

Rainer Ansorge

Mathematical Models of Fluid Dynamics

Modelling, Theory, Basic Numerical Facts –
An Introduction

Rainer Ansorge

Mathematical Models of Fluidynamics

Modelling, Theory, Basic Numerical Facts –
An Introduction



WILEY-VCH GmbH & Co. KGaA

Author

Prof. Dr. Rainer Ansorge

University of Hamburg
Institute for Applied Mathematics

With 30 figures



This book was carefully produced. Nevertheless, authors, editors and publisher do not warrant the information contained therein to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

**Library of Congress Card No.: applied for
British Library Cataloging-in-Publication Data:
A catalogue record for this book is available
from the British Library**

**Bibliographic information published by
Die Deutsche Bibliothek**

Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <<http://dnb.ddb.de>>.

© 2003 WILEY-VCH GmbH & Co. KGaA,
Weinheim

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – nor transmitted or translated into machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

Printed in the Federal Republic of Germany
Printed on acid-free paper

Composition Uwe Krieg, Berlin

Printing Strauss Offsetdruck GmbH,
Mörlenbach

Bookbinding Großbuchbinderei J. Schäffer
GmbH & Co. KG, Grünstadt

ISBN 3-527-40397-3

Dedicated to my family members and to my students

Preface

Mathematical modelling is the process of replacing problems from outside mathematics by mathematical problems. The subsequent mathematical treatment of this model by theoretical and/or numerical procedures works as follows:

1. Transition from the outer-mathematics phenomenon to a mathematical description, what at the same time leads to a translation of problems formulated in terms of the original problem into mathematical problems.

This task forces the scientist or engineer who intends to use mathematical tools to

- cooperate with experts working in the field the original problem belongs to. Thus, he has to learn the language of these experts and to understand their way of thinking (teamwork).
 - create or to accept an idealized description of the original phenomena, i.e. an a-priori-neglection of properties of the original problems which are expected to be of no great relevance with respect to the questions under consideration. These simplifications are useful in order to reduce the degree of complexity of the model as well as of its mathematical treatment.
 - identify structures within the idealized problem and to replace these structures by suitable mathematical structures.
2. Treatment of the mathematical substitute.

This task normally requires

- independent activity of the theoretically working person.
- treatment of the problem by tools of mathematical theory.
- the solution of the particular mathematical problems occurring by the use of these theoretical tools, i.e. differential equations or integral equations, op-

timal control problems or systems of algebraic equations etc. have to be solved. Often numerical procedures are the only way to do this and to answer the particular questions under consideration at least approximately. The error of the approximate solution compared with the unknown so-called exact solution does normally not really affect the answer to the original problem, provided that the numerical method as well as the numerical tools are of sufficiently high accuracy. In this context, it should be realized that the question for an exact solution does not make sense because of the idealizations mentioned above and since the initial data presented with the original problem normally originate from statistics or from experimental measurements.

3. Retranslation of the results.

The qualitative and quantitative statements received from the mathematical model need now to be retranslated into the language by which the original problem was formulated. This means that the results have now to be interpreted with respect to their real-world-meaning. This process again requires teamwork with the experts from the beginning.

4. Model check up.

After retranslation, the results have to be checked with respect to their relevance and accuracy, e.g. by experimental measurements. This work has to be done by the experts of the original problem. If the mathematical results coincide sufficiently well with the results of experiments stimulated by the theoretical forecasts, the mathematical part is over and a new tool to help the physicists, engineers etc. in similar situations is born.

Otherwise – if no trivial logical or computational errors can be found – the model has to be revised. In this situation, the gap between mathematical and real results can only originate from too extensive idealizations in the modelling process.

The development of mathematical models does not only stimulate new experiments and does not only lead to constructive-prognostic – hence technical – tools for physicists or engineers but is also important from the point of view of the theory of cognition: It allows to understand connections between different elements out of the unstructured set of observations or – in other words – to create theories.

Fields of applications of mathematical descriptions are well known since centuries in physics, engineering, also in music etc. In modern biology, medicine, philology and economy mathematical models are used, too, and this even holds for certain fields of arts like oriental ornaments.

This book presents an introduction into models of fluid mechanics, leads to important properties of fluid flows which can theoretically be derived from the models and shows some basic ideas for the construction of effective numerical procedures. Hence, all aspects of theoretical fluid dynamics are addressed, namely modelling, mathematical theory and numerical methods.

We do not expect the reader to be familiar with a lot of experimental experiences. The knowledge of some fundamental principles of physics like conservation of mass, conservation of energy etc. is sufficient. The most important idealization is – contrary to the molecular structure of materials – the assumption of fluid-continua.

Concerning mathematics, it can help to understand the text more easily if the reader is acquainted with some basic elements of

- Linear Algebra
- Calculus
- Partial Differential Equations
- Numerical Analysis
- Theory of Complex Functions
- Functional Analysis.

Functional Analysis does only play a role in the somewhat general theory of discretization algorithms in chapter 6. In this chapter, also the question of the existence of weak entropy solutions of the problems under consideration is discussed. Physicists and engineers are normally not so much interested in the treatment of this problem. Nevertheless, it had to be included into this presentation in order not to leave this question open. The non-existence of a solution shows immediately that a model does not fit the reality if there is a measurable course of physical events. Existence theorems are therefore important not only from the point of view of mathematicians. But, of course, readers who are not acquainted with some functional analytic terminology may skip this chapter.

With respect to models and their theoretical treatment as well as with respect to numerical procedures occurring in sections 4.1, 5.3, 6.4 and chapter 7, a brief introduction into mathematical fluid mechanics like this book can only present basic facts. But the author hopes that this overview will make young scientists interested in this field and that it can help people working in institutes and industries to become more familiar with some fundamental mathematical aspects.

Finally, I wish to thank several colleagues for suggestions, particularly Thomas Sonar, who contributed to chapter 7 when we organized a joint course for graduate students¹, and Dr. Michael Breuss, who read the manuscript carefully. Last but not least I thank the publishers, especially Dr. Alexander Grossmann, for their encouragement.

Rainer Ansorge

Hamburg, September 2002

¹Parts of chapters 1 and 5 are translations from parts of sections 25.1, 25.2, 29.9 of: Ansorge and Oberle: *Mathematik für Ingenieure*, vol. 2, 2nd ed., Berlin, Wiley-VCH 2000.

Contents

Preface	7
1 Ideal Fluids	13
1.1 Modelling by Euler's Equations	13
1.2 Characteristics and Singularities	24
1.3 Potential Flows and (Dynamic) Buoyancy	30
1.4 Motionless Fluids and Sound Propagation	47
2 Weak Solutions of Conservation Laws	51
2.1 Generalization of what will be called a Solution	51
2.2 Traffic Flow Example with Loss of Uniqueness	56
2.3 The Rankine-Hugoniot Condition	62
3 Entropy Conditions	69
3.1 Entropy in Case of Ideal Fluids	69
3.2 Generalization of the Entropy Condition	74
3.3 Uniqueness of Entropy Solutions	80
3.4 The Ansatz due to Kruzkov	92
4 The Riemann Problem	97
4.1 Numerical Importance of the Riemann Problem	97
4.2 The Riemann Problem in the Case of Linear Systems	99
5 Real Fluids	103
5.1 The Navier-Stokes Equations Model	103
5.2 Drag Force and the Hagen-Poiseuille Law	111
5.3 Stokes Approximation and Artificial Time	116
5.4 Foundations of the Boundary Layer Theory; Flow Separation	122
5.5 Stability of Laminar Flows	131

6	Existence Proof for Entropy Solutions by Means of Discretization Procedures	135
6.1	Some Historical Remarks	135
6.2	Reduction to Properties of Operator Sequences	136
6.3	Convergence Theorems	140
6.4	Example	143
7	Types of Discretization Principles	151
7.1	Some General Remarks	151
7.2	The Finite Difference Calculus	156
7.3	The CFL Condition	160
7.4	Lax-Richtmyer Theory	162
7.5	The von Neumann Stability Criterion	168
7.6	The Modified Equation	171
7.7	Difference Schemes in Conservation Form	173
7.8	The Finite Volume Method on Unstructured Grids	176
	Some Extensive Monographs	181
	Index	183

1 Ideal Fluids

1.1 Modelling by Euler's Equations

Physical laws are mainly derived from conservation principles, such as conservation of mass, conservation of momentum, conservation of energy.

Let us consider a fluid (gas or liquid) in motion, i.e. the flow of a fluid¹, let

$$\mathbf{u}(x, y, z, t) = \begin{pmatrix} u_1(x, y, z, t) \\ u_2(x, y, z, t) \\ u_3(x, y, z, t) \end{pmatrix}$$

be the velocity and denote by $\rho = \rho(x, y, z, t)$ the density of this fluid at the point $\mathbf{x} = (x, y, z)$ and at the time instant t^2 .

Let us take out of the fluid at a particular instant t an arbitrary portion of volume $W(t)$ with surface $\partial W(t)$. The particles of the fluid now move, and assume $W(t+h)$ to be the volume at the instant $t+h$ formed by the same particles which had formed $W(t)$ at the time t .

Moreover, let $\varphi = \varphi(x, y, z, t)$ be one of the functions describing a particular state of the fluid at the time t at the point \mathbf{x} such as mass per unit volume (= density), interior energy per volume, momentum per volume etc. Hence, $\int_{W(t)} \varphi d(x, y, z)$ gives the full amount of mass or interior energy, momentum etc. of the volume $W(t)$ under consideration.

¹Flows of other materials are eventually included, too, e.g. flow of cars on highways, provided that the density of cars or particles is sufficiently high.

²Bold letters normally mean vectors or matrices.

We ask for the change of $\int_{W(t)} \varphi d(x, y, z)$ with respect to time, i.e. for

$$\frac{d}{dt} \int_{W(t)} \varphi(x, y, z, t) d(x, y, z). \quad (1.1)$$

We have

$$\begin{aligned} & \frac{d}{dt} \int_{W(t)} \varphi(x, y, z, t) d(x, y, z) \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \left\{ \int_{W(t+h)} \varphi(\tilde{\mathbf{y}}, t+h) d(y_1, y_2, y_3) - \int_{W(t)} \varphi(x, y, z, t) d(x, y, z) \right\}, \end{aligned}$$

where the change from $W(t)$ to $W(t+h)$ is obviously given by the mapping

$$\tilde{\mathbf{y}} = \mathbf{x} + h \cdot \mathbf{u}(\mathbf{x}, t) + \mathbf{o}(h)$$

$$(\tilde{\mathbf{y}} = (y_1, y_2, y_3)^T).$$

The error term $\mathbf{o}(h)$ also depends on \mathbf{x} but keeps the property $\lim_{h \rightarrow 0} \frac{1}{h} \mathbf{o}(h) = \mathbf{0}$ if differentiated with respect to space provided that these spatial derivatives are bounded.

The transformation of the integral taken over the volume $W(t+h)$ to an integral over $W(t)$ by substitution requires the integrand to be multiplied by the determinant of this mapping, i.e. by

$$\begin{aligned} & \begin{vmatrix} (1 + h \partial_x u_1) & h \partial_y u_1 & h \partial_z u_1 \\ h \partial_x u_2 & (1 + h \partial_y u_2) & h \partial_z u_2 \\ h \partial_x u_3 & h \partial_y u_3 & (1 + h \partial_z u_3) \end{vmatrix} + o(h) \\ &= 1 + h \cdot (\partial_x u_1 + \partial_y u_2 + \partial_z u_3) + o(h) \\ &= 1 + h \cdot \operatorname{div} \mathbf{u}(\mathbf{x}, t) + o(h). \end{aligned}$$

Taylor expansion of $\varphi(\tilde{\mathbf{y}}, t + h)$ around (\mathbf{x}, t) therefore leads to

$$\begin{aligned} \frac{d}{dt} \int_{W(t)} \varphi(x, y, z, t) d(x, y, z) \\ = \int_{W(t)} \{ \partial_t \varphi + \varphi \operatorname{div} \mathbf{u} + \langle \mathbf{u}, \nabla \varphi \rangle \} d(x, y, z). \end{aligned} \quad (1.2)$$

Herewith, ∇v denotes the gradient of a scalar function v , and $\langle \cdot, \cdot \rangle$ means the standard scalar product of two vectors out of \mathbb{R}^3 .

The product rule of differentiation gives

$$\varphi \operatorname{div} \mathbf{u} + \langle \mathbf{u}, \nabla \varphi \rangle = \operatorname{div}(\varphi \cdot \mathbf{u}),$$

such that (1.2) leads to the so-called **Reynolds' transport theorem**³

$$\frac{d}{dt} \int_{W(t)} \varphi(x, y, z, t) d(x, y, z) = \int_{W(t)} \{ \partial_t \varphi + \operatorname{div}(\varphi \mathbf{u}) \} d(x, y, z). \quad (1.3)$$

As already mentioned, the dynamics of fluids can be described directly by conservation principles and – as far as gases are concerned – by an additional equation of state.

1. Conservation of mass: If there are no sources or losses of fluid within the subdomain of the flow under consideration, mass remains constant.

Because $W(t)$ and $W(t + h)$ consist of the same particles, they are of the same mass. The mass of $W(t)$ is given by $\int_{W(t)} \rho(x, y, z, t) d(x, y, z)$, and therefore

$$\frac{d}{dt} \int_{W(t)} \rho(x, y, z, t) d(x, y, z) = 0$$

must hold. Taking (1.3) into account (particularly for $\varphi = \rho$), this leads to the requirement

$$\int_{W(t)} \{ \partial_t \rho + \operatorname{div}(\rho \mathbf{u}) \} d(x, y, z) = 0.$$

³Osborne Reynolds (1842 - 1912); Manchester

Since this has to hold for arbitrary $W(t)$, the integrand has to vanish:

$$\boxed{\partial_t \rho + \operatorname{div}(\rho \mathbf{u}) = 0}. \quad (1.4)$$

This equation is called the **continuity equation**.

2. Conservation of momentum: Another conservation principle concerns the momentum of a mass which is defined as

$$\text{mass} \times \text{velocity}.$$

Thus,

$$\int_{W(t)} \rho \mathbf{u} d(x, y, z)$$

gives the momentum of the mass at time t of the volume $W(t)$ and

$$\mathbf{q} = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = \rho \mathbf{u}$$

describes the **density of momentum**.

The principle of conservation of momentum, i.e. Newton's second law

$$\text{force} = \text{mass} \times \text{acceleration},$$

now states that the change of momentum with respect to time equals the sum of all exterior forces acting on the mass of $W(t)$.

In order to describe these exterior forces, we take into account that there exists a certain pressure $p(\mathbf{x}, t)$ at every point \mathbf{x} of the fluid and at every instant t . If \mathbf{n} is considered to be the unit vector normal on the surface $\partial W(t)$ of $W(t)$ and of outward direction, the fluid outside of $W(t)$ leads to a force acting on $W(t)$ which is given by

$$- \int_{\partial W(t)} p \mathbf{n} do \quad (do = \text{area element of } \partial W(t)).$$

Besides the normal forces per unit of the surface area generated by the pressure, there exist also tangential forces which act on the surface by friction of the exterior particles

along the surface. Though this so-called **viscosity** of a fluid leads to a lot of remarkable phenomena, we are going to neglect this property at the first step: Instead of **real fluids** or **viscous fluids** we restrict ourselves in this chapter to so-called **ideal fluids** or **inviscid fluids**. This restriction to ideal fluids, particularly to **ideal gases**, is one of the idealizations mentioned in the preface.

But there are not only exterior forces per unit of the surface area but also exterior forces per unit volume, e.g. the weight.

Let us denote these forces per unit volume by \mathbf{k} such that Newton's second law leads to

$$\frac{d}{dt} \int_{W(t)} \mathbf{q} d(x, y, z) = \int_{W(t)} \mathbf{k}(x, y, z, t) d(x, y, z) - \int_{\partial W(t)} p \cdot \mathbf{n} do.$$

Thus, by Gauss' divergence theorem, we find

$$\begin{aligned} \int_{\partial W(t)} p \mathbf{n} do &= \begin{pmatrix} \int_{\partial W(t)} p n_1 do \\ \int_{\partial W(t)} p n_2 do \\ \int_{\partial W(t)} p n_3 do \end{pmatrix} = \begin{pmatrix} \int_{W(t)} \partial_x p d(x, y, z) \\ \int_{W(t)} \partial_y p d(x, y, z) \\ \int_{W(t)} \partial_z p d(x, y, z) \end{pmatrix} \\ &= \int_{W(t)} \nabla p d(x, y, z). \end{aligned}$$

Together with (1.3),

$$\int_{W(t)} \left\{ \partial_t \mathbf{q} + \begin{pmatrix} \operatorname{div}(q_1 \mathbf{u}) \\ \operatorname{div}(q_2 \mathbf{u}) \\ \operatorname{div}(q_3 \mathbf{u}) \end{pmatrix} - \mathbf{k} + \nabla p \right\} d(x, y, z) = \mathbf{0}$$

follows.

Again, this has to be valid for every arbitrarily chosen volume $W(t)$. If, moreover,

$$\operatorname{div}(q_i \mathbf{u}) = \langle \mathbf{u}, \nabla q_i \rangle + \operatorname{div} \mathbf{u} \cdot q_i$$

is taken into account,

$$\partial_t \mathbf{q} + \langle \mathbf{u}, \nabla \rangle \mathbf{q} + \operatorname{div} \mathbf{u} \cdot \mathbf{q} + \nabla p = \mathbf{k},$$

i.e.

$$\partial_t \mathbf{q} + \frac{1}{\rho} \langle \mathbf{q}, \nabla \rangle \mathbf{q} + \operatorname{div} \left(\frac{1}{\rho} \mathbf{q} \right) \mathbf{q} + \nabla p = \mathbf{k} \quad (1.5)$$

has to be fulfilled.

The number of equations represented by (1.5) equals the spatial dimension of the flow, i.e. the number of components of \mathbf{q} or \mathbf{u} .

By means of the continuity equation, (1.5) can be reformulated as

$$\partial_t \mathbf{u} + \langle \mathbf{u}, \nabla \rangle \mathbf{u} + \frac{1}{\rho} \nabla p = \hat{\mathbf{k}}, \quad (1.6)$$

where the force \mathbf{k} per unit volume has been replaced by the force $\hat{\mathbf{k}} = \frac{1}{\rho} \mathbf{k}$ per unit mass.

Equation (1.4) together with (1.5) or (1.6) are called **Euler's equations**⁴.

$$\rho, E, q_1, q_2, q_3$$

are sometimes called **conservative variables** whereas $\rho, \varepsilon, u_1, u_2, u_3$ are the **primitive variables**. Herewith,

$$E := \rho \varepsilon + \frac{\rho}{2} \|\mathbf{u}\|^2 = \rho \varepsilon + \frac{\|\mathbf{q}\|^2}{2\rho}$$

gives the **total energy** per unit volume where ε stands for the **interior energy** per unit mass, e.g. heat per unit mass.

$$\frac{\rho}{2} \|\mathbf{u}\|^2$$

obviously introduces the **kinetic energy** per unit volume⁵.

$\langle \mathbf{u}, \nabla \rangle \mathbf{u}$ is called the **convection term**.

Remark:

Terms of the form $\partial_t \mathbf{w} + \langle \mathbf{w}, \nabla \rangle \mathbf{w}$ are in the literature often abbreviated by $\frac{D\mathbf{w}}{Dt}$ and are called **material time derivatives** of the vector valued function \mathbf{w} .

⁴Leonhard Euler (1707 - 1783); Basel, Berlin, St. Petersburg

⁵ $\|\cdot\|$: 2-norm.

Next we consider the **First Law of Thermodynamics**, namely the

3. Conservation of energy.

The change per time unit of the total energy E of the mass of a moving fluid volume equals the work which is done per time unit against the exterior forces.

This means in the case of ideal fluids

$$\begin{aligned} \frac{d}{dt} \int_{W(t)} E(x, y, z, t) d(x, y, z) \\ = \int_{W(t)} \langle \rho \hat{\mathbf{k}}, \mathbf{u} \rangle d(x, y, z) - \int_{\partial W(t)} \langle p \mathbf{n}, \mathbf{u} \rangle do \quad ^6. \end{aligned}$$

The relation

$$\int_{\partial W(t)} \langle p \mathbf{n}, \mathbf{u} \rangle do = \int_{\partial W(t)} \langle p \mathbf{u}, \mathbf{n} \rangle do = \int_{W(t)} \operatorname{div}(p \mathbf{u}) d(x, y, z)$$

follows from the Gauss' divergence theorem, so that (1.3) leads to

$$\int_{W(t)} \left\{ \partial_t E + \operatorname{div}(E \cdot \mathbf{u}) + \operatorname{div}(p \cdot \mathbf{u}) - \langle \rho \hat{\mathbf{k}}, \mathbf{u} \rangle \right\} d(x, y, z) = 0 \quad \forall W(t).$$

If \mathbf{k} can be neglected because of small gas weight or if $\hat{\mathbf{k}}$ is the weight of the fluid per unit mass⁷ and \mathbf{u} the velocity of a flow parallel to the earth surface, we get

$$\boxed{\partial_t E + \operatorname{div} \left(\frac{E+p}{\rho} \mathbf{q} \right) = 0}. \quad (1.7)$$

⁶work = force \times length = pressure \times area \times length $\implies \frac{\text{work}}{\text{time}} = \text{force} \times \text{velocity} = \text{pressure} \times \text{area} \times \text{velocity}$.

⁷i.e. $\|\hat{\mathbf{k}}\| = g$ with the earth acceleration g .

Explicitly written and neglecting \mathbf{k} , the equations (1.4), (1.5) and (1.7) read as

$$\begin{aligned}
 \partial_t \rho + \partial_x q_1 & & + \partial_y q_2 & & + \partial_z q_3 & & = 0 \\
 \partial_t q_1 + \partial_x \left(\frac{1}{\rho} q_1 q_1 + p \right) & & + \partial_y \left(\frac{1}{\rho} q_1 q_2 \right) & & + \partial_z \left(\frac{1}{\rho} q_1 q_3 \right) & & = 0 \\
 \partial_t q_2 + \partial_x \left(\frac{1}{\rho} q_2 q_1 \right) & & + \partial_y \left(\frac{1}{\rho} q_2 q_2 + p \right) & & + \partial_z \left(\frac{1}{\rho} q_2 q_3 \right) & & = 0 \\
 \partial_t q_3 + \partial_x \left(\frac{1}{\rho} q_3 q_1 \right) & & + \partial_y \left(\frac{1}{\rho} q_3 q_2 \right) & & + \partial_z \left(\frac{1}{\rho} q_3 q_3 + p \right) & & = 0 \\
 \partial_t E + \partial_x \left(\frac{E+p}{\rho} q_1 \right) & & + \partial_y \left(\frac{E+p}{\rho} q_2 \right) & & + \partial_z \left(\frac{E+p}{\rho} q_3 \right) & & = 0.
 \end{aligned} \tag{1.8}$$

Hence, we are concerned with a system of equations to be used in order to determine the functions ρ , q_1 , q_2 , q_3 , E . But we have to realize that there is an additional sought function, namely the pressure p . In case of constant density ρ^8 , i.e. $\partial_t \rho = 0 \quad \forall (x, y, z)$, only four conservation variables have to be determined such that the five equations (1.8) are sufficient. Otherwise, particularly on event of gas flow, a sixth equation is needed, namely an equation of state. State variables of a gas are

$$\begin{aligned}
 T &= \text{temperature}, & p &= \text{pressure}, & \rho &= \text{density}, & V &= \text{volume}, \\
 \varepsilon &= \text{energy/mass}, & S &= \text{entropy/mass},
 \end{aligned}$$

and a theorem of thermodynamics says that each of these state variables can uniquely be expressed in terms of two of the other state variables. Such relations between three state variables are called **equations of state**. Thus, p can be expressed by ρ and ε (hence by ρ, E, \mathbf{q}); for inviscid so-called γ -gases the relation is given by

$$p = (\gamma - 1) \rho \varepsilon = (\gamma - 1) \left(E - \frac{\|\mathbf{q}\|^2}{2\rho} \right) \tag{1.9}$$

with $\gamma = \text{const} > 1$, such that only the functions ρ , \mathbf{q} , E have to be determined. Herewith, γ is the ratio $\frac{c_p}{c_v}$ of the specific heats. If air is concerned, we have $\gamma \approx 1.4$.

By means of vector valued functions, also systems of differential equations can be described by a single differential equation. In this way and taking (1.9) into account

⁸The case of constant density is not necessarily identical with the case of an **incompressible** flow defined by $\text{div } \mathbf{u} = 0$ because the continuity equation (1.4) is in this situation already fulfilled if the pair (ρ, \mathbf{u}) shows the property $\partial_t \rho + \langle \mathbf{u}, \nabla \rho \rangle = 0$, which does not necessarily imply $\rho = \text{const}$.

as an additional equation, (1.8) can be written as

$$\partial_t \mathbf{V} + \partial_x \mathbf{f}_1(\mathbf{V}) + \partial_y \mathbf{f}_2(\mathbf{V}) + \partial_z \mathbf{f}_3(\mathbf{V}) = \mathbf{0} \quad (1.10)$$

with

$$\mathbf{V} = (\rho, q_1, q_2, q_3, E)^T$$

and with

$$\mathbf{f}_1(\mathbf{V}) = \begin{pmatrix} q_1 \\ \frac{1}{\rho} q_1 q_1 + p \\ \frac{1}{\rho} q_2 q_1 \\ \frac{1}{\rho} q_3 q_1 \\ \frac{E+p}{\rho} q_1 \end{pmatrix}, \mathbf{f}_2(\mathbf{V}) = \begin{pmatrix} q_2 \\ \frac{1}{\rho} q_2 q_1 \\ \frac{1}{\rho} q_2 q_2 + p \\ \frac{1}{\rho} q_3 q_2 \\ \frac{E+p}{\rho} q_2 \end{pmatrix}, \mathbf{f}_3(\mathbf{V}) = \begin{pmatrix} q_3 \\ \frac{1}{\rho} q_3 q_1 \\ \frac{1}{\rho} q_3 q_2 \\ \frac{1}{\rho} q_3 q_3 + p \\ \frac{E+p}{\rho} q_3 \end{pmatrix}.$$

The functions $\mathbf{f}_j(\mathbf{V})$ are called **fluxes**.

If $\mathbf{J} \mathbf{f}_1$, $\mathbf{J} \mathbf{f}_2$, $\mathbf{J} \mathbf{f}_3$ are the **Jacobians**, respectively, (1.10) reads as

$$\partial_t \mathbf{V} + \mathbf{J} \mathbf{f}_1(\mathbf{V}) \cdot \partial_x \mathbf{V} + \mathbf{J} \mathbf{f}_2(\mathbf{V}) \cdot \partial_y \mathbf{V} + \mathbf{J} \mathbf{f}_3(\mathbf{V}) \cdot \partial_z \mathbf{V} = \mathbf{0}. \quad (1.11)$$

Because of their particular meaning in physics, systems of differential equations of type (1.10) are called systems of **conservation laws**, even if they do not originate from physical aspects. Obviously, (1.11) is a quasilinear systems of partial differential equations of first order.

Such systems have to be completed by initial conditions

$$\mathbf{V}(\mathbf{x}, 0) = \mathbf{V}_0(\mathbf{x}), \quad (1.12)$$

where $\mathbf{V}_0(\mathbf{x})$ is a prescribed initial state, and by boundary conditions.

If the flow does not depend on time, it is called a **stationary** or **steady-state** flow and initial conditions do not occur. Otherwise, the flow is called an **instationary** one.

As far as ideal fluids are concerned, it can be assumed that the fluid flow is tangential along the surface of a solid body⁹ fixed in space or along the bank of a river etc. In this case,

$$\langle \mathbf{u}, \mathbf{n} \rangle = 0 \quad (1.13)$$

is one of the boundary conditions where \mathbf{n} are the outward normal unit vectors along the surface of the body.

Often there occur symmetries with respect to space such that the number of unknowns can be reduced, e.g. in the case of rotational symmetry combined with polar coordinates. If there remains only one spatial coordinate, the problem is called **one dimensional**¹⁰. If rectangular space variables are used and if x is the only remaining one, we end up with the system

$$\partial_t \mathbf{V} + \mathbf{J} \mathbf{f}(\mathbf{V}) \cdot \partial_x \mathbf{V} = \mathbf{0} \quad (1.14)$$

with

$$\mathbf{V} = \begin{pmatrix} \rho \\ q \\ E \end{pmatrix}, \quad \mathbf{f}(\mathbf{V}) = \begin{pmatrix} q \\ \frac{1}{\rho} q^2 + p \\ \frac{E+p}{\rho} q \end{pmatrix}$$

and with the equation of state

$$p = (\gamma - 1) \left(E - \frac{q^2}{2\rho} \right).$$

Hence,

$$\mathbf{J} \mathbf{f}(\mathbf{V}) = \begin{pmatrix} 0 & 1 & 0 \\ -\frac{3-\gamma}{2} \frac{q^2}{\rho^2} & (3-\gamma) \frac{q}{\rho} & \gamma - 1 \\ (\gamma - 1) \frac{q^3}{\rho^3} - \gamma \frac{E q}{\rho^2} & \gamma \frac{E}{\rho} - \frac{3(\gamma - 1)}{2} \frac{q^2}{\rho^2} & \gamma \frac{q}{\rho} \end{pmatrix}.$$

⁹e.g. of a wing

¹⁰Two or three dimensional problems are defined analogously.

As can easily be verified, $\lambda_1 = \frac{q}{\rho}$ is a solution of the characteristic equation of $\mathbf{J} \mathbf{f}(\mathbf{V})$ namely of

$$-\lambda^3 + 3 \frac{q}{\rho} \lambda^2 - \left\{ (\gamma^2 - \gamma + 6) \frac{q^2}{2\rho^2} - \gamma(\gamma - 1) \frac{E}{\rho} \right\} \lambda + \left(\frac{\gamma^2}{2} - \frac{\gamma}{2} + 1 \right) \frac{q^3}{\rho^3} - \gamma(\gamma - 1) \frac{E q}{\rho^2} = 0$$

The other two eigenvalues of $\mathbf{J} \mathbf{f}(\mathbf{V})$ are then the roots of

$$\begin{aligned} \lambda^2 - \frac{2q}{\rho} \lambda + (\gamma^2 - \gamma + 2) \frac{q^2}{2\rho^2} - \gamma(\gamma - 1) \frac{E}{\rho} &= 0 : \\ \lambda_{2,3} &= \frac{q}{\rho} \pm \sqrt{\frac{q^2}{\rho^2} + \gamma(\gamma - 1) \frac{E}{\rho} - (\gamma^2 - \gamma + 2) \frac{q^2}{2\rho^2}} \\ &= \frac{q}{\rho} \pm \sqrt{\gamma(\gamma - 1) \frac{E}{\rho} - \gamma(\gamma - 1) \frac{q^2}{2\rho^2}} \\ &= \frac{q}{\rho} \pm \sqrt{\frac{\gamma(\gamma - 1)}{\rho} \left[E - \frac{q^2}{2\rho} \right]} = \frac{q}{\rho} \pm \sqrt{\gamma \frac{p}{\rho}} . \end{aligned} \tag{1.15}$$

Thus, these eigenvalues are real and different from each other ($p > 0$).

Definition:

If all the eigenvalues of a matrix $\mathbf{A}(x, t, \mathbf{V})$ are real and if the matrix can be diagonalized, a system of equations

$$\partial_t \mathbf{V} + \mathbf{A}(x, t, \mathbf{V}) \partial_x \mathbf{V} = 0$$

*is called of **hyperbolic** type at (x, t, \mathbf{V}) . If the eigenvalues are real and different from each other so that \mathbf{A} can certainly be diagonalized, the system is called **strictly hyperbolic**.*

So, obviously, (1.14) is strictly hyperbolic for all (x, t, \mathbf{V}) under consideration.

By the way, because of $\frac{q}{\rho} = u$, the velocity of the flow, and because $\sqrt{\frac{\gamma p}{\rho}}$ describes the local sound velocity \hat{c} (cf. (1.67)), the eigenvalues are

$$\lambda_1 = u, \quad \lambda_2 = u + \hat{c}, \quad \lambda_3 = u - \hat{c},$$

and have equal signs in case of supersonic flow whereas the subsonic flow is characterized by the fact that one eigenvalue shows a different sign than the others.

Definition:

Let $\mathbf{u}(\mathbf{x}, t_0)$ be the velocity field of a flow at the instant t_0 and let \mathbf{x}_0 be an arbitrary point of the particular subset of \mathbb{R}^3 which is occupied by the fluid at this particular instant. If at this moment the system

$$[\mathbf{x}'(s), \mathbf{u}(\mathbf{x}(s), t_0)] = \mathbf{0}^{11}$$

of ordinary differential equations with the initial condition

$$\mathbf{x}(0) = \mathbf{x}_0$$

has a unique solution $\mathbf{x} = \mathbf{x}(s)$ for each of these points \mathbf{x}_0 , each of the curves $\mathbf{x} = \mathbf{x}(s)$ is called a **streamline** at the instant t_0 .

Here, the streamlines may be parametrised by the arc length s with $s = 0$ at the point \mathbf{x}_0 , such that

$$\langle \mathbf{x}', \mathbf{x}' \rangle = 1.$$

Thus, the set of streamlines at an instant t_0 shows a snapshot of the flow at this particular instant. It does not necessarily describe the trajectories along which the fluid particles move when time goes on. Only if the flow is a stationary one, i.e. if \mathbf{u} , ρ , p , $\hat{\mathbf{k}}$ are independent of time, streamlines and trajectories coincide: A particle moves along a fixed streamline in the course of time.

1.2 Characteristics and Singularities

As an introductory example for a more general investigation of conservation laws, let us consider the scalar **Burgers' equation**¹² without exterior forces

$$\partial_t v + \partial_x \left(\frac{1}{2} v^2 \right) = 0, \quad (x, t) \in \Omega, \quad \text{i.e. } x \in \mathbb{R}, t \geq 0, \quad (1.16)$$

which is often studied as a model problem from a theoretical point of view and also as a test problem for numerical procedures.

¹¹where $[\cdot, \cdot]$ is the vector product in \mathbb{R}^3

¹²J. Burgers: Nederl. Akad. van Wetenschappen **43** (1940), 2-12.

Here, the flux is given by $f(v) = \frac{1}{2}v^2$. If

$$v_0(x) = 1 - \frac{x}{2} \tag{1.16a}$$

is chosen as a particular example of an initial condition, the **unique** and **smooth solution**¹³ turns out to be

$$v(x, t) = \frac{2 - x}{2 - t} ,$$

but this solution does only exist locally, namely for $0 \leq t < 2$; for increasing time it runs into a singularity at $t = 2$.

As a matter of fact, classical existence and uniqueness theorems for quasilinear partial differential equations of first order with smooth coefficients ensure the unique existence of a classical smooth solution only in a certain neighbourhood of the initial manifold, provided that also the initial condition is sufficiently smooth.

The occurrence of discontinuities or singularities does not depend on the smoothness of the fluxes: Assume $v(x, t)$ to be a smooth solution of the problem

$$\partial_t v + \partial_x f(v) = 0 \quad \text{for } x \in \mathbb{R}, t \geq 0$$

$$v(x, 0) = v_0(x)$$

in a certain neighbourhood immediately above the x -axis. Obviously, this solution is constant along the straight line

$$x(t) = x_0 + t f'(v_0(x_0)) \tag{1.17}$$

that crosses the x -axis at x_0 where $x_0 \in \mathbb{R}$ is chosen arbitrarily. Its value along this line is therefore $v(x(t), t) = v_0(x_0)$. This can easily be verified by

$$\begin{aligned} \frac{d}{dt} v(x(t), t) &= \partial_t v(x(t), t) + \partial_x v(x(t), t) \cdot x'(t) \\ &= \partial_t v(x, t) + \partial_x v(x, t) \cdot f'(v_0(x_0)) \\ &= \partial_t v(x, t) + \partial_x v(x, t) \cdot f'(v(x, t)) = \partial_t v + \partial_x f(v) = 0 . \end{aligned}$$

These straight lines (1.17), each of which belonging to a particular x_0 , are called the **characteristics** of the given conservation law.

¹³i.e. v is continuously differentiable

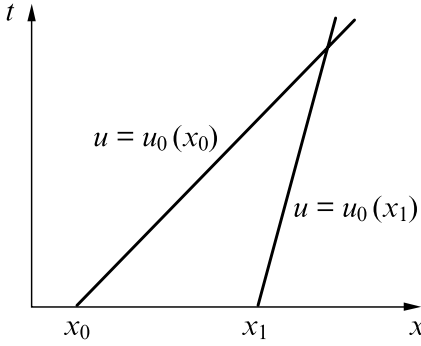


Figure 1: Formation of discontinuities

In case of $x_0 < x_1$, but e.g. $0 < f'(v_0(x_1)) < f'(v_0(x_0))$ ¹⁴, the characteristic through $(x_0, 0)$ intersects the characteristic through $(x_1, 0)$ at an instant $t_1 > 0$, so that at the point of intersection the solution v has to have the value $v_0(x_0)$ as well as the value $v_0(x_1) \neq v_0(x_0)$. Therefore, a discontinuity will occur at the instant t_1 or even earlier. With respect to fluid dynamics, particularly shocks – discussed later on – belong to the various types of discontinuities.

Applied to Burgers' equation (1.16) with the initial function (1.16a), the characteristic through a point x_0 of the x -axis is given by

$$t = 2 \frac{x - x_0}{2 - x_0},$$

such that the discontinuity we had found at $t = 2$ can also immediately be understood by means of Figure 2.

If systems of conservation laws are considered instead of the scalar case, i.e.

$$\mathbf{V}(x, t) \in \mathbb{R}^m, \quad m \in \mathbb{N}, \quad \forall (x, t) \in \Omega,$$

and if only one spatial variable x occurs, a system of characteristics is defined as follows:

Definition:

If $\mathbf{V}(x, t)$ is a solution of (1.10) in the case of only one space variable x , the one-parameter set

$$x_{(i)} = x_{(i)}(t, \chi) \quad (\chi \in \mathbb{R} : \text{set parameter})$$

¹⁴hence also $v_0(x_0) \neq v_0(x_1)$

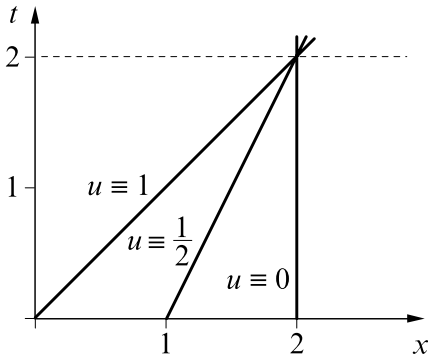


Figure 2: Discontinuity of the solution of the Burgers equation

of real curves defined by the ordinary differential equation

$$\dot{x} = \lambda_i(\mathbf{V}(x, t)) \quad (1.18)$$

for every fixed i ($i = 1, \dots, m$) is called the set of **i-characteristics** of the particular system that belongs to \mathbf{V} . Here, $\lambda_i(\mathbf{V})$ ($i = 1, \dots, m$) are the eigenvalues of the Jacobian $\mathbf{J}f(\mathbf{V})$.

Obviously, this definition coincides in case of $m = 1$ with the previously presented definition of characteristics.

Let us finally and exemplarily study the situation of a system of conservation laws if the system is linear with constant coefficients:

$$\partial_t \mathbf{V} + \mathbf{A} \partial_x \mathbf{V} = \mathbf{0} \quad (1.19)$$

with a constant Jacobian (m, m) -matrix \mathbf{A} . Moreover, let us assume the system to be strictly hyperbolic.

From (1.18), the characteristics turn out to be the set of straight lines given by

$$x_{(i)}(t) = \lambda_i t + x_{(i)}(0), \quad (i = 1, \dots, m),$$

and independent of \mathbf{V} . Hence, for every fixed i , the characteristics belonging to this set are parallel.

If $\mathbf{S} = (\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m)$ is the matrix whose columns consist of the eigenvectors of the Jacobian and if $\mathbf{\Lambda}$ denotes the diagonal matrix consisting of the eigenvalues of \mathbf{A} ,

$$\mathbf{A} = \mathbf{S} \mathbf{\Lambda} \mathbf{S}^{-1}$$

follows.

If new variables $\hat{\mathbf{V}}$ are introduced by

$$\hat{\mathbf{V}} = \mathbf{S}^{-1} \mathbf{V} \quad ,$$

the system gets the form

$$\partial_t \hat{\mathbf{V}} + \mathbf{\Lambda} \partial_x \hat{\mathbf{V}} = \mathbf{0} \quad , \quad \hat{\mathbf{V}}_0 = \mathbf{S}^{-1} \mathbf{V}_0 \quad . \quad (1.21)$$

This is a decoupled system:

$$\partial_t \hat{v}_i + \lambda_i \partial_x \hat{v}_i = 0 \quad , \quad \hat{v}_i(x, 0) = [\mathbf{S}^{-1} \mathbf{V}_0(x)]_i = \hat{v}_{i_0}(x) \quad , \quad (i = 1, \dots, m) \quad .$$

Each of the equations of this system is an independent scalar equation called an **advection equation**. The solution is

$$\hat{v}_i(x, t) = \hat{v}_{i_0}(x - \lambda_i t) \quad (i = 1, \dots, m) \quad . \quad (1.22)$$

Hence, the state at the instant t moves with velocity λ_i into positive or negative x -direction according to the sign of λ_i , respectively. This is called a wave propagation where the propagation velocity is described by λ_i .

Obviously,

$$\mathbf{V} = (\mathbf{s}_1 \ \dots \ \mathbf{s}_m) \begin{pmatrix} \hat{v}_1 \\ \vdots \\ \hat{v}_m \end{pmatrix} = \hat{v}_1 \mathbf{s}_1 + \hat{v}_2 \mathbf{s}_2 + \dots + \hat{v}_m \mathbf{s}_m \quad .$$

Each of the vector valued functions

$$\mathbf{V}_{(i)}(x, t) \quad := \quad \hat{v}_i(x, t) \cdot \mathbf{s}_i$$

solves the system of differential equations because of

$$\begin{aligned} \partial_t \mathbf{V}_{(i)} + \mathbf{S} \mathbf{\Lambda} \mathbf{S}^{-1} \partial_x \mathbf{V}_{(i)} &= \partial_t \hat{v}_i \mathbf{s}_i + \partial_x \hat{v}_i \lambda_i \mathbf{s}_i \\ &= (\partial_t \hat{v}_i + \lambda_i \partial_x \hat{v}_i) \mathbf{s}_i = \mathbf{0} \quad . \end{aligned}$$

$\mathbf{V}_{(i)}$ ($i = 1, \dots, m$) is often called the *solution belonging to the i -th set of characteristics and to the given initial value $\hat{v}_{i_0}(x)$* . Obviously, the vector functions $\mathbf{V}_{(i)}$ ($i = 1, \dots, m$) are linearly independent.

Remark:

The fact that a sufficiently smooth solution of a nonlinear initial value problem does

often only exist in the neighbourhood of the initial manifold – though the correspondent real world process shows a global existence– can't be accepted by a physicist or engineer. He insists the mathematician to present a global solution. This forces the mathematicians to create a more general definition of a solution such that the real world situation can be described in a satisfactory way. Suitable definitions of **weak solutions** will be presented in chapter 2.

Remark:

As far as the scalar linear problem

$$\partial_t v + \partial_x(a v) = 0 \quad , \quad a = \text{const} \quad , \quad (1.23)$$

is concerned, v often describes a concentration, and the flux then is simply given by $f(v) = a v$.

Many physical processes include a further flux of the particular form

$$-\varepsilon \partial_x v \quad (\varepsilon > 0) \quad ,$$

proportional to the concentration descent. The transport phenomenon is then mathematically modelled by

$$\partial_t v + \partial_x(a v - \varepsilon \partial_x v) = 0, \quad \text{i.e.} \quad \partial_t v + a \partial_x v = \varepsilon \partial_{xx} v. \quad (1.24)$$

Because of the **diffusion term** on the right hand side, this equation is of parabolic type. Parabolic equations make their solutions smoother and smoother when time t increases. Therefore, shocks will be smeared as soon as diffusion occurs.

This effect of parabolic equations can easily be demonstrated by examples like the following one: The sum

$$v(x, t) = e^{\frac{a}{2}(x - \frac{a}{2}t)} \sum_{\nu=1}^{\infty} (-1)^{\nu+1} \frac{\sin(\nu x)}{\nu} e^{-\nu^2 t}$$

converges uniformly for $t > 0$ because of the factors $e^{-\nu^2 t}$. This also holds after multiple termwise differentiations of this sum with respect to t as well as with respect to x so that v is a sufficiently smooth function for all $x \in \mathbb{R}$, $t > 0$. It solves (1.24)¹⁵ in the classical understanding, but leads for $t = 0$ to the function

$$v_0(x) = e^{\frac{a}{2}x} \cdot w(x)$$

¹⁵in case of $\varepsilon = 1$

where $w(x)$ represents the Fourier expansion of the 2π -periodic discontinuous function

$$w(x) = \begin{cases} \frac{x}{2} & \text{for } -\pi < x < \pi \\ 0 & \text{for } x = \pm\pi \end{cases} .$$

The curve described by $w(x)$ is sometimes called a saw blade curve.

1.3 Potential Flows and (Dynamic) Buoyancy

Let us try to investigate the forces acting on solid bodies¹⁶ dipped into a fluid flow at a fixed position. For convenience, we restrict ourselves to stationary flows of inviscid fluids. Moreover, it will be assumed that the magnitudes of the velocities are such that the density can be regarded as a constant. Thus, the flow is incompressible and the partial derivatives with respect to t occurring in (1.8) vanish. The first equation in (1.8) if written in primitive variables, i.e. the continuity equation (1.4), then reduces to

$$\operatorname{div} \mathbf{u} = 0, \quad (1.25)$$

where $\mathbf{u} = (u_1(x, y, z), u_2(x, y, z), u_3(x, y, z))^T$ ¹⁷ again stands for the velocity vector of the flow at the space position $\mathbf{x} = (x, y, z)^T$.

The second, third, and fourth equation of (1.8) formulated by means of primitive variables could be written as the vector valued Euler equation (1.6) and lead in case of a stationary flow to

$$\langle \mathbf{u}, \nabla \rangle \mathbf{u} + \frac{1}{\rho} \nabla p = \hat{\mathbf{k}} . \quad (1.26)$$

Let us –moreover– assume that the flow is **irrotational**; this means that the **circulation**

$$Z := \oint_C \mathbf{u}(\mathbf{x}) d\mathbf{x} \quad (1.27)$$

vanishes for every closed contour C within every simply connected subdomain of the flow area. Then \mathbf{u} can be derived from a potential ϕ i.e. there is a scalar function ϕ so that

$$\mathbf{u} = \nabla \phi, \quad (1.28)$$

¹⁶e.g. on a wing of an aircraft

¹⁷For the time being, z means the third space variable; it will later denote the complex variable $x + iy$, but confusion will certainly not occur.

and

$$\operatorname{curl} \mathbf{u} = \mathbf{0} \quad (1.29)$$

holds within this area.

By the way, the vector $\operatorname{curl} \mathbf{u}$ is often called the **vorticity vector** or **angular velocity vector**, and a trajectory of a field of vorticity vectors is called a **vortex line**.

Because of (1.28), a flow of this type is called a **potential flow**, and ϕ is the so-called **velocity potential**.

The fifth equation in (1.8) can be omitted as far as the knowledge of the energy density E is of no interest.

From (1.29), the relation $\langle \mathbf{u}, \nabla \rangle \mathbf{u} = \frac{1}{2} \nabla (\|\mathbf{u}\|^2) - [\mathbf{u}, \operatorname{curl} \mathbf{u}]$ leads together with (1.26) to

$$\nabla \left(\frac{1}{2} \|\mathbf{u}\|^2 + \frac{1}{\rho} p \right) = \hat{\mathbf{k}}. \quad (1.30)$$

Formula (1.30) shows that a fluid flow of the particular type under consideration, namely an approximately stationary, inviscid, incompressible, and irrotational flow can only exist if the exterior forces are the gradient of a scalar function. In other words: these forces have to be conservative, i.e. there has to exist a potential Q with $\hat{\mathbf{k}} = -\nabla Q$, such that (1.30) leads to

$$\nabla \left(\frac{1}{2} \|\mathbf{u}\|^2 + \frac{1}{\rho} p + Q \right) = \mathbf{0}.$$

i.e. to

$$\frac{1}{2} \|\mathbf{u}\|^2 + \frac{1}{\rho} p + Q = \text{const.} \quad (1.31)$$

(1.31) is called the **Bernoulli equation**¹⁸ and is nothing else than the energy conservation law for this particular type of flow. p is called the **static pressure**, whereas the term $\frac{\rho}{2} \|\mathbf{u}\|^2$ i.e. the kinetic energy per volume, is often called the **dynamic pressure**.

As far as an incompressible flow in a circular pipe is concerned, the velocity will necessarily increase as soon as the diameter of the pipe decreases, and – because of

¹⁸Daniel Bernoulli (1700 - 1782); St. Peterburg, Basle

(1.31) – this will lead to a decreasing pressure within the narrow part of the pipe. This phenomenon is called the **hydrodynamic paradoxon**. Applications are carburettors, jet streams etc.

Remark:

A necessary condition for the existence of the irrotational flow was the conservative character of the exterior forces. Let us now assume that –vice versa– these forces are conservative, i.e. $\hat{\mathbf{k}} = -\nabla Q$. Moreover, let us admit that the flow is even compressible in the sense of a particular dependence $\rho = \rho(p)$, called a **barotropic flow**. Integration of (1.26) along a streamline from a constant point P_0 to a variable point P then leads to

$$\int_{P_0}^P \left\{ \frac{1}{2} \nabla (\|\mathbf{u}\|^2) + [\text{curl } \mathbf{u}, \mathbf{u}] + \nabla \Theta + \nabla Q \right\} ds = \mathbf{0},$$

with

$$\Theta := \int \frac{dp}{\rho(p)},$$

hence

$$\nabla \Theta = \frac{1}{\rho} \nabla p,$$

and with

$$ds = \frac{1}{\|\mathbf{u}\|} \mathbf{u} ds \quad (s = \text{arc length}).$$

Because P was arbitrary and because of $\langle [\text{curl } \mathbf{u}, \mathbf{u}], \mathbf{u} \rangle = 0$, this result leads to the **generalized Bernoulli equation**

$$\frac{1}{2} \|\mathbf{u}\|^2 + \Theta + Q = \text{const},$$

which in the case of constant density, i.e. $\Theta = \frac{p}{\rho}$, seems to coincide completely with (1.31). But one has to take into account that the constant at the righthand side may

now change from streamline to streamline. Additionally, we find

$$\begin{aligned}
 \frac{d}{dt}Z &= \frac{d}{dt} \oint_C \mathbf{u}(\mathbf{x}) d\mathbf{x} = \oint_C \frac{d}{dt} \left\langle \mathbf{u}, \frac{d\mathbf{x}}{ds} \right\rangle ds \\
 &= \oint_C \frac{d\mathbf{u}}{dt} d\mathbf{x} + \oint_C \mathbf{u} d\mathbf{u} \\
 &= \oint_C \frac{d\mathbf{u}}{dt} d\mathbf{x} = \oint_C \{ \langle \mathbf{u}, \nabla \rangle \mathbf{u} \} d\mathbf{x} \\
 &= - \oint_C \left\{ \frac{1}{\rho} \nabla p + \nabla Q \right\} d\mathbf{x} \\
 &= - \oint_C \nabla (\Theta + Q) d\mathbf{x} = 0.
 \end{aligned}$$

This is **Kelvin's Theorem**¹⁹, which says that the circulation along a closed curve in an inviscid barotropic flow does not change in time.

If irrotational flows of arbitrary type of incompressibility are under consideration, (1.28) leads together with (1.25) to

$$\Delta\phi = 0 \quad (1.32)$$

whereas the continuity equation (1.4) reads for compressible fluids in the case of stationary flows as

$$\operatorname{div}(\rho\mathbf{u}) = 0.$$

Let us additionally assume the flow to be barotropic. Then, because of

$$\operatorname{div}(\rho\mathbf{u}) = \rho \operatorname{div} \mathbf{u} + \frac{d\rho}{dp} \langle \nabla p, \mathbf{u} \rangle$$

and because of

$$\hat{c} = \sqrt{\frac{dp}{d\rho}} \quad (\text{cf. (1.68)}),$$

(1.26) leads to

$$\begin{aligned}
 &\left(1 - \left(\frac{u_1}{\hat{c}}\right)^2\right) \partial_{xx}\phi + \left(1 - \left(\frac{u_2}{\hat{c}}\right)^2\right) \partial_{yy}\phi + \left(1 - \left(\frac{u_3}{\hat{c}}\right)^2\right) \partial_{zz}\phi \\
 &\quad - 2 \left(\frac{u_1 u_2}{\hat{c}^2} \cdot \partial_y u_1 + \frac{u_2 u_3}{\hat{c}^2} \cdot \partial_z u_2 + \frac{u_3 u_1}{\hat{c}^2} \cdot \partial_x u_3 \right) = 0, \quad (1.33)
 \end{aligned}$$

¹⁹Lord Kelvin of Largs (1824 - 1907); Glasgow

as far as the exterior forces vanish. In this situation, (1.33) generalizes (1.32).

Obviously, (1.33) is a quasilinear partial differential equation of second order for ϕ , certainly elliptic if

$$M := \frac{\|\mathbf{u}\|}{\hat{c}} < 1 \quad , \quad (1.34)$$

i.e. in areas of **subsonic flow**.

Definition: M is called the **Mach number**²⁰.

If particularly a constant flow into x -direction is under consideration which is only disturbed in the neighbourhood of a slim airfoil²¹ with a small angle of attack, products of the values u_i ($i = 2, 3$) with each other and with u_1 can be neglected compared with 1. This leads to $\|\mathbf{u}\|^2 = u_1^2$ and to a shortened version of (1.33), namely to

$$\left(1 - \left(\frac{u_1}{\hat{c}}\right)^2\right) \partial_{xx}\phi + \partial_{yy}\phi + \partial_{zz}\phi = 0 \quad . \quad (1.35)$$

This equation is hyperbolic in areas of **supersonic flow** ($M > 1$), and also the full equation (1.33) is of hyperbolic type in this case.

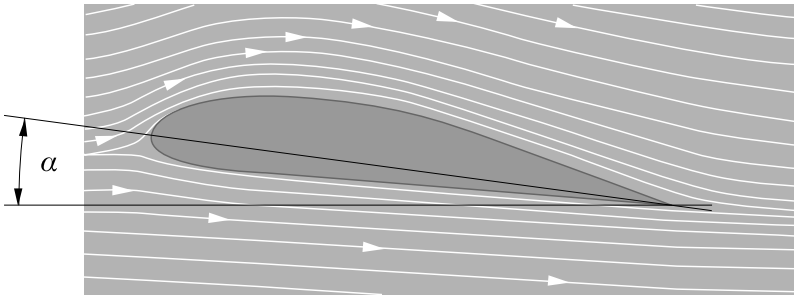


Figure 3: Flow around a wing if the angle of attack is small

Let us extend our idealizations by the assumption that the flow under consideration is a two-dimensional **plane flow**. This means that one of the components of the velocity vector \mathbf{u} in a rectangular coordinate system, e.g. the component in z -direction,

²⁰Ernst Mach (1838 - 1916); Graz, Praha, Vienna

²¹cross-section of a wing or another rigid body in a plane parallel to the direction of the flow

vanishes for all $(x, y) \in \mathbb{R}^2$ and for all $t \geq 0$:

$$u_3 = 0. \quad (1.36)$$

In case of a two-dimensional supersonic flow along a slim airfoile for which (1.35) holds, **Mach's angle** β , determined from

$$|\sin \beta| = \frac{1}{M} < 1$$

describes the angle between the characteristics of the wave equation

$$(1 - M^2)\partial_{xx}\phi + \partial_{yy}\phi = 0$$

and the flow direction given by the direction of the x -axis. The set of all these characteristics is called **Mach's net**. This net plays an important role as far as so-called **methods of characteristics** are used in order to establish efficient numerical procedures.

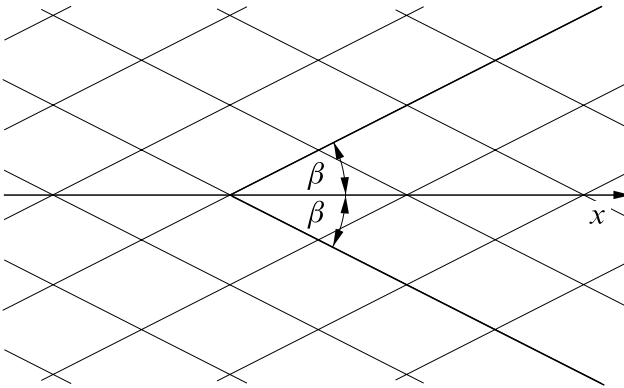


Figure 4: Mach's net in the case of a linearized supersonic flow ($M = \text{const} > 1$)

The two-dimensional case of rotational symmetry, e.g. flow along a projectile, leads correspondingly to **Mach's cone**.

In the two-dimensional plane situation, (1.29) reads as

$$\text{curl } \mathbf{u} = \begin{pmatrix} -\partial_z u_2 \\ \partial_z u_1 \\ \partial_x u_2 - \partial_y u_1 \end{pmatrix} = \mathbf{0}.$$

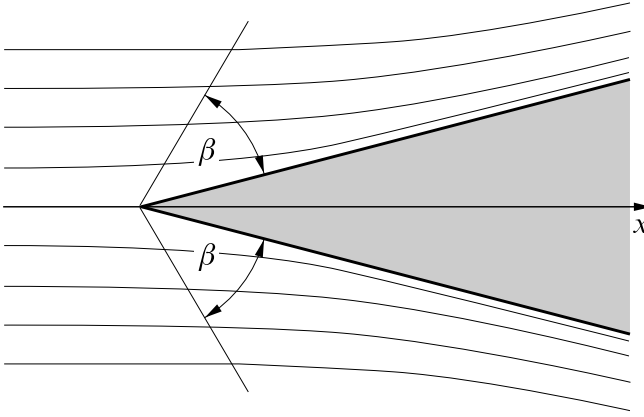


Figure 5: Linearized supersonic flow around a slim rotatory cone

Thus, u_1 and u_2 are independent of the third spatial variable z :

$$\mathbf{u} = \begin{pmatrix} u_1(x, y) \\ u_2(x, y) \\ 0 \end{pmatrix}$$

with

$$\partial_x u_2 - \partial_y u_1 = 0. \quad (1.37)$$

After introduction of the vector

$$\mathbf{v} := \begin{pmatrix} -u_2(x, y) \\ u_1(x, y) \\ 0 \end{pmatrix} \quad (1.38)$$

and by means of the continuity equation (1.25), we receive

$$\text{curl } \mathbf{v} = \begin{pmatrix} 0 \\ 0 \\ \partial_x u_1 + \partial_y u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \text{div } \mathbf{u} \end{pmatrix} = \mathbf{0},$$

such that also \mathbf{v} can be derived from a potential in simply connected parts of the fluid area. In other words, there is a scalar function $\psi = \psi(x, y)$ with

$$-u_2 = \partial_x \psi, \quad u_1 = \partial_y \psi. \quad (1.39)$$

ψ is called the **stream function**.

(1.28) leads to

$$\partial_x \phi = \partial_y \psi, \quad \partial_y \phi = -\partial_x \psi. \quad (1.40)$$

Obviously, (1.40) can be interpreted as the system of **Cauchy–Riemann equations** of the complex function

$$\Omega(z) := \phi(x, y) + i \psi(x, y) \quad (1.41)$$

depending on the complex variable $z = x + i y$.

Ω is called the **complex velocity potential** of the plane potential flow under consideration.

We are going to assume that the first partial derivatives of the functions ϕ and ψ are continuous such that Ω turns out to be a holomorphic function, and we intent to study the forces acting on rigid bodies when dipped into such a fluid flow.

Because we reduced the reality to a plane flow, the rigid body is assumed to be very long with respect to the direction of the third spatial variable; more precisely, it has to be of infinite length from the point of view of mathematics. Therefore, the flow around the contour of the cross section of the body within the (x, y) -plane, e.g. around the contour of an airfoile of a long wing, is of interest.

We stated that the circulation $Z = \oint_C \mathbf{u} \, d\mathbf{x}$ vanishes as long as the closed contour C is the boundary of a simply connected domain within the fluid area in the case of a potential flow. Because of our plane model, C represents a simply closed contour in the (x, y) -plane, i.e. in \mathbb{R}^2 .

But now, a rigid body with its boundary Γ is dipped into the fluid, e.g. a wing. If its airfoil is part of the area surrounded by C , this interior domain of C is no longer a simply connected domain of the fluid area. Hence, the circulation Z around the airfoil does not necessarily vanish but turns out to fulfill

$$\oint_C \mathbf{u} \, d\mathbf{x} = \oint_{\Gamma} \mathbf{u} \, d\mathbf{x}.$$

In order to proof this relation, we take into account that Γ and C can in a first step be connected by two auxiliary lines in such a way that two simply connected domains G_1 and G_2 with the contours C_1 and C_2 , respectively, will occur. Because of $\text{curl } \mathbf{u} = \mathbf{0}$ in G_1 as well as in G_2 , we see in a second step that

$$\oint_{C_1} \mathbf{u} \, d\mathbf{x} = 0 \quad \text{as well as} \quad \oint_{C_2} \mathbf{u} \, d\mathbf{x} = 0$$

hold. This leads to

$$\oint_{C_1} \mathbf{u} dx + \oint_{C_2} \mathbf{u} dx = \mathbf{0}$$

(cf. Figure 6).

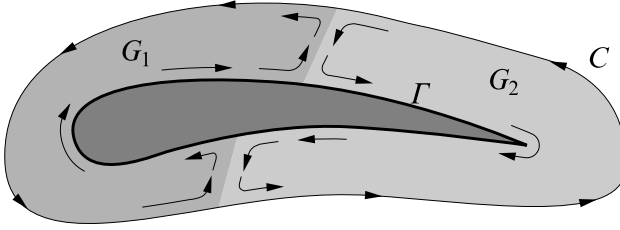


Figure 6: Auxiliary step in the computation of the buoyancy generated by plane potential flows

We realize in a third step that the integrations back and forth along each of the auxiliary lines extinguish each other such that

$$\oint_{C_1} \mathbf{u} dx + \oint_{C_2} \mathbf{u} dx = \oint_C \mathbf{u} dx + \oint_{-\Gamma} \mathbf{u} dx$$

results. Here, $-\Gamma$ denotes the contour of the rigid body if run through in negative direction, and this ends the proof.

Remark:

It has to be mentioned that the proof given is simplified because Γ does not only pass through the interior of the domain occupied by the fluid but is part of its boundary. As a matter of fact, one really has first, roughly speaking, to investigate the case where Γ is replaced by a line Γ_ε of distance ε from Γ and passing only through the fluid, and in a next step one has to study the limit situation $\varepsilon \rightarrow 0$.

Let Γ now be parametrized by $\mathbf{r} = \mathbf{r}(\tau)$, $0 \leq \tau \leq T$. Then

$$Z = \int_0^T \langle \mathbf{u}(\mathbf{r}(\tau)), \dot{\mathbf{r}}(\tau) \rangle d\tau = \int_0^T \{u_1 \dot{x} + u_2 \dot{y}\} d\tau \quad (1.42)$$

holds, and $\dot{\mathbf{r}} = \begin{pmatrix} \dot{x}(\tau) \\ \dot{y}(\tau) \end{pmatrix}$ is a vector tangential to the curve Γ at the point $(x(\tau), y(\tau))$ such that

$$\mathbf{n} := \frac{1}{\sqrt{\dot{x}^2 + \dot{y}^2}} \begin{pmatrix} \dot{y} \\ -\dot{x} \end{pmatrix} \quad (1.43)$$

is a normal unit vector of Γ at this point.

According to (1.13), we assume that the velocity of the flow at the surface of the rigid body is tangential to this surface:

$$\langle \mathbf{u}, \mathbf{n} \rangle = 0 \quad \text{along } \Gamma,$$

i.e. $u_1 \dot{y} = u_2 \dot{x}$. From this and with (1.42), the circulation becomes

$$Z = \int_0^T (u_1 - i u_2) (\dot{x} + i \dot{y}) d\tau,$$

i.e.

$$Z = \int_{\Gamma} w(z) dz, \quad (1.44)$$

with

$$w(z) := u_1(x, y) - i u_2(x, y). \quad (1.45)$$

$w(z)$ is a holomorphic function in the whole domain outside the airfoil because (1.39) yields

$$\partial_x u_1 = \partial_{xy} \psi = \partial_{yx} \psi = \partial_y(-u_2),$$

and because (1.28) leads to

$$\partial_y u_1 = \partial_{xy} \phi = \partial_{yx} \phi = -\partial_x(-u_2),$$

i.e. the Cauchy–Riemann equations for the function $w(z)$ are fulfilled.

Hence, also with respect to the computation of the circulation, it is allowed to integrate along C instead of Γ :

$$Z = \int_C w(z) dz. \quad (1.46)$$

Here, we choose C in such a way that it lies in the annulus between two concentric circles where K_r of radius r surrounds the airfoil and where K_R is a circle with sufficiently great arbitrary radius $R > r$. Without loss of generality, we assume the center of the circles to be the origin of the former (x, y) -plane which now became the complex z -plane.

$w(z)$ can be represented in the annulus – and therefore particularly for every $z \in C$ – by a Laurent series²² around the center $z_0 = 0$ (cf. Figure 7):

$$w(z) = \sum_{\nu=-\infty}^{+\infty} a_\nu z^\nu .$$

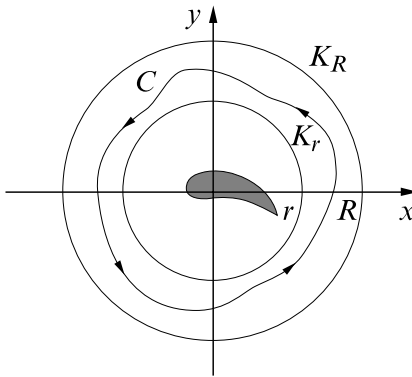


Figure 7: Annulus around an airfoil

We know by experience that the fluid flow is often influenced by the rigid body only in a certain neighbourhood of the contour; the flow around a ship that crosses a calm lake with constant velocity gives an example. From this point of view we are going to assume that the velocity $\mathbf{u} = \mathbf{u}(x, y)$ is constant for $|z| \rightarrow \infty$:

$$\lim_{|z| \rightarrow \infty} \mathbf{u}(x, y) = \begin{pmatrix} u_{1,\infty} \\ u_{2,\infty} \end{pmatrix} \quad (1.47)$$

with constant components $u_{1,\infty}$ and $u_{2,\infty}$ such that

$$\lim_{|z| \rightarrow \infty} w(z) = u_{1,\infty} - i u_{2,\infty} .$$

²²Pierre Alphonse Laurent (1813 - 1854); Le Havre

Because the coefficients a_ν ($\nu = 0, \pm 1, \pm 2, \dots$) are constant numbers which do not depend on R , and because R is allowed to tend to infinity, we find for $|z| \rightarrow \infty$ that all the a_ν with positive index have to vanish.

Hence, the result

$$\lim_{|z| \rightarrow \infty} w(z) = a_0 = u_{1,\infty} - i u_{2,\infty}. \quad (1.48)$$

follows.

On the other hand, Cauchy's formula

$$a_{-1} = \frac{1}{2\pi i} \oint_C w(\zeta) d\zeta,$$

leads to

$$Z = 2\pi i a_{-1}. \quad (1.49)$$

The (twodimensional) force \mathbf{K} to be calculated caused by the flow and acting on the rigid body turns out to become

$$\mathbf{K} = \begin{pmatrix} K_1 \\ K_2 \end{pmatrix} = - \int_0^L p \mathbf{n} ds \quad \left([\mathbf{K}] = \frac{[\text{force}]}{[\text{length}]} \right), \quad (1.50)$$

where \mathbf{n} is the unit normal vector from (1.43), p denotes the pressure along the surface of the body caused by the fluid, s measures the arc length along Γ beginning at an arbitrary point of it, and where L is the total length of Γ . Let us study this force separately from the other exterior forces acting on the body, i.e. let us assume that the sum of these other forces vanishes. This leads to a constant potential U in Bernoulli's equation (1.31) and therefore to a constant **total pressure**

$$p_0 := \frac{\rho}{2} \|\mathbf{u}\|^2 + p.$$

Because Γ is a closed contour, $\int_0^L \mathbf{n} ds = \mathbf{0}$ ²³. Therefore,

$$\mathbf{K} = - \int_0^L \left\{ p_0 - \frac{\rho}{2} \|\mathbf{u}\|^2 \right\} \mathbf{n} ds = \frac{\rho}{2} \int_0^L \|\mathbf{u}\|^2 \mathbf{n} ds,$$

i.e.

$$K_1 = \frac{\rho}{2} \int_0^T \frac{1}{\sqrt{\dot{x}^2 + \dot{y}^2}} \|\mathbf{u}\|^2 \dot{y} \frac{ds}{d\tau} d\tau,$$

$$K_2 = -\frac{\rho}{2} \int_0^T \frac{1}{\sqrt{\dot{x}^2 + \dot{y}^2}} \|\mathbf{u}\|^2 \dot{x} \frac{ds}{d\tau} d\tau.$$

The temporary introduction of the complex number

$$k = K_2 + i K_1 \tag{1.51}$$

yields

$$k = -\frac{\rho}{2} \int_0^T \frac{1}{\sqrt{\dot{x}^2 + \dot{y}^2}} \|\mathbf{u}\|^2 (\dot{x} - i \dot{y}) \frac{ds}{d\tau} d\tau = -\frac{\rho}{2} \int_0^T \|\mathbf{u}\|^2 (\dot{x} - i \dot{y}) d\tau.$$

The complex number within the parantheses on the righthand side of

$$\begin{aligned} \|\mathbf{u}\|^2 (\dot{x} - i \dot{y}) &= (u_1^2 + u_2^2) (\dot{x} - i \dot{y}) = (u_1 - i u_2) (u_1 + i u_2) (\dot{x} - i \dot{y}) \\ &= w(z) (u_1 \dot{x} + u_2 \dot{y} + i u_2 \dot{x} - i u_1 \dot{y}) \end{aligned}$$

is not really complex because the imaginary part vanishes (cf. (1.13)). It can therefore be replaced by its conjugate complex number:

$$\begin{aligned} \|\mathbf{u}\|^2 (\dot{x} + i \dot{y}) &= w(z) (u_1 \dot{x} + u_2 \dot{y} - i u_2 \dot{x} + i u_1 \dot{y}) \\ &= w(z) (u_1 - i u_2) (\dot{x} + i \dot{y}) = w^2(z) \dot{z}. \end{aligned}$$

²³ $\frac{ds}{d\tau} = \sqrt{\dot{x}^2 + \dot{y}^2}$, hence

$$\int_0^L \mathbf{n} ds = \int_0^T \frac{1}{\sqrt{\dot{x}^2 + \dot{y}^2}} \begin{pmatrix} \dot{y} \\ -\dot{x} \end{pmatrix} \frac{ds}{d\tau} d\tau = \begin{pmatrix} y(T) - y(0) \\ -x(T) + x(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This leads to

$$k = -\frac{\rho}{2} \int_0^T w^2(z) \dot{z} d\tau = -\frac{\rho}{2} \int_{\Gamma} w^2(z) dz.$$

Because w is holomorph, also w^2 is a holomorphic function. Hence,

$$k = -\frac{\rho}{2} \int_C w^2(z) dz. \quad (1.52)$$

But within the annulus we have

$$\begin{aligned} w^2(z) &= \left(a_0 + \frac{a_{-1}}{z} + \frac{a_{-2}}{z^2} + \dots \right) \left(a_0 + \frac{a_{-1}}{z} + \frac{a_{-2}}{z^2} + \dots \right) \\ &= a_0^2 + 2 \frac{a_0 a_{-1}}{z} + \frac{A_{-2}}{z^2} + \frac{A_{-3}}{z^3} + \dots \end{aligned}$$

with certain coefficients A_ν ($\nu = -2, -3, \dots$).

Cauchy's residuum formula and (1.52) therefore lead to

$$k = -\frac{\rho}{2} 2a_0 a_{-1} \cdot 2\pi i.$$

Thus, (1.48) and (1.49) yield

$$k = -\rho (u_{1,\infty} - i u_{2,\infty}) Z,$$

and because the circulation Z is real (cf. (1.27)), the comparison of the real and imaginary parts of (1.51) results in

$$\begin{aligned} K_1 &= \rho u_{2,\infty} Z \\ K_2 &= -\rho u_{1,\infty} Z. \end{aligned} \quad (1.53)$$

If particularly $u_{2,\infty} = 0$ but $u_{1,\infty} \neq 0$, i.e. if the undisturbed flow is parallel to the x -axis, and if $Z \neq 0$, we find a **lift** $K_2 \neq 0$, i.e. there is a force acting on the rigid body perpendicular to the direction of the flow.

By suitable construction of wings, the airfoils can lead to $Z < 0$ so that an aircraft can start against its weight, a hydrofoil can lift in the water etc. Of course, also Archimedes' static buoyancy, given by (1.62), has additionally to be taken into account.

Remark:

The formulas (1.53) are called **Kutta–Zhukovsky buoyancy formulas**.²⁴

Remark:

Whereas the first equation in (1.53) describes the buoyancy at least qualitatively in a correct manner, the result $K_1 = 0$ contradicts real experiences: Also in case of an incompressible irrotational stationary flow –as more or less realized by calmly running little rivers– the flow leads to some force affecting the rigid body –e.g. a bridge pier– parallel to the flow. Of course, this contradiction results from one of our idealisations, namely from the assumption of an inviscid and therefore frictionless fluid.

In order to understand what really happens parallel to the flow, we have to reduce these idealizations. This will later be done by the transition from the Euler equations to the so-called Navier-Stokes equations and from (1.13) to the so-called **no-slip condition**

$$\mathbf{u} = \mathbf{0} \text{ along } \Gamma . \quad (1.54)$$

(1.54) expresses the idea that moving viscous fluids leave a monomolecular non-moving layer along impermeable walls of solid bodies because of adhesion. Thus, friction along such surfaces does not mean friction between the fluid and the solid material but always just friction between fluid particles.

Of course, there are also other boundary conditions where **partial slip** on the surfaces occur, e.g. in case of rarefied gas flow, porous walls etc. The tangential component of the flow is then proportional to the local shear stress.

If (1.33) is reduced to the case of an irrotational, stationary, plane flow, i.e. $u_3 = 0$, $\partial_y u_1 = \partial_x u_2$, if there are no exterior forces, and if (u_1, u_2) is for convenience replaced by (u, v) , we get

$$\left(1 - \left(\frac{u}{\hat{c}}\right)^2\right) \partial_x u + \left(1 - \left(\frac{v}{\hat{c}}\right)^2\right) \partial_y v - \frac{u v}{\hat{c}^2} \cdot (\partial_y u + \partial_x v) = 0 . \quad (1.55)$$

The u - v -plane is called the **hodograph plane**.

We assume the equations

$$u = u(x, y)$$

$$v = v(x, y)$$

to be invertible such that x and y can be expressed by u and v , i.e.

$$D := \begin{vmatrix} \partial_x u & \partial_y u \\ \partial_x v & \partial_y v \end{vmatrix} \neq 0 .$$

²⁴Martin Wilhelm Kutta (1867 - 1927); Stuttgart Nikolai Jegorowitsch Zhukovsky (1847 - 1921); Moscow

In other words, it will be assumed that the vectors $\partial_x \mathbf{u}$ and $\partial_y \mathbf{u}$ are linearly independent.

This leads immediately to

$$\partial_x u = D \partial_v y, \quad \partial_y u = -D \partial_v x, \quad \partial_x v = -D \partial_u y, \quad \partial_y v = D \partial_u x .$$

The nonlinear equation (1.55) together with the equation $\partial_y u = \partial_x v$ for the functions $u(x, y)$ and $v(x, y)$ of the irrotational flow can so –after division by D – be transformed into the linear equations

$$\begin{aligned} \left(1 - \frac{v^2}{\hat{c}^2}\right) \partial_u x + \left(1 - \frac{u^2}{\hat{c}^2}\right) \partial_v y + \frac{u v}{\hat{c}^2} (\partial_v x + \partial_u y) &= 0 \\ \partial_v x - \partial_u y &= 0 \end{aligned} \quad (1.56)$$

for the functions $x(u, v)$ and $y(u, v)$.

The transition from the original equations to the linear equations (1.56) is called **method of hodographs** and corresponds to the **Legendre transformation** in the theory of partial differential equations.

In simply connected domains, $\mathbf{x}(\mathbf{u})$ can be derived from a potential Θ , i.e. $x = \Theta_u$, $y = \Theta_v$, because of

$$\text{curl } \mathbf{x}(\mathbf{u}) = (0, 0, \partial_u y - \partial_v x)^T = \mathbf{0} .$$

(1.56) therefore yields

$$\left(1 - \frac{v^2}{\hat{c}^2}\right) \partial_{uu} \Theta + \left(1 - \frac{u^2}{\hat{c}^2}\right) \partial_{vv} \Theta + 2 \frac{u v}{\hat{c}^2} \partial_{uv} \Theta = 0 . \quad (1.57)$$

If polar coordinates ($w = \|\mathbf{u}\|$, α) are used in the hodograph plane, i.e.

$$u = w \cos \alpha, \quad v = w \sin \alpha,$$

$$\partial_u = \cos \alpha \partial_w - \frac{1}{w} \sin \alpha \partial_\alpha,$$

$$\partial_v = \sin \alpha \partial_w + \frac{1}{w} \cos \alpha \partial_\alpha,$$

(1.57) turns to the so-called **hodograph equation**

$$\partial_{ww} \Theta + \frac{1}{w^2} \left(1 - \frac{w^2}{\hat{c}^2}\right) \partial_{\alpha\alpha} \Theta + \frac{1}{w} \left(1 - \frac{w^2}{\hat{c}^2}\right) \partial_w \Theta = 0 \quad (1.58)$$

which does not explicitly contain α .

If we try to solve (1.58) by the ansatz

$$\Theta(w, \alpha) = g(w) \sin(m\alpha) \quad \text{or} \quad \Theta(w, \alpha) = g(w) \cos(m\alpha) \quad (m \in \mathbb{R}), \quad (1.59)$$

we receive for the unknown function $g(w)$ the ordinary differential equation

$$g''(w) + \frac{1}{w} \left(1 - \frac{w^2}{\hat{c}^2}\right) g'(w) - \frac{m^2}{w^2} \left(1 - \frac{w^2}{\hat{c}^2}\right) g(w) = 0. \quad (1.60)$$

Solutions of the hodograph equation of type (1.59) are called **Chapligin solutions**²⁵.

Let $g_m(w)$ be a solution of (1.60) that belongs to a particular m and let $\Theta_m(w, \alpha)$ be the solution of (1.59) that corresponds to this solution.

Examples:

$m = 0$ leads to

$$\frac{g_0''}{g_0'} = -\frac{1}{w} + \frac{w}{\hat{c}^2},$$

hence

$$g_0' = \frac{c_1^{(0)}}{w} e^{\frac{1}{2} \frac{w^2}{\hat{c}^2}}.$$

The power series of $e^{\frac{1}{2} \frac{w^2}{\hat{c}^2}}$ converges uniformly for all values of w . Integration can therefore be done term by term and yields

$$g_0(w) = c_1^{(0)} \left\{ \ln w + \sum_{\nu=1}^{\infty} \frac{\left(\frac{w}{\hat{c}}\right)^{2\nu}}{2^{\nu+1} \nu \nu!} \right\} + c_2^{(0)}.$$

Here, $c_1^{(0)}$ and $c_2^{(0)}$ are arbitrary constants.

Analogously for $m = 1$:

$$g_1(w) = c_1^{(1)} w + c_2^{(1)} \left\{ \frac{1}{w^2} + \frac{w \ln w}{2\hat{c}^2} + \sum_{\nu=0}^{\infty} \frac{w^{2(\nu+1)}}{2^{\nu+2} \hat{c}^{2(\nu+2)}} \right\}.$$

Because (1.58) is linear and homogeneous, also all linear combinations of particular Chapligin solutions solve the equation (1.58). The coefficients of the expansion as well as the constants $c_1^{(m)}$ and $c_2^{(m)}$ have to be chosen in such a way that the expansion fits the given situation at least approximately.

²⁵C.A. Chapligin: Sci. Ann. Univ. Moscow. Math. Phys. **21** (1904), 1-121

1.4 Motionless Fluids and Sound Propagation

Obviously, because viscosity does only play a role in moving fluids, the results of this section are valid for real fluids, too.

Let us—in a first step— consider the case of constant density ρ , approximately realized in liquids.

In this situation, (1.26) leads for motionless fluids, i.e. for $\mathbf{u} = \mathbf{0}$, to the so-called **hydrostatic equation**

$$\hat{\mathbf{k}} = \nabla \left(\frac{p}{\rho} \right) ,$$

and (1.31) becomes

$$U + \frac{p}{\rho} = \text{const} .$$

Let us assume that we do not yet know how the free surface of a motionless liquid behaves if only the force of gravity and a constant exterior (e.g. atmospheric) pressure p_0 affects this liquid. Hence, the force per mass unit is given by

$$\hat{\mathbf{k}} = (0, 0, -g)^T \quad (g = \text{acceleration of gravity}) ,$$

such that $U = gz + \text{const}$, i.e. $\rho g z + p = \text{const}$. Particularly, if (x, y, z_0) is a point of the surface, $\rho g z + p = \rho g z_0 + p_0$ or

$$\rho g (z - z_0) = -(p - p_0) = -\hat{p} \tag{1.61}$$

holds. Here, \hat{p} is the overpressure inside the liquid compared with the exterior pressure.

Because of

$$z_0 = \frac{\text{const} - p_0}{\rho g} ,$$

z_0 is constant, thus independent of (x, y) . In other words, the surface of the liquid is plane or—more precisely— parallel to the earth surface. If $h = z_0 - z$ is the height of an arbitrarily shaped liquid column, (1.61) gives

$$\rho g h q = \hat{p} q ,$$

where q is the base area of the column.

The force affecting this base is given by the righthand side of the equation and does not depend on the form of the column whereas the lefthand side gives the weight of a *cylindrical* column of the fluid of the same base area and of the same height.

This phenomenon is called the **hydrostatic paradox**.

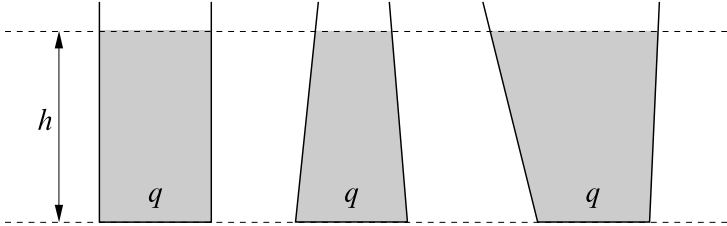


Figure 8: Hydrostatic paradox

We are now going to dip a solid body of volume V and of surface F into a non-moving liquid.

Obviously, an overall force \mathbf{K} with

$$\mathbf{K} = (K_1, K_2, K_3)^T = - \int_F p \mathbf{n} \, do + \mathbf{G}$$

affects the body where p is the interior pressure of the liquid. \mathbf{G} is the weight of the body, and \mathbf{n} denotes the outward directed normal unit vector at the points of the surface. Because of $\mathbf{G} = (0, 0, -G)^T$, (1.61) yields

$$K_1 = \rho g \int_F (z - z_0 - \frac{p_0}{\rho g}) n_1 \, do = \rho g \int_F \langle \mathbf{a}, \mathbf{n} \rangle \, do$$

with $\mathbf{a} := (z - z_0 - \frac{p_0}{\rho g}, 0, 0)^T$. The divergence theorem therefore leads to

$$K_1 = \rho g \int_V \operatorname{div} \mathbf{a} \, dV = \rho g \int_V \frac{\partial(z - z_0)}{\partial x} \, dV = 0.$$

Analogously: $K_2 = 0$. K_3 turns out to be given by

$$K_3 = \rho g \int_V \frac{\partial(z - z_0)}{\partial z} \, dV - G = \rho g V - G,$$

such that

$$\mathbf{K} = (0, 0, \rho g V - G)^T; \quad (1.62)$$

ρ is the density of the liquid (!) so that $\rho g V$ is the weight of the particular part of the liquid which is displaced by the solid body. Hence, the body is affected by a force against its weight direction, and this (static) buoyancy equals the weight of the displaced quantity of the liquid (**Archimedes' principle**)²⁶.

In order to study the sound propagation in a fluid, we assume that the fluid does not move from a macroscopic point of view. Moreover, we will only consider such sound effects which are due to very small disturbances of the density and of the pressure within the fluid. So, we do no longer assume the density to be constant but to be of only small variations.

Let $\hat{\rho}$ and \hat{p} be the averages of the density and of the pressure, respectively. Only these averages are expected to be constant quantities.

The disturbances of the density will be expressed by

$$\rho = \hat{\rho} (1 + \sigma(\mathbf{x}, t))$$

with $|\sigma| \ll 1$ and with only small spatial derivatives of σ ²⁷.

The continuity equation (1.4) then becomes

$$\hat{\rho} \sigma_t + \hat{\rho} \operatorname{div} ((1 + \sigma)\mathbf{u}) = 0 ,$$

thus, after division by $\hat{\rho}$ and taking the assumptions on σ into account,

$$\sigma_t + \operatorname{div} \mathbf{u} = 0 . \quad (1.64)$$

Experiments show that the propagation of sound waves occurs more or less in an adiabatic way, i.e. without gain or loss of heat. The equation of state to be taken into account therefore reads as

$$\rho^{-\gamma} p = \text{const}^{28} \quad (1.65)$$

such that

$$\frac{p}{\hat{p}} = \left(\frac{\rho}{\hat{\rho}} \right)^\gamma = (1 + \sigma)^\gamma .$$

²⁶Achimedes (about 220 BC); Syracuse

²⁷Our results do not necessarily hold if the propagation of *great* variations of the pressure or of the density are under consideration as they can occur in the case of detonations.

²⁸ γ from (1.9)

Because σ is small,

$$(1 + \sigma)^\gamma \approx 1 + \gamma \sigma .$$

Hence

$$p = \hat{p}(1 + \gamma \sigma) ,$$

so that

$$\frac{1}{\rho} \nabla p = \frac{\hat{p} \gamma}{(1 + \sigma) \hat{\rho}} \nabla \sigma . \quad (1.66)$$

In our model of small disturbances, the velocity of the fluid particles and its spatial derivatives are so small that higher order terms of these quantities can be neglected compared with first order terms. In other words, the convection term in (1.6) can be neglected. Also exterior forces do not play a role in our context.

If then (1.66) is put into (1.6), we find

$$\mathbf{u}_t = -\frac{\gamma \hat{p}}{\hat{\rho}} \nabla \sigma ,$$

where $1 + \sigma$ was approximated by 1.

Forming the divergence of this term followed by a change of the sequence of the time derivative and the spatial derivatives, we end up with

$$(\operatorname{div} \mathbf{u})_t = -\frac{\gamma \hat{p}}{\hat{\rho}} \Delta \sigma .$$

This result can be compared with (1.64) when differentiated with respect to t . This comparison yields

$$\sigma_{tt} = \frac{\gamma \hat{p}}{\hat{\rho}} \Delta \sigma ,$$

a wave equation for σ . It shows that the sound waves propagate within the fluid with the velocity

$$\hat{c} = \sqrt{\frac{\gamma \hat{p}}{\hat{\rho}}} . \quad (1.67)$$

Definition: \hat{c} is called the local sound speed.

Remark: Because of (1.65), \hat{c} can also be represented by

$$\hat{c} = \sqrt{\frac{dp}{d\rho}} . \quad (1.68)$$

2 Weak Solutions of Conservation Laws

2.1 Generalization of what will be called a Solution

As already announced, we are going to discuss in this chapter the introduction of a suitable globalization of the definition of a conservation law solution. And we do already know that such a solution can't be expected to be a smooth function because we had already realized that discontinuities can arise even from smooth initial states. Moreover, discontinuities which arose at an instant t_0 can move with respect to space and time as will be seen later. Of course, this also holds if the initial state already shows one or more discontinuities.

On the other hand, discontinuities which existed at an instant t_0 can disappear for $t > t_0$ as far as a suitable definition of global solutions will be used. Examples will be presented later. All these phenomena can occur in scalar problems as well as in case of systems of conservation laws.

We restrict ourselves within this chapter to only one space variable but allow the solution $\mathbf{V} = (v_1, \dots, v_m)^T$ sought for to consist of more than one component. Hence, we are going to treat the problem

$$\partial_t \mathbf{V} + \partial_x \mathbf{f}(\mathbf{V}) = \mathbf{0} \quad \text{on} \quad \Omega = \{(x, t) \mid x \in \mathbb{R}, t \geq 0\}, \quad (2.1)$$

$$\mathbf{V}(x, 0) = \mathbf{V}_0(x),$$

where we will assume for the time being that \mathbf{f} is at least one time differentiable.

Example:

*Another example of this kind – besides (1.14) – is the system of the so-called **shallow water equations**.*

These equations describe the shape and the velocity of surface waves as far as shallow liquids are concerned. Herewith, we consider an idealized situation, namely waves of an inviscid incompressible liquid travelling only along a canal of low depth whose

direction coincides with the x -axis whereas the flow of liquid particles into z -direction can be neglected. This idealization seems to be acceptable if the amplitudes of the surface waves are small compared with the wave length. The width b of the canal is assumed to be constant and we ask for the height $h(x, t)$ of the liquid surface over the bottom of the canal (cf. Figure 9).

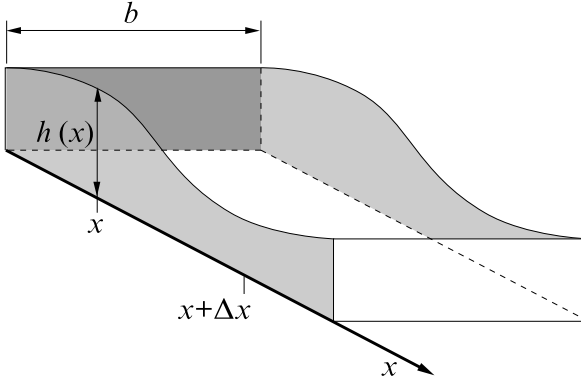


Figure 9: Shallow water flow along a canal

The mass of the fluid within a segment $x \leq \xi \leq x + \Delta x$ of the canal at the instant t is given by

$$m(t) = b \rho \int_x^{x+\Delta x} h(\xi, t) d\xi .$$

With (1.3), if φ is particularly chosen as $\varphi = b \rho h$, conservation of mass leads to

$$\begin{aligned} \frac{d}{dt} m(t) &= b \rho \frac{d}{dt} \int_x^{x+\Delta x} h(\xi, t) d\xi = b \rho \int_x^{x+\Delta x} \{ \partial_t h + \partial_x (uh) \} d\xi \\ &= 0 \quad \forall (x, \Delta x), \end{aligned}$$

i.e

$$\partial_t h + \partial_x (uh) = 0 .$$

This replaces the continuity equation in this particular situation.

The momentum of the mass $m(t)$ reads

$$b \rho \int_x^{x+\Delta x} u(\xi, t) h(\xi, t) d\xi .$$

The forces per length unit occurring from the hydrostatic pressure, namely

$$b p = \int_0^h (h - z) \rho g dz = \rho g \frac{h^2}{2} \quad (g : \text{acceleration of gravity}),$$

act on the liquid in the segment under consideration. The principle of conservation of momentum then leads by means of (1.3), now with $\varphi = b \rho u h$, to

$$\partial_t(\rho h u) + \partial_x(\rho h u^2) + \partial_x p = 0,$$

i.e. to

$$\partial_t(h u) + \partial_x(h u^2 + \frac{h^2}{2} g) = 0.$$

This equation turns out to be Euler's equation for our example.

If the continuity equation as well as the Euler equation is multiplied by g , if $\partial_t(h u)$ is replaced by $\partial_t h u + h \partial_t u$, if a variable $\psi(x, t) := g h(x, t)$ is introduced, if the continuity equation is again taken into account, and finally, if the second equation is divided by ψ , we find the **system of shallow water equations**

$$\partial_t \begin{pmatrix} \psi \\ u \end{pmatrix} + \partial_x \begin{pmatrix} u \psi \\ \frac{u^2}{2} + \psi \end{pmatrix} = \mathbf{0}.$$

Obviously, this is in fact a particular system of the type (1.2).

In order to establish now a suitable generalized definition of a *weak solution*, let us temporarily assume that there is a smooth solution of problem (2.1) on Ω . Put this solution into (2.1), then multiply (2.1) by an arbitrary **test function** $\Phi \in C_0^1(\Omega)$. Here, $C_0^1(\Omega)$ is defined to be the set of all functions continuously differentiable on Ω and with compact support.

Particularly, each of the functions Φ vanishes on the boundary of its support, possibly with the exception of such parts of the boundary which belong to the x axis.

Now integrate by parts over Ω . As a matter of fact, this is only an integration over the compact support of Φ , hence over a closed bounded region.

This integration leads to²⁹

$$\int_{\Omega} \{ \mathbf{V} \partial_t \Phi + \mathbf{f}(\mathbf{V}) \partial_x \Phi \} d(x, t) + \int_{-\infty}^{+\infty} \mathbf{V}_0(x) \Phi(x, 0) dx = \mathbf{0} \quad \forall \Phi \in C_0^1(\Omega). \quad (2.2)$$

²⁹It should be noted that this reformulation of the problem uses the conservation form of the differential equation and can't be generally applied to arbitrary quasilinear partial differential equations.

Obviously, the left hand side of (2.2) does no longer make sense only for sufficiently smooth functions \mathbf{V} but –for instance– also for all $\mathbf{V} \in L_1^{loc}(\Omega)$, i.e. for all vector valued functions \mathbf{V} , whose components can be integrated over compact subsets of Ω in the sense of Lebesgue.

Let us now forget the way we got (2.2), and let us –vice versa– ask for functions $\mathbf{V} \in L_1^{loc}(\Omega)$ which fulfill (2.2). These functions will be called **weak solutions** of the original problem (2.1), and these solutions are no longer necessarily smooth.

On the other hand, our motivation of (2.2) shows that every smooth solution is also a weak solution.

Vice versa, every weak solution \mathbf{V} which is a smooth function in a neighbourhood $D(P_0)$ of a point $P_0 = (x_0, t_0)$, does there not only fulfill (2.2) but also (2.1) in the classic understanding. Indeed, let $\Phi \in C_0^1(\Omega)$ be an arbitrary test function with support $\overline{D(P_0)}$. Then

$$\int_{D(P_0)} \{\mathbf{V} \partial_t \Phi + \mathbf{f}(\mathbf{V}) \partial_x \Phi\} d(x, t) + \int_{-\infty}^{\infty} \mathbf{V}_0(x) \Phi(x, 0) dx = \mathbf{0}$$

holds and leads to

$$\int_{D(P_0)} \{\partial_t \mathbf{V} \cdot \Phi - \partial_t (\mathbf{V} \Phi) + \partial_x \mathbf{f}(\mathbf{V}) \cdot \Phi - \partial_x (\mathbf{f}(\mathbf{V}) \cdot \Phi)\} d(x, t) - \int_{-\infty}^{\infty} \mathbf{V}_0(x) \Phi(x, 0) dx = \mathbf{0}.$$

By the divergence theorem, the last equation can be written as

$$\int_{D(P_0)} \{\partial_t v_i + \partial_x f_i(\mathbf{V})\} \Phi d(x, t) - \int_{\partial D(P_0)} \left(\left\langle \begin{pmatrix} f_i(\mathbf{V}) \\ v_i \end{pmatrix}, \mathbf{n} \right\rangle \right) \Phi d(\partial D(P_0)) - \int_{-\infty}^{\infty} v_{0,i}(x) \Phi(x, 0) dx = 0$$

$$(i = 1, \dots, m),$$

where \mathbf{n} denotes the outward normal unit vector on the boundary $\partial D(P_0)$ of $D(P_0)$ and where $\langle \cdot, \cdot \rangle$ means the standard \mathbb{R}^2 -scalar product.

If $\partial D(P_0)$ does not contain points of the x -axis, both boundary integrals vanish because we have $\Phi \equiv 0$ along $\partial D(P_0)$. Otherwise, the first boundary integral turns into an integral over the particular part of the boundary that belongs to the x -axis because Φ vanishes along the rest of this boundary. In each of these cases we find

$$\int_{D(P_0)} \{\partial_t v_i + \partial_x f_i(\mathbf{V})\} \Phi d(x, t) - \int_{-\infty}^{\infty} \left\{ \left\langle \begin{pmatrix} f_i(\mathbf{V}) \\ v_i \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right\rangle + v_{0,i} \right\} \Phi(x, 0) dx = 0,$$

i.e., for $i = 1, \dots, m$,

$$\int_{D(P_0)} \{\partial_t \mathbf{V} + \partial_x \mathbf{f}(\mathbf{V})\} \Phi(x, t) d(x, t) + \int_{-\infty}^{\infty} \{\mathbf{V}(x, 0) - \mathbf{V}_0(x)\} \Phi(x, 0) dx = \mathbf{0}.$$

Since Φ was arbitrary and \mathbf{V} was assumed to be smooth, this can only hold if

$$\partial_t \mathbf{V} + \partial_x \mathbf{f}(\mathbf{V}) = \mathbf{0}$$

on $D(P_0)$, hence particularly at P_0 , and if $\mathbf{V}(x_0, 0) = \mathbf{V}_0(x_0)$. But because P_0 was chosen arbitrarily in its neighbourhood, our assertion follows.

This concept of a weak solution of (2.1) leads to the advantage that the set of admissible solutions can be extended considerably. In particular, discontinuous solutions can be admitted as far as the Lebesgue integrability will not be disturbed.

But there is also a remarkable disadvantage: The uniqueness of the solution as it was at least guaranteed locally –together with its existence– gets lost. Often there exists more than one weak solution so that a new question arises: By which procedure can the particular weak solution answering the originally formulated real world problem be picked out of the set of all the weak solutions of (2.1) ?

As a matter of fact, the solution or weak solution of a differential equation problem can often only be described approximately by means of a numerical procedure. Numerical procedures normally consist in discrete models of the original problem, i.e. in finite dimensional problems whose solutions are expected to lie in a certain neighbourhood of the unknown solution of the original problem. Particularly, we expect the approximate solution to converge to this unknown solution if the dimension of the discrete model problem tends to infinity. If convergence of the numerical solution

to a weak solution of the original problem can be shown, one has to ensure that this weak solution coincides with the relevant solution instead with one of the other weak solutions.³⁰

For the time being, let us in the next section study an example where the transition to a weak formulation does really lead to a loss of uniqueness of the solution.

2.2 Traffic Flow Example with Loss of Uniqueness

As far as a sufficiently dense car traffic is concerned, the density ρ [cars/ml] as well as the flux f [cars/h] of macroscopic traffic flow models are approximately regarded as functions defined on a continuum, e.g. for $a \leq x \leq b$. The spatial variable x [mile] denotes the positions along a oneway lane which is interpreted as a onedimensional set, and $0 \leq t < \infty$.

If there are no approaches or exits and no crossings along the part of the road under consideration, the conservation of mass, i.e. the constant number of cars, obviously leads within our continuous model to the demand for the validity of the continuity equation

$$\partial_t \rho + \partial_x f = 0 \quad (2.3)$$

with

$$f = v\rho, \quad (2.4)$$

where $v = v(x, t)$ [ml/h] denotes the speed of the cars at position x at instant t .

(2.3) has to be completed by an initial condition

$$\rho(x, 0) = \rho_0(x) \quad (2.5)$$

with a given function ρ_0 as well as by a spatial boundary condition as far as the part of the road under consideration can't be regarded as a lane of infinite length. Moreover, it should be assumed that –hopefully– friction between cars does not occur.

As by experience, ρ depends above all explicitly on the speed of the cars:

$$\rho = \rho(v). \quad (2.6)$$

³⁰Of course, convergence is a basic requirement and must be accompanied by other important properties of the numerical procedure in order to make it an applicable method.

It is an empirically given decreasing and strictly monotone function. Hence, we can also consider v to be a function of ρ .

Let v_f be the mean value of the individual maximal car velocities if the road in front of a driver is more or less empty. This value is assumed to be a constant, e.g. $v_f = 80$ ml/h.

Obviously, the initial density

$$\rho_0(x) = \rho(v(x, 0))$$

is known as soon as the initial speed distribution

$$v(x, 0) = v_0(x) \tag{2.7}$$

is given.

Instead of $\rho(v)$ or of $v(\rho)$, often the graph of

$$f = v(\rho) \cdot \rho =: f(\rho) \tag{2.8}$$

is given empirically and turns out to be a strictly concave function called the **fundamental diagram**.

Thus, this traffic flow model leads to an initial value problem for a scalar conservation law of type (2.1), where \mathbf{V} consists of the only one component ρ such that $\mathbf{f}(\mathbf{V}) = f(\rho)$.

Let ρ^* denote the *jam concentration*, i.e. the density of cars within a standing congestion. Also this value is assumed to be a constant similar to v_f , e.g. $\rho^* = 400$ cars/ml.

A most simple $f(\rho)$ model, later used in order to explain exemplarily certain facts and connections, was given by Greenshields:

$$f(\rho) = v_f \cdot \rho \cdot \left(1 - \frac{\rho}{\rho^*}\right). \tag{2.9}$$

Thus, the flux vanishes or –in other words– there is (almost) no traffic, if the traffic density vanishes or if the road users stand within a traffic jam.

(2.9) is a concave parabola with its maximum at $\left(\frac{\rho^*}{2}, \frac{v_f \rho^*}{4}\right)$ and leads together with (2.4) to the relation

$$\rho = \rho^* \left(1 - \frac{v}{v_f}\right).$$

Hence, in case of Greenshields' model, the graph of $\rho(v)$ turns out to be a straight line.

The more general model for arbitrary fluxes is due to Lighthill & Whitham³¹.

By means of Greenshields' model we are going to start from the top of a traffic jam, i.e.

$$\rho_0(x) = \begin{cases} \rho^* & \text{for } x < 0 \\ 0 & \text{for } x \geq 0 \end{cases} \quad (2.10)$$

for $t_0 = 0$. The top of the traffic jam (red traffic light) is located at $x = 0$, and we identify the forward direction of the traffic with the positive direction of the x -axis.

Remark:

Conservation law problems whose initial values are constant on the left hand side of a discontinuity and constant also on its right hand side are called **Riemann-problems**³².

With (2.9) and (1.17), the characteristics are given by

$$\frac{1}{v_f}(x - x_0) = \begin{cases} -t & \text{for } x_0 < 0 \\ t & \text{for } x_0 \geq 0 \end{cases} .$$

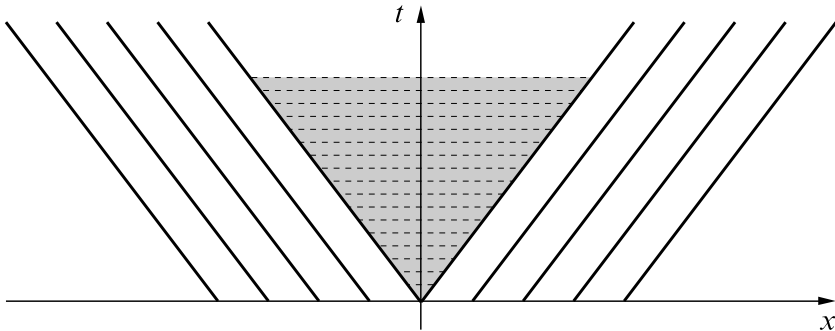


Figure 10: Characteristics in case of the dissolution of a traffic jam

³¹M.J. Lighthill and C.B. Whitham: Proc. Roy. Soc. A **229** (1955), 317-345

³²Bernhard Riemann (1826 - 1866); Göttingen

Obviously, the characteristics do not intersect for $t > 0$ so that we find immediately

$$\rho(x, t) = \begin{cases} \rho^* & \text{for } x < -v_f t \\ 0 & \text{for } x \geq v_f t \end{cases} .$$

Thus, the question is open how the solution will look like within the hatched region of Figure 10, i.e. for

$$-v_f t < x < v_f t, \quad t > 0 .$$

It can easily be verified that

$$\rho_1(x, t) := \begin{cases} \rho^* & \text{for } x < 0 \\ 0 & \text{for } x \geq 0 \end{cases} \quad \forall t \geq 0 \quad (2.11)$$

is a weak solution of

$$\partial_t \rho + f_v \cdot \partial_x \left(\rho \left(1 - \frac{\rho}{\rho^*} \right) \right) = 0 ,$$

i.e. of the Greenshields type of (2.3) together with the initial condition (2.10), because the left hand side of (2.2) if realized for our particular flux f leads to

$$\begin{aligned} & \int_0^\infty \int_{-\infty}^\infty \left[\partial_t \Phi \rho_1 + \partial_x \Phi v_f \rho_1 \left(1 - \frac{\rho_1}{\rho^*} \right) \right] d(x, t) + \int_{-\infty}^\infty \Phi(x, 0) \rho_0(x) dx \\ &= \int_0^\infty \int_{-\infty}^0 \rho^* \partial_t \Phi d(x, t) + \int_{-\infty}^0 \Phi(x, 0) \rho^* dx \\ &= \rho^* \int_{-\infty}^0 \left\{ \int_0^\infty \partial_t \Phi dt + \Phi(x, 0) \right\} dx = \rho^* \int_{-\infty}^0 \{ \Phi(x, 0) - \Phi(x, 0) \} dx = 0 . \end{aligned}$$

In other words, (2.1) is a solution which keeps its initial state: The drivers do not start though the traffic light switches from red to green. This solution describes a special situation where the state along the left hand boundary of the hatched region of Figure 10, i.e. along the characteristic $t = -\frac{x}{v_f}$, is connected with the right hand boundary

$t = \frac{x}{v_f}$ across a line of discontinuity.

Of course, drivers do not behave this way, so that (2.11) is certainly an unreasonable, hence irrelevant solution. As a matter of fact, the drivers **ride impulse** was not yet introduced into the model.

But there is also another weak solution of our problem, namely

$$\rho_2(x, t) := \begin{cases} \rho^* & \text{for } x < -v_f t, \quad t \geq 0 \\ \rho^* \frac{v_f t - x}{2v_f t} & \text{for } -v_f t \leq x \leq v_f t, \quad t > 0 \\ 0 & \text{for } x > v_f t, \quad t \geq 0 \end{cases}, \quad (2.12)$$

as can also easily be verified by elementary integration procedures. Within this solution the states along the left- and the righthand sides of the hatched region of Figure 10 are continuously connected. Herewith, the density decreases so that this type of a solution is called a **rarefaction wave**.

The head of the traffic jam, i.e. the the front end of the still standing chain of cars, withdraws against the direction of the movement of the traffic with speed $-v_f$ whereas the front cars move with speed $+v_f$. Obviously, (2.12) coincides with our experiences within the bounds of Greenshields' idealizations. It therefore has to be regarded as the particular weak solution that fits the circumstances of the given real situation.³³

Of course, the loss of uniqueness can also occur if systems of conservation laws instead of scalar situations are concerned.

As a further application of Greenshields' traffic flow model, let us consider the propagation of the end of a traffic jam opposite to the direction of the movement.

We start from an initial situation at the instant $t = 0$ where the traffic shows maximal flux on the left of $x = 0$, i.e. $\rho = \frac{\rho^*}{2}$, whereas the cars on the right of $x = 1$ did already stop, i.e. $\rho = \rho^*$, and where we have a continuous linear transition in between:

$$\rho_0(x) = \begin{cases} \frac{\rho^*}{2} & \text{for } x < 0 \\ \frac{\rho^*}{2}(1+x) & \text{for } 0 \leq x \leq 1 \\ \rho^* & \text{for } x > 1 \end{cases}. \quad (2.13)$$

³³This is also an example of a situation mentioned earlier, namely that a solution being continuous for $t > 0$ can develop from a discontinuous initial function.

The characteristics then turn out to be

$$\left. \begin{aligned} x &= x_0 && \text{for } x_0 \leq 0 \\ t &= \frac{1}{v_f} \left(1 - \frac{x}{x_0} \right) && \text{for } 0 < x_0 \leq 1 \\ t &= \frac{x - x_0}{v_f} && \text{for } x_0 > 1 \end{aligned} \right\} . \quad (2.14)$$

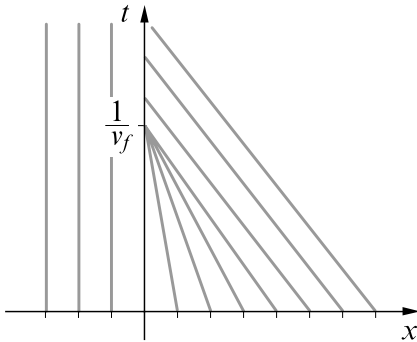


Figure 11: Characteristics at the creation of a traffic jam

This solution $\rho(x, t)$ keeps its continuity for the time $0 \leq t < \frac{1}{v_f}$. Then suddenly a discontinuity occurs at $t = \frac{1}{v_f}$, and this discontinuity moves for $t > \frac{1}{v_f}$ along a not yet known curve Γ . But one knows that this curve takes a course in the left upper half plane because the characteristics arising from the x -axis on the right of $x_0 = 1$ certainly intersect with the charactics arising from $x_0 < 0$. As soon as Γ will be known, the values of ρ given along the characteristics make possible to compute explicitly a solution $\rho(x, t)$. It is discontinuous along Γ but still an element of L_1^{loc} , thus really a weak solution.

The drivers do not expect such a discontinuity to crop up, and this often is one of the reasons of a rear-end collision.

How Γ can be determined will be shown by one of the examples of the next section.

2.3 The Rankine-Hugoniot Condition

We are going to treat the particular case where a certain discontinuity of a weak solution V of (2.1) moves continuously along the x -axis. In other words, the set of points of Ω where these discontinuities occur is a continuous curve Γ in Ω . Here, it can happen that all the components of V along this way are discontinuous or that this only holds for some of them. In gas dynamics, often the density is particularly concerned and forms a shock.

Let $P_0 = (x_0, t_0) \in \Gamma$ with $t_0 > 0$ and assume a bounded neighbourhood $\overline{D(P_0)} \subset \Omega$ of P to exist which does not contain points of the x -axis and where V is smooth outside Γ ³⁴.

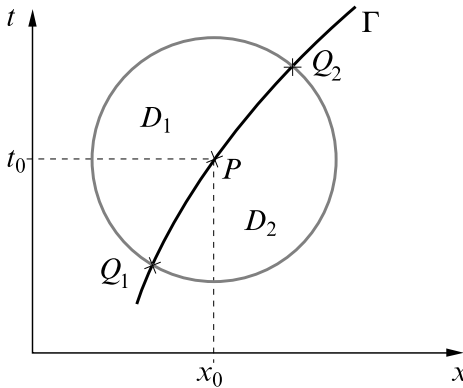


Figure 12: Sketch to explain the proof of the Rankine-Hugoniot condition

We decompose $\overline{D(P_0)}$ into the two parts $\overline{D_1}$ and $\overline{D_2}$ (cf. Figure 12) each of which we are going to treat as closed domains.

Taking the smoothness of V in the open domains D_1 and D_2 into account, we find on $\overline{D_1}$ and also on $\overline{D_2}$ for a particular $\Phi \in C_0^1(\Omega)$ whose compact support coincides

³⁴ The considerations to be formulated here can easily be generalized in order to fit also more complicated situations.

with $\overline{D(P_0)} = \overline{D_1 \cup D_2}$ the relation

$$\begin{aligned}
& - \int_{\Omega} [\mathbf{V} \partial_t \Phi + \mathbf{f}(\mathbf{V}) \partial_x \Phi] d(x, t) - \int_{-\infty}^{\infty} \mathbf{V}_0(x) \Phi(x, 0) dx \\
= & - \int_{D_1} [\partial_t(\mathbf{V} \Phi) + \partial_x(\mathbf{f}(\mathbf{V}) \Phi)] d(x, t) + \int_{D_1} [\partial_t \mathbf{V} \Phi + \partial_x \mathbf{f}(\mathbf{V}) \Phi] d(x, t) \\
& - \int_{D_2} [\partial_t(\mathbf{V} \Phi) + \partial_x(\mathbf{f}(\mathbf{V}) \Phi)] d(x, t) + \int_{D_2} [\partial_t \mathbf{V}_t \Phi + \partial_x \mathbf{f}(\mathbf{V}) \Phi] d(x, t) \\
= & 0 .
\end{aligned}$$

Together with the validity of (2.1) in D_1 and in D_2 ,

$$\int_{D_1} [\partial_t(\mathbf{V} \Phi) + \partial_x(\mathbf{f}(\mathbf{V}) \Phi)] d(x, t) + \int_{D_2} [\partial_t(\mathbf{V} \Phi) + \partial_x(\mathbf{f}(\mathbf{V}) \Phi)] d(x, t) = 0 .$$

The divergence theorem applied to each of the components of each of both integrals then leads to

$$\int_{\partial D_1} \{(\mathbf{V} \Phi) dx - (\mathbf{f}(\mathbf{V}) \Phi) dt\} + \int_{\partial D_2} \{(\mathbf{V} \Phi) dx - (\mathbf{f}(\mathbf{V}) \Phi) dt\} = 0 .$$

Because Φ vanishes on the boundary of its support, i.e. on the boundary of D , we find

$$\int_{Q_1}^{Q_2} \{(\mathbf{V}_\ell \Phi) dx - (\mathbf{f}(\mathbf{V}_\ell) \Phi) dt\} - \int_{Q_1}^{Q_2} \{(\mathbf{V}_r \Phi) dx - (\mathbf{f}(\mathbf{V}_r) \Phi) dt\} = 0 ,$$

where \mathbf{V}_ℓ denotes the lefthand limits of \mathbf{V} along Γ if for every particular fixed value of t the values of x approach Γ from the left, i.e. $\mathbf{V}_\ell(x, t) = \mathbf{V}(x-0, t)$ for $(x, t) \in \Gamma$. Analogously, \mathbf{V}_r means the values of \mathbf{V} if we approach Γ from the right. The integrals have to be understood as line integrals along Γ .

Let

$$[\mathbf{V}] := \mathbf{V}_\ell - \mathbf{V}_r \quad , \quad [\mathbf{f}] := \mathbf{f}_\ell - \mathbf{f}_r := \mathbf{f}(\mathbf{V}_\ell) - \mathbf{f}(\mathbf{V}_r)$$

be the **jumps**.

Then, obviously,

$$\int_{Q_1}^{Q_2} \left\{ [\mathbf{V}] \frac{dx}{dt} - [\mathbf{f}] \right\} \Phi dt = 0,$$

where $\frac{dx}{dt} =: \hat{v}$ denotes the velocity by which the discontinuity moves along the x -axis. Since the test function Φ with support $\overline{D(P_0)}$ was arbitrary and because P_0 was arbitrarily chosen point on Γ , the so-called **Rankine-Hugoniot condition**

$$[\mathbf{V}] \hat{v} = [\mathbf{f}] \tag{2.15}$$

turns out to hold along Γ componentwise under the conditions we had assumed to be fulfilled.

(2.15) is also called **jump condition**, and for every weak solution \mathbf{V} there is a particular velocity: $\hat{v} = \hat{v}(\mathbf{V})$.

Example:

If (2.15) is applied to our one dimensional gas flow where

$$\mathbf{V} = \begin{pmatrix} \rho \\ q \\ E \end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix} q \\ \frac{1}{\rho} q^2 + p \\ \frac{E+p}{\rho} q \end{pmatrix}, \tag{2.16}$$

and if $q = \rho u$ – with the flow velocity u – is taken into account, we find from the first component of the jump condition

$$(\rho_\ell - \rho_r) \hat{v} = q_\ell - q_r = \rho_\ell u_\ell - \rho_r u_r$$

i.e.

$$\rho_\ell (\hat{v} - u_\ell) = \rho_r (\hat{v} - u_r). \tag{2.17}$$

The second component leads to

$$(q_\ell - q_r) \hat{v} = \frac{q_\ell^2}{\rho_\ell} - \frac{q_r^2}{\rho_r} + p_\ell - p_r,$$

i.e.

$$\rho_\ell u_\ell (\hat{v} - u_\ell) = \rho_r u_r (\hat{v} - u_r) + p_\ell - p_r, \tag{2.18}$$

and the third component gives

$$(E_\ell - E_r) \hat{v} = E_\ell \frac{q_\ell}{\rho_\ell} - E_r \frac{q_r}{\rho_r} + p_\ell \frac{q_\ell}{\rho_\ell} - p_r \frac{q_r}{\rho_r},$$

i.e.

$$E_\ell (\hat{v} - u_\ell) = E_r (\hat{v} - u_r) + p_\ell u_\ell - p_r u_r . \quad (2.19)$$

If

$$u_\ell = u_r = \hat{v} , \quad (2.20)$$

(2.17) is fulfilled, and (2.18) then gives $p_\ell = p_r$ such that also (2.19) is realized.

Hence, in this case, the discontinuity 'swims' with the flow; this means that there are no gas particles which cross the point of discontinuity though $\rho_\ell = \rho_r$ or $E_\ell = E_r$ do not necessarily hold.

The situation described by

$$q_\ell \rho_r = q_r \rho_\ell \quad , \quad \rho_\ell \neq \rho_r \quad (2.21)$$

is called **contact discontinuity**.

Otherwise, $u_\ell \neq \hat{v}$, $u_r \neq \hat{v}$ show that there are gas particles which cross the discontinuity because (2.17) causes in this case that $\hat{v} - u_\ell$ and $\hat{v} - u_r$ are both positive or both negative³⁵. This type of discontinuity is called a **shock** and also the curve Γ is called a shock or a shock curve. The particular side of the shock where the particles having already crossed the shock are situated is called the **back side of the shock**, the other part is called the **front side**. If, for instance, $\hat{v} > u_r > 0$, the righthand side is the front side, the lefthand side is the back side.

Assume a weak solution to be smooth outside the discontinuity curve Γ , hence to be outside Γ also a solution of (2.1). The weak formulation (2.2) of the problem under consideration is then equivalent to the demand that the Rankine-Hugoniot condition is fulfilled along Γ and that simultaneously the original version of the initial value problem (2.1) is fulfilled, too. This equivalence follows from the already proven fact that weak solutions are at the same time classic solutions on domains where these weak solutions are smooth.

In case of a linear system

$$\partial_t \mathbf{V} + \partial_x (\mathbf{A}\mathbf{V}) = 0 ,$$

³⁵We omit the unnatural case $\rho < 0$.

the Rankine-Hugoniot condition leads to

$$[\mathbf{V}] \cdot \hat{\mathbf{v}} = [\mathbf{A} \mathbf{V}_\ell - \mathbf{A} \mathbf{V}_r] = \mathbf{A} [\mathbf{V}] .$$

Thus, $\hat{\mathbf{v}}$ is an eigenvalue of \mathbf{A} as far as a discontinuity of the initial function occurs, i.e. as far as $[\mathbf{V}] \neq \mathbf{0}$ turns out to hold for $t = 0$ and, hence, also for increasing values of t . This leads to

$$\dot{x} = \lambda_k \quad \text{for a particular } k$$

since $\dot{x}(t) = \hat{\mathbf{v}}$ is the differential equation for Γ .

But because of (1.18), this is also the differential equation for the k -th set of characteristic curves so that discontinuities move along characteristics as far as linear problems are concerned.

Let us wind up this section by the determination of Γ occurring in Greenshields' traffic flow model in case of a traffic jam increasing opposite to the flow direction.

Because of the relation

$$f_\ell - f_r = v_f \left(\rho_\ell - \frac{\rho_\ell^2}{\rho^*} - \rho_r + \frac{\rho_r^2}{\rho^*} \right)$$

presented in the Greenshields model, and taking (2.15) into account, Γ has to be determined by the differential equation

$$\dot{x} = \frac{f_\ell - f_r}{\rho_\ell - \rho_r} = v_f \left(1 - \frac{\rho_\ell + \rho_r}{\rho^*} \right) .$$

From Figure 11, the initial condition reads as

$$x\left(\frac{1}{v_f}\right) = 0 . \tag{2.22}$$

Moreover, Figure 11 shows

$$\rho = \rho^* \quad \text{for} \quad t > \frac{1}{v_f}$$

along the characteristics coming from the righthand side, and this does not depend on the location of Γ . Along the vertical characteristics coming from the left side, we find analogously

$$\rho = \frac{\rho^*}{2} \quad \text{for} \quad t > \frac{1}{v_f} .$$

Hence $\rho_\ell = \frac{\rho^*}{2}$, and therefore

$$\dot{x} = -\frac{v_f}{2}.$$

Together with (2.22), the shock turns out to be the straight line

$$t = \frac{1}{v_f} (1 - 2x), \quad x \leq 0.$$

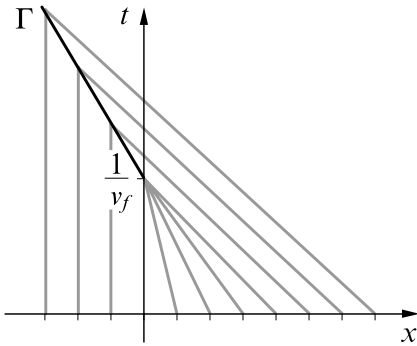


Figure 13: Shock when a traffic jam increases

Knowing Γ and the values of ρ along the characteristic curves, a weak solution of our traffic flow example can now immediately be constructed. Particularly, the propagation of the end of the congestion against the direction of the cars which have not yet come to a halt can so be evaluated:

$$\rho(x, t) = \begin{cases} \frac{\rho^*}{2} & \text{for } 0 \leq t < \frac{1}{v_f}(1 - 2x), \quad -\infty < x \leq 0 \\ \frac{\rho^*}{2} \left(1 + \frac{x}{1 - v_f t}\right) & \text{for } 0 \leq t < \frac{1-x}{v_f}, \quad 0 \leq x < 1 \\ \rho^* & \text{otherwise.} \end{cases} \quad (2.23)$$

This solution shows under the assumptions of the Greenshields idelizations and for the particular initial situation under consideration that the end of the chain of non-moving cars originally located at $x = 1$ will be located at $x = 0$ after $\frac{1}{v_f}$ time units. From this very moment on the end of the traffic jam moves with the **congestion velocity** $\frac{1}{2}v_f$

against the direction of the cars still moving, and the transition from the density $\frac{1}{2}v_f$ of the moving traffic to the density ρ^* of the set of cars which already came to a halt occurs very suddenly, at least from the driver's point of view.

With respect to the first part of our traffic flow model as well as generally, the question arises how in case of non-unique weak solutions of (2.1) the particular solution relevant for answering the real world problem that underlies the mathematical model can be characterized.

We are going to treat this question in the next chapter.

3 Entropy Conditions

3.1 Entropy in Case of Ideal Fluids

The loss of uniqueness caused by the installation of the concept of weak solutions leads to the necessity to formulate a criterion by which the physically relevant weak solution can be picked out of the set of all the weak solutions. We therefore ask for an additional constraint that characterizes this particular weak solution.

In order to get an idea, let us look again to the important application of conservation laws to the description of the dynamics of ideal gases. Besides the state equation, we took into account the conservation of mass, the conservation of momentum and the conservation of energy. The conservation of energy was formulated by the first law of thermodynamics, but what we did not yet respect is the second one and its assertion on the behaviour of the state variable

$$S = \text{entropy/mass} .$$

Without going into details, it should only be mentioned that the entropy describes a measure for the probability of the existence of a particular physical state (**Boltzmann statistics** in thermodynamics³⁶).

The **Second Law of Thermodynamics**³⁷ can be stated as follows:

If there is a closed physical system without supply of energy from outside, every physical process inside the system takes a course such that the entropy does not decrease³⁸, and it increases if possible³⁹.

Figure 14 shows an example for the validity of the second law, namely a shock tube divided by a membrane into two parts one of which is filled by a gas whereas the other

³⁶Ludwig Boltzmann (1844 - 1906); Graz, Vienna, Leipzig, Vienna

³⁷R.J.E. Clausius (1822 - 1888); Zurich, Würzburg, Bonn

³⁸e.g.: cf. R. Ansorge und Th. Sonar: ZAMM 77 (1997), 803-821

³⁹irreversible processes

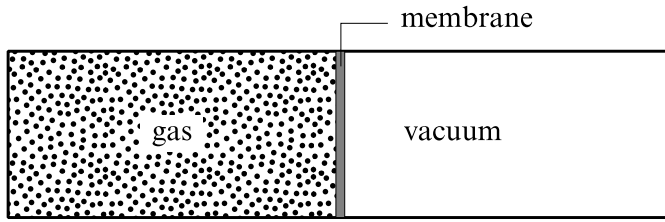


Figure 14: Example for increasing entropy

one is evacuated. When the membrane will be taken away, parts of the gas will begin to move into the evacuated part of the tube though all the conservation principles will also be respected if the gas particles would keep their positions. And as a matter of fact, the Euler equations will indeed also be fulfilled if nothing would happen. But this, obviously, does not describe the physically relevant solution as everybody knows from experience⁴⁰. The reason is that the equidistribution of the particles over the whole tube is a state more probable than the initial situation.

In other words, we expect the second law to be a model for the formulation of a constraint that allows to pick out of the set of weak solutions of a conservation law system the relevant solution, even if problems of other fields of applications such as problems of economy models are concerned. In any case, such a constraint will be called an **entropy condition** and the solution picked out will be called an **entropy solution**, hopefully the only one by the proof of a uniqueness theorem.

Generalizations of the second law in order to find a constraint that does not only work for physical tasks are due to several authors, particularly to Oleinik and Lax who started treating this question in a first step for scalar conservation laws.

To understand their ideas, let us look for the situation in gas dynamics, for convenience in case of a one dimensional flow.

It was already stated in section 1.1 that each of the state variables can uniquely be expressed in terms of two of the other state variables by means of a so-called equation of state; a particular example was demonstrated by (1.9).

One result of the theory of thermodynamics leads to the relation

$$dS = \frac{d\epsilon + p d\tilde{W}}{T} \quad (3.1)$$

⁴⁰By the way, this is an analogue to the situation where a traffic light switches from red to green but the drivers do not start.

where \tilde{W} denotes the specific volume $\tilde{W} := \frac{1}{\rho}$ (volume of mass 1).

If we ask particularly for the connection between the three state variables ϵ , S and $\rho = \frac{1}{\tilde{W}}$, the equation of state belonging to these variables leads to

$$d\epsilon = \partial_S \epsilon dS + \partial_{\tilde{W}} \epsilon d\tilde{W} ,$$

hence with (3.1) to

$$T dS = \partial_S \epsilon dS + \partial_{\tilde{W}} \epsilon d\tilde{W} + p d\tilde{W} = \partial_S \epsilon dS + (\partial_{\tilde{W}} \epsilon + p) d\tilde{W} ,$$

such that

$$\partial_S \epsilon = T, \quad \partial_{\tilde{W}} \epsilon = -p \tag{3.2}$$

follows from the linear independence of dS and $d\tilde{W}$.

Because of

$$\partial_t \epsilon = \partial_S \epsilon \partial_t S + \partial_{\tilde{W}} \epsilon \partial_t \tilde{W} = T \partial_t S - p \partial_{\tilde{W}} ,$$

the relation

$$\left. \begin{array}{l} T \partial_t S = \partial_t \epsilon + p \partial_{\tilde{W}} = \partial_t \epsilon - p \frac{\partial_t \rho}{\rho^2} \\ T \partial_x S = \partial_x \epsilon - p \frac{\partial_x \rho}{\rho^2} . \end{array} \right\} \tag{3.3}$$

follows. Analogously:

Denoting the entropy per volume by $s := \rho S$, the quantity $\partial_t s + \partial_x(us)$ can be expressed as

$$\begin{aligned} \partial_t s + \partial_x(us) &= \partial_t(\rho S) + \partial_x((\rho u) \cdot S) \\ &= \partial_t \rho S + \rho \partial_t S + \partial_x(\rho u) S + \rho u \partial_x S \\ &= [\partial_t \rho + \partial_x(\rho u)] \cdot S + \rho [\partial_t S + u \partial_x S] . \end{aligned}$$

The first term on the right side vanishes because of the continuity equation, and taking (3.3) into account the last equation leads to

$$\partial_t s + \partial_x(us) = \frac{\rho}{T} \left\{ \left(\partial_t \epsilon - p \frac{\partial_t \rho}{\rho^2} \right) + u \left(\partial_x \epsilon - p \frac{\partial_x \rho}{\rho^2} \right) \right\} . \tag{3.4}$$

The conservation law (1.7) concerning the specific total energy E and formulated in the particular situation of a one dimensional flow reads as

$$\partial_t E + \partial_x \left(\frac{E+p}{\rho} q \right) = 0$$

or – with $E = \rho \epsilon + \frac{\rho}{2} u^2$, reformulated by means of primitive variables –

$$\begin{aligned} & \partial_t \rho \epsilon + \rho \partial_t \epsilon + \frac{1}{2} \partial_t \rho u^2 + \rho u \partial_t u \\ & + u \left\{ \partial_x p + \partial_x \rho \epsilon + \rho \partial_x \epsilon + \frac{1}{2} \partial_x \rho u^2 + \rho u \partial_x u \right\} + \left(\rho \epsilon + \frac{\rho}{2} u^2 + p \right) \partial_x u = 0, \end{aligned}$$

i.e.

$$\begin{aligned} & (\partial_t \rho + \partial_x(\rho u)) \left(\epsilon + \frac{u^2}{2} \right) + \rho \left(\partial_t \epsilon + u \partial_x \epsilon + \frac{p}{\rho} \partial_x u \right) \\ & + \rho u \left(\partial_t u + \frac{\partial_x p}{\rho} + u \partial_x u \right) = 0. \quad (3.5) \end{aligned}$$

Also in this equation, the first term vanishes because of the continuity equation. The second conservation law, namely $\partial_t q + \partial_x \left(\frac{q^2}{\rho} + p \right) = 0$, also rewritten by means of primitive variables, leads to

$$\rho \left(\partial_t u + \frac{\partial_x p}{\rho} + u \partial_x u \right) + u (\partial_t \rho + \partial_x(\rho u)) = 0,$$

so that also the third term on the left side of (3.5) disappears, again by means of the continuity equation. Thus, after division by ρ ,

$$\partial_t \epsilon + u \partial_x \epsilon + \frac{p}{\rho} \partial_x u = 0, \quad (3.6)$$

remains, so that the term inside the outer braces in formula (3.4) turns to

$$-\frac{p}{\rho} \partial_x u - \frac{p}{\rho} \frac{\partial_t \rho}{\rho} - \frac{p}{\rho} \frac{u \partial_x \rho}{\rho} = -\frac{p}{\rho^2} (\partial_t \rho + \partial_x(\rho u)) = 0, \quad (3.7)$$

again because of the continuity equation.

Hence, also s fulfills a conservation law:

$$\partial_t s + \partial_x(u s) = 0, \quad (3.8)$$

and this automatically provided that the Euler equations are fulfilled.

But we have to take into account that this result required the differentiability properties to be ensured which we had used in order to derive (3.8). Thus, if the flow is smooth, section 1.1. shows that the total entropy

$$\int_{W(t)} s d(x, y, z)$$

of a volume $W(t)$ arbitrarily taken out of the fluid at an arbitrary instant t stays constant with respect to time. This is known as an **isentropic flow**.

Equation (3.8) can not be expected to hold if the smoothness assumptions are not fulfilled. But what can be expected, because of the second main theorem, is the validity of the inequality

$$\partial_t s + \partial_x (u s) \geq 0 \quad (3.9)$$

in its weak form, i.e.

$$\int_{\Omega} (s \partial_t \Phi + u s \partial_x \Phi) d(x, t) + \int_{\mathbb{R}} s(x, 0) \Phi(x, 0) dx \leq 0$$

$$\forall \Phi \in C_0^1(\Omega) \quad \text{with} \quad \Phi \geq 0. \quad (3.10)$$

Formally and because of its inequality character, (3.9) was multiplied by *nonnegative* test functions, then integrated by parts, and finally this ‘derivation’ of (3.10) should be forgotten.

Equation (3.10) is the entropy condition for the flow under consideration, and the ‘derivation’ shows that it is indeed satisfied automatically provided that the original conservation laws are fulfilled in the classical understanding. This follows from the fact that (3.8) was a direct consequence of the Euler equations.

If the weak solution \mathbf{V} of the Euler equations is discontinuous, hence not necessarily unique, (3.10) is really an *additional* condition.

Remark:

We mention without proof that the physical entropy per volume $-s = -s(\mathbf{V})$ of gas dynamics is strictly convex.

3.2 Generalization of the Entropy Condition

It was Lax⁴¹ whose definition of an entropy solution of a general system of conservation laws (2.1) can in the case of one spatial variable and by imitation of the gas dynamics situation be formulated as follows:

Definition:

A weak solution $\mathbf{V} = (v_1, \dots, v_m)^T$ of the system (2.1), i.e. a solution of (2.2), is called an **entropy solution** of this system if there is a scalar and strictly convex function $\tilde{S} = \tilde{S}(\mathbf{V})$ as well as a scalar function $\tilde{F} = \tilde{F}(\mathbf{V})$ belonging to \tilde{S} so that in every domain where \mathbf{V} is smooth, (2.1) leads automatically to the validity of

$$\partial_t \tilde{S}(\mathbf{V}) + \partial_x \tilde{F}(\mathbf{V}) = 0, \quad (3.11)$$

and so that in the non-smooth situation the **entropy condition**

$$\int_{\Omega} \left\{ \tilde{S}(\mathbf{V}(x, t)) \partial_t \Phi(x, t) + \tilde{F}(\mathbf{V}(x, t)) \partial_x \Phi(x, t) \right\} d(x, t) + \int_{\mathbb{R}} \tilde{S}(\mathbf{V}_0(x)) \Phi(x, 0) dx \geq 0 \quad \forall \Phi \in C_0^1(\Omega), \Phi \geq 0 \quad (3.12)$$

is respected.

Remarks:

- \tilde{S} is often also called the **entropy functional** and \tilde{F} is called the **entropy flux**.
- In gas dynamics we have $\tilde{S} = -s$.
- Sometimes we will more specifically talk of *Lax entropy solutions* because later also some other definitions will be given.
- (3.12) is often abbreviated by the formulation:

$$\partial_t \tilde{S}(\mathbf{V}) + \partial_x \tilde{F}(\mathbf{V}) \leq 0 \quad \text{holds weakly.}$$

After the choice of \tilde{S} , \tilde{F} is already more or less determined by (3.11). Especially in the scalar case $m = 1$ ($\mathbf{V} = v$, $\mathbf{f} = f$), (3.11) leads for smooth solutions v to

$$\tilde{S}'(v) \cdot \partial_t v + \tilde{F}'(v) \partial_x v = 0,$$

⁴¹P. Lax, in: Contributions to Nonlinear Functional Analysis (E. Zarantonello, ed.), Academic Press 1971

and therefore with (2.1) to

$$\tilde{S}'(v) \{-\partial_x f(v)\} + \tilde{F}'(v) \partial_x v = 0 ,$$

i.e. to

$$[-\tilde{S}'(v) f'(v) + \tilde{F}'(v)] \partial_x v = 0 .$$

Because this must hold automatically independent of the choice of $v_o(x)$ ⁴², the relation

$$\tilde{F}'(v) = \tilde{S}'(v) f'(v)$$

has to be fulfilled, i.e.

$$\tilde{F}(v) = \int_0^v \tilde{S}'(\alpha) f'(\alpha) d\alpha + \text{const} . \quad (3.13)$$

Just the constant is still arbitrary.

If $m > 1$ and if the function \mathbf{V} is smooth, (3.11) yields

$$\langle \nabla_{\mathbf{v}} \tilde{S}, \partial_t \mathbf{V} \rangle + \langle \nabla_{\mathbf{v}} \tilde{F}, \partial_x \mathbf{V} \rangle = 0 ,$$

where

$$\nabla_{\mathbf{v}} = (\partial_{v_1}, \partial_{v_2}, \dots, \partial_{v_m})^T .$$

By analogy with $m = 1$, putting (2.1) into this equation, we find

$$\begin{aligned} - \langle \nabla_{\mathbf{v}} \tilde{S}, \mathbf{J} \mathbf{f}(\mathbf{V}) \partial_x \mathbf{V} \rangle + \langle \nabla_{\mathbf{v}} \tilde{F}, \partial_x \mathbf{V} \rangle &= 0 , \text{ i.e.} \\ (\partial_x \mathbf{V})^T \left(\nabla_{\mathbf{v}} \tilde{F} - (\mathbf{J} \mathbf{f}(\mathbf{V}))^T \nabla_{\mathbf{v}} \tilde{S} \right) &= 0 . \end{aligned}$$

Again, because this relation must hold for any choice of \mathbf{V}_0 , \tilde{S} and \tilde{F} have to be connected by

$$\nabla_{\mathbf{v}} \tilde{F} - (\mathbf{J} \mathbf{f}(\mathbf{V}))^T \nabla_{\mathbf{v}} \tilde{S} = \mathbf{0} , \quad (3.14)$$

and this is a system of m differential equations for only two functions \tilde{S} , \tilde{F} to be determined that depends on the m independent variables v_1, \dots, v_m .

⁴²particularly also for initial functions generating smooth solutions

Hence,(3.14) does not necessarily have a solution in case of $m > 2^{43}$ so that the definition of an entropy condition used so far has to be modified suitably. We shall come back to this question in section 3.3.

But if $\mathbf{f}(\mathbf{V})$ is a gradient, too, i.e.

$$\mathbf{f}(\mathbf{V}) = \nabla_{\mathbf{V}} k(\mathbf{V})$$

with a scalar function $k(\mathbf{V})$, (3.14) turns by the particular choice of $\tilde{S}(\mathbf{V}) = \frac{1}{2} \langle \mathbf{V}, \mathbf{V} \rangle$ into the system of differential equations

$$\partial_{v_i} \tilde{F} = \sum_{\nu=1}^m \partial_{v_\nu} v_i k v_\nu = \partial_{v_i} \left(\sum_{\nu=1}^m \partial_{v_\nu} k v_\nu \right) - \partial_{v_i} k \quad (i = 1, \dots, m),$$

and this system can be satisfied by several solutions, e.g. by the particular entropy flux

$$\tilde{F}(\mathbf{V}) = \langle \mathbf{f}(\mathbf{V}), \mathbf{V} \rangle - k(\mathbf{V}) + \text{const} . \quad (3.15)$$

A straight forward formulation of these considerations in case of $m = 1$ leads directly to the particular couple

$$\tilde{S}(v) = \frac{1}{2} v^2, \quad \tilde{F}(v) = \int_0^v f'(\alpha) \alpha d\alpha + \text{const}, \quad (3.16)$$

but –obviously– also an infinite number of other suitable couples exists.

This fact generates a further question, namely for the entropy solution's independence of the choice of the entropy functional \tilde{S} .

In order to answer this question, we cite the following theorem whose proof is more or less just a word-to-word copy of the proof of the Rankine-Hugoniot condition taking (3.12) into account instead of (2.2):

Theorem 3.1:

Assume that there is a couple \tilde{S}, \tilde{F} of functions belonging to a weak solution \mathbf{V} and fulfilling condition (3.14); let the discontinuities of \mathbf{V} form a curve Γ along which the Rankine-Hugoniot condition holds. Then it can be stated that \mathbf{V} is a Lax entropy solution if and only if the inequality

$$[\tilde{S}] \hat{v} \leq [\tilde{F}] \quad (3.17)$$

⁴³As far as gas dynamics is concerned, a suitable entropy flux belonging to $\tilde{S} = -s$ exists though $3 \leq m \leq 5$.

holds where \hat{v} is the velocity occurring in the Rankine-Hugoniot condition (2.15), i.e. the velocity by which the discontinuities move along the x -axis.

As in (2.15), $[\tilde{S}] := \tilde{S}(\mathbf{V}_\ell) - \tilde{S}(\mathbf{V}_r)$ and $[\tilde{F}] := \tilde{F}(\mathbf{V}_\ell) - \tilde{F}(\mathbf{V}_r)$ denote the heights of the jumps.

In the scalar case and together with (2.15), (3.17) yields a particular form of the entropy condition equivalent to (3.12), namely

$$\frac{[\tilde{S}][f]}{[v]} \leq [\tilde{F}], \quad (3.18)$$

i.e –by means of (3.13)–

$$\frac{\tilde{S}(v_\ell) - \tilde{S}(v_r)}{v_\ell - v_r} (f(v_\ell) - f(v_r)) - \int_{v_r}^{v_\ell} \tilde{S}'(\alpha) f'(\alpha) d\alpha \leq 0. \quad (3.19)$$

This leads to the following theorem:

Theorem 3.2:

If the entropy functional $\tilde{S}(v)$ is strictly convex⁴⁴, i.e.

$$\tilde{S}''(v) > 0, \forall v \in \mathbb{R},$$

if also the flux f is strictly convex, and if discontinuities do only occur along smooth curves Γ which do not intersect in the area under consideration, condition (3.19) does only hold if and only if along each of these curves Γ the jump relation

$$v_\ell \geq v_r \quad (3.20)$$

is respected. In particular, this result shows that all entropy conditions formed by means of strictly convex functionals \tilde{S} and their entropy fluxes \tilde{F} are equivalent because these functionals do not play a role within formula (3.20).

PROOF:

The function $\tilde{f}(v)$ defined for every fixed value of v_r by

$$\tilde{f}(v) := \begin{cases} \frac{\tilde{S}(v) - \tilde{S}(v_r)}{v - v_r} (f(v) - f(v_r)) - \int_{v_r}^v \tilde{S}'(\alpha) f'(\alpha) d\alpha & \text{for } v \neq v_r \\ 0 & \text{for } v = v_r \end{cases}$$

⁴⁴as it was up to now assumed to be

is differentiable, and its derivative reads

$$\tilde{f}'(v) = \begin{cases} - \left\{ \tilde{S}'(v) - \frac{\tilde{S}(v) - \tilde{S}(v_r)}{v - v_r} \right\} \cdot \left\{ f'(v) - \frac{f(v) - f(v_r)}{v - v_r} \right\} & \text{for } v \neq v_r \\ 0 & \text{for } v = v_r. \end{cases}$$

The terms within the braces are positive for $v > v_r$ and negative for $v < v_r$ because \tilde{S} as well as f are strictly convex. This leads to

$$\tilde{f}'(v) < 0 \quad \text{for} \quad v \neq v_r$$

so that v_r is the only value for which \tilde{f} vanishes, and \tilde{f} is strictly monotonously decreasing for $v \neq v_r$. Thus,

$$\tilde{f}(v) > 0 \quad \text{for } v < v_r \quad , \quad \tilde{f}(v) = 0 \quad \text{for } v = v_r \quad , \quad \tilde{f}(v) < 0 \quad \text{for } v > v_r . \quad (3.21)$$

Equation (3.19) implies $\tilde{f}(v_\ell) \leq 0$, and because of (3.21) this is really equivalent to

$$v_\ell \geq v_r. \quad (3.22)$$

□

Remark:

The same arguments if applied to our traffic flow example with its strictly concave flux $f(\rho)$ would similarly lead to the equivalence of the entropy condition to

$$\rho_\ell \leq \rho_r ,^{45} \quad (3.23)$$

and the interpretation of this result shows that car drivers endeavour to smooth a discontinuous situation as soon as possible, i.e. to come to an equals sign⁴⁶, or that they intent to achieve a nondecreasing development of the traffic density close to a discontinuity in being. In particular, no car driver stops when he notices that there is a traffic jam in some distance in front of him; he would rather drive to the end of the jam such that the density there will be increased. The end of the jam, i.e. the discontinuity, passes over him against his ride direction when he stops.

The particular weak solution (2.11) of our traffic flow example does obviously not fulfill condition (3.23). In other words, (3.23) represents the driver's ride impulse and completes the rough Lighthill-Whitham traffic flow model⁴⁷.

⁴⁵The strictly concave case can easily be transformed into a problem with strictly convex flux.

⁴⁶The discontinuity at the top of a traffic jam is such a situation.

⁴⁷R. Ansorge: *Transpn. Res.* **24 B** (1990), 133-143

There is also a geometric characterization of the entropy solution provided that the flux f is strictly convex:

The Rankine-Hugoniot condition (2.15) together with the mean value theorem yields

$$\hat{v} = \frac{f(v_\ell) - f(v_r)}{v_\ell - v_r} = f'(v_r + \vartheta(v_\ell - v_r)), \quad (0 < \vartheta < 1),$$

and the strict convexity of f together with (3.20) then leads to

$$f'(v_r) < \hat{v} < f'(v_\ell). \quad (3.24)$$

Thus, if and only if the weak solution is the entropy solution, the ascents of the characteristics (1.17) belonging to this solution cause the characteristics to run into the shock curves Γ for increasing t instead of leaving them (cf. Figure 15).

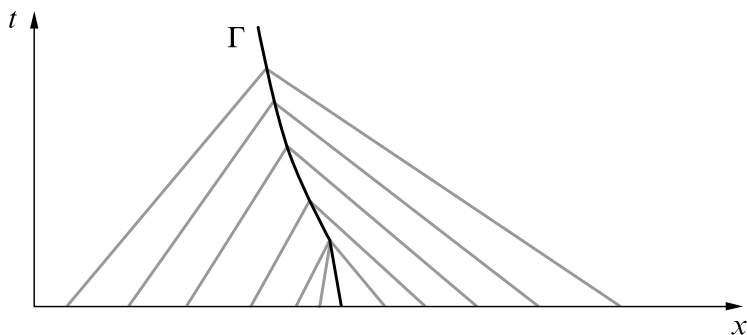


Figure 15: Course of the characteristics in case of entropy solutions of scalar problems as soon as shocks appear

Olga Oleinik did also publish an entropy condition for weak solutions of scalar conservation laws, and this already in 1957⁴⁸. It can be formulated as follows:

Definition:

A weak solution v of the scalar problem (2.1) is called an entropy solution in the sense of Oleinik – or an Oleinik solution – if there is a constant $\mathcal{E} > 0$ so that

$$\frac{v(x+a, t) - v(x, t)}{a} \leq \frac{\mathcal{E}}{t}, \quad \forall a > 0, \forall t > 0, \forall x \in \mathbb{R}. \quad (3.25)$$

Assume the discontinuities of a weak solution of a one-dimensional scalar conservation law to form a curve Γ such that the Rankine-Hugoniot condition (2.15) holds

⁴⁸O. Oleinik: Usp. Math. Nauk. (N.S.) **12** (1957), 3-73

along this curve; if then the point $(x, t) \in \Omega$ ($t > 0$, fixed) moves to Γ from the left, the point $(x + a, t)$, $a > 0$, finally becomes a point right of Γ such that (3.25) states

$$v(x + a, t) - v_\ell \leq \frac{\mathcal{E}}{t} a.$$

If a now tends to zero,

$$v_r - v_\ell \leq 0$$

follows.

Remark:

Thus, an Oleinik solution is a Lax solution, too. If –hopefully– a uniqueness theorem for Lax entropy solutions will be available, the Oleinik solution is also uniquely determined and equals the Lax solution.

This uniqueness problem for one-dimensional scalar conservation laws will be discussed in the next section, and also the question concerning a suitable generalization of the definition of a Lax entropy solution to systems of conservation laws will then be treated.

Remark:

Another access to the idea of an entropy solution for scalar problems ensues from the wellknown uniqueness of the solution $v_\epsilon(x, t)$ of the parabolic problem

$$\begin{aligned} \partial_t v + \partial_x f(v) &= \epsilon \partial_{xx} v, \quad \epsilon > 0 \\ v(x, 0) &= v_0(x), \end{aligned}$$

if the righthand side of (2.1) is enriched analogously to the linear problem (1.24) by a diffusion term where the diffusion parameter ϵ tends to 0⁴⁹.

3.3 Uniqueness of Entropy Solutions

In this section we are going to investigate the uniqueness of the Lax entropy solution. Lax himself gave the outline of the following theorem:

Theorem 3.3:

Let \tilde{v} and v be two Lax entropy solutions of the scalar form of the conservation law (2.1) which are piecewise continuous for every fixed $t \geq 0$, non-smooth only along

⁴⁹N.N. Kuznetsov: USSR Comput. Math. Math.Phys. **16** (1976), 105-119

certain curves Γ , and belonging to initial functions \tilde{v}_0 and v_0 , respectively. The flux f is assumed to be strictly convex. Then, the distance

$$\|\tilde{v}(\cdot, t) - v(\cdot, t)\|_{L_1(I)}$$

is monotonously decreasing with respect to t , where the L_1 -norm has to be understood as taken over an arbitrary but fixed interval I of the x -axis⁵⁰.

PROOF: Decompose for every fixed t the interval I into intervals $(x_\nu, x_{\nu+1})$ ($\nu = 0, \pm 1, \pm 2, \dots$) in such way that $\tilde{v}(x, t) - v(x, t)$ has the sign $(-1)^\nu$ uniformly on $(x_\nu, x_{\nu+1})$. A decomposition of this type exists because $\tilde{v} - v$ was assumed to be piecewise continuous with respect to x . Evidently, x_ν depends on t : $x_\nu = x_\nu(t)$ ($\nu = 0, \pm 1, \pm 2, \dots$). Thus,

$$\|\tilde{v}(\cdot, t) - v(\cdot, t)\|_{L_1(I)} = \sum_{\nu} (-1)^\nu \int_{x_\nu(t)}^{x_{\nu+1}(t)} [\tilde{v}(x, t) - v(x, t)] dx. \quad (3.26)$$

\tilde{v} and v are piecewise smooth on $(x_\nu(t), x_{\nu+1}(t))$, and differentiation with respect to t yields

$$\begin{aligned} \frac{d}{dt} \|\tilde{v}(\cdot, t) - v(\cdot, t)\|_{L_1(I)} &= \sum_{\nu} (-1)^\nu \left\{ [\tilde{v}(x_{\nu+1}(t), t) - v(x_{\nu+1}(t), t)] x'_{\nu+1}(t) \right. \\ &\quad \left. - [\tilde{v}(x_\nu(t), t) - v(x_\nu(t), t)] x'_\nu(t) \right. \\ &\quad \left. + \int_{x_\nu(t)}^{x_{\nu+1}(t)} [\partial_t \tilde{v} - \partial_t v(x, t)] dx \right\}. \end{aligned}$$

With (2.1),

$$\begin{aligned} \int_{x_\nu(t)}^{x_{\nu+1}(t)} \partial_t v(x, t) dx &= - \int_{x_\nu(t)}^{x_{\nu+1}(t)} \partial_x f(v(x, t)) dx \\ &= -f(v(x_{\nu+1}(t), t)) + f(v(x_\nu(t), t)) \end{aligned}$$

follows; analogously for \tilde{v} .

⁵⁰Particularly, if then $\tilde{v}_0 = v_0$ in L_1 , i.e. $\|\tilde{v}(\cdot, 0) - v(\cdot, 0)\|_{L_1(I)} = 0$ for every interval I , $\|\tilde{v}(\cdot, t) - v(\cdot, t)\|_{L_1(I)} = 0$ follows for all $t \geq 0$, i.e. $\tilde{v} = v$ in L_1 at every time level t , and this means uniqueness.

Hence,

$$\begin{aligned}
& \frac{d}{dt} \|\tilde{v}(\cdot, t) - v(\cdot, t)\|_{L_1(I)} \\
&= \sum_{\nu} (-1)^{\nu} \left\{ [\tilde{v}(x_{\nu+1}(t), t) - v(x_{\nu+1}(t), t)] x'_{\nu+1}(t) \right. \\
&\quad - [\tilde{v}(x_{\nu}(t), t) - v(x_{\nu}(t), t)] x'_{\nu}(t) \\
&\quad - [f(\tilde{v}(x_{\nu+1}(t), t)) - f(v(x_{\nu+1}(t), t))] \\
&\quad \left. + [f(\tilde{v}(x_{\nu}(t), t)) - f(v(x_{\nu}(t), t))] \right\}.
\end{aligned} \tag{3.27}$$

We have to distinguish between two different situations:

- i) The function $\tilde{v}(\cdot, t_0) - v(\cdot, t_0)$ is continuous at $x_{\nu}(t_0)$ ⁵¹. Then this function vanishes at $x_{\nu}(t_0)$, i.e.

$$\tilde{v}(x_{\nu}(t_0), t_0) = v(x_{\nu}(t_0), t_0). \tag{3.28}$$

- ii) One of the functions $\tilde{v}(x, t_0)$, $v(x, t_0)$ is discontinuous at $x_{\nu}(t_0)$.

We therefore need to sum up on the righthand side of (3.27) only for the particular numbers ν characterized by the fact that $(x_{\nu}(t), t)$ or $(x_{\nu+1}(t), t)$ lie on curves of discontinuities of \tilde{v} or of v .

Let us at first discuss the case that $(x_{\nu+1}(t), t)$ lies on a curve of discontinuities of \tilde{v} whereas v is smooth at this position. The end point $x_{\nu+1}$ of the interval $(x_{\nu}(t), x_{\nu+1}(t))$ has then to be looked upon as $x_{\nu+1} - 0$, i.e. $\tilde{v}(x_{\nu+1}(t), t) = \tilde{v}_{\ell}$. If, for example, $\tilde{v} - v > 0$ on $(x_{\nu}, x_{\nu+1})$, hence $\tilde{v} - v < 0$ on $(x_{\nu+1}, x_{\nu+2})$, then ν is even and

$$\tilde{v}_{\ell} > v_{\ell} = v_r > \tilde{v}_r \tag{3.29}$$

because $x_{\nu+1}$ has on $(x_{\nu+1}, x_{\nu+2})$ to be looked upon as $x_{\nu+1} + 0$.

This leads to

$$\begin{aligned}
& (-1)^{\nu} \left\{ [\tilde{v}(x_{\nu+1}(t), t) - v(x_{\nu+1}(t), t)] x'_{\nu+1}(t) \right. \\
&\quad \left. - [f(\tilde{v}(x_{\nu+1}(t), t)) - f(v(x_{\nu+1}(t), t))] \right\}
\end{aligned}$$

⁵¹for the arbitrary fixed value of t_0 under consideration

$$= (\tilde{v}_\ell - v_\ell) x'_{\nu+1}(t) - f(u_\ell) + f(v_\ell).$$

Because of the smoothness of v at $(x_{\nu+1}(t), t)$, $x = x_{\nu+1}(t)$ can only be a point of the shock curve of \tilde{v} . The Rankine-Hugoniot condition (2.15) therefore makes it possible to estimate the righthand side of the last equation by

$$\begin{aligned} (\tilde{v}_\ell - v_\ell) \frac{f(\tilde{v}_\ell) - f(\tilde{v}_r)}{\tilde{v}_\ell - \tilde{v}_r} - f(\tilde{v}_\ell) + f(v_\ell) \\ = (\tilde{v}_\ell - v_\ell) \left[\frac{f(\tilde{v}_\ell) - f(\tilde{v}_r)}{\tilde{v}_\ell - \tilde{v}_r} - \frac{f(\tilde{v}_\ell) - f(v_\ell)}{\tilde{v}_\ell - v_\ell} \right] \leq 0, \end{aligned}$$

where it was taken into account that (3.29) yields $\tilde{v}_\ell - v_\ell > 0$, but that

$$\frac{f(\tilde{v}_\ell) - f(\tilde{v}_r)}{\tilde{v}_\ell - \tilde{v}_r} - \frac{f(\tilde{v}_\ell) - f(v_\ell)}{\tilde{v}_\ell - v_\ell} \leq 0.$$

Here, the last inequality results from $\tilde{v}_r < v_\ell < \tilde{v}_\ell$, also following from (3.29), as well as from the strict convexity of f (cf. Fig. 16).

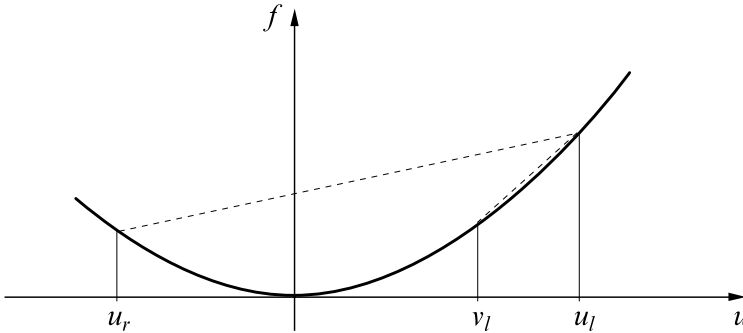


Figure 16: Sketch concerning Lax' uniqueness theorem

A similar result follows if $(x_{\nu+1}(t), t)$ is a point of a curve which is a set of discontinuities of \tilde{v} as well as of v ; in this situation, the inequality

$$\tilde{v}_\ell > v_\ell > v_r > \tilde{v}_r$$

has to be taken into account instead of (3.29), and it is this part of the proof where the jump condition $v_\ell > v_r$ applies to.

The same reflections work for the values of \tilde{v} and v at the position $(x_\nu(t), t)$, namely disappearance or non-positivity of the terms. Here, $x_\nu(t) = x_r$ has to be observed in case of a discontinuity.

The case $\tilde{v} < v$ on $(x_\nu, x_{\nu+1})$ can be treated in the same way. So, indeed,

$$\frac{d}{dt} \|\tilde{v}(\cdot, t) - v(\cdot, t)\|_{L_1(I)} \leq 0$$

follows. □

As already emphasized, in case of systems of conservation laws of more than two equations, pairs (\tilde{S}, \tilde{F}) of an entropy functional \tilde{S} and an entropy flux \tilde{F} connected with each other by (3.14) and appropriate to distinguish an entropy solution from other weak solutions do not necessarily exist.

We continue treating only one-dimensional problems but we try to formulate a suitable analog to (3.24) as far as problems are concerned for which the convexity property of the flux can also suitably be generalized.

Let $N \subset \mathbb{R}^m$ be a subset of m -dimensional vectors and assume the system (2.1) to be strictly hyperbolic for all $\mathbf{V} = (v_1, v_2, \dots, v_m)^T \in N$, i.e. the eigenvalues $\lambda_i(\mathbf{V})$ of the Jacobian $\mathbf{J} \mathbf{f}(\mathbf{V})$ are real and different from each other for all $\mathbf{V} \in N$. Moreover, let $\mathbf{f}(\mathbf{V}) \in C^2(N)$ and

$$\lambda_1(\mathbf{V}) < \lambda_2(\mathbf{V}) < \dots < \lambda_m(\mathbf{V}). \quad (3.30)$$

The eigenvalues keep this order for all $\mathbf{V} \in N$ since otherwise a vector $\mathbf{V} \in N$ with

$$\lambda_i(\mathbf{V}) = \lambda_k(\mathbf{V}) \quad \text{for a pair } (i, k) \quad \text{with } i \neq k$$

would have to exist because of continuity reasons. But this would contradict the strict hyperbolicity on N .

The eigenvectors $s_i(\mathbf{V})$ ($i = 1, 2, \dots, m$) belonging to the eigenvalues $\lambda_i(\mathbf{V})$ ($i = 1, 2, \dots, m$), respectively, are linearly independent.

Let $\mathbf{V}(x, t)$ be a solution of (2.1) which is smooth on a domain $G \subset \Omega$ and with

$$\mathbf{V}(x, t) \in N \quad , \quad \forall (x, t) \in G .$$

The i -characteristics $x = x_{(i)}(t)$ corresponding to this solution (cf. (1.18)) are uniquely determined due to the Picard-Lindelöf theorem if the functions $\lambda_i(\mathbf{V}(x, t))$ fulfill a Lipschitz condition with respect to x ($i = 1, \dots, m$) and as far as initial states $x_{(i)}(t_0)$ with $(x_{(i)}(t_0), t_0) \in G$ are prescribed. Of course, there are also existence and uniqueness theorems with weaker assumptions available, e.g. the existence theorem of Peano combined with the uniqueness condition of Nagumo.

Then, along an i -characteristic,

$$\begin{aligned} \frac{d}{dt} \mathbf{V}(x_{(i)}(t), t) \\ = - [\mathbf{J} \mathbf{f}(\mathbf{V}(x_{(i)}(t), t)) - \lambda_i(\mathbf{V}(x_{(i)}(t), t)) \cdot \mathbf{I}] \partial_x \mathbf{V}(x_{(i)}(t), t), \end{aligned} \quad (3.31)$$

where \mathbf{I} in this context means the $m \times m$ unit matrix.

In case of $m = 1$, the righthand side of (3.31) vanishes as already observed earlier, so that $\mathbf{V} = v$ is constant along a characteristic which therefore becomes a straight line.

Does an analogon exist for $m > 1$? In other words, does a weak solution \mathbf{V} and does a set of i -characteristics $x_{(i)}$ belonging to this solution exist so that the solution is constant along these characteristics which, thus, become straight lines?

A positive answer requires in case of $\partial_x \mathbf{V}(x_{(i)}(t), t) \neq \mathbf{0}$ the validity of the relation

$$\partial_x \mathbf{V}(x_{(i)}(t), t) = \mathbf{s}_i(\mathbf{V}(x_{(i)}(t), t)) \quad \forall t > 0^{52}.$$

Solutions of this type can really be specified: If we introduce so-called **centered waves** $\mathbf{V} = \mathbf{V}^*$, centered at a position $P_0 = (x_0, t_0) \in G$, i.e. waves of the form

$$\mathbf{V}(x, t) = \mathbf{V}^* \left(\frac{x - x_0}{t - t_0} \right) \quad (3.32)$$

with

$$\mathbf{V}^*(\xi) \in N \cap [C^1(\mathbb{R})]^m \quad \forall P = (x, t) \in G, \quad \xi := \frac{x - x_0}{t - t_0}.$$

These centered waves solve (2.1) if they fulfill

$$\partial_t \mathbf{V} + \partial_x \mathbf{f}(\mathbf{V}) = (\partial_x \xi \mathbf{J} \mathbf{f}(\mathbf{V}^*) + \partial_t \xi \mathbf{I}) \mathbf{V}^{*'}(\xi) = \mathbf{0}$$

which can also be described as

$$(\mathbf{J} \mathbf{f}(\mathbf{V}^*(\xi)) - \xi \mathbf{I}) \mathbf{V}^{*'}(\xi) = \mathbf{0}. \quad (3.33)$$

Hence, for one $i \in (1, 2, \dots, m)$, the pair of relations

$$\begin{aligned} \mathbf{V}^{*'}(\xi) &= \mathbf{s}_i(\mathbf{V}^*(\xi)), \\ \xi &= \lambda_i(\mathbf{V}^*(\xi)) \quad (\text{for all possible values of } \xi) \end{aligned} \quad (3.34)$$

must hold.

⁵²Here, the norm of the vectorfield \mathbf{V} is chosen suitably.

In this case, differentiation of the second equation of (3.34) with respect to ξ together with the first equation yields (cf. footnote 51)

$$1 = \langle \nabla_{\mathbf{V}} \lambda_i(\mathbf{V}), \mathbf{s}_i(\mathbf{V}) \rangle. \quad (3.35)$$

Definition:

If $\langle \nabla_{\mathbf{V}} \lambda_i(\mathbf{V}), \mathbf{s}_i(\mathbf{V}) \rangle \neq 0$ holds for all $\mathbf{V} \in N$ with a set $N \subset \mathbb{R}^m$, the i -th characteristic field $(\lambda_i(\mathbf{V}), \mathbf{s}_i(\mathbf{V}))$ is called **genuinely nonlinear** on N .

Remark:

Because the scalar case implies $\lambda(\mathbf{V}) = \lambda(v) = f'(v)$ which yields

$$\nabla_{\mathbf{V}} \lambda(\mathbf{V}) = f''(v),$$

(3.35) can only be fulfilled if $f''(v) \neq 0$, and $s(v)$ in this case is a scalar factor $s(v)$ different from zero. It can be normalized by $s(v) = \frac{1}{|f''(v)|}$ which makes (3.35) equivalent to $f''(v) > 0$, i.e. to strict convexity of the flux (or, if we normalize by $\frac{-1}{|f''(v)|}$, to strict concavity).

We are therefore going to look upon (3.35) as the transition announced earlier of the strict convexity of the flux as defined for scalar problems to the i -th characteristic field in case of $m > 1$.

Let us now assume N to be a simply connected domain in \mathbb{R}^m . Moreover, let $\mathbf{V}^0 \in N$ ($\mathbf{V}^0 \neq \mathbf{0}$) be a given **state** and $\xi_0 := \lambda_i(\mathbf{V}^0)$. If ξ_0 will be used as starting point with initial condition

$$\mathbf{V}^*(\xi_0) = \mathbf{V}^0, \quad (3.36)$$

the second equation in (3.34) will obviously be fulfilled at this point ξ_0 .

The first equation of (3.34) represents a system of ordinary differential equations where (3.36) is used as its initial condition. We know that there is a unique solution of this system in a certain neighbourhood of (ξ_0, \mathbf{V}^0) provided that the assumptions of a standard theorem like the Picard-Lindelöf theorem are fulfilled. This solution runs with respect to \mathbf{V} in N , let's say for $\xi_0 - a \leq \xi \leq \xi_0 + a$ with sufficiently small $a > 0$. Also (3.35) then holds for $\mathbf{V} = \mathbf{V}^*(\xi)$ because of the assumptions.

Since also the first equation of (3.34) was fulfilled, also (3.33) holds. Integration of (3.33) together with (3.36) then yields the fact that the second equation of (3.34) is fulfilled, too, at least for the values of ξ under consideration and if a will be again somewhat diminished if necessary.

We conclude that there is a nonvanishing smooth solution \mathbf{V} , namely a wave centered at P_0 , belonging to a given nonvanishing smooth initial state \mathbf{V}^0 and to a particular integer i , leading to an i -th characteristic field which is genuinely nonlinear on N :

$$\mathbf{V}(x, t) = \mathbf{V}^* \left(\frac{x - x_0}{t - t_0} \right) \quad , \quad \mathbf{V}^*(\lambda_i(\mathbf{V}^0)) = \mathbf{V}^0 \quad (3.37)$$

for all $(x, t) \in \Omega$ with

$$\lambda_i(\mathbf{V}^0) - a \leq \frac{x - x_0}{t - t_0} \leq \lambda_i(\mathbf{V}^0) + a \quad \text{with } a > 0 \text{ sufficiently small.} \quad (3.38)$$

Along each of the straight lines

$$\frac{x - x_0}{t - t_0} = c = \text{const}$$

with

$$\lambda_i(\mathbf{V}^0) - a \leq c \leq \lambda_i(\mathbf{V}^0) + a$$

forming a fan, the solution

$$\mathbf{V}^* \left(\frac{x - x_0}{t - t_0} \right) = \mathbf{V}^*(c)$$

is constant.

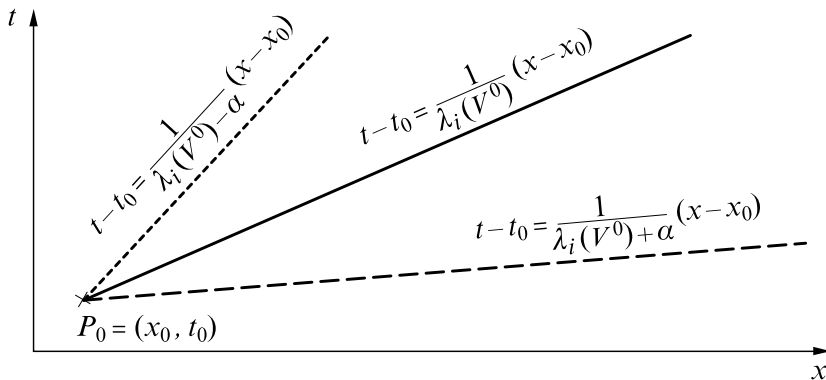


Figure 17: Fan of i -characteristics of centered waves

Thus, along these straight lines given by

$$x_{(i)}(t) = x_0 + c(t - t_0) ,$$

the relation

$$\frac{d}{dt} \mathbf{V}(x_{(i)}(t), t) = \frac{d}{dt} \mathbf{V}^*(c) = \mathbf{0} \quad (3.39)$$

holds. But because $\mathbf{V}^{*'}(c)$ is an eigenvector belonging to $\lambda_i(\mathbf{V}^*(c))$, this also applies to

$$\partial_x \mathbf{V}(x_{(i)}(t), t) = \frac{1}{t - t_0} \mathbf{V}^{*'}(c) .$$

It follows from (1.18) and (3.33) that the straight lines generating the fan and belonging to this particular integer i are the i -characteristics formed by means of the solution (3.32) through the point P_0 , and the centered waves constructed so far are constant along these lines.

Let us now assume that \mathbf{V}^1 is a state in the neighbourhood of \mathbf{V}^0 . Then, by continuity arguments, $\lambda_i(\mathbf{V}^1)$ lies in a neighbourhood of $\lambda_i(\mathbf{V}^0)$ such that

$$|\lambda_i(\mathbf{V}^1) - \lambda_i(\mathbf{V}^0)| \leq a$$

can be fulfilled for all states \mathbf{V}^1 sufficiently close to \mathbf{V}^0 .

If particularly

$$\lambda_i(\mathbf{V}^0) < \lambda_i(\mathbf{V}^1) \leq \lambda_i(\mathbf{V}^0) + a^{53} ,$$

we find the result

$$\mathbf{V}^* \left(\frac{x - x_0}{t - t_0} \right) = \mathbf{V}^0$$

along the straight lines

$$t - t_0 = \frac{1}{\lambda_i(\mathbf{V}^0)} (x - x_0) ,$$

so that the solution $\mathbf{V}^*(\xi)$ is constant along the lines and equals \mathbf{V}^0 if this is its value at ξ_0 . The same arguments show that the solution equals \mathbf{V}^1 along the straight lines

$$t - t_0 = \frac{1}{\lambda_i(\mathbf{V}^1)} (x - x_0)$$

if this is its value prescribed at ξ_0 .

In other words: In case of $\lambda_i(\mathbf{V}^0) < \lambda_i(\mathbf{V}^1)$, the state \mathbf{V}^0 can be **smoothly connected from the right** with the state \mathbf{V}^1 by a centered i -rarefaction wave⁵⁴.

⁵³In case of $\lambda_i(\mathbf{V}^1) < \lambda_i(\mathbf{V}^0)$, change the roles of \mathbf{V}^1 and \mathbf{V}^0 .

⁵⁴cf. also the definition that follows formula (3.43)

Remark:

In the scalar case and if f is strictly convex, $\lambda_i(\mathbf{V}^0) < \lambda_i(\mathbf{V}^1)$ means $f'(v^0) < f'(v^1)$, hence

$$v^0 < v^1 .$$

In other words: If t is fixed, the states decrease for decreasing x , i.e. from the right to the left (rarefaction)⁵⁵. Indeed, if the entropy condition (3.20) is fulfilled, this rarefaction can only be caused by a continuous process because of $v_\ell = v_r$.

Assume now the system (2.1) to have discontinuous solutions besides or instead of smooth rarefaction waves, and let us assume that these discontinuities form certain curves Γ along which the Rankine-Hugoniot condition is satisfied.

We ask again for conditions easy to apply in order to characterize **uniquely** the physically relevant entropy solution⁵⁶.

Here, we restrict ourselves to problems whose characteristic fields are genuinely non-linear on $N = \mathbb{R}^m$ ⁵⁷.

A weak solution will be called an entropy solution if it respects a condition of the following type:

- If initial values are prescribed along a non-characteristic curve, the condition will only be fulfilled by one weak solution.
- The condition coincides with (3.24) as far as scalar problems are concerned.

If $[\mathbf{V}] = \mathbf{V}_\ell - \mathbf{V}_r$ are the jumps along a curve Γ of discontinuities of a weak solution moving along the x -axis with speed \hat{v} , let k be the particular integer with

$$\lambda_k(\mathbf{V}_r) < \hat{v} < \lambda_{k+1}(\mathbf{V}_r) , \quad k \in \{1, 2, \dots, m-1\} , \quad (3.40)$$

or $\hat{v} < \lambda_1(\mathbf{V}_r)$ or $\hat{v} > \lambda_m(\mathbf{V}_r)$ ⁵⁸.

In order to determine the m components of \mathbf{V}_r at $P_0 \in \Gamma$ from the given initial values, only the values of \mathbf{V}_0 along the particular characteristics through P_0 are available which show the steeper ascents

$$\frac{1}{\lambda_k(\mathbf{V}_0)} , \frac{1}{\lambda_{k-1}(\mathbf{V}_0)} , \dots , \frac{1}{\lambda_1(\mathbf{V}_0)} \quad (\text{cf. Fig. 18})^{59} .$$

⁵⁵If the flux f is a concave function like the flux of the traffic flow model, the word 'left' has to be replaced by 'right' and 'decrease' by 'increase'.

⁵⁶Remember that condition (3.12) could not necessarily be applied in case of more than two equations.

⁵⁷analogously to scalar problems, i.e to $m = 1$ where only convex fluxes were taken into account

⁵⁸cf. (3.30)

If exemplarily a system of decoupled equations is under consideration⁶⁰, along each of these k steeper characteristics one and only one of the m components of $\mathbf{V}_r(P_0)$ can be determined. Hence, $m - k$ informations are missing.

If simultaneously

$$\lambda_j(\mathbf{V}_\ell) < \hat{v} < \lambda_{j+1}(\mathbf{V}_\ell) ,$$

holds, corresponding arguments lead to the result that there are only $m - j$ informations available for the computation of the m components of \mathbf{V}_ℓ . Hence, j informations are missing.

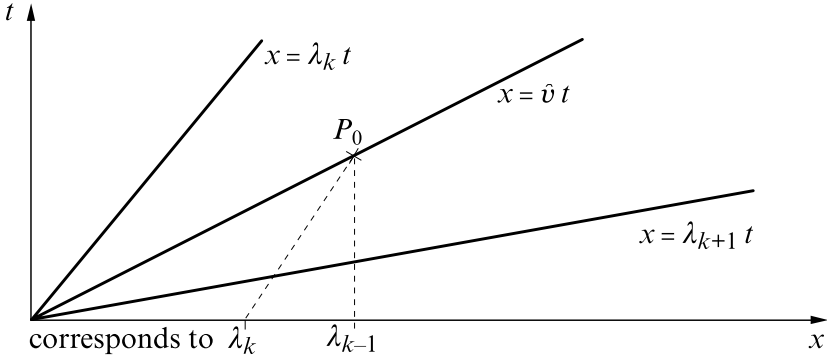


Figure 18: Sketch concerning the completeness of informations needed in order to ensure uniqueness

Thus, $m - k + j$ informations are missing for the unique determination of the two quantities \mathbf{V}_ℓ and \mathbf{V}_r . But m additional informations concerning the relations between \mathbf{V}_ℓ , \mathbf{V}_r and \hat{v} follow from the Rankine-Hugoniot condition (2.15). One of them must be used for the elimination of \hat{v} . Thus, the real number of missing informations is only $1 - k + j$. In other words, **uniqueness** can only be expected for

$$k = j + 1,$$

i.e., if the relations

$$\lambda(\mathbf{V}_r) < \hat{v} < \lambda_{k+1}(\mathbf{V}_r) \quad , \quad \lambda_{k-1}(\mathbf{V}_\ell) < \hat{v} < \lambda_k(\mathbf{V}_\ell) \quad (3.41)$$

hold simultaneously.

⁵⁹The particular choice $\mathbf{0} \in \Gamma$ in Fig. 18 does not lead to a loss of generality.

⁶⁰cf. the linear situation (1.19)

Although this result was only illustrated by an exemplary situation, we accept it as a hint on a general definition of an entropy solution:

Definition:

\mathbf{V} is called an **entropy solution** if there is an index k for which (3.41) holds. A discontinuity of this type is then called a **k -shock** and the inequalities (3.41) are called **entropy inequalities** or **(Lax-) shock conditions**.

(3.41) includes

$$\lambda_k(\mathbf{V}_r) < \hat{v} < \lambda_k(\mathbf{V}_\ell),$$

so that (3.41) really coincides with (3.24) as far as scalar problems are concerned. Thus, both properties claimed earlier with respect to the formulation of an entropy condition for weak solutions of systems of conservation laws are respected.

There are some connections between an i -characteristic field and the so-called i -Riemann-invariant:

Definition:

A scalar function $w = w(\mathbf{V})$ defined for all $\mathbf{V} \in N \subset \mathbb{R}^m$ is called an **i -Riemann invariant** on N , if

$$\langle \nabla_{\mathbf{V}} w(\mathbf{V}), \mathbf{s}_i(\mathbf{V}) \rangle = 0 \quad \forall \mathbf{V} \in N. \quad (3.42)$$

Remark:

Thus, the property of an i -characteristic field to be genuinely nonlinear, coincides with the property of $\lambda_i(\mathbf{V})$ not to be an i -Riemann invariant.

If the graph of a centred i -rarefaction wave \mathbf{V} is situated in N , its i -Riemann-invariants are constant. This follows from (3.34):

$$\partial_x w(\mathbf{V}(x, t)) = \langle \nabla_{\mathbf{V}} w(\mathbf{V}), \mathbf{s}_i(\mathbf{V}) \rangle_{\mathbf{v}=\mathbf{v}^*(x, t)} = 0 \quad (3.43)$$

and similarly $\partial_t w(\mathbf{V}(x, t)) = 0$.

Definition:

If a solution \mathbf{V} of the system (2.1) belonging to any initial function \mathbf{V}_0 is smooth on a region $G \subset \Omega$, and if all the i -Riemann-invariants $w(\mathbf{V}(x, t))$ are constant on G , \mathbf{V} is called an **i -simple wave** or an **i -rarefaction wave** on G .

3.4 The Ansatz due to Kruzkov

The way used by S.N. Kruzkov⁶¹ in order to introduce the idea of a weak solution of problem (2.1) seems to differ from the idea of Lax. In order to understand the situation, it is sufficient to restrict the presentation of Kruzkov's Ansatz to only one-dimensional scalar problems.

Also Kruzkov tries to find his weak solutions in the space $L_1^{\text{loc}}(\Omega)$ as it was already done in the Lax definition of a weak solution. Moreover, also the test functions used in (2.2.) by Lax play a crucial role in Kruzkov's definition, too, with the one restriction that these functions now vanish along the whole boundary of their compact support, i.e. also along parts of the boundary belonging to the x -axis: $\Phi(x, 0) \equiv 0$. This leads to the necessity to pay regard to the initial function in another way than by the second integral in formula (2.2).

Kruzkov calls a function $v \in L_1^{\text{loc}}(\Omega)$ a weak solution of the problem

$$\partial_t v + \partial_x f(v) = 0, \quad v(x, 0) = v_0(x),$$

if the following two conditions are fulfilled:

$$\int_{\Omega} \left\{ \partial_t \Phi(x, t) |v(x, t) - c| + \partial_x \Phi(x, t) \operatorname{sgn}(v(x, t) - c) [f(v(x, t)) - f(c)] \right\} dx \geq 0 \quad (3.44)$$

$$\forall \Phi \in C_0^1(\Omega) \text{ with } \Phi \geq 0 \text{ and with } \Phi(x, 0) \equiv 0, \quad \forall c \in \mathbb{R},$$

and

$$\lim_{t \rightarrow 0} \int_{-R}^R |v(x, t) - v_0(x)| dx = 0 \quad \forall R \in \mathbb{R} \quad (t \in [0, T] \setminus \tilde{\epsilon}) \quad (3.45)$$

for a value of $T > 0$ and for a set $\tilde{\epsilon}$ of measure zero so that $v(x, t)$ is almost everywhere well defined as a function of x for every fixed $t \in [0, T] \setminus \tilde{\epsilon}$.

Kruzkov could show that there is a unique weak solution in $L_1^{\text{loc}}(\Omega)$ in the sense of (3.44), (3.45) provided that the flux f is smooth and strictly convex. Smoothness and strict convexity of the flux were assumptions also made by Oleinik and Lax.

⁶¹Soviet Math. Dokl. **10** (1969), 785-788

Hence, in Kruzkov's definition, the characterization of an entropy solution does not consist of an equation (2.2) to be solved for all test functions together with an additional entropy condition (3.12) but only of an inequality⁶².

One can say that Kruzkov's definition already includes an entropy condition.

Indeed, as far as the scalar problem is concerned, the left side of (3.44) corresponds with (3.12) if $\Phi(x, 0) = 0$ is taken into account, if for every particular constant $c \in \mathbb{R}$ the entropy functional $\tilde{S}(v)$ will be identified with $|v - c|$ and the entropy flux $\tilde{F}(v)$ with $\text{sgn}(v - c) [f(v) - f(c)]$.

If \tilde{S} and \tilde{F} are chosen in this way, also the connection between these two functions demanded by (3.13) is kept piecewise, namely for all values of v with the exception of $v = c$. Moreover, the particular functional \tilde{S} is also convex as entropy functionals were expected to be but no longer strictly convex, unfortunately.

On the other hand, there is a somewhat stronger demand than formulated by Lax, namely that the validity of (3.44) has not only to be ensured for all nonnegative test functions Φ but at the same time also for all $c \in \mathbb{R}$. Thus, (3.45) must now hold for all test *elements*

$$\hat{\Phi} \in \{(\Phi, c) \mid \Phi \in C_0^1(\Omega), \Phi \geq 0, \Phi(x, 0) = 0; c \in \mathbb{R}\}. \quad (3.46)$$

It can be shown by more than one chain of evidence that Kruzkov's weak solution coincides with the Lax-Oleinik entropy solution provided that the flux is smooth and strictly convex. One of these chains follows from numerical aspects: There are finite difference approximations whose numerical solutions converge for decreasing step sizes to the Lax-Oleinik entropy solution but also –as can be proved by a slightly different convergence argument– to the Kruzkov solution⁶³.

It should be noted that there are problems from fields of applications where the assumptions on smoothness and strict convexity of the flux are violated. But Kruzkov together with Panov could show that a unique solution of the problem (3.44), (3.45) does even exist if f is merely continuous⁶⁴. Thus, in the scalar situation, the Kruzkov concept of an entropy solution is more far-reaching.

But one has to pay for the renunciation of convexity and smoothness: It can happen that one of the properties of the entropy solution v_E important from the point of view of numerical methods gets lost, namely that an initial function v_0 with compact

⁶²and of a condition that respects the initial condition in a suitable way

⁶³A non-numerical argument concerning the connection between both definitions can be found in the paper already mentioned in footnote 38.

⁶⁴S.N. Kruzkov und E.Y. Panov: Soviet Math.Dokl. **42** (1991), 316-321

support leads to a compact support of $v_E(\cdot, t)$ (for every fixed $t > 0$) as it is the case if f is smooth and convex. A test example for this loss is realized by

$$f(v) = \frac{|v|^\alpha}{\alpha} \quad \text{with} \quad 0 < \alpha < 1, \quad (3.47)$$

$$v_0(x) = \begin{cases} 0 & \text{for } x < -1 \\ 1 & \text{for } -1 \leq x \leq 0 \\ 0 & \text{for } x > 0 \end{cases} .$$

The solution reads

$$v_E(x, t) = \begin{cases} 0 & \text{for } t > \alpha(x+1) \\ 1 & \text{for } x < t \leq \alpha(x+1) \\ \left(\frac{t}{x}\right)^{\frac{1}{1-\alpha}} & \text{for } t \leq x \end{cases} \quad (3.48)$$

(cf. Fig. 19).

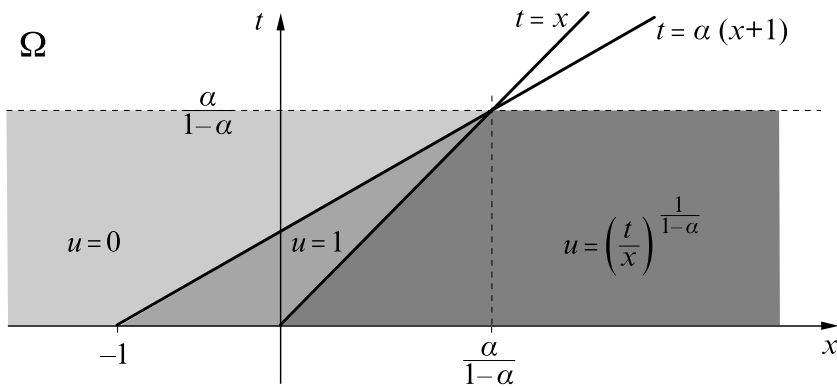


Figure 19: Kruzkov-Panov solution (3.48) for $0 \leq t \leq \frac{\alpha}{1-\alpha}$

Whereas the smooth and strictly convex case behaves hyperbolicly, a merely continuous flux can lead to a more parabolic character of the problem. Numerical procedures for problems of this particular type were under some suitable conditions recently developed for the first time by M. Breuss⁶⁵. Here already the Kruzkov-Panov example (3.47) shows that the important CFL-condition (cf. sections 6.1, 7.3) can not be

⁶⁵M. Breuss: Numerik von Erhaltungsgleichungen in Nicht-Standard-Situationen, PhD-Thesis, Dept. of Mathematics, University of Hamburg, 2001

fulfilled as far as explicit finite difference methods (cf. sections 4.1, 6.4, 7.2) are used. Thus, implicit methods have to be taken into account, and this can really be done successfully.

Examples of mathematical models of real world problems in fluid mechanics where standard assumptions like strict hyperbolicity, strict convexity or smoothness of the fluxes are not necessarily fulfilled occur in the theory of flows through porous media, e.g. if oil production is concerned:

The scalar so-called non-standard **Buckley-Leverett equation**

$$\partial_t v + \partial_x \left(\frac{v^2}{v^2 + a(1-v)^2} \right) = 0, \quad a \in \mathbb{R}$$

models the displacement of an oleic phase by an aqueous phase in a porous medium. In this situation, the flux is neither convex nor concave.

Tveito and Winther⁶⁶ considered theoretically with respect to stability the system

$$\begin{aligned} \partial_t s + \partial_x f(s, c) &= 0 \\ \partial_t (sc + a(c)) + \partial_x (cf(s, c)) &= 0, \end{aligned}$$

which also models the displacement of oil by water in a porous medium. But in this case, the water is thickened by dissolved polymer in order to increase its viscosity. The difficulty of this example arises from the fact that the eigenvalues of the Jacobian are real but can coincide so that the system is hyperbolic but not necessarily strictly hyperbolic.

⁶⁶Proc. 3rd Intern. Conf. Hyperb. Probl. (B. Engquist, B. Gustafsson, ed.), (Chartwell-Bratt, Uppsala), 888-898 (1991)

4 The Riemann Problem

4.1 Numerical Importance of the Riemann Problem

Treating the traffic jam dissolution example we met the first time with a Riemann problem. It consisted of a conservation law of type (2.1) together with an initial function being piecewise constant with certain jumps at isolated positions.

One tool constructed in order to solve differential equations numerically is the so-called **Finite Difference Method (FDM)** where one tries the unknown values of the exact solution at isolated points to be approximated (cf. also chapters 6, 7).

Particularly in case of problem (2.1), we ask for numerical values $V_\nu^{(n)}$ at isolated positions (x_ν, t_n) of the upper half plane Ω expected to be good approximations for the values $V_E(x_\nu, t_n)$ of the exact entropy solution at these points.

A particular possibility for the construction of such FDMs consists of covering the positive time axis in a first step by a not necessarily equidistant **grid** $\{t_n \mid (n = 0, 1, \dots)\}$ of **time step sizes** $\Delta t_n := t_{n+1} - t_n$ together with fixing isolated grid points x_ν ; ($\nu = 0, \pm 1, \pm 2, \dots$) along the x -axis. Thus, by the cartesian product of these two sets, Ω will be covered with a net of grid points parallel to the axes. Also the step sizes $h_\nu := x_{\nu+1} - x_\nu$ don't need necessarily to be equidistant (cf. Fig. 20).

If the computation of the values $V_\nu^{(n+1)}$ along the time level $t = t_{n+1}$ does only use the knowledge of the values $V_\nu^{(n)}$ along the time level before and computed in the very last step, we call the scheme a **one step method**. **Multi step methods** also include approximate values along earlier levels, but we are going to restrict the present survey to one step methods.

Let $t_0 = 0$ be the level to start from with approximate initial values

$$V_\nu^{(0)} := \frac{2}{x_{\nu+1} - x_{\nu-1}} \int_{\frac{1}{2}(x_{\nu-1} + x_\nu)}^{\frac{1}{2}(x_\nu + x_{\nu+1})} V_0(x) dx \quad (\nu = 0, \pm 1, \pm 2, \dots). \quad (4.1)$$

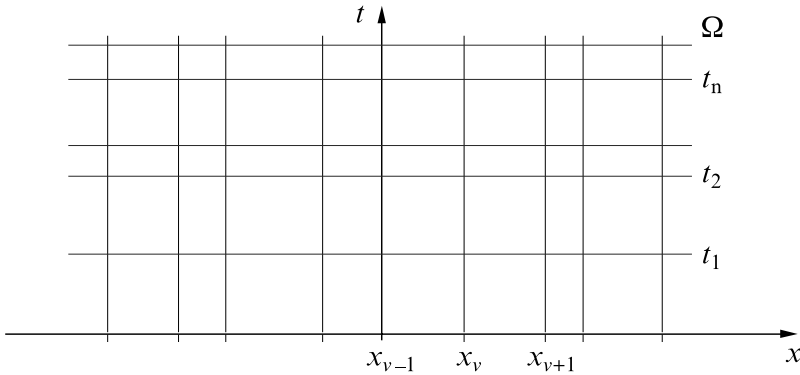


Figure 20: Grid of an FDM

It seems to be reasonable to describe the difference between the approximate solution and the unknown exact solution in terms of the topology of the space forming the basis of the original problem, which is in the particular situation under consideration the space $L_1^{\text{loc}}(\Omega)$. This difference or **truncation error** should be estimated using a suitable norm. Hopefully, these errors will decrease for decreasing step sizes and will tend to zero sufficiently fast if the step sizes tend to zero. If this holds, the method is called **convergent**.

Obviously, a direct comparison of the approximate solution to the exact solution in terms of the topology of the original space can take place if the discrete values $\{\mathbf{V}_{\nu}^{(n)}\}$ of the approximate solution defined only at isolated points can be extended to the inter grid points suitably, e.g. by

$$\hat{\mathbf{V}}(x, t) := \mathbf{V}_{\nu}^{(n)} \quad \text{for} \quad \frac{x_{\nu-1} + x_{\nu}}{2} < x \leq \frac{x_{\nu} + x_{\nu+1}}{2}, \quad t_n \leq t < t_{n+1}, \quad (4.2)$$

$$(\nu = 0, \pm 1, \pm 2, \dots; n = 0, 1, 2, \dots).$$

By this special extension or **reconstruction** the discrete function now becomes a function $\hat{\mathbf{V}}$ constant on each of the described rectangles. Hence, $\hat{\mathbf{V}}$ is a step function on Ω and, thus, an element of $L_1^{\text{loc}}(\Omega)$.

If we restrict $\hat{\mathbf{V}}$ to the time level $t = t_n$, this restricted function $\hat{\mathbf{V}}(x, t_n)$ of the space variable x is a step function, too, from which the one step method together with the reconstruction now generates $\hat{\mathbf{V}}(x, t_{n+1})$.

If $\hat{\mathbf{V}}(x, t_n)$ is used as an initial function for the original problem in order to generate a weak solution of the differential equation in (2.1.) for the region $t \geq t_n$, a Riemann

problem comes about. Its exact solution if particularly restricted to the time level $t = t_{n+1}$ would lead at the grid points along this level to values which could be looked upon as approximations to the weak solution of the original problem at these new level grid points. This idea was for the first time realized by **S.K. Godunov**⁶⁷ and was later often used by other authors as a basic idea for further numerical schemes. Herewith, effective approximate **Riemann solvers** were developed including local linearizations.

We are therefore now going to study the Riemann problem at least in case of linear problems with constant coefficients.

4.2 The Riemann Problem in the Case of Linear Systems

Let us recall problem (1.19), i.e.

$$\partial_t \mathbf{V} + \mathbf{A} \partial_x \mathbf{V} = \mathbf{0} \quad ,$$

with a constant (m,m) -matrix \mathbf{A} , and let us prescribe initial values

$$\mathbf{V}_0(x) = \begin{cases} \mathbf{V}_\ell & \text{for } x < 0 \\ \mathbf{V}_r & \text{for } x > 0 \end{cases} \quad (4.3)$$

with constant vectors \mathbf{V}_ℓ and \mathbf{V}_r different from each other. It causes no loss of generality using the origin as the position of the jump.

Assume the problem to be again strictly hyperbolic so that the eigenvalues λ_i of the matrix \mathbf{A} are real and different from each other. The eigenvectors belonging to these eigenvalues are then obviously real, too, and linearly independent. Thus, they can be used as a basis in \mathbb{R}^m , and in areas where the solution is smooth, (1.21) and (1.22) yield

$$\mathbf{V}(x, t) = \sum_{i=1}^m \hat{v}_{i0}(x - \lambda_i t) \mathbf{s}_i \quad (4.4)$$

with

$$\hat{\mathbf{V}}_0(x) = (\hat{v}_{10}(x), \hat{v}_{20}(x), \dots, \hat{v}_{m0}(x))^T = \mathbf{S}^{-1} \mathbf{V}_0(x) .$$

This leads particularly to

$$\hat{\mathbf{V}}_{0\ell}(x) =: (w_1, w_2, \dots, w_m)^T = \mathbf{S}^{-1} \mathbf{V}_\ell \quad \text{for } x < 0$$

⁶⁷Mat. Sb. **47** (1959), 357-393

and analogously to

$$\hat{\mathbf{V}}_{0_r}(x) =: (\tilde{w}_1, \tilde{w}_2, \dots, \tilde{w}_m)^T = \mathbf{S}^{-1} \mathbf{V}_r \quad \text{for } x > 0$$

with certain constants w_i and \tilde{w}_i .

If the eigenvalues are ordered as in (3.30),

$$\mathbf{V}(x, t) = \sum_{i=1}^{k_0} \tilde{w}_i \mathbf{s}_i + \sum_{i=k_0+1}^m w_i \mathbf{s}_i \quad (4.5)$$

follows from (4.4) provided that for a given position (x, t) the first k_0 eigenvalues are the particular ones for which $x - \lambda_i t > 0$ ($i = 1, 2, \dots, k_0$), whereas $x - \lambda_i t < 0$ ($i = k_0 + 1, k_0 + 2, \dots, m$) holds for the other $m - k_0$ eigenvalues.

Let us now look at the i -characteristics $x = \lambda_i t$ arising from the jump position and let $t_0 > 0$ be fixed. Let us move for this fixed t_0 from the left to the right along a line parallel to the x -axis. This line crosses the k_0 -characteristic at the position $x_0 = \lambda_{k_0} t_0$ (cf. Fig. 21). As soon as this will happen the coefficient \tilde{w}_{k_0} in formula (4.4) has to be replaced by w_{k_0} .

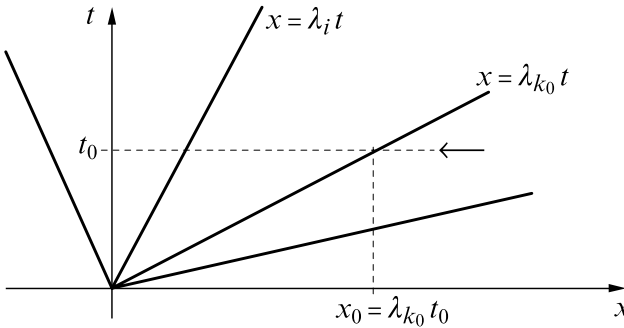


Figure 21: How to solve the linear Riemann problem

The solution \mathbf{V} described by (4.5) then obviously jumps. The height of this jump is

$$[\mathbf{V}]_{k_0}(x_0, t_0) = (w_{k_0} - \tilde{w}_{k_0}) \mathbf{s}_{k_0} \quad (4.6)$$

and leads to

$$\mathbf{A} [\mathbf{V}]_{k_0}(x_0, t_0) = (w_{k_0} - \tilde{w}_{k_0}) \lambda_{k_0} \mathbf{s}_{k_0} = \lambda_{k_0} [\mathbf{V}]_{k_0}(x_0, t_0). \quad (4.7)$$

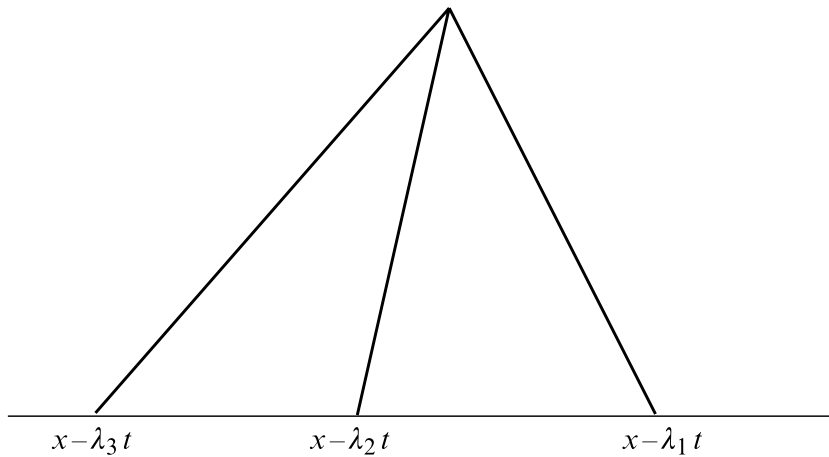


Figure 22: Characteristic cones of the solution of the linear Riemann problem

We know already that discontinuities move along characteristics as far as linear problems are concerned. This makes the k_0 -characteristic line through the jump position a k_0 -shock Γ_{k_0} if $w_{k_0} \neq \tilde{w}_{k_0}$ so that (4.7) just becomes the Rankine-Hugoniot condition (2.15) applied to the special linear situation.

(4.5) and (4.7) yield the result that the solution of our Riemann problem for $t > 0$ can also be formulated as

$$\mathbf{V}(x, t) = \mathbf{V}_\ell - \sum_{\substack{i \\ \lambda_i < \frac{x}{t}}} (w_i - \tilde{w}_i) \mathbf{s}_i \quad (4.8)$$

or as

$$\mathbf{V}(x, t) = \mathbf{V}_r + \sum_{\substack{i \\ \lambda_i > \frac{x}{t}}} (w_i - \tilde{w}_i) \mathbf{s}_i \quad (4.9)$$

(cf. Fig. 22).

5 Real Fluids

5.1 The Navier-Stokes Equations Model

An ideal fluid was defined to be a material without friction between neighbouring particles of different flow velocities, hence without tangential forces along the surfaces of arbitrary volumes $W(t)$ picked out of the fluid flow. The mathematical model of ideal fluids can therefore only describe certain particular situations, e.g. flows of low speeds, low densities, low viscosities or high temperatures etc., and its results can seriously contradict the reality if such properties are not ensured. The forces acting on bridge piers parallel to the flow of calmly flowing small rivers were already mentioned when the Kutta-Zhukovsky formulas were presented in section 1.3.

Thus, if the tangential forces generated by the viscosity play an important role in a real flow situation they can't be neglected in a mathematical model expected to be able to describe the flow in a satisfactory manner.

Let us therefore try to revise the Euler equations model of ideal fluids by taking the viscosity forces additionally into account in order to generate a model of real fluids.

Reminding the reader to the fact that viscosity in our context always means friction between fluid particles⁶⁸, these additional forces can only occur in case of motion, i.e. for non-vanishing velocity of the flow, and lead then to an additional term within the conservation law of momentum:

The righthand side of (1.5) consisted of the forces per unit volume \mathbf{k} only, and the part of these forces falling to the volume $W(t)$ therefore was given by

$$\int_{W(t)} \mathbf{k} d(x, y, z).$$

⁶⁸cf. (1.54) and its explanation

Forces generated by friction of the fluid particles on the outside of this volume but moving along its surface $\partial W(t)$ are obviously proportional to the area of $\partial W(t)$ and have tangential directions, whereas the forces generated by pressure of the outside fluid are proportional to this area, too, but directed perpendicular to $\partial W(t)$.

If $\mathbf{s} = (s_1, s_2, s_3)^T$ are the forces per unit area generated by the outer fluid and acting on the surface $\partial W(t)$ of an inner volume $W(t)$ arbitrarily picked out of the fluid, the relation

$$\mathbf{s} = \mathbf{s}(t, x, y, z, \mathbf{n}), \quad (x, y, z) \in \partial W(t) \quad (5.1)$$

holds where \mathbf{n} is the unit vector normal on $\partial W(t)$ at $(x, y, z) \in \partial W(t)$ with outward direction.

Herewith, \mathbf{s} depends linearly on \mathbf{n} :

$$\mathbf{s} = (\sigma_{ij}) \mathbf{n} \quad (5.2)$$

with $\sigma_{ij} : (x, y, z, t) \rightarrow \sigma_{ij}(x, y, z, t)$.

(5.2) follows from the fact that within the conservation law of momentum, now to be written as

$$\begin{aligned} \int_{W(t)} \left\{ \partial_t \mathbf{q} + \left\langle \frac{1}{\rho} \mathbf{q}, \nabla \right\rangle \mathbf{q} + \operatorname{div} \left(\frac{1}{\rho} \mathbf{q} \right) \mathbf{q} \right\} d(x, y, z) \\ = \int_{W(t)} \mathbf{k} d(x, y, z) + \int_{\partial W(t)} \mathbf{s} do, \end{aligned}$$

the volume terms vanish of a higher order than the surface terms in case of $W(t) \rightarrow 0$.

Here, the **stress tensor** (σ_{ij}) is treated like a real (3, 3)–matrix.

Moreover, **Euler's law of vanishing momentum**, i.e.

$$\int_{W(t)} [\mathbf{r}, \mathbf{k}] d(x, y, z) + \int_{\partial W(t)} [\mathbf{r}, \mathbf{s}] do = \mathbf{0}, \quad (5.3)$$

$\mathbf{r} = (x, y, z)^T$, yields the symmetry of the stress tensor:

$$\sigma_{ij} = \sigma_{ji}. \quad (5.4)$$

Vanishing friction in case of ideal fluids has to be modelled by $\sigma_{ij} = 0$ for $i \neq j$. Taking (5.2) into account, this leads to

$$\sigma_{11} = \sigma_{22} = \sigma_{33} =: -p \quad (\text{pressure})$$

because the vectors \mathbf{s} and \mathbf{n} are parallel in this situation. This is a slightly more general explanation of (1.5).

As far as real fluids are under consideration, it can be expected the terms σ_{ij} ($i \neq j$) to be small, and the terms σ_{ii} will certainly differ from $-p$ only to a small degree. We therefore put

$$\tilde{\sigma}_{ij} = \begin{cases} \sigma_{ii} + p & \text{for } i = j \\ \sigma_{ij} & \text{for } i \neq j, \end{cases} \quad (5.5)$$

so that all the $\tilde{\sigma}_{ij}$ are small.

Because of the symmetry of (σ_{ij}) , also the tensor $(\tilde{\sigma}_{ij})$ is symmetric.

If the exterior forces \mathbf{k} per unit volume are replaced by the forces $\hat{\mathbf{k}}$ per unit mass,

$$\begin{aligned} & \int_{W(t)} \left\{ \partial_t \mathbf{q} + \left\langle \frac{1}{\rho} \mathbf{q}, \nabla \right\rangle \mathbf{q} + \operatorname{div} \left(\frac{1}{\rho} \mathbf{q} \right) \mathbf{q} - \rho \hat{\mathbf{k}} \right\} d(x, y, z) \\ &= - \int_{\partial W(t)} p \mathbf{n} \, do + \int_{\partial W(t)} (\tilde{\sigma}_{ik}) \mathbf{n} \, do \end{aligned}$$

follows.

Obviously, if a fluid does not move or if neighbouring particles do not move with different velocities, friction can't occur. In modelling the flow, we therefore assume the terms $\tilde{\sigma}_{ij}$ ($i, j = 1, 2, 3$) to depend only on the components of the vectors ∇u_i ($i = 1, 2, 3$). Here, we restrict ourselves to so-called **Newtonian Fluids**, i.e. the terms $\tilde{\sigma}_{ij}$ depend on the components ∇u_i homogeneously and linearly:

$$\tilde{\sigma}_{ij} = \sum_{\mu, \nu=1}^3 \alpha_{\mu\nu}^{(i,j)} \partial_{x_\mu} u_\nu. \quad (5.5)$$

Thus, every $\tilde{\sigma}_{ij}$ is a function of the elements of the Jacobian $\mathbf{J}\mathbf{u}$:

$$\tilde{\sigma}_{ij} : \mathbf{J}\mathbf{u} \rightarrow \tilde{\sigma}_{ij}(\mathbf{J}\mathbf{u}), \quad (5.6)$$

and the $\alpha_{\mu\nu}^{(i,j)}$ are constants.

Analogously to the case of ideal fluids where every axis is a main stress axis, we also expect in our model the relation (5.6) to be independent of orthogonal linear transformations, i.e.

$$\mathbf{A}^T (\tilde{\sigma}_{ij}) \mathbf{A} = (\tilde{\sigma}_{ij} (\mathbf{A}^T \mathbf{J} \mathbf{u} \mathbf{A})) , \quad \forall \mathbf{A} \in \mathbb{R}^{(3,3)} \quad \text{with} \quad \mathbf{A}^T \mathbf{A} = \mathbf{I} . \quad (5.7)$$

It is trivial that the matrix $\mathbf{B} := \text{div } \mathbf{u} \cdot \mathbf{I}$ is symmetric, and because it is even a diagonal matrix whose diagonal elements equal each other it shows the property

$$\mathbf{A}^T \mathbf{B} \mathbf{A} = \mathbf{B} .$$

The elements b_{ij} of \mathbf{B} depend on the elements of the Jacobian, i.e. $\mathbf{B} = (b_{ij}(\mathbf{J}\mathbf{u}))$. Because of $\text{div } \mathbf{u} = \text{trace } \mathbf{J}\mathbf{u}$ and because traces of matrices are invariants of orthogonal transformations, also

$$(b_{ij} (\mathbf{A}^T \mathbf{J} \mathbf{u} \mathbf{A})) = (b_{ij}(\mathbf{J}\mathbf{u})) = \mathbf{B}$$

follows.

Thus, \mathbf{B} fulfills all the expected demands on $(\tilde{\sigma}_{ij})$. But also

$$\mathbf{D} := \mathbf{J}\mathbf{u} + (\mathbf{J}\mathbf{u})^T =: \mathbf{D}(\mathbf{J}\mathbf{u})$$

is symmetric, depends on the elements of the Jacobian in a homogeneous and linear way and shows the property

$$\begin{aligned} \mathbf{A}^T \mathbf{D} \mathbf{A} &= \mathbf{A}^T \mathbf{J} \mathbf{u} \mathbf{A} + \mathbf{A}^T (\mathbf{J} \mathbf{u})^T \mathbf{A} \\ &= \mathbf{A}^T \mathbf{J} \mathbf{u} \mathbf{A} + (\mathbf{A}^T \mathbf{J} \mathbf{u} \mathbf{A})^T , \end{aligned}$$

$$\text{i.e.} \quad \mathbf{A}^T \mathbf{D}(\mathbf{J}\mathbf{u}) \mathbf{A} = \mathbf{D}((\mathbf{A}^T \mathbf{J} \mathbf{u} \mathbf{A})) .$$

Thus, also \mathbf{D} fulfills all the demands on $(\tilde{\sigma}_{ij})$, and therefore also $\lambda \mathbf{B} + \eta \mathbf{D}$ with arbitrary constants λ, η does so. It can be shown that these are already all types of matrices which satisfy our demands. Hence,

$$\begin{aligned} \int_{W(t)} \left\{ \partial_t \mathbf{q} + \left\langle \frac{1}{\rho} \mathbf{q}, \nabla \right\rangle \mathbf{q} + \text{div} \left(\frac{1}{\rho} \mathbf{q} \right) \mathbf{q} - \rho \hat{\mathbf{k}} \right\} d(x, y, z) \\ - \int_{\partial W(t)} (-p \mathbf{I} + \lambda \mathbf{B} + \eta \mathbf{D}) \mathbf{n} \, do = \mathbf{0} . \end{aligned} \quad (5.8)$$

By means of the divergence theorem,

$$\int_{\partial W(t)} p \mathbf{n} \, do = \int_{W(t)} \nabla p \, d(x, y, z)$$

and

$$\int_{\partial W(t)} \mathbf{B} \mathbf{n} \, do = \int_{W(t)} \begin{pmatrix} \operatorname{div} \mathbf{b}_1 \\ \operatorname{div} \mathbf{b}_2 \\ \operatorname{div} \mathbf{b}_3 \end{pmatrix} d(x, y, z)$$

follow where the \mathbf{b}_i are the row vectors of \mathbf{B} , i.e.

$$\int_{\partial W(t)} \mathbf{B} \mathbf{n} \, do = \int_{W(t)} \nabla (\operatorname{div} \mathbf{u}) \, d(x, y, z),$$

and similarly

$$\int_{\partial W(t)} \mathbf{D} \mathbf{n} \, do = \int_{W(t)} \begin{pmatrix} \operatorname{div} \mathbf{d}_1 \\ \operatorname{div} \mathbf{d}_2 \\ \operatorname{div} \mathbf{d}_3 \end{pmatrix} d(x, y, z)$$

where

$$\begin{aligned} \operatorname{div} \mathbf{d}_1 &= 2 \partial_x (\partial_x u_1) + \partial_y (\partial_y u_1 + \partial_x u_2) + \partial_z (\partial_z u_1 + \partial_x u_3) \\ &= \Delta u_1 + \partial_{xx} u_1 + \partial_{xy} u_2 + \partial_{xz} u_3 \\ &= \Delta u_1 + \partial_x (\operatorname{div} \mathbf{u}). \end{aligned}$$

Analogous formulas hold for $\operatorname{div} \mathbf{d}_2$ and $\operatorname{div} \mathbf{d}_3$.

The equation (5.8) therefore reads as

$$\int_{W(t)} \left\{ \partial_t \mathbf{q} + \left\langle \frac{1}{\rho} \mathbf{q}, \nabla \right\rangle \mathbf{q} + \operatorname{div} \left(\frac{1}{\rho} \mathbf{q} \right) \mathbf{q} - \rho \hat{\mathbf{k}} \right. \\ \left. + \nabla p - \lambda \nabla (\operatorname{div} \mathbf{u}) - \eta \Delta \mathbf{u} - \eta \nabla (\operatorname{div} \mathbf{u}) \right\} d(x, y, z) = \mathbf{0}$$

for every part $W(t)$ picked arbitrarily out of the volume of the fluid, i.e.

$$\partial_t (\rho \mathbf{u}) + \langle \mathbf{u}, \nabla \rangle (\rho \mathbf{u}) + (\rho \operatorname{div} \mathbf{u}) \mathbf{u} - \rho \hat{\mathbf{k}} + \nabla p - (\lambda + \eta) \nabla (\operatorname{div} \mathbf{u}) - \eta \Delta \mathbf{u} = \mathbf{0}$$

or

$$\begin{aligned} \partial_t \rho \mathbf{u} + \rho \partial_t \mathbf{u} + \langle \mathbf{u}, \nabla \rho \rangle \mathbf{u} + \rho \langle \mathbf{u}, \nabla \rangle \mathbf{u} + (\rho \operatorname{div} \mathbf{u}) \mathbf{u} - \rho \hat{\mathbf{k}} + \nabla p \\ - (\lambda + \eta) \nabla (\operatorname{div} \mathbf{u}) - \eta \Delta \mathbf{u} = 0. \end{aligned}$$

Because of $\rho \operatorname{div} \mathbf{u} + \langle \mathbf{u}, \nabla \rho \rangle = \operatorname{div}(\rho \mathbf{u})$, the relation

$$\{\partial_t \rho + \operatorname{div}(\rho \mathbf{u})\} \mathbf{u} + \rho \partial_t \mathbf{u} + \rho \langle \mathbf{u}, \nabla \rangle \mathbf{u} - \rho \hat{\mathbf{k}} + \nabla p - (\lambda + \eta) \nabla (\operatorname{div} \mathbf{u}) - \eta \Delta \mathbf{u} = 0$$

follows, and taking the continuity equation into account, we conclude after division by ρ with the so-called **Navier⁶⁹ - Stokes⁷⁰ equations**

$$\boxed{\partial_t \mathbf{u} + \langle \mathbf{u}, \nabla \rangle \mathbf{u} - \frac{\lambda + \eta}{\rho} \nabla (\operatorname{div} \mathbf{u}) - \frac{\eta}{\rho} \Delta \mathbf{u} = - \frac{1}{\rho} \nabla p + \hat{\mathbf{k}}} \quad (5.9)$$

Here η is called the **first viscosity coefficient**. Of course, historically there were several types of arguments leading to the Navier - Stokes equations, and this led to the convention not to call λ the **second viscosity coefficient** but $\xi := \lambda + \frac{2}{3}\eta$. These coefficients are constants depending on the fluid material as well as on its temperature. The unit of measurement is $\text{g} \cdot \text{cm}^{-1} \cdot \text{sec}^{-1} = \mathbf{Poise}$ (cf. footnote 71).

In case of incompressible flows characterized by $\operatorname{div} \mathbf{u} = 0$, λ does not occur in the equations, and if also the exterior forces $\hat{\mathbf{k}}$ do not play a role, the Navier-Stokes equations are reduced to

$$\partial_t \mathbf{u} + \langle \mathbf{u}, \nabla \rangle \mathbf{u} - \nu \Delta \mathbf{u} + \frac{1}{\rho} \nabla p = \mathbf{0}, \quad (5.10)$$

where $\nu = \frac{\eta}{\rho}$ is called the **kinematic viscosity**.

By similar arguments, the law describing the conservation of energy takes for viscous flows the form

$$\partial_t E + \operatorname{div}((E + p) \mathbf{u}) - \lambda \cdot \operatorname{div}(\operatorname{div} \mathbf{u} \cdot \mathbf{u}) - \eta \cdot \operatorname{div} \left(\left(\mathbf{J} \mathbf{u} + (\mathbf{J} \mathbf{u})^T \right) \mathbf{u} \right) = 0.$$

Some of the terms within this equation can be simplified, e.g.

$$\operatorname{div}(\operatorname{div} \mathbf{u} \cdot \mathbf{u}) = \langle \nabla (\operatorname{div} \mathbf{u}), \mathbf{u} \rangle + (\operatorname{div} \mathbf{u})^2$$

etc.

⁶⁹Claude Louis Marie Henri Navier (1785 - 1836); Dijon, Paris

⁷⁰George Gabriel Stokes (1819 - 1903); Cambridge

The Navier-Stokes equations model is excellently confirmed by experiments.

An impression of the magnitude of the kinematic viscosity and its sensibility with respect to temperature is given by the following table:

Material	Temperature [$^{\circ}C$]	Kinematic viscosity [cm^2/sec]
Mercury	0	0.00125
	10	0.00123
	20	0.00117
Air	0	0.133
	10	0.140
	20	0.143
Water	0	0.0178
	10	0.0130
	20	0.0101
Machine oil (depending on the brand; approx.)	0	7.34
	20	3.82

In order to study the behaviour of a real ship in the ocean or of a wing in air flow etc. by means of experiments, engineers normally use a small model of the ship dipped into a wave canal or a small model of the wing brought into a wind tunnel, respectively. Let us ask for conditions under which the results of the experiments are in accordance with the reality. To answer this question it is advisable to use dimensionless quantities.

Let us treat for convenience the case of an incompressible flow:

Assume $\mathbf{u}_0 = (u_0, 0, 0)^T$ to be the flow velocity in a great distance of the solid, i.e. of the hull of the ship, of the wing etc., let L be a characteristic length of the solid, e.g. the length of the ship, the length of the wing's airfoil etc., let T be the time the fluid needs to cover the stretch of length L with velocity \mathbf{u}_0 , and let $P = \frac{\rho}{2} u_0^2$ be the dynamic pressure in a great distance of the solid.

Let us consider the particular case where exterior forces per unit volume can be neglected and let us now introduce the dimensionless quantities

$$\tilde{\mathbf{u}} = \frac{1}{u_0} \mathbf{u}, \quad \tilde{t} = \frac{t}{T}, \quad \tilde{p} = \frac{p}{2P}, \quad \tilde{x} = \frac{x}{L}, \quad \tilde{y} = \frac{y}{L}, \quad \tilde{z} = \frac{z}{L}.$$

The incompressible Navier-Stokes equations then turn out to become

$$\partial_{\tilde{t}} \tilde{\mathbf{u}} + \langle \tilde{\mathbf{u}}, \tilde{\nabla} \rangle \tilde{\mathbf{u}} + \tilde{\nabla} \tilde{p} - \frac{1}{\text{Re}} \tilde{\Delta} \tilde{\mathbf{u}} = \mathbf{0}, \quad (5.11)$$

where

$$\text{Re} := \frac{\rho u_0 L}{\eta} = \frac{u_0 L}{\nu} \quad (5.12)$$

is the so-called **Reynolds number** and

$$\tilde{\nabla} := (\partial_{\tilde{x}}, \partial_{\tilde{y}}, \partial_{\tilde{z}})^T, \quad \tilde{\Delta} = \partial_{\tilde{x}\tilde{x}} + \partial_{\tilde{y}\tilde{y}} + \partial_{\tilde{z}\tilde{z}}.$$

Thus, the experiment coincides with real world behaviour of the solid put into the flow if the Reynolds numbers of both situations coincide. This leaves a great freedom with respect to the particular choices of u_0 , L , ν realized in the experiment. The flows are then called **similar**.

In case of $\eta \neq 0$, the Laplace operator occurs in (5.10) so that this system of equations now becomes a parabolic system if instationary flows are concerned or an elliptic system if higher dimensional stationary flow is under consideration, respectively.

The continuity equation as well as the energy conservation law already known from the model of ideal fluid flows remain unaffected by the viscosity and remain valid. This also holds for the additional equation of state as far as gas flows are concerned.

There are only a few situations where the Euler equations or the Navier-Stokes equations can explicitly be solved by wellknown functions. In all the other situations one has to solve them approximately by means of numerical procedures.

Two of the problems which can be treated explicitly will be studied in the next section, and both problems are examples of stationary flows.

Remark:

The Navier-Stokes equations do not take the particular additional buoyancy phenomena into account arising in compressible gas flows from density variations caused by temperature differences between different parts of the gas. In order to respect these phenomenons, the Navier-Stokes equations have to be enriched by additional terms. If the enriched equations are also written in a dimensionless form, further dimensionless parameters besides the Reynolds number occur, particularly the Prandtl number

$$Pr := \frac{\nu}{\lambda} \rho g c_p$$

with the specific heat conductivity λ .

5.2 Drag Force and the Hagen-Poiseuille Law

Viscous flows can be **laminar** or **turbulent**. It is laminar if neighbouring fluid particles move more or less parallel to each other along stable trajectories, whereas turbulent flows are characterized by a disordered flow that superimposes the main flow. Thus, the turbulent flow is an instability phenomenon. When Reynolds studied the flow of liquids through small glass tubes, he noticed that the transition from the laminar to the turbulent state of the flow takes place in a very abrupt way depending on the parameter that is now called the Reynolds number (cf. (5.11)).

The transition to instabilities will briefly be discussed in the fifth section of this chapter. In this section the flows are assumed to be laminar.

Our first simple application of the Navier-Stokes equations concerns the laminar viscous stationary flow in y -direction of small velocity parallel to a given area δF of a plain solid plate where this area is part of the (y, z) -plane. Let the fluid be incompressible in the sense of $\rho = \text{const}$. Because of the friction between the fluid particles moving along and the fluid particles sticking to the rigid body in accordance with the no-slip condition, there is a force acting on δF . Obviously, this force has only one component in y -direction, and the velocity is independent of the space variables y and z :

$$\mathbf{u} = (0, u(x), 0)^T \quad \text{with} \quad u(0) = 0 \quad (\text{no-slip condition}).$$

The undisturbed velocity of the fluid that passes the surface area at a great distance is assumed to be constant, i.e. $\partial_x u(x) \rightarrow 0$ for $x \rightarrow \infty$.

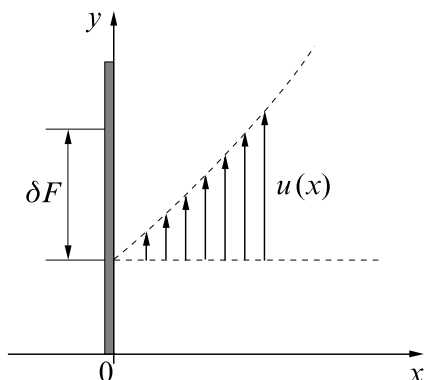


Figure 23: Flow along a plain membrane

Compared with (5.10), the forces per mass unit occurring within the Euler equations are now enriched by the additional term $\nu \Delta \mathbf{u}$ so that the forces per volume unit are now modelled by

$$\mathbf{w} = \begin{pmatrix} 0 \\ \eta \partial_{xx} u(x) \\ 0 \end{pmatrix}. \quad (5.13)$$

The resistance of the relevant area acting against the flow in a direction parallel to itself and arising from the friction or –vice versa– the **drag force** acting on the plate by the friction of the fluid comes therefore up to be

$$\delta F \int_0^\infty \mathbf{w} dx = \begin{pmatrix} 0 \\ -\eta u_x(0) \delta F \\ 0 \end{pmatrix}.$$

Hence, the drag related to the friction of a laminar flow is given by

$$W = \eta \partial_x u(0) \delta F. \quad (5.14)$$

Coming back to Reynolds' experiments, let us study as an example the one-dimensional laminar stationary cylinder-symmetric flow of an incompressible liquid through a circle-cylindrical pipe of radius R where the gravity force is neglected, where other outer forces do not occur and where the no-slip condition $\mathbf{u} = \mathbf{0}$ holds at the interior surface of the pipe (cf. Fig. 24).

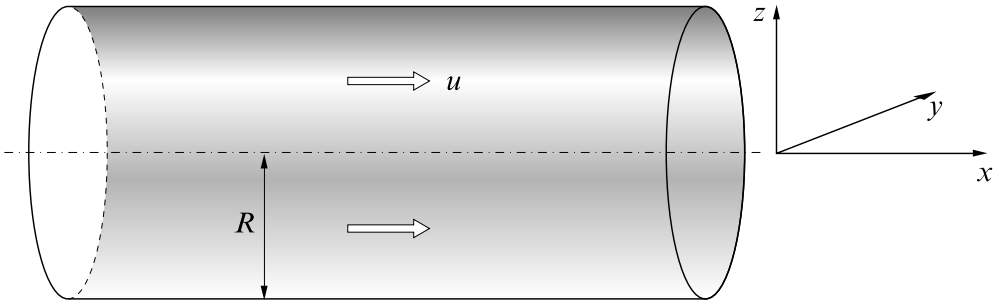


Figure 24: Laminar pipe flow

Because of $\mathbf{u} = (u, 0, 0)^T$ the relations

$$\langle \mathbf{u}, \nabla \rangle \mathbf{u} = (u \partial_x u, 0, 0)^T$$

and

$$\Delta \mathbf{u} = (\Delta u, 0, 0)^T, \quad \operatorname{div} \mathbf{u} = \partial_x u$$

follow.

Taking the continuity equation $\operatorname{div} \mathbf{u} = 0$ into account, we find $\partial_x u \equiv 0$, thus also $\partial_{xx} u \equiv 0$. Hence, (5.8) reduces to

$$-\frac{\eta}{\rho} \begin{pmatrix} \Delta u \\ 0 \\ 0 \end{pmatrix} + \frac{1}{\rho} \begin{pmatrix} \partial_x p \\ \partial_y p \\ \partial_z p \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \quad (5.15)$$

This leads to

$$\partial_y p \equiv \partial_z p \equiv 0, \quad \text{i.e. } p = p(x),$$

and therefore to

$$\Delta u = \partial_{yy} u + \partial_{zz} u = \frac{\partial_x p}{\eta}. \quad (5.16)$$

The left side of (5.16) is represented by a function that depends only on the variables y and z , whereas the right side depends only on x . So, $\partial_x p$ has to be constant.

The particular case of a flow considered here can therefore only occur if p depends linearly on x .

If polar coordinates (r, φ) are introduced in the (y, z) -plane, and if the cylinder symmetry is respected, i.e. $u = u(r)$ independent of φ , the well-known transformation rule

$$\Delta u = \frac{1}{r} \partial_r (r \partial_r u) + \frac{1}{r^2} \partial_{\varphi\varphi} u$$

transforms (5.16) into

$$\partial_r (r \partial_r u) = \frac{\partial_x p}{\eta} r.$$

Because p is independent of (y, z) , hence independent of (r, φ) , integration leads to

$$r \partial_r u = \frac{\partial_x p}{\eta} \frac{r^2}{2} + c_1 \quad (5.17)$$

with an integration constant c_1 .

Besides the no-slip condition $u(R) = 0$, also the boundary condition

$$|(\partial_r u)_{r=0}| < \infty$$

has to be fulfilled so that c_1 has to vanish.

If (5.17) is now integrated,

$$u(r) = \frac{\partial_x p}{2\eta} \frac{r^2}{2} + c_2$$

results where the constant c_2 has to be chosen in such a way that the no-slip condition is fulfilled. This finally yields the so-called **Hagen-Poiseuille law**

$$u(r) = -\frac{\partial_x p}{4\eta} (R^2 - r^2). \quad (5.18)$$

For every fixed x , this law shows a parabolic velocity distribution, and the flow follows the direction of decreasing pressure (cf. Fig 25).

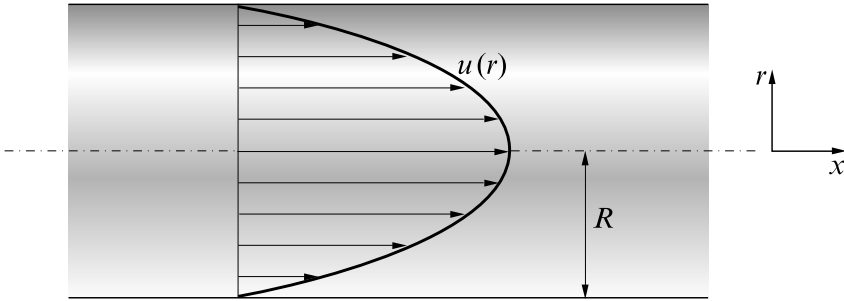


Figure 25: The Hagen-Poiseuille flow through a pipe

The mass that flows through the cross-section Q of the tube per second can now immediately be determined, namely

$$M = \int_Q \rho u dQ = - \int_0^{2\pi} \int_0^R \frac{\rho \partial_x p}{4\eta} (R^2 - r^2) r dr d\varphi = -\frac{\pi \rho \partial_x p}{8\eta} R^4, \quad (5.19)$$

and is proportional to the 4th power of the radius R though the cross-section only grows proportionally to its 2nd power.

Remark: This mass flow was discovered by the prussian civil service hydraulic engineer Hagen⁷¹ by experiments in 1839 without knowledge of (5.18), and independently

⁷¹Gotthilf Heinrich Ludwig Hagen (1797 - 1884); Berlin

also by the french physiologist Poiseuille⁷² in 1841 when he was interested in blood flow. It was also Poiseuille who found formula (5.19). The parabolic velocity distribution was for the first time mentioned in 1845 by Stokes.

The experiments of Hagen and of Poiseuille were realized for small tube diameters (0.015 - 3.00 mm) and confirmed formula (5.19), thus also the law (5.18) now named after them. Reynolds used glas tubes of higher diameters up to 10 mm, he also filled some dyes into the liquids he worked with, and he found out that there is a critical value of the parameter which is now called the Reynolds number, and this **critical Reynolds number** is characteristic for the transition of the flow from laminar to turbulent. He found $Re_{cr} = 1160$, and he recognized that the Hagen-Poiseuille law loses its validity in regions of turbulence: The mean flow velocity decreases because of a higher flow resistance.

The drag force of the laminar Hagen-Poiseuille flow in a part of the cylindric pipe of length L can be computed by means of (5.14) and of (5.18) in the following way:

$$W = \eta 2\pi RL (\partial_r u)_{r=R} = \eta 2\pi RL \frac{\partial_x p}{2\eta} R = \pi R^2 \Delta p, \quad (5.20)$$

where Δp denotes the pressure difference between the two ends of the particular part of the pipe under consideration.

There are also some other cross sections of ducts for which the Navier-Stokes equations can be solved analytically under the same assumptions on the flow as in the Hagen-Poiseuille case (steady, laminar, incompressible, viscous), e.g. for annuli, equilateral triangles etc.⁷³

Situations similar to the flow of a liquid through narrow ducts occur for other flows through narrow gaps or slits, e.g. investigation of the flow of lubricating oils in lubricating research called **tribology**.

As already mentioned, there are only very few examples where the Navier-Stokes equations can be solved exactly. This led to efforts to simplify these equations whenever this seemed to be suitable. Some of the important developments concerning simplifications in case of small and of great Reynolds numbers will briefly be discussed in the next sections. Also for great Reynolds numbers we do not go back to the Euler equations but preserve some of the structures of the Navier-Stokes equations and study other suitable simplifications, and the no-slip condition will also be kept in case of great Reynolds numbers. The case of small Reynolds numbers was particularly investigated by Stokes and led to a model equation nowadays called the **Stokes**

⁷²Jean Louis Marie Poiseuille (1799 - 1869); Paris

⁷³C.Y. Wang: Ann.Rev.Fluid Mech., 159-177 (1991)

approximation (cf. (5.25)), whereas theory as well as practical applications of fluid dynamics for small Reynolds numbers was significantly pressed ahead by Prandtl⁷⁴ in his **boundary layer theory** and his **wing theory**.

5.3 Stokes Approximation and Artificial Time

The difficulties in solving stationary or instationary Navier-Stokes problems or Euler problems are mainly due to the nonlinearity of the convection term $\langle \mathbf{u}, \nabla \rangle \mathbf{u}$.

Another difficulty particularly connected with the case of an incompressible instationary flow is the fact that the pressure p one has to determine besides the velocity \mathbf{u} is not represented in the continuity equation by its time derivative, e.g. by an additional term $\kappa \partial_t p$ ($\kappa = \text{const}$). If this would be the case and if a suitable initial pressure p_0 would be prescribed, the instationary problem could be reformulated as a genuine initial boundary value problem whose differential equation could be written as

$$\partial_t \varphi - L[\varphi] = \mathbf{0} \quad \text{with} \quad \varphi = (u_1, u_2, u_3, p)^T \quad (5.21)$$

and could theoretically as well as numerically be treated by means of appropriate techniques.

Let us assume that the solutions of instationary Navier-Stokes problems converge for $t \rightarrow \infty$ in each case towards the solution of the correspondent stationary problem independent of the initial function \mathbf{u}_0 .

If this also holds for the instationary problem (5.21) consisting of the Navier-Stokes equations and of a continuity equation enriched by a physically unjustified additional time derivative of p , the solution of this manipulated problem must coincide with the stationary solution of the original Navier-Stokes problem for $t \rightarrow \infty$ because of

$$p_t \rightarrow 0 \quad \text{for} \quad t \rightarrow \infty.$$

This leads to the following idea of a method for solving the stationary Navier-Stokes problem:

Introduce an **artificial time** τ and enrich the stationary Navier-Stokes equations as well as the continuity equation by adding linear combinations of certain terms in such a way that after addition of a more or less arbitrary initial function a complete instationary initial or initial boundary value problem comes into being whose solution

⁷⁴Ludwig Prandtl (1875 - 1953); Hannover, Göttingen

is expected to have the property to annul the additional artificial terms for $\tau \rightarrow \infty$. In engineering, this concept is called **method of artificial compressibility** and is occasionally numerically realized.

Such an artificially enriched stationary problem might look as follows:

$$\left. \begin{aligned} -\partial_\tau \mathbf{u} &= \langle \mathbf{u}, \nabla \rangle \mathbf{u} + \nabla p - \frac{1}{Re} \Delta \mathbf{u} + \alpha \nabla \operatorname{div} \mathbf{u} \\ -\partial_\tau p &= \beta^2 \operatorname{div} \mathbf{u} \end{aligned} \right\} \text{ in } \Omega, \quad (5.22)$$

$$\mathbf{u} = \mathbf{u}_0 \quad \text{on } \partial\Omega,$$

$$\mathbf{u} = \mathbf{u}^{[0]}, \quad p = p_0 \quad \text{for } \tau = 0,$$

where $\mathbf{u}^{[0]}$ is assumed to coincide with \mathbf{u}_0 on $\partial\Omega$. $\alpha > 0$ and $\beta \neq 0$ are arbitrary constants. They should be chosen in such a way that the numerical treatment of the artificially enriched problem is going to become a method as effective as possible.

Of course, the expected property of the artificially added terms to vanish for increasing values of τ should be guaranteed by a mathematical proof.

From a numerical point of view, this strategy can lead to an iterative procedure in order to solve approximately the stationary boundary value problem in the following way:

Use the more or less arbitrarily added artificial initial function as initial approximation, i.e.

$$\varphi^{[0]} := \begin{pmatrix} \mathbf{u}^{[0]} \\ p_0 \end{pmatrix}. \quad (5.23)$$

Then discretize (5.22) with respect to the artificial time variable τ using in every step a suitable step size $\Delta_n \tau$, e.g. an equidistant step size $\Delta \tau$.

An example for the semi-discrete approximate equation coming into being is the explicit Euler method

$$\varphi^{[n+1]} = \varphi^{[n]} - \Delta_n \tau \cdot L[\varphi^{[n]}] \quad (n = 0, 1, 2, \dots) \quad (5.24)$$

formulated in the style of (5.21). Here, every $\mathbf{u}^{[n]}$ is expected to fulfill the boundary conditions given in (5.22). $\varphi^{[n]}$ will then not only be interpreted as an approximate solution of the artificial instationary problem on its n th artificial time level but also as an approximate solution of the stationary problem after the n th iteration step. If the iteration converges sufficiently fast, it will be stopped after a certain number of n_0

steps, and $\varphi^{[n_0]}(x, y, z)$ will then be considered to be the final approximate solution of the stationary problem.

At the beginning, this method was used and justified only heuristically by engineers who asked the mathematicians for an also mathematical justification. This justification could then at least partially be given by a proof of the fact that in case of the Stokes approximation the solution of (5.22) really converges for $\tau \rightarrow \infty$ to the solution of the stationary problem⁷⁵. Here, the truncation errors brought in by additional space discretizations are left out of account.

The Stokes approximation is developed from the Navier-Stokes equations by omission of the convective term. This seems to be reasonable in case of incompressible flows for small Reynolds numbers. Namely, if (5.11) is written as

$$Re \cdot \partial_{\tilde{\tau}} \tilde{\mathbf{u}} + Re \cdot \langle \tilde{\mathbf{u}}, \tilde{\nabla} \rangle \tilde{\mathbf{u}} + Re \cdot \nabla \tilde{p} - \tilde{\Delta} \tilde{\mathbf{u}} = \mathbf{0},$$

one can see that the linear diffusion term dominates the convection term.

If we omit the tildes and if we restrict ourselves to the stationary problem, after neglectation of the convection term only

$$Re \cdot \nabla p - \Delta \mathbf{u} = \mathbf{0} \quad (5.25)$$

remains. Finally, we replace $Re \cdot p$ by a new quantity p and enrich the system exemplarily in the way presented in (5.22). The system then turns out to take the form

$$\left. \begin{aligned} \partial_{\tau} \mathbf{u} + \nabla p &= \Delta \mathbf{u} + \alpha \nabla \operatorname{div} \mathbf{u} \\ \partial_{\tau} p + \beta^2 \operatorname{div} \mathbf{u} &= 0 \end{aligned} \right\} \text{ in } \Omega, \quad (5.26)$$

$$\mathbf{u} = \mathbf{0} \quad \text{on } \partial \Omega,$$

$$\mathbf{u} = \mathbf{u}^{[0]}, \quad p = p_0 \quad \text{for } \tau = 0,$$

where also the linear structure of (5.25) was taken into account. Because of this structure, it suffices to show the convergence of \mathbf{u} for increasing values of τ to the stationary solution in case of vanishing exterior forces and of homogeneous boundary conditions.

The velocity part of the stationary solution of (5.26) then obviously vanishes, i.e.

$$\mathbf{u} \equiv \mathbf{0},$$

and the following theorem holds:

⁷⁵St. Fellehner, Dissertation Hamburg 1995.

Theorem 5.1:

The solution \mathbf{u} of (5.26)⁷⁶ converges in the sense of the L_2 -norm for $\tau \rightarrow \infty$ to the zero function.⁷⁷

PROOF: Using the denotations

$$\begin{aligned} \mathbf{J} \mathbf{u} \cdot \mathbf{v} &= (\langle \nabla u_1, \mathbf{v} \rangle, \langle \nabla u_2, \mathbf{v} \rangle, \dots, \langle \nabla u_d, \mathbf{v} \rangle)^{\mathbf{T}}, \\ \langle \mathbf{J} \mathbf{u}, \mathbf{J} \mathbf{v} \rangle &= \sum_{i,j} \partial_{x_j} u_i \partial_{x_j} v_i, \\ (p, q) &= \int_{\Omega} p q \, dx, \quad \|p\|_{\Omega} = \sqrt{(p, p)} \quad \text{in } L_2(\Omega), \\ (\mathbf{u}, \mathbf{v}) &= \int_{\Omega} \langle \mathbf{u}, \mathbf{v} \rangle \, dx, \quad \|\mathbf{u}\|_{\Omega} = \sqrt{(\mathbf{u}, \mathbf{u})} \quad \text{in } L_2^d(\Omega), \\ (\mathbf{J} \mathbf{u}, \mathbf{J} \mathbf{v}) &= \int_{\Omega} \langle \mathbf{J} \mathbf{u}, \mathbf{J} \mathbf{v} \rangle \, dx, \quad \|\mathbf{J} \mathbf{u}\|_{\Omega} = \sqrt{(\mathbf{J} \mathbf{u}, \mathbf{J} \mathbf{u})} \quad \text{in } L_2^{d \times d}(\Omega), \end{aligned}$$

taking

$$\operatorname{div}(\mathbf{J} \mathbf{u} \cdot \mathbf{v}) = \langle \Delta \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{J} \mathbf{u}, \mathbf{J} \mathbf{v} \rangle,$$

into account or –by means of Gauss' divergence theorem– the relation

$$\langle \Delta \mathbf{u}, \mathbf{v} \rangle = \int_{\partial \Omega} \left\langle \frac{\partial \mathbf{u}}{\partial n}, \mathbf{v} \right\rangle \, d\sigma - (\mathbf{J} \mathbf{u}, \mathbf{J} \mathbf{v}), \quad (5.27)$$

and taking finally also into account that the integral along the boundary vanishes, (5.26) leads to

$$\begin{aligned} \frac{1}{2} \frac{d}{d\tau} \left(\|\mathbf{u}\|_{\Omega}^2 + \frac{1}{\beta^2} \|p\|_{\Omega}^2 \right) &= (\mathbf{u}, \mathbf{u}_{\tau}) + \frac{1}{\beta^2} (p, \partial_{\tau} p) \\ &= (\mathbf{u}, -\nabla p) + (\mathbf{u}, \Delta \mathbf{u}) + (\mathbf{u}, \alpha \nabla(\operatorname{div} \mathbf{u})) \\ &\quad + \frac{1}{\beta^2} (p, -\beta^2 \operatorname{div} \mathbf{u}) \quad (5.28) \\ &= -\|\mathbf{J} \mathbf{u}\|_{\Omega}^2 - \alpha \|\operatorname{div} \mathbf{u}\|_{\Omega}^2. \end{aligned}$$

⁷⁶For convenience, we omit considerations concerning the behaviour of p .

⁷⁷Here, it is assumed that there is a unique smooth solution of the artificial instationary problem.

Thus,

$$\sigma(\tau) := \left(\|\mathbf{u}\|_{\Omega}^2 + \frac{1}{\beta^2} \|p\|_{\Omega}^2 \right)$$

decreases monotonously, and this the stronger the greater α ⁷⁸. Hence, $\|\mathbf{u}\|_{\Omega}$ and $\|p\|_{\Omega}$ are bounded so that

$$\sigma_{\infty} := \lim_{\tau \rightarrow \infty} \sigma(\tau)$$

exists.

The Poincaré inequality⁷⁹ then yields in $L_2^d(\Omega)$

$$\|\mathbf{u}\|_{\Omega}^2 \leq c_1 \|\mathbf{J} \mathbf{u}\|_{\Omega}^2$$

with a constant $c_1 > 0$ only depending on the domain Ω . Equation (5.27) therefore leads to

$$\sigma'(\tau) = \frac{d}{d\tau} \left(\|\mathbf{u}\|_{\Omega}^2 + \frac{1}{\beta^2} \|p\|_{\Omega}^2 \right) \leq -c \|\mathbf{u}\|_{\Omega}^2 \quad (5.29)$$

where $c = \frac{2}{c_1} > 0$.

Formula (5.29) implies

$$\sigma(\tau_0) - \sigma_{\infty} \geq c \int_{\tau_0}^{\infty} \|\mathbf{u}\|_{\Omega}^2 d\tau \quad (5.30)$$

so that particularly the integral on the righthand side exists, i.e. $\|\mathbf{u}\|_{\Omega} \in L_2(\mathbb{R}^+)$.

But also

$$\frac{d}{d\tau} \left(\|\mathbf{u}\|_{\Omega}^2 \right)$$

is bounded as can be seen by the following arguments:

Differentiate the differential equation in (5.26) with respect to τ which leads in a first step analogously to (5.29) to

$$\frac{d}{d\tau} \left(\|\partial_{\tau} \mathbf{u}\|_{\Omega}^2 + \frac{1}{\beta^2} \|\partial_{\tau} p\|_{\Omega}^2 \right) \leq -c \|\partial_{\tau} \mathbf{u}\|_{\Omega}^2, \quad (5.31)$$

⁷⁸This can also be used for numerical purposes.

⁷⁹cf. e.g. Dautrey-Lions: *Mathematical Analysis and Numerical Methods for Science and Technology*, vol. 2, p. 126

hence to the boundedness of $\|\partial_\tau \mathbf{u}\|$. Because of

$$\left| \frac{d}{d\tau} \left(\|\mathbf{u}\|_\Omega^2 \right) \right| = 2 |(\mathbf{u}, \partial_\tau \mathbf{u})| \leq 2 \|\mathbf{u}\|_\Omega \|\partial_\tau \mathbf{u}\|_\Omega,$$

the existence of a constant ω with

$$\left| \frac{d}{d\tau} \left(\|\mathbf{u}\|_\Omega^2 \right) \right| \leq \omega \tag{5.32}$$

follows. The inequalities (5.32) and (5.30) then imply

$$\lim_{\tau \rightarrow \infty} \|\mathbf{u}(\tau)\|_\Omega = 0.$$

Otherwise there would exist an $\epsilon > 0$ and a number $\tau_i > \tilde{\tau}$ for every $\tilde{\tau} > 0$ with $\|\mathbf{u}(\tau_i)\|_\Omega^2 \geq \epsilon$, and by means of the inequality

$$-\omega(\tau - \tau_i) \leq \|\mathbf{u}(\tau)\|_\Omega^2 - \|\mathbf{u}(\tau_i)\|_\Omega^2 \leq \omega(\tau - \tau_i),$$

which follows for every $\tau > \tau_i$ from (5.32),

$$\|\mathbf{u}(\tau)\|_\Omega^2 \geq \epsilon - \omega(\tau - \tau_i)$$

would follow. Thus, also

$$J_i := \int_{\tau_i}^{\tau_i + \frac{\epsilon}{\omega}} \|\mathbf{u}(\tau)\|_\Omega^2 d\tau \geq \frac{\epsilon^2}{\omega} - \frac{\omega}{2} \frac{\epsilon^2}{\omega^2} = \frac{\epsilon^2}{2\omega}$$

would result.

Now, after the choice of a suitable subsequence of the sequence $\{J_i\}$ the intervals $\tau_i \leq \tau \leq \tau_i + \frac{\epsilon}{\omega}$ ($i = 1, 2, \dots$) could be assumed to be disjoint and would lead to

$$\sum_i J_i \leq \int_{\tau_1}^{\infty} \|\mathbf{u}(\tau)\|_\Omega^2 d\tau < \infty.$$

This finally would contradict the unboundedness of the partial sums of $\sum_i J_i$. \square

5.4 Foundations of the Boundary Layer Theory; Flow Separation

Let us now look at the other extreme situation of the Navier-Stokes equations written in the dimensionless form (5.11), namely at the case of great Reynolds numbers particularly studied by Prandtl. Here, we omit again the tildes used in (5.11), and because the Reynolds numbers are great, the behaviour of the solutions can be expected not to differ too much from the solutions of the Euler equations⁸⁰

Prandtl published his results in 1904 on the International Congress of Mathematicians in Heidelberg⁸¹. Let us now follow his ideas.

Assume a solid to be put into a viscous flow, let the no-slip condition (1.54) hold along the wall Γ of this solid, i.e. $\mathbf{u}_\Gamma = \mathbf{0}$, omit the influences of exterior forces, and assume the flow to have been irrotational before the solid was put into it.

Thus, the velocity \mathbf{u} could at this time be derived from a potential Φ which fulfills the potential equation $\Delta \Phi = 0$ because of the incompressibility of the fluid. This leads to the fact that there was no difference compared to the frictionless situation.

Prandtl restricted his considerations to a piece Γ of the surface of the solid which can be assumed to be (more or less) plane⁸², and this plane was then used as (x, y) -plane in an (x, y, z) -coordinate system. In a great distance of the plane, theoretically for $z \rightarrow \infty$, Prandtl expected the flow to keep its character as a potential flow so that the Euler equations hold for large values of z .

The idea that the flow could perhaps be irrotational everywhere outside the solid leads for $\mathbf{u} \neq \mathbf{0}$ immediately to a contradiction because the problem

$$\begin{aligned}\Delta \Phi &= 0 && \text{outside the solid,} \\ \mathbf{u}_\Gamma &= \mathbf{0} && \text{no-slip condition}\end{aligned}$$

has only the solution $\mathbf{u} \equiv \mathbf{0}$. We omit the proof and take into account that a jump from the velocity $\mathbf{u} = \mathbf{0}$ along Γ to a potential flow of a velocity $\mathbf{u} \neq \mathbf{0}$ on the

⁸⁰It took many years to make mathematically sure that the Navier-Stokes solutions are close to the Euler solutions for small values of η : K. Nickel showed in *Arch.Rat.Mech.Anal.* **13**, 1-14 (1963), that both solutions can under certain conditions and for $\eta \rightarrow 0$ asymptotically be looked upon as solutions of the Prandtl boundary layer equations (5.48) presented in this section. This also justified subsequently Prandtl's assumption on the small diameter of the boundary layer we are going to use later.

⁸¹Verh. d. III. Int.Math.-Kongr. Heidelberg 1904, S. 484-494, Leipzig: Teubner 1905

⁸²W. Tollmien showed that all these considerations also hold for curved surfaces provided that the curvatures do not vary too much. Cf. Handb. d. Exper.-Physik, **IV**, part I, 248 ff. (1938).

outside of the solid is impossible, too, because the existence of the viscosity term does not allow discontinuities of this type.

Hence, the flow between the surface of the solid and the distant parts of the flow can't be irrotational. The domain of this rotational flow is called the **boundary layer**. Here, it is more or less open what the word *distant* means in this context from the point of view of real applications. In other words: What is the thickness δ of the boundary layer, i.e. from which distance on can the difference between the real flow and the original potential flow be neglected? Of course, the answer depends on a convention. Often, 1% difference between the velocities is an accepted value.

Prandtl assumed the thickness δ to be small compared with the length L used in order to write the Navier-Stokes equations in the dimensionless form (5.11), i.e.

$$\delta \ll 1, \quad (5.33)$$

and he assumed that the transition from the rotational flow to the irrotational potential flow behaves continuously.

Let all the derivatives occurring in the following considerations exist, let them be continuous, and let the dimensionless velocity \mathbf{U} of the distant potential flow be parallel to Γ , i.e. parallel to the (x, y) -plane. Hence,

$$\mathbf{U} = \mathbf{U}(x, y, t) = (U(x, y, t), V(x, y, t), 0)^T.$$

Moreover, the components of \mathbf{U} are assumed to be of the same order of magnitude as the velocity u_0 used for the derivation of the dimensionless form (5.11) of the Navier-Stokes equations so that U and V are of the order 1 of magnitude. Because of the continuous transition from $\mathbf{u} = (u, v, w)^T$ to \mathbf{U} , the assumptions

$$u = \mathcal{O}(1) \quad \text{and} \quad v = \mathcal{O}(1) \quad \text{for} \quad z \rightarrow \delta \quad (5.34)$$

with the Landau symbol \mathcal{O} are reasonable, and this will certainly also be true along all the layers parallel to Γ , particularly for $x, y \rightarrow 1$ where x and y are the dimensionless spatial variables in (5.11).

Concerning the spatial derivatives of the u - and v -components parallel to Γ as well as their time derivatives, Prandtl expected them not to vary very much within the boundary layer, i.e.

$$\partial_t u, \partial_t v, \partial_x u, \partial_y u, \partial_{xx} u, \partial_{yy} u, \partial_x v, \partial_y v, \partial_{xx} v, \partial_{yy} v = \mathcal{O}(1). \quad (5.35)$$

If additionally the continuity equation $\text{div } \mathbf{u} = 0$ is taken into account, (5.35) yields a further result, namely

$$\partial_z w = \mathcal{O}(1) \quad \text{for} \quad z \rightarrow \delta \quad \text{or for} \quad x, y \rightarrow 1, \quad (5.36)$$

so that the no-slip condition leads to

$$|w(x, y, z, t)| = \left| \int_0^z \partial_z w(x, y, \zeta, t) d\zeta \right| \leq \|\partial_z w\|_\infty |z|,$$

i.e.

$$w = \mathcal{O}(\delta). \quad (5.37)$$

Combining assumption (5.33) with the no-slip condition, the mean value theorem leads inside the boundary layer to

$$u(x, y, z, t) \approx \partial_z u(x, y, z, t) \cdot z,$$

and therefore with (5.34) to

$$\partial_z u = \mathcal{O}\left(\frac{1}{\delta}\right). \quad (5.38)$$

Analogously also the relations

$$\partial_z v = \mathcal{O}\left(\frac{1}{\delta}\right), \quad \partial_{zz} u = \mathcal{O}\left(\frac{1}{\delta^2}\right), \quad \partial_{zz} v = \mathcal{O}\left(\frac{1}{\delta^2}\right) \quad (5.39)$$

hold.

We also obtain by the mean value theorem the statement

$$w(x, y, z, t) = w(0, y, z, t) + \partial_x w(\vartheta x, y, z, t) \cdot x, \quad 0 < \vartheta < 1,$$

and because x can grow up to the order 1, (5.37) yields

$$\partial_x w = \mathcal{O}(\delta) \quad (5.40)$$

and analogously

$$\partial_y w, \partial_{xx} w, \partial_{yy} w = \mathcal{O}(\delta), \quad \partial_t w = \mathcal{O}(\delta). \quad (5.41)$$

Moreover, taking (5.36) into account, we additionally find

$$\partial_{zz} w = \mathcal{O}\left(\frac{1}{\delta}\right). \quad (5.42)$$

Let us now write down separately the three components of (5.11) where we are going to note down below each of the terms occurring in the equations their particular orders of magnitude as far as they are already known. Here we follow a representation given by H. Schlichting in the first edition of his book on boundary-layer theory:⁸³

$$-\nabla p =$$

$$\left\{ \begin{array}{l} \partial_t u + u \partial_x u + v \partial_y u + w \partial_z u - \frac{1}{\text{Re}} (\partial_{xx} u + \partial_{yy} u + \partial_{zz} u) \\ 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad \delta \quad \frac{1}{\delta} \quad \frac{1}{\text{Re}} \quad 1 \quad 1 \quad \frac{1}{\delta^2} \\ \partial_t v + u \partial_x v + v \partial_y v + w \partial_z v - \frac{1}{\text{Re}} (\partial_{xx} v + \partial_{yy} v + \partial_{zz} v) \\ 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad \delta \quad \frac{1}{\delta} \quad \frac{1}{\text{Re}} \quad 1 \quad 1 \quad \frac{1}{\delta^2} \\ \partial_t w + u \partial_x w + v \partial_y w + w \partial_z w - \frac{1}{\text{Re}} (\partial_{xx} w + \partial_{yy} w + \partial_{zz} w) \\ \delta \quad 1 \quad \delta \quad 1 \quad \delta \quad \delta \quad 1 \quad \frac{1}{\text{Re}} \quad \delta \quad \delta \quad \frac{1}{\delta} \end{array} \right. \quad (5.43)$$

Of course, the pressure gradient parallel to the plane is assumed to be bounded in all real applications, i.e.

$$\partial_x p, \partial_y p = \mathcal{O}(1).$$

But this implies that the first two equations can only be valid with respect to the order of magnitude if

$$\text{Re} = \mathcal{O}\left(\frac{1}{\delta^2}\right), \quad \text{i.e.} \quad \delta = \mathcal{O}\left(\frac{1}{\sqrt{\text{Re}}}\right), \quad (5.44)$$

and this leads because of the third equation to

$$\partial_z p = \mathcal{O}(\delta). \quad (5.45)$$

The equation (5.44) shows that Prandtl's assumption on a small thickness of the boundary layer for great Reynolds numbers does not lead to a contradiction.

It also shows that the boundary layer theory can *only* be accepted as an approximation for the Navier-Stokes equations in the case of great Reynolds numbers.

⁸³H. Schlichting: Grenzschichttheorie; Karlsruhe, Braun 1951

Because of

$$|p(x, y, z, t) - p(x, y, 0, t)| = \left| \int_0^z \partial_z p(x, y, \zeta, t) d\zeta \right| \leq |z| \|\partial_z p\|_\infty,$$

the pressure differences in z -direction within the boundary layer are of order δ^2 , hence very small for great Reynolds numbers. We therefore put

$$\partial_z p \equiv 0,$$

so that p does approximately not depend on z within the boundary layer:

$$p = p(x, y, t). \quad (5.46)$$

The pressure gradient in the boundary layer thus coincides with the well known pressure gradient of the potential flow outside the boundary layer where the Euler equations hold. Using dimensionless variables and taking $W = 0$ into account, we therefore obtain approximately

$$\partial_t U + U \partial_x U + V \partial_y U = -\partial_x p$$

$$\partial_t V + U \partial_x V + V \partial_y V = -\partial_y p$$

so that $\partial_x p$ and $\partial_y p$ are known quantities.

If now all the terms in (5.43) of order δ will be neglected compared with the terms of order 1, the system of three equations reduces to only two equations, and these equations read in dimensionless form as

$$\partial_t u + u \partial_x u + v \partial_y u + w \partial_z u - \nu \partial_{zz} u = -\frac{1}{\rho} \partial_x p$$

$$\partial_t v + u \partial_x v + v \partial_y v + w \partial_z v - \nu \partial_{zz} v = -\frac{1}{\rho} \partial_y p$$

(5.48) are the **Prandtl Boundary Layer Equations** added by the incompressibility condition

$$\partial_x u + \partial_y v + \partial_z w = 0. \quad (5.49)$$

The velocity component w still remains in the system (5.48) but does now only play the role of a parameter which will later be fixed by (5.49) and by boundary conditions. An example of such boundary conditions reads as

$$u = v = w = 0 \quad \text{for } z = 0$$

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} U(x, y, t) \\ V(x, y, t) \\ 0 \end{pmatrix} \quad \text{for } z \rightarrow \infty.$$

$z \rightarrow \infty$ and $\delta \ll L$ seem to be contradictory but one has to take into account that the boundary layer is of an infinite thickness from a theoretical point of view, and only their effective thickness judged from a physical point of view is very small.

If the boundary layer is affected by sucking off or blowing up parts of it, the boundary conditions have to model these facts, e.g. by replacing the equation $w = 0$ in (5.50) by

$$w = w_0(x, y, t) \quad \text{for } z = 0,$$

provided that the direction of the sucking or blowing coincides with the w -direction and that its flow velocity w_0 is of order $O(\delta)$, i.e.

$$w_0 = \mathcal{O}\left(\frac{1}{\sqrt{\text{Re}}}\right).$$

Otherwise some of the terms expected by Prandtl to be very small can no longer be neglected.

Besides boundary conditions also initial conditions have to be formulated, and this not only with respect to time in case of instationary problems but also in every case with respect to space because the Boundary Layer Equations are of parabolic character. This parabolicity often includes simplifications as far as the construction of approximate analytic solutions of the Navier-Stokes equations is desired. But also with respect to numerical purposes Prandtl's theory can often make life easier.

If the problem to be treated concerns an instationary flow, the task to begin with is the formulation of initial conditions with respect to time, i.e. conditions of the type

$$\begin{aligned} u(x, y, z, 0) &= f_1(x, y, z) \\ v(x, y, z, 0) &= f_2(x, y, z) \end{aligned} \quad ,$$

and in every case also with respect to space, i.e. the stipulation of suitable initial profiles of the u - and v -components of the flow velocity.

There is no general consensus how the initial profiles should look like, but it should be noticed that the influence of the initial profiles decreases for increasing distance from the area where these profiles were prescribed. This follows from the parabolic situation.

These spatial initial conditions may take into consideration that a boundary layer comes into being only after a certain distance from the edges of the solid put into the fluid, e.g. from the front edge of a wing:

$$\begin{aligned}u(x_0, y, z, t) &= g_{11}(y, z, t) \\v(x_0, y, z, t) &= g_{12}(y, z, t)\end{aligned}$$

and

$$\begin{aligned}u(x, y_0, z, t) &= g_{21}(x, z, t) \\v(x, y_0, z, t) &= g_{22}(x, z, t).\end{aligned}$$

Of course, the conditions prescribed should go well together, e.g.

$$f_1(x_0, y, z) = g_{11}(y, z, 0) \tag{5.54}$$

etc.

Many phenomena of flows taking place along the surface of a solid can quantitatively be explained by means of the boundary layer concept. One of them is the separation of the boundary layer flow from a curved surface.

Here, we are going to describe this separation briefly and only qualitatively in case of stationary 2-D flows: Assume an ideal fluid to flow along a curved surface. A lateral flow with main flow velocity $\mathbf{u} = (u_\infty, 0, 0)^T$, $u_\infty > 0$, around a circular cylinder of infinite length whose axis coincides with the z -axis is a 2-D example. The velocity of particles arriving at the surface of the cylinder increases due to Bernoulli's equation (1.31) because they move at first against decreasing pressure. The velocity then begins to decrease because the pressure grows again. Finally the particles move behind the cylinder with their original speed (cf. Fig. 26).

In the case of a viscid flow, a certain amount of the kinetic energy of the particles will be annihilated in the first phase by friction so that there is in the second phase not enough energy available to bring the particles back to their former speed. It can even

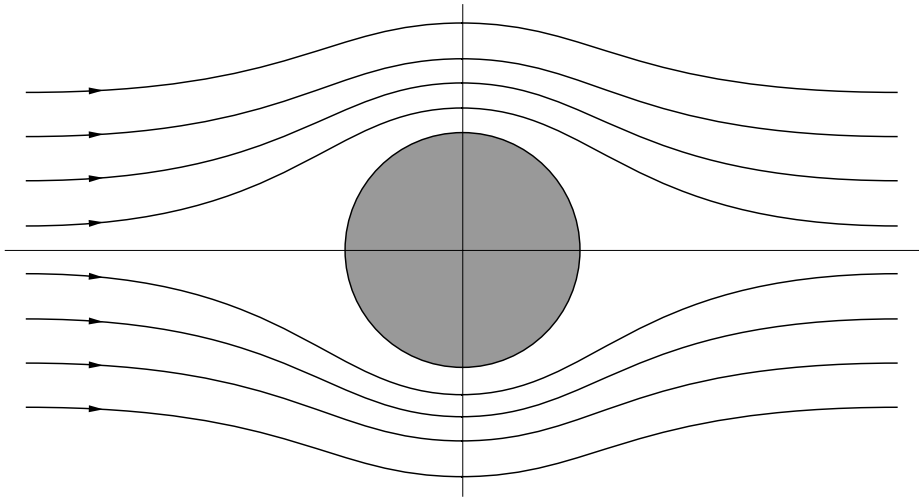


Figure 26: Ideal flow around a cylinder

happen that the particles stop and that the velocity reverses its direction because of the increasing pressure, i.e. separation of the boundary layer takes place.

In the 2-D situation, the point of separation at the curved surface which is in two dimensions a curve Γ can be well described mathematically. For this purpose, let u

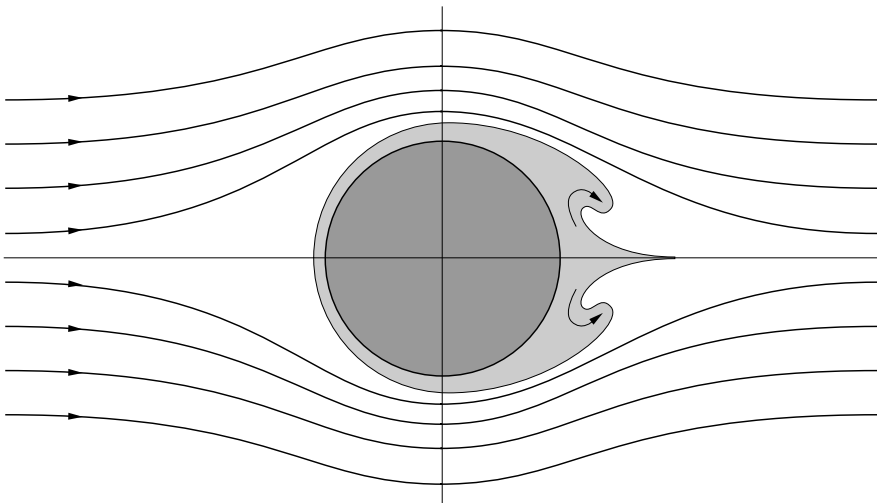


Figure 27: Flow with separation

be the particular component of the flow velocity parallel to Γ and let $\partial_n u$ be the normal derivative of this component. The separation point $P \in \Gamma$ is then obviously given as the point where this normal derivative vanishes (cf. Fig 28):

$$\partial_n u(P) = 0. \quad (5.55)$$

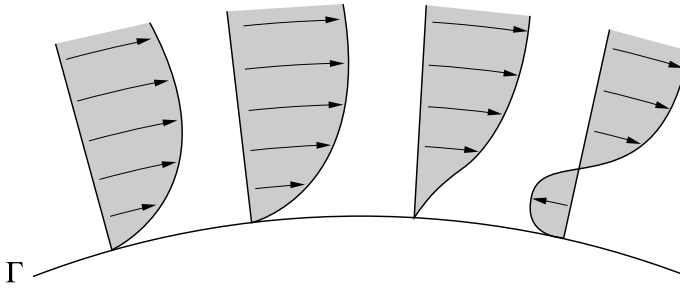


Figure 28: Explanation of (5.55) in case of a 2-D boundary layer flow

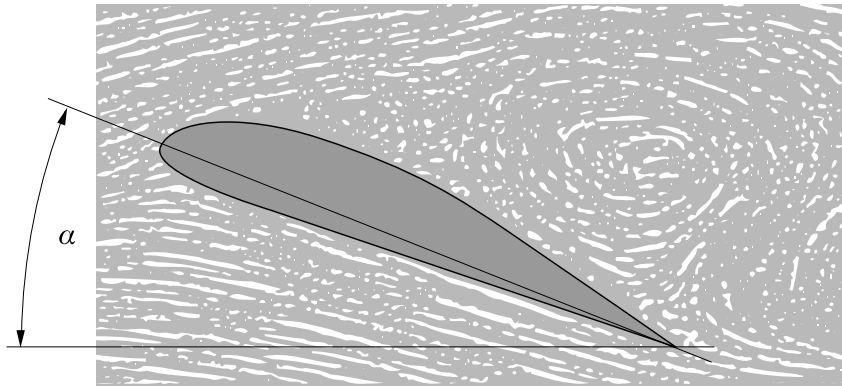


Figure 29: Boundary layer separation on a wing

Beyond the point of separation, the flow loses its buoyancy effect. Hence, one tries to design a wing in such a way that the point of separation is situated as near as possible at the back edge of the wing, e.g. by sucking off the boundary layer.

It should be mentioned that there is no general mathematical definition of separation in case of 3-D flows.

5.5 Stability of Laminar Flows

The beginning of the transition of a flow from the laminar state to a turbulent one can be looked upon as a transition from a stable behaviour of the solution of the Navier-Stokes equations belonging to this flow to instability.

In order to approach this idea, let us for convenience consider a situation easy to survey, namely an incompressible 2-D flow and the knowledge of a stationary average laminar flow in x -direction whose velocity U only depends on the second space variable y :

$$\mathbf{U} = \begin{pmatrix} U(y) \\ 0 \end{pmatrix}.$$

The pressure of this flow is denoted by $P(x, y)$, and (\mathbf{U}, P) is assumed to fulfill the given incompressible Navier-Stokes problem⁸⁴ and to be one part of a superposition whose other part consists of infinitesimal incompressible disturbances $(\tilde{\mathbf{u}}, \tilde{p})$ with

$$\tilde{\mathbf{u}} = \begin{pmatrix} \tilde{u}(x, y, t) \\ \tilde{v}(x, y, t) \end{pmatrix} \quad \text{and} \quad \tilde{p} = \tilde{p}(x, z, t).$$

This second part is not necessarily a stationary flow and, besides (\mathbf{U}, P) , also the complete flow

$$\mathbf{u} = \mathbf{U} + \tilde{\mathbf{u}}, \quad p = P + \tilde{p} \quad (5.56)$$

is assumed to be a solution of the given Navier-Stokes problem⁸⁵.

Because the disturbances are infinitesimal, quadratic terms can be neglected compared with linear terms, so that putting (5.56) into the Navier-Stokes equations (5.10) and into the continuity equation then yields the relations

$$\begin{aligned} \partial_t \tilde{u} + U \partial_x \tilde{u} + \tilde{v} \frac{dU}{dy} + \frac{1}{\rho} \partial_x P + \frac{1}{\rho} \partial_x \tilde{p} &= \nu \left(\frac{d^2 U}{dy^2} + \Delta \tilde{u} \right) \\ \partial_t \tilde{v} + U \partial_x \tilde{v} + \frac{1}{\rho} \partial_y P + \frac{1}{\rho} \partial_y \tilde{p} &= \nu \Delta \tilde{v} \\ \partial_x \tilde{u} + \partial_y \tilde{v} &= 0. \end{aligned} \quad (5.57)$$

⁸⁴It is absolutely problematic to suppose such an average flow to exist.

⁸⁵This assumption is called the **closing**, and the treatment of the instability problem in this way is called the **Reynolds averaging**.

Because the average flow solves the Navier-Stokes equations on its own, (5.57) reduces to the linear equations

$$\begin{aligned} \partial_t \tilde{u} + U \partial_x \tilde{u} + \tilde{v} \frac{dU}{dy} + \frac{1}{\rho} \partial_x \tilde{p} &= \nu \Delta \tilde{u} \\ \partial_t \tilde{v} + U \partial_x \tilde{v} + \frac{1}{\rho} \partial_y \tilde{p} &= \nu \Delta \tilde{v} \\ \partial_x \tilde{u} + \partial_y \tilde{v} &= 0. \end{aligned} \quad (5.58)$$

If the first equation in (5.58) is differentiated with respect to y and the second one with respect to x , the pressure can be eliminated so that there remains only one equation for u and v besides the continuity equation. The no-slip condition along surfaces of solids leads to additional boundary conditions.

Let us now suppose that the disturbances result from the superposition of waves propagating in x -direction where each of them fulfills (5.58). These waves can be expressed by

$$\tilde{\mathbf{u}} = \begin{pmatrix} \varphi(y) \\ \psi(y) \end{pmatrix} e^{i(\alpha x - \gamma t)} \quad (5.59)$$

and are called **Tollmien-Schlichting waves**. Here, $\alpha \in \mathbb{C}$, $\gamma \in \mathbb{C}$. The particular case $\alpha \in \mathbb{R}$ describes the investigation of the **temporal stability**, whereas an investigation of the case $\alpha \in \mathbb{C}$, $\gamma \in \mathbb{R}$ concerns the so-called **spatial stability** not discussed here.

The continuity equation leads to

$$\varphi(y) = \frac{i}{\alpha} \psi'(y),$$

and the equation that remains after the elimination of \tilde{p} then yields

$$(U - \gamma)(\psi'' - \alpha^2 \psi) - U'' \psi = -\frac{i\nu}{\alpha} (\psi^{(4)} - 2\alpha^2 \psi'' + \alpha^4 \psi). \quad (5.60)$$

The linear homogeneous 4th order ordinary differential equation (5.60) for ψ with complex coefficients is called the **Orr-Sommerfeld equation**. Boundary conditions are for instance the no-slip condition $\psi = 0$ for $y = 0$ and the disappearance of the disturbances for $y \rightarrow \infty$.

Thus, for every fixed $\alpha \in \mathbb{R}$, an eigenvalue problem arises where the complex **phase speed of perturbation** γ plays the role of the eigenvalue. The parameter ν^{86} is

⁸⁶or, if the problem is formulated in dimensionless form, $\frac{1}{\text{Re}}$

considered to be already known from the laminar average flow. This laminar flow corresponds to the **trivial solution** $\psi \equiv 0$ of problem (5.60), and one asks for nontrivial solutions bifurcating from it.

If the sign of the imaginary part of γ is negative, the disturbance decreases when time grows, and its amplitude does not depend on time if this imaginary part vanishes. But an unstable situation occurs for

$$\operatorname{sgn}(\operatorname{Im} \gamma) > 0. \quad (5.61)$$

The Orr-Sommerfeld equation can under realistic boundary conditions for U normally not be solved explicitly though it is linear and homogeneous. Numerical procedures known from numerical bifurcation theory have to be used.

The smallest value of the Reynolds number giving rise to the validity of (5.61) is called the **critical Reynolds number** $\operatorname{Re}_{\text{cr}}$ and is of interest because it is the smallest Reynolds number for which the flow can become unstable. The curve $\operatorname{sgn}(\operatorname{Im} \gamma) = 0$ obviously separates the area of instability in the $(\alpha, \operatorname{Re})$ -plane from the area of a stable laminar flow and is called the **curve of neutral stability**.

It should be noted that an instability result due to this linearized theory of disturbances does not necessarily ensure that the real underlying flow can already be considered to be completely turbulent. A measure accepted by engineers is the so-called **e^9 -rule**: The flow is really a turbulent one if the small amplitude of the initial disturbance has increased by the factor $e^9 \approx 8100$.

6 Existence Proof for Entropy Solutions by Means of Discretization Procedures

Though we have already studied some properties of weak solutions of conservation laws, we did not yet prove the existence of this type of solutions, particularly of entropy solutions. The only exceptions were some scalar problems whose solutions could explicitly be specified.

We are now going to attack the existence problem.

6.1 Some Historical Remarks

Existence proofs for solutions of ordinary or partial differential equations are often structured in the following way: In a first step, the differential equation will be discretized as normally done in Numerical Analysis in order to create numerical procedures. Thus, a finite dimensional approximate version of the original problem solvable on a certain grid for every suitable step size parameter h comes into being. If the solution of the discrete problem is only given on a set of discrete points, it will be extended to the inter grid points by interpolation or other techniques in such a way that it belongs to the particular function space the solution or weak solution of the original problem is expected to belong to. This extension is called **reconstruction** (cf. the special case (4.2)). Moreover, we assume that all the computations to be made were theoretically realized without occurrence of round off errors.

If it can be shown in a second step that a sequence \mathcal{L} of these reconstructed approximate solutions determined for every step size parameter h_j of a sequence $\{h_j\}$ of positive step size parameters with $\lim_{j \rightarrow \infty} h_j = 0$ is compact in the function space under consideration, convergent subsequences of \mathcal{L} exist. It can often be shown in a third step that under certain conditions the limits of those subsequences solve the original problem. Thus, solutions exist in this case, but different convergent subsequences can lead to different solutions of the original problem because uniqueness was not yet ensured.

A well known example of an existence proof of this type is the particular proof of Peano's theorem on the existence of a solution of the initial value problem

$$y' = f(x, y), \quad y(x_0) = y_0, \quad f \text{ continuous}$$

where Euler's polygon method is used in order to construct approximate solutions in the space of functions continuous on a compact interval and equipped with the maximum norm.

Another example is the existence proof of Courant, Friedrichs, Lewy where this type of proof was for the first time applied on initial value problems of systems of quasilinear hyperbolic partial differential equations. The proof was published in their famous paper of 1928⁸⁷, and it was in this paper that a condition concerning the step ratio was formulated in order to ensure numerical stability of the numerical procedure, later called the CFL condition. This condition ensures that the region affecting the solution at a point Q where the solution is expected to exist is covered by the region that affects the approximate solution at Q , and this condition is extremely important when treating conservation laws by means of explicit numerical methods⁸⁸.

Of course, because Courant, Friedrichs and Lewy asked for smooth solutions, they could only give conditions for local existence.

Under certain restrictions concerning the initial functions, J. Glimm gave an existence theorem for global weak solutions of conservation laws in a paper that also became famous⁸⁹, and one of his tools was also a certain finite difference procedure.

We are now going to study a general theory on discretizations of partial differential equations that can also help to prove existence theorems for conservation laws as occurring in fluid dynamics⁹⁰.

6.2 Reduction to Properties of Operator Sequences

In 1948, Kantorowitch published his important paper on the connections between numerical mathematics and functional analysis⁹¹.

Particularly in case of linear partial evolution equations, Lax and Richtmyer provided in 1956 a functionalanalytic frame for the construction and investigation of finite dif-

⁸⁷Math. Ann. **100**, 32-74 (1928)

⁸⁸For more details cf. Section 7.3.

⁸⁹Comm. Pure Appl. Math. **18**, 697-715 (1965)

⁹⁰R. Ansorge, ZAMM **73**, 239-253 (1993)

⁹¹Uspeki Mat. **3**, 89-185 (1948)

ference methods⁹² (cf. section 7.4). Their theory was later generalized with respect to different aspects.

Inspired by the Lax-Richtmyer Theory but besides the particular question of the discretization of differential equations, Stummel published in 1972 his general theory on discretization algorithms⁹³. Though it was already rather general, it was still influenced by linear structures underlying the given general problem. This theory was later also generalized into different directions and applied to several types of problems such as ordinary or partial differential equations, integral equations, optimization problems etc. Since then, this theory was called **Functional Analysis of Discretization Algorithms** (Vainikko).

But not yet included into this theory were weak solutions of differential equations or problems whose solutions are not necessarily unique.

Let us therefore in a first step try to describe an imbedding of discretization techniques for problems of this type into the general theory and look for problems given in the following general form:

Let X , Y , X_n , Y_n be topological, usually metric spaces, and let us ask for an element $v \in X$ which fulfills

$$Av = a, \quad A : X \rightarrow Y \quad \text{a given operator.} \quad (6.1)$$

Let the numerical procedure constructed in order to solve this problem also be described in a general form by

$$A_n v_n = a_n, \quad A_n : X_n \rightarrow Y_n, \quad n \in \mathbb{N}, \quad (6.2)$$

where $\{a_n | a_n \in Y_n, n \in \mathbb{N}\}$ is a given sequence approximating $a \in Y$ in a sense still to be specified.

Neither $X_n \subset X_{n+1}$ nor $X_n \subset X$ and neither $Y_n \subset Y_{n+1}$ nor $Y_n \subset Y$ ($n = 1, 2, \dots$) are assumed to hold.

Assume that there are either mappings

$$p_n : X \rightarrow X_n, \quad q_n : Y \rightarrow Y_n$$

$$\begin{array}{ccc} X & \xrightarrow{A} & Y \\ p_n \downarrow & & \downarrow q_n \\ X_n & \xrightarrow{A_n} & Y_n \end{array}$$

⁹²Comm. Pure Appl. Math. **9**, 267-293 (1956)

⁹³Proc. Conf. Num. Anal. Dublin 1972 (ed.: J. Miller), 285-310, Academic Press 1973

or an imbedding of the numerical procedure, just for theoretical purposes, into the frame of the original problem. This can for instance be done in the sense of

$$X_n \subset X, \quad Y_n = Y \quad (n = 1, 2, \dots),$$

but not necessarily

$$X_n \subset X_{n+1} \quad (n = 1, 2, \dots).$$

Just for convenience, let us suppose such an imbedding to be realized, but the given original problem does not necessarily need to be linear, and its solutions are allowed to be weak in a sense to be specified.

Let the original problem be described by

$$\tilde{A}v = w, \tag{6.3}$$

with $\tilde{A} : \tilde{X} \rightarrow Y$, $\tilde{X} \subset X$.

The weak formulation of (6.3) which then will be the problem really to be solved is supposed to be described in the following way:

Let J be an index set; let $\{a(\Phi), \Phi \in J\} \subset Y$ be a given set and $\{A(\Phi), \Phi \in J\}$ a set of operators with joint domain $D \subset X$, $D \rightarrow Y$.

We ask for an element $v \in X$ with

$$A(\Phi)v = a(\Phi) \quad \forall \Phi \in J. \tag{6.4}$$

Definition:

The elements of

$$\Theta := \{v \in X | v \text{ solves (6.4)}\}$$

*are called **weak solutions** of (6.3) if the following implications hold:*

If v solves (6.3), then $v \in \Theta$; if $v \in \tilde{X} \cap \Theta$, it also solves (6.3).

*The elements $\Phi \in J$ are called **test elements**.*

Let the numerical method be described in an also rather general form by

$$\hat{A}_n v_n = \hat{a}_n \quad (n = 1, 2, \dots) \tag{6.5}$$

where $\hat{A}_n: X_n \rightarrow Z$ ($n = 1, 2, \dots$) with a certain topological space Z . Here, $\{\hat{a}_n\} \subset Z$ is a given sequence compact in Z .

Assume

$$\Theta_n := \{v_n \in X_n | v_n \text{ solves (6.5)}\} \neq \emptyset \quad (n = 1, 2, \dots),$$

but every Θ_n may contain more than one element. The solutions v_n of (6.5) are called **approximate solutions** of (6.4).

The numerical method (6.5) is assumed not to depend on test elements because computers cannot understand what test elements are. But we suppose that also this method can be formulated weakly by means of a sequence of operators

$$\{A_n(\Phi) | \Phi \in J, n \in \mathbb{N}\} \text{ with } A_n(\Phi) : X_n \rightarrow Y_n,$$

and of a sequence $\{a_n(\Phi)\}$ with

$$\lim_{n \rightarrow \infty} a_n(\Phi) = a(\Phi) \quad \text{for every fixed } \Phi \in J,$$

so that

$$A_n(\Phi)v_n = a_n(\Phi) \quad \forall \Phi \in J, \forall v_n \in \Theta_n \tag{6.6}$$

holds.

(6.6) is called a **weak formulation of the numerical method**.

Some concepts help to prove convergence results:

Definition:

A pair $[\{C_n\}, C]$, consisting of an operator sequence $\{C_n\}$ and of an operator C is called **asymptotically closed** if the implication

$$v_n \rightarrow v \quad \wedge \quad C_n v_n \rightarrow z \quad \Rightarrow \quad C v = z$$

holds.

Definition:

An operator sequence $\{C_n\}$ is called **asymptotically regular** if the implication

$$\{C_n v_n\} \text{ compact in } Y \quad \Rightarrow \quad \{v_n\} \text{ compact in } X$$

holds.

Definition:

Method (6.5) is called **convergent** if set convergence

$$\Theta_n \rightarrow \Theta$$

is ensured in the following sense:

$\{\Theta_n\}$ is **discretely compact**, i.e. every sequence $\{v_n | v_n \in \Theta_n ; n = 1, 2, \dots\}$ is compact in X , and if v is a limit of a convergent subsequence, $v \in \Theta$ follows.

Remark:

$\Theta_n \rightarrow \Theta$ implies $\Theta \not\subset \emptyset$. In other words: There is a solution of (6.4), i.e. a weak solution of (6.3).⁹⁴

6.3 Convergence Theorems

We can now state a convergence theorem:

Theorem 6.1:

- (i) Let $\{A_n(\Phi)\}$, $A(\Phi)$ be asymptotically closed for every fixed $\Phi \in J$ and let
- (ii) $\{\hat{A}_n\}$ be asymptotically regular.

Then:

$$\Theta_n \rightarrow \Theta.$$

PROOF: $\{\hat{a}_n\}$ compact in $Z \Rightarrow \{\hat{A}_n v_n | v_n \in \Theta_n\}$ compact in Z , and this for every sequence $\{v_n | v_n \in \Theta_n\}$.

Thus, each of the sequences $\{v_n\}$ is compact in X because of assumption (ii), i.e.

$$\exists \{v_n | v_n \in \Theta_n, n \in \mathbb{N}' \subset \mathbb{N}\}, \quad \exists v \in X \quad \text{with} \quad v_n \rightarrow v \quad (n \in \mathbb{N}').$$

Here, \mathbb{N}' and v are independent of Φ .

But because v_n fulfills also $A_n(\Phi)v_n = a_n(\Phi)$, and because of

$$a_n(\Phi) \rightarrow a(\Phi) \quad \forall \Phi \in J,$$

supposition (i) yields $A(\Phi)v = a(\Phi)$, i.e. $v \in \Theta$. □

⁹⁴A proof of $\Theta_n \rightarrow \Theta$ is therefore also an existence proof.

If one is only interested in a particular problem, i.e. in a particular right hand side of equation (6.3), assumption (ii) can obviously be replaced by the weaker requirement

$$(ii^*) \quad \{\Theta_n\} \text{ discretely compact.}$$

If uniqueness of a weak solution is not guaranteed, sometimes additional restrictions are formulated in order to pick out of the set of weak solutions the particular one relevant from the user's point of view. And as in the case of scalar conservation laws, this restriction often takes the form of an inequality the relevant solution has additionally to fulfill. Therefore we are going to call such a restriction generally an **entropy condition**, and we formulate this condition in a general setting as

$$B(\hat{\Phi})v \leq 0 \quad \forall \hat{\Phi} \in \hat{J}, \quad (6.7)$$

where $\{B(\hat{\Phi}) | \hat{\Phi} \in \hat{J}\}$ denotes a set of nonlinear continuous functionals mapping X into \mathbb{R} , and where \hat{J} is a certain index set which may differ from J .

We suppose that there is a uniqueness theorem available⁹⁵:

*There is at most one solution of (6.4) that fulfills also (6.7), called the **entropy solution**.*

If the original problem is completed by an entropy condition in order to characterize the relevant solution, it seems to be reasonable to discretize also this condition by

$$B_n(\hat{\Phi})v_n \leq 0 \quad \forall \hat{\Phi} \in \hat{J} \quad (n = 1, 2, \dots), \quad (6.8)$$

and to accept only such approximate solutions $v_n \in S_n$ which additionally fulfill this discrete entropy version. Herewith, the functionals $B_n(\hat{\Phi})$ are also assumed to be continuous.

We suppose

$\{(6.5), (6.8)\}$ has *at least* one solution $\hat{v}_n \in \Theta_n$ for every fixed $n \in \mathbb{N}$.

(6.8) is called a sequence of **discrete entropy conditions**, but we do not expect a **discrete entropy solution**

$$\hat{v}_n \quad (n \in \mathbb{N} \text{ arbitrary})$$

necessarily to be unique.

⁹⁵Theorem 3.3 is an example.

Definition:

An operator sequence $\{C_n\}$ is called **continuously convergent** to an operator C if the implication

$$v_n \rightarrow v \Rightarrow C_n v_n \rightarrow C v$$

holds⁹⁶⁹⁷. We then write

$$C_n \xrightarrow{c} C.$$

We can now supplement Theorem 6.1 with another important remark:

Theorem 6.2: *Let*

$$\hat{\Theta}_n := \{\hat{v}_n \in \Theta_n \mid \hat{v}_n \text{ solves (6.8)}\} \neq \emptyset,$$

and let

$$B_n(\hat{\Phi}) \xrightarrow{c} B(\hat{\Phi}) \quad \text{be fulfilled for every fixed } \hat{\Phi} \in \hat{J}.$$

Moreover, assume the suppositions of Theorem 6.1 to be fulfilled.

Then there is an entropy solution \hat{v} of $\{(6.4),(6.7)\}$, and

$$\hat{\Theta}_n \rightarrow \{\hat{v}\}.$$

PROOF:

Because of the validity of Theorem 6.1, i.e. because of $\Theta_n \rightarrow \Theta$, there is a subsequence

$$\{\hat{v}_n \mid \hat{v}_n \in \hat{\Theta}_n, n \in \mathbb{N}' \subset \mathbb{N}\}$$

and a weak solution $v \in \Theta$ with $\{\hat{v}_n\} \rightarrow v, n \in \mathbb{N}'$.

The *continuous convergence* therefore yields $B_n(\hat{\Phi})\hat{v}_n \rightarrow B(\hat{\Phi})v$.

From (6.8), $B(\hat{\Phi})v \leq 0$ follows. Consequently, v is an entropy solution, and because the entropy solution was assumed to be unique, v coincides with this entropy solution, i.e. $v = \hat{v}$, and not only a subsequence converges to \hat{v} but the full sequence converges:

$$\lim_{n \rightarrow \infty} \hat{v}_n = \hat{v}. \quad \square$$

⁹⁶P. du Bois-Reymond: Sitz.-Berichte, Preussische Akad. d. Wiss., 359-360 (1886)

⁹⁷R. Courant: Göttinger Nachr., 101-109 (1914)

Conditions for the property of *continuous convergence* on metric spaces follow from a theorem due to W. Rinow⁹⁸ :

Theorem 6.3:

If the operators C, C_n ($n = 1, 2, \dots$) mapping a metric space X into a metric space Y are individually continuous, then $C_n \xrightarrow{c} C$ is equivalent to the simultaneous validity of the two statements:

- (j) $\{C_n\}$ is a sequence of equicontinuous operators ,
- (jj) $C_n \rightarrow C$ holds pointwise on a subset $\tilde{X} \stackrel{c}{\text{dense}} X$.

6.4 Example

As an example, we treat the scalar version of problem (2.1). Thus, let $x \in \mathbb{R}$ be a spatial variable and $t \geq 0$ a time variable, and let Ω be again the upper (x, t) -halfplane.

We ask for solutions $v : \Omega \rightarrow \mathbb{R}$ of the scalar conservation law

$$\begin{aligned} \partial_t v + \partial_x f(v) &= 0, & t > 0 \\ v(x, 0) &= v_0(x) & \text{with } v_0 \in BV(\mathbb{R}) \cap L_\infty(\mathbb{R}), \end{aligned} \tag{6.9}$$

where the flux $f \in C^1$ is strictly convex and has its minimum at zero which can be assumed to be zero: $f(0) = 0$. Let

$$|f'|_\infty^* := \max \{ |f'(v)|, |v| \leq \|v_0\|_{L_\infty} \} < \infty . \tag{6.10}$$

Because of possible discontinuities we are interested in weak global solutions $v \in X := L_1^{\text{loc}}$ of

$$\begin{aligned} A(\Phi)v &:= - \int_\Omega \int \{ v \partial_t \Phi + f(v) \partial_x \Phi \} dx dt \\ &- \int_{-\infty}^{\infty} v_0(x) \Phi(x, 0) dx = 0 \quad \forall \Phi \in J := C_0^1(\Omega) . \end{aligned} \tag{6.11}$$

Let

$$\tilde{S} = \tilde{S}(v, c) , \quad \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$$

⁹⁸Die innere Geometrie der metrischen Räume, p. 78 ff., Berlin-Göttingen-Heidelberg: Springer 1961

be a one parameter set of so-called **entropy functions** continuously differentiable for every fixed $c \in \mathbb{R}$ with respect to v , possibly with an exception at $v = c$.

$\tilde{F} = \tilde{F}(v, c) \quad (\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R})$ is called the **entropy flux**, if

$$\partial_t \tilde{S}(v(x, t), c) + \partial_x \tilde{F}(u(x, t), c) = 0 \quad \text{on } \Omega'$$

holds for every fixed $c \in \mathbb{R}$ and for every pair (v, Ω') , for which v is a weak solution on Ω but even a smooth C^1 -solution on $\Omega' \subset \Omega$.

If a solution v , which is weak on Ω , fulfills on Ω additionally the inequality

$$\partial_t \tilde{S}(v(x, t), c) + \partial_x \tilde{F}(v(x, t), c) \leq 0 \quad \forall c \in \mathbb{R}$$

in a weak sense, i.e.

$$\begin{aligned} - \int_{\Omega} \int \left\{ \partial_t \Phi(x, t) \tilde{S}(v(x, t), c) + \partial_x \Phi(x, t) \tilde{F}(v(x, t), c) \right\} dx dt \\ - \int_{-\infty}^{\infty} \Phi(x, 0) \tilde{S}(v_0(x), c) dx \leq 0 \quad (6.12) \\ \forall c \in \mathbb{R}, \quad \forall \text{ nonnegative } \Phi \in J, \end{aligned}$$

then v is called the **entropy solution**, denoted by \hat{v} .

Thus, test elements in this example are the pairs

$$\hat{\Phi} := (\Phi, c) \quad \text{with } \hat{\Phi} \in \hat{J} := \{(\Phi, c) | \Phi \in J, \Phi \geq 0, c \in \mathbb{R}\},$$

and the integral inequality can therefore be written as

$$B(\hat{\Phi})v \leq 0 \quad \forall \hat{\Phi} \in \hat{J} \quad (6.13)$$

in accordance with the general theory. Particular choices of \tilde{S} in the literature are

$$\begin{aligned} \tilde{S} &= \tilde{S}(v), & \tilde{S} & \text{ independent of } c \text{ but strictly convex,} \\ & & \hat{\Phi} &= \{\Phi \in J | \Phi \geq 0\} \subset J \\ & & & \text{(Lax; cf. (3.11)),} \\ \tilde{S}(v, c) &= |v - c| & & \text{(Kruzkov; cf. (3.44)),} \\ \tilde{S}(v, c) &= \begin{cases} v - c & , \quad v \geq c \\ 0 & , \quad v < c \end{cases} & & \text{(Harten, Hyman, Lax⁹⁹)} \\ & \text{etc.} & & \end{aligned}$$

⁹⁹Comm. Pure Appl. Math. **29**, 297-322 (1976)

Each of these choices then implies a particular choice of \tilde{F} according to (3.13).

Let the numerical procedure used in the general setting be put in concrete form by a $(2k+1)$ -point finite difference scheme:

$$\frac{v_j^{\nu+1} - v_j^\nu}{\Delta t} + \frac{g_{j+\frac{1}{2}}^\nu - g_{j-\frac{1}{2}}^\nu}{\Delta x} = 0, \quad (6.14)$$

where ν counts the number of time steps, j the number of spatial steps, starting from

$$v_j^0 = \frac{1}{\Delta x} \int_{(j-1/2)\Delta x}^{j+1/2\Delta x} v_o(x) dx.$$

Here, the **numerical flux**

$$g_{j+\frac{1}{2}}^\nu := g(v_{j-k+1}^\nu, v_{j-k+2}^\nu, \dots, v_j^\nu, v_{j+1}^\nu, \dots, v_{j+k}^\nu)$$

is assumed to be Lipschitz continuous and **consistent** with the original problem, i.e.

$$g(w, w, \dots, w) = f(w), \quad \forall w \in \mathbb{R}. \quad (6.15)$$

Methods of this form are called **methods in conservation form**.

We restrict the choices of the step sizes by the requirement that the CFL condition, already introduced in the beginning of this chapter, has to be respected. For the problem under consideration this means

$$0 < \lambda := \frac{\Delta t}{\Delta x} = \text{const} \leq \frac{1}{|f'|_\infty^*}. \quad (6.16)$$

Let $\Delta x = O(\frac{1}{n})$ ($n \in \mathbb{N}$), and put

$$v_n(x, t) = v_j^\nu \quad \text{for} \quad \begin{cases} (j - \frac{1}{2})\Delta x \leq x < (j + \frac{1}{2})\Delta x \\ \nu \Delta t \leq t < (\nu + 1)\Delta t \end{cases}$$

$$(j = 0, \pm 1, \pm 2, \dots), (\nu = 1, 2, \dots).$$

If X_n is then chosen as the space of functions defined on Ω and constant on these rectangles, the finite difference scheme can be formulated as

$$\hat{A}_n v_n = 0 \quad (n = 1, 2, \dots) \quad (6.17)$$

with operators $\hat{A}_n : X_n \rightarrow Z := \mathbb{R}$.

Multiplication of (6.14) by test functions $\Phi \in J$ and integration leads to

$$A_n(\Phi) v_n = \tag{6.18}$$

$$\sum_{\nu=0}^{\infty} \sum_j \int_{\nu \Delta t}^{(\nu+1) \Delta t} \int_{(j-\frac{1}{2})\Delta x}^{(j+\frac{1}{2})\Delta x} \Phi(x, t) \left[v_j^{\nu+1} - v_j^{\nu} + \lambda \left(g_{j+\frac{1}{2}}^{\nu} - g_{j-\frac{1}{2}}^{\nu} \right) \right] dx dt = 0$$

and describes the operators $A_n(\Phi)$ as well as the right sides $a_n(\Phi)$, which equal zero in this context.

Using the abbreviation

$$\tilde{S}_j^{\nu}(c) = \tilde{S}(v_j^{\nu}, c),$$

v_n is called a **discrete entropy solution**, denoted by \hat{v}_n , if

$$\frac{\tilde{S}_j^{\nu+1}(c) - \tilde{S}_j^{\nu}(c)}{\Delta t} + \frac{G_{j+\frac{1}{2}}^{\nu}(c) - G_{j-\frac{1}{2}}^{\nu}(c)}{\Delta x} \leq 0 \quad \forall c \in \mathbb{R}, \tag{6.19}$$

where

$$G_{j+\frac{1}{2}}^{\nu}(c) := G(u_{j-k+1}^{\nu}, \dots, v_j^{\nu}, v_{j+1}^{\nu}, \dots, v_{j+k}^{\nu}, c)$$

denotes a Lipschitz continuous **numerical entropy flux** to be specified and consistent with the entropy flux \tilde{F} in the sense of

$$G(w, w, \dots, w, c) = \tilde{F}(w, c) \quad \forall (w, c) \in \mathbb{R}^2. \tag{6.20}$$

For $k = 1$ (3-point case) the **flux splitting** choice

$$G(\alpha, \beta, c) = \tilde{F}_+(\alpha, c) + \tilde{F}_-(\beta, c) \tag{6.21}$$

with

$$\tilde{F}_+(\alpha, c) := \begin{cases} \tilde{F}(\alpha, c) & , \alpha \geq 0 \\ 0 & , \alpha < 0 \end{cases}, \quad \tilde{F}_-(\beta, c) := \begin{cases} 0 & , \beta \geq 0 \\ \tilde{F}(\beta, c) & , \beta < 0 \end{cases}$$

is one of several possible choices of G , where \tilde{F} is known as soon as \tilde{S} is chosen.

The weak formulation of the discrete form of the entropy condition needed for the use of Theorem 6.2 arises from multiplication of (6.19) by nonnegative test functions $\hat{\Phi} \in J$ and integration:

$$\begin{aligned}
 B_n(\hat{\Phi}) v_n &:= \int_{\Omega} \int \hat{\Phi}(x, t) \left\{ \frac{1}{\Delta t} \left[\tilde{S}(v_n(x, t + \Delta t), c) - \tilde{S}(v_n(x, t), c) \right] \right. \\
 &+ \frac{1}{\Delta x} [G(v_n(x, t), v_n(x + \Delta x, t), c) \\
 &- G(v_n(x - \Delta x, t), v_n(x, t), c)] \left. \right\} dx dt \leq 0
 \end{aligned} \tag{6.22}$$

$\forall \hat{\Phi} \in \hat{J}$, i.e.

$$\begin{aligned}
 B_n(\hat{\Phi}) v_n &= - \int_{\Delta t}^{\infty} \int_{-\infty}^{\infty} \frac{\hat{\Phi}(x, t) - \hat{\Phi}(x, t - \Delta t)}{\Delta t} \tilde{S}(v_n(x, t), c) dx dt \\
 &- \frac{1}{\Delta t} \int_0^{\Delta t} \sum_i \int_{x_i - \Delta x/2}^{x_i + \Delta x/2} \hat{\Phi}(x, t) dx \tilde{S}(v_i^0, c) dt \\
 &- \int_{\Omega} \int \frac{\hat{\Phi}(x + \frac{\Delta x}{2}, t) - \hat{\Phi}(x - \frac{\Delta x}{2}, t)}{\Delta x} G(v_n(x - \frac{\Delta x}{2}, t), v_n(x + \frac{\Delta x}{2}, t), c) dx dt \\
 &\leq 0 \quad \forall \hat{\Phi} \in \hat{J}.
 \end{aligned} \tag{6.23}$$

Thus, B and B_n from (6.7) and (6.8), respectively, are put into concrete form provided that a 3-point scheme is under consideration. If then, for instance, the Harten-Hyman-Lax choice of \tilde{S} is taken into account and G is chosen according to (6.21), then $B(\hat{\Phi})$ in (6.12), (6.13) is continuous for every fixed $\hat{\Phi}$. The functionals $B_n(\hat{\Phi})$ ($n = 1, 2, \dots$) in (6.22) are equicontinuous, and pointwise convergence $B_n(\hat{\Phi}) \rightarrow B(\hat{\Phi})$ can easily be shown provided that the joint domain of these functionals is restricted to elements $v \in X^*$; herewith, X^* is the space of functions $v \in X$ restricted to the domain $\text{supp } \hat{\Phi}$ which is equipped with the L_1 -metric on X^* .

Hence, the assumptions (j) and (jj) of Rinow's theorem are fulfilled, i.e.

$$B_n(\hat{\Phi}) \xrightarrow{c} B(\hat{\Phi}),$$

and Theorem 6.2 applies as far as $\hat{\Theta}_n \neq \emptyset^{100}$.

¹⁰⁰cf. footnote 90

The explicit¹⁰¹ **Engquist-Osher 3-point flux splitting scheme**¹⁰², characterized by (6.14) with

$$g_{j+\frac{1}{2}}^{\nu} = g(v_j^{\nu}, v_{j+1}^{\nu}) := f_{-}(v_{j+1}^{\nu}) + f_{+}(v_j^{\nu}) \quad (6.24)$$

where

$$f_{+}(\alpha) = \begin{cases} f(\alpha) & \text{for } \alpha \geq 0 \\ 0 & \text{for } \alpha < 0 \end{cases}, \quad f_{-}(\alpha) = \begin{cases} 0 & \text{for } \alpha \geq 0 \\ f(\alpha) & \text{for } \alpha < 0 \end{cases},$$

together with the Harten-Hyman-Lax choice of \tilde{S} , is an example. (6.24) then describes the operators \hat{A}_n ($n = 1, 2, \dots$) in (6.17)¹⁰³, hence also the operators $A_n(\Phi)$ via (6.18). The method simulates an important property of weak solutions of scalar problems of type (2.1) provided the CFL condition is fulfilled, namely their **monotonicity**, hence its **Total Variation Diminishing** or **TVD** property and, therefore, its **Total Variation Boundedness** or **TVB**.

We interrupt the proof of our convergence theorem for some explanations:

A method used for the numerical treatment of scalar conservation laws of type (2.1), represented in the general form

$$v_j^{n+1} = H(v_{j-k}^n, v_{j-k+1}^n, \dots, v_{j+k}^n),$$

*is called **monotone** if the partial derivatives of the function H with respect to all its variables are nonnegative.*

This imitates the following property of the weak solutions of (2.1): If two initial functions \tilde{v}_0 and v_0 fulfill the inequality

$$\tilde{v}_0(x) \geq v_0(x) \quad \forall x \in \mathbb{R},$$

this inequality holds everywhere in Ω for the weak solutions $\tilde{v}(x, t)$ and $v(x, t)$ belonging to these initial functions, respectively.

TV-boundedness in this context (and formulated not only for scalar equations but also for systems) means that $\sum_j \|V_{j+1}^n - V_j^n\|_{\mathbb{R}^m} \leq K$ for all n , and **TVD methods** introduced by Harten¹⁰⁴ simulate the fact that the total variation of weak solutions of scalar problems does not increase, and it is obvious that TVD implies TVB. The fact that monotonicity of a scheme implies its TVD property was shown in an

¹⁰¹It follows from this explicitness that $\Theta_n \neq \emptyset$, ($n = 1, 2, \dots$).

¹⁰²Math.Comput. **34**, 45-75 (1980)

¹⁰³ $\hat{a}_n = 0$, ($n = 1, 2, \dots$).

¹⁰⁴SIAM J. Numer.Anal. **21**, 1-23 (1984)

appendix to the paper of Harten, Hyman, Lax (cf. footnote 98), written by Barbara Keyfitz.

We continue proving the convergence theorem:

It was LeVeque¹⁰⁵ who showed that Θ_n is discretely compact¹⁰⁶ provided that v_0 is of bounded variation. It is here where we put our requirement $v_0 \in BV(\mathbb{R})$ to use.

Thus, assumption (ii) of Theorem 6.1 is fulfilled because (ii*) is fulfilled.

The property (i), i.e.

$$[\{A_n(\Phi)\}, A(\Phi)] \text{ discretely closed ,}$$

remains to be shown in order to show the convergence of the method by means of Theorem 6.1. But this can in case of TVB methods for scalar equations immediately be deduced from the **Lax-Wendroff Theorem**¹⁰⁷.

This ends the proof for the existence of an entropy solution of our scalar conservation law example.

As a particular result, we notice that a convergent monotone consistent scheme in conservation form converges to the entropy solution of the original problem.

The Lax-Wendroff Theorem referred to in the proof of the convergence theorem reads as follows:

Theorem 6.4:

Let $\lambda = \frac{\Delta t}{\Delta x} = \text{const}$ and let $\Delta t = \mathcal{O}\left(\frac{1}{n}\right)$ ($n \in \mathbb{N}$) for $n \rightarrow \infty$. Define $\mathbf{V}^n \in L_1^{\text{loc}}$ to be the step function

$$\mathbf{V}^n(x, t) := \mathbf{V}_j^n \text{ for } x_j - \frac{\Delta x}{2} \leq x < x_j + \frac{\Delta x}{2} \text{ , } t_n \leq t < t_{n+1} \text{ ,}$$

$$(j = 0, \pm 1, \pm 2, \dots \text{ ; } n = 0, 1, 2, \dots) \text{ on } \Omega = \{(x, t) \mid x \in \mathbb{R} \text{ , } t \geq 0\} \text{ ,}$$

where the numerical values \mathbf{V}_j^n were computed by means of an explicit TV-bounding finite difference one-step scheme whose numerical flux \mathbf{g} is consistent and Lipschitz continuous and which starts from

$$\mathbf{V}_j^0 := \frac{1}{\Delta x} \int_{x_j - \frac{\Delta x}{2}}^{x_j + \frac{\Delta x}{2}} \mathbf{V}_0(\xi) d\xi \quad (j = 0, \pm 1, \pm 2, \dots) \text{ .}$$

¹⁰⁵Numerical Methods for Conservation Laws, p. 164. Basel-Boston-Berlin: Birkhäuser 1990

¹⁰⁶though he used another terminology

¹⁰⁷Comm. Pure Appl. Math. **13**, 217-237 (1960)

Assume that \mathbf{V}^n converges for $n \rightarrow \infty$, i.e. for $\Delta t \rightarrow 0$), to a function \mathbf{V} in the L_1^{loc} -topology, i.e. with respect to the L_1 -norm on each compact subset of Ω and for each of the components of \mathbf{V}^n .

Then \mathbf{V} is a weak solution of (2.1).

We omit the proof of the Lax-Wendroff Theorem but notice that this is not yet a convergence theorem! The theorem only says that the limits of convergent sequences of approximate solutions –**as far as they do exist**– are weak solutions of the original problem.

In order to finish off a proof showing the important property of a method to converge, one has to look for properties which ensure the validity of the assumptions of Theorems 6.1 and 6.2 (as it was realized in the case of the Engquist-Osher scheme).

It should be mentioned that the general convergence theorem of this chapter does of course not only apply in case of initial value problems for scalar conservation laws but also in case of certain nonlinear elliptic boundary value problems numerically treated by means of projection methods etc.¹⁰⁸

¹⁰⁸R. Ansorge and J. Lei, ZAMM **71**, 207-221 (1991)

7 Types of Discretization Principles

7.1 Some General Remarks

We want to describe the principal possibilities of discretizing hyperbolic conservation laws and describe their common features as well as their differences¹⁰⁹. For the sake of simplicity we consider the generic scalar equation

$$L(u) := \partial_t u + \partial_x u + \partial_y u = 0. \quad (7.1)$$

There are three basic classes of discretization methods for (7.1), namely the **Finite Element Method** (FEM), the **Finite Volume Method** (FVM), and the **Finite Difference Method** (FDM) already introduced briefly in section 4.1 and already used for more theoretical purposes in section 6.4. Although these names sound quite definitive there are very many particular schemes behind each of the three classes of methods.

In the FDM we begin by discretizing time

$$\mathbb{G}^n := \{n\Delta t \mid \Delta t > 0, n \in \mathbb{N} \cup \{0\}\}$$

as well as space

$$\mathbb{G}_{i,j} := \{(i\Delta x, j\Delta y) \mid \Delta x, \Delta y > 0, i, j \in \mathbb{Z}\}$$

and replace the differential operators in (7.1) by **difference operators**, for example:

$$\begin{aligned} \partial_t u(i\Delta x, j\Delta y, n\Delta t) &\approx \frac{u(i\Delta x, j\Delta y, (n+1)\Delta t) - u(i\Delta x, j\Delta y, n\Delta t)}{\Delta t} \\ &=: \frac{u_{i,j}^{n+1} - u_{i,j}^n}{\Delta t} \end{aligned}$$

¹⁰⁹cf. also Chapter 6

and

$$\begin{aligned}\partial_x u(i\Delta x, j\Delta y, n\Delta t) &\approx \frac{u(i\Delta x, j\Delta y, n\Delta t) - u((i-1)\Delta x, j\Delta y, n\Delta t)}{\Delta x} \\ &=: \frac{u_{i,j}^n - u_{i-1,j}^n}{\Delta x} \\ \partial_y u(i\Delta x, j\Delta y, n\Delta t) &\approx \frac{u(i\Delta x, j\Delta y, n\Delta t) - u(i\Delta x, (j-1)\Delta y, n\Delta t)}{\Delta y} \\ &=: \frac{u_{i,j}^n - u_{i,j-1}^n}{\Delta y}.\end{aligned}$$

Replacing the derivatives in (7.1) by the differences leads to the difference equation

$$L^h(u) := \frac{u_{i,j}^{n+1} - u_{i,j}^n}{\Delta t} + \frac{u_{i,j}^n - u_{i-1,j}^n}{\Delta x} + \frac{u_{i,j}^n - u_{i,j-1}^n}{\Delta y} \approx 0,$$

where the superscript h indicates a discrete value (very often one encounters $\Delta x = \Delta y$ and calls this value simply h). Since we decided to approximate the spatial derivatives at time level $n\Delta t$ the resulting difference equation is **explicit**, i.e. the solution at the resulting time level $(n+1)\Delta t$ is explicitly computable through

$$u_{i,j}^{n+1} \approx u_{i,j}^n - \frac{\Delta t}{\Delta x} (u_{i,j}^n - u_{i-1,j}^n) - \frac{\Delta t}{\Delta y} (u_{i,j}^n - u_{i,j-1}^n). \quad (7.2)$$

If we had decided to take the spatial differences at $t = (n+1)\Delta t$ we would have ended up with an **implicit** equation, requiring the solution of a system of linear equations at each time step. Note that in (7.2) we still have used the exact solution u of (7.1). Since there will be errors due to the replacement of differential quotients by difference quotients we can not assume (7.2) to hold for u . Thus, (7.2) should be rewritten as

$$U_{i,j}^{n+1} = U_{i,j}^n - \frac{\Delta t}{\Delta x} (U_{i,j}^n - U_{i-1,j}^n) - \frac{\Delta t}{\Delta y} (U_{i,j}^n - U_{i,j-1}^n), \quad (7.3)$$

and (7.3) now defines an approximation $U_{i,j}^n$ to $u_{i,j}^n$ within the bounds of the discretization and other errors. Note further that the representation of the numerical solution within the FDM framework is **pointwise**, i.e. we know pointwise values of approximations of the exact solution.

We conclude with a philosophical remark:

In FDMs one **discretizes the operators** occurring in the differential equation. The discrete solution $U^h := \left(U_{i,j}^n \mid i, j \in \mathbb{Z} ; n \in \mathbb{N} \right)$ consists of point values at the grid points of $\mathbb{G}^n \times \mathbb{G}_{i,j}$.

In the FEM one always starts from a weak (variational) formulation like (2.2). The space $\Omega := \mathbb{R}^2 \times \{t \in \mathbb{R} \mid t \geq 0\}$ is divided by means of a grid into certain **elements**. These elements may be chosen as the space-time *bricks* building the finite difference grid $\mathbb{G}^n \times \mathbb{G}_{i,j}$ but can be far more general. Due to their simplicity and their ability to resolve even complex geometries, triangulations are often used where the elements are triangles or tetrahedra. Applying the notion of weak solutions and choosing a space W of test functions leads to

$$\forall \Phi \in W : \int_{\Omega} \{u \partial_t \Phi + u \partial_x \Phi + u \partial_y \Phi\} dt dx dy = 0 .$$

Replacing the test function space by a finite dimensional subspace W^h leads to

$$\forall \Phi^h \in W^h : \int_{\Omega} \left\{ u \partial_t \Phi^h + u \partial_x \Phi^h + u \partial_y \Phi^h \right\} dt dx dy = 0 .$$

This is a set of so many equations as the dimension of $W^h := \mathbf{span}\{\Phi_1, \dots, \Phi_n\}$. A popular choice is the space of polynomials linear in t and x, y vanishing outside each element $\Sigma \subset \Omega$.

Choosing the approximation

$$U(x, y, t) := \sum_{i=1}^n U_i \Phi_i(x, y, t)$$

for u then leads to a linear system for the unknown coefficients U_i since

$$\begin{aligned}
 \forall j = 1, \dots, n : \quad & \int_{\Omega} \{U \partial_t \Phi_j + U \partial_x \Phi_j + U \partial_y \Phi_j\} dt dx = \\
 & \int_{\Omega} \left\{ \sum_{i=1}^n U_i \Phi_i \partial_t \Phi_j + \sum_{i=1}^n U_i \Phi_i \partial_x \Phi_j + \sum_{i=1}^n U_i \Phi_i \partial_y \Phi_j \right\} dt dx dy \\
 = \quad & \sum_{i=1}^n U_i \left(\int_{\Omega} \Phi_i \partial_t \Phi_j dt dx dy \right) + \sum_{i=1}^n U_i \left(\int_{\Omega} \Phi_i \partial_x \Phi_j dt dx dy \right) \\
 & + \sum_{i=1}^n U_i \left(\int_{\Omega} \Phi_i \partial_y \Phi_j dt dx dy \right) = \\
 = \quad & \sum_{i=1}^n \left(\int_{\Omega} \Phi_i (\partial_t \Phi_j + \partial_x \Phi_j + \partial_y \Phi_j) dt dx dy \right) U_i = 0 .
 \end{aligned}$$

Now, given U_i on the old time level, the U_i on the new time level follow from solving this linear system.

FEMs of the type just described are usually called space-time FEMs. It is often convenient to discretize in space only with finite elements and tackle the time derivative in a finite difference manner. In order to remove the coupling of every element with any other element one can use **discontinuous Galerkin methods**, where discontinuities are allowed at element faces. These types of methods are very close to finite volume methods, but not identical. Also popular are streamline diffusion FEMs which belong to the class of **Petrov-Galerkin** schemes. Petrov-Galerkin methods are characterized by choosing a test space different from the Ansatz space in which the discrete solution is sought. Although some mathematical analysis is available for streamline diffusion methods the numerical results obtained so far can not cope with results obtained by FDMs.

Again we end up with a philosophical remark:

In FEMs one **discretizes the solution** of the differential equation. The discrete solution is made up of polynomials defined piecewise on the elements.

The FVMs as third class of methods share all good properties of the FDM as well as of FEMs. As with FEMs arbitrary cells in space can be used for the discretization, while all of the FD technology is applicable at the same time. Here, one starts with a

type of weak solution which stems directly from gas dynamics, where the conservative equations are formulated as integral balances on so-called control volumes. For our generic model this weak form results from integrating

$$\partial_t u + \partial_x u + \partial_y u = \partial_t u + \langle \nabla, \mathbf{u} \rangle = 0, \quad \mathbf{u} := (u, u)^T,$$

over a bounded control volume $\sigma \subset \mathbb{R}^2$ with boundary $\partial\sigma$ where the unit outer normal vector \mathbf{n} is defined almost everywhere. From Gauss' divergence theorem it then follows

$$\frac{d}{dt} \int_{\sigma} u \, dx \, dy + \oint_{\partial\sigma} \langle \mathbf{u}, \mathbf{n} \rangle \, ds = 0.$$

Now the **cell average** of u on σ is defined as

$$\bar{u}_{\sigma}(t) := \frac{1}{|\sigma|} \int_{\sigma} u(x, y, t) \, dx \, dy,$$

where $|\sigma|$ denotes the area of σ , so that the weak formulation of gas dynamics enrolls itself as being an evolution equation for cell averages

$$\frac{d}{dt} \bar{u}_{\sigma} = - \frac{1}{|\sigma|} \oint_{\partial\sigma} \langle \mathbf{u}, \mathbf{n} \rangle \, ds. \quad (7.4)$$

FVMs are derived from (7.4) by approximating the line integrals by a suitable Gauss formula. The product $\langle \mathbf{u}, \mathbf{n} \rangle$ is replaced by a so-called **numerical flux function** which is essentially a finite difference formula.

Again we conclude with a philosophical remark:

FVMs are discrete **evolution equations for cell averages**. There is freedom in representing the numerical solution.

Some people like to view FVMs as being part of the Petrov-Galerkin family of FEMs with piecewise constant trial functions. The central idea in FVMs is to **adapt FDMs to unstructured grids** like triangulations and to use the computational advantages from FDMs and FEMs.

There are even more discretization methods available than the three classes described here, but they may often be subsumed as particular instances of FDMs, FEMs and FVMs. A popular method in the computation of incompressible viscous flows, and now becoming more and more attractive for compressible fluid flow, is the class of

spectral discretizations. Here the Ansatz space consists of globally defined trigonometric or polynomial functions on the whole domain (or, as in the spectral element method, on macro-elements). Choosing a FEM-like Ansatz leads to **spectral Galerkin** methods. If the differential equation has to be satisfied pointwise at so-called collocation points one talks about **spectral collocation** or **pseudospectral** methods. These methods are very close to FDMs with large stencils.

7.2 The Finite Difference Calculus

Let us now study more intensively the structure of finite difference methods.

Even one century before the invention of calculus through Leibniz and Newton, the foundations were laid for a discrete calculus employing differences of functions. Napier's ¹¹⁰ and Briggs' ¹¹¹ invention of logarithms revolutionized naval navigation and necessitated interpolation formulae for computing intermediate values in tabulated data. It was the genius Thomas Harriot ¹¹² who founded the field of finite difference calculus. This field flourished in the days of Leibniz and Euler and came to enormous influence in the area of the numerical treatment of ordinary and partial differential equations.

To give the reader again an idea of the spirit of the finite difference calculus, we consider the **forward difference operator** Δ defined by

$$\Delta u(x) := u(x + \Delta x) - u(x) ,$$

where Δx denotes an increment different from zero. It is often convenient to introduce the **shift operator** S through

$$Su(x) := u(x + \Delta x)$$

so that $\Delta = S - I$, where I is the identity operator. Assuming a smooth function $x \rightarrow u(x)$, Taylor series expansion results in

$$u(x + \Delta x) = \sum_{\nu=0}^{\infty} \frac{1}{\nu!} \Delta x^{\nu} D^{\nu} u(x),$$

where D^{ν} is defined as $\underbrace{\frac{d}{dx} \circ \frac{d}{dx} \circ \cdots \circ \frac{d}{dx}}_{\nu\text{-times}}$ and $D^0 = I$.

¹¹⁰John Napier (1559 - 1617); Merchiston

¹¹¹Henry Briggs (1556 - 1630); London, Oxford

¹¹²Thomas Harriot (1560 - 1621); Oxford

Utilizing the shift operator, this is nothing but the representation

$$Su(x) = \left(\sum_{\nu=0}^{\infty} \frac{1}{\nu!} \Delta x^{\nu} D^{\nu} \right) u(x),$$

where now, formally, the term in parathensis is exactly $e^{\Delta x D}$, i.e. we have found the representation

$$S = e^{\Delta x D} \tag{7.5}$$

and hence

$$\Delta = e^{\Delta x D} - I. \tag{7.6}$$

Introducing the **central difference operator** δ by means of

$$\delta u(x) := u(x + \Delta x/2) - u(x - \Delta x/2),$$

a further Taylor series

$$u(x \pm \Delta x/2) = \sum_{\nu=0}^{\infty} \frac{1}{\nu!} (\pm \Delta x)^{\nu} \frac{1}{2^{\nu}} D^{\nu} u(x)$$

reveals

$$\delta u(x) = 2 \sum_{\nu=0}^{\infty} \frac{1}{(2\nu+1)!} \left(\frac{\Delta x}{2} D \right)^{2\nu+1} u(x),$$

i.e.

$$\delta = 2 \sinh \left(\frac{\Delta x}{2} D \right). \tag{7.7}$$

Inverting this equation gives a relation between the differential operator D with the central difference operator δ , namely

$$D = \frac{2}{\Delta x} \sinh^{-1} \frac{\delta}{2} = \frac{1}{\Delta x} \left(\delta - \frac{1^2}{2^2 3!} \delta^3 + \frac{1^2 3^2}{2^4 5!} \delta^5 - \dots \right) \tag{7.8}$$

where $\delta^n := \underbrace{\delta \circ \delta \circ \dots \circ \delta}_{n\text{-times}}$. Equation (7.8) is an **exact difference formula** for the operator D .

More equations like (7.8) and (7.6) can be constructed.

We are mainly interested in the applicability of formulae like (7.8). To get an idea, we consider the simple linear transport equation $\partial_t u + \partial_x u = 0$ which we rewrite as

$$\partial_t u = L(D)u := -\partial_x u,$$

where the differential operator D is now partial derivation with respect to x . From (7.5), we deduce

$$u(x, t + \Delta t) = \left(e^{\Delta t \partial_t} \right) u(x, t)$$

and, since $\partial_t = L(D)$, we conclude

$$u_i^{n+1} = \left(e^{\Delta t L(D)} \right) u_i^n. \quad (7.9)$$

We now use (7.8) to eliminate D in terms of δ in (7.9). This gives

$$u_i^{n+1} = \left(e^{\Delta t L(2/\Delta x \sinh^{-1}(\delta/2))} \right) u_i^n. \quad (7.10)$$

Note that (7.10) is not an approximate difference equation but an exact one. Retaining only the linear terms in the exponential function gives

$$U_i^{n+1} = \left(I + \Delta t L(2/\Delta x \sinh^{-1}(\delta/2)) \right) U_i^n.$$

Retaining also only linear terms in the expansion (7.8) and noting that $L(D) = -D = -\partial_x$, leads finally to

$$U_i^{n+1} = \left(I - \Delta t \frac{\delta}{\Delta x} \right) U_i^n.$$

Expanding the central difference operator gives a finite difference method

$$U_i^{n+1} = U_i^n - \frac{\Delta t}{\Delta x} \left(U_{i+1/2}^n - U_{i-1/2}^n \right) \quad (7.11)$$

for the transport equation.

Hence, the finite difference calculus may be directly used to construct finite difference methods for partial differential equations. More constructions of this type can be found in the literature.

Now let the above difference scheme act on the true solution u . Employing again Taylor series expansions

$$\begin{aligned} u_i^{n+1} &= u(x, t) + \Delta t \partial_t u(x, t) + \mathcal{O}(\Delta t^2) \\ u_{i+1/2}^n &= u(x, t) + \frac{\Delta x}{2} \partial_x u(x, t) + \frac{1}{2!} \frac{\Delta x^2}{2^2} \partial_x^2 u(x, t) + \mathcal{O}(\Delta x^3) \\ u_{i-1/2}^n &= u(x, t) - \frac{\Delta x}{2} \partial_x u(x, t) + \frac{1}{2!} \frac{\Delta x^2}{2^2} \partial_x^2 u(x, t) + \mathcal{O}(\Delta x^3) \end{aligned}$$

and inserting them in (7.11) gives

$$u + \Delta t \partial_t u + \mathcal{O}(\Delta t^2) = u - \frac{\Delta t}{\Delta x} (\Delta x \partial_x u + \mathcal{O}(\Delta x^3))$$

which, after rearranging, leads to

$$\partial_t u + \partial_x u = \mathcal{O}(\Delta t, \Delta x^2). \quad (7.12)$$

Thus, our simple difference method (7.11) is of **order one in time** and of **order two in space**. This notion of order is usually called **truncation error** since it results from truncating Taylor series.

Note that we did not say a word about the convergence of (7.11). The property that (7.12) formally tends to $\partial_t u + \partial_x u = 0$ for $\Delta t, \Delta x \rightarrow 0$ is called **consistency**.

If a difference operator is applied to a sequence $(U_i^n \mid i \in \mathbb{Z})$ then it acts on the sequence index, i.e. $\Delta_i U_i^n := \Delta U_i^n = U_{i+1}^n - U_i^n$.

For later reference we introduce the **antidifference operator** Δ^{-1} as follows. If $\Delta u_i = g_i$ holds, then we define

$$u_i = \Delta^{-1} g_i + c,$$

where c denotes an arbitrary constant. The constant appears since after application of Δ we get back $\Delta u_i = g_i$. We first prove

$$\sum_{i=n_0}^{n-1} \Delta u_i = u_n - u_{n_0}, \quad (7.13)$$

$$\Delta \left(\sum_{i=n_0}^{n-1} u_i \right) = u_n. \quad (7.14)$$

Equation (7.13) simply follows from

$$\sum_{i=n_0}^{n-1} \Delta u_i = (u_{n_0+1} - u_{n_0}) + (u_{n_0+2} - u_{n_0+1}) + \dots + (u_{n-1} - u_{n-2}) + (u_n - u_{n-1}),$$

while (7.14) is nothing but

$$\Delta \left(\sum_{i=n_0}^{n-1} u_i \right) = \sum_{i+1=n_0}^{n-1} u_{i+1} - \sum_{i=n_0}^{n-1} u_i = \sum_{i=n_0-1}^{n-1} u_{i+1} - \sum_{i=n_0}^{n-1} u_i .$$

Using formula (7.14), we can immediately deduce

$$\Delta^{-1} u_n = \sum_{i=n_0}^{n-1} u_i + c , \quad (7.15)$$

where n_0 is an arbitrary start index.

We now give a discrete analogue of the product rule, namely

$$\Delta(u_i v_i) = S u_i \Delta v_i + v_i \Delta u_i . \quad (7.16)$$

The proof consists in evaluating the right and left hand side to see that they coincide.

Writing the product rule in the form $v_i \Delta u_i = \Delta(u_i v_i) - u_{i+1} \Delta v_i$ and applying the operator Δ^{-1} results in $\Delta^{-1}(v_i \Delta u_i) = \Delta^{-1} \Delta(u_i v_i) - \Delta^{-1}(u_{i+1} \Delta v_i)$. Following formula (7.15), we finally get the formula for **summation by parts**, which is a discrete analogue to the formula of integration by parts, namely

$$\sum_{i=n_0}^{n-1} v_i \Delta u_i = u_n v_n - \sum_{i=n_0}^{n-1} u_{i+1} \Delta v_i + c . \quad (7.17)$$

7.3 The CFL Condition

In 1928 Courant¹¹³, Friedrichs¹¹⁴ and Hans Lewy published a milestone paper (cf. Section 6.1) which was forgotten in war times and became famous when it was rediscovered during the Manhattan project.

¹¹³Richard Courant (1888-1972); Göttingen, New York

¹¹⁴Kurt Otto Friedrichs (1901-1982); Brunswick, New York

As already emphasized in the last chapter, Courant's interests at that time mainly concerned the proof of existence of solutions to partial differential equations. In that field he was one of the founders of the idea to employ finite difference schemes as approximations to partial differential equations. If he could prove that the difference scheme is consistent and convergent, and furthermore that the limit solution (as Δt and Δx tend to zero) satisfies the original differential equation, then he would have proven existence of a solution of the partial differential equation. Although Courant, Friedrichs and Lewy succeeded with this scheme for the (elliptic) Laplace equation, they encountered serious problems with the (hyperbolic) wave equation. Finally they succeeded in formulating the celebrated CFL condition which is necessary for convergence.

In order to demonstrate the basic behaviour in the context under consideration, let us consider the simple forward explicit difference scheme for the slightly more general transport equation $\partial_t u + a \partial_x u = 0$, $a < 0$, namely

$$U_i^{n+1} = U_i^n - a \frac{\Delta t}{\Delta x} (U_{i+1}^n - U_i^n) = \left(1 + a \frac{\Delta t}{\Delta x}\right) U_i^n - a \frac{\Delta t}{\Delta x} U_{i+1}^n. \quad (7.18)$$

(7.18) is an explicit three-point scheme for the explicit computation of U_i^{n+1} . Viewed in the (x, t) -plane we can identify the **difference stencil** of (7.18) as shown in Fig. 30.

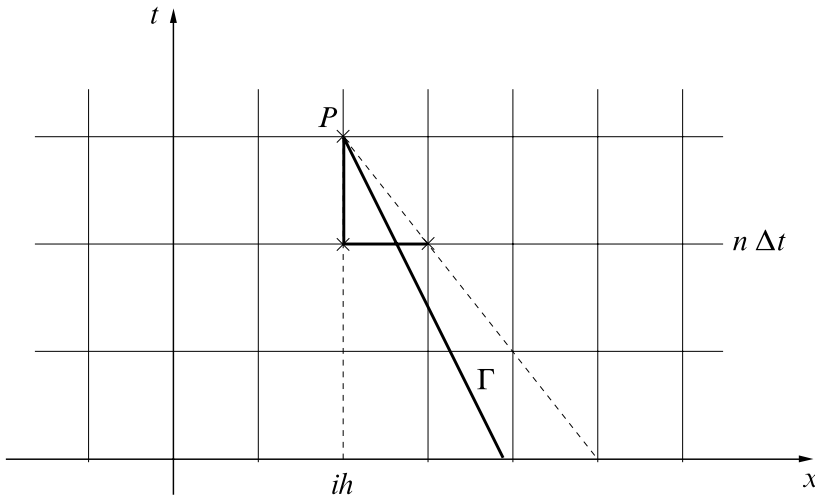


Figure 30: Difference stencil and characteristics

The stencil gives rise to the definition of **numerical characteristics** which are straight lines enveloping the stencil. The numerical characteristics cut an interval off the axis

$t = 0$ which is called **numerical domain of dependence**. Changing the initial datum in this interval changes the numerical solution at point P while perturbations outside this interval would not affect the numerical solution at P . However, there is also the characteristic Γ of the differential equation we would like to approximate. If Γ lies outside the numerical characteristics as shown in Fig. 30, then changing the initial datum between Γ and the numerical characteristic would change the true solution at P but not the numerical solution. Thus, convergence of the difference scheme is not possible. Keeping Γ inside the numerical characteristics ensures that all perturbations of the initial datum which affect the true solution at P will always affect the numerical solution at P , too.

The characteristic Γ has slope

$$\frac{dt}{dx} = \frac{1}{a}$$

and the right numerical characteristic has slope $-\Delta t/\Delta x$. Thus, necessary for convergence is $\Delta t/\Delta x \leq 1/|a|$ which is already the Courant-Friedrichs-Lewy (CFL) condition

$$\Delta t \leq \frac{\Delta x}{|a|}. \quad (7.19)$$

In case of $a > 0$, and replacing the forward difference equation by the backward one, (7.19) again ensures the suitable behaviour of the domains of dependence. This dependence of the type of the difference equation on the sign of a , i.e. on the direction of the movement of the wave, leads to so-called **upwind schemes**.

In the case of a scalar conservation law $\partial_t u + \partial_x f(u) = 0$, the same arguments hold for the quasilinear form $\partial_t u + f'(u)\partial_x u = 0$. Hence, instead of (7.19), explicit finite difference schemes with three-point stencil have to satisfy the **CFL condition**

$$\frac{\Delta t}{\Delta x} \max_i |f'(U_i^n)| \leq 1 \quad (7.20)$$

in each time step n .

7.4 Lax-Richtmyer Theory

We now formalize the treatment of finite difference schemes supplementing chapter 6 in order to present at least the historical central idea of a convergence theory for this particular type of schemes.

Although we have to confine ourselves with linear problems, the Lax-Richtmyer theory allows to gain deep insight into the interplay between the notions of consistency, stability and convergence. Originally, it was formulated for autonomous, linear initial value problems only but was later generalized by several authors so that also non-autonomous and nonlinear problems could be included¹¹⁵.

We consider an **abstract evolution equation** in a Banach space B , where a linear operator $A : D_A \rightarrow B$ acts upon a subspace D_A of B . The evolution equation we want to discretize is then given as

$$\begin{aligned} \frac{d}{dt}u(t) &= A(u(t)), \quad t \in [0, T], \quad T > 0 \\ u(0) &= u_0, \quad u_0 \in D_A. \end{aligned} \quad (7.21)$$

A one-parameter family $u(t)$ with $t \in [0, T]$, $u(t) \in D_A$, is called a **genuine solution** of (7.21), if

$$\lim_{\substack{\Delta t \rightarrow 0 \\ 0 \leq t \leq T}} \left\| \frac{u(t + \Delta t) - u(t)}{\Delta t} - Au(t) \right\|_B = 0. \quad (7.22)$$

Let D be the set of all initial data $u_0 \in B$ so that a unique genuine solution exists with $u(0) = u_0$ and such that convergence in (7.22) is uniform in t ¹¹⁶. Then there is a one-parameter family of linear operators $E_0(t) : D \rightarrow B$ defined by

$$u(t) = E_0(t)u_0$$

which are called **evolution operators** or **solution operators**. We always assume that our problem is properly posed, which in particular implies that $E_0(t)$ is bounded uniformly in t , i.e. there is a constant $K > 0$ such that $\|E_0(t)\|_{D \rightarrow B} < K$ for $0 \leq t \leq T$. Then the Hahn-Banach extension theorem implies the existence of a **generalized solution operator** $E(t) : B \rightarrow B$ with the same bound K as $E_0(t)$. The generalized solution operator has the **semigroup property**

$$\forall s, t \geq 0 : \quad E(s + t) = E(s) E(t)$$

and the closure of the operator A which we also denote by A is the **infinitesimal generator** of this semigroup. One can further show that $E(t)$ commutes with A , so

¹¹⁵cf. R. Ansorge: Survey of equivalence theorems in the theory of difference approximations for partial initial value problems. In: *Topics in Numerical Analysis* (J.J. Miller, ed.), London-New York-San Francisco: Academic Press 1977

¹¹⁶We shall see in a minute that uniform convergence is automatic.

that

$$E(t)Au_0 = AE(t)u_0.$$

Then (7.21) can be written as

$$\lim_{\substack{\Delta t \rightarrow 0 \\ 0 \leq t \leq T}} \left\| E(t) \left(\frac{E(\Delta t) - I}{\Delta t} - A \right) u_0 \right\|_B = 0,$$

which, since $\|E(t)\|_B < K$ for $0 \leq t \leq T$, implies the uniform convergence of (7.22) with respect to t .

In order to discuss fairly general finite difference schemes, let us introduce linear **difference operators**

$$\Lambda_{0/1} = \Lambda_{0/1}(\Delta t, \Delta x_1, \dots, \Delta x_d)$$

and the class of difference schemes

$$\Lambda_1 U^{n+1} = \Lambda_0 U^n.$$

We assume the existence of Λ_1^{-1} and that $\Lambda_1^{-1}\Lambda_0$ is a bounded linear operator defined in all of B . From the CFL condition we know that Δt and the Δx_i can not go to zero independently of each other as far as explicit methods are concerned. Hence we assume relations $\Delta x_i = \varphi_i(\Delta t)$, $i = 1, \dots, d$, and introduce the **discrete evolution operator**

$$C(\Delta t) := \Lambda_1^{-1}(\Delta t, \varphi_1(\Delta t), \dots, \varphi_d(\Delta t)) \cdot \Lambda_0(\Delta t, \varphi_1(\Delta t), \dots, \varphi_d(\Delta t))$$

so that the class of difference schemes under consideration is

$$U^{n+1} = C(\Delta t)U^n. \quad (7.23)$$

The family of operators $C(\Delta t)$ is said to define a **consistent** approximation for the initial value problem if for every $u(t)$ in a class of genuine solutions with initial data $u(0)$ dense in B the **consistency condition**

$$\lim_{\Delta t \rightarrow 0} \lim_{0 \leq t \leq T} \left\| \left(\frac{C(\Delta t) - I}{\Delta t} - A \right) u(t) \right\|_B = 0 \quad (7.24)$$

holds. Since $u(t)$ denotes a genuine solution we can use the fact that due to (7.22) $Au(t)$ is close to $(u(t + \Delta t) - u(t))/\Delta t$ to give the alternative definition

$$\lim_{\Delta t \rightarrow 0} \lim_{0 \leq t \leq T} \left\| \frac{u(t + \Delta t) - C(\Delta t)u(t)}{\Delta t} \right\|_B = 0. \quad (7.25)$$

The quantity

$$\left\| \frac{u(t + \Delta t) - C(\Delta t)u(t)}{\Delta t} \right\|_B := \hat{\varepsilon}(\Delta t, u_0)$$

is called **truncation error**, and if this error is of order p , i.e.

$$\hat{\varepsilon}(\Delta t, u_0) = \mathcal{O}(\Delta t^p),$$

the method is called a **method of order p** . Consistency therefore means

$$\hat{\varepsilon}(\Delta t, u_0) = o(1)$$

which is particularly fulfilled for $p > 0$.

Note that this abstract notion is in complete agreement with our earlier calculations in (7.12).

Consistency guarantees that in the limit $\Delta t \rightarrow 0$ the difference scheme formally reproduces the differential equation. The notion of convergence is concerned with the limit behaviour of the discrete solutions. **Consistency does not necessarily already imply convergence!**

Due to (7.23), iterated application of $C(\Delta t)$ gives

$$U^n = C(\Delta t)U^{n-1} = C(\Delta t) \circ C(\Delta t)U^{n-2} = \dots = C(\Delta t)^n U^0 = C(\Delta t)^n u_0,$$

and this discrete function is expected to be a good approximation to $u(n\Delta t) = E(n\Delta t)u_0$. The family of operators $C(\Delta t)$ provides a **convergent** approximation to the initial value problem if, for fixed $t \in [0, T]$,

$$\lim_{j \rightarrow \infty} \|C(\Delta_j t)^{n_j} u_0 - E(t)u_0\|_B = 0 \tag{7.26}$$

holds for each sequence $(\Delta_j t)_{j=1,2,\dots}$ tending to zero where n_j is chosen to satisfy $\lim_{j \rightarrow \infty} n_j \Delta_j t = t$.

In reaching the discrete solution at time $t = n\Delta t$, the operator $C(\Delta t)$ has to act n times on the initial datum. Thinking of small errors in u_0 or errors due to rounding during the computation, we would not be satisfied with a difference scheme which would amplify these small perturbations until they spoil the solution. The central idea of stability is therefore to put bounds on the possible amplification of initial data. The operator $C(\Delta t)$ is said to define a **stable** approximation if for some $t^* > 0$ the infinite set of operators

$$C(\Delta t)^n, \quad 0 < \Delta t < t^*, \quad 0 \leq n\Delta t \leq T,$$

are uniformly bounded, i.e. if there exists a constant κ with

$$\|C(\Delta t)^n\|_B \leq \kappa, \quad \forall n \in \mathbb{N}, \forall \Delta t$$

with $0 < \Delta t \leq t^*$ and with $0 \leq n\Delta t \leq T$.

With the notions of consistency, stability and convergence we can now formulate the central **Equivalence Theorem** of the Lax-Richtmyer theory:

Theorem 7.1 : *A finite difference scheme approximating a properly posed initial value problem is convergent if and only if it is consistent and stable.*

PROOF (of the direction: *consistency* \wedge *stability* \Rightarrow *convergence*¹¹⁷):

$U(0) = u_0$ implies

$$U^n - u(t) = \sum_{\nu=1}^n C(\Delta t)^{\nu-1} \{ C(\Delta t) u(t_{n-\nu}) - u(t_{n-\nu+1}) \} + u(t_n) - u(t)$$

$(t_\mu = \mu \Delta t)$.

Thus, from the stability property, we obtain

$$\begin{aligned} \|U^n - u(t)\|_B &\leq \sum_{\nu=1}^n \|C(\Delta t)^{\nu-1}\|_B \cdot \|C(\Delta t) u(t_{n-\nu}) - u(t_{n-\nu+1})\|_B \\ &\quad + \|u(t_n) - u(t)\|_B \\ &\leq \kappa \sum_{\nu=1}^n \|C(\Delta t) E_0(t_{n-\nu}) u_0 - E_0(t_{n-\nu+1}) u_0\|_B \\ &\quad + \|u(t_n) - u(t)\|_B, \end{aligned}$$

leading to

$$\|U^n - u(t)\|_B \leq \kappa n \Delta t \cdot \hat{\varepsilon}(\Delta t, u_0) + \|u(t_n) - u(t)\|_B.$$

¹¹⁷From an engineering point of view, this is the more important direction. The complete proof can be found in the original paper of Lax and Richtmyer cited in footnote 92. If consistency and stability are defined properly, theorems of the Lax-Richtmyer type can be proven even in the case of nonlinear difference schemes for nonlinear equations, as already mentioned in footnote 115.

Taking $n \Delta t \leq T$ into account, we find

$$\|U^n - u(t)\|_B \leq \kappa T \hat{\varepsilon}(\Delta t, u_0) + \|u(n \Delta t) - u(t)\|_B .$$

Because of $n \Delta t \rightarrow t$, because of the continuity of u with respect to t , and because of the consistency condition (7.25), the truth of the assertion is obvious. \square

Particularly for $n \Delta t = t$, we get an estimate for the **global discretization error** $\|U^n - u(t)\|_B$, namely

$$\|U^n - u(t)\|_B \leq \kappa t \hat{\varepsilon}(\Delta t, u_0) .$$

Here, round off errors are not included!

Example:

Let us again treat the transport equation $\partial_t u + a \partial_x u = 0$ in the case of $a < 0$.

If then (7.18) will be expressed in terms of the Lax-Richtmyer theory, and if a relation between the step sizes Δt and Δx will be described, we find

$$\|C(\Delta t)\|_\infty \leq \left| 1 + a \frac{\Delta t}{\Delta x} \right| + |a| \frac{\Delta t}{\Delta x} .$$

The CFL condition (7.19) then ensures

$$\|C(\Delta t)\|_\infty \leq \left(1 + a \frac{\Delta t}{\Delta x} \right) + |a| \frac{\Delta t}{\Delta x} = 1 ,$$

hence

$$\|C^n(\Delta t)\|_\infty \leq \|C(\Delta t)\|_\infty^n \leq 1 .$$

Thus, the CFL condition guarantees the numerical stability of the scheme in the sense of the Lax-Richtmyer theory.

And also the other direction is true: Stability in the sense of Lax-Richtmyer requires the CFL condition to be fulfilled, and this holds independently of the norm.

Of course, the upwind scheme for $a > 0$ can be treated in the same way.

The upwinding idea can also be looked upon from the point of view of information theory: Looking into the direction of the traveling information it seems to be unreasonable to use for the computation of a future value of the solution at a given space position such values of the present time level whose positions lie downstream and were not yet influenced by the information.

7.5 The von Neumann Stability Criterion

Consistency of a finite difference scheme is easy to verify by means of Taylor expansions. In contrast, verifying stability using the pure definition may be much harder. In the case of linear initial value problems or problems with periodic boundary data there is a simple mechanism due to von Neumann¹¹⁸.

We follow start with a complex Fourier transform of a discrete function represented as a sequence $U^n = (U_j^n)_{j \in \mathbb{Z}}$ and defined at grid points only. The discrete Fourier transform is defined as

$$\widehat{U}^n(\xi) = \sum_{j \in \mathbb{Z}} U_j^n e^{-ij\xi}, \quad 0 \leq \xi < 2\pi.$$

The importance of the discrete Fourier transform in analyzing stability of finite difference schemes can be seen if it is applied to the shift operator S defined already in section 7.2. For our purposes, we refine the definition by calling $S_+ U^n := S(\Delta x) U^n := (U_{j+1}^n)_{j \in \mathbb{Z}}$ the **forward shift operator**. Then it follows from the definition of the discrete Fourier transform

$$\begin{aligned} \widehat{S_+ U^n}(\xi) &= \sum_{j \in \mathbb{Z}} U_{j+1}^n e^{-ij\xi} = \sum_{j \in \mathbb{Z}} U_j^n e^{-i(j-1)\xi} \\ &= \sum_{j \in \mathbb{Z}} U_j^n e^{-ij\xi} e^{i\xi} = e^{i\xi} \sum_{j \in \mathbb{Z}} U_j^n e^{-ij\xi} \\ &= e^{i\xi} \widehat{U}^n(\xi). \end{aligned}$$

Thus, applying the shift operator to a grid function is, in Fourier space, multiplication with $e^{i\xi}$. The function $e^{i\xi}$ is called the **symbol** of the shift operator S .

Completely analogously one gets for the **backward shift operator** $S_- U^n := S(-\Delta x) U^n$ the symbol $e^{-i\xi}$, i.e.

$$\widehat{S_- U^n}(\xi) = e^{-i\xi} \widehat{U}^n(\xi).$$

Since every linear finite difference scheme is made up of combinations of forward and backward shift operators the way is open to look at discrete Fourier transforms of difference schemes.

¹¹⁸John von Neumann (1903 - 1957) Berlin, Hamburg, Göttingen, Princeton

Example:

Consider again the difference scheme

$$U_i^{n+1} = U_i^n - \frac{\Delta t}{2\Delta x} a (U_{i+1}^n - U_{i-1}^n)$$

as an approximation to $\partial_t u + a\partial_x u = 0$. It is an easy exercise to see that the scheme is consistent of order 2 in space and order 1 in time. Written in terms of shift operators and sequences we get

$$U^{n+1} = \left(I - \frac{\Delta t}{2\Delta x} a (S_+ - S_-) \right) U^n .$$

Now, taking Fourier transform on both sides, it follows that

$$\widehat{U^{n+1}}(\xi) = \left(1 - \frac{\Delta t}{2\Delta x} a (e^{i\xi} - e^{-i\xi}) \right) \widehat{U^n}(\xi).$$

If the symbol $\Theta(\Delta t, \xi) := 1 - \frac{\Delta t}{2\Delta x} a (e^{i\xi} - e^{-i\xi})$ turns out to be larger than one in modulus, this means that any small perturbation in U^n will be amplified in the step computing U^{n+1} . In our case, since $e^{\pm i\xi} = \cos \xi \pm i \sin \xi$, we get

$$\Theta(\Delta t, \xi) = 1 - \frac{\Delta t}{\Delta x} a i \sin \xi .$$

Taking the modulus, this amounts to

$$|\Theta(\Delta t, \xi)|^2 = 1 + \left(\frac{\Delta t}{\Delta x} a \right)^2 \sin^2 \xi \geq 1$$

regardless how small the time step is chosen. The symbol is usually called the **amplification factor** of the difference scheme. The present scheme is unconditionally unstable, since the modulus of the amplification factor is larger than one except for $\xi \in \{0, \pi, 2\pi\}$. Assuming $a > 0$, the upwind difference scheme

$$U_i^{n+1} = U_i^n - \frac{\Delta t}{\Delta x} a (U_i^n - U_{i-1}^n)$$

results in the sequence formulation

$$U^{n+1} = \left(I - \frac{\Delta t}{\Delta x} a (I - S_-) \right) U^n$$

which gives the discrete Fourier transform

$$\widehat{U^{n+1}}(\xi) = \left(1 - \frac{\Delta t}{\Delta x} a (1 - e^{-i\xi}) \right) \widehat{U^n}(\xi) .$$

Let us define $K := \frac{\Delta t}{\Delta x} a$ and compute the amplification factor

$$\Theta(\Delta t, \xi) = \left| 1 - K(1 - e^{-i\xi}) \right| \leq |1 - K| + K \underbrace{|e^{-i\xi}|}_{=1}$$

$$\stackrel{\text{CFL condition}}{\leq} 1 - K + K = 1.$$

Hence, our upwind scheme is stable under the usual CFL condition.

We remark in passing that the same analysis holds in multiple space dimensions. Here, the discrete Fourier transform is defined as

$$\widehat{U}^n(\Xi) = \sum_{\mathbf{k} \in \mathbb{Z}^d} U_{\mathbf{k}}^n e^{i\mathbf{k}^T \Xi}$$

where d is the space dimension and $\Xi = (\xi_1, \dots, \xi_d)^T$.

In the case of systems of equations with n components in d space dimensions, amplification is given by

$$\widehat{U}^{n+1}(\Xi) = \Theta(\Delta t, \Xi) \widehat{U}^n(\Xi)$$

where now Θ is an $n \times n$ matrix, the **amplification matrix**. For reasons of generality we stay with the case of systems in the following.

Recursion reveals the relation

$$\widehat{U}^n(\Xi) = \Theta^n(\Delta t, \Xi) \widehat{U}^0(\Xi)$$

so that stability has to do with the properties of the iterated amplification matrix. In general, we can state the **necessary and sufficient stability criterion** that for some positive Δt^* the matrices (or, in the case of scalar equations, the amplification factor)

$$\Theta^n(\Delta t, \Xi)$$

should be uniformly bounded for $0 < \Delta t < \Delta t^*, 0 \leq n\Delta t \leq T$ and all $\Xi \in \mathbb{Z}^d$. If $\rho_\theta(\Delta t, \Xi)$ denotes the spectral radius of Θ (i.e. the maximum of the moduli of the eigenvalues), then it follows from linear algebra that

$$\rho_\theta^n(\Delta t, \Xi) \leq \|\Theta^n(\Delta t, \Xi)\| \leq \|\Theta(\Delta t, \Xi)\|^n.$$

A necessary condition for stability is therefore the existence of a constant C so that

$$\rho_\theta^n(\Delta t, \Xi) \leq C$$

for $0 < \Delta t < \Delta t^*$, $0 \leq n\Delta t \leq T$ and all $\Xi \in \mathbb{Z}^d$. Without loss of generality we assume $C \geq 1$ and therefore

$$\rho_\theta(\Delta t, \Xi) \leq C^{\frac{1}{n}}, \quad 0 \leq n \leq \frac{T}{\Delta t}.$$

Thus, in particular, we get

$$\rho_\theta(\Delta t, \Xi) \leq C^{\frac{\Delta t}{T}}.$$

Now the expression $C^{\Delta t/T}$ is bounded for $0 < \Delta t < \Delta t^*$ by a linear function $1 + \text{const } \Delta t$.

Thus, a necessary condition for stability is the celebrated **von Neumann condition**

$$|\lambda_i| \leq 1 + \mathcal{O}(\Delta t)$$

for $0 < \Delta t < \Delta t^*$, all $\Xi \in \mathbb{Z}^d$ and $i = 1, \dots, n$, where λ_i denotes the i -th eigenvalue of $\Theta(\Delta t, \Xi)$.

It should be noted that there are several other stability conditions like the famous Kreiss matrix theorem or the much less known criterion of Buchanan¹¹⁹.

7.6 The Modified Equation

Finally, a tool should briefly be explained which can often help to analyse finite difference procedures.

To every finite difference approximation of order $\mathcal{O}(\Delta t^p, \Delta x^q)$ of a given differential equation there is another differential equation to which the difference scheme is a better approximation.

In order to demonstrate this remark, let us start again from the simple upwind scheme

$$U_i^{n+1} = U_i^n - \frac{\Delta t}{\Delta x} a (U_i^n - U_{i-1}^n) \quad (7.27)$$

for the transport equation

$$\partial_t u + a \partial_x u = 0, \quad a > 0. \quad (7.28)$$

¹¹⁹For a thorough treatment see: Richtmyer, R.D. and K.W. Morton, Difference methods for initial-value problems, 2nd ed., New York, London, Sydney: Interscience Publishers 1967

Inserting the Taylor series expansions of a sufficiently smooth function u , i.e.

$$\begin{aligned} u_i^{n+1} &= u + \Delta t \partial_t u + \frac{\Delta t^2}{2} \partial_t^2 u + \frac{\Delta t^3}{6} \partial_t^3 u + \mathcal{O}(\Delta t^4) \\ u_{i-1}^n &= u - \Delta x \partial_x u + \frac{\Delta x^2}{2} \partial_x^2 u - \frac{\Delta x^3}{6} \partial_x^3 u + \mathcal{O}(\Delta x^4), \end{aligned}$$

into the upwind scheme results in

$$\begin{aligned} \partial_t u + \frac{\Delta t}{2} \partial_t^2 u + \frac{\Delta t^2}{6} \partial_t^3 u + \mathcal{O}(\Delta t^3) &= -a \partial_x u + a \frac{\Delta x}{2} \partial_x^2 u - a \frac{\Delta x^2}{6} \partial_x^3 u \\ &\quad + \mathcal{O}(\Delta x^3), \end{aligned}$$

or, after rearrangement of terms, in

$$\begin{aligned} \partial_t u + a \partial_x u &= -\frac{\Delta t}{2} \partial_t^2 u + \frac{a \Delta x}{2} \partial_x^2 u - \frac{\Delta t^2}{6} \partial_t^3 u - \frac{a \Delta x^2}{6} \partial_x^3 u \\ &\quad + \mathcal{O}(\Delta t^3, \Delta x^3). \end{aligned} \tag{7.29}$$

We do now want to replace the time derivatives on the right hand side by space derivatives. In order to do so, we are **not** allowed to use the differential equation (7.28), since our u does not necessarily satisfy this equation but equation (7.29)! Therefore, we have to accomplish our task using (7.29) alone. We take the time derivative of (7.29), i.e.

$$\begin{aligned} \partial_t^2 u + a \partial_t \partial_x u &= -\frac{\Delta t}{2} \partial_t^3 u + \frac{a \Delta x}{2} \partial_t \partial_x^2 u - \frac{\Delta t^2}{6} \partial_t^4 u \\ &\quad + \frac{a \Delta x^2}{6} \partial_t \partial_x^3 u + \mathcal{O}(\Delta t^3, \Delta x^3), \end{aligned}$$

and then the space derivative of (7.29) which we multiply by a ,

$$\begin{aligned} a \partial_t \partial_x u + a^2 \partial_x^2 u &= -a \frac{\Delta t}{2} \partial_t^2 \partial_x u + \frac{a^2 \Delta x}{2} \partial_x^3 u - \frac{a \Delta t^2}{6} \partial_t^3 \partial_x u \\ &\quad - \frac{a^2 \Delta x^2}{6} \partial_x^4 u + \mathcal{O}(\Delta t^3, \Delta x^3). \end{aligned}$$

Subtracting the last equation from the first one gives

$$\begin{aligned} \partial_t^2 u &= a^2 \partial_x^2 u + \Delta t \left(-\frac{1}{2} \partial_t^3 u + \frac{a}{2} \partial_t^2 \partial_x u \right) + \Delta x \left(\frac{a}{2} \partial_t \partial_x^2 u - \frac{a^2}{2} \partial_x^3 u \right) \\ &\quad + \mathcal{O}(\Delta t^3, \Delta x^3). \end{aligned}$$

Some further similar manipulations of (7.29) lead to

$$\begin{aligned}\partial_t^3 u &= -a^3 \partial_x^3 u + \mathcal{O}(\Delta t, \Delta x), \\ \partial_x \partial_t^2 u &= a^2 \partial_x^3 u + \mathcal{O}(\Delta t, \Delta x), \\ \partial_t \partial_x^2 u &= -a \partial_x^3 u + \mathcal{O}(\Delta t, \Delta x).\end{aligned}$$

Substituting these expressions into (7.29) yields the **modified equation**

$$\begin{aligned}\partial_t u + a \partial_x u &= \frac{a \Delta x}{2} \left(1 - \frac{\Delta t}{\Delta x}\right) \partial_x^2 u - \frac{a \Delta x^2}{6} \left(2 \frac{\Delta t^2}{\Delta x^2} - 3 \frac{\Delta t}{\Delta x} + 1\right) \partial_x^3 u \\ &\quad + \mathcal{O}(\Delta x^3, \Delta x^2 \Delta t, \Delta x \Delta t^2, \Delta t^3).\end{aligned}$$

What can be seen from this equation? First of all we see that the truncation error of our difference scheme applied to the transport equation (7.28) is $\mathcal{O}(\Delta t, \Delta x)$, since the first term on the right hand side of the modified equation is

$$\frac{a}{2} \Delta x \partial_x^2 u - \frac{a}{2} \Delta t \partial_x^2 u$$

which goes to zero linearly with time and space increments. Moreover, our upwind scheme is a good approximation to the parabolic differential equation

$$\partial_t u + a \partial_x u = \varepsilon(a, \Delta t) \partial_x^2 u$$

with positive (due to the CFL condition) diffusion coefficient. Since such equations have smooth solutions, we expect our upwind scheme to be well-behaved.

Many further properties of the difference scheme can be deduced from the modified equation, e.g. numerical stability.

7.7 Difference Schemes in Conservation Form

It seems to be obvious that numerical procedures can only be expected to lead to good approximations of the solution of the original problem if as many properties known in advance as possible are imitated by the scheme.

One of these reproductions, already mentioned in chapter 6, is the use of so-called methods of conservation form in order to treat conservation law equations numerically.

This means that the system (2.1) of conservation laws, i.e.

$$\partial_t \mathbf{V} + \partial_x \mathbf{f}(\mathbf{V}) = \partial_t \mathbf{V} + \mathbf{J} \mathbf{f}(\mathbf{V}) \partial_x \mathbf{V} = 0 \quad (7.30)$$

is approximated by an explicit $(2k + 1)$ -step finite difference scheme

$$\frac{1}{\Delta t} (\mathbf{V}_j^{n+1} - \mathbf{V}_j^n) + \frac{1}{\Delta x} (\mathbf{g}_{j+\frac{1}{2}}^n - \mathbf{g}_{j-\frac{1}{2}}^n) = 0 \quad (7.31)$$

with a consistent numerical flux

$$\mathbf{g}_{j+\frac{1}{2}}^n := \mathbf{g}(\mathbf{V}_{j-k+1}^n, \mathbf{V}_{j-k}^n, \dots, \mathbf{V}_j^n, \mathbf{V}_{j+1}^n, \dots, \mathbf{V}_{j+k}^n) \quad (7.32)$$

and not –e.g.– by a method like

$$\frac{1}{\Delta t} (\mathbf{V}_j^{n+1} - \mathbf{V}_j^n) + \mathbf{J} \mathbf{f}(\mathbf{V}_j^n) \frac{1}{2\Delta x} (\mathbf{V}_{j+1}^n - \mathbf{V}_{j-1}^n) = 0.$$

An example of a method of type (7.31) is the 3-point **Lax-Friedrichs Scheme**¹²⁰

$$\mathbf{V}_j^{n+1} - \frac{1}{2} (\mathbf{V}_{j+1}^n + \mathbf{V}_{j-1}^n) + \frac{1}{2} \frac{\Delta t}{\Delta x} \{ \mathbf{f}(\mathbf{V}_{j+1}^n) - \mathbf{f}(\mathbf{V}_{j-1}^n) \} = 0 \quad (7.33)$$

with

$$\mathbf{g}_{j+\frac{1}{2}}^n = -\frac{1}{2} \frac{\Delta x}{\Delta t} (\mathbf{V}_{j+1}^n - \mathbf{V}_j^n) + \frac{1}{2} \{ \mathbf{f}(\mathbf{V}_{j+1}^n) + \mathbf{f}(\mathbf{V}_j^n) \}.$$

If the numerical vectors \mathbf{V}_j^n in (7.33) are replaced by an exact smooth solution $\mathbf{V}(x_j, t_n)$ of (7.30), the left side of (7.33) differs from zero, but its Taylor expansion then leads to a truncation error of order $\mathcal{O}(\Delta t)$ in the sense of the Lax-Richtmyer theory provided that $\lambda = \frac{\Delta t}{\Delta x} = \text{const}$ is prescribed.

Hence, (7.33) is a method of conservation form of order 1. It can be shown that the method is monotone in the meaning of Section 6.4, and, as a matter of fact, **all monotone methods are only of order 1** as can be established by means of the ‘Modified Equation’ of Section 7.6.

But, of course, there are also methods of higher order. One of them is the 3-point **Lax-Wendroff Scheme**¹²¹. Its numerical flux reads in case of scalar 1-dimensional conservation laws of type (2.1) as

$$\mathbf{g}_{j+\frac{1}{2}}^n = \frac{1}{2} \left\{ f(v_{j+1}^n) + f(v_j^n) - \frac{1}{\lambda} (\mu_{j+\frac{1}{2}}^n)^2 (v_{j+1}^n - v_j^n) \right\} \quad (7.34)$$

¹²⁰Proc. Nat. Acad. Sci. USA **68** (1971), 1686-1688

¹²¹Comm. Pure Appl. Math. **13**, 217-237 (1960)

with

$$\mu_{j+\frac{1}{2}}^n = \lambda \frac{f(v_{j+1}^n) - f(v_j^n)}{v_{j+1}^n - v_j^n}, \quad \lambda = \frac{\Delta t}{\Delta x}.$$

By Taylor expansion, it can be verified that this is a method of order 2. Therefore, the scheme seems to be a useful one, at least as long as the smoothness assumptions necessary for the Taylor expansion are guaranteed in a neighborhood of (x_j, t_n) .

But if the exact weak solution shows a shock at a certain position, the numerical solution begins to oscillate close to this position though the scheme respects the conservation character of the original problem.

But what is not respected by the method is the TVD-property of the solutions of the original problem.

Numerical solutions produced by means of TVD-methods do not begin to oscillate because this would lead to an increase of the total variation.

Several attempts were made in the literature to overcome the problem of the low order of monotone methods and the demand for keeping the advantages of higher order methods, e.g. the use of greater step sizes without loss of accuracy in areas where the exact solution is smooth.

One of these ideas, e.g. in case of scalar problems, consists in the construction of so-called **flux limiter methods**:

Let g_H be the numerical flux of a higher order method, e.g. of the Lax-Wendroff scheme, and let g_L be the numerical flux of a low order procedure, e.g. a TVD-scheme.

If $g_H(v_{j+k-1}^n, \dots, v_{j+k}^n)$ is abbreviated by $g_H(v^n; j)$ and analogously $g_L(v^n; j)$, let us create a new method by

$$g(v^n; j) := g_L(v^n; j) + \varphi(v^n; j) \{g_H(v^n; j) - g_L(v^n; j)\}. \quad (7.35)$$

Herewith, φ –called the **limiter**– has to be chosen in such a way that $\varphi \approx 1$ in areas where the exact entropy solution is expected to be smooth, and $\varphi \approx 0$ close to shocks. In other words, φ has to be a sensor for sudden strong increase or decrease of the unknown exact solution expected to occur at positions where the numerical solution increases or decreases strongly with respect to space, e.g.

$$\varphi = \varphi \left(\frac{v_j^n - v_{j-1}^n}{v_{j+1}^n - v_j^n} \right) \quad (7.36)$$

where φ is a sufficiently smooth and bounded function with $\varphi(1) = 1$ and $\varphi(0) = 0$. Of course, difficulties can still make trouble, e.g. if an extremum of the solution lies between x_j and x_{j+1} .

Also other ideas were created in order to get over the problem of non-physical oscillations as far as higher order methods are under consideration. One of them in the case of higher dimensions was the introduction of so-called **ENO schemes** (essentially non-oscillatory schemes)¹²². This idea consists of a piecewise polynomial reconstruction from the piecewise constant values of the approximate solution of the last time step where the stencil for each cell is adaptively chosen in such a way that intensive oscillations near shocks do not occur.

7.8 The Finite Volume Method on Unstructured Grids

Let us finally briefly put the idea of finite volume schemes on unstructured grids introduced in section 7.1 into concrete form. For convenience, we restrict ourselves to the scalar 2-dimensional conservation law (cf. (1.10))

$$\begin{aligned} \partial_t u + \partial_x f_1(u) + \partial_y f_2(u) &= 0 \quad \text{in } \mathbb{R}^2 \times \mathbb{R}^+, \\ u(x, y; 0) &= u_0(x, y) \quad \text{in } \mathbb{R}^2. \end{aligned} \tag{7.37}$$

A so-called **unstructured grid** of $\hat{\Omega} \subset \mathbb{R}^2$ consists of a set of polygons σ_i ($i \in I \subset \mathbb{N}$), e.g. triangles, each of which has the same number m of edges and vertices, so that

$$\hat{\Omega} = \bigcup_{i \in I} \sigma_i,$$

and two polygons are assumed not to intersect or only to intersect along a common edge or at a common vertex. These polygons are often called **cells**. The angles at the vertices have to be bounded uniformly away from zero, and we assume $h = \mathcal{O}(\Delta t)$ where h represents a measure for the maximal diameter of the cells.

Unstructured grids can more easily be adapted to complicated geometries than structured grids, e.g. to the gas flow around flying objects, and refinements of the net in regions where the exact solution is expected to vary more than in neighbouring areas

¹²²cf. Harten, A.: Multi-dimensional ENO schemes for general geometries. ICASE Report 91-76, Langley Research Center: Hampton 1991

can easily be realized. The development of **grid generation techniques**, therefore, is a very important task in this context.

Let $I_i \subset I$ be the index set consisting of the numbers of the particular cells being neighbours of σ_i with a joint edge.

With $u^n(x, y) := u(x, y; n \Delta t)$ ($n \in \mathbb{N}_0$) in the case of a given time step Δt and with an exact solution u assumed for a short time to be smooth, equation (7.37) leads to

$$\frac{u^{n+1}(x, y) - u^n(x, y)}{\Delta t} + \partial_x f_1(u^n(x, y)) + \partial_y f_2(u^n(x, y)) = \mathcal{O}(\Delta t). \quad (7.38)$$

Taking the divergence theorem into account, integration of this equation over the cell σ_i with edges $\partial\sigma_{i,k}$ ($k = 1, 2, \dots, m$) results in

$$\frac{1}{\Delta t} \int_{\sigma_i} \{u^{n+1}(x, y) - u^n(x, y)\} d\sigma_i + \sum_{k \in I_i} \int_{\partial\sigma_{i,k}} \langle \mathbf{f}(u), \mathbf{n}^{i,k} \rangle ds = \mathcal{O}(\Delta t) \quad (7.39)$$

where $\mathbf{f}(u) = (f_1(u), f_2(u))^T$ is assumed to be sufficiently smooth and where $\mathbf{n}^{i,k}$ denotes the outer normal unit vector on $\partial\sigma_{i,k}$ ^{121 122}.

If we replace the functions $u^n(x, y)$, $u^{n+1}(x, y)$ along the cell σ_i by constant approximate values v_i^n , v_i^{n+1} to be calculated, if the term $\mathcal{O}(\Delta t)$ is neglected, if $|\sigma_i|$ denotes the area of σ_i and $s_{i,k}$ the length of the edge $\partial\sigma_{i,k}$, (7.39) leads to

$$v_i^{n+1} = v_i^n - \frac{\Delta t}{|\sigma_i|} \sum_{k \in I_i} \langle \mathbf{f}(\tilde{v}_{i,k}^n), \mathbf{n}^{i,k} \rangle s_{i,k} \quad (7.40)$$

($n = 0, 1, 2, \dots$) where the values $\tilde{v}_{i,k}^n$ represent approximations for u along the edges $\partial\sigma_{i,k}$, respectively, and where v_i^0 is chosen to be the mean value of the initial function on σ_i :

$$v_i^0 = \frac{1}{|\sigma_i|} \int_{\sigma_i} u_0(x, y) d\sigma_i, \quad \forall i \in I.$$

Obviously, from a physical point of view, the term $\langle \mathbf{f}(\tilde{v}_{i,k}^n), \mathbf{n}^{i,k} \rangle s_{i,k}$ describes approximately the flux through the edge $\partial\sigma_{i,k}$.

¹²¹cf. (7.4)

¹²²Here, we assumed that each of the cells has neighbours along each of its edges. If this is not the case, namely along the edges being parts of the boundary of $\hat{\Omega}$, one has to respect suitable boundary conditions.

The question arises, how the values $\tilde{v}_{i,k}^n$ should be chosen.

Because the conservation principle should also be reflected by the approximate values, we aim for

$$\langle \mathbf{f}(\tilde{v}_{i,k}^n), \mathbf{n}^{i,k} \rangle s_{i,k} \approx - \langle \mathbf{f}(\tilde{v}_{k,i}^n), \mathbf{n}^{k,i} \rangle s_{k,i} \quad (7.41)$$

($s_{k,i} = s_{i,k}$).

In order to create a scheme easy to handle, we are going to replace $\tilde{v}_{i,k}^n$ and $\tilde{v}_{k,i}^n$ by a certain mean value of v_i^n and v_k^n together with a replacement of the physical flux by a numerical flux $g_{i,k}(v_i^n, v_k^n)$ that fulfills the relation (7.41) exactly, i.e.

$$g_{i,k}(v_i^n, v_k^n) = -g_{k,i}(v_k^n, v_i^n). \quad (7.42)$$

Moreover,

$$g_{i,k}(w, w) = \langle \mathbf{f}(w), \mathbf{n}^{i,k} \rangle s_{i,k} \quad (7.43)$$

should be fulfilled making the finite volume method now constructed, namely

$$v_i^{n+1} = v_i^n - \frac{\Delta t}{|\sigma_i|} \sum_{k \in I_i} g_{i,k}(v_i^n, v_k^n), \quad (7.44)$$

consistent with (7.39). Finally, we do now forget the assumption on the smoothness of the exact solution because the integration process allows to deal also with weak solutions.

Convergence proofs for FVMs in higher dimensions and on unstructured grids, particularly mathematical proofs for convergence to the entropy solutions, are often not yet available. But what can be done is, to formulate the methods in such a way that they lead to well-known and well working methods if analogously formulated for one-dimensional problems.

A method which could be shown to behave very well in the 1-D situation was the Engquist-Osher flux splitting scheme¹²³ with a numerical flux described by formula (6.24).

If the numerical flux $g_{i,k}$ in (7.43) is chosen accordingly to (6.24) as

$$g_{i,k}(v, \hat{v}) = \left[\hat{f}_{i,k}^-(v) + \hat{f}_{i,k}^+(\hat{v}) \right] \quad (7.45)$$

¹²³cf. section 6.4

with

$$\hat{f}_{i,k}(w) := \langle \mathbf{f}(w), \mathbf{n}^{i,k} \rangle s_{i,k}$$

and with

$$\begin{aligned} \hat{f}_{i,k}^+(v) &= \int_0^v \max \{ \hat{f}'_{i,k}(w), 0 \} dw + \hat{f}_{i,k}(0) \\ \hat{f}_{i,k}^-(\hat{v}) &= \int_0^{\hat{v}} \min \{ \hat{f}'_{i,k}(w), 0 \} dw, \end{aligned} \tag{7.46}$$

it corresponds to the Engquist-Osher scheme in 1-D, and it fulfills the relations (7.42) and (7.43).

After these brief introductions, we stop here, but it should be mentioned that there are lots of schemes available for the numerical treatment of conservation laws like the Euler equations or for the solution of Navier-Stokes problems, scalar or systems, one-dimensional or multi-dimensional. Most of them show advantages compared with other schemes but often certain disadvantages as well. Many additional informations can be found in some of the extensive monographs listed behind this page.

Some Extensive Monographs

Since the aim of this book is to give just an introduction into Mathematical Fluid Dynamics (**MFD**), it is impossible to refer to all aspects of this field, but there is a lot of very good and modern extensive monographs available. Most of them describe parts of MFD like numerical procedures or they treat the basic equations theoretically or they justify the mathematical treatment from a physical or engineering point of view. Some modern books useful for intensified studies are listed here:

- [1] Brower, W.B., *Dynamics of flows in one space dimension*. Boca Raton, London, New York, Washington, D.C., CRC 1999
- [2] Chorin, A.J. and J.E. Marsden, *A Mathematical Introduction to Fluid Mechanics, 2nd ed.*, New York, Berlin, Heidelberg, Tokyo, Springer 1984
- [3] Dubois, Th., F. Jauberteau and R. Temam, *Dynamic multilevel methods and the numerical simulation of turbulence*, Cambridge, University Press 1999
- [4] Emanuel, G., *Analytic fluid dynamics*, Boca Raton, London, New York, Washington, D.C., CRC 2001
- [5] Feistauer, M., *Mathematical methods in fluid dynamics*, Boca Raton, London, New York, Washington, D.C., CRC 1993
- [6] Ferziger, J.H. and M. Peric, *Computational methods for fluid dynamics*, Berlin, Heidelberg, Springer 1999
- [7] Freistühler, H. (ed.), *Analysis of systems of conservation laws*, Boca Raton, London, New York, Washington, D.C., Chapman Hall/CRC 1998
- [8] Frisch, U., *Turbulence: the legacy of A.N. Kolmogoroff*, Cambridge, University Press 1995
- [9] Godlewski, E. and P.-A. Raviart, *Numerical approximation of hyperbolic systems of conservation laws*, Berlin, Heidelberg, New York, Springer 1996
- [10] Holmes, P., J.L. Lumley and G. Berkooz, *Turbulence, coherent structures, dynamical systems and symmetry*, Cambridge, University Press 1996

- [11] Kröner, D., *Numerical schemes for conservation laws*, Chichester, Stuttgart, Wiley Teubner 1997
- [12] LeVeque, R.J., *Numerical methods for conservation laws*, Basel, Boston, Berlin, Birkhäuser 1990
- [13] Majda, A., *Compressible fluid flow and systems of conservation laws in several space variables*, New York, Berlin, Heidelberg, Tokyo, Springer 1984
- [14] Malek, J., J. Necas and M. Rokyta, *Advanced topics in theoretical fluid mechanics*, Boca Raton, London, New York, Washington, D.C., CRC 1999
- [15] Meister, A. and J. Struckmeier, *Hyperbolic partial differential equations*, Braunschweig, Wiesbaden, Friedr. Vieweg 2002
- [16] Richtmyer, R.D. and K.W. Morton, *Difference methods for initial value problems, 2nd ed.*, New York, London, Sydney, John Wiley and Sons 1967
- [17] Schlichting, H. and K. Gersten, *Boundary-layer theory, 8th ed.*, Berlin, Heidelberg, New York, Springer 2000
- [18] Serre, D., *Systems of conservation laws I*, Cambridge, University Press 1999
- [19] Smoller, J., *Shock waves and reaction-diffusion equations*, New York, Heidelberg, Berlin, Springer 1983
- [20] Sonar, Th., *Mehrdimensionale ENO-Verfahren*, Stuttgart, B.G.Teubner 1997
- [21] Spurk, J.H., *Fluid mechanics*, Berlin, Heidelberg, New York, Springer 1997
- [22] Warsi, Z.U.A., *Fluid dynamics, theoretical and computational approaches*, Boca Raton, London, New York, Washington, D.C., CRC 1998

Index

- advection equation 28
- amplification factor 169
- amplification matrix 170
- angular velocity vector 31
- antidifference operator 159
- approximate solutions 139
- Archimedes' principle 49
- artificial compressibility 117
- artificial time 116
- asymptotically closed 139
- asymptotically regular 139

- back side of a shock 65
- backward shift operator 168
- barotropic flow 32
- Bernoulli equation 31
 - , generalized 32
- Bernoulli, Daniel 31
- Bois-Reymond, P. du 142
- Boltzmann, L. 69
- boundary layer 123
- boundary layer equations 126
- boundary layer theory 116
- Breuss, M. 94
- Briggs, H. 156
- Buckley-Leverett equation 95
- buoyancy, dynamic 30
 - , static 43, 49
- Burgers equation 24
- Burgers, J. 24

- Cauchy-Riemann equations 37
- cell average 155
- cells of a FVM 177
- centered waves 85

- central difference operator 157
- CFL condition 136, 145, 161, 162
- Chapligin solutions 46
- Chapligin, C.A. 46
- characteristic heat conductivity 110
- characteristics 24, 25, 27
- characteristics–, numerical 161
- circulation 30
- Clausius, R.J.E. 69
- closing of disturbed laminar flow 131
- collocation method 156
- complex velocity potential 37
- congestion 57
- congestion velocity 67
- conservation laws 21
- conservation of energy 13, 19
 - , of mass 13, 15
 - , of momentum 16
- conservation principles 9, 13
- conservative variables 18
- consistency of numerical methods 145, 159, 164
- contact discontinuity 65
- continuity equation 16
- continuous convergence 142
- convection term 18
- convergence of numerical methods 98, 139, 166
- Courant, R. 136, 142, 160
- critical Reynolds number 115, 133
- curve of neutral stability 133

- density 13, 20
- density of momentum 16
- difference operator 151, 157, 164

- difference stencil 156, 161
- diffusion term 29
- discontinuous Galerkin methods 154
- discrete entropy condition 141
- discrete entropy solutions 141
- discrete evolution operator 164
- discrete Fourier transform 168
- discretely compact 140
- discretization error, global 167
- divergence theorem 17
- drag force 112
- dynamic pressure 31

- elements of the FEM 153
- energy, interior 18
 - , kinetic 18
 - , total 18
- Engquist, B. 148
- Engquist-Osher scheme 148, 179
- ENO schemes 176
- entropy 20, 69
- entropy condition 70, 73, 74, 76–79, 91, 141
- entropy flux 74, 144
 - , numerical 146
- entropy function 144
- entropy functional 74
- entropy inequalities 91
- entropy solution 70, 74, 76, 79, 80, 84, 89, 91, 93, 141
- equation of state 15, 20
- equivalence theorem 163
- e^9 -rule 133
- Euler equations 13, 18
- Euler's law of vanishing momentum 104
- Euler, L. 18, 104
- evolution equation 163
- evolution operator 164
- exact difference formula 157
- explicit difference scheme 152

- Fellehner, St. 118
- finite difference method (FDM) 97, 151
- finite difference methods (FDM) 145
- finite element method (FEM) 151

- finite volume method (FVM) 151
- first law of thermodynamics 19
- flow separation 122
- flow, barotropic 32
 - , incompressible 20
 - , instationary 21
 - , irrotational 30
 - , isentropic 73
 - , laminar 111
 - , plane 34
 - , potential 30, 31
 - , stationary 21
 - , steady-state 21
 - , subsonic 34
 - , supersonic 34
 - , through porous media 95
 - , turbulent 111
- fluid, ideal 17
 - , inviscid 17
 - , real 17
 - , viscous 17
- flux 21
 - , numerical 145, 155
- flux limiter methods 175
- flux splitting 146, 148
- force, conservative 31
- forward difference operator 156
- forward shift operator 168
- Friedrichs, K.-O. 136
- Friedrichs, K.O. 160, 174
- front side of a shock 65
- functional analysis of discretization algorithms 137
- fundamental diagram 57

- γ -gas 20
- gas, ideal 17
- generalized solution operator 163
- genuine solution 163
- genuinely nonlinear 86
- Glimm, J. 136
- global discretization error 167
- Godunov, S.K. 99
- Greenshield' model 58
- grid 97

- , unstructured 155, 176
- grid generation 177
- Hagen, G.H.L. 114
- Hagen-Poiseuille law 114
- Harriot, Th. 156
- Harten, A. 144, 149
- hodograph equation 45
- hodograph plane 44
- hydrodynamic paradoxon 32
- hydrostatic equation 47
- hydrostatic paradoxon 48
- Hyman, J.M. 144
- ideal fluid 17
- ideal gas 17
- implicit finite difference Method 152
- incompressible flow 20
- infinitesimal generator 163
- interior energy 18
- inviscid fluid 17
- irrotational flow 30
- isentropic flow 73
- iterative procedure 117
- Jacobian matrix 21
- jam concentration 57
- jump condition 64
- Kantorowitch, L.W. 136
- Kelvin's theorem 33
- Kelvin, Lord of Largs 33
- Keyfitz, B. 149
- kinetic energy 18
- Kruzkov, S.N. 92, 144
- Kutta, M.W. 44
- Kutta-Zhukovsky buoyancy formula 44
- Kuznetsov, N.N. 80
- laminar flow 111
- Laurent, P.A. 40
- Lax shock conditions 91
- Lax, P. 74, 144, 174, 175
- Lax-Friedrichs scheme 174
- Lax-Richtmyer theory 162
- Lax-Wendroff scheme 174
- Lax-Wendroff theorem 149
- Legendre transformation 45
- Lei, J. 150
- Lewy, H. 136
- lift 43
- Lighthill, M.J. 58
- Lighthill-Whitham model 58
- local sound speed 50
- Mach number 34
- Mach's angle 35
- Mach's cone 35
- Mach's net 35
- Mach, E. 34, 35
- main theorem of thermodynamics, first
 - 19, 69
 - , second 69
- mathematical model 7
- method of artificial compressibility 117
- method of characteristics 35
- method of conservation form 145, 173
- method of hodographs 45
- modified equation 173
- monotone schemes 148
- Morton, K.W. 171
- multistep method 97
- Napier, J. 156
- Navier, C.L.M.H. 108
- Navier-Stokes equations 108, 123
- Neumann, J. v. 168
- Newton's second law 16
- Newtonian fluids 105
- Nickel, K. 122
- no-slip condition 44
- non-standard conservation laws 95
- numerical characteristics 161
- numerical domain of dependence 162
- numerical entropy flux 146
- numerical flux 145, 155, 178
- numerical stability 167
- Oleinik, O. 79
- one step method 97
- order of a numerical procedure 159, 165
- Orr-Sommerfeld equation 132

- Osher, St. 148
 Panov, E.Y. 93
 partial-slip condition 44
 Petrov-Galerkin schemes 154
 phase speed of perturbation 132
 plane flow 37
 Poise 108
 Poiseuille, J.L.M 115
 polymer 95
 potential flow 30, 31
 Prandtl number 110
 Prandtl, L. 110, 116, 122, 123
 pressure 16, 20, 104
 –, dynamic 31
 –, static 31
 –, total 41
 primitive variables 18
 pseudospectral method 156

 quasilinear systems of partial differential equations 21

 Rankine-Hugoniot condition 62, 101
 rarefaction wave 60, 91
 real fluid 17
 reconstruction of functions 98, 135
 Reynolds averaging 131
 Reynolds number 110
 –, critical 115
 Reynolds' transport theorem 15
 Reynolds, O. 15, 110, 131
 Richtmyer, R. 136, 162, 171
 ride impulse 59
 Riemann invariant 91
 Riemann solver 99
 Riemann, B. 58
 Riemann-problem 58
 Rinow, W. 143
 round off errors 167

 scheme
 –, $(2k + 1)$ -point 145
 –, 3-point 174
 Schlichting, H. 125, 132
 second law of thermodynamics 69

 semigroup property 163
 separation of flows 122
 shallow water equation 51
 shift operator 156
 –, backward 168
 –, forward 168
 shock 62, 91
 shock conditions 91
 smooth connection of states 88
 solution operator 163
 solution operator–, generalized 163
 Sommerfeld, A. 132
 Sonar, Th. 69
 sound propagation 47
 sound velocity, local 23, 50
 spatial stability 132
 specific heats 20
 spectral discretization 155
 spectral Galerkin method 156
 stability criterion 170
 stability, numerical 167
 –, of laminar flows 131
 static pressure 31
 step size 98
 Stokes approximation 115
 Stokes, G.G. 108, 115
 stream function 36
 streamline 24
 stress tensor 104
 Stummel, F. 137
 subsonic flow 24
 succing of the boundary layer 127
 summation by parts 160
 supersonic flow 24

 temperature 20
 temporal stability 132
 test elements 93, 138
 test function 53
 Tollmien, W. 122, 132
 Tollmien-Schlichting waves 132
 total energy 18
 total pressure 41
 traffic flow 56
 transport theorem 15

- tribology 115
- truncation error 98, 159, 165
- turbulent flow 111
- TVB schemes 148
- TVD schemes 148
- Tveito, A. 95

- unstructured grids 155, 176
- upwind scheme 162

- variables, conservative 18
 - , primitive 18
- velocity potential 31
- viscosity 17, 95
- viscosity coefficient, first 108
 - , kinematic 108
 - , second 108
- viscous fluid 17
- von Neumann condition 171
- von Neumann stability criterion 168, 170
- vorticity vector 31

- Wang, C.Y. 115
- weak formulation of numerical procedures 139
- weak solution 29, 53, 54, 138
- Wendroff, B. 149
- Whitham, C.B. 58
- wing theory 116
- Winther, R. 95

- Zhukovsky, N.J. 44