \varepsilon(1 - \alpha)g_3 h & 0 & 0 & 0 \end{bmatrix},$$

$$S = \begin{bmatrix} 0 \\ 0 \\ (1 - \rho)(-\varepsilon\varphi_{ik}\partial_k b + \varphi_{i3})\alpha g_3 h + h F_i^D + g_i h\alpha - \rho\alpha C_D |u| u_i / \varepsilon - \varepsilon\rho h\alpha g_3 \partial_i b \\ -\varepsilon(1 - \alpha)g_3 h \partial_i b - h F_i^D / \rho + h(1 - \alpha)g_i - (1 - \alpha)C_D |u| u_i / \varepsilon \end{bmatrix}.$$

Note that compared with the model by Le [45], we have added extra friction terms with the drag coefficient C_D as a leading order turbulence parameterization.

The orientation of the Cartesian coordinate system is shown in Figure 4.1 in which θ is the angle of the x_1 - x_2 plane with the horizontal. The depth-averaged quantities in the above model are constant in the x_3 direction and are the particle volume fraction α , the fluid velocity vector u and the solids velocity vector v . The flow depth is given by h and the bottom topography by b . The constants $\varepsilon = H/L$ and $\rho = \rho^f / \rho^s$ represent the height to length ratio of the flow and the ratio between the fluid density ρ^f and the solids density ρ^s , respectively. The gravity vector is given by $\vec{g} = [g_1, g_2, -g_3]^T$ in which g_3 is the vertical component of the gravity (see Figure 4.1) and C_D is a drag coefficient. The above quantities are all scaled and dimensionless. To obtain the variables in dimensional form, denoted by $(\cdot)^*$, we have used the following scalings: $[x^*, y^*] = L[x, y]$, $t^* = \sqrt{L/g^*}t$, $[u_1^*, u_2^*] = \sqrt{g^*L}[u_1, u_2]$, $[v_1^*, v_2^*] = \sqrt{g^*L}[v_1, v_2]$, $v_T^* = \sqrt{g^*L}v_T$, $[g_1^*, g_2^*, g_3^*]^T = g^*g[\sin(\theta), 0, \cos(\theta)]^T$ with g the gravity constant.

A closure needs to be given for the drag function F^D and we follow Pitman

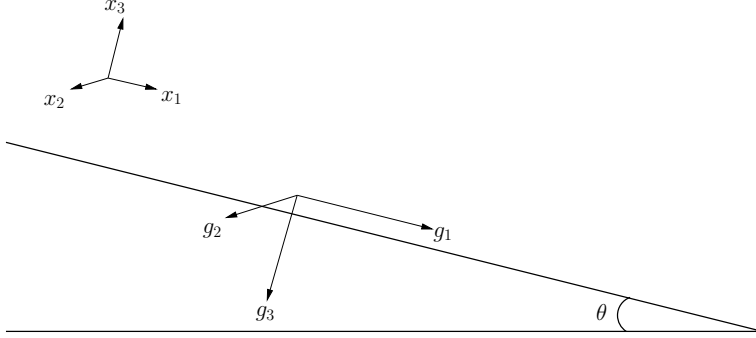


Figure 4.1: Orientation of the coordinate system and the gravity vector.

and Le [62] by taking $F_i^D = \beta(u_i - v_i)$ in which β is given by:

$$\beta = \frac{(1 - \rho)\alpha}{v_T(1 - \alpha)^n}, \quad n = \begin{cases} 3.65 & \text{for } Re_t < 0.2, \\ 4.35Re_t^{-0.03} - 1 & \text{for } 0.2 < Re_t < 1, \\ 4.45Re_t^{-0.1} - 1 & \text{for } 1 < Re_t < 500, \\ 1.39 & \text{for } 500 < Re_t, \end{cases}$$

where $Re_t = d\rho_f v_T / \mu_f$, in which d is the particle diameter, ρ^f the fluid density, μ_f the fluid viscosity and v_T the terminal velocity of an isolated particle falling in the fluid. We remark that as $1 - \rho$ increases, the drag function F^D makes the system (4.1) increasingly stiffer. We are, however, interested in the case where ρ is approximately 0.9. In this situation the model does not have stiff source terms and no special algorithms are needed to deal with stiffness.

The functions φ were introduced by Pitman and Le [62] to relate basal and diagonal shear stresses to the normal stress in the solids phase stress tensor in the 3-dimensional two-phase model before depth-averaging. The functions φ are given by:

$$\begin{aligned} \varphi_{i3} &= -\frac{v_i}{\|v\|} \tan(\phi_{bed}), \quad i = 1, 2, & \varphi_{ii} &= k^\mp, \quad i = 1, 2 \\ \varphi_{12} &= -\text{sign}(\partial_2 v_1) \sin(\phi_{int}) k^\mp, & \varphi_{21} &= -\text{sign}(\partial_1 v_2) \sin(\phi_{int}) k^\mp, \\ k^\mp &= 2 \frac{1 \mp \sqrt{1 - \cos^2(\phi_{int})(1 + \tan^2(\phi_{bed}))}}{\cos^2(\phi_{int})} - 1, \end{aligned}$$

in which the “ $-$ ” in the “ \mp ” applies when $\partial_k v_k > 0$ and the “ $+$ ” applies when $\partial_k v_k < 0$. Furthermore, $\|\cdot\|$ is the Euclidean norm, ϕ_{int} is the internal angle of

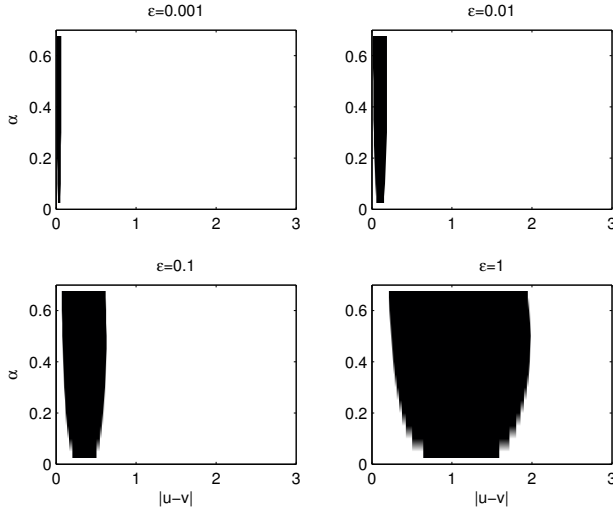


Figure 4.2: *Regimes of hyperbolicity for the depth-averaged model. For the values of α and $|u - v|$ in the shaded area the model is elliptic.*

friction, which measures how layers of solid particles slide over one another and ϕ_{bed} is the basal angle of friction, indicating how easily solid particles slide over the bottom [35].

To determine whether the depth averaged model is hyperbolic, we need to determine their eigenvalues. If all eigenvalues are real and distinct, the model is hyperbolic. Deriving the eigenvalues for the depth-averaged model is not trivial, so eigenvalues are computed numerically for a number of given parameters. Consider the case in which the topography is flat, $b = 0$. We take $h = 1$, $\rho = 0.9$, $g_3 = 1$ and we assume $k^\mp = k^-$. Furthermore, we take $\phi_{bed} = 14.75^\circ$ and $\phi_{int} = 24.5^\circ$ which hold for fine glass particles [7]. For different height to length ratios, ranging from $\varepsilon = 0.001$ to $\varepsilon = 1$, we determine the eigenvalues as a function of the particle volume fraction α and the absolute difference between the phase velocities $|u - v|$. In Figure 4.2 we show for which values of α and $|u - v|$ the depth-averaged model is not hyperbolic (in the shaded areas some of the eigenvalues are not real). We see that the region for which the model is not hyperbolic decreases as ε decreases. In this chapter we are only interested in cases where the model is hyperbolic. When the model is not hyperbolic, a different numerical approach needs to be introduced which is not treated in this chapter.

4.2 The DGFEM discretization

In this section we present a space-time DGFEM formulation for the depth-averaged two-phase flow model. We remark however that the space DGFEM formulation is very similar and for some of the numerical test cases we will apply space DGFEM. For more on space DGFEM we refer to Appendix A and Cockburn and Shu [18].

We consider the solution in the open space-time domain $\mathcal{E} \subset \mathbb{R}^3$ using the definition of the space-time elements as given in Section 2.2.1. Furthermore, in this chapter, we consider the solution only on a fixed grid. We consider linear approximations of $U(t, \bar{x}) \in \mathbb{R}^6$ and test functions $W(t, \bar{x}) \in \mathbb{R}^6$ in the function spaces defined in Section 2.2.2.

4.2.1 Basis functions

Polynomial approximations for the trial function U and the test functions W in each element $\mathcal{K} \in \mathcal{T}_h^n$ are introduced as:

$$U(t, \bar{x})|_{\mathcal{K}} = \hat{U}_m \psi_m(t, \bar{x}) \text{ and } W(t, \bar{x})|_{\mathcal{K}} = \hat{W}_l \psi_l(t, \bar{x}), \quad (4.2)$$

with ψ_m the basis functions, $\bar{x} \in \mathbb{R}^2$, and expansion coefficients \hat{U}_m and \hat{W}_l , respectively, for $m, l = 0, 1, 2, 3$. The basis functions ψ_m are given by $\psi_0 = 1$ and $\psi_m = \varphi_m(t, \bar{x})$ for $m = 1, 2, 3$ where the functions $\varphi_m(x)$ in element \mathcal{K} are related to the basis functions $\hat{\varphi}_m(\xi)$, with $\hat{\varphi}_m(\xi) \in P^1(\hat{\mathcal{K}})$ and ξ the local coordinates in the master element $\hat{\mathcal{K}}$, through the mapping $G_{\mathcal{K}}: \varphi_m = \hat{\varphi}_m \circ G_{\mathcal{K}}^{-1}$.

4.2.2 The weak formulation

Due to the nonconservative products (4.1) cannot be transformed into divergence form. This causes problems once the solution becomes discontinuous, because the weak solution in the classical sense of distributions then does not exist. Consequently, standard space-time DGFEM discretizations cannot be applied. In Chapter 2 we derived a discontinuous Galerkin finite element weak formulation for general hyperbolic equations with nonconservative products and we apply this weak formulation here as well.

We refer to Chapter 2 for the derivation of the weak formulation for (4.1). The main criterium posed on the weak formulation is that the formulation must reduce to that for the conservative system if the system of nonconservative partial differential equations can be transformed into conservative form. The weak formulation for (4.1) is given by:

Find a $U \in V_h$ such that for all $W \in V_h$:

$$\begin{aligned}
0 &= \sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} (-W_{i,0} U_i - W_{i,k} F_{ik} + W_i G_{ikr} V_{r,k} - W_i S_i) dK \\
&+ \sum_{\mathcal{K} \in \mathcal{T}_h^n} \left(\int_{K(t_{n+1}^-)} W_i^L U_i^L dK - \int_{K(t_n^+)} W_i^L U_i^R dK \right) \\
&+ \sum_{\mathcal{S} \in \mathcal{S}^n} \int_{\mathcal{S}} (W_i^L - W_i^R) \widehat{P}_i^{nc} d\mathcal{S} \\
&+ \sum_{\mathcal{S} \in \mathcal{S}^n} \int_{\mathcal{S}} \{W_i\} \left(\int_0^1 G_{ikr}(\phi(\tau; U^L, U^R)) \frac{\partial \phi_r}{\partial \tau}(\tau; U^L, U^R) d\tau \bar{n}_k^L \right) d\mathcal{S}.
\end{aligned} \tag{4.3}$$

The last term makes it different from standard discontinuous Galerkin finite element formulations. It is needed to introduce a measure for the nonconservative product where U is discontinuous. Note that an extra function, $\phi(\tau; U_L, U_R)$, has been introduced to deal with the regularization of U across the discontinuity. In Chapter 3 the effect of the choice of $\phi(\tau; U_L, U_R)$ on the numerical solution was investigated. We concluded that the numerical diffusion has a regularizing effect across discontinuities, which significantly reduces the dependence of the solution on $\phi(\tau; U_L, U_R)$, so that often it does not matter in practice how $\phi(\tau; U_L, U_R)$ is chosen. We adopt a linear path: $\phi(\tau; U_L, U_R) = U_L + \tau(U_R - U_L)$. Furthermore, we use here the NCP numerical flux $\widehat{P}^{nc}(U^L, U^R, \bar{n}^L)$ designed in Chapter 2 for systems containing nonconservative products as a generalization of the HLL flux [74]. The NCP numerical flux $\widehat{P}^{nc}(U^L, U^R, \bar{n}^L)$ reads:

$$\widehat{P}_i^{nc}(U_L, U_R, \bar{n}^L) = \begin{cases} F_{ik}^L \bar{n}_k^L - \frac{1}{2} \int_0^1 G_{ikr}(\bar{\phi}(\tau; U_L, U_R)) \frac{\partial \bar{\phi}_r}{\partial \tau}(\tau; U_L, U_R) d\tau \bar{n}_k^L & \text{if } S_L > 0, \\ \{F_{ik}\} \bar{n}_k^L + \frac{1}{2} (S_R \bar{U}_i^* + S_L \bar{U}_i^* - S_L U_i^L - S_R U_i^R) & \text{if } S_L < 0 < S_R, \\ F_{ik}^R \bar{n}_k^L + \frac{1}{2} \int_0^1 G_{ikr}(\bar{\phi}(\tau; U_L, U_R)) \frac{\partial \bar{\phi}_r}{\partial \tau}(\tau; U_L, U_R) d\tau \bar{n}_k^L & \text{if } S_R < 0, \end{cases} \tag{4.4}$$

with \bar{U}^* given by:

$$\bar{U}_i^* = \frac{S_R U_i^R - S_L U_i^L + (F_{ik}^L - F_{ik}^R) \bar{n}_k^L}{S_R - S_L} - \frac{1}{S_R - S_L} \int_0^1 G_{ikr}(\phi(\tau; U_L, U_R)) \frac{\partial \phi_r}{\partial \tau}(\tau; U_L, U_R) d\tau \bar{n}_k^L. \tag{4.5}$$

Note that the first terms on the right hand side of (4.4) are in each case the upwind or unstable numerical fluxes. The wave speeds S_L and S_R in the numerical flux are usually approximated by the minimum and maximum eigenvalues

of the Jacobian matrix. The characteristic polynomial of the Jacobian matrix of the depth-averaged model, $\partial F/\partial U + G$ is $c(\lambda) = (\lambda - q_v)(\lambda - q_u)p(\lambda)$ in which $p(\lambda) = \lambda^4 + a_1\lambda^3 + a_2\lambda^2 + a_3\lambda + a_4$, where

$$\begin{aligned}
a_1 &= -2(q_u + q_v), \\
a_2 &= q_u^2 + q_v^2 + 4q_uq_v - \varepsilon g_3 h(1 - \alpha + \rho\alpha), \\
&\quad - \frac{1}{2}\varepsilon g_3 h(1 - \rho)(1 + \alpha)(\varphi_{11}n_1^2 + \varphi_{22}n_2^2 + \varphi_{12}n_1n_2 + \varphi_{21}n_1n_2) \\
a_3 &= -2q_uq_v(q_u + q_v) + 2q_v\varepsilon g_3 h(1 - \alpha) + 2\varepsilon\rho g_3\alpha hq_u \\
&\quad + 2q_u(\frac{1}{2}\varepsilon g_3 h(1 + \alpha)(1 - \rho)(\varphi_{11}n_1^2 + \varphi_{22}n_2^2 + \varphi_{12}n_1n_2 + \varphi_{21}n_1n_2)), \\
a_4 &= q_u^2q_v^2 - q_u^2(\frac{1}{2}h\varepsilon g_3(1 - \rho)(1 + \alpha)(\varphi_{11}n_1^2 + \varphi_{22}n_2^2 + \varphi_{12}n_1n_2 + \varphi_{21}n_1n_2)) \\
&\quad + \frac{1}{2}\varepsilon^2 g_3^2 h^2(1 - \rho)(1 - \alpha)(\varphi_{11}n_1^2 + \varphi_{22}n_2^2 + \varphi_{12}n_1n_2 + \varphi_{21}n_1n_2) \\
&\quad - q_v^2\varepsilon g_3 h(1 - \alpha) - q_u^2\varepsilon\rho g_3\alpha h.
\end{aligned} \tag{4.6}$$

Two eigenvalues are $\lambda_1 = q_v$ and $\lambda_2 = q_u$. Since explicitly solving the quartic polynomial $p(\lambda) = 0$ yields rather unwieldy relations, we approximate the remaining four eigenvalues.

We approximate $p(\lambda)$ by $\tilde{p}(\lambda) = (\lambda - q^u - \mathcal{A})(\lambda - q^u + \mathcal{A})(\lambda - q^v - \mathcal{B})(\lambda - q^v + \mathcal{B})$ and expand \tilde{p} as $\tilde{p} = \lambda^4 + a_1\lambda^3 + b_2\lambda^2 + b_3\lambda + b_4$ with coefficients:

$$\begin{aligned}
b_2 &= q_u^2 + q_v^2 + 4q_uq_v - (\mathcal{A}^2 + \mathcal{B}^2), \\
b_3 &= 2q_v\mathcal{A}^2 + 2q_u\mathcal{B}^2 - 2q_vq_u(q_u + q_v), \\
b_4 &= q_v^2q_u^2 - q_u^2\mathcal{B}^2 - q_v^2\mathcal{A}^2 + \mathcal{A}^2\mathcal{B}^2.
\end{aligned} \tag{4.7}$$

Note that by choosing

$$\begin{aligned}
\mathcal{A} &= \sqrt{\varepsilon g_3 h(1 - \alpha)}, \\
\mathcal{B} &= \sqrt{\frac{1}{2}h\varepsilon g_3(1 - \rho)(1 + \alpha)(\varphi_{11}n_1^2 + \varphi_{22}n_2^2 + \varphi_{12}n_1n_2 + \varphi_{21}n_1n_2)},
\end{aligned} \tag{4.8}$$

the coefficients a_i and b_i almost match. We approximate the solutions to $p(\lambda)$ now as $\lambda_{3,4} = q_u \pm \mathcal{A}$ and $\lambda_{5,6} = q_v \pm \mathcal{B}$. The error in the approximation of the roots is then proportional to $p(\lambda_{3,4}) = \mathcal{O}(\varepsilon^2)$ and $p(\lambda_{5,6}) = \mathcal{O}(\varepsilon)$.

As mentioned above, $\phi(\tau; U_L, U_R)$ had to be chosen and we adopted $\phi(\tau; U_L, U_R) = U_L + \tau(U_R - U_L)$. This choice of the path presents us the opportunity to exactly determine the integral due to the nonconservative product

in (4.3):

$$\int_0^1 G_{kr}(\phi(\tau; U_L, U_R)) \frac{\partial \phi_r}{\partial \tau}(\tau; U_L, U_R) d\tau \bar{n}_k^L = \begin{bmatrix} 0 \\ 0 \\ -\varepsilon \rho g_3 \llbracket h \rrbracket \int_0^1 \alpha h d\tau \\ -\varepsilon \rho g_3 \llbracket h \rrbracket \int_0^1 \alpha h d\tau \\ -\varepsilon g_3 \llbracket h \rrbracket \int_0^1 (1 - \alpha) h d\tau \\ -\varepsilon g_3 \llbracket h \rrbracket \int_0^1 (1 - \alpha) h d\tau \end{bmatrix}, \quad (4.9)$$

in which

$$\begin{aligned} \int_0^1 \alpha h d\tau &= \frac{1}{3}(\alpha_L h_L + \frac{1}{2}(\alpha_R h_L + \alpha_L h_R) + \alpha_R h_R), \\ \int_0^1 (1 - \alpha) h d\tau &= \llbracket h \rrbracket - \frac{1}{3}(\alpha_L h_L + \frac{1}{2}(\alpha_R h_L + \alpha_L h_R) + \alpha_R h_R). \end{aligned}$$

4.2.3 Pseudo-time stepping

By replacing U and W in the weak formulation (4.3) by their polynomial expansions (4.2), a system of algebraic equations for the expansion coefficients of U is obtained. For each physical time step, the system can be written as:

$$\mathcal{L}(\hat{U}^n; \hat{U}^{n-1}) = 0. \quad (4.10)$$

This system of coupled non-linear equations is solved by adding a pseudo-time derivative of the primitive variables $V = [h, \alpha, v_i, u_i]^T$, hence (4.10) becomes:

$$\mathcal{M} \frac{\partial \hat{V}^n}{\partial \tau} = -\mathcal{L}(\hat{V}^n; \hat{V}^{n-1}), \quad \mathcal{M} = \int_{\mathcal{K}} \phi \frac{\partial U}{\partial V} d\mathcal{K}, \quad (4.11)$$

which is integrated to steady-state in pseudo-time. Following Van der Vegt and Van der Ven [79], we use the explicit Runge-Kutta method for inviscid flow with Melson correction given by:

Algorithm 1 Five-stage explicit Runge-Kutta scheme:

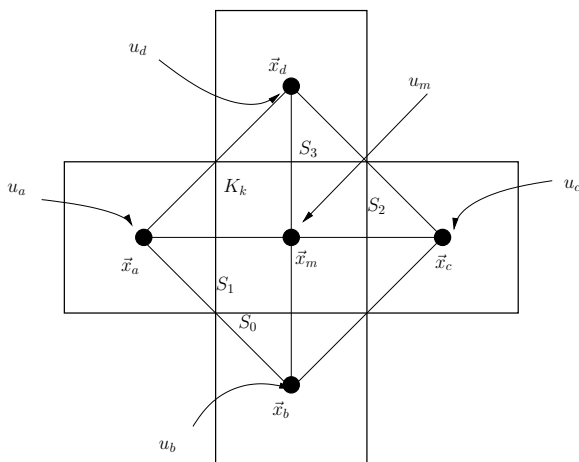
1. Initialize $\hat{Y}^0 = \hat{V}$.

2. For all stages $s = 1$ to 5:

$$(I + \alpha_s \lambda I) \hat{Y}^s = \hat{Y}^0 + \alpha_s \lambda (\hat{Y}^{s-1} - \mathcal{M}^{-1} \mathcal{L}(\hat{Y}^{s-1}; \hat{V}^{n-1})).$$

3. Update $\hat{V} = \hat{Y}^5$.

The coefficient λ is defined as $\lambda = \Delta\tau / \Delta t$, with $\Delta\tau$ the pseudo-time step and Δt the physical time step. The Runge-Kutta coefficients α_s use are [79]: $\alpha_1 = 0.0797151$, $\alpha_2 = 0.163551$, $\alpha_3 = 0.283663$, $\alpha_4 = 0.5$ and $\alpha_5 = 1.0$.

Figure 4.3: *Slope limiter in 2D.*

4.2.4 Slope limiter and discontinuity detector

In numerical discretizations of the weak formulation (4.3), spurious oscillations generally appear near discontinuities. Using the Krivodonova discontinuity detector [43], we apply a slope limiter only near discontinuities to deal with these spurious oscillations. We use the slope limiter given in [51] which we describe briefly here for reasons of clarity.

The idea of the slope limiter is to replace the original polynomial P_0 by a new polynomial P that uses the data u_m of the midpoints of the original element in element K_k and its neighboring elements u_a , u_b , u_c and u_d . Eight polynomials are constructed, 4 Lagrange polynomials, P_i , $i = 1, 2, 3, 4$ and 4 Hermite polynomials P_i , $i = 5, 6, 7, 8$. For the Hermite polynomials we also need the physical gradient of the data in the neighboring elements at the points \vec{x} , i.e., ∇u_a , ∇u_b , ∇u_c and ∇u_d (see Fig. 4.3).

To construct the Lagrange polynomials consider the surface through x_m , x_a and x_b . Name the polynomial through this surface P_1 with $P_1 = \hat{P}_1^a + \hat{P}_1^b x + \hat{P}_1^c y$. The coefficients \hat{P}_1^a , \hat{P}_1^b and \hat{P}_1^c are found by solving:

$$\begin{bmatrix} 1 & x_m & y_m \\ 1 & x_a & y_a \\ 1 & x_b & y_b \end{bmatrix} \begin{bmatrix} \hat{P}_1^a \\ \hat{P}_1^b \\ \hat{P}_1^c \end{bmatrix} = \begin{bmatrix} u_m \\ u_a \\ u_b \end{bmatrix}.$$

In the same way, polynomials P_2 , P_3 and P_4 are constructed by considering the remaining three surfaces.

Each of the four Hermite polynomials are determined by looking at the current element and one of the neighbors, e.g., the first Hermite polynomial, P_5 , is found by looking at the neighboring element sharing face S_0 . In the midpoint x_b , the gradient of the solution is ∇u_b , while the solution in the midpoint of the current element is u_m . The first Hermite polynomial is given by: $P_5 = \hat{P}_5^a + \hat{P}_5^b x + \hat{P}_5^c y$ where:

$$\begin{aligned}\hat{P}_5^a &= u_m - x_b \cdot \nabla u_b, \\ \hat{P}_5^b &= \partial_x u_b \quad \text{in } x_b, \\ \hat{P}_5^c &= \partial_y u_b \quad \text{in } x_b.\end{aligned}$$

In the same way, polynomials P_6 , P_7 and P_8 are constructed by considering the remaining three surfaces.

The linear approximation of the original polynomial is determined just like the Hermite polynomials. In the midpoint x_m , the solution is u_m and the gradient is ∇u_m . The linear approximation is: $P_0 = \hat{P}_0^a + \hat{P}_0^b x + \hat{P}_0^c y$ where:

$$\begin{aligned}\hat{P}_0^a &= u_m - x_m \cdot \nabla u_m, \\ \hat{P}_0^b &= \partial_x u_m \quad \text{in } x_m, \\ \hat{P}_0^c &= \partial_y u_m \quad \text{in } x_m.\end{aligned}$$

Now project P_j , $j = 0, \dots, 8$, onto the DG space and solve for $(\hat{u}_0)_j$, $(\hat{u}_1)_j$ and $(\hat{u}_2)_j$:

$$\begin{bmatrix} \int_{K_k} \psi_0 \psi_0 dK & \int_{K_k} \psi_0 \psi_1 dK & \int_{K_k} \psi_0 \psi_2 dK \\ \int_{K_k} \psi_1 \psi_0 dK & \int_{K_k} \psi_1 \psi_1 dK & \int_{K_k} \psi_1 \psi_2 dK \\ \int_{K_k} \psi_2 \psi_0 dK & \int_{K_k} \psi_2 \psi_1 dK & \int_{K_k} \psi_2 \psi_2 dK \end{bmatrix} \begin{bmatrix} (\hat{u}_0)_j \\ (\hat{u}_1)_j \\ (\hat{u}_2)_j \end{bmatrix} = \begin{bmatrix} \int_{K_k} \psi_0 P_j dK \\ \int_{K_k} \psi_1 P_j dK \\ \int_{K_k} \psi_2 P_j dK \end{bmatrix}.$$

After the polynomial reconstruction is performed, an oscillation indicator is used to assess the smoothness of P_i . The oscillation indicator for the polynomial P_i , $i = 0, \dots, 8$, is defined as $o_i = \|\nabla P_i\|$, with $\|\cdot\|$ the Euclidian norm. The coefficients of the new solution u in element K_k are constructed as the sum of all the polynomials multiplied by a weight, $\hat{u}_q = \sum_{i=0}^8 w_i (\hat{u}_q)_i$, $q = 0, 1, 2$, in which the weights are computed as:

$$w_i = \frac{(\epsilon + o_i(P_i))^{-\gamma}}{\sum_{j=0}^8 (\epsilon + o_j(P_j))^{-\gamma}}, \quad (4.12)$$

where γ is a positive number and ϵ a small number to avoid division by 0. Take for example $\epsilon = 10^{-12}$. The effect of γ and the combination of polynomials (Lagrange and original or Lagrange, original and Hermite) is tested in Section 4.3.2.

The discontinuity detector introduced in Krivodonova et al. [43] defines for each element \mathcal{K}_k^n a measure of the discontinuity \mathcal{I}_k . This will indicate regions where the gradient of a variable \mathcal{V} is large. For the depth-averaged two-phase flow equations, depending on the situation, we choose either $\mathcal{V} = h$ or $\mathcal{V} = \alpha$. The discontinuity detector is given by:

$$\mathcal{I}_k^n = \max(\mathcal{I}_k^n(h), \mathcal{I}_k^n(\alpha)), \quad \mathcal{I}_k^n(\mathcal{V}) = \frac{\sum_{\mathcal{S}_m \in \partial \mathcal{K}_k^n} \int_{\mathcal{S}_m} |\mathcal{V}^R - \mathcal{V}^L| d\mathcal{S}}{h_{\mathcal{K}}^{(p+1)/2} |\partial \mathcal{K}_k^n| \|\mathcal{V}\|_{\infty}}, \quad (4.13)$$

where $h_{\mathcal{K}}$ is the cell measure defined as the radius of the largest circumscribed circle in the element \mathcal{K}_k^n , p the polynomial order, $|\partial \mathcal{K}_k^n|$ the surface area of the element and $\|\cdot\|_{\infty}$ the maximum norm. The solution is estimated [43] to be smooth when $\mathcal{I}_k < 1$ and non-smooth when $\mathcal{I}_k > 1$.

4.3 Verification

4.3.1 Sub- and supercritical flow over a bump

We consider the 1D steady-state solution of sub- and supercritical flow over a bump (see also Chapter 3). This is a popular test case to verify shallow water codes [14, 31, 47, 71, 87] and we extend the test case to the depth-averaged two-phase flow model. For this test case we consider:

$$\begin{aligned} \begin{bmatrix} h(1-\alpha) \\ h\alpha \\ h\alpha v \\ h(1-\alpha)u \\ b \end{bmatrix}_t + \begin{bmatrix} h(1-\alpha)u \\ h\alpha v \\ h\alpha v^2 + \frac{1}{2}\varepsilon(1-\rho)\varphi_{11}g_3h^2\alpha \\ h(1-\alpha)u^2 \\ 0 \end{bmatrix}_x \\ + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ G_{31} & 0 & 0 & 0 & G_{35} \\ G_{41} & 0 & 0 & 0 & G_{45} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} h \\ \alpha \\ v \\ u \\ b \end{bmatrix}_x = \begin{bmatrix} 0 \\ 0 \\ S_3 \\ S_4 \\ 0 \end{bmatrix}, \quad (4.14) \end{aligned}$$

where

$$\begin{aligned} G_{31} &= \varepsilon\rho\alpha g_3 h, & G_{35} &= \varepsilon(1-\rho)\varphi_{11}g_3 h\alpha + \varepsilon\rho h\alpha g_3, \\ G_{41} &= \varepsilon(1-\alpha)g_3 h, & G_{45} &= \varepsilon(1-\alpha)g_3 h, \\ S_3 &= h F^D, & S_4 &= -h F^D/\rho. \end{aligned}$$

Note that we take the given topography to be formally unknown in the system. This leads to a well-balanced scheme (see Chapter 3). Let the upstream variables

be denoted as h_0, α_0, u_0 and v_0 . For both the subcritical and supercritical test case we take $h_0 = 1, u_0 = 1, v_0 = 1$, and $\alpha = 0.3$.

We consider the solution on a domain $x \in [0, 1]$ in which the topography is given by [87]:

$$b(x) = \begin{cases} 0.2 - 20(x - 0.5)^2 & \text{if } 0.4 \leq x \leq 0.6, \\ 0 & \text{otherwise.} \end{cases}$$

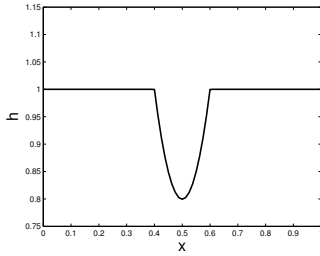
As initial condition we take $h + b = 1, u = u_0, v = v_0$ and $\alpha = \alpha_0$. At the boundaries we define the exterior trace to be the same as the initial condition. For the subcritical test case we take $g_3 = 10^8$ while for the supercritical test case $g_3 = 25$. Other parameters in the model are chosen as: $\varepsilon = 0.01, \rho = 0.9, \theta = 0^\circ, \varphi_{11} = 2(1 - \sqrt{1 - \cos^2(\phi_{int})(1 + \tan^2(\phi_{bed}))}) / \cos^2(\phi_{int}) - 1, \phi_{int} = 24.5^\circ$ and $\phi_{bed} = 14.75^\circ$. In F^D , the parameters are $\rho^f = 1000 \text{ kg m}^{-3}, v_T = 0.143 \text{ m s}^{-1}, d = 10^{-3} \text{ m}$ and $\mu^f = 10^{-3} \text{ kg (ms)}^{-1}$.

We compute the order of convergence by comparing the space-time discontinuous Galerkin finite element solution of (4.14) to an “exact” solution of (4.14). This “exact” solution is found by setting the time-derivative terms in (4.14) to zero and then solving the system of ODE’s with a RK45 method on a grid with 10000 points.

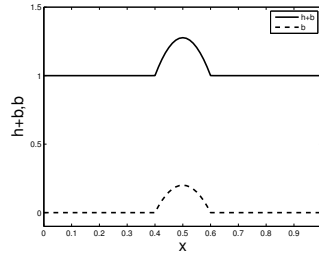
In Figure 4.4 we plot the numerical solutions of the total flow height $h + b$, topography b , flow depth h , particle volume fraction α and the velocities u and v for sub- and supercritical flow. The order of convergence is given for the mixture momentum $h\alpha v + \rho h(1 - \alpha)u$ as well as the topography b in Table 4.1 for sub- and supercritical flow. The reason why we also show the order of convergence for the topography b is because it is taken formally as an unknown in the system (as in Chapter 3) and we show that the topography converges at the same rate as the other unknowns. For a linear polynomial approximation we obtain as expected second order convergence.

4.3.2 The slope limiter

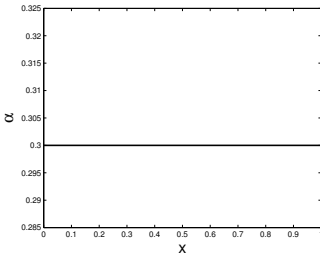
We consider a Riemann problem to test the effect of the polynomials (Lagrange, original and/or Hermite) in the slope limiter and the parameter γ (see (4.12)) in the space-time DGFEM discretization. For this test case we neglect the source terms. Furthermore, we simplify the expressions for $\varphi_{11}, \varphi_{22}, \varphi_{12}$ and φ_{21} by taking $\varphi_{11} = \varphi_{22} = 1$ and $\varphi_{12} = \varphi_{21} = 0$. Other parameters in the model are chosen as $\rho = 1, g = 1, \varepsilon = 1$ and $\theta = 0^\circ$. We consider the solution on a domain $[0, 1] \times [0, 1]$ divided into 32×32 elements. A physical time step of $\Delta t = 0.005$ is used and we consider the solution at final time $T = 0.37$. We consider the



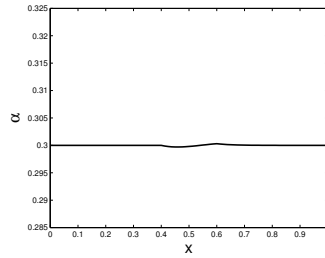
(a) Subcritical flow: flow depth h .



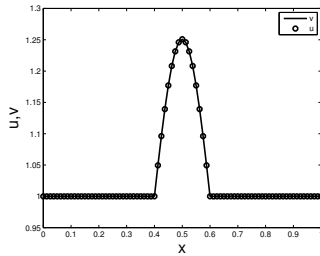
(b) Supercritical flow: flow height $h + b$ and topography b .



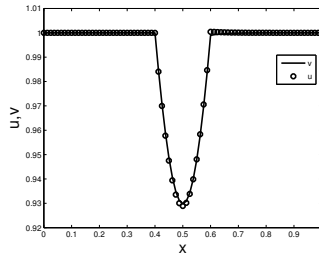
(c) Subcritical flow: particle volume fraction α .



(d) Supercritical flow: particle volume fraction α .



(e) Subcritical flow: velocities u and v .



(f) Supercritical flow: velocities u and v .

Figure 4.4: Steady-state solution on a grid with 80 elements.

Subcritical flow				
N_{cells}	b	p	$h(\alpha v + \rho(1 - \alpha)u)$	p
10	$1.5171 \cdot 10^{-2}$	-	$4.2299 \cdot 10^{-3}$	-
20	$3.7397 \cdot 10^{-3}$	2.0	$1.0484 \cdot 10^{-3}$	2.0
40	$9.3222 \cdot 10^{-4}$	2.0	$2.9977 \cdot 10^{-4}$	1.8
80	$2.3296 \cdot 10^{-4}$	2.0	$8.1480 \cdot 10^{-5}$	1.9
Supercritical flow				
N_{cells}	b	p	$h(\alpha v + \rho(1 - \alpha)u)$	p
10	$1.2910 \cdot 10^{-2}$	-	$4.0446 \cdot 10^{-4}$	-
20	$3.4861 \cdot 10^{-3}$	1.9	$1.4343 \cdot 10^{-4}$	1.5
40	$9.0211 \cdot 10^{-4}$	2.0	$3.7399 \cdot 10^{-5}$	1.9
80	$2.2925 \cdot 10^{-4}$	2.0	$9.4394 \cdot 10^{-6}$	2.0

Table 4.1: Sub- and supercritical flow: L^2 error for the topography b and the total momentum $h(\alpha v + \rho(1 - \alpha)u)$ and the convergence rate p .

following two-dimensional Riemann problem:

$$V(t = 0) = \begin{cases} V^L & \text{for } x < 0.5 \text{ and } y < 0.5, \\ V^R & \text{otherwise,} \end{cases}$$

in which V is the vector of primitive variables and $V^L = [1, 0.4, 0, 0, 0, 0]^T$ and $V^R = [0.5, 0.2, 0, 0, 0, 0]^T$. At the boundaries we set $u_1 = u_2 = v_1 = v_2 = 0$.

The slope limiter is used in element \mathcal{K}_k^n only when the discontinuity detector $\mathcal{I}_k^n > \varepsilon_{kriv}$. In this test case we take $\varepsilon_{kriv} = 1$.

In Figure 4.5 we compare the solution of the volume fraction α along the diagonal $x = y$ as computed using a slope limiter with the combination Lagrange and original polynomials; and, the combination Lagrange, Hermite and original polynomials. For each combination we furthermore compare the solution using $\gamma = 1$ and $\gamma = 10$ in (4.12). We see that the least numerical dissipation is introduced using the combination Lagrange and original polynomials while $\gamma = 1$. Increasing γ to $\gamma = 10$ introduces more smoothing to the solution. Also adding the Hermite polynomials to the combination Lagrange and original polynomials increases the amount of numerical dissipation. This can also be seen in Figure 4.6 where we compare the flow height h calculated using the combination Lagrange and original polynomials with $\gamma = 1$; and, the combination Lagrange, Hermite and original polynomials with $\gamma = 10$. We plot the results per element to show the discontinuities at the element faces which would not be visible with

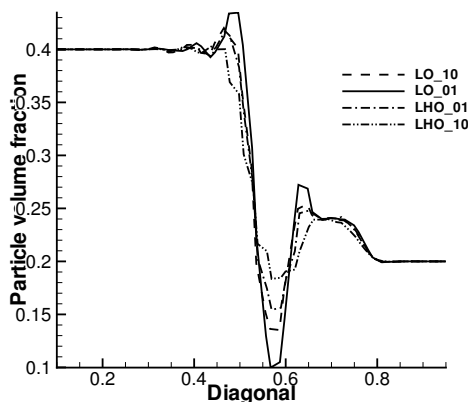
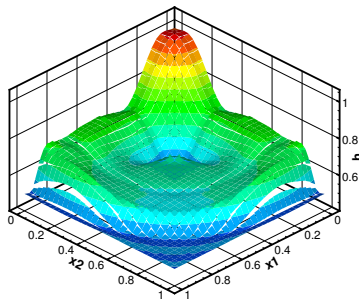
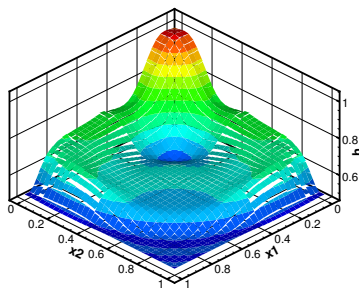


Figure 4.5: *Solution of the volume fraction α at $T = 0.37$ along the diagonal $x = y$ as computed using the combination of Lagrange and original polynomials (LO) in the slope limiter or the combination of Lagrange, Hermite and original polynomials (LHO) with $\gamma = 1$ or $\gamma = 10$.*

post-processing. We remark that without the slope limiter it was not possible to do this test case because α became less than zero in regions around large discontinuities due to undershoots. In Figure 4.7 we indicate the areas where the discontinuity detector detects large discontinuities. In these regions the slope limiter is used. The scheme is robust for a wide range of γ values, but for accuracy reasons γ should be chosen as small as possible, because this minimizes the numerical dissipation. For the Riemann problem presented here, the best combination would be the Lagrange and original polynomials with $\gamma = 1$. As can be seen in Figures 4.5 and 4.6 there is a wave crest which can be captured using the combination Lagrange and original polynomials with $\gamma = 1$ which is not captured using the combination Lagrange, Hermite and original polynomials with $\gamma = 10$ since this combination introduces too much numerical dissipation. For other applications though, more numerical dissipation may be desirable to avoid large over- and undershoots which can be achieved by slightly increasing the value for γ and/or using a different combination of polynomials.



(a) The flow height h as calculated using Lagrange and the original polynomials in the slope limiter. Furthermore, $\gamma = 1$ in (4.12).



(b) The flow height h as calculated using Lagrange, Hermite and the original polynomials in the slope limiter. Furthermore, $\gamma = 10$ in (4.12).

Figure 4.6: *Solution of the flow height h at $T = 0.37$. Too much numerical dissipation is introduced in (b) since there is a wave crest in (a) which is not captured in (b).*

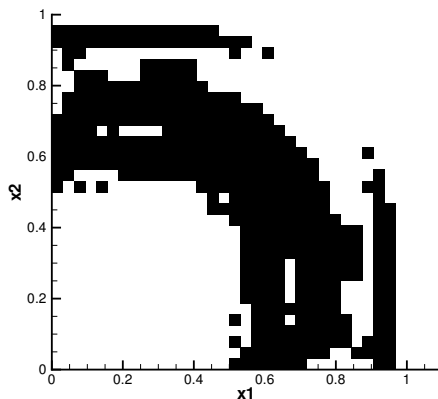


Figure 4.7: *The shaded area indicates where the discontinuity detector has large values and where the slope limiter is used; situation at $T = 0.37$.*

4.4 Validation

In [2, 4, 71] laboratory experiments of shallow water and in [82] laboratory experiments of shallow granular flow through a contraction were compared to numerical results. We will simulate a two-phase flow mixture consisting of solid particles in water in which the density of the solid particles is slightly higher than that of water. We will simulate the flow of this mixture as it enters a contraction. Initially we start with a flow with very low particle volume fraction (5 %) and the flow reaches a steady-state with oblique jumps. We then perturb this steady-state by increasing the particle volume fraction at the inlet to 30 % for a short period. This perturbation was sufficient to perturb the flow with oblique jumps to one with an upstream moving shock as was observed by Akers and Bokhove [2] (see Figure 1.2). We now describe the numerical setup.

In our numerical calculations we consider a channel in the Cartesian coordinate system $(x, y) \in [0, 10] \times [-0.5, 0.5]$. The channel converges from $x = 4$ to $x = 4.7228$ so that $y \in [-0.3, 0.3]$ and diverges from $x = 4.7228$ to $x = 6.1685$ (see Figure 4.8). As initial condition we take $h = 0.2$, $\alpha = 0.05$, $v_1 = u_1 = 0.5$ and $v_2 = u_2 = 0$. Define $h_w = h(1 - \alpha)$. At the inflow boundary we specify $h_w = 0.19$, the x -components of the velocities, $u_1 = 0.5$ and $v_1 = 0.5$, the y -components of the velocities, $u_2 = 0$ and $v_2 = 0$, and the particle volume fraction α . Initially, the inflow condition for the particle volume fraction is $\alpha = 0.05$. For time $20 < t < 35$ we change the inflow condition by increasing the particle

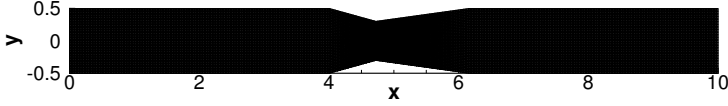


Figure 4.8: *Geometry and mesh of the chute with a contraction.*

volume fraction to $\alpha = 0.3$ after which we decrease the particle volume fraction again to $\alpha = 0.05$. At the walls we follow Ambati and Bokhove [3] and impose:

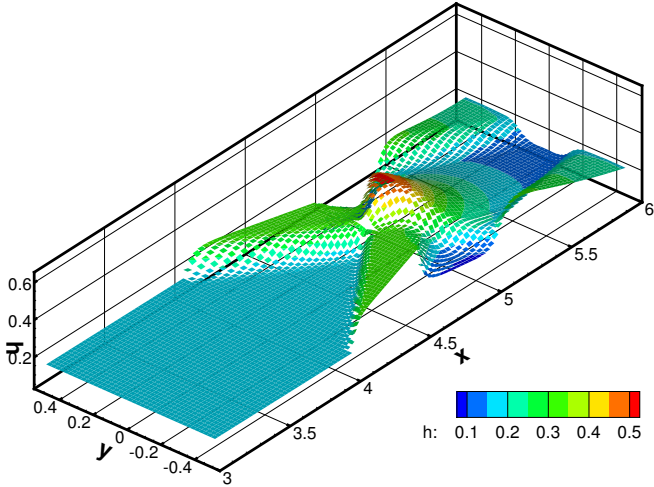
$$\begin{aligned} \mathbf{v}^b \cdot \bar{\mathbf{n}} &= -\mathbf{v}^L \cdot \bar{\mathbf{n}}, & \mathbf{v}^b \cdot \bar{\mathbf{t}} &= \mathbf{v}^L \cdot \bar{\mathbf{t}}, \\ \mathbf{u}^b \cdot \bar{\mathbf{n}} &= -\mathbf{u}^L \cdot \bar{\mathbf{n}}, & \mathbf{u}^b \cdot \bar{\mathbf{t}} &= \mathbf{u}^L \cdot \bar{\mathbf{t}}, \end{aligned} \quad (4.15)$$

where $\bar{\mathbf{t}}$ is the unit tangential vector orthogonal to the normal vector $\bar{\mathbf{n}}$. Furthermore, we extrapolate the void fraction, $\alpha^R = \alpha^L$ and the flow height $h^R = h^L$. At the outflow boundary, all variables are extrapolated, $U^R = U^L$.

There are a number of constants in the depth-averaged flow model. We will consider shallow liquid-solid flows with a height to length ratio of $\varepsilon = 0.2$ as a feasible approximation and for which the liquid to solid density ratio is $\rho = 0.9$. The gravity constant is $g = 1.5$ so that the gravity components are $g_1 = \sin(\theta)g$, $g_2 = 0$ and $g_3 = \cos(\theta)g$ in which θ is the angle of the contraction with respect to the horizontal (see also Figure 4.1). We take $\theta = 0.625^\circ$ for $0 \leq x \leq 7$ and $\theta = 10^\circ$ for $x > 7$ so that the outflow boundary has no effect on the solution in the contraction. To be able to calculate the drag function F^D , we use the following constants: $\rho^f = 1000 \text{ kg m}^{-3}$ and $\mu^f = 10^{-3} \text{ kg (ms)}^{-1}$ while the solid particles are assumed to have a diameter of $d = 10^{-3} \text{ m}$ and $v_T = 0.143 \text{ m s}^{-1}$ [38]. The internal angle of friction is taken to be $\phi_{int} = 24.5^\circ$ and the bed friction angle is $\phi_{bed} = 14.75^\circ$ [7]. The bottom topography is taken constant $b(x, y) = 0$ and the drag coefficient is $C_D = 10^{-4}$.

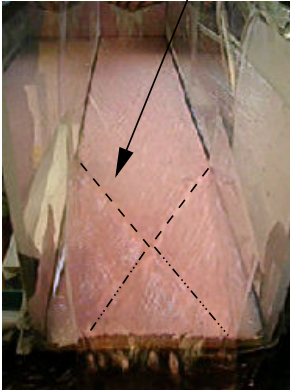
We compute the solution for the depth-averaged model using space DGFEM until $t = 100$ using a CFL number of $\text{CFL} = 0.8$ on a grid with 400 elements in the x -direction and 40 elements in the y -direction. In the slope limiter, a combination of Lagrange, Hermite and original polynomials was used with $\gamma = 10$ to avoid severe over- and undershoots. In Figures 4.9- 4.11 we show the transition of the flow height h from oblique jumps to an upstream steady shock (for a

comparison, see also Figure 1.2). We see the same as observed by Akers and Bokhove [2]. In Figures 4.9a, b and c we see that the first steady-state solution is captured very well. After increasing the particle volume fraction at the inlet, the steady-state is perturbed. The transitional phase in Figures 4.10a and b show the numerical and laboratory results, respectively, and we see different behavior between the results. The second steady-state, an upstream steady shock, however is captured again very well (see Figures 4.11a and b). We remark that if $C_D = 0$, we do not get an upstream steady shock. As expected, an upstream moving shock appears.

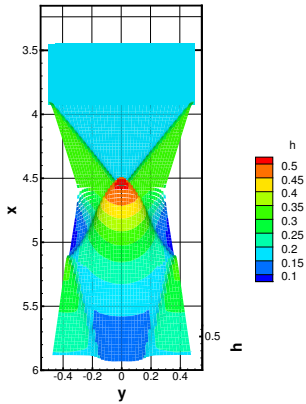


(a) Side view of h .

Oblique jumps

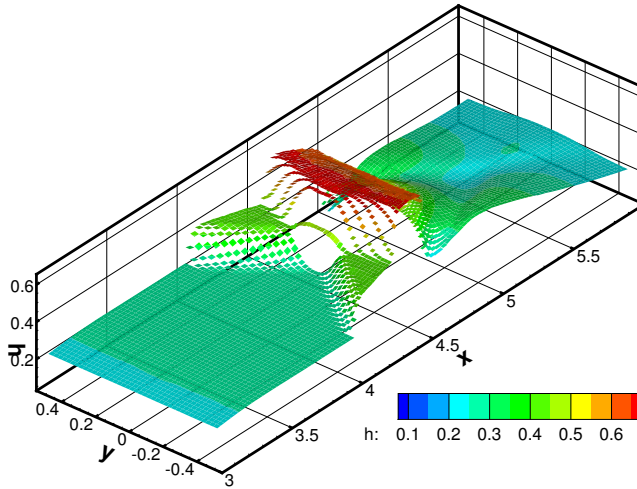


(b) Snapshot from the laboratory experiment.



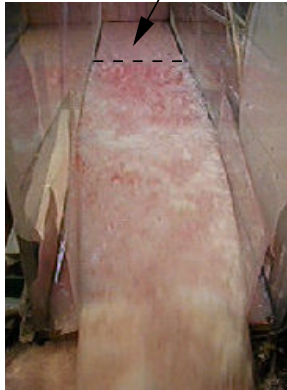
(c) Top view of h .

Figure 4.9: *Oblique jump solution at $t = 22$.*



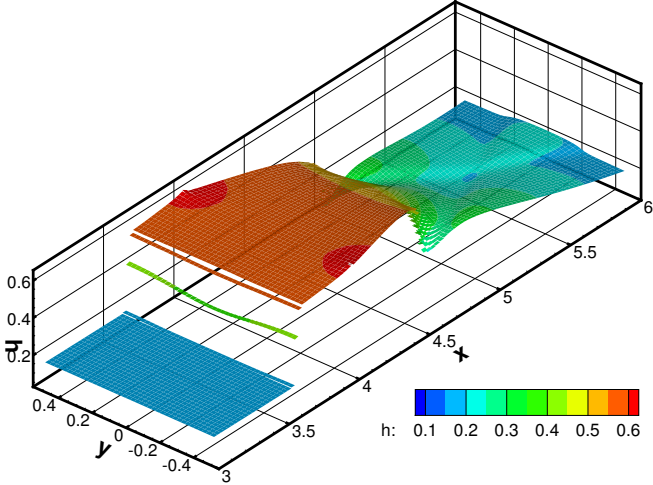
(a) Side view.

Hydraulic jump



(b) Snapshot from the laboratory experiment.

Figure 4.10: *Transition phase at $t = 39$.*



(a) Side view.

Hydraulic jump



(b) Snapshot from the laboratory experiment.

Figure 4.11: *Upstream moving shock at $t = 100$.*

***h*-Multigrid optimization for higher order accurate space-time discontinuous Galerkin discretizations**

Space-time discontinuous Galerkin (DG) discretizations of partial differential equations result in large systems of algebraic equations that need to be solved at each time-step. To efficiently solve these algebraic equations, we combine a pseudo-time integration method with new multigrid techniques. We remark that in this chapter we only consider *h*-multigrid techniques so that with multigrid we mean *h*-multigrid. The main benefits of this multigrid algorithm is that no large global linear system needs to be solved and, through the use of explicit Runge-Kutta type smoothers, the locality of the DG discretization is preserved. A two- and three-level Fourier analysis is used to investigate and optimize the *h*-multigrid algorithm for second and third order accurate space-time DG discretizations of the two-dimensional advection-diffusion equation. We will compare our new optimized schemes with existing *h*-multigrid techniques employing Runge-Kutta type smoothers.

5.1 Space-time DG for the 2D advection-diffusion equation

In this section we summarize the space-time DG method for the 2D advection-diffusion equation. We start by explaining the space-time formulation of the

problem after which the space-time DG weak formulation is discussed. Introducing a polynomial approximation in the weak formulation then results in a system of algebraic equations. Also, we summarize the pseudo-time integration and explicit Runge-Kutta methods that we optimize later in Section 5.4.

5.1.1 Space-time DG formulation

In the space-time DG method, the space and time variables are treated together. A point at time $t = x_0$ with position vector $\bar{x} = (x_1, x_2)$ has Cartesian coordinates $x = (x_0, \bar{x})$ in the open domain $\mathcal{E} \subset \mathbb{R}^3$. At time t , the flow domain $\Omega(t)$ is defined as $\Omega(t) := \{\bar{x} \in \mathbb{R}^2 : (t, \bar{x}) \in \mathcal{E}\}$. By taking t_0 and T as the initial and final time of the evolution of the space-time flow domain, the space-time domain boundary $\partial\mathcal{E}$ consists of the hyper-surfaces $\Omega(t_0) \equiv \{x \in \partial\mathcal{E} : x_0 = t_0\}$, $\Omega(T) \equiv \{x \in \partial\mathcal{E} : x_0 = T\}$ and $\mathcal{Q} \equiv \{x \in \partial\mathcal{E} : t_0 < x_0 < T\}$. With this notation, the 2D advection-diffusion equation can be written as:

$$\begin{cases} u_{,0} + a_k u_{,k} - A_{ks} u_{,s,k} = 0, & k, s = 1, 2, & \text{on } \mathcal{E}, \\ u = u_0, & & \text{on } \Omega(t_0), \\ u = u^b, & & \text{on } \mathcal{Q}, \end{cases} \quad (5.1)$$

where a comma denotes differentiation with respect to the Cartesian coordinate x_k , $a_k \in \mathbb{R}$ are the constant advection coefficients and $A_{ks} \in \mathbb{R}^+$ the constant diffusion coefficients. We take $A_{11} = \nu_x$, $A_{22} = \nu_y$ and $A_{12} = A_{21} = 0$. The initial flow field is denoted by u_0 and the boundary data by u^b . The summation convention is used on repeated indices.

5.1.2 Weak formulation and discretization

The approximation $\Omega_h(t_n)$ of the flow domain $\Omega(t_n)$ is divided into N_n non-overlapping spatial elements $K_j(t_n)$. The space-time elements \mathcal{K}_j^n are constructed by connecting K_j^n with K_j^{n+1} using linear or quadratic interpolation in time. The flow domain \mathcal{E} , limited to the time interval (t_n, t_{n+1}) , defines a space-time slab, \mathcal{E}_h^n . The tessellation \mathcal{T}_h^n of \mathcal{E}_h^n consists of all space-time elements \mathcal{K}_j^n .

Within a space-time slab we distinguish faces connecting space-time slabs, $K_j(t_n^+)$ and $K_j(t_{n+1}^-)$, internal faces \mathcal{S}_I^n and boundary faces \mathcal{S}_B^n . The outward space-time normal vector on a space-time element \mathcal{K}_j^n is denoted by $n = (n_t, \bar{n})$, with n_t the temporal and \bar{n} the spacial part of the outward normal vector. On an internal face $\mathcal{S} \in \mathcal{S}_I$, the traces from the left and right element are denoted by $(\cdot)^L$ and $(\cdot)^R$, respectively. The average operator is defined as $\{\!\{ \cdot \}\!\} = \frac{1}{2}((\cdot)^L + (\cdot)^R)$ and the jump operator as $[\![\cdot]\!]_k = (\cdot)^L \bar{n}_k^L + (\cdot)^R \bar{n}_k^R$.

We consider approximations of $u(x)$ and test functions $v(x)$ in the finite element space W_h defined as:

$$W_h = \{W \in L_2(\mathcal{E}_h) : W|_{\mathcal{K}} \circ G_{\mathcal{K}}^n \in P^p(\hat{\mathcal{K}}), \forall \mathcal{K} \in \mathcal{T}_h\},$$

where $L_2(\mathcal{E}_h)$ is the space of square integrable functions on \mathcal{E}_h , $P^p(\hat{\mathcal{K}})$ the space of polynomials of degree at most p on the reference element $\hat{\mathcal{K}} = (-1, 1)^3$ and $G_{\mathcal{K}}^n$ the iso-parametric mapping from the master element to the space-time element \mathcal{K}_j^n . Furthermore, we also use the following space:

$$V_h = \{V \in (L_2(\mathcal{E}_h))^2 : V|_{\mathcal{K}} \circ G_{\mathcal{K}}^n \in (P^p(\hat{\mathcal{K}}))^2, \forall \mathcal{K} \in \mathcal{T}_h\}.$$

The space-time DG weak formulation of the 2D advection-diffusion equation is: *Find a $u \in W_h$ such that for all $v \in W_h$:*

$$\begin{aligned} & - \sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} (v_{,0} u + v_{,k} (a_k u - A_{ks} u_{,s})) d\mathcal{K} + \sum_{\mathcal{S} \in \mathcal{S}_I^n} \int_{\mathcal{S}} \llbracket v \rrbracket_k a_k \hat{u} d\mathcal{S} \\ & + \sum_{\mathcal{K} \in \mathcal{T}_h^n} \left(\int_{K(t_{n+1}^-)} v^L u^L dK - \int_{K(t_n^+)} v^L u^R dK \right) + \sum_{\mathcal{S} \in \mathcal{S}_B^n} \int_{\mathcal{S}} v^L a_k \hat{u} \bar{n}_k^L d\mathcal{S} \\ & - \sum_{\mathcal{S} \in \mathcal{S}_I^n} \int_{\mathcal{S}} \llbracket v \rrbracket_k A_{ks} \{ \{ u_{,s} - \eta^S \mathcal{R}_s^S \} \} d\mathcal{S} - \sum_{\mathcal{S} \in \mathcal{S}_B^n} \int_{\mathcal{S}} v^L A_{ks} (u_{,s}^L - \eta^S \mathcal{R}_s^S) \bar{n}_k^L d\mathcal{S} \\ & - \sum_{\mathcal{S} \in \mathcal{S}_I^n} \int_{\mathcal{S}} A_{ks} \{ \{ v_{,k} \} \} \llbracket u \rrbracket_s d\mathcal{S} - \sum_{\mathcal{S} \in \mathcal{S}_B^n} \int_{\mathcal{S}} v_{,k}^L A_{ks} (u^L - u^b) \bar{n}_s^L d\mathcal{S} = 0, \end{aligned} \tag{5.2}$$

in which a space-time generalization of the approaches by Bassi and Rebay [10] and Brezzi [13] were followed for the discretization of the viscous flux. The local lifting operator is defined as [41]: *Find an $\mathcal{R}^S \in V_h$ such that for all $w \in V_h$:*

$$\sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} w_s \mathcal{R}_s^S d\mathcal{K} = \begin{cases} \int_{\mathcal{S}} \{ \{ w_s \} \} \llbracket u \rrbracket_s d\mathcal{S} & \text{for } \mathcal{S} \in \mathcal{S}_I^n, \\ \int_{\mathcal{S}} w_s^L (u^L - u^b) \bar{n}_s^L d\mathcal{S} & \text{for } \mathcal{S} \in \mathcal{S}_B^n. \end{cases}$$

The stabilization parameter η^S is constant and should be chosen greater or equal to the number of space-faces of an element. An upwind flux is used for \hat{u} in (5.2). We refer to [69] for a full error analysis and derivation of the space-time DG algorithm for the advection-diffusion equation.

The DG discretization is obtained by approximating the test and trial functions in each element $\mathcal{K} \in \mathcal{T}_h^n$ with polynomial expansions:

$$u(t, \bar{x})|_{\mathcal{K}} = \hat{u}_i \psi_i(t, \bar{x}), \quad v(t, \bar{x})|_{\mathcal{K}} = \hat{v}_j \psi_j(t, \bar{x}), \tag{5.3}$$

with ψ_i the basis functions and \hat{u}_i and \hat{v}_j expansion coefficients, respectively, for $i, j = 0, 1, 2, \dots, N$, where N depends on the polynomial degree of the basis functions in W_h . On the reference element $\hat{\mathcal{K}}$, the basis functions $\hat{\psi}_i(\xi) \in P^p(\hat{\mathcal{K}})$ are a combination of one-dimensional Legendre polynomials. These are related to $\psi_i(x)$ in the space-time element \mathcal{K} through the mapping $G_{\mathcal{K}}^n$: $\psi_i = \hat{\psi}_i \circ (G_{\mathcal{K}}^n)^{-1}$. In this chapter we consider linear and quadratic polynomial approximations.

For the multigrid optimization in Section 5.4, we consider a uniform space-time mesh with elements $\Delta t \times \Delta x \times \Delta y$ and we consider (5.2) only with periodic boundary conditions. We will use the following dimensionless numbers:

$$CFL = \frac{a\Delta t}{h}, \quad Re_x = \frac{a(\Delta x)^2}{\nu_x h}, \quad Re_y = \frac{a(\Delta y)^2}{\nu_y h}, \quad AR = \frac{\Delta y}{\Delta x},$$

in which $h = \Delta x \sqrt{1 + AR^2}$ and $a = \sqrt{a_x^2 + a_y^2}$. Furthermore, we introduce the flow angle γ^{flow} with respect to the x -axis so that $a_x = \cos(\gamma^{flow})a$ and $a_y = \sin(\gamma^{flow})a$.

Then, replacing u and v in (5.2) with their expansions (5.3) and using the fact that the coefficients \hat{v} are arbitrary, we obtain the following discrete system for the vector of expansion coefficients \hat{u}^n at time level n :

$$\mathcal{L}_h(\hat{u}^n; \hat{u}^{n-1}) := \Delta x \Delta y ((\mathcal{L}_x^a + \mathcal{L}_y^a + \mathcal{L}_x^d + \mathcal{L}_y^d + \mathcal{L}^t)\hat{u}^n + \mathcal{L}^{t-1}\hat{u}^{n-1}) = 0. \quad (5.4)$$

The space-time DG discretization can be represented using the following dimensionless stencil notation:

$$\begin{aligned} \bar{\mathcal{L}}_x^a &= \begin{bmatrix} 0 & & \\ L_x^a & D_x^a & 0 \\ & 0 & \end{bmatrix}, \quad \bar{\mathcal{L}}_y^a = \begin{bmatrix} 0 & & \\ 0 & D_y^a & 0 \\ & L_y^a & \end{bmatrix}, \quad \bar{\mathcal{L}}^t = \begin{bmatrix} 0 & & \\ 0 & D^t & 0 \\ & 0 & \end{bmatrix}, \\ \bar{\mathcal{L}}_x^d &= \begin{bmatrix} 0 & & \\ L_x^d & D_x^d & U_x^d \\ & 0 & \end{bmatrix}, \quad \bar{\mathcal{L}}_y^d = \begin{bmatrix} 0 & & \\ 0 & U_y^d & 0 \\ & D_y^d & L_y^d \end{bmatrix}, \quad \bar{\mathcal{L}}^{t-1} = \begin{bmatrix} 0 & & \\ 0 & D^{t-1} & 0 \\ & 0 & \end{bmatrix}, \end{aligned} \quad (5.5)$$

in which the blocks $L_{x,y}^{a,d}$, $D_{x,y}^{a,d}$, $U_{x,y}^d$, $D^{t,t-1} \in \mathbb{R}^{m \times m}$ and m depends on the order of the space-time DG discretization. Note, we assume here that the advection coefficients a_k are all positive. If they are (partly) negative the upwind stencil in \mathcal{L}_x^a and/or \mathcal{L}_y^a must be changed accordingly.

The inviscid part of the stencil in (5.4) depends on the CFL number and can now be written as:

$$\begin{aligned} \mathcal{L}_x^a &= CFL \cos(\gamma^{flow}) \sqrt{1 + AR^2} \bar{\mathcal{L}}_x^a = \Gamma_x^a \bar{\mathcal{L}}_x^a, \\ \mathcal{L}_y^a &= CFL \sin(\gamma^{flow}) \frac{\sqrt{1 + AR^2}}{AR} \bar{\mathcal{L}}_y^a = \Gamma_y^a \bar{\mathcal{L}}_y^a. \end{aligned} \quad (5.6)$$

The viscous part of the stencil depends on the Reynolds and CFL numbers:

$$\mathcal{L}_x^d = \frac{CFL}{Re_x} \bar{\mathcal{L}}_x^d = \Gamma_x^d \bar{\mathcal{L}}_x^d, \quad \mathcal{L}_y^d = \frac{CFL}{Re_y} \bar{\mathcal{L}}_y^d = \Gamma_y^d \bar{\mathcal{L}}_y^d, \quad (5.7)$$

while the stencils related to the time discretization are given by:

$$\mathcal{L}^t = \bar{\mathcal{L}}^t, \quad \mathcal{L}^{t-1} = \bar{\mathcal{L}}^{t-1}. \quad (5.8)$$

5.1.3 Pseudo-time integration and Runge-Kutta methods

The space-time discretization (5.4) is solved using a multigrid algorithm with a Runge-Kutta type smoother. For notational purposes let the system of algebraic equations be denoted as

$$\mathcal{L}_h(\hat{u}^n; \hat{u}^{n-1}) = 0. \quad (5.9)$$

Following van der Vegt and van der Ven [79], to solve the system of coupled equations for the expansion coefficients (5.9), a pseudo time derivative is added to the system:

$$\mathcal{M} \frac{\partial \hat{u}^*}{\partial \tau} = -\mathcal{L}_h(\hat{u}^*; \hat{u}^{n-1}), \quad (5.10)$$

which is integrated to steady-state in pseudo-time. At steady state, $\hat{u}^n = \hat{u}^*$. Here \mathcal{M} is the mass matrix defined as $\mathcal{M}_{ij} = \int_{\mathcal{K}} \psi_i \psi_j d\mathcal{K}$, with ψ_i the basis functions used in the DG discretization.

For the pseudo-time integration we introduce the dimensionless number $\lambda = \Delta\tau/\Delta t$ and use the pseudo-time CFL number, defined as $CFL^\tau = \lambda CFL$ and the pseudo-time Von Neumann number $VN^\tau = \min(CFL^\tau/Re_x, CFL^\tau/Re_y)$. We rewrite (5.10) by setting $\mathcal{L}_h = \Delta x \Delta y \bar{\mathcal{L}}_h$ and $\mathcal{M} = \Delta x \Delta y \Delta t \bar{\mathcal{M}}$:

$$\Delta t \bar{\mathcal{M}} \frac{\partial \hat{u}^*}{\partial \tau} = -\bar{\mathcal{L}}_h(\hat{u}^*; \hat{u}^{n-1}), \quad (5.11)$$

in which $\bar{\mathcal{M}}$ is the dimensionless mass-matrix. To solve (5.11) we consider 5-stage Runge-Kutta methods. For notational purposes, we set $\bar{\mathcal{L}}_h(V^*; u^{n-1}) = \bar{\mathcal{L}}_h(V^*)$. Initialize $\hat{V}^0 = \hat{u}^{n-1}$. Then, an N -stage Runge-Kutta scheme is given by:

$$\hat{V}^j = \left(\hat{V}^0 - \lambda \left(\sum_{l=1}^j \alpha_{j+1,l} \bar{\mathcal{M}}^{-1} \bar{\mathcal{L}}_h(\hat{V}^{l-1}) \right) + \lambda \beta_j \hat{V}^{j-1} \right) / (1 + \beta_j \lambda I), \quad j = 1, \dots, N. \quad (5.12)$$

Return then $\hat{u}^n = \hat{V}^N$. For notational purposes we also write:

$$\hat{u}^n = S_h \hat{u}^{n-1}. \quad (5.13)$$

We require from the 5-stage Runge-Kutta schemes that they are second order accurate in pseudo-time. This requirement gives conditions for the α coefficients [29]. For the β coefficients, which serve as a Melson correction to improve stability for small values of $\lambda \cong 1$, see Melson et al. [55], no additional conditions are imposed.

5.2 h -Multigrid algorithms for linear systems

In this section we will first discuss the algebraic formulation of multigrid techniques for linear systems. Next, we discuss the multigrid error transformation operators. These operators will be analyzed using Fourier analysis in Section 5.3. For more information on multigrid methods, we refer to Hackbusch [27], Trottenberg et al. [77] and Wesseling [84].

5.2.1 Algebraic formulation of multigrid algorithms for linear systems

In a multigrid technique for the solution of linear partial differential equations, we introduce a finite sequence N_c of increasingly coarser meshes \mathcal{G}_{nh} , $n \in \{1, \dots, N_c\}$ to generate coarser approximations of the original problem. For $1 \leq n < m \leq N_c$, restriction operators $R_{nh}^{mh} : \mathcal{G}_{nh} \rightarrow \mathcal{G}_{mh}$ and prolongation operators $P_{mh}^{nh} : \mathcal{G}_{mh} \rightarrow \mathcal{G}_{nh}$ are introduced to transfer the data between the different meshes. Following [42], we define the prolongation operator P_{mh}^{nh} for the space-time DG discretization as:

$$\begin{aligned} \hat{u}_i^{nh} &= (\mathcal{M}_{nh}^{-1})_{il} \left(\int_{\mathcal{K}_{nh}} \psi_l^{nh} \psi_j^{mh} d\mathcal{K} \right) \hat{u}_j^{mh} \\ &= P_{mh}^{nh} \hat{u}_j^{mh}, \quad 1 \leq n < m \leq N_c, \end{aligned} \tag{5.14}$$

where \mathcal{M}_{nh} is the mass matrix of element \mathcal{K}_{nh} . This definition stems from the L_2 -projection of the coarse grid solution U_{mh} in an element \mathcal{K}_{mh} on the corresponding set of fine grid elements $\{\mathcal{K}_{nh}\}$. The restriction operator is defined as $R_{nh}^{mh} = \frac{1}{4}(P_{mh}^{nh})^T$. In (5.14), the embedding of spaces is assumed, i.e. $W_{mh} \subset W_{nh}$ for $n < m$, to ensure that u^{mh} is defined in \mathcal{K}_{nh} .

In a multigrid technique to solve $L_h v_h = f_h$ on \mathcal{G}_h , with L_h a linear discretization operator and f_h a given righthand side, a set of auxiliary problems is solved. At each grid level \mathcal{G}_{nh} , $1 \leq n \leq N_c$, we solve $L_{nh} v_{nh} = f_{nh}$ using an iterative method S_{nh} starting from an initial guess w_{nh}^0 . We assume that each operator L_{nh} is invertible. The multigrid algorithm is defined as follows:

Definition 5.2.1. MG_n -Algorithm.

if $n = N_c$, the coarsest mesh, then

$$w_{nh}^1 = MG_n(w_{nh}^0, f_{nh}, \nu_1, \nu_2) := L_{nh}^{-1} f_{nh}, \quad (5.15)$$

else the multigrid algorithm MG_n is defined as

1. (Pre-smoothing). Let $u_{nh}^0 = w_{nh}^0$ and define $u_{nh}^{\nu_1}$ by

$$u_{nh}^{l+1} = S_{nh} u_{nh}^l, \quad l = 0, 1, \dots, \nu_1 - 1;$$

with S_{nh} the smoothing operator at grid level n given by (5.13).

2. (Restricting). Restrict the residual to the coarser mesh \mathcal{G}_{mh} with $n < m \leq N_c$,

$$r_{mh} = R_{nh}^{mh}(f_{nh} - L_{nh} u_{nh}^{\nu_1}).$$

3. (Coarse-grid solving). At grid level m set $z_{mh}^0 = 0$ and repeat the MG_m -algorithm γ -times

$$z_{mh}^l = MG_m(z_{mh}^{l-1}, r_{mh}, \nu_1, \nu_2), \quad l = 1, \dots, \gamma.$$

4. (Correcting). Let $y_{nh}^0 = u_{nh}^{\nu_1} + P_{mh}^{nh} z_{mh}^\gamma$.

5. (Post-smoothing). Define $y_{nh}^{\nu_2}$ by

$$y_{nh}^{l+1} = S_{nh} y_{nh}^l, \quad l = 0, 1, \dots, \nu_2 - 1.$$

Finally, set $w_{nh}^1 = MG_n(w_{nh}^0, f_{nh}, \nu_1, \nu_2) := y_{nh}^{\nu_2}$ as the result of the MG_n -algorithm.

Using different sequences of meshes \mathcal{G}_{nh} various multigrid cycles, such as the V, W or F-cycle can be constructed.

5.2.2 Multigrid error transformation operator

In order to understand the performance of the multigrid algorithm we need the multigrid error transformation operator. Define the error of the initial guess w_{nh}^0 in the multigrid algorithm at grid level n as $e_{nh}^0 = w_{nh}^0 - v_{nh}$, with v_{nh} the exact solution and the error of the approximation w_{nh}^1 after application of $MG_n(w_{nh}^0, f_{nh}, \nu_1, \nu_2)$ as $e_{nh}^1 = w_{nh}^1 - v_{nh}$. The initial and approximation error are related through the multigrid error transformation operator $e_{nh}^1 = M_{nh} e_{nh}^0$.

Definition 5.2.2. The multigrid error transformation operator M_{nh} at grid level n is defined recursively, with $m > n$, as:

$$M_{nh} = S_{nh}^{\nu_2} (I_{nh} - P_{mh}^{nh} (I_{mh} - M_{mh}^\gamma) L_{mh}^{-1} R_{nh}^{mh} L_{nh}) S_{nh}^{\nu_1},$$

with I_{nh} the identity operator, ν_l , $l = 1, 2$, the number of pre- and post-smoothing iterations and γ is the cycle index.

For a derivation, see e.g. [77, 78]. We are specifically interested in two- and three-level multigrid. For two grid levels ($n = 1, m = 2$) the multigrid error transformation operator is equal to

$$M_h^{2g} = S_h^{\nu_2} (I_h - P_{2h}^h L_{2h}^{-1} R_h^{2h} L_h) S_h^{\nu_1}, \quad (5.16)$$

and for three grid levels ($n = 1, m = 2$) and ($n = 2, m = 4$) we obtain

$$M_h^{3g} = S_h^{\nu_2} (I_h - P_{2h}^h (I_{2h} - M_{2h}^\gamma) L_{2h}^{-1} R_h^{2h} L_h) S_h^{\nu_1} \quad (5.17)$$

with

$$M_{2h} = S_{2h}^{\nu_4} (I_{2h} - P_{4h}^{2h} L_{4h}^{-1} R_{2h}^{4h} L_{2h}) S_{2h}^{\nu_3}. \quad (5.18)$$

We will also investigate the effect of replacing the exact solution on the coarsest mesh in (5.15) with ν_C smoothing steps. This approach is more straightforward for non-linear problems, such as the compressible Navier-Stokes equations.

5.3 Fourier analysis of discrete operators

To analyze the two- and three-level multigrid error transformation operators, (5.16) and (5.17)-(5.18), respectively, we will use discrete Fourier analysis. The analysis closely follows Brandt [12] and Wienands and Joppich [85]. See also Hackbusch [27], Hackbusch and Trottenberg [28], Trottenberg et al. [77], Wesseling [84], and van der Vegt and Rhebergen [78].

5.3.1 Introduction

Assume a finite mesh G_h^N , which is defined in \mathbb{R}^2 as

$$G_h^N := \{(x_1, x_2) = (k_1 h_1, k_2 h_2) \mid k \in G_k^N, h \in (\mathbb{R}^+)^2\},$$

where we use the index set G_k^N given by

$$G_k^N := \{k \in \mathbb{Z}^2 \mid -N_i \leq k_i \leq N_i - 1, N_i \in \mathbb{N}, i = 1, 2\}.$$

We also consider an infinite mesh G_h , which is defined in \mathbb{R}^2 as

$$G_h := \{(x_1, x_2) = (k_1 h_1, k_2 h_2) \mid k \in \mathbb{Z}^2, h \in (\mathbb{R}^+)^2\}.$$

On G_h we define for $v_h, w_h : G_{nh} \rightarrow \mathbb{C}$ the scaled Euclidian inner product as

$$(v_h, w_h)_{G_h} := \lim_{N \rightarrow \infty} \frac{1}{4N^2} \sum_{x \in G_h^N} v_h(x) \overline{w_h(x)}. \quad (5.19)$$

Here an over-bar denotes the complex conjugate and $x = (x_1, x_2)$. In the analysis we will also need the discrete ℓ^2 inner product and norm on G_h , which are defined for $v_h, w_h : G_h \rightarrow \mathbb{C}$, respectively, as $(v_h, w_h)_{\ell^2(G_h)} := \sum_{x \in G_h} v_h(x) \overline{w_h(x)}$ and $\|v_h\|_{\ell^2(G_h)} := (v_h, v_h)_{\ell^2(G_h)}^{\frac{1}{2}}$.

In the multigrid algorithm on each of the meshes G_{nh} , $n \in \{1, \dots, N_c\}$, we solve the linear system $L_{nh} v_{nh}(x) = f_{nh}(x)$, $x \in G_{nh}$, with L_{nh} the matrix resulting from a numerical discretization on the mesh G_{nh} of a (system of) linear partial differential equations with constant coefficients and f_{nh} the right hand side. The linear system on the mesh G_{nh} is described using stencil notation

$$L_{nh} v_{nh}(x) = \sum_{k \in J_n} l_{n,k} v_{nh}(x + kh), \quad x \in G_{nh}, \quad (5.20)$$

with stencil coefficients $l_{n,k} \in \mathbb{R}^{m_k \times m_k}$ and finite index sets $J_n \subset \mathbb{Z}^2$ describing the stencil. In a space-time DG discretization, the 5-point stencil is given by:

$$J_n := \{k = (k_1, k_2) \mid k_1, k_2 \in \{-1, 0, 1\}\}, \quad \text{and} \quad [L]_n = \begin{bmatrix} 0 & l_{0,1} & 0 \\ l_{-1,0} & l_{0,0} & l_{1,0} \\ 0 & l_{0,-1} & 0 \end{bmatrix}_n,$$

in which the stencil coefficients $l_{n,k}$ are generally $m_k \times m_k$ matrices, with $m_k \geq 1$ depending on the order of the discretization. For the advection-diffusion equation, it follows from (5.4)-(5.8):

$$\begin{aligned} l_{0,1} &= \Gamma_y^d U_y^d, & l_{-1,0} &= (\Gamma_x^a L_x^a + \Gamma_x^d L_x^d), \\ l_{0,0} &= (\Gamma_x^a D_x^a + \Gamma_y^a D_y^a + \Gamma_x^d D_x^d + \Gamma_y^d D_y^d + D^t), \\ l_{1,0} &= \Gamma_x^d U_x^d, & l_{0,-1} &= (\Gamma_y^a L_y^a + \Gamma_y^d L_y^d). \end{aligned}$$

On the infinite mesh G_h , we define for $x \in G_h$ the continuous Fourier modes with frequency $\theta = (\theta_1, \theta_2) \in \mathbb{R}^2$ as $\phi_h(\theta, x) := e^{i\theta \cdot x/h}$ with $\theta \cdot x/h := \theta_1 x_1/h_1 + \theta_2 x_2/h_2$, $h \in (\mathbb{R}^+)^2$ and $i = \sqrt{-1}$. We also define the space of bounded infinite grid functions by $\mathcal{F}(G_h) := \{v_h \mid v_h : G_h \rightarrow \mathbb{C} \text{ with } \|v_h\|_{G_h} < \infty\}$. For each $v_h \in \mathcal{F}(G_h)$ there exists a Fourier transformation, hence $v_h(x)$ can be written as a linear combination of Fourier components

$$v_h(x) = \int_{|\theta| \leq \pi} \widehat{v_h}(\theta) e^{i\theta \cdot x/h} d\theta, \quad x \in G_h, \quad (5.21)$$

with $x/h := (x_1/h_1, x_2/h_2) = j \in \mathbb{Z}^2$, and inverse transformation

$$\widehat{v}_h(\theta) = \frac{1}{4\pi^2} \sum_{x \in G_h} v_h(x) e^{-i\theta \cdot x/h}, \quad -\pi \leq \theta_j < \pi,$$

see e.g. [12]. Due to aliasing, Fourier components with $|\hat{\theta}| := \max\{|\theta_1|, |\theta_2|\} \geq \pi$ are not visible on G_h . These modes coincide with $e^{i\theta \cdot x/h}$, where $\theta = \hat{\theta} \pmod{2\pi}$. Hence, the Fourier space

$$\mathcal{F} := \text{span} \left\{ e^{i\theta \cdot x/h} \mid \theta \in \Theta = [-\pi, \pi)^2, x \in G_h \right\} \quad (5.22)$$

contains any bounded infinite grid function.

5.3.2 Fourier symbols of grid operators

In this section we will summarize the Fourier symbols of the basic multigrid operators, namely the fine and coarse grid operators, and the restriction, prolongation and smoothing operator. We will consider the Fourier symbols for the frequencies θ^α , with $\alpha \in \alpha_2 := \{\alpha = (\bar{\alpha}_1, \bar{\alpha}_2) \mid \bar{\alpha}_i \in \{0, 1\}, i = 1, 2\}$. These modes are defined as

$$\begin{aligned} \theta &= \theta^0 \in \Theta_{2h} := [-\pi/2, \pi/2)^2 \\ \theta^\alpha &:= \theta^0 - (\bar{\alpha}_1 \text{sign}(\theta_1), \bar{\alpha}_2 \text{sign}(\theta_2))\pi, \end{aligned}$$

see Figure 5.1.

Fine grid and smoothing operator

Define the fine grid operator as

$$(L_h v_h)(x) = \sum_{k \in J_{L_h}} l_k v_h(x + kh), \quad x \in G_h,$$

with J_{L_h} the stencil of the fine grid operator. The fine grid operator is related to its discrete Fourier transform $\widehat{L_h v_h}(\theta)$ through the relation (see (5.21))

$$(L_h v_h)(x) = \int_{|\theta| \leq \pi} \widehat{L_h v_h}(\theta) e^{i\theta \cdot x/h} d\theta. \quad (5.23)$$

The discrete Fourier transform can be further evaluated into:

$$\widehat{L_h v_h}(\theta) = \widehat{L_h}(\theta) \widehat{v_h}(\theta), \quad \text{with} \quad \widehat{L_h}(\theta) = \sum_{k \in J_{L_h}} l_k e^{i\theta \cdot k}. \quad (5.24)$$

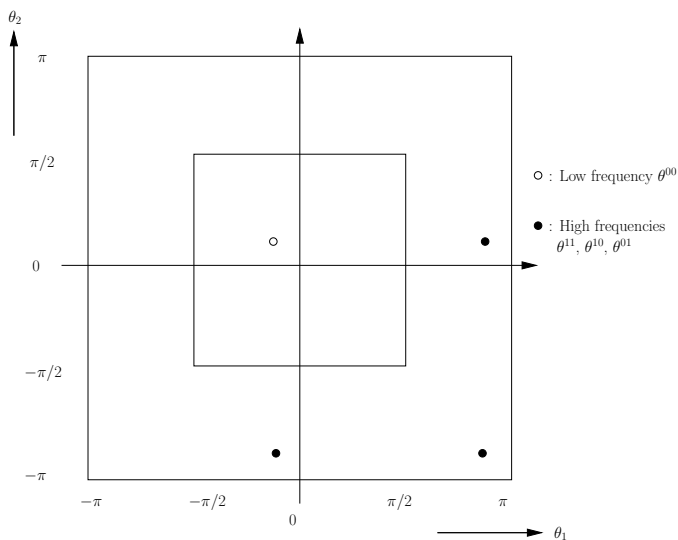


Figure 5.1: *Low and high frequencies Fourier modes in two-level multigrid.*

Similarly, we obtain for the smoothing operator S_h , which is defined as

$$(S_h v_h)(x) = \sum_{k \in J_{S_h}} s_k v_h(x + kh), \quad x \in G_h,$$

with J_{S_h} the stencil of the smoothing operator, the relation

$$(S_h v_h)(x) = \int_{|\theta| \leq \pi} \widehat{S_h v_h}(\theta) e^{i\theta \cdot x/h} d\theta \quad (5.25)$$

where the discrete Fourier transform can be further evaluated into:

$$\widehat{S_h v_h}(\theta) = \widehat{S_h}(\theta) \widehat{v_h}(\theta), \quad \text{with} \quad \widehat{S_h}(\theta) = \sum_{k \in J_{S_h}} s_k e^{i\theta \cdot k}. \quad (5.26)$$

Restriction operator

Define the restriction operator as

$$(R_h^{2h} v_h)(\bar{x}) = \sum_{k \in J_R} r_k v_h(\bar{x} + kh), \quad \bar{x} \in G_{2h}, \quad \bar{x} + kh \in G_h,$$

with J_R the stencil of the restriction operator. The restriction operator is related to its discrete Fourier transform $\widehat{R_h^{2h}v_h}(2\theta^0)$ through the relation

$$(R_h^{2h}v_h)(\bar{x}) = \int_{|\theta^0| \leq \frac{\pi}{2}} \widehat{R_h^{2h}v_h}(2\theta^0) e^{i2\theta^0 \cdot \bar{x}/(2h)} d\theta^0. \quad (5.27)$$

The discrete Fourier transform $\widehat{R_h^{2h}v_h}(2\theta^0)$ can be further evaluated into:

$$\widehat{R_h^{2h}v_h}(2\theta^0) = \sum_{\alpha \in \alpha_2} \widehat{R_h^{2h}}(\theta^\alpha) \widehat{v_h}(\theta^\alpha), \quad \text{with} \quad \widehat{R_h^{2h}}(\theta) = \sum_{k \in J_R} r_k e^{i\theta \cdot k}, \quad (5.28)$$

which shows that the restriction operator couples the fine grid modes θ^α into a coarse grid mode θ^0 .

Coarse grid operator

Define the coarse grid operator for $\bar{x} \in G_{2h}$ as

$$(L_{2h}v_{2h})(\bar{x}) = \sum_{k \in J_{L_{2h}}} l_k v_{2h}(\bar{x} + 2kh), \quad \bar{x} \in G_{2h}, \bar{x} + 2kh \in G_{2h}$$

with $J_{L_{2h}}$ the stencil of the coarse grid operator. The coarse grid operator is related to its discrete Fourier transform $\widehat{L_{2h}v_{2h}}(\theta^0)$ through the relation

$$(L_{2h}v_{2h})(\bar{x}) = \int_{|\theta^0| \leq \frac{\pi}{2}} \widehat{L_{2h}v_{2h}}(\theta^0) e^{i2\theta^0 \cdot \bar{x}/(2h)} d\theta^0. \quad (5.29)$$

The discrete Fourier transform can be further evaluated into:

$$\widehat{L_{2h}v_{2h}}(2\theta^0) = \widehat{L_{2h}}(2\theta^0) \widehat{v_{2h}}(2\theta^0), \quad \text{with} \quad \widehat{L_{2h}}(2\theta^0) = \sum_{k \in J_{L_{2h}}} l_k e^{2i\theta^0 \cdot k}, \quad (5.30)$$

where $\theta^0 \in \Theta_{2h}$. Analogously, we obtain for L_{4h} the symbol

$$\widehat{L_{4h}}(4\theta^0) = \sum_{k \in J_{L_{4h}}} l_k e^{4i\theta^0 \cdot k}.$$

Prolongation operator

The definition of the prolongation operator requires the introduction of subsets of the mesh G_h . Define the meshes G_h^α as

$$G_h^\alpha := \{x = kh \mid k \in \mathbb{Z}^2, k = \alpha \pmod{2}\}.$$

On each mesh G_h^α a separate index set J_P^α is defined as

$$J_P^\alpha := \{\alpha = (\bar{\alpha}_1, \bar{\alpha}_2) \mid \bar{\alpha}_i \in \{0, 1\}, i = 1, 2\}$$

with $J_P = \cup_{\alpha \in \alpha_d} J_P^\alpha$. The prolongation operator on the meshes G_h^α then is equal to

$$(P_{2h}^h v_{2h})(x) = \sum_{k \in J_P^\alpha} p_k v_{2h}(x + kh), \quad x \in G_h^\alpha, x + kh \in G_{2h}.$$

In two dimensions we obtain the meshes

$$\begin{aligned} G_h^{00} &= \{(x_1, x_2) = (k_1 h_1, k_2 h_2) \mid k_1 \text{ even}, k_2 \text{ even}\} \\ G_h^{11} &= \{(x_1, x_2) = (k_1 h_1, k_2 h_2) \mid k_1 \text{ odd}, k_2 \text{ odd}\} \\ G_h^{10} &= \{(x_1, x_2) = (k_1 h_1, k_2 h_2) \mid k_1 \text{ odd}, k_2 \text{ even}\} \\ G_h^{01} &= \{(x_1, x_2) = (k_1 h_1, k_2 h_2) \mid k_1 \text{ even}, k_2 \text{ odd}\}, \end{aligned}$$

with $h_i \in \mathbb{R}^+$, and the sets J_P^α which are equal to

$$\begin{aligned} J^{00} &= \{k \in J_P \mid k_1 \text{ even}, k_2 \text{ even}\}, & J^{11} &= \{k \in J_P \mid k_1 \text{ odd}, k_2 \text{ odd}\}, \\ J^{10} &= \{k \in J_P \mid k_1 \text{ odd}, k_2 \text{ even}\}, & J^{01} &= \{k \in J_P \mid k_1 \text{ even}, k_2 \text{ odd}\}. \end{aligned}$$

The prolongation operator is related to its discrete Fourier transform through the relation

$$(P_{2h}^h v_{2h})(x) = \sum_{\alpha \in \alpha_d} \int_{|\theta^0| \leq \frac{\pi}{2}} \widehat{P_{2h}^h v_{2h}}(\theta^\alpha) e^{i\theta^\alpha \cdot x/h} d\theta^0. \quad (5.31)$$

The discrete Fourier transform can be further evaluated into

$$\widehat{P_{2h}^h v_{2h}}(\theta^\alpha) = \widehat{P_{2h}^h}(\theta^\alpha) \widehat{v_{2h}}(2\theta^0) \quad (5.32)$$

with the Fourier symbol $\widehat{P_{2h}^h}(\theta^\alpha)$ defined as

$$\widehat{P_{2h}^h}(\theta^\alpha) = \sum_{\beta \in \alpha_d} \sum_{k \in J_P^\beta} p_k e^{i\theta^\alpha \cdot k}.$$

5.3.3 Two-grid Fourier analysis

For the two-grid Fourier analysis we define the $2h$ -Fourier harmonics $\mathcal{F}_{2h}(\theta)$ as

$$\mathcal{F}_{2h}(\theta) := \text{span}\{\phi_h(\theta^\alpha, x) \text{ with } \alpha \in \alpha_2\}.$$

The introduction of $2h$ -Fourier harmonics is motivated by the fact that each low frequency $\theta^0 \in \Theta_{2h}$ is coupled to the high frequency modes θ^α with $\alpha \in \alpha_2 \setminus \{0\}$. These modes are not visible on the coarse mesh G_{2h} due to aliasing.

The Fourier symbols $\widehat{L}_h(\theta)$ and $\widehat{L}_{2h}(2\theta)$ can be zero for certain values of θ . These frequencies are removed from the space of $2h$ -Fourier harmonics through the definition of

$$\mathcal{F}^{2g} := \mathcal{F}_{2h} \setminus \cup_{\theta \in \Psi_{2g}} \mathcal{F}_{2h}(\theta)$$

with

$$\Psi_{2g} := \{\theta \in \Theta_{2h} \mid \widehat{L}_{2h}(2\theta^0) = 0 \text{ or } \widehat{L}_h(\theta^\alpha) = 0\}.$$

The error e_h^D after one iteration of a two-grid multigrid cycle is determined by $e_h^D = M_h^{2g} e_h^A$, with e_h^A the initial error and M_h^{2g} the two level multigrid error transformation operator defined by (5.16). The error e_h^D has the Fourier decomposition for $x \in G_h$

$$\begin{aligned} e_h^D(x) &= \int_{|\theta| \leq \pi} \widehat{e}_h^D(\theta) e^{i\theta \cdot x/h} d\theta \\ &= \sum_{\alpha \in \alpha_2} \int_{|\theta^0| \leq \pi/2} (\widehat{M}_h^{2g} e_h^A)(\theta^\alpha) e^{i\theta^\alpha \cdot x/h} d\theta^0. \end{aligned} \quad (5.33)$$

Using the Fourier symbols discussed in Section 5.3.2, the discrete Fourier transform of $\widehat{M}_h^{2g} e_h^A(\theta^\alpha)$ can be further evaluated into

$$\begin{aligned} \widehat{M}_h^{2g} e_h^A(\theta^\alpha) &= (\widehat{S}_h(\theta^\alpha))^{\nu_1 + \nu_2} \widehat{e}_h^A(\theta^\alpha) - \\ &\quad (\widehat{S}_h(\theta^\alpha))^{\nu_2} \widehat{P}_{2h}^h(\theta^\alpha) \widehat{L}_{2h}^{-1}(2\theta^0) \sum_{\beta \in \alpha_2} \widehat{R}_h^{2h}(\theta^\beta) \widehat{L}_h(\theta^\beta) (\widehat{S}_h(\theta^\beta))^{\nu_1} \widehat{e}_h^A(\theta^\beta). \end{aligned} \quad (5.34)$$

In order to simplify notation the following matrices are defined for 2D problems

$$\widehat{L}_h^{2g}(\theta) = \text{diag}(\widehat{L}_h(\theta^{00}), \widehat{L}_h(\theta^{11}), \widehat{L}_h(\theta^{10}), \widehat{L}_h(\theta^{01})) \in \mathbb{C}^{4m \times 4m} \quad (5.35)$$

$$\widehat{S}_h^{2g}(\theta) = \text{diag}(\widehat{S}_h(\theta^{00}), \widehat{S}_h(\theta^{11}), \widehat{S}_h(\theta^{10}), \widehat{S}_h(\theta^{01})) \in \mathbb{C}^{4m \times 4m} \quad (5.36)$$

$$\widehat{R}_h^{2g}(\theta) = (\widehat{R}_h^{2h}(\theta^{00}), \widehat{R}_h^{2h}(\theta^{11}), \widehat{R}_h^{2h}(\theta^{10}), \widehat{R}_h^{2h}(\theta^{01})) \in \mathbb{C}^{m \times 4m} \quad (5.37)$$

$$\widehat{P}_h^{2g}(\theta) = (\widehat{P}_{2h}^h(\theta^{00}), \widehat{P}_{2h}^h(\theta^{11}), \widehat{P}_{2h}^h(\theta^{10}), \widehat{P}_{2h}^h(\theta^{01})^T) \in \mathbb{C}^{4m \times m} \quad (5.38)$$

where diag refers to a diagonal matrix consisting of $m \times m$ blocks with $m \geq 1$. The discrete Fourier transform of the two-level multigrid error transformation operator $\widehat{M}_h^{2g} \in \mathbb{C}^{4m \times 4m}$ then is equal to

$$\widehat{M}_h^{2g}(\theta) = (\widehat{S}_h^{2g}(\theta))^{\nu_2} \left(I^{2g} - \widehat{P}_h^{2g}(\theta) \widehat{L}_{2h}^{-1}(2\theta^0) \widehat{R}_h^{2g}(\theta) \widehat{L}_h^{2g}(\theta) \right) (\widehat{S}_h^{2g}(\theta))^{\nu_1}, \quad (5.39)$$

with $\theta \in \Theta_{2h} \setminus \Psi_{2g}$ and I^{2g} the $4m \times 4m$ identity matrix.

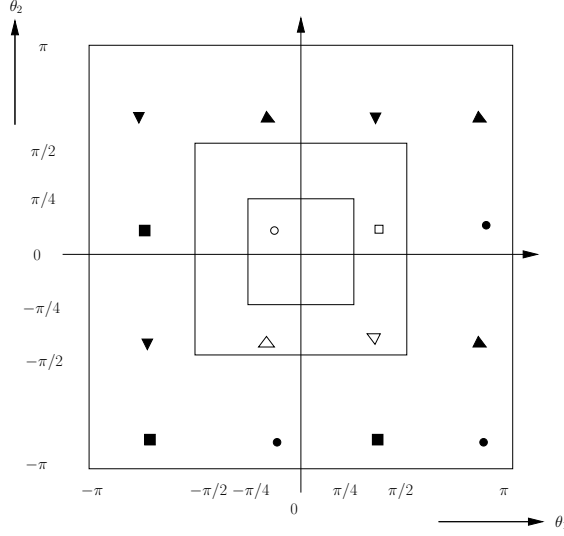


Figure 5.2: *Low, medium and high frequencies Fourier modes in three-level multigrid.*

5.3.4 Three-grid Fourier analysis

For the three-grid Fourier analysis we define the Fourier harmonics $\mathcal{F}_{4h}(\theta)$ as

$$\begin{aligned}
 \mathcal{F}_{4h}(\theta) &:= \text{span}\{\phi_h(\theta_\beta^\alpha, x) \mid \alpha \in \alpha_2, \beta \in \beta_2\}, \quad \text{where} \\
 \theta &= \theta_0^0 \in \Theta_{4h} := [-\pi/4, \pi/4]^2, \\
 \theta_\beta &= \theta_0^0 - (\bar{\beta}_1 \text{sign}(\theta_1), \bar{\beta}_2 \text{sign}(\theta_2))\pi, \\
 \theta_\beta^\alpha &:= \theta_\beta - (\bar{\alpha}_1 \text{sign}((\theta_1)_\beta), \bar{\alpha}_2 \text{sign}((\theta_2)_\beta))\pi, \\
 \beta_2 &:= \{\beta = (\bar{\beta}_1, \bar{\beta}_2) \mid \bar{\beta}_i \in \{0, \frac{1}{2}\}, i = 1, 2\},
 \end{aligned} \tag{5.40}$$

see Figure 5.2. Note that we now have 16 coupled Fourier harmonics all related to θ_{00}^{00} . In the transition from G_{2h} to G_{4h} the modes $\theta_\beta = \theta_0^0$ are not visible due to aliasing. Hence for $\bar{x} \in G_{4h}$ we have

$$\phi_h(\theta_{00}^{00}, \bar{x}) = \phi_h(\theta_{\frac{1}{2}\frac{1}{2}}^{00}, \bar{x}) = \phi_h(\theta_{\frac{1}{2}0}^{00}, \bar{x}) = \phi_h(\theta_{0\frac{1}{2}}^{00}, \bar{x}),$$

and all modes θ_β^α on G_{4h} alias to θ_{00}^{00} resulting in the relation

$$\phi_{2h}(2\theta_{00}^{00}, \bar{x}) = \phi_{4h}(4\theta_{00}^{00}, \bar{x}) = \phi_{4h}(4\theta_\beta^\alpha, \bar{x}) \quad \text{with } \bar{x} \in G_{4h}, \theta_{00}^{00} \in \Theta_{4h}.$$

The Fourier symbols $\widehat{L}_h(\theta)$, $\widehat{L}_{2h}(2\theta)$ and $\widehat{L}_{4h}(4\theta)$ can be zero for certain values of θ . These frequencies are removed from the space of Fourier harmonics through the definition of

$$\mathcal{F}^{3g} := \mathcal{F}_{4h} \setminus \cup_{\theta \in \Psi_{3g}} \mathcal{F}_{4h}(\theta),$$

with

$$\Psi_{3g} := \{\theta \in \Theta_{4h} \mid \widehat{L}_{4h}(4\theta_0^0) = 0 \text{ or } \widehat{L}_{2h}(2\theta_\beta^0) = 0 \text{ or } \widehat{L}_h(\theta_\beta^\alpha) = 0\}.$$

The error e_h^D after one iteration of a three-grid multigrid cycle is determined by $e_h^D = M_h^{3g} e_h^A$, with e_h^A the initial error and M_h^{3g} the three-level multigrid error transformation operator defined by (5.17).

In order to compute the discrete Fourier transform of M_h^{3g} we need some intermediate results. First, all modes $\phi_h(\theta_\beta^\alpha, x)$ on G_h alias to $\phi_{2h}(2\theta_\beta^0, \bar{x})$ on the mesh G_{2h} . In order to compute the effect of the coarse grid correction operator M_{2h} we use (5.34), replace θ^α with $2\theta_\beta^0$, and h with $2h$. The discrete Fourier transform of $\widehat{M_{2h}^{2g} e_{2h}^A}(\theta^\alpha)$ is thus equal to

$$\begin{aligned} \widehat{M_{2h}^{2g} e_{2h}^A}(2\theta_\beta^0) &= (\widehat{S_{2h}}(2\theta_\beta^0))^{\nu_3 + \nu_4} \widehat{e_{2h}^A}(2\theta_\beta^0) - (\widehat{S_{2h}}(2\theta_\beta^0))^{\nu_4} \widehat{P_{4h}^{2h}}(2\theta_\beta^0) \widehat{L_{4h}^{-1}}(4\theta_0^0) \times \\ &\quad \sum_{\beta_1 \in \alpha_2} \widehat{R_{2h}^{4h}}(2\theta_{\beta_1}^0) \widehat{L_{2h}}(2\theta_{\beta_1}^0) (\widehat{S_{2h}}(2\theta_{\beta_1}^0))^{\nu_3} \widehat{e_{2h}^A}(2\theta_{\beta_1}^0). \end{aligned} \quad (5.41)$$

Introduce the coarse grid correction matrix as $\widetilde{M}_{2h} = I_{2h} - M_{2h}^{\gamma_c}$. Using (5.41) we obtain the relation

$$\widetilde{\widehat{M_{2h} e_{2h}^A}}(2\theta_\beta^0) = \sum_{\tilde{\beta}_2 \in \beta_2} \widehat{M}(\beta, \tilde{\beta}_2; \gamma_c) \widehat{e_{2h}^A}(2\theta_{\tilde{\beta}_2}^0) \quad (5.42)$$

where the coefficients $\widehat{M}(\beta, \tilde{\beta}_2; \gamma_c)$ are independent of $\widehat{e_{2h}^A}(2\theta_\beta^0)$.

The error e_h^D has for $x \in G_h$ the Fourier decomposition

$$\begin{aligned} e_h^D(x) &= \int_{|\theta| \leq \pi} \widehat{e_h^D}(\theta) e^{i\theta \cdot x/h} d\theta \\ &= \sum_{\beta \in \beta_2} \sum_{\gamma \in \alpha_2} \int_{|\theta^0| \leq \pi/4} (\widehat{M_h^{3g} e_h^A})(\theta_\beta^\gamma) e^{i\theta_\beta^\gamma \cdot x/h} d\theta_0^0. \end{aligned} \quad (5.43)$$

The discrete Fourier transform of $\widehat{M}_h^{3g} e_h^A(\theta_\beta^\gamma)$ can be further evaluated into

$$\begin{aligned} \widehat{M}_h^{3g} e_h^A(\theta_\beta^\gamma) &= (\widehat{S}_h(\theta_\beta^\gamma))^{\nu_1+\nu_2} \widehat{e}^A(\theta_\beta^\gamma) - (\widehat{S}_h(\theta_\beta^\gamma))^{\nu_2} \widehat{P}_{2h}^h(\theta_\beta^\gamma) \times \\ &\quad \sum_{\widetilde{\beta}_2 \in \beta_2} \widehat{M}(\beta, \widetilde{\beta}_2; \gamma_c) \widehat{L}_{2h}^{-1}(2\theta_{\widetilde{\beta}_2}^0) \sum_{\alpha \in \alpha_2} \widehat{R}_h^{2h}(\theta_{\widetilde{\beta}_2}^\alpha) \widehat{L}_h(\theta_{\widetilde{\beta}_2}^\alpha) (\widehat{S}_h(\theta_{\widetilde{\beta}_2}^\alpha))^{\nu_1} \widehat{e}^A(\theta_{\widetilde{\beta}_2}^\alpha). \end{aligned}$$

In order to simplify notation the following matrices are defined

$$\widehat{L}_h^{2g}(\theta_\beta) = \text{diag}(\widehat{L}_h(\theta_\beta^{00}), \widehat{L}_h(\theta_\beta^{11}), \widehat{L}_h(\theta_\beta^{10}), \widehat{L}_h(\theta_\beta^{01})) \in \mathbb{C}^{4m \times 4m} \quad (5.44)$$

$$\widehat{S}_h^{2g}(\theta_\beta) = \text{diag}(\widehat{S}_h(\theta_\beta^{00}), \widehat{S}_h(\theta_\beta^{11}), \widehat{S}_h(\theta_\beta^{10}), \widehat{S}_h(\theta_\beta^{01})) \in \mathbb{C}^{4m \times 4m} \quad (5.45)$$

$$\widehat{R}_h^{2g}(\theta_\beta) = (\widehat{R}_h^{2h}(\theta_\beta^{00}), \widehat{R}_h^{2h}(\theta_\beta^{11}), \widehat{R}_h^{2h}(\theta_\beta^{10}), \widehat{R}_h^{2h}(\theta_\beta^{01})) \in \mathbb{C}^{m \times 4m} \quad (5.46)$$

$$\widehat{P}_h^{2g}(\theta_\beta) = (\widehat{P}_{2h}^h(\theta_\beta^{00}), \widehat{P}_{2h}^h(\theta_\beta^{11}), \widehat{P}_{2h}^h(\theta_\beta^{10}), \widehat{P}_{2h}^h(\theta_\beta^{01}))^T \in \mathbb{C}^{4m \times m} \quad (5.47)$$

where diag refers to a diagonal matrix consisting of $m \times m$ blocks with $m \in \mathbb{N}$.

We also introduce the matrices

$$\widehat{L}_h^{3g}(\theta) = \text{bdiag}(\widehat{L}_h^{2g}(\theta_{00}), \widehat{L}_h^{2g}(\theta_{\frac{1}{2}\frac{1}{2}}), \widehat{L}_h^{2g}(\theta_{\frac{1}{2}0}), \widehat{L}_h^{2g}(\theta_{0\frac{1}{2}})) \in \mathbb{C}^{16m \times 16m}$$

$$\widehat{S}_h^{3g}(\theta) = \text{bdiag}(\widehat{S}_h^{2g}(\theta_{00}), \widehat{S}_h^{2g}(\theta_{\frac{1}{2}\frac{1}{2}}), \widehat{S}_h^{2g}(\theta_{\frac{1}{2}0}), \widehat{S}_h^{2g}(\theta_{0\frac{1}{2}})) \in \mathbb{C}^{16m \times 16m}$$

$$\widehat{R}_h^{3g}(\theta) = \text{bdiag}(\widehat{R}_h^{2g}(\theta_{00}), \widehat{R}_h^{2g}(\theta_{\frac{1}{2}\frac{1}{2}}), \widehat{R}_h^{2g}(\theta_{\frac{1}{2}0}), \widehat{R}_h^{2g}(\theta_{0\frac{1}{2}})) \in \mathbb{C}^{4m \times 16m}$$

$$\widehat{P}_h^{3g}(\theta) = \text{bdiag}(\widehat{P}_h^{2g}(\theta_{00}), \widehat{P}_h^{2g}(\theta_{\frac{1}{2}\frac{1}{2}}), \widehat{P}_h^{2g}(\theta_{\frac{1}{2}0}), \widehat{P}_h^{2g}(\theta_{0\frac{1}{2}})) \in \mathbb{C}^{16m \times 4m}$$

$$\widehat{Q}_h^{3g}(\theta) = \text{bdiag}(\widehat{L}_{2h}^{-1}(2\theta_{00}), \widehat{L}_{2h}^{-1}(2\theta_{\frac{1}{2}\frac{1}{2}}), \widehat{L}_{2h}^{-1}(2\theta_{\frac{1}{2}0}), \widehat{L}_{2h}^{-1}(2\theta_{0\frac{1}{2}})) \in \mathbb{C}^{4m \times 4m}$$

$$\widehat{U}^{3g}(\theta; \gamma_c) = \widehat{M}(\theta; \gamma_c) \in \mathbb{C}^{4m \times 4m}.$$

The discrete Fourier transform of the error transformation operator for a three-level multigrid cycle $\widehat{M}_h^{3g}(\theta) \in \mathbb{C}^{16m \times 16m}$ then is equal to

$$\widehat{M}_h^{3g}(\theta) = (\widehat{S}_h^{3g}(\theta))^{\nu_2} \left(I^{3g} - \widehat{P}_h^{3g}(\theta) \widehat{U}^{3g}(\theta; \gamma_c) \widehat{Q}_h^{3g}(\theta) \widehat{R}_h^{3g}(\theta) \widehat{L}_h^{3g}(\theta) \right) (\widehat{S}_h^{3g}(\theta))^{\nu_1} \quad (5.48)$$

with $\theta \in \Theta_{4h} \setminus \Psi_{3g}$ and where I^{3g} the $16m \times 16m$ identity matrix. We still need to obtain an explicit expression for $\widehat{U}^{3g}(\theta; \gamma_c)$. On the mesh G_{2h} the modes θ_β^α reduce after the restriction operator to modes $2\theta_\beta^0$, hence using the result of the two-level analysis given by (5.39), the coarse grid error transformation operator is equal to

$$\widehat{M}_{2h}^{2g}(2\theta_\beta) = (\widehat{S}_{2h}^{2g}(2\theta_\beta))^{\nu_4} \left(I^{2g} - \widehat{P}_{2h}^{2g}(2\theta_\beta) \widehat{L}_{4h}^{-1}(4\theta_\beta^0) \widehat{R}_{2h}^{2g}(2\theta_\beta) \widehat{L}_{2h}^{2g}(2\theta_\beta) \right) (\widehat{S}_{2h}^{2g}(2\theta_\beta))^{\nu_3}, \quad (5.49)$$

with $2\theta_\beta \in \Theta_{2h} \setminus \Psi_{2g}$, I^{2g} the $4m \times 4m$ identity matrix and \widehat{L}_{2h}^{2g} , \widehat{S}_{2h}^{2g} , \widehat{R}_{2h}^{2g} and \widehat{P}_{2h}^{2g} given by (5.44)-(5.47), respectively, with h replaced by $2h$. The matrix $\widehat{U}^{3g}(\theta; \gamma_c)$ then is equal to

$$\widehat{U}^{3g}(\theta; \gamma_c) = I^{2g} - (\widehat{M}_{2h}^{2g}(2\theta_\beta))^{\gamma_c}.$$

5.3.5 Spectral radius

The spectral radius of the error transformation operator gives a prediction of the asymptotic rate of convergence of the multigrid method. The asymptotic convergence factor per cycle is defined as

$$\mu = \lim_{m \rightarrow \infty} \left(\sup_{e_h^{(0)} \neq 0} \frac{\|e_h^{(m)}\|_{\ell^2(G_h)}}{\|e_h^{(0)}\|_{\ell^2(G_h)}} \right)^{\frac{1}{m}}$$

where $e_h^{(m)}$ is the error after m applications of the multigrid cycle, hence $e_h^{(0)} = e_h^A$ and $e_h^{(1)} = e_h^D$. It can be shown that the asymptotic convergence factor is equal to

$$\mu = \sup_{\theta \in \Theta_{ng} \setminus \Psi_{ng}} \rho(\widehat{M}^{ng}(\theta)),$$

where ρ is the spectral radius defined as:

$$\rho(M) = \max\{|\lambda| \mid \lambda \in \sigma(M)\}, \text{ where } \sigma(M) = \{\lambda \in \mathbb{C} \mid \lambda \text{ is an eigenvalue of } M\}.$$

Note that a requirement for convergence is that the spectral radius has to satisfy $\mu < 1$. We will optimize our multigrid algorithm by minimizing the spectral radius of the three-level multigrid error transformation operator (5.48).

5.4 Optimizing multigrid

In this section we will discuss the optimization of the Runge-Kutta smoothers. We will consider 5-stage smoothers discussed in Section 5.1.3, distinguishing between RK schemes with only non-zero diagonal terms versus Runge-Kutta smoothers with all coefficients in (5.12) non-zero. For notational purposes, a diagonal 5-stage RK scheme is denoted dRK5, while a full 5-stage RK is denoted as fRK5. We will optimize the coefficients of the smoothers for three-level multigrid computations for a space-time DG discretization of the advection-diffusion equation in two-space dimensions using linear and quadratic basis-functions for steady flows.

To optimize our multigrid algorithm we will search for coefficients in the Runge-Kutta smoothers (5.12) for which the spectral radius of the three-level multigrid error transformation operator (5.48) is minimal. For the optimization, we use the procedures `fminsearch` and `fmincon`, available in Matlab. As constraint in the `fmincon` procedure, we require that the spectral radius of the smoother and the three-level multigrid error transformation operator are less than 1. The optimization will be performed for advection dominated flows in which we fix the Reynolds numbers $Re_x = Re_y = 100$. We also fix the flow angle $\gamma^{flow} = \pi/4$ the aspect ratio $AR = 1$ and $CFL^{\Delta t} = 100$ for steady flows. For the different optimized smoothers, the optimized coefficients are given in Tables 5.1 and 5.2. In Figures 5.3-5.6 we plot the stability domain and eigenvalue spectra of the discretization, smoother and three-level algorithm for the new dRK5 and fRK5 smoothers. We remark that the optimization is performed by approximating the exact solution on the coarsest grid by performing ν_C iterations of the smoother. If, however, we perform the optimization by using the exact solution on the coarsest grid instead, the Runge-Kutta coefficients may be different, but the spectral radius of the multigrid algorithm will be approximately the same.

The optimization results in Tables 5.1 and 5.2 show that by optimizing the Runge-Kutta smoothers a significant reduction in the spectral radius of the multigrid error transformation matrix ρ^{MG} can be obtained. This should result in a significantly improved convergence rate of the multigrid scheme, but needs to be verified on actual computations since only periodic boundary conditions are taken into account in the Fourier analysis. See Section 5.6 on the performance of the multigrid algorithms.

The optimization process has a big impact on the Runge-Kutta coefficients which greatly differ per case. Also, the Runge-Kutta coefficients are very different from commonly used schemes since they are optimized for a fast multigrid convergence and not for time accuracy. It really pays off to optimize close to a realistic flow state, but of course this has its own limitations. Finding a good balance between optimization and more generally applicable algorithms is still an open question. In practice, the Runge-Kutta smoothers use local time stepping and this gives the opportunity to apply locally the best smoother for the actual flow state.

5.5 Rescaling

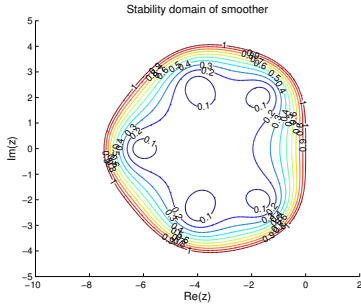
On grids with high aspect ratios, a strong coupling in one direction versus a weak coupling in the other directions are a cause for slow convergence of the multigrid method [77]. This can also be seen in the scaling of the different operators of the discretization. If AR is large, the scaling of \mathcal{L}_x^a and \mathcal{L}_x^d in (5.6) and (5.7) becomes large so that the discretization terms in the x -direction are dominant.

Steady flows, linear basis functions, 3-level multigrid				
	$dRK5$		$fRK5$	
	initial	optimized	initial	optimized
α_{21}	0.0791451	0.05768995298	0.05768995298	0.0578331573
α_{31}	-	-	0.0	-0.0002051554736
α_{32}	0.163551	0.1405960888	0.1405960888	0.1403808301
α_{41}	-	-	0.0	0.0003953470071
α_{42}	-	-	0.0	-0.001195029164
α_{43}	0.283663	0.267958213	0.267958213	0.2681810517
α_{51}	-	-	0.0	0.0001441249202
α_{52}	-	-	0.0	-0.0002608610327
α_{53}	-	-	0.0	-0.0003368070181
α_{54}	0.5	0.5	0.5	0.8473374098
α_{61}	-	-	0.0	0.4115573097
α_{62}	-	-	0.0	-0.003144851878
α_{63}	-	-	0.0	-0.0001096455683
α_{64}	-	-	0.0	0.001555741114
α_{65}	1.0	1.0	1.0	0.5901414466
β_1	0.0791451	0.05768995298	0.05768995298	0.04887040625
β_2	0.163551	0.1405960888	0.1405960888	0.1274785795
β_3	0.283663	0.267958213	0.267958213	0.2287556298
β_4	0.5	0.5	0.5	0.9547064029
β_5	1.0	1.0	1.0	2.52621971
CFL^τ	0.8	0.8	0.8	0.8
VN^τ	0.8	0.8	0.8	0.8
ρ^S	5.943	0.98812	0.98812	0.98914
ρ^{MG}	167.06	0.89151	0.89151	0.81762

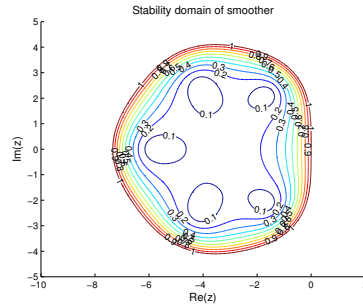
Table 5.1: *Initial and optimized coefficients for the $dRK5$ and $fRK5$ smoothers for 3-level multigrid for steady flows with $CFL^{\Delta t} = 100$ and $Re_x = Re_y = 100$. The flow angle is $\gamma^{flow} = \pi/4$ and the element aspect-ratio $AR = 1$. Here, $\nu_1 = \nu_2 = \nu_3 = \nu_4 = 1$, $\nu_C = 4$ and $\gamma = 1$.*

Steady flows, quadratic basis functions, 3-level multigrid				
	dRK5		fRK5	
	initial	optimized	initial	optimized
α_{21}	0.0791451	0.04865009589	0.04865009589	0.04877436325
α_{31}	-	-	0.0	-0.0002188348438
α_{32}	0.163551	0.130316854	0.130316854	0.1300906122
α_{41}	-	-	0.0	2.608884832e-05
α_{42}	-	-	0.0	2.444376496e-05
α_{43}	0.283663	0.2729621396	0.2729621396	0.2734805705
α_{51}	-	-	0.0	-0.001250385487
α_{52}	-	-	0.0	-0.0007838720635
α_{53}	-	-	0.0	-0.0004890887712
α_{54}	0.5	0.5	0.5	4.412139367
α_{61}	-	-	0.0	0.8097217358
α_{62}	-	-	0.0	0.08435089009
α_{63}	-	-	0.0	-0.01986799007
α_{64}	-	-	0.0	0.01359815476
α_{65}	1.0	1.0	1.0	0.1121972094
β_1	0.0791451	0.04865009589	0.04865009589	0.5551936269
β_2	0.163551	0.130316854	0.130316854	0.1333199239
β_3	0.283663	0.2729621396	0.2729621396	-1.332263675
β_4	0.5	0.5	0.5	-3.649588578
β_5	1.0	1.0	1.0	0.46771792
CFL^τ	0.4	0.4	0.4	0.4
VN^τ	0.4	0.4	0.4	0.4
ρ^S	10.684	0.98974	0.98974	0.9896
ρ^{MG}	124.02	0.90049	0.90049	0.89903

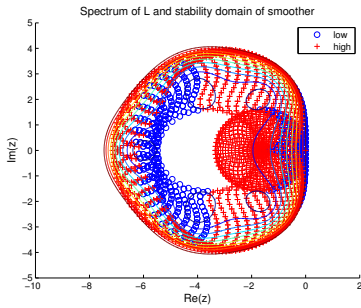
Table 5.2: Initial and optimized coefficients for the dRK5 and fRK5 smoothers for 3-level multigrid for steady flows with $CFL^{\Delta t} = 100$ and $Re_x = Re_y = 100$. The flow angle is $\gamma^{flow} = \pi/4$ and element aspect-ratio $AR = 1$. Here, $\nu_1 = \nu_2 = \nu_3 = \nu_4 = 1$, $\nu_C = 4$ and $\gamma = 1$.



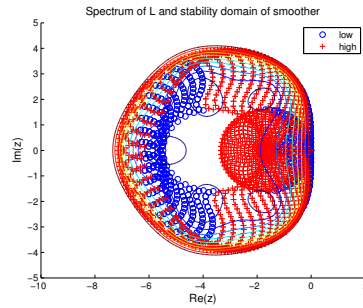
(a) Stability domain dRK5.



(b) Stability domain fRK5.

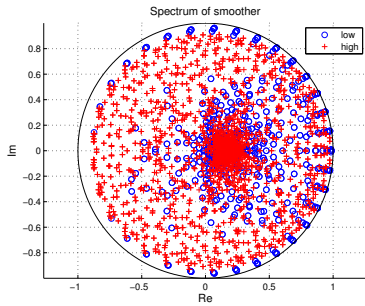


(c) Spectrum of \mathcal{L}_h and stability domain dRK5.

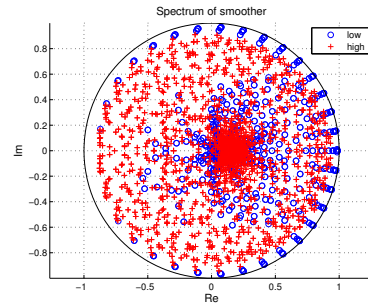


(d) Spectrum of \mathcal{L}_h and stability domain fRK5.

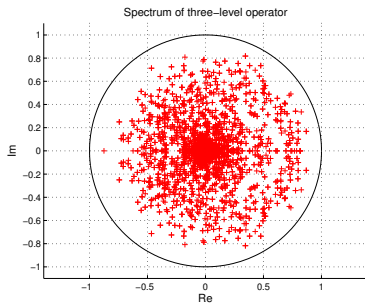
Figure 5.3: Stability domain of the dRK5 smoother (left) and the fRK5 smoother (right) and eigenvalue spectra of \mathcal{L}_h for space-time DG discretizations of the 2D advection-diffusion equation using linear basis functions (steady flow case, $CFL^{\Delta t} = 100$).



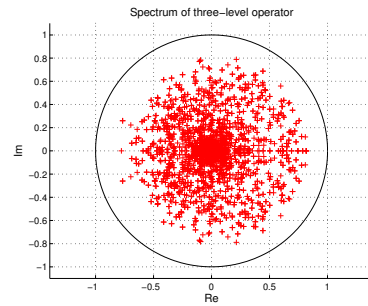
(a) Eigenvalue spectra dRK5 smoother.



(b) Eigenvalue spectra fRK5 smoother.

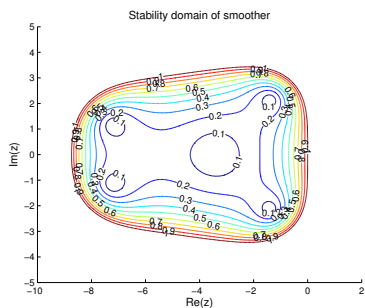


(c) Eigenvalue spectra 3-level MG with dRK5.

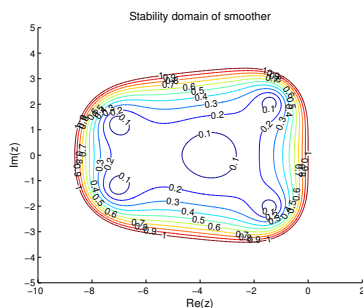


(d) Eigenvalue spectra 3-level MG with fRK5.

Figure 5.4: *Eigenvalue spectra of the dRK5 smoother (left) and the fRK5 smoother (right) for space-time DG discretizations of the 2D advection-diffusion equation using linear basis functions (steady flow case, $CFL^{\Delta t} = 100$).*



(a) Stability domain dRK5.



(b) Stability domain fRK5.

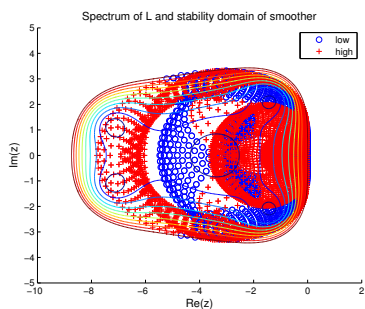
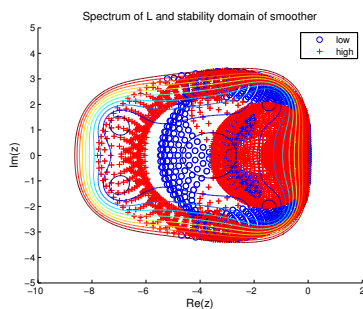
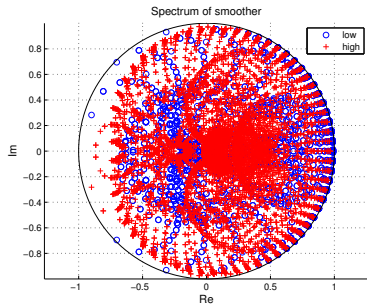
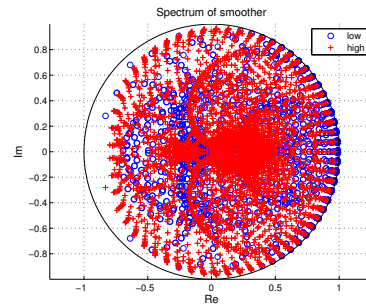
(c) Spectrum of \mathcal{L}_h and stability domain dRK5.(d) Spectrum of \mathcal{L}_h and stability domain fRK5.

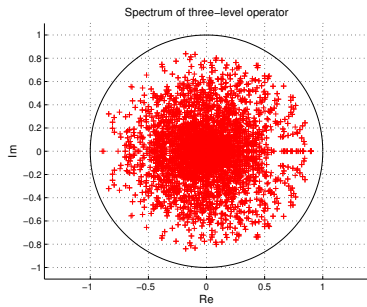
Figure 5.5: Stability domain of the dRK5 smoother (left) and the fRK5 smoother (right) and eigenvalue spectra of \mathcal{L}_h for space-time DG discretizations of the 2D advection-diffusion equation using quadratic basis functions (steady flow case, $CFL^{\Delta t} = 100$).



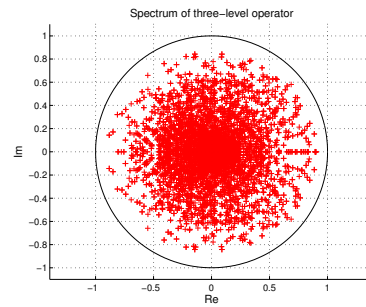
(a) Eigenvalue spectra dRK5 smoother.



(b) Eigenvalue spectra fRK5 smoother.



(c) Eigenvalue spectra 3-level MG with dRK5.



(d) Eigenvalue spectra 3-level MG with fRK5.

Figure 5.6: *Eigenvalue spectra of the dRK5 smoother (left) and the fRK5 smoother (right) for space-time DG discretizations of the 2D advection-diffusion equation using quadratic basis functions (steady flow case, $CFL^{\Delta t} = 100$).*

Similarly, if AR is small, the scaling of \mathcal{L}_y^a and \mathcal{L}_y^d in (5.6) and (5.7) becomes large so that the discretization terms in the y -direction are dominant. If viscosity becomes important, in for instance a boundary layer, the diffusion operators \mathcal{L}_x^d and \mathcal{L}_y^d in (5.7) start to dominate over the advection terms (5.6) resulting also in slower convergence. In this section we introduce a rescaling to the pseudo-time algorithm. This rescaling is possible since we are only interested in fast convergence to steady-state in pseudo-time and not in pseudo-time accuracy. We introduce a rescaling that better balances the different operators for all aspect ratio cells, whether the flow is inviscid or viscous. This in contrast to the pseudo-time algorithm in [40] in which a combination of two Runge-Kutta smoothers was used for flows with inviscid and viscous regions: the EXI Runge-Kutta smoother for inviscid flows and the EXV Runge-Kutta smoother for viscous flows. This scaling is important since it also allows to optimize the multigrid algorithm for only the unit aspect ratio.

Consider again the pseudo-time system (5.11). The reason why the multigrid algorithm performs badly on high aspect ratio cells is related to the scalings of $\mathcal{L}_{x,y}^{a,d}$ in (5.6) and (5.7). Adding a rescaling \mathbb{S} to the l.h.s. of (5.11) balances the different operators:

$$\mathbb{S}\mathcal{M}\frac{\partial \hat{u}^*}{\partial \tau} = -\mathcal{L}_h(\hat{u}^*; \hat{u}^{n-1}). \quad (5.50)$$

There are different options as to how to choose the rescaling \mathbb{S} . We introduce the following rescaling:

$$\mathbb{S} = \frac{\sqrt{1 + AR^2}}{\sqrt{8}} \left(\cos(\gamma^{flow}) + \frac{\sin(\gamma^{flow})}{AR} \right) + \frac{1}{\sqrt{8}} \left(\frac{1}{Re_x} + \frac{1}{Re_y} \right),$$

which is just $(\Gamma_x^a + \Gamma_y^a + \Gamma_x^d + \Gamma_y^d)/CFL$ and scaled such that when $\gamma^{flow} = \pi/4$, $Re_x = Re_y = \infty$ and $AR = 1$, then $\mathbb{S} = 1$.

5.6 Testing multigrid performance

In order to demonstrate the performance of the optimized algorithms we consider the following initial boundary value problem:

$$\begin{aligned} u_t + a_k u_{,k} - \nu(u_{,k})_{,k} &= 0, & (x_1, x_2) \in \Omega = (0, 1)^2, & t \in \mathbb{R}^+, \\ u(x_1, x_2, t) &= g(x_1, x_2), & (x_1, x_2) \in \partial\Omega, & t \in \mathbb{R}^+, \\ u(x_1, x_2, 0) &= 1 - \frac{1}{2}(x_1 + x_2) & (x_1, x_2) \in \Omega, \end{aligned} \quad (5.51)$$

in which $g(x_1, x_2)$ equals at the domain boundary the exact steady state solution of (5.51) given by:

$$u(x_1, x_2) = \frac{1}{2} \left(\frac{\exp(a_1/\nu_1) - \exp(a_1 x_1/\nu_1)}{\exp(a_1/\nu_1) - 1} + \frac{\exp(a_2/\nu_2) - \exp(a_2 x_2/\nu_2)}{\exp(a_2/\nu_2) - 1} \right).$$

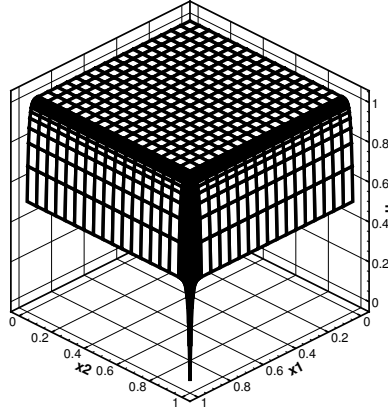


Figure 5.7: Solution to the advection-diffusion equation given by (5.51) on a Shishkin mesh with flow angle $\gamma^{flow} = \pi/4$.

We consider two flow angles with respect to the x -axis, respectively, $\pi/4$ and $\pi/8$. In the discretization we use a Shishkin mesh in which the coordinates (x^u, y^u) of a uniform mesh are mapped onto a mesh suitable for dealing with boundary layers. The mapping is given by:

$$x_i = \begin{cases} 2(1 - \sigma_i)x_i^u, & \text{for } x_i^u < 0.5 \\ 1 + 2\sigma_i(x_i^u - 1), & \text{for } x_i^u \geq 0.5 \end{cases}, \quad i = 1, 2, \quad x = x_1, y = x_2,$$

where $\sigma_i = \min(0.5, 2\nu/a \ln(N_i))$, and where N_i is the number of elements in the x_i direction.

In Figure 5.7 we show the solution on a Shishkin mesh with flow angle $\gamma^{flow} = \pi/4$. In Figures 5.8 and 5.9, we show the convergence results using different smoothers in the 3-level multigrid. In Figure 5.10 we make a zoom of Figure 5.9 closer to the vertical axis. In these figures, one work unit on the fine grid is defined as one full Runge-Kutta step. On the medium grid, one full Runge-Kutta step corresponds to 1/4 work units while on the coarsest grid a full Runge-Kutta step corresponds to 1/16 work units. It follows that one full multigrid cycle then corresponds to $(\nu_1 + \nu_2) + (\nu_3 + \nu_4)/4 + \nu_C/16$ work units.

The parameters in the test cases are the following: we consider one physical time step, with $\Delta t = 100$, $a = \sqrt{2}$, $\nu_x = \nu_y = 0.01$, $N_1 = N_2 = 32$, and two flow angles, viz. $\gamma^{flow} = \pi/4$ and $\gamma^{flow} = \pi/8$. Depending on the stability of the smoother, we use different CFL^τ and VN^τ numbers. For the original EXI-EXV pseudo-time stepping scheme, see [40], we define $Re_i = a\Delta x_i^2/(\nu_i h)$,

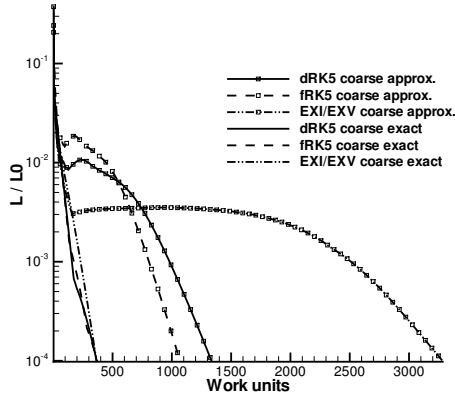
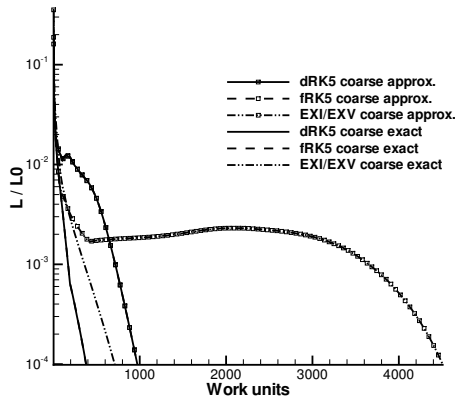
(a) Flow angle $\pi/4$.(b) Flow angle $\pi/8$.

Figure 5.8: Convergence results of space-time DG using linear basis functions for three level multigrid algorithms with different Runge-Kutta smoothers. (dRK5, diagonal explicit 5-stage RK scheme, fRK5 general explicit 5-stage RK scheme, EXI-EXV scheme [40], exact and approximate solution of equations on coarsest mesh).

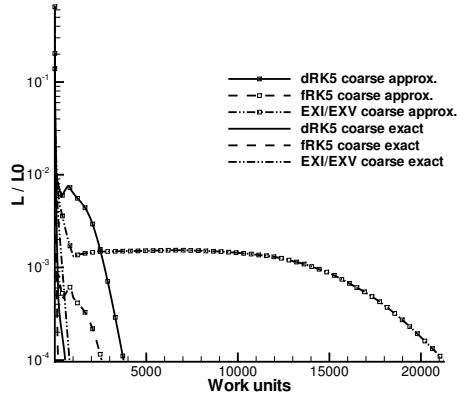
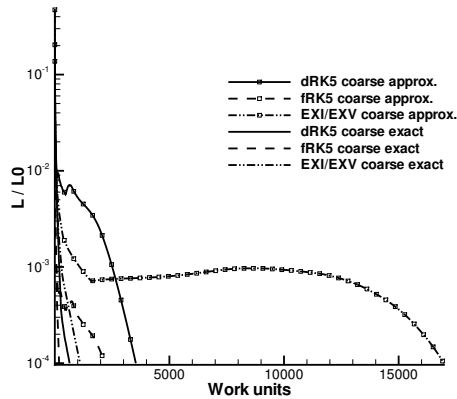
(a) Flow angle $\pi/4$.(b) Flow angle $\pi/8$.

Figure 5.9: Convergence results of space-time DG using quadratic basis functions for three level multigrid algorithms with different Runge-Kutta smoothers. (dRK5, diagonal explicit 5-stage RK scheme, fRK5 general explicit 5-stage RK scheme, EXI-EXV scheme [40], exact and approximate solution of equations on coarsest mesh).

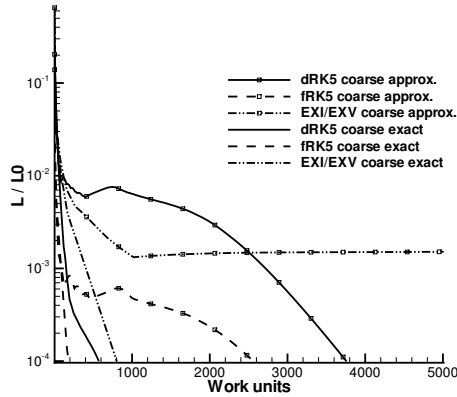
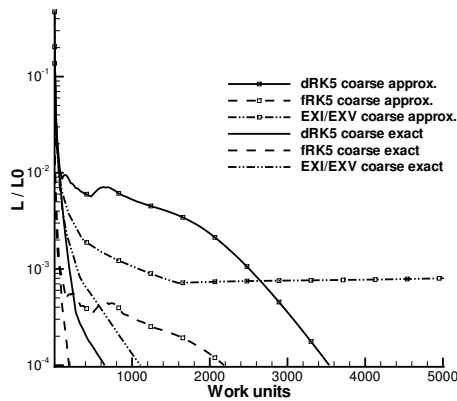
(a) Flow angle $\pi/4$ (zoom).(b) Flow angle $\pi/8$ (zoom).

Figure 5.10: Zoom of the convergence results of space-time DG using quadratic basis functions for three level multigrid algorithms with different Runge-Kutta smoothers. (dRK5, diagonal explicit 5-stage RK scheme, fRK5 general explicit 5-stage RK scheme, EXI-EXV scheme [40], exact and approximate solution of equations on coarsest mesh).

with $h = \Delta x_1 \sqrt{1 + AR^2}$. The EXV scheme was used when $Re_i \leq 1$ and the EXI scheme was used otherwise. For the EXI-EXV smoother in [40], no mass-matrix was used in the pseudo-time calculation. The mass-matrix was, however, used in the computations with the optimized smoothers. For the multigrid computations, $\nu_1 = \nu_2 = \nu_3 = \nu_4 = 1$, $\gamma = 1$ and $\nu_C = 4$ for the approximate calculation on the coarse mesh. Both an exact solution of the linear system on the coarsest mesh and an approximate solution with ν_C pseudo-time iterations on the coarse mesh were tested.

Comparing the results shown in Figures 5.8 and 5.9 we can draw the following conclusions. First of all, we see that using the optimized Runge-Kutta smoothers of Tables 5.1 and 5.2 a big improvement is obtained over the original EXI-EXV smoother. For linear basis functions and a flow angle of $\pi/8$ the number of work units to obtain 4 orders reduction in the residual is reduced from 4507 to 379. For quadratic basis functions and a flow angle of $\pi/4$ the number of work units reduces from 21254 to 184.

For linear basis functions the difference in convergence rate between Runge-Kutta smoothers with only non-zero diagonal terms versus Runge-Kutta smoothers with all coefficients non-zero in (5.12) is negligible. For quadratic basis functions this difference is, however, significant. Using more Runge-Kutta coefficients enlarges the possibilities to optimize the smoother, but the optimization process requires a significantly larger computing time. In order to speed up the optimization process the coefficients of Runge-Kutta schemes with only non-zero diagonal terms are used as initial values.

The effect of solving the equations on the coarsest mesh with high accuracy is very large. Without this the multigrid convergence significantly slows down after a rapid initial decrease of the residual. In particular, for nonlinear problems it is tempting to solve the algebraic system on the coarsest mesh only approximately, because otherwise a global Newton solver is required. The effect of solving the algebraic equations on the coarsest mesh accurately is, however, non-negligible.

The flow angle has a mild effect on the convergence rate. In general if the flow direction is close to one of the mesh lines the convergence rate is the slowest.

It is important not to extrapolate the results for the advection-diffusion equation directly to the Euler and Navier-Stokes equations. In these cases other aspects than optimizing the smoother are also important. In particular, the discretization and solution on the coarse meshes needs further investigation.

Multigrid algorithms for higher order space-time discontinuous Galerkin discretizations of the Euler equations

In Chapter 5 we used Fourier analysis to optimize h -multigrid algorithms for second and third order accurate space-time DG discretizations of the 2D advection-diffusion equation. In this chapter we extend this analysis to three-level p - and hp -multigrid methods. We, however, consider only the EXI Runge-Kutta smoother [79]. To compare the efficiency of the h -, p - and hp -multigrid methods, we present some preliminary results of a more complex test case in which we solve the compressible Euler equations.

6.1 Space-time DG for the Euler equations

In this section we summarize the space-time DG method for the Euler equations. We start by introducing the space-time formulation after which the space-time DG weak formulation is discussed. After introducing a polynomial approximation in the weak formulation a system of algebraic equations is obtained. Finally, we summarize the pseudo-time integration and the explicit EXI Runge-Kutta method. See also [79].

6.1.1 Space-time formulation

In the space-time DG method, the space and time variables are treated together. A point at time $t = x_0$ with position vector $\bar{x} = (x_1, x_2)$ has Cartesian coordinates $x = (x_0, \bar{x})$ in the open domain $\mathcal{E} \subset \mathbb{R}^3$. At time t , the flow domain $\Omega(t)$ is defined as $\Omega(t) := \{\bar{x} \in \mathbb{R}^2 : (t, \bar{x}) \in \mathcal{E}\}$. By taking t_0 and T as the initial and final time of the evolution of the space-time flow domain, the space-time domain boundary $\partial\mathcal{E}$ consists of the hyper-surfaces $\Omega(t_0) \equiv \{x \in \partial\mathcal{E} : x_0 = t_0\}$, $\Omega(T) \equiv \{x \in \partial\mathcal{E} : x_0 = T\}$ and $\mathcal{Q} \equiv \{x \in \partial\mathcal{E} : t_0 < x_0 < T\}$. Let the initial flow field be denoted by U_0 and the boundary data by U^b . Then, using index notation with the summation convention on repeated indices and the comma notation to denote partial differentiation, the Euler equations can be written as:

$$\begin{cases} U_{,0} + F_k(U)_{,k} = 0, & \text{on } \mathcal{E}, \\ U = U_0, & \text{on } \Omega(t_0), \\ U = U^b, & \text{on } \mathcal{Q}, \end{cases} \quad (6.1)$$

with $U \in \mathbb{R}^4$ the vector of conservative variables and $F \in \mathbb{R}^{4 \times 2}$ the inviscid flux given by:

$$U = \begin{bmatrix} \rho \\ \rho u_j \\ \rho E \end{bmatrix}, \quad F_k = \begin{bmatrix} \rho u_k \\ \rho u_j u_k + p \delta_{jk} \\ u_k(\rho E + p) \end{bmatrix}, \quad (6.2)$$

where $j, k = 1, 2$. The Euler equations are completed with the equation of state for a calorically perfect gas: $p = (\gamma - 1)\rho(E - \frac{1}{2}u_i u_i)$, with γ the ratio of specific heats.

6.1.2 Weak formulation and discretization

The approximation $\Omega_h(t_n)$ of the flow domain $\Omega(t_n)$ is divided into N_n non-overlapping spatial elements $K_j(t_n)$. The space-time elements \mathcal{K}_j^n are constructed by connecting K_j^n with K_j^{n+1} using linear interpolation in time. The flow domain \mathcal{E} , limited to the time interval (t_n, t_{n+1}) , defines a space-time slab, \mathcal{E}_h^n . The tessellation \mathcal{T}_h^n of \mathcal{E}_h^n consists of all space-time elements \mathcal{K}_j^n .

Within a space-time slab we distinguish faces connecting space-time slabs, $K_j(t_n^+)$ and $K_j(t_{n+1}^-)$, internal faces \mathcal{S}_I^n and boundary faces \mathcal{S}_B^n . The outward space-time normal vector on a space-time element \mathcal{K}_j^n is denoted by $n = (n_t, \bar{n})$, with n_t the temporal and \bar{n} the spacial part of the outward normal vector. On an internal face $\mathcal{S} \in \mathcal{S}_I$, the traces from the left and right element are denoted by $(\cdot)^L$ and $(\cdot)^R$, respectively. The average operator is defined as $\{\!\{ \cdot \}\!\} = \frac{1}{2}((\cdot)^L + (\cdot)^R)$ and the jump operator as $[\![\cdot]\!]_k = (\cdot)^L \bar{n}_k^L + (\cdot)^R \bar{n}_k^R$.

We consider approximations of $U(x)$ and test functions $V(x)$ in the finite element space W_h defined as:

$$W_h^p = \{W \in (L_2(\mathcal{E}_h))^4 : W|_{\mathcal{K}} \circ G_{\mathcal{K}}^n \in P^p(\hat{\mathcal{K}}), \forall \mathcal{K} \in \mathcal{T}_h\},$$

where $L_2(\mathcal{E}_h)$ is the space of square integrable functions on \mathcal{E}_h , $P^p(\hat{\mathcal{K}})$ the space of polynomials of degree at most p on reference element $\hat{\mathcal{K}} = (-1, 1)^3$ and $G_{\mathcal{K}}^n$ the mapping from the master element to the space-time element \mathcal{K}_j^n .

Considering only fixed grids, the space-time DG weak formulation of the Euler equations is: *Find a $U \in W_h$ such that for all $V \in W_h$:*

$$\begin{aligned} & - \sum_{\mathcal{K} \in \mathcal{T}_h} \int_{\mathcal{K}} (V_{i,0} U_i + V_{i,k} F_{ik}) d\mathcal{K} + \sum_{S \in \mathcal{S}_h^n} \int_S (V_i^L - V_i^R) H_i dS \\ & + \sum_{\mathcal{K} \in \mathcal{T}_h} \left(\int_{K(t_{n+1}^-)} V_i^L U_i^L dK - \int_{K(t_n^+)} V_i^L U_i^R dK \right) + \sum_{S \in \mathcal{S}_B^n} \int_S V_i^L H_i dS = 0, \end{aligned} \quad (6.3)$$

in which $H = H(U^L, U^R, \bar{n}^L)$ is the inviscid numerical flux. Following [79] we use the HLLC approximate Riemann solver given by:

$$\begin{aligned} H_i = \frac{1}{2} (F_{ik}^L \bar{n}_k^L + F_{ik}^R \bar{n}_k^R) & + \frac{1}{2} (|S^M| - |S^L|) U_{i*}^L + |S^L| U_i^L \\ & + \frac{1}{2} (|S^R| - |S^M|) U_{i*}^R - |S^R| U_i^R, \end{aligned}$$

with $F^{L,R} = F(U^{L,R})$. The intermediate states U_*^L and U_*^R are given by:

$$U_*^{L,R} = \frac{S^{L,R} - q^{L,R}}{S^{L,R} - S^M} + \frac{1}{S^{L,R} - S^M} \begin{bmatrix} 0 \\ (p_* - p^{L,R}) \bar{n}_k^L \\ p_* S^M - p^{L,R} q^{L,R} \end{bmatrix},$$

with $q = \bar{n}_k^L u_k$ the normal velocity and p_* the intermediate pressure:

$$p_* = \rho^L (S^L - q^L) (S^M - q^L) + p^L = \rho^R (S^R - q^R) (S^M - q^R) + p^R.$$

The middle wave speed S^M is defined as:

$$S^M = \frac{\rho^R q^R (S^R - q^R) - p^R - \rho^L q^L (S^L - q^L) + p^L}{\rho^R (S^R - q^R) - \rho^L (S^L - q^L)},$$

and the left and right wave speeds as $S^L = \min\{q^L - a^L, q^R - a^R\}$ and $S^R = \max\{q^L + a^L, q^R + a^R\}$, respectively, with $q = \sqrt{\gamma p / \rho}$ the speed of sound. For more information on the derivation of the space-time DG weak formulation, we refer to [79].

The space-time DG system of algebraic equations is obtained by approximating the test and trial functions in (6.3) in each element $\mathcal{K} \in \mathcal{T}_h^n$ with the polynomial expansions:

$$U(t, \bar{x})|_{\mathcal{K}} = \hat{U}_i \psi_i(t, \bar{x}), \quad V(t, \bar{x})|_{\mathcal{K}} = \hat{V}_j \psi_j(t, \bar{x}), \quad (6.4)$$

where ψ_i are the basis functions and \hat{U}_i and \hat{V}_j the expansion coefficients, respectively, for $i, j = 0, 1, 2, \dots, N$, and N depends on the polynomial degree of the basis functions in W_h . On the reference element $\hat{\mathcal{K}}$, the basis functions $\hat{\psi}_i(\xi) \in P^p(\hat{\mathcal{K}})$ are a combination of one-dimensional Legendre polynomials. They are related to $\psi_i(x)$ in element \mathcal{K} through the mapping $G_{\mathcal{K}}^n$: $\psi_i = \hat{\psi}_i \circ (G_{\mathcal{K}}^n)^{-1}$. The system of algebraic equations is written compactly as $\mathcal{L}_h(\hat{U}^n; \hat{U}^{n-1}) = 0$.

6.1.3 Pseudo-time integration and Runge-Kutta methods

Following [79], to solve the system of coupled equations for the expansion coefficients $\mathcal{L}_h(\hat{U}^n; \hat{U}^{n-1}) = 0$, a pseudo-time derivative is added to the system:

$$|K_j| \frac{\partial \hat{U}^*(K_j)}{\partial \tau} = -\frac{1}{\Delta t} \mathcal{L}_h(\hat{U}^*; \hat{U}^{n-1}), \quad (6.5)$$

which is integrated to steady-state in pseudo-time. At steady state, $\hat{U}^n = \hat{U}^*$. To solve (6.5), we investigate the use of h -, p - and hp -multigrid methods with the explicit 5-stage EXI Runge-Kutta smoother [79]. For notational purposes, we set $\mathcal{L}_h(\hat{V}^*; \hat{U}^{n-1}) = \mathcal{L}_h(\hat{V}^*)$. Initialize $\hat{V}^0 = \hat{U}^{n-1}$. The EXI smoother is given by:

$$\hat{V}^j = \left(\hat{V}^0 + \alpha_j \Delta \tau (\hat{V}^{j-1} - \mathcal{L}_h(\hat{V}^{j-1}) / (|K_j| \Delta t)) \right) / (I + \alpha_j \Delta \tau I), \quad j = 1, \dots, 5. \quad (6.6)$$

with $\hat{U}^n = \hat{V}^5$. The Runge-Kutta coefficients at stage j are denoted by α_j and defined as: $\alpha_1 = 0.0791451$, $\alpha_2 = 0.163551$, $\alpha_3 = 0.283663$, $\alpha_4 = 0.5$ and $\alpha_5 = 1.0$. The matrix I represents the identity matrix. For notational purposes we also write (6.6) as:

$$\hat{U}^n = S_{h|p} \hat{U}^{n-1}. \quad (6.7)$$

Here h refers to the grid discretization while p refers to the polynomial order.

6.2 Multigrid methods

In this section we present the h -, p - and hp -multigrid methods for solving the non-linear system of algebraic equations resulting from the space-time DG discretization of the Euler equations.

6.2.1 h -Multigrid algorithm

In an h -multigrid technique we introduce a finite sequence N_c of increasingly coarser meshes G_{nh} , $n \in \{1, \dots, N_c\}$ to generate coarser approximations of the original problem. For $1 \leq n < m \leq N_c$, restriction operators $R_{nh}^{mh} : G_{nh} \rightarrow G_{mh}$ for the solution and $\bar{R}_{nh}^{mh} : G_{nh} \rightarrow G_{mh}$ for the residuals, together with prolongation operators $P_{mh}^{nh} : G_{mh} \rightarrow G_{nh}$ are introduced to transfer the data on the different meshes. Following [42], we define the prolongation operator P_{mh}^{nh} for the space-time DG discretization as:

$$\begin{aligned} \hat{U}_{ki}^{nh} &= (\mathcal{M}_{nh}^{-1})_{il} \left(\int_{\mathcal{K}_{nh}} \psi_l^{nh} \psi_j^{mh} d\mathcal{K} \right) \hat{U}_{kj}^{mh} \\ &= (P_{mh}^{nh})_{ij} \hat{U}_{kj}^{mh}, \quad 1 \leq n < m \leq N_c, \end{aligned} \quad (6.8)$$

where \mathcal{M}_{nh} is the mass matrix of the space-time element $\mathcal{K}_{nh} \in \mathcal{T}_{nh}^n$. The mass matrix \mathcal{M}_{nh} is defined as:

$$(\mathcal{M}_{nh})_{li} = \int_{\mathcal{K}_{nh}} \psi_l^{nh} \psi_i^{nh} d\mathcal{K}. \quad (6.9)$$

The restriction operator for the solution is defined as $R_{nh}^{mh} = (P_{mh}^{nh})^{-1}$, while the restriction operator for the residuals is defined as $\bar{R}_{nh}^{mh} = \frac{1}{4}(P_{mh}^{nh})^T$. In (6.8), the embedding of spaces was assumed, i.e. $W_{mh} \subset W_{nh}$ with $m > n$ to ensure that U^{mh} is defined on \mathcal{K}_{nh} .

In the h -multigrid technique, to solve the non-linear system $\mathcal{N}_h(U_h) = f_h$ on G_h , with \mathcal{N}_h the nonlinear operator and f_h a given righthand side, a set of auxiliary problems is solved. At each grid level G_{nh} , $1 \leq n \leq N_c$, we solve $\mathcal{N}_{nh}(U_{nh}) = f_{nh}$ using the EXI Runge-Kutta method. The h -multigrid Full Approximation Scheme (FAS), is defined as follows:

Definition 6.2.1. hMG_n -Algorithm. Start with the initial guess V_{nh}^0 . if $n = N_c$, the coarsest mesh, then solve (e.g. with a Newton method)

$$V_{nh}^1 = hMG_n(V_{nh}^0, f_{nh}, \nu_1, \nu_2) := \mathcal{N}_{nh}^{-1} f_{nh}, \quad (6.10)$$

else the multigrid algorithm hMG_n at the mesh G_{nh} is defined as

1. (Pre-smoothing). Let $U_{nh}^0 = V_{nh}^0$ and define $U_{nh}^{\nu_1}$ by

$$U_{nh}^{l+1} = S_{nh|p} U_{nh}^l, \quad l = 0, 1, \dots, \nu_1 - 1;$$

with $S_{nh|p}$ the smoothing operator at grid level n given in (6.7).

2. (Restricting). Restrict the residual and solution to the coarser mesh G_{mh} with $n < m \leq N_c$,

$$\begin{aligned} Z_{mh}^0 &= R_{nh}^{mh} U_{nh}^{\nu_1}, \\ r_{mh} &= N_{mh}(R_{nh}^{mh} U_{nh}^{\nu_1}) + \bar{R}_{nh}^{mh}(f_{nh} - \mathcal{N}_{nh} U_{nh}^{\nu_1}). \end{aligned}$$

3. (Coarse-grid solving). At grid level m repeat the hMG_m -algorithm γ -times

$$Z_{mh}^l = hMG_m(Z_{mh}^{l-1}, r_{mh}, \nu_1, \nu_2), \quad l = 1, \dots, \gamma.$$

4. (Correcting). Let $Y_{nh}^0 = U_{nh}^{\nu_1} + P_{mh}^{nh}(Z_{mh}^\gamma - R_{nh}^{mh} U_{nh}^{\nu_1})$.

5. (Post-smoothing). Define Y^{ν_2} by

$$Y_{nh}^{l+1} = S_{nh|p} Y_{nh}^l, \quad l = 0, 1, \dots, \nu_2 - 1.$$

Finally, set $V_{nh}^1 = hMG_n(V_{nh}^0, f_{nh}, \nu_1, \nu_2) := Y_{nh}^{\nu_2}$ as the result of the hMG_n -algorithm.

Using different sequences of meshes G_{nh} various multigrid cycles, such as the V, W or F-cycle can be constructed.

6.2.2 p -Multigrid method

In the p -multigrid method, lower order approximations U^q , $q < p$, on the same grid serve as coarse approximations in the solution of the non-linear system $\mathcal{N}(U^p) = f$. Let N_p be the number of p -levels. To transfer the solution and residuals between the lower-order W_h^q and higher-order W_h^p spaces, $1 \leq q < p \leq N_p$, we follow [9]. The solution $U_i^p \in W_h^p$ can be obtained from $U_i^q \in W_h^q$ by solving:

$$\int_{\mathcal{K}_h} V_i^p U_i^p d\mathcal{K} = \int_{\mathcal{K}_h} V_i^p U_i^q d\mathcal{K}, \quad \forall V^p \in W_h^p. \quad (6.11)$$

Replacing the test and trial functions in (6.11) by their polynomial expansions (6.4) gives:

$$\mathcal{M}_{lm}^p \hat{U}_{im}^p = \left(\int_{\mathcal{K}_h} \psi_l^p \psi_m^p d\mathcal{K} \right) \hat{U}_{im}^p = \left(\int_{\mathcal{K}_h} \psi_l^p \psi_n^q d\mathcal{K} \right) \hat{U}_{in}^q =: (\mathcal{M}_q^p)_{ln} \hat{U}_{in}^q.$$

The prolongation operator \mathcal{P}_q^p then is defined as:

$$\hat{U}_{im}^p = (\mathcal{M}^p)_{ml}^{-1} (\mathcal{M}_q^p)_{ln} \hat{U}_{in}^q = (\mathcal{P}_q^p)_{mn} \hat{U}_{in}^q. \quad (6.12)$$

Similarly, the solution restriction operator \mathcal{R}_p^q is defined as:

$$\hat{U}_{in}^q = (\mathcal{M}^q)_{nl}^{-1} (\mathcal{M}_p^q)_{lm} \hat{U}_{im}^p = (\mathcal{R}_p^q)_{nm} \hat{U}_{im}^p, \quad (6.13)$$

while the residual restriction operator is defined as $\bar{\mathcal{R}}_p^q = (\mathcal{P}_q^p)^T$.

In the p -multigrid technique, to solve the non-linear system $\mathcal{N}^p(U^p) = f^p$ for the p^{th} -order solution U^p , with f a given righthand side, a set of auxiliary problems is solved. For each level n , $1 \leq n \leq N_p$, we solve $\mathcal{N}^n(U^n) = f^n$ using the EXI Runge-Kutta method. The p -multigrid algorithm is defined as follows:

Definition 6.2.2. pMG_p -Algorithm. Start with the initial guess V_p^0 . if $p = 1$, the lowest polynomial order, then solve (e.g. with a Newton method)

$$V_p^1 = pMG_p(V_p^0, f^p, \nu_1^p, \nu_2^p) := \mathcal{N}_p^{-1} f^p, \quad (6.14)$$

else the multigrid algorithm pMG , for polynomial order p is defined as

1. (Pre-smoothing). Let $U_p^0 = V_p^0$ and define $U_p^{\nu_1^p}$ by

$$U_p^{l+1} = S_{h|p} U_p^l, \quad l = 0, 1, \dots, \nu_1^p - 1;$$

with $S_{h|p}$ the smoothing operator in (6.7).

2. (Restricting). Restrict the residual and solution to the lower polynomial order q , $1 \leq q < p$,

$$\begin{aligned} Z_q^0 &= \mathcal{R}_p^q U_p^{\nu_1^p}, \\ r_q &= N_q(\mathcal{R}_p^q U_p^{\nu_1^p}) + \bar{\mathcal{R}}_p^q (f_p - N_p U_p^{\nu_1^p}). \end{aligned}$$

3. (Coarse-grid solving). At the polynomial order q repeat the pMG_q -algorithm γ^q -times

$$Z_q^l = pMG_q(Z_q^{l-1}, r_q, \nu_1^p, \nu_2^p), \quad l = 1, \dots, \gamma^q.$$

4. (Correcting). Let $Y_p^0 = U_p^{\nu_1^p} + P_q^p(Z_q^{\gamma^q} - R_q^p U_p^{\nu_1^p})$.

5. (Post-smoothing). Define $Y_p^{\nu_2^p}$ by

$$Y_p^{l+1} = S_{h|p} Y_p^l, \quad l = 0, 1, \dots, \nu_2^p - 1.$$

Finally, set $V_p^1 = pMG_p(V_p^0, f_p, \nu_1^p, \nu_2^p) := Y_p^{\nu_2^p}$ as the result of the pMG_p -algorithm.

6.2.3 hp -Multigrid method

A combined h - and p -multigrid method is obtained if we replace in the pMG algorithm the $p = 1$ step, (6.14), by calling the hMG algorithm to solve the lowest order problem. Note that by calling the hMG algorithm, the system of non-linear algebraic equations to be solved is very small compared to the original problem and it will therefore be relatively cheap to solve with e.g. a Newton method.

6.3 Fourier analysis for p - and hp -multigrid

In this section we use three-level Fourier analysis to investigate the p - and hp -multigrid algorithm. For this we consider space-time DG discretizations of the 2D advection-diffusion equation using quadratic basis functions. The aim of the Fourier analysis is to calculate the spectral radius which gives a prediction of the asymptotic rate of convergence of the multigrid method. For a three-level Fourier analysis of the h -multigrid algorithm and more details on the notation and derivations, we refer to Chapter 5.

6.3.1 p -Multigrid error transformation operator

The derivation of the p -multigrid error transformation operator is similar to that of the h -multigrid error transformation operator given in Chapter 5. Therefore, we only introduce the definition of the p -multigrid error transformation operator.

Definition 6.3.1. The multigrid error transformation operator $M_{h|p}$ of order p is defined recursively for $p = N^p, N^p - 1, \dots, 1$ as:

$$M_{h|p} = S_{h|p}^{\nu_2^p} (I^p - \mathcal{P}_{p-1}^p (I^{p-1} - M_{h|p-1}^{\gamma^p}) L_{h|p-1}^{-1} \mathcal{R}_p^{p-1} L_{h|p}) S_{h|p}^{\nu_1^p}, \quad (6.15)$$

with $S_{h|p}$ the smoother (6.7), ν_1^p, ν_2^p the number of pre- and post-smoothing iterations and γ^p is the cycle index. When $p = 1$, $M_{h|p-1}^{\gamma^p} = 0$.

6.3.2 Three-level p -multigrid Fourier analysis

To couple the three-level p -multigrid Fourier analysis to the three-level h -multigrid Fourier analysis, we consider the 16-dimensional subspace given by (5.40) in Chapter 5, otherwise it would not be necessary to consider these modes simultaneously. Since the p -multigrid algorithm only acts on one grid, the p -multigrid Fourier analysis is much easier compared to the three-level h -multigrid Fourier analysis.

We introduce the following matrices:

$$\begin{aligned}\hat{L}_{2g}^p(\theta_\beta) &= \text{diag}(\hat{L}_{h|p}(\theta_\beta^{00}), \hat{L}_{h|p}(\theta_\beta^{11}), \hat{L}_{h|p}(\theta_\beta^{10}), \hat{L}_{h|p}(\theta_\beta^{01})) \in \mathbb{C}^{4m_p \times 4m_p} \\ \hat{S}_{2g}^p(\theta_\beta) &= \text{diag}(\hat{S}_{h|p}(\theta_\beta^{00}), \hat{S}_{h|p}(\theta_\beta^{11}), \hat{S}_{h|p}(\theta_\beta^{10}), \hat{S}_{h|p}(\theta_\beta^{01})) \in \mathbb{C}^{4m_p \times 4m_p} \\ \hat{\mathcal{R}}_{2g}^{p|p-1}(\theta_\beta) &= \text{diag}(\hat{\mathcal{R}}_h^{p|p-1}(\theta_\beta^{00}), \hat{\mathcal{R}}_h^{p|p-1}(\theta_\beta^{11}), \hat{\mathcal{R}}_h^{p|p-1}(\theta_\beta^{10}), \hat{\mathcal{R}}_h^{p|p-1}(\theta_\beta^{01})) \in \mathbb{C}^{4m_p \times 4m_p} \\ \hat{\mathcal{P}}_{2g}^{p-1|p}(\theta_\beta) &= \text{diag}(\hat{\mathcal{P}}_h^{p-1|p}(\theta_\beta^{00}), \hat{\mathcal{P}}_h^{p-1|p}(\theta_\beta^{11}), \hat{\mathcal{P}}_h^{p-1|p}(\theta_\beta^{10}), \hat{\mathcal{P}}_h^{p-1|p}(\theta_\beta^{01})) \in \mathbb{C}^{4m_p \times 4m_p},\end{aligned}$$

with $\hat{L}_{h|p}$, $\hat{S}_{h|p}$, $\hat{\mathcal{R}}_h^{p|p-1}$, $\hat{\mathcal{P}}_h^{p-1|p}$, respectively the Fourier symbols of the discretization, smoother, restriction and prolongation operators, see Chapter 5. We also introduce the matrices

$$\begin{aligned}\hat{L}_{3g}^p(\theta) &= \text{diag}(\hat{L}_{2g}^p(\theta_{00}), \hat{L}_{2g}^p(\theta_{\frac{1}{2}\frac{1}{2}}), \hat{L}_{2g}^p(\theta_{\frac{1}{2}0}), \hat{L}_{2g}^p(\theta_{0\frac{1}{2}})) \in \mathbb{C}^{16m_p \times 16m_p} \\ \hat{S}_{3g}^p(\theta) &= \text{diag}(\hat{S}_{2g}^p(\theta_{00}), \hat{S}_{2g}^p(\theta_{\frac{1}{2}\frac{1}{2}}), \hat{S}_{2g}^p(\theta_{\frac{1}{2}0}), \hat{S}_{2g}^p(\theta_{0\frac{1}{2}})) \in \mathbb{C}^{16m_p \times 16m_p} \\ \hat{\mathcal{R}}_{3g}^{p|p-1}(\theta) &= \text{diag}(\hat{\mathcal{R}}_{2g}^{p|p-1}(\theta_{00}), \hat{\mathcal{R}}_{2g}^{p|p-1}(\theta_{\frac{1}{2}\frac{1}{2}}), \hat{\mathcal{R}}_{2g}^{p|p-1}(\theta_{\frac{1}{2}0}), \hat{\mathcal{R}}_{2g}^{p|p-1}(\theta_{0\frac{1}{2}})) \in \mathbb{C}^{16m_p \times 16m_p} \\ \hat{\mathcal{P}}_{3g}^{p-1|p}(\theta) &= \text{diag}(\hat{\mathcal{P}}_{2g}^{p-1|p}(\theta_{00}), \hat{\mathcal{P}}_{2g}^{p-1|p}(\theta_{\frac{1}{2}\frac{1}{2}}), \hat{\mathcal{P}}_{2g}^{p-1|p}(\theta_{\frac{1}{2}0}), \hat{\mathcal{P}}_{2g}^{p-1|p}(\theta_{0\frac{1}{2}})) \in \mathbb{C}^{16m_p \times 16m_p}.\end{aligned}$$

The Fourier symbol of the three-level p -multigrid error transformation operator $\hat{M}_{h|p}(\theta) \in \mathbb{C}^{16m_p \times 16m_p}$ then is equal to

$$\begin{aligned}\hat{M}_{h|p}(\theta) &= (\hat{S}_{3g}^p(\theta))^{\nu^2} \hat{K}^{3g}(\theta) (\hat{S}_{3g}^p(\theta))^{\nu^2}, \\ \hat{K}^{3g}(\theta) &= \hat{I}_{3g}^p - \hat{\mathcal{P}}_{3g}^{p-1|p}(\theta) (\hat{I}_{3g}^{p-1} - (U_{3g}^{p-1})^{\gamma^p}) (\hat{I}_{3g}^{p-1})^{-1}(\theta) \hat{\mathcal{R}}_{3g}^{p|p-1}(\theta) \hat{L}_{3g}^p(\theta),\end{aligned}\tag{6.16}$$

for $\theta \in \Theta_{4h} \setminus \Psi_{3g}$ and $\hat{I}_{3g}^p \in \mathbb{R}^{16m_p \times 16m_p}$ identity matrix. Furthermore, U_{3g} is given by:

$$(U_{3g}^q)^{\gamma^p} = \begin{cases} (\hat{M}_{h|p-1}(\theta))^{\gamma^p}, & \text{if } q = p-1, \\ 0, & \text{if } q = p-2 \text{ and one level } h\text{-multigrid}, \\ \hat{M}^{3g}(\theta), & \text{if } q = p-2 \text{ and three level } h\text{-multigrid}, \end{cases}$$

where $\hat{M}^{3g}(\theta)$ is (5.48), given in Chapter 5, applied to a $(p-2)^{th}$ -order space-time DG discretization.

6.3.3 Results from Fourier analysis

Using the Fourier analysis of the previous sections, we will determine the spectral radius of the different multigrid techniques to solve the system of algebraic

equations resulting from the space-time DG discretization of the 2D advection-diffusion equation. We compare the spectral radii obtained from single grid, h -, p - and hp -multigrid.

We consider quadratic basis functions and the following parameters: $CFL = 100$, $Re_x = Re_y = 100$, $AR = 1$, $\gamma^{low} = \pi/4$, $CFL^\tau = VN^\tau = 1$, $\nu_1^* = \nu_2^* = 5$ and $\nu_C^* = 10$, where $*$ = h, p . Furthermore, $\gamma^* = 1$. The results of the Fourier analysis are given in Table 6.1. The “work” column represents the amount of work needed to converge the solution six orders with respect to its original error. This column is more representative when comparing the efficiency of the different multigrid techniques. This is because the spectral radius only predicts, in the asymptotic regime, the amount of multigrid cycles needed for convergence. However, the amount of work per multigrid cycle differs per technique:

$$\text{work per cycle} = \begin{cases} g^p b^p, & \text{SG,} \\ (g^p b^p + g^{p-1} b^{p-1})(\nu_1^p + \nu_2^p) + g^{p-2} b^{p-2} \nu_C^p, & \text{pMG,} \\ g^p b^p ((c^h + c^{h-1})(\nu_1^h + \nu_2^h) + c^{h-2} \nu_C^h), & \text{hMG,} \\ c^h (g^p b^p + g^{p-1} b^{p-1})(\nu_1^p + \nu_2^p) + \\ \quad g^{p-2} b^{p-2} ((c^h + c^{h-1})(\nu_1^h + \nu_2^h) + c^{h-2} \nu_C^h), & \text{hpMG,} \end{cases}$$

where g^p is the number of Gauss quadrature points in an element in the space-time discretization depending on the polynomial order p , b^p the number of basis functions depending on the polynomial order p and c^h a weighting for the number of cells depending on the grid-level h . We have: $g^p = 9$, $g^{p-1} = 4$, $g^{p-2} = 1$, $b^p = 6$, $b^{p-1} = 4$, $b^{p-2} = 1$, $c^h = 1$, $c^{h-1} = 1/4$ and $c^{h-2} = 1/16$. For example, for the single grid (SG) computation, which has a spectral radius $\rho = 0.99418$, the number of cycles to converge six orders, N_c , follows from $\rho^{N_c} = 10^{-6} \rightarrow N_c = 1389$. The amount of work then equals $9 \cdot 6 \cdot 1389 = 75006$. From Table 6.1 we can obtain the following conclusions. All multigrid methods perform better than a single-grid method and the h -multigrid method has the best theoretical performance. For the non-linear test case in Section 6.4.2, we will, however, need for the h -multigrid method a CFL^τ number which is twice as small compared to the CFL^τ number for the single-grid, p - and hp -multigrid methods. From Table 6.1 it follows that p - and hp -multigrid methods then will perform better than the h -multigrid method, but the h -multigrid method still performs better than single-grid methods. Furthermore, we remark that in the Fourier analysis it does not make much difference if we solve the coarse grid problem in the h -multigrid, or the lowest order stage of the p -multigrid algorithm, exactly or via ν_C iterations of the EXI Runge-Kutta smoother. This also explains why there is hardly any difference in spectral radius between the p - and hp -multigrid techniques.

In Figure 6.1 we plot the spectrum of \mathcal{L}_h , the space-time discretization operator of the 2D advection-diffusion equation, and the stability domain of the EXI

p -levels	h -levels	CFL^τ	ρ	work
1	1	1.0	0.99010	75006
3	1	1.0	0.74318	33370
3	3	1.0	0.73347	32091
1	3	1.0	0.66350	24098
1	3	0.5	0.80364	45360

Table 6.1: *Spectral radii and work units of the different multigrid strategies.*

smoother. The eigenvalue spectra of the single grid, h -, p - and hp - multigrid techniques, are depicted in Figure 6.2.

Besides applying an h -multigrid technique as solver for the lowest-order problem in the p -multigrid algorithm, we could also have applied a p -multigrid technique as solver for the coarse grid problem in the h -multigrid algorithm. We, however, chose not to follow this path due to findings of Yavneh [88] in which it is shown that an enriched discretization on the coarsest grid in h -multigrid is preferred.

6.4 Multigrid algorithms applied to the Euler equations

In this section we apply the h -, p - and hp -multigrid algorithms to space-time DG discretizations of the Euler equations. As a test case we will consider 2D steady subsonic flow around a NACA0012 airfoil. The angle of attack is taken to be $\alpha = 2^\circ$ and we take a far-field Mach number of $Ma = 0.5$. We first discuss the choice of basis functions for our DG discretization after which we discuss the numerical results of a simulation.

6.4.1 Choosing the basis functions

For the NACA0012 test case we noticed that the choice of basis functions in the space-time DG discretization \mathcal{L}_h was important for the stability of the scheme. We therefore calculate the spectrum of the complete discretization \mathcal{L}_h , including boundary conditions and high aspect ratio elements, for a number of choices of basis functions. Since we consider a steady-state problem, the space-time DG discretization is taken to be only first-order accurate in time but third-order accurate in space. The spectrum of \mathcal{L}_h is determined numerically by calculating

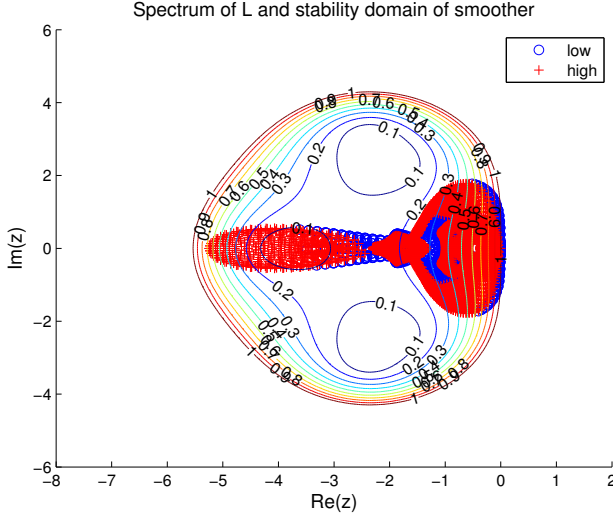


Figure 6.1: Spectrum of \mathcal{L}_h , the space-time discretization operator of the 2D advection-diffusion equation, and the stability domain of the EXI Runge-Kutta method.

the eigenvalues of the following approximation of the Jacobian of \mathcal{L}_h :

$$\frac{\partial \mathcal{L}_h}{\partial U} = \begin{bmatrix} \vdots & & & & & \\ \dots & \frac{\partial \mathcal{L}_{exp.coef=i}^{cell=k}}{\partial U_{exp.coef=i}^{cell=k}} & \frac{\partial \mathcal{L}_{exp.coef=i+1}^{cell=k}}{\partial U_{exp.coef=i+1}^{cell=k}} & \dots & \frac{\partial \mathcal{L}_{exp.coef=i}^{cell=k+1}}{\partial U_{exp.coef=i}^{cell=k+1}} & \frac{\partial \mathcal{L}_{exp.coef=i+1}^{cell=k+1}}{\partial U_{exp.coef=i+1}^{cell=k+1}} & \dots \\ \dots & \frac{\partial \mathcal{L}_{exp.coef=i+1}^{cell=k}}{\partial U_{exp.coef=i+1}^{cell=k}} & \frac{\partial \mathcal{L}_{exp.coef=i+1}^{cell=k}}{\partial U_{exp.coef=i+1}^{cell=k}} & \dots & \frac{\partial \mathcal{L}_{exp.coef=i+1}^{cell=k+1}}{\partial U_{exp.coef=i+1}^{cell=k+1}} & \frac{\partial \mathcal{L}_{exp.coef=i+1}^{cell=k+1}}{\partial U_{exp.coef=i+1}^{cell=k+1}} & \dots \\ \vdots & & & & & & \\ \dots & \frac{\partial \mathcal{L}_{exp.coef=i}^{cell=k+1}}{\partial U_{exp.coef=i}^{cell=k+1}} & \frac{\partial \mathcal{L}_{exp.coef=i+1}^{cell=k+1}}{\partial U_{exp.coef=i+1}^{cell=k+1}} & \dots & \frac{\partial \mathcal{L}_{exp.coef=i}^{cell=k+1}}{\partial U_{exp.coef=i}^{cell=k+1}} & \frac{\partial \mathcal{L}_{exp.coef=i+1}^{cell=k+1}}{\partial U_{exp.coef=i+1}^{cell=k+1}} & \dots \\ \dots & \frac{\partial \mathcal{L}_{exp.coef=i+1}^{cell=k+1}}{\partial U_{exp.coef=i+1}^{cell=k+1}} & \frac{\partial \mathcal{L}_{exp.coef=i+1}^{cell=k+1}}{\partial U_{exp.coef=i+1}^{cell=k+1}} & \dots & \frac{\partial \mathcal{L}_{exp.coef=i+1}^{cell=k+1}}{\partial U_{exp.coef=i+1}^{cell=k+1}} & \frac{\partial \mathcal{L}_{exp.coef=i+1}^{cell=k+1}}{\partial U_{exp.coef=i+1}^{cell=k+1}} & \dots \\ \vdots & & & & & & \end{bmatrix}$$

with

$$\frac{\partial \mathcal{L}_i^k}{\partial U_i^j} = \frac{\mathcal{L}_i^k(U_i^j) - \mathcal{L}_i^k(U_i^j + \varepsilon)}{\varepsilon}, \quad \varepsilon = \sqrt{10^{-13}}.$$

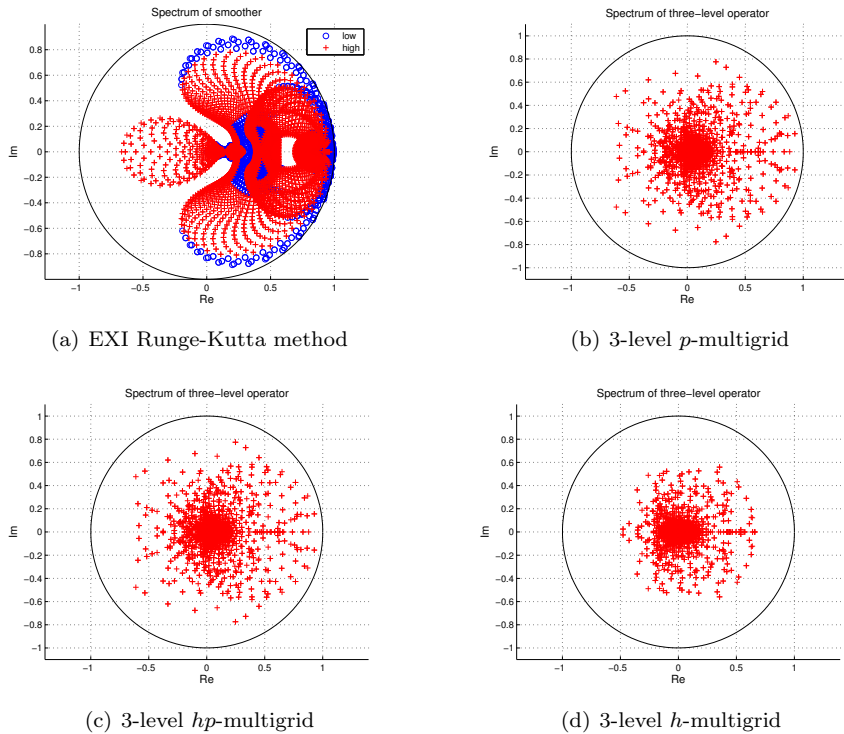


Figure 6.2: Eigenvalue spectra of the different multigrid techniques for the space-time DG discretization of the 2D advection-diffusion test case using quadratic basis functions.

	$\Delta\tau$	Normalized $\Delta\tau$
Option 1	0.0475	1
Option 2	0.0275	0.579
Scaled $\bar{\mathcal{M}} = I$	0.0089	0.187
Orthonormal	0.0179	0.377

Table 6.2: Pseudo-time step $\Delta\tau$ such that the spectrum of \mathcal{L}_h with quadratic basis functions lies within the stability domain of the EXI smoother. We normalized $\Delta\tau$ with respect to the $\Delta\tau$ of the option 1 Legendre basis functions.

Let \mathcal{M} be the mass-matrix (see (6.9)). On a uniform grid we can also write $\mathcal{M} = \Delta t |K| \bar{\mathcal{M}}$ in which $\bar{\mathcal{M}}$ is dimensionless. With this definition we consider the following sets of quadratic basis functions on the reference element $\hat{\mathcal{K}}$:

1. Legendre basis functions, option 1: $\psi_1 = 1, \psi_2 = \xi_1, \psi_3 = \xi_2, \psi_4 = \xi_1\xi_2, \psi_5 = \xi_1^2 - \frac{1}{3}, \psi_6 = \xi_2^2 - \frac{1}{3}$ (using Gram-Schmidt orthonormalization see [32]).
2. Legendre basis functions, option 2: $\psi_1 = 1, \psi_2 = \xi_1, \psi_3 = \xi_2, \psi_4 = \xi_1\xi_2, \psi_5 = \frac{1}{2}(3\xi_1^2 - 1), \psi_6 = \frac{1}{2}(3\xi_2^2 - 1)$ (see [1], pg. 798).
3. Scaled quadratic basis functions such that $\bar{\mathcal{M}} = I$: $\psi_1 = 1, \psi_2 = \sqrt{3}\xi_1, \psi_3 = \sqrt{3}\xi_2, \psi_4 = 3\xi_1\xi_2, \psi_5 = \frac{1}{2}\sqrt{5}(3\xi_1^2 - 1), \psi_6 = \frac{1}{2}\sqrt{5}(3\xi_2^2 - 1)$.
4. Orthonormal quadratic basis functions on the reference element $\hat{\mathcal{K}}$: $\psi_1 = \sqrt{0.5}, \psi_2 = \sqrt{1.5}\xi_1, \psi_3 = \sqrt{1.5}\xi_2, \psi_4 = 1.5\xi_1\xi_2, \psi_5 = \frac{1}{2}\sqrt{2.5}(3\xi_1^2 - 1), \psi_6 = \frac{1}{2}\sqrt{2.5}(3\xi_2^2 - 1)$.

For all sets of basis functions, we plot the spectrum of \mathcal{L}_h and the stability domain of the EXI Runge-Kutta smoother in Figure 6.3. The calculations are performed on a grid with 28×4 elements. For each set the pseudo-time step $\Delta\tau$ was chosen such that the spectrum of \mathcal{L}_h lies within the stability domain of the smoother (see Table 6.2). We see that, depending on the choice of basis functions, the largest pseudo-time step is five times larger than the smallest pseudo-time step. The best choice is to take the Legendre basis functions, option 1. We remark that the Fourier analysis of Section 6.3.3 and the numerical simulations in Section 6.4.2 are done with the Legendre basis functions, option 1.

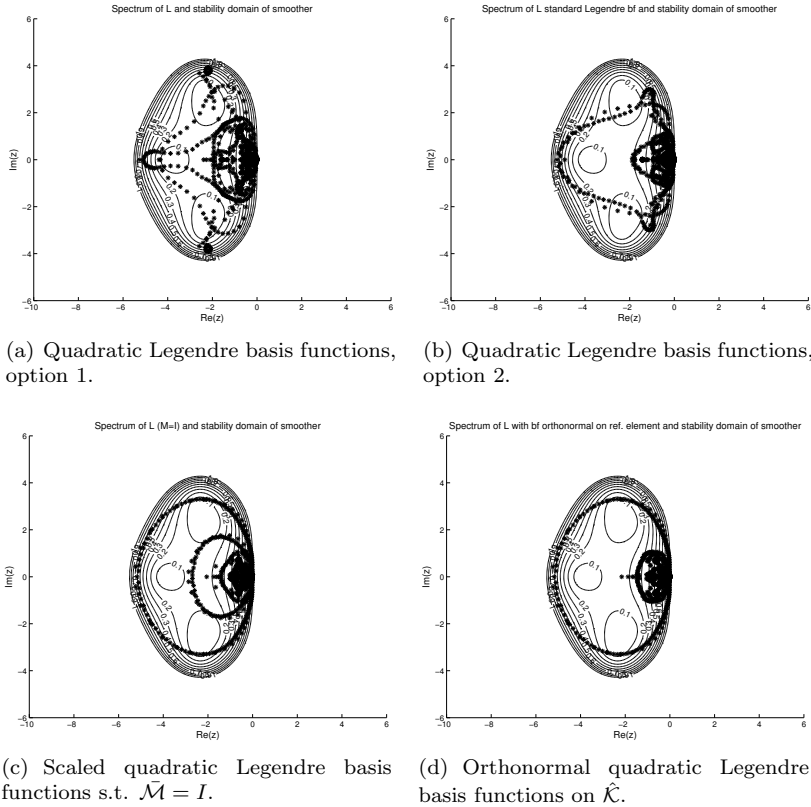


Figure 6.3: *Spectrum of \mathcal{L}_h on the grid around a NACA0012 airfoil.*

6.4.2 Simulations

We consider 2D steady subsonic flow around a NACA0012 airfoil with an angle of attack of $\alpha = 2^\circ$. We take a far-field Mach number of $Ma = 0.5$. As mentioned in Section 6.4.1, since we consider a steady-state flow problem, we use a space-time DG discretization which is only first-order accurate in time but third-order accurate in space. The grid around the airfoil has 448×64 elements and is shown in Figure 6.4. We consider five solution strategies to solve the algebraic system resulting from the space-time DG discretization of the Euler equations:

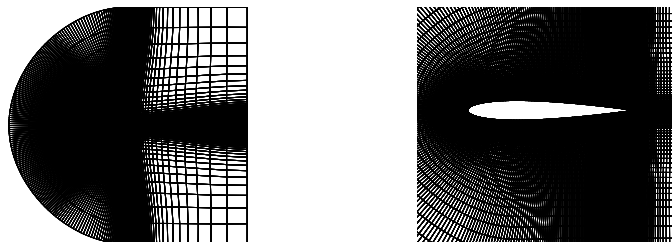
1. Single grid.
2. Three-level h -multigrid with $\nu_1^h = \nu_2^h = \nu_3^h = \nu_4^h = 10$. The coarse grid problem is solved exactly using a matrix-free Newton method.
3. Three-level h -multigrid with $\nu_1^h = \nu_2^h = \nu_3^h = \nu_4^h = 5$. The coarse grid problem is solved approximately taking $\nu_C^h = 20$.
4. Three-level p -multigrid with $\nu_1^p = \nu_2^p = \nu_3^p = \nu_4^p = 5$. The lowest order problem is solved approximately taking $\nu_C^p = 20$.
5. Three-level hp -multigrid with $\nu_1^* = \nu_2^* = \nu_3^* = \nu_4^* = 5$, $* = h, p$. The coarse grid problem is solved approximately taking $\nu_C^h = 20$.

For the single-grid, p - and hp -multigrid computations we used a pseudo-time CFL number of $CFL\tau = 1.6$, while for h -multigrid $CFL\tau = 0.8$. Larger pseudo-time CFL numbers for h -multigrid resulted in unstable calculations. The Mach contours are given in Figure 6.5 while the convergence history plot is given in Figure 6.6.

From Table 6.1 we expect hp -multigrid to be the best solution method and single-grid to be the worst. We see, however, that p -multigrid is slightly better than hp -multigrid. We also see that h -multigrid performs worst.

The hp -multigrid initially shows a significant improvement in the reduction of the residual compared to the single-grid computation. In particular, in the early cycles hp -multigrid is very efficient. In the asymptotic regime, however, single-grid and hp -multigrid have approximately the same convergence rate. The reason for this behavior is unclear yet. For the p -multigrid method, initial convergence is significantly faster than for the single-grid computations, but in the asymptotic regime a comparable convergence history with the single grid computations is obtained.

Furthermore, initially the h -multigrid performs rather well, however, after the high-frequency modes have been smoothed, h -multigrid efficiency quickly deteriorates. A reason for this could be that at some point the coarse-grid



(a) NACA0012 airfoil mesh: 448×64 elements.

(b) Zoom of the mesh around the airfoil.

Figure 6.4: *The computational mesh used for inviscid subsonic flow around a NACA0012 airfoil.*

problem of the h -multigrid algorithm is not solved well with respect to the characteristic components, as was demonstrated in [88]. This problem may be overcome by employing an improved coarse grid operator. This has not been studied in the present work, but is currently under investigation. We also see that there is hardly any difference in solving the coarse grid problem exactly by the Newton method or approximating the solution by performing ν_C smoothing steps. This is contrary to the results obtained in Chapter 5 where we saw a large improvement when solving the coarse grid problem exactly.

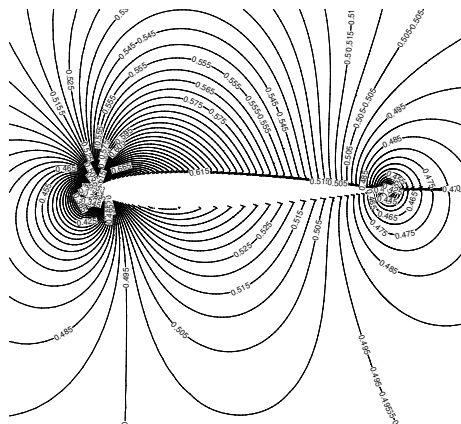


Figure 6.5: Mach contours of inviscid flow around a NACA0012 airfoil.

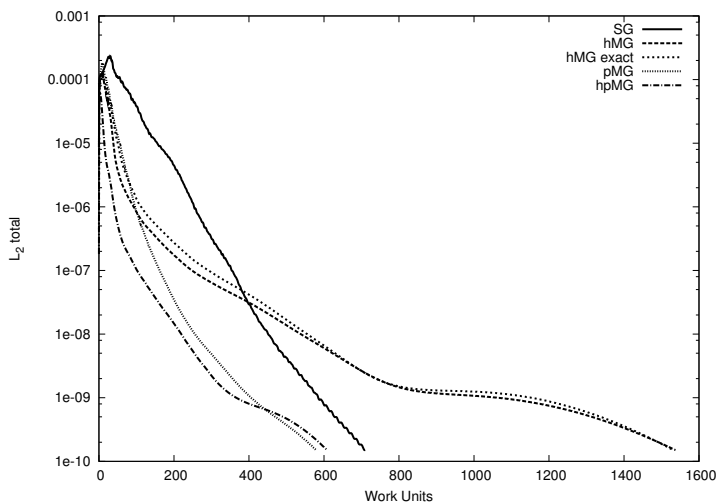


Figure 6.6: Convergence history of single-grid, h -, p - and hp -multigrid techniques (inviscid flow around a NACA0012 airfoil).

An alternative derivation of the discontinuous Galerkin finite element weak formulation

In Chapter 2 we introduced a discontinuous Galerkin (DG) finite element weak formulation for PDE's with nonconservative products. For this, we used the theory of Dal Maso, LeFloch and Murat [54] in which a Borel measure was assigned to the nonconservative product so that even if the solution was discontinuous, the nonconservative product was well defined. In this chapter we follow a similar approach to provide an alternative derivation for DG weak formulations for hyperbolic PDE's and a new DG weak formulation for elliptic/parabolic PDE's.

DG methods for hyperbolic PDE's have been thoroughly investigated by Cockburn and co-workers. Detailed surveys can be found in [16, 17]. In the derivation of the DG weak formulation, commonly two different routes are followed. To explain these different approaches, we consider a space DG discretization of a 1D scalar hyperbolic PDE, $\partial_t u + \partial_x f(u) = 0$, on a domain Ω . Denote by \mathcal{T}_h the tessellation of Ω and let K_k be an element of the tessellation \mathcal{T}_h . The computational domain Ω_h is the union of all elements of the tessellation. Furthermore, let the exact solution u be approximated by $u_h \in W_h$ with W_h a finite element space, and let $v_h \in W_h$ be a test function.

Following the first approach, the weak formulation of the PDE is not considered on the whole domain, but rather on an element. The PDE is multiplied by a test function $v_h \in W_h$, integrated over an element $K \in \mathcal{T}_h$ and u is replaced by its approximation $u_h \in W_h$:

$$0 = \int_K (v_h \partial_t u_h + v_h \partial_x f(u_h)) dK, \quad \forall v_h \in W_h.$$

After integration by parts we obtain:

$$0 = \int_K (v_h \partial_t u_h - f(u_h) \partial_x v_h) dK + \int_{\partial K} v_h^L f(u_h^L) \bar{n}^L d(\partial K), \quad \forall v_h \in W_h, \quad (7.1)$$

where \bar{n}^L is the outward unit normal to the boundary ∂K and u^L is the trace at the element boundary ∂K defined as $u^L = \lim_{\varepsilon \downarrow 0} u(x - \varepsilon \bar{n}^L)$. The problem with this formulation, is that there is no coupling of element K with its neighboring elements. To deal with this, the argument is used that the trace at an element boundary is double valued because there is no direct relation with the data in a neighboring element. This implies that the flux on an element boundary does not have a precise meaning. A numerical flux $\hat{f}(u^L, u^R, \bar{n}^L)$ is therefore introduced by viewing the discontinuity on an element boundary in the normal direction \bar{n}^L as a local Riemann problem. Here u^R is the trace on the element boundary on a neighboring element of K . Equation (7.1) then becomes:

$$0 = \int_K (v_h \partial_t u_h - f(u_h) \partial_x v_h) dK + \int_{\partial K} v_h^L \hat{f}(u_h^L, u_h^R, \bar{n}^L) d(\partial K), \quad \forall v_h \in W_h, \quad (7.2)$$

which is solved for all elements of the tessellation \mathcal{T}_h .

The second approach avoids the problem of coupling the weak formulation on an element to the weak formulation on neighboring elements by immediately considering the whole computational domain. The PDE is multiplied by a test function $v_h \in W_h$, integrated over the element $K \in \mathcal{T}_h$ and u is replaced by its approximation $u_h \in W_h$. After summation over all elements in the tessellation \mathcal{T}_h we obtain:

$$0 = \sum_{K \in \mathcal{T}_h} \int_K (v_h \partial_t u_h + v_h \partial_x f(u_h)) dK, \quad \forall v_h \in W_h. \quad (7.3)$$

Integration by parts then results in

$$0 = \sum_{K \in \mathcal{T}_h} \int_K (v_h \partial_t u_h - f(u_h) \partial_x v_h) dK + \sum_{K \in \mathcal{T}_h} \int_{\partial K} v_h^L f(u_h^L) \bar{n}^L d(\partial K), \quad \forall v_h \in W_h. \quad (7.4)$$

The coupling between the elements is achieved by noting that the summation over the element boundaries can be written as a sum over element faces \mathcal{S} :

$$\sum_{K \in \mathcal{T}_h} \int_{\partial K} v_h^L f(u_h^L) \bar{n}^L d(\partial K) = \sum_{\mathcal{S}} \int_{\mathcal{S}} \left(\frac{1}{2} (v_h^L \bar{n}^L + v_h^R \bar{n}^R) (f(u_h^L) + f(u_h^R)) + \frac{1}{2} (v_h^L + v_h^R) (\bar{n}^L f(u_h^L) + \bar{n}^R f(u_h^R)) \right) d\mathcal{S}. \quad (7.5)$$

Using now the argument that the formulation must be conservative, the following condition is imposed:

$$\int_{\mathcal{S}} v_h \bar{n}^L f(u_h^L) d\mathcal{S} = - \int_{\mathcal{S}} v_h \bar{n}^R f(u_h^R) d\mathcal{S}, \quad \forall v_h \in W_h, \quad (7.6)$$

so that the underlined term in (7.5) vanishes and (7.4) becomes:

$$0 = \sum_{K \in \mathcal{T}_h} \int_K (v_h \partial_t u_h - f(u_h) \partial_x v_h) dK + \sum_S \int_S (v_h^L - v_h^R) \frac{1}{2} (f(u_h^L) + f(u_h^R)) \bar{n}^L dS, \quad (7.7)$$

$\forall v_h \in W_h$, where we used $\bar{n}^L = -\bar{n}^R$. It is well known that the DG weak formulation is unstable in this form. A numerical flux $\hat{f}(u_h^L, u_h^R, \bar{n}^L)$ is therefore introduced to replace the central flux $\frac{1}{2}(f(u_h^L) + f(u_h^R))\bar{n}^L$ in the face integrals. Note that in the derivation of (7.2) the numerical flux plays two rolls, namely coupling of the elements and stabilization. Here, the numerical flux is only introduced for stabilization purposes. We obtain:

$$0 = \sum_{K \in \mathcal{T}_h} \int_K (v_h \partial_t u_h - f(u_h) \partial_x v_h) dK + \sum_S \int_S (v_h^L - v_h^R) \hat{f}(u_h^L, u_h^R, \bar{n}^L) dS, \quad \forall v_h \in W_h. \quad (7.8)$$

The relation (7.6) is necessary for (7.8) and (7.2) to be the same. However, the equality in (7.6) is mathematically not necessarily true so that the argument that the formulation must be conservative is questionable (Take e.g. $f = au$, $a = \text{constant}$: in DG u is discontinuous and so also f).

In the above we see that both routes have their problems. In the first route it is the coupling of the elements while in the second route it is the conservation assumption of the formulation to neglect part of the face integrals. In order to obtain a mathematically more consistent derivation we introduce a new approach based on Borel measures.

For elliptic and (incompletely) parabolic equations also various formulations are presented in the literature. These include the method of Brezzi et al. [13], the LDG method [18], the IP method [21], Bassi et al. [11] and NIPG [65]. An analysis of these formulations is given in [6]. Articles dealing with the space-time DG formulation of elliptic/(incompletely) parabolic equations are e.g. [41, 69]. Using the theory for nonconservative products, it is possible to recover the method of Brezzi et al. [13], but also to define a new weak formulation. We discuss both possibilities in this chapter.

The outline of this chapter is as follows. In Section 7.1 we summarize two theorems on Borel measures. In Section 7.2 we discuss the space-time DG method presented in [41, 79] while in Section 7.3 we present a derivation of the space-time DG weak formulation based on Borel measures. We remark that the space-time DG method is a generalization of the space DG method. We therefore omit the alternative derivation for the space DG method since it is analogous. In Section 7.4 we perform numerical simulations to investigate the accuracy of the different formulations.

7.1 Borel measures in DG

We will derive discontinuous Galerkin finite element formulations for systems of partial differential equations of the form:

$$\mathcal{F}_{ij}(U)_{,j} = U_{i,0} + F_{ik}(U)_{,k} - (A_{ikrs}(U)U_{r,s})_{,k} = 0, \quad \bar{x} \in \mathbb{R}^q, \quad t > 0, \quad (7.9)$$

with $U \in \mathbb{R}^m$, $F \in \mathbb{R}^{m \times q}$, $\mathcal{F} \in \mathbb{R}^{m \times (q+1)}$ and $A \in \mathbb{R}^{m \times q \times m \times q}$; we use the comma notation to denote partial differentiation and the summation convention on repeated indices. Here $(\cdot)_{,0}$ denotes partial differentiation with respect to time and $(\cdot)_{,k}$, ($k = 1, \dots, q$) partial differentiation with respect to the spatial coordinates. In a space-time context, the space and time variables are, however, not explicitly distinguished. A point at time $t = x_0$ with position $\bar{x} = (x_1, x_2, \dots, x_q)$ has Cartesian coordinates $x = (x_0, \bar{x}) \in \mathbb{R}^{q+1}$.

We introduce now some theorems. Let $\Omega \subseteq \mathbb{R}^{q+1}$ with $\Omega = \Omega_u \cup S_u$, where Ω_u is the set of points of approximate continuity and S_u the set of points of approximate jump. If the product $A_{ikrs}(U)U_{r,s}$ in (7.9) cannot be written as $G_{i,k}$, then $A_{ikrs}(U)U_{r,s}$ is a nonconservative product. Following [54, 64], assume a given family of Lipschitz continuous paths $\phi : [0, 1] \times \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ that satisfy, for some $K > 0$ and for all $U^L, U^R \in \mathbb{R}^m$, with U^L and U^R the left and right traces at a discontinuity, and $\tau \in [0, 1]$, the properties:

$$(H1) \quad \phi_r(0; U^L, U^R) = U_r^L, \quad \phi_r(1; U^L, U^R) = U_r^R,$$

$$(H2) \quad \phi_r(\tau; U^L, U^L) = U_r^L,$$

$$(H3) \quad \left| \frac{\partial \phi_r}{\partial \tau}(\tau; U^L, U^R) \right| \leq K |U_r^L - U_r^R|, \text{ a.e. in } [0, 1],$$

$$(H4) \quad \phi_r(\tau; U^L, U^R) = \phi_r(1 - \tau; U^R, U^L).$$

The theory of Dal Maso, LeFloch and Murat [54] (DLM theory) then states (see Chapter 2):

Theorem 7.1.1. *Let $U : \Omega \rightarrow \mathbb{R}^m$ be a bounded function of bounded variation defined on an open subset Ω of \mathbb{R}^{q+1} and $A : \mathbb{R}^m \rightarrow \mathbb{R}^m$ a locally bounded Borel function. Then there exists a unique family of real-valued bounded Borel measures μ_{ik}^A on Ω , $i = 1, 2, \dots, m$, $k = 1, \dots, q$ such that*

1. if B is a Borel subset of Ω_u , then

$$\mu_{ik}^A(B) = \int_B A_{ikrs}(U)U_{r,s} \, d\lambda, \tag{7.10}$$

where λ is the Borel measure;

2. if B is a Borel subset of S_u , then

$$\mu_{ik}^A(B) = \int_{B \cap S_u} \int_0^1 A_{ikrs}(\phi(\tau; U^L, U^R)) \frac{\partial \phi_r}{\partial \tau}(\tau; U^L, U^R) \, d\tau \, n_s^L \, dH^q, \tag{7.11}$$

with U^L and U^R the left and right traces at the discontinuity, where H^q denotes the q -dimensional Hausdorff measure and n^L the outward normal with respect to the left state.

We remark that in the particular case that there is a G such that $G_i(U)_{,k} = A_{ikrs}(U)U_{r,s}$, then (7.11) becomes:

$$\mu_{ik}^A(B) = \int_{B \cap S_u} (G_i(U^R) - G_i(U^L))n_k^L dH^q. \quad (7.12)$$

In the special case of conservative fluxes \mathcal{F}_{ik} , we introduce the following theorem:

Theorem 7.1.2. *Let $U : \Omega \rightarrow \mathbb{R}^m$ be a bounded function of bounded variation defined on an open subset Ω of \mathbb{R}^{q+1} and $\mathcal{F} : \mathbb{R}^m \rightarrow \mathbb{R}^m$ a locally bounded Borel function. Then there exists a unique family of real-valued bounded Borel measures μ_i^f on Ω , $i = 1, 2, \dots, m$ such that*

1. *if B is a Borel subset of Ω_u , then*

$$\mu_i^f(B) = \int_B \mathcal{F}_{ik}(U)_{,k} d\lambda, \quad (7.13)$$

where λ is a Borel measure;

2. *if B is a Borel subset of S_u , then*

$$\mu_i^f(B) = \int_{B \cap S_u} (\mathcal{F}_{ik}(U^R) - \mathcal{F}_{ik}(U^L))n_k^L dH^q, \quad (7.14)$$

with U^L and U^R the left and right traces at the discontinuity, where H^q denotes the q -dimensional Hausdorff measure and n^L the outward normal with respect to the left state.

7.2 Space-time DGFEM weak formulation

In this section we introduce space-time elements, function spaces, trace operators and we discuss the space-time DGFEM weak formulation.

7.2.1 Elements

In the space-time DGFEM method, the space and time variables are not distinguished. A point at time $t = x_0$ with position vector $\bar{x} = (x_1, x_2, \dots, x_q)$ has Cartesian coordinates (x_0, \bar{x}) in the open domain $\mathcal{E} \subset \mathbb{R}^{q+1}$. At time t , the flow domain $\Omega(t)$ is defined as $\Omega(t) := \{\bar{x} \in \mathbb{R}^q : (t, \bar{x}) \in \mathcal{E}\}$. The space-time domain boundary $\partial\mathcal{E}$ consists of the hyper-surfaces $\Omega_0 := \{x \in \partial\mathcal{E} : x_0 = t_0\}$, $\Omega_T := \{x \in \partial\mathcal{E} : x_0 = T\}$ and $\mathcal{Q} := \{x \in \partial\mathcal{E} : t_0 < x_0 < T\}$. The space-time normal vector at $\partial\mathcal{E}$ is defined as $n := (n_0, \bar{n})^T$, with n_0 the temporal component and \bar{n} the spatial component.

The time interval $[t_0, T]$ is partitioned using the time levels $t_0 < t_1 < \dots < T$, where the n^{th} time interval is defined as $I_n = (t^n, t^{n+1})$ with length $\Delta t_n = t_{n+1} - t_n$. The space-time domain \mathcal{E} is then divided into N_t space-time slabs $\mathcal{E}^n := \mathcal{E} \cap I_n$, with boundaries $\Omega(t_n)$, $\Omega(t_{n+1})$ and $\mathcal{Q}^n = \partial\mathcal{E}^n / (\Omega(t_n) \cup \Omega(t_{n+1}))$.

In each space-time slab we introduce a finite element tessellation \mathcal{T}_h^n with space-time elements \mathcal{K} . The flow domain $\Omega(t_n)$ is approximated by $\Omega_h(t_n)$, where $\Omega_h(t_n) \rightarrow \Omega(t_n)$ as $h \rightarrow 0$, with h the radius of the smallest sphere completely containing the largest space-time element. The domain $\Omega_h(t_n)$ is divided into N_n non-overlapping spatial elements $K(t_n)$. At t_{n+1} the spatial elements $K(t_{n+1})$ are obtained by mapping the vertices of the elements $K(t_n)$ to their position at t_{n+1} . The space-time elements $\mathcal{K} \in \mathcal{T}_h^n$ are then obtained by interpolation in time between the elements $K(t_n)$ and $K(t_{n+1})$, such that $\mathcal{E}_h^n = \cup_{\mathcal{K} \in \mathcal{T}_h^n} \mathcal{K} \rightarrow \mathcal{E}^n$ as $h \rightarrow 0$. Furthermore, each space-time element \mathcal{K}_j^n can be mapped to the master element $\hat{\mathcal{K}} \subset \mathbb{R}^{q+1}$ by an isoparametric mapping $G_{\mathcal{K}}$.

The element boundary $\partial\mathcal{K}_j^n$, which is the union of open faces of \mathcal{K}_j^n , consists of three parts: $K_j(t_n^+) = \lim_{\epsilon \downarrow 0} K_j(t_n + \epsilon)$, $K_j(t_{n+1}^-) = \lim_{\epsilon \downarrow 0} K_j(t_{n+1} - \epsilon)$ and $\mathcal{Q}_j^n = \partial\mathcal{K}_j^n / (K_j(t_n^+) \cup K_j(t_{n+1}^-))$. Define the grid velocity $v \in \mathbb{R}^q$ as $v = \Delta \bar{x} / \Delta t$. The outward space-time normal vector at an element boundary point on $\partial\mathcal{K}_j^n$ is given by:

$$n = \begin{cases} (1, \bar{0}) & \text{at } K_j(t_{n+1}^-), \\ (-1, \bar{0}) & \text{at } K_j(t_n^+), \\ (-v_k \bar{n}_k, \bar{n}) & \text{at } \mathcal{Q}_j^n, \end{cases} \quad (7.15)$$

where $\bar{0} \in \mathbb{R}^q$. Note that since the space-time normal vector n has length one, the space component \bar{n} of the space-time normal has a length $|\bar{n}| = 1 / \sqrt{1 + v \cdot v}$. It can be convenient to split the element boundaries into separate faces. We therefore define interior and boundary faces. An interior face is shared by two neighboring elements \mathcal{K}_i^n and \mathcal{K}_j^n , such that $\mathcal{S}_{ij}^n = \mathcal{Q}_i^n \cap \mathcal{Q}_j^n$, and a boundary face is defined as $\mathcal{S}_{B_j}^n = \partial\mathcal{E}^n \cap \mathcal{Q}_j^n$. The set of interior faces in time slab I^n is denoted by \mathcal{S}_I^n and the set of all boundary faces by \mathcal{S}_B^n . Furthermore, we define the time-faces as $\mathcal{S}_{K_j}^{n+1} = \mathcal{K}_j^n \cap \mathcal{K}_j^{n+1}$ where $\mathcal{S}_{K_j}^0 = \mathcal{K}_j^0 \cap \Omega_0$ and $\mathcal{S}_{K_j}^T = \mathcal{K}_j^T \cap \Omega_T$. In space-time slab \mathcal{E}_h^n , the total set of faces is denoted by $\mathcal{S}^n = \mathcal{S}_I^n \cup \mathcal{S}_B^n \cup \mathcal{S}_K^{n+1} \cup \mathcal{S}_K^n$. The total set of space-faces is denoted by $\mathcal{S}_{IB}^n = \mathcal{S}_I^n \cup \mathcal{S}_B^n$.

7.2.2 Function spaces and trace operators

We consider approximations of $U(x, t)$ and functions $V(x, t)$ in the finite element space W_h , which is defined as:

$$W_h = \{W \in (L^2(\mathcal{E}_h))^m : W|_{\mathcal{K}} \circ G_{\mathcal{K}} \in (P^p(\hat{\mathcal{K}}))^m, \forall \mathcal{K} \in \mathcal{T}_h\},$$

where $L^2(\mathcal{E}_h)$ is the space of square integrable functions on \mathcal{E}_h and $P^p(\hat{\mathcal{K}})$ denotes the space of polynomials of degree at most p on the reference element $\hat{\mathcal{K}}$. Here m denotes the dimension of U . We will also use the space

$$V_h = \{V \in (L^2(\mathcal{E}_h))^{m \times q} : V|_{\mathcal{K}} \circ G_{\mathcal{K}} \in (P^p(\hat{\mathcal{K}}))^{m \times q}, \forall \mathcal{K} \in \mathcal{T}_h\}.$$

The trace of a function $f \in W_h$ at the element boundary $\partial\mathcal{K}$ is defined as $f^L = \lim_{\epsilon \downarrow 0} f(x - \epsilon n^L)$ with n^L the unit outward space-time normal at $\partial\mathcal{K}$. The outward normal vector of the adjacent element \mathcal{K}^R is denoted as n^R with $n^R = -n^L$. When only the space components of the outward normal vector are considered we will use the notation \bar{n}^L . A function $f \in W_h$ has a double valued trace at element boundaries $\partial\mathcal{K}$. The traces of a function f at an internal face $\mathcal{S} = \bar{\mathcal{K}}^L \cap \bar{\mathcal{K}}^R$ are denoted by f^L and f^R . The jump of f at an internal face in the direction k of a Cartesian coordinate system is defined as $\llbracket f \rrbracket_k = f^L n_k^L + f^R n_k^R$. The average of f at an internal face is defined as $\{\!\{ f \}\!\} = \frac{1}{2}(f^L + f^R)$. The jump operator satisfies the following product rule on all interior faces for $\forall g \in W_h$ and $\forall f \in W_h$, which can be proven by direct verification:

$$\llbracket g_i f_{ik} \rrbracket_k = \{\!\{ g_i \}\!\} \llbracket f_{ik} \rrbracket_k + \llbracket g_i \rrbracket_k \{\!\{ f_{ik} \}\!\}. \quad (7.16)$$

Element boundary integrals are related to face integrals by:

$$\sum_n \left(\sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\partial\mathcal{K}} g_i^L f_{ik}^L n_k^L d\partial\mathcal{K} \right) = \sum_n \left(\sum_{\mathcal{S} \in \mathcal{S}_I^n} \int_{\mathcal{S}} \llbracket g_i f_{ij} \rrbracket_j d\mathcal{S} + \sum_{\mathcal{S} \in \mathcal{S}_B^n} \int_{\mathcal{S}} g_i^L f_{ij}^L \bar{n}_j^L d\mathcal{S} + \sum_{\mathcal{S} \in \mathcal{S}_R^n} \int_{\mathcal{S}} \llbracket g_i f_{i0} \rrbracket_0 d\mathcal{S} \right), \quad (7.17)$$

with $k = 0, \dots, q$, $j = 1, \dots, q$, $i = 1, \dots, m$. From now on, for notational purposes, we use the same notation for interior and boundary faces by introducing ghost values U^R at the domain boundary.

7.2.3 Weak formulation

We now introduce the space-time DG weak formulation of (7.9). The commonly used derivation of the space-time DG weak formulation is well known and therefore not repeated here. Instead, we refer to [8, 13, 41, 79]. The space-time DG

weak formulation of (7.9) is given by: Find a $U \in W_h$ such that for all $V \in W_h$:

$$\begin{aligned}
 0 = & \sum_n \left(- \sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} (V_{i,0} U_i + V_{i,k} (F_{ik}(U) - A_{ikrs}(U) U_{r,s})) d\mathcal{K} \right. \\
 & + \sum_{S \in \mathcal{S}_{IB}^n} \int_S (V_i^L - V_i^R) \{ \{ F_{ik}(U) - v_k U_i \} \bar{n}_k^L d\mathcal{S} \\
 & - \sum_{S \in \mathcal{S}_{IB}^n} \int_S \llbracket V_i \rrbracket_k \{ \{ A_{ikrs}(U) (U_{r,s} + \mathcal{R}_{rs}) \} \} d\mathcal{S} \\
 & \left. - \sum_{S \in \mathcal{S}_{IB}^n} \int_S \{ \{ V_{i,k} A_{ikrs}(U) \} \} \llbracket U_r \rrbracket_s d\mathcal{S} + \sum_{S \in \mathcal{S}_K^n} \int_S \llbracket V_i \rrbracket_0 \{ \{ U_i \} \} dK \right), \tag{7.18}
 \end{aligned}$$

with the lifting operator \mathcal{R}_{rs} defined as: Find an $\mathcal{R} \in V_h$ such that for all $\varphi \in V_h$:

$$\sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} \varphi_{rs} \mathcal{R}_{rs} d\mathcal{K} = \sum_{S \in \mathcal{S}_{IB}^n} \int_S \llbracket \varphi_{rs} (\hat{U}_r - U_r) \rrbracket_s d\mathcal{S}. \tag{7.19}$$

Note that in the derivation of the lifting operator by Brezzi et al. [13], the numerical flux \hat{U} was introduced.

7.3 Derivation based on Borel measures

Using the Borel measures from Section 7.1, we introduce the definitions of generalized DG flux derivatives and generalized DG diffusion terms. We then derive a new weak formulation for (7.9). Two criteria are imposed on the weak formulation for the hyperbolic part of (7.9), namely, the scheme must be locally element wise conservative, and second, we require causality in time. We will also show, that by introducing a definition of the generalized DG derivative, it is possible to obtain exactly weak formulation (7.18). As mentioned already in the introduction, the space DG derivation follows analogously to that of the space-time DG derivation and is therefore omitted.

7.3.1 Generalized DG flux derivatives

Using Theorem 7.1.2, generalized DG flux derivatives, denoted by $\tilde{D}_k^{DG} \mathcal{F}_{ik}(U)$, $k = 0, \dots, q$, $i = 1, \dots, m$, satisfy:

$$\begin{aligned} \int_{\mathcal{E}_h} V_i \tilde{D}_k^{DG} \mathcal{F}_{ik}(U) d\mathcal{E}_h &= \sum_n \left(\sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} V_i \mathcal{F}_{ik}(U)_{,k} d\mathcal{K} - \sum_{\mathcal{S} \in \mathcal{S}^n} \int_{\mathcal{S}} \hat{V}_i [\mathcal{F}_{ik}(U)]_k d\mathcal{S} \right) \\ &= \int_{\mathcal{E}_h} V_i (\mathcal{F}_{ik}(U)_{,k} + \mathcal{R}_i^f) d\mathcal{E}_h, \quad \forall V \in W_h, \end{aligned} \quad (7.20)$$

where $\mathcal{R}^f \in \mathbb{R}^m$ is the global lifting operator defined as: *Find an $\mathcal{R}^f \in W_h$, such that for all $V \in W_h$:*

$$\int_{\mathcal{E}_h} V_i \mathcal{R}_i^f d\mathcal{E}_h = - \sum_n \left(\sum_{\mathcal{S} \in \mathcal{S}^n} \int_{\mathcal{S}} \hat{V}_i [\mathcal{F}_{ik}(U)]_k d\mathcal{S} \right). \quad (7.21)$$

It follows that $\tilde{D}_k^{DG} \mathcal{F}_{ik}(U) = \mathcal{F}_{ik}(U)_{,k} + \mathcal{R}_i^f$ a.e. on \mathcal{E}_h . Note that in (7.20) and (7.21) we still need to define \hat{V} . This will be discussed in Section 7.3.3.

7.3.2 Generalized DG diffusion terms

Using Theorem 7.1.1, the generalized DG diffusion term, denoted by $(A_{ikrs}(U)U_{r,s})^{DG}$, with $k, s = 1, \dots, q$, $i, r = 1, \dots, m$, satisfies:

$$\begin{aligned} \int_{\mathcal{E}_h} \varphi_{ik} (A_{ikrs}(U)U_{r,s})^{DG} d\mathcal{E}_h &= \\ & \sum_n \left(\sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} \varphi_{ik} d\mu_{ik}^A(\mathcal{K}_j) + \sum_{\mathcal{S} \in \mathcal{S}_{IB}^n} \int_{\mathcal{S}} \hat{\varphi}_{ik} d\mu_{ik}^A(\mathcal{S}) \right) \\ &= \sum_n \left(\sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} \varphi_{ik} A_{ikrs}(U)U_{r,s} d\mathcal{K} \right. \\ & \quad \left. + \sum_{\mathcal{S} \in \mathcal{S}_{IB}^n} \int_{\mathcal{S}} \hat{\varphi}_{ik} \left(\int_0^1 A_{ikrs}(\phi(\tau; U^L, U^R)) \frac{\partial \phi_r}{\partial \tau}(\tau; U^L, U^R) d\tau \bar{n}_s^L \right) d\mathcal{S} \right) \\ &= \int_{\mathcal{E}_h} \varphi_{ik} (A_{ikrs}(U)U_{r,s} + \mathcal{R}_{ik}^A) d\mathcal{E}_h, \quad \forall \varphi \in V_h, \end{aligned} \quad (7.22)$$

where $\mathcal{R}^A \in \mathbb{R}^{m \times q}$ is the global lifting operator defined as: Find an $\mathcal{R}^A \in V_h$, such that for all $\varphi \in V_h$:

$$\int_{\mathcal{E}_h} \varphi_{ik} \mathcal{R}_{ik}^A d\mathcal{E}_h = \sum_n \left(\sum_{S \in \mathcal{S}_{IB}^n} \int_S \hat{\varphi}_{ik} \left(\int_0^1 A_{ikrs}(\phi(\tau; U^L, U^R)) \frac{\partial \phi_r}{\partial \tau}(\tau; U^L, U^R) d\tau \bar{n}_s^L \right) dS \right). \quad (7.23)$$

It follows that $(A_{ikrs}(U)U_{r,s})^{DG} = A_{ikrs}(U)U_{r,s} + \mathcal{R}_{ik}^A$ a.e. on \mathcal{E}_h . In (7.22) and (7.23) we still require the definition of $\hat{\varphi}_{ik}$. This will be discussed in Section 7.3.3.

7.3.3 Space-time discontinuous Galerkin weak formulation

We now derive the space-time discontinuous Galerkin finite element weak formulation. In a discontinuous Galerkin finite element framework we do not determine the weak formulation for (7.9), but for:

$$\tilde{D}_k^{DG} \tilde{\mathcal{F}}_{ik}(U) = 0, \quad x \in \mathbb{R}^{q+1}, \quad x_0 > 0, \quad k = 0, 1, \dots, q, \quad (7.24)$$

with $\tilde{\mathcal{F}} \in \mathbb{R}^{m \times (q+1)}$ given by:

$$\tilde{\mathcal{F}}_{ik}(U) = \begin{cases} U_i, & \text{if } k = 0, \\ F_{ik}(U) - (A_{ikrs}(U)U_{r,s})^{DG}, & \text{if } k = 1, \dots, q. \end{cases} \quad (7.25)$$

We omit the tilde notation from now on.

A new weak formulation based on generalized DG flux derivatives and diffusion terms

Multiplying (7.24) by a test function $V \in W_h$, integrating over the domain \mathcal{E}_h and using the definition of the generalized DG flux derivatives (7.20), we obtain:

$$0 = \sum_n \left(\sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} V_i \mathcal{F}_{ik}(U)_{,k} d\mathcal{K} - \sum_{S \in \mathcal{S}^n} \int_S \hat{V}_i \llbracket \mathcal{F}_{ik}(U) \rrbracket_k d\bar{S} \right), \quad k = 0, \dots, q. \quad (7.26)$$

Using (7.25), we can rewrite (7.26) as:

$$\begin{aligned}
0 &= \sum_n \left(\sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} V_i(U_{i,0} + F_{ik}(U)_{,k} - (A_{ikrs}(U)U_{r,s})_{,k}^{DG}) d\mathcal{K} \right. \\
&\quad - \sum_{S \in \mathcal{S}_K^n} \int_S \hat{V}_i[[U_i]]_0 d\mathcal{S} - \sum_{S \in \mathcal{S}_{IB}^n} \int_S \hat{V}_i[[U_i]]_0 d\mathcal{S} \\
&\quad \left. - \sum_{S \in \mathcal{S}_{IB}^n} \int_S \hat{V}_i[[F_{ik}(U) - (A_{ikrs}(U)U_{r,s})^{DG}]_k] d\mathcal{S} \right) \\
&= \sum_n \left(\sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} V_i(U_{i,0} + F_{ik}(U)_{,k} - (A_{ikrs}(U)U_{r,s})_{,k}^{DG}) d\mathcal{K} \right. \\
&\quad - \sum_{S \in \mathcal{S}_{IB}^n} \int_S \hat{V}_i[[F_{ik}(U) - v_k U_i - (A_{ikrs}(U)U_{r,s})^{DG}]_k] d\mathcal{S} \\
&\quad \left. - \sum_{S \in \mathcal{S}_K^n} \int_S \hat{V}_i[[U_i]]_0 d\mathcal{S} \right), \quad k = 1, \dots, q,
\end{aligned} \tag{7.27}$$

where we used the definition of the space-time normal (7.15). We are left with choosing the numerical flux for the test function V in (7.27) and the test function for the definition of $(A_{ikrs}(U)U_{r,s})^{DG}$ (see (7.22) and (7.23)). These test functions belong to different spaces ($V \in W_h$ and $\varphi \in V_h$) and will also be defined differently. We start with the test function $V \in W_h$. The choice for the numerical flux of the test function $V \in W_h$ follows from requiring that on an element the formulation must be conservative.

Theorem 7.3.1. *If the numerical flux \hat{V} for the test function $V \in W_h$ in (7.27) is defined as $\hat{V} = \{\{V\}\}$, then the DG formulation (7.27) will be locally conservative on an element.*

Proof. Integrating by parts the volume integral in (7.27) we obtain:

$$\begin{aligned}
0 &= \sum_n \left(- \sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} (V_{i,0} U_i + V_{i,k} (F_{ik}(U) - (A_{ikrs}(U)U_{r,s})^{DG})) d\mathcal{K} \right. \\
&\quad + \sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\partial\mathcal{K}} V_i^L (U_i^L n_0^L + (F_{ik}^L - (A_{ikrs} U_{r,s})^{DG}) \bar{n}_k^L) d(\partial\mathcal{K}) \\
&\quad - \sum_{S \in \mathcal{S}_{IB}^n} \int_S \hat{V}_i[[F_{ik}(U) - v_k U_i - (A_{ikrs}(U)U_{r,s})^{DG}]_k] d\mathcal{S} \\
&\quad \left. - \sum_{S \in \mathcal{S}_K^n} \int_S \hat{V}_i[[U_i]]_0 d\mathcal{S} \right),
\end{aligned} \tag{7.28}$$

where $F_{ik}^{L,R} = F_{ik}(U^{L,R})$. Using relations (7.16) and (7.17) and the definition of the normal vector n (7.15), the element boundary integral in (7.28) becomes:

$$\begin{aligned}
 & \sum_n \left(\sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\partial\mathcal{K}} V_i^L (U_i^L n_0^L + (F_{ik}^L - (A_{ikrs} U_{r,s})^{DG}) \bar{n}_k^L) d(\partial\mathcal{K}) \right) \\
 &= \sum_n \left(\sum_{\mathcal{S} \in \mathcal{S}_{IB}^n} \int_{\mathcal{S}} \llbracket V_i (F_{ik}(U) - (A_{ikrs}(U) U_{r,s})^{DG}) \rrbracket_k d\mathcal{S} \right. \\
 & \quad \left. + \sum_{\mathcal{S} \in \mathcal{S}_{IB}^n} \int_{\mathcal{S}} \llbracket V_i U_i \rrbracket_0 d\mathcal{S} + \sum_{\mathcal{S} \in \mathcal{S}_K^n} \int_{\mathcal{S}} \llbracket V_i U_i \rrbracket_0 d\mathcal{S} \right) \\
 &= \sum_n \left(\sum_{\mathcal{S} \in \mathcal{S}_{IB}^n} \int_{\mathcal{S}} \llbracket V_i (F_{ik}(U) - v_k U_i - (A_{ikrs}(U) U_{r,s})^{DG}) \rrbracket_k d\mathcal{S} + \sum_{\mathcal{S} \in \mathcal{S}_K^n} \int_{\mathcal{S}} \llbracket V_i U_i \rrbracket_0 d\mathcal{S} \right) \\
 &= \sum_n \left(\sum_{\mathcal{S} \in \mathcal{S}_{IB}^n} \int_{\mathcal{S}} (\llbracket V_i \rrbracket_k \{ F_{ik}(U) - v_k U_i - (A_{ikrs}(U) U_{r,s})^{DG} \} \right. \\
 & \quad \left. + \llbracket V_i \rrbracket \llbracket F_{ik}(U) - v_k U_i - (A_{ikrs}(U) U_{r,s})^{DG} \rrbracket_k) d\mathcal{S} \right. \\
 & \quad \left. + \sum_{\mathcal{S} \in \mathcal{S}_K^n} \int_{\mathcal{S}} \llbracket V_i \rrbracket_0 \{ U_i \} + \llbracket V_i \rrbracket \llbracket U_i \rrbracket_0 d\mathcal{S} \right).
 \end{aligned} \tag{7.29}$$

Combining (7.28) and (7.29) we obtain:

$$\begin{aligned}
 0 &= \sum_n \left(- \sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} (V_{i,0} U_i + V_{i,k} (F_{ik}(U) - (A_{ikrs}(U) U_{r,s})^{DG})) d\mathcal{K} \right. \\
 & \quad \left. + \sum_{\mathcal{S} \in \mathcal{S}_{IB}^n} \int_{\mathcal{S}} (V_i^L - V_i^R) \{ F_{ik}(U) - v_k U_i \} \bar{n}_k^L d\mathcal{S} \right. \\
 & \quad \left. - \sum_{\mathcal{S} \in \mathcal{S}_{IB}^n} \int_{\mathcal{S}} \llbracket V_i \rrbracket_k \{ (A_{ikrs}(U) U_{r,s})^{DG} \} d\mathcal{S} + \sum_{\mathcal{S} \in \mathcal{S}_K^n} \int_{\mathcal{S}} \llbracket V_i \rrbracket_0 \{ U_i \} d\mathcal{S} \right. \\
 & \quad \left. + \sum_{\mathcal{S} \in \mathcal{S}_{IB}^n} \int_{\mathcal{S}} (\llbracket V_i \rrbracket - \hat{V}_i) \llbracket F_{ik}(U) - v_k U_i - A_{ikrs}(U) U_{r,s}^{DG} \rrbracket_k d\mathcal{S} \right. \\
 & \quad \left. + \sum_{\mathcal{S} \in \mathcal{S}_K^n} \int_{\mathcal{S}} (\llbracket V_i \rrbracket - \hat{V}_i) \llbracket U_i \rrbracket_0 d\mathcal{S} \right).
 \end{aligned} \tag{7.30}$$

The DG formulation is conservative if the last two integrals in (7.30) are zero. This can be achieved by taking $\hat{V}_i = \llbracket V_i \rrbracket$. \square

Consider again (7.27). Integrate by parts the volume integral and using Theorem 7.3.1 to choose the numerical flux for the test function $V \in W_h$, we obtain:

$$\begin{aligned}
0 = & \sum_n \left(- \sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} (V_{i,0} U_i + V_{i,k} (F_{ik}(U) - (A_{ikrs}(U) U_{r,s})^{DG})) d\mathcal{K} \right. \\
& + \sum_{S \in \mathcal{S}_{IB}^n} \int_S (V_i^L - V_i^R) \{F_{ik}(U) - v_k U_i\} \bar{n}_k^L dS \\
& \left. - \sum_{S \in \mathcal{S}_{IB}^n} \int_S \llbracket V_i \rrbracket_k \{ (A_{ikrs}(U) U_{r,s})^{DG} \} dS + \sum_{S \in \mathcal{S}_K^n} \int_S \llbracket V_i \rrbracket_0 \{U_i\} dS \right), \tag{7.31}
\end{aligned}$$

where we used relation (7.15) for the time component of the space-time normal vector and relations (7.16) and (7.17) to write the element boundary integrals as face integrals. To use the definition of the generalized DG diffusion term (7.22) and (7.23), we have to choose $\varphi_{ik} = V_{i,k} \in V_h$. Then, from Theorem 7.3.1, it follows directly that $\hat{\varphi} = \{\{\varphi\}\}$ and we obtain:

$$\begin{aligned}
0 = & \sum_n \left(- \sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} (V_{i,0} U_i + V_{i,k} (F_{ik}(U) - A_{ikrs}(U) U_{r,s})) d\mathcal{K} \right. \\
& + \sum_{S \in \mathcal{S}_{IB}^n} \int_S (V_i^L - V_i^R) \{F_{ik}(U) - v_k U_i\} \bar{n}_k^L dS + \sum_{S \in \mathcal{S}_K^n} \int_S \llbracket V_i \rrbracket_0 \{U_i\} dS \\
& - \sum_{S \in \mathcal{S}_{IB}^n} \int_S \llbracket V_i \rrbracket_k \{A_{ikrs}(U) U_{r,s} + \mathcal{R}_{ik}^A\} dS \\
& \left. + \sum_{S \in \mathcal{S}_{IB}^n} \int_S \{V_{i,k}\} \left(\int_0^1 A_{ikrs}(\phi(\tau; U^L, U^R)) \frac{\partial \phi_r}{\partial \tau}(\tau; U^L, U^R) d\tau \bar{n}_s^L \right) dS \right). \tag{7.32}
\end{aligned}$$

The DG formulation is generally numerically unstable and a stabilizing term is added to the central flux $\{F_{ik} - v_k U_i\} \bar{n}_k^L$, together forming an upwind numerical flux: $H_i = (\{F_{ik} - v_k U_i\} + H_{ik}^{stab}) \bar{n}_k^L$. Furthermore, the central flux on the faces S_K is replaced by an upwind flux, naturally ensuring causality in time:

$$\hat{U} = \begin{cases} U^L & \text{if } n_0^L = 1, \\ U^R & \text{if } n_0^L = -1. \end{cases}$$

We obtain now the DG weak formulation: *Find a $U \in W_h$ such that for all*

$V \in W_h$:

$$\begin{aligned}
 0 = & - \sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}_j} (V_{i,0} U_i + V_{i,k} (F_{ik}(U) - A_{ikrs}(U) U_{r,s})) dK \\
 & + \sum_{\mathcal{S} \in \mathcal{S}_{iB}^n} \int_{\mathcal{S}} (V_i^L - V_i^R) H_i d\mathcal{S} - \sum_{\mathcal{S} \in \mathcal{S}_{iB}^n} \int_{\mathcal{S}} \llbracket V_i \rrbracket \{ \{ A_{ikrs}(U) U_{r,s} + \mathcal{R}_{ik}^A \} \} d\mathcal{S} \\
 & + \sum_{\mathcal{S} \in \mathcal{S}_{iB}^n} \int_{\mathcal{S}} \{ \{ V_{i,k} \} \} \left(\int_0^1 A_{ikrs}(\phi(\tau; U^L, U^R)) \frac{\partial \phi_r}{\partial \tau}(\tau; U^L, U^R) d\tau \bar{n}_s^L \right) d\mathcal{S} \\
 & + \sum_{\mathcal{K} \in \mathcal{T}_h^n} \left(\int_{K_j(t_{n+1}^-)} V_i^L U_i^L dK - \int_{K_j(t_n^+)} V_i^L U_i^R dK \right),
 \end{aligned} \tag{7.33}$$

where we used

$$\sum_{\mathcal{S} \in \mathcal{S}_K^n} \int_{\mathcal{S}} \llbracket V_i \rrbracket_0 \hat{U}_i d\mathcal{S} = \sum_{\mathcal{K} \in \mathcal{T}_h^n} \left(\int_{K_j(t_{n+1}^-)} V_i^L U_i^L dK - \int_{K_j(t_n^+)} V_i^L U_i^R dK \right), \tag{7.34}$$

which follows from (7.17). Due to the introduction of the upwind flux at the faces \mathcal{S}_K , each space-time slab only depends on the previous space-time slab so that the summation over all space-time slabs in (7.33) could be dropped. We see that (7.33) and (7.22) still require an expression for the path ϕ . Since the terms in which the path is required are diffusion terms, the expression is not very important: in the limit of the element size tending to zero, there are no discontinuities in the solution and hence these terms vanish. In the remainder of this chapter we therefore choose a linear path: $\phi(\tau; U^L, U^R) = U^L + \tau(U^R - U^L)$.

Recovering the weak formulation of Brezzi et al. [13]

We show in this section, that using the theory developed in the previous sections it is possible to recover the weak formulation of Brezzi et al. [13] (7.18). For this, we introduce the generalized DG derivative. Using Theorem 7.1.1 and (7.12) with $A_{ikrs} = 1$, the generalized DG derivative of U_i with respect to \bar{x}_k denoted

by $D_k^{DG}U_i$, $k = 1, \dots, q$, $i = 1, \dots, m$, satisfies:

$$\begin{aligned}
\int_{\mathcal{E}_h} \varphi_{ik} D_k^{DG}U_i d\mathcal{E}_h &= \sum_n \left(\sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} \varphi_{ik} d\mu_{ik}^A(\mathcal{K}_j) + \sum_{S \in \mathcal{S}_{IB}^n} \int_S \hat{\varphi}_{ik} d\mu_{ik}^A(S) \right) \\
&= \sum_n \left(\sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} \varphi_{ik} U_{i,k} d\mathcal{K} - \sum_{S \in \mathcal{S}_{IB}^n} \int_S \hat{\varphi}_{ik} \llbracket U_i \rrbracket_k dS \right) \\
&= \sum_n \left(\sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} \varphi_{ik} U_{i,k} d\mathcal{K} - \sum_{S \in \mathcal{S}_{IB}^n} \int_S \{\{\varphi_{ik}\}\} \llbracket U_i \rrbracket_k dS \right) \\
&= \int_{\mathcal{E}_h} \varphi_{ik} (U_{i,k} + \mathcal{R}_{ik}^d) d\mathcal{E}_h, \quad \forall \varphi \in V_h,
\end{aligned} \tag{7.35}$$

where $\mathcal{R}^d \in \mathbb{R}^{m \times q}$ is the global lifting operator defined as: Find an $\mathcal{R}^d \in V_h$, such that for all $\varphi \in V_h$:

$$\sum_n \left(\sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} \varphi_{ik} \mathcal{R}_{ik}^d d\mathcal{K} \right) = - \sum_n \left(\sum_{S \in \mathcal{S}_{IB}^n} \int_S \{\{\varphi_{ik}\}\} \llbracket U_i \rrbracket_k dS \right). \tag{7.36}$$

and where we take $\hat{\varphi}_{ik} = \{\{\varphi_{ik}\}\}$, which is a direct consequence of Theorem 7.3.1. We see that this definition of the global lifting operator \mathcal{R}^d exactly corresponds to that of \mathcal{R} in (7.19) if we choose $\hat{U} = \{\{U\}\}$, as was done in [10]. It follows that $D_k^{DG}U_i = U_{i,k} + \mathcal{R}_{ik}^d$ a.e. on \mathcal{E}_h . Instead of deriving the space-time DG weak formulation for (7.24), we derive now the weak formulation for

$$\tilde{D}_k^{DG} \hat{\mathcal{F}}_{ik}(U) = 0, \quad x \in \mathbb{R}^{q+1}, \quad x_0 > 0, \quad k = 0, 1, \dots, q, \tag{7.37}$$

with $\hat{\mathcal{F}} \in \mathbb{R}^{m \times (q+1)}$ given by:

$$\hat{\mathcal{F}}_{ik}(U) = \begin{cases} U_i, & \text{if } k = 0, \\ F_{ik}(U) - A_{ikrs}(U) D_s^{DG}U_r, & \text{if } k = 1, \dots, q. \end{cases} \tag{7.38}$$

The steps of deriving the weak formulation follows closely those of Section 7.3.3. We therefore only consider some intermediate results. First, (7.31) becomes

$$\begin{aligned}
0 &= \sum_n \left(- \sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} (V_{i,0}U_i + V_{i,k}(F_{ik}(U) - A_{ikrs}(U)D_s^{DG}U_r)) d\mathcal{K} \right. \\
&\quad + \sum_{S \in \mathcal{S}_{IB}^n} \int_S (V_i^L - V_i^R) \{\{F_{ik}(U) - v_k U_i\}\} \bar{n}_k^L dS + \sum_{S \in \mathcal{S}_K^n} \int_S \llbracket V_i \rrbracket_0 \{\{U_i\}\} dS \\
&\quad \left. - \sum_{S \in \mathcal{S}_{IB}^n} \int_S \llbracket V_i \rrbracket_0 \{\{A_{ikrs}(U)D_s^{DG}U_r\}\} dS \right).
\end{aligned} \tag{7.39}$$

Using the definition of the generalized DG derivative (7.35) in (7.39) we obtain exactly (7.18), hence using the Borel measures, we have recovered the weak formulation of Brezzi et al. [13]. As in Section 7.3.3 we introduce numerical fluxes for stabilization purposes. We obtain now the DG weak formulation: Find a $U \in W_h$ such that for all $V \in W_h$:

$$\begin{aligned}
 0 = & - \sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} (V_{i,0} U_i + V_{i,k} (F_{ik} - A_{ikrs}(U) U_{r,s})) dK \\
 & + \sum_{\mathcal{K} \in \mathcal{T}_h^n} \left(\int_{K(t_{n+1}^-)} V_i^L U_i^L dK - \int_{K(t_n^+)} V_i^L U_i^R dK \right) \\
 & + \sum_{\mathcal{S} \in \mathcal{S}_{IB}^n} \int_{\mathcal{S}} (V_i^L - V_i^R) H_i d\mathcal{S} - \sum_{\mathcal{S} \in \mathcal{S}_{IB}^n} \int_{\mathcal{S}} \llbracket V_i \rrbracket_k \{ \{ A_{ikrs}(U) (U_{r,s} + \mathcal{R}_{rs}^d) \} \} d\mathcal{S} \\
 & - \sum_{\mathcal{S} \in \mathcal{S}_{IB}^n} \int_{\mathcal{S}} \{ \{ V_{i,k} A_{ikrs}(U) \} \} \llbracket U_r \rrbracket_s d\mathcal{S}.
 \end{aligned} \tag{7.40}$$

The essential difference between (7.33) and (7.40) is the treatment of the non-linear term $A_{ikrs}(U) U_{r,s}$. In (7.33) this term is treated as a nonconservative product, while in (7.40) $A_{ikrs}(U)$ and $U_{r,s}$ are treated separately. We remark that if the tensor A is constant, (7.33) and (7.40) are the same.

7.3.4 A comment on the derivation for hyperbolic PDE's

We return in this section to the example of the 1D hyperbolic scalar equation in the introduction of this chapter. We will show that we do not need relation (7.6) to obtain (7.7).

Let $\tilde{\partial}_x^{DG} f(u)$ denote the 1D (inviscid) space-DG equivalent of $\tilde{D}_k^{DG} \mathcal{F}_{ik}(U)$ in (7.20). Then, in a DG framework we seek the weak formulation for:

$$\partial_t u + \tilde{\partial}_x^{DG} f(u) = 0, \quad x \in \Omega_h \subset \mathbb{R}, \quad t > 0. \tag{7.41}$$

Multiplying (7.41) by a test function $v_h \in W_h$, integrating over the domain Ω_h and using the definition of the 1D space-DG equivalent of the generalized DG flux derivative (7.20), we obtain (compare with (7.3)):

$$0 = \sum_{K \in \mathcal{T}_h} \int_K v_h (\partial_t u_h + \partial_x f(u_h)) dK - \underbrace{\sum_{\mathcal{S}} \int_{\mathcal{S}} \hat{v}_h (\bar{n}^L f(u_h^L) + \bar{n}^R f(u_h^R)) d\mathcal{S}}_{(7.42)},$$

$\forall v_h \in W_h$. Integrate by parts the volume integral and using Theorem 7.3.1 to define $\hat{v}_h = \frac{1}{2}(v_h^L + v_h^R)$ we find (compare with (7.4)):

$$0 = \sum_{K \in \mathcal{T}_h} \int_K (v_h \partial_t u_h - f(u_h) \partial_x v_h) dK + \sum_{K \in \mathcal{T}_h} \int_{\partial K} v_h^L f(u_h^L) \bar{n}^L d(\partial K) \\ - \sum_{\mathcal{S}} \int_{\mathcal{S}} \frac{1}{2} (v_h^L + v_h^R) (\bar{n}^L f(u_h^L) + \bar{n}^R f(u_h^R)) d\mathcal{S}, \quad \forall v_h \in W_h. \quad (7.43)$$

Using equality (7.5), the sum of the boundary integrals in (7.43) can be written as a sum over face integrals. We immediately see that the underlined term in (7.43) cancels against the underlined term in (7.5). We automatically get (7.7) without having to use the (questionable) relation (7.6).

7.4 Test cases

To compare the original weak formulation (7.40) and the new formulation (7.33) we conduct two test cases. We consider steady state solutions of the compressible Navier-Stokes equations for a cylinder and a NACA0012 airfoil.

The two dimensional Navier-Stokes equations are given by (7.9) in which

$$U = \begin{bmatrix} \rho \\ \rho u_j \\ \rho E \end{bmatrix}, \quad F_k = \begin{bmatrix} \rho u_k \\ \rho u_j u_k + p \delta_{jk} \\ u_k (\rho E + p) \end{bmatrix}, \quad A_{ikrs}(U) = \frac{\partial F_{ik}^v(U, \nabla U)}{\partial (U_{r,s})},$$

with $j, k = 1, 2$, $i, r = 1, \dots, 4$, and where $F_k^v = [0, \tau_{jk}, \tau_{jk} u_j - q_k]^T$. Here ρ is the density, ρu_k the momentum density in the Cartesian coordinate direction x_k , ρE the total energy density and δ the Kronecker delta function. For a full description of the homogeneity tensor A_{ikrs} , we refer to [41]. The total stress tensor τ is defined as $\tau_{jk} = \lambda u_{i,i} \delta_{jk} + \mu (u_{j,k} + u_{k,j})$ with $i = 1, 2$ and the dynamic viscosity coefficient μ is given by Sutherland's law:

$$\frac{\mu}{\mu_\infty} = \frac{T_\infty + T_S}{T + T_S} \left(\frac{T}{T_\infty} \right)^{3/2},$$

where T is the temperature and T_S a constant. The subscript ∞ denotes free-stream values. The second viscosity coefficient λ is related to μ by the Stokes hypothesis: $3\lambda + 2\mu = 0$. The heat flux vector q has components $q_k = -\kappa T_{,k}$, with κ the thermal conductivity coefficient. We assume a calorically perfect gas in thermodynamic equilibrium. The pressure p and internal energy e are given by the equations of state $p = \rho RT$ and $e = c_v T$, where $R = c_p - c_v$ is the specific gas constant and c_p and c_v the specific heats at constant pressure and constant

volume, respectively. The total energy is the sum of the internal and kinetic energy: $E = e + \frac{1}{2}u_i u_i$. Furthermore, the ratio of specific heats is denoted by $\gamma = c_p/c_v$.

We non-dimensionalize the Navier-Stokes equations by choosing the recurrent set $\{\rho_\infty, a_\infty, T_\infty, L\}$ in which L is the characteristic length scale of the problem and $a_\infty = \sqrt{\gamma p_\infty/\rho_\infty}$ is the free-stream speed of sound. After dimensionalization, six dimensionless Pi groups can be formed. In the simulations we set the ratio of specific heats $\gamma = 1.4$, the Prandtl number $Pr = c_p \mu_\infty/\kappa_\infty = 0.72$ and the ratio $\theta_S = T_S/T_\infty = 0.4$. The remaining three dimensionless groups, viz. the angle of attack α , free-stream Mach number $M_\infty = u_\infty/a_\infty$ and free-stream Reynolds number $Re_\infty = \rho_\infty u_\infty L/\mu_\infty$, are specified depending on the test case. By definition of our non-dimensionalization, $\rho_\infty = 1$, $a_\infty = 1$, $T_\infty = 1$ and $L = 1$.

In the implementation of the DG discretization (7.40), to improve the computational efficiency, it is customary to replace the central flux of the diffusion terms $\{\{A_{ikrs}(U)U_{r,s} + \mathcal{R}_{ik}^d\}\}$ by

$$\{\{A_{ikrs}(U)U_{r,s} + \mathcal{R}_{ik}^d\}\} \approx \{\{A_{ikrs}(U)U_{r,s} + \eta(\mathcal{R}_{ik}^d)^S\}\},$$

in which η is a stabilization constant and where the local lifting operator $(\mathcal{R}_{ik}^d)^S$ is an approximation of the global lifting operator. The local lifting operator is defined as: *Find an $(\mathcal{R}^d)^S \in V_h$, such that for all $\varphi \in V_h$:*

$$\sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} \varphi_{ik} (\mathcal{R}_{ik}^d)^S d\mathcal{K} = - \int_S \{\{\varphi_{ik}\}\} [U_i]_k d\mathcal{S}, \quad \mathcal{S} \in \mathcal{S}_{IB}^n. \quad (7.44)$$

Note, the local lifting operator is only non-zero in the elements connecting to the face \mathcal{S} . The local lifting operator is related to the global lifting operator through

$$\sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} \varphi_{ik} \mathcal{R}_{ik}^d d\mathcal{K} = \sum_{\mathcal{S} \in \mathcal{S}_{IB}^n} \sum_{\mathcal{K} \in \mathcal{T}_h^n} \int_{\mathcal{K}} \varphi_{ik} (\mathcal{R}_{ik}^d)^S d\mathcal{K},$$

see [41]. A similar simplification is applied to (7.33).

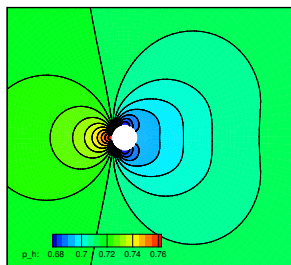
7.4.1 Cylinder

We first consider flow around a circular cylinder for a free-stream Mach number of $M_\infty = 0.3$, a free-stream Reynolds number of $Re_\infty = 40$ and an angle of attack of $\alpha = 0^\circ$. We use a piecewise quadratic polynomial approximation for the solution. To solve the solution to steady-state we employ three level multigrid as described in [42]. We use 10 pre- and post-smoothing steps with 20 pseudo-time steps to approximate the solution on the coarsest grid. We

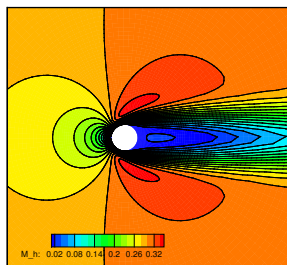
consider a structured O-type mesh with 128×128 elements. We considered the solution after the residual was reduced by six orders. The residual then, for both discretizations, was $6.2 \cdot 10^{-8}$. In Figure 7.1 we show the pressure and Mach contours as well as the absolute difference between the contours calculated with the Brezzi method and the new method. In Figure 7.2 we show the pressure coefficient distribution along the cylinder. The largest absolute difference in pressure is $1.0 \cdot 10^{-6}$ and in Mach number $1.0 \cdot 10^{-6}$.

7.4.2 NACA0012 airfoil

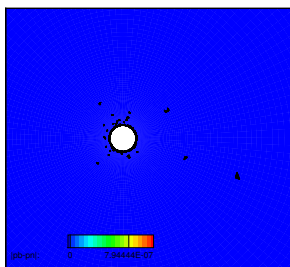
Next, we compute the viscous flow past a NACA0012 airfoil for a free-stream Mach number of $M_\infty = 0.8$, a free-stream Reynolds number of $Re_\infty = 73$ and an angle of attack of $\alpha = 10^\circ$. We use a piecewise linear polynomial approximation for the solution. To solve the solution to steady-state we employ the EXI-EXV pseudo-time stepping scheme of [40]. We use a structured C-type mesh with 320×80 elements. We considered the solution after the residual was reduced by three orders. The residual then, for both discretizations, was $3.3 \cdot 10^{-5}$. In Figure 7.3 we show the pressure coefficient distribution along the NACA0012 airfoil. In Figure 7.4 we show the pressure and Mach contours as well as the absolute absolute difference between the contours calculated with the Brezzi method and the new method. The largest absolute difference in pressure is $1.2 \cdot 10^{-4}$ and in Mach number $1.1 \cdot 10^{-4}$.



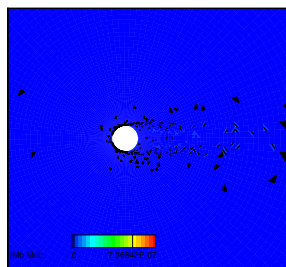
(a) Pressure contours.



(b) Mach contours.



(c) Absolute difference pressure contours.



(d) Absolute difference Mach contours.

Figure 7.1: *Cylinder test case ($Re_\infty = 40$, $M_\infty = 0.3$). Pressure and Mach contours as computed with a piece-wise quadratic polynomial approximation. Left column: the pressure contours and the absolute difference between the pressure contours calculated with the Brezzi method and the new formulation given by (7.33). Right column: the Mach contours and the absolute difference between the Mach contours calculated with the Brezzi method and the new formulation (7.33).*

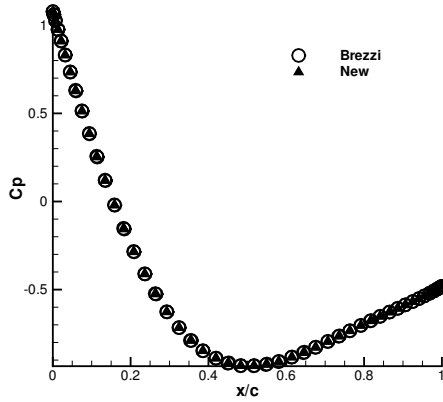


Figure 7.2: *Cylinder test case* ($Re_\infty = 40$, $M_\infty = 0.3$). *Pressure coefficient distribution along the cylinder.*

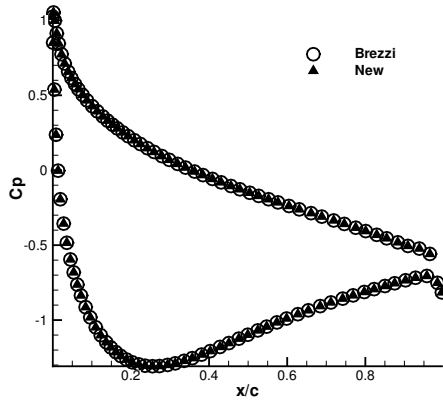
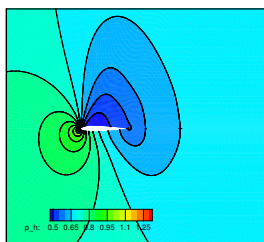
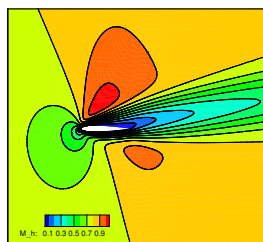


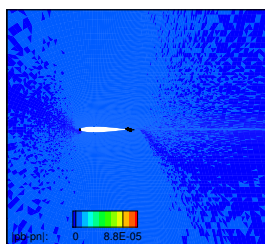
Figure 7.3: *NACA0012 test case* ($Re_\infty = 73$, $M_\infty = 0.8$, $\alpha = 10^\circ$). *Pressure coefficient distribution along the NACA0012 airfoil.*



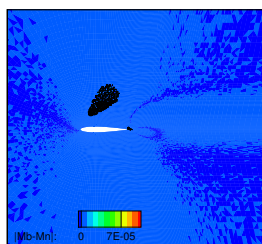
(a) Pressure contours.



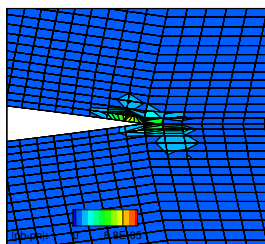
(b) Mach contours.



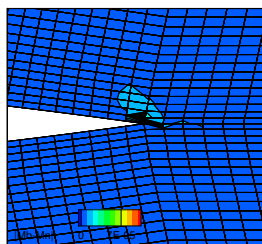
(c) Absolute difference pressure contours.



(d) Absolute difference Mach contours.



(e) Absolute difference pressure contours zoom.



(f) Absolute difference Mach contours zoom.

Figure 7.4: *NACA0012* test case ($Re_\infty = 73$, $M_\infty = 0.8$, $\alpha = 10^\circ$). Pressure and Mach contours as computed with a piece-wise linear polynomial approximation. Left column: the pressure contours and the absolute difference between the pressure contours calculated with the Brezzi method and formulation (7.33). Right column: the Mach contours and the absolute difference between the Mach contours calculated with the Brezzi method and formulation (7.33).

Conclusions and recommendations

In this thesis discontinuous Galerkin (DG) finite element methods aimed at solving hydrodynamic models of two-phase flows were presented. Also, significant attention was given to the development of efficient multigrid techniques for higher order accurate space-time DG discretizations. In this chapter, we draw conclusions and give recommendations for further research.

Nonconservative products. We have derived weak formulations for space- and space-time DG finite element methods for nonconservative hyperbolic partial differential equations. We also introduced a numerical flux for systems with nonconservative products (NCP-flux) suitable for DG finite element methods.

As test cases we considered the shallow water equations with and without dynamic topography (1D and 2D) and a simplified depth-averaged two-phase flow model. For the shallow water equations we considered rest flow over discontinuous topography and showed, both numerically and theoretically, that rest flow is preserved. We also considered subcritical and supercritical flow over a bump. For these test cases we obtained second and third order accuracy for suitable basis functions. We also considered more complex test cases: steady state transcritical flow with a shock, a perturbation of a steady state solution over a discontinuous topography and a dam breaking problem over a rectangular bump. For the two dimensional shallow water equations with dynamic topography, we considered hydraulic and morphological transport through a contraction.

For the simplified depth-averaged two-phase flow model we also considered subcritical and supercritical flow over a bump and again obtained second order accuracy using linear basis functions. A dam-break test case was further used to investigate the effect of the path on the numerical solution. The effect of the path was very small in the numerical solutions. Taking different paths did not lead to relevant changes in the final solution. We did see, however, that for certain paths it is not sufficient to simply use a two-point Gauss integration scheme over the whole domain of integration for the path integral, but higher order integration rules were required. It resulted in significantly larger computational cost which is undesirable.

Finally, we examined the effect of the path across a contact wave and saw that we could not capture the stationary contact discontinuity. By making the mesh such that the contact wave falls within an element we did see that the numerical error made is a full order smaller than if the contact wave falls exactly on a face. The numerical dissipation has a regularizing effect decreasing the effect of the path, but at the moment it is still unclear how to choose the path in case of a contact discontinuity and this is a topic of further research. The regularizing effect due to numerical dissipation across shock-waves is much larger explaining why we did not experience any significant effect of the path in test cases containing shock waves.

Depth-averaged two-phase flows. Recently, a depth-averaged two-phase flow model was introduced by Pitman and Le [62] and Le [45] to model shallow debris flows. We slightly extended this model by including extra friction terms to simulate turbulent friction. The depth-averaged model contains nonconservative products which makes it numerically challenging to solve. In Chapter 2 we developed a discontinuous Galerkin finite element method to deal with nonconservative products which we applied in Chapter 4 to solve the depth-averaged two-phase flow model of Le [45].

The DG finite element discretizations for the depth-averaged model were verified against steady-state flow solutions over a bump and we obtained second order convergence when using linear polynomial approximations. To prevent numerical oscillations, the WENO slope limiter [51] in combination with Krivodonova's discontinuity detector [43] was successfully applied. A Riemann problem solution was shown which could not be solved without the slope limiter due to severe undershoots.

Furthermore, the effect of the choice of the polynomials and the parameter γ in the slope limiter were shown. The scheme is robust for a wide range of γ values, but for accuracy reasons γ should be chosen as small as possible, because this minimizes the numerical dissipation. Also adding the Hermite polynomials to the combination of Lagrange and unlimited polynomials increases the amount

of numerical dissipation. This could be seen in the Riemann problem we investigated where there was a wave crest that could only be captured using the Lagrange and unlimited polynomials and setting $\gamma = 1$. Certain applications with strong gradients, however, need more numerical dissipation to avoid over- and undershoots so that γ may need to be slightly increased. This was necessary e.g. in the validation test case where we used the combination of Lagrange, Hermite and unlimited polynomials with $\gamma = 10$.

Finally, we qualitatively validated the model by showing its ability to capture the changes in a steady state solution with oblique jumps to a short time increase in the number of particles and compare with the results obtained by Akers and Bokhove [2].

Multigrid. In Chapter 5 we have discussed two- and three-level multigrid analysis for linear algebraic systems resulting from a higher order accurate space-time DG discretization of the 2D advection-diffusion equation. This allows the analysis of the convergence rate of a multigrid algorithm, but is also useful to optimize the smoothers in the multigrid algorithm. Optimal smoothers are obtained using a constrained optimization process, which finds smoother coefficients such that the spectral radius of the multigrid error transformation operator is minimal. This can result in a significant improvement in the convergence rate of the multigrid algorithm.

The multigrid optimization has been used to obtain improved algorithms for the solution of the algebraic system resulting from a space-time discontinuous Galerkin discretization of the advection-diffusion equation in two space dimensions. These algorithms have been tested on a two-dimensional problem containing boundary layers. The optimized algorithms showed a significant improvement compared to the original EXI-EXV Runge-Kutta method discussed in [40, 42]. Apart from optimizing the multigrid smoother, also the solution of the algebraic system on the coarsest mesh has a big impact on the multigrid performance.

In Chapter 6 we then compared and analyzed h -, p - and hp -multigrid methods employing the explicit EXI Runge-Kutta smoother for the solution of the space-time DG discretization of the Euler and 2D advection-diffusion equations. From the Fourier theory it is to be expected that hp -multigrid is the most efficient solution technique of the space-time DG discretization. From the numerical simulation of subsonic inviscid flow around a NACA0012 airfoil we saw, however, that the p -multigrid technique shows the best convergence rate. The p -multigrid is the most efficient due to its rapid residual reduction in the initial multigrid cycles and the least deterioration of the asymptotic convergence compared to the other multigrid techniques. The h -multigrid method shows a poor convergence rate after the high-frequency error modes have been

smoothed. A reason for this could be that at some point the coarse-grid problem of the h -multigrid algorithm is not solved well with respect to the characteristic components, as was demonstrated in [88]. This problem may be overcome by employing an improved coarse grid operator. This has not been studied in the present work, but is currently under investigation.

In the numerical computations of “real-life” problems, the choice of basis functions can have a significant effect on the stability of the scheme. For this reason we also investigated the spectrum of the discretization for different sets of basis functions. We found that the un-scaled Legendre basis functions allowed a 5 times larger pseudo-time step than the scaled Legendre basis functions in which the dimensionless mass-matrix on a uniform grid is the identity matrix. For this reason we suggest the use of un-scaled Legendre basis functions.

An alternative discontinuous Galerkin derivation. Defining generalized DG derivatives and DG diffusion terms, based on Borel measures, we have introduced an alternative derivation of the discontinuous Galerkin finite element weak formulation. Using Borel measures, we have also introduced a new DG weak formulation for parabolic/elliptic partial differential equations. We performed two numerical simulations of the compressible Navier-Stokes equations past a cylinder and a NACA0012 airfoil. We saw that with the new DG weak formulation we obtain very similar results as those obtained using the weak formulation of Brezzi et al. [13].

Regarding hyperbolic PDE’s, we showed that two commonly used approaches in deriving the DG weak formulation have their problems. We have introduced a mathematically more consistent derivation based on Borel measures.

Further research. Hydrodynamic two-phase flow models contain many interesting aspects, e.g., the presence of nonconservative products, stiff source terms, flows with free-surfaces and a mixture velocity field that has to be divergence free. In this thesis we have only addressed a few of these aspects and there are still many interesting research topics that need to be addressed before the DG method can be applied to fully 3D models of two-phase flows.

The incompressible Navier-Stokes equations can provide much insight into developing efficient DG methods for equations requiring divergence free velocity fields. The space-time DG method is particularly interesting in this respect since pseudo-compressibility methods in pseudo-time allow simple implementations of the method, but also because of its superior ability of dealing with moving grids needed for free-surface boundary conditions. However, computationally, the space-time DG method is expensive. We have set an initial step to improve the efficiency of solving the space-time DG discretization by developing and analyzing h -multigrid methods for 2D advection-diffusion equations. We have shown

that significant improvements in computational efficiency are possible. More research needs to be done to develop and analyze h -, p - and/or hp -multigrid methods specifically for equations requiring divergence free velocity fields. A good candidate for analysis would be the Oseen equations.

As stepping stones for fully 3D two-phase flow models, intermediate two-phase flow models can also be of interest and aid in the development. Examples include the depth-averaged two-phase flow model as addressed in this thesis, but also hydrostatic 3D two-phase flow models. The depth-averaged model was used to test the DG finite element method for nonconservative products. A hydrostatic 3D two-phase flow model, on the other hand, requires the development of a DG finite element method capable of dealing with moving boundaries on deforming meshes. Furthermore, such a model can also be used to test the optimized multigrid techniques for space-time DG discretizations of incompressible flows.

APPENDIX A

Derivation of the space DGFEM weak formulation for hyperbolic nonconservative partial differential equations

In this Appendix we derive a space DGFEM weak formulation for hyperbolic nonconservative partial differential equations (see also e.g. Cockburn and Shu [18] for more on the Runge-Kutta discontinuous Galerkin method for conservative hyperbolic systems). As opposed to the derivation of the weak formulation for space-time DGFEM in Chapter 2, we now only consider fixed grids. We first introduce the function spaces after which we derive the weak formulation.

The space DGFEM weak formulation

Let $\Omega \subset \mathbb{R}^q$ be the bounded flow domain approximated by Ω_h such that $\Omega_h \rightarrow \Omega$ as $h \rightarrow 0$, with h the radius of the smallest sphere completely containing the largest element K_j . Consider approximations of $U(x, t)$ and the test function $V(x, t)$ in the finite element space defined as:

$$W_h = \{V \in (L^2(\Omega_h))^m : V|_{K_j} \circ F_K \in (P^p(\hat{K}))^m\}, \quad (\text{A.1})$$

where m denotes the dimension of U .

The weak formulation for space DGFEM can be derived in a similar manner as that for space-time DGFEM, except that now we consider fixed grids.

Before discussing the space DGFEM weak formulation for equations containing nonconservative products, we first introduce as a reference the space DGFEM weak formulation for equations in conservative form (see e.g. Tassi, Bokhove and Vionnet [71]).

Consider partial differential equations in conservative form:

$$U_{i,0} + H_{ik,k} = 0, \quad \bar{x} \in \mathbb{R}^q, \quad t > 0, \quad (\text{A.2})$$

where $U \in \mathbb{R}^m$ and $H \in \mathbb{R}^m \times \mathbb{R}^q$. Using the approach discussed in Tassi, Bokhove and Vionnet [71], the space DG formulation for (A.2) can be stated as:

Find a $U \in W_h$ such that for all $V \in W_h$:

$$0 = \sum_j \int_{K_j} (V_i U_{i,0} - V_{i,k} H_{ik}) dK + \sum_{S \in \mathcal{S}_I} \int_S \llbracket V_i \rrbracket_k \{ \{ H_{ik} \} \} dS + \sum_{S \in \mathcal{S}_B} \int_S V_i^L H_{ik}^L \bar{n}_k^L dS. \quad (\text{A.3})$$

Note that at this point no numerical fluxes have been introduced yet into the DG formulation. We now continue with equations containing nonconservative products. Let $U \in W_h$ (see (A.1)). We know that the numerical solution is continuous on an element and discontinuous across a face, so, using Theorem 2.1.2, U is a weak solution to (2.3) if:

$$0 = \int_{\Omega_h} V_i U_{i,0} dK + \int_{\Omega_h} V_i d\bar{\mu}_i \quad (\text{A.4})$$

$$= \sum_j \int_{K_j} V_i (U_{i,0} + D_{ikr} U_{r,k}) dK + \sum_{S \in \mathcal{S}_I} \int_S \widehat{V}_i \left(\int_0^1 D_{ikr}(\phi(\tau; U^L, U^R)) \frac{\partial \phi_r}{\partial \tau}(\tau; U^L, U^R) d\tau \bar{n}_k^L \right) dS, \quad (\text{A.5})$$

where $V \in W_h$ is an arbitrary test function. Furthermore, \widehat{V} is the value (numerical flux) of the test function V on a face \mathcal{S} . Note that Theorem 2.1.2 is applied to nonconservative products in space-time where space and time variables are not explicitly distinguished. In space DGFEM this is the case and we only need the space part of the measure in Theorem 2.1.2. This measure is denoted in (A.4) as $\bar{\mu}_i$. The crucial point in obtaining the DG formulation is the choice of the numerical flux for the test function V . Using $D_{ikr} = \partial F_{ik} / \partial U_r + G_{ikr}$,

(A.5) can be rewritten as:

$$0 = \sum_j \int_{K_j} V_i(U_{i,0} + F_{ik,k} + D_{ikr}U_{r,k}) dK - \sum_{S \in \mathcal{S}_I} \int_S \widehat{V}_i \llbracket F_{ik} \rrbracket_k dS + \sum_{S \in \mathcal{S}_I} \int_S \widehat{V}_i \left(\int_0^1 D_{ikr}(\phi(\tau; U^L, U^R)) \frac{\partial \phi_r}{\partial \tau}(\tau; U^L, U^R) d\tau \bar{n}_k^L \right) dS \quad (\text{A.6})$$

We choose the numerical flux for V such that if there exists a Q such that $G_{ikr} = \partial Q_{ik} / \partial U_r$, then the DG formulation for the system containing nonconservative products reduces to the conservative space DGFEM weak formulation given by (A.3) with $H_{ik} = F_{ik} + Q_{ik}$.

Theorem A.1. *If the numerical flux \widehat{V} for the test function V in (A.6) is defined as $\widehat{V} = \{\{V\}\}$, then the weak formulation (A.6) will reduce to the conservative space DGFEM formulation (A.3) when there exists a Q such that $G_{ikr} = \partial Q_{ik} / \partial U_r$ so that $H_{ik} = F_{ik} + Q_{ik}$.*

Proof Assume there is a Q such that $G_{ikr} = \partial Q_{ik} / \partial U_r$. We immediately see:

$$\int_0^1 G_{ikr}(\phi(\tau; U^L, U^R)) \frac{\partial \phi_r}{\partial \tau}(\tau; U^L, U^R) d\tau \bar{n}_k^L = -\llbracket Q_{ik} \rrbracket_k. \quad (\text{A.7})$$

Integrating by parts the volume integral in (A.6) we obtain:

$$0 = \sum_k \int_{K_k} (V_i U_{i,0} - V_{i,k} (F_{ik} + Q_{ik})) dK + \sum_k \int_{\partial K_k} V_i^L (F_{ik}^L + Q_{ik}^L) \bar{n}_k^L d(\partial K) - \sum_{S \in \mathcal{S}_I} \int_S \widehat{V}_i \llbracket F_{ik} + Q_{ik} \rrbracket_k dS. \quad (\text{A.8})$$

We write $H_{ik} = F_{ik} + Q_{ik}$. Use relations (2.12) and (2.13) to write the element boundary integrals as face integrals:

$$\begin{aligned} \sum_j \int_{\partial K_j} V_i^L H_{ik}^L \bar{n}_k^L d(\partial K) &= \sum_{S \in \mathcal{S}_I} \int_S \llbracket V_i H_{ik} \rrbracket_k dS + \sum_{S \in \mathcal{S}_B} \int_S V_i^L H_{ik}^L \bar{n}_k^L dS \\ &= \sum_{S \in \mathcal{S}_I} \int_S \left(\{\{V_i\}\} \llbracket H_{ik} \rrbracket_k + (V_i^L - V_i^R) \{\{H_{ik}\}\} \bar{n}_k^L \right) dS \\ &\quad + \sum_{S \in \mathcal{S}_B} \int_S V_i^L H_{ik}^L \bar{n}_k^L dS. \end{aligned} \quad (\text{A.9})$$

Combining (A.8) and (A.9) we obtain:

$$0 = \sum_j \int_{K_j} V_i U_{i,0} - V_{i,k} H_{ik} dK + \sum_{\mathcal{S} \in \mathcal{S}_I} \int_{\mathcal{S}} (\{\{V_i\}\} [H_{ik}]_k + (V_i^L - V_i^R) \{\{H_{ik}\}\} \bar{n}_k^L) d\mathcal{S} \\ + \sum_{\mathcal{S} \in \mathcal{S}_B} \int_{\mathcal{S}} V_i^L H_{ik}^L \bar{n}_k^L d\mathcal{S} - \sum_{\mathcal{S} \in \mathcal{S}_I} \int_{\mathcal{S}} \widehat{V}_i [H_{ik}]_k d\mathcal{S}. \quad (\text{A.10})$$

The term $\{\{V_i\}\} [H_{ik}]_k$ is set to zero in the space DG formulation for conservative systems arguing that the formulation must be conservative. For a general nonconservative system we can not use this argument. Instead, we note that by taking $\widehat{V} = \{\{V\}\}$ on the faces \mathcal{S} , the contribution $\int_{\mathcal{S}} \{\{V_i\}\} [H_{ik}]_k d\mathcal{S}$ cancels with $-\int_{\mathcal{S}} \widehat{V}_i [H_{ik}]_k d\mathcal{S}$. We now obtain the weak formulation given by (A.3). \square

Theorem A.1 allows us to finalize the derivation of the DGFEM weak formulation, similar to the space-time DG formulation, to:

Find a $U \in W_h$ such that for all $V \in W_h$:

$$0 = \sum_j \int_{K_j} (V_i U_{i,0} - V_{i,k} F_{ik} + V_i G_{ikr} U_{r,k}) dK + \sum_{\mathcal{S}} \int_{\mathcal{S}} (V_i^L - V_i^R) \widehat{P}_i^{nc} d\mathcal{S} + \\ \sum_{\mathcal{S}} \int_{\mathcal{S}} \{\{V_i\}\} \left(\int_0^1 G_{ikr}(\phi(\tau; U^L, U^R)) \frac{\partial \phi_r}{\partial \tau}(\tau; U^L, U^R) d\tau \bar{n}_k^L \right) d\mathcal{S}. \quad (\text{A.11})$$

Note that we combined the fluxes at interior and boundary faces by using a ghost value U^R at the boundary.

The three-dimensional two-phase flow model

In this Appendix we present the three-dimensional two-phase flow model as derived by Jackson [38]. By depth-averaging this model, Pitman and Le [62] and Le [45] derived a depth-averaged two-phase flow model for shallow two-phase flows.

The three-dimensional two-phase flow model

Assume that the only fluid's stress is the fluid's pressure. Furthermore, the densities ρ^f and ρ^s of both phases are assumed to be constant. The three-dimensional model consists of two continuity equations and two momentum equations. To write the equations in compact form, we use the summation convention on repeated indices. The continuity equations are given by:

$$\begin{aligned}\partial_t((1 - \alpha)) + \partial_k((1 - \alpha)u_k) &= 0, \\ \partial_t(\alpha) + \partial_k(\alpha v_k) &= 0,\end{aligned}\tag{B.1}$$

and the momentum equations are:

$$\begin{aligned}\partial_t((1 - \alpha)\rho^f u_i) + \partial_k((1 - \alpha)\rho^f u_i u_k) &= -(1 - \alpha)\partial_k(\delta_{ik}p^f) - F_i^D + (1 - \alpha)\rho^f g_i, \\ \partial_t(\alpha\rho^s v_i) + \partial_k(\alpha\rho^s v_i v_k + T_{ik}^s) &= -\alpha\partial_k(\delta_{ik}p^f) + F_i^D + \alpha\rho^s g_i.\end{aligned}\tag{B.2}$$

Here, $i, k = 1, 2, 3$. The Cartesian coordinate system we consider is at an angle θ with respect to the horizontal (see Figure 4.1). In these equations α is the particle volume fraction, u the fluid velocity vector, v the solids velocity vector, \vec{g} the gravity vector, T^s the solids stress tensor, p^f is the fluid pressure, F^D the generalized drag force and δ represents the Kronecker delta function.

Bibliography

- [1] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. Dover Publications, New York, 1965.
- [2] B. Akers and O. Bokhove. Hydraulic flow through a channel contraction: Multiple steady states. *Phys. Fluids*, 20:056601 1–15, 2008.
- [3] V.R. Ambati and O. Bokhove. Space-time discontinuous Galerkin discretization of rotating shallow water equations. *J. Comput. Phys.*, 225:1233–1261, 2007.
- [4] V.R. Ambati and O. Bokhove. Space-time discontinuous Galerkin finite element method for shallow water flows. *J. Comput. Appl. Math.*, 204:452–462, 2007.
- [5] T.B. Anderson and R. Jackson. A Fluid Mechanical Description of Fluidized Beds. *Ind. Eng. Chem. Fundam.*, 6:527–539, 1967.
- [6] D. Arnold, F. Brezzi, B. Cockburn, and D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39:1749–1779, 2002.
- [7] N.J. Balmforth and R.R. Kerswell. Granular collapse in two dimensions. *J. Fluid Mech.*, 538:399–428, 2005.
- [8] F. Bassi, A. Crivellini, S. Rebay, and M. Savini. Discontinuous Galerkin solution of the Reynolds-averaged Navier-Stokes and $k-\omega$ turbulence model equations. *Computers & Fluids*, 34:507–540, 2004.

- [9] F. Bassi, A. Ghidoni, S. Rebay, and P. Tesini. High-order accurate p -multigrid discontinuous Galerkin solution of the Euler equations. *Int. J. Numer. Meth. Fluids*, 60:847–865, 2009.
- [10] F. Bassi and S. Rebay. A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations. *J. Comput. Phys.*, 131:267–279, 1997.
- [11] F. Bassi, S. Rebay, G. Mariotti, S. Pedinotti, and M. Savini. A high-order accurate discontinuous finite element method for inviscid and viscous turbomachinery flows. *in: R. Decuyper and G. Dibelius (Eds.), Second European Conference on Turbomachinery, Fluid Dynamics and Thermodynamics, Technologisch Instituut, Antwerpen*, pages 99–108, 1997.
- [12] A. Brandt. Rigorous quantitative analysis of multigrid, I: Constant coefficients two-level cycle with L_2 -norm. *SIAM J. Numer. Anal.*, 31:1695–1730, 1994.
- [13] F. Brezzi, G. Manzini, D. Marini, P. Pietra, and A. Russo. Discontinuous Galerkin approximations for elliptic problems. *Numer. Meth. Part. Diff. Eq.*, 16:365–378, 2000.
- [14] M. Castro, J.M. Gallardo, and C. Parés. High order finite volume schemes based on reconstruction of states for solving hyperbolic systems with non-conservative products. Application to shallow-water systems. *Math. Comput.*, 75:1103–1134, 2006.
- [15] M.-C. Chiou, Y. Wang, and K. Hutter. Influence of obstacles on rapid granular flows. *Acta Mechanica*, 175:105–122, 2005.
- [16] B. Cockburn. *Discontinuous Galerkin methods for convection dominated problems*. Lecture notes in computational science and engineering, Vol. 9, Springer-Verlag, Berlin, 1999.
- [17] B. Cockburn, G.E. Karniadakis, and C.W. Shu, editors. *Discontinuous Galerkin methods. Theory, computation and applications*. Lecture notes in computational science and engineering, Vol. 11, Springer-Verlag, Berlin, 2000.
- [18] B. Cockburn and C.W. Shu. The Runge-Kutta discontinuous Galerkin method for conservation laws V. *J. Comput. Phys.*, 141:199–224, 1998.
- [19] B. Cockburn and C.W. Shu. Runge-Kutta discontinuous Galerkin methods for convection-dominated problems. *J. Sci. Comput.*, 16:173–261, 2001.

- [20] R.P. Denlinger and R.M. Iverson. Flow of variably fluidized granular masses across three-dimensional terrain 2. Numerical predictions and experimental tests. *J. Geophys. Res.*, 106:553–566, 2001.
- [21] J. JR. Douglas and T. Dupont. *Interior penalty procedures for elliptic and parabolic Galerkin methods*. Lecture notes in physics, 58 Springer-Verlag, 1976.
- [22] D. A. Drew and R.T. Lahey. *Particulate Two-phase Flow*, chapter 16. Butterworth-Heinemann, Boston, 1993.
- [23] H. Enwald, E. Peirano, and A.E. Almstedt. Eulerian Two-phase Flow Theory Applied to Fluidization. *Int. J. Multiphase Flow*, 22:21–66, 1996.
- [24] K.J. Fidkowski, T.A. Oliver, J. Lu, and L. Darmofal. p -multigrid solution of high-order discontinuous Galerkin discretizations of the compressible Navier-Stokes equations. *J. Comput. Phys.*, 207:92–113, 2005.
- [25] S. Gottlieb and C.W. Shu. Total variation diminishing Runge-Kutta schemes. *Math. Comp.*, 67:73–85, 1998.
- [26] J.M.N.T. Gray, Y.-C. Tai, and S. Noelle. Shock waves, dead zones and particle-free regions in rapid granular free-surface flows. *J. Fluid Mech.*, 491:161–181, 2003.
- [27] W. Hackbusch, editor. *Multi-grid methods and applications*. Springer-Verlag, Berlin, 1985.
- [28] W. Hackbusch and U. Trottenberg, editors. *Multigrid methods: proceedings of the conference held at Köln-Porz, November 23-27, 1981*. Springer-Verlag, Berlin, 1982.
- [29] E. Hairer, S.P. Norsett, and G. Wanner, editors. *Solving Ordinary Differential Equations I*. Springer-Verlag, Berlin, 1993.
- [30] P.W. Hemker, W. Hoffmann, and M.H. van Raalte. Fourier two-level analysis for discontinuous Galerkin discretization with linear elements. *Numer. Linear Algebra Appl.*, 11:473–491, 2004.
- [31] D.D. Houghton and A. Kasahara. Nonlinear Shallow Fluid Flow Over an Isolated Ridge. *Comm. on Pure and Applied Math.*, 21:1–23, 1968.
- [32] <http://mathworld.wolfram.com/LegendrePolynomial.html>.
- [33] <http://volcanoes.usgs.gov/images/pglossary/lahar.php>.
- [34] <http://www.liceng.dk/LIC/Services/SlurryAndSediment/index.shtml>.

- [35] K. Hutter and K.R. Rajagopal. On flows of granular materials. *Continuum Mech. Thermodyn.*, 6:81–139, 1994.
- [36] K. Hutter, Y. Wang, and S.P. Pudasaini. The Savage-Hutter avalanche model: how far can it be pushed? *Phil. Trans. R. Soc. A*, 363:1507–1528, 2005.
- [37] R.M. Iverson and R.P. Denlinger. Flow of variably fluidized granular masses across three-dimensional terrain 1. Coulomb mixture theory. *J. Geophys. Res.*, 106:537–552, 2001.
- [38] R. Jackson. *The dynamics of fluidized particles*. Cambridge University Press, 2000.
- [39] A. Jameson. Time dependent calculations using multigrid, with applications to unsteady flows past airfoils and wings. *AIAA paper 91-1596*, 1991.
- [40] C.M. Klaij, J.J.W. van der Vegt, and H. van der Ven. Pseudo-time stepping methods for space-time discontinuous Galerkin discretizations of the compressible Navier-Stokes equations. *J. Comput. Phys.*, 219:622–643, 2006.
- [41] C.M. Klaij, J.J.W. van der Vegt, and H. van der Ven. Space-time discontinuous Galerkin method for the compressible Navier-Stokes equations. *J. Comput. Phys.*, 217:589–611, 2006.
- [42] C.M. Klaij, M.H. van Raalte, H. van der Ven, and J.J.W. van der Vegt. h -Multigrid for space-time discontinuous Galerkin discretizations of the compressible Navier-Stokes equations. *J. Comput. Phys.*, 227:1024–1045, 2007.
- [43] L. Krivodonova, J. Xin, J.F. Remacle, N. Chevaugeon, and J.E. Flaherty. Shock detection and limiting with discontinuous Galerkin methods for hyperbolic conservation laws. *Appl. Numer. Math.*, 48:323–338, 2004.
- [44] E.J. Kubatko, J.J. Westerink, and C. Dawson. An unstructured grid morphodynamic model with a discontinuous Galerkin method for bed evolution. *Ocean Modelling*, 15:71–89, 2006.
- [45] L.H. Le. *New models for geophysical flows*. Ph. D. dissertation, State University of New York at Buffalo, 2006.
- [46] P.G. LeFloch. Shock waves for nonlinear hyperbolic systems in nonconservative form, Report 593. *Institute for Mathematics and its Applications, Minneapolis, MN*, 1989.

- [47] R.J. LeVeque. Balancing source terms and flux gradients in high-resolution Godunov methods: the quasi-steady wave-propagation algorithm. *J. Comput. Phys.*, 146:346–365, 1998.
- [48] J. Ling, P.V. Skudarnov, C.X. Lin, and M.A. Ebadian. Numerical investigations of liquid-solid slurry flows in a fully developed turbulent flow region. *International Journal of Heat and Fluid Flow*, 24:389–398, 2003.
- [49] P. Lucas, H. Bijl, and H. van Zuijlen. Efficient unsteady high Reynolds number flow computations on unstructured grids. *submitted to Comput. Fluids*, 2009.
- [50] H. Luo, J.D. Baum, and R. Löhner. A p -multigrid discontinuous Galerkin method for the Euler equations on unstructured grids. *J. Comput. Phys.*, 211:767–783, 2006.
- [51] H. Luo, J.D. Baum, and R. Löhner. A Hermite WENO-based limiter for discontinuous Galerkin method on unstructured grids. *J. Comput. Phys.*, 225:686–713, 2007.
- [52] J.J. Major and R.M. Iverson. Debris-flow deposition: Effects of pore-fluid pressure and friction concentrated at flow margins. *GSA Bulletin*, 111:1424–1434, 1999.
- [53] B.S. Mascarenhas, B.T. Helenbrook, and H.L. Atkins. Application of p -multigrid to discontinuous Galerkin formulations of the Euler equations. *AIAA J.*, 47:1200–1208, 2009.
- [54] G. Dal Maso, P. G. LeFloch, and F. Murat. Definition and weak stability of nonconservative products. *J. Math. Pures Appl.*, 74:483–548, 1995.
- [55] N.D. Melson, M.D. Sanetrik, and H.L. Atkins. Time-accurate Navier-Stokes calculations with multigrid acceleration. *in Proc. 6th Copper mountain confer. on multigrid methods*, pages 423–437, 1993.
- [56] C.R. Nastase and D.J. Mavriplis. High-order discontinuous Galerkin methods using an hp -multigrid approach. *J. Comput. Phys.*, 213:330–357, 2006.
- [57] C. Parés. Numerical methods for nonconservative hyperbolic systems: a theoretical framework. *SIAM J. Numer. Anal.*, 44:300–321, 2006.
- [58] C. Parés and M. Castro. On the well-balance property of Roe’s method for nonconservative hyperbolic systems. Applications to shallow-water systems. *ESIAM: Mathematical Modeling and Numerical Analysis*, 38:821–852, 2004.

- [59] A.K. Patra, A.C. Bauer, C.C. Nichita, E.B. Pitman, M.F. Sheridan, M. Bursik, B. Rupp, A. Webber, A.J. Stinton, L.M. Namikawa, and C.S. Renschler. Parallel adaptive numerical simulation of dry avalanches over natural terrain. *J. Volc. Geothermal Research*, 139:1–21, 2005.
- [60] A.K. Patra, C.C. Nichita, A.C. Bauer, E.B. Pitman, M. Bursik, and M.F. Sheridan. Parallel adaptive discontinuous Galerkin approximation for thin layer avalanche modeling. *Computers and Geosciences*, 32:912–926, 2006.
- [61] L. Pesch and J.J.W. van der Vegt. A discontinuous Galerkin finite element discretization of the Euler equations for compressible and incompressible fluids. *J. Comput. Phys.*, 227:5426–5446, 2008.
- [62] E.B. Pitman and L. Le. A two-fluid model for avalanche and debris flows. *Phil. Trans. R. Soc. A*, 363:1573–1601, 2005.
- [63] O. Pouliquen and Y. Forterre. Friction law for dense granular flows: application to the motion of a mass down a rough inclined plane. *J. Fluid Mech.*, 453:133–151, 2002.
- [64] S. Rhebergen, O. Bokhove, and J.J.W. van der Vegt. Discontinuous Galerkin finite element methods for hyperbolic nonconservative partial differential equations. *J. Comput. Phys.*, 227:1887–1922, 2008.
- [65] B. Riviere, M.F. Wheeler, and V. Girault. Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems. Part I. *Comput. Geosci.*, 3:337–360, 1999.
- [66] R. Saurel and R. Abgrall. A Multiphase Godunov Method for Compressible Multifluid and Multiphase Flows. *J. Comput. Phys.*, 150:425–467, 1999.
- [67] S.B. Savage and K. Hutter. The dynamics of avalanches of granular materials down a rough incline. *J. Fluid Mech.*, 199:177–215, 1989.
- [68] K. Shahbazi, D.J. Mavriplis, and N.K. Burgess. Multigrid algorithms for high-order discontinuous Galerkin discretizations of the compressible Navier-Stokes equations. *J. Comput. Phys.*, 228:7917–7940, 2009.
- [69] J.J. Sudirham, J.J.W. van der Vegt, and R.M.J. van Damme. Space-time discontinuous Galerkin method for advection-diffusion problems on time-dependent domains. *Appl. Numer. Math.*, 56:1491–1518, 2006.
- [70] Y.C. Tai, S. Noelle, J.M.N.T. Gray, and K. Hutter. Shock-capturing and front-tracking methods for granular avalanches. *J. Comput. Phys.*, 175:269–301, 2002.

- [71] P.A. Tassi, O. Bokhove, and C.A. Vionnet. Space discontinuous Galerkin method for shallow water flows - kinetic and HLLC flux, and potential vorticity generation. *Advances in Water Resources*, 30:998–1015, 2007.
- [72] P.A. Tassi, S. Rhebergen, C.A. Vionnet, and O. Bokhove. A discontinuous Galerkin finite element model for river bed evolution under shallow flows. Additional appendices. <http://eprints.eemcs.utwente.nl/9962/>, pages 1–45, 2007.
- [73] P.A. Tassi, S. Rhebergen, C.A. Vionnet, and O. Bokhove. A discontinuous Galerkin finite element model for river bed evolution under shallow flows. *Comput. Methods Appl. Mech. Engrg.*, 197:2930–2947, 2008.
- [74] E.F. Toro. *Riemann Solvers and Numerical Methods for Fluid dynamics*. Springer-Verlag, 1997.
- [75] I. Toumi. A Weak Formulation of Roe’s Approximate Riemann Solver. *J. Comput. Phys.*, 102:360–373, 1992.
- [76] I. Toumi and A. Kumbaro. An Approximate Linearized Riemann Solver for a Two-Fluid Model. *J. Comput. Phys.*, 124:286–300, 1996.
- [77] U. Trottenberg, C.W. Oosterlee, and A. Schüller. *Multigrid*. Academic Press, London, 2001.
- [78] J.J.W. van der Vegt and S. Rhebergen. *Multigrid optimization using discrete Fourier analysis*. VKI Lecture Notes, 2009.
- [79] J.J.W. van der Vegt and H. van der Ven. Space-Time Discontinuous Galerkin Finite Element Method with Dynamic Grid Motion for Inviscid Compressible Flows I. General Formulation. *J. Comput. Phys.*, 182:546–585, 2002.
- [80] M.H. van Raalte and P.W. Hemker. Two-level multigrid analysis for the convection-diffusion equation discretized by a discontinuous Galerkin method. *Numer. Linear Algebra Appl.*, 12:563–584, 2005.
- [81] B.G.M. van Wachem, J.C. Schouten, and C.M. van den Bleek. Comparative Analysis of CFD Models of Dense Gas-Solid Systems. *AIChE Journal*, 47:1035–1051, 2001.
- [82] A.W. Vreman, M. Al-Tarazi, J.A.M. Kuipers, M. van Sint Annaland, and O. Bokhove. Supercritical shallow granular flow through a contraction, experiment, theory and simulation. *J. Fluid Mech.*, 578:233–269, 2007.

- [83] Y. Wang, K. Hutter, and S.P. Pudasaini. The Savage-Hutter theory: A system of partial differential equations for avalanche flows of snow, debris and mud. *Z. Angew. Math. Mech.*, 84:507–527, 2004.
- [84] P. Wesseling. *An introduction to multigrid methods*. Wiley, Chicester, 1991.
- [85] R. Wienands and W. Joppich. *Practical Fourier analysis for multigrid methods*. Chapman & Hall/CRC, 2005.
- [86] R. Wienands and C.W. Oosterlee. On three-grid Fourier analysis for multigrid. *SIAM J. Sci. Comput.*, 23:651–671, 2001.
- [87] Y Xing and C.W. Shu. High order well-balanced finite volume WENO schemes and discontinuous Galerkin methods for a class of hyperbolic systems with source terms. *J. Comput. Phys.*, 214:567–598, 2006.
- [88] I. Yavneh. Coarse-grid correction for nonelliptic and singular perturbation problems. *SIAM J. Sci. Comput.*, 19:1682–1699, 1998.
- [89] W.P. Ziemer. *Weakly Differentiable Functions: Sobolev Spaces and Functions of Bounded Variation*. Springer-Verlag (New York Inc), 1989.

Summary

The first research topic in this thesis is the development of space- and space-time discontinuous Galerkin (DG) finite element methods for hydrodynamic models of shallow (liquid-solid) two-phase flows. Many interesting aspects of these models arise, e.g., the presence of nonconservative products, stiff source terms, flows with free-surfaces and, in the three-dimensional (3D) model, the velocity field of the two-phase mixture has to be divergence free. This thesis provides some of the tools which are necessary for solving hydrodynamic models of two-phase flows with space- and/or space-time DG finite element methods.

An important research topic is nonconservative products which are present in many two-phase flow models. A large part of this thesis is therefore devoted to developing a general method which can be applied to partial differential equations containing nonconservative products. For this, we combine the theory of Dal Maso, LeFloch and Murat, in which a definition is given for nonconservative products even where the solution field is discontinuous. This theory also provides the mathematical foundation for a new DG finite element method. For this new DG method, we show standard $(p + 1)$ -order convergence results using p^{th} order basis-functions for test-cases of which we know the exact solution. We also show its ability to deal with more complex test cases. Finally, we apply the method to a depth-averaged two-phase flow model of which the numerical results are qualitatively validated against results obtained from a laboratory experiment.

The second topic of this thesis is multigrid. The use of multigrid is of great importance to obtain efficient solvers for fully 3D two-phase flow models. As an initial step to improve the efficiency of solving the space-time DG discretization, we have developed, analyzed and tested optimized multigrid methods using

explicit Runge-Kutta type smoothers for the 2D advection-diffusion equation.

Many physical models describing fluid motion contain second (and higher) order derivatives. Obtaining a DG discretization for these higher order derivatives is non-trivial and many different DG methods exist to deal with these terms. As final topic of this thesis we introduce an alternative derivation of DG methods based on Borel measures. This alternative derivation gives a consistent treatment of derivative terms by assigning a measure to derivatives when the flow field is discontinuous. We investigate the various DG weak formulations arising from this technique by considering the 2D compressible Navier-Stokes equations for the viscous flows over a cylinder and a NACA0012 airfoil.

Samenvatting

Het eerste onderwerp van onderzoek in dit proefschrift is het ontwikkelen van ruimte- en ruimte-tijd discontinue Galerkin (DG) eindige elementen methodes voor hydrodynamische modellen voor ondiepe (vloeistof-vaste deeltjes) twee-fasen stromingen. Deze modellen bevatten veel interessante aspecten, zoals de aanwezigheid van niet-conservatieve producten, stijve bron termen, stromingen met vrije oppervlaktes, en in het drie dimensionale (3D) model moet het snelheidsveld van het twee-fasen mengsel divergentie vrij zijn. In dit proefschrift ontwikkelen we een aantal technieken die nodig zijn voor het kunnen oplossen van hydrodynamische modellen voor twee-fasen stromingen met behulp van de ruimte- en/of ruimte-tijd DG eindige elementen methode.

Niet-conservatieve producten, die voorkomen in veel twee-fasen modellen, zijn een belangrijk onderzoeksonderwerp. Een groot gedeelte van dit proefschrift is daarom gewijd aan het ontwikkelen van een algemene methode die toegepast kan worden op partiële differentiaal vergelijkingen die niet-conservatieve producten bevatten. Hiervoor combineren we de theorie van Dal Maso, LeFloch en Murat, waarin een definitie wordt gegeven voor niet-conservatieve producten zelfs daar waar het oplossingsveld discontinue is. Deze theorie legt het wiskundige fundament voor een nieuwe DG eindige elementen methode. Voor deze nieuwe methode laten we standaard $(p + 1)$ -orde convergentie resultaten zien bij het gebruik van p^{de} orde basis-functies voor problemen waarvan de exacte oplossing bekend is. We laten ook zien dat deze methode geschikt is voor complexere test cases. Uiteindelijk passen we deze methode toe op een diepte-gemiddeld twee-fasen model waarbij de resultaten kwalitatief gevalideerd worden met experimenten uitgevoerd in het laboratorium.

Het tweede onderwerp van dit proefschrift is multigrid. Het gebruik van

multigrid, of andere efficiënte oplosmethodes, is van groot belang voor het oplossen van de volledige 3D twee-fasen modellen. Als eerste stap om de efficiëntie van de methode te verbeteren voor het oplossen van de ruimte-tijd DG discretizatie, hebben we geoptimaliseerde multigrid methodes met expliciete Runge-Kutta smoothers ontwikkeld, geanalyseerd en getest voor de 2D advection-diffusie vergelijking.

Veel fysische modellen die stromingen van vloeistoffen beschrijven, bevatten tweede (of hogere) orde afgeleides. Het afleiden van een DG discretizatie voor deze hogere orde afgeleides is niet triviaal en er bestaan vele verschillende DG methodes voor het omgaan met deze termen. Tenslotte introduceren wij in dit proefschrift een alternatieve afleiding voor DG methodes gebaseerd op Borel maten. Deze alternatieve afleiding geeft een consistente manier voor het omgaan met afgeleides door een maat toe te kennen aan afgeleides als het stromingsveld discontinue is. Wij onderzoeken de verschillende DG zwakke formuleringen die ontstaan door deze techniek door de 2D compressibele Navier-Stokes vergelijkingen te beschouwen voor visceuze stromingen om een cylinder en een NACA0012 vliegtuigvleugel.

Acknowledgments

The research presented in this thesis has been carried out over the past four years in the Numerical Analysis and Computational Mechanics (NACM) group at the University of Twente. I have had the pleasure of working together with a very enjoyable group and I would like to thank you all.

First of all I would like to thank Jaap van der Vegt and Onno Bokhove for giving me the opportunity of letting me do this PhD and the continuous motivation during this project. Jaap, working with you has been a great pleasure. You have always shown much interest in my work and helped in improving anything I delivered. You have always been very optimistic in always seeing the bright side of any results I presented you. Onno, thank you for all the enthusiasm that you have shown in my work and for giving me the freedom and confidence in everything I was doing.

Many thanks goes to Chris Klaij. After supervising me on my master's thesis, I could still count on your advice and many discussions during my PhD. Thanks also to Bernard, Fedderik and Arek for the many useful Monday morning meetings.

Thank you to my friends and colleagues Vijaya Ambati, Pablo Tassi, Bob Peeters, Domokos Sármany and Ivan Lakhturov. Thanks for the great company both at the office and outside. And Ivan, thanks for all the competition in table-tennis, pool, swimming, wake-boarding, climbing,...

Thanks to all members of NACM and AAMP for the friendly and enjoyable years. In particular I would like to thank Ruud van Damme for all his advice and help over the years. Also Henk Sollie for his help with C++ and hpGEM and Mike for his help with linear solvers. Finally, a special thanks to everyone I've shared my office with: Natanael, Hadi, Henk, Tim, Bob, Ivan, Lie, Marcel, Alyona, Vijaya and Domokos.

Sander Rhebergen, 2010

List of publications

Journal publications:

- S. Rhebergen, O. Bokhove and J.J.W. van der Vegt, Discontinuous Galerkin finite element method for shallow two-phase flows, *Comput. Methods Appl. Mech. Engrg.* 198:819-830, 2009.
- P.A. Tassi, S. Rhebergen, C.A. Vionnet and O. Bokhove, A discontinuous Galerkin finite element model for river bed evolution under shallow flows, *Comput. Methods Appl. Mech. Engrg.* 197:2930-2947, 2008.
Additional appendices are available on <http://eprints.eemcs.utwente.nl>.
- S. Rhebergen, O. Bokhove and J.J.W. van der Vegt, Discontinuous Galerkin finite element methods for hyperbolic nonconservative partial differential equations, *J. Comput. Phys.* 227:1887-1922, 2008.

Lecture notes:

- J.J.W. van der Vegt and S. Rhebergen, Multigrid optimization using discrete Fourier analysis, *Von Karman Institute lecture notes*, 2009.

Submitted:

- S. Rhebergen, J.J.W. van der Vegt and H. van der Ven, Multigrid optimization for space-time discontinuous Galerkin discretizations of advection dominated flows, *Book chapter ADIGMA*, Springer, 2009.

In preparation:

- S. Rhebergen and J.J.W. van der Vegt, An alternative derivation of the discontinuous Galerkin finite element weak formulation.
- J.J.W. van der Vegt and S. Rhebergen, Multigrid optimization for higher order accurate space-time discontinuous Galerkin discretizations.

About the author

The author was born in Kagondo, Tanzania on October 19, 1981. He attended Mananga Primary School in Tshaneni, Swaziland from 1986-1993, before moving to the Netherlands. In 2000 he got his VWO diploma at the RSG 't Slingerbos in Harderwijk.

From 2000-2005 he studied Applied Mathematics, with a minor in Aeronautical Engineering, at the University of Twente. A technical training was performed at the Aerodynamics group at the University of Delft. He implemented and tested the ESDIRK method in the Numecca software package. The research for his master thesis "*Approximate Riemann solver for viscous flows*" was performed at the University of Twente under the supervision of dr.ir. C.M. Klaij, prof.dr.ir. J.J.W. van der Vegt and prof.dr.ir. B.J. Geurts.

In 2005 he started his Ph.D.-research in the Numerical Analysis and Computational Mechanics group at the University of Twente under the supervision of prof.dr.ir. J.J.W. van der Vegt and dr.ir. O. Bokhove. The result of his research is contained in this thesis.