# Numerical Hydrodynamics
# or
# Computational Fluid Dynamics

**P.W. Hemker**

Lecture notes by
D. Benden, P. Brouwer, R. Mirani, A. Schrijvers

# Contents

CHAPTER 1

# Physical Background

## 1. Introduction

In hydrodynamics we study the flow of fluids. Fluids are divided into liquids, gases and plasma. A liquid keeps its volume but not its shape, a flowing gas or plasma can change both shape and volume.

The difference between a gas and a plasma is that a gas consists of neutral atoms or molecules while a plasma consists of, usually because of high temperatures, electrically charged particles, like ions or electrons.The hydrodynamics of plasma (magneto-hydrodynamics) is not discussed in the numerical part of this class.

In hydrodynamics we don't study the behavior of individual particles,but the behavior of all particles in a volume element. Therefore it is important that a physical small volume element still contains many particles, in that case the variation of the amount of particles is relatively small. This is formulated in the following relation

$$|\,N - \langle N\rangle\,| \ll \langle N\rangle\,.$$

Now we can define useful statistical variables of state. A flow is characterised by its velocityfield $\vec{v}(\vec{x}, t)$ and at least two state variables, for example pressure and density. Thus, we can describe a flow in 3 dimensional space by five state variables, for example mass, energy and the 3 components of the momentum. Five hydrodynamic equations correspond with these five state variables.

CHAPTER 2

# Equations of fluid flow

## 1. Conservation of mass

The equation for the conservation of mass, can be expressed in a Lagrangian coordinate system, and in a Eulerian coordinate system. The first one is a coordinate system that moves along with the flow. The equation of the conservation of mass is then:

$$\frac{dM}{dt} = 0. \tag{1}$$

The Eulerian coordinate system is a coordinate system with respect to a given point or volume in space. Material is moving with respect to this coordinate system. The equation of the conservation of mass then becomes:

$$\frac{\partial}{\partial t}\int_V \rho dV = -\oint_S \rho\vec{v}\cdot\vec{n}dS. \tag{2}$$

With the help of the Divergence Theorem, the right-handside of equation (2) can be rewritten as:

$$-\oint_S \rho\vec{v}\cdot\vec{n}dS = -\int_V \mathrm{div}(\rho\vec{v})dV. \tag{3}$$

The left side of (2) and the right side of (3) written in differentials gives the Continuity Equation:

$$\frac{\partial\rho}{\partial t} + \mathrm{div}(\rho\vec{v}) = 0. \tag{4}$$

## 2. Conservation of momentum

In this section, the following conventions will be used:

$$\nabla\varphi = \nabla_j\varphi = \begin{pmatrix} \varphi_x \\ \varphi_y \\ \varphi_z \end{pmatrix}$$

and

$$\mathrm{div}\,\vec{w} = \nabla\cdot\vec{w} = \nabla_j w_j = \frac{\partial}{\partial x}w_x + \frac{\partial}{\partial y}w_y + \frac{\partial}{\partial z}w_z.$$

The notation $\nabla_j w_j$is called the Einstein summation convention.
The equations for the conservation of momentum in the Langrangian coordinate system are:

$$\frac{d}{dt}(m\vec{v}_i) = \vec{F}_i. \tag{5}$$

In the Eulerian coordinate sytem, this becomes:

$$\frac{\partial}{\partial t}\int_V (\rho\vec{v}_i)dV + \oint_S (\rho\vec{v}_i)\vec{v}_j \cdot \vec{n}_j dS = \int_V \rho\vec{f}_i dV + \oint_S \vec{\vec{\sigma}}_{ij} \cdot \vec{n}_j dS, \tag{6}$$

with $\sigma_{ij} = -P\delta_{ij} + \tau_{ij}$, which is called the stress tensor. The first part , $-P\delta_{ij}$, is the pressure part, with $\delta_{ij}$ the Kronecker delta, because the pressure is the same in all directions. The second part, $\tau_{ij}$, is the proper stress tensor, which vanishes for a frictionless fluid. The proper stress tensor is symmetric, so also $\sigma_{ij}$ is symmetric. In equation (6) , $\int_V \rho\vec{f}_i dV$ describes the force on the inside, and $\oint_S \vec{\vec{\sigma}}_{ij} \cdot \vec{n}_j dS$ the force on the surface of the volume. Again using the Divergence Theorem , and assume that $\vec{f}_i = 0$ , equation (6) becomes:

$$\frac{\partial}{\partial t}(\rho v_i) + \nabla_j(\rho v_i v_j + \delta_{ij}P) = \nabla_k \tau_{ik} \tag{7}$$

These equations are known as the Navier-Stokes equations. A 'Newtonian' fluid is characterised by a linear relation between the stress and the deformation :

$$\tau_{ij} = \mu(\nabla_i v_j + \nabla_j v_i) + \delta_{ij}\lambda\nabla_k v_k \tag{8}$$

with $\mu$ and $\lambda$ the coefficients of viscosity.
$\nabla_i v_j$ in (8) is called the deformation rate tensor. It can be written in a symmetric part, $e_{ij}$, and an antisymmetric part, $\omega_{ij}$ ,as follows:

$$\nabla_i v_j = \underbrace{\frac{1}{2}(\nabla_j v_i + \nabla_i v_j)}_{e_{ij}} + \underbrace{\frac{1}{2}(\nabla_i v_j - \nabla_j v_i)}_{\omega_{ij}} \tag{9}$$

$e_{ij}$is the strain rate tensor or deformation tensor, and $\omega_{ij}$ the rotation tensor. The term $\omega_{ij}$ does not appear in (8), because rotation has no influence on the force on the surface. The strain rate tensor can be divided futher in a shear rate part ( the off-diagonal elements), and a normal strain part ( the diagonal elements). Because $e_{ij}$ is symmetric, it can be written in diagonal form, in the eigenvector coordinate system. Along the eigenvectors, the fluid expands or shrinks, depending in the eigenvalues belonging to the eigenvectors. Eigenvalues bigger than 1 imply expansion, and vica versa.

## 3. Conservation of energy

In the Lagrangian coordinate system, the equation of the conservation of energy is:

$$\frac{d}{dt}E = \frac{d}{dt}Q + \frac{d}{dt}W, \tag{10}$$

where Q represents heat, and W mechanical energy. In the Eulerian coordinate system this equation becomes:

$$\frac{\partial}{\partial t}\int_V (\rho e)dV + \oint_S (\rho e)\vec{v}_j\vec{n}_j dS = -\oint_S \vec{q}_j\vec{n}_j dS + \oint_S \vec{v}_i\vec{\vec{\sigma}}_{ij}\vec{n}_j + \int_V \tilde{q}dV \tag{11}$$

where $\frac{d}{dt}Q = -\oint_S \vec{q}_j\vec{n}_j dS$ represents the heat that flows into the volume, $\vec{q}$ is the heat flux, and $\tilde{q}$ is the heat generation within the volume. $\frac{d}{dt}W = \oint_S \vec{v}_i\sigma_{ij}\vec{n}_j$ represents mechanical energy: the force exerted on the surface of the volume.

Using the Divergence Theorem, equation (11) can be written in differential form as:

$$\frac{\partial}{\partial t}(\rho e) + \nabla_j(\rho e v_j) = -\nabla_j q_j + \nabla_j(v_i \sigma_{ij}) + \tilde{q}. \tag{12}$$

Fouriers Law represents the relation between the heat flux and the temperature:

$$\vec{q_j} = -\kappa \nabla_i T \tag{13}$$

In this equation, $\kappa$ is the heat conduction coefficient, and $\nabla_i T$ the temperature gradient.

## 4. The system of equations

Combining the differential forms for the conservation laws in sections 1.1-1.3, we find the following system of equations:

$$\left\{ \begin{array}{lll} \frac{\partial}{\partial t}(\rho) & + \nabla_j(\rho v_j & ) = 0 \\ \frac{\partial}{\partial t}(\rho v_i) & + \nabla_j(\rho v_i v_j - \sigma_{ij} & ) = \rho f_i \\ \frac{\partial}{\partial t}(\rho e) & + \nabla_j(\rho e v_j \ - v_k \sigma_{kj} + q_j & ) = \tilde{q} \end{array} \right. \tag{14}$$

Noting that $\sigma_{ij} = -P\delta_{ij} + \tau_{ij}$, and assuming that $f_i = 0$ and $\tilde{q} = 0$, equation (14) can be rewritten as:

$$\left\{ \begin{array}{lll} \frac{\partial}{\partial t}(\rho) & + \nabla_j(\rho v_j & ) = 0 \\ \frac{\partial}{\partial t}(\rho v_i) & + \nabla_j(\rho v_i v_j + \delta_{ij} P & ) = \nabla_j \tau_{ij} \\ \frac{\partial}{\partial t}(\rho e) & + \nabla_j(\rho(e + P/\rho) v_j & ) = \nabla_j(v_k \tau_{kj} - q_j) \end{array} \right. \tag{15}$$

In this equation, the div-terms in the lhs determine the convective flux, and the div-terms in the rhs the diffusive flux.
With $\tau_{ij}$ replaced by (8) and $q_j$ by (13), the equations in (15) can be written as:

$$\left\{ \begin{array}{lll} \frac{\partial}{\partial t}(\rho) & + \nabla_j(\rho v_j & ) = 0 \\ \frac{\partial}{\partial t}(\rho v_i) & + \nabla_j(\rho v_i v_j + \delta_{ij} P & ) = \nabla_j[\mu(\nabla_i v_j + \nabla_j v_i) + \delta_{ij}\lambda \nabla_k v_k] \\ \frac{\partial}{\partial t}(\rho e) & + \nabla_j(\rho(e + P/\rho) v_j & ) = \nabla_j(v_k \tau_{kj} + \kappa \nabla_j T) \end{array} \right. \tag{16}$$

In the above equations the unknown variables are $\rho, v_x, v_y, v_z, e, P$ and $T$. In order to close the above system, we need two independant additional relations between these unknowns. For a perfect gas we have the following relations where $T$ and $P$ can be written as functions of $e$, $v^2$ and $\rho$:

$$\begin{array}{l} T = \frac{1}{C_v}(e - \frac{1}{2}v^2) \\ P = (\gamma - 1)\rho(e - \frac{1}{2}v^2) \end{array}$$

In general , this results in a difficult, nonlinear system of equations. There are some special cases, in which simplifications can be made. In the next section, some of these special cases will be discussed.

## 5. Special cases

**5.1. Incompressible flow.** When a fluid is incompressible, for example water, it means that $\rho$ is constant in time. When $\frac{\delta}{\delta t}\rho = 0$, also $\nabla_j v_j = 0$, according to the first equation in (15). Or in words, the divergence of $\vec{v}$ is zero for an incompressible fluid. The system of equations then becomes:

$$\left\{ \begin{array}{l} \nabla_j v_j = 0 \\ \rho \frac{\partial}{\partial t}(v_i) + \rho v_j \nabla_j v_i + \nabla_j \delta_{ij} P = \nabla_j \tau_{ij} \\ \rho \frac{\partial}{\partial t}(e) + \rho \nabla_j(e) + v_j \nabla_j P = \nabla_j(v_k \tau_{kj} - q_j) \end{array} \right. \tag{17}$$

**5.2. Euler equations for inviscid flow.** Now assume that the fluid is heat conduction free, and viscosity free. Heat conduction free means $\kappa = 0$, and so also $q_j = -\kappa \nabla_j T = 0$. Viscosity free means that $\lambda = \mu = 0$. When $\lambda$ and $\mu$ are zero, so also will be $\tau_{ij}$. With $\tau_{ij} = 0$ and $q_j = 0$, the equations in (15) become:

$$\left\{ \begin{array}{lll} \frac{\partial}{\partial t}(\rho) & + \nabla_j(\rho v_j & ) = 0 \\ \frac{\partial}{\partial t}(\rho v_i) & + \nabla_j(\rho v_i v_j + \delta_{ij} P & ) = 0 \\ \frac{\partial}{\partial t}(\rho e) & + \nabla_j(\rho(e + P/\rho) v_j) = 0 & \end{array} \right. \tag{18}$$

These equations are called the 'Euler Equations of inviscid flow'.

**5.3. The Stokes equations.** When a fluid has constant viscosity, that means $\nabla_j \mu = 0$ and $\nabla_j \lambda = 0$, then $\nabla_j \tau_{ij}$ in (15) can be written as:

$$\begin{aligned} \nabla_j \tau_{ij} &= \nabla_j(\mu(\nabla_i v_j + \nabla_j v_i)) + \nabla_j \delta_{ij} \lambda \nabla_j v_j \\ &= \mu \nabla_j \nabla_i v_j + \mu \nabla_j \nabla_j v_i + \lambda \nabla_i \nabla_j v_j \\ &= \mu \Delta v_i + (\mu + \lambda) \nabla_i \mathrm{div} \vec{v}. \end{aligned}$$

If the fluid is also incompressible, div $\vec{v} = 0$, and the flow is steady, $\frac{\partial}{\partial t}(\rho v_i) = 0$, then the second equation in (15) becomes:

$$\rho \nabla_j (v_i v_j) + \nabla_i P = \mu \Delta v_i.$$

For a fluid with a low velocity, this equation becomes:

$$\begin{aligned} \mathrm{div} \vec{v} &= 0 \\ \mu \Delta v_i - \nabla_i P &= 0 \end{aligned}$$

These equations are known as the Stokes equations. They are linear systems of equations, $L\vec{u} = 0$. In matrix form, these equations are:

$$\begin{pmatrix} \mu\Delta & 0 & 0 & -\frac{\partial}{\partial x_1} \\ 0 & \mu\Delta & 0 & -\frac{\partial}{\partial x_2} \\ 0 & 0 & \mu\Delta & -\frac{\partial}{\partial x_3} \\ -\frac{\partial}{\partial x_1} & -\frac{\partial}{\partial x_2} & -\frac{\partial}{\partial x_3} & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ P \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

with $\det(L) = (-\nabla_j)(-\nabla_j)(\mu\Delta)^2 = \mu^2 \Delta^3$. This shows that the matrix L is a $6^{\text{th}}$ order elliptic operator.

**5.4. The convection diffusion equation.** Looking at the equations in (14), you can see that there is a lot of similarity between the three equations. When you write the conservation variables as $(\rho, \rho v_i, \rho e) = \vec{u}$, the general structure of the equations can be written as:

$$\frac{\partial}{\partial t}\vec{u} + \mathrm{div}\ J_c(\vec{u}) = \mathrm{div}\ J_d(\vec{u}). \tag{19}$$

In this equation, $\mathrm{div} J_c(\vec{u})$ represents the convective flux, and $\mathrm{div} J_d(\vec{u})$ the diffusive flux.

The convection diffusion equation is:

$$\frac{\partial}{\partial t}(\rho u) + \mathrm{div}\ (\rho u \vec{v}) = \mathrm{div}\ (d \quad \mathrm{grad} u), \tag{20}$$

in which $u$ is the mass fraction of species in the medium with density $\rho$ and velocity field $\vec{v}$. Another example of a convection diffusion equation is the heat equation, without viscous dissipation:

$$\frac{\partial}{\partial t}u + v_j \nabla_j u = \frac{\kappa}{\rho C_v}\Delta u. \tag{21}$$

CHAPTER 2

# Basic equations

## 1. Conservation laws

We now first formulate three basic conservation laws: conservation of mass, momentum and energy:

**1.1. Conservation of mass.** Suppose we have a flow without creation or annihilation of particles. With $\rho$ the mass density we express the conservation of mass by the equation of continuity

$$\partial_t \rho + \nabla \cdot (\rho \vec{v}) = 0. \tag{22}$$

Integrating this equation over a fixed volume $V$ and using Gauss' theorem, it is easy to see the conservation of mass.

$$\frac{\partial}{\partial t} \int_V \rho \, dV + \int_S \rho \vec{v} \cdot \vec{n} \, dS = 0, \tag{23}$$

with $S = \partial V$ the surface of volume $V$ and $\vec{n}$ the outward unit normal on the surface $S$. Notice from (22) that for an incompressible flow $\nabla \cdot v = 0$.

**1.2. Conservation of momentum.** The acceleration of a fluid element is determined by the forces on this element. Here we have to distinguish between external forces $\vec{f}$ like gravity and internal forces like pressure $p$ on the surface. This brings us to the equation of motion :

$$\partial_t \vec{v} + \vec{v} \cdot \nabla \vec{v} = \rho^{-1} (\vec{f} - \nabla p). \tag{24}$$

Assuming frictionless flow and absence of external forces, so that the pressure is the only force working on a volume, the law of conservation of momentum is given by

$$\rho \left( \partial_t \vec{v} + \vec{v} \cdot \nabla \vec{v} \right) = -\nabla p. \tag{25}$$

Integration of (25) over a fixed volume $V$ gives the equation :

$$\frac{\partial}{\partial t} \int_V \rho \vec{v} \, dV + \int_S \rho \vec{v} (\vec{v} \cdot \vec{n}) \, dS + \int_S p \vec{n} \, dS = 0, \tag{26}$$

with $\vec{n}$ the outward unit normal on the surface $S = \partial V$. Equation (26) shows the change in momentum in volume $V$.

**1.3. Conservation of energy.** Under the assumption of a nonviscous, non-heat-conducting fluid and abscence of body forces, the law of conservation of energy is given by

$$\partial_t(\frac{1}{2}\rho v^2 + \rho e) + \nabla\cdot(\frac{1}{2}\rho v^2 + \rho e + p)\vec{v} = \vec{f}\cdot\vec{v}\,. \tag{27}$$

with $e$ the specific internal energy and $\frac{1}{2}\rho v^2 + \rho e$ the total energy. Again integration over a fixed volume $V$ gives:

$$\frac{\partial}{\partial t}\int_V (\frac{1}{2}\rho v^2 + \rho e)dV = -\int_S (\frac{1}{2}\rho v^2 + \rho h)\vec{v}\cdot\vec{n}dS + \int_V \vec{f}\cdot\vec{v}dV\,. \tag{28}$$

with $h = e + \frac{p}{\rho}$ the specific enthalpy. Equation (28) shows us that the total energy in volume $V$ changes due to energy exchange over surface $S$ and due to external work, like gravity.

## 2. Non-ideal fluids

**2.1. Law of entropy.** Writing the conservation of energy from (27) without external forces as

$$(\partial_t\frac{1}{2}\rho v^2 + \nabla\cdot(\frac{1}{2}\rho v^2 + p)\vec{v}) + (\partial_t\rho e + \nabla\cdot\rho e\vec{v}) = 0 \tag{29}$$

we see that for ideal (adiabatic) fluids we have

$$\partial_t\rho e + \nabla\cdot(\rho e\vec{v}) = \rho\dot{e} \tag{30}$$

and, using (25),

$$\partial_t(\frac{1}{2}\rho v^2) + \nabla\cdot(\frac{1}{2}\rho v^2 + p)\vec{v} = p\nabla\cdot\vec{v}, \tag{31}$$

such that

$$\rho\dot{e} + p\nabla\cdot\vec{v} = 0. \tag{32}$$

For non-ideal fluids (non adiabatic) we may have heat transport and viscosity. These are included in the first law of hydrodynamics, this is a generalisation of (32):

$$\rho\dot{e} + P : \nabla\vec{v} = -\nabla\cdot\vec{J}_q \tag{33}$$

with $P \equiv p\mathbf{I} + \Pi$ the pressure tensor, $\Pi$ the viscosity tensor and $\vec{J}_q$ the flow of heath. So we find

$$\rho\dot{e} + p\nabla\cdot\vec{v} = \rho T\dot{s} = -\nabla\cdot\vec{J}_q - \Pi : \nabla\vec{v}, \tag{34}$$

with $T$ the temperature and $s$ the specific entropy. Using the second law of thermodynamics:

$$\rho\dot{s} \geq -\nabla\cdot(T^{-1}\vec{J}_q) \tag{35}$$

we find for the entropy production $\sigma = \partial_t(\rho s) + \nabla\cdot(\rho s\vec{v} + T^{-1}\vec{J}_q)$

$$T\sigma = -\vec{J}_q\cdot\frac{\nabla T}{T} - \Pi : \nabla\vec{v}. \tag{36}$$

## 3. Navier-Stokes equations

Newtonian fluids are characterised by a linear relation between the viscous pressure tensor $\Pi$ and the velocity field:

$$\Pi_{ij} = -2\eta \left( \frac{1}{2}(\partial_i v_j + \partial_j v_i) - \frac{1}{3}\partial_k v_k \delta_{ij} \right) - \zeta \partial_k v_k \delta_{ij}, \tag{37}$$

with $\eta$ and $\zeta$ the first and second viscosity coefficients. We can find the second viscosity or volume-viscosity in relativistic gases and systems with internal degrees of freedom. For the classical mono-atomic gas, we have $\zeta = 0$, such that this system has no second viscosity.

In the equation of motion, the divergence of the pressure tensor is an internal force, so we get

$$\rho \dot{\vec{v}} = \vec{f} - \nabla p - \nabla \cdot \Pi. \tag{38}$$

If we assume that the vicosity coefficients do not depend on the spatial coordinates, we can, using (37), write the equation of motion as

$$\rho \dot{\vec{v}} = \vec{f} - \nabla p + \eta \nabla^2 \vec{v} + (\frac{1}{3}\eta + \zeta) \nabla (\nabla \cdot \vec{v}). \tag{39}$$

This equation is called the Navier-Stokes equation. For non-compressible fluids we can write

$$\rho \dot{\vec{v}} = \vec{f} - \nabla p + \eta \nabla^2 \vec{v}\,, \tag{40}$$

and if $\nabla \wedge \vec{v} = 0$, we can write

$$\rho \dot{\vec{v}} = \vec{f} - \nabla p + \mu \nabla^2 \vec{v} \tag{41}$$

with effective viscosity $\mu \equiv \frac{4}{3}\eta + \zeta$.

CHAPTER 3

# Numerical solution by FDM

In the transformation of the PDE into a finite set of equations, it is convenient to distinguish three stages:

- The discretisation of the domain $\Omega$.
- The discretisation of the solution $u$.
- The discretisation of the equation.

For the finite difference method (FDM), the domain $\Omega$ is represented by a finite subset of points $\{x_i\} \subset \Omega$. These points are the so called *nodal points* of the grid.

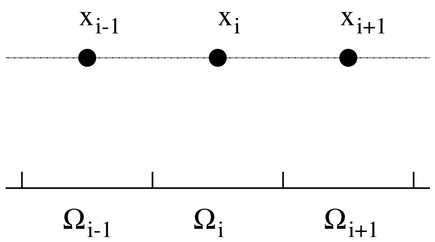



FIGURE 1. Set of nodal points $x_i$ and the intervals $\Omega_i$ around these points.

The function $u$ is represented by a set of function values $u_i$ that approximate $u(x_i)$. The discrete equations are obtained by replacing the differentials in the PDE by finite differences in the discrete equations.

## 1. Analysis of a FDM for a time-dependent problem

We will show and analyse a number of characteristic problems that are encountered in the treatment of time-dependent problems. We do this by means of the simple hyperbolic problem on some domain $\Omega \subset \mathbf{R}$:

$$\left\{ \begin{array}{lcl} \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) & = & 0 \\ u(t_0, x) & = & u_0(x) \end{array} \right. \tag{42}$$

As a grid we take an ordered set of nodal points $\{x_i\}_{i=0}^{N}$. Integration of the PDE in (42) over the interval $\Omega_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ with $x_{i-\frac{1}{2}} = \frac{x_i - x_{i-1}}{2}$ and $x_{i+\frac{1}{2}} = \frac{x_{i+1} - x_i}{2}$, yields for $i = 1, ..., N-1$

$$\frac{\partial}{\partial t} \int_{\Omega_i} u \, d\Omega + f(u(t,x))|_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} = 0. \tag{43}$$

The discrete approximation is represented by $\{u_i\}$, with $u_i = u_i(t)$ an approximation of the mean value of $u$ over the interval $\Omega_i$. So we get

$$\Delta x \, \tfrac{\partial}{\partial t} u_i + f_{i+\frac{1}{2}} - f_{i-\frac{1}{2}} \quad = \quad 0 \tag{44}$$

with $\Delta x = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$. In order to find a system of equations for the unknown $u_i(t)$, we also have to relate $f_{i+\frac{1}{2}}$ with the values $\{u_i\}$. We do this by

$$f_{i+\frac{1}{2}} \quad = \quad w f(u_i) + (1-w) f(u_{i+1}) \tag{45}$$

with $w \in [0,1]$. This gives us a system of ordinary differential equations

$$\frac{\partial}{\partial t} u_i(t) = \frac{w}{\Delta x} f(u_{i-1}) + \frac{1-2\,w}{\Delta x} f(u_i) + \frac{w-1}{\Delta x} f(u_{i+1}) \tag{46}$$

for the unknown $u_i(t)$ with $i = 1, ..., N-1$. We write this system briefly as

$$\frac{\partial}{\partial t} u_h = g(u_h). \tag{47}$$

There are many ways to numerically solve this system. In the next section, we study a few simple (but representative) ones.

**1.1. Backward-, forward Euler and Cranck-Nicolson.** We can realise the time discretisation of (47) simply by time-stepping with steps $\Delta t$. The left side of the equation can now be discretised as

$$\frac{\partial}{\partial t} u_i \;\rightarrow\; \frac{u_i(t^{n+1}) - u_i(t^n)}{\Delta t} \tag{48}$$

It is more complicated to discretise the right side of equation (47). Three popular methods, each using a different discretisation of the right side, are the *forward Euler method*, the *backward Euler method*, and the *Cranck Nicolson method*:

| | |
|---|---|
| Forward Euler | $u(t^{n+1}) \;=\; u(t^n) + \Delta t\, g(u(t^n))$ |
| Backward Euler | $u(t^{n+1}) \;=\; u(t^n) + \Delta t\, g(u(t^{n+1}))$ |
| Cranck-Nicolson | $u(t^{n+1}) \;=\; u(t^n) + \frac{\Delta t}{2}\, \big(g(u(t^n)) + g(u(t^{n+1}))\big)$ |

We combine these three methods in one parametrised formula:

$$u(t^{n+1}) = u(t^n) + \Delta t\,\big((1-\theta) g(u(t^n)) + \theta g(u(t^{n+1}))\big) \tag{49}$$

with $\theta \in [0,1]$. Notice that for $\theta = 0$ we get the *forward Euler method*, for $\theta = 1$ we get the *backward Euler method*, and for $\theta = \frac{1}{2}$, we get the *Cranck-Nicolson method*.

Now, the full discretised system reads

$$\begin{array}{rcccll} u_i^{n+1} & = & u_i^n & + & \frac{\Delta t}{\Delta x}\,\theta & \big(w f_{i-1}^{n+1} + (1-2\,w) f_i^{n+1} + w f_{i+1}^{n+1}\big) \\ & & & + & \frac{\Delta t}{\Delta x}\,(1-\theta) & \big(w f_{i-1}^{n} + (1-2\,w) f_i^{n} + w f_{i+1}^{n}\big) \end{array} \tag{50}$$

with $f_i^n = f(u_i(t^n))$, and two discretisation parameters, $w$ and $\theta$:

$$w \in [0,1] \quad \rightarrow \quad \text{discretisation of } \frac{\partial}{\partial x}$$

$$\theta \in [0,1] \quad \rightarrow \quad \text{discretisation of } \frac{\partial}{\partial t}$$

In general, to compute $\{u_i^{n+1}\}$ from $\{u_i^n\}$, we have to solve a system of nonlinear equations. Because it is difficult to analyse nonlinear equations, we simplify the problem by linearising it. We set $\frac{df}{du} = f'$ constant, and we define $\lambda = f' \frac{\Delta t}{\Delta x}$. Now equation (50) becomes

$$\begin{aligned} u_i^{n+1} = u_i^n &+ \lambda\,\theta \left(w u_{i-1}^{n+1} + (1-2\,w) u_i^{n+1} + w u_{i+1}^{n+1}\right) \\ &+ \lambda\,(1-\theta) \left(w u_{i-1}^{n} + (1-2\,w) u_i^{n} + w u_{i+1}^{n}\right) \end{aligned} \tag{51}$$

The analytic solution for the simplified original problem $u_t + f' u_x = 0$ is simply $\psi(x - f't)$ for an arbitrary initial condition $\psi(x) = u(0,x)$. To analyze the discrete system (51) we neglect the boundary conditions and we consider the problem for $\{u_i\}_{-\infty}^{\infty}$. So we write the initial condition as

$$u_i^0 = u(0, x_i) = \int \hat{u}(\omega) e^{\mathbf{i}\omega(ih)} d\omega \tag{52}$$

for $i = -\infty \ldots \infty$. Since we now have a linear equation, we only have to study a single Fourier mode

$$u_i^0 = e^{\mathbf{i}\omega(ih)} \tag{53}$$

for some arbitrary $\omega$. (An arbitrary initial function is just a linear combination of such Fourier modes.) Now, using equation (51), we can ("Ansatz") write $u_i^n$ as

$$u_i^n = \gamma^n e^{\mathbf{i}\omega(ih)}, \tag{54}$$

where $\gamma$ is the complex number that satisfies

$$\begin{aligned} \gamma - 1 = \;& \lambda\theta \left(w\gamma e^{-\mathbf{i}\omega h} + (1-2w)\gamma - (1-w)\gamma e^{\mathbf{i}\omega h}\right) + \\ & \lambda(1-\theta) \left(w e^{-\mathbf{i}\omega h} + (1-2w) - (1-w) e^{\mathbf{i}\omega h}\right). \end{aligned} \tag{55}$$

We can compute $\gamma$ and find:

$$\gamma = \frac{1 + \lambda(1-\theta)((1-2w) + w e^{-\mathbf{i}\omega h} - (1-w)e^{\mathbf{i}\omega h})}{1 - \lambda\theta((1-2w) + w e^{-\mathbf{i}\omega h} - (1-w)e^{\mathbf{i}\omega h}}). \tag{56}$$

$\gamma$ is the *amplification factor* for the time-integration: $|u_i^n| = |\gamma|^n$. The solution increases for $t \to \infty$ if $|\gamma| > 1$, and it decreases for $|\gamma| < 1$. We see that

$$|\gamma^2| = \frac{(1 + \lambda(1-\theta)(1-2w)(1-\cos(\omega h)))^2 + (\lambda(1-\theta)\sin(\omega h))^2}{(1 - \lambda\theta(1-2w)(1-\cos(\omega h)))^2 + (\lambda\theta\sin(\omega h))^2}. \tag{57}$$

So, we can describe the stability of the method by $|\gamma|$:

$$\begin{array}{ll} |\gamma| < 1 & \text{unconditional stability} \\ |\gamma| = 1 & \text{neutral stability} \\ |\gamma| > 1 & \text{unconditional instability} \end{array} \tag{58}$$

Now we can study the effect of the parameter $w$ (space discretisation) and the parameter $\theta$ (time-discretisation). Furthermore we can see what is the effect of the parameter $\lambda$.

**1.2. Forward, backward and central differencing.** We analyse the stability for three different space discretisations: the central, backward and forward differencing.

- If for central discretisation in space, $w = \frac{1}{2}$, we see that

$$|\gamma|^2 = \frac{1 + (\lambda(1-\theta)\sin(\omega h))^2}{1 + (\lambda\theta\sin(\omega h))^2} \tag{59}$$

and we find

$$\begin{array}{llll} |\gamma| < 1 & \text{if} & \theta > \frac{1}{2} & \text{Backward Euler } (\theta = 1) \\ |\gamma| = 1 & \text{if} & \theta = \frac{1}{2} & \text{Cranck-Nicolson } (\theta = \frac{1}{2}) \\ |\gamma| > 1 & \text{if} & \theta < \frac{1}{2} & \text{Forward Euler } (\theta = 0) \end{array} \tag{60}$$

- If for backward discretisation in space, $w = 1$, we see that

$$|\gamma|^2 = \frac{(1 - \lambda(1-\theta)(1-\cos(\omega h)))^2 + (\lambda(1-\theta)\sin(\omega h))^2}{(1 + \lambda\theta(1-\cos(\omega h)))^2 + (\lambda\theta\sin(\omega h))^2} \tag{61}$$

and we find the condition for stability

$$|\gamma| \leq 1 \quad \text{if and only if} \quad \lambda(\lambda(1-2\theta) - 1) \leq 0 \tag{62}$$

Since $\lambda = f'\frac{\Delta t}{\Delta x}$, we find depending on the sign of $f'$:
  - For $f' > 0$, we get stability if and only if $\lambda(1-2\theta) \leq 1$.
    * If $\theta \geq \frac{1}{2}$, the method is stable.
    * If $\theta < \frac{1}{2}$, we only get stability for $|\lambda|$ small enough (small $\Delta t$).
  - For $f' < 0$, we get stability if and only if $\lambda(1-2\theta) \geq 1$.
    * If $\theta > \frac{1}{2}$, we only get stability for $|\lambda|$ large enough (small $\Delta x$).
    * If $\theta \leq \frac{1}{2}$, the method is instable.

- If for forward discretisation in space, $w = 0$, we see that

$$|\gamma|^2 = \frac{(1 + \lambda(1-\theta)(1-\cos(\omega h)))^2 + (\lambda(1-\theta)\sin(\omega h))^2}{(1 - \lambda\theta(1-\cos(\omega h)))^2 + (\lambda\theta\sin(\omega h))^2} \tag{63}$$

and we find the condition for stability

$$|\gamma| \leq 1 \quad \text{if and only if} \quad \lambda(\lambda(1-2\theta) + 1) \leq 0 \tag{64}$$

Since $\lambda = f'\frac{\Delta t}{\Delta x}$, we find depending on the sign of $f'$:
  - For $f' > 0$, we get stability if and only if $\lambda(1-2\theta) \leq -1$.
    * If $\theta > \frac{1}{2}$, we only get stability for $\lambda$ large enough (small $\Delta x$).
    * If $\theta \leq \frac{1}{2}$, the method is instable.
  - For $f' < 0$, we get stability if and only if $\lambda(1-2\theta) \geq -1$.
    * If $\theta \geq \frac{1}{2}$, the method is stable.
    * If $\theta < \frac{1}{2}$, we only get stability for $\lambda$ small enough (small $\Delta t$).

In the backward and forward discretisation, we recognise a symmetry. If $f' > 0$, then $w = 1$ corresponds with upwind discretisation, and $w = 0$ with downwind discretisation; if $f' < 0$ it is the other way. Conclusion: We notice that the *upwind discretisation* is unconditionally stable when $\theta \geq \frac{1}{2}$, and for $\theta < \frac{1}{2}$ stable if the time-step is small enough. The *downwind discretisation* is instable for all reasonable cases.

This is a general finding: good implicit methods (where linear systems have to be solved in each timestep) can be unconditionally stable; good explicit methods can be stable for small enough time steps.

**1.3. The modified equation.** We consider again the linear problem

$$\left\{ \begin{array}{lcl} \frac{\partial u}{\partial t} + f' \frac{\partial}{\partial x} u & = & 0 \\ u(t_0, x) & = & u_0(x) \end{array} \right. \tag{65}$$

and its discretisation (51). We can write this discretisation also as

$$\begin{array}{lclll} u_i^{n+1} - u_i^n & = & -\lambda\theta & \frac{1}{2} & (u_{i+1}^{n+1} - u_{i-1}^{n+1}) + \\ & & -\lambda\theta & (\frac{1}{2} - w) & (u_{i+1}^{n+1} - 2u_i^{n+1} + u_{i-1}^{n+1}) + \\ & & -\lambda(1-\theta) & \frac{1}{2} & (u_{i+1}^{n} - u_{i-1}^{n}) + \\ & & -\lambda(1-\theta) & (\frac{1}{2} - w) & (u_{i+1}^{n} - 2u_i^{n} + u_{i-1}^{n}). \end{array} \tag{66}$$

In order to compute the local trucation error, we expand the function $u(x,t)$ in a Taylor series around the point $x = x_i$ and $t = t_n + \theta\Delta t$. Substitution of these series in the discretisation yields

$$\frac{\partial u}{\partial t} + f' \frac{\partial}{\partial x} u = f' \Delta x ((\theta - \frac{1}{2})\lambda + (w - \frac{1}{2})) \frac{\partial^2 u}{\partial x^2} + O(\Delta x^2) + O(\Delta t^2), \tag{67}$$

with $\lambda = f' \frac{\Delta t}{\Delta x}$. So we may see (66) as a second order discretisation of a *modified equation*

$$\frac{\partial u}{\partial t} + f' \frac{\partial u}{\partial x} = f' \Delta x ((\theta - \frac{1}{2})\lambda + (w - \frac{1}{2})) \frac{\partial u^2}{\partial x^2}. \tag{68}$$

For this reason,

$$\epsilon_{num} = f' \Delta x ((\theta - \frac{1}{2})\lambda + (w - \frac{1}{2})) \frac{\partial u^2}{\partial x^2} \tag{69}$$

is also called the *numerical diffusion* of the difference scheme. Notice that the stability of the difference scheme as studied before, corresponds exactly with a positive numerical diffusion. The numerical diffusion, however, can cause problems in a convection diffusion problem, if the diffusion coefficient in the equation is (much) smaller than the numerical diffusion.

## 2. General FDMs

The FDM can also be used if the problem has more space dimensions. Suppose we have the two-dimensional problem

$$u_t + au_x + bu_y = f, \tag{70}$$

on a rectangular grid $x_{i,j}$. We approximate the function $u(x, y, t)$ by nodal values $u^n_{i,j}$. The space discretisation in the e.g. $x$-direction is

$$\frac{\partial u}{\partial x} \to \frac{u_{i+1,j} - u_{i-1,j}}{2\Delta x} \tag{71}$$

However it is more difficult to use FDM's on non-uniform grids.

## 3. Staggered grids

For systems of equations all methods mentioned above are easy to generalise. All functions can simply be replaced by vector functions. However, it is also possible to approximate the different (dependent) variables by different representations. One way is to approximate the different components of the solution on different point sets. This is the case e.g. with *staggered grids.*

We saw already that the Stokes equations can be written as

$$\begin{pmatrix} \mu\Delta & 0 & -\nabla_1 \\ 0 & \mu\Delta & -\nabla_2 \\ \nabla_1 & -\nabla_2 & 0 \end{pmatrix} \cdot \begin{pmatrix} u \\ v \\ p \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{72}$$

We now can use a staggered grid to represent $u$, $v$ and $p$. Together with the three operators, this has been illustrated in Figure 2.

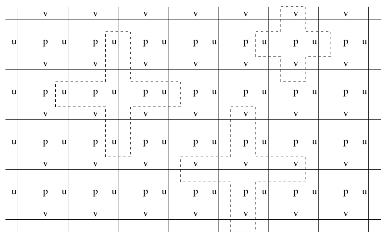


FIGURE 2. Staggered grid to represent $u$, $v$ and $p$.

CHAPTER 4

# Other discretisation principles

We saw already that in the transformation of the PDE into a finite set of equations, it is convenient to distinguish a number of stages:

1. The discretisation of the domain $\Omega$.
2. The discretisation of the solution $u$.
3. The discretisation of the equation.

## 1. Discretisation domain

Suppose we have some discretisation of the domain $\Omega$ by means of nodal points $\{x_i\}$ or a partitioning $\{\Omega_e\}$ e.g. a triangularisation. Notice that we sometimes use both approaches at the same time: using a triangulation we can use the triangles as $\{\Omega_e\}$ or the triangle vertices as $\{x_i\}$. Now we will have to define how to discretise the solution and the equation.

## 2. Discretisation solution

In many numerical results, the approximate solution for a PDE is represented by a finite collection of values $u_i$ in nodal points $x_i$. To get a function $u_h(x)$ defined on the whole domain, we have to interpolate the solution between these nodal points. To this end we can write $u_h$ as

$$u_h(x) = \sum_{j=1}^{N} a_j \phi_j(x) \tag{73}$$

with $\{\phi_j\}$ a collection of *basis functions* on the domain. We can choose e.g. piecewise polynomial basisfunctions with $\phi_i(x_j) = \delta_{ij}$, or use a spectral method, writing the approximation as

$$u_h(x) = \sum_{j=1}^{N} a_j \sin(\pi j x) + b_j \cos(\pi j x) \tag{74}$$

## 3. Discretisation equation

**3.1. Weighted residual methods.** Suppose we have a problem of the form

$$\mathcal{N}(u) = f \tag{75}$$

and to solve for $u$, we want to approximate the solution by a function $u_h$, defined as

$$u_h(x) = \sum_{j=1}^{N} a_j \phi_j(x), \tag{76}$$

with fixed basisfunctions $\{\phi_j\}$ and with $a_1, ..., a_N$ unknown parameters. Now we can discretise the problem by choosing $N$ new functions $w_i$, and construct the $N$ equations

$$\int_\Omega \mathcal{N}(u_h(x)) w_i(x) d\Omega = \int_\Omega f(x) w_i(x) d\Omega, \tag{77}$$

with $i = 1, ..., N$. This procedure is called the weighted residual discretisation. There are many possible choiches for the testfunctions $\{w_i\}$. We give three examples:

3.1.1. *Collocation method.* In the collocation method, we choose the testfunctions $\{w_i\}$ as the Dirac delta function:

$$w_i = \delta_{x_i} \tag{78}$$

To do this, the basisfunctions $\{\phi_i\}$ should be smooth enough, so that $N(u_h(x))$ can be computed for $x = x_i$.

3.1.2. *Finite volume method.* In the finite volume method, we choose the testfunctions $\{w_e\}$ as the characteristic functions over the cells $\Omega_e$. In this approach, the hyperbolic conservation laws can be preserved for each cell.

3.1.3. *Finite Element Methods (FEM).* In the finite element method, we choose the testfunctions $\{w_i\}$ equal to the basisfunctions $\{\phi_i\}$. When applied to symmetric positive definite elliptic PDEs, these methods can also be regarded as variational methods.

**3.2. Variational methods.** Suppose we have the problem (Poisson equation)

$$\begin{array}{lll} \Delta u = f & \text{on} & \Omega \\ u = 0 & \text{on} & \Gamma = \bar{\Omega}/\Omega \end{array} \tag{79}$$

This equation can be written in a variational form: find $u$ such that

$$\int_\Omega \nabla u . \nabla w d\Omega = \int_\Omega f w d\Omega \tag{80}$$

for all functions $w$ on $\Omega$. Using the approximate solution of the form

$$u_h(x) = \sum_{j=1}^{N} a_j \phi_j(x) \tag{81}$$

and, furthermore, choosing the basis functions $\{\phi_i\}$ for $w_j$ (FEM), we get

$$\sum_{j=1}^{N} a_j \int_\Omega \nabla \phi_j \nabla \phi_i d\Omega = \int_\Omega f \phi_i d\Omega \tag{82}$$

for all $i = 1, ..., N$. This is a discrete set of equations, that can be solved for $\{a_j\}$.

We get the same discrete problem by defining

$$J(u) = \int_\Omega |\nabla u|^2 d\Omega - 2\int_\Omega f u d\Omega \tag{83}$$

and minimize $J(\sum_{j=1}^{N} a_j\phi_j)$ over $\{a_j\}$. Notice that the true solution of (80) is the continuous and differentiable function satisfying $u = 0$ on $\Gamma$ (in fact an $H_0^1(\Omega)$-function), that minimizes $J(u)$! So, discretizing we look for a minimizing function in a subset of $H_0^1(\Omega)$.

CHAPTER 5

# Classical FDM for hyperbolic equations

In this chapter, we show and analyse a number of characteristic problems that are encountered in the treatment of time-dependent hyperbolic problems. We do this again by means of the simple convection equation.

$$u_t + f(u)_x = 0 \tag{84}$$

We analyse the (simple) situation with $f(u) = au$ for some $a \in \mathbf{R}$. Then we know the exact solution for this problem:

$$u(x,t) = u(x - at, 0). \tag{85}$$

With Fourier analysis in mind, we take for some fixed $\omega \in \mathbf{R}$ the initial function equal to

$$u(x,0) = e^{\mathbf{i}\omega x}, \tag{86}$$

and, therefore, we take the discretised initial function equal to

$$u_i^0 = e^{\mathbf{i}\omega i\Delta x}. \tag{87}$$

Depending on the scheme we use (linear with constant coefficients), we now get a solution at time $n\Delta t$ of the form

$$u_i^n = \gamma^n e^{\mathbf{i}\omega i\Delta x}. \tag{88}$$

If we consider the exact solution at time $n\Delta t$, we see what $\gamma$ should be:

$$\gamma = e^{-\mathbf{i}\omega a\Delta t}. \tag{89}$$

Using this result, we distinguish two types of errors for a scheme:

- the dissipation error: $-\log(|\gamma|)$,
- the dispersion error: $\frac{arg(\gamma)}{-\omega a\Delta t} - 1$.

We see directly that for the exact solution, both errors are zero. Furthermore we call a scheme stable if $|\gamma| < 1 + k\Delta t$ for some $k$, because for such a scheme, the growth of the error is limited: Integrating over a finite interval $[0,T]$, the solution remains bounded for $\Delta t \to 0$, because $\lim_{\Delta t\to 0}(1 + k\Delta t)^{N/\Delta t} = exp(kT)$. Notice that this definition of stability is slightly weaker than in Chapter 3.

## 1. Classical schemes

### 1.1. The central scheme.

$$u_i^{n+1} = u_i^n + a\frac{\Delta t}{\Delta x}\frac{u_{i-1}^n - u_{i+1}^n}{2} \tag{90}$$

The central scheme is an explicit (forward Euler) method, with $|\gamma| > 1$:

$$\gamma = 1 - \mathbf{i}\frac{a\Delta t}{\Delta x}\sin(\omega\Delta t). \tag{91}$$

So this method is not stable using the definition of stability from Chapter 3. It is, however, in the sense of this chapter.

### 1.2. The Leap-frog scheme.

$$u_i^{n+1} = u_i^n - \frac{a\Delta t}{\Delta x}\frac{u_{i+1}^n - u_{i-1}^n}{2}. \tag{92}$$

In this central two-level method, an approximation should be available on two time levels.

### 1.3. Lax-Friedrichs scheme.

$$u_i^{n+1} = \frac{u_{i+1}^n + u_{i-1}^n}{2} + a\frac{\Delta t}{\Delta x}\frac{u_{i-1}^n - u_{i+1}^n}{2}. \tag{93}$$

The Lax-Friedrichs scheme is only stable if

$$\lambda := \frac{a\Delta t}{\Delta x} < 1. \tag{94}$$

This is called the CFL-condition. Writing a Taylor sequence for this method, we get the modified equation

$$u_t + au_x = -\frac{\Delta t}{2}(a^2 - (\frac{\Delta x}{\Delta t})^2)u_{xx} + O(\Delta x^2 + \Delta t^2). \tag{95}$$

Here we see that this scheme solves a modified equation with a diffusion term. We see that this scheme is only stable if the modified equation has a positive diffusion term.

### 1.4. Lax-Wendroff scheme.

$$u_i^{n+1} = u_i^n - a\frac{\Delta t}{\Delta x}\frac{u_{i+1}^n - u_{i-1}^n}{2} + (a\frac{\Delta t}{\Delta x})^2\frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{2}. \tag{96}$$

The Lax-Wendroff scheme is also only stable if the CFL-condition is satisfied:

$$\lambda := \frac{a\Delta t}{\Delta x} < 1. \tag{97}$$

Writing a Taylor sequence for this scheme, we see a dispersion term in the modified equation:

$$u_t + au_x = \frac{a\Delta x^2}{6}(\lambda^2 - 1)u_{xxx} + O(\Delta x^3). \tag{98}$$

We conclude that the dispersion $\mu < 0$ for $\lambda > 0$. If we fourier transform the dispersion equation $u_t + au_x = \mu u_{xxx}$, we get

$$i\tau\hat{u} + ai\xi\hat{u} = \mu(i\xi)^3\hat{u}\,. \tag{99}$$

Hence we get

$$\tau + (a + \mu\xi^2)\xi = 0. \tag{100}$$

We conclude that using the Lax-Wendroff scheme, high frequency components ($\xi >> 1$) move faster, and low frequency components move slower.

So, in the numerical representation of a discontinuity applying the Lax-Wendroff scheme, the high frequency components lay behind.

**1.5. Beam-Warming scheme.**

$$u_i^{n+1} = u_i^n - \frac{a\Delta t}{\Delta x}\frac{3u_i^n - 4u_{i-1}^n + u_{i-2}^n}{2} + (\frac{a\Delta t}{\Delta x})^2\frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{2}. \tag{101}$$

Writing a Taylor sequence for the Beam-Warming scheme, we see a positive dispersion term in the modified equation:

$$u_t + au_x = \frac{a\Delta x^2}{6}(2 - 3\lambda + \lambda^2)u_{xxx} + O(\Delta x^3). \tag{102}$$

This is a dispersive equation, and $\mu > 0$ if the CFL-number $\lambda < 1$. Again using a fourier transformation, we conclude that, in the numerical representation of a discontinuity, the high frequency components will move ahead.

CHAPTER 5

# Explicit difference schemes

## 1. The differential equation

The differential equation, that is to be solved by some explicit difference schemes is:

$$u_t + f_x(u) = u_t + au_x = 0, \qquad \text{with } a > 0. \tag{103}$$

This is the one-dimensional convection equation without diffusion. In the numerical approximations below it will be solved for $a = 1$ on the domain $t \geq 0$ and $x \geq 0$, with the initial condition $u = 0$ at $t = 0$ and the boundary condition $u = 1$ at $x = 0$. So $x = 0$ is the inflow boundary.
The explicit difference schemes that will be used, are the Lax-Friedrichs scheme, the Lax-Wendroff scheme, the Beam-Warming scheme and the Leapfrog scheme. The original explicit central scheme, is:

$$u_i^{n+1} = u_i^n + \frac{\Delta t}{2\Delta x}(f_{i-1}^n - f_{i+1}^n) = u_i^n + \frac{a\Delta t}{2\Delta x}(u_{i-1}^n - u_{i+1}^n). \tag{104}$$

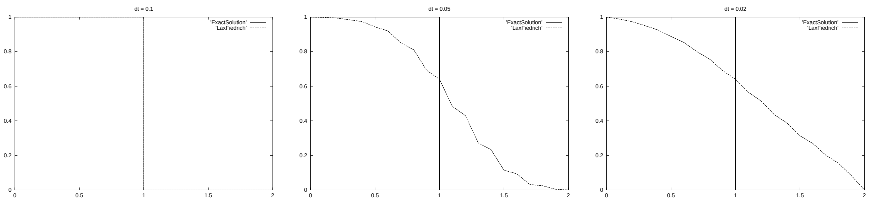
From this scheme, the other schemes are easily derived. The numerical approximations that will be computed by means of different schemes, is the approximation of $u(t, x)$ at $t = 1$.

## 2. The Lax-Friedrich scheme

A slight change in the central explicit scheme (104) gives:

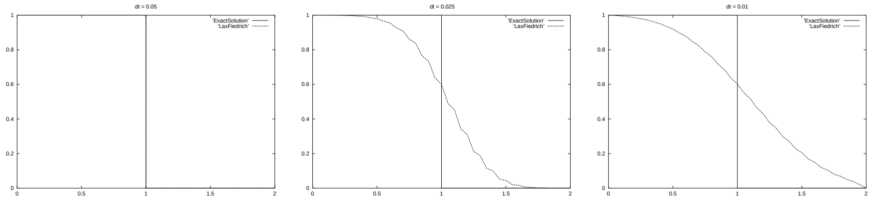
$$u_i^{n+1} = \frac{1}{2}(u_{i-1}^n + u_{i+1}^n) + \frac{a\Delta t}{2\Delta x}(u_{i-1}^n - u_{i+1}^n). \tag{105}$$

This scheme is called the Lax-Fiedrich scheme. Using this scheme for a numerical approximation of (103) , with $\Delta x = 0.1$ , $t = 1$ and different $\Delta t$, we find the following results:




The value of $\Delta t$ for the first graph is 0.1 for the second 0.05 and for the third 0.02. The exact solution is given by the solid line. The dotted line is the numerical approximation, obtained with the Lax-Fiedrich scheme. Notice that for $\Delta t = 0.1$ and $\Delta x = 0.1$, the numerical solution perfectly matches the exact solution. For other values of $\Delta t$, we see that there is numerical diffusion.
The same calculations, can be made for $\Delta x = 0.05$. Then for $\Delta t = 0.05$ ( the first

graph) and $\Delta t = 0.025$ ( the second one) and $\Delta t = 0.01$ ( the last one) the results are.



Notice that the magnitude of $\frac{\Delta t}{\Delta x}$ for these graphs, is the same as those for the first three graphs. The results obtained with the smaller value for $\Delta x$ are better than those obtained with the larger value. This can be seen with the modified equation:

$$u_t + au_x = \frac{\Delta x^2}{2\Delta t}\left(1 - \left(\frac{\Delta t}{\Delta x}a\right)^2\right)u_{xx} + \cdots \tag{106}$$

In this equation, $\frac{\Delta x^2}{2\Delta t}\left(I - \left(\frac{\Delta t}{\Delta x}a\right)^2\right)u_{xx}$ represents the numerical diffusion. Simplified, it looks like $u_t + au_x = \nu u_{xx}$, with

$$\nu = \frac{\Delta x^2}{2\Delta t}(1 - \lambda^2). \tag{107}$$

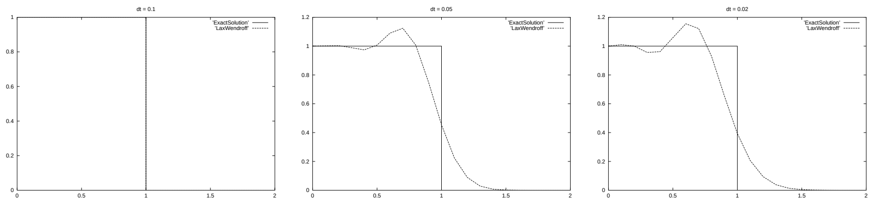
This is a diffusive equation if $\lambda < 1$. You can see immediately, that for a constant value of $\frac{\Delta t}{\Delta x}$, and thus a constant value for $\lambda$ , the numerical diffusion term is smaller for a smaller value of $\Delta x$.

## 3. The Lax-Wendroff scheme

An improved scheme is the one-step second order accurate explicit scheme based on three nodal points in the first time level. This scheme is the Lax-Wendroff scheme. It looks like:

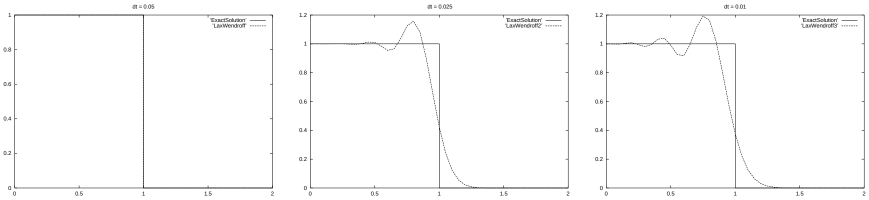
$$u_i^{n+1} = u_i^n - \frac{a\Delta t}{\Delta x}\frac{1}{2}(u_{i+1}^n - u_{i-1}^n) + \left(\frac{a\Delta t}{\Delta x}\right)^2\frac{1}{2}(u_{i+1}^n - 2u_i^n + u_{i-1}^n). \tag{108}$$

A numerical approximation of (103) according to this scheme results in the graphs below:



Again $\Delta x = 0.1$ , $t = 1$ and $\Delta t = 0.1$ for the first graph, 0.05 for the second one, and 0.02 for the third one. Again for $\Delta t = \Delta x$, the numerical solution is the same as the exact solution. For $\Delta t \neq \Delta x$ we see again the numerical diffusion. You can see the exact solution as an infinit sum of sines and cosines, in which each sine/cosine travels with its own speed. In the numerical approximation with the Lax-Wendroff scheme, some of them are going to fast, other ones to slow.

For $\Delta x = 0.05$, and $\Delta t = 0.05$, 0.025 and 0.01 the results are:



If you look closely to the results with $\Delta x = 0.1$ and the results with $\Delta x = 0.05$, you can see that the sine/cosine behavior is more extreme for $\Delta x = 0.05$.
The modified equation for the Lax-Wendroff scheme is:

$$u_t + au_x = \frac{\Delta x^2}{6} a \left( \frac{\Delta t^2}{\Delta x^2} a^2 - 1 \right) u_{xxx} + \cdots \tag{109}$$

With

$$\mu = \frac{\Delta x^2}{6} (\lambda^2 - 1), \tag{110}$$

this equation can be written as $u_t + au_x = \mu u_{xxx}$ . This is the dispersive equation and $\mu < 0$ if $\lambda < 1$. The dispersion equation can easily be analysed by Fourier analysis. Assume the solution to be of the form $u(t,x) = e^{j\omega x} e^{j\alpha(\omega)t}$, then we see $\alpha + a\omega = -\mu\omega^3$ or,

$$u(t,x) = e^{j\omega x} e^{-j\omega t(a+\mu\omega^2)}, \tag{111}$$

i.e. this particular mode in the solution moves with a phase velocity $(a + \mu\omega^2)$. Now it is clear, that rapidly oscillating modes move faster in the positve direction than other modes if $\mu > 0$, and slower if $\mu < 0$. This behaviour is obvious in the results of the Lax-Wendroff scheme.

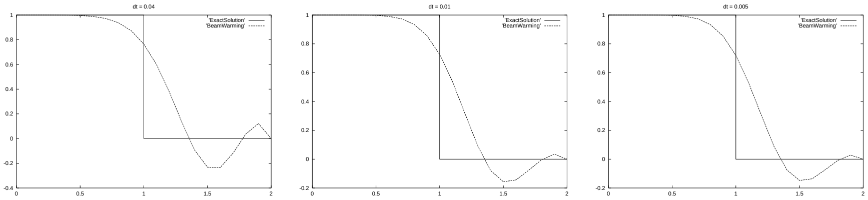
## 4. The Beam-Warming scheme

The Beam-Warming scheme is similar to the Lax-Wendroff scheme, but with a one-sided second order approximation of the first derivative. So the scheme looks like:

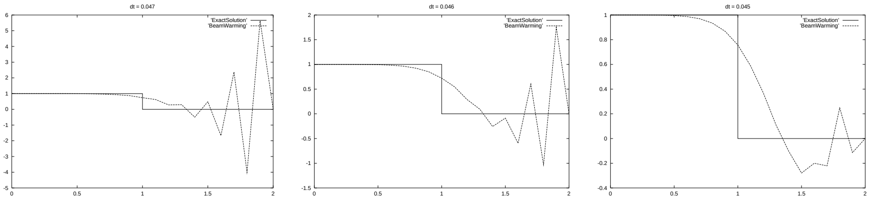
(112)

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{2\Delta x} a (3u_i^n - 4u_{i-1}^n + u_{i-2}^n) + \frac{1}{2} \left( \frac{\Delta t}{\Delta x} \right)^2 a^2 (u_{i+1}^n - 2u_i^n + u_{i-1}^n).$$

For an update around $x = 0$ ( i.e. for $i = 1$) this scheme gives a problem because then $u[i-2] = u[-1]$ , and $u[-1]$ is not defined.( Remember that the boundary condition $u = 1$ at $x = 0$.) The best results are found by taking $u[-1] = 1$.
Solving (103) with this update scheme, leads immediately to extreem errors, when taking $\Delta x = 0.1$, and $\Delta t$ about the same magnitude. Just do one or two iterations by hand for $\Delta x = 0.1$ and $\Delta t = 0.1$ for example. You will see that after only two updates, the errors are already huge. For smaller $\Delta t$, the results become better. By trial and error, you can find that for $\Delta x = 0.1$, $\Delta t$ has to be smaller than 0.047, in order to get sensible results. For $\Delta t = 0.04$, 0.01 and 0.005 the results are in the graphs below:

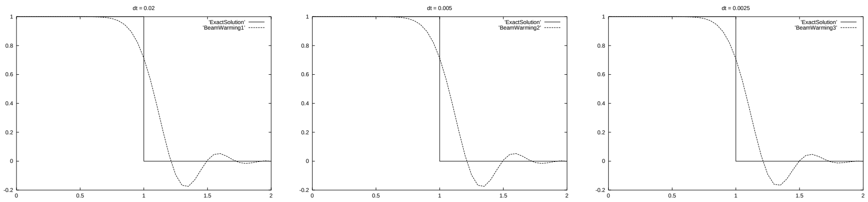



What can be seen immediately, is that it doesn't really matter how small you take $\Delta t$. The results for $\Delta t = 0.04$ are almost the same as those for $\Delta t = 0.001$. For even smaller values of $\Delta t$ , only the processor time becomes longer, and the result doesn't change much. $\Delta t = 0.047$ is about the critical magnitude. Bigger values for $\Delta t$ result in huge errors. The results for $\Delta t = 0.047$, $0.046$ and $\Delta t = 0.045$ are in the next three graphs:




These graphs show, that after $\Delta t$ is smaller than some specific value, of course depending on the value of $\Delta x$, the result converges very quickly to the best numerical approximation that can be achieved by the Beam-Warming scheme.

Again, the results for $\Delta x = 0.05$ are for $\Delta t = 0.02$ (left), $\Delta t = 0.005$ (middle) and $\Delta t = 0.0025$ (right):




Again, there is hardly any difference between the three results, obtained with a different $\Delta t$. If you compare the results with the results obtained with $\Delta x = 0.1$, again the sine/cosine behavior is more extreme for smaller values of $\Delta x$.
The modified equation for the Beam-Warming scheme is

$$u_t + au_x = \frac{\Delta x^2}{6} a \left( 2 - \frac{3\Delta t}{\Delta x} a + \frac{\Delta t^2}{\Delta x^2} a^2 \right) u_{xxx}. \tag{113}$$

With

$$\mu = \frac{\Delta x^2}{6}(2 - 3\lambda + \lambda^2), \tag{114}$$

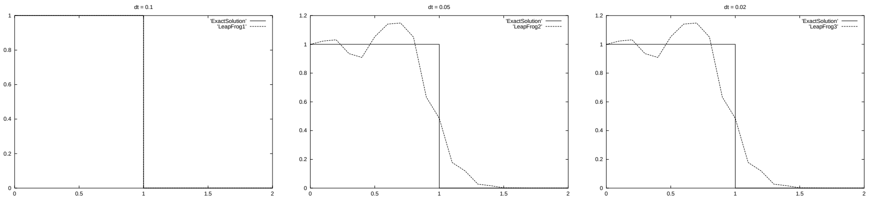
this equation becomes: $u_t + au_x = \mu u_{xxx}$ . This is again a dispersive equation, with $\mu > 0$ if $\lambda < 1$. Because $\mu > 0$, the high frequency modes will now move faster in the positive direction. This is also very obvious in the results.

## 5. The Leapfrog scheme

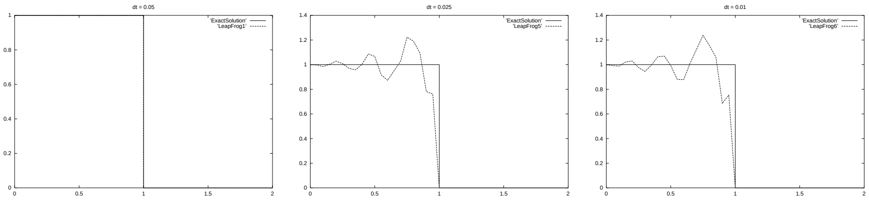
The Leapfrog scheme uses two time levels for an update. The scheme is:

$$u_i^{n+1} = u_i^{n-1} - \frac{\Delta t}{\Delta x} 2(u_{i+1}^n - u_{i-1}^n). \tag{115}$$

For $n = 0$, the values of $u_i^0$ will be used as well for $u_i^{n-1}$ as for $u_i^n$. For $\Delta x = 0.1$ and $\Delta t = 0.1,\ 0.05$ and $0.02$ the results are:




For $\Delta x = 0.05$ and $\Delta t = 0.05,\ 0.025$ and $0.01$ the results are:




What can be seen again, is that for $\Delta x = \Delta t$, the numerical solution is exactly the same as the exact solution.

CHAPTER 6

# Schemes in conservation form

## 1. Basis

In this section we want to study the nonlinear hyperbolic system of equations

$$u_t + f(u)_x = 0. \tag{116}$$

with $u \in \Re$ and $f : \Re^n \to \Re^n$. To solve equation (116) we derive the following numerical scheme :

$$\begin{aligned} 0 &= \int\int_{\Omega} (u_t + f(u)_x)\, d\Omega \\ &= \oint_{\partial\Omega} (n_t u + n_x f(u)\, d\Gamma) \\ &= u_i^{n+1}\Delta x - u_i^n \Delta x - f_{i+\frac{1}{2}}\Delta t + f_{i-\frac{1}{2}}\Delta t \\ &= u_i^{n+1} - u_i^n - \frac{\Delta t}{\Delta x}\left(f_{i+\frac{1}{2}} - f_{i-\frac{1}{2}}\right). \end{aligned} \tag{117}$$

This gives us the time-stepping scheme

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x}\left(f_{i+\frac{1}{2}} - f_{i-\frac{1}{2}}\right) \tag{118}$$

**1.1. The numerical flux function.** To calculate the new value $u_i^{n+1}$ we have to estimate the flux over the edges using the values $\{u_1^n, \ldots, u_N^n\}$, thus

$$f_{i+\frac{1}{2}} = f(\ldots, u_{i-1}, u_i, u_{i+1}, u_{i+2}, \ldots).$$

This expression is called the *numerical flux function.* Important is that the flux at a cell wall computed for cell $\Omega_i$ is the same as the corresponding computed flux in the neighbouring cell, i.e. the amount that flows out of a cell has to be equal to the amount that flows into the next cell.At each cell wall we distinguish the (approximated) value of the solution at the left and the right side. E.g. in cell $\Omega_i$ we approximate the solution by the (mean value) $u_i$ So to compute the flux we have to approximately solve a Riemann Problem.

## 2. The Riemann problem for linear hyperbolic systems

The Riemann problem for (116) is the initial value problem with the initial condition

$$u_0 = \begin{cases} u_L, & x < 0 \\ u_R, & x > 0 \end{cases}$$

where $u_L$ and $u_r$ are constants.

First we study the linear case of the the Riemann problem :

$$\vec{q}_t + A\vec{q}_x = 0,$$

where $A := \frac{\partial f}{\partial q}$ is a matrix independent of $q$, $x$ or $t$. If $A$ is a diagonal matrix one has a system of $n$ independent scalar equations. If (116) is a hyperbolic system, by definition $\frac{\partial f}{\partial q}$ has a complete set of real eigenvalues. In that case we use the eigenvalue decomposision to solve the Riemann problem. Notice that, since we have a hyperbolic system, by definition $A$ can only have real eigenvalues. We can write $A$ as $A = R\Lambda R^{-1}$ with $R$, the matrix with the column eigenvectors, nonsingular, and $\Lambda$ a diagonal matrix which contains the eigenvalues of $A$. Multiplying our equation by $R^{-1}$ we get

$$(R^{-1}\vec{q})_t + R^{-1}R\Lambda(R^{-1}\vec{q})_x = 0,$$

With $\vec{w} = R^{-1}\vec{q}$ we get the following system of $n$ independent scalar equations

$$w_t + \Lambda w_x = 0 \text{ with } \vec{q} = R\vec{w}.$$

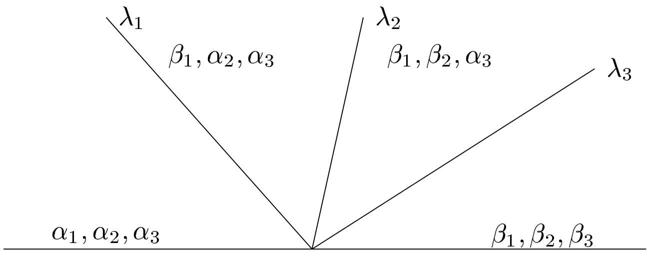
The diagonal matrix $\Lambda$ shows the different velocities for the different components.

$$\vec{q} = \sum_i w_i(t,x)\vec{R}_i = \sum_i w_i(x - \lambda_i t)\vec{R}_i$$

with $w_i(x),\ i = 1, ..., n$, determined by the initial conditions

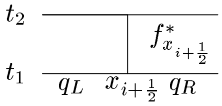
$$\vec{q}_0(x) = \sum_i w(x)\vec{R}_i.$$

This is shown in the next example with a state space of dimension 3. **Example This example shows the effect of the eigenvalues $\lambda_1, \lambda_2, \lambda_3$**




**We see that the component characterized by $\alpha_1$ and $\beta_1$ travels to the left with speed $\lambda_1$, $\alpha_2$ and $\beta_2$ with speed $\lambda_2$ to the right and $\alpha 3$ and $\beta_3$ with speed $\lambda_3$ to the right.**

■ To solve the Riemann problem in the linear case, we use the given values $u_L$ and $u_R$ respectively to the left and to the right of the cell boundary. From these values we can derive the contributions $\alpha_i$ and $\beta_i$ in the direction of the eigenvector $R_i$.



$$\vec{q}_L = \sum_i \alpha_i \vec{R}_i$$
$$\vec{q}_R = \sum_i \beta_i \vec{R}_i$$

Splitting the matrix $\Lambda$ in a part $\Lambda^+$ with the positive and a part $\Lambda^-$ with the negative eigenvalues, we get for $A$

$$A^+ = R\Lambda^+ R^{-1}, \quad A^- = R\Lambda^- R^{-1},$$

and thus

$$A = A^+ + A^-.$$

We compute the flux at the interface :

$$f^*_{x_{i+\frac{1}{2}}} = f(\vec{q}_B) \text{ with,} \tag{119}$$
$$\begin{aligned} \vec{q}_B &= \sum_{\lambda_i>0} \alpha_i \vec{R}_i + \sum_{\lambda_i<0} \beta_i \vec{R}_i \\ &= \sum_{\lambda_i>0} \vec{R}_i (R^{-1}\vec{q}_L)_i + \sum_{\lambda_i<0} \vec{R}_i (R^{-1}\vec{q}_R)_i \end{aligned}$$
(120)

Since $f$ is a linear function of the form $f(\vec{u}) = A\vec{u}$ we get :

$$\begin{aligned} f(\vec{q}_B) = A\vec{q}_B &= \sum_{\lambda_i>0} AR_i(R^{-1}q_L)_i + \sum_{\lambda_i<0} AR_i(R^{-1}q_R)_i && (121)\\ &= \sum_{\lambda_i>0} R\Lambda_i(R^{-1}q_L)_i + \sum_{\lambda_i<0} R\Lambda_i(R^{-1}q_R)_i \\ &= R\Lambda^+R^{-1}\vec{q}_L + R\Lambda^-R^{-1}\vec{q}_R \\ &= A^+\vec{q}_L + A^-\vec{q}_R. && (122) \end{aligned}$$

This gives us the exact solution

$$f(\vec{q}_L, \vec{q}_R) = A^+\vec{q}_L + A^-\vec{q}_R, \tag{123}$$

of the Riemann problem. In the linear case this exact solution can serve as an numerical flux function.

**2.1. The Courant-Isaacson-Rees scheme.** With de numerical flux function (123) the numerical scheme (118) is called the Courant-Isaacson-Rees scheme.

$$\begin{aligned} u_i^{n+1} &= u_i^n - \frac{\Delta t}{\Delta x}\left(f(u_i, u_{i+1}) - f(u_{i-1}, u_i)\right) && (124)\\ &= u_i^n + \frac{\Delta t}{\Delta x}\left(A^+u_{i-1} + (A^- - A^+)u_i - A^-u_{i+1}\right) \end{aligned}$$

**2.2. Artificial diffusion.** We can write the numerical flux function (123) in a different way,

$$f(u,v) = \frac{1}{2}f(u) + \frac{1}{2}f(v) - \frac{1}{2}\|A(v-u),$$

with

$$\|A = A^{+} - A^{-}.$$

Now we can write the Courant-Isaacson-Rees scheme (124) also as

$$u_i^{n+1} - u_i^n + \frac{\Delta t}{\Delta x}\left(\frac{f(u_i)+f(u_{i+1})}{2} - \frac{f(u_{i-1})+f(u_i)}{2}\right) = \frac{\Delta t}{2\Delta x}\|A(u_{i+1} - 2u_i + u_i - 1) \tag{125}$$

This can be seen as in the central difference approximation (second order approximation) of

$$u_t + f(u)_x \simeq \|A u_{xx}\Delta x. \tag{126}$$

We see that the Courant-Isaacson-Rees scheme does not approximate problem (116) up to second order but it rather adds artificial diffusion to the problem. Actualy it is a first order approximation of problem (116)

CHAPTER 7

# The Riemann Problem

## 1. Definition of the problem

In this section [**1**] we discuss the solution of an $n$-dimensional non-linear system of hyperbolic conservation laws of the form

$$q_t + f(q)_x = 0, \tag{127}$$

where $q = (q_1, q_2, ..., q_n) \in \mathbf{R}^n$ and $f(q) : \mathbf{R}^n \to \mathbf{R}^n$, and the initial value for $q$ is given by:

$$q(0,x) = \begin{cases} q_L & \text{if } x < 0, \\ q_R & \text{if } x > 0. \end{cases}$$

Since $f(q) : \mathbf{R}^n \to \mathbf{R}^n$, an alternative form for writing this is in the $n \times n$ matrix form $A(q) = \frac{\partial f}{\partial q}(q)$ and (127) will become

$$q_t + A(q)q_x = 0, \tag{128}$$

where $A(q)$ can be written as:

$$A(q) = R(q)\Lambda(q)R(q)^{-1}. \tag{129}$$

Here $\Lambda(q)$ is the diagonalized form of $A(q)$ and $R(q)$ is composed of the eigenvectors $\vec{R}_k$ of $A(q)$. Because the system is hyperbolic, $\Lambda(q)$ contains the $n$ real eigenvalues $\lambda_k(q) \in \mathbf{R}$ of $A(q)$ on its diagonal.
There are now two distinct cases that can apply to each eigenvalue-eigenvector pair:

- the *genuinly non-linear case*: $(\vec{\nabla}\lambda_k(q), \vec{R}_k(q)) \neq 0$
- the *linearly degenerate case*: $(\vec{\nabla}\lambda_k(q), \vec{R}_k(q)) = 0$

where

$$\vec{\nabla}\lambda_k(q) \equiv \left(\frac{\partial \lambda_k}{\partial q_1}, \frac{\partial \lambda_k}{\partial q_2}, ..., \frac{\partial \lambda_k}{\partial q_n}\right)^T.$$

In the following sections we discuss the methods to solve these two cases. The genuinly non-linear case will be dealt with in section 4. The degenerate case is treated in section 5.

## 2. Property of the solution $\tilde{q}(t,x)$

Before we continue, first an important property of the solution of the Riemann problem, defined as $\tilde{q}(t,x)$, will be derived. This property is needed in section (4). An alternative way for writing equation (127) is

$$\frac{\partial q}{\partial(\alpha t)}\frac{\partial(\alpha t)}{\partial t} + \frac{\partial f(q)}{\partial q}\frac{\partial q}{\partial(\alpha x)}\frac{\partial(\alpha x)}{\partial x} = 0 \tag{130}$$

and therefore

$$\alpha\left(\frac{\partial q}{\partial(\alpha t)}+\frac{\partial f(q)}{\partial q}\frac{\partial q}{\partial(\alpha x)}\right)=0 \tag{131}$$

This means, that if $\tilde{q}(t,x)$ is a solution of (127), so is $\tilde{q}(\alpha t,\alpha x)$. However, since $\tilde{q}(t,x)$ is a **unique** solution

$$q(\alpha_1 t,\alpha_1 x)=q(\alpha_2 t,\alpha_2 x) \qquad \forall\alpha_1,\alpha_2\in\mathbf{R}^+. \tag{132}$$

This means, that $q(t,x)$ can be written as $q(t,x)=q(\frac{x}{t})$.

## 3. The Rankine Hugoniot relations

When dealing with a discontinuity in the solution for $q(t,x)$ the integral form of (1.1) has to be employed, meaning

$$\int_{\partial G} qn_t+f(q)n_x d\sigma=0 \tag{133}$$

where $G$ is an arbitrary area in the $(x,t)$-space and $\vec{n}=(n_x,n_t)$ is the outward unit normal to $\partial G$. Since any area can be used for $\partial G$ it will suffice to consider a small area along the diagonal discontinuity in the $(x,t)$-space. This integration will result in the equation

$$f(q_R)-f(q_L)+\frac{\Delta x}{\Delta t}(q_L-q_R)=0 \tag{134}$$

with $\frac{\Delta x}{\Delta t}=\lambda_k(q_L)=\lambda_k(q_R)$. This is the *Rankine Hugoniot* relation, which a discontinuous solution of the hyperbolic conservation law should satisfy.

## 4. The genuinly non-linear case

In the case

$$(\vec{\nabla}\lambda_k(q),\vec{R}_k(q))\neq 0, \tag{135}$$

for simplicity we scale the eigenvectors in such a way that

$$(\vec{\nabla}\lambda_k(q),\vec{R}_k(q))=1. \tag{136}$$

Now we can consider the decomposition (129) and choose an arbitrary $k\in\{1,2,...,n\}$. For this $k$ we consider a path through $q_L$ in the state space that is everywhere tangent to the eigenvector $\vec{R}_k(q)$. We parametrize this path by considering the differential equation

$$\frac{dq(\xi)}{d\xi}=\vec{R}_k(q(\xi)) \tag{137}$$

with

$$q(\xi_L)=q_L. \tag{138}$$

The solution of (137) is denoted by $\tilde{q}(\xi)$ or $\Gamma_k(q)$. The variable $\xi$ used here is an abstract variable. We first assume that $q_R=\tilde{q}(\xi_R)$ is located somewhere on $\Gamma_k(q)$. The eigenvalue $\lambda_k(q)$ changes over the path $\Gamma_k(q)$. This follows from the equation:

$$\frac{\partial\lambda_k(\tilde{q}(\xi))}{\partial\xi}=\frac{\partial\lambda_k(\tilde{q}(\xi))}{\partial\tilde{q}(\xi)}\frac{\partial\tilde{q}(\xi)}{\partial\xi}=\frac{\partial\lambda_k(\tilde{q}(\xi))}{\partial\tilde{q}(\xi)}\vec{R}_k(\tilde{q}(\xi)) \tag{139}$$

$$=\vec{\nabla}\lambda_k\cdot\vec{R}_k(\tilde{q}(\xi))=1. \tag{140}$$

Therefore

$$\lambda_k(\tilde{q}(\xi)) = \xi. \tag{141}$$

This describes the value of $\lambda_k(\tilde{q}(\xi))$ for different $\xi$ on $\Gamma_k$. It follows:

$$\lambda_k(\tilde{q}(\xi_L)) = \xi_L \tag{142}$$

and

$$\lambda_k(\tilde{q}(\xi_R)) = \lambda_k(q_R) = \xi_R. \tag{143}$$

Keeping the result of section 2 in mind, we continue the derivation of a solution for equation (127).

$$\frac{\partial q(t,x)}{\partial t} + \frac{\partial f(q(t,x))}{\partial x} = \frac{\partial q(t,x)}{\partial t} + A(q(t,x))\frac{\partial q(t,x)}{\partial x} = 0 \tag{144}$$

This can now be written as

$$\frac{\partial q(\frac{x}{t})}{\partial(\frac{x}{t})}\frac{\partial(\frac{x}{t})}{\partial t} + A(q(t,x))\frac{\partial q(\frac{x}{t})}{\partial(\frac{x}{t})}\frac{\partial(\frac{x}{t})}{\partial x} \tag{145}$$

$$= -\frac{x}{t^2}\frac{\partial q(\frac{x}{t})}{\partial\frac{x}{t}} + \frac{1}{t}A(q(t,x))\frac{\partial q(\frac{x}{t})}{\partial\frac{x}{t}} \tag{146}$$

$$= -\frac{x}{t^2}\vec{R}_k(q(\frac{x}{t})) + \frac{1}{t}A(q(t,x))\vec{R}_k(q(\frac{x}{t})) \tag{147}$$

$$= -\frac{x}{t^2}\vec{R}_k(q(\frac{x}{t})) + \frac{1}{t}\lambda_k(q(\frac{x}{t}))\vec{R}_k(q(\frac{x}{t})) \tag{148}$$

The transition (146) to (147) holds if $q(\frac{x}{t}) = \tilde{q}(\frac{x}{t})$ and therefore, an obvious solution for (148) is $\lambda_k(\tilde{q}(\frac{x}{t})) = \frac{x}{t}$. Hence we find for $q(t,x)$, the solution of (127),

$$q(t,x) = \begin{cases} q_L & \text{if } \frac{x}{t} < \lambda_k(q_L) \\ \tilde{q}(\frac{x}{t}) & \text{if } \lambda_k(q_L) < \frac{x}{t} < \lambda_k(q_R) \\ q_R & \text{if } \frac{x}{t} > \lambda_k(q_R) \end{cases} \tag{149}$$

Therefore, if $\lambda(q_L) = \xi_L < \frac{x}{t} < \xi_R = \lambda_k(q_R)$ there will be a smooth solution, referred to as a *rarefaction wave*. However, there won't be a domain for $q(t,x)$ described by $\tilde{q}(\frac{x}{t})$ if $\lambda_k(q_L) > \lambda_k(q_R)$, since then, according to (149) there won't be a domain $\lambda_k(q_L) < \frac{x}{t} < \lambda_k(q_R)$.
If $\lambda(q_L) > \lambda(q_R)$ the characteristics from the left run faster (to the right) then the characteristics from the right and no smooth solution exists. A discontinuity appears that moves according to the *Rankine-Hugoniot relations*, derived in section 3. In this case there is no smooth solution and a *shock wave* appears.

## 5. The linearly degenerate case

In the *linearly degenerate* case we still consider (137) with initial condition (138), but (135) no longer applies. Instead, the relationship between the eigenvectors and -values is now given by

$$(\vec{\nabla}\lambda_k(q), \vec{R}_k(q)) = 0 \qquad \forall q \in \mathbf{R}^n. \tag{150}$$

Once again we define $\tilde{q}(\xi)$ as the solution of (137) and (138). In analogy with (140) we now find:

$$\frac{\partial\lambda(\tilde{q}(\xi))}{\partial\xi} = \vec{\nabla}\lambda_k(\tilde{q}(\xi)) \cdot \vec{R}_k(\tilde{q}(\xi)) = 0 \tag{151}$$

Therefore $\lambda_k(\tilde{q}(\xi))$ is constant along the path $\tilde{q}(\xi)$. This means that $q_L$ and $q_R$ are connected if

$$\lambda_k(q_L) = \lambda_k(\tilde{q}(\xi)) = \lambda_k(q_R) \qquad \forall \xi \in [\xi_L, \xi_R]. \tag{152}$$

For $q(t,x)$ now follows:

$$q(t,x) = \begin{cases} q_L & \frac{x}{t} = \xi < \lambda_k(q_L) = \lambda_k(q_R) \\ q_R & \frac{x}{t} = \xi > \lambda_k(q_R) = \lambda_k(q_L) \end{cases}$$

This implies a discontinuity in $q(t,x)$ along the line $\frac{x}{t} = \lambda_k(\tilde{q}(\xi))$, referred to as a *contact discontinuity*. This contact discontinuity should satisfy the Rankine Hugoniot relations, derived in section 3.

## 6. Riemann invariants

In the previous sections we assumed, that $q_L$ and $q_R$ are directly connected by a $k$-path $\Gamma_k$. In the general case, however, it will take a combination of **all** of the $n$ paths $\Gamma_i$ (with $i \in \{1, 2, \ldots, n\}$) to connect the point $q_L$ to $q_R$. This combined path will be referred to as $\Gamma$[1].
Along a path $\Gamma_i$ the value of $\vec{q}$ will change from $q_{i-1}$ to $q_i$ and will consist of two sectors where the solution is constant at values of respectively $q_{i-1}$ and $q_i$, seperated by a discontinuity or rarefaction wave. Since there are $n$ paths $\Gamma_k$ the total solution for $\vec{q}$ along the path $\Gamma$ will consist of $n+1$ sectors where the solution is constant at a value $q_i$ separated by a number of discontinuities and rarefaction waves, one for each eigenvalue.

$$q_L = q_0 \to q_1 \to q_2 \to \ldots \to q_{n-1} \to q_n = q_R.$$

The values of $q_0 = q_L$ and $q_n = q_R$ are known by definition, the values $q_i$ with $i \in \{1, 2, \ldots, n-1\}$ need to be calculated. To do this, the simple wave and contact discontinuities separating the values $q_i$ need to be constructed.
Quantities that remain constant along the $k$-paths are useful for constructing these simple waves and contact discontinuities.
In the **linear case** we already found such quantities, the eigenvalues $\lambda_k$. The equivalent of such quantities in the **non-linear case** are called the *Riemann invariants*. These Riemann invariants generally are smooth functions $\psi_k^i : \mathbf{R}^n \to \mathbf{R}$ defined on the state-space $\mathbf{R}^n$. The property that they are constant along a path $\Gamma_k$ is expressed by the formula

$$(\vec{\nabla}\psi_k^i(q), \vec{R}_k(q)) = 0. \qquad \forall q \in \mathbf{R}^n. \tag{153}$$

In general we may expect $n-1$ $k$-Riemann invariants $\psi_k^i$ along each $k$-path, whose gradients are linearly independent in the state-space $\mathbf{R}^n$.
The Riemann invariants $\psi_k^i$ will adhere to

$$\psi_k^i(q_{k-1}) = \psi_k^i(q_k) \quad i \in \{1, 2, \ldots, n-1\}.$$

Since there are $n$ $k$-paths with $n-1$ Riemann invariants each, there is a total of $n(n-1)$ of such relations. From these $n(n-1)$ relations it is generally possible to derive the values of the $n-1$ unknown vectors $\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_{n-1}$.

[1]In the linear case it is possible to connect any two points in the state-space by a path $\Gamma$ since the eigenvectors span the state space $\mathbf{R}^n$. If linearisation is possible and the distance between $q_L$ and $q_R$ is not large, the same may hold in the non-linear case. However, in the general case, for a non-linear problem a solution is not guaranteed.

**6.1. An example of the calculation of Riemann invariants.** [**2**] Consider a system described by the conservation laws

$$v_t - u_x = 0, \qquad u_t + p(v)_x = 0, \qquad t > 0, x \in \mathbf{R} \tag{154}$$

where $p' < 0$ and $p'' > 0$[2]. The definition

$$\vec{U} = (v, u), \qquad \vec{F}(\vec{U}) = (-u, p(v)), \tag{155}$$

is made, so that (154) can be written in the form

$$\vec{U}_t + \vec{F}(\vec{U})_x = 0. \tag{156}$$

Equation (156) is a hyperbolic system since the Jacobian matrix,

$$dF = \begin{pmatrix} 0 & -1 \\ p'(v) & 0 \end{pmatrix} \tag{157}$$

has real and distinct eigenvalues

$$\lambda_1 \equiv -\sqrt{-p'(v)} < 0 < \sqrt{-p'(v)} \equiv \lambda_2. \tag{158}$$

The left eigenvector, corresponding to $\lambda_1$, is $\vec{r_1} = (1, \sqrt{-p'(v)})^T$. Since $n = 2$ it should be possible to find $n - 1 = 1$ Riemann invariant, $\psi_1$. This Riemann invariant should satisfy the equation $\vec{r_1} \cdot \vec{\nabla}\psi_1 = 0$, or

$$\psi_{1v} + \sqrt{-p'(v)}\psi_{1u} = 0, \tag{159}$$

and is given explicitly by

$$\psi_1(v, u) = u - \int^v \sqrt{-p'(y)}dy. \tag{160}$$

A similar calculation shows that the Riemann invariant, corresponding to the eigenvector $\vec{r_2} = (-1, \sqrt{-p'(v)})^T$, is given by

$$\psi_2(v, u) = u + \int^v \sqrt{-p'(y)}dy \tag{161}$$

**6.2. Osher's approximate Riemann solver.** Implementing a numerical method for the Euler equations, based on a finite-volume principle, requires the repeated computation of the flux between neighbouring cells. This makes solving the Riemann problem a CPU intensive task, requiring, amongst other things, a solution for the Rankine Hugoniot relations at discontinuities in the solution for $q(t, x)$. An interesting alternative is an *approximate Riemann solver*, which provides a solution of slightly inferior quality, but in a significantly shorter amount of time. The approximate Riemann solver discussed in this section is the one developed by Osher [**3**].

In Osher's scheme, a variable $\vec{U}$ is defined in the $n$-dimensional phase space (with $n$=3 for the one-dimensional Euler equations[3]). In this $n$-dimensional phase-space, the numerical flux of Osher's scheme for a system of conservative laws is:

$$f^{*(O)}_{i+\frac{1}{2}} = \frac{1}{2}(f_i + f_{i+1}) - \frac{1}{2}\int_{\vec{U}_i}^{\vec{U}_{i+1}} |A(\vec{U})| d\vec{U}. \tag{162}$$

[2] If $p(v)$ is chosen $p(v) = kv^{-\gamma}$, then the isentropic gas dynamics equations are recovered.

[3] An example of a possible choice for these three dimensions is $\rho$, $u$ and $E$.

To calculate the integral in this equation a path has to be chosen from $\vec{U}_i$ to $\vec{U}_{i+1}$. Such a path can be constructed by the eigenvectors $\vec{R}_k(\vec{U})$ of the Jacobian matrix $\frac{df(\vec{U})}{d\vec{U}}$. If the state vector $\vec{U}$ is defined as

$$\vec{U} = \begin{pmatrix} \rho \\ \rho u \\ \rho E \end{pmatrix} \tag{163}$$

then for the Euler equation the flux is

$$\vec{f}(\vec{U}) = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho u H \end{pmatrix}, \tag{164}$$

where in the case of a perfect gas ($p$=pressure, $H$=enthalpy and $\gamma$ is the ratio of specific heats)

$$\gamma = \frac{c_p}{c_v}, \tag{165}$$

$$p = (\gamma - 1)(\rho E - \frac{1}{2}\rho u^2), \tag{166}$$

$$H = \frac{\rho E + p}{\rho} \tag{167}$$

then $A(\vec{U}) = \frac{df(\vec{U})}{d\vec{U}}$ is

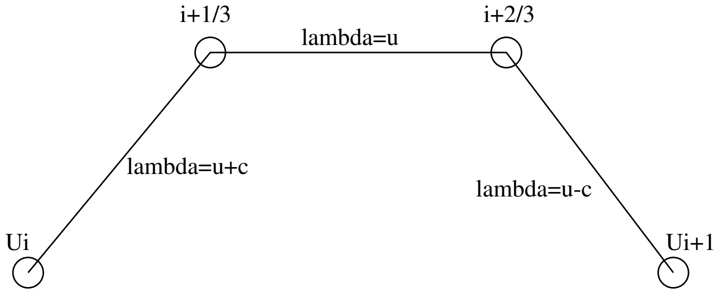
$$A(\vec{U}) = \begin{pmatrix} 0 & 1 & 0 \\ -(3-\gamma)\frac{u^2}{2} & (3-\gamma)u & \gamma - 1 \\ (\gamma-1)u^3 - \gamma u E & \gamma E - 3\frac{\gamma-1}{2}u^2 & \gamma u \end{pmatrix}. \tag{168}$$

Now the eigenvectors of $A(\vec{U})$, $\vec{R}_k(\vec{U})$, and their associated eigenvalues $\lambda_k(\vec{U})$ can be calculated.
Using these eigenvectors a path can now be constructed from $\vec{U}_i$ to $\vec{U}_{i+1}$. After the order of using the eigenvectors is selected, this decomposition of the integration path is unique.
However, since the eigenvectors, $\vec{R}_k(\vec{U})$, are **functions of the phase-space** and not constant vectors, constructing the path is not a straightforward task. This is where the Riemann invariants come in useful. With the method described in section 6.1 $(n-1)$ Riemann invariants can be defined along each path. Below we give a table listing all the eigenvalues and their associated Riemann invariants:

| $\lambda$ | $\psi$ |
|---|---|
| $u+c$ | $\sigma$<br>$u - \frac{2c}{\gamma-1}$ |
| $u$ | $u$<br>$p$ |
| $u-c$ | $\sigma$<br>$u + \frac{2c}{\gamma-1}$ |

Here,

$$\sigma = c_v ln\frac{p}{\rho^\gamma} \tag{169}$$




Integration paths in phase space for Osher's scheme
and the one-dimensional Euler equations

denotes the logarithm of the entropy and

$$c = \sqrt{\frac{\gamma p}{\rho}} \tag{170}$$

denotes the speed of sound.
Now the order in which to follow the eigenvectors is selected. Osher chose the order of decreasing eigenvalues here, but any other order is possible too.
The unique position of the intermediate points $i+\frac{1}{3}$ and $i+\frac{2}{3}$ in the phase-space can be determined from the constancy of two Riemann invariants along each of the three simple wave paths. They provide six relations for the six unknowns $\vec{U}_{i+\frac{1}{3}}$ and $\vec{U}_{i+\frac{2}{3}}$ (or in other words: $\rho_{i+\frac{1}{3}}$,$u_{i+\frac{1}{3}}$,$E_{i+\frac{1}{3}}$, $\rho_{i+\frac{2}{3}}$,$u_{i+\frac{2}{3}}$ and $E_{i+\frac{2}{3}}$, according to (163)). These six relations and the equations they've been derived from (using (166),(167),(169) and (170)) have been summarized in the table below:

| | |
|---|---|
| $\sigma_i=\sigma_{i+\frac{1}{3}}$ | $\frac{\rho_i E_i-\frac{1}{2}\rho_i u_i^2}{\rho_i^\gamma} = \frac{\rho_{i+\frac{1}{3}}E_{i+\frac{1}{3}}-\frac{1}{2}\rho_{i+\frac{1}{3}}u^2_{i+\frac{1}{3}}}{\rho^\gamma_{i+\frac{1}{3}}}$ |
| $u_i - 2\frac{c_i}{\gamma-1} = u_{i+\frac{1}{3}} - 2\frac{c_{i+\frac{1}{3}}}{\gamma-1}$ | $u_i - 2\sqrt{\frac{\gamma}{\gamma-1}\frac{\rho_i E_i-\frac{1}{2}\rho_i u_i^2}{\rho_i}}$ $= u_{i+\frac{1}{3}} - 2\sqrt{\frac{\gamma}{\gamma-1}\frac{\rho_{i+\frac{1}{3}}E_{i+\frac{1}{3}}-\frac{1}{2}\rho_{i+\frac{1}{3}}u^2_{i+\frac{1}{3}}}{\rho_{i+\frac{1}{3}}}}$ |
| $u_{i+\frac{1}{3}} = u_{i+\frac{2}{3}}$ | $u_{i+\frac{1}{3}} = u_{i+\frac{2}{3}}$ |
| $p_{i+\frac{1}{3}} = p_{i+\frac{2}{3}}$ | $\rho_{i+\frac{1}{3}}E_{i+\frac{1}{3}} - \frac{1}{2}\rho_{i+\frac{1}{3}}u^2_{i+\frac{1}{3}} = \rho_{i+\frac{2}{3}}E_{i+\frac{2}{3}} - \frac{1}{2}\rho_{i+\frac{2}{3}}u^2_{i+\frac{2}{3}}$ |
| $\sigma_{i+\frac{2}{3}}=\sigma_{i+1}$ | $\frac{\rho_{i+\frac{2}{3}}E_{i+\frac{2}{3}}-\frac{1}{2}\rho_{i+\frac{2}{3}}u^2_{i+\frac{2}{3}}}{\rho^\gamma_{i+\frac{2}{3}}} = \frac{\rho_{i+1}E_{i+1}-\frac{1}{2}\rho_{i+1}u^2_{i+1}}{\rho^\gamma_{i+1}}$ |
| $u_i + 2\frac{c_{i+\frac{2}{3}}}{\gamma-1} = u_{i+1} + 2\frac{c_{i+1}}{\gamma-1}$ | $u_{i+\frac{2}{3}} + 2\sqrt{\frac{\gamma}{\gamma-1}\frac{\rho_{i+\frac{2}{3}}E_{i+\frac{2}{3}}-\frac{1}{2}\rho_{i+\frac{2}{3}}u^2_{i+\frac{2}{3}}}{\rho_{i+\frac{2}{3}}}}$ $= u_{i+1} + 2\sqrt{\frac{\gamma}{\gamma-1}\frac{\rho_{i+1}E_{i+1}-\frac{1}{2}\rho_{i+1}u^2_{i+1}}{\rho_{i+1}}}$ |

Before the result for the six unknowns are given, first the following definitions are made:

$$A = \frac{\rho_i E_i - \frac{1}{2}u_i^2\rho_i}{\rho_i^\gamma},$$

$$B = u_i - 2\frac{c_i}{\gamma - 1},$$

$$C = \frac{\rho_{i+1}E_{i+1} - \frac{1}{2}u_{i+1}^2\rho_{i+1}}{\rho_{i+1}^{\gamma}},$$

$$D = u_{i+1} + 2\frac{c_{i+1}}{\gamma - 1}.$$

The six unknowns in $\vec{U}_{i+\frac{1}{3}}$ and $\vec{U}_{i+\frac{2}{3}}$ are:

$$\rho_{i+\frac{2}{3}} = \left(\frac{\frac{D-B}{2}}{\sqrt{\frac{\gamma}{\gamma-1}C}(1+(\frac{A}{C})^{\frac{1}{2\gamma}})}\right)^{\frac{2}{\gamma-1}}, \tag{171}$$

$$u_{i+\frac{1}{3}} = u_{i+\frac{2}{3}} = D - 2\sqrt{\frac{\gamma}{\gamma-1}C\rho_{i+\frac{2}{3}}^{\gamma-1}}, \tag{172}$$

$$\rho_{i+\frac{1}{3}} = (\frac{C}{A})^{\frac{1}{\gamma}}\rho_{i+\frac{2}{3}}, \tag{173}$$

$$E_{i+\frac{2}{3}} = \frac{C\rho_{i+\frac{2}{3}}^{\gamma} + \frac{1}{2}\rho_{i+\frac{2}{3}}u_{i+\frac{2}{3}}^2}{\rho_{i+\frac{2}{3}}}, \tag{174}$$

$$E_{i+\frac{1}{3}} = \frac{\rho_{i+\frac{2}{3}}E_{i+\frac{2}{3}} - \frac{1}{2}\rho_{i+\frac{2}{3}}u_{i+\frac{2}{3}}^2 + \frac{1}{2}\rho_{i+\frac{1}{3}}u_{i+\frac{1}{3}}^2}{\rho_{i+\frac{1}{3}}}. \tag{175}$$

Now the entire integration path along the eigenvectors $\vec{R}_k(\vec{U})$, $k = 1, 2, 3$, is defined. Along each subpath $\Gamma(j)$, the integration of $|A(\vec{U})|$ in equation (162) takes on the simple expression:

$$\int_{\Gamma(j)} |A(\vec{U})|d\vec{U} = \int_{\Gamma(j)} |A(\vec{U})|\vec{R}_j(\vec{U})dw = \int_{\Gamma(j)} |\lambda_j(\vec{U})|R_j(\vec{U})dw, \tag{176}$$

meaning that the numerical flux can now be written as:

$$f_{i+\frac{1}{2}}^{*(O)} = \frac{1}{2}(f_i + f_{i+1}) - \frac{1}{2}\sum_j \int_{\Gamma(j)} |\lambda_j|\vec{R}_j dw. \tag{177}$$

When calculating the integration along each subpath $\Gamma(j)$ a distinction has to be made between genuinly non-linear waves, for which the eigenvalues are not constant along the simple wave associated with the $j$-th eigenvector, that is for which

$$\frac{d\lambda_j}{dw} = \frac{d\lambda_j}{dU} \cdot \vec{R}_j \neq 0 \tag{178}$$

and the linear degenerate waves that satisfy

$$\frac{d\lambda_j}{dw} = \frac{d\lambda_j}{dU} \cdot \vec{R}_j = 0. \tag{179}$$

This has been discussed earlier in section 1.
The eigenvalue $\lambda = u$ remains constant along its associated eigenvector (after all, $\psi = u$ is a Riemann invariant of this eigenvector) and therefore the integration along it will result in:

$$\int_{U_{1+\frac{1}{3}}}^{U_{i+\frac{2}{3}}} |A(\vec{U})|d\vec{U} = sgn(u)\left[f(\vec{U}_{i+\frac{2}{3}}) - f(\vec{U}_{i+\frac{1}{3}})\right]. \tag{180}$$

In the genuinly non-linear case, the non-zero gradient of $\lambda$ means, that somewhere there exists a unique **sonic point**, defined by $\lambda(\vec{U}^*) = 0$, at which $\lambda$ changes sign.

The location of this sonic point, $\vec{U}^*$, can be derived by solving three equations for the three unknowns $\rho^*$, $u^*$ and $E^*$. When integrating along the path associated with $\lambda = u + c$ these three equations are

$$u^* = -c^* \tag{181}$$

$$s^* = s_i \tag{182}$$

$$u^* - 2\frac{c^*}{\gamma - 1} = u_i - 2\frac{c_i}{\gamma - 1} \tag{183}$$

and the solution is (using $A = \frac{\rho_i E_i - \frac{1}{2}\rho_i u_i^2}{\rho_i^\gamma}$ and $B = u_i - 2\frac{c_i}{\gamma-1}$):

$$\rho^* = \left(\frac{-B}{2\sqrt{\frac{\gamma}{\gamma-1}A} + \sqrt{\gamma(\gamma-1)A}}\right)^{\frac{2}{\gamma-1}}, \tag{184}$$

$$u^* = -\sqrt{\gamma(\gamma-1)A{\rho^*}^{\gamma-1}}, \tag{185}$$

$$E^* = A{\rho^*}^{\gamma-1} + \frac{1}{2}{u^*}^2. \tag{186}$$

When integrating along the path associated with $\lambda = u - c$ the equations are

$$u^* = c^* \tag{187}$$

$$s^* = s_{i+1} \tag{188}$$

$$u^* + 2\frac{c^*}{\gamma - 1} = u_{i+1} + 2\frac{c_{i+1}}{\gamma - 1} \tag{189}$$

and the solution is (using $C = \frac{\rho_{i+1} E_{i+1} - \frac{1}{2}\rho_{i+1} u_{i+1}^2}{\rho_{i+1}^\gamma}$ and $D = u_{i+1} + 2\frac{c_{i+1}}{\gamma-1}$):

$$\rho^* = \left(\frac{D}{2\sqrt{\frac{\gamma}{\gamma-1}C} + \sqrt{\gamma(\gamma-1)C}}\right)^{\frac{2}{\gamma-1}}, \tag{190}$$

$$u^* = \sqrt{\gamma(\gamma-1)C{\rho^*}^{\gamma-1}}, \tag{191}$$

$$E^* = C{\rho^*}^{\gamma-1} + \frac{1}{2}{u^*}^2. \tag{192}$$

When performing the integration along the path associated with $\lambda = u + c$ there are now two distinct cases:

1. $\lambda_i \cdot \lambda_{i+\frac{1}{3}} >= 0$, i.e. $\lambda$ does not change sign along the path.
2. $\lambda_i \cdot \lambda_{i+\frac{1}{3}} < 0$, i.e. $\lambda$ changes sign along the path.

In the first case, the same approach as in (180) applies. In the second case, the integration path has to be divided into the integration from $\vec{U}_i$ to $\vec{U}^*$ and the integration from $\vec{U}^*$ to $\vec{U}_{i+1}$. Now the approach from (180) can be applied to both subpaths. This finally results in

$$\int_{\vec{U}_i}^{\vec{U}_{i+\frac{1}{3}}} |A| d\vec{U} = \begin{cases} f(\vec{U}_{i+\frac{1}{3}}) - f(\vec{U}_i) & \lambda(\vec{U}_i) >= 0, \lambda(\vec{U}_{i+\frac{1}{3}}) >= 0 \\ 2f(\vec{U}^*_{i+\frac{1}{3}}) - f(\vec{U}_i) - f(\vec{U}_{i+\frac{1}{3}}) & \lambda(\vec{U}_i) >= 0, \lambda(\vec{U}_{i+\frac{1}{3}}) < 0 \\ f(\vec{U}_i) + f(\vec{U}_{i+\frac{1}{3}}) - 2f(\vec{U}^*_{i+\frac{1}{3}}) & \lambda(\vec{U}_i) < 0, \lambda(\vec{U}_{i+\frac{1}{3}}) >= 0 \\ f(\vec{U}_i) - f(\vec{U}_{i+\frac{1}{3}}) & \lambda(\vec{U}_i) < 0, \lambda(\vec{U}_{i+\frac{1}{3}}) < 0. \end{cases}$$

The same applies to performing the integration along the path associated with $\lambda = u - c$. Here the result for the integral is:

$$\int_{\vec{U}_{i+\frac{2}{3}}}^{\vec{U}_{i+1}} |A| d\vec{U} = \begin{cases} f(\vec{U}_{i+1}) - f(\vec{U}_{i+\frac{2}{3}}) & \lambda(\vec{U}_{i+\frac{2}{3}}) >= 0, \lambda(\vec{U}_{i+1}) >= 0 \\ 2f(\vec{U}^*_{i+1}) - f(\vec{U}_{i+\frac{2}{3}}) - f(\vec{U}_{i+1}) & \lambda(\vec{U}_{i+\frac{2}{3}}) >= 0, \lambda(\vec{U}_{i+1}) < 0 \\ f(\vec{U}_{i+\frac{2}{3}}) + f(\vec{U}_{i+1}) - 2f(\vec{U}^*_{i+1}) & \lambda(\vec{U}_{i+\frac{2}{3}}) < 0, \lambda(\vec{U}_{i+1}) >= 0 \\ f(\vec{U}_{i+\frac{2}{3}}) - f(\vec{U}_{i+1}) & \lambda(\vec{U}_{i+\frac{2}{3}}) < 0, \lambda(\vec{U}_{i+1}) < 0. \end{cases}$$

Finally, as mentioned in (180), the integration along the path associated with $\lambda = u$ will result in:

$$\int_{\vec{U}_{i+\frac{1}{3}}}^{\vec{U}_{i+\frac{2}{3}}} |A| d\vec{U} = \begin{cases} f(\vec{U}_{i+\frac{2}{3}}) - f(\vec{U}_{i+\frac{1}{3}}) & \lambda(\vec{U}_{i+\frac{1}{3}}) = \lambda(\vec{U}_{i+\frac{2}{3}}) \geq 0 \\ f(\vec{U}_{i+\frac{1}{3}}) - f(\vec{U}_{i+\frac{2}{3}}) & \lambda(\vec{U}_{i+\frac{1}{3}}) = \lambda(\vec{U}_{i+\frac{2}{3}}) < 0. \end{cases}$$

**6.3. Computing the numerical flux.** When written down explicitly, the formula for the numerical flux turns into a rather simple expression. For instance, when there are no sonic points along the integration paths and all eigenvalues are greater than zero, the formula for the numerical flux becomes:

$$f^{*(O)}_{i+\frac{1}{2}} = f_i.$$

Another example is when there is a sonic point in the integral from $U_i$ to $U_{i+\frac{1}{3}}$, with the eigenvector at $U_{i+\frac{1}{3}}$ greater than zero, $u < 0$ along the integration path from $U_{i+\frac{1}{3}}$ to $U_{i+\frac{2}{3}}$ and the eigenvalue, $\lambda = u - c$, along the final integration path smaller than zero:

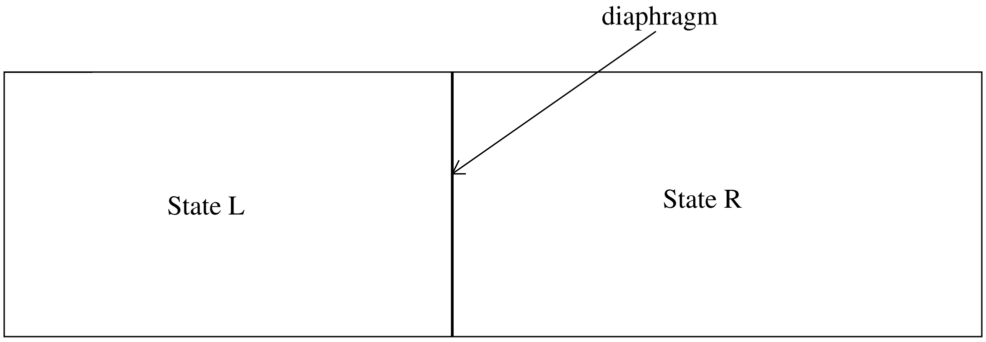
$$f^{*(O)}_{i+\frac{1}{2}} = f^*.$$

In the general case, the result for $f^{*(O)}_{i+\frac{1}{2}}$ depends only on the sign of the eigenvalues at the points $\vec{U}_i$, $\vec{U}_{i+\frac{1}{3}}$, $\vec{U}_{i+\frac{2}{3}}$ and $\vec{U}_{i+1}$.

| $\lambda(\vec{U}_i) = u + c$ | $\lambda(\vec{U}_{i+\frac{1}{3}}) = u + c$ | $\lambda(\vec{U}_{i+\frac{1}{3}}) = u$ | $\lambda(\vec{U}_{i+\frac{2}{3}}) = u - c$ | $\lambda(\vec{U}_{i+1}) = u - c$ |
|---|---|---|---|---|
| + | + | + | + | + |
| + | + | + | + | - |
| + | + | + | - | + |
| + | + | + | - | - |
| + | + | - | - | - |
| - | + | - | - | - |
| + | - | - | - | - |
| - | - | - | - | - |

Since $c$, described by (170), is always greater than zero, depending on the sign of the eigenvalue $\lambda = u$, **either** the integration path associated with the eigenvalue $\lambda = u + c$ or the one associated with $\lambda = u - c$ has the potential of containing a sonic point. As a consequence of this, as shown in the table above, there are only eight distinct cases that can occur. An assumption that is made here, is that none of the eigenvalues are equal to zero. Although this special case is straightforward to deal with, it has not been added to the table for reasons of simplicity. The formula for the numerical flux, using the cases in the table above, is now:

$$f^{*(O)}_{i+\frac{1}{2}} = \frac{1}{2} f_i + \frac{1}{2} f_{i+1} - \frac{1}{2} \sum_{k=0}^{2} sign(\lambda_k)[f_{i+\frac{k+1}{3}} - f_{i+\frac{k}{3}}]$$




initial state at t=0

$$= \frac{1}{2}f_i + \frac{1}{2}f_{i+1} - \frac{1}{2}(-)[f^* - f_i] - \frac{1}{2}(+)[f_{i+1} - f^*].$$

Here, $(+)$ and $(-)$ denote the sign of the eigenvalues between the sonic point and respectively $\vec{U}_{i+1}$ and $\vec{U}_i$. Note that this rule still applies if the sonic point is located outside of the integration path.

**6.4. The shock tube problem.** In the shock tube problem, also called the *Riemann problem*, a diaphragm, located at $x = x_0$, separating a gas at two different initial conditions $\rho$, $u$ and $p$ in a long, one-dimensional tube, bursts at time $t = 0$. The initial conditions are

$$\rho = \rho_L, \quad u = u_L, \quad p = p_L \qquad x < x_0 \quad t = 0 \tag{193}$$

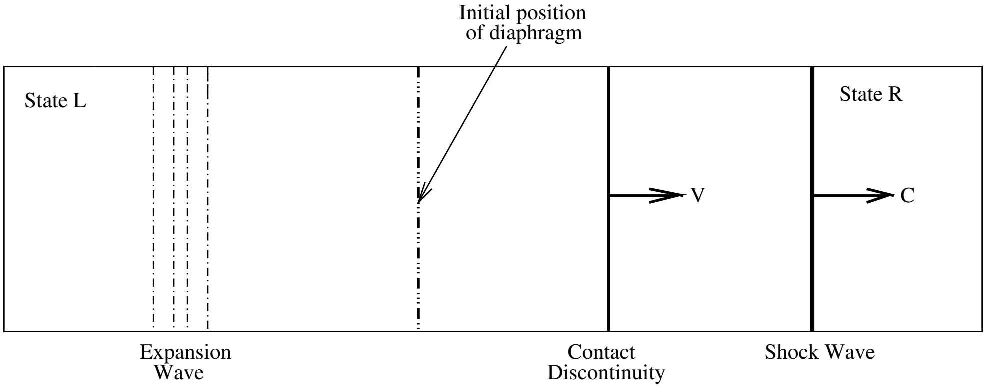
$$\rho = \rho_R, \quad u = u_R, \quad p = p_R \qquad x > x_0 \quad t = 0 \tag{194}$$

with $p_L > p_R$.

If an infinitely long tube is considered, which avoids reflections at the tube ends, and viscous effects along the tube walls are ignored, the solution to the Euler equations can be obtained on the basis of simple waves separating regions of uniform conditions.

At the bursting of the diaphragm, at $t = t_0$, a shock wave and a contact discontinuity will propagate to the right of the tube and an expansion wave will propagate to the left.

- Shock wave: The shock wave is a pressure discontinuity that separates the region with the undisturbed gas in state $\vec{U}_R$ from the region with the disturbed low-pressure gas. It propagates to the right and since it moves in regions of uniform conditions, it will have a constant *propagation speed*.
- Contact discontinuity: The contact discontinuity separates the region with the disturbed low-pressure gas from the region with the disturbed high-pressure gas. Since it moves in regions of uniform conditions, like the shock wave, it propagates at a constant velocity $V = u$. The pressure and velocities normal to its surface are continuous.
- Expansion wave: The expansion wave propagates to the left and separates the disturbed high pressure gas from the undisturbed high-pressure gas in state $\vec{U}_L$ on its left. The flow through the expansion wave varies continuously. The entropy remains constant.



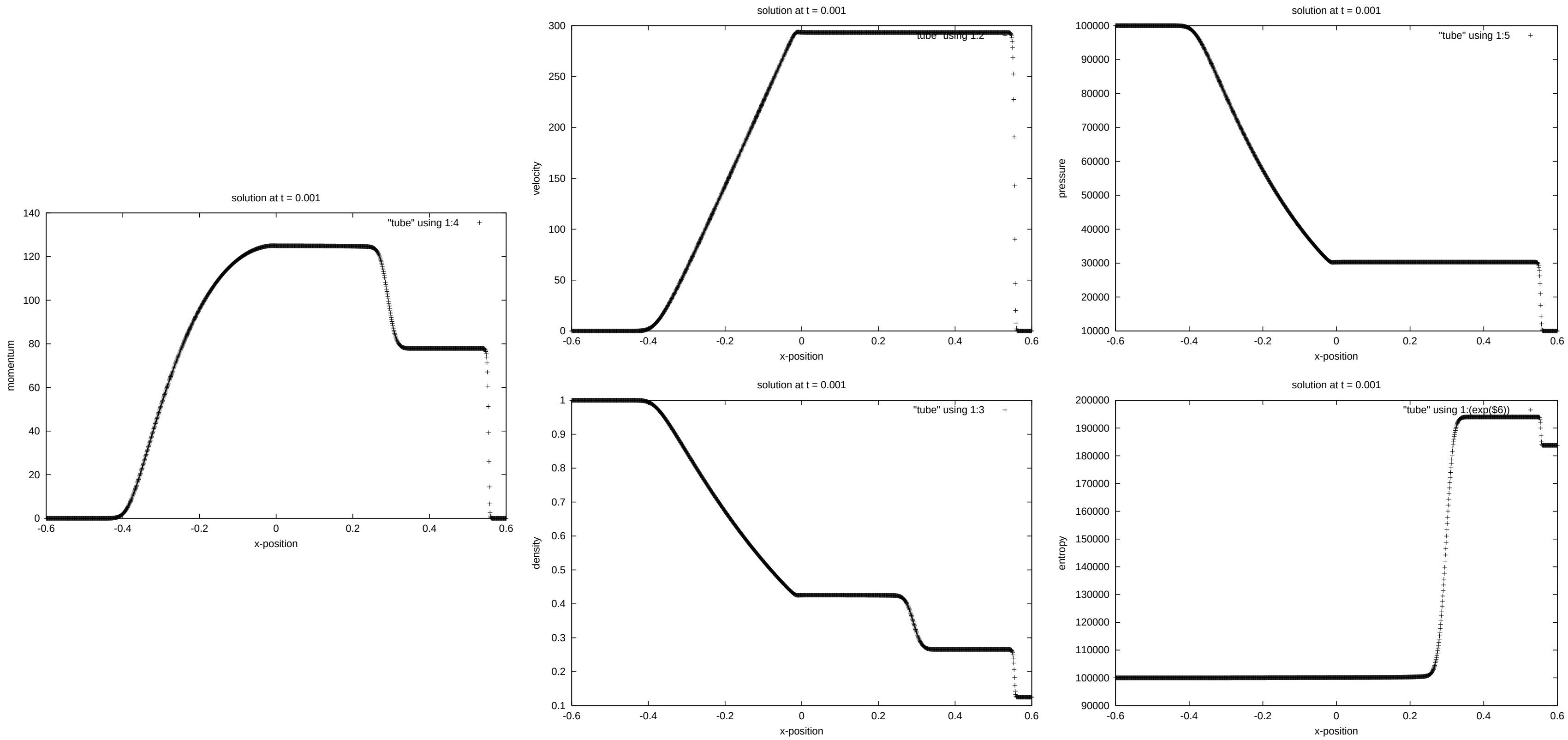

Flow at t>0

**6.5. Simulation of the shock tube problem.** Using Osher's approximate Riemann solver, described in section 6.2, it is possible to find a solution for the Euler equations that approaches the exact solution to an acceptable degree. For the gas in the tube, air is used, which has a value $\gamma = 1.4$ for the ratio of specific heats. The constant variable $x_0$ is set at $x_0 = 0$. For the initial conditions, the values shown below in SI units[4] are used:

$$p_L = 10^5, \quad \rho_L = 1.0, \quad u_L = 0, \quad p_R = 10^4, \quad \rho = 0.125, \quad u_R = 0,$$

corresponding with an initial pressure ratio of 10. These test data correspond to those applied by Sod (1978).

[4] A more extensive discussion of this example can be found in section 16.6.3 of Hirsch's book[**3**].



solution at t = 0.001
"tube" using 1:4
momentum
x-position
solution at t = 0.001
"tube" using 1:2
velocity
x-position
solution at t = 0.001
"tube" using 1:5
pressure
x-position
solution at t = 0.001
"tube" using 1:3
density
x-position
solution at t = 0.001
"tube" using 1:(exp($6))
entropy
x-position

# Bibliography


[1] P.W. Hemker: *Eight Lectures on Discretisation Principles and Hyperbolic Conservation Laws*, 13 december 1999

[2] J. Smoller: *Shock Waves and Reaction Diffusion Equations, Volume 258 of Grundlehren der Mathematischen Wissenschaften*, Springer Verlag, 1983

[3] C. Hirsch: *Numerical Computation of Internal and External Flows, Volume 2: Computational Methods for Inviscid and Viscous Flows*, J.Wiley, 1988