NTNU
Norwegian University of
Science and Technology

Stable Finite Volume Methods for the Isentropic Euler Equations

Ivar Andreas Ø Kaland

Master of Science in Physics and Mathematics
Submission date: February 2016
Supervisor: Ulrik Skre Fjordholm, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

Preface

I am truly grateful to my supervisor Ulrik Skre Fjordholm for all his support and help that he's given me. As was the case with my project before, his guidance was essential to this thesis. I would also like to thank my friends in Trondheim for making my time as enjoyable as it's been, as well as my family and girlfriend for their support.

Contents

1 Introduction

The Euler equations of gas dynamics describe how the mass, moment and pressure of a moving gas are related. They are named after Leonard Euler and correspond to the case where the Navier-Stokes equations has zero viscosity and heat conduction[9]. In one dimension, they are given by

$$\begin{aligned} \rho_t + (\rho w)_x &= 0, \\ (\rho w)_t + (\rho w^2 + p)_x &= 0, \\ E_t + (w(E+p))_x &= 0. \end{aligned} \tag{1.1}$$

Here ρ denotes the density of a non-viscous fluid or a gas, w is its velocity in the x-direction and p is the pressure. In cases where the viscous effects are negligible, such as in gas dynamics, these equations are important. The Euler equations are hyperbolic conservation equations. This thesis will deal with the isentropic case with ideal polytropic gasses. In this case, the entropy in the system remains constant. The system (1.1) is then reduced to the isentropic Euler equations

$$\begin{aligned} \rho_t + (\rho v)_x &= 0, \\ (\rho v)_t + (\rho v^2 + \kappa\rho^\gamma)_x &= 0. \end{aligned}$$

The isentropic Euler equations have applications within the field of acoustics, since acoustic waves have varying pressure and density, but the entropy remains constant[11]. Some examples of other hyperbolic conservation laws include the nonlinear shallow water equations and Buckley-Leverett equation, or the linear advection equation. Nonlinear systems of conservation equations tend to be much more problematic to solve analytically than linear equations. It's also hard to find stability results of numerical methods, although Harten, Osher, Lax and Tadmor and several others have done a lot of work in this field[10, 11, 14, 16, 18]. Among the central concepts developed was entropy conservation and entropy stability. If schemes satisfy certain entropy conditions, then we can obtain entropy stability and convergence towards the physically relevant solution. For the isentropic Euler equations, the relevant mathematical entropy is the physical energy of the solution.

The main goal of the thesis is to develop energy conserving and energy stable schemes for the isentropic Euler equations. To handle the discontinuities in the solution and dissipation of energy, we develop a diffusion operator. Together, our energy conservative scheme and this diffusion operator will make up an energy stable scheme for the isentropic Euler equations. The energy conservative method will be second-order accurate, while the energy stable one is only first-order accurate.

The rest of Section 1 will contain an introduction into systems of hyperbolic equations and the Euler equations. The equations will be derived, as well as the isentropic case. Section 2 introduces the theory of finite volume schemes and introduce a few popular schemes. We develop our scheme in Section 3. In Section 3.4, the energy conservative scheme is developed as the theory behind it is discussed. We then perform numerical experiments to test our method. After this, we develop energy stable schemes in Section 3.7 by adding numerical viscosity to the energy conservative scheme, again followed by numerical experiments. The scheme will be compared with some more common methods, such as the Lax-Friedrichs and Rusanov scheme.

Throughout the discussion, we will use the notation u_x to denote the partial derivative of u with respect to x, that is $u_x = \frac{\partial u}{\partial x}$. The speed components of the gas will be denoted by variables w and ω, respectively for the x component and the y component. Furthermore we will later make use of the following identities and notation.

$$\begin{aligned} [\![jk]\!]_{i+1/2} &= [\![j]\!]_{i+1/2}\,\overline{k}_{i+1/2} + [\![k]\!]_{i+1/2}\,\overline{j}_{i+1/2} \\ [\![j^2]\!]_{i+1/2} &= 2\cdot[\![j]\!]_{i+1/2}\,\overline{j}_{i+1/2}, \end{aligned} \tag{1.2}$$

where

$$\begin{aligned} [\![j]\!]_{i+1/2} &= j_{i+1} - j_i \\ \overline{j}_{i+1/2} &= \frac{j_i + j_{i+1}}{2}. \end{aligned} \tag{1.3}$$

1.1 Hyperbolic systems of conservation laws

Hyperbolic systems of conservation laws are time-dependent systems of partial differential equations. In one space dimension they take the form

$$u_t(x,t) + f_x(u(x,t)) = 0 \tag{1.4}$$

where $u : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}^m$ is the m-dimensional vector of some quantity that is conserved, such as energy, mass, density, momentum and so on. The variables of u are commonly called the *state variables*. The function $f(u) : \mathbb{R}^m \to \mathbb{R}^m$ is usually called a *flux function*, and describes the rate of flow, or flux, of all the state variables at (x,t). For the Euler equations, our conserved quantities are mass, momentum and energy. In short, if u_n denotes the nth conserved state variable we can write

$$\begin{aligned} \tfrac{\partial}{\partial t}u_1(x,t) + \tfrac{\partial}{\partial x}f_1(u(x,t)) &= 0 \\ &\cdots \\ &\cdots \\ &\cdots \\ \tfrac{\partial}{\partial t}u_m(x,t) + \tfrac{\partial}{\partial x}f_m(u(x,t)) &= 0 \end{aligned} \tag{1.5}$$

with $u = (u_1, ..., u_m) \in \mathbb{R}^m$ as the conserved quantities and $f = (f_1, ..., f_m) : \mathbb{R}^m \to \mathbb{R}^m$ being the flux functions. The nth state variable can be written on the form u_1, while the rate of flow for the nth state variable can be written as $f_n(u(x,t))$. The equations in (1.5) also need some initial data to be solved.

The system (1.4) is hyperbolic, which means that the Jacobian matrix of f, denoted as $J = f'(u)$, has real eigenvalues for each value of u and is diagonalizable, that is, there exists some invertible matrix M such that the matrix $M^{-1}JM$ is diagonal.

In two space dimensions, the system of equations is instead given by

$$u_t(x,y,t) + f(u(x,y,t))_x + g(u(x,y,t))_y = 0 \tag{1.6}$$

with $u : \mathbb{R}^2 \times \mathbb{R}_+ \to \mathbb{R}^m$ and $f, g : \mathbb{R}^m \to \mathbb{R}^m$. Equations (1.4) and (1.6) are written in their differential form.

Similarly, we can write the conservation equations (1.4) in an integral form. The total value of a conserved variable over the interval $[x_1, x_2]$ at time t is given by

$$u_{total} = \int_{x_1}^{x_2} u(x,t)dx.$$

With the assumption that none of the quantity u will be created or destroyed in the domain, the only change in u at some point will be caused by the flow of u in and out of the border. The rate of flow is measured by the flux f, so the change of u over time will be given by

$$\frac{d}{dt}\int_{x_1}^{x_2} u(x,t)\,dx + f(x_2,t) - f(x_1,t) = 0.$$

Thus, the change of the mass in $[x_1, x_2]$ is given by

$$\int_{x_1}^{x_2} u_t\,dx + \int_{x_1}^{x_2} f(u)_x\,dx = 0,$$

which is the integral form of a conservation equation. x_1 and x_2 can be any values on the valid domain, so this relation must be valid everywhere.

The results presented above is similar for more space dimensions and can readily be expanded to account for n spatial dimensions. Writing these equations in their differential form, we assume that the partial derivatives are defined everywhere. However, it is a fundamental feature of nonlinear conservation laws that discontinuities (shocks) easily can develop, even from smooth initial data. If we write a problem in its weak form, we allow for discontinuities in our solution.

1.2 Weak form

Solutions to nonlinear conservation tend to develop shocks, i.e. be discontinuous at some point. The differential forms of a conservation equation will not make sense in if we have discontinuities in our solution, since the values of the partial derivatives are undefined there. In other words, u is not a classical solution. This problem is solved by writing the equations in the weak form, and considering weak solutions to the problem. Let $\varphi(x,t)$ be a smooth test function in $C_c^\infty(\mathbb{R}\times\mathbb{R}_+)$ and let $u : \mathbb{R}\times\mathbb{R}_+ \longrightarrow \mathbb{R}$ be a smooth solution of the hyperbolic conservation law. We proceed to multiply our equation (1.4) with φ, and integrate over space and time

$$\begin{aligned} 0 &= \int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}_+} (u_t + f(u)_x)\varphi(x,t)\,dt\,dx \\ &= \int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}_+} u\varphi_t + f(u)\varphi_x\,dt\,dx + \int\limits_{\mathbb{R}} u(x,0)\varphi(x,0)\,dx - 0 \end{aligned}$$

Rearranging this gives us

$$\int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}_+} (\varphi_t u + \varphi_x f(u))\,dx\,dt + \int\limits_{\mathbb{R}} (u_0(x)\varphi(x,0)\,dx = 0 \quad \forall\varphi\in C_c^\infty(\mathbb{R}\times\mathbb{R}_+) \tag{1.7}$$

We say that (1.7) is the weak form of (1.4), and if this holds, then $u(x,t)$ is a weak solution. It's important to note that this does not require u to be differentiable, even though it might be. For hyperbolic systems, the weak solution based on distributions will not guarantee uniqueness. To get the solution we want, it is necessary to place some additional selection criterion on it.

1.3 Parabolic Regularization

The physically relevant solutions of (1.4) are the ones that we find as $u = \lim_{\epsilon\to 0} u^{\epsilon}$, with

$$u_t^{\varepsilon} + f(u^{\varepsilon})_x = \varepsilon u_{xx}^{\varepsilon} \quad (\varepsilon > 0). \tag{1.8}$$

The right side term can also be expressed differently as

$$u_t^{\varepsilon} + f(u^{\varepsilon})_x = \varepsilon (P u_x^{\varepsilon})_x, \tag{1.9}$$

where $P(u, u_x)$ is the viscosity matrix. Such solutions are called *vanishing viscosity* solutions. For each ϵ, there exists exactly one solution to (1.8), and it is always smooth[2]. Viscosity and entropy are closely related. Note that by a sign difference in definition, mathematical entropy is non-increasing rather than non-decreasing in time. For conservation laws, we should see the entropy being conserved in smooth regions of the solution. In areas with shocks, the physical entropy would increase across the discontinuity, and so we expect the mathematical entropy to be dissipated (decrease) around shocks.

We begin with looking at the solutions where the entropy is conserved. We define a convex function $\eta(u) : \mathbb{R}^n \to \mathbb{R}$ to be a function that measures the entropy present in u. When the solution is smooth, the entropy must be a conserved variable in such areas. By multiplying (1.4) by $\eta'(u)$

$$\begin{aligned} 0 &= \eta'(u)u_t + \eta'(u)f(u)_x \\ &= \eta(u)_t + \eta'(u)f(u)_x \\ &= \eta(u)_t + \eta'(u)f'(u)u_x \end{aligned}$$

We are interested in finding conditions such that the entropy η itself is conserved, that is

$$\eta(u)_t + q(u)_x = 0. \tag{1.10}$$

As we see, this will be satisfied only if such a q exists that

$$q'(u) = \eta'(u)f'(u). \tag{1.11}$$

Around any shocks, the mathematical entropy must dissipate and decrease; the equality does not hold. In such parts of the domain, the entropy equality is invalid. By multiplying (1.8) by $\eta'(u^\varepsilon)$

$$\begin{aligned}
0 &= \eta'(u^\varepsilon)u^\varepsilon_t + \eta'(u^\varepsilon)f(u^\varepsilon)_x - \varepsilon\eta'(u^\varepsilon)u^\varepsilon_{xx} \\
&= \eta'(u^\varepsilon)u^\varepsilon_t + \eta'(u^\varepsilon)f(u^\varepsilon)_x - \varepsilon\eta(u^\varepsilon)_{xx} + \varepsilon((\eta'(u^\varepsilon)u^\varepsilon_x)_x - \eta'(u^\varepsilon)u^\varepsilon_{xx}) \\
&= \eta'(u^\varepsilon)u^\varepsilon_t + \eta'(u^\varepsilon)f(u^\varepsilon)_x - \varepsilon\eta(u^\varepsilon)_{xx} + \varepsilon\eta''(u^\varepsilon_x)(u^\varepsilon_x)^2 \\
&\geq \eta'(u^\varepsilon)u^\varepsilon_t + \eta'(u^\varepsilon)f(u^\varepsilon)_x - \varepsilon\eta(u^\varepsilon)_{xx}
\end{aligned}$$

Rearranging this gives us that

$$\eta(u^\varepsilon)_t + q(u^\varepsilon)_x \leq \varepsilon\eta(u^\varepsilon)_{xx}$$

As $\varepsilon \to 0$, we end up with the condition that

$$\eta(u)_t + q(u)_x \leq 0 \tag{1.12}$$

This does not make sense in the differential form whenever we have discontinuous solutions u. It has to be interpreted in the weak form

$$\int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}_+} \eta(u)\varphi_t + q(u)\varphi_x \, dxdt + \int\limits_{\mathbb{R}} \eta(u_0(x))\varphi(x,0)dx \geq 0, \ \forall\varphi \in C_c^\infty\left(\mathbb{R}\times\mathbb{R}_+\right), \tag{1.13}$$

with $\varphi \geq 0$. Equation (1.13) is the entropy inequality in its weak form. In contrast to the entropy equality, which only works well for smooth solutions and assumes that all the entropy must be conserved, the entropy inequality allows for entropy to dissipate around shocks, which is what we want. This can correspond to energy changing into forms which are not accounted for in our model, such as heat.

1.4 Euler Equation

1.4.1 Introduction

The Euler equations model adiabatic and inviscid flow. As mentioned before, the equations themselves can be seen as Navier-Stokes equations where you operate with the assumption that there is no viscosity or heat conduction. Fey, Courant, Friedrichs and many others have written about the Euler equations[5, 6, 4]. The variables that are found in the Euler equations are the density of the gas ρ, its velocity w, its total energy E, as well as its gas pressure p. Our first conserved variable is the density ρ. The density is transported along the velocity field, so we have the flux

$$f(x,t) = \rho(x,t)w(x,t).$$

Thus, the differential form can be written as

$$\rho_t + (\rho w)_x = 0. \tag{1.14}$$

Equation (1.14) is otherwise known as the continuity equation, and the density is a preserved variable of the Euler Equations.

Next is the momentum. We proceed in a similar manner. The momentum is given by ρw. Since the momentum is also transported by the velocity field, this contribution to the flux function will be

$$f_{partial}(x,t) = (\rho w)(x,t)w(x,t) = \rho w^2$$

In accordance with Newton's laws, there are also other forces that are working on the fluid. Assuming no outside forces, then the sole force will be proportional to the pressure gradient ∇p. In one dimension, this simplifies to p_x. The full flux function for the momentum is therefore

$$f(x,t) = (\rho w^2)(x,t) + p(x,t)$$

Thus, the momentum equation in its differential form is

$$(\rho w)_t + (\rho w^2 + p)_x = 0, \tag{1.15}$$

also known as the momentum equation.

The last conserved variable of the Euler equations is the total energy E of the system. As with the momentum, its flux function will have contributions from the gas being transported by the velocity field, as well as being affected by the pressure. It has flux function

$$f(x,t) = w(E + p(x,t)).$$

Therefore, the equation of conserved energy in its differential form is

$$E_t + [w(E+p)]_x = 0. \tag{1.16}$$

In summary, the system of Euler equations models the conservation of its state variables mass, momentum and energy

$$\begin{bmatrix} \rho \\ \rho w \\ E \end{bmatrix}_t + \begin{bmatrix} \rho w \\ \rho w^2 + p \\ w(E+p) \end{bmatrix}_x = 0.$$

1.4.2 Thermodynamic relations

To solve the Euler equations, we wish to find a way to express the pressure p in terms of our state variables. To do this, we look at some of the thermodynamic relations that will apply for us. In this thesis, we will look at polytropic, ideal gases. A number of the relations that applies for such gases will be used in later sections.

An ideal gas will obey the ideal gas law

$$p = R\rho T, \tag{1.17}$$

where R is the specific gas constant. A good approximation for the internal energy of an ideal polytropic gas is

$$e(T) = c_v T, \tag{1.18}$$

making e proportional to the temperature T. The constant c_v is called the specific heat at constant volume. Rearranging (1.17), we see that p/ρ will also only depend on the temperature. We introduce a quantity h, such that

$$h(T) = e + \frac{p}{\rho} = (c_v + R)T = c_p T, \tag{1.19}$$

where c_p is called the specific heat at constant pressure. h is called the enthalpy of the system, and we see that it also depends only on T. For molecules with α degrees of freedom, the specific heats c_v and c_p are given by

$$\begin{aligned} c_v &= \frac{\alpha}{2}R, \\ c_p &= \left(1 + \frac{\alpha}{2}\right)R. \end{aligned}$$

This is well known from kinetic theory[13], and by definition the specific heats c_p and c_v are constant. We notice that the ratio between the specific heats c_v and c_p is independent of R

$$c_p/c_v = \frac{\alpha + 2}{\alpha} = \gamma,$$

and we call this ratio γ the adiabatic exponent. Note that γ does not depend on R. Common values of γ are $\gamma = 5/3$ and $\gamma = 7/5$ for monatomic and diatomic gases, respectively. The ideal gas law (1.17) states that

$$T = \frac{p}{R\rho}.$$

Inserting this into the expression for the internal energy, we obtain

$$e = \frac{\alpha R}{2}\frac{p}{R\rho} = \frac{\alpha p}{2\rho}.$$

Since $\frac{2}{\alpha} = \gamma - 1$, we have

$$e = \frac{p}{(\gamma - 1)\rho}. \tag{1.20}$$

This relation between e and p is important and will be used several times throughout the text. A common way to decompose the energy E is to divide it into its internal and kinetic energy terms

$$E = \rho e + \frac{1}{2}\rho w^2 = \frac{p}{\gamma - 1} + \frac{1}{2}\rho w^2. \tag{1.21}$$

In thermodynamics, entropy is a quantity that measures how unavailable a system's thermal energy is for conversion into mechanical work. It is often interpreted as the degree of disorder in the system. By definition, for a reversible process we have that the change in entropy ds is

$$ds = d\left(\ln T^{c_v} - \ln \rho^R\right). \tag{1.22}$$

According to the second law of thermodynamics, the entropy is non-decreasing, so $ds \geq 0$. In the isentropic case, then $ds = 0$. So then, by integrating (1.22) we get

$$\begin{aligned} s &= c_v \ln\left(\frac{T}{\rho^{R/c_v}}\right) \qquad (1.23)\\ &= c_v \ln(\frac{p}{\rho^\gamma}) + \text{const}. \end{aligned}$$

If we solve this for p, we get

$$p = \frac{1}{e^{\text{const}/c_v}} e^{s/c_v} \rho^\gamma$$

In the isentropic case, s is constant everywhere, so e^{s/c_v} is constant. Thus we define the constant κ as

$$\kappa = e^{\frac{s}{c_v} - \frac{\text{const}}{c_v}}.$$

We end up with the following isentropic relation between p and ρ for a polytropic, ideal gas

$$p = \kappa \rho^\gamma. \tag{1.24}$$

This expression for p also gives a useful term for the speed of sound c for a gas. The density and pressure varies within an acoustic wave, while the entropy remains constant, and the expression for the sound speed is

$$c = \sqrt{\left.\frac{\partial p}{\partial \rho}\right|_{s=const}} = \sqrt{\gamma \kappa \rho^{\gamma-1}},$$

with $p = p(\rho)$ such as in (1.24).

1.4.3 The isentropic Euler equations

By inserting the expression (1.24) for p and (1.20) for e into (1.21), the expression for the energy reduces to an expression that depends only on ρ and ρw, the state variables. Then the Energy equation (1.16) is redundant, and the Euler equations are reduced to

$$\begin{bmatrix} \rho \\ \rho w \end{bmatrix}_t + \begin{bmatrix} \rho w \\ \rho w^2 + p \end{bmatrix}_x = 0, \tag{1.25}$$

which is the isentropic Euler equations. What lets us drop the conservation of energy is the fact that it is implicit in the assumption that s is constant. In systems where the entropy really is constant, solutions would remain smooth, and conservation of energy would automatically be satisfied.

Nonlinear systems of equations will, even with arbitrarily regular data, tend to develop shocks; The isentropic Euler equations are no exceptions. When we have shocks in our solution, the entropy is not conserved. Across real shocks in gas dynamics, the entropy will increase, rather than remain constant. If we operate with the assumption that the system entropy is conserved around shocks, we will see that numerical schemes will develop huge oscillations around shocks unless we add some form of dissipation. The added dissipation is corresponds to energy changing forms into for example heat, which is not accounted for in the Euler equations.

We will see later that a shock is the correct vanishing-viscosity solution to the isentropic equations only if the energy increases across the shock, which reflects the the outside work that has been done on the system. Therefore, we can use the system energy as the entropy function η introduced in Section 1.3. Where the solution is smooth, the entropy equality (1.10) will apply, while the entropy inequality (1.12) will be a sufficient condition across shocks.

1.4.4 Hyperbolicity of the isentropic Euler equations

As we mentioned in Section 1.1, the Euler equations are hyperbolic. To verify that the Isentropic Euler equations are indeed a hyperbolic system, we find the eigenvalues of $f(u) = \begin{bmatrix} \rho w \\ \rho w^2 + p \end{bmatrix}$. The Jacobian is given by

$$J = f'(u) = \begin{bmatrix} 0 & 1 \\ -w^2 + \gamma\kappa\rho^{\gamma-1} & 2w \end{bmatrix} = 0, \tag{1.26}$$

so

$$\det(J - \lambda I) = \begin{vmatrix} -\lambda & 1 \\ -w^2 + \gamma\kappa\rho^{\gamma-1} & 2w - \lambda \end{vmatrix} = 0.$$

The eigenvectors are

$$\lambda = \frac{2w \pm \sqrt{4w^2 - 4\left(w^2 - \gamma\kappa\rho^{\gamma-1}\right)}}{2} = w \pm \sqrt{\gamma\kappa\rho^{\gamma-1}} = w \pm \sqrt{c}.$$

In other words, since the eigenvectors are real and distinct, the system is hyperbolic.

1.4.5 Entropy pair

The entropy of interest for the isentropic Euler equations is the total energy of the system[3], which for an isentropic system is (1.21)

$$\eta(u) = \frac{1}{2}\rho w^2 + \rho e(\rho). \tag{1.27}$$

Using this entropy makes sense, since the energy remains constant in an isentropic system. It is this entropy that will yield the physically relevant solutions to (1.7).

We proceed to prove that $\eta(u)$ is convex. In such a case the Hessian of $\eta(u)$ will be semi-positive definite, that is $H = \eta''(u) \geq 0$. The Hessian of $\eta(u)$ is given by

$$H = \begin{bmatrix} \frac{p'(\rho)}{\rho} + \frac{w^2}{\rho} & -\frac{w}{\rho} \\ -\frac{w}{\rho} & \frac{1}{\rho} \end{bmatrix}.$$

It is well known that a symmetric matrix M is positive definite if and only if there exists a non-singular matrix A such that

$$AA^T = M.$$

Since

$$\begin{bmatrix} \sqrt{\frac{p'(\rho)}{\rho}} & \frac{w}{\sqrt{\rho}} \\ 0 & \frac{-1}{\sqrt{\rho}} \end{bmatrix} \begin{bmatrix} \sqrt{\frac{p'(\rho)}{\rho}} & 0 \\ \frac{w}{\sqrt{\rho}} & \frac{-1}{\sqrt{\rho}} \end{bmatrix} = \begin{bmatrix} \frac{p'(\rho)}{\rho} & -\frac{w}{\rho} \\ -\frac{w}{\rho} & \frac{1}{\rho} \end{bmatrix},$$

then the Hessian must be positive definite and $\eta(u)$ is strictly convex. Next we verify that the entropy in question is indeed valid. Tadmor[18, 17] and Lax and Friedrichs[7] have stated that the entropy functions η will be exactly those whose positive Hessians H symmetrize (1.4) on the left, that is

$$HJ = [HJ]^T ,$$

with J from (1.26). HJ is

$$HJ = \begin{bmatrix} p''(\rho) & -\frac{w}{\rho} \\ -\frac{w}{\rho} & \frac{1}{\rho} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -w^2 + p'(\rho) & 2w \end{bmatrix} = \begin{bmatrix} \frac{w^3}{\rho} - \frac{w}{\rho} p'(\rho) & \frac{p'(\rho)}{\rho} - \frac{w^2}{\rho} \\ \frac{p'(\rho)}{\rho} - \frac{w^2}{\rho} & \frac{w}{\rho} \end{bmatrix},$$

which is indeed symmetric.

To find the entropy flux function q, recall that we wanted to impose condition (1.11) on a function q so that that the entropy equality holds. We would then have that

$$q'^T = \begin{bmatrix} \frac{p'(\rho)}{\gamma-1} - \frac{w^2}{2} \\ w \end{bmatrix}^T \begin{bmatrix} 0 & 1 \\ p'(\rho) - w^2 & 2w \end{bmatrix} = \begin{bmatrix} -w^3 + wp'(\rho) \\ \frac{3}{2} w^2 + \frac{p'(\rho)}{\gamma-1} \end{bmatrix}^T .$$

Integrating q' leaves us with the entropy flux function of the isentropic Euler equations,

$$q = w\left(\frac{1}{2}\rho w^2 + \rho e(\rho) + p(\rho)\right) = w\left(\frac{1}{2}\rho w^2 + \frac{\gamma}{(\gamma-1)}\kappa\rho^\gamma\right).$$

In accordance with the discussion in Section 1.3, η and q satisfies the entropy equality (1.10). Summarizing, we have that the entropy pair of interest to us is

$$(\eta(u), q(u)) = \left(\frac{1}{2}\rho w^2 + e(\rho),\, w\left(\frac{1}{2}\rho w^2 + \rho e(\rho) + p(\rho)\right)\right) \tag{1.28}$$

for the isentropic Euler equations.

1.4.6 The two-dimensional isentropic Euler equations

The two dimensional isentropic Euler equations will take the form

$$\begin{bmatrix} \rho \\ \rho w \\ \rho \omega \end{bmatrix}_t + \begin{bmatrix} \rho w \\ \rho w^2 + p(\rho) \\ \rho w \omega \end{bmatrix}_x + \begin{bmatrix} \rho \omega \\ \rho w \omega \\ \rho \omega^2 + p(\rho) \end{bmatrix}_y = 0. \tag{1.29}$$

This is a straightforward extension of the one-dimensional case, using (1.6) with

$$u = \begin{bmatrix} \rho \\ \rho w \\ \rho \omega \end{bmatrix}, \ f(u) = \begin{bmatrix} \rho w \\ \rho w^2 + p(\rho) \\ \rho w \omega \end{bmatrix}, \ g(u) = \begin{bmatrix} \rho \omega \\ \rho w \omega \\ \rho \omega^2 + p(\rho) \end{bmatrix}.$$

The extra requirement here is that we also require that the momentum is conserved in the second spatial dimension. We stated the definition of hyperbolicity for systems with one spatial dimension in Section 1.1. For two dimensions, the linear combination of the two Jacobians $J_1 = f'(u)$ and $J_2 = g'(u)$ has real eigenvalues for all values of u and the Jacobian matrices of f and g is diagonalizable. That is, for $\xi = [\xi_1, \xi_2]^T \in \mathbb{R}^2$, $|\xi| = 1$, $J_\xi = \xi_1 J_1 + \xi_2 J_2$ is diagonalizable and has real eigenvalues. The eigenvalues of f and g is given by $\lambda_{1,2,3}$ and $\mu_{1,2,3}$, respectively

$$\lambda_1 = w - c, \quad \lambda_2 = w, \quad \lambda_3 = w + c$$
$$\mu_1 = \omega - c, \quad \mu_2 = \omega \quad \mu_3 = \omega + c$$

with $c = \sqrt{\gamma \frac{p(\rho)}{\rho}}$ being the speed of sound for the gas in question.

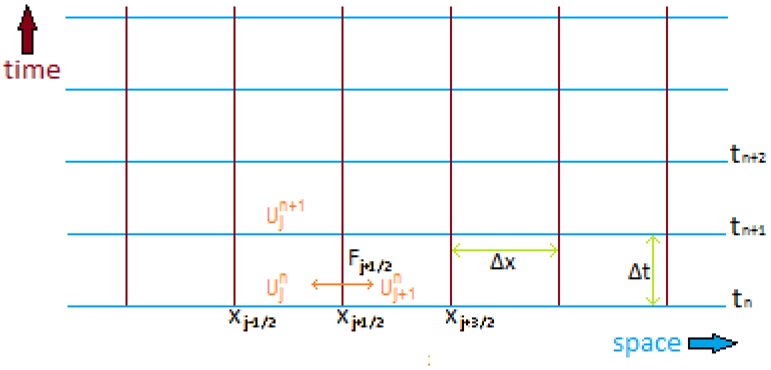


Figure 1: Discretization mesh for the finite volume methods

2 Finite volume methods

As we discussed, our solution is likely to form shocks, so we want a solution that accounts for u not being continuous everywhere. With finite difference schemes, series expansions are used to find an approximation, and in doing so, one will also assume continuous derivatives.

Finite volume methods are among the best suited methods for solving nonlinear conservation equations. With such a method, instead of looking at point values , we partition the domain into control volumes.

The approximated average values are over cells $C_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}})$. Time levels is denoted by $t_n = n\Delta t$, while space is discretized as $x_j = x_L + (j + \frac{1}{2})\Delta x$ for $j = 0, ..., N$, where $\Delta x = \frac{x_R - x_L}{N+1}$. Thus $x_{j-1/2} = x_j - \Delta x/2 = x_L + j\Delta x$ for $j = 0, ..., N+1$. We use these averages to approximate the solution using our cell averages

$$U_i^n \approx \frac{1}{\Delta x} \int\limits_{x_{i-1/2}}^{x_{i+1/2}} u(x, t_n)\, dx.$$

An illustration of our grid is given in Figure 1. Starting with our equation, $u_t + f(u)_x = 0$, we integrate over C_i and divide by Δx

$$\begin{aligned}
0 &= \frac{1}{\Delta x}\int\limits_{x_{i-1/2}}^{x_{i+1/2}} (u_t + f(u)_x)\,dx \\
&= \frac{d}{dt}(\frac{1}{\Delta x}\int\limits_{x_{i-1/2}}^{x_{i+1/2}} u(x,t)\,dx) + \frac{1}{\Delta x}(f(u(x_{i+\frac{1}{2}},t)) - f(u(x_{i-\frac{1}{2}},t))) \\
&\approx \frac{d}{dt}U_i^n(t) + \frac{f_{i+1/2} - f_{i-1/2}}{\Delta x}.
\end{aligned}$$

Here, U_i^n has been used as an approximation to the cell average,

$$U_i^n \approx \frac{1}{\Delta x}\int\limits_{x_{i-1/2}}^{x_{i+1/2}} u(x,t_n)\,dx,$$

and $f(u(x_{i\pm\frac{1}{2}},t))$ has been abbreviated as $f_{i\pm1/2}$. Proceeding by integrating over $t \in [t_n, t_{n+1}]$, the equation is

$$0 \approx U_i^{n+1} - U_i^n + \int\limits_{t^n}^{t^{n+1}} \frac{f_{i+1/2} - f_{i-1/2}}{\Delta x}\,dt = U_i^{n+1} - U_i^n + \frac{\Delta t}{\Delta x}(F_{i+1/2} - F_{i-1/2}).$$

$F_{i\pm1/2}^n = \frac{1}{\Delta t}\int\limits_{t_n}^{t_{n+1}} f_{i\pm1/2}\,dt$ is called the numerical flux. Using this notation, we write our original equation as

$$\underbrace{\frac{d}{dt}U_i}_{u_t} + \underbrace{\frac{F_{i+1/2} - F_{i-1/2}}{\Delta x}}_{f(u)_x} = 0.$$

A slight rearrangement yields the formula

$$\frac{d}{dt}U_i = -\frac{1}{\Delta x}(F_{i+1/2} - F_{i-1/2}). \tag{2.1}$$

By either explicitly calculating $F_{i\pm1/2}^n$ or by approximating it, different finite volume schemes can be constructed.

Godunov observed[8] that each cell in the finite volume scheme make up *Riemann problems*, that is, problems of the form

$$U_{Riemann} = \begin{cases} U_t + f(U)_x = 0 & (x \in \mathbb{R},\, t > t^n) \\ U(x,t^n) = \begin{cases} U_j^n & \text{if } x < x_{j+1/2} \\ U_{j+1}^n & \text{if } x > x_{j+1/2} \end{cases} \end{cases} \tag{2.2}$$

These problems can be solved exactly (like in Godunov's method), or by approximations, like in the methods presented in Section 2.1.

Solutions of Riemann problems are a sets of at most m waves that emanate out from $x = x_{i+1/2}$. The speed of propagation s for each wave is bound by the eigenvalues of $f'(U_i)$ and $f'(U_{i+1})$. Δt is selected so that the waves of each cell will not interact with each other. By denoting $s_{max} = \max_k |\lambda_k(U(x,t))|$, $x \in \mathbb{R}$ as the largest wave speed encountered, we have the condition that

$$s_{max}\frac{\Delta t}{\Delta x} \leq 1,$$

called the *CFL-condition* (CFL is short for Courant-Friedrichs-Lewy).

The finite volume scheme (2.1) can also be extended to account for two (or more) spatial dimensions. Our cell will then be the two dimensional area $C_{i,j} = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}) \times [y_{i-\frac{1}{2}}, y_{i+\frac{1}{2}})$, and

$$U_{i,j} = \frac{1}{\Delta x \Delta y}\int_{C_{i,j}} U(x,y,t)\,dxdy.$$

This gives us the corresponding scheme for two dimensions.

$$\frac{d}{dt}U_i = -\frac{\Delta t}{\Delta x}\left((F_{i+1/2,j} - F_{i-1/2,j}) + (G_{i,j+1/2} - G_{i,j-1/2})\right). \tag{2.3}$$

2.1 Popular Volume Schemes

Three of the most common finite volume methods are Lax-Friedrichs method, Rusanov's method and Roe's method. They are popular choices as they are fairly simple to implement. The fluxes all contain the term $\frac{1}{2}[f(U_{i+1})+f(U_i)]$, but their diffusion operators vary.

2.1.1 Lax-Friedrichs

Lax-Friedrichs method can by found by considering the rarefaction problems where we would have m waves traveling to the left and m waves to the right. The Lax-Friedrichs method uses maximum wave speeds given by $s^l_{i+1/2} = -\frac{\Delta x}{\Delta t}$ and $s^r_{i+1/2} = \frac{\Delta x}{\Delta t}$, so that $-s^l_{j+1/2} = s^r_{j+1/2} = s_{j+1/2}$. The flux function for this method is given by

$$F^n_{i+1/2} = \frac{1}{2}\left(f(U^n_{i+1}) + f(U^n_i)\right) - \frac{\Delta x}{2\Delta t}(U^n_{i+1} - U^n_i) \tag{2.4}$$

with numerical viscosity $\frac{\Delta x}{\Delta t}$ having fixed magnitude. The results we get by using this method are diffusive.

2.1.2 Rusanov

We get Rusanov's method if we instead decide to choose the speed of propagation for our waves locally as

$$-s^l_{i+1/2} = s^r_{i+1/2} = s_{i+1/2} = \max_k \left\{ \left|\lambda_k(U_i)\right|, \left|\lambda_k(U_{i+1})\right| \right\}.$$

For a system of equations, Rusanov's method is given by

$$F_{i+1/2} = \frac{1}{2}(f(U_{i+1}) + f(U_i)) - \frac{1}{2}s_{i+1/2}(U_{i+1} - U_i) \tag{2.5}$$

2.1.3 Roe

Another popular approximate Riemann solver is found in the Roe scheme. Roe suggested taking the nonlinear flux function and linearizing it locally. We do this by replacing f' with a constant $\hat{A}$. Therefore, we have that

$$f(u)_x = f'(u)u_x \approx \hat{A}_{i+1/2}U_x,$$

where $\hat{A}$ can be chosen in several ways. We want to find a matrix $\hat{A}$ to be diagonalizable with real eigenvalues, so that the system remains hyperbolic. Furthermore, we want it to be chosen so that the method is consistent with the original conservation law. For a linear system, we have that

$$\begin{aligned} F(U^n_i, U^n_{i+1}) &= \hat{A}U^n_i + \hat{A}^-(U^n_{i+1} + U^n_i) \qquad (2.6) \\ &= \hat{A}U^n_{i+1} - \hat{A}^+(U^n_{i+1} + U^n_i), \end{aligned}$$

where $A^+ = R\Lambda^+R^{-1}$, $A^- = R\Lambda^-R^{-1}$, R denotes the matrix of the eigenvectors of $\hat{A}$ and $\Lambda$is the diagonal matrix consisting of the eigenvalues of $\hat{A}$. Roe's method for systems of equations is given by the average of the terms in (2.6)

$$F_{i+1/2} = \frac{1}{2}(f(U_{i+1}) + f(U_i)) - \frac{1}{2}R\left|\Lambda\right|R^{-1}(U_{i+1} - U_i), \tag{2.7}$$

since $|A| = A^+ - A^- = R\left|\Lambda\right|R^{-1}$. We will discuss the construction of $\frac{1}{2}R\left|\Lambda\right|R^{-1}$ more in Section 3.7. It is worth noting that Roe's method does not always approximate the physical solution correctly. Indeed, the approximate Riemann solution can be very different from the real Riemann solution if there are more than one strong shock in the Riemann problem. One example is near the point where two shocks collide.

Roe found the Roe average

$$\tilde{U} = \begin{bmatrix} \frac{\rho_i + \rho_{i+1}}{2} \\ \frac{w_i\sqrt{\rho_i} + w_{i+1}\sqrt{\rho_{i+1}}}{\sqrt{\rho_i} + \sqrt{\rho_{i+1}}} \end{bmatrix} \tag{2.8}$$

for the Euler equations [15]. Using this average rather than an arithmetic one will guarantee exact resolution of single shocks.

3 Numerical Methods

3.1 Introduction

After looking at the isentropic Euler equations and nonlinear conservation equations in general, we proceed to develop our own scheme that preserves the energy. The semi-discrete finite volume method to (1.4) was then given by (2.1),

$$\frac{d}{dt}U_i = -\frac{1}{\Delta x}(F_{i+1/2} - F_{i-1/2}). \tag{3.1}$$

A finite volume scheme is said to be *conservative* if it can be written in this form[11]. A conservative volume scheme implies that the two flows from two adjoint cells C_i and C_{i+1} at point $x_{i+1/2}$ balance each other out. That is, they are both given as $F_{i+1/2}$.

Furthermore, if a flux is such that $F(u,u) = f(u)$ the flux F is called *consistent* with f. So when the solution in two cells C_i and C_{i+1} are equal, then the flux is equal to the conservation equation (1.1). A finite volume scheme with a consistent flux is a *consistent scheme.*

In this chapter, we will first introduce the concepts of entropy preservation and stability for discrete schemes. We will develop an energy conservative scheme in Section 3.4 and then develop it further so we get energy stability in Section 3.7. Numerical experiments are performed along each step. At the end, the methods are extended to two spatial dimensions.

3.1.1 Convergence

It's essential that our scheme is converging towards the correct solution as we refine our grid. In addition it is known that the weak formulation of conservation equations can have many different solutions, but we are only interested in the physically relevant one. This will require us to put extra conditions on the problem. The following theorem clarifies when we will get a weak solution of the conservation law.

Theorem 3.1. (Lax-Wendroff[11]). *Let $U^{(k)}$ be numerical approximations computed by a consistent and conservative scheme on a sequence of grids with mesh sizes $\Delta t^{(k)}$and $\Delta x^{(k)}$. Furthermore, let $\Delta t^{(k)}, \Delta x^{(k)} \to 0$ as $k \to \infty$. If $U^{(k)}$ converges boundedly to u almost everywhere as $k \to 0$, then the solution u is a weak solution of (1.4).*

In other words, if the conservative scheme $U^{(k)}$ converges and is consistent, it is a weak solution of the conservation law. However, note that the converse does not necessarily hold.

Now, to make sure that the solution is approaching the physical solution that we want to find, we look for numerical methods which is satisfying the entropy inequality (1.12). The discrete form of this is

$$\frac{d}{dt}\eta(U_i(t)) + \frac{1}{\Delta x}\left(Q_{i+1/2} - Q_{i-1/2}\right) \leq 0 \quad \text{for all } i \in \mathbb{Z} \text{ and } t > 0 \qquad (3.2)$$

with the numerical flux being $Q_{i+1/2} = Q(U_i^n, U_{i+1}^n)$ which is consistent with q and continuous, and $Q : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$. The following Theorem from Osher tells us when a scheme satisfies the entropy inequality.

Theorem 3.2. (Osher[14]) *Let $U^{(k)}$ be a sequence of numerical approximations computed by a consistent and conservative scheme that converges boundedly a.e. to a function u, like in Theorem 3.1. Now assume that an entropy flux Q exists that is consistent with q, so that (3.2) is satisfied for every $U^{(k)}$. Then u is a weak solution of $u_t + f(u)_x = 0$ that satisfies the entropy condition for the entropy pair (η, q).*

In other words, if we can find a such a scheme, it will be a weak solution that satisfies the discrete entropy inequality (3.2).

3.2 Choosing our time discretization

The entropy preserving method that we will develop is second-order accurate, so as to minimize the truncation error, it makes sense to use a second-order accurate time discretization too. So far we have only used the very simple first-order accurate forward Euler method. If we instead use the second-order Runge-Kutta method known as Heun's method, we can expect to get better accuracy when we look at how the energy is preserved. Let

$$\varsigma(U_i^k) = -\frac{1}{\Delta x}\left(F_{i+1/2}^k - F_{i-1/2}^k\right).$$

Then the Forward Euler is given by

$$U_i^{k+1} = U_i^k + \varsigma(U_i^k).$$

Heun's method calculates the intermediate value U_i^*, followed by the final solution at the next integration point, so

$$\begin{aligned} U_i^* &= U_i^k + \Delta t_k \varsigma(U_i^k) \\ U_i^{k+1} &= U_i^k + \frac{\Delta t_k}{2}\left[\varsigma(U_i^k) + \varsigma(U_i^*)\right]. \end{aligned}$$

As we add numerical viscosity in Section 3.6 and onwards, we will be dealing with first-order accurate methods. For these methods, it makes sense to either

1. Use forward Euler to discretize the time.
2. Reconstruct the diffusion operators, so that we again get methods that are second-order accurate.

We will not be reconstructing the diffusion operators, so we will only be using Heun's method when studying the energy conservative properties of the scheme we develop. When we later add diffusion, we will use forward Euler discretization.

3.3 Entropy Conservation

Tadmor explored how to construct schemes that satisfies the entropy condition, first for scalar conservation laws in 1984[16] and then for systems of conservation laws in 1987[18]. If we have an entropy pair(η, q), then a scheme is entropy conservative if it satisfies the discrete entropy equality

$$\frac{d}{dt}\eta(U_i(t) + \frac{1}{\Delta x}\left(Q_{i+1/2} - Q_{i-1/2}\right) = 0. \tag{3.3}$$

Summing up for $i = k, k+1, ..., l-1, l$, we get

$$\frac{d}{dt}\left(\Delta x \sum_{i=k}^{l} \eta(U_i(t))\right) = Q_{l+1/2} - Q_{k-1/2}$$

for some chosen endpoints k, l. If the entropy flux is zero on the boundary, then the total entropy in the measured area will be preserved. If a scheme satisfies the discrete entropy inequality

$$\frac{d}{dt}\eta(U_i(t)) + \frac{1}{\Delta x}\left(Q_{i+1/2} - Q_{i-1/2}\right) \leq 0, \tag{3.4}$$

it is said to be entropy stable. As mentioned in Section 1.3, this allows for the entropy to decrease around the shocks. Tadmor found an explicit condition that would ensure entropy conservation in a scheme[16, 18].

Any method with more numerical viscosity than an entropy conservative method is entropy stable[18]. Hence, if we first develop an entropy conservative method, we can later add numerical viscosity to it to ensure entropy stability (we do this in Section 3.6). This will be our approach. We recall that the relevant entropy for the isentropic Euler equations is the energy of the system. Therefore, we will use the terms *energy conservative* and *energy stable* when we talk about the entropy of our scheme in particular.

We want our solution of (1.4) to satisfy the entropy inequality. This implies that the measured entropy is non-increasing, so

$$\frac{d}{dt}\int_R \eta(u)dx \leq 0.$$

For a numerical scheme that instead satisfies the discrete entropy inequality, we get a discrete stability estimate,

$$\frac{d}{dt}\Delta x \sum_i \eta(U_i) \leq 0. \tag{3.5}$$

In the case where $\eta(u)$ is preserved, the stability estimate will be

$$\frac{d}{dt}\Delta x \sum_i \eta(U_i) = 0.$$

Using the same procedure as in Section 1.3 on (2.1), we get

$$\begin{aligned} \eta'(U)^T\left(\frac{d}{dt}U_i\right) &= -\eta'(U_i)^T\left(\frac{1}{\Delta x}\left(F_{i+1/2}-F_{i-1/2}\right)\right) \\ \frac{d}{dt}\eta_i(U_i) &= -\eta'(U_i)^T\left(\frac{1}{\Delta x}\left(F_{i+1/2}-F_{i-1/2}\right)\right). \end{aligned} \tag{3.6}$$

Similarly to the non-discrete case, we will need put the requirement that $\eta'(U_i)\left(\left(F_{i+1/2}-F_{i-1/2}\right)\right) = \left(Q_{i+1/2}-Q_{i-1/2}\right)$ for (3.3) to hold. Here, $Q_{i+1/2} = Q(U_i, U_{i+1})$ is the numerical entropy flux and is consistent with q.

$\eta'(u)$ is central in the theory behind entropy conservation and stability, and is called the entropy variables. It is usually denoted by

$$v = v(u) = \eta'(u). \tag{3.7}$$

What kind of criteria can we place on $F_{i+1/2}$ to ensure that

$$\eta'(U_i)\left(\left(F_{i+1/2}-F_{i-1/2}\right)\right) = \left(Q_{i+1/2}-Q_{i-1/2}\right)?$$

We make use of the following theorem from Tadmor [18]. First, note that the entropy potential is $\psi(u) = v(u)^T f(u) - q(u)$ by definition.

Theorem 3.3. (Tadmor[18]). *Solutions computed by the scheme with consistent numerical flux $F_{i+1/2}$ satisfy the discrete entropy equality (3.3) with numerical entropy flux*

$$\tilde{Q}_{i+1/2} = \overline{V}^T_{i+1/2}F_{i+1/2} - \overline{\psi}_{i+1/2} \tag{3.8}$$

if and only if

$$[\![V]\!]_{i+1/2}^{T} F_{i+1/2} = [\![\psi]\!]_{i+1/2} \tag{3.9}$$

Here, $\tilde{Q}$ is consistent with q.

This can be verified by rearranging (3.6) and using the definition of the entropy potential.

3.3.1 Example: Burgers' Equation

If we use Burgers' equation as an example, that is, a function of form (1.4), where $f = \frac{1}{2}u^2$, we can choose our entropy function as $\eta = u^2$. Then $q'(u)$ is given by

$$q'(u) = \eta'(u)f'(u) = 2uu = 2u^2 \Rightarrow q(u) = \frac{2}{3}u^3.$$

The entropy potential by definition will be

$$\psi(u) = 2u\frac{u^2}{2} - \frac{2}{3}u^3 = \frac{1}{3}u^3.$$

Now, using (3.9) $[\![2U]\!]_{i+1/2} F_{i+1/2} = [\![\frac{1}{3}U^3]\!]_{i+1/2}$. Then,

$$F_{i+1/2} = \frac{1}{6}\frac{[\![U^3]\!]_{i+1/2}}{[\![U]\!]_{i+1/2}} = \frac{1}{6}\frac{U_{i+1}^3 - U_i^3}{U_{i+1} - U_i} = \frac{U_{i+1}^2 + U_iU_{i+1} + U_i^2}{6}. \tag{3.10}$$

With this, we have found an entropy conservative flux function of Burgers' equation.

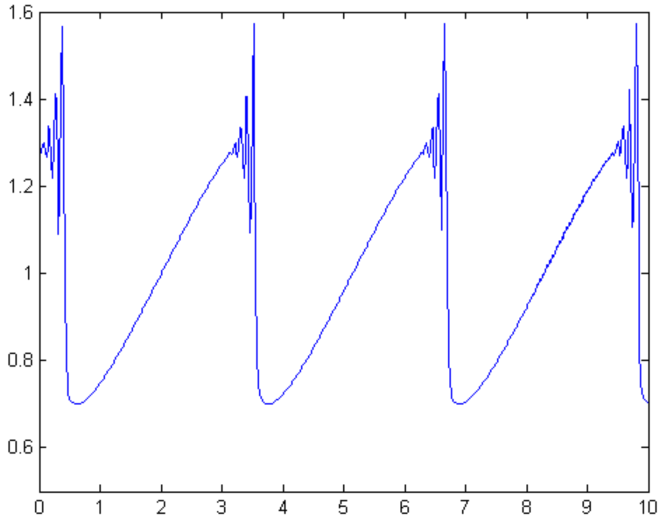


Figure 2: Solution u of (3.11) at $t = 2$, using the flux (3.10). This flux lacks numerical viscosity and gives large oscillations.

Plotting the solution of u at time $t = 2$ with the initial value problem

$$u_0(x) = 1 + 0.3\sin(2x), \tag{3.11}$$

we see that large oscillations develop around discontinuities, as seen in Figure 2. This is expected, and to avoid oscillations we must add numerical viscosity. We continue the example in Section 3.6.

3.4 Developing the Energy Conservative Scheme(EEC)

As mentioned in Section 1.4.5, the relevant entropy function of the isentropic Euler equations to be the total energy of the system. As already explored in Section 3.3 schemes satisfying (3.3) will be entropy conservative. We proceed to develop such a method for the isentropic Euler equations.

3.4.1 The discretization of F

We wish to find a discretization of the numerical flux $F_{i+1/2}$, denoted by $\tilde{F}_{i+1/2} = \begin{bmatrix} \tilde{F}^1_{i+1/2} \\ \tilde{F}^2_{i+1/2} \end{bmatrix} = \begin{bmatrix} \{\rho w\} \\ \{\rho w^2\} + \{\kappa\rho^\gamma\} \end{bmatrix}$, where we let $\{\}$ denote the discretization of a quantity. The idea of our scheme is to find such a discretization $\tilde{F}$ for the isentropic Euler equations (1.25) so that (3.9) holds. For Polytropic gases, p is given by (1.24). Recall that the entropy pair of the isentropic Euler equation (1.28) for a polytropic gas is

$$\eta = \frac{1}{2}\rho w^2 + \kappa\frac{\rho^\gamma}{(\gamma-1)}, \qquad q = w\left(\frac{1}{2}\rho w^2 + \frac{\gamma}{(\gamma-1)}\kappa\rho^\gamma\right).$$

As mentioned in Section 3.3, our entropy variables (3.7) are then defined as

$$v(u) = \eta'(u) = \begin{bmatrix} \kappa\frac{\gamma}{(\gamma-1)}\rho^{\gamma-1} - \frac{w^2}{2} \\ w \end{bmatrix}. \tag{3.12}$$

Again, by definition, the entropy potential of the isentropic Euler equations are

$$\begin{aligned} \psi(u) &= \left(\kappa\frac{\gamma\rho^{\gamma-1}}{(\gamma-1)} - \frac{w^2}{2}\right)w + w\left(w\rho + \kappa\cdot\rho^\gamma\right) - w\left(\frac{1}{2}w\rho + \rho\kappa\frac{\rho^\gamma}{(\gamma-1)} + \kappa\rho^\gamma\right) \\ &= \kappa w\rho^\gamma \end{aligned}$$

We wish (3.9) to hold, so that

$$0 = \llbracket V \rrbracket^T_{i+1/2}\cdot\tilde{F}_{i+1/2} - \llbracket \psi \rrbracket_{i+1/2}.$$

Filling in, we have

$$\begin{aligned} 0 &= \left\llbracket \kappa\frac{\gamma\rho^{\gamma-1}}{(\gamma-1)} - \frac{1}{2}w^2 \right\rrbracket_{i+1/2} \tilde{F}^1_{i+1/2} + \llbracket w \rrbracket\,\tilde{F}^2_{i+1/2} - \llbracket \psi \rrbracket_{i+1/2} \\ &= \kappa\frac{\gamma}{(\gamma-1)}\llbracket \rho^{\gamma-1} \rrbracket_{i+1/2}\tilde{F}^1_{i+1/2} - \overline{w}_{i+1/2}\llbracket w \rrbracket_{i+1/2}\tilde{F}^1_{i+1/2} + \llbracket w \rrbracket\,\tilde{F}^2_{i+1/2} - \llbracket \psi \rrbracket_{i+1/2}. \end{aligned}$$

To simplify, we introduce the term $\Omega = \kappa\frac{\gamma}{\gamma-1}$. We make use of the following identities from (1.2). We then get

$$\begin{aligned} 0 &= \Omega\llbracket \rho^{\gamma-1} \rrbracket_{i+1/2}\tilde{F}^1_{i+1/2} - \overline{w}_{i+1/2}\llbracket w \rrbracket_{i+1/2}\tilde{F}^1_{i+1/2} + \llbracket w \rrbracket_{i+1/2}\tilde{F}^2_{i+1/2} - \kappa\llbracket w\rho^\gamma \rrbracket_{i+1/2} \\ &= \Omega\llbracket \rho^{\gamma-1} \rrbracket_{i+1/2}\tilde{F}^1_{i+1/2} - \kappa\overline{w}_{i+1/2}\llbracket \rho^\gamma \rrbracket_{i+1/2} \\ &\quad + \llbracket w \rrbracket_{i+1/2}\left(\tilde{F}^2_{i+1/2} - \overline{w}_{i+1/2}\tilde{F}^1_{i+1/2} - \kappa\overline{\rho^\gamma}_{i+1/2}\right) \end{aligned}$$

The idea behind the scheme that we are developing is to manipulate the terms so that they are expressed in terms of jumps over the same variable. To be able to do this, we must factorize $\llbracket \rho^{\gamma-1} \rrbracket_{i+1/2}$ and $\llbracket \rho^\gamma \rrbracket_{i+1/2}$ into common terms. Recall that $\gamma - 1 = \frac{2}{\alpha}$, where α is the degrees of freedom of the gas. We can factorize

$$\begin{aligned} 0 &= 2\Omega\overline{\rho^{1/\alpha}}_{i+1/2}\left\llbracket\rho^{1/\alpha}\right\rrbracket_{i+1/2}\tilde{F}^1_{i+1/2} - \kappa\overline{w}_{i+1/2}\left\llbracket\rho^{(2+\alpha)/\alpha}\right\rrbracket_{i+1/2} \\ &\quad + \llbracket w\rrbracket_{i+1/2}\left(\tilde{F}^2_{i+1/2} - \overline{w}_{i+1/2}\tilde{F}^1_{i+1/2} - \kappa\overline{\rho^{\gamma}}_{i+1/2}\right)_{i+1/2} \end{aligned}$$

The second term on the right side will vary different depending on the values of alpha. Commonly, α takes the value of either 3 or 5, so that $\gamma = 5/3$ or $\gamma = 7/5$. If we can find a function such that

$$\left\llbracket\rho^{n/\alpha}\right\rrbracket_{i+1/2} = \Gamma(n,\alpha)_{i+1/2}\left\llbracket\rho^{1/\alpha}\right\rrbracket_{i+1/2}, \tag{3.13}$$

we can group $\llbracket\rho^{\gamma-1}\rrbracket_{i+1/2}$ and $\llbracket\rho^{\gamma}\rrbracket_{i+1/2}$ in terms of $\left\llbracket\rho^{1/\alpha}\right\rrbracket_{i+1/2}$. We proceed to look at for a pattern in the difference identities.

$$\left\llbracket\rho^{2/\alpha}\right\rrbracket_{i+1/2} = \left\llbracket\rho^{1/\alpha}\right\rrbracket_{i+1/2}\overline{\rho^{1/\alpha}}_{i+1/2} + \overline{\rho^{1/\alpha}}_{i+1/2}\left\llbracket\rho^{1/\alpha}\right\rrbracket_{i+1/2} = 2\left\llbracket\rho^{1/\alpha}\right\rrbracket_{i+1/2}\overline{\rho^{1/\alpha}}_{i+1/2},$$

$$\left\llbracket\rho^{3/\alpha}\right\rrbracket_{i+1/2} = \llbracket\rho\rrbracket_{i+1/2}\overline{\rho^{2/\alpha}}_{i+1/2} + \left\llbracket\rho^{2/\alpha}\right\rrbracket_{i+1/2}\overline{\rho^{1/\alpha}}_{i+1/2},$$

$$\left\llbracket\rho^{n/\alpha}\right\rrbracket_{i+1/2} = \left\llbracket\rho^{1/\alpha}\right\rrbracket_{i+1/2}\overline{\rho^{(n-1)/\alpha}}_{i+1/2} + \left\llbracket\rho^{(n-1)/\alpha}\right\rrbracket_{i+1/2}\overline{\rho^{1/\alpha}}_{i+1/2}.$$

We notice that there is a function

$$\Gamma(n,\alpha)_{i+1/2} = \Gamma(n-1,\alpha)_{i+1/2}\overline{\rho^{1/\alpha}}_{i+1/2} + \overline{\rho^{n-1/\alpha}}_{i+1/2},\ \Gamma(1,\alpha) = 1$$

that satisfies (3.13). The two differences we wanted to group together are then written as

$$\begin{aligned} \left\llbracket\rho^{\frac{2}{\alpha}}\right\rrbracket_{i+1/2} &= \Gamma(2,\alpha)_{i+1/2}\left\llbracket\rho^{1/\alpha}\right\rrbracket_{i+1/2} = 2\overline{\rho^{1/\alpha}}_{i+1/2}\left\llbracket\rho^{1/\alpha}\right\rrbracket_{i+1/2}, \\ \left\llbracket\rho^{\frac{\alpha+2}{\alpha}}\right\rrbracket_{i+1/2} &= \Gamma(2+\alpha,\alpha)_{i+1/2}\left\llbracket\rho^{1/\alpha}\right\rrbracket_{i+1/2}. \end{aligned}$$

We then get

$$\begin{aligned} 0 &= \left\llbracket\rho^{1/\alpha}\right\rrbracket_{i+1/2}\left(2\Omega\overline{\rho^{1/\alpha}}_{i+1/2}\tilde{F}^1_{i+1/2} - \kappa\overline{w}_{i+1/2}\Gamma(2+\alpha,\alpha)\right) \\ &\quad + \llbracket w\rrbracket_{i+1/2}\left(\tilde{F}^2_{i+1/2} - \overline{w}\tilde{F}^1_{i+1/2} - \kappa\overline{\rho^{\gamma}}_{i+1/2}\right). \end{aligned}$$

We want to choose such $(\tilde{F}^1_{i+1/2}, \tilde{F}^2_{i+1/2})$that

$$\begin{aligned} \left(2\Omega\overline{\rho^{1/\alpha}}_{i+1/2}\tilde{F}^1_{i+1/2} - \kappa\overline{w}_{i+1/2}\Gamma(2+\alpha)_{i+1/2}\right) &= 0 \\ \left(\tilde{F}^2_{i+1/2} - \overline{w}_{i+1/2}\tilde{F}^1_{i+1/2} - \kappa\overline{\rho^{\gamma}}_{i+1/2}\right) &= 0 \end{aligned}$$

This gives us

$$\tilde{F}_{i+1/2} = \begin{bmatrix} \tilde{F}^1_{i+1/2} \\ \tilde{F}^2_{i+1/2} \end{bmatrix} = \begin{bmatrix} \frac{\gamma-1}{\gamma} \frac{\Gamma(2+\alpha)_{i+1/2}}{2\overline{\rho^{1/\alpha}}_{i+1/2}} \overline{w}_{i+1/2} \\ \overline{w}_{i+1/2} \tilde{F}^1_{i+1/2} + \kappa \overline{\rho^\gamma}_{i+1/2} \end{bmatrix}$$

We end up with a scheme on the form

$$\frac{d}{dt} U_i = \frac{-1}{\Delta x} \left(\tilde{F}_{i+1/2} - \tilde{F}_{i-1/2} \right) \tag{3.14}$$

Which is our Explicit Energy Conserving scheme, shortened as EEC.

3.4.2 Two-dimensional EEC Scheme

We proceed to find a discretization for the two-dimensional system (1.29). The entropy function for this system is

$$\eta(u) = \rho e(\rho) + \frac{1}{2}(w^2 + \omega^2),$$

which gives us the entropy variables

$$v(u) = \begin{bmatrix} \kappa \frac{\gamma}{(\gamma-1)} \rho^{\gamma-1} - \frac{1}{2} \frac{w^2+\omega^2}{\rho^2} \\ w \\ \omega \end{bmatrix}. \tag{3.15}$$

By definition, the entropy potentials are calculated to be

$$\psi^x = \kappa \rho^\gamma w,$$
$$\psi^y = \kappa \rho^\gamma \omega.$$

The method remains the same, and we end up with similar results to what we got in Section (3.4.1)

$$\begin{aligned} 0 \;=\; & \Omega \left\llbracket \rho^{\gamma-1} \right\rrbracket_{i+1/2} \tilde{F}^1_{i+1/2} - \kappa \overline{w}_{i+1/2} \llbracket \rho^\gamma \rrbracket_{i+1/2} \\ & + \llbracket w \rrbracket_{i+1/2} \left(\tilde{F}^2_{i+1/2} - \overline{w}_{i+1/2} \tilde{F}^1_{i+1/2} - \kappa \overline{\rho^\gamma}_{i+1/2} \right) + \llbracket \omega \rrbracket_{i+1/2} \left(\tilde{F}^3_{i+1/2} - \overline{\omega} \tilde{F}^1_{i+1/2} \right) \end{aligned}$$

$$\tilde{F}_{i+1/2} = \begin{bmatrix} \tilde{F}^1_{i+1/2} \\ \tilde{F}^2_{i+1/2} \\ \tilde{F}^3_{i+1/2} \end{bmatrix} = \begin{bmatrix} \frac{\gamma-1}{\gamma} \frac{\Gamma(2+\alpha)_{i+1/2}}{2\overline{\rho^{1/\alpha}}_{i+1/2}} \overline{w}_{i+1/2} \\ \overline{w} \tilde{F}^1_{i+1/2} + \kappa \overline{\rho^\gamma}_{i+1/2} \\ \overline{\omega} \tilde{F}^1_{i+1/2} \end{bmatrix}$$

Similarly, for G

$$\begin{aligned} 0 &= \Omega \left[\!\left[\rho^{\gamma-1}\right]\!\right]_{i+1/2} \tilde{G}^1_{i+1/2} - \kappa\overline{\omega}_{i+1/2} \left[\!\left[\rho^{\gamma}\right]\!\right]_{i+1/2} + \left[\!\left[w\right]\!\right]_{i+1/2} \left(\tilde{G}^2_{i+1/2} - \overline{w}\tilde{G}^1_{i+1/2}\right) \\ &\quad + \left[\!\left[\omega\right]\!\right]_{i+1/2} \left(\tilde{G}^3_{i+1/2} - \overline{\omega}_{i+1/2}\tilde{G}^1_{i+1/2} - \kappa\overline{\rho^{\gamma}}_{i+1/2}\right) \end{aligned}$$

$$\tilde{G}_{i+1/2} = \begin{bmatrix} \tilde{G}^1_{i+1/2} \\ \tilde{G}^2_{i+1/2} \\ \tilde{G}^3_{i+1/2} \end{bmatrix} = \begin{bmatrix} \frac{\gamma-1}{\gamma}\frac{\Gamma(2+\alpha)_{i+1/2}}{2\overline{\rho^{1/\alpha}}_{i+1/2}}\overline{\omega}_{i+1/2} \\ \overline{w}_{i+1/2}\tilde{G}^1_{i+1/2} \\ \overline{\omega}_{i+1/2}\tilde{G}^1_{i+1/2} + \kappa\overline{\rho^{\gamma}}_{i+1/2} \end{bmatrix}$$

Which satisfies the entropy equality for two dimensions, that is

$$\frac{d}{dt}\eta(U_i(t)) + \frac{1}{\Delta x}\left(Q^x_{i+1/2,j} - Q^x_{i-1/2,j}\right) + \frac{1}{\Delta y}\left(Q^y_{i,j+1/2} - Q^y_{i,j-1/2}\right) = 0$$

Thus, our two-dimensional scheme ends up with the form

$$\frac{d}{dt}U_i = -\frac{1}{\Delta x}\left(\tilde{F}_{i+1/2} - \tilde{F}_{i-1/2}\right) - \frac{1}{\Delta y}\left(\tilde{G}_{i+1/2} - \tilde{G}_{i-1/2}\right)$$

3.4.3 Scheme Analysis

The properties of the scheme that we found is denoted as the *Explicit Entropy Conservative* scheme, abbreviated as *EEC*.

Theorem 3.4. *The EEC scheme (3.14) is consistent with the isentropic Euler equations (1.25). It is second order accurate and energy preserving, i.e. it satisfies the entropy equality (3.3) and has the numerical entropy flux*

$$\tilde{Q}_{i+1/2} = \overline{V}^T_{i+1/2}F_{i+1/2} - \overline{\psi}_{i+1/2} = \kappa\frac{\gamma}{(\gamma-1)}\overline{\rho^{\gamma-1}}\tilde{F}_1 - \frac{1}{2}\overline{w^2}\tilde{F}_1 + \overline{w}^2\tilde{F}_1 + \kappa\overline{w}\overline{\rho^{\gamma}} - \kappa\overline{w\rho^{\gamma}},$$

where $\tilde{Q}$ is consistent with q.

Proof. We chose $\tilde{Q}$ so that it satisfies (3.8) so it follows that the EEC scheme satisfies the energy equality according to Theorem 3.3. The accuracy is shown by truncation error analysis. Let F^i_x denote the approximation of the derivative of f^i, that is f at the ith space step. Then, from (3.14) we have that

$$F^i_x = \frac{\tilde{F}_{i+1/2} - \tilde{F}_{i-1/2}}{\Delta x}$$

Taylor expansion of $\tilde{F}_{i+1/2}$ and $\tilde{F}_{i-1/2}$ gives

$$\begin{aligned}
\tilde{F}_{i+1/2} &= f^i + f^i_x \frac{\Delta x}{2} + \frac{1}{2} f^i_{xx} \left(\frac{\Delta x}{2}\right)^2 + \frac{1}{6} f^i_{xxx} \left(\frac{\Delta x}{2}\right)^3 \\
\tilde{F}_{i-1/2} &= f^i - f^i_x \frac{\Delta x}{2} + \frac{1}{2} f^i_{xx} \left(\frac{\Delta x}{2}\right)^2 - \frac{1}{6} f^i_{xxx} \left(\frac{\Delta x}{2}\right)^3
\end{aligned}$$

Then the difference between the real f^i_x and F^i_x is then

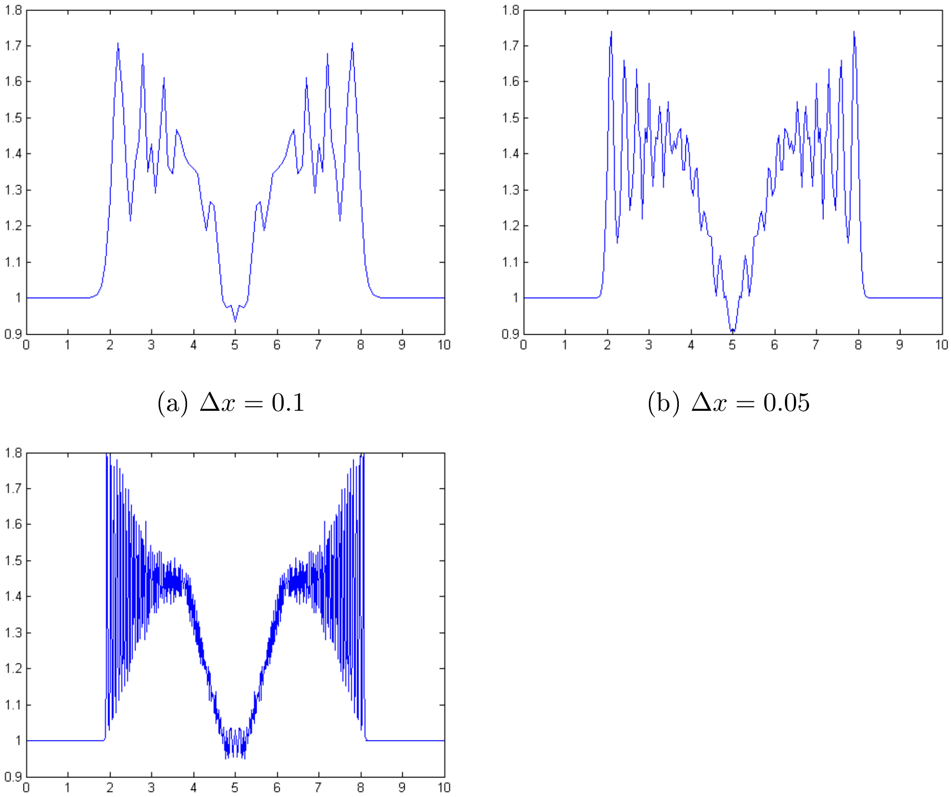
$$F^i_x - f^i_x = \frac{1}{24} f^i_{xxx} \Delta x^2 + \mathcal{O}(\Delta x^3),$$

which means that our scheme is second order accurate. The consistency follows from the definition. □

3.5 Numerical Experiments

Our first test problem is given by the initial conditions

$$\begin{aligned} \rho(x,0) &= \begin{cases} 2 & \text{if } 4 < x < 6 \\ 1 & \text{elsewhere} \end{cases}, \\ w(x,0) &= 0. \end{aligned} \tag{3.16}$$



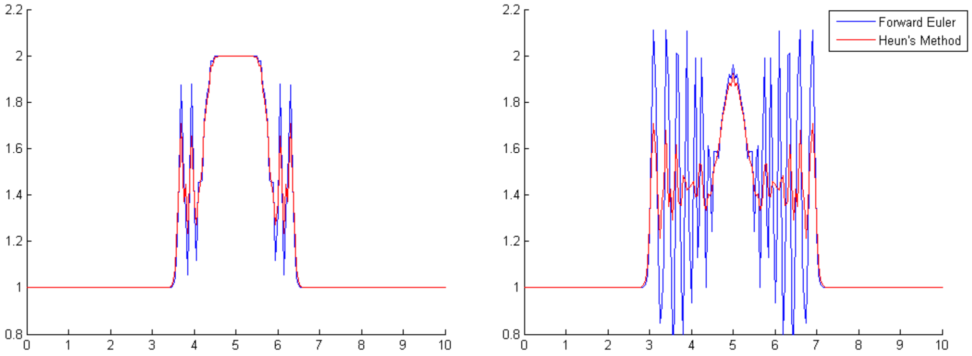
(a) $\Delta x = 0.1$

(b) $\Delta x = 0.05$

(c) $\Delta x = 0.01$

Figure 3: The density ρ from solutions of (3.16) at time $t = 1$, with increasing number of points.

The solution of this problem should see shocks develop in both directions. We notice that the shocks are resolved as they should, but we see large oscillations. As illustrated in Figure 3, the frequency of these oscillations increase as we refine our mesh in space; they have a period of order Δx. To get rid of the oscillations we will later add a dissipating factor to the developed scheme.



(a) $t = 0.2$

(b) $t = 0.5$

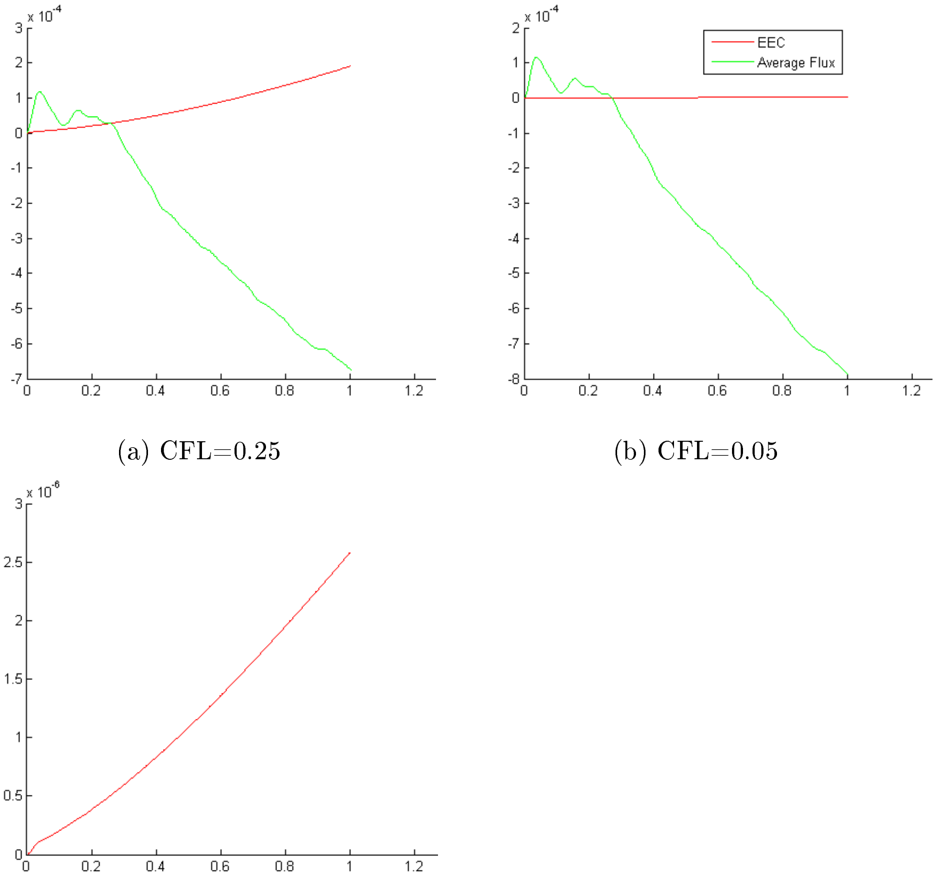
Figure 4: The density ρ from solutions of the EEC scheme, using the Forward Euler and Heun's methods time discretization.

Figure 4 shows the importance of using the Heun's method, which is of second-order accuracy when we compute the solution. As seen in the figure, oscillations will grow at a much faster pace when we only use the first-order accurate Forward Euler method for time discretization.

For our next experiment we introduce the flux

$$F_{i+1/2}^{avg} = \frac{1}{2}\left(f(U_i) + f(U_{i+1})\right) \tag{3.17}$$

to compare our scheme with, which we will just call the average flux. This is just the first term from the methods presented in Section 2.1, namely (2.4), (2.5) and (2.7). For simplicity, we will denote the finite volume scheme with flux $F_{i+1/2}^{avg}$ as the average flux scheme.



(a) CFL=0.25

(b) CFL=0.05

(c) CFL=0.05

Figure 5: Relative change in energy η^{rel}_{change} over time for the EEC method and the average flux without their dissipating terms.

We wish to test the energy conserving properties of the EEC method (3.14). To do this we introduce a measure for relative change in energy over time

$$\eta^{rel}_{change} = \frac{\|\eta^n_{tot} - \eta^0_{tot}\|_{L^1}}{\|\eta^0_{tot}\|_{L^1}},$$

using η^n_{tot} as the total energy in our scheme at time step n, $\eta^n_{tot} = \sum_i \eta(U^n_i)$. As long as the entropy flux q is zero on the border, the entropy should be conserved with our scheme, and η^{rel}_{change} should have a bound that is negligible. Figure 5 illustrates this well, with a sharp reduction in “lost” energy for our method as we reduce the CFL-number. On the other hand, a scheme using the flux (3.17) sees an increase in η^{rel}_{change}, as the method using such a flux is not energy conserving.

CFL	Runtime	η^{rel}_{change}
0.25	0.084035	1.34711e-4
0.20	0.101188	7.23383e-5
0.15	0.133578	3.32239e-5
0.10	0.199802	1.15557e-5
0.05	0.393697	2.10832e-6
0.025	0.781972	4.30241e-7
0.01	1.939399	5.95635e-8

(a) EEC scheme using Heun's method

CFL	Runtime	η^{rel}_{change}
0.25	0.041048	8.3892e-2
0.20	0.046873	4.5319e-2
0.15	0.061111	2.2807e-2
0.10	0.090872	1.0484e-2
0.05	0.183622	3.7439e-3
0.025	0.360909	1.6157e-3
0.01	0.905467	5.9518e-4

(b) EEC scheme using the forward Euler's method

Table 1: Runtimes and change in energy η^{rel}_{change} for the EEC scheme.

Table 1 measures the runtime of the EEC scheme and its relative change in energy for different CFL numbers. The runtime is measured in seconds. The importance of using a second order time discretization becomes apparent when we compare the values of the two tables, as the values of η^{rel}_{change} in Table 1a are several order of magnitude smaller than that of Table 1b .

CFL	Runtime	η^{rel}_{change}
0.25	0.024229	2.4255e-4
0.20	0.029376	2.8123e-4
0.15	0.039802	3.0552e-4
0.10	0.059685	3.1523e-4
0.05	0.121595	3.1861e-4
0.025	0.239345	3.1881e-4
0.01	0.603769	3.1851e-4

(a) Average Flux Scheme with Heun's method

CFL	Runtime	η^{rel}_{change}
0.25	0.033777	1.7319e-2
0.20	0.042700	1.7267e-2
0.15	0.063292	1.7260e-2
0.10	0.091174	1.7298e-2
0.05	0.181621	1.7272e-2
0.025	0.354508	1.7259e-2
0.01	0.887423	1.7252e-2

(b) Roe's Scheme with Heun's method

Table 2: Runtimes and change in energy η^{rel}_{change} for the average flux scheme and Roe's method. The relative energy does not decrease with smaller CFL-numbers, meaning that the schemes are not energy preserving.

As we can see, neither of these schemes well conserve the energy as we decrease the CFL-number. The average flux scheme and Roe's method are computationally cheap to implement, and when we compare the runtimes of Table 2 with those of Table 1, we see that the EEC method is somewhat slower. However, neither the Average flux scheme or Roe's scheme conserves the energy. Computation of η^{rel}_{change} using Lax-Friedrichs method and Rusanov's method yielded similar results. Both schemes had lower computational cost, but no Energy preservation. It is no big surprise that Lax-Friedrichs scheme, Rusanov's Scheme and Roe's scheme are not energy conserving, as they contain diffusion operators which dissipates energy around shocks. Thus, we can expect the energy to be decreasing around shocks.

As stated in Theorem 3.4, the EEC scheme is second-order accurate. If we let the L_1-error of u with spatial grid-size Δx be denoted as $E^{\Delta x}$, then $E^{\Delta x}$ is given by $E^{\Delta x} = \sum_i |U_i - u_i| \Delta x$, where U_i and u_i are the approximation and real solution of u at point $x = i\Delta x$, respectively. A scheme with accuracy k should have its error be given by $E^{\Delta x} = \mathcal{O}(\Delta x^k)$. In other words, the expression for $E^{\Delta x}$ is readily given as

$$E^{\Delta x} = C\Delta x^k,$$

with C being some constant. If we then use different spatial mesh sizes Δx_1 and Δx_2 and evaluate $E^{\Delta x}$ for each of them, we can use two of these expressions to eliminate C by division

$$\frac{E^{\Delta x_1}}{E^{\Delta x_2}} = \left(\frac{\Delta x_1}{\Delta x_2}\right)^k.$$

By taking the logarithm of both sides, we get the expression for k,

$$k = \left(\frac{E^{\Delta x_1}/E^{\Delta x_2}}{\Delta x_1/\Delta x_2}\right).$$

When we want to find the rate of convergence by numerical experiments, test problems such as (3.16) are not well suited because of the strong shocks that exist in the solution from the start. Instead, we choose initial value data that will stay smooth for low values of t. One such test problem is

$$\rho(x,0) = \begin{cases} 1 + \sin(2x - \frac{\pi}{2}) & \text{if } \pi < x < 2\pi \\ 1 & \text{elsewhere} \end{cases}, w(x,0) = 0. \tag{3.18}$$

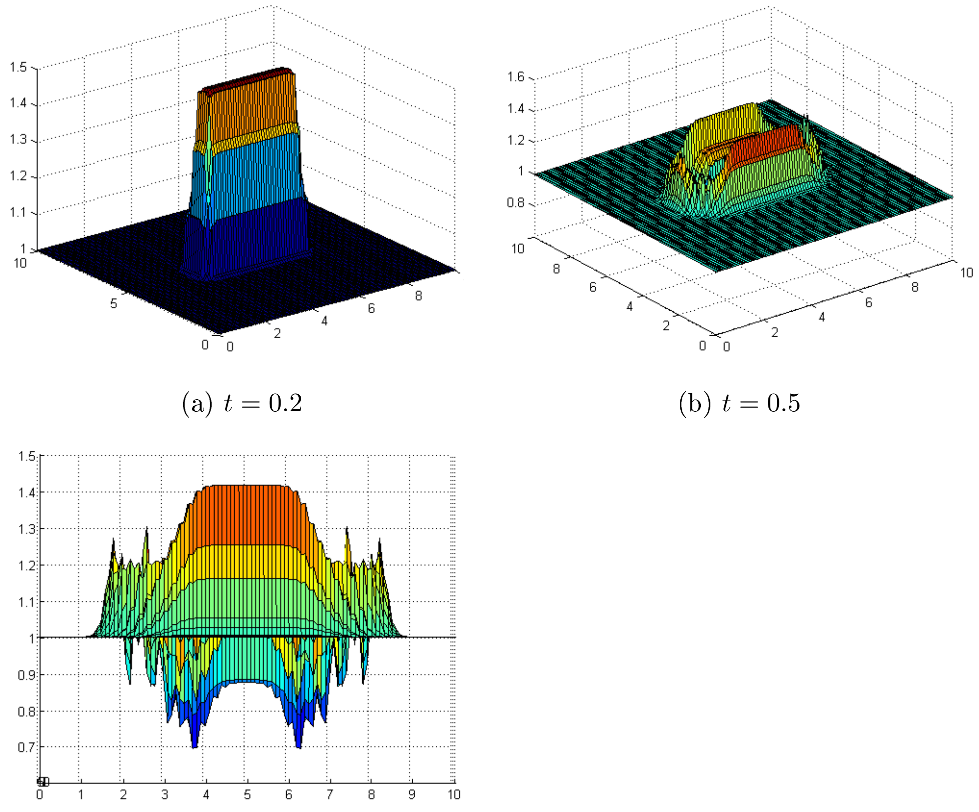
Δx	L_1 error of ρ	Order of Accuracy
0.16	0.112761	—
0.08	0.048675	1.212006
0.04	0.019098	1.349730
0.02	0.005593	1.771820
0.01	0.004954	0.175025

Table 3: Order of Accuracy for the EEC Scheme for test problem (3.18).

Table 3 shows the order of accuracy that was computed with the EEC scheme for test problem (3.18) at time $t = 0.5$. We see that the order of convergence is approaching $k = 2$, but as the Δx gets smaller, MATLAB has difficulties correctly computing the L_1error, and we get up with a $k \approx 0$. When the accuracy is of second order, it is hard to compute solutions and L_1-errors that are accurate enough to get a good measure of k. Still, the pattern was showing that k is tending towards 2.

Finally, we compute the solution to the two-dimensional test problem

$$\begin{aligned} \rho(x,0) &= \begin{cases} 2 & \text{if } 3 < x < 7,\ 4.5 < y < 5.5 \\ 1 & \text{elsewhere} \end{cases}, \\ w(x,0) &= 0, \\ \omega(x,0) &= 0. \end{aligned} \tag{3.19}$$



(a) $t = 0.2$

(b) $t = 0.5$

(c) $t = 0.8$

Figure 6: The density ρ from solutions of (3.19), using the two-dimensional EEC scheme.

The solution is plotted in Figure 6. As we can see, the solution behaves similarly to the one-dimensional scheme. In Section 3.8, we present a solution to the same problem with an energy stable scheme.

3.6 Numerical Viscosity

The definition of an energy stable method is, as was discussed in Section 3.3, one that satisfies the entropy inequality (3.4), that is

$$\frac{d}{dt}\eta(U_i(t)) + \frac{1}{\Delta x}\left(Q_{i+1/2} - Q_{i-1/2}\right) \leq 0.$$

As discussed first in Section 1.3 and then later in Section 1.4.3, we need our energy to be dissipated at the shocks. The EEC-method is energy conservative, so there is no dissipation. We wish to find some diffusion operator to add to our existing scheme, so that our new scheme will be on the form $F_{i+1/2} = \tilde{F}_{i+1/2} - D_{i+1/2}\llbracket V \rrbracket_{i+1/2}$, where $D_{i+1/2}\llbracket V \rrbracket_{i+1/2}$ represents our numerical viscosity. The requirements for entropy stability was stated by Tadmor[18] in the following theorem, together with an entropy dissipation estimate.

Theorem 3.5. (Tadmor[18]). *Assume that we have a positive semi-definite matrix D so that we get*

$$\llbracket V \rrbracket^T_{i+1/2} D_{i+1/2} \llbracket V \rrbracket_{i+1/2} \geq 0 \quad \forall V_i, V_{i+1}. \tag{3.20}$$

A conservative scheme which has more numerical viscosity than an entropy conservative one is entropy stable. That is, if an entropy conservative scheme has flux $\tilde{F}_{i+1/2}$, then the flux of the entropy stable scheme will be given by $F_{i+1/2} = \tilde{F}_{i+1/2} - D_{i+1/2}\llbracket V \rrbracket_{i+1/2}$. Furthermore, an entropy stable scheme with numerical flux F has the entropy dissipation estimate

$$\begin{aligned}\frac{d}{dt}\eta(U_i) + \frac{1}{\Delta x}\left(Q_{i+1/2} - Q_{i-1/2}\right) \;=\; & -\frac{1}{4\Delta x}\llbracket V \rrbracket^T_{i+1/2} D_{i+1/2}\llbracket V \rrbracket_{i+1/2} \\ & -\frac{1}{4\Delta x}\llbracket V \rrbracket^T_{i-1/2} D_{i-1/2}\llbracket V \rrbracket_{i-1/2} \leq 0\end{aligned}$$

Here

$$Q_{i+1/2} = \tilde{Q}_{i+1/2} - \frac{1}{2}\overline{V}^T_{i+1/2} D_{i+1/2}\llbracket V \rrbracket_{i+1/2},$$

where $\tilde{Q}$ is the numerical entropy flux (3.8) and Q is consistent with q .

Proof. Let $\tilde{F}_{i+1/2}$ be a entropy conserving flux, that is

$$\llbracket V \rrbracket_{i+1/2}\tilde{F}_{i+1/2} = \llbracket \psi \rrbracket_{i+1/2}.$$

By definition

$$F_{i+1/2} = \tilde{F}_{i+1/2} - D_{i+1/2} \llbracket V \rrbracket_{i+1/2}, \quad D_{i+1/2} \geq 0.$$

$$\begin{aligned} 0 &= \frac{dU_i}{dt} + \frac{F_{i+1/2} - F_{i-1/2}}{\Delta x} \\ &= \frac{dU_i}{dt} + \frac{\tilde{F}_{i+1/2} - \tilde{F}_{i-1/2}}{\Delta x} \\ &\quad - \left(\frac{D_{i+1/2} \llbracket V \rrbracket_{i+1/2} - D_{i-1/2} \llbracket V \rrbracket_{i-1/2}}{\Delta x} \right) \end{aligned} \tag{3.21}$$

Multiplying with $\eta'(U_i) = V_i$ by (3.21)

$$\begin{aligned} 0 &= \frac{d}{dt}\eta(U_i) + \frac{\tilde{Q}_{i+1/2} - \tilde{Q}_{i-1/2}}{\Delta x} - \frac{V_i D_{i+1/2} \llbracket V \rrbracket_{i+1/2}}{2\Delta x} + \frac{V_i D_{i-1/2} \llbracket V \rrbracket_{i-1/2}}{2\Delta x} \\ &= \frac{d}{dt}\eta(U_i) + \frac{\tilde{Q}_{i+1/2} - \tilde{Q}_{i-1/2}}{\Delta x} - \frac{\bar{V}^T_{i+1/2} D_{i+1/2} \llbracket V \rrbracket_{i+1/2} - \bar{V}^T_{i-1/2} D_{i-1/2} \llbracket V \rrbracket_{i-1/2}}{2\Delta x} \\ &\quad + \frac{1}{4\Delta x} \left(\llbracket V \rrbracket^T_{i-1/2} D_{i+1/2} \llbracket V \rrbracket_{i+1/2} - \llbracket V \rrbracket^T_{i-1/2} D_{i-1/2} \llbracket V \rrbracket_{i-1/2} \right) \end{aligned}$$

If we then let $Q_{i+1/2} = \tilde{Q}_{i+1/2} - \frac{1}{4}\bar{V}^T_{i+1/2} D_{i+1/2} \llbracket V \rrbracket_{i+1/2}$, we have

$$\begin{aligned} 0 &= \frac{d}{dt}\eta(U_i) + \frac{Q_{i+1/2} - Q_{i-1/2}}{\Delta x} \\ &\quad + \frac{1}{2\Delta x} \left(\bar{V}^T_{i+1/2} D_{i+1/2} \llbracket V \rrbracket_{i+1/2} - \bar{V}^T_{i-1/2} D_{i-1/2} \llbracket V \rrbracket_{i-1/2} \right) \\ &\geq \frac{d}{dt}\eta(U_i) + \frac{Q_{i+1/2} - Q_{i-1/2}}{\Delta x} \end{aligned}$$

□

So if (3.20) is satisfied, then F also satisfies the entropy inequality and we can find a stable scheme. If we can find the right matrix D, we can create a modification to the EEC method from Section (3.4) that satisfies the entropy inequality.

3.7 Adding the diffusion operator

The construction of stable diffusion operators is not trivial. One choice would be to use the Rusanov diffusion operator

$$\left|s_{i+1/2}\right| I \left[\!\left[U\right]\!\right]_{i+1/2},$$

where $s_{i+1/2} = \max_k \left\{\left|\lambda_k(U_i)\right|, \left|\lambda_k(U_{i+1})\right|\right\}$ is the same as before. The diffusion operator that we will use with our method is a modified version of the diffusion operators from Roe's method. The diffusion operators are modified so that we get entropy stable operators.

3.7.1 Roe's diffusion operator

Since $\partial u = \frac{\partial u}{\partial v}\partial v$ by the chain rule, we have that

$$\left[\!\left[U\right]\!\right]_{i+1/2} = \left[U_V\right]_{i+1/2} \left[\!\left[V\right]\!\right]_{i+1/2}, \tag{3.22}$$

where $\left[U_V\right]_{i+1/2}$ is the the matrix given by $\frac{\partial u}{\partial v}$ and evaluated in $U_{i+1/2}$. Recall from (3.12) that the entropy variables for the isentropic Euler equations are

$$v = \begin{bmatrix} \kappa\frac{\gamma}{(\gamma-1)}\rho^{\gamma-1} - \frac{1}{2}\frac{m^2}{\rho^2} \\ \frac{m}{\rho} \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}.$$

Writing u in terms of v, we have

$$u(v) = \begin{bmatrix} \rho \\ \rho u \end{bmatrix} = \begin{bmatrix} \left(\frac{v_1 + \frac{v_2^2}{2}}{\Omega}\right)^{\beta} \\ v_2\left(\frac{v_1 + \frac{v_2^2}{2}}{\Omega}\right)^{\beta} \end{bmatrix}.$$

with $\Omega = \frac{\kappa\gamma}{\gamma-1}$ and $\beta = \frac{1}{\gamma-1}$. We can now calculate the change-of-variables matrix

$$U_V = \frac{\rho^{2-\gamma}}{\gamma\kappa}\begin{bmatrix} 1 & w \\ w & w^2 + \kappa\gamma\rho^{\gamma-1} \end{bmatrix}.$$

According to the eigenvector scaling theorem and the results in Barth[1], we can choose R to be the scaled matrix of eigenvectors of f such that $RR^T = U_v$. We see that we must then have

$$R = \sqrt{\frac{\rho^{2-\gamma}}{2\gamma\kappa}}\begin{bmatrix} 1 & 1 \\ w - \sqrt{\kappa\gamma\rho^{\gamma-1}} & w + \sqrt{\kappa\gamma\rho^{\gamma-1}} \end{bmatrix}.$$

Then, if we use Roe's diffusion operator from (2.7) and (3.22), we get

$$R\left|\Lambda\right| R^{-1} \left[\!\left[U\right]\!\right]_{i+1/2} = R\left|\Lambda\right| R^{T} \left[\!\left[V\right]\!\right]_{i+1/2},$$

with

$$|\Lambda| = \begin{bmatrix} w - \sqrt{\kappa\gamma\rho^{\gamma-1}} & 0 \\ 0 & w + \sqrt{\kappa\gamma\rho^{\gamma-1}} \end{bmatrix}.$$

Letting $a = w - \sqrt{\kappa\gamma\rho^{\gamma-1}}$ and $b = w + \sqrt{\kappa\gamma\rho^{\gamma-1}}$, this matrix can be directly calculated to be

$$R\,|\Lambda|\,R^T = \begin{bmatrix} |a| + |b| & |a|\,a + |b|\,b \\ |a|\,a + |b|\,b & |a|\,a^2 + |b|\,b^2 \end{bmatrix}.$$

Note that by evaluating U_V in the arithmetic average from (1.3) instead of the Roe average (2.8), the shock resolving properties of Roe's method is lost. We call the new scheme given by $\hat{F}_{i+1/2} = \tilde{F}_{i+1/2} - R\,|\Lambda|\,R^T\,[\![V]\!]_{i+1/2}$ the *Entropy Stable with Roe Diffusion operator* scheme, abbreviated as *ESRD.*

3.7.2 Two-dimensional diffusion operator

The two-dimensional operator is constructed in an analog manner to the one dimensional case. We proceed to construct the modified version of Roe's diffusion operator. The entropy variables are given by (3.15),

$$v = \begin{bmatrix} \kappa\frac{\gamma}{(\gamma-1)}\rho^{\gamma-1} - \frac{1}{2}\frac{m^2}{\rho^2} \\ w \\ \omega \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}.$$

$u(v)$ is also similarly given as

$$u(v) = \begin{bmatrix} \rho \\ \rho w \\ \rho\omega \end{bmatrix} = \begin{bmatrix} \left(\frac{v_1 + \frac{v_2^2}{2} + \frac{v_3^2}{2}}{\Omega}\right)^{\beta} \\ v_2\left(\frac{v_1 + \frac{v_2^2}{2} + \frac{v_3^2}{2}}{\Omega}\right)^{\beta} \\ v_3\left(\frac{v_1 + \frac{v_2^2}{2} + \frac{v_3^2}{2}}{\Omega}\right)^{\beta} \end{bmatrix}$$

with Ω as before. The matrix U_V is calculated to be

$$U_v = \frac{\rho^{2-\gamma}}{\gamma\kappa}\begin{bmatrix} 1 & w & \omega \\ w & w^2 + \Upsilon & w\omega \\ \omega & w\omega & \omega^2 + \Upsilon \end{bmatrix},$$

with $\Upsilon = \kappa\gamma\rho^{\gamma-1}$. We then find the scaled matrix of eigenvectors of f and g from (1.29) such that $R^x\left(R^x\right)^T = R^y\left(R^y\right)^T = U_V$. They are

$$R^x = \sqrt{\frac{\rho^{2-\gamma}}{2\gamma\kappa}}\begin{bmatrix} 1 & 0 & 1 \\ w-\sqrt{\Upsilon} & 0 & w+\sqrt{\Upsilon} \\ \omega & \sqrt{2\Upsilon} & \omega \end{bmatrix},$$

$$R^y = \sqrt{\frac{\rho^{2-\gamma}}{2\gamma\kappa}}\begin{bmatrix} 1 & 0 & 1 \\ w & -\sqrt{2\Upsilon} & w \\ \omega-\sqrt{\Upsilon} & 0 & \omega+\sqrt{\Upsilon} \end{bmatrix}.$$

The diagonal matrix consisting of the eigenvalues of f is

$$\left|\Lambda^x\right| = \begin{bmatrix} \left|w-\sqrt{\Upsilon}\right| & 0 & 0 \\ 0 & w & 0 \\ 0 & 0 & w+\sqrt{\Upsilon} \end{bmatrix},$$

and similarly for g. Then the calculation is straight forward, and as described in (2.3) of Section 2.1 and Section 2.1.3. The matrices $R^x\left|\Lambda^x\right|\left(R^x\right)^T$ and $R^y\left|\Lambda^y\right|\left(R^y\right)^T$ are calculated in a direct manner, like its one-dimensional counterpart.

3.7.3 Scheme Analysis

Theorem 3.6. *The ESRD scheme is consistent with the isentropic Euler equations (1.25). It is energy stable and it satisfies the energy dissipation estimate*

$$\begin{aligned}\frac{d}{dt}\eta(Ui) + \frac{1}{\Delta x}\left(Q_{i+1/2} - Q_{i-1/2}\right) &= \frac{-1}{4\Delta x}\left(\llbracket V\rrbracket^T_{i+1/2}\left(R|\Lambda|R^T\right)_{i+1/2}\llbracket V\rrbracket_{i+1/2}\right) \\ &\quad \frac{-1}{4\Delta x}\left(\llbracket V\rrbracket^T_{i-1/2}\left(R|\Lambda|R^T\right)_{i-1/2}\llbracket V\rrbracket_{i-1/2}\right) \\ &\leq 0,\end{aligned}$$

where $Q_{i+1/2} = \tilde{Q}_{i+1/2} - \frac{1}{2}\overline{V}^T_{i+1/2}R|\Lambda|R^T\llbracket V\rrbracket_{i+1/2}$ *is consistent with* Q. $\tilde{Q}$ *is (3.8), the energy flux of the EEC scheme. The scheme is first-order accurate.*

Proof. The difference between the numerical viscosity matrices of the energy preserving EEC scheme and the ESRD scheme is $D_{i+1/2} = P_{i+1/2} - \tilde{P}_{i+1/2} = R|\Lambda|R^T$. This is a positive symmetric matrix, and hence satisfies the stability criterion (3.20) for all V_i and V_{i+1}. The energy dissipation estimate then follows from Theorem 3.5 □

3.8 Numerical Experiments

We continue testing on a problem similar to (3.16),

$$\begin{aligned} \rho(x,0) &= \begin{cases} 2 & \text{if } 3 < x < 5 \\ 1 & \text{elsewhere} \end{cases}, \\ w(x,0) &= 0. \end{aligned} \tag{3.23}$$

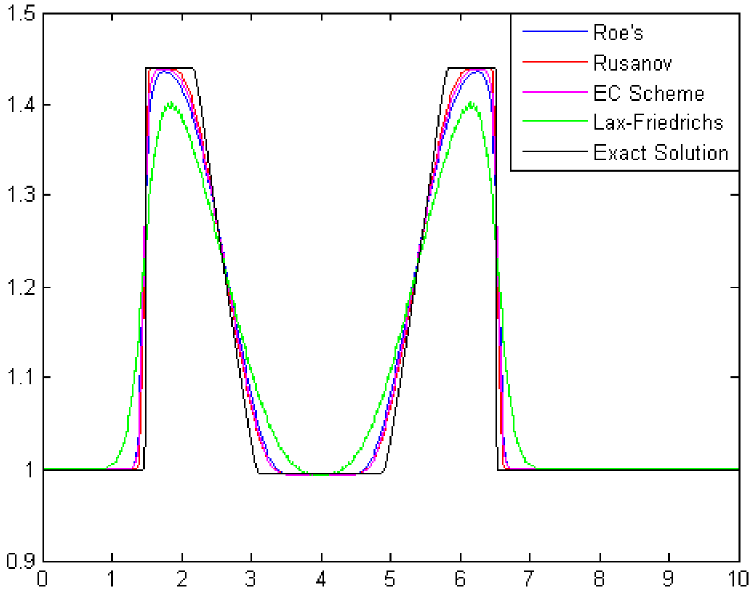
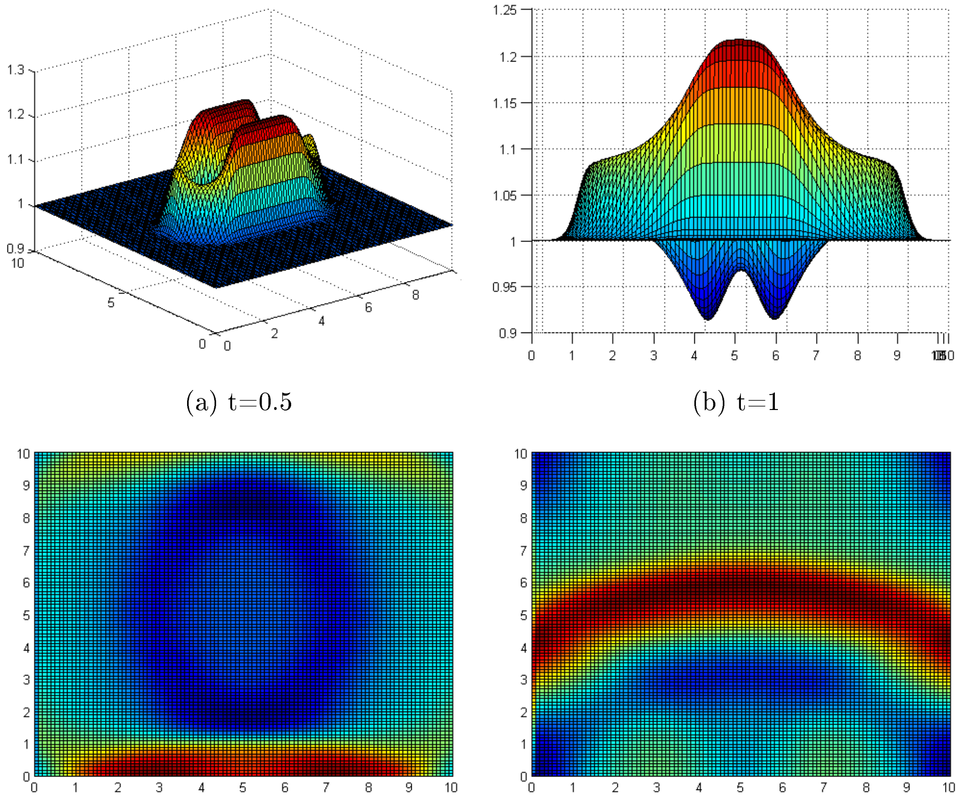


Figure 7: The density ρ from solutions of the ESRD scheme plotted together with several common schemes and a reference solution.

We test the stability of our scheme. Figure 7 shows the ESRD scheme compared with popularly used volume schemes, as well as a high resolution reference solution.

We can see that Rusanov's scheme as well as the ESRD scheme are the schemes that most closely approximate the solution, with Roe's scheme following closely behind. Lax-Friedrichs scheme does not approximate wave speeds locally, and we see that it is quite diffusive.



(a) t=0.5

(b) t=1

(c) t=3

(d) t=6

Figure 8: The density ρ from solutions of (3.19), using the the ESRD scheme. We have used the reflective boundary conditions that $\omega \to -\omega$ at the boundary $y = 0$. That is, $y = 0$ is acting as if it was a solid wall.

In Figure 8, we plotted the solution of the two dimensional problem (3.19) solved by the ESRD scheme. We observe that we don't get growing oscillations like in Figure (6). The method is behaving as expected.

Δx	L_1 error of ρ	Order of Accuracy
0.16	0.415281	—
0.08	0.234966	0.821639
0.04	0.126008	0.898930
0.02	0.069927	0.849602
0.01	0.037054	0.916208
0.005	0.018034	1.038915
0.0025	0.009134	0.981343
0.00125	0.004237	1.108375
0.000625	0.002110	1.005735

Table 4: Order of Accuracy for the ESRD Scheme for test problem (3.18).

The ESRD scheme should only be first-order accurate, because of the added Roe diffusion operator. We examine the rate of convergence for the ESRD scheme by using the same test problem and conditions as in Section 3.5, that is, (3.18) at $t = 0.5$. Table 4 shows that the rate of convergence quickly tend towards a rate $k \approx 1$. As the ESRD method is only first-order accurate, the order of accuracy was much simpler to accurately compute than for the EEC method. This is because small inaccuracies in the real computed solution matters less than it would when we have a second-order accurate method.

3.9 Conclusion

We developed an entropy conservative and entropy stable finite volume scheme for the isentropic Euler schemes for polytropic, ideal gases with any degrees of freedom α. The schemes were developed for one space dimension, and the generalized to account for the two-dimensional case. Numerical experiments showed that the EEC scheme has a decent computational cost, even though it's higher than that of the popularly used methods. It is also second order accurate. The ESRD scheme adds dissipation around shocks to avoid growing oscillations, and is stable, but only first-order accurate.

Further possible work to be done includes making the stable method by using reconstruction[12] to make the method second-order accurate. One could also investigate the positivity preservation of the density variable ρ.

References

[1] T. J. Barth. Numerical methods for gasdynamic systems on unstructured meshes. *An introduction to recent developments in theory and numerics for conservation laws*, 1997.

[2] Stefano Bianchini and Alberto Bressan. Vanishing viscosity solutions of nonlinear hyperbolic systems. *Annals of Mathematics*, 161:223–342, 2005.

[3] Gui-Qiang Chen. Euler equations and related hyperbolic conservation laws. *Handbook of Differential Equations*, 2:1–104, 2006.

[4] R. Courant and K. O. Friedrichs. *Supersonic Flow and Shock Waves.* Springer, 1948.

[5] M. Fey. Multidimensional upwinding i. the method of transport for solving the euler equations. *J. Comput. Phys.*, 143:159–180, 1998.

[6] M. Fey. Multidimensional upwinding ii. decomposition of the euler equations into advection equations. *J. Comput. Phys.*, 143:181–199, 1998.

[7] K.O. Friedrichs and P.D. Lax. Systems of conservation laws with a convex extention. *Proc. Nat. Acad. Sci U.S.A.*, 49:357–393, 1971.

[8] Sergei Godunov. A difference scheme for numerical solution of discontinuous solution of hydrodynamic equations. *Math. Sbornik*, 47:271–306, 1959.

[9] T. Kato. Remarks on zero viscosity limit for nonstationary navier-stokes flows with boundary. *Math. Sci. Res. Inst. Publ.*, 2:85–98, 1984.

[10] Randall J. LeVeque. *Numerical Methods for Conservation Laws.* Birkhauser Verlag, 1992.

[11] Randall J LeVeque. *Finite Volume Methods for Hyperbolic Problems.* Cambridge University Press, 2002.

[12] W. Li. *Efficient Implementation of High-Order Accurate Numerical Methods on Unstructured Grids.* Springer-Verlag, 2014.

[13] Katate Masatsuka. *I Do Like CFD.* Katate Masatsuka, 2013.

[14] S. Osher. Riemann solvers, the entropy condition, and difference approximations. *SIAM J. Numer. Anal.*, 21:217–235, 1984.

[15] Philip L. Roe. Approximate riemann solvers, parameter vectors, and difference schemes. *Journal of Computational physics*, 43(2):357–372, 1981.

[16] E. Tadmor. Numerical viscosity and entropy conditions for conservative difference schemes. *Math. Comp.*, 168:369–381, 1984.

[17] E. Tadmor. Entropy functions for symmetric systems of conservation laws. *J. Math. Anal. Appl.*, 121:355–359, 1987.

[18] E. Tadmor. The numerical viscosity of entropy stable schemes for systems of conservation laws. *Math. Comp.*, 49(I):91–103, 1987.