Weak Implementation of Boundary Conditions for the Finite-Volume Method

Fredrik Fryklund

LUND UNIVERSITY

Faculty of Engineering
Centre for Mathematical Sciences
Numerical Analysis

CENTRUM SCIENTIARUM MATHEMATICARUM

LUND UNIVERSITY

MASTER'S THESIS

Weak Implementation of Boundary Conditions for the Finite-Volume Method

Author:

Fredrik Fryklund
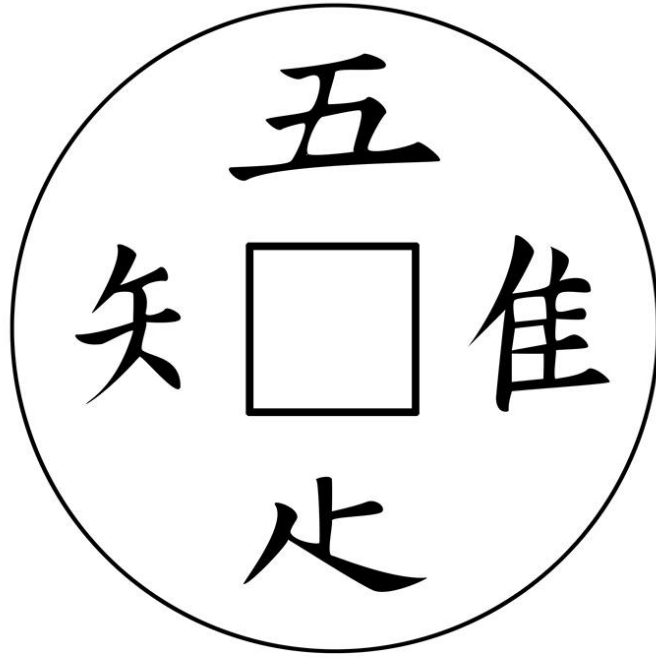
Supervisor:

Associate Professor

Philipp Birken

A thesis submitted in fulfilment of the requirements

for the degree of Graduate student

in the

Faculty of Engineering

Centre for Mathematical Sciences

Numerical Analysis

Examinator:

Professor Gustaf Söderlind

November 2014

五　□　矢　隹　火

(characters arranged around a central square forming 吾唯足知)

"われ ただ たる (を) しる。"

"Ware tada taru (wo) shiru."

"I know only satisfaction."

Inscription on a well,
Ryōan-ji, Kyoto, Japan.

LUND UNIVERSITY

Abstract

Faculty of Engineering
Centre for Mathematical Sciences
Numerical Analysis

Graduate student

Weak Implementation of Boundary Conditions
for the Finite-Volume Method

by Fredrik Fryklund

The Euler equations consist of conservation laws and describe a fluid in motion without viscous forces and heat conduction. They are non-linear and the solutions are often discontinuous. Therefore proofs of convergence are hard to give and the existing results are lacking. A new concept are weakly imposed characteristic boundary conditions, where a priori given boundary data only enter the scheme via ingoing characteristics. Thus a numerical solution for the boundary points is obtained as well. In combination with the node-centred finite-volume method, approximations of two-dimensional steady conservation laws can be made stable. Weakly imposed characteristic boundary conditions are compared to weakly imposed prescribed fluxes for the steady Euler equations, where the residuals converged to 10^{-11} and 10^{-3}, respectively. The performance of these boundary conditions is investigated further for different grid sizes and the unsteady Euler equations.

Populärvetenskaplig Sammanfattning

Att simulera en fluid (vätska eller gas) i rörelse med hjälp av datorer är mycket viktigt för forskning, men likväl för industri och samhälle. Istället för storskaliga experiment kan händelseförloppen undersökas genom datorberäkningar och en fluids beteende kan förutses.

Datorer kan inte lösa problemen exakt, utan ger approximativa lösningar. Att det introduceras fel i beräkningarna är inget problem om felen kan uppskattas på förhand, eftersom lösningens precision då kan kontrolleras. När en dator ska lösa ekvationerna som beskriver en fluid i rörelse tas en lösning fram som sedan uppdateras till en ny lösning. För högpresterande datoralgoritmer blir lösningen mer exakt desto längre koden får arbeta, vilket innebär att uppdateringarna blir mindre och mindre. För att kunna lita på datorns lösning måste det på förhand matematiskt bevisas att beräkningarna ger relevanta resultat. Dock är ekvationerna som beskriver en fluid i rörelse komplicerade, därmed finns få bevis som garanterar att datorns lösning är den som söks.

Denna rapport studerar en metod som introduceras i artikeln [1], där ekvationerna som beskriver en fluid förenklas med syfte att göra den matematiska analysen möjlig. Författarna hävdar att deras metod medför att fel, till följd av att lösningen är approximativ, inte ackumuleras utom kontroll. Dessutom ska datorberäkningar för de icke-förenklade ekvationerna med den nya metoden ge en lösning som blir mer exakt desto fler uppdateringar som görs.

Ofta är delar av lösningen känd. Till exempel är fluidens temperatur vid en kontaktyta känd, eller så är hastigheten vid ett inflöde bestämt på förhand. Vanligast är att datorns lösning ska uppfylla dessa värden exakt, men den nya metoden kräver inte det. Tanken är att det inte behövs, eftersom resten av lösningen ändå är en approximation. Resultatet av arbetet i [1] ger en mer flexibel metod. Deras slutsats verifieras i denna rapport och metoden förklaras mer detaljerat med ett större pedagogiskt fokus.

Vid datorberäkningar måste simuleringsområdet begränsas. För t.ex. en flygplansvinge är det intressant hur luften beter sig precis runt vingen, men inte vad som händer hundratals meter bort. Dessutom har en dator begränsat minne. Begränsningen av simuleringsområdet åstadkoms genom att införa en så kallad numerisk rand runt området man är intresserad av. I verkligheten är inte området begränsat, därför måste den numeriska randen modelleras som om den inte fanns. Det innebär att fluider som strömmar mot den bara ska passera rakt igenom. Ett omfattande experimentellt arbete har utförts baserat på metoden i [1]. I denna rapport visar vi att metoden är väl lämpad för att

modellera numeriska ränder. Detta möjliggör att simuleringsområdet kan minskas ytterligare, därmed behöver datorkraften endast användas för att simulera fluiden i de mest intressanta områdena.

Flödesproblem är mycket vanliga och har stor inverkan inom många områden. Genom att simulera fluiders rörelser kan man designa sina produkter i datorer eller undersöka fysikaliska fenomen. Istället för att tillverka hundratalet olika flygplansmodeller för att testa vilken som ger störst lyftkraft kan experimenten göras i datorer och på så vis minskar resursförbrukningen. Resultatet blir bränslesnålare flygplan, effektivare rotorblad till vindkraftverk eller någon av andra hundratals olika tillämpningsområden. Att metoden fungerar gör att den kan inspirera hur liknande problem ska hanteras och en djupare förståelse kan utvecklas.

Acknowledgements

Foremost, I would like to express my gratitude to my project advisor associate Prof. Philipp Birken who has guided, supported and encouraged me through the entire project. His enthusiasm for the subject has fed my curiosity and strengthened my determination to pursue further study in the field.

Besides my adviser, I would like to thank Prof. Gustaf Söderlind for all the inspirational conversations in his office and for the career- and study guidance.

Further, no visualisation of the data would have been possible without the help from Felicia Brisc from Klimacampus at the University of Hamburg. My sincerest thanks for not making this thesis entirely without figures.

This thesis is not just the result of my most recent project, but the final chapter of the journey I started five years ago moving to Lund. To all wonderful people I have spent seconds to days with: thank you for all the memories we have shared.

Also, my sincerest gratitude to my friends and family for always being there. Their warmth, constant support and perspective mean the world to me.

Finally, all my love to Jennifer for making all the difference by just being herself and for lifting my gaze upwards.

Contents

Abbreviations

BC   **B**oundary **C**ondition

IC   **I**nitial **C**ondition

ODE   **O**rdinary **D**ifferential **E**quation

PDE   **P**artial **D**ifferential **E**quation

Symbols

c	Speed of sound	m/s
E	Total energy	J $\left(m^2\,kg/s^2\right)$
E_{ij}	The edge the polygons Ω_i and Ω_j share	
$\mathcal{E}(\Omega_i)$	Set of edges for the polygon Ω_i	
$\mathbf{f}^N$	Numerical flux function	
H	Enthalpy	J $\left(m^2\,kg/s^2\right)$
m	Linear momentum	$kg\,m/s$
Ma	Mach number	
$\mathcal{N}(\Omega_i)$	Set of nodes for the polygon Ω_i	
$\mathbf{n}$	Outward pointing normal vector	
$\hat{\mathbf{n}}$	$\mathbf{n}$ normalised by $\|\mathbf{n}\|_{L^2(\Omega)}$	
p	State variable: pressure	Pa $\left(N/m^2\right)$
$\mathcal{T}(\Omega)$	Dual grid representation of Ω	
$\mathbf{u}$	Vector of state variables	
$\bar{\mathbf{u}}$	Discrete representation of $\mathbf{u}$ on $\mathcal{T}(\Omega)$	
v_1	State variable: velocity in x-direction	m/s
v_2	State varibale: velocity in y-direction	m/s
α	x-component of $\hat{\mathbf{n}}$	
β	y-component of $\hat{\mathbf{n}}$	
γ	Adiabatic index	
Δk_{ij}	Length of E_{ij} in k-direction	m
ρ	State variable: density	kg/m^3
Ω	Computational domain	$\Omega \subset \mathbb{R}^2$
$\boldsymbol{v}$	Veclocity vector	

Chapter 1

Introduction

The *Euler equations* are a set of *conservation laws* where *density*, *momentum* and *energy* of a moving inviscid fluid are related. This coupled system of hyperbolic partial differential equations (PDE) is well known and used for modelling a wide range of problems where the viscous forces and heat conduction are negligible. Therefore the study of the Euler equations is a relevant matter in several fields of study, with many real world applications.

The analysis of the Euler equations leaves much to be desired and generalised solutions are difficult to obtain. Similar problems are faced by the numerical analysis, as proofs of convergence for the numerical methods pose quite a challenge. Convergence proofs give mathematical credibility to numerical methods and consist of two key ingredients: *consistency* and *stability*. The latter may be defined in several ways, but the main property of stability is that the numerical scheme does not grow or magnify errors out of proportion. Errors are an ubiquitous factor in numerical methods and naturally so, since the methods are based on approximations. For numerical methods, the question is not whether one achieves the correct solution or not, but rather if the errors can be controlled.

Often are parts of the solution or specific requirements for it at the boundary given a priori, known as boundary conditions (BC). The main concept behind *weakly imposed* BCs, as opposed to *strongly imposed* BCs, is to not expect the BCs to be satisfied exactly. By implementing the BCs *weakly*, the values for the boundary points are never set explicitly, even though they are given as BCs. Instead a numerical solution for these points will be obtained as well and the BCs only influence the solution via boundary fluxes.

The importance of BCs for the well-posedness of PDEs is well known and adequate modelling is crucial in order to capture expected physical phenomena. In the article *Finite volume methods, unstructured meshes and strict stability for hyperbolic problems* by Jan Nordström, Karl Forsberg, Carl Adamsson and Peter Eliasson [1], weak BCs are shown to make the node-centred finite-volume method (FVM) , applied to a linear system, strictly stable in the sense that the L^2-norm of the problem's physical quantities does not grow over time. The FVM is a space-discretisation method, often used for the Euler equations, which produces a discrete solution represented on a finite amount of volume elements [2]. In [1] the BCs are imposed by specifying the ingoing characteristics with given boundary data, while local numerical data is used for the outgoing ones. Characteristics are closely related to other fluid properties, such as the *Mach number*, and may be understood as rays in the spatial domain, along which states propagate unchanged.

1.1 Aim of thesis

This thesis aims to validate and reproduce, to some extent, the results that Adamsson et al. present in [1]. They claim weak implementation of characteristic BCs is preferable to *strongly* implemented BCs for FVM, applied to linear hyperbolic conservation laws discretised on unstructured node-centred meshes, since strict stability can be obtained.

The analysis presented in [1] will be reviewed and presented in a less condensed format. This includes proving strict stability for the node-centred FVM, applied to a two-dimensional hyperbolic system of equations with constant coefficient matrices and weak characteristic BCs. The linearised two-dimensional Euler equations are studied in detail.

The relevance of the analysis for the nonlinear Euler equations will be evaluated with several numerical experiments, where the nonlinear Euler equations will be given weak characteristic BCs by prescribing the ingoing characteristics. First and foremost the aim is to reproduce the result in [1] where the residual of the fluxes for NACA0012 wing-profile simulations goes to 10^{-14}. Thereafter, the characteristic BCs' performance will be compared to the performance of prescribed fluxes, instead of strongly imposed BCs. Both are implemented weakly. If the weak characteristic BCs or the weak prescribed fluxes perform well, it might be possible to decrease the size of the computational domain and thereby reducing the amount of calculations. Finally both BCs will be applied to the Shu-vortex test case, where the analytical solution is known and eventual artifacts are studied.

1.2 Thesis overview

First, the main problem from [1] is presented in chapter 2 and it is shown to be hyperbolic and may be understood as a conservation law. Thereafter the Euler equations are presented and their different forms are discussed, where one of them is formulated as the main problem. The next chapter covers space-discretisation methods and the concept behind the solution scheme. The node-centred FVM is discussed in detail and the different aspects of BCs are presented. Chapter 4 follows the outline of [1], but not as condensed. Here, a matrix-vector representation of the discretised main problem is found and conditions for strict stability are given. In chapter 5 the scheme and theory presented in the previous chapter are applied to the linearised Euler equations. The BCs for the nonlinear Euler equations are derived and implementation is discussed. Thereafter the numerical experiments are considered and results are presented. Finally conclusions and thoughts about further research are presented in chapter 6.

Chapter 2

Hyperbolic systems of equations

In this chapter the mathematical properties of the main problem are presented, as formulated in [1]. The given problem is shown to be a *coupled hyperbolic system of PDEs* that may be understood as a *conservation law* and the concept of characteristics is introduced. The system of conservation laws, known as the *Euler equations*, is described in detail since it will be used for numerical validation. It is also shown how one may formulate the Euler equations to obtain the main problem.

2.1 The main problem

The problem Adamsson et al. study in [1] is formulated as

$$\mathbf{u}_t + \mathbf{A}\mathbf{u}_x + \mathbf{B}\mathbf{u}_y = \mathbf{0}, \quad (x,y) \in \Omega \subset \mathbb{R}^2,\ t \in \mathbb{R}_0^+, \tag{2.1}$$

where the subscripts t, x and y are the condensed notation for partial derivatives. It is given that Ω is bounded, $\mathbf{u} \in \mathbb{R}^l$ and $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{l \times l}$ are constant, square and symmetric matrices. Furthermore, suitable boundary- and initial conditions (IC) are assumed to be given and that the BCs are such that the matrix $(\mathbf{A}\,dy - \mathbf{B}\,dx)$ is positive semi-definite at every point on the boundary at any time t, where dx and dy are infinitesimal arc lengths in the x- and y-direction, respectively. That is

$$\mathbf{u}^T(\mathbf{A}\,dy - \mathbf{B}\,dx)\mathbf{u} \geq 0, \quad (x,y) \in \partial\Omega,\ t \in \mathbb{R}_0^+, \tag{2.2}$$

whose importance will become apparent in subsequent chapters when more specific requirements for the BCs are stated.

2.2 Conservation laws

Conservation laws describe a variety of physical phenomena where some quantities in an isolated system change over time. Change either comes via the boundary of the system or implies transformation into another quantity. A formal definition of conservation laws from [3] is stated below.

Definition 1 (Conservation Law). Let Ω be an open subset of $\mathbb{R}^l$, and let $\mathbf{f}_j$, $1 \leq j \leq d$, be d smooth functions from Ω into $\mathbb{R}^l$; the general form of a system of conservation laws in several space variables is

$$\frac{\partial \mathbf{u}}{\partial t} + \sum_{j=1}^{d} \frac{\partial}{\partial x_j} \mathbf{f}_j(\mathbf{u}) = \mathbf{0}, \quad \mathbf{x} = [x_1, \ldots, x_d]^T \in \mathbb{R}^d, \quad t \geq 0, \tag{2.3}$$

where

$$\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_l \end{bmatrix} \tag{2.4}$$

is a vector-valued function from $\mathbb{R}^d \times [0, \infty)$ into Ω. The set Ω is called the set of states and the functions

$$\mathbf{f}_j = \begin{bmatrix} f_{1j} \\ \vdots \\ f_{lj} \end{bmatrix} \tag{2.5}$$

are called the flux-functions. One says that (2.3) is written in *conservative form*.

Some conservation laws are also systems of hyperbolic PDEs, but before considering this notion diagonalisable systems need to be defined:

Definition 2. (Diagonalisable system) A matrix $\mathbf{A}$ is said to be diagonalisable if $\mathbf{A}$ can be expressed as

$$\mathbf{A} = \mathbf{K}\boldsymbol{\Lambda}\mathbf{K}^{-1} \quad \text{or} \quad \boldsymbol{\Lambda} = \mathbf{K}^{-1}\mathbf{A}\mathbf{K}, \tag{2.6}$$

in terms of a diagonal matrix $\boldsymbol{\Lambda}$ and an invertible matrix $\mathbf{K}$. The diagonal elements of $\boldsymbol{\Lambda}$ are the eigenvalues λ_i of $\mathbf{A}$ and the columns $\mathbf{K}^{(i)}$ of $\mathbf{K}$ are the right eigenvectors of $\mathbf{A}$ corresponding to the eigenvalues λ_i, that is

$$\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n), \; \mathbf{K} = [\mathbf{K}^1, \ldots, \mathbf{K}^m], \; \mathbf{A}\mathbf{K}^i = \lambda_i \mathbf{K}^i. \tag{2.7}$$

A system of type (2.3) is said to be diagonalisable if the coefficient matrix $\mathbf{A}$ is diagonalisable.

This allows the following definition, as written in [3]:

Definition 3. (Hyperbolic System) For all $j = 1, \ldots, d$, let

$$\mathbf{A}_j(\mathbf{u}) = \left(\frac{\partial f_{ij}}{\partial u_k}(\mathbf{u}) \right)_{1 \leq i,k \leq l} \tag{2.8}$$

be the Jacobian matrix of $\mathbf{f}_j(\mathbf{u})$. The system (2.3) is called *hyperbolic* if, for any $\mathbf{u} \in \Omega$ and any $\mathbf{w} \in \mathbb{R}^d$, $\mathbf{w} \neq \mathbf{0}$, the matrix

$$\mathbf{A}(\mathbf{u}, \mathbf{w}) = \sum_{j=1}^{d} w_j \mathbf{A}_j(\mathbf{u}) \tag{2.9}$$

has l real eigenvalues and is diagonalisable, see definition 2. Further the system is called *strictly hyperbolic* if the matrix $\mathbf{A}$ has p distinct eigenvalues $\lambda_1(\mathbf{u}, \mathbf{w}) < \lambda_2(\mathbf{u}, \mathbf{w}) < \ldots < \lambda_l(\mathbf{u}, \mathbf{w})$ and p linearly independent corresponding right eigenvectors $r_1(\mathbf{u}, \mathbf{w}), \ldots, r_l(\mathbf{u}, \mathbf{w})$, i.e.,

$$\mathbf{A}(\mathbf{u}, \mathbf{w}) \, \mathbf{r}_k(\mathbf{u}, \mathbf{w}) = \lambda_k(\mathbf{u}, \mathbf{w}) \, \mathbf{r}_k(\mathbf{u}, \mathbf{w}), \quad 1 \leq k \leq l. \tag{2.10}$$

It is easily seen that (2.3) can be formulated with the matrices $\mathbf{A}_j(\mathbf{u})$, since $\frac{\partial}{\partial x_j} \mathbf{f}_j(\mathbf{u}) = \mathbf{A}_j(\mathbf{u}) \frac{\partial \mathbf{u}}{\partial x_j}$, thus the hyperbolic property depends on the structure of the flux functions $\boldsymbol{f}_j$. The main problem (2.1) can be retrieved from the general formulation of conservation laws (2.3) if $j = 1, 2$ and $x_1 = x$, $x_2 = y$, $\mathbf{A}_1(\mathbf{u}) = \mathbf{A}(\mathbf{u})$ and $\mathbf{A}_2(\mathbf{u}) = \mathbf{B}(\mathbf{u})$. Since the matrices $\mathbf{A}$ and $\mathbf{B}$ are symmetric, their linear sum will also be symmetric and hence diagonalisable. The symmetric property also implies that all eigenvalues are real [4]. Therefore the main problem (2.1) is a system of hyperbolic PDEs and may be understood as a conservation law. The system must not necessarily be strictly hyperbolic, which will be the case with the Euler equations.

2.3 The Euler equations

The Euler equations is a system of coupled quasi-linear, first order PDEs of hyperbolic character. They describe the inviscid flow of a compressible fluid, where no heat conduction is present, and captures phenomena such as shock waves. They are derived by formulating equations for the conservation of *mass*, *momentum* and *energy* in an arbitrary bounded volume, called *control volume*. If one also consider viscous forces and heat conduction, then the mixed parabolic-hyperbolic system of nonlinear second order PDEs called the *Navier-Stokes* equations is obtained. In some aspects a first order system is easier to handle, for both analytical and numerical reasons. Also, the Euler equations

are preferable from a modelling perspective when the viscous effects are negligible, such as steady air flow around a wing profile [2].

Definition 4. The follwing system of quasi-linear first order PDEs are the Euler equations in conservative form:

$$\partial_t \rho + \nabla \cdot \mathbf{m} = 0, \tag{2.11}$$

$$\partial_t m_i + \sum_{j=1}^{d} \partial_{x_j}(m_i v_j + p\, \delta_{ij}) = 0, \qquad i = 1, \ldots, d \tag{2.12}$$

$$\partial_t(\rho E) + \nabla \cdot (H\mathbf{m}) = 0, \tag{2.13}$$

where δ_{ij} is the Kronecker delta and ρ, m_i, v_i, p, E and H are density, momentum in direction i per unit volume, the velocity in the i direction, pressure, total energy per unit mass and enthalpy [2].

Just by looking at a fluid in motion one realises that describing it is complicated. The forming and breaking of waves, eddies, shocks etc. clearly indicates that one should expect a discontinuous solution. In fact, even if the initial state for the Euler equations is smooth, singularities will develop in finite time. Therefore classical solutions are too much to ask for. With the introduction of the notion of weak solution and additional conditions, based on physics not included in the Euler equations, both uniqueness and existence are guaranteed in 1D. Sadly, similar results for general existence and uniqueness do not exist for the Euler equations in several space-variables [5, 6].

2.3.1 Forms of the Euler equations

By introducing the vector of conservative variables $\mathbf{u} = [\rho, m_1, m_2, \ldots, m_d, \rho E]^T$ and the convective flux vectors $\mathbf{f}_i^c(\mathbf{u}) = [m_i, m_i v_1, m_i v_2, \ldots, m_i v_i + p, \ldots, m_i v_d, \rho H v_i]^T$ with $\mathbf{f}^c(\mathbf{u}) = [\mathbf{f}_1^c(\mathbf{u}), \mathbf{f}_2^c(\mathbf{u}), \ldots, \mathbf{f}_d^c(\mathbf{u})]$ (2.11), (2.12) and (2.13) can be written as

$$\mathbf{u}_t + \nabla \cdot \mathbf{f}^c(\mathbf{u}) = \mathbf{0}. \tag{2.14}$$

Since the main problem (2.1) will be analysed it is desirable to express the Euler equations in this matrix-vector form, which is achieved by a series of similarity transformations. First, note that for the two-dimensional Euler equations $\nabla \cdot \mathbf{f}^c(\mathbf{u}) = \frac{\partial \mathbf{f}_1^c(\mathbf{u})}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial x} + \frac{\partial \mathbf{f}_2^c(\mathbf{u})}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial y}$. Set $\mathbf{A}(\mathbf{u}) = \frac{\partial \mathbf{f}_1^c(\mathbf{u})}{\partial \mathbf{u}}$ and $\mathbf{B}(\mathbf{u}) = \frac{\partial \mathbf{f}_2^c(\mathbf{u})}{\partial \mathbf{u}}$ where

$$\mathbf{A}(\mathbf{u}) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{\gamma-3}{2}v_1^2 + \frac{\gamma-1}{2}v_2^2 & (3-\gamma)v_1 & -(\gamma-1)v_2 & \gamma-1 \\ -v_1v_2 & v_2 & v_1 & 0 \\ -\gamma v_1 E + (\gamma-1)v_1\,|\boldsymbol{v}|^2 & \gamma E - \frac{\gamma-1}{2}(v_2^2 + 3v_1^2) & -(\gamma-1)v_1v_2 & \gamma v_1 \end{bmatrix},$$
(2.15)

$$\mathbf{B}(\mathbf{u}) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ -v_1v_2 & v_2 & v_1 & 0 \\ \frac{\gamma-3}{2}v_2^2 + \frac{\gamma-1}{2}v_1^2 & -(\gamma-1)v_1 & (3-\gamma)v_2 & \gamma-1 \\ -\gamma v_2 E + (\gamma-1)v_2\,|\boldsymbol{v}|^2 & -(\gamma-1)v_1v_2 & \gamma E - \frac{\gamma-1}{2}(v_1^2 + 3v_2^2) & \gamma v_2 \end{bmatrix},$$
(2.16)

with γ being the adiabatic index and $\boldsymbol{v} = [v_1, v_2]^T$ [7]. Then one obtaines the *conservative matrix form* of the Euler equations

$$\mathbf{u}_t + \mathbf{A}(\mathbf{u})\mathbf{u}_x + \mathbf{B}(\mathbf{u})\mathbf{u}_y = \mathbf{0}. \tag{2.17}$$

Note that the matrices $\mathbf{A}(\mathbf{u})$ and $\mathbf{B}(\mathbf{u})$ are neither symmetric nor constant. The former property may be obtained my transforming the conservative matrix form (2.17) to *primitive form*, which in turn may be symmetrised. The Euler equations in primitive form is an alternative, but equivalent representation, where the relations are expressed for the quantities *density*, *velocity* and *pressure*. Let the subscript p denote primitive form, then the vector of primitive variables in two dimensions is $\mathbf{u}_p = [\rho, v_1, v_2, p]^T$ and related to $\mathbf{u}$ by $\partial\mathbf{u} = \mathbf{M}\partial\mathbf{u}_p$, where ∂ denotes differentiating with respect to t, x or y and

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ v_1 & \rho & 0 & 0 \\ v_2 & 0 & \rho & 0 \\ \frac{|\boldsymbol{v}|^2}{2} & \rho v_1 & \rho v_2 & \frac{1}{\gamma-1} \end{bmatrix}, \tag{2.18}$$

$$\mathbf{M}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{v_1}{\rho} & \frac{1}{\rho} & 0 & 0 \\ -\frac{v_2}{\rho} & 0 & \frac{1}{\rho} & 0 \\ \frac{\gamma-1}{2}\,|\boldsymbol{v}|^2 & -(\gamma-1)v_1 & -(\gamma-1)v_2 & \gamma-1 \end{bmatrix}. \tag{2.19}$$

Multiplication of (2.17) with $\mathbf{M}^{-1}$ from the left gives

$$M^{-1}\frac{\partial u}{\partial t} + M^{-1}A(u)M\,M^{-1}\frac{\partial u}{\partial x} + M^{-1}B(u)M\,M^{-1}\frac{\partial u}{\partial y} = 0$$

$$\Leftrightarrow$$

$$\frac{\partial u_p}{\partial t} + A_p(u_p)\frac{\partial u_p}{\partial x} + B_p(u_p)\frac{\partial u_p}{\partial y} = 0$$

(2.20)

where the matrices $A_p(u_p)$ and $B_p(u_p)$ are

$$A_p(u_p) = M^{-1}AM = \begin{bmatrix} v_1 & \rho & 0 & 0 \\ 0 & v_1 & 0 & 1/\rho \\ 0 & 0 & v_1 & 0 \\ 0 & \gamma p & 0 & v_1 \end{bmatrix}, \quad B_p(u_p) = M^{-1}BM = \begin{bmatrix} v_2 & 0 & \rho & 0 \\ 0 & v_2 & 0 & 0 \\ 0 & 0 & v_2 & 1/\rho \\ 0 & 0 & \gamma p & v_2 \end{bmatrix}.$$

(2.21)

The two-dimensional Euler equations in primitive form (2.20) can be symmetrised simultaneously, meaning both $A_p(u_p)$ and $B_p(u_p)$ may be symmetrised by the transformation matrices

$$S = \begin{bmatrix} \frac{\rho\sqrt{\gamma}}{c} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{\rho c}{\sqrt{\gamma}} & 0 & 0 & \rho c\sqrt{\frac{\gamma-1}{\gamma}} \end{bmatrix}, \quad S^{-1} = \begin{bmatrix} \frac{c}{\rho\sqrt{\gamma}} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{-c}{\rho\sqrt{\gamma(\gamma-1)}} & 0 & 0 & \frac{1}{\rho c}\sqrt{\frac{\gamma}{\gamma-1}} \end{bmatrix},$$

(2.22)

derived in [8]. Following the same procedure as when obtaining (2.20) gives

$$S^{-1}\frac{\partial u_p}{\partial t} + S^{-1}A_p(u_p)S\,S^{-1}\frac{\partial u_p}{\partial x} + S^{-1}B_p(u_p)S\,S^{-1}\frac{\partial u_p}{\partial y} = 0$$

$$\Leftrightarrow$$

$$\frac{\partial u_s}{\partial t} + A_s(u_s)\frac{\partial u_s}{\partial x} + B_s(u_s)\frac{\partial u_s}{\partial y} = 0$$

(2.23)

where $\partial u_s = S^{-1}\partial u_p$ and

$$\mathbf{A}_s(\mathbf{u}_s) = \begin{bmatrix} v_1 & \frac{c}{\sqrt{\gamma}} & 0 & 0 \\ \frac{c}{\sqrt{\gamma}} & v_1 & 0 & c\sqrt{\frac{\gamma-1}{\gamma}} \\ 0 & 0 & v_1 & 0 \\ 0 & c\sqrt{\frac{\gamma-1}{\gamma}} & 0 & v_1 \end{bmatrix}, \quad \mathbf{B}_s(\mathbf{u}_s) = \begin{bmatrix} v_2 & 0 & \frac{c}{\sqrt{\gamma}} & 0 \\ 0 & v_2 & 0 & 0 \\ \frac{c}{\sqrt{\gamma}} & 0 & v_2 & c\sqrt{\frac{\gamma-1}{\gamma}} \\ 0 & 0 & c\sqrt{\frac{\gamma-1}{\gamma}} & v_2 \end{bmatrix},$$

$$(2.24)$$

with $\mathbf{A}_s(\mathbf{u}_s) = \mathbf{S}^{-1}\mathbf{A}_p(\mathbf{u}_p)\mathbf{S}$ and $\mathbf{B}_s(\mathbf{u}_s) = \mathbf{S}^{-1}\mathbf{B}_p(\mathbf{u}_p)\mathbf{S}$. As will be shown later this matrix structure will be of use for stability analysis.

The Euler equations in symmetrised primitive form (2.23) are not yet of the same form as (2.1), since the respective matrices $\mathbf{A}_s(\mathbf{u}_s)$ and $\mathbf{B}_s(\mathbf{u}_s)$ are not constant. They may however be made constant by linearising the system of equations around a solution, which makes the analysis for stability considerably simpler, or in some cases simply possible. To linearise, choose a solution $\tilde{\mathbf{u}}_s = \tilde{\mathbf{u}}_s(\tilde{x}, \tilde{y})$ to the Euler equations (2.24) satisfying possible initial- and boundary conditions. Then fix $\tilde{\mathbf{u}}_s$, by specifying $(\tilde{x}, \tilde{y}) \in \Omega$, then the constant matrices for (2.23) are given by $\mathbf{A}_s = \mathbf{A}_s(\tilde{\mathbf{u}}_s)$ and $\mathbf{B}_s = \mathbf{B}_s(\tilde{\mathbf{u}}_s)$ [2]. To distinguish the constant matrices from the non-constant, the argument $\mathbf{u}$ is always included for the latter.

The Euler equations can also be formulated in integral form, where expression 2.14 is integrated over some $\Omega \subset \mathbb{R}^d$. Assuming that Ω is constant over time and by applying the *Divergence Theorem*, with $\hat{\mathbf{n}} = [\hat{n}_1, \hat{n}_2, \ldots, \hat{n}_d]^T$ as the outward unit normal to the boundary $\partial\Omega$, it follows from theorem 3 in appendix B that

$$\int_\Omega \mathbf{u}_t \, d\mathbf{x} + \int_\Omega \nabla \cdot \mathbf{f}^c(\mathbf{u}) \, d\mathbf{x} = \frac{d}{dt}\int_\Omega \mathbf{u} \, d\mathbf{x} + \int_{\partial\Omega} \mathbf{f}^c(\mathbf{u}) \cdot \hat{\mathbf{n}} \, ds = \mathbf{0}, \qquad (2.25)$$

where ds is an infinitesimal arc length element counted counter clockwise around Ω. This can be interpreted as the total time variation of the l quantities $u_1, \ldots, u_l$ over the domain Ω and the flux in and out of the volume must balance to zero.

An important aspect of the conservative matrix form (2.17) is its *conservative property*, meaning it may be expressed as (2.25). This follows from the property $\nabla \cdot [\mathbf{A}\mathbf{u}, \mathbf{B}\mathbf{u}]^T = \mathbf{A}\mathbf{u}_x + \mathbf{B}\mathbf{u}_y$ which justifies

$$\int_\Omega \mathbf{u}_t \, d\mathbf{x} + \int_\Omega \nabla \cdot [\mathbf{A}\mathbf{u}, \mathbf{B}\mathbf{u}]^T \, d\mathbf{x} = \frac{d}{dt}\int_\Omega \mathbf{u} \, d\mathbf{x} + \int_{\partial\Omega} [\mathbf{A}\mathbf{u}, \mathbf{B}\mathbf{u}]^T \cdot \hat{\mathbf{n}} \, ds = \mathbf{0}. \qquad (2.26)$$

Observe that the primitive form and the symmetrised primitive form do not have this property [7].

2.4 Characteristics

Characteristic variables play an important role for hyperbolic systems, as will be shown a characteristic representation states the number of BCs that are needed. In addition it may also decouple the equations and present a more preferable system. However, this requires that the system is *diagonalisable*, see definition 2, which allows the formulation of characteristic variables:

Definition 5. If a $\mathbf{K}$ in the sense of definition 2 exists, then the characteristic variables are defined as

$$\mathbf{K}^{-1}\mathbf{u} = \mathbf{u}_c, \tag{2.27}$$

with $\mathbf{u}$ from (2.3).

Characteristics are curves, or rays along which a solution $\mathbf{u}(\mathbf{x}, t)$ to a hyperbolic PDE propagates unchanged from its ICs, meaning that the PDE becomes an ODE along each ray [9]. Values at the boundary are also carried into the domain along characteristics, meaning the BCs heavily influence the solution. Instead of a formal definition, two examples will introduce the concept.

Example 1. Consider the one-dimensional linear advection equation with given initial conditions, defined on the domain $-\infty < x < \infty$, $t \geq 0$:

$$\begin{aligned} u_t + au_x &= 0, \\ u(x, 0) &= u_0(x). \end{aligned} \tag{2.28}$$

It is easily checked that $u(x, t) = u_0(x - at)$ is a solution to (2.28). Now consider the characteristics: each line $x_0 = x - at$ defines a curve in the $x-t$ plane that solves the ODE $x'(t) = a$, $x(0) = x_0$. These lines are the *characteristics* of the equation (2.28), which means that the solution $u_0(x - at)$ is constant along each line given by $x(t) = x_0 + at$. This is shown by differentiating $u(x(t), t)$ along one of the characteristics:

$$\begin{aligned} \frac{d}{dt}u(x(t), t) &= \frac{\partial}{\partial t}u(x(t), t) + \frac{\partial}{\partial x}u(x(t), t)x'(t) \\ &= u_t + au_x = 0. \end{aligned} \tag{2.29}$$

The direction and speed of the propagation are determined by size and sign of a, where $a < 0$ gives to the left and $a > 0$ to the right.

For a hyperbolic system of PDEs the eigenvalues of the matrix $\mathbf{A}$, from definition 3, give the direction and propagation speed of the characteristics [9]. Consider the following example.

Example 2. Let $\mathbf{A} \in \mathbb{R}^{l \times l}$ be a constant and diagonalisable matrix and consider the following one-dimensional linear equation with given ICs, defined on the same domain as the previous example:

$$\mathbf{u}_t + \mathbf{A}\mathbf{u}_x = \mathbf{0}, \tag{2.30}$$

which may be written in characteristic form by diagonalisation as in definition 5:

$$\mathbf{u}_{ct} + \mathbf{\Lambda}\mathbf{u}_{cx} = \mathbf{0}. \tag{2.31}$$

Since $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_l)$ equation (2.30) is reduced to a decoupled system of one-dimensional linear advection equations

$$\frac{\partial u_{ci}}{\partial t} + \lambda_i \frac{\partial u_{ci}}{\partial x} = 0, \quad i = 1, 2, \ldots, l, \tag{2.32}$$

where λ_i has the role of a in (2.28).

The signs of the eigenvalues of the system matrix $\mathbf{A}$ determine the direction of propagation for the characteristics. The ingoing characteristics, corresponding to $\lambda_i < 0$, propagate into the computational domain with data from the boundary. The outgoing characteristics, $\lambda_i > 0$ propagate out through the computational domain and carry information about the state in the domain. Therefore the ingoing characteristics must be prescribed, which is done by giving u_{ci} as BCs for those i with $\lambda_i < 0$. The outgoing on the other hand, should not be prescribed, since they are part of the solution that is not known a priori [2]. How these are dealt with is further explained in section 3.4.

The characteristics for the Euler equations intersect after a finite time, and thereby contradict the definition of characteristics. These intersections give rise to discontinuities that are present in the physics as well, such as shocks and the breaking of waves [10, 11].

Chapter 3

Space-discretisation methods

When solving PDEs by numerical means, one must consider some sort of discretisation. There are numerous methods available, often supported by convergence theory, that produce a discrete solution represented on some sort of grid. For conservation laws, the FVM is preferable and will be the space-discretisation method of choice in the *method of lines*. The node-centred FVM is a specific FVM that acts upon *control volumes*, which are the cells on a new grid, called the *dual grid*, constructed from an unstructured triangular grid referred to as the *primary grid*. The character of the BCs, that can be implemented in numerous ways for numerical methods, has a huge impact on the solution as will be discussed from the view point of flow problems.

3.1   The method of lines

One numerical method for solving PDEs is the method of lines. It consists of discretising the system of PDEs in all dependent variables, except for the one with a given IC. This variable is for now left continuous and is usually the time-variable, while the spatial variables are discretised. This means that expressions such as differentiating with respect to space reduce to algebraic expressions. Therefore, the former system of PDEs can now be considered as an initial value problem for a system of ODEs with respect to time. There exist several solution methods for systems of ODEs that reduce them to a system consisting entirely of algebraic expressions, which may be solved with an iterative method. Note that not only discretisation errors with respect to time and space are introduced, but also iteration errors when solving the system of algebraic equations [2, 5]. This thesis will focus on the discretisation of the space part in the method of lines procedure.

3.2 The grid

In order to compute a numerical solution to a given system of PDEs, the computational domain Ω must be discretised into cells, due to the finite storage space of a computer. The discretised computational domain will be referred to as $\mathcal{T}(\Omega)$, which is the set containing the cell-elements Ω_i. There are mainly two kind of grids suitable for implementation of numerical methods: *structured* and *unstructured*, where the former has high regularity. Unstructured grids are more adaptable to complex geometries and will be considered in this thesis [12].

More formally, the computational domain $\Omega \in \mathbb{R}^d$ is discretised into a grid $\mathcal{T}(\Omega)$. The grid consists of cells Ω_i, each with a centre of mass $\mathbf{d}_i$. Let $\mathcal{N}(\Omega_i)$ and $\mathcal{E}(\Omega_i)$ be the set of nodes $\mathbf{c}_i$ and edges E_{ij} corresponding to the cell Ω_i, then the elements $\Omega_i \in \mathcal{T}(\Omega)$ are such that

- $\overline{\Omega} = \bigcup_i \overline{\Omega}_i$,

- $\overline{\Omega}_i \cap \overline{\Omega}_j = \{\mathbf{x}\} \Rightarrow \mathbf{x} \in \mathcal{N}(\Omega_i) \cap \mathcal{N}(\Omega_j)$,

- $\overline{\Omega_i} \cap \overline{\Omega_j} \neq \varnothing$ and $\overline{\Omega_i} \cap \overline{\Omega_j} \neq \{\mathbf{x}\} \Rightarrow \overline{\Omega}_i \cap \overline{\Omega}_j = E_{ij}$ with $E_{ij} = \mathcal{E}(\Omega_i) \cap \mathcal{E}(\Omega_j)$.

Cells sharing at least one edge with the computational domain's boundary will be referred to as *boundary cells* and the edge will be called the *exterior edge*. Other cells will be referred to as interior cells. The notation will later be relaxed: $\mathcal{N}(\Psi)$ and $\mathcal{E}(\Psi)$ may sometimes be the sets of indices for the nodes and edges of Ψ, respectively. It will be obvious to the reader from the context what the notation represents.

3.2.1 The unstructured node-centred grid

The cells of an unstructured two-dimensional grid are here triangles, with some kind of shape condition, e.g. they should not have too small angles. In this thesis the *unstructured node-centred grid* will be considered. Obviously it is an unstructured grid consisting of triangles, but the information, or where there exists a discrete solution, are at the vertices of the triangles. The node-centred FVM, presented in the next section, gives a discretised solution that corresponds to averages over cells. But the current grid, also called the *primary grid* [2], is not used. Instead each node on the primary grid is paired with a cell Ω_i from the new set of cells $\mathcal{T}(\Omega)$ called the *dual grid*. First the barycentres, or mass centres $\mathbf{d}$, of the triangles on the primary grid are found, then they are connected with straight lines to the midpoints of the sides of the corresponding triangle-cell on the primary grid. Now each node $\mathbf{c}_i$, originally a vertex

on the primary grid, is the centre of one polygon-shaped cell Ω_i on the dual grid, created by the connection of the mass centres and the sides of the primary cells, see figures 3.1 and 3.2. Each dual cell has a positively oriented $\mathcal{C}^1$-boundary with an outward facing unit normal $\hat{\mathbf{n}}_i$. It is this grid the node-centred FVM acts upon.
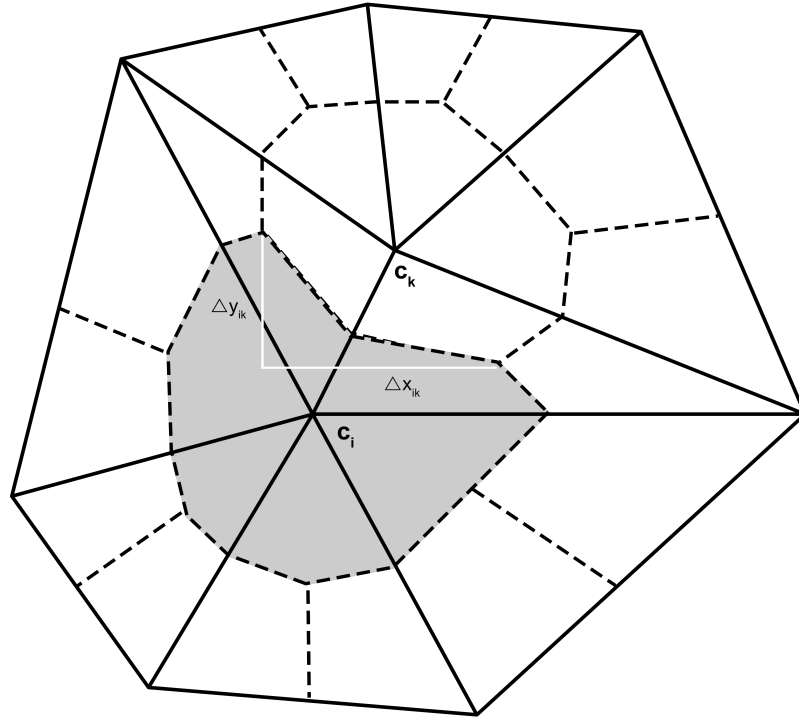

FIGURE 3.1: Example of a dual grid. The primary grid consists of the solid lines and the dual grid of the dotted lines. The shadowed area represents a cell on the dual grid and the white lines denote the components of E_{ij}. Made by M. C. Roberts. 2014.

3.3 The finite-volume method

The FVM is a space-discretisation method for the spatial domain, closely connected to flow problems [2]. The main concept is the division of the computational domain into finite volumes, also called control volumes, on which conservation laws must hold. Since the Euler equations are derived from conservation laws, the FVM is highly suitable. For more ingoing treatment of the FVM, please consult [3].
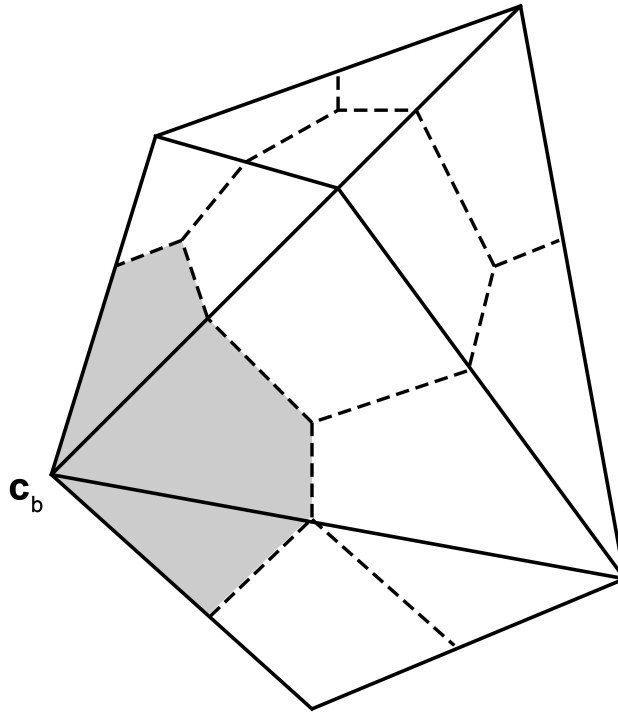
FIGURE 3.2: Example of a dual grid close to the boundary. Made by M. C. Roberts. 2014.

3.3.1 Space-discretisation

On each cell Ω_i of the dual grid $\mathcal{T}(\Omega)$, $\mathbf{u}(\cdot, t)$ is approximated as a constant $\bar{\mathbf{u}}_i(t) \in \mathbb{R}^l$ equal to the cell average of $\mathbf{u}$ over the cell Ω_i at time t:

$$\bar{\mathbf{u}}_i(t) := \frac{1}{|\Omega_i|} \int\limits_{\Omega_i} \mathbf{u}(\mathbf{x}, t) \, d\mathbf{x}, \tag{3.1}$$

where $|\Omega_i|$ denotes the area of Ω_i. Recall the integral form of the Euler equations (2.25), but replace the continuous solution $\mathbf{u}$ with (3.1). Integration over Ω_i instead of Ω yields the cell-wise discretisation

$$|\Omega_i| \frac{d\bar{\mathbf{u}}_i(t)}{dt} + \oint\limits_{\partial\Omega_i} \mathbf{f}^c(\bar{\mathbf{u}}_i) \cdot \hat{\mathbf{n}} \, ds = \mathbf{0}, \quad \forall \Omega_i \in \mathcal{T}(\Omega), \tag{3.2}$$

assuming that the cells do not change over time. Approximating the boundary integral of the convective fluxes in (3.2) is not as straight forward as taking the mean over the control volume. The reason is that the discrete solution $\bar{\mathbf{u}}(\mathbf{x}, t)$ is considered constant over the control volume Ω_i, and therefore the approximation is not continuous over $\partial\Omega_i$,

unless the neighbouring cells satisfy $\bar{\mathbf{u}}_i(t) = \bar{\mathbf{u}}_j(t), \forall \Omega_i, \Omega_j \in \mathcal{T}(\Omega)$. This would imply a constant solution, which generally is not the case. Hence a discrete approximation of the flux over the boundary of the cell is required [3].

Since the cells Ω_i are polygons the integral of the flux over $\partial \Omega_i$ can be rewritten as

$$\oint_{\partial \Omega_i} \mathbf{f}^c(\bar{\mathbf{u}}_i) \cdot \hat{\mathbf{n}} \, ds = \sum_{\mathcal{E}(\Omega_i)} \int_{E_{ij}} \mathbf{f}^c(\bar{\mathbf{u}}_i) \cdot \hat{\mathbf{n}} \, ds, \quad \forall \Omega_i \in \mathcal{T}(\Omega), \tag{3.3}$$

where the shortened notation $\mathcal{E}(\Omega_i)$ instead of $E_{ij} \in \mathcal{E}(\Omega_i)$ for the sum's limit is used. Note that E_{ij} is the edge the cells Ω_i and Ω_j share and is therefore piecewise linear, see figure 3.1. Now a method for approximation of the piecewise integration is needed. Assume there exists a numerical flux function $\mathbf{f}^N$ that approximates the flux over an edge E_{LR}, based on the states $\bar{\mathbf{u}}_L$ and $\bar{\mathbf{u}}_R$ on the left and the right of the edge. The numerical flux function must satisfy the following [2]:

1. $\mathbf{f}^N(\bar{\mathbf{u}}_L, \bar{\mathbf{u}}_R, \hat{\mathbf{n}}_{E_{LR}}) = -\mathbf{f}^N(\bar{\mathbf{u}}_L, \bar{\mathbf{u}}_R, -\hat{\mathbf{n}}_{E_{LR}})$,

2. $\mathbf{f}^N = \mathbf{f}^N(\bar{\mathbf{u}}_L, \bar{\mathbf{u}}_R, \hat{\mathbf{n}}_{E_{LR}})$ is consistent, meaning it is Lipschitz continuous in the first two arguments and that $\mathbf{f}^N(\mathbf{u}, \mathbf{u}, \hat{\mathbf{n}}_{E_{LR}}) = \mathbf{f}^c(\mathbf{u}) \cdot \hat{\mathbf{n}}$.

The last requirement $\mathbf{f}^N(\mathbf{u}, \mathbf{u}, \hat{\mathbf{n}}_{E_{LR}}) = \mathbf{f}^c(\mathbf{u}) \cdot \hat{\mathbf{n}}$ means the solution to the Euler problem as argument will give a numerical flux function equal to the flux $\mathbf{f}^c(\mathbf{u})$ [2].

Let $\mathbf{f}^N(\bar{\mathbf{u}}_i, \bar{\mathbf{u}}_j, \hat{\mathbf{n}}_{E_{ij}})$ replace $\mathbf{f}^c(\bar{\mathbf{u}}) \cdot \hat{\mathbf{n}}$ in (3.3). The piecewise integral on the right hand side is approximated with a Gaussian quadrature, with a Gauss point at the middle of each edge E_{ij} [2]. With $|E_{ij}| = \int_{E_{ij}} ds$ the final form that satisfies the method of lines procedure is obtained:

$$\frac{d\bar{\mathbf{u}}_i(t)}{dt} = -\frac{1}{|\Omega_i|} \sum_{\mathcal{E}(\Omega_i)} |E_{ij}| \, \mathbf{f}^N(\bar{\mathbf{u}}_i, \bar{\mathbf{u}}_j, \hat{\mathbf{n}}_{E_{ij}}), \quad \forall \Omega_i \in \mathcal{T}(\Omega). \tag{3.4}$$

Following the method of lines expression (3.4) it is clear how the numerical algorithm will proceed: Some time marching method will replace the differentiation with respect to time on the left-hand side of (3.4). Then an update is obtained by the sum of all numerical flux functions over all edges. The final solution may be modified with so called *piecewise reconstruction*, which gives the FVM an order of 2 [2].

3.3.2   Convergence of the finite-volume method

The intention of this subsection is to present a brief and non-strict introduction of the concepts behind existence, uniqueness and convergence for the FVM. This part simply functions as orientation and should be read as such.

The FVM seeks for weak solutions, briefly mentioned in section 2.3, that are piecewise classical solutions with discontinuities separated by smooth regions. If the initial data is bounded in $L^\infty(\Omega)$-norm and $\mathbf{u}$ is a classical solution on each cell Ω_i, then $\mathbf{u}$ is a weak solution to the Euler equations if it satisfies the *Rankine-Hugoniot jump condition*. This conditions ensures that mass, momentum and energy are conserved across the discontinuities and thus the conservation laws still hold. The weak solutions are however not unique, which is remedied by applying an additional *entropy condition* [11].

The convergence theory of the FVM for nonlinear systems is very limited. It is possible to show that a scalar hyperbolic conservation law converges, thus a linear system can be diagonalised and become a system of scalar equations. The same does not hold for nonlinear systems of equations, where diagonalisation does not decouple the system. Hence the convergence theory for nonlinear systems is very limited [2].

3.4   Boundary conditions

For flow problems, the BCs give information about how the fluid reacts at the edges of the computational domain or when in contact with obstacles in its path. Even small objects can have huge influence on the fluid's behaviour downstream and huge gradients are often present in the direct vicinity of the obstacles. Therefore high grid resolution and adequate modelling of the BCs are crucial in order to simulate a physical fluid's motion. Further, one often requires that the BCs are bounded in $L^2(\Omega)$-norm and that the ICs state the same conditions at the boundary of the domain as the BCs do at the initial time.

A crude division of BCs gives two groups: physical BCs and numerical BCs. For the physical BCs, the physics behind the given problem are of great importance, since one aims to model some physical phenomena or interaction. An example is the modelling of the interface between metal and air for the system of conservation laws known as Maxwell's equations. The numerical BCs on the other hand are purely numerical constructions needed for limiting an infinite domain. These boundaries are often placed at such a distance from fluid-object interactions that they may be considered to be in the undisturbed free flow, also referred to as the *far-field*. These boundaries need to

be modelled so that the flow quantities can leave the domain without being reflected or create other numerical artifacts. That is, behave as there was no boundary at all. The corresponding situation for Maxwell's equations would be electromagnetic waves propagating into the distance from an antenna.

The numerical boundary can also be put closer to the object-fluid interaction, which can be considered as the opposite of the undisturbed free flow and is called the *near-field*. The same conditions as for modelling a boundary in the far-field applies, but larger gradients are expected. As an example, consider the wing-profile in figure 3.3A which should be given adequate physical BCs that prevent the flow of air to pass through it. To preform computations the domain must be made finite by limiting it with some numerical boundary. In figure 3.3B the finite computational domain is enclosed by the circle, which is given numerical BCs to model the near-field.

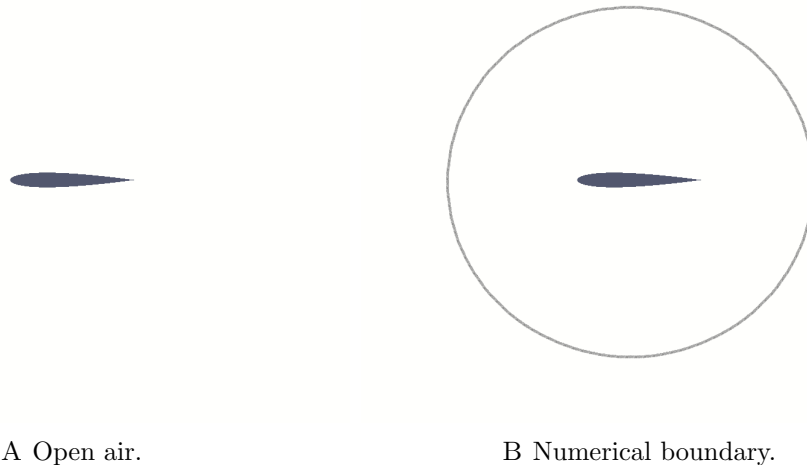A Open air.   B Numerical boundary.

FIGURE 3.3: A two-dimensional wing-profile with and without a numerical boundary.

BCs can be imposed numerically in a numerous of different ways. One approach that will be considered is *prescribed fluxes*, where some $\tilde{\mathbf{u}}$ is choosen based on the BCs or other data if it is a numerical boundary. Then the flux for an exterior edge in (3.4) is evaluated with a numerical flux function $\mathbf{f}^N$, with $\tilde{\mathbf{u}}$ and the local numerical solution $\bar{\mathbf{u}}_i$ as arguments [2].

In [1] the method of *strongly imposed* BCs is also considered, where a numerical solution for the boundary nodes is not obtained. Instead, the boundary values are replaced with some values given as BCs, hence the BCs will be satisfied exactly. This method will not be discussed further in this paper, as the method of weakly imposed prescribed fluxes will be used as a comparison to weakly imposed characteristic BCs, which is described below. The weak implementation of BCs is explained in chapter 4.

3.4.1 Fixed wall

To model solid obstacles the *fixed wall* BC is used, which may look different depending on the system of equations. For the Euler equations the fixed wall boundary condition is modelled with a *slip condition*, meaning that no mass should pass through the obstacle. Mathematically it is expressed as $v_n = \boldsymbol{v} \cdot \hat{\mathbf{n}} = 0$. Note that no further conditions are given for the solution $\mathbf{u}$ at a fixed wall. The slip condition is a physical fact for non-permeable materials, a clear distinction from numerical constructions such as the near- and far-field.

The fixed wall BC is imposed by inserting $v_n = 0$ into $\mathbf{f}^c(\mathbf{u}) \cdot \hat{\mathbf{n}}$ in (3.2), giving that the contribution to $|\Omega_i|\frac{d\mathbf{u}_i(t)}{dt}$ in (3.4) from the fixed wall boundary is

$$\mathbf{f}^c(\mathbf{u}) \cdot \hat{\mathbf{n}} = [0, n_1 p, \dots, n_d p, 0]^T. \tag{3.5}$$

The value of p depends on the local numerical solution. The flux given by (3.5) then replaces the numerical flux function $\mathbf{f}^N$ at boundaries with the fixed wall BC.

3.4.2 Inflow and outflow

The numerical boundary needs to be divided in sections that are given inflow- or outflow BCs. These boundaries will be considered to be in the far-field, meaning that reference values, denoted with ˜, for the density, velocities and pressure will be used as BCs to model the undisturbed flow. The near-field can also be implemented in this way, but the accuracy may be suffering since the choice of reference values becomes more difficult.

One way to impose the BCs is with the concept of characteristics, see section 2.4. By prescribing the *ingoing* characteristics the information on the boundaries is transferred into the computational domain. The *outgoing* characteristics should not be prescribed, since that information is coming from the computational domain and thus a part of the solution, which is not known a priori. The outgoing characteristics are therefore given values based on the current local numerical solution. The total flux over an exterior edge is a combination of in- and outgoing characteristic and this way of imposing BCs is called characteristic BCs.

The total number of characteristics for the Euler equations is equal to the number of eigenvalues λ of the matrix $\mathbf{A}dy - \mathbf{B}dx$ [9]. In two dimensions they are:

$$\lambda_1 = v_n - c,$$
$$\lambda_2 = v_n,$$
$$\lambda_3 = v_n,$$
$$\lambda_4 = v_n + c,$$

derived in appendix A. Recall that negative eigenvalues correspond to ingoing characteristics and vice versa. For supersonic flows $(Ma > 1)$ the fluid travels faster than information in the medium, meaning the that the information is transported in one direction only [2]. Mathematically this is a consequence of $c < |v_n|$ with $v_n < 0$, resulting in $\lambda = v_n + c < 0$ and the number of outgoing characteristics being zero. Thus the Mach number is of great importance as it governs the signs of the eigenvalues, hence the number of ingoing and outgoing characteristics, see table 3.1 for a summation.

	$Ma < 1$	$Ma > 1$
Inflow	3	4
Outflow	1	0

TABLE 3.1: The number of ingoing characteristics at an inflow boundary for the two-dimensional Euler equations for different Mach numbers.

Chapter 4

Weakly imposed boundary conditions

The goal is to obtain a *strictly stable* approximation, meaning that a discrete solution does not grow over time. A theorem for strict stability for the approximation given by the node-centred FVM applied to the main problem (2.1), with weakly imposed characteristic BCs, is eventually presented. This chapter follows the outline of [1], but offers a more detailed description of the analysis.

4.1 Analysis of the node-centred finite-volume method with weak boundary conditions

This thesis defines strict stability in the same way as the article [1]:

Definition 6. (Strict stability) An approximation is called *strictly stable* if

$$\frac{d}{dt}\|\bar{\mathbf{u}}\|_\Omega^2 \leq 0, \tag{4.1}$$

where $\|\cdot\|_\Omega$ is some norm and $\bar{\mathbf{u}}$ is an approximate solution.

The expression (4.1) can be seen as a rate of change of energy and as previously mentioned stability is one of two requirements for convergence. For the continuous case the $L^2(\Omega)$-norm is used:

$$\|\mathbf{u}\|_\Omega^2 = \int_\Omega \mathbf{u}^T \mathbf{u}\, dxdy \tag{4.2}$$

and for the discrete case consider the discrete equivalence

$$\|\bar{\mathbf{u}}\|_{\mathbf{C}}^2 = \bar{\mathbf{u}}^T \mathbf{C} \bar{\mathbf{u}}, \tag{4.3}$$

where $\mathbf{C}$ is a diagonal matrix with the areas $|\Omega_i|$, $\Omega_i \in \mathcal{T}(\Omega)$, as elements.

To obtain a strictly stable approximation in sense of definition 6 the following approach is used:

1. An estimate of the form (4.1) for a continuous solution to the main problem (2.1) is obtained, which gives an idea of what to aim for in the discrete case.

2. The main problem (2.1) is approximated with the node-centred FVM.

3. Weakly imposed characteristic BCs allows the formulation of a theorem for strict stability.

4.1.1 The continuous problem

Consider the main problem, repeated here for convenience: $\mathbf{u} \in \mathbb{R}^l$ and $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{l \times l}$

$$\mathbf{u}_t + \mathbf{A}\mathbf{u}_x + \mathbf{B}\mathbf{u}_y = \mathbf{0}, \quad (x, y) \in \Omega \subset \mathbb{R}^2, \quad t \in \mathbb{R}_0^+, \tag{4.4}$$

where $\mathbf{A}$ and $\mathbf{B}$ are constant, square and symmetric with suitable BCs and ICs. The BCs are such that $(\mathbf{A}dy - \mathbf{B}dx)$ is a positive semi-definite matrix on $\partial\Omega$ at any time t.

Let $\mathbf{u}$ be a solution to the main problem (4.4), insert into (4.2) and differentiate with respect to time:

$$\begin{aligned} \frac{d}{dt}\|\mathbf{u}\|_{\Omega}^2 &= \int_{\Omega} \frac{\partial}{\partial t}(\mathbf{u}^T\mathbf{u}) \, dx \, dy = \int_{\Omega} 2\,\mathbf{u}^T\mathbf{u}_t \, dx \, dy = \int_{\Omega} 2\,\mathbf{u}^T(-\mathbf{A}\mathbf{u}_x - \mathbf{B}\mathbf{u}_y) \, dx \, dy \\ &= \int_{\Omega} \frac{\partial}{\partial x}(-\mathbf{u}^T\mathbf{A}\mathbf{u}) - \frac{\partial}{\partial y}(\mathbf{u}^T\mathbf{B}\mathbf{u}) \, dx \, dy = -\oint_{\partial\Omega} \mathbf{u}^T\mathbf{A}\mathbf{u} \, dy + \oint_{\partial\Omega} \mathbf{u}^T\mathbf{B}\mathbf{u} \, dx \\ &= -\oint_{\partial\Omega} \mathbf{u}^T(\mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\hat{\mathbf{y}}) \cdot \hat{\mathbf{n}} \, \mathbf{u} \, ds, \end{aligned} \tag{4.5}$$

where $\hat{\mathbf{n}} = [dy, -dx]^T/ds$ and $ds = \sqrt{dx^2 + dy^2}$. In the first step, Ω is assumed to be constant over time and Leibniz's rule of integration is used, see (3) in appendix B. Thereafter (4.4) is inserted and rewritten, followed by applying *Green's Theorem*, see theorem 2 in appendix B. Observe that the expression $(\mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\hat{\mathbf{y}}) \cdot \hat{\mathbf{n}}$ is not well-defined since the matrix products $\mathbf{A}\hat{\mathbf{x}}$ and $\mathbf{B}\hat{\mathbf{y}}$ require $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{2 \times 2}$, which generally will not

be the case. The expression will still be used to be consistent with [1], therefore the matrices $\mathbf{A}$ and $\mathbf{B}$ are treated as scalars when matrix multiplication is not well-defined:

$$(\mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\hat{\mathbf{y}}) \cdot \hat{\mathbf{n}} = \left(\begin{bmatrix} \mathbf{A} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \mathbf{B} \end{bmatrix} \right) \cdot \hat{\mathbf{n}}. \tag{4.6}$$

Since $(\mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\hat{\mathbf{y}}) \cdot \hat{\mathbf{n}} = (\mathbf{A}dy - \mathbf{B}dx)$ it is positive semi-definite and satisfies

$$\mathbf{u}^T (\mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\hat{\mathbf{y}}) \cdot \hat{\mathbf{n}}\, \mathbf{u} \geq 0, \quad (x,y) \in \partial\Omega,\, t \in \mathbb{R}_0^+, \tag{4.7}$$

due to the given BCs for the main problem (4.4). One obtaines

$$\frac{d}{dt}\|\mathbf{u}\|_\Omega^2 = - \oint\limits_{\partial\Omega} \mathbf{u}^T (\mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\hat{\mathbf{y}}) \cdot \hat{\mathbf{n}}\, \mathbf{u}\, ds \leq 0, \tag{4.8}$$

which is the continuous equivalence of the estimate (4.1). More specific conditions for the positive semi-definite property of $(\mathbf{A}dy - \mathbf{B}dx)$ are given later in this chapter. The goal now is to achieve the discrete equivalence of boundary integral in (4.8). This form is clearly attractive since it only depends on the flux over the boundary, making the treatment of the BCs crucial.

4.1.2 The semi-discrete problem

Before considering an approximation of the main problem (2.1), the sorting of the vector representing the discrete solution must be discussed. Let $\bar{\mathbf{u}}$ be the projection of $\mathbf{u}$ on a given dual grid $\mathcal{T}(\Omega)$ with n grid points. The vector $\bar{\mathbf{u}}$ is of length $N = ln$, arranged in such a fashion that the first n elements are the discretisation of the first variable in $\mathbf{u}$ on $\mathcal{T}(\Omega)$, the elements $n+1, \ldots, 2n$ are the discretisation of the second variable and so on up until the elements $nl - n, \ldots, nl$ which are the discretisation of the last variable. Let $\bar{\mathbf{u}}^k$ denote the discretisation of the k:th variable in $\mathbf{u}$, then

$$\bar{\mathbf{u}} = \begin{bmatrix} \bar{\mathbf{u}}^1 \\ \bar{\mathbf{u}}^2 \\ \vdots \\ \bar{\mathbf{u}}^l \end{bmatrix}. \tag{4.9}$$

The elements of $\bar{\mathbf{u}}$ are cell averages of the elements in $\mathbf{u}$, which motivates the following notation: $\bar{\mathbf{u}}_i$ is a vector of length l and corresponds to the average of $\mathbf{u}$ over the cell Ω_i. That is, $\bar{\mathbf{u}}_i = [\bar{\mathbf{u}}(i), \bar{\mathbf{u}}(n+1+i), \ldots, \bar{\mathbf{u}}(n+l-n+i)]^T$. In general the overbar $^-$ denotes a discretisaion on $\mathcal{T}(\Omega)$. As will be seen later in this chapter, this specific way of sorting

$\bar{\mathbf{u}}$ yields intuitive expressions for the approximation of the main problem (2.1) and the discrete equivalence of (4.8).

The goal is to obtain the semi-discretisation

$$(\mathbf{I} \otimes \mathbf{P})\bar{\mathbf{u}}_t + (\mathbf{A} \otimes \mathbf{Q}_x)\bar{\mathbf{u}} + (\mathbf{B} \otimes \mathbf{Q}_y)\bar{\mathbf{u}} + \mathbf{b} = \mathbf{0}, \tag{4.10}$$

of the main problem (2.1), where $\otimes$ is the *Kronecker product*, see definition 8 in appendix B. The matrices $\mathbf{Q}_x, \mathbf{Q}_y \in \mathbb{R}^{n \times n}$ are discrete differential operators and the vector $\mathbf{b} \in \mathbb{R}^N$ handles the BCs. These terms are properly derived below, and as will be shown it is crucial that $\mathbf{Q}_x + \mathbf{Q}_x^T$ and $\mathbf{Q}_y + \mathbf{Q}_y^T$ are diagonal matrices in order to obtain strict stability.

To obtain the semi-discretisation (4.10), the node-centred FVM from chapter 2 is used. Integration of the main problem (4.4) over a control volume $\Omega_i \in \mathcal{T}(\Omega)$ can be written as

$$\frac{d\bar{\mathbf{u}}_i(t)}{dt}|\Omega_i| + \mathbf{A} \oint_{\partial\Omega_i} \mathbf{u}\, dy - \mathbf{B} \oint_{\partial\Omega_i} \mathbf{u}\, dx = \mathbf{0}, \tag{4.11}$$

where the boundary line integrals need to be approximated. This is done, as shown in section 3.3.1, by approximating the flux over each edge E_{ij} of the polygon Ω_i with a numerical flux function $\mathbf{f}^N(\bar{\mathbf{u}}_i, \bar{\mathbf{u}}_j, \hat{\mathbf{n}}_{ij}(k))$. Observe that the line integrals are not integrated with respect to ds, but rather dx and dy, resulting in the argument $\hat{\mathbf{n}}_{ij}(k)$ where $\hat{\mathbf{n}}_{ij} = [\Delta y_{ij}, -\Delta x_{ij}]^T/\Delta s_{ij}$ with $\Delta s_{ij} = \sqrt{\Delta y_{ij}^2 + \Delta x_{ij}^2}$. The scalars Δx_{ij} and Δy_{ij} are x- and y-components of the edge vector E_{ij}, which in turn is piecewise linear. The integrals are as before approximated with Gaussian quadratures, with a Gauss point at the middle of the edges E_{ij}. Thus expression (4.11) can be rewritten as

$$\frac{d\bar{\mathbf{u}}_i(t)}{dt}|\Omega_i| + \mathbf{A} \sum_{j \in \mathcal{E}(\Omega_i)} \mathbf{f}^N(\bar{\mathbf{u}}_i, \bar{\mathbf{u}}_j, \hat{\mathbf{n}}_{ij}(1))\Delta s_{ij} - \mathbf{B} \sum_{j \in \mathcal{E}(\Omega_i)} \mathbf{f}^N(\bar{\mathbf{u}}_i, \bar{\mathbf{u}}_j, \hat{\mathbf{n}}_{ij}(2))\Delta s_{ij} = \mathbf{0}. \tag{4.12}$$

In [1] the analysis is performed with the numerical flux function

$$\mathbf{f}^N(\bar{\mathbf{u}}_i, \bar{\mathbf{u}}_j, \hat{\mathbf{n}}_{ij}(k)) = \frac{\bar{\mathbf{u}}_i + \bar{\mathbf{u}}_j}{2}\hat{\mathbf{n}}_{ij}(k), \quad k = 1, 2, \tag{4.13}$$

which means that the x- or y-directed flux over an edge is the mean value of the states on each side of the edge, times the normal in that direction. This specific choice of numerical flux function will give that $\mathbf{Q}_x + \mathbf{Q}_x^T$ and $\mathbf{Q}_y + \mathbf{Q}_y^T$ are diagonal.

4.1.3 Weak implementation of boundary conditions

The flux in (4.12) will be evaluated for three different cells. In each case the vector $\mathbf{Z}_i \in \mathbb{R}^n$, associated with the cell Ω_i, is constructed and replaces the numerical flux function. These vectors will later be the rows of the matrices $\mathbf{Q}_x$ and $\mathbf{Q}_y$. The method will be explained assuming $l = 1$ and will only treat the term containing the factor $\mathbf{A}$. The term with factor $\mathbf{B}$ is handled analogously and once the scalar case is fully explained, the case for any $l \in \mathbb{N}$ will follow.

Consider the x-directed flux, corresponding to the term with A, in (4.12) for an interior cell $\Omega_i \in \mathcal{T}(\Omega)$:

$$flux = A \sum_{j \in \mathcal{E}(\Omega_i)} \frac{\bar{u}_i + \bar{u}_j}{2} \Delta y_{ij} = A \left(\bar{u}_i \sum_{j \in \mathcal{E}(\Omega_i)} \frac{\Delta y_{ij}}{2} + \sum_{j \in \mathcal{E}(\Omega_i)} \frac{\Delta y_{ij}}{2} \bar{u}_j \right). \tag{4.14}$$

The sum of Δy_{ij} over $\partial \Omega_i$ is identically equal to zero, since it is the line integral over a closed curve [13]. The flux for an interior cell is thus the weighted sum of the neighbouring states, with the corresponding halved arclength as weights:

$$flux = A \sum_{j \in \mathcal{E}(\Omega_i)} \frac{\Delta y_{ij}}{2} \bar{u}_j. \tag{4.15}$$

The sum can be replaced with the vectors $\bar{\mathbf{u}}$ and $\mathbf{Z}_i \in \mathbb{R}^{1 \times n}$ with the weights as elements:

$$\mathbf{Z}_i(j) = \frac{\Delta y_{ij}}{2}, \quad \mathbf{Z}_j(i) = -\mathbf{Z}_i(j) = -\frac{\Delta y_{ij}}{2}, \quad \mathbf{Z}_i(i) = 0 \tag{4.16}$$

for neighbouring cells Ω_j. For non-neighbouring cells $\mathbf{Z}_i(j) = 0$.

Now consider an exterior cell Ω_b, for some $b \in \mathcal{N}(\mathcal{T}(\Omega))$, with no BCs for the given exterior edge. The flux over the interior edges of the cell will be calculated as above, and the flux over the exterior edge will be approximated with the nodal value $\bar{u}_b$ of Ω_b times the corresponding length Δy_{bb}:

$$flux = A \left(\sum_{j \in \mathcal{E}(\Omega_b) \setminus b} \frac{\bar{u}_b + \bar{u}_j}{2} \Delta y_{bj} + \bar{u}_b \Delta y_{bb} \right). \tag{4.17}$$

Note that the sum is not over a closed loop, giving

$$\sum_{j \in \mathcal{E}(\Omega_b) \setminus b} \Delta y_{bj} = -\Delta y_{bb} \tag{4.18}$$

and therefore expression (4.17) for the flux can be written as

$$flux = A \left(\sum_{j \in \mathcal{E}(\Omega_b) \backslash b} \frac{\Delta y_{bj}}{2} \bar{u}_j + \frac{\Delta y_{bb}}{2} \bar{u}_b \right). \tag{4.19}$$

The elements of the corresponding vector $\mathbf{Z}_b$ are as for the interior cells, with the exception that

$$\mathbf{Z}_b(b) = \frac{\Delta y_{bb}}{2}. \tag{4.20}$$

With the introduction of $\mathbf{Z}_i$, both (4.17) and (4.14) can be expressed as

$$flux \text{ for } \Omega_i = A \mathbf{Z}_i^T \bar{\mathbf{u}}. \tag{4.21}$$

Finally, consider a situation where BCs of the form $\bar{u}_b = \bar{g}_b$ for an exterior cell Ω_b are given. Above, the nodal value of Ω_b was used to calculate the flux over the exterior edge, but if BCs are given those values can be used instead:

$$flux = A \left(\sum_{j \in \mathcal{E}(\Omega_b) \backslash b} \frac{\bar{u}_b + \bar{u}_j}{2} \Delta y_{bj} + \bar{g}_b \Delta y_{bb} \right). \tag{4.22}$$

Observe that $\bar{g}_b$ has not replaced $\bar{u}_b$ when calculating the flux for the interior edges. The BCs are only introduced when calculating the boundary flux, thus $\bar{u}_b$ is not removed from the scheme, even though it could be set a priori. As a consequence, the numerical scheme will solve for $\bar{u}_b$ and the method is called *weakly imposed boundary conditions* [1].

Once again the sum is not over a closed loop, giving

$$flux = A \left(\sum_{\Delta y_j \in \mathcal{E}(\Omega_b) \backslash b} \frac{\Delta y_{bj}}{2} \bar{u}_j - \frac{\Delta y_{bb}}{2} \bar{u}_b + \bar{g}_b \Delta y_{bb} \right). \tag{4.23}$$

This flux can not in its current form be expressed by some $\mathbf{Z}_i$ and $\bar{\mathbf{u}}$. To remedy this, introduce the *penalty term*

$$b_{Ab} = A(\bar{g}_b - \bar{u}_b)\Delta y_{bb} \tag{4.24}$$

which allows the flux to be written as the sum from (4.21) plus the penalty term:

$$flux = A \left(\sum_{\Delta y_j \in \mathcal{E}(\Omega_b) \backslash b} \frac{\bar{u}_j}{2} \Delta y_{bj} + \frac{\bar{u}_b}{2} \Delta y_{bb} \right) + b_{Ab} = A \mathbf{Z}_b^T \bar{\mathbf{u}} + b_{Ab} \tag{4.25}$$

where the elements of $\mathbf{Z}_b$ are as for an exterior cell without given BCs.

4.1.4 Flux for all cells

By introducing the matrix

$$\mathbf{Q}_x = \begin{bmatrix} \mathbf{Z}_1^T \\ \mathbf{Z}_2^T \\ \vdots \\ \mathbf{Z}_n^T \end{bmatrix} \tag{4.26}$$

and the vector

$$\mathbf{b}_{Ai} = \begin{cases} A(\bar{g}_i - \bar{u}_i)\Delta y_{ii} & \text{at boundary with BCs,} \\ 0 & \text{otherwise} \end{cases} \tag{4.27}$$

it is possible to express the flux in the x-direction for all cells as a matrix-vector multiplication plus a penalty term:

$$\text{flux for all cells} = A\mathbf{Q}_x\bar{\mathbf{u}} + \mathbf{b}_A, \tag{4.28}$$

where the subscript x denotes the direction of the flux and $\mathbf{b}_A \in \mathbb{R}^n$. Since the rows of $\mathbf{Q}_x$ are the vectors $\mathbf{Z}_i^T$, $\forall i \in \mathcal{N}(\mathcal{T}(\Omega))$ one has $\mathbf{Q}_x(ij) = \mathbf{Z}_i(j)$, thus

$$\mathbf{Q}_x(ij) = \frac{\Delta y_{ij}}{2}, \quad \mathbf{Q}_x(ji) = -\mathbf{Q}_x(ij) = -\frac{\Delta y_{ij}}{2}, \tag{4.29}$$

meaning that $\mathbf{Q}_x$ is skew-symmetric if the off-diagonal elements alone are considered. Further, the diagonal elements are

$$\mathbf{Q}_x(ii) = \begin{cases} \frac{\Delta y_{ii}}{2} & \text{if } \Omega_i \text{ is a boundary cell,} \\ 0 & \text{otherwise} \end{cases} \tag{4.30}$$

motivating the introduction of $\mathbf{Y} := \mathbf{Q}_x + \mathbf{Q}_x^T$, where $\mathbf{Y}$ is a diagonal matrix with the elements

$$\mathbf{Y}(ii) = \begin{cases} \Delta y_{ii} & \text{if } \Omega_i \text{ is a boundary cell,} \\ 0 & \text{otherwise.} \end{cases} \tag{4.31}$$

By treating the sum corresponding to y-directed flow analogously, the matrix $\mathbf{Q}_y$ and the penalty term $\mathbf{b}_B$ are obtained, where

$$\mathbf{b}_{Bi} = \begin{cases} -B(\bar{g}_i - \bar{u}_i)\Delta x_{ii} & \text{at boundary with BCs,} \\ 0 & \text{otherwise,} \end{cases} \tag{4.32}$$

and $\mathbf{Q}_y$ is as $\mathbf{Q}_x$ but with $-\frac{\Delta x_{ij}}{2}$ instead of $\frac{\Delta y_{ij}}{2}$ as elements. It follows that $\mathbf{X} := \mathbf{Q}_y + \mathbf{Q}_y^T$ is a diagonal matrix, with the elements

$$\mathbf{X}(ii) = \begin{cases} -\Delta x_{ii} & \text{if } \Omega_i \text{ is a boundary cell,} \\ 0 & \text{otherwise.} \end{cases} \tag{4.33}$$

The existence of $\mathbf{Y}$ and $\mathbf{X}$ is, as previously stated, crucial for obtaining the estimate (4.1).

The flux for every cell may now be formulated as a sum of matrix-vector multiplications and penalty terms:

$$\text{flux} = A\mathbf{Q}_x\bar{\mathbf{u}} + B\mathbf{Q}_y\bar{\mathbf{u}} + \mathbf{b}_A + \mathbf{b}_B \tag{4.34}$$

Then, by introduction of the matrix $\mathbf{P}$ containing the areas of all cells in $\mathcal{T}(\Omega)$:

$$\mathbf{P} = \text{diag}(|\Omega_1|, |\Omega_2|, \ldots, |\Omega_n|), \tag{4.35}$$

the first term in (4.12) can be given for all cells by the vector $\mathbf{P}\bar{\mathbf{u}}_t$. Combining all these expressions gives a single one for (4.12) for all $\Omega_i \in \mathcal{T}(\Omega)$, which is the semi-discretisation (4.10) of the scalar version of the main problem (2.1):

$$\mathbf{P}\bar{\mathbf{u}}_t + A\mathbf{Q}_x\bar{\mathbf{u}} + B\mathbf{Q}_y\bar{\mathbf{u}} + \mathbf{b} = \mathbf{0} \tag{4.36}$$

where

$$\mathbf{b}_i = \mathbf{b}_{Ai} + \mathbf{b}_{Bi} = \begin{cases} (A\hat{\mathbf{x}} + B\hat{\mathbf{y}}) \cdot \mathbf{n}_i(\bar{g}_i - \bar{u}_i) & \text{at boundary with BCs,} \\ 0 & \text{otherwise,} \end{cases} \tag{4.37}$$

with $\mathbf{n}_i = [\Delta y_{ii}, -\Delta x_{ii}]^T$. Note that the BCs are only present in the penalty term, which contains the difference between the solution on the boundary and the given BCs.

4.1.5 The general case

For any $l \in \mathbb{N}$, equation (4.12) can with some modifications be treated analogously to $l = 1$. Focus once again on the flux in the x-direction, using the same numerical flux function as for (4.12) for a single cell Ω_i. The same vector $\mathbf{Z}_i$ as for the case with $l = 1$ will be used to get a matrix-vector representation of the flux, but this time as a

component in the matrix

$$
\tilde{\mathbf{Z}}_i =
\begin{bmatrix}
\mathbf{Z}_i^T & \mathbf{0}^T & \dots & \mathbf{0}^T \\
\mathbf{0}^T & \mathbf{Z}_i^T & \dots & \mathbf{0}^T \\
\vdots & \vdots & \ddots & \vdots \\
\mathbf{0}^T & \mathbf{0}^T & \dots & \mathbf{Z}_i^T
\end{bmatrix},
\tag{4.38}
$$

where $\mathbf{0}$ is the zero-vector of length n. The matrix $\tilde{\mathbf{Z}}_i$ functions as $\mathbf{Z}_i^T$ does when $l = 1$, but the flux must now be evaluated for each of the l entries in $\bar{\mathbf{u}}_i$, hence the number of $\mathbf{Z}_i^T$ entries in $\tilde{\mathbf{Z}}_i^T$ is l. The flux in the x-direction for an interior cell Ω_i in (4.12) can then be written as

$$
\text{flux in x-direction} = \mathbf{A}\tilde{\mathbf{Z}}_i\bar{\mathbf{u}} =
\begin{bmatrix}
a_{11}\mathbf{Z}_i^T & a_{12}\mathbf{Z}_i^T & \cdots & a_{1l}\mathbf{Z}_i^T \\
a_{21}\mathbf{Z}_i^T & a_{22}\mathbf{Z}_i^T & \cdots & a_{2l}\mathbf{Z}_i^T \\
\vdots & \vdots & \ddots & \vdots \\
a_{l1}\mathbf{Z}_i^T & a_{l2}\mathbf{Z}_i^T & \cdots & a_{ll}\mathbf{Z}_i^T
\end{bmatrix}
\bar{\mathbf{u}} = (\mathbf{A} \otimes \mathbf{Z}_i^T)\bar{\mathbf{u}}, \quad (4.39)
$$

where $\otimes$ denotes the Kronecker product, see appendix A.

If $(\mathbf{A} \otimes \boldsymbol{\xi}^T)\bar{\mathbf{u}}$ with $\boldsymbol{\xi} = \mathbf{Z}_i$ gives the flux for an interior cell, then one should be able to assign another operator to $\boldsymbol{\xi}$ in order to obtain the flux for all cells. When $l = 1$ the matrices $\mathbf{Q}_x$ and $\mathbf{Q}_y$ could be used as discrete differentiation operators. Now each of the l quantities in $\boldsymbol{u}_x$ and $\boldsymbol{u}_y$ must be approximated for each cell, which is achieved by letting $\mathbf{Q}_x$ and $\mathbf{Q}_y$ act on each quantity. This is possible due to the structure of $\bar{\mathbf{u}}$ and by letting $\boldsymbol{\xi} = \mathbf{Q}_x$ or $\boldsymbol{\xi} = \mathbf{Q}_y$ for the y-directed flow. To see this, consider the k:th element of the vector $\mathbf{A}\mathbf{u}_x$ which for any k may be written as

$$
\begin{bmatrix}
a_{k1} & a_{k2} & \cdots & a_{kl}
\end{bmatrix}
\begin{bmatrix}
u_{1\,x} \\
u_{2\,x} \\
\vdots \\
u_{l\,x}
\end{bmatrix}
= a_{k1}u_{1\,x} + a_{k2}u_{2\,x} + \dots + a_{kl}u_{l\,x}.
\tag{4.40}
$$

The discrete version of 4.40 is $u_{1\,x}, u_{2\,x}, \dots, u_{l\,x}$ replaced with $\bar{\mathbf{u}}^1, \bar{\mathbf{u}}^2, \dots, \bar{\mathbf{u}}^l$. The differentiation with respect to x on each cell is obtained by applying the discrete differential operator $\mathbf{Q}_x$. So the corresponding discretised k:th element would be

$$
\begin{bmatrix}
a_{k1}\mathbf{Q}_x & a_{k2}\mathbf{Q}_x & \cdots & a_{kl}\mathbf{Q}_x
\end{bmatrix}
\begin{bmatrix}
\bar{\mathbf{u}}^1 \\
\bar{\mathbf{u}}^2 \\
\vdots \\
\bar{\mathbf{u}}^l
\end{bmatrix}
= a_{k1}\mathbf{Q}_x\bar{\mathbf{u}}^1 + a_{k2}\mathbf{Q}_x\bar{\mathbf{u}}^2 + \dots + a_{kl}\mathbf{Q}_x\bar{\mathbf{u}}^l
\tag{4.41}
$$

Thus one obtains the discrete x-directed fluxes for all states in $\mathbf{u}$ for all cells by setting $\boldsymbol{\xi} = \mathbf{Q}_x$, which gives $(\mathbf{A} \otimes \mathbf{Q}_x)\bar{\mathbf{u}}$. With the same reasoning, one realises that $(\mathbf{I} \otimes \mathbf{P})\bar{\mathbf{u}}_t$ gives the vector $\bar{\mathbf{u}}$ differentiated with respect to t, where each element is multiplied with the area it is the cell average over. Finally one obtains the full matrix-vector representation of the discretised main problem:

$$(\mathbf{I} \otimes \mathbf{P})\bar{\mathbf{u}}_t + (\mathbf{A} \otimes \mathbf{Q}_x)\bar{\mathbf{u}} + (\mathbf{B} \otimes \mathbf{Q}_y)\bar{\mathbf{u}} + \mathbf{b} = \mathbf{0}, \tag{4.42}$$

where the boundary terms are collected in the penalty term $\mathbf{b} \in \mathbb{R}^N$ with elements such that

$$\mathbf{b}_i = \mathbf{b}_{Ai} + \mathbf{b}_{Bi} = \begin{cases} (\mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\hat{\mathbf{y}}) \cdot \mathbf{n}_i(\bar{\mathbf{g}}_i - \bar{\mathbf{u}}_i) & \text{at boundary with BCs}, \\ \mathbf{0} & \text{otherwise}. \end{cases} \tag{4.43}$$

Further, it has been shown that $\mathbf{Q}_x + \mathbf{Q}_x^T$ and $\mathbf{Q}_y + \mathbf{Q}_y^T$ are diagonal matrices, an important result that now will be utilised.

4.2 Stability analysis

Now that the approximation (4.42) of the main problem (2.1) is derived, it remains to achieve strict stability in sense of definition 6. The norm $\|\bar{\mathbf{u}}\|_{\mathbf{I}\otimes\mathbf{P}}^2 = \bar{\mathbf{u}}^T(\mathbf{I} \otimes \mathbf{P})\bar{\mathbf{u}}$, which is the discrete equivalence of the $L^2(\Omega)$-norm, differentiated with respect to time in combination with some $\bar{\mathbf{u}}$ satisfying (4.42) yields

$$\begin{aligned} \frac{d}{dt}\|\bar{\mathbf{u}}\|_{\mathbf{I}\otimes\mathbf{P}}^2 &= \bar{\mathbf{u}}_t^T(\mathbf{I} \otimes \mathbf{P})\bar{\mathbf{u}} + \bar{\mathbf{u}}^T(\mathbf{I} \otimes \mathbf{P})\bar{\mathbf{u}}_t = ((\mathbf{I} \otimes \mathbf{P})\bar{\mathbf{u}}_t)^T\bar{\mathbf{u}} + \bar{\mathbf{u}}^T(\mathbf{I} \otimes \mathbf{P})\bar{\mathbf{u}}_t \\ &= (-(\mathbf{A} \otimes \mathbf{Q}_x)\bar{\mathbf{u}} - (\mathbf{B} \otimes \mathbf{Q}_y)\bar{\mathbf{u}} - \mathbf{b})^T\bar{\mathbf{u}} + \bar{\mathbf{u}}^T(-(\mathbf{A} \otimes \mathbf{Q}_x)\bar{\mathbf{u}} - (\mathbf{B} \otimes \mathbf{Q}_y)\bar{\mathbf{u}} - \mathbf{b}) \\ &= -\bar{\mathbf{u}}^T(\mathbf{A} \otimes \mathbf{Q}_x^T)\bar{\mathbf{u}} - \bar{\mathbf{u}}^T(\mathbf{A} \otimes \mathbf{Q}_x)\bar{\mathbf{u}} - \bar{\mathbf{u}}^T(\mathbf{B} \otimes \mathbf{Q}_y^T)\bar{\mathbf{u}} - \bar{\mathbf{u}}^T(\mathbf{B} \otimes \mathbf{Q}_y^T)\bar{\mathbf{u}} - \mathbf{b}^T\bar{\mathbf{u}} - \bar{\mathbf{u}}^T\mathbf{b} \\ &= -\bar{\mathbf{u}}^T(\mathbf{A} \otimes (\mathbf{Q}_x + \mathbf{Q}_x^T))\bar{\mathbf{u}} - \bar{\mathbf{u}}^T(\mathbf{B} \otimes (\mathbf{Q}_y + \mathbf{Q}_y^T))\bar{\mathbf{u}} - 2\bar{\mathbf{u}}^T\mathbf{b}, \end{aligned} \tag{4.44}$$

due to the properties of the Kronecker product, see appendix B, and the assumption that the cells Ω_i are constant over time.

Recall the diagonal matrices $\mathbf{Y} := \mathbf{Q}_x + \mathbf{Q}_x^T = \mathbf{Y}$ and $\mathbf{X} = \mathbf{Q}_y + \mathbf{Q}_y^T$ with the elements (4.31) and (4.33). Let m be the number of non-zero elements on the diagonal, which also is the amount of boundary nodes. This gives

$$\frac{d}{dt}\|\bar{\mathbf{u}}\|_{\mathbf{I}\otimes\mathbf{P}}^2 = -\bar{\mathbf{u}}^T(\mathbf{A} \otimes \mathbf{Y})\bar{\mathbf{u}} - \bar{\mathbf{u}}^T(\mathbf{B} \otimes \mathbf{X})\bar{\mathbf{u}} - 2\bar{\mathbf{u}}^T\mathbf{b} \tag{4.45}$$

where all elements in $\bar{\mathbf{u}}$ corresponding to interior nodes are multiplied with the zeros in $\mathbf{X}$ and $\mathbf{Y}$. In fact, the diagonal structure of $\mathbf{X}$ and $\mathbf{Y}$ gives

$$\boldsymbol{\phi}^T \mathbf{Y} \boldsymbol{\phi} = \sum_{i=1}^{m} \Delta y_i \phi_{ii}^2, \quad \boldsymbol{\phi}^T \mathbf{X} \boldsymbol{\phi} = -\sum_{i=1}^{m} \Delta x_{ii} \phi_i^2 \tag{4.46}$$

for some $\boldsymbol{\phi} \in \mathbb{R}^N$. Consider once again example (4.41), but with $\mathbf{Y}$ instead of $\mathbf{Q_x}$,

$$\begin{bmatrix} a_{k1}\mathbf{Y} & a_{k2}\mathbf{Y} & \cdots & a_{kl}\mathbf{Y} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{u}}^1 \\ \bar{\mathbf{u}}^2 \\ \vdots \\ \bar{\mathbf{u}}^l \end{bmatrix} = a_{k1}\mathbf{Y}\bar{\mathbf{u}}^1 + a_{k2}\mathbf{Y}\bar{\mathbf{u}}^2 + \ldots + a_{kl}\mathbf{Y}\bar{\mathbf{u}}^l, \tag{4.47}$$

where the diagonal matrix $\mathbf{Y}$ decouples the cells, since $\mathbf{Y}\bar{\mathbf{u}}^j$ just gives $\bar{\mathbf{u}}^j$ with separately scaled elements. As a consequence, the expression can be treated cell-wise, since one only gets linear combinations of the elements belonging to the same cell. The same holds when multiplying with $\bar{\mathbf{u}}^k$, as only elements belonging to the same cells are multiplied. The previous result may be written as

$$\bar{\mathbf{u}}^T (\mathbf{A} \otimes \mathbf{Y}) \bar{\mathbf{u}} = \sum_{i \in \mathcal{E}(\Omega)} \left(\bar{\mathbf{u}}_i^T \mathbf{A} \bar{\mathbf{u}}_i \Delta y_{ii} \right). \tag{4.48}$$

This encourages a formulation of $\frac{d}{dt}\|\bar{\mathbf{u}}\|_{\mathbf{I} \otimes \mathbf{P}}^2$ based on cell-wise contributions. Applying the same logic to $\bar{\mathbf{u}}^T (\mathbf{B} \otimes \mathbf{X}) \bar{\mathbf{u}}$, combined with the fact that no reformulation for the penalty term is needed, gives that expression (4.44) can be formulated as

$$\frac{d}{dt}\|\bar{\mathbf{u}}\|_{\mathbf{I} \otimes \mathbf{P}}^2 = -\sum_{i \in \mathcal{E}(\Omega)} \left(\bar{\mathbf{u}}_i^T \mathbf{A} \bar{\mathbf{u}}_i \Delta y_{ii} - \bar{\mathbf{u}}_i^T \mathbf{B} \bar{\mathbf{u}}_i \Delta x_{ii} + 2\bar{\mathbf{u}}_i^T \mathbf{b}_i \right). \tag{4.49}$$

Recall expression (4.6) and consider the discrete version

$$(\mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\hat{\mathbf{y}}) \cdot \hat{\mathbf{n}}_i = (\mathbf{A}\Delta y_{ii} - \mathbf{B}\Delta x_{ii}) / \Delta s_i \tag{4.50}$$

where $\Delta s_i = \sqrt{\Delta x_{ii}^2 + \Delta y_{ii}^2}$ and $\hat{\mathbf{n}}_i = [\Delta y_{ii}, -\Delta x_{ii}]^T / \Delta s_i$. By reformulating $\mathbf{b}$ from (4.43) to

$$\mathbf{b}_i = \mathbf{b}_{Ai} + \mathbf{b}_{Bi} = \begin{cases} (\mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\hat{\mathbf{y}}) \cdot \hat{\mathbf{n}}_i(\bar{\mathbf{g}}_i - \bar{\mathbf{u}}_i) & \text{at boundary with BCs,} \\ \mathbf{0} & \text{otherwise,} \end{cases} \tag{4.51}$$

where the unit normal is used instead, the initial expression (4.44) can finally be written as

$$\frac{d}{dt}\|\bar{\mathbf{u}}\|^2_{\mathbf{I}\otimes\mathbf{P}} = -\sum_{i\in\mathcal{E}(\Omega)} \left(\bar{\mathbf{u}}_i^T (\mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\hat{\mathbf{y}})\cdot\hat{\mathbf{n}}_i\,\bar{\mathbf{u}}_i + 2\bar{\mathbf{u}}_i^T\mathbf{b}_i\right)\Delta s_i, \tag{4.52}$$

which corresponds naturally to the continuous version (4.5).

4.2.1 Strict stability

Before obtaining zero as an upper bound for the estimate of $\frac{d}{dt}\|\bar{\mathbf{u}}\|^2_{\mathbf{I}\otimes\mathbf{P}}$ the BCs need to be discussed, since the expression (4.52) clearly only depends on boundary values. The analysis is based on BCs given in weak characteristic form, however it will be explicitly shown in chapter 5 that strict stability is achieved for the two-dimensional Euler equations with fixed wall BCs as well. As before, the continuous case is first analysed and suggests how to handle the discrete estimate.

Consider the continuous estimate (4.5) again and let the columns of the matrix $\mathbf{K}$ consist of the normed eigenvectors for $(\mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\hat{\mathbf{y}})\cdot\hat{\mathbf{n}}$ and $\mathbf{\Lambda}$ be a diagonal matrix consisting of the eigenvalues of $(\mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\hat{\mathbf{y}})\cdot\hat{\mathbf{n}}$ such that

$$(\mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\hat{\mathbf{y}})\cdot\hat{\mathbf{n}} = \mathbf{K}\mathbf{\Lambda}\mathbf{K}^T. \tag{4.53}$$

Let $\mathbf{\Lambda} = \mathbf{\Lambda}^+ + \mathbf{\Lambda}^-$, where $\mathbf{\Lambda}^+$ and $\mathbf{\Lambda}^-$ are diagonal matrices containing the positive and negative eigenvalues, respectively. The other entries are replaced with zeroes. In the same fashion, let $\mathbf{K}^+$ and $\mathbf{K}^-$ contain the normed eigenvectors corresponding to the respective positive or negative eigenvalues, the other columns replaced with zero vectors. Let the characteristic forms of $\mathbf{u}$ and $\mathbf{g}$ be $\mathbf{u}_c = \mathbf{K}^T\mathbf{u}$ and

$$\mathbf{K}^T\mathbf{g} = \mathbf{g}_c = \mathbf{g}_c^+ + \mathbf{g}_c^-, \tag{4.54}$$

where $\mathbf{K}^{+T}\mathbf{g} = \mathbf{g}_c^+$ and $\mathbf{K}^{-T}\mathbf{g} = \mathbf{g}_c^-$. According to section 3.4 $\mathbf{g}_c^-$ must be given as a boundary condition since $\mathbf{K}^-$ corresponds to the ingoing characteristics, while $\mathbf{g}_c^+$ on the other hand contains information for the outgoing characteristics and will not be prescribed. As already discussed, local data based on the current numerical solution is used instead, meaning $\mathbf{g}_c^+ = \mathbf{K}^{+T}\mathbf{u}$.

Expression (4.5) can now be rewritten, by the use of the relations (4.53) and (4.54), as

$$
\begin{aligned}
\frac{d}{dt}\|\mathbf{u}\|_\Omega^2 &= -\oint_{\partial\Omega} \mathbf{u}^T(\mathbf{A}\hat{\mathbf{x}}+\mathbf{B}\hat{\mathbf{y}})\cdot\hat{\mathbf{n}}\,\mathbf{u}\,ds \\
&= -\oint_{\partial\Omega} \mathbf{u}^T\mathbf{K}\boldsymbol{\Lambda}\mathbf{K}^T\mathbf{u}\,ds \\
&= -\oint_{\partial\Omega} \mathbf{u}^T\mathbf{K}\boldsymbol{\Lambda}^+\mathbf{K}^T\mathbf{u}\,ds - \oint_{\partial\Omega} \mathbf{u}^T\mathbf{K}\boldsymbol{\Lambda}^-\mathbf{K}^T\mathbf{u}\,ds \\
&= -\oint_{\partial\Omega} \mathbf{u}_c^T\boldsymbol{\Lambda}^+\mathbf{u}_c\,ds - \oint_{\partial\Omega} \mathbf{g}_c^T\boldsymbol{\Lambda}^-\mathbf{g}_c\,ds
\end{aligned}
\tag{4.55}
$$

where the relations $\boldsymbol{\Lambda}^-\mathbf{K} = \boldsymbol{\Lambda}\mathbf{K}^- = \boldsymbol{\Lambda}^-\mathbf{K}^-$ are used. Recall that the BCs for the main problem (2.1) are such that $\mathbf{u}^T(\mathbf{A}\hat{\mathbf{x}}+\mathbf{B}\hat{\mathbf{y}})\cdot\hat{\mathbf{n}}\,\mathbf{u} \geq 0$. The first integral, with the integrand $\mathbf{u}_c^T\boldsymbol{\Lambda}^+\mathbf{u}_c$, is unconditionally less or equal to zero, meaning that one has to state the ingoing characteristics $\mathbf{g}_c^-$ such that positive semi-definiteness is obtained. Then $\frac{d}{dt}\|\mathbf{u}\|_{\mathbf{I}\otimes\mathbf{P}}^2$ gets zero as an upper bound.

The estimate (4.1) is almost at hand. The introduction of $\boldsymbol{\Lambda}$ and $\mathbf{K}$ allows a formulation of (4.52) similar to the continuous case (4.55):

$$
\begin{aligned}
\frac{d}{dt}\|\bar{\mathbf{u}}\|_{\mathbf{I}\otimes\mathbf{P}}^2 &= -\sum_{i\in\mathcal{E}(\Omega)} \left(\bar{\mathbf{u}}_i^T(\mathbf{A}\hat{\mathbf{x}}+\mathbf{B}\hat{\mathbf{y}})\cdot\hat{\mathbf{n}}_i\,\bar{\mathbf{u}}_i + 2\bar{\mathbf{u}}_i^T\mathbf{b}_i\right)\Delta s_i \\
&= -\sum_{i\in\mathcal{E}(\Omega)} \left(\bar{\mathbf{u}}_i^T(\mathbf{A}\hat{\mathbf{x}}+\mathbf{B}\hat{\mathbf{y}})\cdot\hat{\mathbf{n}}_i\,\bar{\mathbf{u}}_i + 2\bar{\mathbf{u}}_i^T(\mathbf{A}\hat{\mathbf{x}}+\mathbf{B}\hat{\mathbf{y}})\cdot\hat{\mathbf{n}}_i(\bar{\mathbf{g}}_i-\bar{\mathbf{u}}_i)\right)\Delta s_i \\
&= -\sum_{i\in\mathcal{E}(\Omega)} \left(\bar{\mathbf{u}}_i^T\,\mathbf{K}_i\boldsymbol{\Lambda}_i\mathbf{K}_i^T\bar{\mathbf{u}}_i + 2\bar{\mathbf{u}}_i^T\mathbf{K}_i\boldsymbol{\Lambda}_i\mathbf{K}_i^T(\bar{\mathbf{g}}_i-\bar{\mathbf{u}}_i)\right)\Delta s_i \\
&= -\sum_{i\in\mathcal{E}(\Omega)} \left(\bar{\mathbf{u}}_{ci}^T\boldsymbol{\Lambda}_i^+\bar{\mathbf{u}}_{ci} + \bar{\mathbf{u}}_{ci}^T\boldsymbol{\Lambda}_i^-\bar{\mathbf{u}}_{ci} + 2\bar{\mathbf{u}}_{ci}^T\boldsymbol{\Lambda}_i^-(\bar{\mathbf{g}}_{ci}-\bar{\mathbf{u}}_{ci})\right)\Delta s_i \\
&= -\sum_{i\in\mathcal{E}(\Omega)} \left(\bar{\mathbf{u}}_{ci}^T\boldsymbol{\Lambda}_i^+\bar{\mathbf{u}}_{ci} + \bar{\mathbf{g}}_{ci}^T\boldsymbol{\Lambda}_i^-\bar{\mathbf{g}}_{ci}\right)\Delta s_i + \mathbf{R}
\end{aligned}
\tag{4.56}
$$

where $\mathbf{R} = \sum_{i\in\mathcal{E}(\Omega)}\left(\bar{\mathbf{u}}_{ci}^T\boldsymbol{\Lambda}_i^-\bar{\mathbf{u}}_{ci} + \bar{\mathbf{g}}_{ci}^T\boldsymbol{\Lambda}_i^-\bar{\mathbf{g}}_{ci} - 2\bar{\mathbf{u}}_{ci}^T\boldsymbol{\Lambda}_i^-\bar{\mathbf{g}}_{ci}\right)\Delta s_i$. The subscript i is required for $\mathbf{K}$ and $\boldsymbol{\Lambda}$ since $\hat{\mathbf{n}}_i$ makes the eigenvectors and eigenvalues local for each edge E_{ii}. Note that the contributions from the outgoing characteristics have been omitted since $\mathbf{K}_i^{+\,T}(\bar{\mathbf{g}}_i-\bar{\mathbf{u}}_i) = \mathbf{0}$, $\forall i\in\mathcal{E}(\Omega)$.

The weak characteristic BCs applied via the penalty term may be formulated in an even more general sense. Instead of the penalty term (4.51), consider

$$
\mathbf{b}_i = \begin{cases} \frac{\delta}{2}\boldsymbol{\Lambda}_i^-(\bar{\mathbf{g}}_{ci}-\bar{\mathbf{u}}_{ci}) & \text{at boundary with BCs,} \\ \mathbf{0} & \text{otherwise,} \end{cases}
\tag{4.57}
$$

where the influence of the BCs can be regulated by $\delta \in \mathbb{R}^+$. One still obtains expression (4.56) with the new penalty term, with the exception that

$$\mathbf{R} = \sum_{i \in \partial\Omega} \left((\delta - 1)\bar{\mathbf{u}}_{ci}^T \mathbf{\Lambda}_i^- \bar{\mathbf{u}}_{ci} + \bar{\mathbf{g}}_{ci}^T \mathbf{\Lambda}_i^- \bar{\mathbf{g}}_{ci} - \delta \bar{\mathbf{u}}_{ci}^T \mathbf{\Lambda}_i^- \bar{\mathbf{g}}_{ci} \right) \Delta s_i, \qquad (4.58)$$

which motives the following theorem:

Theorem 1 (Strict stability). The discrete approximation (4.12) of the main problem (4.4) is strictly stable in the sense of definition 6 if $\mathbf{R}$ from (4.58) is non-positive.

The theorem comes naturally if one studies (4.56)

$$\frac{d}{dt}\|\bar{\mathbf{u}}\|_{\mathbf{I}\otimes\mathbf{P}}^2 = - \sum_{i \in \mathcal{E}(\Omega)} \left(\bar{\mathbf{u}}_{ci}^T \mathbf{\Lambda}_i^+ \bar{\mathbf{u}}_{ci} + \bar{\mathbf{g}}_{ci}^T \mathbf{\Lambda}_i^- \bar{\mathbf{g}}_{ci} \right) \Delta s_i + \mathbf{R} \qquad (4.59)$$

where the sum will always be negative, due to the positive semi-definiteness given as BCs, see (4.7). Therefore the sign of $\mathbf{R}$ states whether (4.1) is obtained, meaning that the approximation is strictly stable. If $\mathbf{g}_c \neq 0$, then $\mathbf{R} \leq 0$ if $\delta = 2$ since it gives

$$\mathbf{R} = \sum_{i \in \partial\Omega} (\bar{\mathbf{u}}_{ci} - \bar{\mathbf{g}}_{ci})^T \mathbf{\Lambda}_i^- (\bar{\mathbf{u}}_{ci} - \bar{\mathbf{g}}_{ci}) \Delta s_i \leq 0. \qquad (4.60)$$

Also note that $\delta = 2$ gives the same $\mathbf{R}$ as (4.56). For $g = 0$, $\mathbf{R} \leq 0$ is satisfied for all $\delta \geq 1$.

Chapter 5

Validation & comparison

In this chapter, the node-centred FVM with weakly imposed BCs for the two-dimensional Euler equations in symmetrised primitive form is shown to be strictly stable in sense of definition 6. Thereafter the characteristic BCs for the nonlinear Euler equations are considered, followed by numerical experiments for the nonlinear case. These experiments aim to validate the connection between the stability analysis for constant coefficients matrices and nonlinear Euler equations. The weak characteristic BCs are compared to the weak prescribed fluxes described in chapter 3. The stability of the scheme is not only of interest, but also if it performs better than with weak prescribed fluxes.

5.1 Analysis of the linearised Euler equations in two dimensions

Strict stability, in the sense of definition 6, will be shown for the symmetric Euler equations in primitive form, see (2.23), with linearised matrices. For simplicity, let the boundary of the domain be split in two parts such that $\partial\Omega = \partial\Omega_C \cup \partial\Omega_F$. The segment Ω_C has characteristic BCs and the segment $\partial\Omega_F$ is given the fixed wall BC, meaning $v_n = \boldsymbol{v} \cdot \hat{\mathbf{n}} = 0$.

First the continuous case will be studied, followed by the discretisation. Consider (4.55), with some solution $\mathbf{u}$ to linearised (2.23), before imposing the fixed wall BC:

$$\frac{d}{dt}\|\mathbf{u}\|_\Omega^2 = -\oint_{\partial\Omega_C} \mathbf{u}_c^T \boldsymbol{\Lambda}^+ \mathbf{u}_c + \mathbf{g}_c^T \boldsymbol{\Lambda}^- \mathbf{g}_c \, ds - \oint_{\partial\Omega_F} \mathbf{u}_c^T \boldsymbol{\Lambda} \mathbf{u}_c \, ds, \qquad (5.1)$$

39

where the relation $(\mathbf{A}_s\hat{\mathbf{x}} + \mathbf{B}_s\hat{\mathbf{y}}) \cdot \hat{\mathbf{n}} = \mathbf{K}\boldsymbol{\Lambda}\mathbf{K}^T$ is used with $\hat{\mathbf{n}} = [\alpha, \beta]^T$ as the outward pointing unit normal. In order to evaluate the contribution from the solid wall, $\mathbf{K}\boldsymbol{\Lambda}\mathbf{K}^T$ must be investigated further. The eigenvectors and eigenvalues for $(\mathbf{A}_s\hat{\mathbf{x}} + \mathbf{B}_s\hat{\mathbf{y}}) \cdot \hat{\mathbf{n}}$ are derived in appendix A, as well as the matrices $\mathbf{K}$ (A.17) and $\boldsymbol{\Lambda}$ (A.18).

For the two-dimensional Euler equation in symmetrised primitive form, the matrices $\mathbf{A}_s(\mathbf{u}_s)$ and $\mathbf{B}_s(\mathbf{u}_s)$ are as given in chapter 2. A constant coefficient analysis will be considered, meaning a solution $\tilde{\mathbf{u}}$ evaluated for some $(x, y) \in \Omega$ is given as argument for $\mathbf{A}_s = \mathbf{A}_s(\tilde{\mathbf{u}})$ and $\mathbf{B}_s = \mathbf{B}_s(\tilde{\mathbf{u}})$. This results in

$$
\mathbf{A}_s\alpha + \mathbf{B}_s\beta = \begin{bmatrix} \tilde{v}_n & \alpha\tilde{c}/\sqrt{\gamma} & \beta\tilde{c}/\sqrt{\gamma} & 0 \\ \alpha\tilde{c}/\sqrt{\gamma} & \tilde{v}_n & 0 & \alpha\tilde{c}\sqrt{(\gamma-1)/\gamma} \\ \beta\tilde{c}/\sqrt{\gamma} & 0 & \tilde{v}_n & \beta\tilde{c}\sqrt{(\gamma-1)/\gamma} \\ 0 & \alpha\tilde{c}\sqrt{(\gamma-1)/\gamma} & \beta\tilde{c}\sqrt{(\gamma-1)/\gamma} & \tilde{v}_n, \end{bmatrix} \tag{5.2}
$$

and the diagonal matrix $\boldsymbol{\Lambda}$ is

$$
\boldsymbol{\Lambda} = \begin{bmatrix} \tilde{v}_n - \tilde{c} & 0 & 0 & 0 \\ 0 & \tilde{v}_n & 0 & 0 \\ 0 & 0 & \tilde{v}_n & 0 \\ 0 & 0 & 0 & \tilde{v}_n + \tilde{c} \end{bmatrix}, \tag{5.3}
$$

where the diagonal elements are the eigenvalues of $\mathbf{A}_s\alpha + \mathbf{B}_s\beta$. The specific choice of $\tilde{\mathbf{u}}$ is such that it satisfies the symmetrised Euler equation in primitive form, the ICs and the BCs on $\partial\Omega_F$, meaning that $\tilde{\mathbf{u}}$ is evaluated for some $(x, y) \in \partial\Omega_F$. Thus the superscript $\sim$ represents reference values satisfying $\tilde{v}_n = 0$. As will be shown this leads to strict stability, see definition 6.

The columns of $\mathbf{K}$ are the orthonormal set of eigenvectors to $\mathbf{A}_s\alpha + \mathbf{B}_s\beta$ and thus the characteristic form of $\mathbf{u}_s$ is

$$
\mathbf{u}_c = \mathbf{K}^T\mathbf{u}_s = \begin{bmatrix} \frac{1}{\sqrt{2\gamma}} & -\frac{\alpha}{\sqrt{2}} & -\frac{\beta}{\sqrt{2}} & \sqrt{\frac{\gamma-1}{2\gamma}} \\ -\sqrt{\frac{\gamma-1}{2}} & 0 & 0 & \frac{1}{\sqrt{\gamma}} \\ 0 & -\beta & \alpha & 0 \\ \frac{1}{\sqrt{2\gamma}} & \frac{\alpha}{\sqrt{2}} & \frac{\beta}{\sqrt{2}} & \sqrt{\frac{\gamma-1}{2\gamma}} \end{bmatrix} \begin{bmatrix} \frac{c}{\sqrt{\gamma}} \\ v_1 \\ v_2 \\ 0 \end{bmatrix} = \frac{1}{\rho c\sqrt{2}} \begin{bmatrix} p - \rho c v_n \\ \theta(p - \rho c^2) \\ \rho c\sqrt{2}\,v_t \\ p + \rho c v_n \end{bmatrix} \tag{5.4}
$$

where the relations

$$\begin{cases} p & = \frac{\rho c^2}{\gamma}, \\ \theta & = \sqrt{\frac{2}{\gamma-1}}, \\ v_t & = \boldsymbol{v} \cdot [-\beta, \alpha]^T \end{cases} \tag{5.5}$$

have been used. Imposing the slip condition on $\mathbf{u}_c$ for the integral over $\partial\Omega_F$ in (5.1) gives the integrand

$$\mathbf{u}_c^T \boldsymbol{\Lambda} \mathbf{u}_c = \frac{1}{\rho^2 c^2 2} \begin{bmatrix} p, & \theta(p - \rho c^2), & \rho c\sqrt{2}\, v_t, & p, \end{bmatrix} \begin{bmatrix} -\tilde{c} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \tilde{c} \end{bmatrix} \begin{bmatrix} p \\ \theta(p - \rho c^2) \\ \rho c\sqrt{2}\, v_t \\ p \end{bmatrix} = 0. \tag{5.6}$$

Hence the contribution from the solid wall to $\frac{d}{dt}\|\mathbf{u}\|_\Omega^2$ is identically equal to zero and therefore (4.5) is obtained.

Now consider the discrete version (4.56), but handling the boundary segments $\partial\Omega_C$ and $\partial\Omega_F$ separately

$$\frac{d}{dt}\|\bar{\mathbf{u}}\|_{\mathbf{I}\otimes\mathbf{P}}^2 = -\sum_{i\in\partial\Omega_C} \left(\bar{\mathbf{u}}_{ci}^T \boldsymbol{\Lambda}_i^+ \bar{\mathbf{u}}_{ci} + \bar{\mathbf{g}}_{ci}^T \boldsymbol{\Lambda}_i^- \bar{\mathbf{g}}_{ci}\right) \Delta s_i + \mathbf{R} - \sum_{i\in\partial\Omega_F} \left(\bar{\mathbf{u}}_{ci}^T \boldsymbol{\Lambda}_i \bar{\mathbf{u}}_{ci} + 2\bar{\mathbf{u}}_i^T \mathbf{b_i}\right) \Delta s_i \tag{5.7}$$

where $\mathbf{R}$ is of the form given in (4.58). The penalty term in the sum over $\partial\Omega_F$ is

$$\mathbf{b}_i = \mathbf{K}_i \boldsymbol{\Lambda}_i (\bar{\mathbf{g}}_{ci} - \bar{\mathbf{u}}_{ci}) \tag{5.8}$$

where the slip condition is imposed on $\bar{\mathbf{g}}_c$. Since $\tilde{\mathbf{u}}$ is chosen such that $\tilde{v}_n = 0$, one has that

$$\boldsymbol{\Lambda}_i(\bar{\mathbf{g}}_{ci} + \bar{\mathbf{u}}_{ci}) = \frac{1}{\bar{\rho}_i \bar{c}_i \sqrt{2}} \begin{bmatrix} -\tilde{c} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \tilde{c} \end{bmatrix} \left(\begin{bmatrix} \bar{p}_i \\ \bar{\theta}(\bar{p}_i - \bar{\rho}_i \bar{c}_i^2) \\ \bar{\rho}_i \bar{c}_i \sqrt{2}\, \bar{v}_{ti} \\ \bar{p}_i \end{bmatrix} - \begin{bmatrix} \bar{p}_i - \bar{\rho}_i \bar{c}_i \bar{v}_{ni} \\ \bar{\theta}_i(\bar{p}_i - \bar{\rho}_i \bar{c}_i^2) \\ \bar{\rho}_i \bar{c}_i \sqrt{2}\, \bar{v}_{ti} \\ \bar{p} + \bar{\rho}\bar{c}\bar{v}_{ni} \end{bmatrix} \right) = -\frac{\tilde{c}\bar{v}_{ni}}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}. \tag{5.9}$$

Then, the summation term for the fixed wall becomes

$$\begin{aligned} \bar{\mathbf{u}}_{ci}^T \boldsymbol{\Lambda}_i \bar{\mathbf{u}}_{ci} + 2\bar{\mathbf{u}}_i^T \mathbf{b}_i &= \frac{\tilde{c}}{2\bar{\rho}_i^2 \bar{c}_i^2} ((\bar{p}_i + \bar{\rho}_i \bar{c}_i \bar{v}_{ni})^2 - (\bar{p}_i - \bar{\rho}_i \bar{c}_i \bar{v}_{ni})^2) \\ &\quad - \frac{\tilde{c}\bar{v}_{ni}}{\bar{\rho}_i \bar{c}_i} ((\bar{p}_i + \bar{\rho}_i \bar{c}_i \bar{v}_{ni}) + (\bar{p}_i - \bar{\rho}_i \bar{c}_i \bar{v}_{ni})) = 0, \end{aligned} \tag{5.10}$$

and as in the continuous case the contribution from the solid wall is identically equal to zero. Thus (5.7) becomes (4.56) because of the choice of $\tilde{\mathbf{u}}$ evaluated at some point on

$\partial\Omega_F$. If $\mathbf{R} \leq 0$, strict stability is achieved in the sense of definition 6. It is possible to state the penalty term (5.8) like (4.57), where a scalar $\delta \in \mathbb{R}^+$ regulates the influence of the BCs. For the fixed wall, the diagonal matrix gives that the first and last element in $\bar{\mathbf{g}}_{ci} - \bar{\mathbf{u}}_{ci}$ alone have to be considered and that the resulting vector may be written as

$$\mathbf{b}_i = \begin{cases} \delta\left[\bar{\mathbf{u}}_{ci}(1) - \bar{\mathbf{u}}_{ci}(4),\, 0,\, 0,\, \bar{\mathbf{u}}_{ci}(1) - \bar{\mathbf{u}}_{ci}(4)\right]^T & \text{at boundary with BC,} \\ \mathbf{0} & \text{otherwise,} \end{cases} \tag{5.11}$$

where $\mathbf{K}$ has been left out and all scalars are gathered in δ. Then, the contribution to (5.7) from the fixed wall $\partial\Omega_F$ becomes identically zero for all $i \in \mathcal{E}(\partial\Omega)$ if $\delta = \frac{\tilde{c}}{2}$.

5.2 Boundary conditions for the nonlinear Euler equations in two dimensions

Since the system matrices $\mathbf{A}(\mathbf{u})$ and $\mathbf{B}(\mathbf{u})$ for the nonlinear Euler equations (2.17) are not constant, one does not obtain the same BCs as for the linearised case. Recall that only the conservative formulation gives the integral form (2.25) needed for the FVM. The matrices still need to be in symmetrised primitive form, see (2.23), to express them with eigenvectors and eigenvalues. From this form the negative eigenvalues can be singled out and the ingoing characteristics can be prescribed. Thus in order to obtain the characteristic BCs, similarity transforms with the matrices $\mathbf{M}$ and $\mathbf{S}$, see (2.18) and (2.22) respectively, must be applied to the integrand of the boundary integral in (2.26):

$$\begin{aligned}
\oint_{\partial\Omega} [\mathbf{A}(\mathbf{u})\mathbf{u}, \mathbf{B}(\mathbf{u})\mathbf{u}]^T \cdot \hat{\mathbf{n}}\, ds &= \oint_{\partial\Omega} (\mathbf{A}(\mathbf{u})\hat{\mathbf{x}} + \mathbf{B}(\mathbf{u})\hat{\mathbf{y}}) \cdot \hat{\mathbf{n}}\, \mathbf{u}\, ds \\
&= \oint_{\partial\Omega} \mathbf{MSS}^{-1}\mathbf{M}^{-1}\left(\mathbf{A}(\mathbf{u})\hat{\mathbf{x}} + \mathbf{B}(\mathbf{u})\hat{\mathbf{y}}\right) \cdot \hat{\mathbf{n}}\, \mathbf{MSS}^{-1}\mathbf{M}^{-1}\, \mathbf{u}\, ds \\
&= \oint_{\partial\Omega} \mathbf{Q}\left(\mathbf{A}_s(\mathbf{u}_s)\hat{\mathbf{x}} + \mathbf{B}_s(\mathbf{u}_s)\hat{\mathbf{y}}\right) \cdot \hat{\mathbf{n}}\, \mathbf{Q}^{-1}\, \mathbf{u}\, ds \\
&= \oint_{\partial\Omega} \mathbf{QK\Lambda K}^T\mathbf{Q}^{-1}\, \mathbf{u}\, ds \\
&= \oint_{\partial\Omega} \mathbf{QK\Lambda}^+\mathbf{K}^T\mathbf{Q}^{-1}\, \mathbf{u}\, ds + \oint_{\partial\Omega} \mathbf{QK\Lambda}^-\mathbf{K}^T\mathbf{Q}^{-1}\, \mathbf{u}\, ds,
\end{aligned} \tag{5.12}$$

where $\mathbf{Q} = \mathbf{MS}$. The integrands in the last row of (5.12) are the characteristic BCs for the nonlinear Euler equations. As discussed in section 3.4.2 the structures of $\mathbf{\Lambda}^+$ and $\mathbf{\Lambda}^-$ depend on whether the boundary is an inflow or outflow. For a boundary with the

inflow BC the fluxes are given by

$$\mathbf{QK\Lambda^+K}^T\mathbf{Q}^{-1}\mathbf{u} = \frac{\rho(v_n + c)}{2\gamma} \begin{bmatrix} 1 \\ v_1 + \alpha c \\ v_2 + \beta c \\ v_n c + \frac{|\boldsymbol{v}|^2}{2} + \frac{c^2}{\gamma - 1} \end{bmatrix}, \tag{5.13}$$

$$\mathbf{QK\Lambda^-K}^T\mathbf{Q}^{-1}\mathbf{u} = \frac{\rho}{2\gamma} \begin{bmatrix} v_n(2\gamma - 1) - c \\ v_1(v_n(2\gamma - 1) - c) - \alpha(cv_n - c^2) \\ v_2(v_n(2\gamma - 1) - c) - \beta(cv_n - c^2) \\ \frac{(1-\gamma)(2v_n^2 c + |v_n|^2 c) + 2\gamma v_n(c^2 + \gamma|v_n|^2) + v_n|v_n|^2(1-3\gamma) - 2c^3}{2(\gamma - 1)} \end{bmatrix} \tag{5.14}$$

and the corresponding for outflow are

$$\mathbf{QK\Lambda^+K}^T\mathbf{Q}^{-1}\mathbf{u} = \frac{\rho}{2\gamma} \begin{bmatrix} v_n(2\gamma - 1) + c \\ v_1(v_n(2\gamma - 1) + c) + \alpha(cv_n + c^2) \\ v_2(v_n(2\gamma - 1) + c) + \beta(cv_n + c^2) \\ \frac{(\gamma-1)(2v_n^2 c + |v_n|^2 c) + 2\gamma v_n(c^2 + \gamma|v_n|^2) + v_n|v_n|^2(1-3\gamma) + 2c^3}{2(\gamma - 1)} \end{bmatrix}, \tag{5.15}$$

$$\mathbf{QK\Lambda^-K}^T\mathbf{Q}^{-1}\mathbf{u} = \frac{\rho(v_n - c)}{2\gamma} \begin{bmatrix} 1 \\ v_1 - \alpha c \\ v_2 - \beta c \\ -v_n c + \frac{|\boldsymbol{v}|^2}{2} + \frac{c^2}{\gamma - 1} \end{bmatrix}. \tag{5.16}$$

If the boundary has the fixed wall BC, then $\mathbf{QK\Lambda K}^T\mathbf{Q}^{-1}\mathbf{u} = [0, n_1 p, \ldots, n_d p, 0]^T$ is used instead. It is the same flux as used in section 3.4.1. This is no surprise, since

$$\mathbf{QK\Lambda K}^T\mathbf{Q}^{-1}\mathbf{u} = [\mathbf{A}(\mathbf{u})\mathbf{u}, \mathbf{B}(\mathbf{u})\mathbf{u}]^T \cdot \hat{\mathbf{n}} = \mathbf{f}^c(\mathbf{u}) \cdot \hat{\mathbf{n}}, \tag{5.17}$$

where $\mathbf{u}$ must satisfy $v_n = 0$.

5.3 Implementation

The nonlinear Euler equations with weak characteristic BCs are solved numerically with the node-centred FVM described in section 3.3.1. Implementation of the weak characteristic BCs for the nonlinear Euler equations is quite straight forward. The in- and outflow fluxes over the edges are set to be the combination of in- and outgoing characteristics, where reference data for the undisturbed free flow is used to evaluate $\mathbf{QK\Lambda^-K}^T\mathbf{Q}^{-1}\mathbf{u}$ and the current local numerical solution for $\mathbf{QK\Lambda^+K}^T\mathbf{Q}^{-1}\mathbf{u}$. For fixed walls the same flux function as for weak prescribed fluxes is used, see (3.5). The implementation of

weak prescribed fluxes, which will be used for comparison, is similar. The in- and out-flow fluxes over the exterior edges are evaluated with a numerical flux function that takes reference data and the current local numerical solution as arguments.

Investigating the performance of the node-centred FVM for the main problem, with the method described in chapter 4, is not within the scope of this thesis and therefore it is not implemented. However, the implementation of the scheme is the same as for the nonlinear case. The corresponding expressions for the in- and outflow BCs (5.13) to (5.14) and (5.15) to (5.16) are found in appendix C. One may choose some other numerical flux function than (4.13), as long as the discrete differential operators $\mathbf{Q}_x$ and $\mathbf{Q}_y$ are skew-symmetric for the interior points. Otherwise the estimate of $\frac{d}{dt}\|\bar{\mathbf{u}}\|^2_{\mathbf{I}\otimes\mathbf{P}}$ also considers the interior points and does not obtain the form (4.56), meaning that theorem 1 may not apply [1].

5.4   Numerical experiments

In [1] the relevance of the analysis of the linearised Euler equations for the full nonlinear problem (2.17) was investigated by simulating the steady airflow at Ma 0.5 around a NACA0012 wing-profile for the nonlinear system. These results were compared to those obtained with strongly imposed BCs, see [1], and the weakly imposed characteristic BCs were superior. The 10:th-logarithm of the norm of the residual converged after $3 \cdot 10^5$ time steps to an order 10^{-14} , which is the smallest number that can be represented by a double-precision binary floating-point, while the strongly imposed BCs converged to 10^{-5} after 10^5 time steps. Therefore the main objective is to reproduce this result, where the weakly imposed characteristic BCs excels compared some other BCs. Instead of strongly imposed BCs, the weak prescribed fluxes are used for comparison, see 3.4. Further, the $L^2(\Omega)$-norm of the numerical solution will be measured to investigate if strict stability, see (4.1), holds for the numerical solution for the nonlinear Euler equations, even though no such claim has been made.

In addition, the weakly implemented characteristic BCs' performance are in this thesis compared to weakly implemented prescribed fluxes'. First the NACA0012 wing-profile mentioned above is considered, but on smaller grids. If the size of the computational domain can be decreased, fewer unknowns have to be solved for. This is possible if the BCs are a satisfactory model of the near-field, therefore the smaller domains are made such that they have the same resolution as the central part of the larger domain.

For the Shu-vortex, the analytical solution is known and thus gives yet another tool to measure the BCs' performance. This test case also investigates how the different BCs

handle large gradients and due to the simplicity of the problem, numerical artifacts are easily identified.

The numerical experiments were performed with the Tau 2D inhouse code of the University of Kassel, a numerical solver for the two-dimensional Navier-Stokes and Euler equations. It uses the node-centred FVM and takes several input arguments specifying what methods and parameters to use. For all simulations, AUSMDV will be used as numerical flux function. It approximates the flux directly as a sum of several fluxes from other numerical flux functions. The different properties of these functions give in total a function that is more than its parts.

In chapter 3 piecewise reconstruction was briefly mentioned. For some of the numerical experiments *linear TVD reconstruction* will be used. On each cell the constant solution is reconstructed into a linear solution, based on the neighbouring states. The slope is modified with a limiter called the *Barth-Jesperson* limiter. As a result the order of the method is increased from one to two.

When implicit Euler is used for time marching, a system of equations has to be solved to obtain an update of the solution. This is done with a *Jacobian free Newton-Krylov method*, where the system of equations is linearised and the given multiplication of a Jacobian matrix with a vector is approximated with a finite-difference. The linear systems are solved using GMRES.

The speed of convergence depends heavily on the matrix for the system of equations given by implicit Euler. By using *preconditioning*, the linear system is transformed into an equivalent one, but with a structure that gives faster convergence. In the numerical experiments *ILU preconditioning* with physical renumbering will be used. For more information on AUSMDV, reconstruction, the Jacobian free Newton-Krylov method, GMRES and ILU preconditioning, please refer to [2] and the references therein.

The numerical results were visualised with Paraview, an open-source, multi-platform data analysis and visualisation application. The plots consist for the most part of the visualisation of pressure or velocity data with contour plots. For the pressure contour plots the difference in pressure between two isolines is 0.5 Pa. The grids were generated by the grid generation tool that is part of the Tau 2D code's toolbox. These grids are the primary grid, then Tau 2D code's solver constructs the dual grid. The weak prescribed fluxes were already implemented in the Tau 2D code. The results it provides for reasonably posed problems are trusted and will be used as reference.

5.4.1 The NACA0012 wing-profile

The NACA0012 wing-profile, compared to wings on real life air planes, is very simple. As can be seen in figure 5.1 it is symmetric around its central line and it is well suited for numerical experiments due to its uncomplicated structure. For further information about the NACA0012 wing-profile, visit NASA's website on the subject [14].

The simulations were performed on a grid of radius 20 and on a smaller one with radius 1, with a wing length of 1, both consisting of triangles. The input parameters for the grid generation tool can be found in table 5.1. In figure 5.1 the area around the wing-profile on the radius 20 grid is shown. The numerical experiments for the NACA0012 wing-profile were performed for the steady Euler equations, with the implicit Euler method with a matrix free solver for time-marching and AUSMDV as numerical flux function. The input parameters for the Tau 2D code-solver can be found in table 5.2. The wing is given the fixed wall BC, the left- and right sides of the numerical boundary are set to be an inflow and an outflow, respectively.
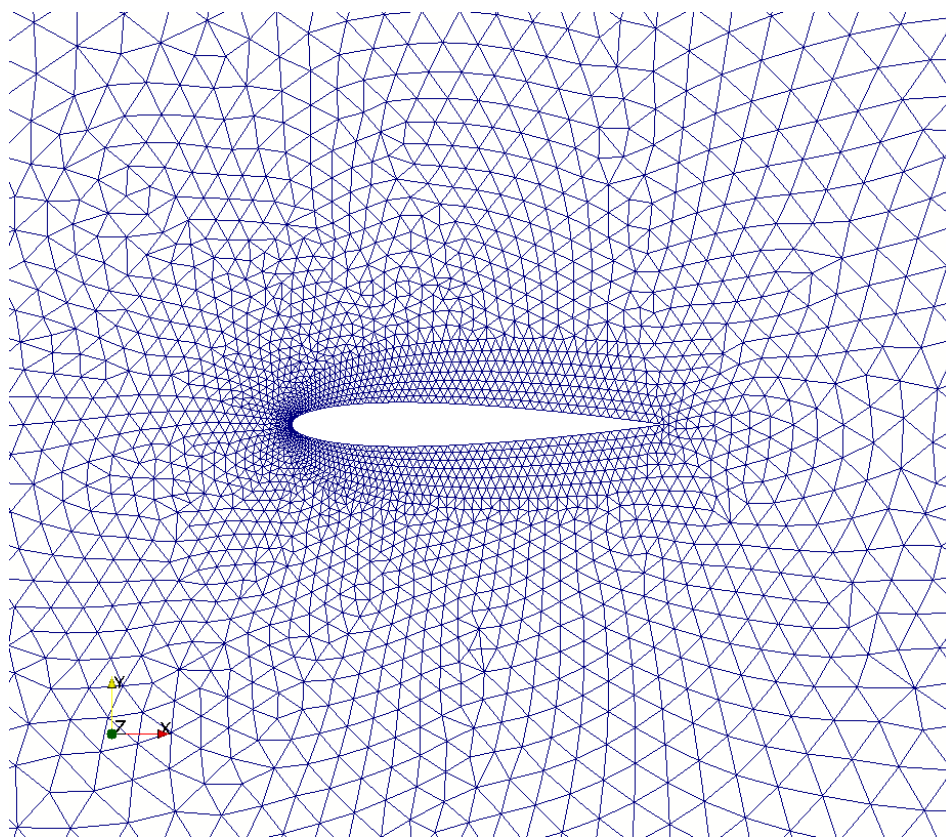
FIGURE 5.1: The NACA0012 wing-profile used for the simulations. Here on an unstructured primary grid.

Streching	0
Max_No_of_steps	100
Orientation	Right, Right
Number of wing nodes	120
Number of boundary nodes for $R = 20$	40
Total number of nodes for $R = 20$	4794
Number of boundary nodes for $R = 1$	40
Total number of nodes for $R = 1$	2404

TABLE 5.1: The NACA0012 wing-profile's grid parameters for the Tau 2D Code's grid generation toolbox. R denotes radius of the computational domain.

Preconditioner	1
preconupdates	0
physical renumbering	1
matrix free solver	0
Krylov dimensions	40
CFL number	5.0
CFL max number	500
Angle of attack	1.25
Reference value ρ	1
Reference value v_1	0.999762
Reference value v_2	0.0218149
Reference value p	0.988631

TABLE 5.2: Input for the NACA0012 wing-profile simulations.

Reconstruction	Piecewise linear
Reconstruction type	TVD
Limiter	Barth Jespersen

TABLE 5.3: Input for the NACA0012 wing-profile simulations with linear reconstruction.

The results given by the Tau 2D code for simulations at Ma 0.5 are found in figure 5.2. There the 10:th-logarithm of the residuals from simulations with weak characteristic BCs given by (5.13) to (5.16) and weak prescribed fluxes from 3.4, respectively, are plotted against the number of time steps taken. For the weak prescribed fluxes, -3 is reached

after approximately 0.05×10^6 time steps and thereafter it converges very slowly. In total 4×10^6 time steps are needed to obtain -11.

The weak characteristic BCs on the other hand arrives at -3 after the same amount of time steps as the weak prescribed fluxes, but continues to converge asymptotically until 2.5×10^5 time steps are taken. Then the curve flattens and the logarithm of the residual converges to -11. Thus weak characteristic BCs are superior compared to the weak prescribed fluxes, since the convergence is too slow once -3 is reached. Still -14 was not obtained for the weak characteristic BCs, but the code might not be able to provide solutions of precision 10^{-14}. In fact, the precision one can expect is the entire algorithm's condition number times 10^{-14}. The condition number is unknown, but if it is 1 then the condition number for the algorithm with one of the weak BCs is around 10^{11}. The possibility exists that the residuals actually converged to 10^{-11} and are not limited by the code's precision. Still, a precision of 10^{-11} is in many cases redundant since real world measurements are often cruder.

The result in [1] is confirmed with figure 5.2 and the main aim of the thesis is thus fulfilled. Further, convergence for a scheme treating the steady nonlinear Euler equations can indeed be shown by obtaining strict stability for the linearised problem, as claimed in [1]. But as will be shown below, this holds only for subsonic velocities.
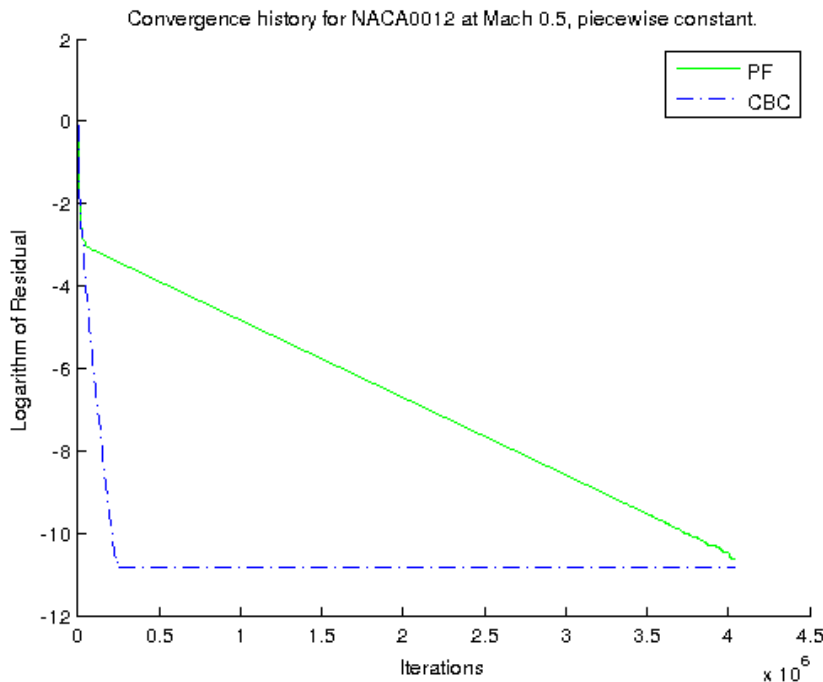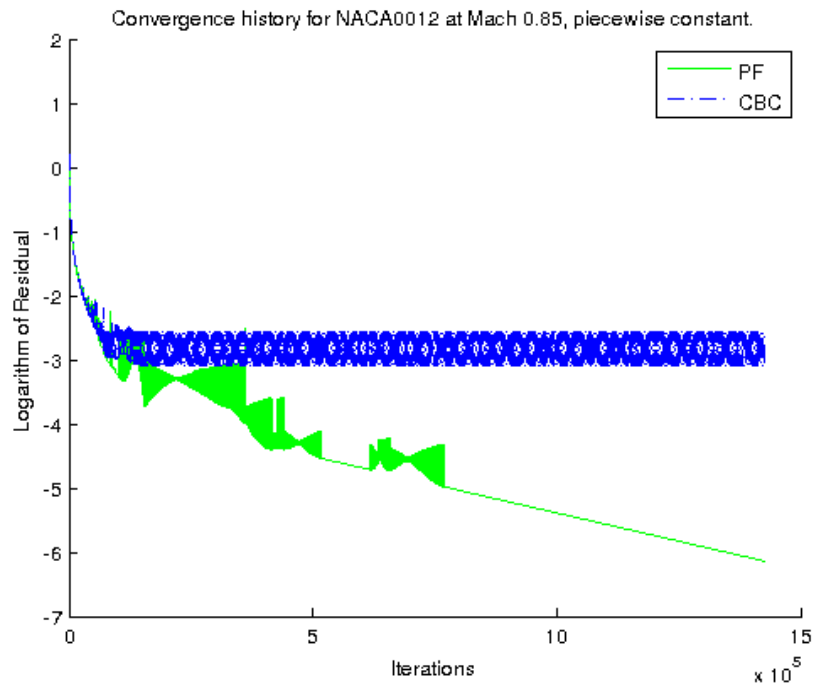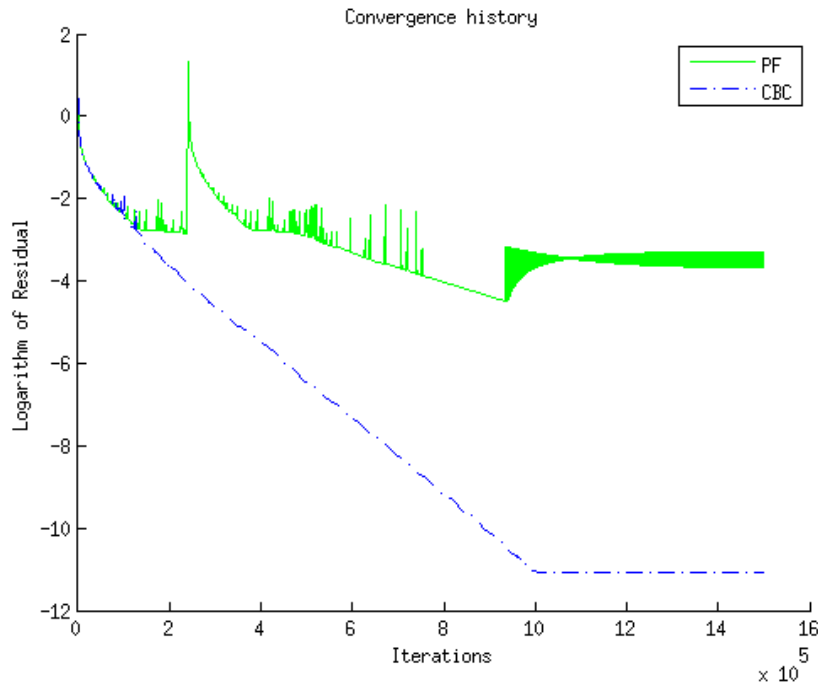

FIGURE 5.2: The logarithm of the residual for weak characteristic BCs (blue, .-) and weak prescribed fluxes (green, -) plotted against the number of time steps taken.

If the simulations described above instead are preformed at Ma 0.85, the results become very different, see figure 5.3. As before, the 10:th-logarithm of the residuals of the results given by simulations with weak characteristic BCs and weak prescribed fluxes are plotted against the number of time steps taken. Both curves oscillate heavily, the former around -2.7 and the latter until it reaches -7.5. Thereafter it converges asymptotically, but slowly. Even though the residual for simulations with weak prescribed fluxes decreases as more time steps are taken, the oscillatory behaviour makes the results unpredictable. Also, the convergence speed is too slow.

A Mach number of 0.85 is in the transonic range, thus shocks along the wing-profile are expected [10]. At the shocks the difference between the cell-wise constant solutions on each side of a shock may be of great magnitude. Hence the solver might have trouble finding a solution along the shocks, which influences the entire solution and could cause the oscillating residuals.

FIGURE 5.3: The logarithm of the residual for weak characteristic BCs (blue, .-) and weak prescribed fluxes (green, -) plotted against the number of time steps taken.

The FVM is of order one, but if piecewise linear reconstruction is used, the order increases to two [2]. Simulations with piece wise constant solutions for Ma 0.85 were done with the input parameters from the tables 5.2 and 5.3, where the latter is for the reconstruction.

In figure 5.4 the 10:th-logarithm of the residuals for the simulations with weak characteristic BCs and the weak prescribed fluxes, respectively, are plotted against the number

of time steps taken. For curve corresponding to the weak prescribed fluxes first decrease to 2×10^5 time steps, then peak only to quite unsteady go past -4 and finally oscillate around -3 after 10×10^5 time steps. The results from simulations with weak characteristic BCs decrease linearly with a slope around -1.1 and oscillate during the first 2×10^5 time steps. After 10×10^5 time steps the slope suddenly becomes zero and the residual converges to -11.

FIGURE 5.4: The logarithm of the residual for weak characteristic BCs (blue, .-) and weak prescribed fluxes (green, -) plotted against the number of time steps taken.

The increment of the order from one to two thus has a huge impact on the convergence of the residuals. The oscillatory behaviour, for the residual given by simulations with weak BCs, is gone and the curve in figure 5.2 behaves as the corresponding curve in figure 5.4, but reaches -11 after 7.5×10^5 time steps more. The analysis in chapter 4 only consideres piecewise constant solutions, meaning no claim about the convergence of the residual for the node-centred FVM with piecewise linear solutions has been made. Further it is remarkable that the performance for simulations with weak prescribed fluxes is still lacking.

So it has been shown that the weak characteristic BCs are favourable over weak prescribed fluxes, due to the convergence of the residual. But what about the actual results? The solutions obtained after 10000 time steps for the different weak BCs are plotted in the figures 5.5 and 5.6, which contain only the relevant part of the large circular computational domain of radius 20. The velocity is Ma 0.85 and linear reconstruction was

used, see table 5.3. The scale gives the data range for the pressure, which measures from 0.4939 to 1.625 Pa. No differences are discernible between the simulation with weak prescribed fluxes in 5.5 and the simulations with weak characteristic BCs in 5.6. The result given by the Tau 2D code with weak prescribed fluxes on such a large computational domain is considered to be correct and will be referred to as the *reference case*. It therefore validates the result given by the weakly imposed characteristic BCs. Another important observation are the expected shocks above and beneath the wing profile, which could be an explanation for the poor performance of the weak BCs at Ma 0.85 with piecewise constant solutions. The shocks are identified by the high density of isolines, which implies huge gradients.

No big gradients are close to the edges in the figures 5.5 and 5.6, implying that the situation on the boundary at least is just as uneventful. To test the performance of the weak BCs, the radius is decreased in order to place the boundary of the domain in the vicinity of the wing-profile.
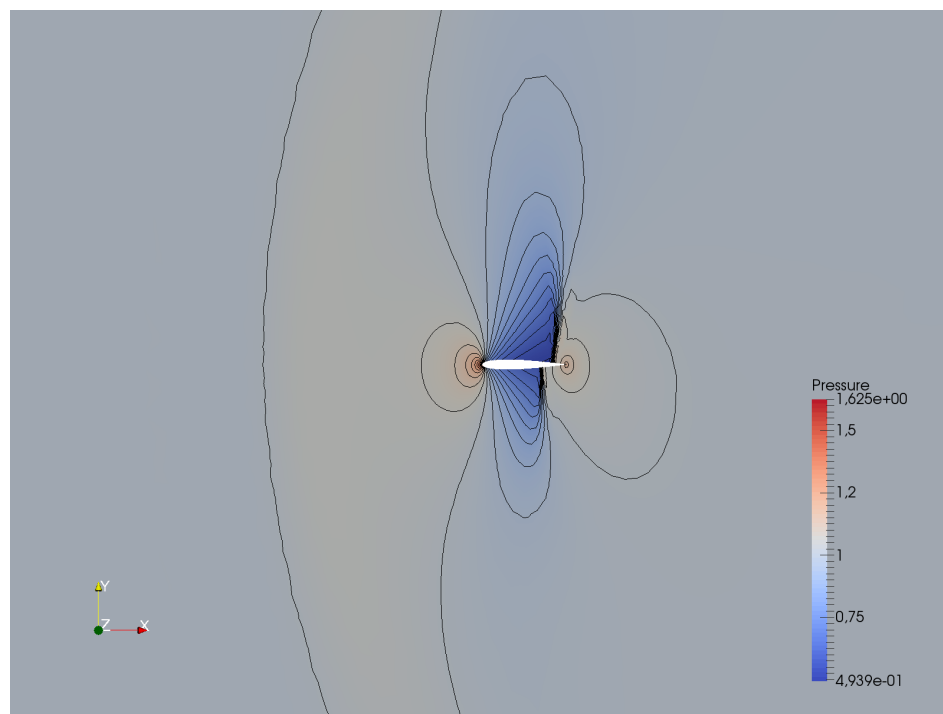
FIGURE 5.5: The NACA0012 wing-profile with pressure isolines on the radius 20 grid with weak prescribed fluxes.

The figures 5.7 and 5.8 show the pressure distribution for weak prescribed fluxes and weak characteristic BCs on the smaller grids with radius 1. The velocity is set to Ma 0.85 and the input parameters are found in the tables 5.2 and 5.3, since piecewise linear solutions are considered. For reference, the field around the wing profile in the reference case 5.5 is enlarged and shown in 5.9.
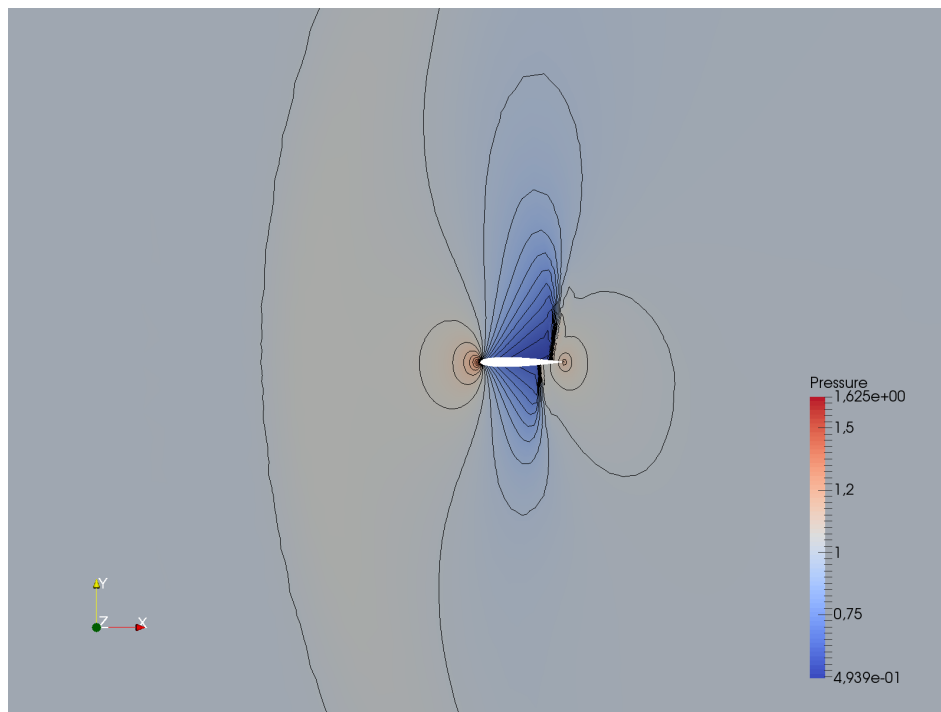
FIGURE 5.6: The NACA0012 wing-profile with pressure isolines on the radius 20 grid with weak characteristic BCs.
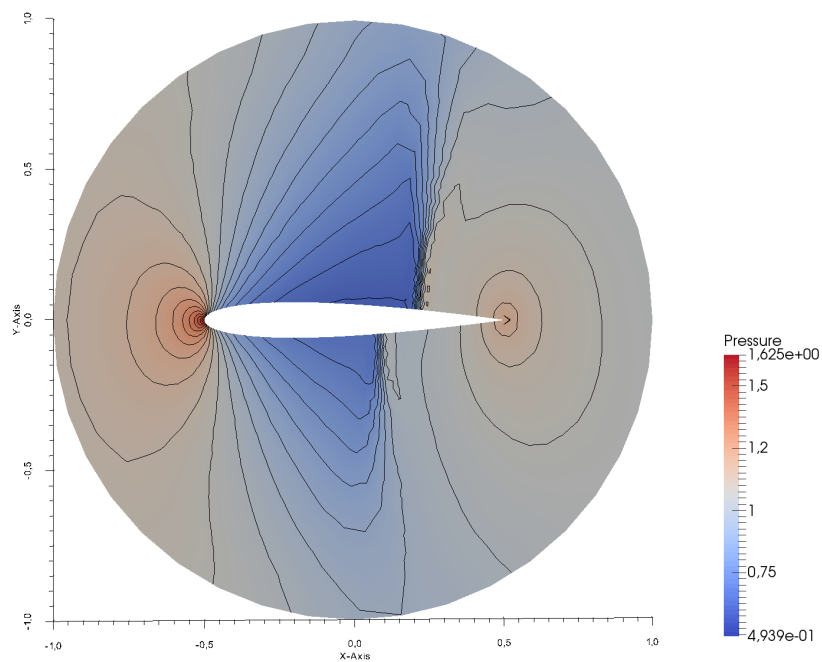
FIGURE 5.7: The NACA0012 wing-profile with pressure isolines on the radius 1 grid with weak prescribed fluxes.
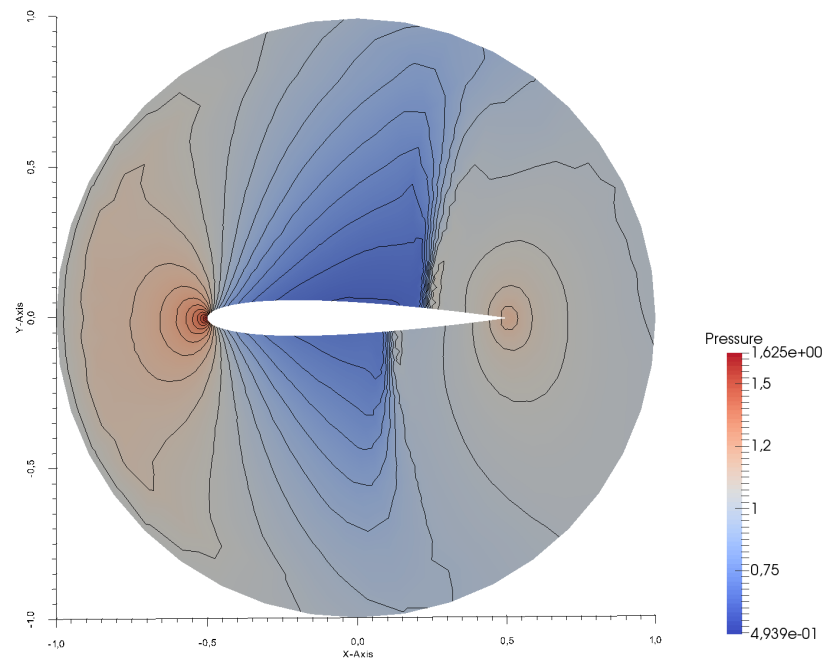
FIGURE 5.8: The NACA0012 wing-profile with pressure isolines on the radius 1 grid with weak characteristic BCs.
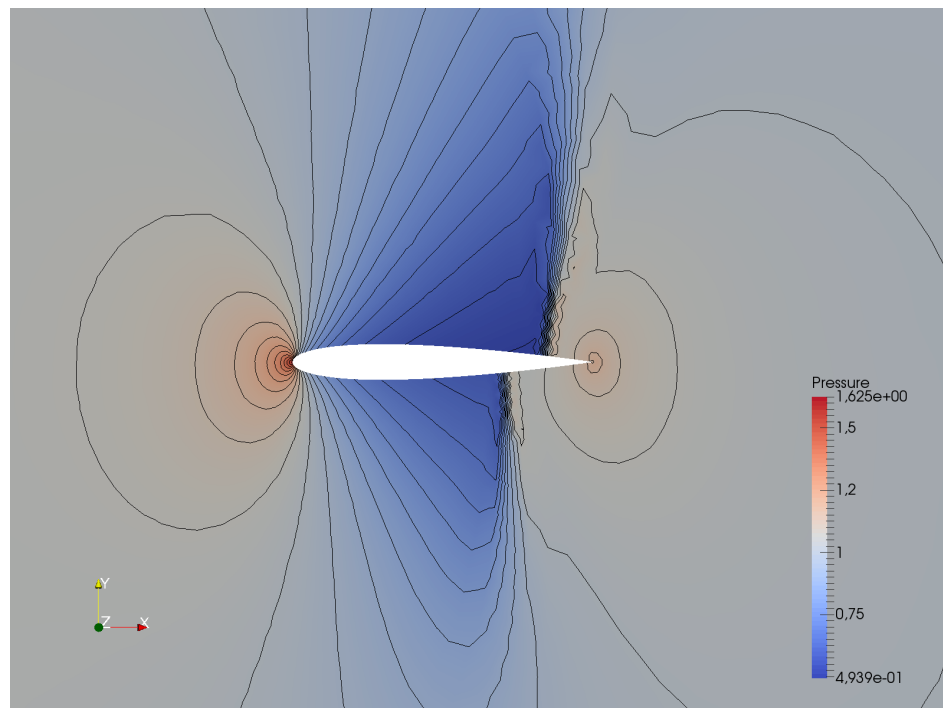

FIGURE 5.9: The enlarged image of the NACA0012 wing-profile with pressure isolines on the radius 20 grid with weak prescribed fluxes.

Both the weak prescribed fluxes and the weak characteristic BCs have larger fields with higher pressure at the front and back of the wing than the reference case. The field of increased pressure is slightly larger for the weak characteristic BCs. Another observation is that the weak characteristic BCs deform the isolines close to the boundary. The isolines on top of the wing in figures 5.7 and 5.9 are mostly straight, but with weak characteristic BCs they are slightly bent along the curvature of the boundary.

Shocks above and beneath the wing-profile are visible for all simulations. The distances from the shocks and to the back of the wing are summarised in table 5.4. For both the weak prescribed fluxes and weak characteristic BCs the distance is greater than for the reference case, but the weak characteristic BCs are slightly closer. None of the cases have as big gradients at the shock locations as the reference case, an observation based on the fact that the isolines are fewer at these points. Still, both the weak characteristic BCs and the weak prescribed fluxes have the same amount of isolines at the shock locations, indicating that their gradients at these points are approximately of the same size.

The locations of the shocks are important, since shocks are directly linked to the amount of force that opposes an object's motion through the air, called drag [10]. The weak characteristic BCs performed better when it came to the positioning of the shocks, as it was closer to the reference case than the weak prescribed fluxes. Thus it captures the flow of air in the direct vicinity of the wing better and therefore is favourable. The weak prescribed fluxes on the other hand still perform above expectations. If the deviation from the result given by the full scale is worth the price depends on what one seeks to simulate. Smaller systems require fewer operations in order to obtain a solution. In addition, it is possible to use the same amount of triangles on the small grid as for the full scale problem. Then computational power is not spent on obtaining a solution for far away cells that are not of interest.

	Above	Below
weak characteristic BCs	0.2984	0.4194
weak prescribed fluxes	0.3140	0.4380
Reference	0.1923	0.3231

TABLE 5.4: Distance from the shock to the back of the wing. The length of the wing is 1.

During the simulations for Ma 0.5 with piecewise constant solutions $\|\bar{\mathbf{u}}\|^2_{\mathbf{I}\otimes\mathbf{P}}$ was measured and the result is presented in figure 5.10. The squared discrete $L^2(\Omega)$-norm increases over the last 6000 time steps. But an increase from 75600.75 to 75605.5 corresponds to 0.063‰ of 75600.75, meaning that numerical errors should not be ruled out. Therefore it is not entirely clear whether $\|\bar{\mathbf{u}}\|^2_{\mathbf{I}\otimes\mathbf{P}} \leq 0$ holds for the nonlinear case, eventhough this observation indicates that it might hold true. Little can be said about this

result, as no theory supporting any connection between the stability of the two problems is available.
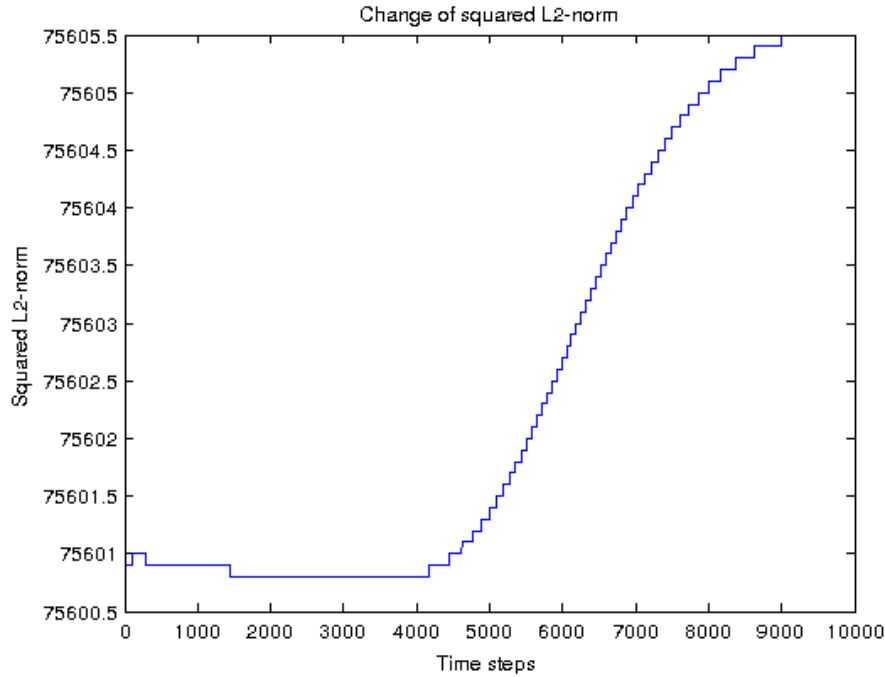

FIGURE 5.10: The change of $\|\bar{\mathbf{u}}\|_{\mathbf{I}\otimes\mathbf{P}}^2$ for the NACA0012 wing-profile with weak characteristic BCs, plotted against the number of time steps taken.

5.4.2 The Shu-vortex

In the Shu-vortex test case a vortex, which is designed to be isentropic, is travelling by convection from the middle to the right in a two-dimensional wind tunnel [2]. In an ideal case, the vortex would simply be absorbed by the boundary when leaving the domain, but instead artifacts may be created, which are pure numerical effects. For the two-dimensional Euler equations the exact solution is known, making the Shu-vortex case ideal for the study of the BCs' performance.

The initial state is the free stream reference values $(\hat{\rho}, \hat{v}_x, \hat{v}_y, \hat{p}) = (1, 0.5, 0, 1)$ perturbed by a vortex, centred at $(\tilde{x}, \tilde{y}) = (0, 0)$, given by the formulas

$$
\begin{aligned}
\delta v_1 &= -\frac{\zeta}{2\pi}(y - \tilde{y})e^{\phi(1-r^2)}, \\
\delta v_2 &= \frac{\zeta}{2\pi}(x - \tilde{x})e^{\phi(1-r^2)}, \\
\delta T &= -\frac{\zeta^2(\gamma - 1)}{16\phi\gamma\pi^2}(x - \tilde{x})e^{2\phi(1-r^2)}.
\end{aligned}
\tag{5.18}
$$

where T is temperature. The scalars ζ and ϕ regulate the speed of the flow in the vortex respectively its size. If the numerical scheme performs well, the isolines should be intact circles around the vortex's centre as it travels. Therefore the evaluation of the BC's performance is based on how the isolines deform and if new are created. If the BCs work perfectly, the vortex will simply be absorbed by the domain as if it just passed through it.

The simulations were done with AUSMDV as numerical flux function, time integration was performed with the explicit Euler method and piecewise reconstruction was used to obtain a second order solution. The parameters for the simulations with the Tau 2D code are summarised in table 5.5. The state after one time step of the Shu-vortex with a surface pressure plot can be seen in figure 5.12 and the grid, made up of 23142 triangle elements, in figure 5.11. The input for the grid generation tool is found in table 5.6. Simulations for both weak characteristic BCs and weak prescribed fluxes were made with 10000 time steps, which corresponds 10 seconds and gives the vortex enough time to leave the domain. The results are presented with a pressure contour profile to give a clear picture of the vortex-boundary interaction. Two contour plots for the x-directed velocity will also be presented for simulations with weak characteristic BCs, sampled at 5000 respectively 6500 time steps. The top and bottom part of the domain are given the fixed wall BC, while the left- and right sides are set to be inflow and outflow, respectively.
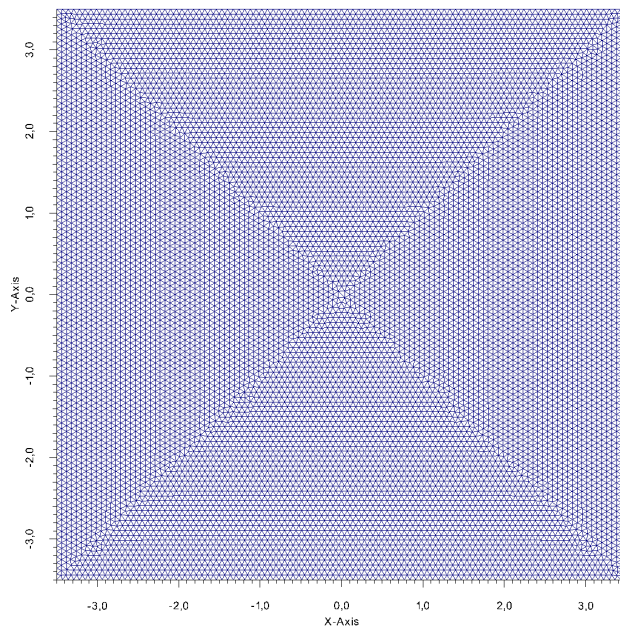
FIGURE 5.11: The square primary grid used for the Shu-vortex simulations.

The simulations for the Shu-vortex were sampled at 5000, 6500, 7500, 8500 and 9500 time steps to capture the process of the vortex passing through the boundary. The weak
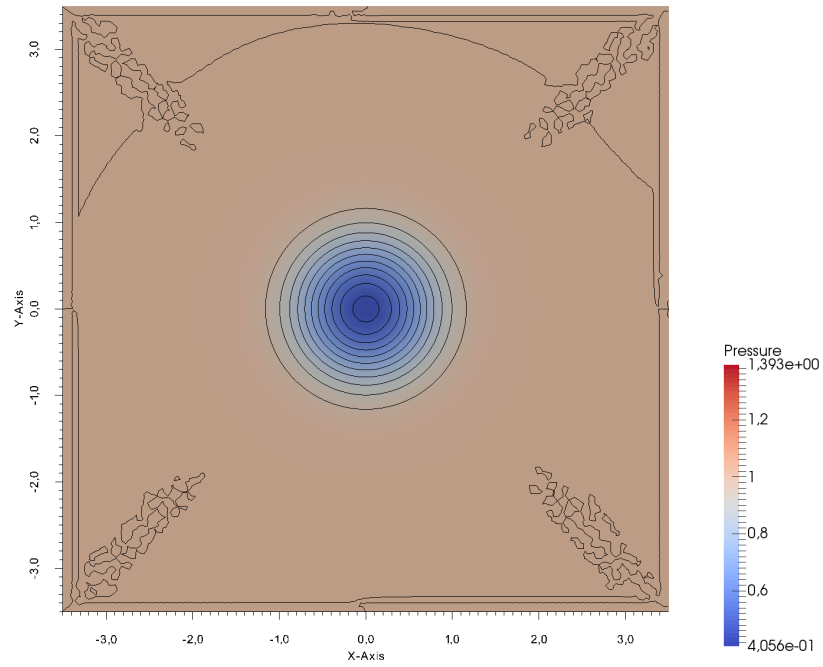
FIGURE 5.12: Contour plot of the pressure for the Shu-vortex on a square domain with weak characteristic BCs after one time step.
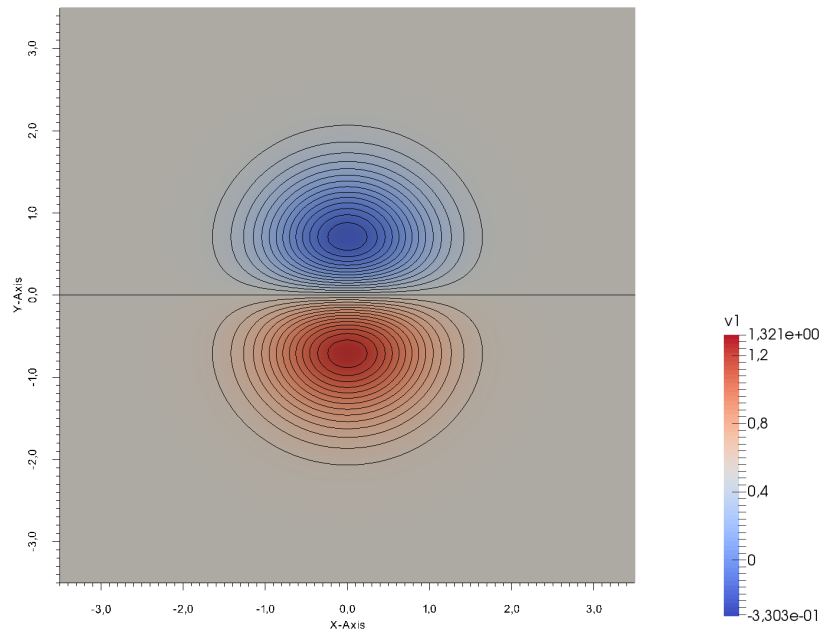

FIGURE 5.13: Contour plot of the x-directed velocity for the Shu-vortex on a square domain with weak characteristic BCs after one time step.

Reconstruction type	TVD
Limiter	Barth Jespersen
Number of timesteps	10000
Mach number	0.42258
Reference value ρ	1
Reference value v_1	0.5
Reference value v_2	0
Reference value p	1

TABLE 5.5: Input for the Shu-vortex simulations.

Streching	0
Max_No_of_steps	100
Orientation	Right
Number of boundary nodes	400
Total number of nodes	11772

TABLE 5.6: The Shu-vortex's grid parameters for the Tau 2D code's grid generation toolbox.

prescribed fluxes' performance is shown in the figures 5.14 to 5.18, the corresponding sequence for the weak characteristic BCs is figure 5.19 to figure 5.23. For all pressure plots the scale goes from 0.4056 to 1.393 Pa. After 5000 time steps the vortex has just reached the boundary, and the part of the isolines not yet in contact with it are symmetric around the vortex's eye, see 5.14 and 5.19. The BCs have not yet played an important role since the vortex has just reached the boundary, but the results give credibility to the solver.

In figure 5.14 the isolines around the vortex's eye deform as they come in contact with the boundary. After 1500 time steps, a field of higher pressure is located above the vortex, as seen in 5.15. The outer isolines around the voxtex's eye have deformed further by spreading out along the boundary. Another observation is that the effects are not symmetrical around the x-axis. In figures 5.16, 5.17 and 5.18 the vortex leaves the domain, but the field of higher pressure from time step 6500 has developed and spread into the domain. There linger isolines for a low pressure field in the vortex's wake.

After 5000 time steps, for the simulation with weak characteristic BC, a high pressure field has developed around the location $(0, -1.0)$, see figure 5.19. It has deformed the isolines closest to the boundary, pushing them up along the edge. The high pressure field grows and intensifies over the next 1500 time steps, totally deforming the vortex's isolines in the direction tangential to the boundary, see figure 5.20. The intensity of the high pressure field is higher than the corresponding field for the weak prescribed fluxess after 6500 time steps. In figure 5.21 the state after 7500 time steps is shown. The high

pressure field artifact and what remains of the vortex are almost equivalent in size and intensity. The vortex's eye is no longer at the x-axis, instead this is now the equilibrium between the vortex and the high pressure field. Both regions have lost intensity and the isolines are fewer, compared to the state at 6500 time steps. In the figures 5.22 and 5.23 the samples at 8500 respective 9500 time steps are shown and it is evident that the number of isolines are much fewer than previously. Two low intensity fields of high respective low pressure remain, which are the ones introduced after 5000 time steps. The fields have now spread, but as indicated by the colour their intensity is lower than for the weak prescribed fluxes after 9500 time steps. In fact, the isolines are fewer for the weak characteristic BCs at this point.

In the Shu-vortex case, the problem is not of steady state type and it is apparent that neither of the BCs simply absorbs the vortex. The weak characteristic BCs give a solution farther from the exact solution than the weak prescribed fluxes when the vortex passes through the boundary. But once the vortex has left the domain, the weak characteristic BCs give a solution closer to the exact one. Judging by the results from the Shu-vortex simulations, the weak prescribed fluxes seem preferable since they create artifacts of lesser magnitude.

The artifact consisting of the high pressure area for the weak characteristic BCs after 6500 could be explained if one studies figure 5.25 where the x-directed velocity is plotted. Here the result does not look quite as bad as the corresponding pressure profile in figure 5.20. One clearly sees that the shape of the isolines is not symmetric around the x-axis anymore. The eye of the vortex is the middle ground between states, but the values along the x-axis and thus in the eye are disturbed by the flux function as the vortex comes in contact with the boundary. Compare figure 5.24, where v_1 is plotted for the simulations with weak characteristic BCs after 5000 time steps, with 5.13. In the latter, the isoline along the x-axis is undisturbed, but that is not the case after 5000 time steps. The same isoline deviates around -11.82 degrees from the x-axis, thus it is possible that the boundary fluxes disturb the symmetry too such an extent that huge artifacts are created. Still, the vortex travels faster than the imbalance can spread, thus just a few isolines linger as remains after 10000 time steps. However, this is just a hypothesis that has not been pursued any further. It has yet to be investigated how weak characteristic BCs may be improved to enhance their performance for unsteady problems.

As for the simulations for the NACA0012 wing-profile, $\|\bar{\mathbf{u}}\|_{\mathbf{I} \otimes \mathbf{P}}$ was measured for each time step for the simulations with weak characteristic BCs, but with linear reconstruction. The results are shown in figure 5.26, where $\|\bar{\mathbf{u}}\|_{\mathbf{I} \otimes \mathbf{P}}$ actually decreases for the first 5000 time steps. At this time the vortex has reached the boundary and $\|\bar{\mathbf{u}}\|_{\mathbf{I} \otimes \mathbf{P}}$ increases until it has left. The increment is just with 7‰, which is small but one hundred
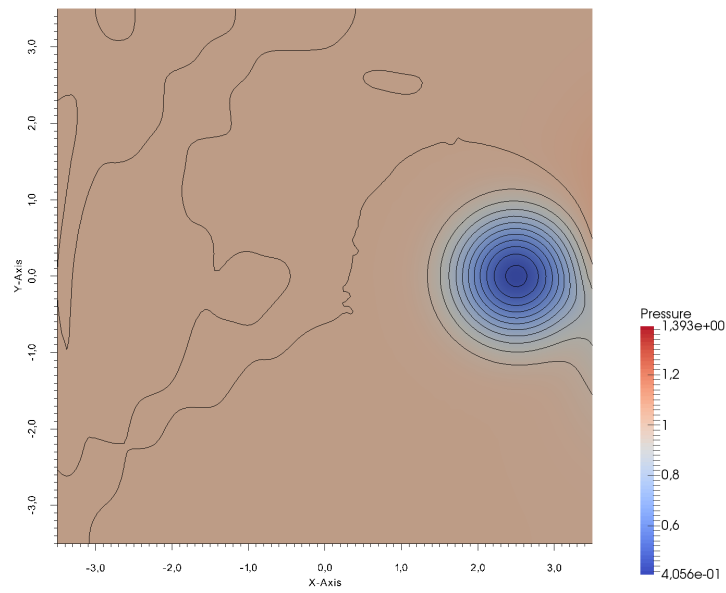
FIGURE 5.14: Contour plot of the pressure for the Shu-vortex on a square domain with weak prescribed fluxes after 5000 time steps.
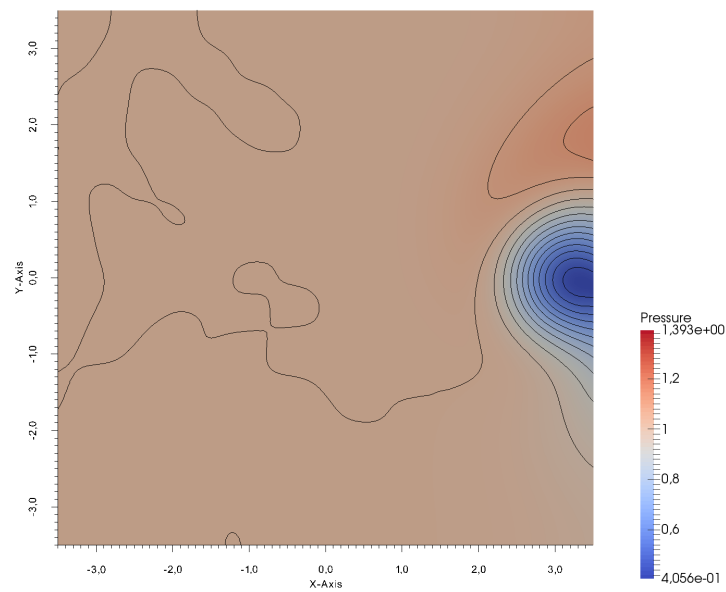

FIGURE 5.15: Contour plot of the pressure for the Shu-vortex on a square domain with weak prescribed fluxes after 6500 time steps.
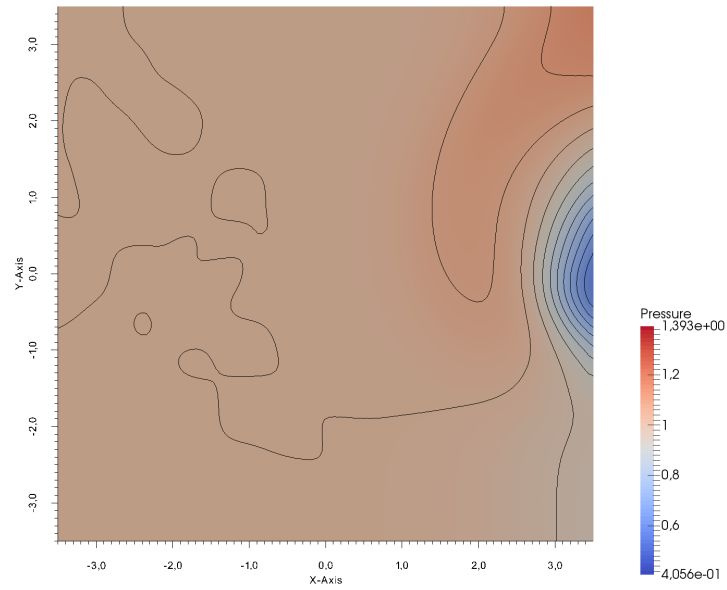
FIGURE 5.16: Contour plot of the pressure for the Shu-vortex on a square domain with weak prescribed fluxes after 7500 time steps.
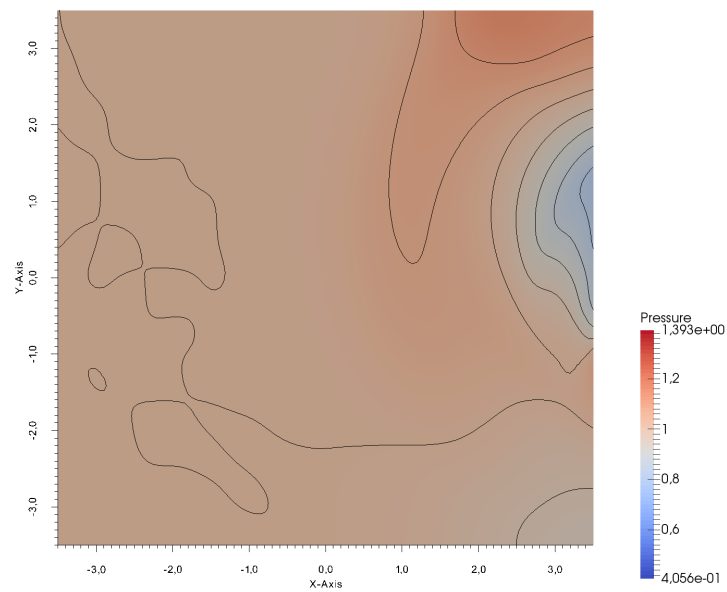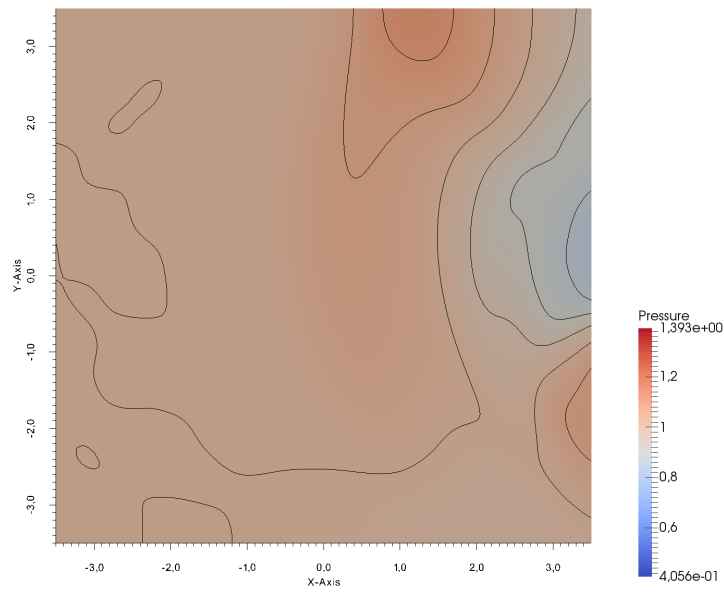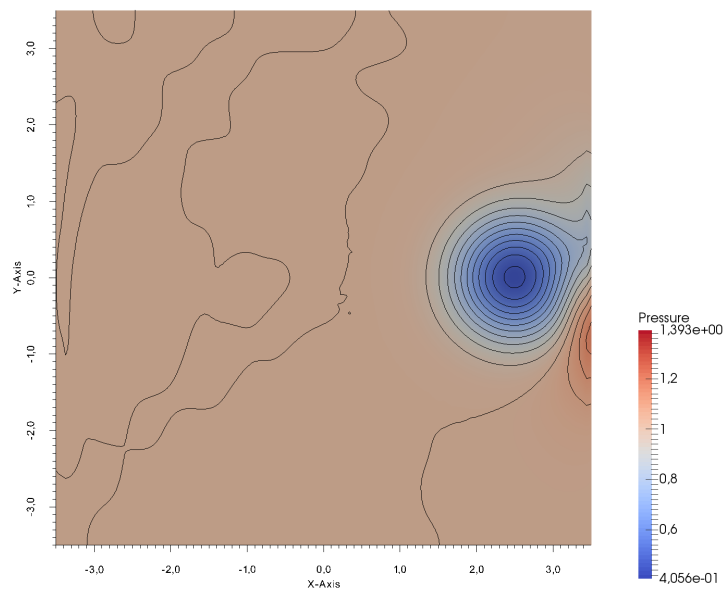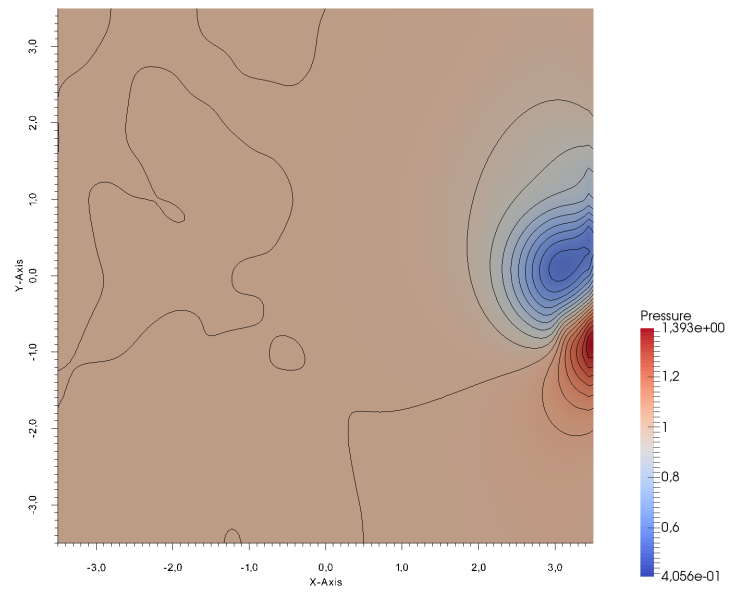

FIGURE 5.17: Contour plot of the pressure for the Shu-vortex on a square domain with weak prescribed fluxes after 8500 time steps.

FIGURE 5.18: Contour plot of the pressure for the Shu-vortex on a square domain with weak prescribed fluxes after 9500 time steps.

FIGURE 5.19: Contour plot of the pressure for the Shu-vortex on a square domain with weak characteristic BCs after 5000 time steps.

FIGURE 5.20: Contour plot of the pressure for the Shu-vortex on a square domain with weak characteristic BCs after 6500 time steps.
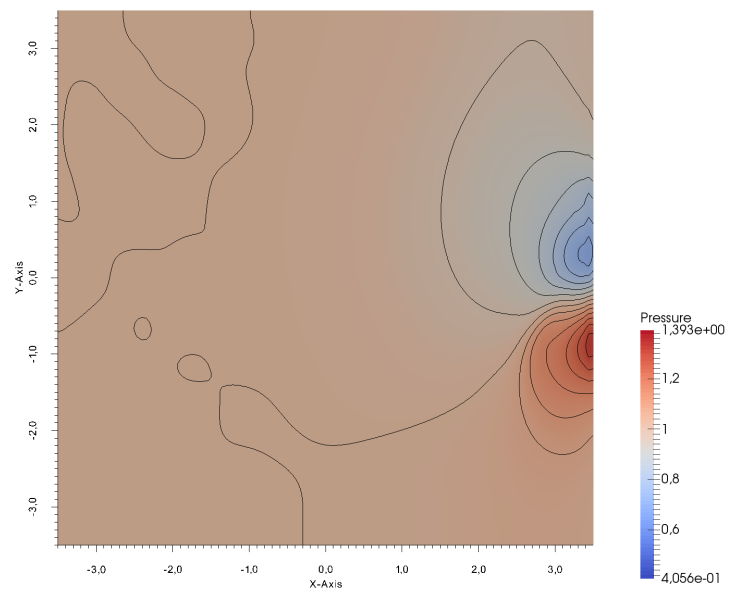

FIGURE 5.21: Contour plot of the pressure for the Shu-vortex on a square domain with weak characteristic BCs after 7500 time steps.
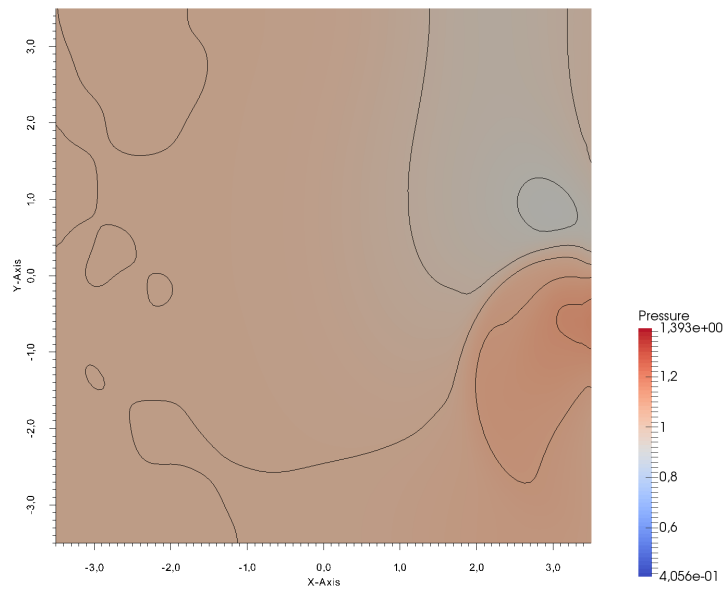
FIGURE 5.22: Contour plot of the pressure for the Shu-vortex on a square domain with weak characteristic BCs after 8500 time steps.
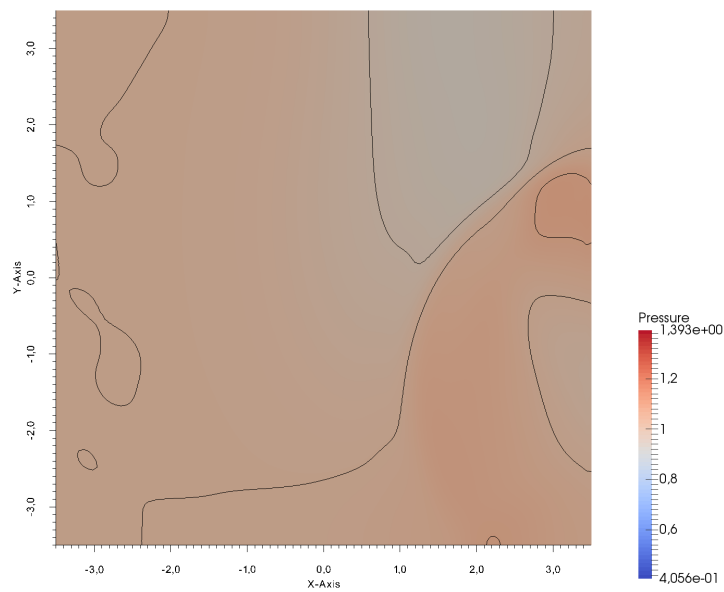

FIGURE 5.23: Contour plot of the pressure for the Shu-vortex on a square domain with weak characteristic BCs after 9500 time steps.
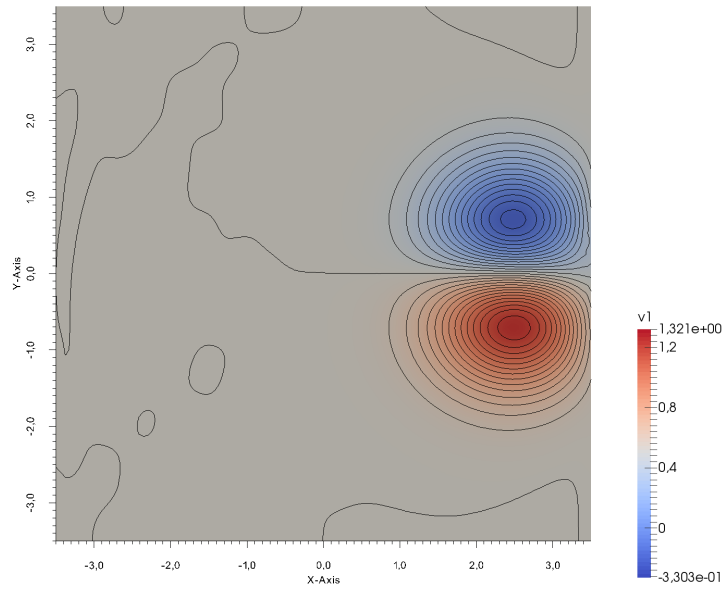
FIGURE 5.24: Contour plot of the x-directed velocity for the Shu-vortex on a square domain with weak characteristic BCs after 5000 time steps.
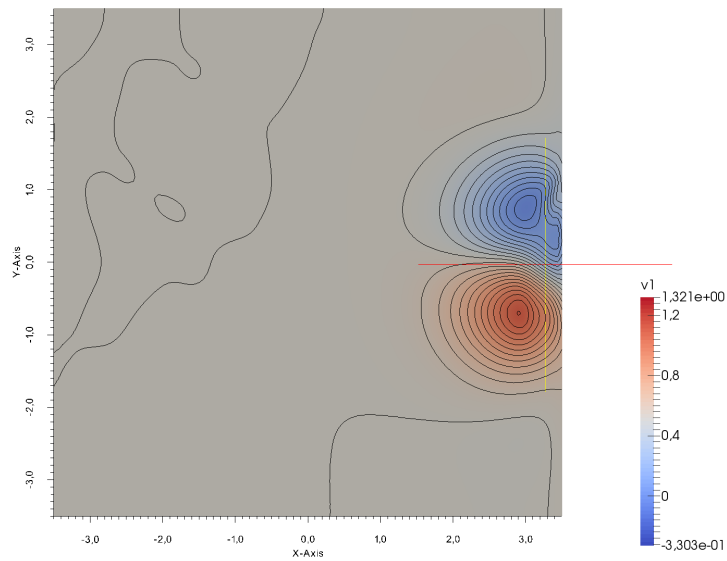

FIGURE 5.25: Contour plot of the x-directed velocity for the Shu-vortex on a square domain with weak characteristic BCs after 65000 time steps. A red line has been placed along the x-axis to emphasise the non-symmetry.

greater in magnitude than for the NACA0012 wing-profile. Once again it is hard to say if the numerical solution to the nonlinear Euler equations obeys $\frac{d}{dt}\|\bar{\mathbf{u}}\|_{\mathbf{I}\otimes\mathbf{P}} \leq 0$. But this expectation could be unmotivated. The local numerical solution at the boundary after 6000 time steps presumably is of greater magnitude than the reference data used for ingoing characteristics. If that is the case, then the BCs (4.7) might no be satisfied. However, no such analysis has been made for the Euler equations (2.17).
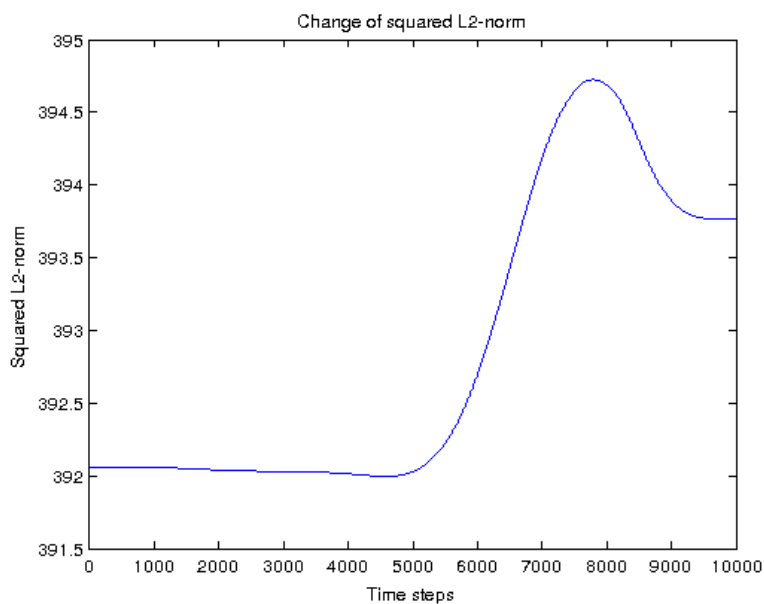

FIGURE 5.26: The change of $\|\bar{\mathbf{u}}\|^2_{\mathbf{I}\otimes\mathbf{P}}$ for the Shu-vortex with weak characteristic BCs, plotted against the number of time steps taken.

Chapter 6

Conclusions & future work

6.1 Conclusions

It has been shown that strict stability, in sense of definition 6, gives a residual that converges to 10^{-11} for the steady Euler equations with weak characteristic BCs at Ma 0.5. As in [1], the weak characteristic BCs excels compared to some other BCs, in this case the weak prescribed fluxes which converged much slower. The residual for the weak characteristic BCs may not have reached 10^{-14} as in [1]. Whether this is due to the condition number of the algorithm or not is unknown at this point. Regardless, the performance is so much better than the weak prescribed fluxes' and the obtained precision is for many real world applications redundant.

To obtain the same result for Ma 0.85 the order of the method had to be increased with linear reconstruction, possibly due to the presence of shocks. This result is remarkable since the analysis in chapter 4 did not consider reconstruction. Further, the weak prescribed fluxes preformed worse than at Ma 0.5 and Ma 0.85 without linear reconstruction.

The results given by simulations with weak characteristic BCs were validated on the grid with radius 20. There the solution was identical to the one given by simulations with prescribed fluxes. On the smaller domain of radius 1, the weak characteristic BCs are preferable based on the observation that the placement of the shocks is better.

The ability to absorb the incoming vortex is lacking for both the weak BCs in the Shu-vortex simulations. The weak prescribed fluxes did not create numerical artifacts of the same great magnitude as the weak characteristic BCs, but still leaves much to be desired as the vortex is deformed in the boundary's tangential direction. For both the

weak BCs, the symmetry of the vortex is disturbed as it reaches the outflow, this could be an indication of the need for more delicate methods for handling the BCs.

Concerning the increase of the $L^2(\Omega)$-norm over time, further investigation has to be made as the increment is very small and round off errors should not yet be ruled out.

6.2   Future work

An important next step would be to mathematically prove the connection between the stability of the linearised problem and the convergence of the residual of the nonlinear problem. In [15] and [16] J. Nordström and M. Svärd briefly mention that if a nonlinear initial-boundary value problem with some perturbation is linearised around a solution **u** and is well-posed for all **u**, then the original problem is well-posed. Thus investigation of this statement would be a natural next step. Further, based on the numerical experiments it might be possible to obtain a theorem similar to 1 for the node-centred FVM with weak characteristic BCs and linear reconstruction.

It would be interesting to implement a solver for problems of the form (4.42) and investigate how other choices of numerical flux function affect the algorithm. Then the numerical solution given by such a solver for the Euler equations can be compared with the solution the Tau 2D code provides for the nonlinear Euler equations.

The weak characteristic BCs and how they perform with unsteady flow problems need to be investigated further. Simulations for other test cases should be made with different strength of the variation over time. Finally, the current implementation can be modified so that it adapts to the size of v_n, giving the correct number of ingoing and outgoing characteristics depending on the Mach number, see table 3.1.

Appendix A

Eigenvectors & eigenvalues for the Euler equations

Consider the matrix

$$\mathbf{C} = \alpha\mathbf{A}_s + \beta\mathbf{B}_s = \begin{bmatrix} v_n & \frac{\alpha c}{\sqrt{\gamma}} & \frac{\beta c}{\sqrt{\gamma}} & 0 \\ \frac{\alpha c}{\sqrt{\gamma}} & v_n & 0 & \alpha c\sqrt{\frac{\gamma-1}{\gamma}} \\ \frac{\beta c}{\sqrt{\gamma}} & 0 & v_n & \beta c\sqrt{\frac{\gamma-1}{\gamma}} \\ 0 & \alpha c\sqrt{\frac{\gamma-1}{\gamma}} & \beta c\sqrt{\frac{\gamma-1}{\gamma}} & v_n \end{bmatrix}, \tag{A.1}$$

where $v_n = [v_1, v_2]^T \cdot \hat{\mathbf{n}} = \alpha v_1 + \beta v_2$, with $\hat{\mathbf{n}} = [\alpha, \beta]^T$ as the unit normal vector. The eigenvalues λ_i, $i = 1, 2, 3, 4$ are given by

$$\det(C - \lambda_i\mathbf{I}) = \begin{bmatrix} v_n - \lambda_i & \frac{\alpha c}{\sqrt{\gamma}} & \frac{\beta c}{\sqrt{\gamma}} & 0 \\ \frac{\alpha c}{\sqrt{\gamma}} & v_n - \lambda_i & 0 & \alpha c\sqrt{\frac{\gamma-1}{\gamma}} \\ \frac{\beta c}{\sqrt{\gamma}} & 0 & v_n - \lambda_i & \beta c\sqrt{\frac{\gamma-1}{\gamma}} \\ 0 & \alpha c\sqrt{\frac{\gamma-1}{\gamma}} & \beta c\sqrt{\frac{\gamma-1}{\gamma}} & v_n - \lambda_i \end{bmatrix}$$

$$= (v_n - \lambda_i)^4 + (\alpha\beta c^2)^2\frac{\gamma-1}{\gamma^2} - (v_n - \lambda_i)^2(\alpha^2 + \beta^2)c^2\frac{\gamma-1}{\gamma} - (\alpha\beta c^2)^2\frac{\gamma-1}{\gamma^2}$$

$$= (v_n - \lambda_i)^2((v_n - \lambda_i)^2 - c^2)\frac{\gamma-1}{\gamma}. \tag{A.2}$$

Setting $\det(C - \lambda_i \mathbf{I}) = 0$ gives the eigenvalues

$$
\begin{aligned}
\lambda_1 &= v_n - c, \\
\lambda_2 &= v_n, \\
\lambda_3 &= v_n, \\
\lambda_4 &= v_n + c.
\end{aligned}
\tag{A.3}
$$

Let $\mathbf{K}_i = (k_1, k_2, k_3, k_4)^\intercal \in \mathbb{R}^4$ be the eigenvector corresponding to the eigenvalue λ_i for the matrix $\mathbf{C}$ above. The eigenvectors are found by solving $\mathbf{CK} = \lambda \mathbf{K}$ for $\mathbf{K}$.

$\underline{\lambda_1 = v_n - c :}$

$$
\mathbf{CK}_1 = \lambda_1 \mathbf{K}_1 \Leftrightarrow
\begin{cases}
\sqrt{\gamma}\, k_1 + \alpha k_2 + \beta k_3 & = 0, & \text{(A.4a)} \\
\alpha k_1 + \gamma k_2 + \alpha \sqrt{\gamma - 1}\, k_4 & = 0, & \text{(A.4b)} \\
\beta k_1 + \gamma k_3 + \beta \sqrt{\gamma - 1}\, k_4 & = 0, & \text{(A.4c)} \\
\alpha \sqrt{\gamma - 1}\, k_2 + \beta \sqrt{\gamma - 1}\, k_3 + \sqrt{\gamma}\, k_4 & = 0. & \text{(A.4d)}
\end{cases}
$$

By dividing A.4d with $\sqrt{\gamma - 1}$ and substract from A.4a a relation for k_1 and k_4 is obtained. Then rewrite A.4a and A.4d to get an expression for k_1 respectively k_4, in relation to k_2 and k_3. Keep A.4c as it is, resulting in the following set of equations

$$
\begin{cases}
k_1 & = \frac{1}{\sqrt{\gamma - 1}}, & \text{(A.5a)} \\[4pt]
k_1 & = -\frac{1}{\sqrt{\gamma}}(\alpha k_2 + \beta k_3), & \text{(A.5b)} \\[4pt]
k_4 & = -\sqrt{\frac{\gamma - 1}{\gamma}}(\alpha k_2 + \beta k_3), & \text{(A.5c)} \\[4pt]
0 & = \alpha k_1 + \sqrt{\gamma}\, k_2 + \alpha \sqrt{\gamma - 1}\, k_4. & \text{(A.5d)}
\end{cases}
$$

Substituting A.5b and A.5c into A.5d gives the relation $\beta^2 k_2 = \alpha \beta k_3$. Set $k_1 = \xi \in \mathbb{R}$, giving

$$
\mathbf{K}_1 = \xi[1, -\alpha\sqrt{\gamma}, -\beta\sqrt{\gamma}, \sqrt{\gamma - 1}\,]^T.
\tag{A.6}
$$

Let $\hat{\mathbf{K}}_1$ be the normalisation of $\mathbf{K}_1$, finally giving

$$
\hat{\mathbf{K}}_1 = \frac{\mathbf{K}_1}{|\mathbf{K}_1|} = \left[\frac{1}{\sqrt{2\gamma}}, -\frac{\alpha}{\sqrt{2}}, -\frac{\beta}{\sqrt{2}}, \sqrt{\frac{\gamma - 1}{2\gamma}} \right]^T,
\tag{A.7}
$$

for $\xi = \frac{1}{\sqrt{2\gamma}}$.

$\underline{\lambda_{2,3} = v_n}$:

$$\mathbf{CK}_{2,3} = \lambda_{2,3}\mathbf{K}_{2,3} \Leftrightarrow \begin{cases} \dfrac{\alpha c}{\sqrt{\gamma}}k_2 + \dfrac{\beta c}{\sqrt{\gamma}}k_3 & = 0 \qquad \text{(A.8a)} \\[2mm] \dfrac{\alpha c}{\sqrt{\gamma}}k_1 + \dfrac{\alpha c\sqrt{\gamma-1}}{\sqrt{\gamma}}k_4 & = 0 \qquad \text{(A.8b)} \\[2mm] \dfrac{\beta c}{\sqrt{\gamma}}k_1 + \dfrac{\beta c\sqrt{\gamma-1}}{\sqrt{\gamma}}k_4 & = 0 \qquad \text{(A.8c)} \\[2mm] \dfrac{\alpha c\sqrt{\gamma-1}}{\sqrt{\gamma}}k_2 + \dfrac{\beta c\sqrt{\gamma-1}}{\sqrt{\gamma}}k_3 & = 0 \qquad \text{(A.8d)} \end{cases}$$

There are two degrees of freedom available, so setting $k_1 = \xi$ and $k_2 = \zeta$ gives

$$\begin{cases} k_1 & = \xi & \text{(A.9a)} \\[2mm] k_2 & = \zeta & \text{(A.9b)} \\[2mm] k_3\beta & = -\alpha\zeta & \text{(A.9c)} \\[2mm] k_4 & = -\dfrac{\xi}{\sqrt{\gamma-1}} & \text{(A.9d)} \end{cases}$$

If $\beta = 0$, then $\zeta = 0$, therefore A.9c can be divided by β and two eigenvectors are obtained:

$$\mathbf{K}_2 = \left[\xi_2, \zeta_2, -\frac{\alpha}{\beta}\zeta_2, -\frac{\xi_2}{\sqrt{\gamma-1}}\right]^T, \qquad \text{(A.10)}$$

$$\mathbf{K}_3 = \left[\xi_3, \zeta_3, -\frac{\alpha}{\beta}\zeta_3, -\frac{\xi_3}{\sqrt{\gamma-1}}\right]^T. \qquad \text{(A.11)}$$

To get the normalised vectors $\hat{\mathbf{K}}_2$ and $\hat{\mathbf{K}}_3$ the parameter $\xi_{2,3}$ must satisfy $\xi_{2,3} = \pm\sqrt{\frac{\gamma-1}{\gamma}\left(1 - \frac{\zeta_{2,3}^2}{\beta^2}\right)}$, since

$$|K_{2,3}| = \sqrt{\xi_{2,3}\frac{\gamma}{\gamma-1} + \frac{\zeta_{2,3}^2}{\beta^2}}. \qquad \text{(A.12)}$$

To obtain the same set of eigenvectors as [1], choose $\zeta_2 = 0$ which gives

$$\hat{\mathbf{K}}_2 = \left[-\sqrt{\frac{\gamma-1}{\gamma}}, 0, 0, \frac{1}{\sqrt{\gamma}}\right]^T \qquad \text{(A.13)}$$

and $\hat{\mathbf{K}}_2 \cdot \hat{\mathbf{K}}_3 = 0$ implies $\xi_3 = 0$ and therefore $\zeta_3 = \pm\beta$ resulting in

$$\hat{\mathbf{K}}_3 = [0, -\beta, \alpha, 0]^T. \qquad \text{(A.14)}$$

$\underline{\lambda_4 = v_n + c}$

$$\mathbf{C}\mathbf{K}_4 = \lambda_4 \mathbf{K}_4 \;\Leftrightarrow\; \begin{cases} -\sqrt{\gamma}\,k_1 + \alpha k_2 + \beta k_3 & = 0 \quad (\text{A.15a})\\[2mm] \alpha k_1 + -\gamma k_2 + \alpha\sqrt{\gamma-1}\,k_4 & = 0 \quad (\text{A.15b})\\[2mm] \beta k_1 + -\gamma k_3 + \beta\sqrt{\gamma-1}\,k_4 & = 0 \quad (\text{A.15c})\\[2mm] \alpha\sqrt{\gamma-1}\,k_2 + \beta\sqrt{\gamma-1}\,k_3 + -\sqrt{\gamma}\,k_4 & = 0 \quad (\text{A.15d}) \end{cases}$$

Following the same routine as for $\hat{\mathbf{K}}_1$ gives the normalised eigenvector

$$\hat{\mathbf{K}}_4 = \left[\frac{1}{\sqrt{2\gamma}}, \frac{\alpha}{\sqrt{2}}, \frac{\beta}{\sqrt{2}}, \sqrt{\frac{\gamma-1}{2\gamma}} \right]^T. \tag{A.16}$$

Let $\hat{\mathbf{K}}_1$, $\hat{\mathbf{K}}_2$, $\hat{\mathbf{K}}_3$ and $\hat{\mathbf{K}}_4$ be the columns of the matrix $\mathbf{K} \in \mathbb{R}^{4\times4}$ such that

$$\mathbf{K} = \begin{bmatrix} \frac{1}{\sqrt{2\gamma}} & -\sqrt{\frac{\gamma-1}{\gamma}} & 0 & \frac{1}{\sqrt{2\gamma}} \\[2mm] -\frac{\alpha}{\sqrt{2}} & 0 & -\beta & \frac{\alpha}{\sqrt{2}} \\[2mm] -\frac{\beta}{\sqrt{2}} & 0 & \alpha & \frac{\beta}{\sqrt{2}} \\[2mm] \sqrt{\frac{\gamma-1}{2\gamma}} & \frac{1}{\sqrt{\gamma}} & 0 & \sqrt{\frac{\gamma-1}{2\gamma}} \end{bmatrix} \tag{A.17}$$

and let $\mathbf{\Lambda} \in \mathbb{R}^{4\times4}$ be a diagonal matrix with the eigenvalues of $\alpha\mathbf{A}_s + \beta\mathbf{B}_s$ as elements:

$$\mathbf{\Lambda} = \begin{bmatrix} v_n - c & 0 & 0 & 0 \\ 0 & v_n & 0 & 0 \\ 0 & 0 & v_n & 0 \\ 0 & 0 & 0 & v_n + c \end{bmatrix}. \tag{A.18}$$

Appendix B

Definitions & theorems

Theorem 2 (Green's Theorem). Let Ω be an open set in $\mathbb{R}^2$ and $P, Q \in \mathcal{C}^1(\Omega)$. If the compact set $D \subset \Omega$ has a boundary ∂D, consisting of piecewise positively oriented $\mathcal{C}^1$-curves, then

$$\int_{\partial D} P \, dx + Q \, dy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dxdy. \tag{B.1}$$

[13]

Theorem 3 (Leibniz's rule of integration).

$$\frac{d}{dt} \int_{h(t)}^{g(t)} F(r, t) \, dr = \int_{h(t)}^{g(t)} \partial_t F(r, t) \, dr + F(r, t)|_{r=g(t)} \frac{d}{dt} g(t) - F(r, t)|_{r=h(t)} \frac{d}{dt} h(t) \tag{B.2}$$

Definition 7. (Positive semi-definite) A matrix $\mathbf{M} \in \mathbb{R}^{n \times n}$ is called positive semi-definite if

$$\mathbf{x}^T \mathbf{M} \mathbf{x} \geq 0 \tag{B.3}$$

for all $\mathbf{x} \in \mathbb{R}^n$.

Definition 8. (Kronecker Product) If $\mathbf{A}$ is a $m \times n$ matrix and $\mathbf{B}$ is a $p \times q$ matrix, then the *Kronecker product* $\mathbf{A} \otimes \mathbf{B}$ is a $mp \times nq$ block matrix of the form

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11} \mathbf{B} & \cdots & a_{1n} \mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1} \mathbf{B} & \cdots & a_{mn} \mathbf{B} \end{bmatrix}. \tag{B.4}$$

The Kronecker product has the following properties:

- If the matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ and $\mathbf{D}$ are of such size that the matrix products $\mathbf{AC}$ and $\mathbf{BD}$ are defined, then

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}, \tag{B.5}$$

- $(\mathbf{A} \otimes \mathbf{B})^T = \mathbf{A}^T \otimes \mathbf{B}^T$,

- $(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$,

- $\mathbf{A} \otimes \mathbf{B} + \mathbf{A} \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} + \mathbf{C})$.

Appendix C

Boundary terms for the linearised Euler equations

The characteristic BCs are given for the linearised two-dimensional Euler equations in primitive form, since it is more common than the symmetrised equations. The vectors corresponding to $\boldsymbol{\Lambda}^+$ and $\boldsymbol{\Lambda}^-$ are evaluated with the local numerical solution and reference data, respectively. The matrix $\mathbf{S}$ is as in (2.22) and for all the following cases it is assumed that $Ma < 1$. At an inflow boundary on has

$$\boldsymbol{\Lambda}^+ = \mathrm{diag}(0,\, 0,\, 0,\, v_n + c), \quad \boldsymbol{\Lambda}^- = \mathrm{diag}(v_n - c,\, v_n,\, v_n,\, 0), \tag{C.1}$$

which give

$$\mathbf{SK\Lambda}^+\mathbf{K}^T\mathbf{u}_s = \frac{(v_n + c)(\gamma v_n + c)}{2\gamma} \begin{bmatrix} \rho/c \\ \alpha \\ \beta \\ \rho c \end{bmatrix},$$

$$\mathbf{SK\Lambda}^-\mathbf{K}^T\mathbf{u}_s = \begin{bmatrix} \frac{\rho(-\gamma v_n^2 + c(v_n(3\gamma-1)-c))}{2\gamma c} \\ \frac{\gamma(\alpha v_1^2(\alpha^2+2\beta^2)+\beta^2 v_2(2\beta v_1 - \alpha v_2)) - c\alpha(v_n(\gamma+1)-c)}{2\gamma} \\ \frac{\gamma(\beta v_2^2(\beta^2+2\alpha^2)+\alpha^2 v_1(2\alpha v_2 - \beta v_1)) - c\beta(v_n(\gamma+1)-c)}{2\gamma} \\ \frac{-\rho c(v_n-c)(\gamma v_n - c)}{2\gamma} \end{bmatrix}. \tag{C.2}$$

The corresponding $\boldsymbol{\Lambda}^+$ and $\boldsymbol{\Lambda}^-$ at an outflow are

$$\boldsymbol{\Lambda}^+ = \mathrm{diag}(0,\, v_n,\, v_n,\, v_n + c), \quad \boldsymbol{\Lambda}^- = \mathrm{diag}(v_n - c,\, 0,\, 0,\, 0) \tag{C.3}$$

which give

$$
\mathbf{S}\mathbf{K}\boldsymbol{\Lambda}^{+}\mathbf{K}^{T}\mathbf{u}_s =
\begin{bmatrix}
\frac{\rho(\gamma v_n^2 + c(v_n(3\gamma-1)+c))}{2\gamma c} \\
\frac{\gamma(\alpha v_1^2(\alpha^2+2\beta^2)+\beta^2 v_2(2\beta v_1 - \alpha v_2))+c\alpha(v_n(\gamma+1)+c)}{2\gamma} \\
\frac{\gamma(\beta v_2^2(\beta^2+2\alpha^2)+\alpha^2 v_1(2\alpha v_2 - \beta v_1))+c\beta(v_n(\gamma+1)+c)}{2\gamma} \\
\frac{\rho c(v_n+c)(\gamma v_n+c)}{2\gamma}
\end{bmatrix},
$$

$$
\mathbf{S}\mathbf{K}\boldsymbol{\Lambda}^{-}\mathbf{K}^{T}\mathbf{u}_s = \frac{(v_n-c)(\gamma v_n - c)}{2\gamma}
\begin{bmatrix}
-\rho/c \\
\alpha \\
\beta \\
-\rho c
\end{bmatrix}.
$$

$$(C.4)$$

Bibliography

[1] C. Adamsson J. Nordström, K. Forsberg and P. Eliasson. Finite volume methods, unstructured meshes and strict stability for hyperbolic problems. *Applied Numerical Mathematics*, 45(12):453–473, June 2003.

[2] P. Birken. *Numerical Methods for the Unsteady Compressible Navier-Stokes Equations, Habilitation Thesis.* University of Kassel, 2012.

[3] E. Godlewski and P.-A. Raviart. *Numerical Approximation of Hyperbolic Systems of Conservation Laws.* Springer, New York, Berlin, Heidelberg, 1996.

[4] A. Holst and V. Ufnarovski. *Matrix Theory.* Lund University, Sweden, 2011.

[5] W. E. Scheisser. *The Numerical Method of Lines.* Academic Press, San Diego, 1991.

[6] G.-Q. Chen and D. Wang. *The Cauchy Problem for the Euler Equations for Compressible Fluids in Handbook of mathematical fluid dynamics, volume I.* Elsevier, 2002.

[7] H.-O. Kreiss and J. Lorenz. *Initial Boundary Value Problems and the Navier-Stokes Equations.* Academic Press, New York, 1989.

[8] S. Abarbanel and D. Gottlieb. Optimal time splitting for two- and three-dimensionsal navier-stokes equations with mixed derivativs. *Journal of Computational Physics*, 41(1):1–33, June 1981.

[9] R. J. LeVque. *Numerical Methods for Conservation Laws.* Birkhäuser Verlag, Basel, 1990.

[10] F. M. White. *Fluid Mechanics.* McGraw-Hill, seventh edition, 2011.

[11] T. Barth and M. Ohlberger. *Chapter 15. Finite Volume Methods: Foundation and Analysis in Encyclopedia of Computational Mechanics, volume I.* John Wiley & Sons, 2004.

[12] G.-H. Y J. Tu and C. Liu. *Computational Fluid Dynamics: A Practical Approach, 2nd edition.* Elsevier Science Limited, UK, 2012.

[13] A. Persson and L.-C. Böiers. *Analys i flera variabler.* Studentlitteratur AB, Lund, 3:6 edition, 2005.

[14] NASA Langley Research Center. Turbulence model numerical analysis 2d naca 0012 airfoil validation case, 2014. URL `http://turbmodels.larc.nasa.gov/naca0012numerics_val.html`.

[15] J. Nordström and M. Svärd. Well-posed boundary conditions for the navier-stokes equations. *Siam Journal on Numerical Analysis*, 43:1231–1255, 2005.

[16] M. Svärd and J. Nordström. Review of summation-by-parts schemes for initial-boundary-value problems. Report, 2013.