Discontinuous Finite-Element Approximations of Two-Dimensional Vorticity Equations with Linear Elliptic Inversions and Circulation

Master's Thesis

Erik Bernsen
Numerical Analysis and Computational Mechanics
Department of Applied Mathematics
University of Twente
The Netherlands

12th November

Contents

1 Preface

This report contains the work which I did for my master's thesis. I want to thank everyone who made it possible for me to finish this work. In particular I want to express my gratitude to Vijaya Ambati and my supervisor, Onno Bokhove, for their help during all the months I spend working on this project. Without all their help, advice, motivating words and enthusiasm I probably would not have succeeded in finishing this report at all.

2 Abstract

We are interested in solving an advection equation for the generalized vorticity coupled with a generalized elliptical equation for the streamfunction. These generalized equations include the incompressible 2D Euler equations, the quasi-geostrophic equations and the rigid lid equations as special cases.

Several kinds of boundary conditions are considered. Particularly interesting are the no normal flow boundary conditions. By imposing a condition on the circulation for each no normal flow boundary we can ensure that the energy and enstrophy of the system are conserved.

A mixed discontinuous/continuous Galerkin method is used to numerically solve the equations. The advection equation for the generalized vorticity is discretized using the discontinuous Galerkin method while the elliptical equation is discretized using the continuous Galerkin method.

The numerical method which we used has the remarkable property that the energy of the numerical solution is conserved. The enstrophy of the numerical solution is L_2-stable and in some cases even conserved.

We implemented the numerical method in C++ and ran some tests to validate the code and the method.

3 Introduction

3.1 Background

It was the Norwegian hydro-dynamicist V. Bjerknes who, in 1904, first realized that weather predictions could be based on the complete set of hydrodynamic and thermodynamic equations. However, at that moment, it was practically impossible to do such a prediction because of the enormous amount of calculations involved.

The first attempt to actually use these set of equations for weather predictions was made even before computers existed. In 1922 Richardson used a finite difference method to solve simplified versions of the hydrodynamic and thermodynamic equations. His results were not very promising. Not only would it take too much time to do practical weather predictions (it took him 6 weeks to make an six hour weather prediction for only a small part of the Earth), the predictions themselves were not very accurate at all.

Only after the computer had been invented, it became possible to do serious numerical weather predictions. In 1948 Jule Charney[1] derived a simplified mathematical model based on the quasi-geostrophic equations (see Section 7.2) which would be able to predict the large scale motions in the atmosphere. The first numerical experiments [2] of his model were conducted in 1950 and the results were very promising. Some of the predictions for periods of time greater than 24 hours were already even better than the older, non mathematical, methods.

When computers became faster and faster the mathematical models could become more and more complicated. Instead of using the simplified models of Charney, it became possible to use models based on the hydrostatic primitive equations. These primitive equations describe the relation between velocity field, pressure field, temperature and other relevant meteorological variables. However, in this report we will consider, among others, the quasi-geostrophic equations.

A very important role in the quasi-geostrophic equations (and in the hydrostatic primitive equations as well) is played by the Coriolis force. This is the force that moving fluid particles feel due to the rotation of the Earth. A consequence of this Coriolis force is the famous law of Buys Ballot which states that the direction of the wind on the Northern hemisphere around areas of high pressure is clockwise. These winds around areas of high and low pressure generate vorticity. Vorticity represents the amount of spinning around of the wind with respect to the Earth.

In the quasi-geostrophic approximation it is assumed that the velocity field is almost non-divergent. This approximation is justified by the fact that the atmosphere of the Earth is very thin compared to the horizontal distances. Then it is possible to represent the velocity field using a stream function. Lines on which the stream function has a constant value are the streamlines of a flow. Thus the vorticity and the stream function are important variables in the quasi-geostrophic equations.

Although the quasi-geostrophic equations are not used any more in today's weather predictions models they still play an important role in understanding the behavior of the atmosphere.

The quasi-geostrophic equations are not the only equations we are interested in. We are also interested in solving the rigid lid equations and the incompressible 2D Euler equations. We will see that the structure of these three sets of equations is very similar and can be captured in one generalized vorticity streamfunction formulation.

3.2 Goals

The goal of this report is to extend the work done by Liu and Shu in [9]. In [9] a numerical method to solve the incompressible 2D Euler equations in vorticity streamfunction formulation is discussed. This method has some very nice properties such as conservation of numerical energy and enstrophy. For the boundary conditions only the case of a domain with only one simply connected Dirichlet boundary is considered. We want to extend the method described in [9] in the following ways:

- We want to consider a more general equation than the incompressible 2D Euler equations. The rigid lid equations and quasi-geostrophic equations are a special case of our more general equations.
- We want to consider more general boundary conditions. Of particularly interest is the case for which the boundary of the domain is not simply connected.
- All the above things should be done in such a way that the properties of energy conservation and enstrophy stability of the scheme in [9] still hold.

4 Generalized Vorticity Stream Function Formulation

We are interested in solving the equation for the generalized vorticity, ω, with a generalized elliptical equation for the stream function ψ in a two-dimensional domain. The elliptical equation is assumed to be linear in ψ.

This coupled system in a domain $\Omega \subset \mathbb{R}^2$ is defined as follows

$$\begin{aligned} \partial_t \omega + \nabla \cdot (\omega \vec{u}) &= 0 && (1) \\ \vec{u} &= A \nabla^{\perp} \psi && (2) \\ \nabla \cdot (A \nabla \psi) - B\psi + C &= \omega && (3) \end{aligned}$$

with $A(x, y) > 0$ a continuous function, $B(x, y) \geq 0$, $C(x, y) \in \mathbb{R}$, $\nabla = [\partial_x, \partial_y]^T$ and the two dimensional curl operator $\nabla^{\perp} = [-\partial_y, \partial_x]^T$. These equations describe the motion of a fluid in a 2D domain. The velocity of the fluid is represented by $\vec{u} = [u, v]^T$ and the generalized vorticity is given by ω. We use the term generalized vorticity to emphasize that ω is defined by (3) which is different from the usual definition of vorticity $\partial_x v - \partial_y u = \nabla \cdot (A \nabla \psi)$. This generalized vorticity-stream function formulation has the advantage to include at least three (geophysical) systems of interest:

- the 2D time-dependent incompressible Euler equations in vorticity stream-function formulation,
- the quasi-geostrophic equations,
- the rigid lid equations.

A more detailed explanation of these equations is found in Section 7.

5 Boundary Conditions

In addition to the equations (1)–(3) it is necessary to specify boundary conditions. We will consider three kinds of boundary conditions: Neumann boundary conditions, periodic boundary conditions and no normal flow boundary conditions. A mixture of these boundary conditions is allowed as well. In the following the boundary of Ω will be denoted by $\partial\Omega$.

5.1 Neumann Boundary Conditions

In the case of Neumann boundary conditions the tangential component of the velocity vanishes on the boundary, thus $\vec{u} \cdot \hat{\tau} = \nabla\psi \cdot \hat{n} = 0$ on $\partial\Omega_N$. Here $\partial\Omega_N \subset \partial\Omega$ is the part of the boundary with Neumann boundary conditions, $\hat{n}$ denotes the normal vector of $\partial\Omega_N$ directed outward and $\hat{\tau}$ is the unit vector tangential to $\partial\Omega_N$ and oriented counterclockwise given by

$$\hat{\tau} = [-n_y, n_x]\,. \tag{4}$$

In some cases it is necessary to prescribe a value for ω on the boundary. For instance on the parts of the boundary with inflow, or $\vec{u} \cdot \hat{n} < 0$, a value of ω should be prescribed. In the case of outflow, or $\vec{u} \cdot \hat{n} > 0$, it is sometimes also necessary to prescribe a value for ω depending on the numerical method. In Section 8.1 we will see that a numerical flux is introduced. If we choose the central flux then we always need to prescribe a value for ω on the boundary. In the case of an upwind flux this is only necessary for the parts of the boundary with inflow.

5.2 Periodic Boundary Conditions

For these boundary conditions the solution to (1)–(3) is periodic in x. Thus

$$\omega(x + L, y) = \omega(x, y) \tag{5}$$
$$\vec{u}(x + L, y) = \vec{u}(x, y) \tag{6}$$
$$\psi(x + L, y) = \psi(x, y) \tag{7}$$

for some constant $L \in \mathbb{R}$. In this case it is also necessary that the functions A, B and C in (3) are periodic in space, hence

$$A(x + L, y) = A(x, y) \tag{8}$$
$$B(x + L, y) = B(x, y) \tag{9}$$
$$C(x + L, y) = C(x, y). \tag{10}$$

In this case it is sufficient to solve the equation on the domain $\Omega \cap [0, L] \times \mathbb{R}$. The boundary conditions for $x = 0$ and $x = L$ are given by (5)–(7), which are called periodic boundary conditions.

The same can be done if the solution is periodic in y, or doubly periodic, that is periodic in both x and y. The portion of the boundary which have periodic boundary conditions will be denoted by $\partial\Omega_P \subset \partial\Omega$.

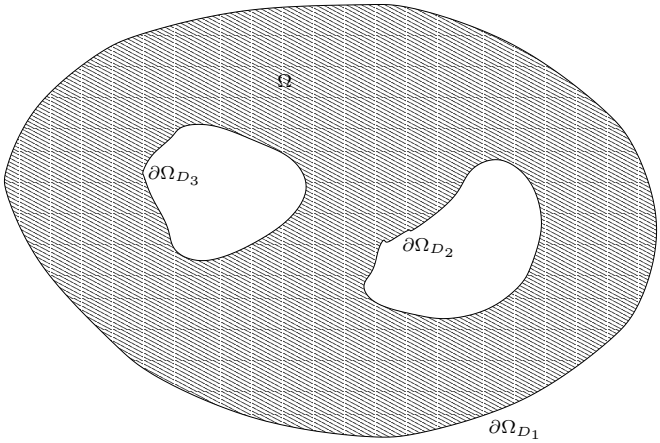


Figure 1: An example of a domain, Ω, with no normal flow boundary conditions. The boundary is partitioned into three separate parts: $\partial\Omega_{D_1}$, $\partial\Omega_{D_2}$ and $\partial\Omega_{D_3}$. Each of these three parts is a simply connected set and not connected to one of the other no normal flow boundaries.

5.3 No Normal Flow Boundary Conditions

When the domain boundary is rigid then there is no flow across the boundary. Hence,

$$\vec{u} \cdot \hat{n} = 0. \tag{11}$$

This is called a no normal flow boundary condition. We denote the part of the boundary with no normal flow boundary conditions with $\partial\Omega_D$. On a simply connected subset of $\partial\Omega_D$ the value of ψ is independent of x and y. This can be seen by using (2) and (4) to obtain from (11)

$$\frac{\partial \psi}{\partial \tau} = \nabla\psi \cdot \hat{\tau} \overset{(4)}{=} \nabla^{\perp}\psi \cdot \hat{n} \overset{(2)}{=} \frac{1}{A}\vec{u} \cdot \hat{n} \overset{(11)}{=} 0 \tag{12}$$

We assume that $\partial\Omega_D$ can be partitioned into N separate simply connected subsets $\partial\Omega_{D_1}, \partial\Omega_{D_2}, \ldots, \partial\Omega_{D_N}$ with $\partial\Omega_{D_i} \cup \partial\Omega_{D_j}$ not a simply connected set for $i \neq j$ (see Figure 1). Now $\psi|_{\partial\Omega_{D_i}}$ is independent of x and y, but may depend on t, thus we have

$$\psi|_{\partial\Omega_{D_i}} = \mathcal{C}_i(t) \qquad (i = 1, 2, \ldots, N) \tag{13}$$

for some function $\mathcal{C}_i(t)$. We will choose $\mathcal{C}_i(t)$ in such a way that

$$\frac{d}{dt}\int_{\partial\Omega_{D_i}} \vec{u} \cdot \hat{\tau}\, d\Gamma = 0 \quad (i = 1, 2, \ldots, N) \tag{14}$$

with $d\Gamma$ a line element along the boundary $\partial\Omega_i$. This choice of $\mathcal{C}_i(t)$ ensures that energy and enstrophy are conserved quantities (see Section 6).

In the case that we combine the no-normal-flow boundary conditions with boundary conditions which allow in- or outflow (such as Neumann boundary conditions) then (14) does not ensure conservation of energy. In this case we should use a different condition for determining $\mathcal{C}_i$. We do not consider this case in this report.

Another way of writing (14) is

$$\int_{\partial\Omega_{D_i}} \vec{u} \cdot \hat{\tau}\, d\Gamma = \mathcal{K}_i \quad (i = 1, 2, \ldots, N) \tag{15}$$

with $\mathcal{K}_i \in \mathbb{R}$ a given constant value. This constant $\mathcal{K}_i$ is called the circulation.

If $B = 0$ in (3) then ψ is only defined up to a constant. In this case we can use the Dirichlet boundary condition $\psi = 0$ on $\partial\Omega_{D_1}$ and thus take $\mathcal{C}_1 = 0$ instead of using (14) to determine $\mathcal{C}_1(t)$.

6 Energy and Enstrophy Conservation

In this section we prove that in the general formulation, for certain boundary conditions, energy and enstrophy are conserved quantities. We assume that we have a combination of no flow and periodic boundary conditions. If there are Neumann boundary conditions then, in general, energy and enstrophy will not be conserved due to the in- or outflow at the relevant portions of the boundary.

6.1 Energy Conservation

The total energy $E(t)$ of the system (1)–(3) is defined by

$$E(t) = \frac{1}{2}\int_\Omega A\left|\nabla\psi\right|^2 + B\left|\psi\right|^2 d\Omega. \tag{16}$$

Note that we required $A(x,y) > 0$ and $B(x,y) \geq 0$ (Section 4) to ensure that $E(t) \geq 0$. In the next theorem we see that this energy is a conserved quantity if the proper boundary conditions are used.

Theorem 6.1 (Energy Conservation). *The total energy, $E(t)$, as given in (16), of the system (1)–(3) subject to periodic boundary conditions (Section 5.2) or no normal flow boundary conditions (Section 5.3) is a conserved quantity:*

$$\frac{d}{dt}E(t) = 0. \tag{17}$$

Proof. If we multiply (1) with ψ, integrate over the domain, Ω, and integrate by parts the second term we obtain

$$\int_\Omega \psi\partial_t\omega\, d\Omega + \int_{\partial\Omega} \omega\psi\vec{u}\cdot\hat{n}\, d\Gamma - \int_\Omega \omega\vec{u}\cdot\nabla\psi\, d\Omega = 0. \tag{18}$$

The second term of (18) vanishes because of the periodic and no flow boundary conditions. The third term vanishes because $\vec{u} \perp \nabla\psi$. Thus we have

$$\int_\Omega \psi\partial_t\omega\, d\Omega = 0.$$

Using (3) and integrating by parts results in

$$\begin{aligned} 0 &= \int_\Omega \psi\partial_t\left(\nabla\cdot A\nabla\psi - B\psi + C\right) d\Omega \\ &= \int_{\partial\Omega} A\psi\partial_t\left(\nabla\psi\cdot\hat{n}\right) d\Gamma - \frac{d}{dt}E(t). \end{aligned} \tag{19}$$

The first term on the right hand side of this equation vanishes because of the periodic and no flow boundary conditions. In the case of no normal flow boundary

conditions this can be seen as follows

$$\begin{aligned}\int_{\partial\Omega_D} A\psi\partial_t\left(\nabla\psi\cdot\hat{n}\right)\,d\Gamma &= \sum_{i=1}^{N}\int_{\partial\Omega_{D_i}} A\psi\partial_t\left(\nabla\psi\cdot\hat{n}\right)\,d\Gamma \\ &= \sum_{i=1}^{N}\mathcal{C}_i(t)\frac{d}{dt}\int_{\partial\Omega_{D_i}}\vec{u}\cdot\hat{\tau}\,d\Gamma = 0 \end{aligned} \tag{20}$$

because of (2), (14) and $\psi|_{\partial\Omega_{D_i}} = \mathcal{C}_i(t)$ $(i = 1, \ldots, N)$. Now (19) and (20) yield (17). □

6.2 Enstrophy Conservation

Enstrophy is a measure for the averaged generalized vorticity. It is defined by

$$S(t) = \frac{1}{2}\int_\Omega A\omega^2 d\Omega. \tag{21}$$

The following theorem shows that the total enstrophy is also a conserved quantity.

Theorem 6.2 (Enstrophy Conservation). *The total enstrophy, S, as defined in (21), of the system (1)–(3) subject to periodic boundary conditions (Section 5.2) or no normal flow boundary conditions (Section 5.3) is a conserved quantity:*

$$\frac{d}{dt}S(t) = 0. \tag{22}$$

Proof. If we multiply (1) by $A\omega$ and integrate over the domain, Ω, we obtain

$$\int_\Omega A\omega\partial_t\omega + A\omega\nabla\cdot(\omega\vec{u})\,d\Omega = 0. \tag{23}$$

For the first term of (23) we can write

$$\int_\Omega A\omega\partial_t\omega\,d\Omega = \frac{d}{dt}S(t). \tag{24}$$

Using integration by parts (IP), the no normal flow and periodic boundary

conditions (BC) and (2) we can see that the second term of (23) vanishes:

$$\begin{aligned}
\int_\Omega A\omega\nabla\cdot(\omega\vec{u})\,d\Omega &\overset{(IP)}{=} \int_{\partial\Omega} A\omega^2\vec{u}\cdot\hat{n}\,d\Gamma - \int_\Omega \omega\vec{u}\cdot\nabla\left(A\omega\right)\,d\Omega \\
&\overset{(BC)}{=} -\int_\Omega \omega\vec{u}\cdot\nabla\left(A\omega\right)\,d\Omega \\
&= -\int_\Omega A\omega\vec{u}\cdot\nabla\omega\,d\Omega - \int_\Omega \omega^2\vec{u}\cdot\nabla A\,d\Omega \\
&= -\int_\Omega A\omega\vec{u}\cdot\nabla\omega\,d\Omega - \int_\Omega A\omega^2\vec{u}\cdot\frac{\nabla A}{A}\,d\Omega \\
&\overset{(2)}{=} -\int_\Omega A\omega\vec{u}\cdot\nabla\omega\,d\Omega - \int_\Omega A\omega^2\nabla\cdot\vec{u}\,d\Omega \\
&= -\int_\Omega A\omega\nabla\cdot(\omega\vec{u})\,d\Omega.
\end{aligned}$$

Hence, we find (22). □

7 Applications

In this section we consider the following three applications of the system (1)–(3):

- The 2D time-dependent incompressible Euler equations in vorticity stream-function formulation [7]. In this case we have

  $$A = 1, \qquad B = 0 \quad \text{and} \quad C = 0$$

  in the system (1)–(3). See Section 7.1 for further details.

- The quasi-geostrophic equations [11]. In this case we take

  $$A = 1, \qquad B = F \quad \text{and} \quad C = \eta_B$$

  in the system (1)–(3). Here $F \geq 0$ is a constant and $\eta_B = \eta_B(x, y)$ is a function. In Section 7.2 we will give more details on these equations.

- The rigid lid equations [8]. In this case we take

  $$A = \frac{1}{H}, \qquad B = 0 \quad \text{and} \quad C = f$$

  in the system (1)–(3). Here $f \in \mathbb{R}$ is a constant and $H = H(x, y)$ is a function of $(x, y) \in \Omega$. In Section 7.3 we see how these equations are related to the shallow water equations.

In the following sections we will work out the details for these three applications.

7.1 Incompressible Euler Equations

In this section we consider the Euler equations in vorticity stream function formulation [7]. We derive these equations from the standard incompressible Euler equations and we show that they are a special case of (1)–(3).

Consider the incompressible 2D Euler equations on $\Omega \subset \mathbb{R}^2$:

$$\partial_t \vec{u} + (\vec{u} \cdot \nabla)\, \vec{u} = -\nabla p/\rho \tag{25}$$

$$\nabla \cdot \vec{u} = 0. \tag{26}$$

In these equations $\vec{u} = [u, v]^T$ denotes the velocity field, p is the pressure field and ρ is the density. We have assumed that the fluid is incompressible and ρ is a constant.

We will define the vorticity as follows

$$\omega = \nabla^\perp \cdot \vec{u} = \partial_x v - \partial_y u. \tag{27}$$

Note that this equals the z-component of the curl of $[u, v, 0]^T$

$$[\partial_x, \partial_y, \partial_z]^T \times [u, v, 0]^T = [0, 0, \partial_x v - \partial_y u]^T. \tag{28}$$

We introduce the stream function, ψ, defined as

$$\vec{u} = \nabla^{\perp}\psi. \tag{29}$$

From (26) it follows that given the velocity field $\vec{u}$, such a function ψ exists and is defined up to a constant. By combining (27) and (29) we can express the vorticity in terms of the stream function

$$\omega = \nabla^{\perp} \cdot \vec{u} = \nabla^{\perp} \cdot \nabla^{\perp}\psi = \nabla^2\psi \equiv \Delta\psi. \tag{30}$$

To obtain the Euler equation in vorticity stream function formulation we take the 2D curl of (25)

$$\nabla^{\perp} \cdot (\partial_t u + (\vec{u} \cdot \nabla)\,\vec{u}) = \nabla^{\perp} \cdot \frac{\nabla p}{\rho}.$$

Using the identities $\nabla^{\perp} \cdot ((\vec{u} \cdot \nabla)\vec{u}) = (\nabla \cdot \vec{u} + \vec{u} \cdot \nabla)\left(\nabla^{\perp} \cdot \vec{u}\right)$ and $\nabla^{\perp} \cdot \nabla p = 0$ we can simplify the above equation to obtain

$$\partial_t\omega + \nabla \cdot (\omega\vec{u}) = 0. \tag{31}$$

The Euler equations in vorticity stream function formulation are now given by (29), (30) and (31). We see that the advantage of the vorticity stream function formulation is that the number of unknowns is reduced by one: instead of solving for three variables (u, v, p) we need to solve for only two (ω, ψ). Note that the Euler equations in vorticity stream function formulation are a special case of (1)–(3) if we take $A \equiv 1$, $B \equiv 0$ and $C \equiv 0$.

7.2 Barotropic Quasi-Geostrophic Equations

The barotropic quasi-geostrophic equations can be used to describe the motions in the atmosphere and oceans of the Earth in mid-latitudes. First we consider the shallow water equations in Section 7.2.1. In Section 7.2.2 we derive the quasi-geostrophic equations from the shallow water equations.

7.2.1 Shallow Water Equations

The shallow water equations describe the motion of a fluid in a thin layer of fluid on a rotating (x, y)-plane. The fluid moves between a rigid bottom at $z = h_B(x, y)$ and a free surface at $z = h(x, y, t)$ (See Figure 2). The fluid is incompressible, and hence the density, ρ, is a constant. The (x, y)-plane is assumed to rotate around the vertical z-axis. Further we assume that the gravity is directed along the z-axis as well. The shallow water equations are derived from the 3D Euler equations, which are given by

$$\begin{aligned} \partial_t\vec{u} + (\vec{u} \cdot \nabla)\,\vec{u} + w\partial_z\vec{u} + f\,[-v, u]^T &= -\nabla p/\rho && (32)\\ \partial_t w + \vec{u} \cdot \nabla w + w\partial_z w &= -\partial_z p/\rho + g && (33)\\ \nabla \cdot \vec{u} + \partial_z w &= 0 && (34) \end{aligned}$$

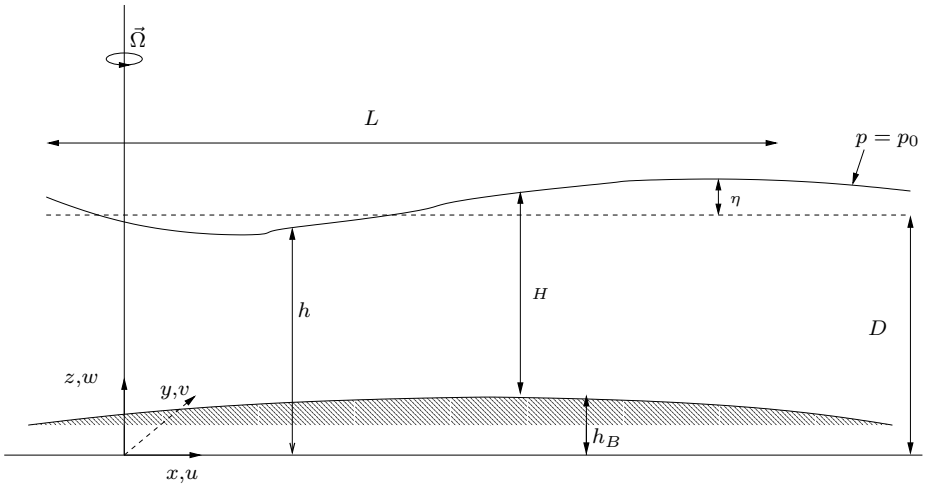


Figure 2: Variables used in the shallow water equations and quasi-geostrophic equations.

with p the pressure, g the gravitational acceleration, $\vec{u} = [u, v]^T$ the horizontal velocity components and w the vertical velocity component. The term $f\,[-v, u]^T$ is due to the Coriolis force which is caused by to the rotation of the plane. The parameter f is called the Coriolis parameter. If the axis of rotation is directed along the z-axis, then this parameter is given by $f = 2\Omega$ with Ω the rate of rotation of the plane. The pressure at the free surface, $z = h(x, y, t)$, is given by the constant p_0, e.g. the atmospheric pressure.

We assume that the horizontal velocity components are independent of z and that the momentum equation in the vertical direction reduces to the hydrostatic approximation

$$\partial_z p = \rho g. \tag{35}$$

Both assumptions are justified by the fact that the layer of fluid is thin [11]. That means, the vertical scales of motion are small compared to the horizontal scales of motion. Using the hydrostatic approximation (35) and the boundary condition, $p\big|_{z=h} = p_0$, for the pressure at the free surface we find

$$p = p_0 + \rho g(h - z). \tag{36}$$

Using these assumptions the 3D Euler equations (32)–(34) can be simplified to

$$\partial_t \vec{u} + (\vec{u} \cdot \nabla)\,\vec{u} + f\,[-v, u]^T = -g\nabla h \tag{37}$$

$$\partial_z p = \rho g \tag{38}$$

$$\nabla \cdot \vec{u} + \partial_z w = 0. \tag{39}$$

From (39) it follows that the vertical component of the velocity can be written as

$$w = -z\nabla \cdot \vec{u} + \tilde{w}(x, y, t).$$

To determine $\tilde{w}(x, y, t)$ we note that the upper and lower boundaries of the fluid are material surfaces. For a surface implicitly defined by $S(x, y, z, t) = 0$ this means that $DS/Dt = 0$ for all points on the surface. The material derivative DS/Dt is defined as $DS/Dt = \partial_t S + \vec{u} \cdot \nabla S + w\partial_z S$. Applying this to the upper boundary, with $S = z - h$, and lower boundary, with $S = z - h_B$, we obtain

$$0 = \frac{D(z-h)}{Dt}\Big|_{z=h} = -\partial_t h - \vec{u} \cdot \nabla h - h\nabla \cdot \vec{u} + \tilde{w} \tag{40}$$

$$0 = \frac{D(z-h_B)}{Dt}\Big|_{z=h_B} = -\vec{u} \cdot \nabla h_B - h_B \nabla \cdot \vec{u} + \tilde{w}, \tag{41}$$

respectively. Subtracting (40) from (41) we find

$$0 = \partial_t H + \vec{u} \cdot \nabla H + H\nabla \cdot \vec{u} \tag{42}$$

with $H = h - h_B$ the depth of the fluid.

The shallow water equations are now given by the momentum balance (37) and mass balance (42). This results in

$$\partial_t \vec{u} + (\vec{u} \cdot \nabla)\, \vec{u} + f\,[-v, u]^T = -g\,(\nabla H + \nabla h_B) \tag{43}$$

$$\partial_t H + \nabla \cdot (H\vec{u}) = 0. \tag{44}$$

Now we can introduce the potential vorticity. The potential vorticity is defined by

$$\xi = \frac{\nabla^\perp \cdot \vec{u} + f}{H}. \tag{45}$$

It has the nice property that it is a materially conserved quantity, that means

$$\frac{D\xi}{Dt} = 0 \tag{46}$$

with $D/Dt = \partial_t + \vec{u} \cdot \nabla$ the material derivative. This can be seen as follows:

$$\begin{aligned}\partial_t \xi + \vec{u} \cdot \nabla \xi &= \partial_t \left(\frac{\nabla^\perp \cdot \vec{u} + f}{H} \right) + \vec{u} \cdot \nabla \left(\frac{\nabla^\perp \cdot \vec{u} + f}{H} \right) \\ &= \frac{\nabla^\perp \cdot \partial_t \vec{u} + \vec{u} \cdot \nabla \left(\nabla^\perp \cdot \vec{u} \right)}{H} - \frac{\nabla^\perp \cdot \vec{u} + f}{H^2} \left(\partial_t H + \vec{u} \cdot \nabla H \right).\end{aligned}$$

Using the identity $\nabla^\perp \cdot ((\vec{u} \cdot \nabla)\, \vec{u}) \equiv (\nabla \cdot \vec{u} + \vec{u} \cdot \nabla) \left(\nabla^\perp \cdot \vec{u} \right)$, (43) and (44) we find

$$\begin{aligned}\partial_t \xi + \vec{u} \cdot \nabla \xi &= \frac{\nabla^\perp \cdot \left(\partial_t \vec{u} + (\vec{u} \cdot \nabla)\, \vec{u} + f[-v, u]^T \right)}{H} \\ &- \frac{\nabla^\perp \cdot \vec{u} + f}{H^2} \left(\partial_t H + \vec{u} \cdot \nabla H + H\nabla \cdot \vec{u} \right) \\ &= -\frac{g\nabla^\perp \cdot (\nabla H + \nabla h_B)}{H} - \frac{\nabla^\perp \cdot \vec{u} + f}{H^2} \left(\partial_t H + \nabla \cdot (\vec{u} H) \right) = 0.\end{aligned}$$

7.2.2 Deriving the quasi-geostrophic Equations

In [11] the quasi geostrophic equations are derived from the shallow water equations using a scaling analysis. We will follow this analysis here. First we define the deviation, η, of the free surface from its rest value

$$H = D + \eta - h_B \tag{47}$$

with D a characteristic value for the rest depth of the fluid. Then we introduce the dimensionless variables x', y', $\vec{u}'$, t' and η':

$$[x, y] = L[x', y'], \quad t = Tt', \quad \vec{u} = U\vec{u}', \quad \eta = N_0\eta', \quad \nabla = \nabla'/L.$$

We assume that L, T, U and N_0 characterize the horizontal length scales, time scales, velocity scales and scale of deviation from the mean surface, respectively. After substituting these dimensionless variables in (43) and (44), we obtain

$$\frac{U}{T}\partial_{t'}\vec{u}' + \frac{U^2}{L}\left(\vec{u}'\cdot\nabla'\right)\vec{u} + Uf\left[-v', u'\right]^T + g\frac{N_0}{L}\nabla'\eta' = 0 \tag{48}$$

$$\frac{N_0}{T}\partial_{t'}\eta' + \frac{U}{L}\left(D + N_0\eta' - h_B\right)\nabla\cdot\vec{u}' + \frac{U}{L}\vec{u}'\cdot\nabla'\left(N_0\eta' - h_B\right) = 0, \tag{49}$$

respectively. Now we introduce the scaling:

$$T = \frac{L}{U}, \quad N_0 = \frac{fUL}{g}, \quad \epsilon \equiv \frac{U}{fL} \ll 1, \quad \frac{h_B}{D} = \epsilon\eta'_B$$

with $\eta'_B = O(1)$. After multiplying (48) by $1/fU$ and (49) by $F/N_0 f$, with $F = f^2L^2/gD$, the rotational Froude number, we obtain

$$\epsilon\left(\partial_{t'}\vec{u}' + \left(\vec{u}'\cdot\nabla\right)\vec{u}'\right) + \left[-v', u'\right]^T + \nabla\eta' = 0 \tag{50}$$

$$\epsilon F\left(\partial_{t'}\eta' + \nabla'\cdot\left(\vec{u}'\eta'\right)\right) + \left(1 - \epsilon\eta'_B\right)\left(\nabla'\cdot\vec{u}'\right) - \epsilon\vec{u}'\cdot\nabla'\eta'_B = 0 \tag{51}$$

For the rest of this section we drop the primes for the dimensionless variables. We write finally the unknowns $\vec{u}$ and η as

$$\vec{u} = \vec{u}_0 + \epsilon\vec{u}_1 + \epsilon^2\vec{u}_2 + O(\epsilon^3) \tag{52}$$

$$\eta = \eta_0 + \epsilon\eta_1 + \epsilon^2\eta_2 + O(\epsilon^3). \tag{53}$$

Substituting (52) and (53) in (50) and (51) gives:

$$\vec{0} = \vec{K}_1 + \epsilon\vec{K}_2 + O(\epsilon^2) \tag{54}$$

and

$$0 = K_3 + \epsilon K_4 + O(\epsilon^2) \tag{55}$$

with

$$\vec{K}_1 \equiv \left[-v_0, u_0\right]^T + \nabla\eta_0 \tag{56}$$

$$\vec{K}_2 \equiv \partial_t\vec{u}_0 + \left(\vec{u}_0\cdot\nabla\right)\vec{u}_0 + \left[-v_1, u_1\right]^T + \nabla\eta_1 \tag{57}$$

$$K_3 \equiv \nabla\cdot\vec{u}_0 \tag{58}$$

$$K_4 \equiv F\left(\partial_t\eta_0 + \nabla\cdot\left(\eta_0\vec{u}_0\right)\right) + \nabla\cdot\vec{u}_1 - \eta_B\left(\nabla\cdot\vec{u}_0\right) - \vec{u}_0\cdot\nabla\eta_B. \tag{59}$$

Clearly we must take $\vec{K}_1 = \vec{K}_2 = \vec{0}$ and $K_3 = K_4 = 0$. From $\vec{K}_1 = \vec{0}$ and (56) we obtain the geostrophic balance condition

$$\vec{u}_0 = \nabla^\perp \eta_0. \tag{60}$$

Note that from this relation the condition $K_3 = 0$ is satisfied automatically. Using the equality in (60) we can simplify the condition $K_4 = 0$ to obtain

$$F\left(\partial_t \eta_0 + \vec{u}_0 \cdot \nabla \eta_0\right) + \nabla \cdot \vec{u}_1 - \vec{u}_0 \cdot \nabla \eta_B = 0. \tag{61}$$

We can eliminate η_1 from $\vec{K}_2 = \vec{0}$ by taking the 2D curl on both sides of this equation and using the identity $\nabla^\perp \cdot \left(\left(\vec{u}_0 \cdot \nabla\right)\vec{u}_0\right) \equiv \left(\nabla \cdot \vec{u}_0 + \vec{u}_0 \cdot \nabla\right)\left(\nabla^\perp \cdot \vec{u}_0\right)$

$$\begin{aligned} 0 &= \nabla^\perp \cdot \left(\partial_t \vec{u}_0\right) + \nabla^\perp \cdot \left(\left(\vec{u}_0 \cdot \nabla\right)\vec{u}_0\right) + \nabla^\perp \cdot \left[-v_1, u_1\right]^T + \nabla^\perp \cdot \nabla \eta_0 \\ &= \partial_t \Delta \eta_0 + \left(\vec{u}_0 \cdot \nabla\right)\Delta \eta_0 + \nabla \cdot \vec{u}_1. \end{aligned} \tag{62}$$

Now we can eliminate $\nabla \cdot \vec{u}_1$ from (61) and (62) and obtain

$$\partial_t \Delta \eta_0 + (\vec{u}_0 \cdot \nabla)\Delta \eta_0 = F\left(\partial_t \eta_0 + (\vec{u}_0 \cdot \nabla)\eta_0\right) - (\vec{u_0} \cdot \nabla)\eta_B \tag{63}$$

or

$$\partial_t(\Delta \eta_0 - F\eta_0 + \eta_B) + (\vec{u}_0 \cdot \nabla)(\Delta \eta_0 - F\eta_0 + \eta_B) = 0. \tag{64}$$

After renaming $\psi = \eta_0$ we arrive at the quasi-geostrophic equations

$$\partial_t \omega + \nabla \cdot (\omega \vec{u}) = 0 \tag{65}$$

$$\vec{u} = \nabla^\perp \psi \tag{66}$$

$$\omega = \Delta \psi - F\psi + \eta_B. \tag{67}$$

If we take $A = 1$, $B = F$ and $C = \eta_B$ in (1)–(3) we see the equivalence between (1)–(3) and (65)–(67).

7.3 Rigid Lid Equations

The rigid lid equations can be deduced from the shallow water equations. Just like the quasi-geostrophic equations and shallow water equations they describe the motion of thin layer of fluid on a rotating plane. To deduce the rigid lid equations we assume that the surface of the fluid is at a constant height. Hence $H = h - h_B(x, y) = H(x, y)$ is fixed in time. The justification for this assumption is that variations in h are usually much smaller than variations in h_B. The rigid lid equations are given in [8] as

$$\partial_t \xi + \vec{u} \cdot \nabla \xi = 0 \tag{68}$$

$$\vec{u} = \frac{1}{H}\nabla^\perp \psi \tag{69}$$

$$\xi = \frac{\nabla^\perp \cdot \vec{u} + f}{H}. \tag{70}$$

These equations follow from equation (44), (45) and (46) with $H = H(x, y)$.

Now we will show that, if we take $\omega = H\xi$, $A = 1/H > 0$, $B = 0$ and $C = f$, the equations (68)–(70) are equivalent to (1)–(3). It is obvious that with this choice of A (2) and (69) are equivalent. From (70) we can derive (3):

$$\begin{aligned}
\omega &= H\xi \\
&= \nabla^{\perp} \cdot \vec{u} + f \\
&= \partial_x \left(\frac{1}{H} \partial_x \psi \right) - \partial_y \left(-\frac{1}{H} \partial_y \psi \right) + f \\
&= \nabla \cdot \left(\frac{1}{H} \nabla \psi \right) + f.
\end{aligned}$$

Finally, we can derive (1) from (68) as follows

$$\begin{aligned}
0 &= \partial_t \xi + \vec{u} \cdot \nabla \xi \\
&= \partial_t \big(\frac{\nabla^{\perp} \cdot \vec{u} + f}{H} \big) + \vec{u} \cdot \nabla \big(\frac{\nabla^{\perp} \cdot \vec{u} + f}{H} \big) \\
&= \frac{1}{H} \partial_t \omega + \vec{u} \cdot \big(\frac{\nabla \omega}{H} - \frac{\omega}{H} \frac{\nabla H}{H} \big) \\
&= \frac{1}{H} \left(\partial_t \omega + \vec{u} \cdot \big(\nabla \omega - \frac{\nabla H}{H} \omega \big) \right) \\
&\overset{(69)}{=} \frac{1}{H} \left(\partial_t \omega + \vec{u} \cdot \nabla \omega + \omega \nabla \cdot \vec{u} \right) \\
&= \frac{1}{H} \big(\partial_t \omega + \nabla \cdot (\omega \vec{u}) \big),
\end{aligned}$$

where we assumed that $H \neq 0$.

8 Finite Element Method

We use a finite element method to solve the generalized vorticity equations (1)–(3). The vorticity equation (1) is solved using a discontinuous Galerkin method, while the elliptic equation for the stream function (3) is solved using a continuous Galerkin method.

8.1 The Discontinuous Galerkin Space Discretization

In this section we discretize (1) using the discontinuous Galerkin method. The numerical approximation of $\vec{u}$ is defined as

$$\vec{u}_h = A\nabla^{\perp}\psi_h \tag{71}$$

with ψ_h a numerical approximation for ψ. For this moment, we assume that by using (3) we can calculate ψ_h given an approximation ω_h of ω. In Section 8.2 we show how this can be done using a continuous Galerkin method.

In order to derive the discontinuous Galerkin discretization we need to define a mesh. Assume that a mesh of N_e elements is given by

$$\mathcal{T}_h = \{K_1, K_2, \ldots, K_{N_e}\} \tag{72}$$

on Ω. We use a set of discontinuous test functions defined by

$$V_h^k = \left\{v \middle| v|_{K_i} \in S^k(K_i) \quad (i = 1, \ldots, N_e)\right\} \tag{73}$$

with $S^k(K)$ a N_k-dimensional space of continuous functions $v_h : K \to \mathbb{R}$. These functions should be order k accurate. Note that a test function $v \in V_h^k$ *is* continuous on each element $K \in \mathcal{T}_h$ but in general discontinuous across element boundaries. In Appendix B.1 we make an explicit choice for the spaces $S^k(K_i)$, but for the description of the method this is not necessary.

IF we multiply (1) with a test function $v_h \in V_h^k$ and integrate over $K \in \mathcal{T}_h$, we get

$$\int_K v_h \partial_t \omega dK + \int_K v_h \nabla \cdot (\omega \vec{u})\, dK = 0. \tag{74}$$

Integrating by parts the second term of this equation results in

$$\int_K v_h \partial_t \omega dK + \int_{\partial K} \omega v \vec{u} \cdot \hat{n} d\Gamma - \int_K \omega \vec{u} \cdot \nabla v_h dK = 0 \tag{75}$$

with ∂K the boundary of the element K. Characteristic for the Galerkin method is that the approximation space for the solution ω of (1)–(3) and the space of test functions, V_h^k, are taken equal. Thus we replace ω with $\omega_h \in V_h^k$ to obtain

$$\int_K v_h \partial_t \omega_h dK + \int_{\partial K} \omega_h v_h \vec{u}_h \cdot \hat{n} d\Gamma - \int_K \omega_h \vec{u}_h \cdot \nabla v_h dK = 0. \tag{76}$$

Note that we also replaced ψ and $\vec{u}$ with approximations ψ_h and $\vec{u}_h$. These approximations follow from (2) and (3) with ω replaced by ω_h. In Section 8.2 we discuss how to compute these approximations numerically.

The second term of (76) causes a problem. In order to calculate this term we need to know the value of ω_h and v_h on the boundary, ∂K, of an element K. But the functions ω_h and v_h usually are not continuous across element boundaries although ω is. To solve this problem we take for the value of v_h on the boundary simply the value of v_h on the inside of the element. We use the $^-$ symbol to denote the value of a function on the inside of the element and $^+$ for the value on the outside of an element. The value $\omega_h \vec{u}_h \cdot \hat{n}$ on ∂K is replaced by a numerical flux $\hat{f}(\omega_h^+, \omega_h^-, \vec{u}_h \cdot \hat{n})$.

If ω_h *is* continuous, thus $\omega_h^+ = \omega_h^- = \omega_h$ then this numerical flux should of course equal the original value $\omega_h \vec{u}_h \cdot \hat{n}$, thus we obtain the following property of $\hat{f}$

$$\hat{f}(\omega_h, \omega_h, \vec{u}_h \cdot \hat{n}) = \omega_h \vec{u}_h \cdot \hat{n} \qquad \text{for each} \qquad \omega_h \in \mathbb{R}. \tag{77}$$

Another property of the numerical flux should be that for two neighboring elements $K_A \in \mathcal{T}_h$ and $K_B \in \mathcal{T}_h$ the flux from K_A to K_B should be opposite to the flux from K_B to K_A. This condition is expressed as

$$\hat{f}(\omega_h^+, \omega_h^-, \vec{u}_h \cdot \hat{n}) = -\hat{f}(\omega_h^-, \omega_h^+, -\vec{u}_h \cdot \hat{n}). \tag{78}$$

Note that we don't use $\vec{u}_h^- \cdot \hat{n}$ and $\vec{u}_h^+ \cdot \hat{n}$ in the numerical flux because $\vec{u}_h \cdot \hat{n}$ is continuous across element boundaries. To see this observe that we have

$$\vec{u}_h \cdot \hat{n} = A \nabla^\perp \psi_h \cdot \hat{n} = A \nabla \psi_h \cdot \hat{\tau}.$$

In Section 8.2 we enforce ψ_h to be a continuous function by using a continuous Galerkin approximation. Due to the continuity of ψ_h and A we have that $\nabla \psi_h \cdot \hat{\tau}$ is continuous across element boundaries.

Some possible choices of $\hat{f}$ given ω^+, ω^- and the normal velocity $u_n = \vec{u}_h \cdot \hat{n}$ are:

Central Flux

$$\hat{f}(\omega^+, \omega^-, u_n) = \frac{\omega^+ + \omega^-}{2} u_n. \tag{79}$$

Upwind Flux

$$\hat{f}(\omega^+, \omega^-, u_n) = \begin{cases} \omega^+ u_n & \text{if } u_n < 0 \\ \omega^- u_n & \text{if } u_n > 0 \end{cases}. \tag{80}$$

Lax-Friedrichs Flux

$$\hat{f}(\omega^+, \omega^-, u_n) = \frac{1}{2} \left(u_n \left(\omega^+ + \omega^- \right) - \alpha \left(\omega^+ - \omega^- \right) \right) \tag{81}$$

with α the maximum of $|\vec{u}_h \cdot \hat{n}|$. This can be the global maximum over the whole domain or the local maximum over one face only.

Note that with all these choices conditions (77) and (78) are satisfied.

After replacing v_h with v_h^- and $\omega_h \vec{u}_h \cdot \hat{n}$ with $\hat{f}(\omega_h^+, \omega_h^-, \vec{u}_h \cdot \hat{n})$ in (76) we find the weak formulation of (1). Find $\omega_h \in V_h^k$ such that for each $v_h \in V_h^k$

$$\int_K v_h \partial_t \omega_h dK + \int_{\partial K} v_h^- \hat{f}(\omega_h^+, \omega_h^-, \vec{u}_h \cdot \hat{n}) d\Gamma - \int_K \omega_h \vec{u}_h \cdot \nabla v_h dK = 0. \quad (82)$$

The weak formulation can be written as a system of equations. In order to do this we assume that we have a (local) basis for $S^k(K)$ given by

$$\left\{ \phi_1^{(K)}, \phi_2^{(K)}, \ldots, \phi_{N_k}^{(K)} \right\}. \quad (83)$$

These are the basisfunctions for the element K. Since functions in V_h^k are not required to be continuous across element boundaries we can define the basisfunctions elementwise. We assume that these basis functions satisfy the following properties

$$\int_K \phi_1^{(K)} \phi_i^{(K)} dK = 0 \quad \text{if} \quad i \neq 1, \quad (84)$$

$$\int_K \phi_1^{(K)} v_h dK = \frac{1}{|K|} \int_K v_h dK \qquad \text{for each} \qquad v_h \in S^k(K) \quad (85)$$

$$\nabla \phi_1 = 0. \quad (86)$$

In Appendix B.1 we will make an explicit choice for these basis functions. Using these basis functions we can replace ω_h in (82) with

$$\omega_h|_K = \sum_{i=1}^{N_k} \omega_i^{(K)} \phi_i^{(K)}.$$

Note that because of the properties (84) and (85) the mean of ω_h over an element K is given by $\omega_1^{(K)}$.

If we take for the testfunctions, v_h, in (82) the basisfunctions

$$v_h|_K = \phi_i^{(K)} \quad (i \geq 1)$$

for each element $K \in \mathcal{T}_h$ and each $i = 1, 2, \ldots, N_k$, then (82) holds for an arbitrarily testfunction $v_h \in V_h^k$.

The following spatial discretization of the vorticity equation is now obtained

$$\sum_{j=1}^{N_k} \frac{d\omega_j^{(K)}}{dt} \int_K \phi_i^{(K)} \phi_j^{(K)} dK + \int_{\partial K} \phi_i^{(K)} \hat{f}(\omega_h^+, \omega_h^-, \vec{u}_h \cdot \hat{n}) d\Gamma -$$

$$\sum_{j=1}^{N_k} \omega_j^{(K)} \int_K \phi_j^{(K)} \vec{u}_h \cdot \nabla \phi_i^{(K)} dK = 0 \quad (i \geq 1).$$

Using (84)–(86), we rewrite this last equation as

$$\begin{aligned}
&\frac{d\omega_1^{(K)}}{dt}\int_K \left(\phi_1^{(K)}\right)^2 dK + \int_{\partial K} \phi_1^{(K)} \hat{f}(\omega_h^+, \omega_h^-, \vec{u}_h \cdot \hat{n}) d\Gamma = 0 \\
&\sum_{j=2}^{N_k} \frac{d\omega_j^{(K)}}{dt}\int_K \phi_i^{(K)}\phi_j^{(K)} dK + \int_{\partial K} \phi_i^{(K)} \hat{f}(\omega_h^+, \omega_h^-, \vec{u}_h \cdot \hat{n}) d\Gamma - \\
&\qquad \sum_{j=1}^{N_k} \omega_j^{(K)} \int_K \phi_j^{(K)} \vec{u}_h \cdot \nabla \phi_i^{(K)} dK = 0 \quad (i \geq 2).
\end{aligned} \tag{87}$$

8.2 Elliptical Inversion: Stream Function

In this section we show how to approximate ψ by a continuous function $\psi_h \in V_h^k$ given an approximation ω_h of ω using (3) and the continuous Galerkin method[10]. By using a continuous Galerkin method instead of a discontinuous we ensure that the velocity is continuous across element boundaries. We used this property of the velocity in Section 8.1 in the numerical flux. First, we derive the weak formulation corresponding to (3).

8.2.1 Weak Formulation

We define the space of test functions, $H^1(\Omega)$, as

$$\begin{aligned}
&H^1(\Omega) = \{\, w \in L^2(\Omega) \big| \partial_x w, \partial_y w \in L^2(\Omega), \\
&\frac{dw}{d\tau}\big|_{\partial\Omega i} = 0 \ (i = 1, \ldots, N) \text{ for no normal flow boundary conditions}, \\
&w(L, y) = w(0, y) \text{ for } x\text{-periodic boundary conditions}, \\
&w(x, L) = w(x, 0) \text{ for } y\text{-periodic boundary conditions } \}\,.
\end{aligned}$$

Note that the values of the test functions on the boundary are such that $\psi \in H^1(\Omega)$. Multiplying (3) by a test function $w \in H_0^1(\Omega)$ and integrating over the domain results in

$$\int_\Omega \nabla \cdot (A \nabla \psi)\, w d\Omega - \int_\Omega B\psi w d\Omega + \int_\Omega C w d\Omega = \int_\Omega \omega w d\Omega. \tag{88}$$

Integrating the first term by parts results in

$$\begin{aligned}
\int_\Omega \nabla \cdot (A \nabla \psi)\, w d\Omega &= \int_{\partial\Omega} A w \nabla\psi \cdot \hat{n}\, d\Gamma - \int_\Omega A \nabla\psi \cdot \nabla w d\Omega \\
&= \int_{\partial\Omega_N} A w \nabla\psi \cdot \hat{n}\, d\Gamma + \int_{\partial\Omega_P} A w \nabla\psi \cdot \hat{n}\, d\Gamma \\
&\quad + \sum_{i=1}^{N} \int_{\partial\Omega_{D_i}} A w \nabla\psi \cdot \hat{n}\, d\Gamma - \int_\Omega A \nabla\psi \cdot \nabla w d\Omega.
\end{aligned}$$

The boundary integral over $\partial\Omega_N$ vanishes since $\nabla\psi \cdot \hat{n} = 0$ for Neumann boundary conditions. Further, for periodic boundary conditions the boundary integral

over $\partial\Omega_P$ vanishes because of (5)–(7) and (8). Using (15) and the fact that $w \in W_h^k$ is constant on parts of the boundary with no normal flow boundary conditions, we can write for the boundary integral over $\partial\Omega_{D_i}$ $(i = 1, 2, \ldots, N)$

$$\begin{aligned}\int_{\Omega_{D_i}} Aw\nabla\psi \cdot \hat{n}\, d\Gamma &= \int_{\partial\Omega_{D_i}} w\vec{u} \cdot \hat{\tau}\, d\Gamma \\ &= w\big|_{\partial\Omega_{D_i}} \int_{\partial\Omega_{D_i}} \vec{u} \cdot \hat{\tau}\, d\Gamma \\ &\overset{(15)}{=} w\big|_{\partial\Omega_{D_i}} \mathcal{K}_i.\end{aligned}$$

The weak formulation then becomes: *Find* $\psi \in H^1(\Omega)$ *such that for each* $w \in H^1(\Omega)$ *the following equation holds*

$$\int_\Omega A\nabla\psi \cdot \nabla w d\Omega + \int_\Omega B\psi w d\Omega = \int_\Omega (C - \omega)\, w d\Omega + \sum_{i=1}^{N} w\big|_{\partial\Omega_{D_i}} \mathcal{K}_i. \quad (89)$$

Note that the left hand side of this equation is symmetric in ψ and w. This is an important property as it allows us to use an efficient method to solve the involving systems numerically.

8.2.2 Discretized Weak Formulation

For the discretized form of (89) we must first define a function space, W_h^k, in which we are going to approximate ψ and the test functions w. Contrary to the discontinuous Galerkin method, the basis functions must be continuous here. We define W_h^k as

$$W_h^k = V_h^k \cap C(\Omega) \cap H^1(\Omega) \quad (90)$$

with $C(\Omega)$ the space of continuous functions on Ω. Note that the functions in this space are continuous, but their derivatives in general are not. As we will see in Section 8.3 it is important that $W_h^k \subset V_h^k$ in order to prove conservation of numerical energy.

The discretized weak formulation is obtained by replacing ψ, ω and w in (89) by $\psi_h \in W_h^k$, $\omega_h \in V_h^k$ and $w_h \in W_h^k$. The discretized weak formulation is now given by: *Find* $\psi_h \in W_h^k$ *such that for each* $w_h \in W_k^h$ *the equation*

$$\int_\Omega A\nabla\psi_h \cdot \nabla w_h d\Omega + \int_\Omega B\psi_h w_h d\Omega = \int_\Omega (C - \omega_h)\, w_h d\Omega + \sum_{i=1}^{N} w_h\big|_{\partial\Omega_{D_i}} \mathcal{K}_i. \quad (91)$$

holds.

The weak formulation can be written as a system of equations. In order to do that assume that a basis for W_h^k consisting of L_k basisfunctions is given by

$$\{\varphi_1, \varphi_2, \ldots, \varphi_{L_k}\}. \quad (92)$$

In the case of the discontinuous Galerkin method we could determine a set of basis functions for each element separately, due to the fact that test functions are discontinuous across element boundaries in that case. In the continuous Galerkin case we can't use the element wise approach. In Appendix B.2, we show how to choose these functions explictly. For ψ_h, we write

$$\psi_h = \sum_{i=1}^{L_k} \psi_i \varphi_i. \tag{93}$$

Substituting this in (91) and taking $w_h = \varphi_i$ for $i = 1, 2, \ldots N_k$ we obtain the following system of equations

$$\begin{aligned} &\sum_{j=1}^{L_k} \psi_j \int_\Omega A \nabla \varphi_j \cdot \nabla \varphi_i d\Omega + \sum_{j=1}^{L_k} \int_\Omega B \psi_j \varphi_i \varphi_j d\Omega = \\ &\int_\Omega (C - \omega_h)\, \varphi_i d\Omega + \sum_{l=1}^{N} \varphi_i\big|_{\partial\Omega_{D_l}} \mathcal{K}_l. \quad (i = 1, 2, \ldots, L_k). \end{aligned} \tag{94}$$

We rewrite (94) as a matrix equation

$$M \left[\psi_1, \psi_2, \ldots, \psi_{L_k}\right]^T = b \tag{95}$$

with M a $L_k \times L_k$ matrix given by

$$M_{ij} = \int_\Omega A \nabla \varphi_i \cdot \nabla \varphi_j + B \varphi_i \varphi_j d\Omega \tag{96}$$

and right hand side $b \in \mathbb{R}^{L_k}$ given by

$$b_i = \int_\Omega (C - \omega_h) \varphi_i d\Omega + \sum_{l=1}^{N} \varphi_i\big|_{\partial\Omega_{D_l}} \mathcal{K}_l. \tag{97}$$

Note that M is symmetric, $M = M^T$, and M is semi-positive definite, $M \geq 0$. These are important properties as they allow us to use the preconditioned conjugate gradient method[3]. The semi-positive definiteness of M can be seen as follows

$$\begin{aligned} &\left[\psi_1, \psi_2, \ldots, \psi_{L_k}\right]^T M \left[\psi_1, \psi_2, \ldots, \psi_{L_k}\right] = \\ &\sum_{i=1}^{L_k} \sum_{j=1}^{L_k} \psi_i \psi_j \int_\Omega A \nabla \varphi_i \cdot \nabla \varphi_j + B \varphi_i \varphi_j \, d\Omega = \\ &\qquad\qquad \int_\Omega A |\nabla \psi_h|^2 + B |\psi_h|^2 \, d\Omega \geq 0, \end{aligned} \tag{98}$$

since $A > 0$ and $B \geq 0$.

If $B \equiv 0$ then $M \geq 0$. This can be seen by observing that if $\left[\psi_1, \psi_2, \ldots, \psi_{L_k}\right]^T$ is a solution of (95) then $[\psi_1 + K, \psi_2 + K, \ldots, \psi_{L_k} + K]$ is also a solution for

any constant $K \in \mathbb{R}$. In this case it is necessary to add an extra condition on ψ_h. For instance we could demand $\int_\Omega \psi_h d\Omega = 0$ or we could fix the value of ψ_h for a certain point $(x_0, y_0) \in \Omega$: $\psi_h(x_0, y_0) = 0$.

On the other hand, if $B \not\equiv 0$ then we have $M > 0$. This can be seen by observing that if

$$[\psi_1, \psi_2, \ldots, \psi_L]^T M [\psi_1, \psi_2, \ldots, \psi_L] = 0 \tag{99}$$

then

$$\int_\Omega A|\nabla\psi_h|^2 d\Omega = 0 \quad \text{and} \quad \int_\Omega B|\psi_h|^2 d\Omega = 0. \tag{100}$$

Now, if $\int_\Omega A|\nabla\psi_h|^2 d\Omega = 0$ then, since $A > 0$, $\nabla\psi_h \equiv 0$ and thus $\psi_h = K$ for a certain constant $K \in \mathbb{R}$. But then

$$0 = \int_\Omega B|\psi_h|^2 d\Omega = K^2 \int_\Omega B d\Omega \tag{101}$$

is only satisfied for $K = 0$ since $B \geq 0$ and $B \not\equiv 0$. Thus we have that if (99) holds then $[\psi_1, \psi_2, \ldots, \psi_{L_k}] = 0$, and hence $M > 0$. In this case the matrix M is invertible and we don't need to impose further conditions on ψ_h.

8.3 Conservation of Numerical Energy and Enstrophy

We define the numerical energy, E_h, as the energy associated with the numerical solution of (1)–(3):

$$E_h(t) = \frac{1}{2} \int_\Omega A|\nabla\psi_h|^2 + B|\psi_h|^2 d\Omega. \tag{102}$$

The next theorem is the equivalent of Theorem 6.1 for the numerical solution. It states that the energy of the numerical solution is conserved exactly.

Theorem 8.1 (Numerical Energy Conservation). *Consider the numerical solution of (1)–(3), obtained by the space discretization described in Section 8.1 and Section 8.2 and subject to periodic boundary conditions (see Section 5.2) or no normal flow boundary conditions (see Section 5.3). The energy associated with this numerical solution, E_h, is a conserved quantity:*

$$\frac{d}{dt} E_h(t) = 0. \tag{103}$$

Proof. First we differentiate (91) with respect to time to obtain

$$\int_\Omega A\nabla w_h \cdot \nabla\left(\partial_t\psi_h\right) d\Omega + \int_\Omega B w_h \partial_t\psi_h d\Omega = -\int_\Omega \partial_t\omega_h w_h d\Omega. \tag{104}$$

Note that we used the fact that the test functions are time independent. Evaluating this equation at an arbitrary time $t = t_a$ and taking $w_h = \psi_h\big|_{t=t_a} \in W_h^k(\Omega)$ results in

$$\left[\int_\Omega A\nabla\psi_h \cdot \partial_t\nabla\psi_h d\Omega + \int_\Omega B\psi_h\partial_t\psi_h d\Omega\right]_{t=t_a} = -\left[\int_\Omega \psi_h\partial_t\omega_h d\Omega\right]_{t=t_a} \tag{105}$$

and thus

$$\left[\frac{d}{dt}E_h\right]_{t=t_a} = -\left[\int_\Omega \psi_h \partial_t \omega_h d\Omega\right]_{t=t_a}. \tag{106}$$

Now we use (82) with $v_h = \psi_h\big|_{t=t_a} \in W_h^k(\Omega) \subset V_h^k(\Omega)$ to obtain

$$\begin{aligned}\left[\int_\Omega \psi_h \partial_t \omega_h d\Omega\right]_{t=t_a} &= \sum_{K\in\mathcal{T}_h}\left[\int_K \psi_h \partial_t \omega_h dK\right]_{t=t_a} \\ &= \left[\sum_{K\in\mathcal{T}_h}\int_K \omega_h \vec{u}_h\cdot\nabla\psi_h dK - \sum_{K\in\mathcal{T}_h}\int_{\partial K}\psi_h \hat{f}(\omega_h^+,\omega_h^-,\nabla\psi_h\cdot\hat{\tau})d\Gamma\right]_{t=t_a} \\ &= 0. \end{aligned} \tag{107}$$

The first term vanishes since $\vec{u}_h \cdot \nabla\psi_h = 0$. The second term vanishes because ψ_h is continuous across boundaries and because of (78). Since t_a was chosen arbitrarily we obtain (103). □

Note that an important step in this proof consists of using ψ_h as a test function in the discontinuous Galerkin discretization (82). Therefore it is essential that the space of test functions for the continuous Galerkin discretization has the property that $W_h^k(\Omega) \subset V_h^k$.

The numerical enstrophy, S_h, is defined as the enstrophy of the numerical solution of (1)–(3):

$$S_h(t) = \frac{1}{2}\int_\Omega A\omega_h^2 d\Omega. \tag{108}$$

The numerical enstrophy is not conserved but it is possible to proof that it is L_2 stable under certain conditions.

Theorem 8.2 (L_2-stability of Enstrophy). *Consider the numerical solution of (1)–(3), obtained by the space discretization described in Section 8.1 and Section 8.2 and subject to periodic boundary conditions (see Section 5.2) or no normal flow boundary conditions (see Section 5.3). Let the numerical flux, $\hat{f}$, obey the following condition*

$$\omega_h^+ \geq \omega_h^- \Rightarrow \hat{f}(\omega_h^+,\omega_h^-,\vec{u}_h\cdot\hat{n}) \leq \overline{\omega}_h \vec{u}_h\cdot\hat{n}. \tag{109}$$

with $\bar{\omega} = (\omega^+ + \omega^-)/2$ and let A be be bounded on the domain, Ω, thus there exist $A_{min}, A_{max} \in \mathbb{R}$ such that for each $(x,y) \in \mathbb{R}^2$

$$A_{min} \leq A(x,y) \leq A_{max}. \tag{110}$$

Then the enstrophy of this numerical solution is a L_2-stable quantity:

$$\frac{d}{dt}S_h(t) \leq 0. \tag{111}$$

In the special case that A is independent of (x, y) and the numerical flux, $\hat{f}$, is the central flux given by (79), then we have exact conservation of numerical enstrophy

$$\frac{d}{dt}S_h(t) = 0. \tag{112}$$

Proof. Using (82) with $v_h = \omega_h$ we can write

$$\begin{aligned}
\frac{d}{dt}S(t) &= \int_\Omega A\omega_h \partial_t \omega_h d\Omega \\
&\leq A_{max}\int_\Omega \omega_h \partial_t \omega_h d\Omega \qquad (113)\\
&= A_{max}\sum_{K\in\mathcal{T}_h}\int_K \omega_h \partial_t \omega_h dK \\
&\overset{(82)}{=} A_{max}\sum_{K\in\mathcal{T}_h}\left(\int_K \omega_h \vec{u}_h\cdot\nabla\omega_h dK - \int_{\partial K}\omega_h^- \hat{f}(\omega_h^+,\omega_h^-,\vec{u}_h\cdot\hat{n}d\Gamma\right).
\end{aligned}$$

After integrating by parts this results in

$$\begin{aligned}
\frac{d}{dt}S(t) &\leq A_{max}\sum_{K\in\mathcal{T}_h}\left(\int_{\partial K}\omega_h^-\left(\omega_h^-\vec{u}_h\cdot\hat{n} - \hat{f}(\omega_h^+,\omega_h^-,\vec{u}_h\cdot\hat{n})\right)d\Gamma\right. \\
&\quad \left. - \int_K \omega_h\nabla\cdot(\vec{u}_h\omega_h)\,dK\right) \qquad (114)\\
&\leq \frac{A_{max}}{A_{min}}\sum_{K\in\mathcal{T}_h}\left(\int_{\partial K}A\omega_h^-\left(\omega_h^-\vec{u}_h\cdot\hat{n} - \hat{f}(\omega_h^+,\omega_h^-,\vec{u}_h\cdot\hat{n})\right)d\Gamma\right. \\
&\quad \left. - \int_K A\omega_h\nabla\cdot(\vec{u}_h\omega_h)\,dK\right). \qquad (115)
\end{aligned}$$

Using integration by parts (IP) and (71) we get

$$\begin{aligned}
&\int_K A\omega_h\nabla\cdot(\omega_h\vec{u}_h)\,dK \overset{(IP)}{=} \int_{\partial K}A(\omega_h^-)^2\vec{u}_h\cdot\hat{n}\,d\Gamma - \int_K \omega_h\vec{u}_h\cdot\nabla\left(A\omega_h\right)\,dK \\
&= \int_{\partial K}A(\omega_h^-)^2\vec{u}_h\cdot\hat{n}\,d\Gamma - \int_K A\omega_h\vec{u}_h\cdot\nabla\omega_h\,dK - \int_K \omega_h^2\vec{u}_h\cdot\nabla A\,dK \\
&= \int_{\partial K}A(\omega_h^-)^2\vec{u}_h\cdot\hat{n}\,d\Gamma - \int_K A\omega_h\vec{u}_h\cdot\nabla\omega_h\,dK - \int_K A\omega_h^2\vec{u}_h\cdot\frac{\nabla A}{A}\,dK \\
&\overset{(71)}{=} \int_{\partial K}A(\omega_h^-)^2\vec{u}_h\cdot\hat{n}\,d\Gamma - \int_K A\omega_h\vec{u}_h\cdot\nabla\omega_h\,dK - \int_K A\omega_h^2\nabla\cdot\vec{u}_h\,dK \\
&= \int_{\partial K}A(\omega_h^-)^2\vec{u}_h\cdot\hat{n}\,d\Gamma - \int_K A\omega_h\nabla\cdot(\omega_h\vec{u}_h)\,dK
\end{aligned}$$

and thus

$$\int_K A\omega\nabla\cdot(\omega_h\vec{u}_h)\,dK = \frac{1}{2}\int_{\partial K}A(\omega_h^-)^2\vec{u}_h\cdot\hat{n}\,d\Gamma. \tag{116}$$

Now we can write

$$\begin{aligned}\frac{d}{dt}S(t) &\leq \frac{A_{max}}{A_{min}}\sum_{K\in\mathcal{T}_h}\int_{\partial K} A\omega_h^-\left(\frac{\omega_h^-\vec{u}_h\cdot\hat{n}}{2}-\hat{f}(\omega_h^+,\omega_h^-,\vec{u}_h\cdot\hat{n})\right)d\Gamma \\ &= \frac{A_{max}}{A_{min}}\sum_{K\in\mathcal{T}_h}\left(\int_{\partial K}\frac{1}{2}A[\omega_h]\left(\hat{f}(\omega_h^+,\omega_h^-,\vec{u}_h\cdot\hat{n})-\overline{\omega}_h\vec{u}_h\cdot\hat{n}\right)d\Gamma\right. \\ &\quad \left.+\int_{\partial K}A\left(\overline{\omega_h^2}\vec{u}_h\cdot\hat{n}-\overline{\omega}_h\hat{f}(\omega_h^+,\omega_h^-,\vec{u}_h\cdot\hat{n})\right)d\Gamma\right) \qquad (117)\end{aligned}$$

with $\overline{\omega}_h=(\omega_h^++\omega_h^-)/2$, $\overline{\omega_h^2}=\left(\left(\omega_h^+\right)^2+\left(\omega_h^-\right)^2\right)/2$ and $[\omega_h]=\omega_h^+-\omega_h^-$. The second term vanishes after summing over all elements. The first term is lesser than or equal to zero because of (109). If $\omega_h^+\geq\omega_h^-$ then this follows immediately from (109) and $[\omega_h]=\omega_h^+-\omega_h^-\geq 0$. If $\omega_h^+\leq\omega_h^-$ then it follows from

$$\begin{aligned}\omega_h^+\leq\omega_h^- &\overset{(109)}{\Rightarrow} f(\omega_h^-,\omega_h^+,-\vec{u}_h\cdot\hat{n})\leq-\bar{\omega}_h\vec{u}_h\cdot\hat{n} \\ &\overset{(78)}{\Rightarrow} -f(\omega_h^+,\omega_h^-,\vec{u}_h\cdot\hat{n})\leq-\bar{\omega}_h\vec{u}_h\cdot\hat{n} \\ &\Rightarrow f(\omega_h^+,\omega_h^-,\vec{u}_h\cdot\hat{n})\geq\bar{\omega}_h\vec{u}_h\cdot\hat{n}.\end{aligned}$$

and $[\omega_h]=\omega_h^+-\omega_h^-\leq 0$. This proves the first statement of the theorem (111).

To prove the second statement we assume that A is independent of $(x,y)\in\Omega$ and that $\hat{f}$ is given by (79). Because $(x,y)\in\Omega$ the inequalities in (113), (114) and (115) become equalities. Thus we have

$$\begin{aligned}\frac{d}{dt}S(t) &= \sum_{K\in\mathcal{T}_h}\left(\int_{\partial K}\frac{1}{2}A[\omega_h]\left(\hat{f}(\omega_h^+,\omega_h^-,\vec{u}_h\cdot\hat{n})-\overline{\omega}_h\vec{u}_h\cdot\hat{n}\right)d\Gamma\right. \\ &\quad \left.+\int_{\partial K}A\left(\overline{\omega_h^2}\vec{u}_h\cdot\hat{n}-\overline{\omega}_h\hat{f}(\omega_h^+,\omega_h^-,\vec{u}_h\cdot\hat{n})\right)d\Gamma\right).\end{aligned}$$

The second term still vanishes after summing over all the elements. The first term vanishes because of (79). Hence we find (112). □

Essentially assumption (109) says that the flux which we use in our scheme should be at least as upwind as the central flux. By this we mean that $\hat{\omega}\equiv\hat{f}(\omega_h^+,\omega_h^-,\vec{u}_h\cdot\hat{n})/\vec{u}_h\cdot\hat{n}$ should be at least as close to ω_h^+ as to ω_h^- if $\vec{u}_h\cdot\hat{n}\leq 0$ and $\hat{\omega}$ should be at least as close to ω_h^- as to ω_h^+ if $\vec{u}_h\cdot\hat{n}\geq 0$. Or equivalently, $|\hat{\omega}-\omega^+|\leq|\hat{\omega}-\omega^-|$ if $\vec{u}_h\cdot\hat{n}\leq 0$ and $|\hat{\omega}-\omega^+|\geq|\hat{\omega}-\omega^-|$ if $\vec{u}_h\cdot\hat{n}\geq 0$. We can easily be check that this requirement is equivalent to (109). We can also check that the central flux, upwind flux and Lax-Friedrichs flux all satisfy this condition.

8.4 An Alternative Weak Formulation

In the previous section we saw that numerical enstrophy is conserved if A is constant and we use a central flux. It is possible to use an alternative weak

formulation for which enstrophy is always conserved if we take the central flux for $\hat{f}$. It is then not necessary that A is constant.

This alternative weak formulation differs from the weak formulation described in Section 8.1 and Section 8.2 in the choice of the spaces of test functions. For the discontinuous Galerkin method we define the space of test functions as $V^k_{h,alternative} = \left\{v/A \middle| v \in V^k_h\right\}$. The space of test functions for the elliptical part is the same as before $W^k_h = V^k_h \cap C(\Omega)$. Now we obtain our alternative weak formulation by multiplying (1) by Av with $v \in V^k_{h,alternative}$ and (3) by $w \in W^k_h$ and integrating over one element $K \in \mathcal{T}_h$ and the whole domain, respectively. After integrating by parts and introducing the numerical flux, $\hat{f}$, the weak formulation becomes: *Find* $\omega_h \in V^k_{h,alternative}$ *and* $\psi_h \in W^k_h$ *such that*

$$\int_K Av\partial_t\omega_h dK - \int_K \omega_h\vec{u}_h \cdot \nabla\left(Av\right) dK + \int_{\partial K} Av\hat{f}(\omega_h^+, \omega_h^-, \vec{u}_h \cdot \hat{n})d\Gamma = 0 \tag{118}$$

$$\vec{u}_h = A\nabla\psi_h \tag{119}$$

$$\int_\Omega A\nabla\psi_h \cdot \nabla w d\Omega + \int_\Omega B\psi_h w d\Omega = \int_\Omega (C - \omega_h) w d\Omega = 0 \tag{120}$$

for each $K \in \mathcal{T}_h$, $v \in V^k_{h,alternative}$ *and* $w \in W^k_h$.

Conservation of numerical energy can be proven by differentiating (120) with respect to time and then filling in $w = \psi_h$. This results in

$$\frac{d}{dt}E(t) = -\sum_{K\in\mathcal{T}_h}\int_K \psi_h\partial_t\omega_h dK = \sum_{K\in\mathcal{T}_h}\int_K A\frac{\psi_h}{A}\partial_t\omega_h dK. \tag{121}$$

Since $\psi_h/A \in V^k_{h,alternative}$ we can use (118) with $v = \psi_h/A$ to obtain

$$\begin{aligned}\frac{d}{dt}E(t) &= \sum_{K\in\mathcal{T}_h}\int_K \omega_h\vec{u}_h \cdot \nabla\left(A\frac{\psi_h}{A}\right) dK + \int_{\partial K}\psi_h\hat{f}(\omega_h^+, \omega_h^-, \vec{u}_h\cdot\hat{n})d\Gamma \\ &= \sum_{K\in\mathcal{T}_h}\int_K \omega_h\vec{u}_h \cdot \nabla\psi_h dK + \int_{\partial K}\psi_h\hat{f}(\omega_h^+, \omega_h^-, \vec{u}_h\cdot\hat{n})d\Gamma.\end{aligned}$$

The first term of this equation vanishes because $\vec{u}_h \perp \nabla\psi_h$ and the second term vanishes after summing over all the elements because of (78).

The proof of conservation of enstrophy is simplified a lot since we don not need A_{max} and A_{min} any longer. Fill in $v_h = \omega_h$ in (118) to obtain

$$\begin{aligned}\frac{dS}{dt} &= \int_\Omega A\omega_h\partial_t\omega_h d\Omega \\ &= \sum_{K\in\mathcal{T}_h}\left(\int_{\partial K} A\omega_h\hat{f}(\omega_h^+, \omega_h^-, \vec{u}_h\cdot\hat{n})d\Gamma - \int_K A\omega_h\vec{u}_h\cdot\nabla\omega_h dK\right).\end{aligned}$$

Using the same method as in the proof of Theorem 8.2 we can show that, if we assume (109), we have

$$\frac{d}{dt}S(t) \leq 0. \tag{122}$$

The advantage of this alternative weak formulation is that in the proof of enstrophy stability we don't need A_{min} and A_{max} any more, and thus enstrophy is exactly conserved if we take a central flux, even if A is not constant.

8.5 Time Stepping Procedure

In this section we will describe the algorithm to perform one time step. Thus given an approximation $\omega_h(t)$ at time t we will show how to obtain an approximation for $\omega_h(t+\Delta t)$ at $t+\Delta t$.

First we repeat the equations (1)–(3) in discretized form as we derived them in Section 8.1 and Section 8.2. Find $\omega_h \in V_h^k$ and $\psi_h \in W_h^k(\Omega)$ such that

$$\int_K v_h \partial_t \omega_h dK + \int_{\partial K} v_h^- \hat{f}(\omega_h^+, \omega_h^-, \vec{u}_h \cdot \hat{n}) d\Gamma \tag{123}$$

$$- \int_K \omega_h \vec{u}_h \cdot \nabla v_h dK = 0$$

$$u_h = A\nabla^\perp \psi_h \tag{124}$$

$$\int_\Omega A\nabla\psi_h \cdot \nabla w_h d\Omega + \int_\Omega B\psi_h w_h d\Omega = \int_\Omega (C - \omega_h)\, w_h d\Omega \tag{125}$$

holds for each $K \in \mathcal{T}_h$, $v_h \in V_h^k(K)$ and $w_h \in W_h^k(\Omega)$.

Note that using these equations we can write $d\omega_h/dt$ as a function of ω_h only

$$\frac{d\omega_h}{dt} = L(\omega_h). \tag{126}$$

Given ω_h we first solve the elliptical part of the system of equations. That is we calculate ψ_h using the boundary conditions and (125). Then we can calculate $\vec{u}_h$ from ψ_h using (124). And finally we calculate $d\omega_h/dt$ from ω_h, $\vec{u}_h$ and ψ_h using the hyperbolic equation (124).

8.5.1 Runge-Kutta Method

In discontinuous Galerkin methods explicit Runge Kutta methods are often used for the time integration. We use the method described used in [9] and described in [13] to obtain a 3-th order (in time) approximation of $\omega_h(t+\Delta t)$. This method is the third order Runge Kutta method given by:

$$\begin{aligned}
\omega^{(0)} &= \omega_h(t) \\
\omega^{(1)} &= \omega^{(0)} + \Delta t L(\omega^{(0)}) \\
\omega^{(2)} &= \frac{3}{4}\omega^{(0)} + \frac{1}{4}\omega^{(1)} + \frac{1}{4}\Delta t L(\omega^{(1)}) \\
\omega^{(3)} &= \frac{1}{3}\omega^{(0)} + \frac{2}{3}\omega^{(2)} + \frac{2}{3}\Delta t L(\omega^{(2)}) \\
\omega_h(t+\Delta t) &= \omega^{(3)}.
\end{aligned}$$

	0^{th} order	1^{th} order
1D mesh	$\Delta t\|u\|/\Delta x \leq \sqrt{3}$	$\Delta t\|u\|/\Delta x \leq \frac{\sqrt{3}}{6}$
2D mesh	$\Delta t\left(\|v\|/\Delta y + \|u\|/\Delta x\right) \leq \sqrt{3}$	–

Table 1: Conditions on the time step Δt.

A nice property of this scheme is that it is total variation diminishing. The total variation of a function is a measure for the oscillations in this function. According to [12] total variation diminishing schemes have the advantage of high-order accuracy in smooth regions, and few oscillations around discontinuities, at least for the one dimensional case.

8.5.2 CFL Condition

The Runge Kutta scheme given in Section 8.5.1 imposes a condition on the size of the time step. This is called the CFL condition. If the time step is taken too large then the scheme becomes unstable. The maximum size of the time step depends on the maximum velocity and the size of the elements. Higher velocities and smaller elements imply that the time step should be taken smaller. In Appendix E, a stability analysis has been performed for the vorticity equation with a fixed velocity field. This stability analysis has been done for a regular one dimensional mesh with 0-th or 1-th order basis functions and for a regular two dimensional mesh with 0-th order basis functions. The result of this analysis is a constraint on the time step for which the scheme is still stable. The results are shown in Table 1. Here Δt is the size of the time step, Δx and Δy are the length and width of each element and u and v are the horizontal and vertical velocity.

In the general case of a non regular grid and higher order basisfunctions we will use a constraint on the time step of the following form

$$\Delta t \leq CFL \min_{K \in \mathcal{T}_h} \frac{d_K}{\max_{(x,y)\in K} ||\vec{u}_h||}$$

with d_K the diameter of an element. In the case that the elements are quadrilaterals we will define this diameter as the minimum distance of the center of the element to one of its four edges. If an element is defined by the four vertices $\vec{q}_i$ $(i = 1, 2, 3, 4)$ then the center of the element is defined by by $\vec{q}_c = 1/4 \sum_{i=1}^{4} \vec{q}_i$. The parameter CFL depends on the time discretization scheme which we use and the order of the basisfunctions. We did not obtain an analytical expression for this parameter CFL in the case of higher order basisfunctions, however.

9 Verification Examples

In this section we present some examples which are used to test the numerical method. We implemented the finite element method described in Section 8 using C++. The details of the numerical method, such as the choice of basisfunctions and the evaluation of the integrals appearing in the weak formulation (124)–(125), are given in Appendices A–D. We We did not implement all the features of the method. For the continuous Galerkin discretization only first order basis functions are implemented. However, for the discontinuous Galerkin discretization the higher order basisfunction are implemented. Further the basisfunctions for the continuous and discontinuous Galerkin approximations (see Appendix B) are not chosen in such a way that $W_h^k \subset V_h^k$. This means that Theorem 8.1 does not apply in this case. However if we take the order of the basisfunctions for the discontinuous Galerkin discretization higher than those for the continuous Galerkin discretization then we *do* have $W_h^k \subset V_h^{k+1}$. This property allows us to verify Theorem 8.1 to some extent.

9.1 Example 1

In the first example we compare our results (qualitatively) with the results for the Euler equations in [9]. Since we use the Euler equations we have

$$A = 1 \qquad B = 0 \qquad C = 0. \tag{127}$$

We are going to solve for the initial conditions given by

$$\omega(x, y, 0) = \begin{cases} \delta \cos(x) - \frac{1}{\rho}\mathrm{sech}^2\left(\frac{y-\pi/2}{\rho}\right) & \text{if } y \leq \pi \\ \delta \cos(x) + \frac{1}{\rho}\mathrm{sech}^2\left(\frac{3\pi/2-y}{\rho}\right) & \text{otherwise} \end{cases} \tag{128}$$

with $\rho = \pi/15$ and $\delta = 0.05$. The domain is given by $\Omega = [0, 2\pi] \times [0, 2\pi]$. On all boundaries we take periodic boundary conditions. The exact solution for this problem appears to be unknown.

We calculate the solution for different orders $k = 1, 2, 3$. Note that these are only the orders for the discontinuous Galerkin basisfunctions. For the continuous Galerkin method we only implemented first order basisfunctions. We use a mesh consisting of $N \times N$ rectangles with $N = 64$. For the time step we take $\Delta t = 0.005$. For the numerical flux we take the upwind flux. The results for $t = 6$ and $t = 8$ are shown in figures 3 to 8. We see that our results resemble the ones in [9] very much.

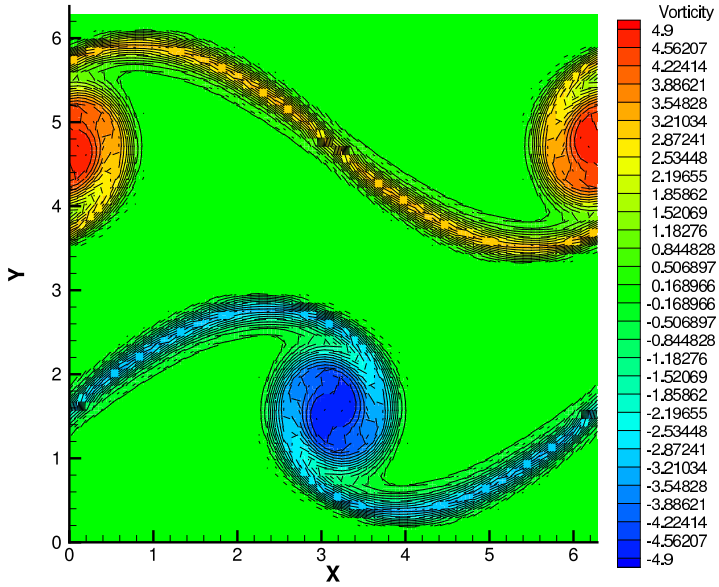


Figure 3: Contours of ω_h for example 1, $t = 6$, $k = 1$, mesh=64×64, $\Delta t = 0.005$, 30 contours from -4.9 to 4.9.

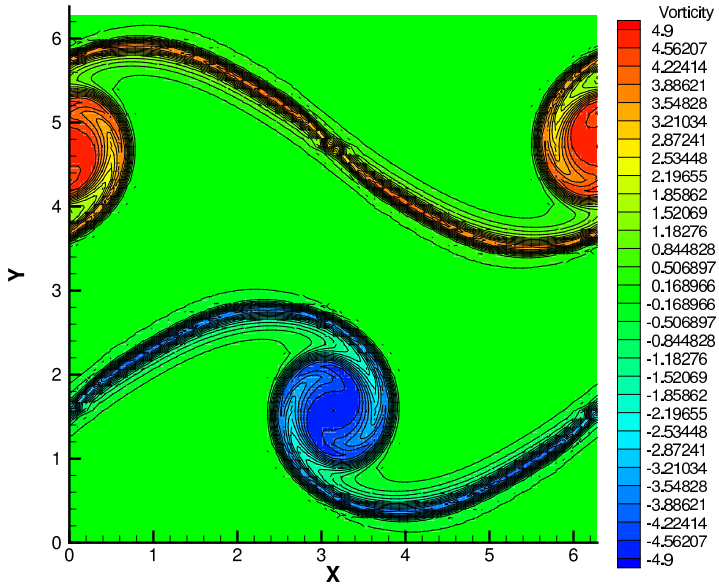


Figure 4: Contours of ω_h for example 1, $t = 6$, $k = 2$, mesh=64×64, $\Delta t = 0.005$, 30 contours from -4.9 to 4.9.

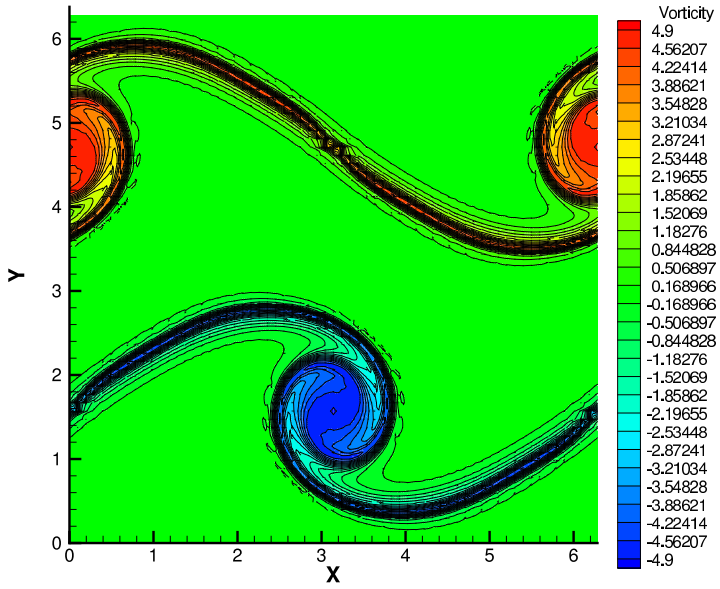


Figure 5: Contours of ω_h for example 1, $t = 6$, $k = 3$, $mesh = 64 \times 64$, $\Delta t = 0.005$, 30 contours from -4.9 to 4.9.

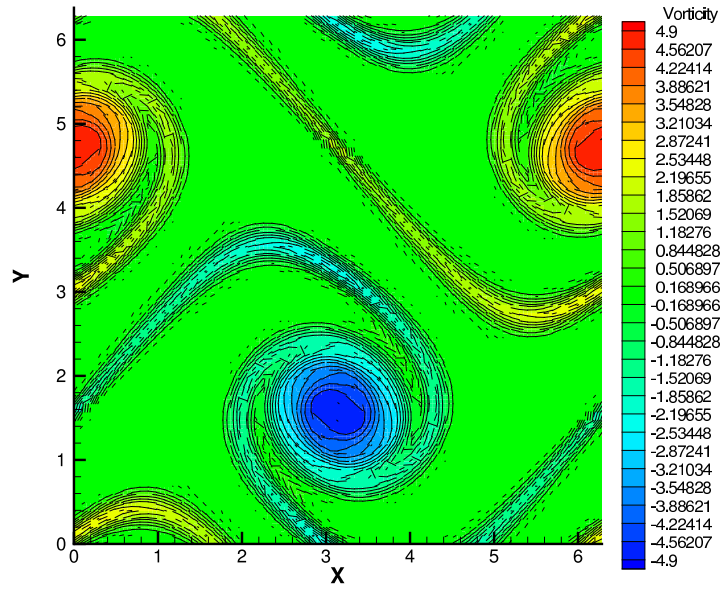


Figure 6: Contours of ω_h for example 1, $t = 8$, $k = 1$, mesh=64×64, $\Delta t = 0.005$, 30 contours from -4.9 to 4.9.

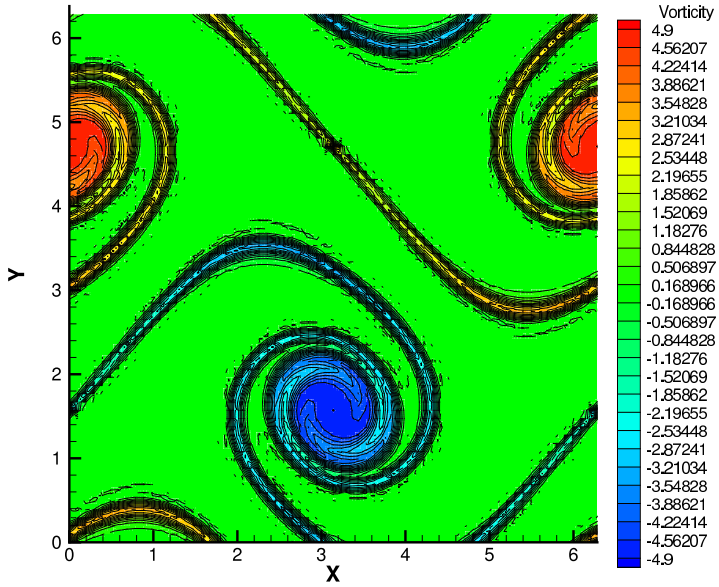


Figure 7: Contours of ω_h for example 1, $t = 8$, $k = 2$, mesh=64×64, $\Delta t = 0.005$, 30 contours from -4.9 to 4.9.

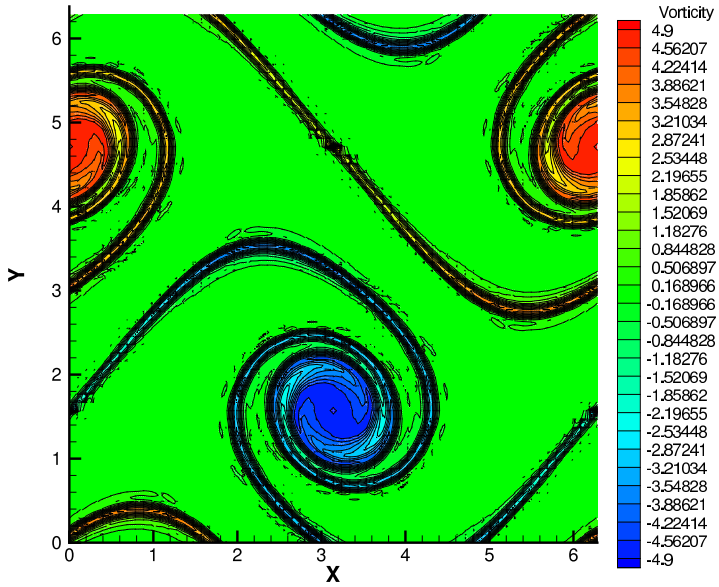


Figure 8: Contours of ω_h for example 1, $t = 8$, $k = 3$, mesh=64×64, $\Delta t = 0.005$, 30 contours from -4.9 to 4.9.

9.2 Example 2

This is another example which is also used in [9] for the Euler equations. Again we have

$$A = 1 \qquad B = 0 \qquad C = 0 \tag{129}$$

and

$$\Omega = [0, 2\pi] \times [0, 2\pi]. \tag{130}$$

The initial condition for ω is given by

$$\omega(x, y, 0) = \begin{cases} -1 & \text{if } \frac{\pi}{2} \leq x \leq \frac{3\pi}{2} \text{ and } \frac{\pi}{4} \leq y \leq \frac{3\pi}{4}, \\ 1 & \text{if } \frac{\pi}{2} \leq x \leq \frac{3\pi}{2} \text{ and } \frac{5\pi}{4} \leq y \leq \frac{7\pi}{4}, \\ 0 & \text{otherwise.} \end{cases} \tag{131}$$

On all boundaries we take periodic boundary conditions again.

We calculate the solution for different orders $k = 1, 2, 3$. We use a mesh consisting of $N \times N$ rectangles with $N = 64$. For the numerical flux we take the upwind flux. For the time step we take $\Delta t = 0.005$. The results for $t = 5$ and $t = 10$ are shown in figures 9 to 14. We see that our results again resemble the results in [9] very much.

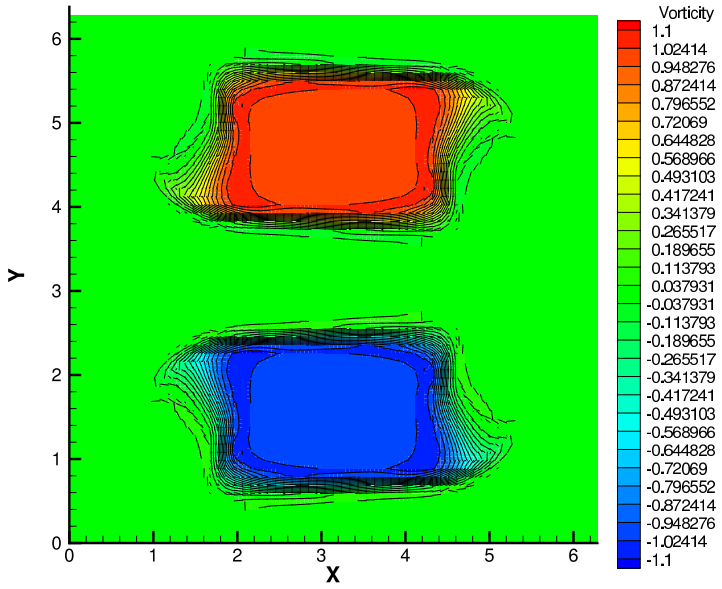


Figure 9: Contours of ω_h for example 2, $t = 5$, $k = 1$, mesh=64×64, $\Delta t = 0.005$, 30 contours from -1.1 to 1.1.

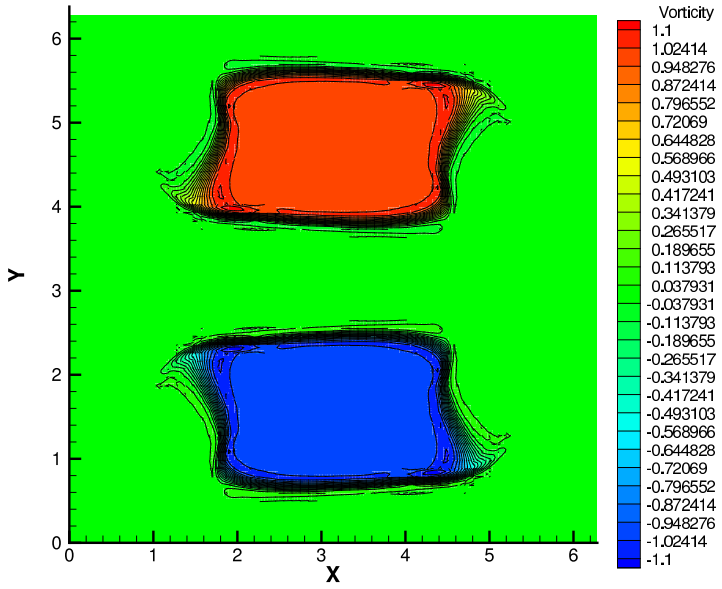


Figure 10: Contours of ω_h for example 2, $t = 5$, $k = 2$, mesh=64×64, $\Delta t = 0.005$, 30 contours from -1.1 to 1.1.

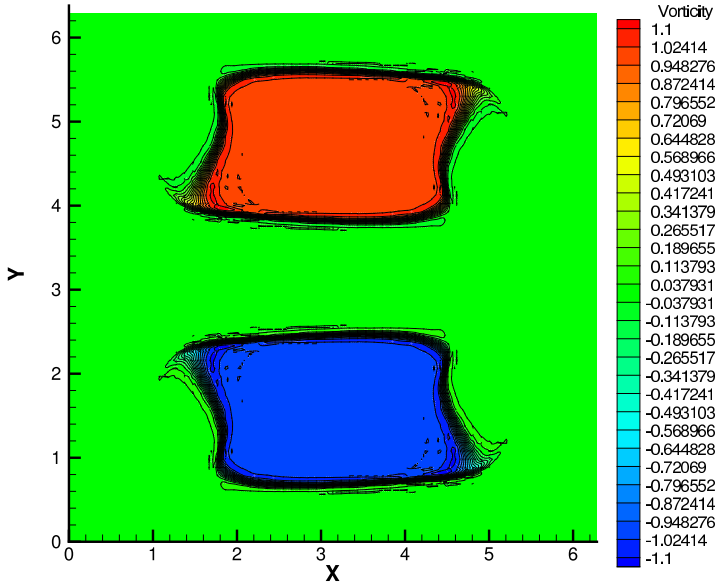


Figure 11: Contours of ω_h for example 2, $t = 5$, $k = 3$, mesh=64×64, $\Delta t = 0.005$, 30 contours from -1.1 to 1.1.

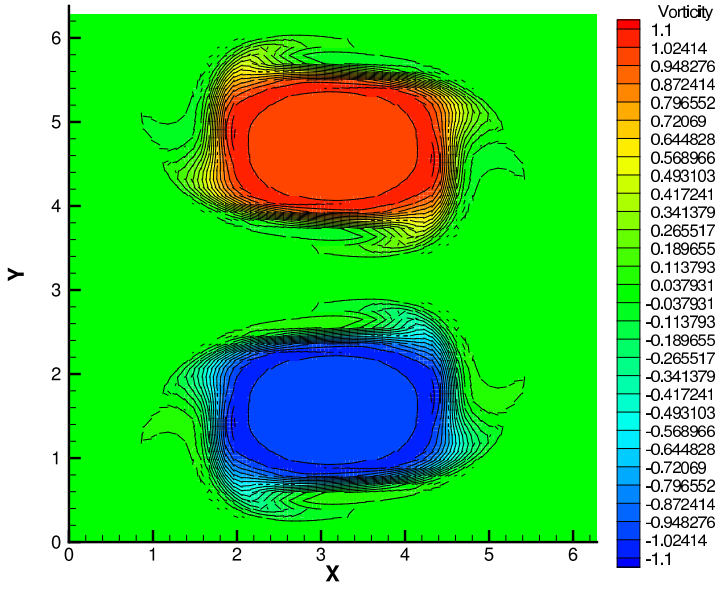


Figure 12: Contours of ω_h for example 2, $t = 10$, $k = 1$, mesh=64×64, $\Delta t = 0.005$, 30 contours from -1.1 to 1.1.

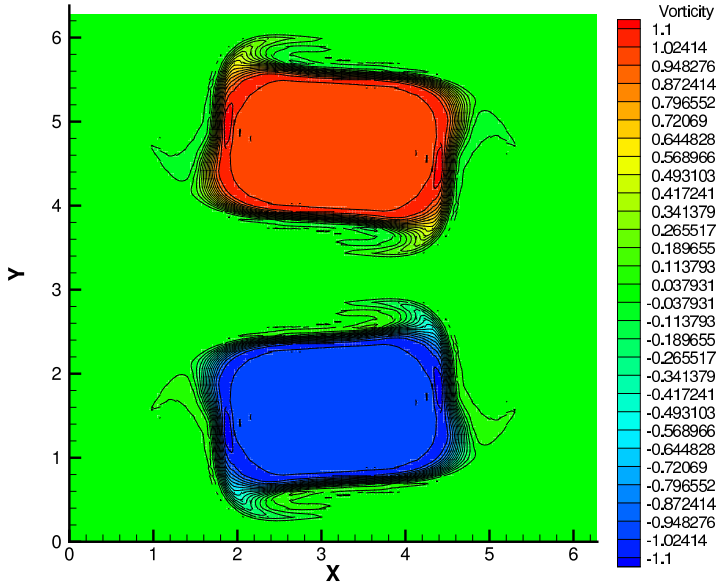


Figure 13: Contours of ω_h for example 2, $t = 10$, $k = 2$, mesh=64×64, $\Delta t = 0.005$, 30 contours from -1.1 to 1.1.

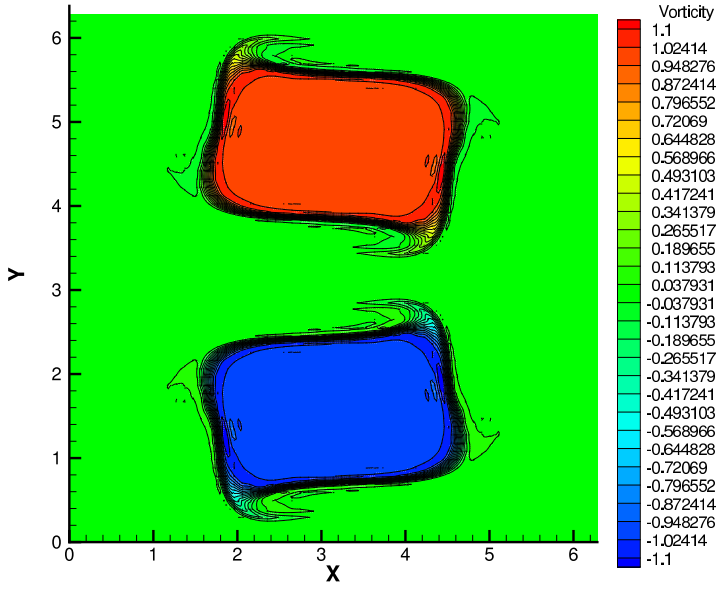


Figure 14: Contours of ω_h for example 2, $t = 10$, $k = 3$, mesh=64×64, $\Delta t = 0.005$, 30 contours from -1.1 to 1.1.

9.3 Example 3

The third example concerns an exact solution and thus we can investigate our numerical results more easily. First we show a method to construct a non-stationary solution from a stationary one. Assume that $(\omega', \psi', \vec{u}')$ satisfy the following equations

$$\begin{aligned} \nabla \cdot (\omega' \vec{u}') &= 0 && (132)\\ \vec{u}' &= A' \nabla^{\perp} \psi' && (133)\\ \nabla \cdot (A' \nabla \psi') - B' \psi' + C' &= \omega'. && (134) \end{aligned}$$

Here $\omega' = \omega'(x, y)$, $\psi' = \psi'(x, y)$, $\vec{u}' = \vec{u}'(x, y)$, $B' = B'(y) \leq 0$, $C' = C'(y)$ and A' is a constant $A' > 0$. The domain is given by $\Omega = \mathbb{R} \times I$ for some interval I. Note that (132)–(134) have the same form as (1)–(3) for stationary solutions.

Now define

$$\begin{aligned} \omega(x, y, t) &= \omega'(x - ct, y)\\ \vec{u}(x, y, t) &= \vec{u}'(x - ct, y) + [c, 0]^T\\ \psi(x, y, t) &= \psi'(x - ct, y) - \frac{cy}{A}\\ A &= A'\\ B(y) &= B'(y)\\ C(y) &= C'(y) - \frac{cyB'(y)}{A'}. \end{aligned}$$

Here $c \in \mathbb{R}$ is the speed of the traveling wave. Then equations (1)–(3) are satisfied. For (1) this can be seen as follows

$$\begin{aligned} &\partial_t \omega(x, y, t) + \nabla \cdot \big(\vec{u}(x, y, t)\omega(x, y, t)\big) =\\ &- c\partial_x \omega'(x - ct, y) + \nabla \cdot \big((\vec{u}'(x - ct, y) + [c, 0]^T)\omega'(x - ct, y)\big) =\\ &\qquad \nabla \cdot (\vec{u}'(x - ct, y)\omega'(x - ct, y)) = 0. \end{aligned}$$

Equation (2) is satisfied because

$$\begin{aligned} &\vec{u}(x, y, t) = \vec{u}'(x - ct, y) + [c, 0]^T = A' \nabla^{\perp} \psi'(x - ct, y) + [c, 0]^T =\\ &A' \nabla^{\perp} \psi'(x - ct, y) - c\nabla^{\perp} y = A' \nabla^{\perp} \left(\psi'(x - ct, y) - \frac{cy}{A}\right) = A \nabla^{\perp} \psi(x, y, t). \end{aligned}$$

And, finally, (3) is satisfied because

$$\begin{aligned} &\nabla \cdot (A \nabla \psi(x, y, t)) - B(y)\psi(x, y, t) + C(y) = \nabla \cdot \left(A' \nabla \left(\psi'(x - ct, y) - \frac{cy}{A'}\right)\right)\\ &- B'(y) \left(\psi'(x - ct, y) - \frac{cy}{A'}\right) + C'(y) - \frac{cyB'(y)}{A'} =\\ &\nabla \cdot \left(A' \nabla \psi'(x - ct, y) - A'[0, \frac{c}{A'}]^T\right) - B'(y)\psi'(x - ct, y) + C'(y) =\\ &\nabla \cdot (A \nabla \psi'(x - ct, y)) - B(y)\psi'(x - ct, y) + C'(y) =\\ &\qquad \omega'(x - ct, y) = \omega(x, y, t). \end{aligned}$$

Now we will consider an explicit choice for $(\omega', \psi', \vec{u}')$ which satisfy (132)–(134). We take

$$\begin{aligned}
\omega'(x,y) &= -2\sin(x)\sin(y) \\
\vec{u}'(x,y) &= [-\sin(x)\cos(y), \cos(x)\sin(y)]^T \\
\psi'(x,y) &= \sin(x)\sin(y) \\
A' &= 1 \\
B'(y) &= 0 \\
C'(y) &= 0.
\end{aligned}$$

Using the method described above we can construct the following non-stationary solution of (1)–(3) with $A = 1$, $B = 0$ and $C = 0$:

$$\omega(x,y,t) = -2\sin(x+t)\sin(y) \tag{135}$$
$$\vec{u}(x,y,t) = [-\sin(x+t)\cos(y) - 1, \cos(x+t)\sin(y)]^T \tag{136}$$
$$\psi(x,y,t) = \sin(x+t)\sin(y) + y. \tag{137}$$

Thus we took for the speed of the traveling wave $c = -1$. Thus we are going to solve (1)–(3) for the initial condition

$$\omega(x,y,0) = -2\sin(x)\sin(y). \tag{138}$$

For the domain we take

$$\Omega = [0, 2\pi] \times [0, 2\pi] \tag{139}$$

with periodic boundary conditions on the right and left boundary. Thus the part of the boundary with periodic boundary conditions is given by $\partial\Omega_P = \{0, 2\pi\} \times [0, 2\pi]$ and the boundary conditions on $\partial\Omega_P$ are given by (8)–(10) with $L = 2\pi$. For the upper and lower boundary we can take no normal flow boundary conditions (see Section 5.3). For the lower boundary, $\partial\Omega_{lower} = [0, 2\pi] \times \{0\}$, we can take the following (Dirichlet) boundary condition

$$\psi\big|_{y=0} = 0.$$

This is possible because $B = 0$ and thus ψ_h is only determined up to a constant. For the upper boundary we prescribe a value for the circulation, $\mathcal{K}_{upper}$. Thus the boundary condition for the upper boundary is given by

$$\int_{\partial\Omega_{upper}} \vec{u} \cdot \hat{\tau}\, d\Gamma = \mathcal{K}_{upper} \tag{140}$$

with $\partial\Omega_{upper} = [0, 2\pi] \times \{2\pi\}$. The value of the circulation, $\mathcal{K}_{upper}$, can be determined using the exact solution to evaluate (140) at $t = 0$. This results in $\mathcal{K}_{upper} = 2\pi$.

We run the program both for central flux and upwind flux, and for first, and second order basisfunctions for the discontinuous Galerkin discretization.

In the case of first order basisfunctions we use $CFL = 0.25$ and in the case of second order basisfunctions we use $CFL = 0.125$. We compute the solution on a regular grid for 8×8, 16×16, 32×32 and 64×64 elements. For all these runs the L_1 and L_∞ errors at $t = 4\pi$ and order of convergence for ω and ψ are shown in Tables 2 to 5.

k	Mesh	L^1 error	order	L^∞ error	order
1	8×8	1.01E+01	-	6.20E-01	-
	16×16	2.12E+00	2.25	2.28E-01	1.44
	32×32	3.49E-01	2.60	6.68E-02	1.77
	64×64	6.02E-02	2.53	1.89E-02	1.82
2	8×8	1.47E+00	-	2.49E-01	-
	16×16	2.72E-01	2.43	8.83E-02	1.50
	32×32	6.62E-02	2.04	3.56E-02	1.31
	64×64	1.69E-02	1.97	1.15E-02	1.63

Table 2: L^1 and L^∞ errors in ω at $t = 4\pi$ for first order ($k = 1$) and second order ($k = 2$) basisfunctions for the discontinuous Galerkin discretization using an upwind flux.

In Table 2 we see that the order of convergence of the L^1 error in the vorticity for first order ($k = 1$) basisfunctions and using an upwind flux is around 2.5. This result lies in between the order of convergence on irregular grids (around 2.2) and regular grids (around 3.0) which Liu and Shu found in [9]. The test case for which they calculated these order was a stationary solution which may explain the differences with our results.

If we use second order ($k = 2$) basisfunction for the discontinuous Galerkin discretization then we have some very unexpected results. Since the basisfunctions for the continuous Galerkin discretization are first order in both cases the order of convergence for $k = 2$ is not necessary higher than for $k = 1$. However we would expect the order of convergence for $k = 2$ to be not smaller than for $k = 1$. The results in Table 2 show that this is not the case: For $k = 2$ the order of convergence is significantly lower (both for L^∞ and L^1 errors) than for $k = 1$. We don't have a good explanation for these results.

Comparing the orders of convergence for the central flux (Table 3) to those for the upwind flux (Table 2) we see that the central flux has a much lower order of convergence in the case of first order basisfunction ($k = 1$). If we use second order basisfunctions ($k = 2$) then the orders of convergence for the central flux are not very much lower than those for an upwind flux.

Tables 4 and 5 show that the orders of convergence for the streamfunction are around 2 in all cases except the case of first order basisfunctions in combination with an upwind flux. Note that the negative effects of the central flux (compared to the upwind flux) are less severe for the streamfunction than for the vorticity.

We will now verify Theorem 8.1 and Theorem 8.2. To do this we must realize that Theorem 8.1 implies that numerical energy is conserved, independent of

k	Mesh	L^1 error	order	L^∞ error	order
1	8×8	1.10E+01	-	1.23E+00	-
	16×16	1.80E+00	2.61	2.64E-01	2.22
	32×32	7.51E-01	1.26	1.36E-01	0.95
	64×64	3.32E-01	1.18	7.54E-02	0.85
2	8×8	1.08E+01	-	2.10E+00	-
	16×16	2.21E+00	2.28	5.94E-01	1.82
	32×32	5.19E-01	2.09	1.42E-01	2.06
	64×64	1.39E-01	1.90	5.01E-02	1.50

Table 3: L^1 and L^∞ errors in ω at $t = 4\pi$ for first order ($k = 1$) and second order ($k = 2$) basisfunctions for the discontinuous Galerkin discretization using a central flux.

k	Mesh	L^1 error	order	L^∞ error	order
1	8×8	6.18E+00	-	4.75E-01	-
	16×16	1.28E+00	2.27	1.07E-01	2.15
	32×32	2.02E-01	2.66	1.82E-02	2.56
	64×64	3.15E-02	2.68	3.12E-03	2.54
2	8×8	1.42E+00	-	1.30E-01	-
	16×16	2.31E-01	2.61	2.39E-02	2.44
	32×32	5.27E-02	2.13	5.51E-03	2.12
	64×64	1.29E-02	2.03	1.35E-03	2.02

Table 4: L^1 and L^∞ errors in ψ at $t = 4\pi$ for first order ($k = 1$) and second order ($k = 2$) (discontinuous Galerkin) basisfunctions using an upwind flux.

the triangularisation which is used. The theorem also assumes that time is not discretized. Thus rather than increasing the number of elements, we should increase the number of time steps and check that the numerical energy, E_h, approaches a time independent function.

We will demonstrate that Theorem 8.1 and Theorem 8.2 appear to hold, even on a very coarse grid. We calculate E_h and S_h on a regular 3×3 grid. We will do this for different time steps: $CFL = 0.5, 0.25, 0.125, 0.0625$. We will use first order basisfunctions as well as second order basisfunctions for the discontinuous Galerkin discretization and we will use a central flux and an upwind flux. The total energy E_h and enstrophy S_h as a function of time are shown in Figure 15 to 18. Note that the results for second order basisfunctions and $CFL = 0.5$ are not given because this time step appears to be to large for the method to remain stable.

In Figure 15(a) and Figure 16(a) we see the consequence of the fact that for first order discontinuous Galerkin basisfunctions we don't have $W_h^k \subset V_h^k$: The numerical energy, E_h, is not a conserved quantity, even if we make the time step

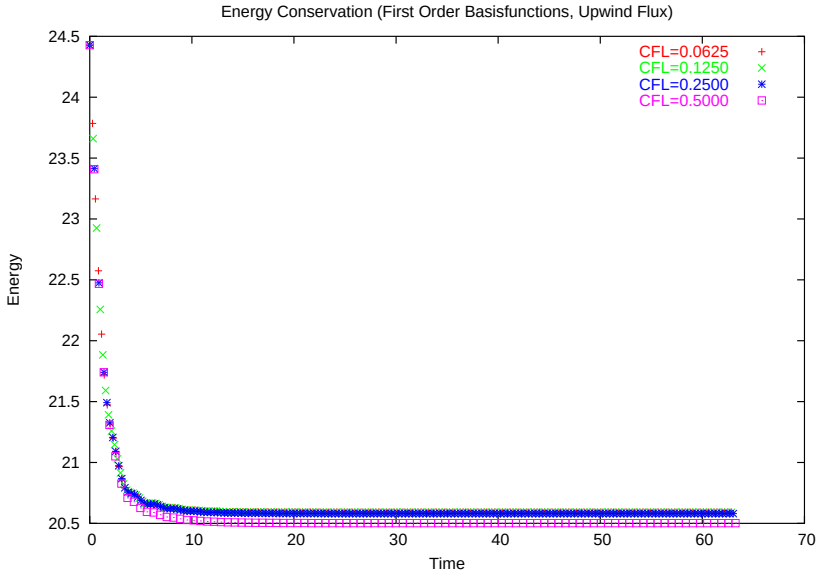
k	Mesh	L^1 error	order	L^∞ error	order
1	8×8	4.50E+00	-	2.85E-01	-
	16×16	7.57E-01	2.57	5.22E-02	2.45
	32×32	2.08E-01	1.86	1.44E-02	1.85
	64×64	5.13E-02	2.01	3.54E-03	2.02
2	8×8	3.63E+00	-	2.80E-01	-
	16×16	3.43E-01	3.40	3.34E-02	3.06
	32×32	5.76E-02	2.57	6.10E-03	2.45
	64×64	1.32E-02	2.12	1.39E-03	2.13

Table 5: L^1 and L^∞ errors in ψ at $t = 4\pi$ for first order ($k = 1$) and second order ($k = 2$) (discontinuous Galerkin) basisfunctions using a central flux.

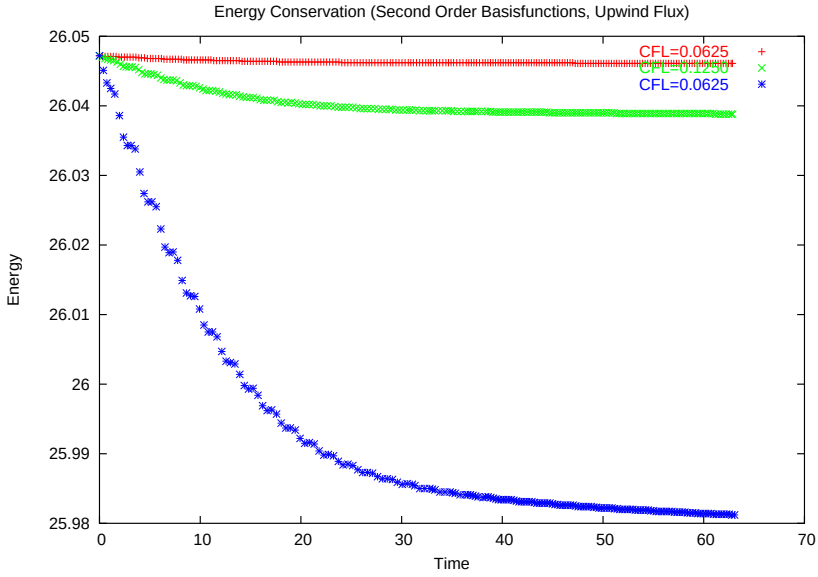
size smaller and smaller. In Figure 15(b) and Figure 16(b) we used second order discontinuous Galerkin basisfunctions and in this case we do have $W_h^k \subset V_h^k$. As a consequence the numerical energy E_h approaches a constant function if the time step is made small enough.

Enstrophy is not a conserved quantity if we use an upwind flux. However, according to Theorem 8.2 the enstrophy should be a decreasing quantity. This can be seen in Figure 17(a) and Figure 17(b).

If we use a central flux then enstrophy is conserved as can be seen in Figure 18(a) and Figure 18(b). Note that enstrophy is conserved in the case of second order as well as first order discontinuous Galerkin basisfunctions.

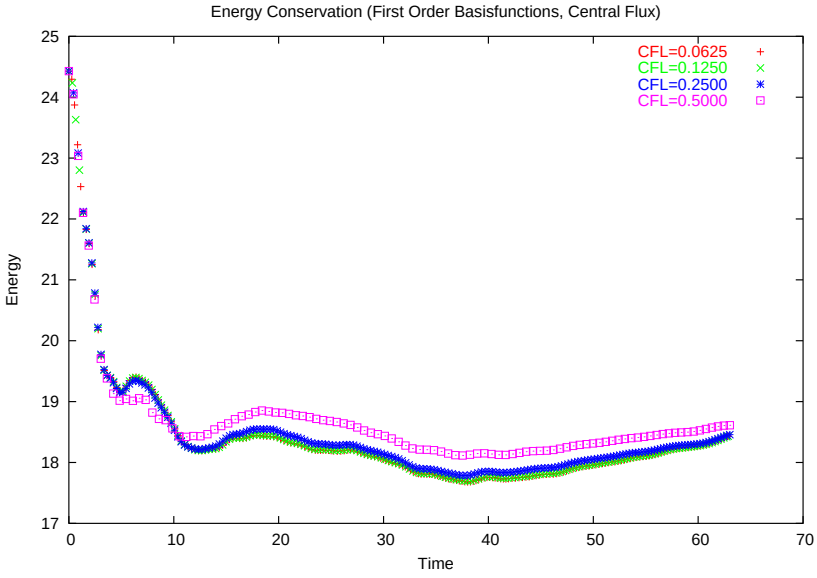


(a)

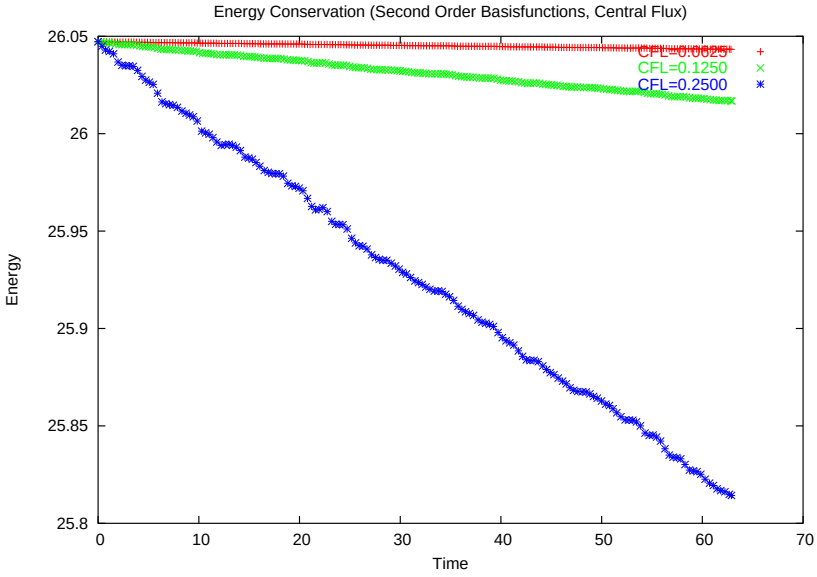


(b)

Figure 15: The total amount of numerical energy E_h on a 3×3 grid using an upwind flux for first and second order basisfunctions and for different time steps ($CFL = 0.5$, $CFL = 0.25$, $CFL = 0.125$, $CFL = 0.0625$).

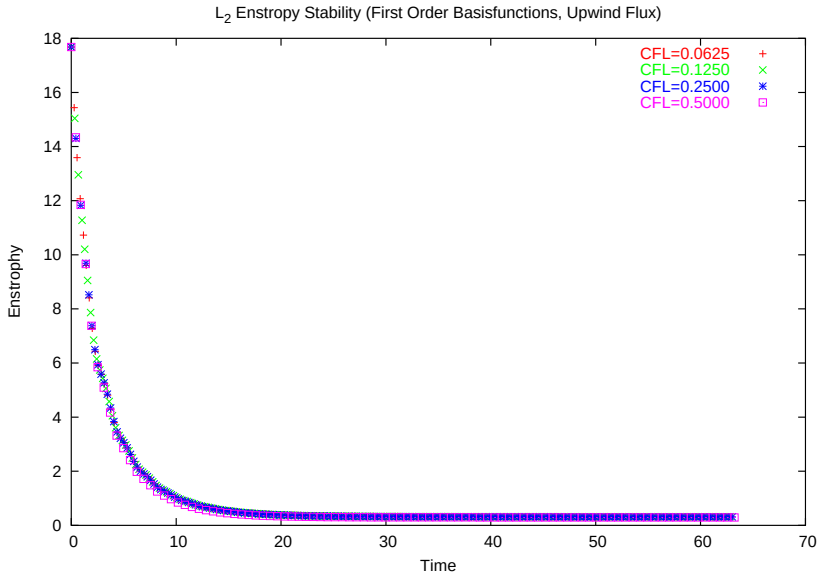


(a)

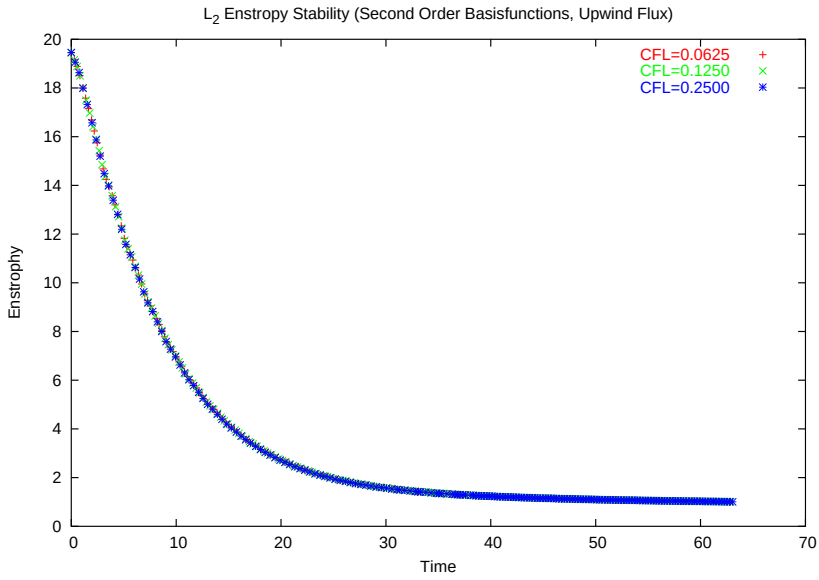


(b)

Figure 16: The total amount of numerical energy E_h on a 3×3 grid using an central flux for first and second order basisfunctions and for different time steps ($CFL = 0.5$, $CFL = 0.25$, $CFL = 0.125$, $CFL = 0.0625$).

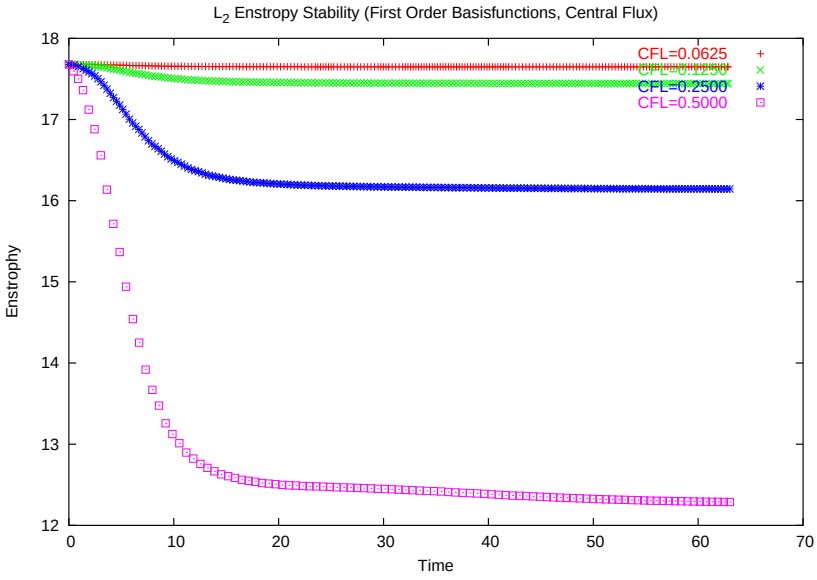


(a)

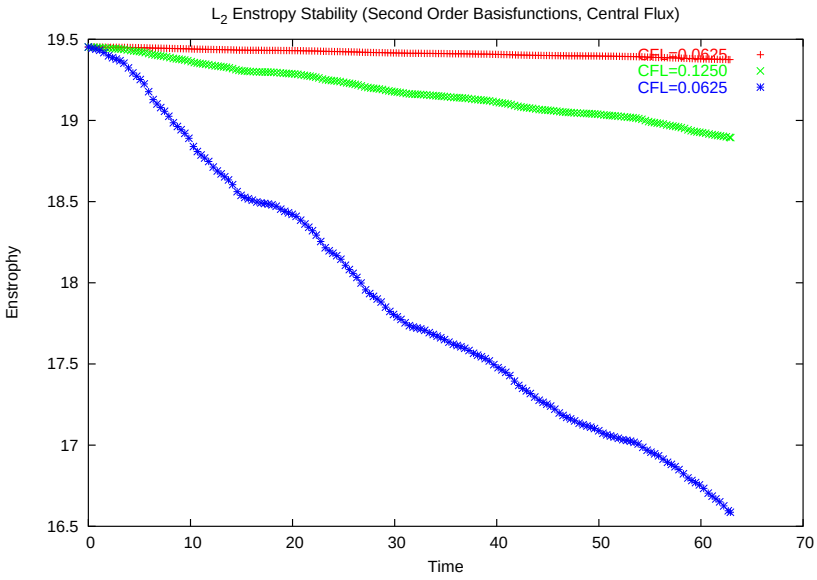


(b)

Figure 17: The total amount of numerical enstrophy S_h on a 3×3 grid using an upwind flux for first and second order basisfunctions and for different time steps ($CFL = 0.5$, $CFL = 0.25$, $CFL = 0.125$, $CFL = 0.0625$).



(a)



(b)

Figure 18: The total amount of numerical enstrophy S_h on a 3×3 grid using an central flux for first and second order basisfunctions and for different time steps ($CFL = 0.5$, $CFL = 0.25$, $CFL = 0.125$, $CFL = 0.0625$).

9.4 Example 4: A Traveling Wave

Here we discuss a traveling wave solution. Assume that $\psi' = \psi'(x, y)$ satisfies the equation

$$\nabla \cdot (A'\nabla\psi') - B'\psi' + C' = \lambda\psi' \quad \text{on } \Omega \tag{141}$$

with $B' = B'(y) \geq 0$, $C' = C'(y) \in \mathbb{R}$, $\lambda \neq 0$, A' a positive constant and the domain given by $\Omega = \mathbb{R} \times I$ for some interval $I \subset \mathbb{R}$. We can then construct a solution of (1)–(3) with $A = A'$, $B = B'$ and $C = C' - \beta y$ for any $\beta \in \mathbb{R}$. We assume that the streamfunction is of the following form

$$\psi(x, y, t) = \psi'(x - ct, y) \tag{142}$$

with $c \in \mathbb{R}$ the (constant) wave speed which will be determined later on. Using this assumption and (3) we obtain the following expression for ω

$$\begin{aligned}\omega(x, y, t) &= \nabla \cdot (A\nabla\psi(x, y, t)) - B(y)\psi(x, y, t) + C(y) \\ &= \nabla \cdot (A'\nabla\psi'(x - ct, y)) - B'(y)\psi'(x - ct, y) + C'(y) - \beta y \\ &\overset{(141)}{=} \lambda\psi'(x - ct, y) - \beta y. \end{aligned} \tag{143}$$

The velocity field follows from (2) and (142)

$$\vec{u}(x, y, t) = A\nabla^\perp\psi(x, y, t) = A'\nabla^\perp\psi'(x - ct, y). \tag{144}$$

We should now choose the wave speed c such that (1) holds. Writing out (1) results in

$$\begin{aligned} 0 &= \partial_t\omega(x, y, t) + \nabla \left(\vec{u}(x, y, t)\omega(x, y, t)\right) \\ &\overset{(144)}{=} \partial_t\omega(x, y, t) + A'\nabla^\perp\psi'(x - ct, y) \cdot \nabla\omega(x, y, t) \\ &\overset{(143)}{=} -\lambda c\partial_x\psi'(x - ct, y) + A'\nabla^\perp\psi'(x - ct, y) \cdot \nabla \left(\lambda\psi'(x - ct, y, t) - \beta y\right) \\ &= -\left(\lambda c + \beta A'\right)\partial_x\psi'(x - ct, y). \end{aligned}$$

Thus if we take for the wave speed

$$c = -\frac{A'\beta}{\lambda} \tag{145}$$

then (1) is satisfied.

We will now run our implementation of the numerical method for the following explicit choices of A', B', C', λ, β and ψ':

$$A' = 2 \qquad B'(y) = 1 \qquad C'(y) = 0 \tag{146}$$

$$\lambda = -5 \qquad \beta = 1/2 \qquad \psi'(x, y) = \sin(x)\sin(y). \tag{147}$$

Note that for these choices (141) is satisfied. Thus an exact solution of (1)–(3) is given by (143), (144) and (142).

k	Mesh	L^1 error	order	L^∞ error	order
1	8×8	2.32E+01	-	1.61E+00	-
	16×16	5.19E+00	2.16	4.55E-01	1.82
	32×32	1.04E+00	2.32	1.02E-01	2.15
2	8×8	9.04E+00	-	8.72E-01	-
	16×16	2.12E+00	2.09	2.15E-01	2.02
	32×32	5.41E-01	1.97	6.07E-02	1.82

Table 6: L^1 and L^∞ errors in ω for first order ($k = 1$) and second order ($k = 2$) (discontinuous Galerkin) basisfunctions using an upwind flux.

For the domain we take $\Omega = [0, 2\pi] \times [0, 2\pi]$ with periodic boundary conditions on the left and right boundaries. On the upper and lower boundaries we will take no-flow boundary conditions. Since we don't have $B = 0$ in this example we can't use Dirichlet boundary conditions for one of these boundaries as we did in the previous example. The boundary conditions for the upper boundary, $\partial\Omega_{upper} = [0, 2\pi] \times \{2\pi\}$, and lower boundary, $\partial\Omega_{upper} = [0, 2\pi] \times \{2\pi\}$, are thus given by

$$\int_{\partial\Omega_{upper}} \vec{u}_h \cdot \hat{n}\, d\Gamma = \mathcal{K}_{upper}$$
$$\int_{\partial\Omega_{lower}} \vec{u}_h \cdot \hat{n}\, d\Gamma = \mathcal{K}_{lower}.$$

The values for the circulations $\mathcal{K}_{upper}$ and $\mathcal{K}_{lower}$ can be calculated by using the exact solution. This results in $\mathcal{K}_{upper} = \mathcal{K}_{lower} = 0$.

We will run the program for different grid sizes to determine the order of convergence. We will also do runs for different orders of basisfunctions. We will use first order basisfunctions in the discontinuous Galerkin discretization, with $CFL = 0.25$, and second order basisfunctions, with $CFL = 0.125$.

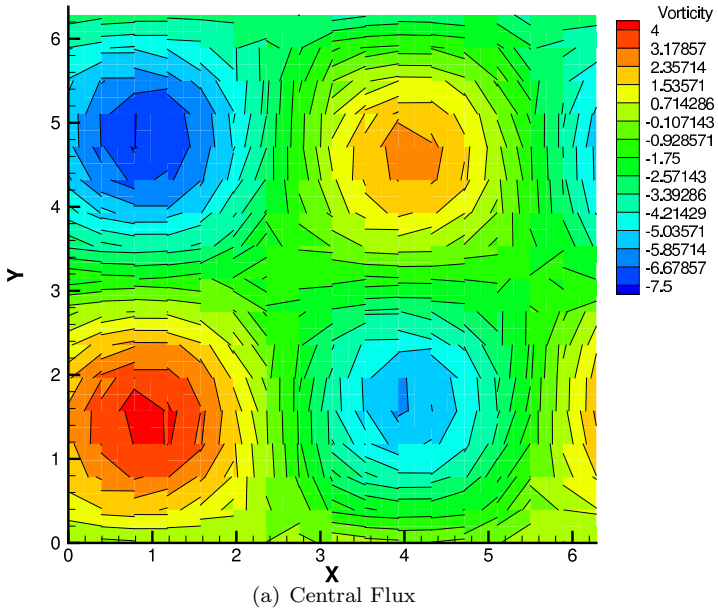
Figure 19 shows the numerical approximation ω_h at $t = 4\pi$ for a 16×16 grid using a central and upwind flux respectively. These pictures clearly illustrate that the discontinuities across element boundaries are much larger when using a central flux compared to using an upwind flux.

The L^1 and L^∞ errors and the orders of convergence for the different runs are presented in Tables 6 to 9.

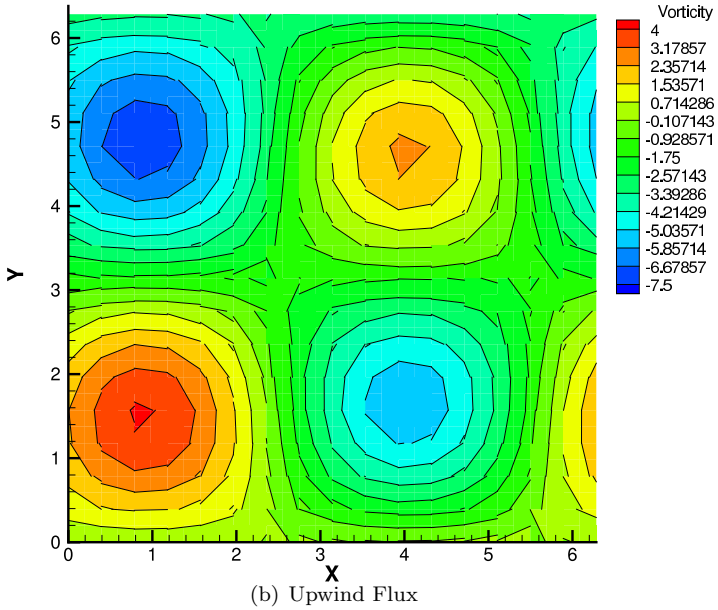
Table 6 shows the same strange behavior for the orders of convergence for the vorticity for $k = 2$ compared to $k = 1$ that we saw in the previous example: Using higher order basisfunctions in the discontinuous Galerkin discretization leads to a lower order of convergence!

In this case the difference between using an upwind flux and a central flux are dramatic. In the case of first order basisfunctions ($k = 1$) the L^∞ error doesn't decrease at all!

The orders of convergence for the errors in the streamfunction when using an



(a) Central Flux



(b) Upwind Flux

Figure 19: Value of ω_h at $t = 4\pi$ for example 4 on a 16×16 grid with first order basisfunctions ($k = 1$) and using an central flux and upwind flux respectively

k	Mesh	L^1 error	order	L^∞ error	order
1	8×8	2.59E+01	-	3.47E+00	-
	16×16	4.90E+00	2.40	8.93E-01	1.96
	32×32	3.58E+00	0.45	1.10E+00	-0.30
2	8×8	1.17E+01	-	1.64E+00	-
	16×16	2.40E+00	2.28	3.37E-01	2.28
	32×32	7.75E-01	1.63	1.85E-01	0.86

Table 7: L^1 and L^∞ errors in ω for first order ($k = 1$) and second order ($k = 2$) (discontinuous Galerkin) basisfunctions using a central flux.

k	Mesh	L^1 error	order	L^∞ error	order
1	8×8	4.86E+00	-	3.25E-01	-
	16×16	1.04E+00	2.22	7.61E-02	2.09
	32×32	2.02E-01	2.36	1.51E-02	2.33
2	8×8	1.85E+00	-	1.27E-01	-
	16×16	4.49E-01	2.04	3.18E-02	2.00
	32×32	1.13E-01	1.99	7.94E-03	2.00

Table 8: L^1 and L^∞ errors in ψ for first order ($k = 1$) and second order ($k = 2$) (discontinuous Galerkin) basisfunctions using an upwind flux.

k	Mesh	L^1 error	order	L^∞ error	order
1	8×8	1.59E+00	-	1.53E-01	-
	16×16	4.42E-01	1.85	3.40E-02	2.17
	32×32	1.93E-01	1.20	1.35E-02	1.33
2	8×8	2.16E+00	-	1.45E-01	-
	16×16	4.59E-01	2.23	3.22E-02	2.17
	32×32	1.08E-01	2.08	7.65E-03	2.07

Table 9: L^1 and L^∞ errors in ψ for first order ($k = 1$) and second order ($k = 2$) (discontinuous Galerkin) basisfunctions using a central flux.

upwind flux (Table 8 are very similar to the results for this case in the previous example (Table 4).

For the central scheme the orders of convergence in the streamfunction compared to (Table 8) are much worse in the case of first order ($k = 1$) basisfunction. In the case of second order basisfunctions the results for this example are similar to those we found in the previous example.

9.5 Example 5: Stuart Vortex

The Stuart vortex[5][6] is a stationary solution of (1)–(3) for the Euler equations. Thus we will take in (1)–(3)

$$A = 1, \qquad B = 0 \quad \text{and} \quad C = 0. \tag{148}$$

A stationary solution is a solution for which ψ, ω and $\vec{u}$ are time independent. A sufficient conditions for a solution to be stationary[4] is

$$\omega = f(\psi). \tag{149}$$

for some function f. This can be seen by substituting (149) in (1)

$$\begin{aligned} 0 &= \partial_t \omega + \nabla \cdot (\omega \vec{u}) \\ &= f'(\psi)\partial_t \psi + \vec{u} \cdot \nabla f(\psi) + f(\psi)\nabla \cdot \vec{u} \\ &= f'(\psi)\partial_t \psi + \vec{u} \cdot f' \nabla \psi \\ &= f'(\psi)\partial_t \psi \end{aligned}$$

and thus, if $f'(\psi) \neq 0$, we have

$$\partial_t \psi = 0. \tag{150}$$

For the function f in (149) we take

$$f(\psi) = \exp(-2\psi).$$

In this case a solution for the stream function is given by

$$\psi = \log\left(a\cosh(y) + \sqrt{a^2-1}\cos(x)\right) \tag{151}$$

with $a \geq 1$. We take $a = 1.5$. For the domain we take

$$\begin{aligned} \Omega &= \left\{(x,y) \in \mathbb{R}^2 \middle| 0 \leq x \leq 2\pi, \psi(x,y) \leq \log\left(ac + a + \sqrt{a^2-1}\right)\right\} \\ &= \left\{(x,y) \in \mathbb{R}^2 \middle| 0 \leq x \leq 2\pi, |y| \leq \operatorname{acosh}\left(c + 1 + \frac{\sqrt{a^2-1}\,(1-\cos(x))}{a}\right)\right\} \end{aligned}$$

with some $c \geq 0$. We will take $c = 2.0$. The upper and lower boundary of this domain coincide with contour lines of the streamfunction. On the left and right boundaries of the domain we have $\nabla\psi \cdot \hat{n} = 0$. Thus we can impose Neumann boundary conditions on the left ($x = 0$) and right ($x = 2\pi$) boundaries. We assume that $\omega+ = \omega^-$ at these boundaries.

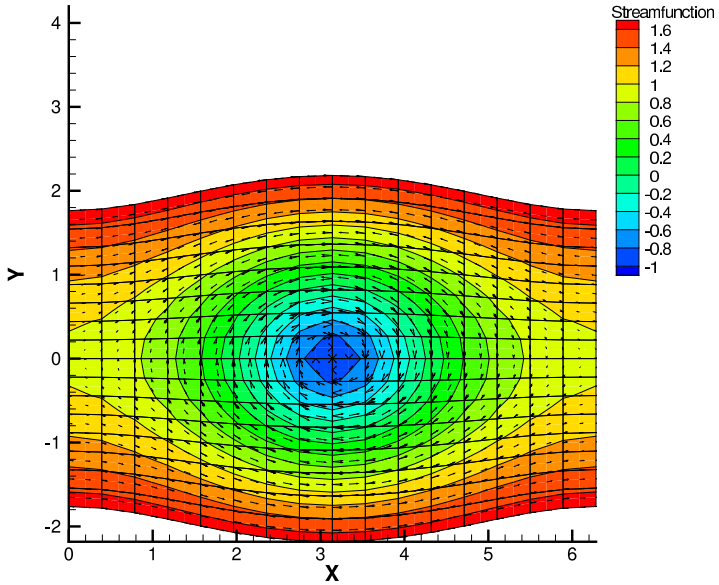
We calculate the numerical solution for 8×8, 16×16, 32×32 and 64×64 grids using an upwind flux and first order basisfunctions. The time step is determined by $CFL = 0.25$.

The velocity field and streamfunction at $t = 0$ and $t = 4$, respectively, are plotted in Figure 20. This figure shows that the velocity field coincides with the contours of the streamfunction. In Figure 21 we zoomed in at the center of the

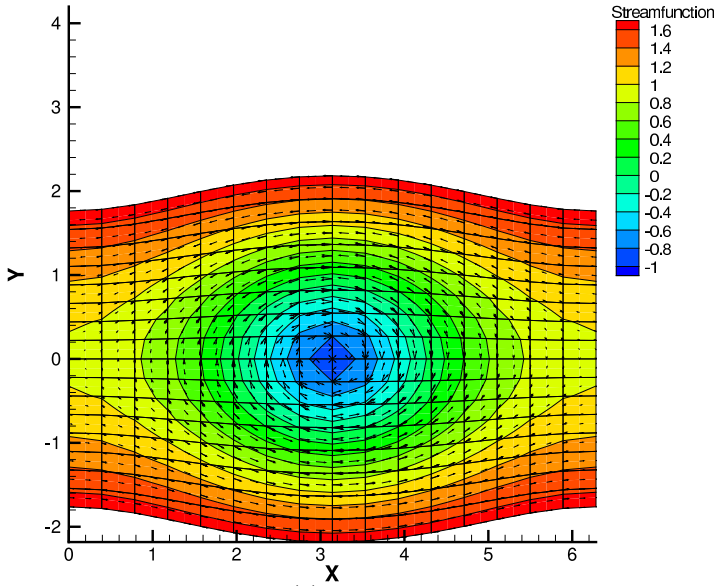
domain at $t = 0$. We see that although the velocity field is discontinuous the normal component of the velocity field across element boundaries is continuous. In Figure 22 we plotted the vorticity field and velocity field at $t = 0$ and $t = 4$, respectively. We see that the largest error in the vorticity field is located at the sharp peak in the center of the domain.

Figures 23 to 24 show the L^1 and L^∞ errors at $t = 4$ in the vorticity field and streamfunction, respectively. As expected, the numerical solution converges to the exact solution as we increase the number of elements. Note that the L^∞ error in the vorticity first decreases a bit and only later starts to increase. The reason for this is that for the initial condition we used the L^2 projection of the exact solution on the space of testfunctions V_h^k. The result is that the approximation of ω differs from the exact value of ω most in the center of the domain at the sharp peak in the vorticity. The exact value of the vorticity will be lower here than the value of the numerical approximation. As we saw the height of the sharp peak in the numerical solution decreases over time. During the first time steps the result of this is that the L^∞ error in the vorticity actually decreases.

Because we use Neumann boundary conditions at the left and right boundaries of the domain we can not expect that Theorem 8.1 and Theorem 8.2 hold. However, for the exact solution the energy and enstrophy are conserved. Figure 25 shows that if we increase the number of elements then dE_h/dt and dS_h/dt approach zero.



(a) $t = 0$



(b) $t = 4$

Figure 20: The streamfunction and velocity field at $t = 0$ and $t = 4$ for the Stuart vortex computed on a 16×16 grid.

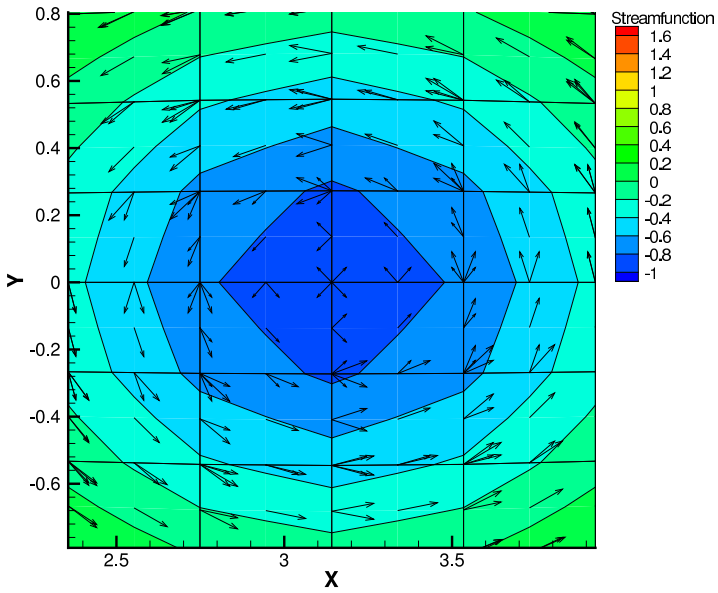
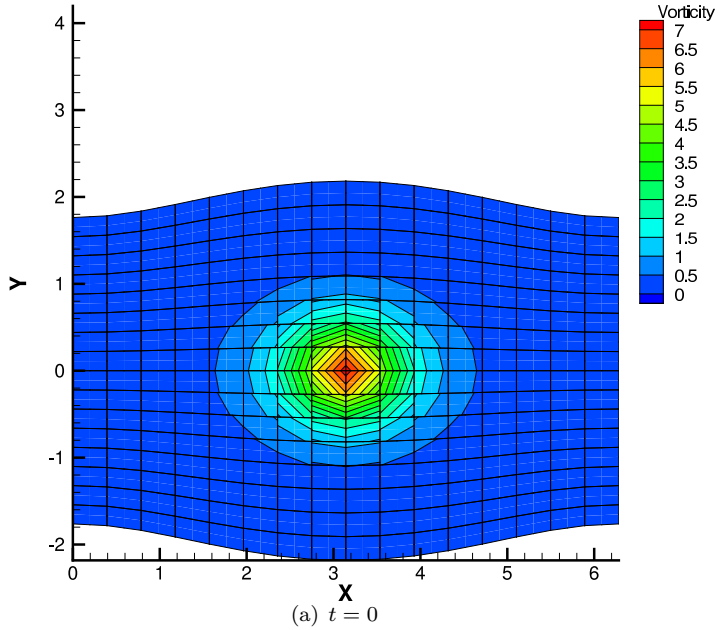
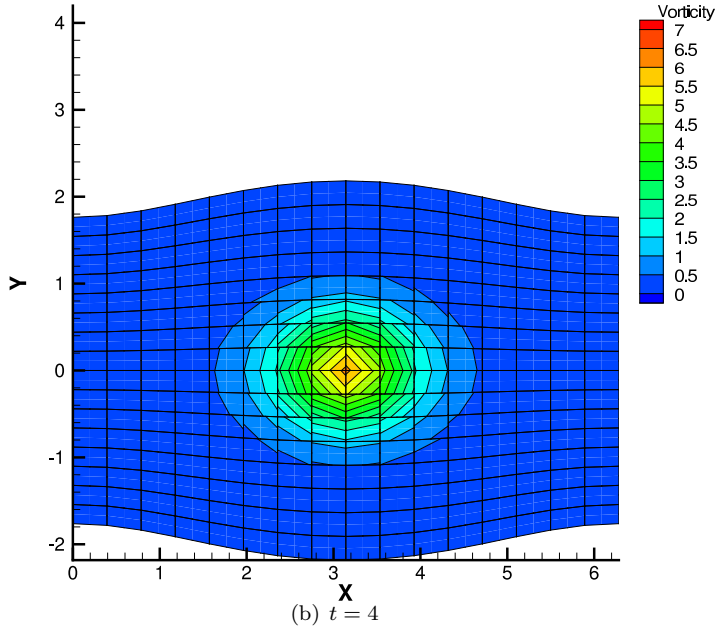


Figure 21: The streamfunction and velocity field at $t = 0$ zoomed in at the center of the domain for the Stuart vortex. The tangential component of the velocity field is discontinuous at element boundaries but the normal component is continuous.

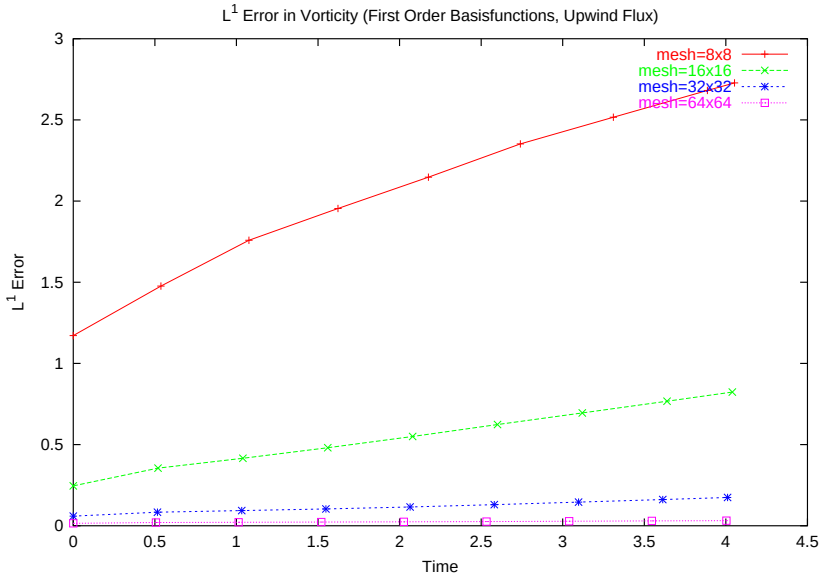


(a) $t = 0$

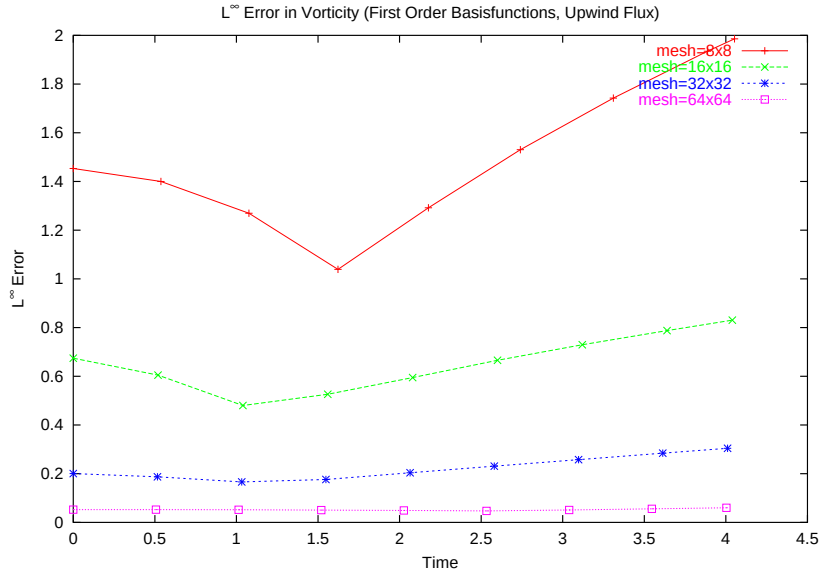


(b) $t = 4$

Figure 22: The vorticity field and velocity field at $t = 0$ and $t = 4$ for the Stuart vortex computed on a 16×16 grid.

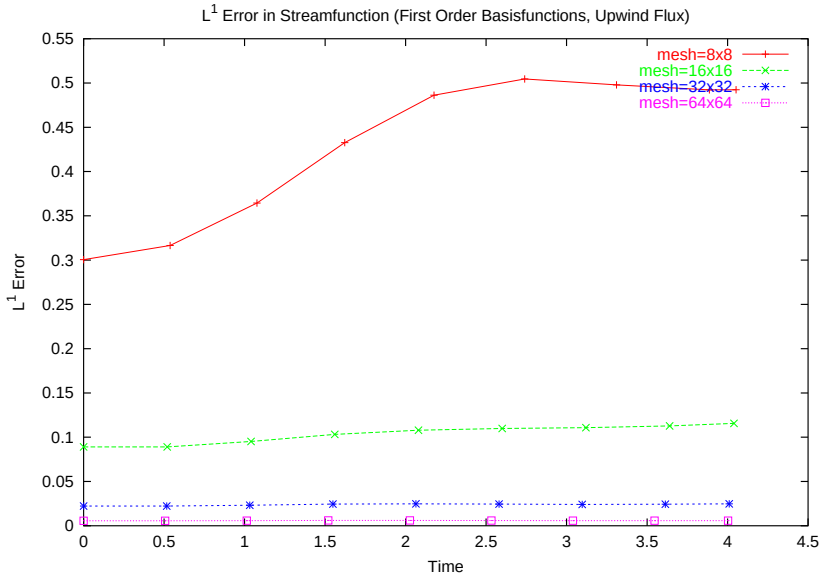


(a)

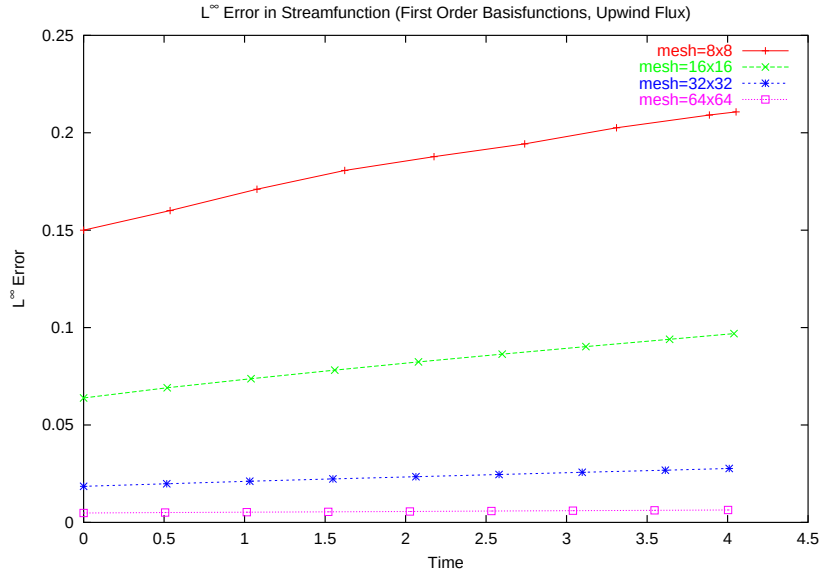


(b)

Figure 23: The L^1 and L^∞ errors in the vorticity field and at $t = 4$ for the Stuart vortex (Section 9.5) as a function of time for different grid sizes.

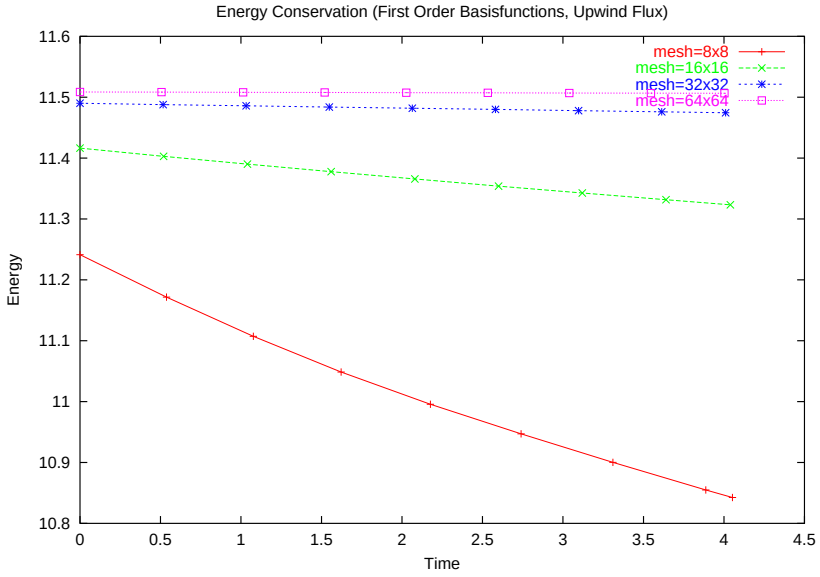


(a)

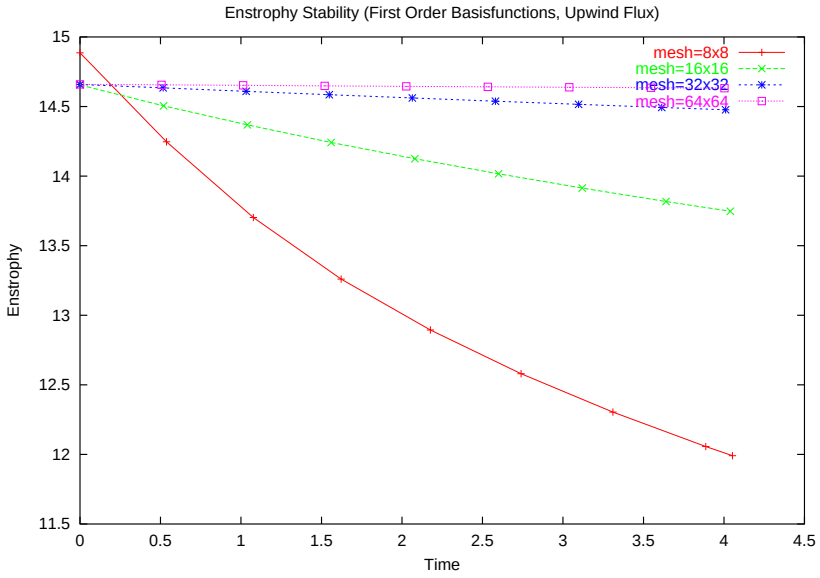


(b)

Figure 24: The L^1 and L^∞ errors in the streamfunction and at $t = 4$ for the Stuart vortex (Section 9.5) as a function of time for different grid sizes.



(a)



(b)

Figure 25: The numerical energy, E_h, and enstrophy, S_h, of the Stuart vortex (Section 9.5) as a function of time for different grid sizes.

10 Conclusions and Discussion

First of all, we reached the goals which we set in Section 3.2. We have been able to extend the method described in [9] for the generalized vorticity streamfunction formulation (1)–(3). These equations include the Euler equations, quasi-geostrophic equations and rigid lid equations as special cases.

Our method can also handle more general boundary conditions than the Dirichlet and periodic boundary conditions which were considered in [9]. If we have a domain with more than one no normal flow boundary conditions (i.e. a domain with islands) then the boundary conditions used in [9] are not sufficient. At least, no if we insist on having exact conservation of numerical energy. The boundary conditions which we introduced in Section 5.3 allow us to use domains with more than one no normal flow boundary condition. We have done this in such a way that numerical energy and enstrophy are a conserved and L_2 stable quantity, respectively. If we use a central flux and the function A in (2) is constant then we even showed that the numerical enstrophy is a conserved quantity as well.

We implemented and tested our method in C++. In the case of first order basisfunctions and an upwind flux we found that the orders of convergence for the L^1 error in both the vorticity and the streamfunction are a little bit higher than two. If we use higher order basisfunctions ($k = 2$) for the discontinuous Galerkin discretization, but first order basisfunctions for the continuous Galerkin discretization then the order of convergence is around two and lower than for the case $k = 1$. We don't have a good explanation for this behavior. Our tests show that the orders of convergence for the central flux are much lower than those for the upwind flux.

This result is a little bit paradoxical: The central flux seems to be a better flux than the upwind flux considering that numerical enstrophy is conserved in the first case and only L_2 stable in the second case (assuming that A is constant of course). On the other hand, considering the orders of convergence, the upwind flux definitely wins from the central flux.

We have the following suggestions for improvements and extension to the work we presented in this report:

- The basisfunctions which we choose in Appendix B don't have the property that $W_h^k(\Omega) \subset V_h^k$, which is an important requirement for Theorem 8.1. We had to take the basisfunctions for the discontinuous Galerkin discretization one order higher than those of the continuous Galerkin discretization to obtain the results of Theorem 8.1. Suggestion: We should define the k-th order basisfunction for the discontinuous Galerkin discretization in a different way, such that we do have $W_h^l(\Omega) \subset V_h^k$.

- At this moment we only implemented first order basisfunctions for the continuous Galerkin discretization. Suggestion: Implement higher order basisfunctions for the continuous Galerkin discretization.

- Except for the test in Section 9.4 we only performed tests for the case

$A = 1$, $B = 0$ and $C = 0$ in (3). We should do more test for more general values of A, B and C. In particular we want to find a test case for the rigid lid equations with $A = 1/H$ not a constant function.

- We assumed that A is a continuous function. One can think of cases in which it is useful to allow A to be discontinuous. For instance we might want to try to use some kind of step function for $A = 1/H$ in the rigid lid equations. If we assume that A is not continuous then we have to take into account the following: The velocity $\vec{u}_h = A\nabla^\perp\psi_h$ will not be continuous across element boundaries any longer and thus we can't use a numerical flux of the form $\hat{f}(\omega^+, \omega^-, \vec{u}_h \cdot \hat{n})$. Instead we probably would have to use a numerical flux of the form $\hat{f}(\omega^+, \omega^-, A^+, A^-, \nabla\psi_h \cdot \hat{\tau})$. We have to define this flux such that conservation of numerical energy (Theorem 8.1) and L_2 stability of numerical enstrophy (Theorem 8.2) still hold.

- The no normal flow boundary conditions which we used don't imply conservation of energy and enstrophy if they are used in combination with in- or outflow boundaries. It would be interesting to extend these no normal flow boundary conditions in such a way that energy is a conserved quantity.

  Another extension to the boundary conditions which we used might be to consider moving boundaries instead of fixed boundaries.

- We saw that using a central scheme has the advantage that numerical enstrophy is conserved in some cases, but the orders of convergence are low. We want of course to have a high order of convergence *and* conservation of numerical enstrophy. We have two suggestions to try to achieve this. First we could try to find a different time discretization scheme for which the central flux behaves better. Another option is to introduce a different kind of numerical flux. At this moment the numerical flux is defined pointwise: $\hat{f}(\omega^+, \omega^-, \vec{u}_h \cdot \hat{n})$ is evaluated at a certain point on the boundary of an element. We suggest to consider fluxes which are defined on a whole face. Thus, instead of approximating $\omega_h\vec{u}_h \cdot \hat{n}$ at one point, this new numerical flux should approximate the integral over a face of $\omega_h\vec{u}_h \cdot \hat{n}$. Perhaps it is then possible to define a numerical flux which has a higher order of convergence than the numerical flux and also conserves numerical enstrophy exactly.

- Our final question is the following: We can increase the accuracy of the numerical solution in two different ways. We can increase the number of elements or we can increase the order of the basisfunctions. Which method is more efficient?

A Reference Element

In this section we will introduce the reference element. The reference element is used in the definition of the basis function for the discontinuous and continuous Galerkin methods. It can also be used to transform the integrals over an element appearing in the weak formulation (124)–(125) to integrals over the reference element. We assume that we have a mesh $\mathcal{T}_h$ which consists of quadrilateral shaped elements. In that case we can define the reference element as

$$\hat{K} = [-1, 1] \times [-1, 1]. \tag{152}$$

We will now define a mapping from the reference element to a real element $K \in \mathcal{T}_h$. First we define the following functions

$$\begin{aligned} N_1(\xi, \eta) &= \frac{1}{4}(1-\xi)(1-\eta) \\ N_2(\xi, \eta) &= \frac{1}{4}(1+\xi)(1-\eta) \\ N_3(\xi, \eta) &= \frac{1}{4}(1+\xi)(1+\eta) \\ N_4(\xi, \eta) &= \frac{1}{4}(1-\xi)(1+\eta). \end{aligned}$$

Now suppose the vertices of the element $K \in \mathcal{T}_h$ are, in anti-clockwise direction, given by $\vec{q}_i = \begin{bmatrix} x_i & y_i \end{bmatrix}^T$ $(i = 1, 2, 3, 4)$. We can then define the mapping $t_K : \hat{K} \mapsto K$ as follows (see Figure 26)

$$t_K(\xi, \eta) = \sum_{i=1}^{4} \vec{q}_i N_i(\xi, \eta). \tag{153}$$

Using this transformation we can write an integral of a function f over an element $K \in \mathcal{T}_h$ as an integral over the reference element

$$\int_K f \, dK = \int_{-1}^{1} \int_{-1}^{1} f \circ t_K |J_K| \, d\xi \, d\eta \tag{154}$$

with $|J_K|$ the determinant of the Jacobian matrix given by

$$\begin{aligned} J_K(\xi, \eta) &= \begin{bmatrix} \partial_\xi x & \partial_\xi y \\ \partial_\eta x & \partial_\eta y \end{bmatrix} = \\ &\frac{1}{4} \begin{bmatrix} (x_2 - x_1)(1-\eta) - (x_4 - x_3)(1+\eta) & (y_2 - y_1)(1-\eta) - (y_4 - y_3)(1+\eta) \\ (x_3 - x_2)(1+\xi) - (x_1 - x_4)(1-\xi) & (y_3 - y_2)(1+\xi) - (y_1 - y_4)(1-\xi) \end{bmatrix}. \end{aligned} \tag{155}$$

The gradient of a function defined on an element can be expressed in terms of the gradient of that function on the reference element as follows

$$\nabla f = \begin{bmatrix} \partial_x f \\ \partial_y f \end{bmatrix} = \begin{bmatrix} \partial_\xi f \partial_x \xi + \partial_\eta f \partial_x \eta \\ \partial_\xi f \partial_y \xi + \partial_\eta f \partial_y \eta \end{bmatrix} = J_K^{-1} \begin{bmatrix} \partial_\xi f \\ \partial_\eta f \end{bmatrix}. \tag{156}$$

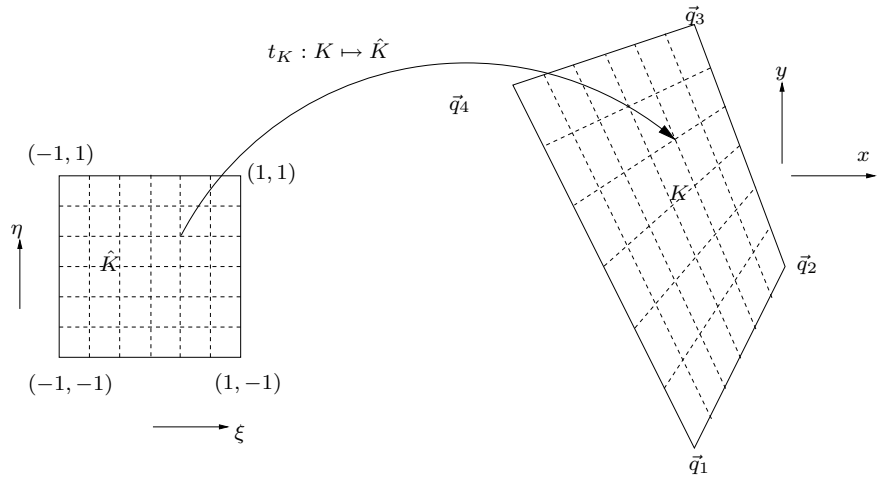


Figure 26: The transformation t_K is used to map the reference element $\hat{K}$ to a real element K.

B Basisfunctions

In this section we will define the basis function for the discontinuous and continuous Galerkin methods introduced in Section 8.1 and Section 8.2. We will make an explicit choice for the basisfunctions given in (83) and (92). First we define the basisfunction for the continuous Galerkin method in Appendix B.2 and then in Appendix B.1 we will define basisfunctions for the continuous Galerkin method.

B.1 k-th Order Basisfunctions for the Discontinuous Galerkin Discretization

Here we explicitly define the space $S^k(K)$ introduced in Section 8.1. We define this space in terms of its basisfunctions, which were introduced in (83). We define k-th order basisfunctions on the reference element as

$$\hat{\phi}_s(\xi, \eta) = \xi^i \eta^j$$

with $s = (i+j+1)(i+j)/2 + j + 1$ for $i \geq 0$, $j \geq 0$ and $i + j \leq k$. The number of basisfunctions per element, N_k, is given by $N_k = k(k+1)/2 + k + 1$. For $k = 0$ we have only one basis function:

$$\hat{\phi}_1(\xi, \eta) = 1.$$

For $k = 1$ a slope is added and we have the following two additional basisfunction

$$\hat{\phi}_2(\xi, \eta) = \xi$$
$$\hat{\phi}_3(\xi, \eta) = \eta$$

And for $k = 2$ we have the following basisfunctions in addition to the previous three

$$\hat{\phi}_4(\xi, \eta) = \xi^2$$
$$\hat{\phi}_5(\xi, \eta) = \xi\eta$$
$$\hat{\phi}_6(\xi, \eta) = \eta^2.$$

The basisfunction given in (83) are now defined by

$$\hat{\phi}_i^{(K)} = \hat{\phi}_i \circ t_K^{-1} - C_i$$

with

$$C_i = \begin{cases} \int_K \hat{\phi}_i \, dK/|K| & \text{if } i = 2, 3, \ldots, N_k \\ 0 & \text{if } i = 1 \end{cases} \tag{157}$$

with $|K| = \int_K dK$ the area of an element. Note that with this definition the properties (84)–(86) are satisfied.

B.2 k-th Order Basisfunctions for the Continuous Galerkin Discretization

In this section we define the basisfunctions introduced in (92) used for the continuous Galerkin discretization. Before we define the *global* basisfunctions on the whole domain, we first define the *local* basisfunctions on the reference element. We define the k-th order basisfunctions $\hat{\varphi}_{ij} : \hat{K} \mapsto \mathbb{R}$ as

$$\varphi_{ij}(\xi, \eta) = p_i(\xi) p_j(\eta) \tag{158}$$

for $i = 0, 1, \ldots, k$, $j = 0, 1, \ldots, k$. Here p_i and p_j are polynomials defined by

$$p_i(\gamma) = \frac{\prod\limits_{j=0,\, j\neq i}^{k} (\gamma - \bar{\gamma}_j)}{\prod\limits_{j=0,\, j\neq i}^{k} (\bar{\gamma}_i - \bar{\gamma}_j)} \tag{159}$$

with $\bar{\gamma}_j$ the roots of the polynomials p_i. More precisely, the roots of p_i are given by $\bar{\gamma}_j$ for $j \in \{0, 1, 2, \ldots, k\} \setminus \{i\}$. For the case $k = 1$ we choose these roots as $\bar{\gamma}_0 = -1$ and $\bar{\gamma}_1 = 1$ and the local basisfunctions are then given by:

$$\hat{\varphi}_{00}(\xi, \eta) = \frac{\xi - 1}{2} \frac{\eta - 1}{2}$$
$$\hat{\varphi}_{01}(\xi, \eta) = \frac{\xi - 1}{2} \frac{\eta + 1}{2}$$
$$\hat{\varphi}_{10}(\xi, \eta) = \frac{\xi + 1}{2} \frac{\eta - 1}{2}$$
$$\hat{\varphi}_{11}(\xi, \eta) = \frac{\xi + 1}{2} \frac{\eta + 1}{2}$$

Note that each basisfunction equals one at exactly one vertex of the element and vanishes at all the other vertices of the element. In the more general case of $k \geq 1$ we demand that the roots $\bar{\gamma}_i$ have the following properties:

- All roots should be distinct: $\bar{\gamma}_j \neq \bar{\gamma}_i$ for $i \neq j$.
- All roots should be in the interval $[-1, 1]$: $\bar{\gamma}_i \in [-1, 1]$ for each $i = 0, 1, 2, \ldots, k$.
- The roots should be distributed symmetrically around $\gamma = 0$: if $\gamma = \bar{\gamma}_i$ for some $i \in \{0, 1, 2, \ldots, k\}$ then a $j \in \{0, 1, 2, \ldots, k\}$ must exist with $\bar{\gamma}_j = -\bar{\gamma}_i$.
- $\gamma = 1$ and $\gamma = -1$ must be roots.

For each of the basisfunction $\hat{\varphi}_{ij}$ we have that $\hat{\varphi}_{ij}(\bar{\gamma}_i, \bar{\gamma}_j) = 1$ and $\hat{\varphi}_{ij}(\bar{\gamma}_v, \bar{\gamma}_w) = 0$ if $v \neq i$ or $w \neq j$. We have three kinds of basisfunctions

- basisfunctions $\hat{\varphi}_{vw}$ for which $(\bar{\gamma}_v, \bar{\gamma}_w)$ is on a vertex of the element. We say that these basisfunctions correspond to a vertex.
- basisfunctions $\hat{\varphi}_{vw}$ for which $(\bar{\gamma}_v, \bar{\gamma}_w)$ is on the boundary of the element, but not on a vertex of the element. We say that these basisfunctions correspond to a face.
- basisfunctions $\hat{\varphi}_{vw}$ for which $(\bar{\gamma}_v, \bar{\gamma}_w)$ lies in the interior of an element. We say that these basisfunctions correspond to an element.

Note that each element has four basisfunctions of the first two kinds because we demanded that $\gamma = -1$ and $\gamma = 1$ must be roots.

For the construction of the *global* basisfunctions, introduced in (92), we use a mapping

$$\mathcal{M}^{(K)} : \{0, 1, 2, \ldots, k\} \times \{0, 1, 2, \ldots, k\} \mapsto \{1, 2, \ldots, L_k\}$$

with L_k the number of *global* basisfunctions. This mapping $\mathcal{M}^{(K)}$ maps for each element $K \in \mathcal{T}_h$ the indices (i, j) of a local basisfunction to the index of the corresponding global basisfunction. This mapping depends of the structure of the mesh and the boundary conditions. Given this mapping we will define the global basisfunctions, restricted to one element, as

$$\varphi_i\big|_K = \sum_{v=0}^{k} \sum_{w=0}^{k} \delta_{i\mathcal{M}^{(K)}(v,w)} \hat{\varphi}_{vw} \circ t_K^{-1}$$

with δ_{vw} the Kronecker delta, defined by

$$\delta_{vw} = \begin{cases} 1 & \text{if } v = w \\ 0 & \text{if } w \neq v \end{cases}.$$

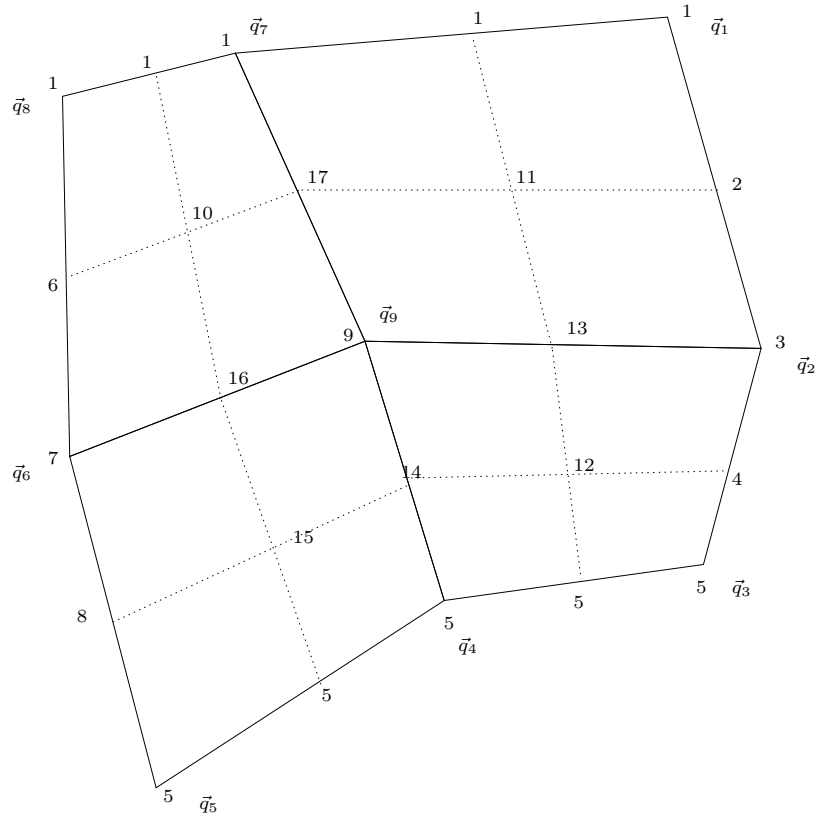


Figure 27: An example of the mappings $\mathcal{M}^{(K)}$.

In Figure 27 it is illustrated how these mappings $\mathcal{M}^{(K)}$ $(K \in \mathcal{T}_h)$ could be defined for 2-nd order basisfunctions and for a very small mesh consisting of only four elements. Each intersection of two lines corresponds to a coordinate $(\xi, \eta) = (\gamma_v, \gamma_w)$ and is labeled with the value of $\mathcal{M}^{(K)}(v, w)$. Note that the mapping is defined in such a way that the resulting global basisfunctions are continuous. Particularly interesting are the upper and lower boundary of the mesh. We assumed no-normal-flow boundaries here and therefore the testfunctions should be constant on these boundaries. This is achieved by mapping all local basisfunction on this boundary to the same global basisfunction index.

C Elemental Integrals

In this section we will work out the details of the integrals appearing in the weak formulation (87), (96) and (97). We will use the mapping t_K described in Appendix A to reduce the integrals over an element $K \in \mathcal{T}_h$ to integrals over the

reference element $\hat{K}$. Integrals over the reference element can be approximated by the Gauss rules described in Appendix D.

First we consider the integrals over the whole domain appearing in (96) and (97). To compute these integrals we note that $\varphi_i\big|_K$ $(i = 1, 2, \ldots, L_k)$ can be written as

$$\varphi_i\big|_K = \begin{cases} \sum_{(p,q)\in\mathcal{M}^{-1}} \hat{\varphi}_{pq} \circ t_K^{-1} & \text{if} \quad \mathcal{M}^{-1} \neq \emptyset \\ 0 & \text{if} \quad \mathcal{M}^{-1} = \emptyset. \end{cases}$$

Here $\mathcal{M}^{-1}$ is used as a notation for

$$\mathcal{M}^{-1} = \left(\mathcal{M}^{(K)}\right)^{-1}(i) = \left\{(p,q)\middle|\mathcal{M}^{(K)}(p,q) = i\right\}.$$

Thus it is sufficient to calculate the integrals appearing in (96) and (97) only over one element instead of over the whole domain and only for local basisfunctions $\varphi_i\big|_K = \hat{\varphi}_{pq} \circ t_K^{-1}$:

$$\int_K A\nabla\left(\hat{\varphi}_{pq} \circ t_K^{-1}\right) \cdot \nabla\left(\hat{\varphi}_{vw} \circ t_K^{-1}\right) dK = \int_{-1}^{1}\int_{-1}^{1} A \circ t_K \begin{bmatrix} \partial_\xi \hat{\varphi}_{pq} \\ \partial_\eta \hat{\varphi}_{pq} \end{bmatrix}^T \left(J_K^{-1}\right)^T J_K^{-1} \begin{bmatrix} \partial_\xi \hat{\varphi}_{vw} \\ \partial_\eta \hat{\varphi}_{vw} \end{bmatrix} |J_K|\, d\xi\, d\eta$$

$$\int_K B\left(\hat{\varphi}_{pq} \circ t_K^{-1}\right)\left(\hat{\varphi}_{vw} \circ t_K^{-1}\right) dK = \int_{-1}^{1}\int_{-1}^{1} (B \circ t_K)\, \hat{\varphi}_{vw}\hat{\varphi}_{pq}|J_K|\, d\xi\, d\eta$$

$$\int_K (C - \omega_h)\left(\hat{\varphi}_{pq} \circ t_K^{-1}\right) dK = \int_{-1}^{1}\int_{-1}^{1} \left((C \circ t_K) - \sum_{i=1}^{N_k} \omega_i^{(K)}\left(\hat{\phi}_i - C_i\right)\right) \hat{\varphi}_{pq}|J_K|\, d\xi\, d\eta$$

for $v, w, p, q \in \{0, 1, 2, \ldots, k\}$. The coefficients C_i $(i = 2, 3, \ldots, k)$ introduced in (157) can be written as

$$C_i = \int_K \left(\hat{\phi}_i \circ t_K^{-1}\right)\left(\hat{\phi}_1 \circ t_K^{-1}\right) dK/|K| = \int_{-1}^{1}\int_{-1}^{1} \hat{\phi}_i|J_K|\, d\xi\, d\eta/|K|$$

with $|K| = \int_{-1}^{1}\int_{-1}^{1} |J_K|\, d\xi\, d\eta$. For the integrals in (87) over the interior of an

element we can write

$$\int_K \phi_i^{(K)} \phi_j^{(K)} \, dK = \int_{-1}^{1} \int_{-1}^{1} \left(\hat{\phi}_i - C_i \right) \left(\hat{\phi}_j - C_j \right) |J_K| \, d\xi \, d\eta$$
$$\int_K \phi_1^{(K)} \phi_1^{(K)} \, dK = \int_{-1}^{1} \int_{-1}^{1} |J_K| \, d\xi \, d\eta$$
$$\int_K \phi_j^{(K)} \vec{u}_h \cdot \nabla \phi_1^{(K)} \, dK = 0$$

and

$$\int_K \phi_j^{(K)} \vec{u}_h \cdot \nabla \phi_i^{(K)} \, dK = \int_K \phi_j^{(K)} A \nabla^{\perp} \psi_h \cdot \nabla \phi_i^{(K)} \, dK$$
$$= \int_{-1}^{1} \int_{-1}^{1} \left(\hat{\phi}_j - C_j \right) (A \circ t_K) \left(\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} J_K^{-1} \begin{bmatrix} \partial_\xi \psi_h \\ \partial_\eta \psi_h \end{bmatrix} \right)^T$$
$$J_K^{-1} \begin{bmatrix} \partial_\xi \hat{\phi}_i \\ \partial_\eta \hat{\phi}_i \end{bmatrix} |J_K| \, d\xi \, d\eta = \int_{-1}^{1} \int_{-1}^{1} \left(\hat{\phi}_j - C_j \right) (A \circ t_K)$$
$$\begin{bmatrix} \partial_\xi \psi_h \\ \partial_\eta \psi_h \end{bmatrix}^T \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} \partial_\xi \hat{\phi}_i \\ \partial_\eta \hat{\phi}_i \end{bmatrix} \, d\xi \, d\eta$$

for $i, j = 2, 3, \ldots, N_k$. We can use the Gauss rules presented in Appendix D to evaluate these integrals numerically.

D Gauss Integration Rules

We use Gauss integration rules to approximate the integrals over elements and faces given in Appendix C. For integrals over faces we can use the following N-point Gauss rule

$$\int_{-1}^{1} f(\gamma) d\gamma \approx \sum_{i=1}^{N} w_i f(\gamma_i) \tag{160}$$

with w_i and γ_i the Gauss weights and points. These Gauss weights and points are given in Table 10 for different values of N. The approximation given in (160) is exact for $(2N-1)$-th or lower order polynomials.

For integrals over the reference element we use the following Gauss rule

$$\int_{-1}^{1} \int_{-1}^{1} f(\xi, \eta) d\xi \, d\eta \approx \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j f(\gamma_i, \gamma_j). \tag{161}$$

The values of w_i and γ_i are again given in Table 10. This Gauss rule is exact for functions $f(\xi, \eta)$ which are $(2N-1)$-th or lower order polynomials in ξ and η.

N	Order	i	γ_i	w_i
1	1	1	0	2
2	3	1	−0.57735027	1
		2	0.57735027	1
3	5	1	−0.77459667	0.55555555
		2	0	0.88888889
		3	0.77459667	0.55555555
4	7	1	−0.86113631	0.34785485
		2	−0.33998104	0.65214515
		3	0.33998104	0.65214515
		4	0.86113631	0.34785485
5	9	1	−0.90617985	0.23692689
		2	−0.53846931	0.47862867
		3	0	0.56888889
		4	0.53846931	0.47862867
		5	0.90617985	0.23692689
6	11	1	−0.93246951	0.17132449
		2	−0.66120939	0.36076157
		3	−0.23861918	0.46791393
		4	0.23861918	0.46791393
		5	0.66120939	0.36076157
		6	0.93246951	0.17132449
7	13	1	−0.94910791	0.12948497
		2	−0.74153119	0.27970539
		3	−0.40584515	0.38183005
		4	0	0.41795918
		5	0.40584515	0.38183005
		6	0.74153119	0.27970539
		7	0.94910791	0.12948497
8	15	1	−0.96028986	0.10122854
		2	−0.79666648	0.22238103
		3	−0.52553241	0.31370665
		4	−0.18343464	0.36268378
		5	0.18343464	0.36268378
		6	0.52553241	0.31370665
		7	0.79666648	0.22238103
		8	0.96028986	0.10122854

Table 10: Gauss points and weights

E Stability Analysis

E.1 1D Stability Analysis for 0-th order basis functions

In this section we perform a stability analysis for the 1D equation

$$\partial_t \omega + u \partial_x \omega = 0 \tag{162}$$

with the velocity, u, a constant independent of time and space. We assume that we have an equidistant grid. The elements are thus given by

$$K_j = [x_j, x_{j+1}] \tag{163}$$

with $x_j = x_0 + j\Delta x$. We will perform the stability analysis for 0-th order basis functions and a central flux. In that case the weak formulation is given by

$$\int_{x_j}^{x_{j+1}} \frac{d\bar{\omega}_j}{dt}\, dx + u \frac{\bar{\omega}_{j+1,k} - \bar{\omega}_{j-1,k}}{2} = 0 \tag{164}$$

with $\bar{\omega}_j$ the value of ω_h in the element K_j. This equation becomes

$$\Delta x \frac{d\bar{\omega}_j}{dt} + u \frac{\bar{\omega}_{j+1,k} - \bar{\omega}_{j-1,k}}{2} = 0 \tag{165}$$

Now assume that the solution is of the form

$$\bar{\omega}_j^n = \lambda^n \exp(iPj\Delta x) \tag{166}$$

with n denoting the time step: $t = t_0 + n\Delta t$. Then we can write

$$\frac{d\bar{\omega}_{j,k}^n}{dt} = i \left(\frac{u \sin(Pj\Delta x)}{\Delta x} \right) \bar{\omega}_{j,k}^n \equiv L \bar{\omega}_{j,k}^n \tag{167}$$

Applying the Runge-Kutta scheme given in Section 8.5.1 results in

$$\bar{\omega}_{i,j}^{n+1} = \left(1 + \Delta t L + \frac{1}{2}\Delta t^2 L^2 + \frac{1}{6}\Delta t^3 L^3 \right) \bar{\omega}_{i,j}^n \equiv R \bar{\omega}_{i,j}^n. \tag{168}$$

To ensure stability we must require that $|R|^2 \leq 1$. We have

$$\begin{aligned} 1 &\geq |R|^2 \\ &= \left| 1 + \Delta t L + \frac{1}{2}\Delta t^2 L^2 + \frac{1}{6}\Delta t^3 L^3 \right|^2 \\ &= \left(1 - \frac{1}{2}\mu^2 \right)^2 + \left(\mu - \frac{1}{6}\mu^3 \right)^2 \\ &= 1 + \frac{1}{36}\mu^4 \left(\mu^2 - 3 \right) \end{aligned}$$

with $i\mu = \Delta t L = i\Delta t \frac{u \sin(Pj\Delta x)}{\Delta x}$. For $1 \geq |R|^2$ we must demand that $-\sqrt{3} \leq \mu \leq \sqrt{3}$. Thus to ensure that the scheme is stable the following restriction must be placed on the time step:

$$\Delta t \frac{|u|}{\Delta x} \leq \sqrt{3}. \tag{169}$$

E.2 1D Stability Analysis for 1-th order basis functions

In this section we perform a stability analysis for the 1D equation

$$\partial_t \omega + u \partial_x \omega = 0 \tag{170}$$

with the velocity, u, a constant independent of time and space. We assume that we have an equidistant rectangular grid. The elements are thus given by

$$K_j = [x_j, x_{j+1}] \tag{171}$$

with $x_j = x_0 + j\Delta x$. We will perform the stability analysis for 1-st order basis functions and a central flux. In that case the basis functions on the element K_j are given by

$$\bar{\phi}_j = 1 \tag{172}$$

and

$$\hat{\phi}_j = 2\frac{x - x_j}{\Delta x} - 1. \tag{173}$$

The numerical approximation for the vorticity can then be written as

$$\omega_h\big|_{K_j} = \bar{\omega}_j \bar{\phi}_j + \hat{\omega}_j \hat{\phi}_j. \tag{174}$$

The weak formulation is given by: Find ω_h such that for each test function v_h the following equation holds:

$$\int_{x_j}^{x_{j+1}} v_h \frac{d\omega_h}{dt}\, dx - \int_{x_j}^{x_{j+1}} \omega u \partial_x v_h \, dx + v_h\big|_{x_{j+1}} \omega_h\big|_{x_{j+1}} - v_h\big|_{x_j} \omega_h\big|_{x_j} = 0. \tag{175}$$

Now using a central flux for $\omega_h\big|_{x_{j+1}}$ and $\omega_h\big|_{x_j}$, using $\omega_h\big|_{K_j} = \bar{\omega}_j + \hat{\omega}_j \hat{\phi}_j$ and taking $v_h\big|_{K_j} = \bar{\phi}_j$ and $v_h\big|_{K_j} = \hat{\phi}_j$ respectively we find

$$\begin{aligned} \Delta x \frac{d\bar{\omega}_j}{dt} + \frac{u}{2}\left(\bar{\omega}_{j+1} - \hat{\omega}_{j+1} + 2\hat{\omega}_j - \bar{\omega}_{j-1} - \hat{\omega}_{j+1}\right) &= 0 \\ \frac{\Delta x}{3}\frac{d\hat{\omega}_j}{dt} - 2u\bar{\omega}_j + \frac{u}{2}\left(\bar{\omega}_{j+1} + 2\bar{\omega}_j + \bar{\omega}_{j-1} + \hat{\omega}_{j-1} - \hat{\omega}_{j+1}\right) &= 0. \end{aligned}$$

Now assume that the solution is of the form

$$\begin{bmatrix} \bar{\omega}_j^n \\ \hat{\omega}_j^n \end{bmatrix} = \lambda^n \begin{bmatrix} \exp(iPj\Delta x) \\ \exp(iQj\Delta x) \end{bmatrix} \tag{176}$$

with n denoting the time: $t = t_0 + n\Delta t$. Then we can write

$$\frac{d}{dt}\begin{bmatrix} \bar{\omega}_j \\ \hat{\omega}_j \end{bmatrix} = L \begin{bmatrix} \bar{\omega}_j \\ \hat{\omega}_j \end{bmatrix} \tag{177}$$

with

$$L \equiv \frac{u}{\Delta x}\begin{bmatrix} -i\sin(P\Delta x) & -1+\cos(Q\Delta x) \\ 3(1-\cos(P\Delta x)) & 3i\sin(Q\Delta x) \end{bmatrix}. \tag{178}$$

Applying the Runge-Kutta scheme given in Section 8.5.1 results in

$$\begin{bmatrix} \bar{\omega}_j^{n+1} \\ \hat{\omega}_j^{n+1} \end{bmatrix} = \left(1 + \Delta t L + \frac{1}{2}\Delta t^2 L^2 + \frac{1}{6}\Delta t^3 L^3\right)\begin{bmatrix} \bar{\omega}_j^{n} \\ \hat{\omega}_j^{n} \end{bmatrix} \equiv R \begin{bmatrix} \bar{\omega}_j^{n} \\ \hat{\omega}_j^{n} \end{bmatrix}. \tag{179}$$

To ensure stability we must require that for each eigenvalue, μ, of R we have $\mu \leq 1$. The eigenvalues of L are given by

$$\begin{aligned} \nu_{1,2} &= i\frac{u}{\Delta x}\left(-\frac{1}{2}\sin(P\Delta x) + \frac{3}{2}\sin(Q\Delta x)\right) \\ &\pm i\frac{u}{\Delta x}\frac{1}{2}\underbrace{\sqrt{\left(\sin(P\Delta x) + 3\sin(Q\Delta x)\right)^2 + 12(\cos(P\Delta x)-1)(\cos(Q\Delta x)-1)}}_{\geq 0}. \end{aligned} \tag{180}$$

The the eigenvalues of R are thus given by

$$\mu_{1,2} = 1 + \Delta t\nu_{1,2} + \frac{1}{2}\Delta t^2\nu_{1,2}^2 + \frac{1}{6}\Delta t^3\nu_{1,2}^3. \tag{181}$$

The absolute value of these eigenvalues is now given by

$$\begin{aligned} |\mu_{1,2}|^2 &= \left(1 - \frac{1}{2}\Delta t^2|\nu_{1,2}|^2\right)^2 + |\nu_{1,2}|^2\left(\Delta t - \frac{1}{6}\Delta t^3|\nu_{1,2}|^2\right)^2 \\ &= 1 + \frac{\Delta t^4|\nu_{1,2}|^4}{36}\left(\Delta t^2|\nu_{1,2}|^2 - 3\right). \end{aligned}$$

To ensure that $|\mu_{1,2}| \leq 1$ we must have

$$\Delta t|\nu_{1,2}| \leq \sqrt{3}. \tag{182}$$

From (180) it follows that $|\nu_{1,2}| \leq 6|u|/\Delta x$ thus to ensure the stability of the numerical scheme we can choose

$$\Delta t \leq \frac{\sqrt{3}}{6}\frac{\Delta x}{|u|}. \tag{183}$$

E.3 2D Stability Analysis for 0-th order basis functions

In this section we perform a stability analysis for the equation

$$\partial_t\omega + \vec{u}\cdot\nabla\omega = 0 \tag{184}$$

with the velocity $\vec{u} = [u, v]$ a constant vector, independent of time and space. We assume that we have a rectangular grid. The elements are thus given by

$$K_{j,k} = [x_j, x_{j+1}] \times [y_k, y_{k+1}] \tag{185}$$

with $x_j = x_0 + j\Delta x$ and $y_k = y_0 + k\Delta y$. We will perform the stability analysis for 0-th order basis functions and a central flux. In that case the weak formulation is given by

$$\int_{x_j}^{x_{j+1}} \int_{y_k}^{y_{k+1}} \frac{d\bar{\omega}_{j,k}}{dt}\, dy\, dx + u\Delta y \frac{\bar{\omega}_{j+1,k} - \bar{\omega}_{j-1,k}}{2} + v\Delta x \frac{\bar{\omega}_{j,k+1} - \bar{\omega}_{j,k-1}}{2} = 0 \quad (186)$$

with $\bar{\omega}_{j,k}$ the value of ω_h in the element $K_{j,k}$. This equation becomes

$$\Delta x \Delta y \frac{d\bar{\omega}_{j,k}}{dt} + u\Delta y \frac{\bar{\omega}_{j+1,k} - \bar{\omega}_{j-1,k}}{2} + v\Delta x \frac{\bar{\omega}_{j,k+1} - \bar{\omega}_{j,k-1}}{2} = 0. \quad (187)$$

Now assume that the solution is of the form

$$\bar{\omega}_{j,k}^n = \lambda^n \exp(i(Pj\Delta x + Qk\Delta y)) \quad (188)$$

with n denoting the time: $t = t_0 + n\Delta t$. Then we can write

$$\frac{d\bar{\omega}_{j,k}^n}{dt} = i\left(\frac{u\sin(Pj\Delta x)}{\Delta x} + \frac{v\sin(Qk\Delta y)}{\Delta y}\right)\bar{\omega}_{j,k}^n \equiv L\bar{\omega}_{j,k}^n \quad (189)$$

Applying the Runge-Kutta scheme given in Section 8.5.1 results in

$$\bar{\omega}_{i,j}^{n+1} = \left(1 + \Delta t L + \frac{1}{2}\Delta t^2 L^2 + \frac{1}{6}\Delta t^3 L^3\right)\bar{\omega}_{i,j}^n \equiv R\bar{\omega}_{i,j}^n. \quad (190)$$

To ensure stability we must require that $|R|^2 \leq 1$. We have

$$\begin{aligned} 1 &\geq |R|^2 \\ &= \left|1 + \Delta t L + \frac{1}{2}\Delta t^2 L^2 + \frac{1}{6}\Delta t^3 L^3\right|^2 \\ &= \left(1 - \frac{1}{2}\mu^2\right)^2 + \left(\mu - \frac{1}{6}\mu^3\right)^2 \\ &= 1 + \frac{1}{36}\mu^4\left(\mu^2 - 3\right) \end{aligned}$$

with $i\mu = \Delta t L = i\Delta t\left(\frac{v\sin(Qk\Delta y)}{\Delta y} + \frac{u\sin(Pj\Delta x)}{\Delta x}\right)$. Thus for values of μ with $-\sqrt{3} \leq \mu \leq \sqrt{3}$ we have $|R|^2 \leq 1$. To ensure that the scheme is stable we must then impose the following restriction on the time step size:

$$\mu \leq \sqrt{3} \qquad \text{and} \qquad \mu \geq -\sqrt{3}. \quad (191)$$

Thus the following should hold

$$-\sqrt{3} \leq \Delta t\left(\frac{v\sin(Qk\Delta y)}{\Delta y} + \frac{u\sin(P\Delta x)}{\Delta x}\right) \leq \sqrt{3}. \quad (192)$$

or

$$\Delta t\left(\frac{|v|}{\Delta y} + \frac{|u|}{\Delta x}\right) \leq \sqrt{3}. \quad (193)$$

References

[1] J.G. Charney. On the scale of atmospheric motions. *Geophysical Publications*, 17(2):251–265, 1948.

[2] J.G. Charney, R. Fjortoft, and J. von Neumann. Numerical integration of the barotropic vorticity equation. *Tellus*, 2(1):237–254, 1950.

[3] G.H¿ Golub and C.F.van Loan. *Matrix Computations*. Johns Hopkins, 1996.

[4] D. Holm, J.E. Marsden, and T. Ratiu. Non-lineair stability of the kelvin-stuart cat's eyes flow. *American Mathematical Society*, 23:171–186, 1986.

[5] J.T.Stuart. On finite amplitude ascillations in laminar mixing layers. *Jounral of FLuid Mechanics*, 29:417–440, 1967.

[6] Lord Kelvin. On a disturbing infinity in lord rayleigh's solution for waves in a plane vortex stratum. *Nature*, 23:45–46, 1880.

[7] L. D. Landau and E. M. Lifschitz. *Fluid Mechanics*. Pergamon Press, 1959.

[8] P.H. LeBlond and L.A. Mysak. *Waves in the Ocean*. Elsevier Scientific Publishing Company, 1978.

[9] J. Liu and C. Shu. A high order discontinuous galerkin method for 2d incompressible flows. *Journal of Computational Physics*, 160:577–596, 2000.

[10] Brigitte Lucquin and Olivier Pironneau. *Introduction to Scientific Computing*. Wiley, 1998.

[11] J. Pedlosky. *Geophysical Fluid Dynamics*. Springer-Verlag, 1979.

[12] C. Shu. Total-variation-diminishing time discretizations. *SIAM J. Sci. Stat. Comput.*, 9(6):1073–1084, November 1988.

[13] C. Shu and S. Osher. Efficient implementation of essentially non-oscillatory shock-capturing schemes. *Journal of Computational Physics*, 77:439–471, 1988.