

Eleuterio F. Toro

# Riemann Solvers and Numerical Methods for Fluid Dynamics

---

A Practical Introduction  
*Third Edition*



# Riemann Solvers and Numerical Methods for Fluid Dynamics

Third Edition

Eleuterio F. Toro

# Riemann Solvers and Numerical Methods for Fluid Dynamics

A Practical Introduction

Third Edition



Professor Eleuterio F. Toro  
University of Trento  
Italy  
eleuteriofrancisco.toro@unitn.it

ISBN 978-3-540-25202-3                    e-ISBN 978-3-540-49834-6  
DOI 10.1007/978-3-540-49834-6  
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2009921818

© Springer-Verlag Berlin Heidelberg 2009

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

*Cover design:* eStudio Calamar S.L.

Printed on acid-free paper

Springer is part of Springer Science+Business Media ([www.springer.com](http://www.springer.com))

In memory of two remarkable persons:  
my mother *Olga Muñoz de Toro* (Chile, 1930–2003) and  
my father-in-law *Fritz Freund* (Germany, 1926–2008).

---

## Preface to the First Edition

In 1917, the British scientist L. F. Richardson made the first reported attempt to predict the weather by solving partial differential equations numerically, by hand! It is generally accepted that Richardson's work, though unsuccessful, marked the beginning of Computational Fluid Dynamics (CFD), a large branch of scientific computing today. His work had the four distinguishing characteristics of CFD: a PRACTICAL PROBLEM to solve, a MATHEMATICAL MODEL to represent the problem in the form of a set of partial differential equations, a NUMERICAL METHOD and a COMPUTER, human beings in Richardson's case. Eighty years on and these four elements remain the pillars of modern CFD. It is therefore not surprising that the generally accepted definition of CFD as *the science of computing numerical solutions to partial differential or integral equations that are models for fluid flow phenomena*, closely embodies Richardson's work.

COMPUTERS have, since Richardson's era, developed to unprecedented levels and at an ever decreasing cost. The range of application areas giving rise to PRACTICAL PROBLEMS to be solved numerically has increased dramatically. In addition to the traditional demands from meteorology, oceanography, some branches of physics and from a range of engineering disciplines, there are at present fresh demands from a dynamic and fast-moving manufacturing industry, whose traditional build-test-fix approach is rapidly being replaced by the use of quantitative methods, at all levels. The need for new materials and for decision-making under environmental constraints are increasing sources of demands for mathematical modelling, numerical algorithms and high-performance computing. MATHEMATICAL MODELS have improved, though the basic equations of continuum mechanics, already available more than a century before Richardson's first attempts at CFD, are still the bases for modelling fluid flow processes. Progress is required at the level of thermodynamics, equations of state, and advances into the modelling of non-equilibrium and multiphase flow phenomena. NUMERICAL METHODS are perhaps the success story of the last eighty years, the last twenty being perhaps the most productive. This success is firmly based on the pioneering works of

## VIII Preface to the First Edition

scientists such as von Neumann, whose research on *stability* explained and resolved the difficulties experienced by Richardson. This success would have been impossible without the contributions from Courant, Friedrichs, Richtmyer, Lax, Oleinik, Wendorff, Godunov, Rusanov, van Leer, Harten, Roe, Osher, Colella, Yee, and many others. The net result is: more accurate, more efficient, more robust and more sophisticated numerical methods are available for ambitious practical applications today.

Due to the massive demands on CFD and the level of sophistication of numerical methods, new demands on *education and training* of the scientists and engineers of the present and the future have arisen. This book is an attempt to contribute to the training and education in numerical methods for fluid dynamics and related disciplines.

The contents of this book were developed over a period of many years of involvement in research on numerical methods, application of the methods to solve practical problems and teaching scientist and engineers at the post-graduate level. The starting point was a module for a Masters Course in Computational Fluid Dynamics at the College of Aeronautics, Cranfield, UK. The material was also part of short courses and lectures given at Cranfield, UK; the Ernst Mach Institute, Freiburg, Germany; the Shock Wave Research Centre, Tohoku University, Sendai, Japan; the Department of Mathematics and the Department of Civil and Environmental Engineering, University of Trento, Italy; the Department of Mathematics, Technical University Federico Santa Maria, Chile; the Department of Mechanics, Technical University of Aachen, Germany; and the Manchester Metropolitan University (MMU), Manchester, UK.

This book is about modern shock-capturing numerical methods for solving time-dependent hyperbolic conservation laws, with smooth and discontinuous solutions, in general multidimensional geometries. The approach is comprehensive, practical and, in the main, informal. All necessary items of information for the practical implementation of all methods studied here, are provided in detail. All methods studied are illustrated through practical numerical examples; numerical results are compared with exact solutions and in some cases with reliable experimental data.

Most of the book is devoted to a coherent presentation of *Godunov methods*. The developments of Godunov's approach over the last twenty years have led to a mature numerical technology, that can be utilised with confidence to solve practical problems in established as well as new areas of application. Godunov methods rely on the solution of the Riemann problem. The exact solution is presented in detail, so as to aid the reader in applying the solution methodology to other hyperbolic systems. We also present a variety of approximate Riemann solvers; again, the amount of detail supplied will hopefully aid the reader who might want to apply the methodologies to solve other problems. Other related methods such as the Random Choice Method and the Flux Vector Splitting Method are also included. In addition, we study *centred* (non-upwind) shock-capturing methods. These schemes are much less sophisticated

than Godunov methods, and offer a cheap and simple alternative. High-order extensions of these methods are constructed for scalar PDEs, along with their Total Variation Diminishing (TVD) versions. Most of these TVD methods are then extended to one-dimensional non-linear systems. Techniques to deal with PDEs with source terms are also studied, as are techniques for multidimensional systems in general geometries.

The presentation of the schemes for non-linear systems is carried out through the time-dependent Euler equations of gas dynamics. Having read the relevant chapters/sections, the reader will be sufficiently well equipped to extend the techniques to other hyperbolic systems, and to advection-reaction-diffusion PDEs.

There are at least two ways of utilising this book. First, it can be used as a means for *self-study*. In the presentation of the concepts, the emphasis has been placed on clarity, sometimes sacrificing mathematical rigour. The typical reader in mind is a graduate student in a department of engineering, physics, applied mathematics or computer science, embarking on a research topic that involves the implementation of numerical methods, from first principles, to solve advection-reaction-diffusion problems. The contents of this book may also be useful to numerical analysts beginning their research on algorithms, as elementary background reading. Such users may benefit from a comprehensive self-study of all the contents of the book, in a period of about two months, perhaps including the practical implementation and testing of most numerical methods presented. Another class of readers who may benefit from self-studying this book are scientists and engineers in industry and research laboratories. At the cost of some repetitiveness, each chapter is almost self-contained and has plenty of cross-referencing, so that the reader may decide to start reading this book in the middle or jump to the last chapter.

This book can also be used as a *teaching aid*. Academics involved in the teaching of numerical methods may find this work a useful reference book. Selected chapters or sections may well form the bases for a final year undergraduate course on numerical methods for PDEs. In a mathematics or computer science department, the contents may include: some sections of chapter 1, chapters 2, 5, 13, some sections of chapter 14, chapter 15 and some sections of chapter 16. In a department of engineering or physics, one may include chapters 3, 4, 6, 7, 8, 9, 10, 11, 12, 15, 16 and 17. A postgraduate course may involve most of the contents of this book, assuming perhaps a working knowledge of compressible fluid dynamics. Short courses for training engineers and scientists in industry and research laboratories can also be based on most of the contents of this book.

Eleuterio Toro  
Manchester, UK  
March 1997.

## ACKNOWLEDGEMENTS FOR FIRST EDITION

The author is indebted to many colleagues who over the years have kindly arranged short and extended visits to their establishments, have organised seminars, workshops and short courses. These events have strongly influenced the development of this book. Special thanks are due to Dr. W. Heilig (Freiburg, Germany); Professors V. Casulli and A. Armanini (Trento, Italy); Professor K. Takayama (Sendai, Japan); Professor J. Figueroa (Valparaiso, Chile) and Professor J. Ballmann (Aachen, Germany). The final stages of the typing of the material were carried out while the author was at MMU, Manchester, UK; the support provided is gratefully acknowledged. Thanks are also due to my former and current PhD students, for their comments on my lectures, on the contents of this book and for their contribution to research in relevant areas. Special thanks are due to Stephen Billett (Centre for Atmospheric Science, University of Cambridge, UK), Caroline Lowe (Department of Aerospace Science, Cranfield University, UK), Nikolaos Nikiforakis (DAMTP, University of Cambridge, UK), Ed Boden (Cranfield and MMU), Ms Wei Hu, Mathew Ivings and Richard Millington (MMU). Thanks are due to my son Marcelo, who helped with the type-setting, and to David Ingram and John Nuttall for their help with the intricacies of Latex. The author thanks Springer-Verlag, and specially Miss Erdmuthe Raufelder in Heidelberg, for their professional assistance.

The author gratefully acknowledges useful feedback on the manuscript from Stephen Billett (University of Cambridge, UK), John Clarke (Cranfield University, UK), Jean-Marc Moschetta (SUPAERO, Toulouse, France), Claus-Dieter Munz (University of Stuttgart, Germany), Jack Pike (Bedford, UK), Ning Qin (Cranfield University, UK), Tsutomu Saito (Cray Japan), Charles Saurel (University of Provence, France), Trevor Randall (MMU, Manchester, UK), Peter Swebey (University of Reading, UK), Marcelo Toro (IBM Glasgow, UK), Helen Yee (NASA Ames, USA) and Clive Woodley (DERA, Fort Halstead, UK).

The permanent encouragement from Brigitte has been an immensely valuable support during the writing of this book. Thank you Brigitte!

---

## Preface to the Third Edition

More than a decade has elapsed since the publication of the first edition of this book. During this period the response from readers the world over has been overwhelming, from students and academics to senior researchers. This decade has also witnessed a significant increase in the use of numerical methods, not only in the traditional areas such as physics and industrial processes, but also in biology, economics, social sciences and in inter-disciplinary research areas. We also observe a new trend in mathematical modelling and numerical simulation. Numerical methods are steadily moving from being a simulation tool for engineering design in technology, to being an indispensable instrument in science, for studying and understanding phenomena of the most varied kind. The expectation is that a simulation will represent the solution of the actual mathematical model. Such expectation implies the need for more and more high-quality research on new and more accurate numerical methods and the need for better training of scientists at the undergraduate and post-graduate levels at universities and higher education institutions. I expect that this new edition of the book will continue to play a role in such endeavours.

In this edition I have included three new chapters, chapters 18 to 20. Chapter 18 is about a multi-dimensional extension of the centred FORCE flux studied in Chapter 7, for one-dimensional systems; in a sense, this new chapter is a response to the increasing role of the so-called centred methods. These have the advantage of avoiding the direct solution of the classical Riemann problem, in the conventional manner. As a result, the applicability of these centred schemes is more general than that of conventional upwind methods; this feature is specially useful when having to solve complicated systems, for which the solution of the Riemann problem may be difficult or impossible to obtain. The new FORCE scheme applies to two and three space dimensions on general structured and unstructured meshes and can be extended to high order of accuracy in space and time in the frameworks of finite volume and discontinuous Galerkin finite element methods. The second new addition, chapter 19, is about the high-order, or generalized, Riemann problem, the

Cauchy problem for hyperbolic balance laws whose initial conditions are two smooth vectors separated by a discontinuity at the origin. The solution of the generalized Riemann problem serves as the building block for constructing fully discrete, one-step Godunov-type schemes of arbitrary order of accuracy in both space and time. This chapter is effectively a generalization of the material of this book studied in previous chapters and responds to the general trend of improving the accuracy, in space and time, of numerical methods for solving evolutionary partial differential equations. These high-order numerical schemes can be constructed in the frameworks of finite volume and discontinuous Galerkin finite element methods. The third new chapter, chapter 20, contains an introduction to these high-order methods in the framework of finite volumes.

This edition of the book contains a substantially revised version of the HLLC Riemann solver of chapter 10. This responds to many communications received from readers and to new developments of the technique and its use for very ambitious scientific and technological applications. Some modifications to chapter 21 have also been carried out, as well general corrections to errors pointed out by readers.

This book, in spite of being introductory in nature, continues to be a book used mainly by researchers, and is perhaps too advanced for teaching undergraduate students. As a result, a new more elementary book is being written in collaboration with Enrico Bertolazzi and Gianluca Vignoli. This new book, to be published by Springer in 2009, is specifically designed for teaching undergraduate students in Science and Engineering, with plenty of exercises, case studies and miniprojects. On the other hand, in order to respond to research needs for better and more sophisticated numerical methods, we are currently preparing a new book, in collaboration with Claus-Dieter Munz, Vladimir Titarev and Michael Dumbser. This book will deal with advanced, high-order, finite volume and discontinuous Galerkin numerical methods for structured and unstructured meshes in multiple space dimensions, to be published by Springer in 2009.

I gratefully acknowledge the contribution of some collaborators to the preparation of this third edition. In particular I thank two of my former PhD students, Dr. Vladimir Titarev, now at Cranfield University, UK and Dr. Cristóbal Castro, now at the Technical University of Munich, Germany. I also thank two former post-doctoral fellows, Dr. Martin Käser, now at the Technical University of Munich, Germany and Dr. Michael Dumbser, now a colleague in my group at Trento University, Italy. Thanks are also due to colleague Enrico Bertolazzi and to visiting scholars María Nofuentes (Universidad de Córdoba, Spain) and Arturo Hidalgo (Universidad Politécnica de Madrid, Spain), who kindly helped in various ways in the preparation of the material.

I find it also appropriate to recall that the initial stages in the preparation of the first edition of this book took place at Cranfield University, formerly Cranfield Institute of Technology, United Kingdom. During the late 80s and early 90s, a very dynamic and stimulating group of scientists from various backgrounds, existed in the Aerodynamics Department, within the School of Aerospace Sciences. Those years at Cranfield greatly influenced the contents and scope of this book, for which I thank former senior and junior colleagues and students, John Clarke (FRS), Phil Roe, Jack Pike, Smadar Karni, James Quirk, Nikos Nikiforakis, Stephen Billett, Alan Dowes, Caroline Lowe, William Speares, Michael Spruce, Ed Boden, Mauricio Cáceres, and many others.

Finally, I thank Springer Verlag for their permanent and positive encouragement and effective support.

Eleuterio Toro  
University of Trento, Italy  
December 2008

---

# Contents

Preface to the First Edition .....	VII
Preface to the Third Edition .....	XI
<b>1 The Equations of Fluid Dynamics.....</b>	<b>1</b>
1.1 The Euler Equations .....	2
1.1.1 Conservation–Law Form .....	3
1.1.2 Other Compact Forms .....	4
1.2 Thermodynamic Considerations .....	5
1.2.1 Units of Measure .....	5
1.2.2 Equations of State (EOS) .....	6
1.2.3 Other Variables and Relations .....	7
1.2.4 Ideal Gases .....	11
1.2.5 Covolume and van der Waal Gases .....	13
1.3 Viscous Stresses .....	15
1.4 Heat Conduction .....	17
1.5 Integral Form of the Equations .....	18
1.5.1 Time Derivatives .....	19
1.5.2 Conservation of Mass .....	20
1.5.3 Conservation of Momentum.....	21
1.5.4 Conservation of Energy .....	23
1.6 Submodels.....	25
1.6.1 Summary of the Equations .....	25
1.6.2 Flow with Area Variation .....	27
1.6.3 Axi–Symmetric Flows .....	28
1.6.4 Cylindrical and Spherical Symmetry .....	29
1.6.5 Plain One–Dimensional Flow .....	29
1.6.6 Steady Compressible Flow .....	31
1.6.7 Viscous Compressible Flow .....	33
1.6.8 Free–Surface Gravity Flow .....	33
1.6.9 The Shallow Water Equations .....	35

1.6.10	Incompressible Viscous Flow . . . . .	38
1.6.11	The Artificial Compressibility Equations . . . . .	39
<b>2</b>	<b>Notions on Hyperbolic Partial Differential Equations . . . . .</b>	<b>41</b>
2.1	Quasi–Linear Equations: Basic Concepts . . . . .	42
2.2	The Linear Advection Equation . . . . .	47
2.2.1	Characteristics and the General Solution . . . . .	47
2.2.2	The Riemann Problem . . . . .	49
2.3	Linear Hyperbolic Systems . . . . .	50
2.3.1	Diagonalisation and Characteristic Variables . . . . .	51
2.3.2	The General Initial–Value Problem . . . . .	52
2.3.3	The Riemann Problem . . . . .	55
2.3.4	The Riemann Problem for Linearised Gas Dynamics . . . . .	58
2.3.5	Some Useful Definitions . . . . .	60
2.4	Conservation Laws . . . . .	61
2.4.1	Integral Forms of Conservation Laws . . . . .	62
2.4.2	Non–Linearities and Shock Formation . . . . .	66
2.4.3	Characteristic Fields . . . . .	77
2.4.4	Elementary–Wave Solutions of the Riemann Problem . . . . .	83
<b>3</b>	<b>Some Properties of the Euler Equations . . . . .</b>	<b>87</b>
3.1	The One–Dimensional Euler Equations . . . . .	87
3.1.1	Conservative Formulation . . . . .	87
3.1.2	Non–Conservative Formulations . . . . .	91
3.1.3	Elementary Wave Solutions of the Riemann Problem . . . . .	94
3.2	Multi–Dimensional Euler Equations . . . . .	103
3.2.1	Two–Dimensional Equations in Conservative Form . . . . .	104
3.2.2	Three–Dimensional Equations in Conservative Form . . . . .	108
3.2.3	Three–Dimensional Primitive Variable Formulation . . . . .	109
3.2.4	The Split Three–Dimensional Riemann Problem . . . . .	111
3.3	Conservative Versus Non–Conservative Formulations . . . . .	112
<b>4</b>	<b>The Riemann Problem for the Euler Equations . . . . .</b>	<b>115</b>
4.1	Solution Strategy . . . . .	116
4.2	Equations for Pressure and Particle Velocity . . . . .	119
4.2.1	Function $f_L$ for a Left Shock . . . . .	120
4.2.2	Function $f_L$ for Left Rarefaction . . . . .	122
4.2.3	Function $f_R$ for a Right Shock . . . . .	123
4.2.4	Function $f_R$ for a Right Rarefaction . . . . .	124
4.3	Numerical Solution for Pressure . . . . .	125
4.3.1	Behaviour of the Pressure Function . . . . .	125
4.3.2	Iterative Scheme for Finding the Pressure . . . . .	127
4.3.3	Numerical Tests . . . . .	129
4.4	The Complete Solution . . . . .	132
4.5	Sampling the Solution . . . . .	136

4.5.1	Left Side of Contact: $S = x/t \leq u_*$ . . . . .	137
4.5.2	Right Side of Contact: $S = x/t \geq u_*$ . . . . .	137
4.6	The Riemann Problem in the Presence of Vacuum . . . . .	139
4.6.1	Case 1: Vacuum Right State . . . . .	140
4.6.2	Case 2: Vacuum Left State . . . . .	142
4.6.3	Case 3: Generation of Vacuum . . . . .	142
4.7	The Riemann Problem for Covolume Gases . . . . .	143
4.7.1	Solution for Pressure and Particle Velocity. . . . .	144
4.7.2	Numerical Solution for Pressure . . . . .	147
4.7.3	The Complete Solution. . . . .	147
4.7.4	Solution Inside Rarefactions . . . . .	148
4.8	The Split Multi-Dimensional Case . . . . .	149
4.9	FORTRAN Program for Exact Riemann Solver . . . . .	151
<b>5</b>	<b>Notions on Numerical Methods . . . . .</b>	<b>163</b>
5.1	Discretisation: Introductory Concepts . . . . .	163
5.1.1	Approximation to Derivatives . . . . .	164
5.1.2	Finite Difference Approximation to a PDE . . . . .	165
5.2	Selected Difference Schemes. . . . .	168
5.2.1	The First Order Upwind Scheme . . . . .	168
5.2.2	Other Well-Known Schemes . . . . .	172
5.3	Conservative Methods . . . . .	174
5.3.1	Basic Definitions . . . . .	175
5.3.2	Godunov's First-Order Upwind Method . . . . .	177
5.3.3	Godunov's Method for Burgers's Equation . . . . .	181
5.3.4	Conservative Form of Difference Schemes . . . . .	184
5.4	Upwind Schemes for Linear Systems . . . . .	187
5.4.1	The CIR Scheme . . . . .	188
5.4.2	Godunov's Method . . . . .	190
5.5	Sample Numerical Results . . . . .	194
5.5.1	Linear Advection . . . . .	194
5.5.2	The Inviscid Burgers Equation . . . . .	196
5.6	FORTRAN Program for Godunov's Method . . . . .	196
<b>6</b>	<b>The Method of Godunov for Non-linear Systems . . . . .</b>	<b>213</b>
6.1	Bases of Godunov's Method . . . . .	213
6.2	The Godunov Scheme . . . . .	216
6.3	Godunov's Method for the Euler Equations . . . . .	218
6.3.1	Evaluation of the Intercell Fluxes . . . . .	219
6.3.2	Time Step Size . . . . .	221
6.3.3	Boundary Conditions . . . . .	222
6.4	Numerical Results and Discussion . . . . .	225
6.4.1	Numerical Results for Godunov's Method . . . . .	226
6.4.2	Numerical Results from Other Methods . . . . .	228

<b>7 Random Choice and Related Methods</b>	237
7.1 Introduction	237
7.2 RCM on a Non-Staggered Grid	238
7.2.1 The Scheme for Non-Linear Systems	239
7.2.2 Boundary Conditions and the Time Step Size	243
7.3 A Random Choice Scheme of the Lax-Friedrichs Type	244
7.3.1 Review of the Lax-Friedrichs Scheme	244
7.3.2 The Scheme	245
7.4 The RCM on a Staggered Grid	247
7.4.1 The Scheme for Non-Linear Systems	247
7.4.2 A Deterministic First-Order Centred Scheme (FORCE)	247
7.4.3 Analysis of the FORCE Scheme	249
7.5 Random Numbers	250
7.5.1 Van der Corput Pseudo-Random Numbers	250
7.5.2 Statistical Properties	251
7.5.3 Propagation of a Single Shock	253
7.6 Numerical Results	255
7.7 Concluding Remarks	256
<b>8 Flux Vector Splitting Methods</b>	265
8.1 Introduction	265
8.2 The Flux Vector Splitting Approach	266
8.2.1 Upwind Differencing	266
8.2.2 The FVS Approach	268
8.3 FVS for the Isothermal Equations	270
8.3.1 Split Fluxes	271
8.3.2 FVS Numerical Schemes	272
8.4 FVS Applied to the Euler Equations	273
8.4.1 Recalling the Equations	274
8.4.2 The Steger-Warming Splitting	276
8.4.3 The van Leer Splitting	277
8.4.4 The Liou-Steffen Scheme	278
8.5 Numerical Results	280
8.5.1 Tests	280
8.5.2 Results for Test 1	280
8.5.3 Results for Test 2	281
8.5.4 Results for Test 3	281
8.5.5 Results for Test 4	282
8.5.6 Results for Test 5	282
<b>9 Approximate-State Riemann Solvers</b>	293
9.1 Introduction	293
9.2 The Riemann Problem and the Godunov Flux	294
9.2.1 Tangential Velocity Components	296
9.2.2 Sonic Rarefactions	296

9.3	Primitive Variable Riemann Solvers (PVRS) . . . . .	297
9.4	Approximations Based on the Exact Solver . . . . .	301
9.4.1	A Two–Rarefaction Riemann Solver (TRRS) . . . . .	301
9.4.2	A Two–Shock Riemann Solver (TSRS) . . . . .	303
9.5	Adaptive Riemann Solvers . . . . .	304
9.5.1	An Adaptive Iterative Riemann Solver (AIRS) . . . . .	304
9.5.2	An Adaptive Noniterative Riemann Solver (ANRS) . . . . .	305
9.6	Numerical Results . . . . .	306
<b>10</b>	<b>The HLL and HLLC Riemann Solvers</b> . . . . .	315
10.1	Introduction . . . . .	315
10.2	The Riemann Problem . . . . .	317
10.2.1	The Godunov Flux . . . . .	317
10.2.2	Integral Relations . . . . .	318
10.3	The HLL Approximate Riemann Solver . . . . .	320
10.4	The HLLC Approximate Riemann Solver . . . . .	322
10.4.1	Useful Relations . . . . .	322
10.4.2	The HLLC Flux for the Euler Equations . . . . .	324
10.4.3	Multidimensional and Multicomponent Flow . . . . .	326
10.5	Wave–Speed Estimates . . . . .	327
10.5.1	Direct Wave Speed Estimates . . . . .	328
10.5.2	Pressure–Based Wave Speed Estimates . . . . .	329
10.6	Summary of HLLC Fluxes . . . . .	331
10.7	Contact Waves and Passive Scalars . . . . .	333
10.8	Numerical Results . . . . .	334
10.9	Closing Remarks . . . . .	336
<b>11</b>	<b>The Riemann Solver of Roe</b> . . . . .	345
11.1	Bases of the Roe Approach . . . . .	346
11.1.1	The Exact Riemann Problem and the Godunov Flux . . . . .	346
11.1.2	Approximate Conservation Laws . . . . .	347
11.1.3	The Approximate Riemann Problem and the Intercell Flux . . . . .	349
11.2	The Original Roe Method . . . . .	351
11.2.1	The Isothermal Equations . . . . .	352
11.2.2	The Euler Equations . . . . .	354
11.3	The Roe–Pike Method . . . . .	358
11.3.1	The Approach . . . . .	358
11.3.2	The Isothermal Equations . . . . .	359
11.3.3	The Euler Equations . . . . .	363
11.4	An Entropy Fix . . . . .	366
11.4.1	The Entropy Problem . . . . .	366
11.4.2	The Harten–Hyman Entropy Fix . . . . .	367
11.4.3	The Speeds $u_*$ , $a_{*L}$ , $a_{*R}$ . . . . .	370
11.5	Numerical Results and Discussion . . . . .	372

11.5.1	The Tests .....	372
11.5.2	The Results .....	373
11.6	Extensions.....	373
<b>12</b>	<b>The Riemann Solver of Osher.....</b>	<b>377</b>
12.1	Osher's Scheme for a General System .....	378
12.1.1	Mathematical Bases .....	378
12.1.2	Osher's Numerical Flux .....	380
12.1.3	Osher's Flux for the Single-Wave Case .....	381
12.1.4	Osher's Flux for the Inviscid Burgers Equation .....	383
12.1.5	Osher's Flux for the General Case .....	384
12.2	Osher's Flux for the Isothermal Equations .....	385
12.2.1	Osher's Flux with P-Ordering .....	386
12.2.2	Osher's Flux with O-Ordering .....	389
12.3	Osher's Scheme for the Euler Equations .....	392
12.3.1	Osher's Flux with P-Ordering .....	393
12.3.2	Osher's Flux with O-Ordering .....	397
12.3.3	Remarks on Path Orderings .....	402
12.3.4	The Split Three-Dimensional Case .....	403
12.4	Numerical Results and Discussion .....	404
12.5	Extensions.....	406
<b>13</b>	<b>High-Order and TVD Methods for Scalar Equations .....</b>	<b>413</b>
13.1	Introduction .....	413
13.2	Basic Properties of Selected Schemes .....	415
13.2.1	Selected Schemes .....	415
13.2.2	Accuracy .....	417
13.2.3	Stability .....	418
13.3	WAF-Type High Order Schemes .....	420
13.3.1	The Basic waf Scheme.....	420
13.3.2	Generalisations of the waf Scheme .....	423
13.4	MUSCL-Type High-Order Methods .....	426
13.4.1	Data Reconstruction .....	426
13.4.2	The MUSCL-Hancock Method (MHM) .....	429
13.4.3	The Piece-Wise Linear Method (PLM) .....	432
13.4.4	The Generalised Riemann Problem (GRP) Method .....	434
13.4.5	Slope-Limiter Centred (SLIC) Schemes .....	436
13.4.6	Other Approaches .....	439
13.4.7	Semi-Discrete Schemes .....	439
13.4.8	Implicit Methods .....	439
13.5	Monotone Schemes and Accuracy .....	440
13.5.1	Monotone Schemes .....	440
13.5.2	A Motivating Example .....	443
13.5.3	Monotone Schemes and Godunov's Theorem .....	447
13.5.4	Spurious Oscillations and High Resolution .....	448

13.5.5 Data Compatibility . . . . .	449
13.6 Total Variation Diminishing ( TVD) Methods . . . . .	451
13.6.1 The Total Variation . . . . .	452
13.6.2 TVD and Monotonicity Preserving Schemes . . . . .	453
13.7 Flux Limiter Methods . . . . .	456
13.7.1 TVD Version of the WAF Method . . . . .	456
13.7.2 The General Flux-Limiter Approach . . . . .	464
13.7.3 TVD Upwind Flux Limiter Schemes . . . . .	469
13.7.4 TVD Centred Flux Limiter Schemes . . . . .	474
13.8 Slope Limiter Methods . . . . .	480
13.8.1 TVD Conditions . . . . .	480
13.8.2 Construction of TVD Slopes . . . . .	481
13.8.3 Slope Limiters . . . . .	482
13.8.4 Limited Slopes Obtained from Flux Limiters . . . . .	484
13.9 Extensions of TVD Methods . . . . .	486
13.9.1 TVD Schemes in the Presence of Source Terms . . . . .	486
13.9.2 TVD Schemes in the Presence of Diffusion Terms . . . . .	486
13.10 Numerical Results for Linear Advection . . . . .	487
<b>14 High-Order and TVD Schemes for Non-Linear Systems</b> . . . . .	493
14.1 Introduction . . . . .	493
14.2 CFL and Boundary Conditions . . . . .	495
14.3 Weighted Average Flux (waf) Schemes . . . . .	496
14.3.1 The Original Version of WAF . . . . .	496
14.3.2 A Weighted Average State Version . . . . .	498
14.3.3 Rarefactions in State Riemann Solvers . . . . .	499
14.3.4 TVD Version of WAF Schemes . . . . .	501
14.3.5 Riemann Solvers . . . . .	503
14.3.6 Summary of the WAF Method . . . . .	503
14.4 The MUSCL-Hancock Scheme . . . . .	504
14.4.1 The Basic Scheme . . . . .	504
14.4.2 A Variant of the Scheme . . . . .	506
14.4.3 TVD Version of the Scheme . . . . .	507
14.4.4 Summary of the MUSCL-Hancock Method . . . . .	510
14.5 Centred TVD Schemes . . . . .	511
14.5.1 Review of the FORCE Flux . . . . .	512
14.5.2 A Flux Limiter Centred (FLIC) Scheme . . . . .	512
14.5.3 A Slope Limiter Centred (SLIC) Scheme . . . . .	514
14.6 Primitive-Variable Schemes . . . . .	515
14.6.1 Formulation of the Equations and Primitive Schemes . . . . .	515
14.6.2 A waf-Type Primitive Variable Scheme . . . . .	517
14.6.3 A MUSCL-Hancock Primitive Scheme . . . . .	520
14.6.4 Adaptive Primitive-Conservative Schemes . . . . .	522
14.7 Some Numerical Results . . . . .	523
14.7.1 Upwind TVD Methods . . . . .	523

14.7.2	Centred TVD Methods . . . . .	524
<b>15</b>	<b>Splitting Schemes for PDEs with Source Terms . . . . .</b>	<b>531</b>
15.1	Introduction . . . . .	531
15.2	Splitting for a Model Equation . . . . .	533
15.3	Numerical Methods Based on Splitting . . . . .	535
15.3.1	Model Equations . . . . .	535
15.3.2	Schemes for Systems . . . . .	536
15.4	Remarks on ODE Solvers . . . . .	537
15.4.1	First-Order Systems of ODEs . . . . .	537
15.4.2	Numerical Methods . . . . .	539
15.4.3	Implementation Details for Split Schemes . . . . .	540
15.5	Concluding Remarks . . . . .	541
<b>16</b>	<b>Methods for Multi-Dimensional PDEs . . . . .</b>	<b>543</b>
16.1	Introduction . . . . .	543
16.2	Dimensional Splitting . . . . .	544
16.2.1	Splitting for a Model Problem . . . . .	544
16.2.2	Splitting Schemes for Two-Dimensional Systems . . . . .	545
16.2.3	Splitting Schemes for Three-Dimensional Systems . . . . .	547
16.3	Practical Implementation of Splitting Schemes in Three Dimensions . . . . .	549
16.3.1	Handling the Sweeps by a Single Subroutine . . . . .	549
16.3.2	Choice of Time Step Size . . . . .	551
16.3.3	The Intercell Flux and the TVD Condition . . . . .	552
16.4	Unsplit Finite Volume Methods . . . . .	555
16.4.1	Introductory Concepts . . . . .	555
16.4.2	Accuracy and Stability of Multidimensional Schemes . . . . .	558
16.5	A MUSCL-Hancock Finite Volume Scheme . . . . .	561
16.6	WAF-Type Finite Volume Schemes . . . . .	563
16.6.1	Two-Dimensional Linear Advection . . . . .	564
16.6.2	Three-Dimensional Linear Advection . . . . .	567
16.6.3	Schemes for Two-Dimensional Nonlinear Systems . . . . .	570
16.6.4	Schemes for Three-Dimensional Nonlinear Systems . . . . .	573
16.7	Non-Cartesian Geometries . . . . .	574
16.7.1	Introduction . . . . .	574
16.7.2	General Domains and Coordinate Transformation . . . . .	575
16.7.3	The Finite Volume Method for Non-Cartesian Domains	578
<b>17</b>	<b>Multidimensional Test Problems . . . . .</b>	<b>585</b>
17.1	Explosions and Implosions . . . . .	586
17.1.1	Explosion Test in Two-Space Dimensions . . . . .	587
17.1.2	Explosion Test in Three Space Dimensions . . . . .	590
17.2	Shock Wave Reflection from a Wedge . . . . .	591

<b>18 FORCE Fluxes in Multiple Space Dimensions</b>	597
18.1 Introduction	597
18.2 Review of FORCE in One Space Dimension	600
18.2.1 FORCE and Related Fluxes	600
18.2.2 Monotonicity and Numerical Viscosity	602
18.3 FORCE Schemes on Cartesian Meshes	605
18.3.1 The Two-Dimensional Case	605
18.3.2 The Three-Dimensional Case	609
18.4 Properties of the FORCE Schemes	610
18.4.1 One-Dimensional Interpretation	610
18.4.2 Some Numerical Experiments	611
18.4.3 Analysis in Multiple Space Dimensions	613
18.5 FORCE Schemes on General Meshes	617
18.6 Sample Numerical Results	621
18.7 Concluding Remarks	621
<b>19 The Generalized Riemann Problem</b>	625
19.1 Introduction	625
19.2 Statement of the Problem	629
19.3 The Cauchy–Kowalewski Theorem	631
19.3.1 Series Expansions and Analytic Functions	632
19.3.2 Illustration of the Cauchy–Kowalewski Theorem	633
19.3.3 The Cauchy–Kowalewski Method	633
19.4 A Method of Solution	635
19.4.1 The Leading Term	636
19.4.2 Higher-Order Terms	637
19.4.3 Source Terms	640
19.4.4 Summary: Numerical Flux and Numerical Source	640
19.4.5 Some remarks	642
19.5 Examples	642
19.5.1 The Linear Advection Equation	643
19.5.2 Linear Advection with a Source Term	645
19.5.3 Non-Linear Equation with a Source Term	646
19.5.4 The Burgers Equation with a Source Term	648
19.6 Other Solvers	651
19.7 Concluding Remarks	653
<b>20 The ADER Approach</b>	655
20.1 Introduction	655
20.2 The Finite Volume Method	657
20.2.1 The Framework	657
20.2.2 The Numerical Flux	658
20.2.3 The Numerical Source	659
20.2.4 Reconstruction	660
20.3 Second-Order Scheme for a Model Equation	663

## XXIV Contents

20.3.1 Numerical Flux and Numerical Source .....	663
20.3.2 The Scheme .....	666
20.4 Schemes of Arbitrary Accuracy .....	667
20.4.1 The Numerical Flux .....	667
20.4.2 The Numerical Source .....	668
20.4.3 Summary .....	668
20.5 Sample Numerical Results .....	669
20.5.1 Long–Time Advection of Smooth Profiles .....	669
20.5.2 Convergence Rates .....	672
20.6 Concluding Remarks .....	673
<b>21 Concluding Remarks .....</b>	<b>679</b>
21.1 Summary of Numerical Aspects .....	679
21.2 Potential Applications .....	681
21.3 Current Research Topics .....	685
21.4 The NUMERICA Library .....	686
<b>References .....</b>	<b>687</b>
<b>Index .....</b>	<b>719</b>

---

# The Equations of Fluid Dynamics

In this chapter we present the governing equations for the dynamics of a compressible material, such as a gas, along with closure conditions in the form of equations of state. Equations of state are statements about the nature of the material in question and require some notions from Thermodynamics. There is no attempt to provide an exhaustive and rigorous derivation of the equations of continuum mechanics; such a task is beyond the scope of this book. Instead, we give a fairly self-contained summary of the equations and the Thermodynamics in a manner that is immediately useful to the main purpose of this book, namely the detailed treatment of Riemann solvers and numerical methods.

The presentation of the equations is unconventional. We first introduce the differential form of the Euler equations along with basic physical quantities and thermodynamic relations leading to equations of state. Then the effects of viscous diffusion and heat transfer are added to the Euler equations. After this, the fundamental integral form of the equations is introduced; conventionally, this is the starting point for presenting the governing equations. This chapter contains virtually all of the necessary background on Fluid Dynamics that is required for a fruitful study of the rest of the book. It also contains useful information for those wishing to embark on complex practical applications. A hierarchy of submodels is also presented. This covers four systems of hyperbolic conservation laws for which Riemann solvers and upwind methods are directly applicable, namely (i) the time-dependent Euler equations, (ii) the steady supersonic Euler equations, (iii) the shallow water equations and (iv) the artificial compressibility equations associated with the incompressible Navier–Stokes equations. Included in the hierarchy are also some simpler models such as linear systems and scalar conservation laws.

Some remarks on notation are in order. A Cartesian frame of reference  $(x, y, z)$  is chosen and the time variable is denoted by  $t$ . Transformation to other coordinate systems is carried out using the chain rule in the usual way, see Sect. 16.7.2 of Chap. 16. Any quantity  $\phi$  that depends on space and time will be written as  $\phi(x, y, z, t)$ . In most situations the governing equations will

be partial differential equations (PDEs). Naturally, these will involve partial derivatives for which we use the notation

$$\phi_t \equiv \frac{\partial \phi}{\partial t}, \quad \phi_x \equiv \frac{\partial \phi}{\partial x}, \quad \phi_y \equiv \frac{\partial \phi}{\partial y}, \quad \phi_z \equiv \frac{\partial \phi}{\partial z}.$$

We also recall some basic notation involving scalars and vectors. The *dot product* of two vectors  $\mathbf{A} = (a_1, a_2, a_3)$  and  $\mathbf{B} = (b_1, b_2, b_3)$  is the *scalar quantity*

$$\mathbf{A} \cdot \mathbf{B} = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

Given a scalar quantity  $\phi$  that depends on the spatial variables  $x, y, z$  the *gradient operator*  $\nabla$  as applied to  $\phi$  is the vector

$$\text{grad } \phi \equiv \nabla \phi \equiv \left( \frac{\partial \phi}{\partial x}, \frac{\partial \phi}{\partial y}, \frac{\partial \phi}{\partial z} \right).$$

The *divergence operator* applies to vectors and the result is a scalar quantity; for a vector  $\mathbf{A}$ , the divergence of  $\mathbf{A}$  is

$$\text{div } \mathbf{A} \equiv \nabla \cdot \mathbf{A} \equiv \frac{\partial a_1}{\partial x} + \frac{\partial a_2}{\partial y} + \frac{\partial a_3}{\partial z}.$$

## 1.1 The Euler Equations

In this section we consider the time-dependent Euler equations. These are a system of non-linear hyperbolic conservation laws that govern the dynamics of a compressible material, such as gases or liquids at high pressures, for which the effects of body forces, viscous stresses and heat flux are neglected.

There is some freedom in choosing a set of variables to describe the flow under consideration. A possible choice is the so called *primitive variables* or *physical variables*, namely,  $\rho(x, y, z, t)$  = density or mass density,  $p(x, y, z, t)$  = pressure,  $u(x, y, z, t)$  =  $x$ -component of velocity,  $v(x, y, z, t)$  =  $y$ -component of velocity,  $w(x, y, z, t)$  =  $z$ -component of velocity. The velocity vector is  $\mathbf{V} = (u, v, w)$ . An alternative choice is provided by the so called *conserved variables*. These are the mass density  $\rho$ , the  $x$ -momentum component  $\rho u$ , the  $y$ -momentum component  $\rho v$ , the  $z$ -momentum component  $\rho w$  and the total energy per unit mass  $E$ . Physically, these conserved quantities result naturally from the application of the fundamental laws of conservation of mass, Newton's Second Law and the law of conservation of energy. Computationally, there are some advantages in expressing the governing equations in terms of the conserved variables. This gives rise to a large class of numerical methods called *conservative methods*, which will be studied later in this book. We next state the equations in terms of the conserved variables under the assumption that the quantities involved are sufficiently smooth to allow for the operation of differentiation to be defined. Later we remove this smoothness constraint to allow for solutions containing discontinuities, such as shock waves.

### 1.1.1 Conservation-Law Form

The five governing conservation laws are

$$\rho_t + (\rho u)_x + (\rho v)_y + (\rho w)_z = 0 , \quad (1.1)$$

$$(\rho u)_t + (\rho u^2 + p)_x + (\rho uv)_y + (\rho uw)_z = 0 , \quad (1.2)$$

$$(\rho v)_t + (\rho uv)_x + (\rho v^2 + p)_y + (\rho vw)_z = 0 , \quad (1.3)$$

$$(\rho w)_t + (\rho uw)_x + (\rho vw)_y + (\rho w^2 + p)_z = 0 , \quad (1.4)$$

$$E_t + [u(E + p)]_x + [v(E + p)]_y + [w(E + p)]_z = 0 . \quad (1.5)$$

Here  $E$  is the total energy per unit volume

$$E = \rho \left( \frac{1}{2} \mathbf{V}^2 + e \right) , \quad (1.6)$$

where

$$\frac{1}{2} \mathbf{V}^2 = \frac{1}{2} \mathbf{V} \cdot \mathbf{V} = \frac{1}{2} (u^2 + v^2 + w^2)$$

is the *specific kinetic energy* and  $e$  is the *specific internal energy*. One generally refers to the full system (1.1)–(1.5) as the Euler equations, although strictly speaking the Euler equations are just (1.2)–(1.4).

The conservation laws (1.1)–(1.5) can be expressed in a very compact notation by defining a column vector  $\mathbf{U}$  of conserved variables and flux vectors  $\mathbf{F}(\mathbf{U})$ ,  $\mathbf{G}(\mathbf{U})$ ,  $\mathbf{H}(\mathbf{U})$  in the  $x$ ,  $y$  and  $z$  directions, respectively. The equations now read

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_y + \mathbf{H}(\mathbf{U})_z = \mathbf{0} , \quad (1.7)$$

with

$$\left. \begin{aligned} \mathbf{U} &= \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ E \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ \rho uw \\ u(E + p) \end{bmatrix}, \\ \mathbf{G} &= \begin{bmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ \rho vw \\ v(E + p) \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} \rho w \\ \rho uw \\ \rho vw \\ \rho w^2 + p \\ w(E + p) \end{bmatrix}. \end{aligned} \right\} \quad (1.8)$$

It is important to note that  $\mathbf{F} = \mathbf{F}(\mathbf{U})$ ,  $\mathbf{G} = \mathbf{G}(\mathbf{U})$ ,  $\mathbf{H} = \mathbf{H}(\mathbf{U})$ ; that is, the flux vectors are to be regarded as functions of the conserved variable vector  $\mathbf{U}$ . Any set of PDEs written in the form (1.7) is called a system of *conservation laws*. As partial derivatives are involved we say that (1.7) is a system of conservation laws in differential form. The differential formulation assumes smooth solutions, that is, partial derivatives are assumed to exist. There are other ways of expressing conservation laws in which the smoothness assumption is relaxed to include discontinuous solutions.

### 1.1.2 Other Compact Forms

An even more compact way of expressing equations (1.1)–(1.5) involves tensors. First note that the spatial derivatives in (1.1) can be expressed in terms of the divergence operator e.g.

$$\operatorname{div}(\rho \mathbf{V}) = \nabla \cdot (\rho \mathbf{V}) = (\rho u)_x + (\rho v)_y + (\rho w)_z .$$

Thus equation (1.1) for conservation of mass can be written as

$$\rho_t + \nabla \cdot (\rho \mathbf{V}) = 0 . \quad (1.9)$$

As the divergence operator may also be applied to tensors, the three momentum equations for conservation of momentum can be written in compact form as

$$(\rho \mathbf{V})_t + \nabla \cdot (\rho \mathbf{V} \otimes \mathbf{V} + p \mathbf{I}) = 0 , \quad (1.10)$$

where  $\mathbf{V} \otimes \mathbf{V}$  is the tensor product and  $\mathbf{I}$  is the unit tensor. These are given respectively by

$$\mathbf{V} \otimes \mathbf{V} = \begin{pmatrix} u^2 & uv & uw \\ vu & v^2 & vw \\wu & wv & w^2 \end{pmatrix} , \quad \mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

The conservation of energy equation can be written as

$$E_t + \nabla \cdot [(E + p)\mathbf{V}] = 0 . \quad (1.11)$$

In fact the complete system of equations (1.9)–(1.11) can be written in divergence form as

$$\mathbf{U}_t + \nabla \cdot \mathbf{H} = \mathbf{0} , \quad (1.12)$$

where  $\mathbf{H}$  is the tensor

$$\mathbf{H} = \begin{bmatrix} \rho u & \rho u^2 + p & \rho vu & \rho wu & u(E + p) \\ \rho v & \rho uv & \rho v^2 + p & \rho wv & v(E + p) \\ \rho w & \rho uw & \rho vw & \rho w^2 + p & w(E + p) \end{bmatrix} . \quad (1.13)$$

Note that the rows of the tensor  $\mathbf{H}$  are the flux vectors  $\mathbf{F}$ ,  $\mathbf{G}$  and  $\mathbf{H}$ , understood as row vectors. For computational purposes it is the compact conservative form (1.7)–(1.8) of equations (1.1)–(1.5) that is most useful. In Chap. 3 we study some mathematical properties of the Euler equations and in Chap. 4 we solve exactly the Riemann problem for the one-dimensional Euler equations for ideal and covolume gases. Numerical methods for the Euler equations are discussed in Chaps. 6–12, 14 and 16.

## 1.2 Thermodynamic Considerations

The stated governing partial differential equations (1.1)–(1.5) for the dynamics of a compressible material are insufficient to completely describe the physical processes involved. There are more unknowns than equations and thus closure conditions are required. Physically, such conditions are statements related to the nature of the material or medium. Relation (1.6) defines the total energy  $E$  in terms of the velocity vector  $\mathbf{V}$  involved in equations (1.1)–(1.5) and a new variable  $e$ , the specific internal energy. One therefore requires another relation defining  $e$  in terms of quantities already given, such as pressure and density, as a closure condition. For some applications, or when more physical effects are added to the basic equations (1.1)–(1.5), other variables, such as temperature for instance, may need to be introduced.

Central to providing closure conditions is a discussion of the fundamentals of Thermodynamics. This introduces new physical variables and provides relations between variables. Under certain conditions the governing equations may be approximated so as to make such discussion of Thermodynamics unnecessary; two important examples are incompressible flows and isentropic flows [112]. The specific internal energy  $e$  has an important role in the *First Law of Thermodynamics*, while the entropy  $s$  is intimately involved with the *Second Law of Thermodynamics*. Entropy plays a fundamental role not just in establishing the governing equations but also at the level of their mathematical properties and the designing of numerical methods to solve them. In addition to the basic thermodynamic variables density  $\rho$ , pressure  $p$ , temperature  $T$ , specific internal energy  $e$  and entropy  $s$ , one may define other new variables that are combinations of these.

### 1.2.1 Units of Measure

A brief discussion of physical quantities and their units of measure is essential. We consistently adopt, unless otherwise stated, the *International System of Units* or *SI Units*. Three basic quantities are length ( $l$ ), mass ( $m$ ) and time ( $t$ ). The unit of measure of length is: one metre = 1 m. Submultiples are: one decimetre =  $10^{-1}$  m, one centimetre =  $10^{-2}$  m, one millimetre =  $10^{-3}$  m. Multiples are:  $10^1$  m,  $10^2$  m and  $10^3$  m = one kilometre. From length one can establish the units of measure of area: one square metre =  $1 \text{ m}^2$  and the units of measure of volume: one cubic metre =  $1 \text{ m}^3$ . The unit of measure of mass is: one kilogram = 1 kg. A useful submultiple is: one gram = 1 g =  $10^{-3}$  kg. As density is  $\rho = m/V$ , where  $V$  is the total volume of the system, the unit of measure of density is: one kilogram per cubic metre =  $1 \text{ kg/m}^3$  =  $1 \text{ kg m}^{-3}$ . The unit of measure of time is: one second = 1 s. The unit of measure of speed is: one metre per second =  $1 \text{ m/s} = 1 \text{ m s}^{-1}$ . The unit of measure of acceleration is: one metre per second per second =  $1 \text{ m/s}^2 = 1 \text{ m s}^{-2}$ . The unit of measure of force is: one Newton = 1 N. The Newton N is defined as the force required to give a mass of 1 kg an acceleration of  $1 \text{ m s}^{-2}$ . Newton's Second Law states

that  $\text{force} = \text{constant} \times \text{mass} \times \text{acceleration}$ . The value of the unit of force is then chosen so as to make  $\text{force} = 1$  when  $\text{constant} = 1$ . Therefore the unit of force (N) is:  $1 \text{ kg m s}^{-2}$ . We now give the unit of measure of pressure. Pressure  $p$  is the magnitude of force per unit area and therefore its unit of measure is: one Newton per square metre  $= 1 \text{ N m}^{-2} \equiv 1 \text{ Pa}$  : one Pascal. Two common units of pressure are 1 bar  $= 10^5 \text{ Pa}$  and 1 atm (atmosphere)  $= 101\,325 \text{ Pa}$ . An important rule in manipulating physical quantities is *dimensional consistency*. For example, in the expression  $\rho u^2 + p$  in the momentum equation (1.2), the dimensions of  $\rho u^2$  must be the same as those of (pressure)  $p$ . This is easily verified.

To introduce the unit of measure of energy we first recall the concept of *Work*. Work ( $W$ ) is done when a force produces a motion and is measured as  $W = \text{force} \times \text{distance moved in the direction of the force}$ . The unit of measure of work is: one Joule  $= 1 \text{ J}$ . One Joule is the work done when the point of application of a force of 1 N moves through a distance of 1 m in the direction of the force. As energy is the capacity to perform work, the unit of measure of energy is also one Joule. The temperature  $T$  will be measured in terms of the *Thermodynamic Scale* or the *Absolute Scale*, in which the unit of measure is: one kelvin  $= 1 \text{ K}$ .

Thermodynamic properties of a system that are proportional to the mass  $m$  of the system are called *extensive* properties. Examples are the total energy  $E$  and the total volume  $V$  of a system. Properties that are independent of  $m$  are called *intensive* properties; examples are temperature  $T$  and pressure  $p$ . Extensive properties may be converted to their *specific* (intensive) values by dividing that property by its mass  $m$ . For instance, from the total volume  $V$  we obtain the *specific volume*  $v = V/m$  (the reader is warned that  $v$  is also used for velocity component). As  $\rho = m/V$ , the specific volume is the reciprocal of density. The units of measure of other quantities will be given as they are introduced.

### 1.2.2 Equations of State (EOS)

A system in thermodynamic equilibrium can be completely described by the basic thermodynamic variables pressure  $p$  and specific volume  $v$ . A family of states in thermodynamic equilibrium may be described by a curve in the  $p$ - $v$  plane, each characterised by a particular value of a variable temperature  $T$ . Systems described by the  $p$ - $v$ - $T$  variables are usually called  $p$ - $v$ - $T$  systems. There are physical situations that require additional variables. Here we are only interested in  $p$ - $v$ - $T$  systems. In these, one can relate the variables via the *thermal equation of state*

$$T = T(p, v) . \quad (1.14)$$

Two more possible relations are

$$p = p(T, v) , \quad v = v(T, p) .$$

The  $p$ - $v$ - $T$  relationship changes from substance to substance. For *thermally ideal gases* one has the simple expression

$$T = \frac{pv}{R}, \quad (1.15)$$

where  $R$  is a constant which depends on the particular gas under consideration.

The First Law of Thermodynamics states that for a non-adiabatic system the change  $\Delta e$  in internal energy  $e$  in a process will be given by  $\Delta e = \Delta W + \Delta Q$ , where  $\Delta W$  is the work done on the system and  $\Delta Q$  is the heat transmitted to the system. Taking the work done as  $dW = -pdv$  one may write

$$dQ = de + pdv. \quad (1.16)$$

The internal energy  $e$  can also be related to  $p$  and  $v$  via a *caloric equation of state*

$$e = e(p, v). \quad (1.17)$$

Two more possible ways of expressing the  $p$ - $v$ - $e$  relationship are

$$p = p(v, e), \quad v = v(e, p).$$

For a *calorically ideal gas* one has the simple expression

$$e = \frac{pv}{\gamma - 1} = \frac{p}{\rho(\gamma - 1)}, \quad (1.18)$$

where  $\gamma$  is a constant that depends on the particular gas under consideration.

The thermal and caloric equations of state for a given material are closely related. Both are necessary for a complete description of the thermodynamics of a system. Choosing a thermal EOS does restrict the choice of a caloric EOS but does not determine it. Note that for the Euler equations (1.1)–(1.5) one only requires a caloric EOS, e.g.  $p = p(\rho, e)$ , unless temperature  $T$  is needed for some other purpose, in which case a thermal EOS needs to be given explicitly.

### 1.2.3 Other Variables and Relations

The *entropy s* results as follows. We first introduce an integrating factor  $1/T$  so that the expression

$$de + pdv = \left( \frac{\partial e}{\partial v} + p \right) dv + \frac{\partial e}{\partial p} dp$$

in (1.16) becomes an exact differential. Then the Second Law of Thermodynamics introduces a new variable  $s$ , called entropy, via the relation

$$Tds = de + pdv. \quad (1.19)$$

For any process the change in entropy is  $\Delta s = \Delta s_0 + \Delta s_i$ , where  $\Delta s_0$  is the entropy carried into the system through the boundaries of the system and  $\Delta s_i$  is the entropy generated in the system during the process. Examples of entropy-generating mechanisms are heat transfer and viscosity, such as may operate within the internal structure of shock waves. The Second Law of Thermodynamics states that  $\Delta s_i > 0$  in any irreversible process. Only in a reversible process is  $\Delta s_i = 0$ .

Another variable of interest is the *specific enthalpy*  $h$ . This is defined in terms of other thermodynamic variables, namely

$$h = e + pv . \quad (1.20)$$

One can also establish various relationships amongst the basic thermodynamic variables already defined. For instance from (1.19)

$$de = Tds - pdv , \quad (1.21)$$

that is to say, one may choose to express the internal energy  $e$  in terms of the variables appearing in the differentials, i.e.

$$e = e(s, v) . \quad (1.22)$$

Also, taking the differential of (1.20) we have  $dh = de + pdv + vdp$ , which by virtue of (1.21) becomes

$$dh = Tds + vdp , \quad (1.23)$$

and thus we can choose to define  $h$  in terms of  $s$  and  $p$ , i.e.

$$h = h(s, p) . \quad (1.24)$$

Relations (1.22) and (1.24) are called *canonical equations of state* and, unlike the thermal and caloric equations of state (1.14) and (1.17), each of these provides a complete description of the Thermodynamics. For instance, given (1.22) in which  $e$  is a function of  $s$  and  $v$  (independent variables) the pressure  $p$  and temperature  $T$  follow as

$$p = - \left( \frac{\partial e}{\partial v} \right)_s , \quad T = \left( \frac{\partial e}{\partial s} \right)_v . \quad (1.25)$$

Relations (1.25) follow from comparing

$$de = \left( \frac{\partial e}{\partial s} \right)_v ds + \left( \frac{\partial e}{\partial v} \right)_s dv$$

with equation (1.21). It is conventional in Thermodynamics to specify clearly the independent variables in partial differentiation, as changes of variables often take place. In (1.25), obviously the independent variables are  $s$  and  $v$ , as is also indicated in (1.22). For instance, the first partial derivative in (1.25)

means differentiation of  $e$  with respect to  $v$  while holding  $s$  constant; the second partial derivative in (1.25) means differentiation of  $e$  with respect to  $s$  while holding  $v$  constant. In a similar manner, equation (1.24) (where  $s$  and  $p$  are the independent variables) produces  $T$  and  $v$  from relation (1.23) and

$$dh = \left( \frac{\partial h}{\partial s} \right)_p ds + \left( \frac{\partial h}{\partial p} \right)_s dp .$$

Hence,

$$T = \left( \frac{\partial h}{\partial s} \right)_p , v = \left( \frac{\partial h}{\partial p} \right)_s . \quad (1.26)$$

The Helmholtz *free energy*  $f$  is defined as

$$f = e - Ts . \quad (1.27)$$

A corresponding canonical EOS is

$$f = f(v, T) ,$$

from which one can obtain

$$s = - \left( \frac{\partial f}{\partial T} \right)_v , p = - \left( \frac{\partial f}{\partial v} \right)_T . \quad (1.28)$$

Two more quantities can be defined if a thermal EOS  $v = v(p, T)$  is given. These are the *volume expansivity*  $\alpha$  (or expansion coefficient) and the *isothermal compressibility*  $\beta$ , namely

$$\alpha = \frac{1}{v} \left( \frac{\partial v}{\partial T} \right)_p , \beta = - \frac{1}{v} \left( \frac{\partial v}{\partial p} \right)_T . \quad (1.29)$$

Using equations (1.28) and (1.27) we obtain

$$\left( \frac{\partial s}{\partial v} \right)_T = \left( \frac{\partial p}{\partial T} \right)_v = \frac{\alpha}{\beta} ,$$

from which it can be shown that

$$\left( \frac{\partial e}{\partial v} \right)_T = \frac{\alpha T - \beta p}{\beta} . \quad (1.30)$$

The *heat capacity at constant pressure*  $c_p$  and the *heat capacity at constant volume*  $c_v$  (specific heat capacities) are now introduced. In general, when an addition of heat  $dQ$  changes the temperature by  $dT$  the ratio  $c = dQ/dT$  is called the *heat capacity* of the system. For a process at constant pressure relation (1.16) becomes

$$dQ = de + d(pv) = dh ,$$

where definition (1.20) has been used. The heat capacity  $c_p$  at constant pressure becomes  $c_p = dQ/dT = dh/dT$ . From (1.23), since  $dp = 0$ ,  $dh = Tds$ . Assuming  $h = h(T, p)$  we obtain

$$c_p = \left( \frac{\partial h}{\partial T} \right)_p = T \left( \frac{\partial s}{\partial T} \right)_p . \quad (1.31)$$

The heat capacity  $c_v$  at constant volume may be written, following a similar argument, as

$$c_v = \left( \frac{\partial e}{\partial T} \right)_v = T \left( \frac{\partial s}{\partial T} \right)_v . \quad (1.32)$$

*The speed of sound* is another variable of fundamental interest. For flows in which particles undergo unconstrained thermodynamic equilibrium one defines a new state variable  $a$ , called the equilibrium speed of sound or just speed of sound. Given a caloric equation of state

$$p = p(\rho, s) , \quad (1.33)$$

one defines the speed of sound  $a$  as

$$a = \sqrt{\left( \frac{\partial p}{\partial \rho} \right)_s} . \quad (1.34)$$

This basic definition can be transformed in various ways using established thermodynamic relations. For instance, given a caloric EOS in the form  $h = h(p, \rho)$ , from (1.23) we can write

$$\left( \frac{\partial h}{\partial p} \right)_\rho dp + \left( \frac{\partial h}{\partial \rho} \right)_p d\rho = T ds + \frac{1}{\rho} \left[ \left( \frac{\partial p}{\partial \rho} \right)_s d\rho + \left( \frac{\partial p}{\partial s} \right)_\rho ds \right] .$$

Setting  $ds = 0$  and using definition (1.34) we obtain

$$a^2 = - \frac{\left( \frac{\partial h}{\partial \rho} \right)_p}{\left( \frac{\partial h}{\partial p} \right)_\rho - \frac{1}{\rho}} .$$

For a thermally ideal gas  $h = h(T)$  and thus

$$a^2 = \frac{\left( \frac{\partial h}{\partial T} \right)_p \left( \frac{\partial T}{\partial \rho} \right)_p}{\left( \frac{\partial h}{\partial T} \right)_p \left( \frac{\partial T}{\partial p} \right)_\rho - \frac{1}{\rho}} .$$

From (1.31)  $\left( \frac{\partial h}{\partial T} \right)_p = c_p$  and, if the thermal EOS (1.15) is acceptable, we obtain

$$a = \sqrt{\gamma(T)RT} = \sqrt{\frac{\gamma p}{\rho}} . \quad (1.35)$$

For a general material the caloric EOS is a functional relationship involving the variables  $p$ - $\rho$ - $e$ . One may also use the specific volume  $v = 1/\rho$  instead of density  $\rho$ . The derived expression for the speed of sound  $a$  depends on the choice of independent variables. Two possible choices and their respective expressions for  $a$  are

$$\left. \begin{aligned} p &= p(\rho, e) , \quad a = \sqrt{\frac{p}{\rho^2} p_e + p_\rho} , \\ e &= e(\rho, p) , \quad a = \sqrt{\frac{p}{\rho^2 e_p} - \frac{e_\rho}{e_p}} , \end{aligned} \right\} \quad (1.36)$$

where subscripts denote partial derivatives.

#### 1.2.4 Ideal Gases

We consider gases obeying the ideal thermal EOS

$$pV = n\mathcal{R}T , \quad (1.37)$$

where  $V$  is the volume,  $\mathcal{R} = 8.134 \times 10^3 \text{ J kilomole}^{-1}\text{K}^{-1}$ , called the *Universal Gas Constant*, and  $T$  is the temperature measured in degrees kelvin (K). Two more universal constants are now introduced. Recall that a mole of a substance is numerically equal to  $\omega$  gram and contains  $6.02 \times 10^{23}$  particles of that substance, where  $\omega$  is the relative atomic mass (RAM) or relative molecular mass (RMM); 1 kilomole =  $\omega$  kg. One kilomole of a substance contains  $N_A = 6.02 \times 10^{26}$  particles of that substance. The constant  $N_A$  is called the *Avogadro Number*. Sometimes this number is given in terms of one mole. The *Boltzmann Constant*  $k$  is now defined as  $k = \mathcal{R}N_A$ ;  $n$  in (1.37) is the number of kilomoles in volume  $V$ , that is  $n = N/N_A$  and  $N$  is the number of molecules. On division by the mass  $m = n\omega$  we have

$$pv = RT , \quad R = \frac{\mathcal{R}}{\omega} , \quad (1.38)$$

where  $R$  is called the *Specific Gas Constant* or simply *Gas Constant*. Solving for  $v$  we write the ideal gas thermal equation of state as

$$v = v(T, p) = \frac{RT}{p} . \quad (1.39)$$

The volume expansivity  $\alpha$  and the isothermal compressibility  $\beta$  defined by (1.29) become

$$\alpha = \frac{1}{T} , \quad \beta = \frac{1}{p} . \quad (1.40)$$

Substitution of these into (1.30) gives

$$\left( \frac{\partial e}{\partial v} \right)_T = 0 .$$

This means that if the ideal thermal EOS (1.39) is assumed, then it follows that the internal energy  $e$  is a function of temperature alone, that is  $e = e(T)$ . In the particular case in which

$$e = c_v T , \quad (1.41)$$

where the specific heat capacity  $c_v$  is a constant, one speaks of a *calorically ideal gas*, or a *polytropic gas*.

It is possible to relate  $c_p$  and  $c_v$  via the general expression

$$c_p = c_v + \frac{\alpha^2 T v}{\beta} . \quad (1.42)$$

For a thermally ideal gas equations (1.39) and (1.40) apply and thus

$$c_p - c_v = R . \quad (1.43)$$

A necessary condition for thermal stability is  $c_v > 0$  and for mechanical stability  $\beta > 0$  [35]. From (1.42) the following inequalities result

$$c_p > c_v > 0 . \quad (1.44)$$

The *ratio of specific heats*  $\gamma$ , or *adiabatic exponent*, is defined as

$$\gamma = \frac{c_p}{c_v} , \quad (1.45)$$

which if used in conjunction with (1.43) gives

$$c_p = \frac{\gamma R}{\gamma - 1} , \quad c_v = \frac{R}{\gamma - 1} . \quad (1.46)$$

For a calorically ideal gas (polytropic gas)  $\gamma$  is a constant and for a thermally ideal gas  $\gamma$  is a function of temperature, i.e.  $\gamma = \gamma(T)$ .

In order to determine the caloric EOS (1.41) we need to determine the specific heat capacities,  $c_v$  in particular. Molecular Theory and the principle of equipartition of energy [439] can also provide an expression for the specific internal energy of a molecule. In general a molecule, however complex, has  $M$  degrees of freedom, of which three are translational. Other possible degrees of freedom are rotational and vibrational. From Molecular Theory it can be shown that if the energy associated with any degree of freedom is a quadratic function in the appropriate variable expressing that degree of freedom, then the mean value of the energy is  $\frac{1}{2}kT$  where  $k$  is the *Boltzmann constant*. Moreover, from the principle of equipartition of energy this is the same for each degree of freedom. Therefore, the mean total energy of a molecule is  $\bar{e} = \frac{1}{2}MkT$ , and for  $N$  molecules we have

$$N\bar{e} = \frac{1}{2}MNkT ,$$

from which the specific internal energy is

$$e = \frac{1}{2}MRT .$$

Use of (1.32) gives directly

$$c_v = \left( \frac{\partial e}{\partial T} \right)_v = \frac{1}{2}MR ,$$

and thus we obtain

$$c_p = \frac{M+2}{2}R .$$

The ratio of specific heats becomes

$$\gamma = \frac{M+2}{M} . \quad (1.47)$$

From the thermal EOS for ideal gases (1.39) we have

$$e = \frac{1}{2}Mpv .$$

But from (1.43) and (1.46)

$$M = \frac{2}{\gamma - 1}$$

and hence

$$e = \frac{pv}{(\gamma - 1)} = \frac{p}{(\gamma - 1)\rho} , \quad (1.48)$$

which is the expression for the specific internal energy advanced in (1.18).

The theoretical expressions for  $c_p$ ,  $c_v$  and  $\gamma$  in terms of  $R$  and  $M$  are found to be very accurate for monatomic gases, for which  $M = 3$  (three translational degrees of freedom). For polyatomic gases rotational and vibrational degrees of freedom contribute to  $M$  but now the expressions might be rather inaccurate when compared with experimental data. A strong dependence on  $T$  is observed. However, the inequality  $1 < \gamma < \frac{5}{3}$ , predicted from (1.47) for the limiting values  $M = 3$  and  $M = \infty$ , holds true.

### 1.2.5 Covolume and van der Waal Gases

A very simple generalisation of the ideal-gas thermal EOS,  $pv = RT$ , is the so-called *covolume equation of state*

$$p(v - b) = RT , \quad (1.49)$$

where  $b$  is called the *covolume* and in SI units has dimensions of  $\text{m}^3\text{kg}^{-1}$ . This EOS applies to dense gases at high pressure for which the volume occupied by the molecules themselves is no longer negligible. There is therefore a reduction in the volume available to molecular motion. This type of correction to the ideal gas EOS is said to have first been suggested by Clausius. Hirn is credited with first having written down EOS (1.49). Sometimes, this equation is also called the Noble–Abel EOS. In the study of propulsion systems, gaseous combustion products at very high densities are reasonably well described by the covolume EOS. In its simplest version the covolume  $b$  is a constant and is determined experimentally or from equilibrium thermochemical calculations. Corner [141] reports on good experimental results for a range of solid propellants and observes that  $b$  changes very little, usually in the range  $0.9 \times 10^{-3} \leq b \leq 1.1 \times 10^{-3}$ . The best values of  $b$  lead to errors of no more than 2% and thus there is some justification in using (1.49) with  $b$  constant. A more accurate covolume EOS defines  $b$  as a function of  $\rho$ , i.e.  $b = b(\rho)$ . Such dependence of  $b$  on  $\rho$  can be given in either tabular or algebraic form. A simple example of an algebraic form is  $b(\rho) = \exp^{-0.4\rho}$ , for  $\rho < 2 \text{ g cm}^{-3}$ . The thermal covolume EOS (1.49) leads to a caloric covolume EOS  $e = e(p, \rho)$  with a corresponding sound speed  $a$ . These are given by

$$e = \frac{p(1 - b\rho)}{\rho(\gamma - 1)}, \quad a = \left[ \frac{\gamma p}{(1 - b\rho)\rho} \right]^{\frac{1}{2}}, \quad (1.50)$$

where  $\gamma$  is the ratio of specific heats as before.

The covolume EOS (1.49) can be further corrected to account for the forces of attraction between molecules, the van der Waal forces. These are neglected in both the ideal and covolume equations of state. Accounting for such forces results in a reduction of the pressure by an amount  $c/v^2$ , where  $c$  is a quantity that depends on the particular gas under consideration. Thus from (1.49) the pressure is corrected as

$$p = \frac{RT}{v - b} - \frac{c}{v^2}.$$

Then we can write

$$(p + \frac{c}{v^2})(v - b) = RT. \quad (1.51)$$

This is generally known as the *van der Waal's equation of state* for real gases.

General background on Thermodynamics and equations of state can be found in virtually any textbook on Thermodynamics or Gas Dynamics. We particularly recommend the book by Sears and Salinger [439], Chap. 1 of the book on Gas Dynamics by Becker [35] and the book by Clarke and McChesney [119]. The review paper by Menikoff and Plohr [349] is highly relevant to the themes of this book. A useful reference on equations of state for combustion problems is [573].

So far, we have presented the Euler equations for the dynamics of a compressible medium along with some elementary notions on Thermodynamics

so that a closed system is obtained. Given initial and boundary conditions the conservation equations can be solved. In this book we are interested in numerical methods to solve the governing equations.

### 1.3 Viscous Stresses

Here we augment the Euler equations (1.7) by adding the physical effects of viscosity. Strictly speaking it is only the momentum equations in (1.7) that are modified. The stresses in a fluid, given by a tensor  $\mathbf{S}$ , are due to the effects of the thermodynamic pressure  $p$  and the viscous stresses. Thus the stress tensor can be written as

$$\mathbf{S} = -p\mathbf{I} + \Pi, \quad (1.52)$$

where  $p\mathbf{I}$  is the spherically symmetric tensor due to  $p$ ,  $\mathbf{I}$  is the unit tensor as in (1.10) and  $\Pi$  is the viscous stress tensor. It is desirable to express  $\mathbf{S}$  in terms of flow variables already defined. For the pressure contribution this has already been achieved by defining  $p$  in terms of other thermodynamic variables via an equation of state. Recall that equations of state are approximate statements about the nature of a material. In defining the viscous stress contribution  $\Pi$  one may resort to the *Newtonian approximation*, whereby  $\Pi$  is related to the derivatives of the velocity field  $\mathbf{V} = (u, v, w)$  via the *deformation tensor*

$$\mathbf{D} = \begin{bmatrix} u_x & \frac{1}{2}(v_x + u_y) & \frac{1}{2}(w_x + u_z) \\ \frac{1}{2}(u_y + v_x) & v_y & \frac{1}{2}(w_y + v_z) \\ \frac{1}{2}(u_z + w_x) & \frac{1}{2}(v_z + w_y) & w_z \end{bmatrix}.$$

The Newtonian assumption is an idealisation in which the relationship between  $\Pi$  and  $\mathbf{D}$  is linear and homogeneous, that is  $\Pi$  will vanish only if  $\mathbf{D}$  vanishes, and the medium is isotropic with respect to this relation; an isotropic medium is that in which there are no preferred directions. By denoting the stress tensor by

$$\Pi = \begin{bmatrix} \tau^{xx} & \tau^{xy} & \tau^{xz} \\ \tau^{yx} & \tau^{yy} & \tau^{yz} \\ \tau^{zx} & \tau^{zy} & \tau^{zz} \end{bmatrix}, \quad (1.53)$$

the Newtonian approximation becomes

$$\Pi = 2\eta\mathbf{D} + (\eta_b - \frac{2}{3}\eta)(\operatorname{div}\mathbf{V})\mathbf{I}, \quad (1.54)$$

or in full

$$\left. \begin{aligned} \tau^{xx} &= \frac{4}{3}\eta u_x - \frac{2}{3}\eta(v_y + w_z) + \eta_b \operatorname{div} \mathbf{V}, \\ \tau^{yy} &= \frac{4}{3}\eta v_y - \frac{2}{3}\eta(w_z + u_x) + \eta_b \operatorname{div} \mathbf{V}, \\ \tau^{zz} &= \frac{4}{3}\eta w_z - \frac{2}{3}\eta(u_x + v_y) + \eta_b \operatorname{div} \mathbf{V}, \\ \tau^{xy} &= \tau^{yx} = \eta(u_y + v_x), \\ \tau^{yz} &= \tau^{zy} = \eta(v_z + w_y), \\ \tau^{zx} &= \tau^{xz} = \eta(w_x + u_z). \end{aligned} \right\} \quad (1.55)$$

In the Newtonian relationship (1.54) there are two scalar quantities that are still undetermined, these are the *coefficient of shear viscosity*  $\eta$  and the *coefficient of bulk viscosity*  $\eta_b$ . Approximate expressions for these are obtained from experimentation and results from Molecular Theory. In particular, for monatomic gases Molecular Theory based on the hard sphere assumption gives  $\eta_b = 0$ , which is found to agree well with experiment. For polyatomic gases  $\eta_b \neq 0$  and appropriate values for  $\eta_b$  are to be given experimentally. Concerning the coefficient of shear viscosity  $\eta$ , it is observed that, as long as temperatures are not too high,  $\eta$  depends strongly on temperature and only slightly on pressure. Again, Molecular Theory and experimentation suggest that  $\eta$  be proportional to  $T^n$ ; in fact  $n = \frac{1}{2}$  in Molecular Theory. A relatively accurate relation between  $\eta$  and  $T$  is the Sutherland formula

$$\eta = C_1 \left[ 1 + \frac{C_2}{T} \right]^{-1} \sqrt{T}, \quad (1.56)$$

where  $C_1$  and  $C_2$  are two experimentally adjustable constants. When  $T$  is measured in kelvin,  $\eta$  has the units of  $\text{kg m}^{-1}\text{s}^{-1}$ . For the case of air one has

$$C_1 = 1.46 \times 10^{-6}, \quad C_2 = 112 \text{ K}.$$

Sutherland's formula describes the dependence of  $\eta$  on  $T$  rather well for a wide range of temperatures, provided no dissociation or ionisation take place. These phenomena occur at very high temperatures where the dependence of  $\eta$  on pressure  $p$ , in addition to temperature  $T$ , cannot be neglected. Useful background on high temperature gas dynamics is found in the book by Anderson [9] and in the book by Clarke and McChesney [119].

In summary, the Navier–Stokes equations (momentum equations) can now be written in differential conservation law form as

$$(\rho \mathbf{V})_t + \nabla \cdot (\rho \mathbf{V} \otimes \mathbf{V} + p \mathbf{l} - \Pi) = \mathbf{0}, \quad (1.57)$$

where  $\Pi$  is given by (1.55) with  $\eta_b = 0$ . Compare with the Euler equations (1.10).

## 1.4 Heat Conduction

Influx of energy contributes to the rate of change of total energy  $E$ . We denote by  $\mathbf{Q} = (q_1, q_2, q_3)^T$  the energy flux vector, which results from (i) heat flow due to temperature gradients, (ii) diffusion processes in gas mixtures and (iii) radiation. Here we only consider effect (i) above.  $\mathbf{Q}$  is identical to the heat flux vector caused by temperature gradients. In a similar manner to that in which viscous stresses were related to gradients of the velocity vector  $\mathbf{V}$ , one can relate  $\mathbf{Q}$  to gradients of temperature  $T$  via Fourier's heat conduction law

$$\mathbf{Q} = -\kappa \nabla T , \quad (1.58)$$

where  $\kappa$  is a positive scalar quantity called the *coefficient of thermal conductivity* or just *thermal conductivity*, and is yet to be determined. Note the analogy between  $\eta$  and  $\kappa$ . This analogy between  $\eta$  and  $\kappa$  goes further in that  $\kappa$ , just as  $\eta$ , depends on  $T$  but only slightly on pressure  $p$ . In fact, Molecular Theory says that  $\kappa$  is directly proportional to  $\eta$ . Under the assumption that the specific heat at constant pressure  $c_p$  is constant, the dimensionless quantity

$$P_r \equiv \frac{c_p \eta}{\kappa} \quad (1.59)$$

is a constant, and is called the Prandtl number. For monatomic gases  $P_r$  is very nearly constant. For air in the temperature range  $200 \text{ K} \leq T \leq 1000 \text{ K}$   $P_r$  differs only slightly from its mean value of 0.7. A formula attributed to Eucken [35] relates  $P_r$  to the ratio of specific heats  $\gamma$  via

$$P_r = \frac{4\gamma}{9\gamma - 5} , \quad (1.60)$$

to account for departures from calorically ideal gas behaviour.

When the effects of viscosity and heat conduction are added to the basic Euler equations (1.7) one has the Navier–Stokes equations with heat conduction

$$\mathbf{U}_t + \mathbf{F}_x^a + \mathbf{G}_y^a + \mathbf{H}_z^a = \mathbf{F}_x^d + \mathbf{G}_y^d + \mathbf{H}_z^d , \quad (1.61)$$

where  $\mathbf{U}$  is the vector of conserved variables, the flux vectors  $\mathbf{F}^a$ ,  $\mathbf{G}^a$  and  $\mathbf{H}^a$  are the inviscid fluxes (a stands for advection) for the Euler equations as given by (1.8) and the respective flux vectors  $\mathbf{F}^d$ ,  $\mathbf{G}^d$  and  $\mathbf{H}^d$  (d stands for diffusion) due to viscosity and heat conduction are

$$\left. \begin{aligned} \mathbf{F}^d &= \begin{bmatrix} 0 \\ \tau^{xx} \\ \tau^{xy} \\ \tau^{xz} \\ u\tau^{xx} + v\tau^{xy} + w\tau^{xz} - q_1 \end{bmatrix}, \\ \mathbf{G}^d &= \begin{bmatrix} 0 \\ \tau^{yx} \\ \tau^{yy} \\ \tau^{yz} \\ u\tau^{yx} + v\tau^{yy} + w\tau^{yz} - q_2 \end{bmatrix}, \\ \mathbf{H}^d &= \begin{bmatrix} 0 \\ \tau^{zx} \\ \tau^{zy} \\ \tau^{zz} \\ u\tau^{zx} + v\tau^{zy} + w\tau^{zz} - q_3 \end{bmatrix}. \end{aligned} \right\} \quad (1.62)$$

The form of the equations given by (1.61) splits the effect of advection on the left-hand side from those of viscous diffusion and heat conduction on the right-hand side. For numerical purposes, the particular form of the equations adopted depends largely on the numerical technique to be used to solve the equations. One possible approach is to split the advection effects from those of viscous diffusion and heat conduction during a small time interval  $\Delta t$ , in which case form (1.61) is perfectly adequate. An alternative form is obtained by combining the fluxes due to advection, viscous diffusion and heat conduction into new fluxes so that the governing equations look formally like a homogeneous system (zero right-hand side) of conservation laws

$$\left. \begin{aligned} \mathbf{U}_t + \mathbf{F}_x + \mathbf{G}_y + \mathbf{H}_z &= \mathbf{0}, \\ \mathbf{F} &= \mathbf{F}^a - \mathbf{F}^d, \quad \mathbf{G} = \mathbf{G}^a - \mathbf{G}^d, \quad \mathbf{H} = \mathbf{H}^a - \mathbf{H}^d. \end{aligned} \right\} \quad (1.63)$$

This form is only justified if the numerical method employed actually exploits the coupling of advection, viscosity and heat conduction when defining numerical approximations to the flux vectors  $\mathbf{F}$ ,  $\mathbf{G}$ , and  $\mathbf{H}$  in (1.63).

## 1.5 Integral Form of the Equations

The actual derivation of the governing equations, such as the Euler and Navier–Stokes equations stated earlier, is based on integral relations on control volumes and their boundaries. The differential form of the equations results from further assumptions on the flow variables (smoothness). In the absence of viscous diffusion and heat conduction one obtains the Euler equations. These admit discontinuous solutions and the smoothness assumption that leads to

the differential form no longer holds true. Thus one must return to the more fundamental integral form involving integrals over control volumes and their boundaries. From a computational point of view there is another good reason for returning to the integral form of the equations. Discretised domains result naturally in finite control volumes or computational cells. Local enforcement of the fundamental equations in these volumes lead to *Finite Volume* numerical methods.

### 1.5.1 Time Derivatives

Before proceeding to the derivation of the equations in integral form we review some preliminary concepts that are needed. Consider a scalar field function  $\phi(x, y, z, t)$ , then the time rate of change of  $\phi$  as registered by an observer moving with the fluid velocity  $\mathbf{V} = (u, v, w)$  is given by

$$\frac{D\phi}{Dt} = \frac{\partial\phi}{\partial t} + \mathbf{V} \cdot \operatorname{grad}\phi . \quad (1.64)$$

This time derivative  $D/Dt$  following a particle is usually called the *substantial derivative* or material derivative. The first term  $\partial\phi/\partial t$  in (1.64) denotes the partial derivative of  $\phi$  with respect to time and represents the *local rate of change of  $\phi$* ; the second term in (1.64) is the *convective rate of change*. The operator  $D/Dt$  can also be applied to vectors in a component-wise manner, in which case equation (1.64) is to be interpreted as a vector equation. In particular, we can obtain the substantial derivative of  $\mathbf{V} = (u, v, w)$ , that is

$$\frac{D\mathbf{V}}{Dt} = \frac{\partial\mathbf{V}}{\partial t} + \mathbf{V} \cdot \operatorname{grad}\mathbf{V} , \quad (1.65)$$

which in full becomes

$$\begin{pmatrix} \frac{Du}{Dt} \\ \frac{Dv}{Dt} \\ \frac{Dw}{Dt} \end{pmatrix}^T = \begin{pmatrix} \frac{\partial u}{\partial t} \\ \frac{\partial v}{\partial t} \\ \frac{\partial w}{\partial t} \end{pmatrix}^T + (u, v, w) \cdot \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial v}{\partial x} & \frac{\partial w}{\partial x} \\ \frac{\partial u}{\partial y} & \frac{\partial v}{\partial y} & \frac{\partial w}{\partial y} \\ \frac{\partial u}{\partial z} & \frac{\partial v}{\partial z} & \frac{\partial w}{\partial z} \end{bmatrix} .$$

The symbol  $( )^T$  denotes transpose of  $( )$ . Actually, equation (1.65) is the acceleration vector of an element of a moving fluid. Let us now consider

$$\Psi(t) = \int \int \int_V \phi(x, y, z, t) dV , \quad (1.66)$$

where the integrand  $\phi$  is any scalar field function and the volume of integration  $V$  is enclosed by a piece-wise smooth boundary surface  $A$  that moves with the material under consideration. It can be shown that the material derivative of  $\Psi$  is given by

$$\frac{D\Psi}{Dt} = \int \int \int_V \frac{\partial \phi}{\partial t} dV + \int \int_A (\mathbf{n} \cdot \phi \mathbf{V}) dA , \quad (1.67)$$

where  $\mathbf{n} = (n_1, n_2, n_3)$  is the outward pointing unit vector normal to the surface  $A$ . The proof of statement (1.67) is based on a three dimensional generalisation of the following result:

$$\frac{d}{d\alpha} \int_{\xi_1(\alpha)}^{\xi_2(\alpha)} f(\xi, \alpha) d\xi = \int_{\xi_1(\alpha)}^{\xi_2(\alpha)} \frac{\partial f}{\partial \alpha} d\xi + f(\xi_2, \alpha) \frac{d\xi_2}{d\alpha} - f(\xi_1, \alpha) \frac{d\xi_1}{d\alpha} . \quad (1.68)$$

Expression (1.67) can be generalised to vectors  $\Psi(x, y, z, t)$  as follows:

$$\frac{D\Psi}{Dt} = \int \int \int_V \frac{\partial \Phi}{\partial t} dV + \int \int_A \Phi(\mathbf{n} \cdot \mathbf{V}) dA . \quad (1.69)$$

The first term on the right hand side of (1.67) represents the local contribution of the field  $\phi$  to the time rate of change of  $\Psi(t)$ . The second term is the contribution due to the motion of the surface moving at the fluid velocity  $\mathbf{V}$ . The surface integral may be transformed to a volume integral by virtue of Gauss's theorem. This states that for any differentiable vector field  $\Phi = (\phi_1, \phi_2, \phi_3)$  and a volume  $V$  with smooth bounding surface  $A$  the following identity holds

$$\int \int \int_A (\mathbf{n} \cdot \Phi) dA = \int \int \int_V \operatorname{div} \Phi dV . \quad (1.70)$$

Gauss's theorem also applies to differentiable scalar fields and tensor fields.

### 1.5.2 Conservation of Mass

The law of conservation of mass can now be stated in integral form by identifying the scalar  $\phi$  in (1.66) and (1.67) as the density  $\rho$ . In this case  $\Psi(t)$  in (1.66) becomes the total mass in the volume  $V$ . By assuming that *no mass is generated or annihilated within  $V$*  and recalling that the surface  $A$  moves with the fluid velocity, which means that no mass flows across the surface, we have  $D\Psi/Dt = 0$ , or from (1.67)

$$\int \int \int_V \frac{\partial \rho}{\partial t} dV + \int \int_A \mathbf{n} \cdot (\rho \mathbf{V}) dA = 0 .$$

This is the integral form of the law of conservation of mass corresponding to the differential form (1.1). This integral conservation law may be generalised to include sources of mass, which will then appear as additional integral terms. A useful reinterpretation of the integral form results if we rewrite it as

$$\int \int \int_V \frac{\partial \rho}{\partial t} dV = - \int \int_A \mathbf{n} \cdot (\rho \mathbf{V}) dA . \quad (1.71)$$

If now  $V$  is a fixed control volume independent of time  $t$  then the left hand side of (1.71) becomes

$$\int \int \int_V \frac{\partial \rho}{\partial t} dV = \frac{d}{dt} \int \int \int_V \rho dV ,$$

and thus is the time-rate of change of the mass enclosed by the volume  $V$ . The right hand side of (1.71) is the net mass inflow, per unit time, over the mass outflow. That is, the mass enclosed by the control volume  $V$ , in the absence of sources or sinks, can only change by virtue of mass flow through the boundary of the control volume  $V$ . Thus we rewrite (1.71) as

$$\frac{d}{dt} \int \int \int_V \rho dV = - \int \int_A \mathbf{n} \cdot (\rho \mathbf{V}) dA . \quad (1.72)$$

Here  $\rho \mathbf{V}$  is the mass flow vector and  $\mathbf{n} \cdot (\rho \mathbf{V})$  is its normal component through the surface  $A$  with outward unit normal vector  $\mathbf{n}$ . For computational purposes this is the formulation of the integral form of the law of conservation of mass that is most useful.

The integral form is actually equivalent to the differential form (1.1) of the law conservation of mass if we assume sufficient smoothness of the flow variables in (1.72). Then we can apply Gauss's theorem and write

$$\int \int_A \mathbf{n} \cdot (\rho \mathbf{V}) dA = \int \int \int_V \operatorname{div}(\rho \mathbf{V}) dV .$$

Then (1.71) becomes

$$\int \int \int_V \left[ \frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{V}) \right] dV = 0 .$$

As  $V$  is arbitrary it follows that the integrand must vanish, that is

$$\rho_t + (\rho u)_x + (\rho v)_y + (\rho w)_z = 0 ,$$

which is (1.1). As pointed out earlier, the Euler equations (1.7) admit discontinuous solutions, such as shock waves and contact surfaces. Hence the differential form (1.7) is not valid in general. The integral form (1.72), however, remains valid.

### 1.5.3 Conservation of Momentum

The differential form of the law of conservation of momentum for the inviscid case was stated as equations (1.2)–(1.5), or in more compact form involving tensors as equation (1.10), the Euler equations. Equations (1.61) contain the momentum equations augmented by the effects of viscosity, which gives the Navier–Stokes equations, and heat conduction. As done for the mass equation, we now provide the foundations for the law of conservation of momentum, derive its integral form in quite general terms and show that under appropriate smoothness assumptions the differential form is implied by the

integral form. A control volume  $V$  with bounding surface  $A$  is chosen and the total momentum in  $V$  is given by

$$\Psi(t) = \int \int \int_V \rho \mathbf{V} \, dV . \quad (1.73)$$

The law of conservation of momentum results from the direct application of Newton's law: *the time rate of change of the momentum in  $V$  is equal to the total force acting on the volume  $V$ .* The total force is divided into surface forces  $f_S$  and volume forces  $f_V$  given by

$$f_S = \int \int_A \mathbf{S} \, dA , \quad f_V = \int \int \int_V \rho \mathbf{g} \, dV . \quad (1.74)$$

Here  $\mathbf{g}$  is the specific volume–force vector and may account for inertial forces, gravitational forces, electromagnetic forces and so on.  $\mathbf{S}$  is the stress vector, which is given in terms of a stress tensor  $\mathbf{S}$  as  $\mathbf{S} = \mathbf{n} \cdot \mathbf{S}$ . The stress tensor  $\mathbf{S}$  can be split into a spherically symmetric part  $-p\mathbf{l}$  due to pressure  $p$ , and a viscous part  $\Pi$  given by (1.52)–(1.53). Application of Newton's Law gives

$$\frac{D\Psi}{Dt} = f_S + f_V ,$$

which by virtue of (1.69) with  $\Phi \equiv \rho \mathbf{V}$  gives

$$\int \int \int_V \frac{\partial}{\partial t} (\rho \mathbf{V}) \, dV = - \int \int_A \mathbf{V} (\mathbf{n} \cdot \rho \mathbf{V}) \, dA + f_S + f_V .$$

Regarding  $V$  as a fixed volume in space, independent of time, we write

$$\frac{d}{dt} \int \int \int_V (\rho \mathbf{V}) \, dV = - \int \int_A \mathbf{V} (\mathbf{n} \cdot \rho \mathbf{V}) \, dA + f_S + f_V , \quad (1.75)$$

which may be interpreted as saying that the *time rate of change of momentum within the fixed control volume  $V$  is due to the net momentum inflow over momentum outflow, given by the first term in (1.75), plus surface and volume forces.* Substituting  $\mathbf{S}$  from (1.52) into (1.74) and writing all surface terms into a single integral we have

$$\left. \begin{aligned} \frac{d}{dt} \int \int \int_V (\rho \mathbf{V}) \, dV &= - \int \int_A [\mathbf{V} (\mathbf{n} \cdot \rho \mathbf{V}) + p \mathbf{n} - \mathbf{n} \cdot \Pi] \, dA \\ &\quad + \int \int \int_V \rho \mathbf{g} \, dV . \end{aligned} \right\} \quad (1.76)$$

This general statement is valid even for the case of discontinuous solutions. The differential conservation law form can now be derived from (1.76) under the assumption that the integrand in the surface integral is sufficiently smooth so that Gauss's theorem may be invoked. Consider the first term of the integrand of the surface integral in (1.76)

$$\mathbf{V}(\mathbf{n} \cdot \rho \mathbf{V}) = \mathbf{n} \cdot \rho \mathbf{V} \otimes \mathbf{V},$$

where  $\mathbf{V} \otimes \mathbf{V}$  is the tensor in (1.10). The three columns of the left hand side are

$$\begin{aligned} u \cdot (\mathbf{n} \cdot \rho \mathbf{V}) &= \mathbf{n} \cdot [\rho u^2, \rho uv, \rho uw]^T, \\ v \cdot (\mathbf{n} \cdot \rho \mathbf{V}) &= \mathbf{n} \cdot [\rho uv, \rho v^2, \rho vw]^T, \\ w \cdot (\mathbf{n} \cdot \rho \mathbf{V}) &= \mathbf{n} \cdot [\rho uw, \rho vw, \rho w^2]^T. \end{aligned}$$

Application of Gauss's theorem to each of the surface terms gives

$$\begin{aligned} \int \int \int_V \frac{\partial}{\partial t} (\rho \mathbf{V}) dV &= - \int \int \int_V [\operatorname{div}(\rho \mathbf{V} \otimes \mathbf{V}) + \operatorname{grad} p - \operatorname{div} \Pi] dV \\ &\quad + \int \int \int_V \rho \mathbf{g} dV. \end{aligned}$$

As this is valid for any arbitrary volume  $V$  the integrand must vanish, i.e.

$$\frac{\partial}{\partial t} (\rho \mathbf{V}) + \operatorname{div} [\rho \mathbf{V} \otimes \mathbf{V} + p \mathbf{l} - \Pi] = \rho \mathbf{g}. \quad (1.77)$$

This is the differential form of the momentum equation, including a source term due to volume forces. When the viscous stresses are identically zero,  $\Pi \equiv \mathbf{0}$  and the volume forces are neglected, we obtain the Euler equations (1.10). If the viscous stresses are given by (1.55) under the Newtonian assumption we obtain the Navier–Stokes equations (1.57) in differential conservation law form.

#### 1.5.4 Conservation of Energy

As done for mass and momentum we now consider the total energy  $\Psi(t)$  in a control volume  $V$ , that is

$$\Psi(t) = \int \int \int_V E dV. \quad (1.78)$$

*The time rate of change of total energy  $\Psi(t)$  is equal to the work done, per unit time, by all the forces acting on the volume plus the influx of energy per unit time into the volume.* Recall that a force  $\mathbf{f}$  acting on a point moving with velocity  $\mathbf{V}$  produces the work  $\mathbf{V} \cdot \mathbf{f}$  per unit time. The surface and volume forces in (1.74) respectively give rise to the following terms:

$$E_{\text{surf}} = - \int \int_A p(\mathbf{V} \cdot \mathbf{n}) dA + \int \int_A \mathbf{V} \cdot (\mathbf{n} \cdot \Pi) dA, \quad (1.79)$$

$$E_{\text{volu}} = \int \int \int_V \rho(\mathbf{V} \cdot \mathbf{g}) dV. \quad (1.80)$$

The first term in (1.79) corresponds to the work done by the pressure while the second term corresponds to the work done by the viscous stresses.  $E_{\text{volu}}$  in (1.80) is the work done by the volume force  $\mathbf{g}$ . To account for the influx of energy into the volume we denote the energy flow vector by  $\mathbf{Q} = (q_1, q_2, q_3)$ ; the flow of energy per unit time across a surface element  $dA$  is given by  $-(\mathbf{n} \cdot \mathbf{Q}) dA$ . This gives another term,

$$E_{\text{infl}} = - \int \int_A (\mathbf{n} \cdot \mathbf{Q}) dA , \quad (1.81)$$

to be included in the equation of balance of energy, which now reads

$$\frac{D\Psi}{Dt} = E_{\text{surf}} + E_{\text{volu}} + E_{\text{infl}} . \quad (1.82)$$

The left hand side of (1.82) can be transformed via (1.67) with the definition  $\phi \equiv E$  and the result is

$$\frac{D\Psi(t)}{Dt} = \int \int \int_V \frac{\partial}{\partial t} E dV + \int \int_A (\mathbf{n} \cdot E \mathbf{V}) dA . \quad (1.83)$$

As done for the laws of conservation of mass and momentum we now reinterpret the volume  $V$  as fixed in space and independent of time and rewrite (1.82)–(1.83) as

$$\frac{d}{dt} \int \int \int_V E dV = - \int \int_A (\mathbf{n} \cdot E \mathbf{V}) dA + E_{\text{surf}} + E_{\text{volu}} + E_{\text{infl}} ,$$

which in full becomes

$$\left. \begin{aligned} \frac{d}{dt} \int \int \int_V E dV &= - \int \int_A [\mathbf{n} \cdot (E \mathbf{V} + p \mathbf{V} + \mathbf{Q}) - \mathbf{V} \cdot (\mathbf{n} \cdot \Pi)] dA \\ &\quad + \int \int \int_V \rho(\mathbf{V} \cdot \mathbf{g}) dV . \end{aligned} \right\} \quad (1.84)$$

Thus the time rate of change of total energy enclosed in the volume  $V$  equals the net flow of energy through the boundary surface  $A$  plus the forces  $E_{\text{surf}}$ ,  $E_{\text{volu}}$  and  $E_{\text{infl}}$  as given by (1.79)–(1.81).

The differential form of the conservation of energy law (1.84) can now be derived by assuming sufficient smoothness and applying Gauss's theorem to all surface integrals. Direct application of Gauss's theorem to the first term of (1.79) and to (1.81) gives

$$\begin{aligned} - \int \int_A \mathbf{n} \cdot (p \mathbf{V}) dA &= - \int \int \int_V \text{div}(p \mathbf{V}) dV , \\ - \int \int_A \mathbf{n} \cdot \mathbf{Q} dA &= - \int \int \int_V \text{div} \mathbf{Q} dV . \end{aligned}$$

The second term of (1.79) can be transformed via Gauss's theorem by first observing that  $\mathbf{V} \cdot (\mathbf{n} \cdot \Pi) = \mathbf{n} \cdot (\mathbf{V} \cdot \Pi)$ . This follows from the symmetry of the viscous stress tensor  $\Pi$ , see (1.53), (1.55). Hence

$$\int \int_A \mathbf{V} \cdot (\mathbf{n} \cdot \Pi) dA = \int \int \int_V \operatorname{div}(\mathbf{V} \cdot \Pi) dV .$$

Substitution of these volume integrals into the integral form of the law of conservation of energy (1.84) gives

$$\int \int \int_V \{E_t + \operatorname{div}[(E+p)\mathbf{V} - \mathbf{V} \cdot \Pi + \mathbf{Q}] - \rho(\mathbf{V} \cdot \mathbf{g})\} dV = 0 .$$

Since  $V$  is arbitrary the integrand must vanish identically, that is

$$E_t + \operatorname{div}[(E+p)\mathbf{V} - \mathbf{V} \cdot \Pi + \mathbf{Q}] = \rho(\mathbf{V} \cdot \mathbf{g}) . \quad (1.85)$$

This is the differential form of the law of conservation of energy with a source term accounting for the effect of body forces; if these are neglected we obtain the homogeneous energy equation contained in (1.61). When viscous and heat conduction effects are neglected we obtain the energy equation (1.5) or (1.11) corresponding to the compressible Euler equations.

## 1.6 Submodels

In this section we consider simplified versions, or submodels, of the governing equations and their closure conditions. Compressible submodels will include flows with area variation; flows with axial symmetry; flows with cylindrical and spherical symmetry; plane one-dimensional flow and further simplifications of this to include linearised and scalar submodels; the one-dimensional version of the Navier–Stokes equations. Incompressible submodels will include free-surface gravity flows and the derivation of the shallow water equations as a special case; we also study various formulations of the incompressible Navier–Stokes equations.

### 1.6.1 Summary of the Equations

Before proceeding with the study of particular situations we summarise the general laws of conservation of mass, momentum and total energy. In differential conservation law form these read

$$\rho_t + \nabla \cdot (\rho \mathbf{V}) = 0 , \quad (1.86)$$

$$\frac{\partial}{\partial t}(\rho \mathbf{V}) + \nabla \cdot [\rho \mathbf{V} \otimes \mathbf{V} + pI - \Pi] = \rho \mathbf{g} , \quad (1.87)$$

$$E_t + \nabla \cdot [(E+p)\mathbf{V} - \mathbf{V} \cdot \Pi + \mathbf{Q}] = \rho(\mathbf{V} \cdot \mathbf{g}) . \quad (1.88)$$

where  $\mathbf{g} = (g_1, g_2, g_3)$  is a body force vector. The integral form of the conservation laws is given by

$$\frac{d}{dt} \int \int \int_V \rho dV = - \int \int_A \mathbf{n} \cdot (\rho \mathbf{V}) dA , \quad (1.89)$$

$$\left. \begin{aligned} \frac{d}{dt} \int \int \int_V (\rho \mathbf{V}) dV &= - \int \int_A [\mathbf{V}(\mathbf{n} \cdot \rho \mathbf{V}) + p\mathbf{n} - \mathbf{n} \cdot \Pi] dA \\ &\quad + \int \int \int_V \rho \mathbf{g} dV , \end{aligned} \right\} \quad (1.90)$$

$$\left. \begin{aligned} \frac{d}{dt} \int \int \int_V E dV &= - \int \int_A [\mathbf{n} \cdot (E \mathbf{V} + p \mathbf{V} + \mathbf{Q}) - \mathbf{V} \cdot (\mathbf{n} \cdot \Pi)] dA \\ &\quad + \int \int \int_V \rho (\mathbf{V} \cdot \mathbf{g}) dV , \end{aligned} \right\} \quad (1.91)$$

where  $V$  is the total volume of an element of fluid and  $A$  is its boundary. Computationally,  $V$  will be a finite volume or computational cell. When body forces are included via a source term vector but viscous and heat conduction effects are neglected we have the Euler equations

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_y + \mathbf{H}(\mathbf{U})_z = \mathbf{S}(\mathbf{U}) , \quad (1.92)$$

$$\left. \begin{aligned} \mathbf{U} &= \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ E \end{bmatrix} , \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ \rho uw \\ u(E + p) \end{bmatrix} , \\ \mathbf{G} &= \begin{bmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ \rho vw \\ v(E + p) \end{bmatrix} , \quad \mathbf{H} = \begin{bmatrix} \rho w \\ \rho uw \\ \rho vw \\ \rho w^2 + p \\ w(E + p) \end{bmatrix} . \end{aligned} \right\} \quad (1.93)$$

Here  $\mathbf{S} = \mathbf{S}(\mathbf{U})$  is a *source* or *forcing term*. Body forces such as gravity may be represented in  $\mathbf{S}$ . Injection of mass, momentum or energy may also be included in  $\mathbf{S}$ . Usually,  $\mathbf{S}(\mathbf{U})$  is a prescribed *algebraic* function of the flow variables and does not involve derivatives of these, but there are exceptions. Equations (1.92) are said to be *inhomogeneous*. When  $\mathbf{S}(\mathbf{U}) \equiv \mathbf{0}$  one speaks of *homogeneous* equations. There are other situations in which source terms  $\mathbf{S}(\mathbf{U})$  arise as a consequence of approximating the homogeneous equations in (1.92) to model situations with particular geometric features. In this case the source term is of geometric character, but we shall still call it a source term.

Sometimes it is convenient to express the equations in terms of the *primitive* or *physical* variables  $\rho$ ,  $u$ ,  $v$ ,  $w$  and  $p$ . By expanding derivatives in (1.92), using the mass equation into the momentum equations and in turn using these into the energy equation one can re-write the three-dimensional Euler equations for ideal gases with a body-force source term as

$$\left. \begin{aligned} \rho_t + u\rho_x + v\rho_y + w\rho_z + \rho(u_x + v_y + w_z) &= 0 , \\ u_t + uu_x + vu_y + wu_z + \frac{1}{\rho}p_x &= g_1 , \\ v_t + uv_x + vv_y + wv_z + \frac{1}{\rho}p_y &= g_2 , \\ w_t + uw_x + vw_y + ww_z + \frac{1}{\rho}p_z &= g_3 , \\ p_t + up_x + vp_y + wp_z + \gamma p(u_x + v_y + w_z) &= 0 . \end{aligned} \right\} \quad (1.94)$$

Next we consider simplifications of the Euler and Navier–Stokes equations augmented by source terms to account for additional flow physics.

### 1.6.2 Flow with Area Variation

Flows with area variation arise naturally in the study of fluid flow phenomena in ducts, pipes, shock tubes and nozzles. One may start from the three dimensional homogeneous version of (1.92) to produce, under the assumption of smooth area variations, a one-dimensional system with geometric source terms. Denote the cross-sectional area of the nozzle by  $A = A(x, t)$ , where  $x$  denotes distance along the nozzle and  $t$  denotes time. It is assumed that the area varies smoothly with space and time and its variation is due to both *translation* and *deformation*. Assume that the speed of translation is given by  $c(x, t)$ .

Most of the presentation that follows is motivated by some useful remarks by Professor Tim Swafford (Swafford, 1998, private communication) on the first edition of this book. He pointed out the correct derivation of the equations for the case in which the area depends on time, and also provided the following key references: Varner *et. al.* [567], Chessor [107] and Warsi [575].

The governing equations read

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{S}(\mathbf{U}) , \quad (1.95)$$

where

$$\mathbf{U} = \begin{bmatrix} A\rho \\ A\rho u \\ AE \end{bmatrix} , \quad \mathbf{F} = \begin{bmatrix} A\rho(u - c) \\ A[\rho u(u - c) + p] \\ A[(u - c)E + up] \end{bmatrix} , \quad \mathbf{S} = \begin{bmatrix} 0 \\ pA_x \\ -pA_t \end{bmatrix} . \quad (1.96)$$

In what follows we assume  $c = 0$ . Manipulation of equations (1.95)–(1.96) lead to the following alternative form

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{S}(\mathbf{U}) , \quad (1.97)$$

where now

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ E \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ u(E + p) \end{bmatrix}, \quad \mathbf{S} = -\frac{1}{A} \frac{dA}{dt} \begin{bmatrix} \rho \\ \rho u \\ (E + p) \end{bmatrix}. \quad (1.98)$$

Here the source term vector contains the differential operator

$$\frac{dA}{dt} = A_t + uA_x, \quad (1.99)$$

which expresses the time variation of the area  $A(x, t)$  along particle paths  $dx/dt = u$ . The upwind nature of the coefficient  $\frac{dA}{dt}$  in the source term vector suggests possible discretisation procedures.

Yet another form of equations (1.95)–(1.96), for the case in which the area is independent of time and  $c = 0$ , is the following

$$(A\mathbf{U})_t + [A\mathbf{F}(\mathbf{U})]_x = -A\mathbf{S}(\mathbf{U})_x, \quad (1.100)$$

where

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ E \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 \\ u(E + p) \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} 0 \\ p \\ 0 \end{bmatrix}. \quad (1.101)$$

Note that the momentum flux does not include the pressure term  $p$ . For details see Ben-Artzi and Falcovitz [38].

### 1.6.3 Axi-Symmetric Flows

Here we consider domains that are symmetric around a coordinate direction. We choose this coordinate to be the  $z$ -axis and is called the axial direction. The second coordinate is  $r$ , which measures distance from the axis of symmetry  $z$  and is called the radial direction. There are two components of velocity, namely  $u(r, z)$  and  $v(r, z)$ . These are respectively the radial ( $r$ ) and axial ( $z$ ) components of velocity. Then the three dimensional (inhomogeneous) conservation laws (1.92) are approximated by a two dimensional problem with geometric source terms  $\mathbf{S}(\mathbf{U})$ , namely

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_r + \mathbf{G}(\mathbf{U})_z = \mathbf{S}(\mathbf{U}), \quad (1.102)$$

where

$$\left. \begin{aligned} \mathbf{U} &= \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ E \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ u(E + p) \end{bmatrix}, \\ \mathbf{G} &= \begin{bmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ v(E + p) \end{bmatrix}, \quad \mathbf{S} = -\frac{1}{r} \begin{bmatrix} \rho u \\ \rho u^2 \\ \rho uv \\ u(E + p) \end{bmatrix}. \end{aligned} \right\} \quad (1.103)$$

An alternative form of (1.102)–(1.103) is

$$\bar{\mathbf{U}}_t + \bar{\mathbf{F}}(\bar{\mathbf{U}})_r + \bar{\mathbf{G}}(\bar{\mathbf{U}})_z = \bar{\mathbf{S}}(\bar{\mathbf{U}}), \quad (1.104)$$

where

$$\left. \begin{aligned} \bar{\mathbf{U}} &= r\mathbf{U}, \quad \bar{\mathbf{F}} = r\mathbf{F}, \quad \bar{\mathbf{G}} = r\mathbf{G}, \\ \bar{\mathbf{S}} &= (0, -p, 0)^T. \end{aligned} \right\} \quad (1.105)$$

From a numerical point of view this form of the equations has its attractions (Pike and Roe, 1989, private communication).

#### 1.6.4 Cylindrical and Spherical Symmetry

Cylindrical and spherically symmetric wave motion arises naturally in the theory of explosion waves in water, air and other media. In these situations the multidimensional equations may be reduced to essentially one-dimensional equations with a geometric source term vector  $\mathbf{S}(\mathbf{U})$  to account for the second and third spatial dimensions. We write

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_r = \mathbf{S}(\mathbf{U}), \quad (1.106)$$

where

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ E \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ u(E + p) \end{bmatrix}, \quad \mathbf{S} = -\frac{\alpha}{r} \begin{bmatrix} \rho u \\ \rho u^2 \\ u(E + p) \end{bmatrix}. \quad (1.107)$$

Here  $r$  is the radial distance from the origin and  $u$  is the radial velocity. When  $\alpha = 0$  we have plane one-dimensional flow; when  $\alpha = 1$  we have cylindrically symmetric flow, an approximation to two-dimensional flow. This is a special case of equations (1.102)–(1.103) when no axial variations are present ( $v = 0$ ). For  $\alpha = 2$  we have spherically symmetric flow, an approximation to three-dimensional flow. Approximations (1.106)–(1.107) can easily be solved numerically to a high degree of accuracy by a good one-dimensional numerical method. These accurate one-dimensional solutions can then be very useful in partially validating two and three dimensional numerical solutions of the full models, see Sect. 17.1 of Chap. 17.

#### 1.6.5 Plain One-Dimensional Flow

We first consider the one-dimensional time dependent case

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0}, \quad (1.108)$$

where

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ E \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ u(E + p) \end{bmatrix}. \quad (1.109)$$

These equations also result from equations (1.106)–(1.107) with  $\alpha \equiv 0$  and  $r$  replaced by  $x$ . Under suitable physical assumptions they produce even simpler mathematical models. In all the submodels studied so far we have assumed some thermodynamic closure condition given by an Equation of State (EOS).

The **isentropic equations** result under the assumption that the entropy  $s$  is constant everywhere, which is a simplification of the thermodynamics. Now the EOS is

$$p = p(\rho) \equiv C\rho^\gamma, \quad C=\text{constant}. \quad (1.110)$$

This makes the energy equation redundant and we have the  $2 \times 2$  system

$$\left. \begin{aligned} \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x &= \mathbf{0}, \\ \mathbf{U} &= \begin{bmatrix} \rho \\ \rho u \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \end{bmatrix}, \end{aligned} \right\} \quad (1.111)$$

with the pressure  $p$  given by the simple EOS (1.110).

The **isothermal equations** are even a simpler model than the isentropic equations, still non-linear. These may be viewed as resulting from the isentropic equations (1.111) with the EOS (1.110) further simplified to

$$p = p(\rho) \equiv a^2\rho, \quad (1.112)$$

where  $a$  is a non-zero constant propagation speed of sound. The isothermal equations are

$$\left. \begin{aligned} \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x &= \mathbf{0}, \\ \mathbf{U} &= \begin{bmatrix} \rho \\ \rho u \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + \rho a^2 \end{bmatrix}. \end{aligned} \right\} \quad (1.113)$$

More submodels may be obtained by writing the isentropic equations as

$$\rho_t + \rho u_x + u\rho_x = 0, \quad (1.114)$$

$$u_t + uu_x = -\frac{1}{\rho}p_x. \quad (1.115)$$

These result from (1.94) with the appropriate simplifications.

The **inviscid Burgers equation** is a scalar (single equation) non-linear equation given by

$$u_t + uu_x = 0, \quad (1.116)$$

and can be obtained from the momentum equation (1.115) by neglecting density, and thus pressure, variations. In conservative form equation (1.116) reads

$$u_t + \left( \frac{u^2}{2} \right)_x = 0. \quad (1.117)$$

**The Linearised Equations of Gas Dynamics** are obtained from (1.114)–(1.115) by considering small disturbances  $\bar{u}$ ,  $\bar{\rho}$  to a motionless gas. Set  $u = \bar{u}$  and  $\rho = \bar{\rho} + \rho_0$ , where  $\rho_0$  is a constant density value. Recall that  $p = p(\rho)$  and neglecting products of small quantities we have

$$p = p(\rho_0) + \bar{\rho} \frac{\partial p}{\partial \rho}(\rho_0),$$

that is,  $p = p_0 + a^2 \bar{\rho}$  with  $p_0 = p(\rho_0)$ , and

$$a^2 = \frac{\partial p}{\partial \rho}(\rho_0) = \text{constant}. \quad (1.118)$$

Substituting into (1.114)–(1.115) and neglecting squares of small quantities we obtain the linear equations

$$\bar{\rho}_t + \rho_0 \bar{u}_x = 0, \quad (1.119)$$

$$\bar{u}_t + \frac{a^2}{\rho_0} \bar{\rho}_x = 0. \quad (1.120)$$

Elimination of  $\bar{u}$  gives

$$\bar{\rho}_{tt} = a^2 \bar{\rho}_{xx}, \quad (1.121)$$

which is the linear second-order wave equation for  $\bar{\rho}(x, t)$ . In matrix form system (1.119)–(1.120) reads

$$\mathbf{W}_t + \mathbf{A} \mathbf{W}_x = \mathbf{0}, \quad (1.122)$$

$$\mathbf{W} = \begin{bmatrix} \rho \\ u \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & \rho_0 \\ a^2/\rho_0 & 0 \end{bmatrix}, \quad (1.123)$$

where bars have been dropped. The coefficient matrix  $\mathbf{A}$  is now constant and thus the system (1.122) is a linear system with constant coefficients, the linearised equations of gas dynamics.

The **linear advection**, sometimes called linear convection, equation

$$u_t + au_x = 0, \quad (1.124)$$

where  $a$  is a constant speed of wave propagation, is a further simplification to (1.121). This is also known as the *one-way wave equation* and plays a major role in the designing, analysing and testing of numerical methods for wave propagation problems.

### 1.6.6 Steady Compressible Flow

The steady, or time independent, homogeneous three dimensional Euler equations (1.92) are

$$\mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_y + \mathbf{H}(\mathbf{U})_z = \mathbf{0} . \quad (1.125)$$

In the steady regime it is important to identify *subsonic* and *supersonic flow*. To this end we recall the definition of Mach number  $M$

$$M = \left[ \frac{(u^2 + v^2 + w^2)}{a^2} \right]^{\frac{1}{2}} , \quad (1.126)$$

where  $a$  is the speed of sound; for ideal gases  $a = \sqrt{(\gamma p / \rho)}$ . Supersonic flow requires  $M > 1$ , while for subsonic flow we have  $M < 1$ . For sonic flow  $M = 1$ . Computationally, the three-dimensional equations may be treated via the method of dimensional splitting, which in essence reduces the three-dimensional problem to a sequence of *augmented* two-dimensional problems, see Chap. 16. The basic approach therefore relies on the two-dimensional case. In differential conservation form we have

$$\mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_y = \mathbf{0} , \quad (1.127)$$

$$\mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ u(E + p) \end{bmatrix} , \quad \mathbf{G} = \begin{bmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ v(E + p) \end{bmatrix} . \quad (1.128)$$

As discontinuous solutions such as shock waves and slip surfaces are to be admitted, we replace the differential form (1.127)–(1.128) by the more general integral conservation form

$$\oint (\mathbf{F} dy - \mathbf{G} dx) = \mathbf{0} . \quad (1.129)$$

The integral is to be evaluated over the boundary of the appropriate control volume. In numerical methods this will be a computational cell.

**Steady linearised models** can be obtained from the steady Euler equations (1.127)–(1.128). An interesting submodel is the small perturbation, two-dimensional steady supersonic equations

$$u_x - a^2 v_y = 0 , \quad v_x - u_y = 0 , \quad (1.130)$$

with

$$a^2 = \frac{1}{M_\infty^2 - 1} . \quad (1.131)$$

$M_\infty = \text{constant}$  denotes the free-stream Mach number and  $u(x, y)$ ,  $v(x, y)$  are small perturbations of the  $x$  and  $y$  velocity components respectively. In matrix form these equations read

$$\mathbf{W}_x + \mathbf{A} \mathbf{W}_y = \mathbf{0} , \quad (1.132)$$

with

$$\mathbf{W} = \begin{bmatrix} u \\ v \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & -a^2 \\ -1 & 0 \end{bmatrix}. \quad (1.133)$$

See Sect. 18.2.2 of Chapter 18 for more information about steady supersonic flow.

### 1.6.7 Viscous Compressible Flow

The one-dimensional version of the compressible Navier–Stokes equations with heat conduction can be obtained from (1.61) by setting  $v = w = 0$ . The result is

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{S}, \quad (1.134)$$

where

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ E \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ u(E + p) \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} 0 \\ \frac{4}{3}(\eta u_x)_x \\ \frac{4}{3}(\eta uu_x)_x - (\kappa T_x)_x \end{bmatrix}. \quad (1.135)$$

**Burgers's equation** is the viscous version of (1.117) corresponding to a scalar non-linear simplification of (1.135), namely

$$u_t + \left( \frac{u^2}{2} \right)_x = \alpha u_{xx}, \quad (1.136)$$

where  $\alpha$  is a coefficient of viscosity. A linearised form of this is

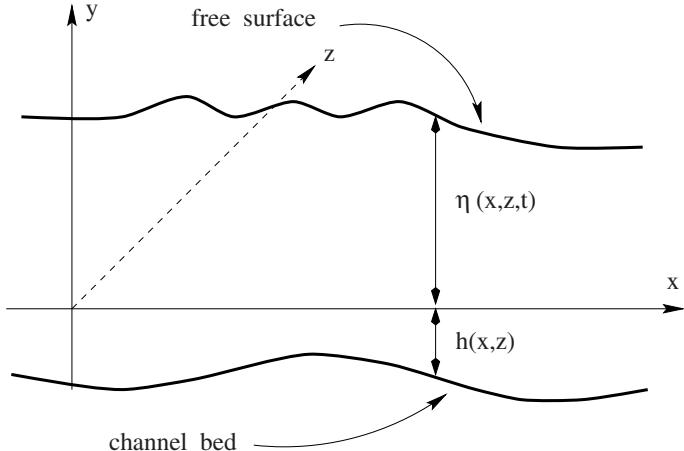
$$u_t + au_x = \alpha u_{xx}, \quad (1.137)$$

which is the viscous version of the linear advection equation (1.124).

Next we consider two examples of incompressible flow. The first concerns inviscid incompressible flow with body forces, with special reference to free-surface gravity flow, as in oceans and rivers for example. The second example concerns the incompressible viscous equations with body forces and heat conduction neglected.

### 1.6.8 Free-Surface Gravity Flow

Consider the flow of water in a channel and assume the water to be incompressible, non-viscous, non-heat conducting and subject to gravitational forces. We adopt the convention that the horizontal plane is given by the coordinates  $x$  and  $z$  and that the vertical direction is given by  $y$ . Denote the body force vector by  $\mathbf{g} = (g_1, g_2, g_3) \equiv (0, -g, 0)$  where  $g$  is the acceleration due to gravity, assumed constant. Two important boundaries of the three dimensional domain are the bottom of the channel, denoted by  $y = -h(x, z)$  and assumed fixed in time, and the free surface under gravity  $y = \eta(x, z, t)$ ,



**Fig. 1.1.** Flow in a channel with free surface under gravity

which depends on space and time. Fig. 1.1 illustrates the situation for a vertical plane  $z = \text{constant}$ . The state  $y = 0$ ,  $u = 0$ ,  $w = 0$  corresponds to the rest position of equilibrium.

For an incompressible fluid the law of conservation of mass can be shown [112] to produce

$$\operatorname{div} \mathbf{V} = u_x + v_y + w_z = 0 . \quad (1.138)$$

From the mass equation in (1.86) we write

$$\frac{D\rho}{Dt} + \rho(\operatorname{div} \mathbf{V}) = 0 ,$$

and from (1.138) it follows that

$$\frac{D\rho}{Dt} = \rho_t + \mathbf{V} \cdot \operatorname{grad} \rho = 0 .$$

That is, the mass density following the fluid is constant. If in addition we assume the water to be homogeneous, no variations in space, then it follows that  $\rho$  does not change with time. Practical examples of incompressible non-homogeneous fluids arise in Oceanography in the study of stratified flows. From the momentum equations in (1.94) we write

$$\left. \begin{aligned} \frac{Du}{Dt} &\equiv u_t + uu_x + vu_y + wu_z = -\frac{1}{\rho}p_x , \\ \frac{Dv}{Dt} &\equiv v_t + uv_x + vv_y + wv_z = -\frac{1}{\rho}p_y - g , \\ \frac{Dw}{Dt} &\equiv w_t + uw_x + vw_y + ww_z = -\frac{1}{\rho}p_z . \end{aligned} \right\} \quad (1.139)$$

There are now four equations and four unknowns, namely  $u$ ,  $v$ ,  $w$  and  $p$ . In principle, given initial and boundary conditions one should be able to solve

(1.138)–(1.139) for  $u$ ,  $v$ ,  $w$  and  $p$  as functions of space  $x$ ,  $y$ ,  $z$  and time  $t$ . Boundary conditions are required at the bottom  $y = -h(x, z)$  and on the free surface under gravity  $y = \eta(x, z, t)$ . Two boundary conditions are given for the free surface

$$\left. \begin{array}{l} \frac{D}{Dt}(\eta - y) = 0 \\ p = p_{\text{atm}} \end{array} \right\} \quad y = \eta(x, z, t) . \quad (1.140)$$

For the bottom boundary one takes

$$\frac{D}{Dt}(h + y) = 0 , \quad y = -h(x, z) . \quad (1.141)$$

In equation (1.140),  $p_{\text{atm}}$  is the atmospheric pressure, which for convenience is assumed to be zero. Equation (1.141) states that the normal component of velocity vanishes, i.e. there is no flow through the bottom of the channel. In spite of the strong physical assumptions made, the free-surface problem as stated remains a very difficult problem to solve numerically. Analytical solutions are impossible to obtain. Further approximations result in more tractable mathematical models. For general background on the topic see the book on water waves by Stoker [465].

### 1.6.9 The Shallow Water Equations

The shallow water equations are an approximation to the full free-surface problem and result from the assumption that the vertical component of the acceleration  $Dv/Dt$  can be neglected. Inserting  $Dv/Dt = 0$  in the second of equations (1.139) gives

$$p = \rho g(\eta - y) . \quad (1.142)$$

This is called the *hydrostatic pressure relation*. Differentiation of  $p$  with respect to  $x$  and  $z$  gives

$$p_x = \rho g \eta_x , \quad (1.143)$$

$$p_z = \rho g \eta_z , \quad (1.144)$$

i.e. both  $p_x$  and  $p_z$  are independent of  $y$  and so the  $x$  and  $z$  components of the acceleration  $Du/Dt$  and  $Dw/Dt$  in (1.139) are independent of  $y$ . Thus, the  $x$  and  $z$  components of velocity are also independent of  $y$  for all  $t$  if they were at a given time,  $t = 0$ , say. Hence, the first and third equations in (1.139), making use of (1.143)–(1.144), become

$$u_t + uu_x + wu_z = -g\eta_x , \quad (1.145)$$

$$w_t + uw_x + ww_z = -g\eta_z . \quad (1.146)$$

An important step in deriving the shallow water equations now follows; we integrate the continuity equation (1.138) with respect to  $y$  (vertical direction) to obtain

$$\int_{-h}^{\eta} u_x \, dy + \int_{-h}^{\eta} w_z \, dy + v|_{y=-h}^{\eta} = 0. \quad (1.147)$$

By expanding the boundary conditions (1.140) and (1.141) according to the definition of total derivative  $D/Dt$ , see (1.64), we obtain

$$(\eta_t + u\eta_x + w\eta_z - v)|_{y=\eta} = 0, \quad (1.148)$$

$$(uh_x + wh_z + v)|_{y=-h} = 0, \quad (1.149)$$

which, if used in (1.147), give

$$\left. \begin{aligned} & \int_{-h}^{\eta} u_x \, dy + \int_{-h}^{\eta} w_z \, dy + \eta_t \\ & + (u|_{y=\eta})\eta_x + (w|_{y=\eta})\eta_z + (u|_{y=-h})h_x + (w|_{y=-h})h_z = 0. \end{aligned} \right\} \quad (1.150)$$

Equation (1.150) can finally be expressed as

$$\eta_t + \frac{\partial}{\partial x} \int_{-h}^{\eta} u \, dy + \frac{\partial}{\partial z} \int_{-h}^{\eta} w \, dy = 0. \quad (1.151)$$

This follows by using the relations

$$\begin{aligned} \frac{\partial}{\partial x} \int_{-h(x,z)}^{\eta(x,z,t)} u \, dy &= \int_{-h}^{\eta} u_x \, dy + (u|_{y=\eta})\eta_x + (u|_{y=-h})h_x, \\ \frac{\partial}{\partial z} \int_{-h(x,z)}^{\eta(x,z,t)} w \, dy &= \int_{-h}^{\eta} w_z \, dy + (w|_{y=\eta})\eta_z + (w|_{y=-h})h_z. \end{aligned}$$

Equation (1.151) can be simplified further. This follows from the observation that both  $u$  and  $w$  are independent of  $y$  and so equation (1.151) becomes

$$\eta_t + [u(\eta + h)]_x + [w(\eta + h)]_z = 0. \quad (1.152)$$

The governing two-dimensional shallow water equations are (1.145), (1.146) and (1.152). We now express these equations in conservation law form. Since  $h(x, z)$  is independent of time  $t$  we have  $h_t = 0$  and so equation (1.152) can be re-written as

$$(\eta + h)_t + [u(\eta + h)]_x + [w(\eta + h)]_z = 0, \quad (1.153)$$

which if multiplied by  $u$  and added to (1.145), premultiplied by  $\eta + h$ , gives

$$[u(\eta + h)]_t + [u^2(\eta + h)]_x + [uw(\eta + h)]_z = -g(\eta + h)\eta_x. \quad (1.154)$$

In an analogous way equations (1.146) and (1.153) give

$$[w(\eta + h)]_t + [uw(\eta + h)]_x + [w^2(\eta + h)]_z = -g(\eta + h)\eta_z. \quad (1.155)$$

The right-hand side term of (1.154) can be re-written as

$$-g(\eta + h)\eta_x = g(\eta + h)h_x - \frac{1}{2}g[(\eta + h)^2]_x \quad (1.156)$$

and so (1.154) becomes

$$(\phi u)_t + (\phi u^2 + \frac{1}{2}\phi^2)_x + (\phi uw)_z = g\phi h_x . \quad (1.157)$$

Similarly, equation (1.155) becomes

$$(\phi w)_t + (\phi uw)_x + (\phi w^2 + \frac{1}{2}\phi^2)_z = g\phi h_z , \quad (1.158)$$

where

$$\phi = gH , H = \eta + h . \quad (1.159)$$

$H$  is the total depth of the fluid and  $\phi$  is called the *geopotential*. The two-dimensional shallow water equations are now written in compact conservation form with source terms, i.e. equations (1.153), (1.157) and (1.158) can be expressed as

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_z = \mathbf{S}(\mathbf{U}) , \quad (1.160)$$

with

$$\left. \begin{aligned} \mathbf{U} &= \begin{bmatrix} \phi \\ \phi u \\ \phi w \end{bmatrix} , \mathbf{F} = \begin{bmatrix} \phi u \\ \phi u^2 + \frac{1}{2}\phi^2 \\ \phi uw \end{bmatrix} , \\ \mathbf{G} &= \begin{bmatrix} \phi w \\ \phi uw \\ \phi w^2 + \frac{1}{2}\phi^2 \end{bmatrix} , \mathbf{S} = \begin{bmatrix} 0 \\ g\phi h_x \\ g\phi h_z \end{bmatrix} . \end{aligned} \right\} \quad (1.161)$$

In equations (1.160)  $\mathbf{U}$  is the vector of conserved variables,  $\mathbf{F}(\mathbf{U})$  and  $\mathbf{G}(\mathbf{U})$  are flux vectors and  $\mathbf{S}(\mathbf{U})$  is the source term vector. For many applications there will be additional terms in the vector  $\mathbf{S}(\mathbf{U})$  to account for Coriolis forces, wind forces, bottom friction, etc. Numerical solution procedures will deal essentially with the homogeneous part of (1.160). The numerical treatment of the source terms is a relatively standard process as is the treatment of the two-dimensional homogeneous problem. Both can be dealt with via splitting schemes; for details see Chaps. 15 and 16. From a computational point of view, most of the effort goes into devising schemes for the basic homogeneous one-dimensional system

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0} , \quad (1.162)$$

with  $\mathbf{U}$  and  $\mathbf{F}(\mathbf{U})$  given by

$$\mathbf{U} = \begin{bmatrix} \phi \\ \phi u \end{bmatrix} , \mathbf{F} = \begin{bmatrix} \phi u \\ \phi u^2 + \frac{1}{2}\phi^2 \end{bmatrix} . \quad (1.163)$$

The conservation laws (1.162)–(1.163) can be written in integral form as

$$\oint (\mathbf{U} dx - \mathbf{F}(\mathbf{U}) dt) = \mathbf{0} . \quad (1.164)$$

Equations (1.164) admit discontinuous solutions while equations (1.162)–(1.163) do not.

See Sect. 18.2.1 of Chapter 18 for more information about the shallow water equations. See also the textbook [528].

### 1.6.10 Incompressible Viscous Flow

We assume the fluid to be incompressible, homogeneous, non-heat conducting and viscous, with constant coefficient of viscosity  $\eta$ . Body forces are also neglected. We study three mathematical formulations of the governing equations in Cartesian coordinates and restrict our attention to the two-dimensional case.

The **primitive variable formulation** of the incompressible two dimensional Navier–Stokes equations is given by

$$u_x + v_y = 0 , \quad (1.165)$$

$$u_t + uu_x + vu_y + \frac{1}{\rho} p_x = \nu [u_{xx} + u_{yy}] , \quad (1.166)$$

$$v_t + uv_x + vv_y + \frac{1}{\rho} p_y = \nu [v_{xx} + v_{yy}] , \quad (1.167)$$

where

$$\nu = \frac{\eta}{\rho} \quad (1.168)$$

is the *kinematic viscosity* and  $\eta$  is the coefficient of *shear viscosity*. We have a set of three equations (1.165)–(1.167) for the three unknowns  $u, v, p$ , the primitive variables. In principle, given a domain along with initial and boundary conditions for the equations one should be able to solve this problem using the primitive variable formulation.

The **stream–function vorticity formulation** is another way of expressing the incompressible Navier–Stokes equations. This formulation is attractive for the two-dimensional case but not so much in three dimensions, in which the role of a stream function is replaced by that of a *vector potential*. The magnitude of the vorticity vector can be written as

$$\zeta = v_x - u_y . \quad (1.169)$$

Introducing a stream function  $\psi$  we have

$$u = \psi_y , \quad v = -\psi_x . \quad (1.170)$$

By combining equations (1.166) and (1.167), so as to eliminate pressure  $p$ , and using (1.169) we obtain

$$\zeta_t + u\zeta_x + v\zeta_y = \nu [\zeta_{xx} + \zeta_{yy}] , \quad (1.171)$$

which is called the *vorticity transport equation*. This is an advection–diffusion equation of *parabolic* type. In order to solve (1.171) one requires the solution for the stream function  $\psi$ , which is in turn related to the vorticity  $\zeta$  via

$$\psi_{xx} + \psi_{yy} = -\zeta . \quad (1.172)$$

This is called the *Poisson equation* and is of *elliptic type*. There are numerical schemes to solve (1.171)–(1.172) using the *apparent* decoupling of the parabolic–elliptic problem (1.165)–(1.167) to transform it into the parabolic equation (1.171) and the elliptic equation (1.172). A relevant observation, from the numerical point of view, is that the advection terms of the left hand side of equation (1.171) can be written in conservative form and hence we have

$$\zeta_t + (u\zeta)_x + (v\zeta)_y = \nu [\zeta_{xx} + \zeta_{yy}] . \quad (1.173)$$

This follows from the fact that  $u_x + v_y = 0$ , which was also used to obtain (1.171) from (1.166)–(1.167).

### 1.6.11 The Artificial Compressibility Equations

The artificial compressibility formulation is yet another approach to formulate the incompressible Navier–Stokes equations and was originally put forward by Chorin [109] for the steady case. Consider the two-dimensional equations (1.165)–(1.167) in non-dimensional form

$$u_x + v_y = 0 , \quad (1.174)$$

$$u_t + uu_x + vu_y + p_x = \alpha [u_{xx} + u_{yy}] , \quad (1.175)$$

$$v_t + uv_x + vv_y + p_y = \alpha [v_{xx} + v_{yy}] , \quad (1.176)$$

where the following non-dimensionalisation has been used

$$u \leftarrow u/V_\infty , \quad v \leftarrow v/V_\infty , \quad p \leftarrow \frac{p}{\rho_\infty V_\infty^2} ,$$

$$x \leftarrow x/L , \quad y \leftarrow y/L , \quad t \leftarrow tV_\infty/L ,$$

$$\alpha = 1/R_{eL} , \quad R_{eL} = \frac{V_\infty L}{\nu_\infty} .$$

Multiplying (1.174) by the non-zero parameter  $c^2$  and adding an *artificial compressibility* term  $p_t$  the first equation reads

$$p_t + (uc^2)_x + (vc^2)_y = 0 .$$

By using equation (1.174) the advective terms in (1.175)–(1.176) can be written in *conservative form*, so that the modified system becomes

$$p_t + (uc^2)_x + (vc^2)_y = 0 , \quad (1.177)$$

$$u_t + (u^2 + p)_x + (uv)_y = \alpha [u_{xx} + u_{yy}] , \quad (1.178)$$

$$v_t + (uv)_x + (v^2 + p)_y = \alpha [v_{xx} + v_{yy}] . \quad (1.179)$$

These equations can be written in compact form as

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_y = \mathbf{S}(\mathbf{U}) , \quad (1.180)$$

where

$$\left. \begin{aligned} \mathbf{U} &= \begin{bmatrix} p \\ u \\ v \end{bmatrix} , \quad \mathbf{F} = \begin{bmatrix} c^2 u \\ u^2 + p \\ uv \end{bmatrix} , \\ \mathbf{G} &= \begin{bmatrix} c^2 v \\ uv \\ v^2 + p \end{bmatrix} , \quad \mathbf{S} = \begin{bmatrix} 0 \\ \alpha(u_{xx} + u_{yy}) \\ \alpha(v_{xx} + v_{yy}) \end{bmatrix} . \end{aligned} \right\} \quad (1.181)$$

Equations (1.180)–(1.181) are called the *artificial compressibility equations*. Here  $c^2$  is the *artificial compressibility factor*, a constant parameter. The source term vector in this case is a function of second derivatives. Note that the modified equations are equivalent to the original equations in the *steady state* limit.

The left-hand side of the artificial compressibility equations form a non-linear hyperbolic system. The Riemann problem can be defined and solved exactly or approximately. See Chaps. 4, 8, 9, 10, 11 and 12 on approaches to solve the Riemann problem. Once a Riemann solver is available one can deploy Godunov-type numerical methods to solve the equations with general initial data. See Chaps. 6, 7, 13, 14 and 16 for possible numerical methods. The topic of numerical methods for the artificial compressibility equations is not pursued in this textbook; the interested reader is referred to [514] for details on exact and approximate Riemann solvers for the artificial compressibility equations. Further information about the artificial compressibility equations is found in Sect. 18.2.3 of Chapter 18.

---

## Notions on Hyperbolic Partial Differential Equations

In this chapter we study some elementary properties of a class of hyperbolic Partial Differential Equations (PDEs). The selected aspects of the equations are those thought to be essential for the analysis of the equations of fluid flow and the implementation of numerical methods. For general background on PDEs we recommend the book by John [272] and particularly the one by Zachmanoglou and Thoe [596]. The discretisation techniques studied in this book are strongly based on the underlying Physics and mathematical properties of PDEs. It is therefore justified to devote some effort to some fundamentals on PDEs. Here we deal almost exclusively with *hyperbolic* PDEs and *hyperbolic conservation laws* in particular. There are three main reasons for this: (i) The equations of compressible fluid flow reduce to hyperbolic systems, the Euler equations, when the effects of viscosity and heat conduction are neglected. (ii) Numerically, it is generally accepted that the hyperbolic terms of the PDEs of fluid flow are the terms that pose the most stringent requirements on the discretisation techniques. (iii) The theory of hyperbolic systems is much more advanced than that for more complete mathematical models, such as the Navier–Stokes equations. In addition, there has in recent years been a noticeable increase in research and development activities centred on the theme of hyperbolic problems, as these cover a wide range of areas of scientific and technological interest. A good source of up-to-date work in this field is found in the proceedings of the series of meetings on Hyperbolic Problems, see for instance [87], [184], [213]. See also [326]. Other relevant publications are those of Godlewski and Raviart [215], Hörmander [258] and Tveito and Winther [551].

We restrict ourselves to some of the basics on hyperbolic PDEs and choose an informal way of presentation. The selected topics and approach are almost exclusively motivated by the theme of the book, namely the Riemann problem and high-resolution upwind and centred numerical methods.

## 2.1 Quasi–Linear Equations: Basic Concepts

In this section we study systems of first–order partial differential equations of the form

$$\frac{\partial u_i}{\partial t} + \sum_{j=1}^m a_{ij}(x, t, u_1, \dots, u_m) \frac{\partial u_j}{\partial x} + b_i(x, t, u_1, \dots, u_m) = 0, \quad (2.1)$$

for  $i = 1, \dots, m$ . This is a system of  $m$  equations in  $m$  unknowns  $u_i$  that depend on space  $x$  and a time–like variable  $t$ . Here  $u_i$  are the *dependent variables* and  $x, t$  are the *independent variables*; this is expressed via the notation  $u_i = u_i(x, t)$ ;  $\partial u_i / \partial t$  denotes the partial derivative of  $u_i(x, t)$  with respect to  $t$ ; similarly  $\partial u_i / \partial x$  denotes the partial derivative of  $u_i(x, t)$  with respect to  $x$ . We also make use of subscripts to denote partial derivatives. System (2.1) can also be written in matrix form as

$$\mathbf{U}_t + \mathbf{A}\mathbf{U}_x + \mathbf{B} = \mathbf{0}, \quad (2.2)$$

with

$$\mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} a_{11} & \dots & a_{1m} \\ a_{21} & \dots & a_{2m} \\ \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mm} \end{bmatrix}. \quad (2.3)$$

If the entries  $a_{ij}$  of the matrix  $\mathbf{A}$  are all constant and the components  $b_j$  of the vector  $\mathbf{B}$  are also constant then system (2.2) is *linear with constant coefficients*.

If  $a_{ij} = a_{ij}(x, t)$  and  $b_i = b_i(x, t)$  the system is *linear with variable coefficients*. The system is still linear if  $\mathbf{B}$  depends linearly on  $\mathbf{U}$  and is called *quasi–linear* if the coefficient matrix  $\mathbf{A}$  is a function of the vector  $\mathbf{U}$ , that is  $\mathbf{A} = \mathbf{A}(\mathbf{U})$ . Note that quasi–linear systems are in general systems of non–linear equations. System (2.2) is called *homogeneous* if  $\mathbf{B} = \mathbf{0}$ . For a set of PDEs of the form (2.2) one needs to specify the range of variation of the independent variables  $x$  and  $t$ . Usually  $x$  lies in a subinterval of the real line, namely  $x_l < x < x_r$ ; this subinterval is called the *spatial domain* of the PDEs, or just *domain*. At the values  $x_l, x_r$  one also needs to specify *Boundary Conditions (BCs)*. In this Chapter we assume the domain is the full real line,  $-\infty < x < \infty$ , and thus no boundary conditions need to be specified. As to variations of time  $t$  we assume  $t_0 < t < \infty$ . An *Initial Condition (IC)* needs to be specified at the initial time, which is usually chosen to be  $t_0 = 0$ .

Two scalar ( $m = 1$ ) examples of PDEs of the form (2.1) are the linear advection equation

$$\frac{\partial u}{\partial t} + a \frac{\partial u}{\partial x} = 0 \quad (2.4)$$

and the inviscid Burgers equation

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = 0, \quad (2.5)$$

both introduced in Sect. 1.6.2 of Chap. 1. In the linear advection equation (2.4) the coefficient  $a$  (a constant) is the wave propagation speed. In the Burgers equation  $a = a(u) = u$ .

**Definition 2.1 (Conservation Laws).** *Conservation laws are systems of partial differential equations that can be written in the form*

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0}, \quad (2.6)$$

where

$$\mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix}, \quad \mathbf{F}(\mathbf{U}) = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{bmatrix}. \quad (2.7)$$

$\mathbf{U}$  is called the vector of conserved variables,  $\mathbf{F} = \mathbf{F}(\mathbf{U})$  is the vector of fluxes and each of its components  $f_i$  is a function of the components  $u_j$  of  $\mathbf{U}$ .

**Definition 2.2 (Jacobian Matrix).** *The Jacobian of the flux function  $\mathbf{F}(\mathbf{U})$  in (2.6) is the matrix*

$$\mathbf{A}(\mathbf{U}) = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} = \begin{bmatrix} \frac{\partial f_1}{\partial u_1} & \dots & \frac{\partial f_1}{\partial u_m} \\ \frac{\partial f_2}{\partial u_1} & \dots & \frac{\partial f_2}{\partial u_m} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial u_1} & \dots & \frac{\partial f_m}{\partial u_m} \end{bmatrix}. \quad (2.8)$$

The entries  $a_{ij}$  of  $\mathbf{A}(\mathbf{U})$  are partial derivatives of the components  $f_i$  of the vector  $\mathbf{F}$  with respect to the components  $u_j$  of the vector of conserved variables  $\mathbf{U}$ , that is  $a_{ij} = \partial f_i / \partial u_j$ .

Note that conservation laws of the form (2.6)–(2.7) can also be written in quasi–linear form (2.2), with  $\mathbf{B} \equiv \mathbf{0}$ , by applying the chain rule to the second term in (2.6), namely

$$\frac{\partial \mathbf{F}(\mathbf{U})}{\partial x} = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} \frac{\partial \mathbf{U}}{\partial x}.$$

Hence (2.6) becomes

$$\mathbf{U}_t + \mathbf{A}(\mathbf{U})\mathbf{U}_x = \mathbf{0},$$

which is a special case of (2.2). The scalar PDEs (2.4) and (2.5) can be expressed as conservation laws, namely

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0, \quad f(u) = au, \quad (2.9)$$

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0, \quad f(u) = \frac{1}{2}u^2. \quad (2.10)$$

**Definition 2.3 (Eigenvalues).** *The eigenvalues  $\lambda_i$  of a matrix  $\mathbf{A}$  are the solutions of the characteristic polynomial*

$$|\mathbf{A} - \lambda \mathbf{I}| = \det(\mathbf{A} - \lambda \mathbf{I}) = 0 , \quad (2.11)$$

where  $\mathbf{I}$  is the identity matrix. The eigenvalues of the coefficient matrix  $\mathbf{A}$  of a system of the form (2.2) are called the eigenvalues of the system.

Physically, eigenvalues represent speeds of propagation of information. Speeds will be measured positive in the direction of increasing  $x$  and negative otherwise.

**Definition 2.4 (Eigenvectors).** *A right eigenvector of a matrix  $\mathbf{A}$  corresponding to an eigenvalue  $\lambda_i$  of  $\mathbf{A}$  is a vector  $\mathbf{K}^{(i)} = [k_1^{(i)}, k_2^{(i)}, \dots, k_m^{(i)}]^T$  satisfying  $\mathbf{A}\mathbf{K}^{(i)} = \lambda_i \mathbf{K}^{(i)}$ . Similarly, a left eigenvector of a matrix  $\mathbf{A}$  corresponding to an eigenvalue  $\lambda_i$  of  $\mathbf{A}$  is a vector  $\mathbf{L}^{(i)} = [l_1^{(i)}, l_2^{(i)}, \dots, l_m^{(i)}]$  such that  $\mathbf{L}^{(i)}\mathbf{A} = \lambda_i \mathbf{L}^{(i)}$ .*

For the scalar examples (2.9)–(2.10) the eigenvalues are trivially found to be  $\lambda = a$  and  $\lambda = u$  respectively. Next we find eigenvalues and eigenvectors for a system of PDEs.

*Example 2.5 (Linearised Gas Dynamics).* The linearised equations of Gas Dynamics, derived in Sect. 1.6.2 of Chap. 1, are the  $2 \times 2$  linear system

$$\left. \begin{aligned} \frac{\partial \rho}{\partial t} + \rho_0 \frac{\partial u}{\partial x} &= 0 , \\ \frac{\partial u}{\partial t} + \frac{a^2}{\rho_0} \frac{\partial \rho}{\partial x} &= 0 , \end{aligned} \right\} \quad (2.12)$$

where the unknowns are the density  $u_1 = \rho(x, t)$  and the speed  $u_2 = u(x, t)$ ;  $\rho_0$  is a constant reference density and  $a$  is the sound speed, a positive constant. When written in the matrix form (2.2) this system reads

$$\mathbf{U}_t + \mathbf{A}\mathbf{U}_x = \mathbf{0} , \quad (2.13)$$

with

$$\mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \equiv \begin{bmatrix} \rho \\ u \end{bmatrix} , \quad \mathbf{A} = \begin{bmatrix} 0 & \rho_0 \\ a^2/\rho_0 & 0 \end{bmatrix} . \quad (2.14)$$

The eigenvalues of the system are the zeros of the characteristic polynomial

$$|\mathbf{A} - \lambda \mathbf{I}| = \det \begin{bmatrix} 0 - \lambda & \rho_0 \\ a^2/\rho_0 & 0 - \lambda \end{bmatrix} = 0 .$$

That is,  $\lambda^2 = a^2$ , which has two real and distinct solutions, namely

$$\lambda_1 = -a , \quad \lambda_2 = +a . \quad (2.15)$$

We now find the right eigenvectors  $\mathbf{K}^{(1)}$ ,  $\mathbf{K}^{(2)}$  corresponding to the eigenvalues  $\lambda_1$  and  $\lambda_2$ .

The eigenvector  $\mathbf{K}^{(1)}$  for eigenvalue  $\lambda = \lambda_1 = -a$  is found as follows: we look for a vector  $\mathbf{K}^{(1)} = [k_1, k_2]^T$  such that  $\mathbf{K}^{(1)}$  is a right eigenvector of  $\mathbf{A}$ , that is  $\mathbf{A}\mathbf{K}^{(1)} = \lambda_1\mathbf{K}^{(1)}$ . Writing this in full gives

$$\begin{bmatrix} 0 & \rho_0 \\ a^2/\rho_0 & 0 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} -ak_1 \\ -ak_2 \end{bmatrix},$$

which produces two linear algebraic equations for the unknowns  $k_1$  and  $k_2$

$$\rho_0 k_2 = -ak_1, \quad \frac{a^2}{\rho_0} k_1 = -ak_2. \quad (2.16)$$

The reader will realise that in fact these two equations are equivalent and so effectively we have a single linear algebraic equation in two unknowns. This gives a one-parameter family of solutions. Thus we select an arbitrary non-zero parameter  $\alpha_1$ , a scaling factor, and set  $k_1 = \alpha_1$  in any of the equations to obtain  $k_2 = -\alpha_1 a / \rho_0$  for the second component and hence the first right eigenvector becomes

$$\mathbf{K}^{(1)} = \alpha_1 \begin{bmatrix} 1 \\ -a/\rho_0 \end{bmatrix}. \quad (2.17)$$

The eigenvector  $\mathbf{K}^{(2)}$  for eigenvalue  $\lambda = \lambda_2 = +a$  is found in a similar manner. The resulting algebraic equations for  $\mathbf{K}^{(2)}$  corresponding to the eigenvalue  $\lambda_2 = +a$  are

$$\rho_0 k_2 = ak_1, \quad \frac{a^2}{\rho_0} k_1 = ak_2. \quad (2.18)$$

By denoting the second scaling factor by  $\alpha_2$  and setting  $k_1 = \alpha_2$  we obtain

$$\mathbf{K}^{(2)} = \alpha_2 \begin{bmatrix} 1 \\ a/\rho_0 \end{bmatrix}. \quad (2.19)$$

Taking the scaling factors to be  $\alpha_1 = \rho_0$  and  $\alpha_2 = \rho_0$  gives the right eigenvectors

$$\mathbf{K}^{(1)} = \begin{bmatrix} \rho_0 \\ -a \end{bmatrix}, \quad \mathbf{K}^{(2)} = \begin{bmatrix} \rho_0 \\ a \end{bmatrix}. \quad (2.20)$$

**Definition 2.6 (Hyperbolic System).** A system (2.2) is said to be hyperbolic at a point  $(x, t)$  if  $\mathbf{A}$  has  $m$  real eigenvalues  $\lambda_1, \dots, \lambda_m$  and a corresponding set of  $m$  linearly independent right eigenvectors  $\mathbf{K}^{(1)}, \dots, \mathbf{K}^{(m)}$ . The system is said to be strictly hyperbolic if the eigenvalues  $\lambda_i$  are all distinct.

Note that strict hyperbolicity implies hyperbolicity, because real and distinct eigenvalues ensure the existence of a set of linearly independent eigenvectors. The system (2.2) is said to be *elliptic* at a point  $(x, t)$  if none of the eigenvalues  $\lambda_i$  of  $\mathbf{A}$  are real. Both scalar examples (2.9)–(2.10) are trivially hyperbolic. The linearised gas dynamic equations (2.12) are also hyperbolic, since  $\lambda_1$  and  $\lambda_2$  are both real at any point  $(x, t)$ . Moreover, as the eigenvalues are also distinct this system is strictly hyperbolic.

*Example 2.7 (The Cauchy–Riemann Equations).* An example of a first-order system of the form (2.2) with  $t$  replaced by  $x$  and  $x$  replaced by  $y$  is the Cauchy–Riemann equations

$$\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} = 0, \quad \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} = 0, \quad (2.21)$$

where  $u_1 = u(x, y)$  and  $u_2 = v(x, y)$ . These equations arise in the study of *analytic functions* in Complex Analysis [379]. When written in matrix notation (2.2) equations (2.21) become

$$\mathbf{U}_x + \mathbf{A}\mathbf{U}_y = \mathbf{0}, \quad (2.22)$$

with

$$\mathbf{U} = \begin{bmatrix} u \\ v \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \quad (2.23)$$

The characteristic polynomial  $|\mathbf{A} - \lambda\mathbf{I}| = 0$  gives  $\lambda^2 + 1 = 0$ , which has no real solutions for  $\lambda$  and thus the system is *elliptic*.

*Example 2.8 (The Small Perturbation Equations).* In Sect. 1.6.2 of Chap. 1, the small perturbation steady equations were introduced

$$u_x - a^2 v_y = 0, \quad v_x - u_y = 0, \quad (2.24)$$

with

$$a^2 = \frac{1}{M_\infty^2 - 1}. \quad (2.25)$$

$M_\infty = \text{constant}$  denotes the free-stream Mach number and  $u(x, y)$ ,  $v(x, y)$  are small perturbations of the  $x$  and  $y$  velocity components respectively. In matrix notation these equations read

$$\mathbf{U}_x + \mathbf{A}\mathbf{U}_y = \mathbf{0}, \quad (2.26)$$

with

$$\mathbf{U} = \begin{bmatrix} u \\ v \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & -a^2 \\ -1 & 0 \end{bmatrix}. \quad (2.27)$$

The character of these equations depends entirely on the value of the Mach number  $M_\infty$ . For subsonic flow  $M_\infty < 1$  the characteristic polynomial has complex solutions and thus the equations are of elliptic type. For supersonic flow  $M_\infty > 1$  and the system is strictly hyperbolic, with eigenvalues

$$\lambda_1 = -a, \quad \lambda_2 = +a. \quad (2.28)$$

It is left to the reader to check that the corresponding right eigenvectors are

$$\mathbf{K}^{(1)} = \alpha_1 \begin{bmatrix} 1 \\ 1/a \end{bmatrix}, \quad \mathbf{K}^{(2)} = \alpha_2 \begin{bmatrix} 1 \\ -1/a \end{bmatrix}, \quad (2.29)$$

where  $\alpha_1$  and  $\alpha_2$  are two non-zero scaling factors. By taking the values  $\alpha_1 = \alpha_2 = a$  we obtain the following expressions for the right eigenvectors

$$\mathbf{K}^{(1)} = \begin{bmatrix} a \\ 1 \end{bmatrix}, \quad \mathbf{K}^{(2)} = \begin{bmatrix} a \\ -1 \end{bmatrix},$$

## 2.2 The Linear Advection Equation

A general, time-dependent linear advection equation in three space dimensions reads

$$u_t + a(x, y, z, t)u_x + b(x, y, z, t)u_y + c(x, y, z, t)u_z = 0, \quad (2.30)$$

where the unknown is  $u = u(x, y, z, t)$  and  $a, b, c$  are variable coefficients. If the coefficients are sufficiently smooth one can express (2.30) as a conservation law with source terms, namely

$$u_t + (au)_x + (bu)_y + (cu)_z = u(a_x + b_y + c_z). \quad (2.31)$$

In this section we study in detail the initial-value problem (IVP) for the special case of the linear advection equation, namely

$$\left. \begin{aligned} \text{PDE: } & u_t + au_x = 0, \quad -\infty < x < \infty, \quad t > 0. \\ \text{IC: } & u(x, 0) = u_0(x), \end{aligned} \right\} \quad (2.32)$$

where  $a$  is a constant wave propagation speed. The initial data at time  $t = 0$  is a function of  $x$  alone and is denoted by  $u_0(x)$ . We warn the reader that for systems we shall use a different notation for the initial data. Generally, we shall not be explicit about the conditions  $-\infty < x < \infty; t > 0$  on the independent variables when stating an initial-value problem. The PDE in (2.32) is the simplest hyperbolic PDE and in view of (2.9) is also the simplest hyperbolic conservation law. It is a very useful model equation for the purpose of studying numerical methods for hyperbolic conservation laws, in the same way as the linear, first-order ordinary differential equation

$$\frac{dx}{dt} = \beta, \quad x = x(t), \quad \beta = \text{constant}, \quad (2.33)$$

is a popular model equation for analysing numerical methods for Ordinary Differential Equation (ODEs). Two useful references on ordinary differential equations are Brown [81] and Lambert [296]. In Sect. 15.4 of Chap. 15 we study numerical methods for ODEs in connection with source terms in inhomogeneous PDEs.

### 2.2.1 Characteristics and the General Solution

We recall the definition of *characteristics* or *characteristic curves* in the context of a scalar equation such as that in (2.32). Characteristics may be defined as curves  $x = x(t)$  in the  $t$ - $x$  plane along which the PDE becomes an ODE. Consider  $x = x(t)$  and regard  $u$  as a function of  $t$ , that is  $u = u(x(t), t)$ . The rate of change of  $u$  along  $x = x(t)$  is

$$\frac{du}{dt} = \frac{\partial u}{\partial t} + \frac{dx}{dt} \frac{\partial u}{\partial x} . \quad (2.34)$$

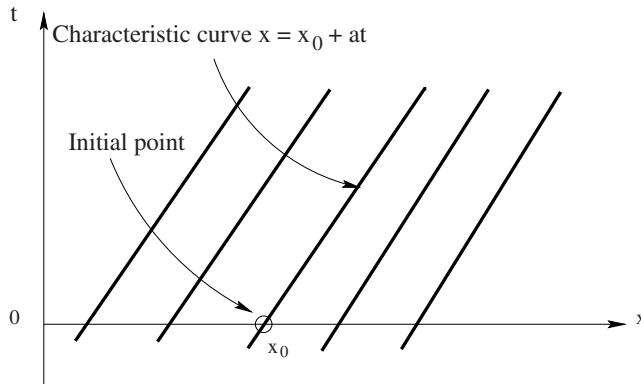
If the characteristic curve  $x = x(t)$  satisfies the ODE

$$\frac{dx}{dt} = a , \quad (2.35)$$

then the PDE in (2.32), together with (2.34) and (2.35), gives

$$\frac{du}{dt} = \frac{\partial u}{\partial t} + a \frac{\partial u}{\partial x} = 0 . \quad (2.36)$$

Therefore the rate of change of  $u$  along the characteristic curve  $x = x(t)$  satisfying (2.35) is zero, that is,  $u$  is *constant along the curve*  $x = x(t)$ . The speed  $a$  in (2.35) is called the *characteristic speed* and according to (2.35) it is the slope of the curve  $x = x(t)$  in the  $t$ - $x$  plane. In practice it is more common to use the  $x$ - $t$  plane to sketch the characteristics, in which case the slope of the curves in question is  $1/a$ . The family of characteristic curves  $x = x(t)$  given by



**Fig. 2.1.** Picture of characteristics for the linear advection equation for positive characteristic speed  $a$ . Initial condition at time  $t = 0$  fixes the initial position  $x_0$

the ODE (2.35) are illustrated in Fig. 2.1 for  $a > 0$  and are a one-parameter family of curves. A particular member of this family is determined when an *initial condition* (IC) at time  $t = 0$  for the ODE (2.35) is added. Suppose we set

$$x(0) = x_0 , \quad (2.37)$$

then the single characteristic curve that passes through the point  $(x_0, 0)$ , according to (2.35) is

$$x = x_0 + at . \quad (2.38)$$

This is also illustrated in Fig. 2.1. Now we may regard the initial position  $x_0$  as a parameter and in this way we reproduce the full one-parameter family

of characteristics. The fact that the curves are *parallel* is typical of linear hyperbolic PDEs with constant coefficients.

Recall the conclusion from (2.36) that  $u$  remains constant along characteristics. Thus, if  $u$  is given the initial value  $u(x, 0) = u_0(x)$  at time  $t = 0$ , then along the whole characteristic curve  $x(t) = x_0 + at$  that passes through the initial point  $x_0$  on the  $x$ -axis, the solution is

$$u(x, t) = u_0(x_0) = u_0(x - at). \quad (2.39)$$

The second equality follows from (2.38). The interpretation of the solution (2.39) of the PDE in (2.32) is this: *given an initial profile  $u_0(x)$ , the PDE will simply translate this profile with velocity  $a$  to the right if  $a > 0$  and to the left if  $a < 0$ .* The shape of the initial profile *remains unchanged*. The model equation in (2.32) under study contains some of the basic features of wave propagation phenomena, where *a wave is understood as some recognisable feature of a disturbance that travels at a finite speed*.

### 2.2.2 The Riemann Problem

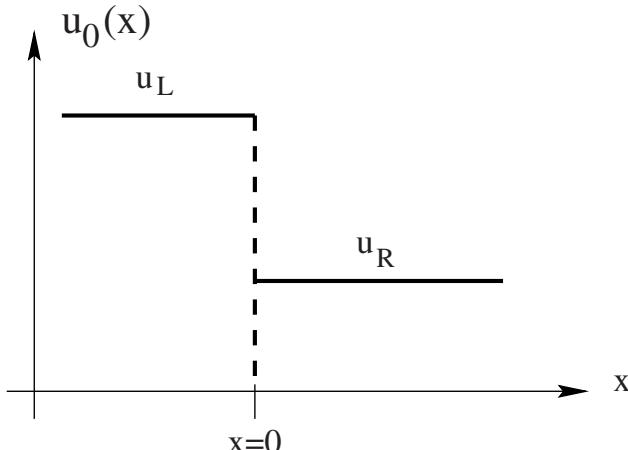
By using geometric arguments we have constructed the analytical solution of the general IVP (2.32) for the linear advection equation. This is given by (2.39) in terms of the initial data  $u_0(x)$ . Now we study a special IVP called the Riemann problem

$$\left. \begin{array}{l} \text{PDE:} \quad u_t + au_x = 0 \\ \text{IC:} \quad u(x, 0) = u_0(x) = \begin{cases} u_L & \text{if } x < 0, \\ u_R & \text{if } x > 0, \end{cases} \end{array} \right\} \quad (2.40)$$

where  $u_L$  (left) and  $u_R$  (right) are two constant values, as shown in Fig. 2.2. Note that the initial data has a discontinuity at  $x = 0$ . IVP (2.40) is the simplest initial-value problem one can pose. The trivial case would result when  $u_L = u_R$ . From the previous discussion on the solution of the general IVP (2.32) we expect any point on the initial profile to propagate a distance  $d = at$  in time  $t$ . In particular, we expect the initial discontinuity at  $x = 0$  to propagate a distance  $d = at$  in time  $t$ . This particular characteristic curve  $x = at$  will then separate those characteristic curves to the left, on which the solution takes on the value  $u_L$ , from those curves to the right, on which the solution takes on the value  $u_R$ ; see Fig. 2.3. So the solution of the Riemann problem (2.40) is simply

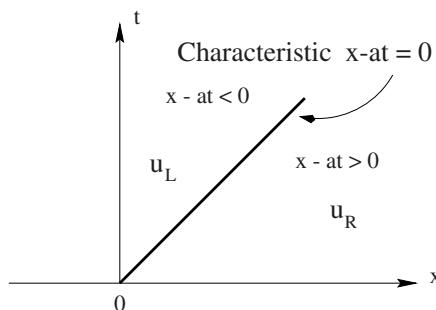
$$u(x, t) = u_0(x - at) = \begin{cases} u_L & \text{if } x - at < 0, \\ u_R & \text{if } x - at > 0. \end{cases} \quad (2.41)$$

Solution (2.41) also follows directly from the general solution (2.39), namely  $u(x, t) = u_0(x - at)$ . From (2.40),  $u_0(x - at) = u_L$  if  $x - at < 0$  and  $u_0(x - at) = u_R$  if  $x - at > 0$ .



**Fig. 2.2.** Illustration of the initial data for the Riemann problem. At the initial time the data consists of two constant states separated by a discontinuity at  $x = 0$

$u_R$  if  $x - at > 0$ . The solution of the Riemann problem can be represented in the  $x-t$  plane, as shown in Fig. 2.3. Through any point  $x_0$  on the  $x$ -axis one can draw a characteristic. As  $a$  is constant these are all parallel to each other. For the solution of the Riemann problem the characteristic that passes through  $x = 0$  is significant. This is the only one across which the solution changes.



**Fig. 2.3.** Illustration of the solution of the Riemann problem in the  $x-t$  plane for the linear advection equation with positive characteristic speed  $a$

## 2.3 Linear Hyperbolic Systems

In the previous section we studied in detail the behaviour and the general solution of the simplest PDE of hyperbolic type, namely the linear advection

equation with constant wave propagation speed. Here we extend the analysis to sets of  $m$  hyperbolic PDEs of the form

$$\mathbf{U}_t + \mathbf{A}\mathbf{U}_x = \mathbf{0} , \quad (2.42)$$

where the coefficient matrix  $\mathbf{A}$  is constant. From the assumption of hyperbolicity  $\mathbf{A}$  has  $m$  real eigenvalues  $\lambda_i$  and  $m$  linearly independent eigenvectors  $\mathbf{K}^{(i)}$ ,  $i = 1, \dots, m$ .

### 2.3.1 Diagonalisation and Characteristic Variables

In order to analyse and solve the general IVP for (2.42) it is found useful to transform the dependent variables  $\mathbf{U}(x, t)$  to a new set of dependent variables  $\mathbf{W}(x, t)$ . To this end we recall the following definition

**Definition 2.9 (Diagonalisable System).** A matrix  $\mathbf{A}$  is said to be *diagonalisable* if  $\mathbf{A}$  can be expressed as

$$\mathbf{A} = \mathbf{K}\Lambda\mathbf{K}^{-1} \text{ or } \Lambda = \mathbf{K}^{-1}\mathbf{A}\mathbf{K} , \quad (2.43)$$

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

$$\Lambda = \begin{bmatrix} \lambda_1 & \dots & 0 \\ 0 & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & \lambda_m \end{bmatrix} , \quad \mathbf{K} = [\mathbf{K}^{(1)}, \dots, \mathbf{K}^{(m)}] , \quad \mathbf{A}\mathbf{K}^{(i)} = \lambda_i\mathbf{K}^{(i)} . \quad (2.44)$$

A system (2.42) is said to be *diagonalisable* if the coefficient matrix  $\mathbf{A}$  is diagonalisable. Based on the concept of diagonalisation one often defines a hyperbolic system (2.42) as a system with real eigenvalues and diagonalisable coefficient matrix.

### Characteristic variables

The existence of the inverse matrix  $\mathbf{K}^{-1}$  makes it possible to define a new set of dependent variables  $\mathbf{W} = (w_1, w_2, \dots, w_m)^T$  via the transformation

$$\mathbf{W} = \mathbf{K}^{-1}\mathbf{U} \text{ or } \mathbf{U} = \mathbf{K}\mathbf{W} , \quad (2.45)$$

so that the linear system (2.42), when expressed in terms of  $\mathbf{W}$ , becomes *completely decoupled*, in a sense to be defined. The new variables  $\mathbf{W}$  are called *characteristic variables*. Next we derive the governing PDEs in terms of the characteristic variables, for which we need the partial derivatives  $\mathbf{U}_t$  and  $\mathbf{U}_x$  in equations (2.42). Since  $\mathbf{A}$  is constant,  $\mathbf{K}$  is also constant and therefore these derivatives are

$$\mathbf{U}_t = \mathbf{K}\mathbf{W}_t, \quad \mathbf{U}_x = \mathbf{K}\mathbf{W}_x.$$

Direct substitution of these expressions into equation (2.42) gives

$$\mathbf{K}\mathbf{W}_t + \mathbf{A}\mathbf{K}\mathbf{W}_x = \mathbf{0}.$$

Multiplication of this equation from the left by  $\mathbf{K}^{-1}$  and use of (2.43) gives

$$\mathbf{W}_t + \mathbf{A}\mathbf{W}_x = \mathbf{0}. \quad (2.46)$$

This is called the *canonical form* or *characteristic form* of system (2.42). When written in full this system becomes

$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix}_t + \begin{bmatrix} \lambda_1 & \dots & 0 \\ 0 & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & \lambda_m \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix}_x = \mathbf{0}. \quad (2.47)$$

Clearly the  $i$ -th PDE of this system is

$$\frac{\partial w_i}{\partial t} + \lambda_i \frac{\partial w_i}{\partial x} = 0, \quad i = 1, \dots, m \quad (2.48)$$

and involves the *single unknown*  $w_i(x, t)$ ; the system is therefore *decoupled* and is identical to the linear advection equation in (2.32); now the characteristic speed is  $\lambda_i$  and there are  $m$  characteristic curves satisfying  $m$  ODEs

$$\frac{dx}{dt} = \lambda_i, \quad \text{for } i = 1, \dots, m. \quad (2.49)$$

### 2.3.2 The General Initial-Value Problem

We now study the IVP for the PDEs (2.42). The initial condition is now denoted by superscript  $(0)$ , namely

$$\mathbf{U}^{(0)} = (u_1^{(0)}, \dots, u_m^{(0)})^T,$$

rather than by subscript  $0$ , as done for the scalar case. We find the general solution of the IVP by first solving the corresponding IVP for the canonical system (2.46) or (2.47) in terms of the characteristic variables  $\mathbf{W}$  and initial condition  $\mathbf{W}^{(0)} = (w_1^{(0)}, \dots, w_m^{(0)})^T$  such that

$$\mathbf{W}^{(0)} = \mathbf{K}^{-1}\mathbf{U}^{(0)} \text{ or } \mathbf{U}^{(0)} = \mathbf{K}\mathbf{W}^{(0)}.$$

The solution of the IVP for (2.46) is direct. By considering each unknown  $w_i(x, t)$  satisfying (2.48) and its corresponding initial data  $w_i^{(0)}$  we write its solution immediately as

$$w_i(x, t) = w_i^{(0)}(x - \lambda_i t), \quad \text{for } i = 1, \dots, m. \quad (2.50)$$

Compare with solution (2.39) for the scalar case. The solution of the general IVP in terms of the original variables  $\mathbf{U}$  is now obtained by transforming back according to (2.45), namely  $\mathbf{U} = \mathbf{K}\mathbf{W}$ .

*Example 2.10 (Linearised Gas Dynamics Revisited).* As a simple example we now study the general IVP for the linearised equations of Gas Dynamics (2.12), namely

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}_t + \begin{bmatrix} 0 & \rho_0 \\ a^2/\rho_0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}_x = \mathbf{0}, \quad u_1 \equiv \rho, \quad u_2 \equiv u,$$

with initial condition

$$\begin{bmatrix} u_1(x, 0) \\ u_2(x, 0) \end{bmatrix} = \begin{bmatrix} u_1^{(0)}(x) \\ u_2^{(0)}(x) \end{bmatrix}.$$

We define characteristic variables

$$\mathbf{W} = (w_1, w_2)^T = \mathbf{K}^{-1} \mathbf{U},$$

where  $\mathbf{K}$  is the matrix of right eigenvectors and  $\mathbf{K}^{-1}$  is its inverse, both given by

$$\mathbf{K} = \begin{bmatrix} \rho_0 & \rho_0 \\ -a & a \end{bmatrix}, \quad \mathbf{K}^{-1} = \frac{1}{2a\rho_0} \begin{bmatrix} a & -\rho_0 \\ a & \rho_0 \end{bmatrix}.$$

Since  $\lambda_1 = -a$  and  $\lambda_2 = a$ , in terms of the characteristic variables we may write

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}_t + \begin{bmatrix} -a & 0 \\ 0 & a \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}_x = \mathbf{0},$$

or in full

$$\frac{\partial w_1}{\partial t} - a \frac{\partial w_1}{\partial x} = 0, \quad \frac{\partial w_2}{\partial t} + a \frac{\partial w_2}{\partial x} = 0.$$

The initial condition satisfies

$$\begin{bmatrix} w_1^{(0)} \\ w_2^{(0)} \end{bmatrix} = \mathbf{K}^{-1} \begin{bmatrix} u_1^{(0)} \\ u_2^{(0)} \end{bmatrix},$$

or in full

$$w_1^{(0)}(x) = \frac{1}{2a\rho_0} \left[ a u_1^{(0)}(x) - \rho_0 u_2^{(0)}(x) \right],$$

$$w_2^{(0)}(x) = \frac{1}{2a\rho_0} \left[ a u_1^{(0)}(x) + \rho_0 u_2^{(0)}(x) \right].$$

Each equation involves a single independent variable and is a linear advection equation of the form (2.48). The solution for  $w_1$  and  $w_2$  in terms of their initial data  $w_1^{(0)}$ ,  $w_2^{(0)}$ , according to (2.50) is

$$w_1(x, t) = w_1^{(0)}(x + at), \quad w_2(x, t) = w_2^{(0)}(x - at),$$

or in full

$$w_1(x, t) = \frac{1}{2a\rho_0} \left[ a u_1^{(0)}(x + at) - \rho_0 u_2^{(0)}(x + at) \right],$$

$$w_2(x, t) = \frac{1}{2a\rho_0} \left[ au_1^{(0)}(x - at) + \rho_0 u_2^{(0)}(x - at) \right] .$$

This is the solution in terms of the characteristic variables. In order to obtain the solution to the original problem we transform back using  $\mathbf{U} = \mathbf{K}\mathbf{W}$ . This gives the final solution as

$$\begin{aligned} u_1(x, t) &= \left. \begin{aligned} &\frac{1}{2a} \left[ au_1^{(0)}(x + at) - \rho_0 u_2^{(0)}(x + at) \right] \\ &+ \frac{1}{2a} \left[ au_1^{(0)}(x - at) + \rho_0 u_2^{(0)}(x - at) \right] \end{aligned} \right\}, \\ u_2(x, t) &= \left. \begin{aligned} &-\frac{1}{2\rho_0} \left[ au_1^{(0)}(x + at) - \rho_0 u_2^{(0)}(x + at) \right] \\ &+ \frac{1}{2\rho_0} \left[ au_1^{(0)}(x - at) + \rho_0 u_2^{(0)}(x - at) \right] \end{aligned} \right\}. \end{aligned}$$

**Exercise 2.11.** Find the solution of the general IVP for the Small Perturbation Equations (2.24) using the above methodology.

**Solution 2.12.** (Left to the reader).

We return to the expression  $\mathbf{U} = \mathbf{K}\mathbf{W}$  in (2.45) used to recover the solution to the original problem. When written in full this expression becomes

$$\begin{aligned} u_1 &= w_1 k_1^{(1)} + w_2 k_1^{(2)} + \dots + w_m k_1^{(m)}, \\ u_i &= w_1 k_i^{(1)} + w_2 k_i^{(2)} + \dots + w_m k_i^{(m)}, \\ u_m &= w_1 k_m^{(1)} + w_2 k_m^{(2)} + \dots + w_m k_m^{(m)}, \end{aligned}$$

or

$$\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix} = w_1 \begin{bmatrix} k_1^{(1)} \\ k_2^{(1)} \\ \vdots \\ k_m^{(1)} \end{bmatrix} + w_2 \begin{bmatrix} k_1^{(2)} \\ k_2^{(2)} \\ \vdots \\ k_m^{(2)} \end{bmatrix} + \dots + w_m \begin{bmatrix} k_1^{(m)} \\ k_2^{(m)} \\ \vdots \\ k_m^{(m)} \end{bmatrix}, \quad (2.51)$$

or more succinctly

$$\mathbf{U}(x, t) = \sum_{i=1}^m w_i(x, t) \mathbf{K}^{(i)}. \quad (2.52)$$

This means that the function  $w_i(x, t)$  is the coefficient of  $\mathbf{K}^{(i)}$  in an *eigenvector expansion* of the vector  $\mathbf{U}$ . But according to (2.50),  $w_i(x, t) = w_i^{(0)}(x - \lambda_i t)$  and hence

$$\mathbf{U}(x, t) = \sum_{i=1}^m w_i^{(0)}(x - \lambda_i t) \mathbf{K}^{(i)}. \quad (2.53)$$

Thus, given a point  $(x, t)$  in the  $x$ - $t$  plane, the solution  $\mathbf{U}(x, t)$  at this point depends only on the initial data at the  $m$  points  $x_0^{(i)} = x - \lambda_i t$ . These are the

intersections of the characteristics of speed  $\lambda_i$  with the  $x$ -axis. The solution (2.53) for  $\mathbf{U}$  can be seen as the superposition of  $m$  waves, each of which is advected independently without change in shape. The  $i$ -th wave has shape  $w_i^{(0)}(x)\mathbf{K}^{(i)}$  and propagates with speed  $\lambda_i$ .

### 2.3.3 The Riemann Problem

We study the Riemann problem for the hyperbolic, constant coefficient system (2.42). This is the special IVP

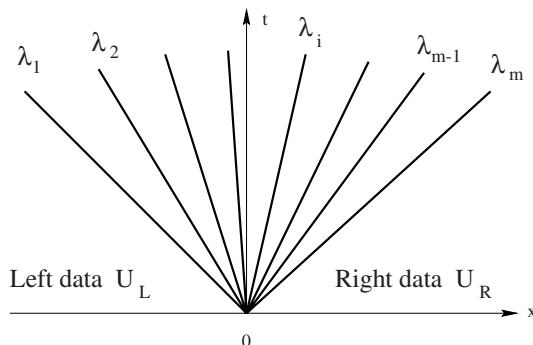
$$\left. \begin{array}{l} \text{PDEs: } \mathbf{U}_t + \mathbf{A}\mathbf{U}_x = \mathbf{0}, \quad -\infty < x < \infty, \quad t > 0, \\ \text{IC: } \mathbf{U}(x, 0) = \mathbf{U}^{(0)}(x) = \begin{cases} \mathbf{U}_L & x < 0, \\ \mathbf{U}_R & x > 0 \end{cases} \end{array} \right\} \quad (2.54)$$

and is a generalisation of the IVP (2.32). We assume that the system is *strictly hyperbolic* and we order the real and distinct eigenvalues as

$$\lambda_1 < \lambda_2 < \dots < \lambda_m. \quad (2.55)$$

### The General Solution

The structure of the solution of the Riemann problem (2.54) in the  $x-t$  plane is depicted in Fig. 2.4. It consists of  $m$  waves emanating from the origin, one for each eigenvalue  $\lambda_i$ . Each wave  $i$  carries a *jump discontinuity* in  $\mathbf{U}$  propagating with speed  $\lambda_i$ . Naturally, the solution to the left of the  $\lambda_1$ -wave is simply the initial data  $\mathbf{U}_L$  and to the right of the  $\lambda_m$ -wave is  $\mathbf{U}_R$ . The task at hand is to find the solution in the wedge between the  $\lambda_1$  and  $\lambda_m$  waves. As the eigenvectors  $\mathbf{K}^{(1)}, \dots, \mathbf{K}^{(m)}$  are *linearly independent*, we



**Fig. 2.4.** Structure of the solution of the Riemann problem for a general  $m \times m$  linear hyperbolic system with constant coefficients

can expand the data  $\mathbf{U}_L$ , constant left state, and  $\mathbf{U}_R$ , constant right state, as linear combinations of the set  $\mathbf{K}^{(1)}, \dots, \mathbf{K}^{(m)}$ , that is

$$\mathbf{U}_L = \sum_{i=1}^m \alpha_i \mathbf{K}^{(i)}, \quad \mathbf{U}_R = \sum_{i=1}^m \beta_i \mathbf{K}^{(i)}, \quad (2.56)$$

with constant coefficients  $\alpha_i, \beta_i$ , for  $i = 1, \dots, m$ . Formally, the solution of the IVP (2.54) is given by (2.53) in terms of the initial data  $w_i^{(0)}(x)$  for the characteristic variables and the right eigenvectors  $\mathbf{K}^{(i)}$ . Note that each of the expansions in (2.56) is a special case of (2.53). In terms of the characteristic variables we have  $m$  scalar Riemann problems for the PDEs

$$\frac{\partial w_i}{\partial t} + \lambda_i \frac{\partial w_i}{\partial x} = 0, \quad (2.57)$$

with initial data obtained by comparing (2.56) with (2.53), that is

$$w_i^{(0)}(x) = \begin{cases} \alpha_i & \text{if } x < 0, \\ \beta_i & \text{if } x > 0, \end{cases} \quad (2.58)$$

for  $i = 1, \dots, m$ . From the previous results, see equation (2.50), we know that the solutions of these scalar Riemann problems are given by

$$w_i(x, t) = w_i^{(0)}(x - \lambda_i t) = \begin{cases} \alpha_i & \text{if } x - \lambda_i t < 0, \\ \beta_i & \text{if } x - \lambda_i t > 0. \end{cases} \quad (2.59)$$

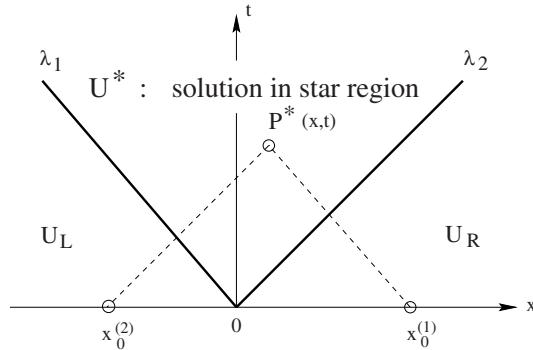
For a given point  $(x, t)$  there is an eigenvalue  $\lambda_I$  such that  $\lambda_I < \frac{x}{t} < \lambda_{I+1}$ , that is  $x - \lambda_i t > 0 \forall i$  such that  $i \leq I$ . We can thus write the final solution to the Riemann problem (2.54) in terms of the original variables as

$$\mathbf{U}(x, t) = \sum_{i=I+1}^m \alpha_i \mathbf{K}^{(i)} + \sum_{i=1}^I \beta_i \mathbf{K}^{(i)}, \quad (2.60)$$

where the integer  $I = I(x, t)$  is the maximum value of the sub-index  $i$  for which  $x - \lambda_i t > 0$ .

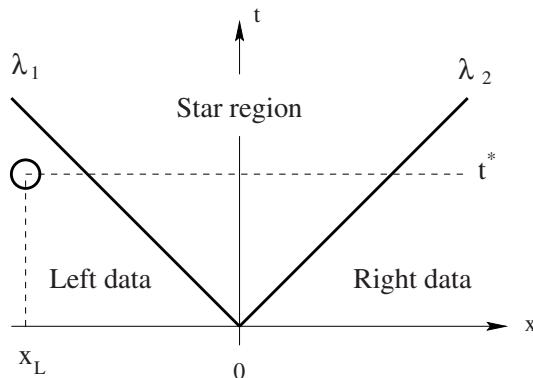
## The Solution for a $2 \times 2$ System

As an example consider the Riemann problem for a general  $2 \times 2$  linear system. From the origin  $(0, 0)$  in the  $(x, t)$  plane there will be two waves travelling with speeds that are equal to the characteristic speeds  $\lambda_1$  and  $\lambda_2$  ( $\lambda_1 < \lambda_2$ ); see Fig. 2.5. The solution to the left of  $dx/dt = \lambda_1$  is simply the data state  $\mathbf{U}_L = \alpha_1 \mathbf{K}^{(1)} + \alpha_2 \mathbf{K}^{(2)}$  and to the right of  $dx/dt = \lambda_2$  the solution is the constant data state  $\mathbf{U}_R = \beta_1 \mathbf{K}^{(1)} + \beta_2 \mathbf{K}^{(2)}$ . The wedge between the  $\lambda_1$  and  $\lambda_2$  waves is usually called the *Star Region* and the solution there is denoted by  $\mathbf{U}^*$ ; its value is *due to the passage of two waves emerging from*



**Fig. 2.5.** Structure of the solution of the Riemann problem for a  $2 \times 2$  linear system with constant coefficients

*the origin of the initial discontinuity.* From the point  $P^*(x, t)$  we trace back the characteristics with speeds  $\lambda_1$  and  $\lambda_2$ . These are parallel to those passing through the origin. The characteristics through  $P^*$  pass through the initial points  $x_0^{(2)} = x - \lambda_2 t$  and  $x_0^{(1)} = x - \lambda_1 t$ . The coefficients in the expansion (2.60) for  $\mathbf{U}(x, t)$  are thus determined. The solution at a point  $P^*$  has the form (2.60). It is a question of choosing the correct coefficients  $\alpha_i$  or  $\beta_i$ . Select a time  $t^*$  and a point  $x_L$  to the left of the slowest wave so  $\mathbf{U}(x_L, t^*) = \mathbf{U}_L$ , see Fig. 2.6. The solution at the starting point  $(x_L, t^*)$  is obviously



**Fig. 2.6.** The Riemann problem solution found by travelling along dashed horizontal line  $t = t^*$

$$\mathbf{U}_L = \sum_{i=1}^2 \alpha_i \mathbf{K}_i = \alpha_1 \mathbf{K}^{(1)} + \alpha_2 \mathbf{K}^{(2)},$$

i.e. all coefficients are  $\alpha$ 's, that is, the point  $(x_L, t^*)$  lies to the left of every wave. As we move to the right of  $(x_L, t^*)$  on the horizontal line  $t = t^*$  we cross the wave  $dx/dt = \lambda_1$ , hence  $x - \lambda_1 t$  changes from negative to positive, see (2.59), and therefore the coefficient  $\alpha_1$  above changes to  $\beta_1$ . Thus the solution in the entire *Star Region*, between the  $\lambda_1$  and  $\lambda_2$  waves, is

$$\mathbf{U}^*(x, t) = \beta_1 \mathbf{K}^{(1)} + \alpha_2 \mathbf{K}^{(2)} . \quad (2.61)$$

As we continue moving right and cross the  $\lambda_2$  wave the value  $x - \lambda_2 t$  changes from negative to positive and hence the coefficient  $\alpha_2$  in (2.60) and (2.61) changes to  $\beta_2$ , i.e the solution to the right of the fastest wave of speed  $\lambda_2$  is, trivially,

$$\mathbf{U}_R = \beta_1 \mathbf{K}^{(1)} + \beta_2 \mathbf{K}^{(2)} .$$

*Remark 2.13.* From equation (2.56) it is easy to see that the jump in  $\mathbf{U}$  across the whole wave structure in the solution of the Riemann problem is

$$\Delta \mathbf{U} = \mathbf{U}_R - \mathbf{U}_L = (\beta_1 - \alpha_1) \mathbf{K}^{(1)} + \dots + (\beta_m - \alpha_m) \mathbf{K}^{(m)} . \quad (2.62)$$

It is an eigenvector expansion with coefficients that are the strengths of the waves present in the Riemann problem. The wave strength of wave  $i$  is  $\beta_i - \alpha_i$  and the jump in  $\mathbf{U}$  across wave  $i$ , denoted by  $(\Delta \mathbf{U})_i$ , is

$$(\Delta \mathbf{U})_i = (\beta_i - \alpha_i) \mathbf{K}^{(i)} . \quad (2.63)$$

When solving the Riemann problem, sometimes it is more useful to expand the total jump  $\Delta \mathbf{U} = \mathbf{U}_R - \mathbf{U}_L$  in terms of the eigenvectors and unknown wave strengths  $\delta_i = \beta_i - \alpha_i$ .

### 2.3.4 The Riemann Problem for Linearised Gas Dynamics

As an illustrative example we apply the methodology described in the previous section to solve the Riemann problem for the linearised equations of Gas Dynamics (2.12)

$$\mathbf{U}_t + \mathbf{A} \mathbf{U}_x = \mathbf{0} ,$$

with

$$\mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \equiv \begin{bmatrix} \rho \\ u \end{bmatrix} , \quad \mathbf{A} = \begin{bmatrix} 0 & \rho_0 \\ a^2/\rho_0 & 0 \end{bmatrix} .$$

The eigenvalues of the system are

$$\lambda_1 = -a , \quad \lambda_2 = +a ,$$

and the corresponding right eigenvectors are

$$\mathbf{K}^{(1)} = \begin{bmatrix} \rho_0 \\ -a \end{bmatrix} , \quad \mathbf{K}^{(2)} = \begin{bmatrix} \rho_0 \\ a \end{bmatrix} .$$

First we decompose the left data state  $\mathbf{U}_L = [\rho_L, u_L]^T$  in terms of the right eigenvectors according to equation (2.56), namely

$$\mathbf{U}_L = \begin{bmatrix} \rho_L \\ u_L \end{bmatrix} = \alpha_1 \begin{bmatrix} \rho_0 \\ -a \end{bmatrix} + \alpha_2 \begin{bmatrix} \rho_0 \\ a \end{bmatrix}.$$

Solving for the unknown coefficients  $\alpha_1$  and  $\alpha_2$  we obtain

$$\alpha_1 = \frac{a\rho_L - \rho_0 u_L}{2a\rho_0}, \quad \alpha_2 = \frac{a\rho_L + \rho_0 u_L}{2a\rho_0}.$$

Similarly, by expanding the right-hand data  $\mathbf{U}_R = [\rho_R, u_R]^T$  in terms of the eigenvectors and solving for the coefficients  $\beta_1$  and  $\beta_2$  we obtain

$$\beta_1 = \frac{a\rho_R - \rho_0 u_R}{2a\rho_0}, \quad \beta_2 = \frac{a\rho_R + \rho_0 u_R}{2a\rho_0}.$$

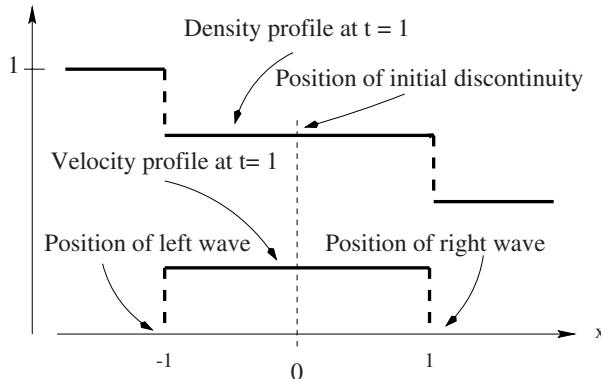
Now by using equation (2.61) we find the solution in the star region as

$$\mathbf{U}^* = \begin{bmatrix} \rho^* \\ u^* \end{bmatrix} = \beta_1 \begin{bmatrix} \rho_0 \\ -a \end{bmatrix} + \alpha_2 \begin{bmatrix} \rho_0 \\ a \end{bmatrix}.$$

After some algebraic manipulations we obtain the solution explicitly as

$$\left. \begin{aligned} \rho_* &= \frac{1}{2}(\rho_L + \rho_R) - \frac{1}{2}(u_R - u_L)\rho_0/a, \\ u_* &= \frac{1}{2}(u_L + u_R) - \frac{1}{2}(\rho_R - \rho_L)a/\rho_0. \end{aligned} \right\} \quad (2.64)$$

Fig. 2.7 illustrates the solution for  $\rho(x, t)$  and  $u(x, t)$  at time  $t = 1$  for the parameter values  $\rho_0 = 1$ ,  $a = 1$  and initial data  $\rho_L = 1$ ,  $u_L = 0$ ,  $\rho_R = \frac{1}{2}$  and  $u_R = 0$ . The two symmetric waves that emerge from the initial position of the discontinuity carry a discontinuous jump in both density  $\rho$  and velocity  $u$ .



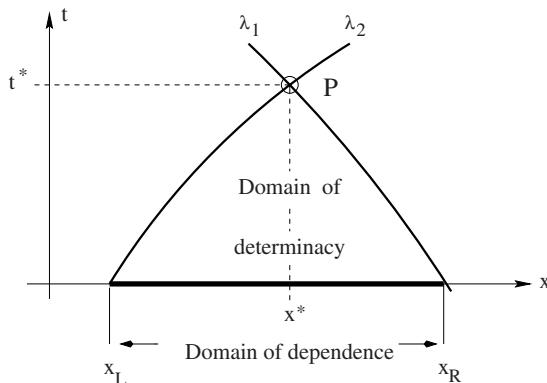
**Fig. 2.7.** Density and velocity solution profiles at time  $t=1$

*Remark 2.14.* The exact solution (2.64) can be very useful in testing numerical methods for systems with discontinuous solutions.

### 2.3.5 Some Useful Definitions

Next we recall some standard definitions associated with hyperbolic systems.

**Definition 2.15 (Domain of Dependence).** *Recall that for the linear advection equation the solution at a given point  $P = (x^*, t^*)$  depends solely on the initial data at a single point  $x_0$  on the  $x$ -axis. This point is obtained by tracing back the characteristic passing through the point  $P = (x^*, t^*)$ . As a matter of fact, the solution at  $P = (x^*, t^*)$  is identical to the value of the initial data  $u_0(x)$  at the point  $x_0$ . One says that the domain of dependence of the point  $P = (x^*, t^*)$  is the point  $x_0$ . For a  $2 \times 2$  system the domain of dependence is an interval  $[x_L, x_R]$  on the  $x$ -axis that is subtended by the characteristics passing through the point  $P = (x^*, t^*)$ .*



**Fig. 2.8.** Domain of dependence of point  $P$  and corresponding domain of determinacy, for a  $2 \times 2$  system

Fig. 2.8 illustrates the domain of dependence for a  $2 \times 2$  system with characteristic speeds  $\lambda_1$  and  $\lambda_2$ , with  $\lambda_1 < \lambda_2$ . In general, the characteristics of a hyperbolic system are curved. For a larger system the domain of dependence is determined by the slowest and fastest characteristics and is always a bounded interval, as the characteristic speeds for hyperbolic systems are always finite.

**Definition 2.16 (Domain of Determinacy).** *For a given domain of dependence  $[x_L, x_R]$ , the domain of determinacy is the set of points  $(x, t)$ , within the domain of existence of the solution  $\mathbf{U}(x, t)$ , in which  $\mathbf{U}(x, t)$  is solely determined by initial data on  $[x_L, x_R]$ .*

In Fig. 2.8 we illustrate the domain of determinacy of an interval  $[x_L, x_R]$  for the case of a  $2 \times 2$  system with characteristic speeds  $\lambda_1$  and  $\lambda_2$ , with  $\lambda_1 < \lambda_2$ .

**Definition 2.17 (Range of Influence).** Another useful concept is that of the range of influence of a point  $Q = (x_0, 0)$  on the  $x$ -axis. It is defined as the set of points  $(x, t)$  in the  $x-t$  plane in which the solution  $\mathbf{U}(x, t)$  is influenced by initial data at the point  $Q = (x_0, 0)$ .

Fig. 2.9 illustrates the range of influence of a point  $Q = (x_0, 0)$  for the case of a  $2 \times 2$  system with characteristic speeds  $\lambda_1$  and  $\lambda_2$ , with  $\lambda_1 < \lambda_2$ .

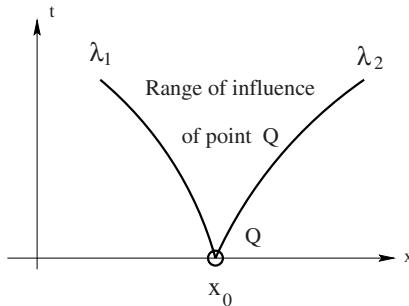


Fig. 2.9. Range of influence of point  $Q$  for a  $2$  by  $2$  system

## 2.4 Conservation Laws

The purpose of this section is to provide the reader with a succinct presentation of some mathematical properties of hyperbolic conservation laws. We restrict our attention to those properties thought to be essential to the development and application of numerical methods for conservation laws. In Chap. 1 we applied the physical principles of conservation of mass, momentum and energy to derive time-dependent, multidimensional non-linear systems of conservations laws. In this section we restrict ourselves to simple model problems. In Sect. 2.1 we advanced the formal definition of a system of  $m$  conservation laws

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0}, \quad (2.65)$$

where  $\mathbf{U}$  is the vector of conserved variables and  $\mathbf{F}(\mathbf{U})$  is the vector of fluxes. This system is hyperbolic if the Jacobian matrix

$$\mathbf{A}(\mathbf{U}) = \frac{\partial \mathbf{F}}{\partial \mathbf{U}}$$

has real eigenvalues  $\lambda_i(\mathbf{U})$  and a complete set of linearly independent eigenvectors  $\mathbf{K}^{(i)}(\mathbf{U})$ ,  $i = 1, \dots, m$ , which we assume to be ordered as

$$\lambda_1(\mathbf{U}) < \lambda_2(\mathbf{U}) < \dots < \lambda_m(\mathbf{U}),$$

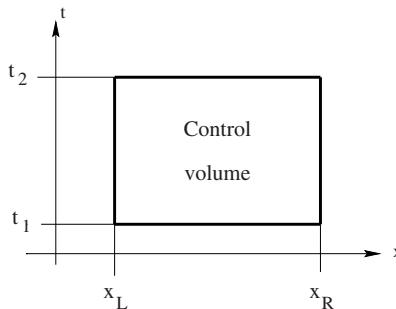
$$\mathbf{K}^{(1)}(\mathbf{U}), \mathbf{K}^{(2)}(\mathbf{U}), \dots, \mathbf{K}^{(m)}(\mathbf{U}).$$

It is important to note that now eigenvalues and eigenvectors depend on  $\mathbf{U}$ , although sometimes we shall omit the argument  $\mathbf{U}$ .

### 2.4.1 Integral Forms of Conservation Laws

As discussed in Sect. 1.5 of Chap. 1, conservation laws may be expressed in differential and integral form. There are two good reasons for considering the integral form (s) of the conservation laws: (i) the derivation of the governing equations is based on physical conservation principles expressed as integral relations on control volumes, (ii) the integral formulation requires less smoothness of the solution, which paves the way to extending the class of admissible solutions to include discontinuous solutions.

The integral form has variants that are worth studying in detail. Consider a one-dimensional time dependent system, such as the Euler equations introduced in Sect. 1.1 of Chap. 1. Choose a control volume  $V = [x_L, x_R] \times [t_1, t_2]$  on the  $x-t$  plane as shown in Fig. 2.10. The integral form, see Sect. 1.5, of the



**Fig. 2.10.** A control volume  $V = [x_L, x_R] \times [t_1, t_2]$  on  $x-t$  plane

equation for conservation of mass in one space dimension is

$$\frac{d}{dt} \int_{x_L}^{x_R} \rho(x, t) dx = f(x_L, t) - f(x_R, t) ,$$

where  $f = \rho u$  is the flux. For the complete system we have

$$\frac{d}{dt} \int_{x_L}^{x_R} \mathbf{U}(x, t) dx = \mathbf{F}(\mathbf{U}(x_L, t)) - \mathbf{F}(\mathbf{U}(x_R, t)) , \quad (2.66)$$

where  $\mathbf{F}(\mathbf{U})$  is the flux vector. This is one version of the integral form of the conservation laws: *Integral Form I*. The corresponding differential form reads as (2.65). Another version of the integral form of the conservation laws is obtained by integrating (2.66) in time between  $t_1$  and  $t_2$ , with  $t_1 \leq t_2$ . See Fig. 2.10. Clearly,

$$\int_{t_1}^{t_2} \left[ \frac{d}{dt} \int_{x_L}^{x_R} \mathbf{U}(x, t) dx \right] dt = \int_{x_L}^{x_R} \mathbf{U}(x, t_2) dx - \int_{x_L}^{x_R} \mathbf{U}(x, t_1) dx$$

and thus (2.66) becomes

$$\left. \begin{aligned} \int_{x_L}^{x_R} \mathbf{U}(x, t_2) dx &= \int_{x_L}^{x_R} \mathbf{U}(x, t_1) dx + \int_{t_1}^{t_2} \mathbf{F}(\mathbf{U}(x_L, t)) dt \\ &\quad - \int_{t_1}^{t_2} \mathbf{F}(\mathbf{U}(x_R, t)) dt , \end{aligned} \right\} \quad (2.67)$$

which we call: *Integral Form II* of the conservation laws.

Another version of the integral form of the conservation laws is obtained by integrating (2.65) in any domain  $V$  in  $x-t$  space and using Green's theorem. The result is

$$\oint [\mathbf{U} dx - \mathbf{F}(\mathbf{U}) dt] = \mathbf{0} , \quad (2.68)$$

where the line integration is performed along the boundary of the domain, in an anticlock-wise manner. We call this version *Integral Form III* of the conservation laws. Note that *Integral Form II* of the conservation laws is a special case of *Integral Form III*, in which the control volume  $V$  is the rectangle  $[x_L, x_R] \times [t_1, t_2]$ .

A fourth integral form results from adopting a more mathematical approach for extending the concept of solution of (2.65) to include discontinuities. See Chorin and Marsden [112]. A *weak* or *generalized* solution  $\mathbf{U}$  is required to satisfy the integral relation

$$\int_0^{+\infty} \int_{-\infty}^{+\infty} [\phi_t \mathbf{U} + \phi_x \mathbf{F}(\mathbf{U})] dx dt = - \int_{-\infty}^{+\infty} \phi(x, 0) \mathbf{U}(x, 0) dx , \quad (2.69)$$

for all *test functions*  $\phi(x, t)$  that are continuously differentiable and have *compact support*. A function  $\phi(x, t)$  has compact support if it vanishes outside some bounded set. Note that in (2.69) the derivatives of  $\mathbf{U}(x, t)$  and  $\mathbf{F}(\mathbf{U})$  have been passed on to the test function  $\phi(x, t)$ , which is sufficiently smooth to admit these derivatives.

*Remark 2.18.* The integral forms (2.66)–(2.69) corresponding to (2.65) are valid for any system (2.65), not just for the Euler equations.

## Examples of Conservation Laws

*Scalar* conservation laws ( $m = 1$ ) in differential form read

$$u_t + f(u)_x = 0 , \quad f(u) : \text{flux function.} \quad (2.70)$$

To be able to solve for the conserved variable  $u(x, t)$  the flux function  $f(u)$  must be a completely determined algebraic function of  $u(x, t)$ , and possibly some extra parameters of the problem. As seen in Sect. 2.2 the linear advection

equation is the simplest example, in which the flux function is  $f(u) = au$ , a linear function of  $u$ .

The **inviscid Burgers's equation** has flux  $f(u) = \frac{1}{2}u^2$ , a quadratic function of  $u$ . Another example of a conservation law is the **traffic flow equation**

$$\rho_t + f(\rho)_x = 0, \quad f(\rho) = u_m(1 - \frac{\rho}{\rho_m})\rho. \quad (2.71)$$

Here the conserved variable  $\rho(x, t)$  is a density function (density of motor vehicles),  $u_m$  and  $\rho_m$  are parameters of the problem, namely the maximum speed of vehicles and the maximum density, both positive constants. For details on the traffic flow equation see Whitham [582], Zachmanoglou and Thoe [596], Toro [528] and Haberman [232]. An example of practical interest in oil-reservoir simulation is the **Buckley-Leverett equation**

$$u_t + f(u)_x = 0, \quad f(u) = \frac{u^2}{u^2 + b(1-u)^2}, \quad (2.72)$$

where  $b$  is a parameter of the problem. More details of this equation are found in LeVeque [308].

**Systems of conservation laws** are constructed, as obvious examples, from linear systems

$$\mathbf{U}_t + \mathbf{A}\mathbf{U}_x = \mathbf{0},$$

with constant coefficient matrix  $\mathbf{A}$ . The required *conservation-law form* is obtained by defining the flux function as the product of the coefficient matrix  $\mathbf{A}$  and the vector  $\mathbf{U}$ , namely

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0}, \quad \mathbf{F}(\mathbf{U}) = \mathbf{A}\mathbf{U}. \quad (2.73)$$

Trivially, the Jacobian matrix is  $\mathbf{A}$ .

*Example 2.19 (Isothermal Gas Dynamics).* The isothermal equations of Gas Dynamics, see Sect. 1.6.2 of Chap. 1, are one example of a non-linear system of conservation laws. These are

$$\left. \begin{aligned} \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x &= \mathbf{0}, \\ \mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} &\equiv \begin{bmatrix} \rho \\ \rho u \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \equiv \begin{bmatrix} \rho u \\ \rho u^2 + a^2 \rho \end{bmatrix}, \end{aligned} \right\} \quad (2.74)$$

where  $a$  is positive, constant speed of sound. The Jacobian matrix is found by first expressing  $\mathbf{F}$  in terms of the components  $u_1 \equiv \rho$  and  $u_2 \equiv \rho u$  of the vector  $\mathbf{U}$  of conserved variables, namely

$$\mathbf{F}(\mathbf{U}) = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \equiv \begin{bmatrix} u_2 \\ u_2^2/u_1 + a^2 u_1 \end{bmatrix}.$$

According to (2.8) the Jacobian matrix is

$$\mathbf{A}(\mathbf{U}) = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} = \begin{bmatrix} 0 & 1 \\ -(u_2/u_1)^2 + a^2 & 2u_2/u_1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ a^2 - u^2 & 2u \end{bmatrix} .$$

It is left to the reader to verify that the eigenvalues of  $\mathbf{A}$  are

$$\lambda_1 = u - a , \quad \lambda_2 = u + a \quad (2.75)$$

and that the right eigenvectors are

$$\mathbf{K}^{(1)} = \begin{bmatrix} 1 \\ u - a \end{bmatrix} , \quad \mathbf{K}^{(2)} = \begin{bmatrix} 1 \\ u + a \end{bmatrix} , \quad (2.76)$$

where the scaling factors for  $\mathbf{K}^{(1)}$  and  $\mathbf{K}^{(2)}$  have been taken to be unity. The isothermal equations of Gas Dynamics are thus hyperbolic.

*Example 2.20 (Isentropic Gas Dynamics).* Another non-linear example of a system of conservation laws are the isentropic equations of Gas Dynamics

$$\left. \begin{aligned} \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x &= \mathbf{0} , \\ \mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} &\equiv \begin{bmatrix} \rho \\ \rho u \end{bmatrix} , \quad \mathbf{F} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \equiv \begin{bmatrix} \rho u \\ \rho u^2 + p \end{bmatrix} , \end{aligned} \right\} \quad (2.77)$$

together with the closure condition, or equation of state (EOS),

$$p = C\rho^\gamma , \quad C = \text{constant} . \quad (2.78)$$

See Sect. 1.6.2 of Chap. 1.

**Exercise 2.21.** (i) Find the Jacobian matrix, the eigenvalues and the right eigenvectors for the isentropic equations (2.77)–(2.78). (ii) Show that for a generalized isentropic EOS,  $p = p(\rho)$ , the system is hyperbolic if and only if  $p'(\rho) > 0$ , that is, the pressure must be a monotone increasing function of  $\rho$ . (iii) Show that the sound speed has the general form

$$a = \sqrt{p'(\rho)} .$$

**Solution 2.22.** The eigenvalues are

$$\lambda_1 = u - a , \quad \lambda_2 = u + a , \quad (2.79)$$

and the right eigenvectors are

$$\mathbf{K}^{(1)} = \begin{bmatrix} 1 \\ u - a \end{bmatrix} , \quad \mathbf{K}^{(2)} = \begin{bmatrix} 1 \\ u + a \end{bmatrix} , \quad (2.80)$$

with the sound speed  $a$  as claimed.

### 2.4.2 Non-Linearities and Shock Formation

Here we study some distinguishing features of non-linear hyperbolic conservation laws, such as wave steepening and shock formation. We restrict our attention to the initial-value problem for scalar non-linear conservation laws, namely

$$u_t + f(u)_x = 0, \quad u(x, 0) = u_0(x). \quad (2.81)$$

A corresponding integral form of the conservation law is

$$\frac{d}{dt} \int_{x_L}^{x_R} u(x, t) dx = f(u(x_L, t)) - f(u(x_R, t)). \quad (2.82)$$

The flux function  $f$  is assumed to be a function of  $u$  only, which under certain circumstances is an inadequate representation of the physical problem being modelled. Relevant physical phenomena of our interest are shock waves in compressible media. These have viscous dissipation and heat conduction, in addition to pure advection. A more appropriate flux function for a model conservation law would also include a dependence on  $u_x$ , so that the modified conservation law would read

$$u_t + f(u)_x = \alpha u_{xx}, \quad (2.83)$$

with  $\alpha$  a positive coefficient of viscosity. The conservation law in (2.81) may be rewritten as

$$u_t + \lambda(u)u_x = 0, \quad (2.84)$$

where

$$\lambda(u) = \frac{df}{du} = f'(u) \quad (2.85)$$

is the *characteristic speed*. In the system case this corresponds to the eigenvalues of the Jacobian matrix. For the linear advection equation  $\lambda(u) = a$ , constant. For the inviscid Burgers equation  $\lambda(u) = u$ , that is, the characteristic speed depends on the solution and is in fact identical to the conserved variable. For the traffic flow equation  $\lambda(u) = u_m(1 - \frac{2u}{\rho_m})$ .

The behaviour of the flux function  $f(u)$  has profound consequences on the behaviour of the solution  $u(x, t)$  of the conservation law itself. A crucial property is *monotonicity* of the characteristic speed  $\lambda(u)$ . There are essentially three possibilities:

- $\lambda(u)$  is a monotone *increasing* function of  $u$ , i.e.

$$\frac{d\lambda(u)}{du} = \lambda'(u) = f''(u) > 0 \quad (\text{convex flux})$$

- $\lambda(u)$  is a monotone *decreasing* function of  $u$ , i.e.

$$\frac{d\lambda(u)}{du} = \lambda'(u) = f''(u) < 0 \quad (\text{concave flux})$$

- $\lambda(u)$  has extrema, for some  $u$ , i.e.

$$\frac{d\lambda(u)}{du} = \lambda'(u) = f''(u) = 0 \text{ (non-convex, non-concave flux).}$$

In the case of non-linear systems of conservation laws the character of the flux function is determined by the Equation of State. One speaks of convex, or otherwise, equations of state. See the review paper by Menikoff and Plohr [349]. For the inviscid Burgers equation  $\lambda'(u) = f''(u) = 1 > 0$ , the flux is convex. For the traffic flow equation  $\lambda'(u) = f''(u) = -2u_m/\rho_m < 0$ , the flux is concave.

**Exercise 2.23.** Analyse the character of the flux function for the Buckley–Leverett equation and show that it is non-convex, non-concave.

**Solution 2.24.** (Left to the reader).

We study the *inviscid* IVP (2.81) and for the moment we assume that the initial data  $u(x, 0) = u_0(x)$  is smooth. For some finite time the solution  $u(x, t)$  will remain smooth. We rewrite the IVP as

$$\left. \begin{aligned} u_t + \lambda(u)u_x &= 0, \quad \lambda(u) = f'(u), \\ u(x, 0) &= u_0(x). \end{aligned} \right\} \quad (2.86)$$

Note that the PDE in (2.86) is a non-linear extension of the linear advection equation in (2.32) in which the characteristic speed is  $\lambda(u) = a = \text{constant}$ . We construct solutions to IVP (2.86) following characteristic curves, in much the same way as performed for the linear advection equation.

### Construction of Solutions on Characteristics

Consider characteristic curves  $x = x(t)$  satisfying the IVP

$$\frac{dx}{dt} = \lambda(u), \quad x(0) = x_0. \quad (2.87)$$

Then, by regarding both  $u$  and  $x$  to be functions of  $t$  we find the total derivative of  $u$  along the curve  $x(t)$ , namely

$$\frac{du}{dt} = u_t + \lambda(u)u_x = 0. \quad (2.88)$$

That is,  $u$  is constant along the characteristic curve satisfying the IVP (2.87) and therefore the slope  $\lambda(u)$  is constant along the characteristic. Hence the characteristic curves are straight lines. The value of  $u$  along each curve is the value of  $u$  at the initial point  $x(0) = x_0$  and we write

$$u(x, t) = u_0(x_0). \quad (2.89)$$

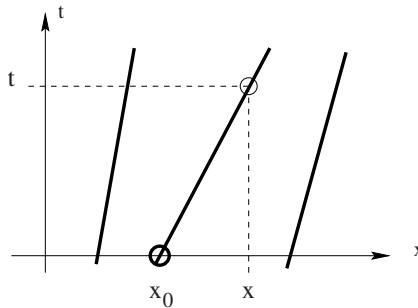
Fig. 2.11 shows a typical characteristic curve emanating from the initial point  $x_0$  on the  $x$ -axis. The slope  $\lambda(u)$  of the characteristic may then be evaluated at  $x_0$  so that the solution characteristics curves of IVP (2.87) are

$$x = x_0 + \lambda(u_0(x_0))t . \quad (2.90)$$

Relations (2.89) and (2.90) may be regarded as the analytical solution of IVP (2.86). Note that the point  $x_0$  depends on the given point  $(x, t)$ , see Fig. 2.11, and thus  $x_0 = x_0(x, t)$ . The solution given by (2.89) and (2.90) is implicit, which is more apparent if we substitute  $x_0$  from (2.90) into (2.89) to obtain

$$u(x, t) = u_0(x - \lambda(u_0(x_0))t) . \quad (2.91)$$

Note that this solution is identical in form to the solution (2.39) of the linear advection equation in (2.32).



**Fig. 2.11.** Typical characteristic curves for a non-linear hyperbolic conservation law

Next we verify that relations (2.89) and (2.90) actually define the solution. From (2.89) we obtain the  $t$  and  $x$  derivatives

$$u_t = u'_0(x_0) \frac{\partial x_0}{\partial t} , \quad u_x = u'_0(x_0) \frac{\partial x_0}{\partial x} . \quad (2.92)$$

From (2.90) the  $t$  and  $x$  derivatives are found to be

$$\left. \begin{aligned} \lambda(u_0(x_0)) + [1 + \lambda'(u_0(x_0))u'_0(x_0)t] \frac{\partial x_0}{\partial t} &= 0 , \\ [1 + \lambda'(u_0(x_0))u'_0(x_0)t] \frac{\partial x_0}{\partial x} &= 1 . \end{aligned} \right\} \quad (2.93)$$

From (2.93) we obtain

$$\frac{\partial x_0}{\partial t} = -\frac{\lambda(u_0(x_0))}{1 + \lambda'(u_0(x_0))u'_0(x_0)t} \quad (2.94)$$

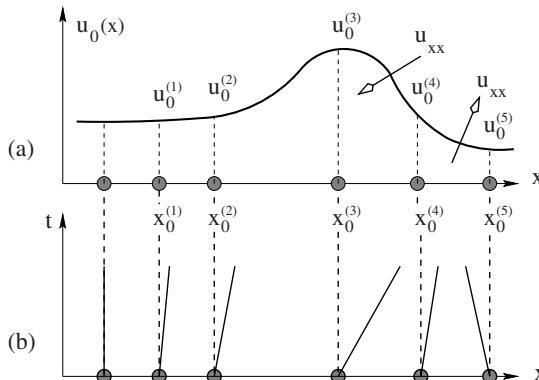
and

$$\frac{\partial x_0}{\partial x} = \frac{1}{1 + \lambda'(u_0(x_0))u'_0(x_0)t}. \quad (2.95)$$

Substitution of (2.94)–(2.95) into (2.92) verifies that  $u_t$  and  $u_x$  satisfy the PDE in (2.86).

### Wave Steepening

Recall that in the case of the linear advection equation, in which the characteristic speed is  $\lambda(u) = a = \text{constant}$ , the solution consists of the initial data  $u_0(x)$  translated with speed  $a$  *without distortion*. In the non-linear case the characteristic speed  $\lambda(u)$  is a function of the solution itself. Distortions are therefore produced; this is a distinguishing feature of non-linear problems.



**Fig. 2.12.** Wave steepening in a convex, non-linear hyperbolic conservation law: (a) initial condition and (b) corresponding picture of characteristics

To explain the wave distortion phenomenon we consider initial data  $u_0(x)$  as shown in Fig. 2.12. A smooth initial profile is shown in Fig. 2.12a along with five initial points  $x_0^{(i)}$  and their corresponding initial data values  $u_0^{(i)} = u_0(x_0^{(i)})$ . For the moment let us assume that the flux function  $f(u)$  is convex, that is  $\lambda'(u) = f''(u) > 0$ . In this case the characteristic speed is an increasing function of  $u$ . Fig. 2.12b shows the characteristics  $x^{(i)}(t)$  emanating from the initial points  $x_0^{(i)}$  and carrying the constant initial values  $u_0^{(i)}$  along them. Given the assumed convex character of the flux, higher values of  $u_0(x)$  will travel faster than lower values of  $u_0(x)$ . There are two intervals on the  $x$ -axis where distortions are most evident. These are the intervals  $I_E = [x_0^{(1)}, x_0^{(3)}]$  and  $I_C = [x_0^{(3)}, x_0^{(5)}]$ . In  $I_E$  the value  $u_0^{(3)}$  will propagate faster than  $u_0^{(2)}$  and this in turn will propagate faster than  $u_0^{(1)}$ . The orientation of the respective characteristics in Fig. 2.12b makes this situation clear. At a later time the initial data in  $I_E$  will have been transformed into a broader

and flatter profile. We say that  $I_E$  is an *expansive* region. In the expansive region the characteristic speed increases as  $x$  increases, that is  $\lambda_x > 0$ . By contrast the interval  $I_C$  is *compressive* and  $\lambda_x < 0$ ; the value  $u_0^{(3)}$  will propagate faster than  $u_0^{(4)}$  and this in turn will propagate faster than  $u_0^{(5)}$ , as shown by the orientation of the respective characteristics in Fig. 2.12b. The compressive region will tend to get steeper and narrower as time evolves. The wave steepening mechanism will eventually produce folding over of the solution profile, with corresponding crossing of characteristics, and triple-valued solutions. Note that the compressive and expansive character of the data just described reverses for the case of a concave flux,  $\lambda'(u) = f''(u) < 0$ . Before crossing of characteristics the single-valued solution may be found following characteristics, as described previously. When characteristics first intersect we say that the wave breaks; the derivative  $u_x$  becomes infinite and this happens at a precise breaking time  $t_b$  given by

$$t_b = \frac{-1}{\lambda_x(x_0)}. \quad (2.96)$$

This is confirmed by equations (2.94)–(2.95). Breaking first occurs on the characteristic emanating from  $x = x_0$  for which  $\lambda_x(x_0)$  is negative and  $|\lambda_x(x_0)|$  is a maximum. For details see Whitham [582].

This is an anomalous situation that may be rescued by going back to the physical origins of the equations and questioning the adequacy of the model furnished by (2.81). The improved model equation (2.83) says that the time rate of change of  $u$  is not just due to the advection term  $f(u)_x$  but is a competing balance between advection and the diffusion term  $\alpha u_{xx}$ . As shown in Fig. 2.12a in the interval  $[x_0^{(3)}, x_0^{(4)}]$  the *wave steepening* effect of  $f(u)_x$  is opposed by the *wave-easing* effect of  $\alpha u_{xx}$ , which is negative there. In the interval  $[x_0^{(4)}, x_0^{(5)}]$  the role of these contradictory effects is reversed. The more complete description of the physics does not allow folding over of the solution. But rather than working with the more complete, and therefore more complex, viscous description of the problem, it is actually possible to insist on using the inviscid model (2.81) by allowing discontinuities to be formed as a *process of increasing compression*, namely shock waves. Further details are found in Lax [301], Whitham [582] and Smoller [451].

## Shock Waves

Shock waves in air are small transition layers of very rapid changes of physical quantities such as pressure, density and temperature. The transition layer for a strong shock is of the same order of magnitude as the mean-free path of the molecules, that is about  $10^{-7}$  m. Therefore replacing these waves as mathematical discontinuities is a reasonable approximation. Very weak shock waves such as sonic booms, are an exception; the discontinuous approximation

here can be very inaccurate indeed, see Whitham [582]. For a discussion on shock thickness see Landau and Lifshitz [297], pp. 337–341.

We therefore insist on using the simplified model (2.81) but in its integral form, e.g. (2.82). Consider a solution  $u(x, t)$  such that  $u(x, t)$ ,  $f(u)$  and their derivatives are continuous everywhere except on a line  $s = s(t)$  on the  $x$ – $t$  plane across which  $u(x, t)$  has a jump discontinuity. Select two fixed points  $x_L$  and  $x_R$  on the  $x$ –axis such that  $x_L < s(t) < x_R$ . Enforcing the conservation law in integral form (2.82) on the control volume  $[x_L, x_R]$  leads to

$$f(u(x_L, t)) - f(u(x_R, t)) = \frac{d}{dt} \int_{x_L}^{s(t)} u(x, t) dx + \frac{d}{dt} \int_{s(t)}^{x_R} u(x, t) dx .$$

Direct use of formula (1.68) of Chap. 1 yields

$$\left. \begin{aligned} f(u(x_L, t)) - f(u(x_R, t)) &= [u(s_L, t) - u(s_R, t)] S \\ &\quad + \int_{x_L}^{s(t)} u_t(x, t) dx + \int_{s(t)}^{x_R} u_t(x, t) dx , \end{aligned} \right\}$$

where  $u(s_L, t)$  is the limit of  $u(s(t), t)$  as  $x$  tends to  $s(t)$  from the left,  $u(s_R, t)$  is the limit of  $u(s(t), t)$  as  $x$  tends to  $s(t)$  from the right and  $S = ds/dt$  is the speed of the discontinuity. As  $u_t(x, t)$  is bounded the integrals vanish identically as  $s(t)$  is approached from left and right and we obtain

$$f(u(x_L, t)) - f(u(x_R, t)) = [u(s_L, t) - u(s_R, t)] S . \quad (2.97)$$

This algebraic expression relating the jumps  $\Delta f = f(u(x_R, t)) - f(u(x_L, t))$ ,  $\Delta u = u(x_R, t) - u(x_L, t)$  and the speed  $S$  of the discontinuity is called the *Rankine–Hugoniot condition* and is usually expressed as

$$\Delta f = S \Delta u . \quad (2.98)$$

For the scalar case considered here one can solve for the speed  $S$  as

$$S = \frac{\Delta f}{\Delta u} . \quad (2.99)$$

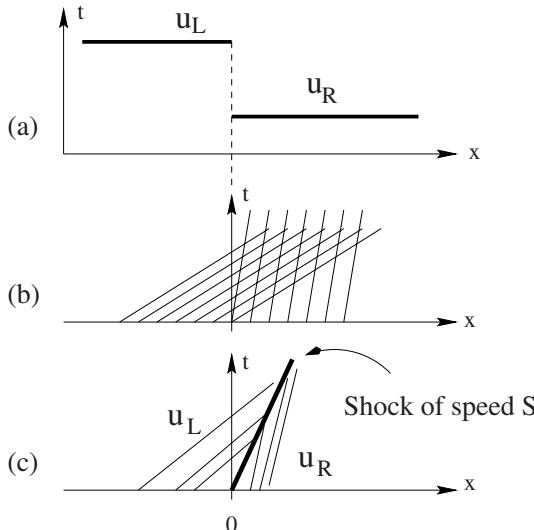
Therefore, in order to admit discontinuous solutions we may formulate the problem in terms of PDEs, which are valid in smooth parts of the solution, and the Rankine–Hugoniot conditions across discontinuities.

## Two Examples of Discontinuous Solutions

Consider the following initial–value problem for the inviscid Burgers equation

$$\left. \begin{aligned} u_t + f(u)_x &= 0 , \quad f(u) = \frac{1}{2} u^2 , \\ u(x, 0) &= u_0(x) = \begin{cases} u_L & \text{if } x < 0 , \\ u_R & \text{if } x > 0 . \end{cases} \end{aligned} \right\} \quad (2.100)$$

First assume that  $u_L > u_R$ . As the flux is convex  $\lambda'(u) = f''(u) > 0$  the characteristic speeds on the left are greater than those on the right, that is  $\lambda_L \equiv \lambda(u_L) > \lambda_R \equiv \lambda(u_R)$ . Based on the discussion about Fig. 2.12 the initial data in IVP (2.100) is the extreme case of compressive data. Crossing of characteristics takes place immediately, as illustrated in Fig. 2.13b. The



**Fig. 2.13.** (a) Compressive discontinuous initial data (b) picture of characteristics and (c) solution on  $x$ - $t$  plane

discontinuous solution of the IVP is

$$u(x, t) = \begin{cases} u_L & \text{if } x - St < 0 , \\ u_R & \text{if } x - St > 0 , \end{cases} \quad (2.101)$$

where the speed of the discontinuity is found from (2.99) as

$$S = \frac{1}{2}(u_L + u_R) . \quad (2.102)$$

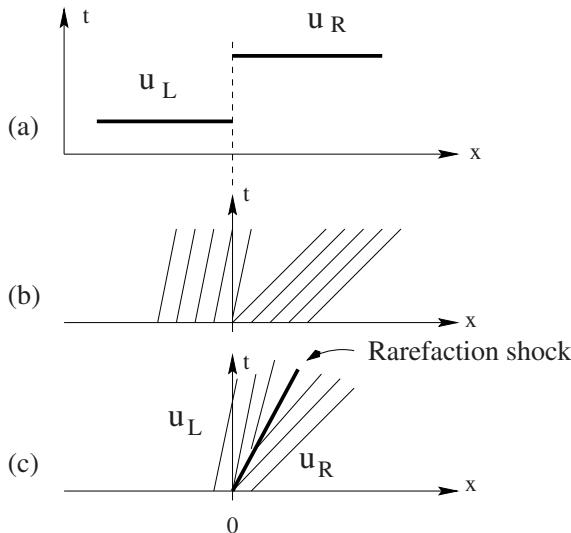
This discontinuous solution is a shock wave and is compressive in nature as discussed previously and as observed in Fig. 2.13a; it satisfies the following condition

$$\lambda(u_L) > S > \lambda(u_R) , \quad (2.103)$$

which is called the *entropy condition*. More details are found in Chorin and Marsden [112], LeVeque [308], Smoller [451], Whitham [582].

Now we assume that  $u_L < u_R$  in the IVP (2.100). This data is the extreme case of expansive data, for convex  $f(u)$ . A possible mathematical solution has identical form as solution (2.101)–(2.102) for the compressive data case. See

Fig. 2.14. However, this solution is physically incorrect. The discontinuity has not arisen as the result of compression,  $\lambda_L < \lambda_R$ ; the characteristics diverge from the discontinuity. This solution is called a *rarefaction shock*, or *entropy-violating shock*, and does not satisfy the entropy condition (2.103); it is therefore rejected as a physical solution. Compare Figs. 2.13 and 2.14; in the compressive case characteristics run into the discontinuity. Given the



**Fig. 2.14.** (a) Expansive discontinuous initial data (b) picture of characteristics and (c) rarefaction shock solution on  $x$ - $t$  plane

*expansive* character of the data and based on the discussion on Fig. 2.12, it would be more reasonable to expect the initial data to break up immediately and to broaden with time. This actually gives another solution to be discussed next.

### Rarefaction Waves

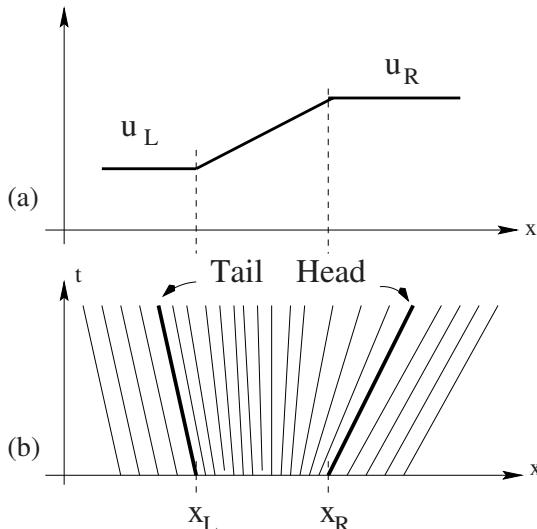
Reconsider the IVP (2.100) with general convex flux function  $f(u)$

$$\left. \begin{aligned} u_t + f(u)_x &= 0 , \\ u(x, 0) &= u_0(x) = \begin{cases} u_L & \text{if } x < 0 , \\ u_R & \text{if } x > 0 , \end{cases} \end{aligned} \right\} \quad (2.104)$$

and *expansive* initial data,  $u_L < u_R$ . As discussed previously, the entropy-violating solution to this problem is

$$u(x, t) = \begin{cases} u_L & \text{if } x - St < 0, \\ u_R & \text{if } x - St > 0, \end{cases} \quad S = \frac{\Delta f}{\Delta u}. \quad (2.105)$$

Amongst the various other reasons for rejecting this solution as a physical solution, instability stands out as a prominent argument. By instability it is meant that *small perturbations of the initial data lead to large changes in the solution*. As a matter of fact, under small perturbations, the whole character of the solution changes completely, as we shall see.

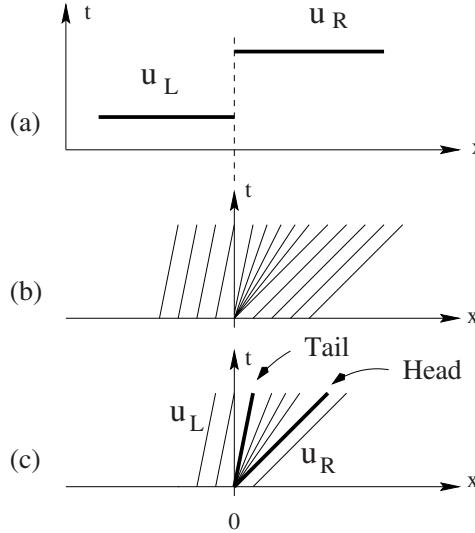


**Fig. 2.15.** Non-centred rarefaction wave: (a) expansive smooth initial data, (b) picture of characteristics on  $x-t$  plane

Let us modify the initial data in (2.104) by replacing the discontinuous change from  $u_L$  to  $u_R$  by a linear variation of  $u_0(x)$  between two fixed points  $x_L < 0$  and  $x_R > 0$ . Now the initial data reads

$$u_0(x) = \begin{cases} u_L & \text{if } x \leq x_L, \\ u_L + \frac{u_R - u_L}{x_R - x_L}(x - x_L) & \text{if } x_L < x < x_R, \\ u_R & \text{if } x \geq x_R, \end{cases} \quad (2.106)$$

and is illustrated in Fig. 2.15a. The corresponding picture of characteristics emanating from the initial time  $t = 0$  is shown in Fig. 2.15b. The solution  $u(x, t)$  to this problem is found by following characteristics, as discussed previously, and consists of two constant states,  $u_L$  and  $u_R$ , separated by a region of *smooth transition* between the data values  $u_L$  and  $u_R$ . This is called a *rarefaction wave*. The right edge of the wave is given by the characteristic emanating



**Fig. 2.16.** Centred rarefaction wave: (a) expansive discontinuous initial data (b) picture of characteristics (c) entropy satisfying (rarefaction) solution on  $x$ - $t$  plane

from  $x_R$

$$x = x_R + \lambda(u_R)t \quad (2.107)$$

and is called the *Head* of the rarefaction. It carries the value  $u_0(x_R) = u_R$ . The left edge of the wave is given by the characteristic emanating from  $x_L$

$$x = x_L + \lambda(u_L)t \quad (2.108)$$

and is called the *Tail* of the rarefaction. It carries the value  $u_0(x_L) = u_L$ .

As we assume convexity,  $\lambda'(u) = f''(u) > 0$ , larger values of  $u_0(x)$  propagate faster than lower values and thus the wave spreads and flattens as time evolves. The spreading of waves is a typical non-linear phenomenon not seen in the study of linear hyperbolic systems with constant coefficients. The entire solution is

$$\left. \begin{array}{ll} u(x, t) = u_L & \text{if } \frac{x-x_L}{t} \leq \lambda_L, \\ \lambda(u) = \frac{x-x_L}{t} & \text{if } \lambda_L < \frac{x-x_L}{t} < \lambda_R, \\ u(x, t) = u_R & \text{if } \frac{x-x_R}{t} \geq \lambda_R. \end{array} \right\} \quad (2.109)$$

No matter how small the size  $\Delta x = x_R - x_L$  of the interval over which the discontinuous data in IVP (2.104) has been spread over, the structure of the above rarefaction solution remains unaltered and is entirely different from the rarefaction shock solution (2.105), for which small changes to the data lead to large changes in the solution. Thus the rarefaction shock solution is unstable. From the above construction the rarefaction solution is stable and as  $x_L$  and  $x_R$  approach zero from below and above respectively, the discontinuous data at  $x = 0$  in IVP (2.104) is reproduced. Therefore, the limiting case is to be

interpreted as follows:  $u_0(x)$  takes on all the values between  $u_L$  and  $u_R$  at  $x = 0$  and consequently  $\lambda(u_0(x))$  takes on all the values between  $\lambda_L$  and  $\lambda_R$  at  $x = 0$ . As higher values propagate faster than lower values the initial data disintegrates immediately giving rise to a rarefaction solution. This limiting rarefaction in which all characteristics of the wave emanate from a single point is called a *centred rarefaction wave*. The solution is

$$\left. \begin{array}{ll} u(x, t) = u_L & \text{if } \frac{x}{t} \leq \lambda_L, \\ \lambda(u) = \frac{x}{t} & \text{if } \lambda_L < \frac{x}{t} < \lambda_R, \\ u(x, t) = u_R & \text{if } \frac{x}{t} \geq \lambda_R, \end{array} \right\} \quad (2.110)$$

and is illustrated in Fig. 2.16.

Now we have at least two solutions to the IVP (2.104). Thus, having extended the concept of solution to include discontinuities, extra spurious solutions are now part of this extended class. The question is how to distinguish between a physically correct solution and a spurious solution. The anticipated answer is that a physical discontinuity, in addition to the Rankine–Hugoniot condition (2.98), also satisfies the entropy condition (2.103).

### The Riemann Problem for the Inviscid Burgers Equation

We finalise this section by giving the solution of the Riemann problem for the inviscid Burgers equation, namely

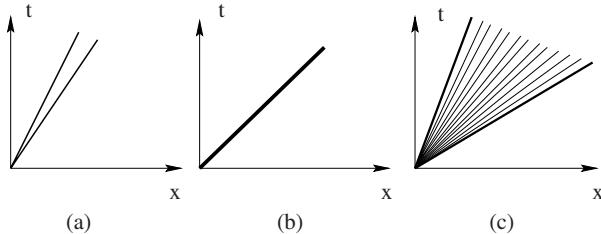
$$\left. \begin{array}{l} \text{PDE : } u_t + (\frac{u^2}{2})_x = 0, \\ \text{IC : } u(x, 0) = \begin{cases} u_L, & x < 0, \\ u_R, & x > 0. \end{cases} \end{array} \right\} \quad (2.111)$$

From the previous discussion the exact solution is a *single wave* emanating from the origin as shown in Fig. 2.17a. In view of the entropy condition this wave is either a shock wave, when  $u_L > u_R$ , or a rarefaction wave, when  $u_L \leq u_R$ . The complete solution is

$$\left. \begin{array}{l} u(x, t) = \begin{cases} u_L & \text{if } x - St < 0 \\ u_R & \text{if } x - St > 0 \end{cases} \end{array} \right\} \text{ if } u_L > u_R, \\ \left. \begin{array}{l} S = \frac{1}{2}(u_L + u_R) \\ u(x, t) = \begin{cases} u_L & \text{if } \frac{x}{t} \leq u_L \\ \frac{x}{t} & \text{if } u_L < x/t < u_R \\ u_R & \text{if } x/t \geq u_R \end{cases} \end{cases} \right\} \text{ if } u_L \leq u_R. \quad (2.112)$$

Fig. 2.17 shows the solution of the Riemann problem for the inviscid Burgers equation. Fig. 2.17a depicts the structure of the general solution and consists of a single wave, Fig. 2.17b shows the case in which the solution is a shock wave and Fig. 2.17c shows the case in which it is a rarefaction wave.

Some of the studied notions for scalar conservations laws extend quite directly to systems of hyperbolic conservations laws, as we see in the next section.



**Fig. 2.17.** Solution of the Riemann problem for the inviscid Burgers equation: (a) structure of general solution (single wave, shock or rarefaction), (b) solution is a shock wave and (c) solution is a rarefaction wave

### 2.4.3 Characteristic Fields

Consider a hyperbolic system of conservation laws of the form (2.65) with real eigenvalues  $\lambda_i(\mathbf{U})$  and corresponding right eigenvectors  $\mathbf{K}^{(i)}(\mathbf{U})$ . The characteristic speed  $\lambda_i(\mathbf{U})$  defines a *characteristic field*, the  $\lambda_i$ -field. Sometimes one also speaks of the  $\mathbf{K}^{(i)}$ -field, that is the characteristic field defined by the eigenvector  $\mathbf{K}^{(i)}$ .

**Definition 2.25 (Linearly degenerate fields).** A  $\lambda_i$ -characteristic field is said to be linearly degenerate if

$$\nabla \lambda_i(\mathbf{U}) \cdot \mathbf{K}^{(i)}(\mathbf{U}) = 0, \quad \forall \mathbf{U} \in \Re^m, \quad (2.113)$$

where  $\Re^m$  is the set of real-valued vectors of  $m$  components.

**Definition 2.26 (Genuinely nonlinear fields).** A  $\lambda_i$ -characteristic field is said to be genuinely nonlinear if

$$\nabla \lambda_i(\mathbf{U}) \cdot \mathbf{K}^{(i)}(\mathbf{U}) \neq 0, \quad \forall \mathbf{U} \in \Re^m. \quad (2.114)$$

The symbol ‘.’ denotes the dot product in *phase space*.  $\nabla \lambda_i(\mathbf{U})$  is the gradient of the eigenvalue  $\lambda_i(\mathbf{U})$ , namely

$$\nabla \lambda_i(\mathbf{U}) = \left( \frac{\partial}{\partial u_1} \lambda_i, \frac{\partial}{\partial u_2} \lambda_i, \dots, \frac{\partial}{\partial u_m} \lambda_i \right)^T.$$

The *phase space* is the space of vectors  $\mathbf{U} = (u_1, \dots, u_m)$ ; for a  $2 \times 2$  system we speak of the *phase plane*  $u_1 - u_2$ . Note that for a linear system (2.42) the eigenvalues  $\lambda_i$  are constant and therefore  $\nabla \lambda_i(\mathbf{U}) = \mathbf{0}$ . Hence all characteristic fields of a linear hyperbolic system with constant coefficients are linearly degenerate.

**Exercise 2.27.** Show that both characteristic fields of the isothermal equations of Gas Dynamics (2.74) are genuinely non-linear.

**Solution 2.28.** First we write the eigenvalues (2.75) in terms of the conserved variables, namely

$$\lambda_1 = \frac{u_2}{u_1} - a, \quad \lambda_2 = \frac{u_2}{u_1} + a,$$

$$\nabla \lambda_1(\mathbf{U}) = \left( -\frac{u}{\rho}, \frac{1}{\rho} \right)^T, \quad \nabla \lambda_2(\mathbf{U}) = \left( -\frac{u}{\rho}, \frac{1}{\rho} \right)^T.$$

Therefore

$$\nabla \lambda_1(\mathbf{U}) \cdot \mathbf{K}^{(1)}(\mathbf{U}) = -\frac{a}{\rho} \neq 0,$$

$$\nabla \lambda_2(\mathbf{U}) \cdot \mathbf{K}^{(2)}(\mathbf{U}) = \frac{a}{\rho} \neq 0$$

and thus both characteristic fields are *genuinely non-linear*, as claimed.

*Example 2.29 (Detonation Analogue).* In the study of detonation waves in high energy solids it is found useful to devise mathematical objects that preserve some of the basic physical features of detonation phenomena but are simpler to analyse than more comprehensive models. Fickett [191] proposed a system that is essentially the inviscid Burgers equation plus a reaction model. He called the system *detonation analogue*. Clarke and colleagues [118] pointed out that this analogue is also exceedingly useful for numerical purposes. Writing the system in *conservation-law form* one has the *inhomogeneous system* with a source term, namely

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{S}(\mathbf{U}), \quad (2.115)$$

$$\mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \equiv \begin{bmatrix} \rho \\ \alpha \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \frac{1}{2}(\rho^2 + \alpha Q) \\ 0 \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} 0 \\ 2\sqrt{1-\alpha} \end{bmatrix}. \quad (2.116)$$

The parameter  $Q$  plays the role of *heats of reaction* and  $\alpha$  is a *reaction progress variable*. The mathematical character of the system is determined solely by the homogeneous part,  $\mathbf{S} = \mathbf{0}$ . The Jacobian matrix is

$$\mathbf{A}(\mathbf{U}) = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} = \begin{bmatrix} u_1 & \frac{1}{2}Q \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \rho & \frac{1}{2}Q \\ 0 & 0 \end{bmatrix}.$$

Simple calculations show that the eigenvalues are

$$\lambda_1 = 0, \quad \lambda_2 = \rho \quad (2.117)$$

and the right eigenvectors are

$$\mathbf{K}^{(1)} = \begin{bmatrix} 1 \\ -2\rho/Q \end{bmatrix}, \quad \mathbf{K}^{(2)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \quad (2.118)$$

The detonation analogue is therefore hyperbolic.

**Exercise 2.30.** Check that the  $\lambda_1$ -field is linearly degenerate and that the  $\lambda_2$ -field is genuinely non-linear.

**Solution 2.31.** (Left to the reader).

## Rankine–Hugoniot Conditions

Given a system of hyperbolic conservation laws

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0} \quad (2.119)$$

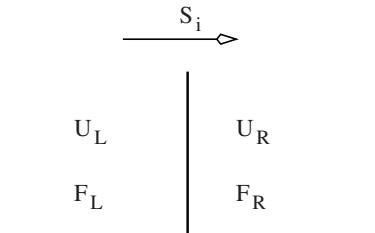
and a discontinuous wave solution of speed  $S_i$  associated with the  $\lambda_i$ -characteristic field, the Rankine–Hugoniot conditions state

$$\Delta \mathbf{F} = S_i \Delta \mathbf{U}, \quad (2.120)$$

with

$$\Delta \mathbf{U} \equiv \mathbf{U}_R - \mathbf{U}_L, \quad \Delta \mathbf{F} \equiv \mathbf{F}_R - \mathbf{F}_L, \quad \mathbf{F}_L = \mathbf{F}(\mathbf{U}_L), \quad \mathbf{F}_R = \mathbf{F}(\mathbf{U}_R),$$

where  $\mathbf{U}_L$  and  $\mathbf{U}_R$  are the respective states immediately to the left and right of the discontinuity. Fig. 2.18 illustrates the Rankine–Hugoniot conditions. Note that unlike the scalar case, see (2.99), it is generally not possible to



**Fig. 2.18.** Illustration of the Rankine–Hugoniot conditions for a single discontinuity of speed  $S_i$  connecting two constant states  $\mathbf{U}_L$  and  $\mathbf{U}_R$  via a system of conservation laws

solve for the speed  $S_i$ . For a linear system with constant coefficients

$$\mathbf{U}_t + \mathbf{A}\mathbf{U}_x = \mathbf{0},$$

with eigenvalues  $\lambda_i$ , for  $i = 1, \dots, m$ , the Rankine–Hugoniot conditions across the wave of speed  $S_i \equiv \lambda_i$  read

$$\Delta \mathbf{F} = \mathbf{A} \Delta \mathbf{U} = \lambda_i (\Delta \mathbf{U})_i. \quad (2.121)$$

See (2.63). Actually, these conditions provide a technique for finding the solution of the Riemann problem for linear hyperbolic system with constant coefficients.

**Exercise 2.32.** Solve the Riemann problem for the linearised equations of Gas Dynamics using the Rankine–Hugoniot conditions across each wave.

**Solution 2.33.** The structure of the solution is depicted in Fig. 2.5. The unknowns are  $\rho_*$  and  $u_*$  in the *Star Region*. Recall that the vector  $\mathbf{U}$  and the coefficient matrix  $\mathbf{A}$  are given by

$$\mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \equiv \begin{bmatrix} \rho \\ u \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & \rho_0 \\ a^2/\rho_0 & 0 \end{bmatrix}$$

and the eigenvalues are

$$\lambda_1 = -a, \quad \lambda_2 = +a.$$

Application of the Rankine–Hugoniot conditions across the  $\lambda_1$ –wave of speed  $S_1 = \lambda_1$  gives

$$\begin{bmatrix} 0 & \rho_0 \\ \frac{a^2}{\rho_0} & 0 \end{bmatrix} \begin{bmatrix} \rho_* - \rho_L \\ u_* - u_L \end{bmatrix} = -a \begin{bmatrix} \rho_* - \rho_L \\ u_* - u_L \end{bmatrix}.$$

Expanding and solving for  $u_*$  gives

$$u_* = u_L - (\rho_* - \rho_L) \frac{a}{\rho_0}.$$

For the  $\lambda_2$ –wave of speed  $S_2 = \lambda_2$  we obtain

$$u_* = u_R + (\rho_* - \rho_R) \frac{a}{\rho_0}.$$

The simultaneous solution of these two linear algebraic equations for the unknowns  $\rho_*$  and  $u_*$  is

$$\rho_* = \frac{1}{2}(\rho_L + \rho_R) - \frac{1}{2}(u_R - u_L)\rho_0/a,$$

$$u_* = \frac{1}{2}(u_L + u_R) - \frac{1}{2}(\rho_R - \rho_L)a/\rho_0,$$

which is the solution (2.64) obtained using a different technique based on eigenvector expansion of the initial data. The technique that makes use of the Rankine–Hugoniot conditions is more direct.

## Generalised Riemann Invariants

For a general quasi-linear hyperbolic system

$$\mathbf{W}_t + \mathbf{A}(\mathbf{W})\mathbf{W}_x = \mathbf{0}, \tag{2.122}$$

with

$$\mathbf{W} = [w_1, w_2, \dots, w_m]^T,$$

we consider the wave associated with the  $i$ –characteristic field with eigenvalue  $\lambda_i$  and corresponding right eigenvector

$$\mathbf{K}^{(i)} = \left[ k_1^{(i)}, k_2^{(i)}, \dots, k_m^{(i)} \right]^T.$$

The vector of dependent variables  $\mathbf{W}$  here is some suitable set, which may be the set conserved variables, for instance. Recall that any system of conservation laws may always be expressed in quasi-linear form via the Jacobian matrix, see (2.6) and (2.8).

The *Generalised Riemann Invariants* are relations that hold true, for certain waves, *across* the wave structure and lead the following  $(m - 1)$  ordinary differential equations

$$\frac{dw_1}{k_1^{(i)}} = \frac{dw_2}{k_2^{(i)}} = \frac{dw_3}{k_3^{(i)}} = \dots = \frac{dw_m}{k_m^{(i)}}. \quad (2.123)$$

They relate ratios of changes  $dw_s$  of a quantity  $w_s$  to the respective component  $k_s^{(i)}$  of the right eigenvector  $\mathbf{K}^{(i)}$  corresponding to a  $\lambda_i$ -wave family. For a detailed discussion see the book by Jeffrey [269].

*Example 2.34 (Linearised Gas Dynamics revisited).* Here we find the Generalised Riemann Invariants for the linearised equations of Gas Dynamics. The dependent variables are

$$\mathbf{W} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \equiv \begin{bmatrix} \rho \\ u \end{bmatrix}$$

and the right eigenvectors are

$$\mathbf{K}^{(1)} = \begin{bmatrix} \rho_0 \\ -a \end{bmatrix}, \quad \mathbf{K}^{(2)} = \begin{bmatrix} \rho_0 \\ a \end{bmatrix}.$$

Across the  $\lambda_1$ -wave we have

$$\frac{d\rho}{\rho_0} = \frac{du}{-a},$$

which leads to

$$du + \frac{a}{\rho_0} d\rho = 0.$$

After integration this produces

$$I_L(\rho, u) = u + \frac{a}{\rho_0} \rho = \text{constant}. \quad (2.124)$$

The constant of integration is obtained by evaluating  $I_L(\rho, u)$  at a reference state. Across the  $\lambda_2$ -wave we have

$$\frac{d\rho}{\rho_0} = \frac{du}{a},$$

which leads to

$$I_R(\rho, u) = u - \frac{a}{\rho_0} \rho = \text{constant}. \quad (2.125)$$

Again the constant of integration is obtained by evaluating  $I_R(\rho, u)$  at a reference state.

**Exercise 2.35.** Solve the Riemann problem for the linearised equations of Gas Dynamics using the Generalised Riemann Invariants.

**Solution 2.36.** Application of  $I_L(\rho, u)$  across the left wave connecting the states  $\mathbf{W}_L$  and  $\mathbf{W}_*$  gives

$$u_* + \frac{a}{\rho_0} \rho_* = u_L + \frac{a}{\rho_0} \rho_L .$$

Similarly, application of  $I_R(\rho, u)$  across the right wave connecting the states  $\mathbf{W}_R$  and  $\mathbf{W}_*$  gives

$$u_* - \frac{a}{\rho_0} \rho_* = u_R - \frac{a}{\rho_0} \rho_R$$

and the simultaneous solution for the unknowns  $\rho_*$  and  $u_*$  gives

$$\left. \begin{aligned} \rho_* &= \frac{1}{2}(\rho_L + \rho_R) - \frac{1}{2}(u_R - u_L)\rho_0/a , \\ u_* &= \frac{1}{2}(u_L + u_R) - \frac{1}{2}(\rho_R - \rho_L)a/\rho_0 , \end{aligned} \right\}$$

which is the same solution (2.64) obtained from applying other techniques.

**Exercise 2.37.** Solve the Riemann problem for the Small Perturbation Equations (2.24) using the following techniques:

- by expanding the initial data  $\mathbf{U}_L$  and  $\mathbf{U}_R$  in terms of the eigenvectors, see (2.56).
- by expanding the total jump  $\Delta\mathbf{U}$  in terms of the eigenvectors, see (2.62).
- by using the Rankine-Hugoniot Conditions across each wave, see (2.121).
- by applying the Generalised Riemann Invariants, see (2.123).

**Solution 2.38.** Use of any of the suggested techniques will give the general solution

$$\left. \begin{aligned} u_* &= \frac{1}{2}(u_L + u_R) + \frac{1}{2}a(v_R - v_L) , \\ v_* &= \frac{1}{2}(v_L + v_R) + \frac{1}{2a}(u_R - u_L) . \end{aligned} \right\}$$

*Example 2.39 (Isentropic Gas Dynamics Revisited).* For this example the eigenvalues are  $\lambda_1 = u - a$  and  $\lambda_2 = u + a$ , with  $a = \sqrt{p'(\rho)} = \sqrt{\frac{\gamma p}{\rho}}$  defining the sound speed. The corresponding right eigenvectors are given by

$$\mathbf{K}^{(1)} = \begin{bmatrix} 1 \\ u - a \end{bmatrix} , \quad \mathbf{K}^{(2)} = \begin{bmatrix} 1 \\ u + a \end{bmatrix} .$$

Across the left  $\lambda_1 = u - a$  wave we have

$$\frac{d\rho}{1} = \frac{d(\rho u)}{u - a} ,$$

which after expanding differentials yields

$$du + \frac{a}{\rho} d\rho = 0 .$$

On exact integration we obtain

$$I_L(\rho, \rho u) = u + \int \frac{a}{\rho} d\rho = \text{constant} . \quad (2.126)$$

Across the right  $\lambda_2 = u + a$  wave we obtain

$$I_R(\rho, \rho u) = u - \int \frac{a}{\rho} d\rho = \text{constant} . \quad (2.127)$$

As the sound speed  $a$  is a function of  $\rho$  alone we can evaluate the integral term above exactly as

$$\int \frac{a}{\rho} d\rho = \frac{2a}{\gamma - 1}$$

by first noting that

$$a = \sqrt{p'(\rho)} = \sqrt{C\gamma\rho^{\gamma-1}} = \sqrt{C\gamma}\rho^{\frac{\gamma-1}{2}} .$$

Then the *left* and *right* Riemann Invariants become

$$\left. \begin{aligned} I_L(\rho, \rho u) &= u + \frac{2a}{\gamma-1} = \text{constant across the } \lambda_1\text{-wave,} \\ I_R(\rho, \rho u) &= u - \frac{2a}{\gamma-1} = \text{constant across the } \lambda_2\text{-wave.} \end{aligned} \right\} \quad (2.128)$$

Generalised Riemann Invariants provide a powerful tool of analysis of hyperbolic conservation laws.

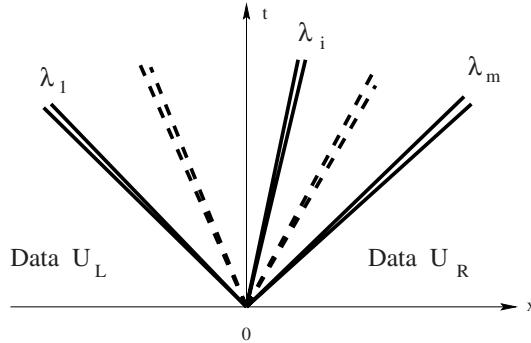
#### 2.4.4 Elementary–Wave Solutions of the Riemann Problem

The Riemann problem for a general  $m \times m$  non-linear hyperbolic system with data  $\mathbf{U}_L, \mathbf{U}_R$  is the IVP

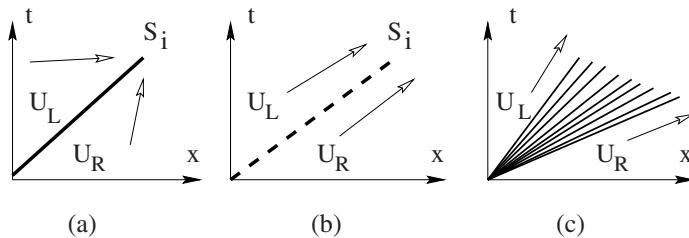
$$\left. \begin{aligned} \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x &= \mathbf{0}, \\ \mathbf{U}(x, 0) &= \mathbf{U}^{(0)}(x) = \begin{cases} \mathbf{U}_L & \text{if } x < 0, \\ \mathbf{U}_R & \text{if } x > 0. \end{cases} \end{aligned} \right\} \quad (2.129)$$

The similarity solution  $\mathbf{U}(x/t)$  of (2.129) consists of  $m + 1$  constant states separated by  $m$  waves, as depicted by the  $x-t$  picture of Fig. 2.19. For each eigenvalue  $\lambda_i$  there is a wave family. For linear systems with constant coefficients each wave is a discontinuity of speed  $S_i = \lambda_i$  and defines a linearly degenerate field.

For non-linear systems the waves may be discontinuities such as shock waves and contact waves, or smooth transition waves such as rarefactions. The possible types of waves present in the solution of the Riemann problem depends crucially on closure conditions. For the Euler equations we shall



**Fig. 2.19.** Structure of the solution of the Riemann problem for a system of non-linear conservation laws



**Fig. 2.20.** Elementary wave solutions of the Riemann problem: (a) shock wave of speed  $S_i$ , (b) contact discontinuity of speed  $S_i$  and (c) rarefaction wave

only consider Equations of State such that the only waves present are shocks, contacts and rarefactions. Suppose that the initial data states  $\mathbf{U}_L$ ,  $\mathbf{U}_R$  are connected by a single wave, that is, the solution of the Riemann problem consists of a single non-trivial wave; all other waves have zero strength. This assumption is entirely justified as we can always solve the Riemann problem with general data and then select the constant states on either side of a particular wave as the initial data for the Riemann problem. If the wave is a discontinuity then the wave is a shock wave or a contact wave.

### Shock Wave

For a shock wave the two constant states  $\mathbf{U}_L$  and  $\mathbf{U}_R$  are connected through a single jump discontinuity in a *genuinely non-linear field*  $i$  and the following conditions apply

- the Rankine–Hugoniot conditions

$$\mathbf{F}(\mathbf{U}_R) - \mathbf{F}(\mathbf{U}_L) = S_i(\mathbf{U}_R - \mathbf{U}_L). \quad (2.130)$$

- the entropy condition

$$\lambda_i(\mathbf{U}_L) > S_i > \lambda_i(\mathbf{U}_R). \quad (2.131)$$

Fig. 2.20a depicts a shock wave of speed  $S_i$ . The characteristic  $dx/dt = \lambda_i$  on both sides of the wave run into the shock wave, which illustrates the compressive character of a shock.

### Contact Wave

For a *contact wave* the two data states  $\mathbf{U}_L$  and  $\mathbf{U}_R$  are connected through a single jump discontinuity of speed  $S_i$  in a *linearly degenerate field*  $i$  and the following conditions apply

- the Rankine–Hugoniot conditions

$$\mathbf{F}(\mathbf{U}_R) - \mathbf{F}(\mathbf{U}_L) = S_i(\mathbf{U}_R - \mathbf{U}_L). \quad (2.132)$$

- constancy of the Generalised Riemann Invariants across the wave

$$\frac{dw_1}{k_1^{(i)}} = \frac{dw_2}{k_2^{(i)}} = \frac{dw_3}{k_3^{(i)}} = \cdots = \frac{dw_m}{k_m^{(i)}}. \quad (2.133)$$

- the parallel characteristic condition

$$\lambda_i(\mathbf{U}_L) = \lambda_i(\mathbf{U}_R) = S_i. \quad (2.134)$$

Fig. 2.20b depicts a contact discontinuity. Characteristics on both sides of the wave run parallel to it.

### Rarefaction Wave

For a rarefaction wave the two data states  $\mathbf{U}_L$  and  $\mathbf{U}_R$  are connected through a *smooth transition* in a *genuinely non-linear field*  $i$  and the following conditions are met

- constancy of the Generalised Riemann Invariants across the wave

$$\frac{dw_1}{k_1^{(i)}} = \frac{dw_2}{k_2^{(i)}} = \frac{dw_3}{k_3^{(i)}} = \cdots = \frac{dw_m}{k_m^{(i)}}. \quad (2.135)$$

- divergence of characteristics

$$\lambda_i(\mathbf{U}_L) < \lambda_i(\mathbf{U}_R). \quad (2.136)$$

Fig. 2.20c depicts a rarefaction wave. Characteristics on the left and right of the wave diverge as do characteristics inside the wave.

*Remark 2.40.* The solution of the general Riemann problem contains  $m$  waves of any of the above type, namely: shock waves, contact discontinuities and rarefaction waves. In solving the general Riemann problem we shall enforce these conditions by discriminating the particular type of wave present.

For further study we recommend the following references: Lax [301], Whitham [582], Chorin and Marsden [112], Courant and Friedrichs [143], Smoller [451] and LeVeque [308]. See also the papers by Lax [302] and [300].

The introductory concepts of this chapter will be used to analyse some of the properties of the Euler equations in Chap. 3. For the time-dependent one dimensional Euler equations we solve the Riemann problem exactly in Chap. 4, while in Chaps. 9 to 12 we present approximate Riemann solvers.

## Some Properties of the Euler Equations

In this chapter we apply the mathematical tools presented in Chap. 2 to analyse some of the basic properties of the time-dependent Euler equations. As seen in Chap. 1, the Euler equations result from neglecting the effects of viscosity, heat conduction and body forces on a compressible medium. Here we show that these equations are a system of *hyperbolic* conservations laws and study some of their mathematical properties. In particular, we study those properties that are essential for finding the solution of the Riemann problem in Chap. 4. We analyse the eigenstructure of the equations, that is, we find eigenvalues and eigenvectors; we study properties of the characteristic fields and establish basic relations across rarefactions, contacts and shock waves. It is worth remarking that the process of finding eigenvalues and eigenvectors usually involves a fair amount of algebra as well as some familiarity with basic physical quantities and their relations. For very complex systems of equations finding eigenvalues and eigenvectors may require the use of symbolic manipulators. Useful background reading for this chapter is found in Chaps. 1 and 2.

### 3.1 The One-Dimensional Euler Equations

Here we study the one-dimensional time-dependent Euler equations with an ideal Equation of State, using conservative and non-conservative formulations. The basic structure of the solution of the Riemann problem is outlined along with a detailed study of the elementary waves present in the solution. We provide the foundations for finding the exact solution of the Riemann problem in Chap. 4.

#### 3.1.1 Conservative Formulation

The conservative formulation of the Euler equations, in differential form, is

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0}, \quad (3.1)$$

where  $\mathbf{U}$  and  $\mathbf{F}(\mathbf{U})$  are the vectors of conserved variables and fluxes, given respectively by

$$\mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} \rho \\ \rho u \\ E \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ u(E + p) \end{bmatrix}. \quad (3.2)$$

Here  $\rho$  is density,  $p$  is pressure,  $u$  is particle velocity and  $E$  is total energy per unit volume

$$E = \rho\left(\frac{1}{2}u^2 + e\right), \quad (3.3)$$

where  $e$  is the *specific internal energy* given by a caloric Equation of State (EOS)

$$e = e(\rho, p). \quad (3.4)$$

For ideal gases one has the simple expression

$$e = e(\rho, p) = \frac{p}{(\gamma - 1)\rho}, \quad (3.5)$$

with  $\gamma = c_p/c_v$  denoting the *ratio of specific heats*. From the EOS (3.5) and using equation (1.36) of Chap. 1 we write the *sound speed*  $a$  as

$$a = \sqrt{(p/\rho^2 - e_\rho)/e_p} = \sqrt{\frac{\gamma p}{\rho}}. \quad (3.6)$$

The conservation laws (3.1)–(3.2) may also be written in quasi-linear form

$$\mathbf{U}_t + \mathbf{A}(\mathbf{U})\mathbf{U}_x = \mathbf{0}, \quad (3.7)$$

where the coefficient matrix  $\mathbf{A}(\mathbf{U})$  is the *Jacobian matrix*

$$\mathbf{A}(\mathbf{U}) = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} = \begin{bmatrix} \partial f_1 / \partial u_1 & \partial f_1 / \partial u_2 & \partial f_1 / \partial u_3 \\ \partial f_2 / \partial u_1 & \partial f_2 / \partial u_2 & \partial f_2 / \partial u_3 \\ \partial f_3 / \partial u_1 & \partial f_3 / \partial u_2 & \partial f_3 / \partial u_3 \end{bmatrix}.$$

**Proposition 3.1 (Jacobian Matrix).** *The Jacobian matrix  $\mathbf{A}$  is*

$$\mathbf{A}(\mathbf{U}) = \begin{bmatrix} 0 & 1 & 0 \\ -\frac{1}{2}(\gamma - 3)(\frac{u_2}{u_1})^2 & (3 - \gamma)(\frac{u_2}{u_1}) & \gamma - 1 \\ -\frac{\gamma u_2 u_3}{u_1^2} + (\gamma - 1)(\frac{u_2}{u_1})^3 & \frac{\gamma u_3}{u_1} - \frac{3}{2}(\gamma - 1)(\frac{u_2}{u_1})^2 & \gamma(\frac{u_2}{u_1}) \end{bmatrix}.$$

*Proof.* First we express all components  $f_i$  of the flux vector  $\mathbf{F}$  in terms of the components  $u_i$  of the vector  $\mathbf{U}$  of conserved variables, namely  $u_1 \equiv \rho$ ,  $u_2 \equiv \rho u$ ,  $u_3 \equiv E$ . Obviously  $f_1 = u_2 \equiv \rho u$ . To find  $f_2$  and  $f_3$  we first need

to express the pressure  $p$  in terms of the conserved variables. From (3.3) and (3.5) we find

$$p = (\gamma - 1)[u_3 - \frac{1}{2}(u_2^2/u_1)] .$$

Thus the flux vector can be written as

$$\mathbf{F}(\mathbf{U}) = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} \equiv \begin{bmatrix} u_2 \\ \frac{1}{2}(3-\gamma)\frac{u_2^2}{u_1} + (\gamma-1)u_3 \\ \gamma\frac{u_2}{u_1}u_3 - \frac{1}{2}(\gamma-1)\frac{u_3^3}{u_1^2} \end{bmatrix} .$$

By direct evaluation of all partial derivatives we arrive at the sought result.

**Exercise 3.2.** Write the Jacobian matrix  $\mathbf{A}(\mathbf{U})$  in terms of the sound speed  $a$  and the particle velocity  $u$ .

**Solution 3.3.**

$$\mathbf{A}(\mathbf{U}) = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2}(\gamma-3)u^2 & (3-\gamma)u & \gamma-1 \\ \frac{1}{2}(\gamma-2)u^3 - \frac{a^2 u}{\gamma-1} & \frac{3-2\gamma}{2}u^2 + \frac{a^2}{\gamma-1} & \gamma u \end{bmatrix} . \quad (3.8)$$

Often, the Jacobian matrix is also expressed in terms of the *total specific enthalpy*  $H$ , which is related to the *specific enthalpy*  $h$  and other variables, namely

$$H = (E + p)/\rho \equiv \frac{1}{2}u^2 + h , \quad h = e + p/\rho . \quad (3.9)$$

The Jacobian matrix may also be written as

$$\mathbf{A}(\mathbf{U}) = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2}(\gamma-3)u^2 & (3-\gamma)u & \gamma-1 \\ (\gamma-1)u^3 - \gamma u E/\rho & \gamma E/\rho - \frac{3}{2}(\gamma-1)u^2 & \gamma u \end{bmatrix} . \quad (3.10)$$

**Proposition 3.4 (The Homogeneity Property).** *The Euler equations (3.1)–(3.2) with the ideal-gas EOS (3.5) satisfy the homogeneity property*

$$\mathbf{F}(\mathbf{U}) = \mathbf{A}(\mathbf{U})\mathbf{U} . \quad (3.11)$$

*Proof.* The proof of this property is immediate. By multiplying the Jacobian matrix (3.8) by the vector  $\mathbf{U}$  in (3.2) we identically reproduce the vector  $\mathbf{F}(\mathbf{U})$  of fluxes in (3.2).

This remarkable property of the Euler equations forms the basis for numerical schemes of the *Flux Vector Splitting* type studied in Chap. 8. Note that the relationship between the flux  $\mathbf{F}$ , the coefficient matrix  $\mathbf{A}$  and the conserved variables  $\mathbf{U}$  for the Euler equations is identical to that for linear systems with constant coefficients, see Sect. 2.4 of Chap. 2. This property is also satisfied by the Euler equations with an Equation of State that is slightly more general than (3.5). See Steger and Warming [463] for details.

**Proposition 3.5.** *The eigenvalues of the Jacobian matrix  $\mathbf{A}$  are*

$$\lambda_1 = u - a, \quad \lambda_2 = u, \quad \lambda_3 = u + a \quad (3.12)$$

and the corresponding right eigenvectors are

$$\mathbf{K}^{(1)} = \begin{bmatrix} 1 \\ u - a \\ H - ua \end{bmatrix}, \quad \mathbf{K}^{(2)} = \begin{bmatrix} 1 \\ u \\ \frac{1}{2}u^2 \end{bmatrix}, \quad \mathbf{K}^{(3)} = \begin{bmatrix} 1 \\ u + a \\ H + ua \end{bmatrix}. \quad (3.13)$$

*Proof.* Use of the expression (3.8) for  $\mathbf{A}$  and the *characteristic polynomial*

$$|\mathbf{A} - \lambda \mathbf{I}| = 0,$$

lead to

$$(\lambda - u)(\gamma u - \lambda)[(2u - \gamma u - \lambda)] + (\lambda - u)[-a^2 - (\gamma - 1)u^2 + (\gamma - 1)\gamma u^2] + \Delta = 0,$$

where

$$\Delta = \frac{1}{2}(\gamma u - \lambda)(1 - \gamma)u^2 - \frac{1}{2}(\gamma - 1)u^2[(1 - 2\gamma)\lambda + \gamma u].$$

Manipulations show that  $\Delta$  also contains the common factor  $(\lambda - u)$ , which implies that  $\lambda_2 = u$  is a root of the characteristic polynomial and thus an eigenvalue of  $\mathbf{A}$ . After cancelling  $(\lambda - u)$  the remaining terms give

$$\lambda^2 - 2u\lambda + u^2 - a^2 = 0,$$

with real roots

$$\lambda_1 = u - a, \quad \lambda_3 = u + a.$$

Therefore the eigenvalues are:  $\lambda_1 = u - a$ ,  $\lambda_2 = u$ ,  $\lambda_3 = u + a$  as claimed. To find the right eigenvectors we look, see Sect. 2.1 of Chap. 2, for a vector  $\mathbf{K} = [k_1, k_2, k_3]^T$  such that

$$\mathbf{A}\mathbf{K} = \lambda\mathbf{K}.$$

By substituting  $\lambda = \lambda_i$  in turn, solving for the components of the vector  $\mathbf{K}$  and selecting appropriate values for the scaling factors we find the desired result.

The eigenvalues are all real and the eigenvectors  $\mathbf{K}^{(1)}$ ,  $\mathbf{K}^{(2)}$ ,  $\mathbf{K}^{(3)}$  form a complete set of *linearly independent eigenvectors*. We have thus proved that the time-dependent, one-dimensional Euler equations for ideal gases are *hyperbolic*. In fact these equations are *strictly hyperbolic*, because the eigenvalues are all real and *distinct*, as long as the sound speed  $a$  remains positive. Hyperbolicity remains a property of the Euler equations for more general equations of state, as we shall see in Chap. 4 for covolume gases.

### 3.1.2 Non-Conservative Formulations

The Euler equations (3.1)–(3.2) may be formulated in terms of variables other than the conserved variables. For smooth solutions all formulations are equivalent. For solutions containing shock waves however, non-conservative formulations give incorrect shock solutions. This point is addressed via the shallow water equations and the isothermal equations in Sect. 3.3 of this chapter. In spite of this, non-conservative formulations have some advantages over their conservative counterpart, when analysing the equations, for instance. Also, from the numerical point of view, there has been a recent revival of the idea of using schemes for non-conservative formulations of the equations. See e.g. Karni [278] and Toro [508], [517].

#### Primitive-Variable Formulations

For smooth solutions the equations may be formulated, and solved, using variables other than the conserved variables. For the one-dimensional case one possibility is to choose a vector  $\mathbf{W} = (\rho, u, p)^T$  of *primitive* or *physical* variables, with  $p$  given by the equation of state. Expanding derivatives in the first of equations (3.1)–(3.2), the mass equation, we obtain

$$\rho_t + u\rho_x + \rho u_x = 0 . \quad (3.14)$$

By expanding derivatives in the second of equations (3.1)–(3.2), the momentum equation, we obtain

$$u [\rho_t + u\rho_x + \rho u_x] + \rho \left[ u_t + uu_x + \frac{1}{\rho} p_x \right] = 0 .$$

Use of (3.14) followed by division through by  $\rho$  gives

$$u_t + uu_x + \frac{1}{\rho} p_x = 0 . \quad (3.15)$$

In a similar manner, the energy equation in (3.1)–(3.2) can be rearranged so as to use (3.14) and (3.15). The result is

$$p_t + \rho a^2 u_x + u p_x = 0 . \quad (3.16)$$

Thus, in quasi-linear form we have

$$\mathbf{W}_t + \mathbf{A}(\mathbf{W})\mathbf{W}_x = \mathbf{0}, \quad (3.17)$$

where

$$\mathbf{W} = \begin{bmatrix} \rho \\ u \\ p \end{bmatrix}, \quad \mathbf{A}(\mathbf{W}) = \begin{bmatrix} u & \rho & 0 \\ 0 & u & 1/\rho \\ 0 & \rho a^2 & u \end{bmatrix}. \quad (3.18)$$

**Proposition 3.6.** *The system (3.17)–(3.18) has real eigenvalues*

$$\lambda_1 = u - a, \quad \lambda_2 = u, \quad \lambda_3 = u + a, \quad (3.19)$$

with corresponding right eigenvectors

$$\mathbf{K}^{(1)} = \alpha_1 \begin{bmatrix} 1 \\ -a/\rho \\ a^2 \end{bmatrix}, \quad \mathbf{K}^{(2)} = \alpha_2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{K}^{(3)} = \alpha_3 \begin{bmatrix} 1 \\ a/\rho \\ a^2 \end{bmatrix}. \quad (3.20)$$

where  $\alpha_1, \alpha_2, \alpha_3$  are scaling factors, or normalisation parameters, see Sect. 2.1 of Chap. 2. The left eigenvectors are

$$\left. \begin{array}{l} \mathbf{L}^{(1)} = \beta_1(0, 1, -\frac{1}{\rho a}), \\ \mathbf{L}^{(2)} = \beta_2(1, 0, -\frac{1}{a^2}), \\ \mathbf{L}^{(3)} = \beta_3(0, 1, \frac{1}{\rho a}), \end{array} \right\} \quad (3.21)$$

where  $\beta_1, \beta_2, \beta_3$  are scaling factors.

*Proof.* (Left to the reader).

**Exercise 3.7.** Verify that by choosing appropriate normalisation parameters  $\alpha_1, \alpha_2, \alpha_3$  and  $\beta_1, \beta_2, \beta_3$  in (3.20) and (3.21) respectively, the left and right eigenvectors  $\mathbf{L}^{(j)}$  and  $\mathbf{K}^{(j)}$  of  $\mathbf{A}(\mathbf{W})$  are bi-orthonormal, that is

$$\mathbf{L}^{(j)} \cdot \mathbf{K}^{(i)} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases} \quad (3.22)$$

## Characteristic Equations

Recall that the eigenvalues  $\lambda_1 = u - a, \lambda_2 = u, \lambda_3 = u + a$  define characteristic directions  $dx/dt = \lambda_i$  for  $i = 1, 2, 3$ . For a set of partial differential equations (3.17) a *characteristic equation* says that in a direction  $dx/dt = \lambda_i$ ,  $\mathbf{L}^{(i)} \cdot d\mathbf{W} = 0$ , or in full

$$\mathbf{L}^{(i)} \cdot \begin{bmatrix} d\rho \\ du \\ dp \end{bmatrix} = 0 . \quad (3.23)$$

By expanding (3.23) for  $\mathbf{L}^{(1)}, \mathbf{L}^{(2)}, \mathbf{L}^{(3)}$  we obtain the *characteristic equations*

$$\left. \begin{array}{l} dp - \rho a \, du = 0 \text{ along } dx/dt = \lambda_1 = u - a , \\ dp - a^2 \, d\rho = 0 \text{ along } dx/dt = \lambda_2 = u , \\ dp + \rho a \, du = 0 \text{ along } dx/dt = \lambda_3 = u + a . \end{array} \right\} \quad (3.24)$$

These differential relations hold true *along characteristic* directions. For numerical purposes, linearisation of these equations provides ways of solving the Riemann problem for the Euler equations, approximately; see Sect. 9.3 of Chap. 9.

### Entropy Formulation

The entropy  $s$  can be written as

$$s = c_v \ln\left(\frac{p}{\rho^\gamma}\right) + C_0 , \quad (3.25)$$

where  $C_0$  is a constant. From this equation we obtain

$$p = C_1 \rho^\gamma e^{s/c_v} , \quad (3.26)$$

where  $C_1$  is a constant. Now, if in the primitive-variable formulation (3.17) we use entropy  $s$  instead of pressure  $p$  we have the new vector of unknowns

$$\mathbf{W} = (\rho, u, s)^T , \quad (3.27)$$

and a corresponding new way of expressing the governing equations.

**Proposition 3.8.** *The entropy  $s$  satisfies the following PDE*

$$s_t + us_x = 0 . \quad (3.28)$$

*Proof.* From (3.25) and the expression (3.6) for the sound speed  $a$  we have

$$s_t = \frac{c_v}{p} [p_t - a^2 \rho_t] , \quad s_x = \frac{c_v}{p} [p_x - a^2 \rho_x] .$$

But from (3.16)  $p_t = -\rho a^2 u_x - u p_x$ , and hence  $s_t + us_x = 0$ , as claimed.

The significance of the result is that

$$s_t + us_x = \frac{ds}{dt} = 0 , \quad (3.29)$$

and so in regions of smooth flow, the entropy  $s$  is constant along particle paths  $dx/dt = u$ . Hence, along a particle path one has the *isentropic law* given by

$$p = C\rho^\gamma, \quad (3.30)$$

where  $C = C(s_0)$  is a function of the *initial entropy*  $s_0$  and is constant along the path so long as the flow remains smooth; see Sect. 1.6.2 of Chap. 1. In general of course,  $C$  changes from path to path. When solving the Riemann problem the initial entropy can be computed on the initial data of the Riemann problem, which is piece-wise constant. If  $C$  is the same constant throughout the flow domain we speak of *isentropic flow*, or sometimes, *homentropic flow*. This leads to the special set of governing equations (1.109)–(1.110) presented in Chap. 1. The governing equations for the entropy formulation, written in quasi-linear form, are

$$\mathbf{W}_t + \mathbf{A}(\mathbf{W})\mathbf{W}_x = \mathbf{0}, \quad (3.31)$$

with

$$\mathbf{A}(\mathbf{W}) = \begin{bmatrix} u & \rho & 0 \\ a^2/\rho & u & \frac{1}{\rho} \frac{\partial p}{\partial s} \\ 0 & 0 & u \end{bmatrix}. \quad (3.32)$$

**Proposition 3.9.** *The eigenvalues of system (3.31)–(3.32) are*

$$\lambda_1 = u - a, \lambda_2 = u, \lambda_3 = u + a \quad (3.33)$$

and the corresponding right eigenvectors are

$$\mathbf{K}^{(1)} = \begin{bmatrix} 1 \\ -a/\rho \\ 0 \end{bmatrix}, \quad \mathbf{K}^{(2)} = \begin{bmatrix} -\frac{\partial p}{\partial s} \\ 0 \\ a^2 \end{bmatrix}, \quad \mathbf{K}^{(3)} = \begin{bmatrix} 1 \\ a/\rho \\ 0 \end{bmatrix}. \quad (3.34)$$

*Proof.* (Left to the reader).

### 3.1.3 Elementary Wave Solutions of the Riemann Problem

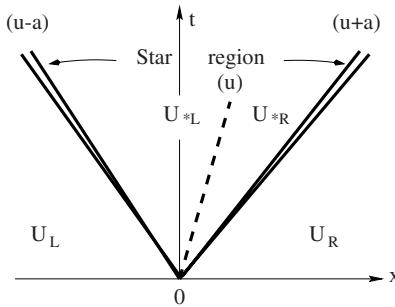
Here we describe the structure of the solution of the Riemann problem as a set of *elementary waves* such as rarefactions, contacts and shock waves, see Sect. 2.4.4 of Chapt. 4. Each of these elementary waves are studied in detail. Basic relations across these waves are established. Such relations will be used in Chapt. 4 to connect all unknown states to the data states and thus find the complete solution of the Riemann problem.

The Riemann problem for the one-dimensional, time dependent Euler equations (3.1)–(3.2) with data  $(\mathbf{U}_L, \mathbf{U}_R)$  is the IVP

$$\left. \begin{aligned} \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x &= \mathbf{0}, \\ \mathbf{U}(x, 0) &= \mathbf{U}^{(0)}(x) = \begin{cases} \mathbf{U}_L & \text{if } x < 0, \\ \mathbf{U}_R & \text{if } x > 0. \end{cases} \end{aligned} \right\} \quad (3.35)$$

The physical analogue of the Riemann problem is the *shock-tube problem* in Gas Dynamics, in which the velocities  $u_L$  and  $u_R$  on either side of the diaphragm, here idealised by an initial discontinuity, are zero. Shock tubes and shock-tube problems have played, over a period of more than 100 years, a fundamental role in fluid dynamics research.

The structure of the similarity solution  $\mathbf{U}(x/t)$  of (3.35) is as depicted in Fig. 3.1. There are three waves *associated* with the three characteristic fields



**Fig. 3.1.** Structure of the solution of the Riemann problem in the  $x$ - $t$  plane for the time-dependent, one dimensional Euler equations. There are three wave families associated with the eigenvalues  $u - a$ ,  $u$  and  $u + a$

corresponding to the eigenvectors  $\mathbf{K}^{(i)}$ ,  $i = 1, 2, 3$ . We choose the convention of representing the outer waves, when their character is unknown, by a pair of rays emanating from the origin and the middle wave by a dashed line. Each wave family is shown along with the corresponding eigenvalue. The three waves separate four constant states. From left to right these are  $\mathbf{U}_L$  (left data state);  $\mathbf{U}_{*L}$  between the 1-wave and the 2-wave;  $\mathbf{U}_{*R}$  between the 2-wave and the 3-wave and  $\mathbf{U}_R$  (right data state). As we shall see the waves present in the solution are of three types: rarefaction waves, contact discontinuities and shock waves. In order to identify the types we analyse the characteristic fields for  $\mathbf{K}^{(i)}$ ,  $i = 1, 2, 3$ ; see Sects. 2.4.3 and 2.4.4 of Chap. 2.

**Proposition 3.10.** *The  $\mathbf{K}^{(2)}$ -characteristic field is linearly degenerate and the  $\mathbf{K}^{(1)}$ ,  $\mathbf{K}^{(3)}$  characteristic fields are genuinely non-linear.*

*Proof.* For the  $\mathbf{K}^{(2)}$ -characteristic field we have

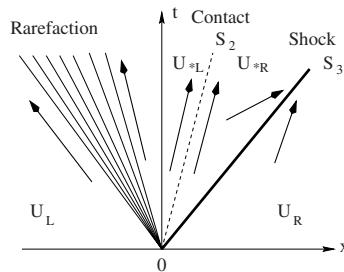
$$\nabla \lambda_2(\mathbf{U}) = [\partial \lambda_2 / \partial u_1, \partial \lambda_2 / \partial u_2, \partial \lambda_2 / \partial u_3] = [-u/\rho, 1/\rho, 0].$$

Hence

$$\nabla \lambda_2 \cdot \mathbf{K}^{(2)} = [-u/\rho, 1/\rho, 0] \cdot \begin{bmatrix} 1 \\ u \\ \frac{1}{2}u^2 \end{bmatrix} = 0$$

and therefore the  $\mathbf{K}^{(2)}$  characteristic field is linearly degenerate as claimed. The proof that the  $\mathbf{K}^{(1)}$  and  $\mathbf{K}^{(3)}$  characteristic fields are *genuinely nonlinear* is left to the reader.

The wave associated with the  $\mathbf{K}^{(2)}$  characteristic field is a *contact discontinuity* and those associated with the  $\mathbf{K}^{(1)}$ ,  $\mathbf{K}^{(3)}$  characteristic fields will either be rarefaction waves (smooth) or shock waves (discontinuities), see Sect. 2.4.4 of Chapt. 4. Of course one does not know in advance what types of waves will be present in the solution of the Riemann problem. The only exception is the middle wave, which is always a contact discontinuity. Fig. 3.2 shows a



**Fig. 3.2.** Structure of the solution of the Riemann problem in the  $x$ - $t$  plane for the time-dependent, one dimensional Euler equations, in which the left wave is a rarefaction, the middle wave is a contact discontinuity and the right wave is a shock wave

particular case in which the left wave is a rarefaction, the middle wave is a contact and the right wave is a shock wave. For each wave we have drawn a pair of arrows, one on each side, to indicate the characteristic directions of the corresponding eigenvalue. For the rarefaction wave we have

$$\lambda_1(\mathbf{U}_L) \leq \lambda_1(\mathbf{U}_{*L}) .$$

The eigenvalue  $\lambda_1(\mathbf{U})$  increases monotonically as we cross the rarefaction wave from left to right and the characteristics on either side diverge from the wave; compare with Fig. 2.20 of Chap. 2. For the shock wave, characteristics run into the wave and we have

$$\lambda_3(\mathbf{U}_{*R}) > S_3 > \lambda_3(\mathbf{U}_R) ,$$

which is the *entropy condition*. See Sect. 2.4.4 of Chap. 2.  $S_3$  is the speed of the 3-shock. For the contact wave we have

$$\lambda_2(\mathbf{U}_{*L}) = \lambda_2(\mathbf{U}_{*R}) = S_2 ,$$

where  $S_2$  is the speed of the contact wave; the characteristics are parallel to the contact wave. Recall that this is what happens for *all characteristic fields* in linear hyperbolic systems with constant coefficients. Next we study each type of waves separately.

## Contact Discontinuities

The contact discontinuity in the solution of the Riemann problem for the Euler equations can be analysed by utilising the eigenstructure of the equations. In particular the Generalised Riemann Invariants will reveal which quantities change across the wave. Recall that for a general  $m \times m$  hyperbolic system, such as (3.1)–(3.2) or (3.7), with

$$\mathbf{W} = [w_1, w_2, \dots, w_m]^T ,$$

and right eigenvectors

$$\mathbf{K}^{(i)} = \left[ k_1^{(i)}, k_2^{(i)}, \dots, k_m^{(i)} \right] ,$$

the  $i$ -th Generalised Riemann Invariants are the  $(m - 1)$  ODEs

$$\frac{dw_1}{k_1^{(i)}} = \frac{dw_2}{k_2^{(i)}} = \frac{dw_3}{k_3^{(i)}} = \dots = \frac{dw_m}{k_m^{(i)}} .$$

Using the eigenstructure (3.12)–(3.13) of the conservative formulation (3.1)–(3.2), for the  $\mathbf{K}^{(2)}$ -wave we have

$$\frac{d\rho}{1} = \frac{d(\rho u)}{u} = \frac{dE}{\frac{1}{2}u^2} . \quad (3.36)$$

Manipulation of these equalities gives

$$p = \text{constant}, u = \text{constant}$$

*across the contact wave.* The same result follows directly by inspection of the eigenvector  $\mathbf{K}^{(2)}$  in (3.20) for the primitive-variable formulation (3.17)–(3.18): the wave jumps in  $\rho$ ,  $u$  and  $p$  are proportional to the corresponding components of the eigenvector. These are zero for the velocity and pressure. The jump in  $\rho$  is in general non-trivial. To conclude: *a contact wave is a discontinuous wave across which both pressure and particle velocity are constant but density jumps discontinuously as do variables that depend on density, such as specific internal energy, temperature, sound speed, entropy, etc.*

## Rarefaction Waves

Rarefaction waves in the Euler equations are associated with the  $\mathbf{K}^{(1)}$  and  $\mathbf{K}^{(3)}$  characteristic fields. Inspection of the eigenvectors (3.20) for the primitive-variable formulation reveals that  $\rho$ ,  $u$  and  $p$  change across a rarefaction wave. We now utilise the Generalised Riemann Invariants for the eigenstructure (3.33)–(3.34) of the entropy formulation (3.31)–(3.32).

**Proposition 3.11.** *For the Euler equations the Generalised Riemann Invariants across 1 and 3 rarefactions are*

$$\left. \begin{aligned} I_L(u, a) &= u + \frac{2a}{\gamma-1} = \text{constant} \\ s &= \text{constant} \end{aligned} \right\} \text{across } \lambda_1 = u - a, \quad (3.37)$$

$$\left. \begin{aligned} I_R(u, a) &= u - \frac{2a}{\gamma-1} = \text{constant} \\ s &= \text{constant} \end{aligned} \right\} \text{across } \lambda_3 = u + a. \quad (3.38)$$

*Proof.* Across a wave associated with  $\lambda_1 = u - a$  wave we have

$$\frac{d\rho}{1} = \frac{du}{-a/\rho} = \frac{ds}{0}.$$

Two meaningful relations are

$$u + \int \frac{a}{\rho} d\rho = \text{constant} \text{ and } s = \text{constant}. \quad (3.39)$$

Similarly, across the  $\lambda_3 = u + a$  wave we have

$$u - \int \frac{a}{\rho} d\rho = \text{constant} \text{ and } s = \text{constant}. \quad (3.40)$$

In order to reproduce (3.37) and (3.38) we need to evaluate the integrals in (3.39) and (3.40). First we note that by inspection of the eigenvectors  $\mathbf{K}^{(1)}$  and  $\mathbf{K}^{(3)}$  the condition of constant entropy across the respective waves is immediate. We may therefore use the isentropic law (3.30) with the constant  $C$  evaluated at the appropriate data state (constant). Thus the integral is as found for the isentropic equations in Sect. 2.4.3 of Chap. 2, that is

$$\int \frac{a}{\rho} d\rho = \frac{2a}{\gamma-1},$$

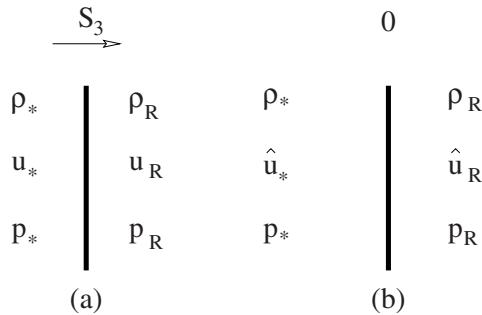
and thus equations (3.37)–(3.38) are reproduced.

To summarise: *a rarefaction wave is a smooth wave associated with the 1 and 3 fields across which  $\rho$ ,  $u$  and  $p$  change. The wave has a fan-type shape and is enclosed by two bounding characteristics corresponding to the Head and the Tail of the wave. Across the wave the Generalised Riemann Invariants apply.* The solution within the rarefaction will be given in Chap. 4, where the full solution of the Riemann problem is presented.

## Shock Waves

Details on the Physics of shock waves are found in any book on Gas Dynamics. We particularly recommend Becker [35], Anderson [10], Landau and Lifshitz [297]. The specialised book by Zeldovich and Raizer [599] is highly recommended.

In the context of the one-dimensional Euler equations, shock waves are discontinuous waves associated with the genuinely non-linear fields 1 and 3. All three quantities  $\rho$ ,  $u$  and  $p$  change across a shock wave. Consider the  $\mathbf{K}^{(3)}$  characteristic field and assume the corresponding wave is a right-facing shock wave travelling at the constant speed  $S_3$ ; see Fig. 3.3. In terms of the primitive variables we denote the state ahead of the shock by  $\mathbf{W}_R = (\rho_R, u_R, p_R)^T$  and the state behind the shock by  $\mathbf{W}_* = (\rho_*, u_*, p_*)^T$ . We are interested in deriving relations, across the shock wave, between the various quantities involved. Central to the analysis is the application of the Rankine–Hugoniot conditions. It is found convenient to transform the problem to a new frame



**Fig. 3.3.** Right-facing shock wave: (a) stationary frame of reference, shock has speed  $S_3$ ; (b) frame of reference moves with speed  $S_3$ , so that the shock has zero speed

of reference moving with the shock so that in the new frame the shock speed is zero. Fig. 3.3 depicts both frames of reference. In the transformed frame (b) the states ahead and behind the shock have changed by virtue of the transformation. Densities and pressures remain unaltered while velocities have changed to the *relative velocities*  $\hat{u}_R$  and  $\hat{u}_*$  given by

$$\hat{u}_* = u_* - S_3, \quad \hat{u}_R = u_R - S_3. \quad (3.41)$$

Application of the Rankine–Hugoniot conditions in the frame in which the shock speed is zero gives

$$\rho_* \hat{u}_* = \rho_R \hat{u}_R, \quad (3.42)$$

$$\rho_* \hat{u}_*^2 + p_* = \rho_R \hat{u}_R^2 + p_R, \quad (3.43)$$

$$\hat{u}_* (\hat{E}_* + p_*) = \hat{u}_R (\hat{E}_R + p_R). \quad (3.44)$$

By using the definition of total energy  $E$  and introducing the *specific internal energy*  $e$  the left-hand side of (3.44) may be written as

$$\hat{u}_* \rho_* \left[ \frac{1}{2} \hat{u}_*^2 + (e_* + p_*/\rho_*) \right]$$

and the right-hand side of (3.44) as

$$\hat{u}_R \rho_R \left[ \frac{1}{2} \hat{u}_R^2 + (e_R + p_R / \rho_R) \right] .$$

Now we use the *specific enthalpy*  $h$  and write

$$h_* = e_* + p_* / \rho_* , \quad h_R = e_R + p_R / \rho_R . \quad (3.45)$$

Use of equations (3.42) and (3.44) leads to

$$\frac{1}{2} \hat{u}_*^2 + h_* = \frac{1}{2} \hat{u}_R^2 + h_R . \quad (3.46)$$

By using (3.42) into (3.43) we write

$$\rho_* \hat{u}_*^2 = (\rho_R \hat{u}_R) \hat{u}_R + p_R - p_* = (\rho_* \hat{u}_*) \frac{\rho_* \hat{u}_*}{\rho_R} + p_R - p_* .$$

After some manipulations we obtain

$$\hat{u}_*^2 = \left( \frac{\rho_R}{\rho_*} \right) \left[ \frac{p_R - p_*}{\rho_R - \rho_*} \right] . \quad (3.47)$$

In a similar way we obtain

$$\hat{u}_R^2 = \left( \frac{\rho_*}{\rho_R} \right) \left[ \frac{p_R - p_*}{\rho_R - \rho_*} \right] . \quad (3.48)$$

Substitution of (3.47)–(3.48) into (3.46) gives

$$h_* - h_R = \frac{1}{2} (p_* - p_R) \left[ \frac{\rho_* + \rho_R}{\rho_* \rho_R} \right] . \quad (3.49)$$

Assuming the specific internal energy  $e$  is given by the the caloric equation of state (3.4), it is then more convenient to rewrite the energy equation (3.49) using (3.45). We obtain

$$e_* - e_R = \frac{1}{2} (p_* + p_R) \left[ \frac{\rho_* - \rho_R}{\rho_* \rho_R} \right] . \quad (3.50)$$

Note that up to this point no assumption on the general caloric EOS (3.4) has been made. In what follows, we derive shock relations that apply to ideal gases in which the ideal caloric EOS (3.5) is assumed. By using (3.5) into (3.50) and performing some algebraic manipulations one obtains

$$\frac{\rho_*}{\rho_R} = \frac{\left( \frac{p_*}{p_R} \right) + \left( \frac{\gamma-1}{\gamma+1} \right)}{\left( \frac{\gamma-1}{\gamma+1} \right) \left( \frac{p_*}{p_R} \right) + 1} . \quad (3.51)$$

This establishes a useful relation between the density ratio  $\rho_*/\rho_R$  and the pressure ratio  $p_*/p_R$  across the shock wave.

We now introduce Mach numbers

$$M_R = u_R/a_R, \quad M_S = S_3/a_R, \quad (3.52)$$

where  $M_R$  is the Mach number of the flow ahead of the shock, in the original frame;  $M_S$  is the *shock Mach number*. Manipulation of equations (3.48), (3.51) and (3.52) leads to expressions for the density and pressure ratios across the shock as functions of the relative Mach number  $M_R - M_S$ , namely

$$\frac{\rho_*}{\rho_R} = \frac{(\gamma + 1)(M_R - M_S)^2}{(\gamma - 1)(M_R - M_S)^2 + 2}, \quad (3.53)$$

$$\frac{p_*}{p_R} = \frac{2\gamma(M_R - M_S)^2 - (\gamma - 1)}{(\gamma + 1)}. \quad (3.54)$$

The shock speed  $S_3$  can be related to the density and pressure ratios across the shock wave. In terms of the pressure ratio (3.54) we first note the following relationship

$$M_R - M_S = -\sqrt{\left(\frac{\gamma + 1}{2\gamma}\right)\left(\frac{p_*}{p_R}\right) + \left(\frac{\gamma - 1}{2\gamma}\right)}.$$

This leads to an expression for the shock speed as a function of the pressure ratio across the shock, namely

$$S_3 = u_R + a_R \sqrt{\left(\frac{\gamma + 1}{2\gamma}\right)\left(\frac{p_*}{p_R}\right) + \left(\frac{\gamma - 1}{2\gamma}\right)}. \quad (3.55)$$

Note that as the shock strength tends to zero, the ratio  $p_*/p_R$  tends to unity and the shock speed  $S_3$  approaches the characteristic speed  $\lambda_3 = u_R + a_R$ , as expected. We can also obtain an expression for the *particle velocity*  $u_*$  behind the shock wave. From (3.42) we relate  $u_*$  to the density ratio across the shock, namely

$$u_* = (1 - \rho_R/\rho_*)S_3 + u_R\rho_R/\rho*. \quad (3.56)$$

*Example 3.12 (Shock Wave).* Consider a shock wave of shock Mach number  $M_S = 3$  propagating into the atmosphere with conditions (ahead of the shock)  $\rho_R = 1.225 \text{ kg/m}^3$ ,  $u_R = 0 \text{ m/s}$ ,  $p_R = 101325 \text{ Pa}$ . Assume the process is suitably modelled by the ideal gas EOS (3.5) with  $\gamma = 1.4$ . From the definition of sound speed (3.6) we obtain  $a_R = 340.294 \text{ m/s}$ . As the shock Mach number  $M_S = 3$  is assumed (a parameter) then equation (3.52) gives the shock speed as  $S = 1020.882 \text{ m/s}$ . From equation (3.53) we obtain  $\rho_* = 4.725 \text{ kg/m}^3$ . From equation (3.54) we obtain  $p_* = 1047025 \text{ Pa}$  and from equation (3.56) we obtain  $u_* = 756.2089 \text{ m/s}$ .

*Remark 3.13.* Shock relations (3.53), (3.54) and (3.56) define a state

$$(\rho_*, u_*, p_*)^T$$

behind a shock for given initial conditions  $(\rho_R, u_R, p_R)^T$  ahead of the shock and a chosen shock Mach number  $M_S$ , or equivalently a shock speed  $S_3$ . The shock is associated with the 3-wave family. These relations can be useful in setting up test problems involving a single shock wave to test numerical methods.

The analysis for a 1-shock wave (left facing) travelling with velocity  $S_1$  is entirely analogous. The state ahead of the shock (left side now) is denoted by  $\mathbf{W}_L = (\rho_L, u_L, p_L)^T$  and the state behind the shock (right side) by  $\mathbf{W}_* = (\rho_*, u_*, p_*)^T$ . As done for the 3-shock we transform to a stationary frame of reference. The relative velocities are

$$\hat{u}_L = u_L - S_1, \quad \hat{u}_* = u_* - S_1. \quad (3.57)$$

Mach numbers are

$$M_L = u_L/a_L, \quad M_S = S_1/a_L. \quad (3.58)$$

The density and pressure ratio relationship is

$$\frac{\rho_*}{\rho_L} = \frac{\left(\frac{p_*}{p_L}\right) + \left(\frac{\gamma-1}{\gamma+1}\right)}{\left(\frac{\gamma-1}{\gamma+1}\right)\left(\frac{p_*}{p_L}\right) + 1}. \quad (3.59)$$

In terms of the relative Mach number  $M_L - M_S$  the density and pressure ratios across the left shock can be expressed as follows

$$\frac{\rho_*}{\rho_L} = \frac{(\gamma + 1)(M_L - M_S)^2}{(\gamma - 1)(M_L - M_S)^2 + 2}, \quad (3.60)$$

$$\frac{p_*}{p_L} = \frac{2\gamma(M_L - M_S)^2 - (\gamma - 1)}{(\gamma + 1)}. \quad (3.61)$$

The shock speed  $S_1$  can be obtained from either (3.60) or (3.61). In terms of the pressure ratio (3.61) we have

$$M_L - M_S = \sqrt{\left(\frac{\gamma + 1}{2\gamma}\right)\left(\frac{p_*}{p_L}\right) + \left(\frac{\gamma - 1}{2\gamma}\right)},$$

which leads to

$$S_1 = u_L - a_L \sqrt{\left(\frac{\gamma + 1}{2\gamma}\right)\left(\frac{p_*}{p_L}\right) + \left(\frac{\gamma - 1}{2\gamma}\right)}. \quad (3.62)$$

Note that as the shock strength tends to zero, the ratio  $p_*/p_L$  tends to unity and the shock speed  $S_1$  approaches the characteristic speed  $\lambda_1 = u_L - a_L$ , as expected. The particle velocity behind the left shock is

$$u_* = (1 - \rho_L/\rho_*)S_1 + u_L\rho_L/\rho_* . \quad (3.63)$$

Shock relations (3.60), (3.61) and (3.63) define a state  $(\rho_*, u_*, p_*)^T$  behind a shock for given initial conditions  $(\rho_L, u_L, p_L)^T$  ahead of the shock and a chosen shock Mach number  $M_S$ , or equivalently a shock speed  $S_1$ . The shock is associated with the 1-wave family.

### 3.2 Multi-Dimensional Euler Equations

In the previous section we analysed the one-dimensional, time-dependent Euler equations. Here we study a few basic properties of the two and three dimensional cases. In differential conservation-law form the three-dimensional equations are

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_y + \mathbf{H}(\mathbf{U})_z = \mathbf{0} , \quad (3.64)$$

with

$$\left. \begin{aligned} \mathbf{U} &= \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ E \end{bmatrix} , \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ \rho uw \\ u(E + p) \end{bmatrix} , \\ \mathbf{G} &= \begin{bmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ \rho vw \\ v(E + p) \end{bmatrix} , \quad \mathbf{H} = \begin{bmatrix} \rho w \\ \rho uw \\ \rho vw \\ \rho w^2 + p \\ w(E + p) \end{bmatrix} . \end{aligned} \right\} \quad (3.65)$$

Here  $E$  is the total energy per unit volume

$$E = \rho \left( \frac{1}{2} \mathbf{V}^2 + e \right) , \quad (3.66)$$

where  $\frac{1}{2}\mathbf{V}^2 = \frac{1}{2}\mathbf{V} \cdot \mathbf{V} = \frac{1}{2}(u^2 + v^2 + w^2)$  is the *specific kinetic energy* and  $e$  is *specific internal energy* given by a caloric equation of state (3.4).

The corresponding integral form of the conservation laws (3.64) is given by

$$\frac{d}{dt} \int \int \int_V \mathbf{U} dV + \int \int_A \mathcal{H} \cdot \mathbf{n} dA = \mathbf{0} , \quad (3.67)$$

where  $V$  is a *control volume*,  $A$  is the boundary of  $V$ ,  $\mathcal{H} = (\mathbf{F}, \mathbf{G}, \mathbf{H})$  is the tensor of fluxes,  $\mathbf{n}$  is the outward unit vector normal to the surface  $A$ ,  $dA$  is an area element and  $\mathcal{H} \cdot \mathbf{n} dA$  is the flux component normal to the boundary  $A$ . The conservation laws (3.67) state that the time-rate of change of  $\mathbf{U}$  inside volume  $V$  depends only on the *total flux* through the surface  $A$ , the boundary of the control volume  $V$ . Numerical methods of the finite volume type, see Sect. 16.7.3 of Chap. 16, are based on this formulation of the equations. For

details of the derivation of integral form of the conservation laws see Sects. 1.5 and 1.6.1 of Chap. 1.

In the next section we study some properties of the two-dimensional Euler equation in conservation form

### 3.2.1 Two-Dimensional Equations in Conservative Form

The two-dimensional version of the Euler equations in differential conservative form is

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_y = \mathbf{0}, \quad (3.68)$$

with

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ E \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ u(E + p) \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ v(E + p) \end{bmatrix}. \quad (3.69)$$

#### Eigenstructure

Here we find the Jacobian matrix of the  $x$ -split equations, its eigenvalues and corresponding right eigenvectors. We also study the types of characteristic fields present.

**Proposition 3.14.** *The Jacobian matrix  $\mathbf{A}(\mathbf{U})$  corresponding to the flux  $\mathbf{F}(\mathbf{U})$  is given by*

$$\mathbf{A}(\mathbf{U}) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -u^2 + \frac{1}{2}(\gamma - 1)\mathbf{V}^2 & (3 - \gamma)u & -(\gamma - 1)v & \gamma - 1 \\ -uv & v & u & 0 \\ u \left[ \frac{1}{2}(\gamma - 1)\mathbf{V}^2 - H \right] & H - (\gamma - 1)u^2 & -(\gamma - 1)uv & \gamma u \end{bmatrix}. \quad (3.70)$$

The eigenvalues of  $\mathbf{A}$  are

$$\lambda_1 = u - a, \quad \lambda_2 = \lambda_3 = u, \quad \lambda_4 = u + a, \quad (3.71)$$

with corresponding right eigenvectors

$$\left. \begin{aligned} \mathbf{K}^{(1)} &= \begin{bmatrix} 1 \\ u - a \\ v \\ H - au \end{bmatrix}, & \mathbf{K}^{(2)} &= \begin{bmatrix} 1 \\ u \\ v \\ \frac{1}{2}\mathbf{V}^2 \end{bmatrix}, \\ \mathbf{K}^{(3)} &= \begin{bmatrix} 0 \\ 0 \\ 1 \\ v \end{bmatrix}, & \mathbf{K}^{(4)} &= \begin{bmatrix} 1 \\ u + a \\ v \\ H + ua \end{bmatrix}. \end{aligned} \right\} \quad (3.72)$$

*Proof.* Exercise.

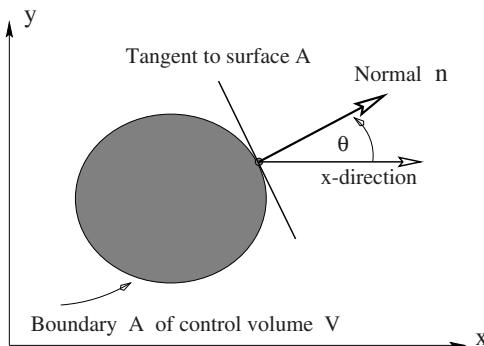
### Rotational Invariance

We next prove an important property, called the *rotational invariance* of the Euler equations. The property allows the proof of hyperbolicity in time for the two-dimensional equations (3.68)–(3.69) and can also be used for computational purposes to deal with domains that are not aligned with the Cartesian directions, see Sect. 16.7.3 of Chap. 16. We first note that outward unit vector  $\mathbf{n}$  normal to the surface  $A$  in the two-dimensional case is given by

$$\mathbf{n} \equiv (n_1, n_2) \equiv (\cos \theta, \sin \theta), \quad (3.73)$$

where  $\theta$  is the angle formed by  $x$ -axis and the normal vector  $\mathbf{n}$ ;  $\theta$  is measured in an anticlockwise manner and lies in the range  $0 \leq \theta \leq 2\pi$ . Fig. 3.4 depicts the situation. The integrand of the surface integral in (3.67) becomes

$$(\mathbf{F}, \mathbf{G}) \cdot \mathbf{n} = \cos \theta \mathbf{F}(\mathbf{U}) + \sin \theta \mathbf{G}(\mathbf{U}). \quad (3.74)$$



**Fig. 3.4.** Control volume  $V$  on  $x$ - $y$  plane; boundary of  $V$  is  $A$ , outward unit normal vector is  $\mathbf{n}$  and  $\theta$  is angle between the  $x$ -direction and  $\mathbf{n}$

**Proposition 3.15 (Rotational Invariance).** *The two-dimensional Euler equations (3.68)–(3.69) satisfy the rotational invariance property*

$$\cos \theta \mathbf{F}(\mathbf{U}) + \sin \theta \mathbf{G}(\mathbf{U}) = \mathbf{T}^{-1} \mathbf{F}(\mathbf{T}\mathbf{U}), \quad (3.75)$$

for all angles  $\theta$  and vectors  $\mathbf{U}$ . Here  $\mathbf{T} = \mathbf{T}(\theta)$  is the rotation matrix and  $\mathbf{T}^{-1}(\theta)$  is its inverse, namely

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & \sin \theta & 0 \\ 0 & -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{T}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta & 0 \\ 0 & \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (3.76)$$

*Proof.* First we calculate  $\hat{\mathbf{U}} = \mathbf{T}\mathbf{U}$ . The result is

$$\hat{\mathbf{U}} = \mathbf{T}\mathbf{U} = [\rho, \rho\hat{u}, \rho\hat{v}, E]^T,$$

with  $\hat{u} = u \cos \theta + v \sin \theta$ ,  $\hat{v} = -u \sin \theta + v \cos \theta$ . Next we compute  $\hat{\mathbf{F}} = \mathbf{F}(\hat{\mathbf{U}})$  and obtain

$$\hat{\mathbf{F}} = \mathbf{F}(\hat{\mathbf{U}}) = [\rho\hat{u}, \rho\hat{u}^2 + p, \rho\hat{u}\hat{v}, \hat{u}(E + p)]^T.$$

Now we apply  $\mathbf{T}^{-1}$  to  $\mathbf{F}(\hat{\mathbf{U}})$ . The result is easily verified to be

$$\mathbf{T}^{-1}\hat{\mathbf{F}} = \begin{bmatrix} \rho\hat{u} \\ \cos \theta (\rho\hat{u}^2 + p) - \sin \theta (\rho\hat{u}\hat{v}) \\ \sin \theta (\rho\hat{u}^2 + p) + \cos \theta (\rho\hat{u}\hat{v}) \\ \hat{u}(E + p) \end{bmatrix} = \cos \theta \mathbf{F} + \sin \theta \mathbf{G}.$$

This is clearly satisfied for the first and fourth components. Further manipulation show that it is also satisfied for the second and third flux components and the proposition is thus proved.

### Hyperbolicity in Time

Here we use the rotational invariance property of the two-dimensional time dependent Euler equations to show that the equations are *hyperbolic in time*.

**Definition 3.16 (Hyperbolicity in time).** *System (3.68)–(3.69) is hyperbolic in time if for all admissible states  $\mathbf{U}$  and real angles  $\theta$ , the matrix*

$$\mathbf{A}(\mathbf{U}, \theta) = \cos \theta \mathbf{A}(\mathbf{U}) + \sin \theta \mathbf{B}(\mathbf{U}) \quad (3.77)$$

*is diagonalisable. Here  $\mathbf{A}(\mathbf{U})$  and  $\mathbf{B}(\mathbf{U})$  are respectively the Jacobian matrices of the fluxes  $\mathbf{F}(\mathbf{U})$  and  $\mathbf{G}(\mathbf{U})$  in (3.68).*

**Proposition 3.17.** *The two-dimensional Euler equations (3.68)–(3.69) are hyperbolic in time.*

*Proof.* We want to prove that the matrix  $\mathbf{A}(\mathbf{U}, \theta)$  in (3.77) is *diagonalisable*, see Sect. 2.3.2 of Chap. 2. That is we want to prove that there exist a diagonal matrix  $\mathbf{\Lambda}(\mathbf{U}, \theta)$  and a non-singular matrix  $\mathbf{K}(\mathbf{U}, \theta)$  such that

$$\mathbf{A}(\mathbf{U}, \theta) = \mathbf{K}(\mathbf{U}, \theta) \mathbf{\Lambda}(\mathbf{U}, \theta) \mathbf{K}^{-1}(\mathbf{U}, \theta). \quad (3.78)$$

By differentiating (3.75) with respect to  $\mathbf{U}$  we have

$$\mathbf{A}(\mathbf{U}, \theta) = \cos \theta \mathbf{A}(\mathbf{U}) + \sin \theta \mathbf{B}(\mathbf{U}) = \mathbf{T}(\theta)^{-1} \mathbf{A}(\mathbf{T}(\theta)\mathbf{U}) \mathbf{T}(\theta).$$

But the matrix  $\mathbf{A}(\mathbf{U})$  is diagonalisable, it has four linearly independent eigenvectors  $\mathbf{K}^{(i)}(\mathbf{U})$  given by (3.72). Therefore we can write

$$\mathbf{A}(\mathbf{U}) = \mathbf{K}(\mathbf{U})\Lambda(\mathbf{U})\mathbf{K}^{-1}(\mathbf{U}),$$

where  $\mathbf{K}(\mathbf{U})$  is the non-singular matrix the columns of which are the right eigenvectors  $\mathbf{K}^{(i)}(\mathbf{U})$ ,  $\mathbf{K}^{-1}(\mathbf{U})$  is its inverse and  $\Lambda(\mathbf{U})$  is the diagonal matrix with the eigenvalues  $\lambda_i(\mathbf{U})$  given by (3.71) as the diagonal entries. Then we have

$$\begin{aligned}\mathbf{A}(\mathbf{U}, \theta) &= \mathbf{T}(\theta)^{-1} \{ \mathbf{K}(\mathbf{T}(\theta)\mathbf{U}) \Lambda(\mathbf{T}(\theta)\mathbf{U}) \mathbf{K}^{-1}(\mathbf{T}(\theta)\mathbf{U}) \} \mathbf{T}(\theta) \\ &= \{ \mathbf{T}(\theta)^{-1} \mathbf{K}(\mathbf{T}(\theta)\mathbf{U}) \} \Lambda(\mathbf{T}(\theta)\mathbf{U}) \{ \mathbf{T}(\theta)^{-1} \mathbf{K}(\mathbf{T}(\theta)\mathbf{U}) \}^{-1}.\end{aligned}$$

Hence the condition for hyperbolicity holds by taking

$$\mathbf{K}(\mathbf{U}, \theta) = \mathbf{T}^{-1}(\theta) \mathbf{K}(\mathbf{T}(\theta)\mathbf{U}), \quad \Lambda(\mathbf{U}, \theta) = \Lambda(\mathbf{T}(\theta)\mathbf{U}).$$

We have thus proved that the time-dependent, two dimensional Euler equations are hyperbolic in time, as claimed.

## Characteristic Fields

Next we analyse the characteristic fields associated with the four eigenvectors given by (3.72).

**Proposition 3.18 (Types of Characteristic Fields).** *For  $i = 1$  and  $i = 4$  the  $\mathbf{K}^{(i)}(\mathbf{U})$  characteristic fields are genuinely non-linear, while for  $i = 2$  and  $i = 3$  they are linearly degenerate.*

*Proof.* The proof that the fields  $i = 2$  and  $i = 3$  are linearly degenerate is trivial. Clearly

$$\nabla \lambda_2 = \nabla \lambda_3 = (-u/\rho, 1/\rho, 0, 0).$$

By inspecting  $\mathbf{K}^{(2)}(\mathbf{U})$  and  $\mathbf{K}^{(3)}(\mathbf{U})$  it is obvious that

$$\nabla \lambda_2 \cdot \mathbf{K}^{(2)}(\mathbf{U}) = \nabla \lambda_3 \cdot \mathbf{K}^{(3)}(\mathbf{U}) = 0$$

and therefore the 2 and 3 characteristic fields are linearly degenerate as claimed. The proof for  $i = 1, 4$  involves some algebra. The result is

$$\nabla \lambda_1 \cdot \mathbf{K}^{(1)}(\mathbf{U}) = -\frac{(\gamma+1)a}{2\rho} \neq 0, \quad \nabla \lambda_4 \cdot \mathbf{K}^{(4)}(\mathbf{U}) = \frac{(\gamma+1)a}{2\rho} \neq 0$$

and thus the 1 and 4 characteristic fields are *genuinely non-linear* as claimed.

In the context of the Riemann problem we shall see that across the 2 and 3 waves both pressure  $p$  and *normal* velocity component  $u$  are constant. The 2 field is associated with a contact discontinuity, across which density jumps discontinuously. The 3 field is associated with a shear wave across which the *tangential* velocity component jumps discontinuously. The 1 and 4 characteristic fields are associated with shock waves and rarefaction waves.

### 3.2.2 Three-Dimensional Equations in Conservative Form

Here we extend previous results proved for the two-dimensional equations, to the time-dependent three dimensional Euler equations. Proofs are omitted, they involve elementary but tedious algebra.

#### Eigenstructure

The Jacobian matrix  $\mathbf{A}$  corresponding to the flux  $\mathbf{F}(\mathbf{U})$  in (3.64) is given by

$$\mathbf{A} = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ \hat{\gamma}H - u^2 - a^2 & (3 - \gamma)u & -\hat{\gamma}v & -\hat{\gamma}w & \hat{\gamma} \\ -uv & v & u & 0 & 0 \\ -uw & w & 0 & u & 0 \\ \frac{1}{2}u[(\gamma - 3)H - a^2] & H - \hat{\gamma}u^2 & -\hat{\gamma}uv & -\hat{\gamma}uw & \gamma u \end{bmatrix}, \quad (3.79)$$

where

$$H = (E + p)/\rho = \frac{1}{2}\mathbf{V}^2 + \frac{a^2}{(\gamma - 1)}, \quad \mathbf{V}^2 = u^2 + v^2 + w^2, \quad \hat{\gamma} = \gamma - 1. \quad (3.80)$$

The  $x$ -split one-dimensional system is hyperbolic with real eigenvalues

$$\lambda_1 = u - a, \quad \lambda_2 = \lambda_3 = \lambda_4 = u, \quad \lambda_5 = u + a. \quad (3.81)$$

The matrix of corresponding right eigenvectors is

$$\mathbf{K} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 \\ u - a & u & 0 & 0 & u + a \\ v & v & 1 & 0 & v \\ w & w & 0 & 1 & w \\ H - ua & \frac{1}{2}\mathbf{V}^2 & v & w & H + ua \end{bmatrix}. \quad (3.82)$$

We also give the expression for the inverse matrix of  $\mathbf{K}$ , namely

$$\mathbf{K}^{-1} = \frac{(\gamma - 1)}{2a^2} \begin{bmatrix} H + \frac{a}{\hat{\gamma}}(u - a) & -(u + \frac{a}{\hat{\gamma}}) & -v & -w & 1 \\ -2H + \frac{4}{\hat{\gamma}}a^2 & 2u & 2v & 2w & -2 \\ -\frac{2va^2}{\hat{\gamma}} & 0 & \frac{2a^2}{\hat{\gamma}} & 0 & 0 \\ -\frac{2wa^2}{\hat{\gamma}} & 0 & 0 & \frac{2a^2}{\hat{\gamma}} & 0 \\ H - \frac{a}{\hat{\gamma}}(u + a) & -u + \frac{a}{\hat{\gamma}} & -v & -w & 1 \end{bmatrix}. \quad (3.83)$$

#### Rotational Invariance

We now state the rotational invariance property for the three-dimensional case.

**Proposition 3.19.** *The time-dependent three dimensional Euler equations are rotationally invariant, that is they satisfy*

$$\cos \theta^{(y)} \cos \theta^{(z)} \mathbf{F}(\mathbf{U}) + \cos \theta^{(y)} \sin \theta^{(z)} \mathbf{G}(\mathbf{U}) + \sin \theta^{(y)} \mathbf{H}(\mathbf{U}) = \mathbf{T}^{-1} \mathbf{F}(\mathbf{T}\mathbf{U}) , \quad (3.84)$$

for all angles  $\theta^{(y)}$ ,  $\theta^{(z)}$  and vectors  $\mathbf{U}$ . Here  $\mathbf{T} = \mathbf{T}(\theta^{(y)}, \theta^{(z)})$  is the rotation matrix

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \cos \theta^{(y)} \cos \theta^{(z)} & \cos \theta^{(y)} \sin \theta^{(z)} & \sin \theta^{(y)} 0 \\ 0 & -\sin \theta^{(z)} & \cos \theta^{(z)} & 0 & 0 \\ 0 & -\sin \theta^{(y)} \cos \theta^{(z)} & -\sin \theta^{(y)} \sin \theta^{(z)} & \cos \theta^{(y)} 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} , \quad (3.85)$$

and is the product of two rotation matrices, namely

$$\mathbf{T} = \mathbf{T}(\theta^{(y)}, \theta^{(z)}) = \mathbf{T}^{(y)} \mathbf{T}^{(z)} , \quad (3.86)$$

with

$$\left. \begin{aligned} \mathbf{T}^{(y)} \equiv \mathbf{T}^{(y)}(\theta^{(y)}) &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \cos \theta^{(y)} & 0 & \sin \theta^{(y)} 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -\sin \theta^{(y)} 0 & \cos \theta^{(y)} 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} , \\ \mathbf{T}^{(z)} \equiv \mathbf{T}^{(z)}(\theta^{(z)}) &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \cos \theta^{(z)} & \sin \theta^{(z)} 0 & 0 \\ 0 & -\sin \theta^{(z)} & \cos \theta^{(z)} 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} . \end{aligned} \right\} \quad (3.87)$$

More details of the rotational invariance and related properties of the three-dimensional Euler equations are found in Billett and Toro [64].

### 3.2.3 Three-Dimensional Primitive Variable Formulation

As done for the one-dimensional Euler equations, we can express the two and three dimensional equations in terms of primitive variables.

**Proposition 3.20.** *The three-dimensional, time-dependent Euler equations can be written in terms of the primitive variables  $\mathbf{W} = (\rho, u, v, w, p)^T$  as*

$$\left. \begin{aligned} \rho_t + u\rho_x + v\rho_y + w\rho_z + \rho(u_x + v_y + w_z) &= 0 , \\ u_t + uu_x + vu_y + wu_z + \frac{1}{\rho}p_x &= 0 , \\ v_t + uv_x + vv_y + wv_z + \frac{1}{\rho}p_y &= 0 , \\ w_t + uw_x + vw_y + ww_z + \frac{1}{\rho}p_z &= 0 , \\ p_t + up_x + vp_y + wp_z + \rho a^2(u_x + v_y + w_z) &= 0 . \end{aligned} \right\} \quad (3.88)$$

*Proof.* To prove this result one follows the same steps as for the one-dimensional case leading to equations (3.14)–(3.16).

Equations (3.88) can be written in quasi-linear form as

$$\mathbf{W}_t + \mathbf{A}(\mathbf{W})\mathbf{W}_x + \mathbf{B}(\mathbf{W})\mathbf{W}_y + \mathbf{C}(\mathbf{W})\mathbf{W}_z = \mathbf{0} , \quad (3.89)$$

where the coefficient matrices  $\mathbf{A}(\mathbf{W})$ ,  $\mathbf{B}(\mathbf{W})$  and  $\mathbf{C}(\mathbf{W})$  are given by

$$\mathbf{A}(\mathbf{W}) = \begin{bmatrix} u & \rho & 0 & 0 & 0 \\ 0 & u & 0 & 0 & 1/\rho \\ 0 & 0 & u & 0 & 0 \\ 0 & 0 & 0 & u & 0 \\ 0 & \rho a^2 & 0 & 0 & u \end{bmatrix} , \quad (3.90)$$

$$\mathbf{B}(\mathbf{W}) = \begin{bmatrix} v & \rho & 0 & 0 & 0 \\ 0 & v & 0 & 0 & 0 \\ 0 & 0 & v & 0 & 1/\rho \\ 0 & 0 & 0 & v & 0 \\ 0 & 0 & \rho a^2 & 0 & v \end{bmatrix} , \quad (3.91)$$

$$\mathbf{C}(\mathbf{W}) = \begin{bmatrix} w & \rho & 0 & 0 & 0 \\ 0 & w & 0 & 0 & 0 \\ 0 & 0 & w & 0 & 0 \\ 0 & 0 & 0 & w & 1/\rho \\ 0 & 0 & 0 & \rho a^2 & w \end{bmatrix} . \quad (3.92)$$

**Proposition 3.21.** *The eigenvalues of the coefficient matrix  $\mathbf{A}(\mathbf{W})$  in (3.90) are given by*

$$\lambda_1 = u - a , \lambda_2 = \lambda_3 = \lambda_4 = u , \lambda_5 = u + a . \quad (3.93)$$

with corresponding right eigenvectors

$$\left. \begin{aligned} \mathbf{K}^{(1)} &= \begin{bmatrix} \rho \\ -a \\ 0 \\ 0 \\ \rho a^2 \end{bmatrix}, \quad \mathbf{K}^{(2)} = \begin{bmatrix} 1 \\ 0 \\ v \\ w \\ 0 \end{bmatrix}, \quad \mathbf{K}^{(3)} = \begin{bmatrix} \rho \\ 0 \\ 1 \\ w \\ 0 \end{bmatrix}, \\ \mathbf{K}^{(4)} &= \begin{bmatrix} \rho \\ 0 \\ v \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{K}^{(5)} = \begin{bmatrix} \rho \\ a \\ 0 \\ 0 \\ \rho a^2 \end{bmatrix}. \end{aligned} \right\} \quad (3.94)$$

*Proof.* The proof involves the usual algebraic steps for finding eigenvalues and eigenvectors. See Sect. 2.1 of Chap. 2.

### 3.2.4 The Split Three-Dimensional Riemann Problem

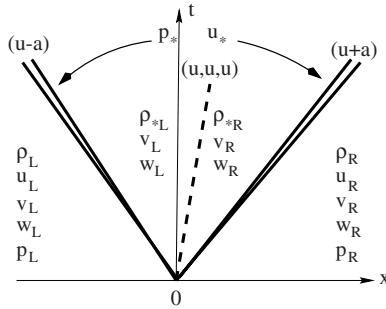
When solving numerically the two or three dimensional Euler equations by most methods of the upwind type in current use, one requires the solution of *split* Riemann problems. The  $x$ -split, three-dimensional Riemann problem is the IVP

$$\left. \begin{aligned} \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x &= \mathbf{0}, \\ \mathbf{U}(x, 0) &= \mathbf{U}^{(0)}(x) = \begin{cases} \mathbf{U}_L & \text{if } x < 0, \\ \mathbf{U}_R & \text{if } x > 0, \end{cases} \end{aligned} \right\} \quad (3.95)$$

where

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ E \end{bmatrix}, \quad \mathbf{F}(\mathbf{U}) = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ \rho uw \\ u(E + p) \end{bmatrix}. \quad (3.96)$$

The structure of the similarity solution is shown in Fig. 3.5 and is almost identical to that for the one-dimensional case shown in Fig. 3.1. Both pressure and *normal* particle velocity  $u$  are constant in the *Star Region*, across the middle wave. There are two new characteristic fields associated with  $\lambda_3 = u$  and  $\lambda_4 = u$ , arising from the multiplicity 3 of the eigenvalue  $u$ ; these correspond to two *shear waves* across which the respective tangential velocity components  $v$  and  $w$  change discontinuously. For the two-dimensional case we proved in Sect. 3.2.1 that the  $\lambda_3$ -field is linearly degenerate. This result is also true for the  $\lambda_4$ -field in three dimensions. The 1 and 5 characteristic fields are genuinely non-linear and are associated with rarefactions or shock waves, just as in the one-dimensional case. By inspecting the eigenvectors  $\mathbf{K}^{(1)}$  and  $\mathbf{K}^{(5)}$  in (3.94) we see immediately that the Generalised Riemann Invariants across 1 and 5 rarefaction waves give no change in the tangential velocity components  $v$  and  $w$  across these waves, see Fig. 3.5. In fact this is



**Fig. 3.5.** Structure of the solution of the three-dimensional split Riemann problem

also true when these waves are shock waves. Consider a right shock wave of speed  $S$  associated with the 5 field. By transforming to a frame of reference in which the shock speed is zero and applying the Rankine–Hugoniot conditions we obtain the same relations (3.42)–(3.44) as in the one-dimensional case plus two extra relations involving  $v$  and  $w$ . The three relevant relations are

$$\rho_*(u_* - S) = \rho_R(u_R - S), \quad (3.97)$$

$$\rho_*(u_* - S)(v_* - S) = \rho_R(u_R - S)(v_R - S), \quad (3.98)$$

$$\rho_*(u_* - S)(w_* - S) = \rho_R(u_R - S)(w_R - S). \quad (3.99)$$

Application of the shock condition (3.97) into equations (3.98) and (3.99) gives directly  $v_* = v_R$  and  $w_* = w_R$ . A similar analysis for a left shock wave gives an equivalent result. Hence the tangential velocity components  $v$  and  $w$  remain constant across the non-linear waves 1 and 5, irrespective of their type.

Therefore finding the solution of the Riemann problem for the split three-dimensional equations is fundamentally the same as finding the solution for the corresponding one-dimensional Riemann problem. The solution for the extra variables  $v$  and  $w$  could not be simpler: it consists of single jump discontinuities across the shear waves from the values  $v_L, w_L$  on the left data state to the values  $v_R, w_R$  on the right data state. This simple behaviour of the tangential velocity components in the solution of split Riemann problems is sometimes incorrectly modelled by some *approximate* Riemann solvers.

### 3.3 Conservative Versus Non–Conservative Formulations

The specific purpose of this section is first to make the point that under the assumption of smooth solutions, *conservative* and non–conservative formulations are not unique. It is vitally important to scrutinise the conservative formulations carefully, as these may be conservative purely in a mathematical

sense. The key question is to see what the *conserved quantities* are in the formulation and whether the conservation statements they imply make *physical sense*. The second point of interest here is to make the reader aware of the fact that in the presence of shock waves, formulations that are *conservative purely in a mathematical sense* will produce wrong shock speeds and thus wrong solutions. We illustrate these points through the one-dimensional shallow water equations, see Sect. 1.6.3 of Chap. 1,

$$\begin{bmatrix} \phi \\ \phi u \end{bmatrix}_t + \begin{bmatrix} \phi u \\ \phi u^2 + \frac{1}{2}\phi^2 \end{bmatrix}_x = \mathbf{0}. \quad (3.100)$$

They express the physical laws of conservation of mass and momentum. Under the assumption of smooth solutions we can expand derivatives so as to write the equations in primitive-variable form

$$\phi_t + u\phi_x + \phi u_x = 0, \quad (3.101)$$

$$u_t + uu_x + \phi_x = 0. \quad (3.102)$$

It is tempting to derive new conservation–law forms of the shallow water equations starting from equations (3.101)–(3.102). One possibility is to keep the mass equation as in (3.101) and re–write the momentum equation (3.102) as

$$u_t + \left(\frac{1}{2}u^2 + \phi\right)_x = 0. \quad (3.103)$$

Now we have an alternative *conservative* form of the shallow water equations, namely

$$\begin{bmatrix} \phi \\ u \end{bmatrix}_t + \begin{bmatrix} \phi u \\ \frac{1}{2}u^2 + \phi \end{bmatrix}_x = \mathbf{0}. \quad (3.104)$$

Mathematically, see Chap. 2, this is a system of conservation laws. It expresses conservation of mass, as in (3.101), and *conservation of particle speed u*. Physically, this second conservation law does not make sense. A critical question is this : *can we use the conservation–law form (3.104) for the shallow water equations.* The anticipated answer is : *yes we can, if and only if solutions are smooth.* In the presence of shock waves formulations (3.100) and (3.104) lead to different solutions, as we now demonstrate.

Without loss of generality we consider a right facing shock wave in which the state ahead of the shock is given by the variables  $\phi_R, u_R$ .

**Proposition 3.22.** *A right-facing shock wave solution of (3.100) has shock speed*

$$\left. \begin{aligned} S &= u_R + Q/\phi_R, \\ Q &= \left[\frac{1}{2}(\phi_* + \phi_R)\phi_*\phi_R\right]^{\frac{1}{2}}, \end{aligned} \right\} \quad (3.105)$$

while a right-facing shock wave solution of (3.104) has speed

$$\left. \begin{aligned} \hat{S} &= u_R + \hat{Q}/\phi_R, \\ \hat{Q} &= \left[\frac{2}{\phi_* + \phi_R}\right]^{\frac{1}{2}}\phi_*\phi_R. \end{aligned} \right\} \quad (3.106)$$

*Proof.* This is left to the reader as an exercise. Use contents of Chap. 2 and those of Sect. 3.1.3 of Chap. 3.

*Remark 3.23.* Clearly the shock speeds  $S$  and  $\hat{S}$  are equivalent only when  $\phi_* \equiv \phi_R$ , that is when the shock wave is trivial. In general

$$\hat{S} \leq S \quad (3.107)$$

and thus shock solutions of the *new* (incorrect) conservation laws (3.104) are slower than shock solutions of the *conventional* (correct) conservation laws (3.100). Note also that the conservative form (3.104) is non-unique.

Consider now the isothermal equations of Gas Dynamics, see Sect. 1.6.2 of Chap. 1. In conservation-law form these equations read

$$\begin{bmatrix} \rho \\ \rho u \end{bmatrix}_t + \begin{bmatrix} \rho u \\ \rho u^2 + a^2 \rho \end{bmatrix}_x = \mathbf{0}, \quad (3.108)$$

where the sound speed  $a$  is constant. These conservation laws state that mass and momentum are conserved, which is in accord with the laws of conservation of mass and momentum studied in Chap. 1. Let us now assume that solutions are sufficiently smooth so that partial derivatives exist; we expand derivatives and after some algebraic manipulations obtain the primitive-variable formulation

$$\rho_t + u\rho_x + \rho u_x = 0, \quad (3.109)$$

$$u_t + uu_x + \frac{a^2}{\rho} \rho_x = 0. \quad (3.110)$$

This is a perfectly acceptable formulation, valid for smooth flows.

*New conservation laws* can be constructed, starting from the primitive formulation (3.109)–(3.110) above. One such possible system of conservation laws is

$$\begin{bmatrix} \rho \\ u \end{bmatrix}_t + \begin{bmatrix} \rho u \\ \frac{1}{2}u^2 + a^2 \ln \rho \end{bmatrix}_x = \mathbf{0}. \quad (3.111)$$

Mathematically, these equations are a set of conservation laws, see Sects. 2.1 and 2.4 of Chap. 2. Physically however, they are useless, they state that mass and *velocity* are conserved !

**Exercise 3.24.** Using the contents of Sect. 3.1.3 for isolated shock waves, compare the shock solutions of the two conservative formulations (3.108) and (3.111). Which gives the fastest shock ? Find other *conservative* formulations corresponding to (3.109)–(3.110).

---

## The Riemann Problem for the Euler Equations

In his classical paper of 1959, Godunov [216] presented a conservative extension of the first-order upwind scheme of Courant, Isaacson and Rees [144] to non-linear systems of hyperbolic conservation laws. The key ingredient of the scheme is the solution of the Riemann problem. The purpose of this chapter is to provide a detailed presentation of the complete, exact solution to the Riemann problem for the one-dimensional, time-dependent Euler equations for ideal and covolume gases, including vacuum conditions. The methodology can then be applied to other hyperbolic systems.

The exact solution of the Riemann problem is useful in a number of ways. First, it represents the solution to a system of hyperbolic conservation laws subject to the simplest, non-trivial, initial conditions; in spite of the simplicity of the initial data the solution of the Riemann problem contains the fundamental physical and mathematical character of the relevant set of conservation laws. The solution of the general IVP may be seen as resulting from non-linear superposition of solutions of local Riemann problems [212]. In the case of the Euler equations the Riemann problem includes the so called *shock-tube problem*, a basic physical problem in Gas Dynamics. For a detailed discussion on the shock-tube problem the reader is referred to the book by Courant and Friedrichs [143]. The exact Riemann problem solution is also an invaluable reference solution that is useful in assessing the performance of numerical methods and to check the correctness of programs in the early stages of development. The Riemann problem solution, exact or approximate, can also be used locally in the method of Godunov and high-order extensions of it; this is the main role we assign to the Riemann problem here. A detailed knowledge of the exact solution is also fundamental when utilising, assessing and developing approximate Riemann solvers.

There is no exact closed-form solution to the Riemann problem for the Euler equations, not even for ideal gases; in fact not even for much simpler models such as the isentropic and isothermal equations. However, it is possible to devise *iterative* schemes whereby the solution can be computed numerically to any desired, practical, degree of accuracy. Key issues in designing an exact

Riemann solver are: the variables selected, the equations used, the number of equations and the technique for the iterative procedure, the initial guess and the handling of unphysical iterates, such as negative pressure. Godunov is credited with the first exact Riemann solver for the Euler equations [216]. By today's standards Godunov's first Riemann solver is cumbersome and computationally inefficient. Later, Godunov [218] proposed a second exact Riemann solver. Distinct features of this solver are: the equations used are simpler, the variables selected are more convenient from the computational point of view and the iterative procedure is rather sophisticated. Much of the work that followed contains the fundamental features of Godunov's second Riemann solver. Chorin [110], independently, produced improvements to Godunov's first Riemann solver. In 1979, van Leer [559] produced another improvement to Godunov's first Riemann solver resulting in a scheme that is similar to Godunov's second solver. Smoller [451] proposed a rather different approach; later, Dutt [179] produced a practical implementation of the scheme. Gottlieb and Groth [222] presented another Riemann solver for ideal gases; of the schemes they tested, theirs is shown to be the most efficient. Toro [498] presented an exact Riemann solver for ideal and covolume gases of comparable efficiency to that of Gottlieb and Groth. More recently, Schleicher [430] and Pike [384] have also presented new exact Riemann solvers which appear to be the fastest to date. For gases obeying a general equation of state the reader is referred to the pioneering work of Colella and Glaz [135]. Other relevant publications are that of Menikoff and Plohr [349] and that of Saurel, Larini and Loraud [429].

In this chapter we present a solution procedure of the Riemann problem for the Euler equations for both ideal and covolume gases. The methodology is presented in great detail for the ideal gas case. We then address the issue of *vacuum* and provide an exact solution for the three cases that can occur. Particular emphasis is given to the *sampling* of the solution; this will be useful to provide the complete solution and to utilise it in numerical methods such as the Godunov method [216] and Glimm's method or Random Choice Method [212], [110]. The necessary background for this chapter is found in Chaps. 1, 2 and 3.

## 4.1 Solution Strategy

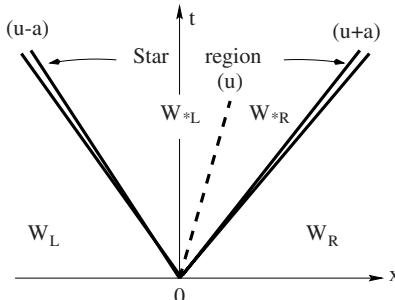
The Riemann problem for the one-dimensional time-dependent Euler equations is the Initial Value Problem (IVP) for the conservation laws

$$\left. \begin{aligned} \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x &= \mathbf{0}, \\ \mathbf{U} &= \begin{bmatrix} \rho \\ \rho u \\ E \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ u(E + p) \end{bmatrix}, \end{aligned} \right\} \quad (4.1)$$

with initial conditions (IC)

$$\mathbf{U}(x, 0) = \mathbf{U}^{(0)}(x) = \begin{cases} \mathbf{U}_L & \text{if } x < 0 \\ \mathbf{U}_R & \text{if } x > 0. \end{cases} \quad (4.2)$$

The domain of interest in the  $x-t$  plane are points  $(x, t)$  with  $-\infty < x < \infty$  and  $t > 0$ . In practice one lets  $x$  vary in a finite interval  $[x_L, x_R]$  around the point  $x = 0$ . In solving the Riemann problem we shall frequently make use of the vector  $\mathbf{W} = (\rho, u, p)^T$  of primitive variables, rather than the vector  $\mathbf{U}$  of conserved variables, where  $\rho$  is density,  $u$  is particle velocity and  $p$  is pressure. The Riemann problem (4.1)–(4.2) is the simplest, non-trivial, IVP for (4.1). Data consists of just two constant states, which in terms of primitive variables are  $\mathbf{W}_L = (\rho_L, u_L, p_L)^T$  to the left of  $x = 0$  and  $\mathbf{W}_R = (\rho_R, u_R, p_R)^T$  to the right of  $x = 0$ , separated by a discontinuity at  $x = 0$ . Physically, in the context of the Euler equations, the Riemann problem is a slight generalisation of the so called *shock-tube problem*: two stationary gases ( $u_L = u_R = 0$ ) in a tube are separated by a diaphragm. The rupture of the diaphragm generates a nearly *centred* wave system that typically consists of a rarefaction wave, a contact discontinuity and a shock wave. This physical problem is reasonably well approximated by solving the shock-tube problem for the Euler equations. In the Riemann problem the particle speeds  $u_L$  and  $u_R$  are allowed to be non-zero, but the structure of the solution is the same as that of the shock-tube problem. In general, given



**Fig. 4.1.** Structure of the solution of the Riemann problem on the  $x-t$  plane for the one-dimensional time-dependent Euler equations

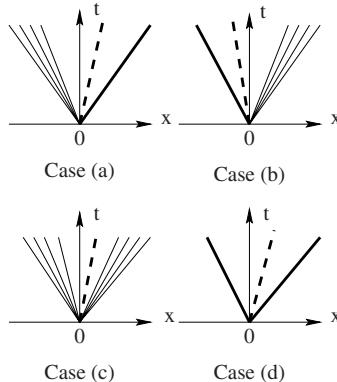
the conservation equations (4.1) for the dynamics, it is left to the statements about the material, the equation of state, to determine not only the structure of the solution of the Riemann problem but also the mathematical character of the equations. In this chapter we restrict our attention to *ideal gases* obeying the caloric Equation of State (EOS)

$$e = \frac{p}{(\gamma - 1)\rho}, \quad (4.3)$$

and *covolume gases* obeying

$$e = \frac{p(1 - b\rho)}{(\gamma - 1)\rho}, \quad (4.4)$$

where  $\gamma$  is the ratio of specific heats, a constant, and  $b$  is the covolume, also a constant. See Sects. 1.2.4 and 1.2.5 of Chap. 1. For the case in which no



**Fig. 4.2.** Possible wave patterns in the solution of the Riemann problem: (a) left rarefaction, contact, right shock (b) left shock, contact, right rarefaction (c) left rarefaction, contact, right rarefaction (d) left shock, contact, right shock

*vacuum* is present the exact solution of the Riemann problem (4.1), (4.2) has three waves, which are *associated* with the eigenvalues  $\lambda_1 = u - a$ ,  $\lambda_2 = u$  and  $\lambda_3 = u + a$ ; see Fig. 4.1. Note that the speeds of these waves are not, in general, the characteristics speeds given by the eigenvalues. The three waves separate four constant states, which from left to right are:  $\mathbf{W}_L$  (data on the left hand side),  $\mathbf{W}_{*L}$ ,  $\mathbf{W}_{*R}$  and  $\mathbf{W}_R$  (data on the right hand side).

The unknown region between the left and right waves, the *Star Region*, is divided by the middle wave into the two subregions *Star Left* ( $\mathbf{W}_{*L}$ ) and *Star Right* ( $\mathbf{W}_{*R}$ ). As seen in Sect. 3.1.3 of Chap. 3, the middle wave is always a contact discontinuity while the left and right (non-linear) waves are either shock or rarefaction waves. Therefore, according to the type of non-linear waves there can be four possible wave patterns, which are shown in Fig. 4.2. There are two possible variations of these, namely when the left or right non-linear wave is a *sonic rarefaction wave*; these two cases are only of interest when utilising the solution of the Riemann problem in Godunov-type methods. For the purpose of constructing a solution scheme for the Riemann problem it is sufficient to consider the four patterns of Fig. 4.2.

An analysis based on the eigenstructure of the Euler equations, Sect. 3.1.3 Chap. 3, reveals that both pressure  $p_*$  and particle velocity  $u_*$  between the left and right waves are constant, while the density takes on the two constant values  $\rho_{*L}$  and  $\rho_{*R}$ . Here we present a solution procedure which makes use of the constancy of pressure and particle velocity in the *Star Region* to derive a

single, algebraic non-linear equation for pressure  $p_*$ . In summary, the main physical quantities sought are  $p_*$ ,  $u_*$ ,  $\rho_{*L}$  and  $\rho_{*R}$ .

## 4.2 Equations for Pressure and Particle Velocity

Here we establish equations and solution strategies for computing the pressure  $p_*$  and the particle velocity  $u_*$  in the *Star Region*.

**Proposition 4.1 (solution for  $p_*$  and  $u_*$ ).** *The solution for pressure  $p_*$  of the Riemann problem (4.1), (4.2) with the ideal gas Equation of State (4.3) is given by the root of the algebraic equation*

$$f(p, \mathbf{W}_L, \mathbf{W}_R) = f_L(p, \mathbf{W}_L) + f_R(p, \mathbf{W}_R) + \Delta u = 0 , \quad \Delta u \equiv u_R - u_L , \quad (4.5)$$

where the function  $f_L$  is given by

$$f_L(p, \mathbf{W}_L) = \begin{cases} (p - p_L) \left[ \frac{A_L}{p + B_L} \right]^{\frac{1}{2}} & \text{if } p > p_L \text{ (shock)} , \\ \frac{2a_L}{(\gamma-1)} \left[ \left( \frac{p}{p_L} \right)^{\frac{\gamma-1}{2\gamma}} - 1 \right] & \text{if } p \leq p_L \text{ (rarefaction)} , \end{cases} \quad (4.6)$$

the function  $f_R$  is given by

$$f_R(p, \mathbf{W}_R) = \begin{cases} (p - p_R) \left[ \frac{A_R}{p + B_R} \right]^{\frac{1}{2}} & \text{if } p > p_R \text{ (shock)} , \\ \frac{2a_R}{(\gamma-1)} \left[ \left( \frac{p}{p_R} \right)^{\frac{\gamma-1}{2\gamma}} - 1 \right] & \text{if } p \leq p_R \text{ (rarefaction)} , \end{cases} \quad (4.7)$$

and the data-dependent constants  $A_L$ ,  $B_L$ ,  $A_R$ ,  $B_R$  are given by

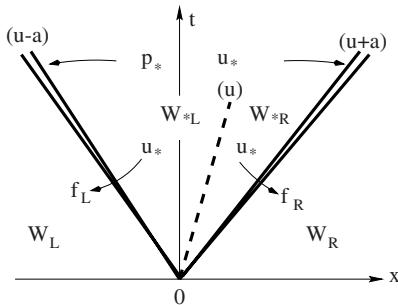
$$\left. \begin{aligned} A_L &= \frac{2}{(\gamma+1)\rho_L} , & B_L &= \frac{(\gamma-1)}{(\gamma+1)}p_L , \\ A_R &= \frac{2}{(\gamma+1)\rho_R} , & B_R &= \frac{(\gamma-1)}{(\gamma+1)}p_R . \end{aligned} \right\} \quad (4.8)$$

The solution for the particle velocity  $u_*$  in the Star Region is

$$u_* = \frac{1}{2}(u_L + u_R) + \frac{1}{2}[f_R(p_*) - f_L(p_*)] . \quad (4.9)$$

*Remark 4.2.* Before proceeding to prove the above statements we make some useful remarks. Once (4.5) is solved for  $p_*$  the solution for  $u_*$  follows as in (4.9) and the remaining unknowns are found by using standard gas dynamics relations studied in Chap. 3. The function  $f_L$  governs relations across the left non-linear wave and serves to connect the unknown particle speed  $u_*$  to the known state  $\mathbf{W}_L$  on the left side, see Fig. 4.3; the relations depend on the type

of wave (shock or rarefaction). The arguments of  $f_L$  are the pressure  $p$  and the data state  $\mathbf{W}_L$ . Similarly, the function  $f_R$  governs relations across the right wave and connects the unknown  $u_*$  to the right data state  $\mathbf{W}_R$ ; its arguments are  $p$  and  $\mathbf{W}_R$ . For convenience we shall often omit the data arguments of the functions  $f$ ,  $f_L$  and  $f_R$ . The sought pressure  $p_*$  in the *Star Region* is the root of the algebraic equation (4.5),  $f(p) = 0$ . A detailed analysis of the pressure function  $f(p)$  reveals a particularly simple behaviour and that for *physically relevant data* there exists a unique solution to the equation  $f(p) = 0$ .



**Fig. 4.3.** Strategy for solving the Riemann problem via a pressure function. The particle velocity is connected to data on the left and right via functions  $f_L$  and  $f_R$

*Proof.* Here we derive expressions for  $f_L$  and  $f_R$  in equation (4.5). We do this by considering each non-linear wave separately.

#### 4.2.1 Function $f_L$ for a Left Shock

We assume the left wave is a shock moving with speed  $S_L$  as shown in Fig. 4.4a; pre-shock values are  $\rho_L$ ,  $u_L$  and  $p_L$  and post-shock values are  $\rho_{*L}$ ,  $u_*$  and  $p_*$ .

As done in Sect. 3.1.3 of Chap. 3, we transform the equations to a frame of reference moving with the shock, as depicted in Fig. 4.4b. In the new frame the shock speed is zero and the *relative velocities* are

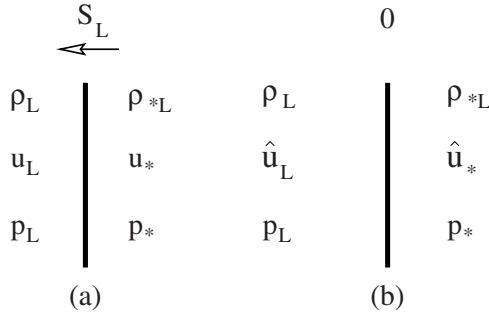
$$\hat{u}_L = u_L - S_L, \quad \hat{u}_* = u_* - S_L. \quad (4.10)$$

The Rankine–Hugoniot Conditions, see Sect. 3.1.3 of Chap. 3, give

$$\rho_L \hat{u}_L = \rho_{*L} \hat{u}_*, \quad (4.11)$$

$$\rho_L \hat{u}_L^2 + p_L = \rho_{*L} \hat{u}_*^2 + p_*, \quad (4.12)$$

$$\hat{u}_L (\hat{E}_L + p_L) = \hat{u}_* (\hat{E}_{*L} + p_*). \quad (4.13)$$



**Fig. 4.4.** Left wave is a shock wave of speed  $S_L$ : (a) stationary frame, shock speed is  $S_L$  (b) frame of reference moving with speed  $S_L$ , shock speed is zero

We introduce the *mass flux*  $Q_L$ , which in view of (4.11) may be written as

$$Q_L \equiv \rho_L \hat{u}_L = \rho_{*L} \hat{u}_* . \quad (4.14)$$

From equation (4.12)

$$(\rho_L \hat{u}_L) \hat{u}_L + p_L = (\rho_{*L} \hat{u}_*) \hat{u}_* + p_* .$$

Use of (4.14) and solving for  $Q_L$  gives

$$Q_L = -\frac{p_* - p_L}{\hat{u}_* - \hat{u}_L} . \quad (4.15)$$

But from equation (4.10)  $\hat{u}_L - \hat{u}_* = u_L - u_*$  and so  $Q_L$  becomes

$$Q_L = -\frac{p_* - p_L}{u_* - u_L} , \quad (4.16)$$

from which we obtain

$$u_* = u_L - \frac{(p_* - p_L)}{Q_L} . \quad (4.17)$$

We are now close to having related  $u_*$  to data on the left hand side. We seek to express the right hand side of (4.17) purely in terms of  $p_*$  and  $\mathbf{W}_L$ , which means that we need to express  $Q_L$  as a function of  $p_*$  and the data on the left hand side. We substitute the relations

$$\hat{u}_L = \frac{Q_L}{\rho_L} , \quad \hat{u}_* = \frac{Q_L}{\rho_{*L}} ,$$

obtained from (4.14) into equation (4.15) to produce

$$Q_L^2 = -\frac{\frac{p_* - p_L}{u_* - u_L}}{\frac{1}{\rho_{*L}} - \frac{1}{\rho_L}} . \quad (4.18)$$

As seen in Sect. 3.1.3 of Chap. 3, the density  $\rho_{*L}$  is related to the pressure  $p_*$  behind the left shock via

$$\rho_{*L} = \rho_L \left[ \frac{\left(\frac{\gamma-1}{\gamma+1}\right) + \left(\frac{p_*}{p_L}\right)}{\left(\frac{\gamma-1}{\gamma+1}\right) \left(\frac{p_*}{p_L}\right) + 1} \right]. \quad (4.19)$$

Substitution of  $\rho_{*L}$  into (4.18) yields

$$Q_L = \left[ \frac{p_* + B_L}{A_L} \right]^{\frac{1}{2}}, \quad (4.20)$$

which in turn reduces (4.17) to

$$u_* = u_L - f_L(p_*, \mathbf{W}_L), \quad (4.21)$$

with

$$f_L(p_*, \mathbf{W}_L) = (p_* - p_L) \left[ \frac{A_L}{p_* + B_L} \right]^{\frac{1}{2}}$$

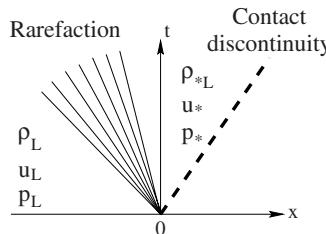
and

$$A_L = \frac{2}{(\gamma + 1)\rho_L}, \quad B_L = \frac{(\gamma - 1)}{(\gamma + 1)}p_L.$$

Thus, the sought expression for  $f_L$  for the case in which the left wave is a shock wave has been obtained.

#### 4.2.2 Function $f_L$ for Left Rarefaction

Now we derive an expression for  $f_L$  for the case in which the left wave is a rarefaction wave, as shown in Fig. 4.5. The unknown state  $\mathbf{W}_{*L}$  is now connected to the left data state  $\mathbf{W}_L$  using the isentropic relation and the Generalised Riemann Invariants for the left wave. As seen in Sect. 3.1.2 of



**Fig. 4.5.** Left wave is a rarefaction wave that connects the data state  $\mathbf{W}_L$  with the unknown state  $\mathbf{W}_{*L}$  in the star region to the left of the contact discontinuity

Chap. 3, the isentropic law

$$p = C\rho^\gamma, \quad (4.22)$$

where  $C$  is a constant, may be used across rarefactions.  $C$  is evaluated at the *initial* left data state by applying the isentropic law, namely

$$p_L = C \rho_L^\gamma ,$$

and so the constant  $C$  is

$$C = p_L / \rho_L^\gamma ,$$

from which we write

$$\rho_{*L} = \rho_L \left( \frac{p_*}{p_L} \right)^{\frac{1}{\gamma}} . \quad (4.23)$$

In Sect. 3.1.3 of Chap. 3 we showed that across a left rarefaction the Generalised Riemann Invariant  $I_L(u, a)$  is constant. By evaluating the constant on the left data state we write

$$u_L + \frac{2a_L}{\gamma - 1} = u_* + \frac{2a_{*L}}{\gamma - 1} , \quad (4.24)$$

where  $a_L$  and  $a_{*L}$  denote the sound speed on the left and right states bounding the left rarefaction wave. See Fig. 4.5.

Substitution of  $\rho_{*L}$  from (4.23) into the definition of  $a_{*L}$  gives

$$a_{*L} = a_L \left( \frac{p_*}{p_L} \right)^{\frac{\gamma-1}{2\gamma}} , \quad (4.25)$$

and equation (4.24) leads to

$$u_* = u_L - f_L(p_*, \mathbf{W}_L) , \quad (4.26)$$

with

$$f_L(p_*, \mathbf{W}_L) = \frac{2a_L}{(\gamma - 1)} \left[ \left( \frac{p_*}{p_L} \right)^{\frac{\gamma-1}{2\gamma}} - 1 \right] .$$

This is the required expression for the function  $f_L$  for the case in which the left wave is a rarefaction wave.

### 4.2.3 Function $f_R$ for a Right Shock

Here we find the expression for the function  $f_R$  for the case in which the right wave is a shock wave travelling with speed  $S_R$ . The situation is entirely analogous to the case of a left shock wave. Pre-shock values are  $\rho_R$ ,  $u_R$  and  $p_R$  and post-shock values are  $\rho_{*R}$ ,  $u_*$  and  $p_*$ . In the transformed frame of reference moving with the shock, the shock speed is zero and the *relative* velocities are

$$\hat{u}_R = u_R - S_R , \quad \hat{u}_* = u_* - S_R . \quad (4.27)$$

Application of the Rankine–Hugoniot conditions gives

$$\left. \begin{aligned} \rho_{*R} \hat{u}_* &= \rho_R \hat{u}_R , \\ \rho_{*R} \hat{u}_*^2 + p_* &= \rho_R \hat{u}_R^2 + p_R , \\ \hat{u}_*(\hat{E}_{*R} + p_*) &= \hat{u}_R(\hat{E}_R + p_R) . \end{aligned} \right\} \quad (4.28)$$

Now the *mass flux* is defined as

$$Q_R \equiv -\rho_{*R} \hat{u}_* = -\rho_R \hat{u}_R . \quad (4.29)$$

By performing algebraic manipulations similar to those for a left shock we derive the following expression for the mass flux

$$Q_R = \left[ \frac{p_* + B_R}{A_R} \right]^{\frac{1}{2}} . \quad (4.30)$$

Hence the particle velocity in the *Star Region* satisfies

$$u_* = u_R + f_R(p_*, \mathbf{W}_R) , \quad (4.31)$$

with

$$\begin{aligned} f_R(p_*, \mathbf{W}_R) &= (p_* - p_R) \left[ \frac{A_R}{p_* + B_R} \right]^{\frac{1}{2}} , \\ A_R &= \frac{2}{(\gamma + 1)\rho_R} , \quad B_R = \frac{(\gamma - 1)}{(\gamma + 1)} p_R . \end{aligned}$$

This is the sought expression for  $f_R$  for the case in which the right wave is a shock wave.

#### 4.2.4 Function $f_R$ for a Right Rarefaction

The derivation of the function  $f_R$  for the case in which the right wave is a rarefaction wave is carried out in an entirely analogous manner to the case of a left rarefaction. The isentropic law gives

$$\rho_{*R} = \rho_R \left( \frac{p_*}{p_R} \right)^{\frac{1}{\gamma}} \quad (4.32)$$

and the Generalised Riemann Invariant  $I_R(u, a)$  for a right rarefaction gives

$$u_* - \frac{2a_{*R}}{\gamma - 1} = u_R - \frac{2a_R}{\gamma - 1} . \quad (4.33)$$

Using (4.32) into the definition of sound speed  $a_{*R}$  gives

$$a_{*R} = a_R \left( \frac{p_*}{p_R} \right)^{\frac{\gamma-1}{2\gamma}} , \quad (4.34)$$

which if substituted into (4.33) leads to

$$u_* = u_R + f_R(p_*, \mathbf{W}_R) , \quad (4.35)$$

with

$$f_R(p_*, \mathbf{W}_R) = \frac{2a_R}{\gamma - 1} \left[ \left( \frac{p_*}{p_R} \right)^{\frac{\gamma-1}{2\gamma}} - 1 \right].$$

The functions  $f_L$  and  $f_R$  have now been determined for all four possible wave patterns of Fig. 4.2. Now by eliminating  $u_*$  from equations (4.21) or (4.26) and (4.31) or (4.35) we obtain a single equation

$$f(p_*, \mathbf{W}_L, \mathbf{W}_R) \equiv f_L(p_*, \mathbf{W}_L) + f_R(p_*, \mathbf{W}_R) + \Delta u = 0, \quad (4.36)$$

which is the required equation (4.5) for the pressure. This proves the first part of the proposition. Assuming this single non-linear algebraic equation is solved (numerically) for  $p_*$  then the solution for the particle velocity  $u_*$  can be found from equation (4.21) if the left wave is a shock ( $p_* > p_L$ ) or from equation (4.26) if the left wave is a rarefaction ( $p_* \leq p_L$ ) or from equation (4.31) if the right wave is a shock ( $p_* > p_R$ ) or from equation (4.35) if the right wave is a rarefaction wave ( $p_* \leq p_R$ ). It can also be found from a mean value as

$$u_* = \frac{1}{2}(u_L + u_R) + \frac{1}{2}[f_R(p_*) - f_L(p_*)],$$

which is equation (4.9), and the proposition has thus been proved.

## 4.3 Numerical Solution for Pressure

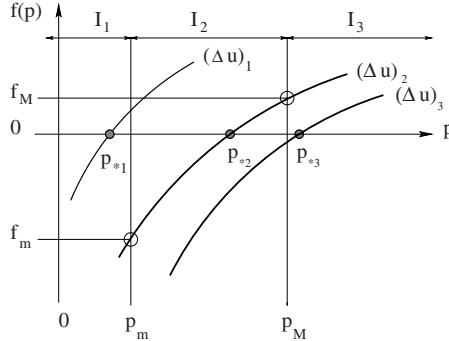
The unknown pressure  $p_*$  in the *Star Region* is found by solving the single algebraic equation (4.5),  $f(p) = 0$ , numerically. Any standard technique can be used. See Maron and Lopez [337] for background on numerical methods for algebraic equations. The behaviour of the pressure function  $f(p)$  plays a fundamental role in finding its roots numerically.

### 4.3.1 Behaviour of the Pressure Function

Given data  $\rho_L$ ,  $u_L$ ,  $p_L$  and  $\rho_R$ ,  $u_R$ ,  $p_R$  the pressure function  $f(p)$  behaves as shown in Fig. 4.6. It is monotone and concave down as we shall demonstrate. The first derivatives of  $f_K$  ( $K=L,R$ ) with respect to  $p$  are

$$f'_K = \begin{cases} \left( \frac{A_K}{B_K + p} \right)^{1/2} \left[ 1 - \frac{p - p_K}{2(B_K + p)} \right] & \text{if } p > p_K \text{ (shock)}, \\ \frac{1}{\rho_K a_K} \left( \frac{p}{p_K} \right)^{-(\gamma+1)/2\gamma} & \text{if } p \leq p_K \text{ (rarefaction)}. \end{cases} \quad (4.37)$$

As  $f' = f'_L + f'_R$  and by inspection  $f'_K > 0$ , the function  $f(p)$  is monotone as claimed. The second derivatives of the functions  $f_K$  are



**Fig. 4.6.** Behaviour of the pressure function in the solution of the Riemann problem

$$f''_K = \begin{cases} -\frac{1}{4} \left( \frac{A_K}{B_K + p} \right)^{1/2} \left[ \frac{4B_K + 3p + p_K}{(B_K + p)^2} \right] & \text{if } p > p_K \text{ (shock)} , \\ -\frac{(\gamma+1)a_K}{2\gamma^2 p_K^2} \left( \frac{p}{p_K} \right)^{-(3\gamma+1)/2\gamma} & \text{if } p \leq p_K \text{ (rarefaction)} . \end{cases} \quad (4.38)$$

Since  $f'' = f''_L + f''_R$  and  $f''_K < 0$  the function  $f(p)$  is concave down as anticipated. From equations (4.37) and (4.38) it can be seen that  $f'_K \rightarrow 0$  as  $p \rightarrow \infty$  and  $f''_K \rightarrow 0$  as  $p \rightarrow \infty$ . This behaviour of  $f_K$ , and thus of  $f(p)$ , has implications when devising iteration schemes to find the zero  $p_*$  of  $f(p) = 0$ . The velocity difference  $\Delta u = u_R - u_L$  and the pressure values  $p_L$ ,  $p_R$  are the most important parameters of  $f(p)$ . With reference to Fig. 4.6 we define

$$p_{\min} = \min(p_L, p_R), \quad p_{\max} = \max(p_L, p_R),$$

$$f_{\min} = f(p_{\min}), \quad f_{\max} = f(p_{\max}).$$

For given  $p_L$ ,  $p_R$  it is the velocity difference  $\Delta u$  which determines the value of  $p_*$ . Three intervals  $I_1$ ,  $I_2$  and  $I_3$  can be identified:

$$\left. \begin{array}{ll} p_* \text{ lies in } I_1 = (0, p_{\min}) & \text{if } f_{\min} > 0 \text{ and } f_{\max} > 0 , \\ p_* \text{ lies in } I_2 = [p_{\min}, p_{\max}] & \text{if } f_{\min} \leq 0 \text{ and } f_{\max} \geq 0 , \\ p_* \text{ lies in } I_3 = (p_{\max}, \infty) & \text{if } f_{\min} < 0 \text{ and } f_{\max} < 0 . \end{array} \right\} \quad (4.39)$$

For sufficiently large  $\Delta u$ , as  $(\Delta u)_1$  in Fig. 4.6, the solution  $p_*$  is as  $p_{*1}$ , which lies in  $I_1$  and thus  $p_* < p_L$ ,  $p_* < p_R$ ; so the two non-linear waves are rarefaction waves. For  $\Delta u$  as  $(\Delta u)_2$  in Fig. 4.6  $p_* = p_{*2}$  lies between  $p_L$  and  $p_R$  and hence one non-linear wave is a rarefaction wave and the other is a shock wave. For sufficiently small values of  $\Delta u$ , as  $(\Delta u)_3$  in Fig. 4.6,  $p_* = p_{*3}$  lies in  $I_3$ , that is  $p_* > p_L$ ,  $p_* > p_R$ , which means that both non-linear waves are shock waves. The interval where  $p_*$  lies is identified by noting the signs of  $f_{\min}$  and  $f_{\max}$ ; see 4.39.

Another observation on the behaviour of  $f(p)$  is this: in  $I_1$  both  $f'(p)$  and  $f''(p)$  vary rapidly; this may lead to numerical difficulties when searching for the root of  $f(p) = 0$ . As  $p$  increases the shape of  $f(p)$  tends to resemble that of a straight line. For non-vacuum initial data  $\mathbf{W}_L, \mathbf{W}_R$  there exists a unique positive solution  $p_*$  for pressure, provided  $\Delta u$  is *sufficiently small*. As a matter of fact, even for the case in which the data states are *non-vacuum states*, values of  $\Delta u$  larger than a critical value  $(\Delta u)_{\text{crit}}$  lead to vacuum in the solution of the Riemann problem. The critical value can be found analytically in terms of the initial data. Clearly for a positive solution for pressure  $p_*$  we require  $f(0) < 0$ . Direct evaluation of  $f(p)$  gives the *pressure positivity condition*

$$(\Delta u)_{\text{crit}} \equiv \frac{2a_L}{\gamma - 1} + \frac{2a_R}{\gamma - 1} > u_R - u_L . \quad (4.40)$$

*Vacuum is created* by the non-linear waves if this condition is violated. The structure of the solution in this case is different from that depicted in Fig. 4.1 and so is the method of solution, as we shall see in Sect. 4.6 of this chapter.

### 4.3.2 Iterative Scheme for Finding the Pressure

Given the particularly simple behaviour of the pressure function  $f(p)$  and the availability of analytical expressions for the derivative of  $f(p)$  we use a Newton-Raphson [337] iterative procedure to find the root of  $f(p) = 0$ . Suppose a guess value  $p_0$  for the true solution  $p_*$  is available; since  $f(p)$  is a smooth function we can find an approximate value of  $f(p)$  at a neighbouring point  $p_0 + \delta$  via a Taylor expansion

$$f(p_0 + \delta) = f(p_0) + \delta f'(p_0) + O(\delta^2) . \quad (4.41)$$

If the  $p_0 + \delta$  is a solution of  $f(p) = 0$  then

$$f(p_0) + \delta f'(p_0) = 0 , \quad (4.42)$$

and so the *corrected* value  $p_1 = p_0 + \delta$  is

$$p_1 = p_0 - \frac{f(p_0)}{f'(p_0)} . \quad (4.43)$$

The above procedure generalises to

$$p_{(k)} = p_{(k-1)} - \frac{f(p_{(k-1)})}{f'(p_{(k-1)})} , \quad (4.44)$$

where  $p_{(k)}$  is the  $k$ -th iterate. The iteration procedure is stopped whenever the relative pressure change

$$CHA = \frac{|p_{(k)} - p_{(k-1)}|}{\frac{1}{2} [p_{(k)} + p_{(k-1)}]} , \quad (4.45)$$

is less than a prescribed small tolerance  $TOL$ . Typically  $TOL = 10^{-6}$ .

In order to implement the iteration scheme (4.44) we need a guess value  $p_0$  for the pressure. Given the benign behaviour of  $f(p)$  the choice of  $p_0$  is not too critical. An inadequate choice of  $p_0$  results in a large number of iterations to achieve convergence. A difficulty that requires special handling in the Newton–Raphson method arises when the root is close to zero (strong rarefaction waves) and the guess value  $p_0$  is too large: the next iterate for pressure can be negative. This is due to the rapid variations of the first and second derivatives of  $f(p)$  near  $p = 0$ . We illustrate the effect of the initial guess value by considering four possible choices. Three of these are approximations to the solution  $p_*$  for pressure, see Chap. 9 for details. One such approximation is the so called *Two–Rarefaction* approximation

$$p_{\text{TR}} = \left[ \frac{a_L + a_R - \frac{1}{2}(\gamma - 1)(u_R - u_L)}{a_L/p_L^{\frac{\gamma-1}{2\gamma}} + a_R/p_R^{\frac{\gamma-1}{2\gamma}}} \right]^{\frac{2\gamma}{\gamma-1}}, \quad (4.46)$$

and results from the exact function (4.5) for pressure under the assumption that the two non-linear waves are rarefaction waves. If the solution actually consists of two rarefactions then  $p_{\text{TR}}$  is exact and no iteration is required. A second guess value results from a linearised solution based on primitive variables. This is

$$\left. \begin{aligned} p_0 &= \max(TOL, p_{\text{PV}}), \\ p_{\text{PV}} &= \frac{1}{2}(p_L + p_R) - \frac{1}{8}(u_R - u_L)(\rho_L + \rho_R)(a_L + a_R). \end{aligned} \right\} \quad (4.47)$$

A third guess value is given by a *Two–Shock* approximation

$$\left. \begin{aligned} p_0 &= \max(TOL, p_{\text{TS}}), \\ p_{\text{TS}} &= \frac{g_L(\hat{p})p_L + g_R(\hat{p})p_R - \Delta u}{g_L(\hat{p}) + g_R(\hat{p})}, \\ g_K(p) &= \left( \frac{A_K}{p+B_K} \right)^{\frac{1}{2}}, \end{aligned} \right\} \quad (4.48)$$

where  $A_K$  and  $B_K$  given by (4.8). Here  $\hat{p}$  is an estimate of the solution; the value  $\hat{p} = p_0$  given by (4.47) works well. Note that approximate solutions may predict, incorrectly, a negative value for pressure, even when condition (4.40) is satisfied. Thus in order to avoid negative guess values we introduce the small positive constant  $TOL$ , as used in the iteration procedure. As a fourth guess value we utilise the arithmetic mean of the data, namely

$$p_0 = \frac{1}{2}(p_L + p_R). \quad (4.49)$$

Next, we carry out some tests on the effect of the various guess values for  $p_0$  on the convergence of the Newton–Raphson iterative scheme for finding the pressure  $p_*$ .

### 4.3.3 Numerical Tests

Five Riemann problems are selected to test the performance of the Riemann solver and the influence of the initial guess for pressure. The tests are also used to illustrate some typical wave patterns resulting from the solution of the Riemann problem. Table 4.1 shows the data for all five tests in terms of primitive variables. In all cases the ratio of specific heats is  $\gamma = 1.4$ . The source code for the exact Riemann solver, called HE-E1RPEXACT, is part of the library *NUMERICA* [519]; a listing is given in Sect. 4.9.

Test 1 is the so called Sod test problem [453]; this is a very mild test and its solution consists of a left rarefaction, a contact and a right shock. Fig. 4.7 shows solution profiles for density, velocity, pressure and specific internal energy across the complete wave structure, at time  $t = 0.25$  units. Test 2, called the *123 problem*, has solution consisting of two strong rarefactions and a trivial stationary contact discontinuity; the pressure  $p_*$  is very small (close to vacuum) and this can lead to difficulties in the iteration scheme to find  $p_*$  numerically. Fig. 4.8 shows solution profiles. Test 2 is also useful in assessing the performance of numerical methods for low density flows, see Einfeldt et al. [182]. Test 3 is a very severe test problem, the solution of which contains a left rarefaction, a contact and a right shock; this test is actually the left half of the blast wave problem of Woodward and Colella [584], Fig. 4.9 shows solution profiles. Test 4 is the right half of the Woodward and Colella problem; its solution contains a left shock, a contact discontinuity and a right rarefaction, as shown in Fig. 4.10. Test 5 is made up of the right and left shocks emerging from the solution to tests 3 and 4 respectively; its solution represents the collision of these two strong shocks and consists of a left facing shock (travelling very slowly to the right), a right travelling contact discontinuity and a right travelling shock wave. Fig. 4.11 shows solution profiles for Test 5.

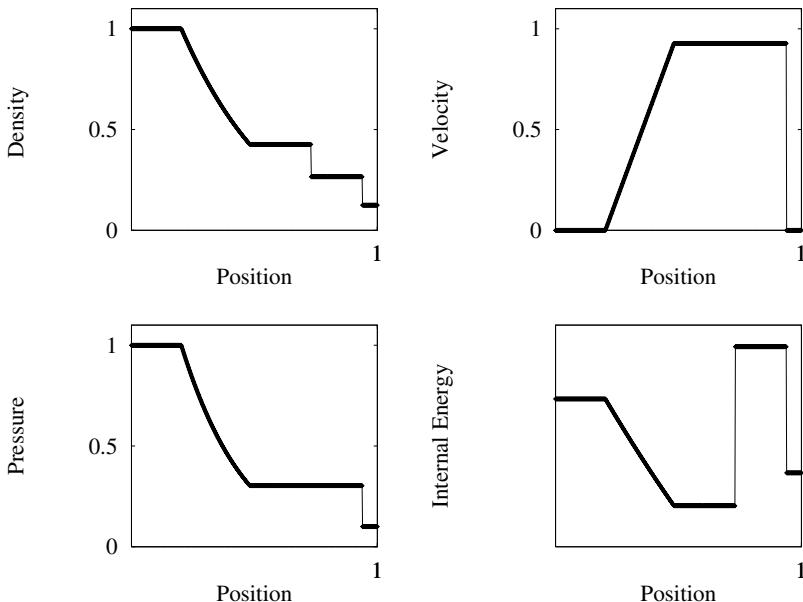
Test	$\rho_L$	$u_L$	$p_L$	$\rho_R$	$u_R$	$p_R$
1	1.0	0.0	1.0	0.125	0.0	0.1
2	1.0	-2.0	0.4	1.0	2.0	0.4
3	1.0	0.0	1000.0	1.0	0.0	0.01
4	1.0	0.0	0.01	1.0	0.0	100.0
5	5.99924	19.5975	460.894	5.99242	-6.19633	46.0950

Table 4.1. Data for five Riemann problem tests

Table 4.2 shows the computed values for pressure in the *Star Region* by solving the pressure equation  $f(p) = 0$  (equation 4.5) by a Newton–Raphson method. This task is carried out by the subroutine STARPU, which is contained in the FORTRAN 77 program given in Sect. 4.9 of this chapter.

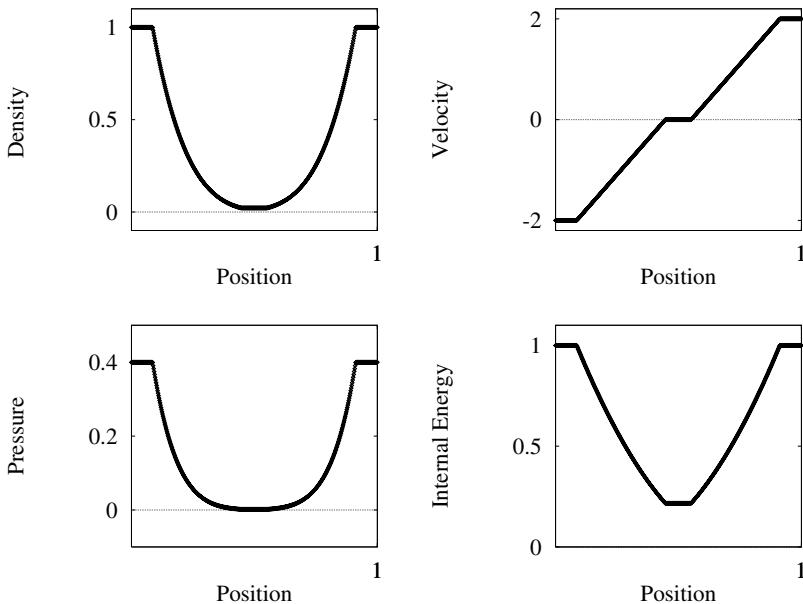
Test	$p_*$	$p_{\text{TR}}$	$p_{\text{PV}}$	$p_{\text{TS}}$	$\frac{1}{2}(p_L + p_R)$
1	0.30313	0.30677(3)	0.55000(5)	0.31527(3)	0.55(5)
2	0.00189	exact(1)	TOL(8)	TOL(8)	0.4(9)
3	460.894	912.449(5)	500.005(4)	464.108(3)	500.005(4)
4	46.0950	82.9831(5)	50.005(4)	46.4162(3)	50.005(4)
5	1691.64	2322.65(4)	781.353(5)	1241.21(4)	253.494(6)

Table 4.2 Guess values  $p_0$  for iteration scheme. Next to each guess is the required number of iterations for convergence (in parentheses).



**Fig. 4.7.** Test 1: Exact solution for density, velocity, pressure and specific internal energy at time  $t = 0.25$  units

The exact, converged, solution for pressure is given in column 2. Columns 3 to 6 give the guess values  $p_{\text{TR}}$ ,  $p_{\text{PV}}$ ,  $p_{\text{TS}}$  and the arithmetic mean value of the data. The number in parentheses next to each guess value is the number of iterations required for convergence for a tolerance  $TOL = 10^{-6}$ . For Test 1,  $p_{\text{TR}}$  and  $p_{\text{TS}}$  are the best guess values for  $p_0$ . For Test 2,  $p_{\text{TR}}$  is actually the exact solution (two rarefactions). By excluding Test 2,  $p_{\text{TS}}$  is the best guess overall. Experience in using hybrid schemes suggests that a combination of two or three approximations is bound to provide a suitable guess value for  $p_0$  that is both accurate and efficient. In the FORTRAN 77 program provided in Sect. 4.9 of this chapter, the subroutine STARTE contains a hybrid scheme involving



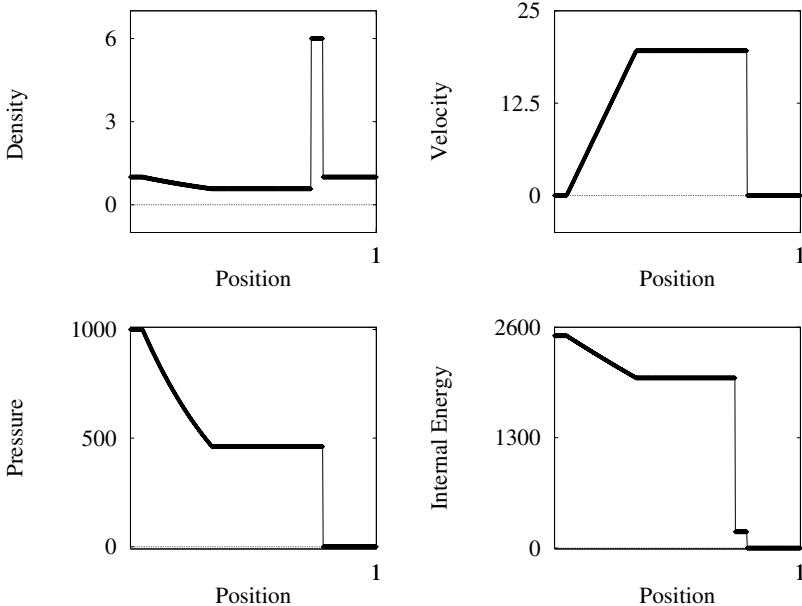
**Fig. 4.8.** Test 2: Exact solution for density, velocity, pressure and specific internal energy at time  $t = 0.15$  units

$p_{PV}$ ,  $p_{TR}$  and  $p_{TS}$ . In a typical application of the exact Riemann solver to a numerical method, the overwhelming majority of Riemann problems will consist of *nearby states* which can be accurately approximated by the simple value  $p_{PV}$ .

Having found  $p_*$ , the solution  $u_*$  for the particle velocity follows from (4.9) and the density values  $\rho_{*L}$ ,  $\rho_{*R}$  follow from appropriate wave relations, as detailed in the next section. Table 4.3 shows exact solutions for pressure  $p_*$ , speed  $u_*$ , densities  $\rho_{*L}$  and  $\rho_{*R}$  for tests 1 to 5. These quantities may prove of some use for initial testing of programs.

Test	$p_*$	$u_*$	$\rho_{*L}$	$\rho_{*R}$
1	0.30313	0.92745	0.42632	0.26557
2	0.00189	0.00000	0.02185	0.02185
3	460.894	19.5975	0.57506	5.99924
4	46.0950	-6.19633	5.99242	0.57511
5	1691.64	8.68975	14.2823	31.0426

Table 4.3. Exact solution for pressure, speed and densities for tests 1 to 5.



**Fig. 4.9.** Test 3: Exact solution for density, velocity, pressure and specific internal energy at time  $t = 0.012$  units

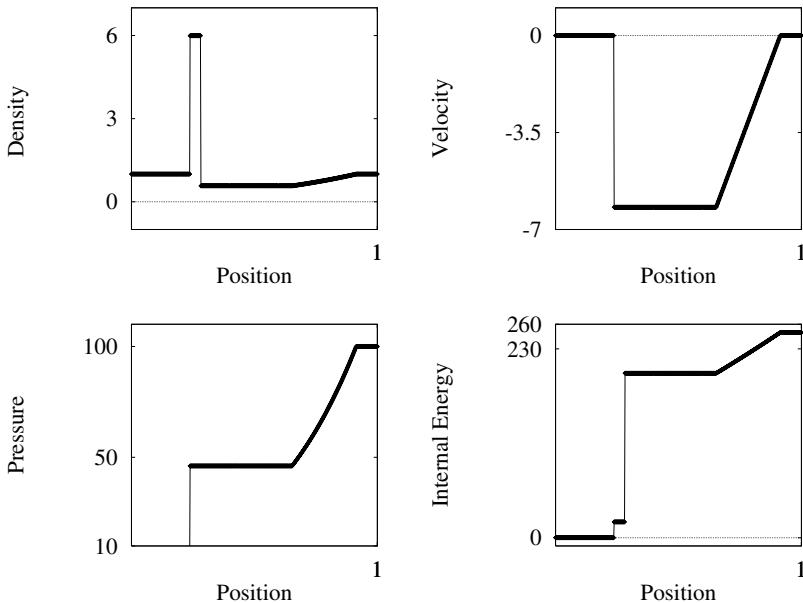
## 4.4 The Complete Solution

So far, we have an algorithm for computing the pressure  $p_*$  and particle velocity  $u_*$  in the *Star Region*. We still do not know the sought values  $\rho_{*L}$  and  $\rho_{*R}$  for the density in this region; these are computed by identifying the types of non-linear waves, and can be done by comparing the pressure  $p_*$  to the pressures  $p_L$  and  $p_R$ , and then applying the appropriate conditions across the respective waves. Another pending task is to determine completely the left and right waves. For shock waves we only need to find the density behind the wave and the shock speed. For rarefaction waves there is more work involved: we need  $\rho$  behind the wave, equations for the *Head* and *Tail* of the wave and the full solution inside the rarefaction fan.

There are two cases. First we consider the case in which the sampling point  $(x, t)$  lies to the left of the contact discontinuity, as in Fig. 4.12. Again, there are two possibilities; these are now studied separately.

### Left Shock Wave

A left shock wave, see Fig. 4.12a, is identified by the condition  $p_* > p_L$ . We know  $p_*$  and  $u_*$ . From the pressure ratio, see Sect. 3.1.3 of Chap. 3, we find the density according to



**Fig. 4.10.** Test 4: Exact solution for density, velocity, pressure and specific internal energy at time  $t = 0.035$  units

$$\rho_{*L} = \rho_L \left[ \frac{\frac{p_*}{p_L} + \frac{\gamma-1}{\gamma+1}}{\frac{\gamma-1}{\gamma+1} \frac{p_*}{p_L} + 1} \right]. \quad (4.50)$$

The shock speed  $S_L$  is also a function of the pressure  $p_*$ . From (4.10) and (4.14) we deduce the shock speed as

$$S_L = u_L - Q_L/\rho_L, \quad (4.51)$$

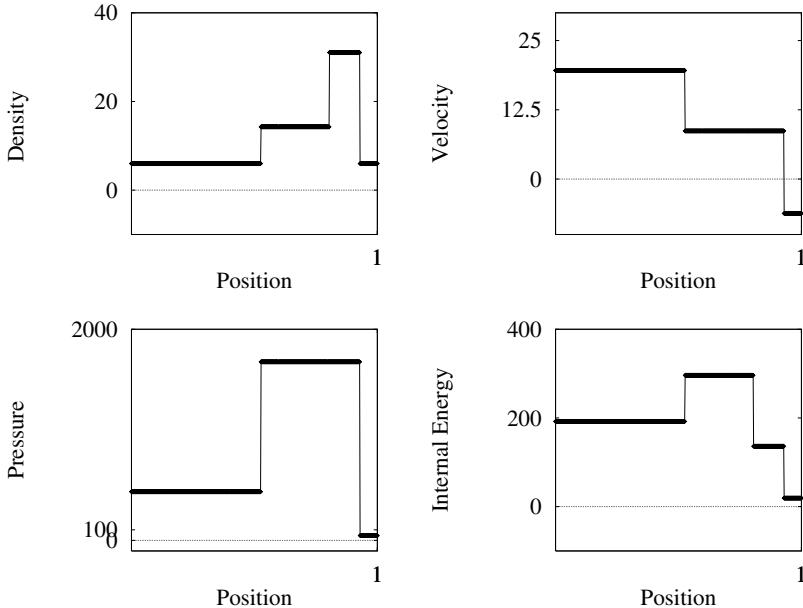
where the mass flux  $Q_L$  is given by (4.20). More explicitly, see Sect. 3.1.3 of Chap. 3, one has

$$S_L = u_L - a_L \left[ \frac{\gamma+1}{2\gamma} \frac{p_*}{p_L} + \frac{\gamma-1}{2\gamma} \right]^{\frac{1}{2}}. \quad (4.52)$$

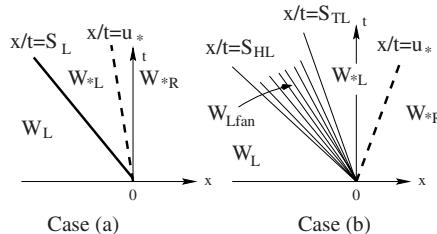
We have therefore completely determined the solution for the entire region to the left of the contact discontinuity in the case in which the left wave is a shock wave.

### Left Rarefaction Wave

A left rarefaction wave, see Fig. 4.12b, is identified by the condition  $p_* \leq p_L$ . The pressure  $p_*$  and the particle velocity  $u_*$  in the *Star Region* are known. The density follows from the isentropic law as



**Fig. 4.11.** Test 5: Exact solution for density, velocity, pressure and specific internal energy at time  $t = 0.035$  units



**Fig. 4.12.** Sampling the solution at a point to the left of the contact: (a) left wave is a shock (b) left wave is a rarefaction

$$\rho_{*L} = \rho_L \left( \frac{p_*}{p_L} \right)^{\frac{1}{\gamma}}. \quad (4.53)$$

The sound speed behind the rarefaction is

$$a_{*L} = a_L \left( \frac{p_*}{p_L} \right)^{\frac{\gamma-1}{2\gamma}}. \quad (4.54)$$

The rarefaction wave is enclosed by the *Head* and the *Tail*, which are the characteristics of speeds given respectively by

$$S_{HL} = u_L - a_L, \quad S_{TL} = u_* - a_{*L}. \quad (4.55)$$

We now find the solution for  $\mathbf{W}_{\text{Lfan}} = (\rho, u, p)^T$  inside the left rarefaction fan. This is easily obtained by considering the characteristic ray through the origin  $(0, 0)$  and a general point  $(x, t)$  inside the fan. The slope of the characteristic is

$$\frac{dx}{dt} = \frac{x}{t} = u - a ,$$

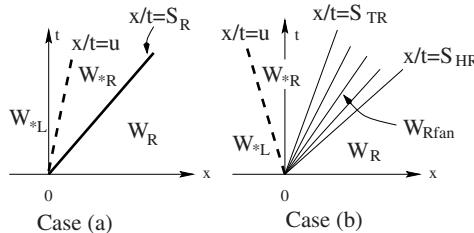
where  $u$  and  $a$  are respectively the sought particle velocity and sound speed at  $(x, t)$ . Also, use of the Generalised Riemann Invariant  $I_L(u, a)$  yields

$$u_L + \frac{2a_L}{\gamma - 1} = u + \frac{2a}{\gamma - 1} .$$

The simultaneous solution of these two equations for  $u$  and  $a$ , use of the definition of the sound speed  $a$  and the isentropic law give the result

$$\mathbf{W}_{\text{Lfan}} = \begin{cases} \rho = \rho_L \left[ \frac{2}{(\gamma+1)} + \frac{(\gamma-1)}{(\gamma+1)a_L} (u_L - \frac{x}{t}) \right]^{\frac{2}{\gamma-1}} , \\ u = \frac{2}{(\gamma+1)} \left[ a_L + \frac{(\gamma-1)}{2} u_L + \frac{x}{t} \right] , \\ p = p_L \left[ \frac{2}{(\gamma+1)} + \frac{(\gamma-1)}{(\gamma+1)a_L} (u_L - \frac{x}{t}) \right]^{\frac{2\gamma}{\gamma-1}} . \end{cases} \quad (4.56)$$

Next we consider the solution at a point  $(x, t)$  to the right of the contact discontinuity for the two possible wave configurations of Fig. 4.13.



**Fig. 4.13.** Sampling the solution at a point to the right of the contact: (a) right wave is a shock (b) right wave is a rarefaction

### Right Shock Wave

A right shock wave, see Fig. 4.13a, is identified by the condition  $p_* > p_R$ . We know the pressure  $p_*$  and the particle velocity  $u_*$ . The density  $\rho_{*R}$  is found to be

$$\rho_{*R} = \rho_R \left[ \frac{\frac{p_*}{p_R} + \frac{\gamma-1}{\gamma+1}}{\frac{\gamma-1}{\gamma+1} \frac{p_*}{p_R} + 1} \right] \quad (4.57)$$

and the shock speed is

$$S_R = u_R + Q_R / \rho_R , \quad (4.58)$$

with the mass flux  $Q_R$  given by (4.30). More explicitly we have

$$S_R = u_R + a_R \left[ \frac{(\gamma+1)}{2\gamma} \frac{p_*}{p_R} + \frac{(\gamma-1)}{2\gamma} \right]^{\frac{1}{2}} . \quad (4.59)$$

### Right Rarefaction Wave

A right rarefaction wave, see Fig. 4.13b, is identified by the condition  $p_* \leq p_R$ . The pressure  $p_*$  and velocity  $u_*$  in the *Star Region* are known. The density is found from the isentropic law as

$$\rho_{*R} = \rho_R \left( \frac{p_*}{p_R} \right)^{\frac{1}{\gamma}} , \quad (4.60)$$

from which the sound speed follows as

$$a_{*R} = a_R \left( \frac{p_*}{p_R} \right)^{\frac{\gamma-1}{2\gamma}} . \quad (4.61)$$

The speeds for the *Head* and *Tail* are given respectively by

$$S_{HR} = u_R + a_R , \quad S_{TR} = u_* + a_{*R} . \quad (4.62)$$

The solution for  $\mathbf{W}_{R\text{fan}}$  inside a right rarefaction fan is found in an analogous manner to the case of a left rarefaction fan. The solution is

$$\mathbf{W}_{R\text{fan}} = \begin{cases} \rho = \rho_R \left[ \frac{2}{(\gamma+1)} - \frac{(\gamma-1)}{(\gamma+1)a_R} (u_R - \frac{x}{t}) \right]^{\frac{2}{\gamma-1}} , \\ u = \frac{2}{(\gamma+1)} \left[ -a_R + \frac{(\gamma-1)}{2} u_R + \frac{x}{t} \right] , \\ p = p_R \left[ \frac{2}{(\gamma+1)} - \frac{(\gamma-1)}{(\gamma+1)a_R} (u_R - \frac{x}{t}) \right]^{\frac{2\gamma}{\gamma-1}} . \end{cases} \quad (4.63)$$

## 4.5 Sampling the Solution

We have developed a solver to find the exact solution of the complete wave structure of the Riemann problem at any point  $(x, t)$  in the relevant domain of interest  $x_L < x < x_R$ ;  $t > 0$ , with  $x_L < 0$  and  $x_R > 0$ . We now provide a solution sampling procedure which, apart from being a summary of the solution, may also prove of practical use when programming the solution algorithm. Suppose we wish to evaluate the solution at a general point  $(x, t)$ .

We denote the solution of the Riemann problem at  $(x, t)$  in terms of the vector of primitive variables  $\mathbf{W} = (\rho, u, p)^T$ . As the solution  $\mathbf{W}$  is a similarity solution we perform the sampling in terms of the speed  $S = x/t$ . When the solution at a specified time  $t$  is required the solution profiles are only a function of space  $x$ . In sampling the complete solution there are two cases to consider.

#### 4.5.1 Left Side of Contact: $S = x/t \leq u_*$

As shown in Fig. 4.12 there are two possible wave configurations. Fig. 4.12a shows the case in which the left wave is a shock wave. In this case the complete solution on the left hand side of the contact wave is

$$\mathbf{W}(x, t) = \begin{cases} \mathbf{W}_{*L}^{\text{sho}} & \text{if } S_L \leq \frac{x}{t} \leq u_* , \\ \mathbf{W}_L & \text{if } \frac{x}{t} \leq S_L , \end{cases} \quad (4.64)$$

where  $S_L$  is the shock speed given by (4.52),  $\mathbf{W}_{*L}^{\text{sho}} = (\rho_{*L}^{\text{sho}}, u_*, p_*)^T$  with  $\rho_{*L}^{\text{sho}}$  given by (4.50) and  $\mathbf{W}_L$  is the left data state. If the left wave is a rarefaction, as depicted by Fig. 4.12b, then the complete solution on the left hand side of the contact consists of three states, namely

$$\mathbf{W}(x, t) = \begin{cases} \mathbf{W}_L & \text{if } \frac{x}{t} \leq S_{HL} , \\ \mathbf{W}_{\text{Lfan}} & \text{if } S_{HL} \leq \frac{x}{t} \leq S_{TL} , \\ \mathbf{W}_{*L}^{\text{fan}} & \text{if } S_{TL} \leq \frac{x}{t} \leq u_* , \end{cases} \quad (4.65)$$

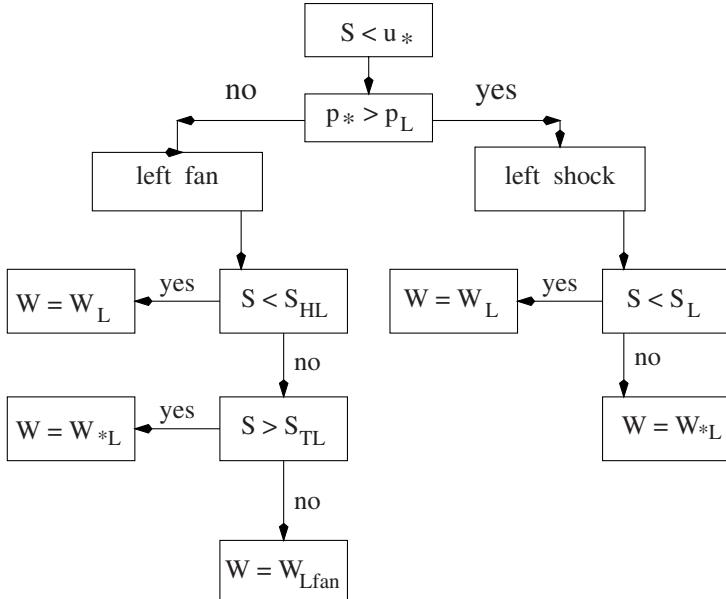
where  $S_{HL}$  and  $S_{TL}$  are the speeds of the head and tail of the rarefaction given by (4.55),  $\mathbf{W}_{*L}^{\text{fan}} = (\rho_{*L}^{\text{fan}}, u_*, p_*)^T$  with  $\rho_{*L}^{\text{fan}}$  given by (4.53),  $\mathbf{W}_{\text{Lfan}}$  is the state inside the rarefaction fan given by (4.56) and  $\mathbf{W}_L$  is the left data state. Fig. 4.14 shows a flow chart for sampling the solution at any point  $(x, t)$  to the left of the contact discontinuity. An analogous flow chart results from the case in which the point  $(x, t)$  lies to the right of the contact discontinuity; the reader is encouraged to draw the flow chart for this case.

#### 4.5.2 Right Side of Contact: $S = x/t \geq u_*$

As shown in Fig. 4.13 there are two possible wave configurations. Fig. 4.13a shows the case in which the right wave is a shock wave. In this case the complete solution on the right hand side of the contact wave is

$$\mathbf{W}(x, t) = \begin{cases} \mathbf{W}_{*R}^{\text{sho}} & \text{if } u_* \leq \frac{x}{t} \leq S_R , \\ \mathbf{W}_R & \text{if } \frac{x}{t} \geq S_R , \end{cases} \quad (4.66)$$

where  $S_R$  is the shock speed given by (4.59),  $\mathbf{W}_{*R}^{\text{sho}} = (\rho_{*R}^{\text{sho}}, u_*, p_*)^T$  with  $\rho_{*R}^{\text{sho}}$  given by (4.57) and  $\mathbf{W}_R$  is the right data state. If the right wave is a



**Fig. 4.14.** Flow chart for sampling the solution at a point  $(x, t)$  to the left of the contact discontinuity  $\frac{dx}{dt} = u_*; S = x/t$

rarefaction, as depicted by Fig. 4.13b, then the complete solution on the right hand side of the contact consists of three states, namely

$$\mathbf{W}(x, t) = \begin{cases} \mathbf{W}_{*R}^{\text{fan}} & \text{if } u_* \leq \frac{x}{t} \leq S_{TR}, \\ \mathbf{W}_{R\text{fan}} & \text{if } S_{TR} \leq \frac{x}{t} \leq S_{HR}, \\ \mathbf{W}_R & \text{if } \frac{x}{t} \geq S_{HR}, \end{cases} \quad (4.67)$$

where  $S_{HR}$  and  $S_{TR}$  are the speeds of the head and tail of the rarefaction given by (4.62),  $\mathbf{W}_{*R}^{\text{fan}} = (\rho_{*R}^{\text{fan}}, u_*, p_*)^T$  with  $\rho_{*R}^{\text{fan}}$  given by (4.60),  $\mathbf{W}_{R\text{fan}}$  is the state inside the right rarefaction fan given by (4.63) and  $\mathbf{W}_R$  is the right data state.

**Exercise 4.3.** Write a flow chart for sampling the solution at any point  $(x, t)$  to the right of the contact discontinuity  $\frac{dx}{dt} = u_*$ .

**Solution 4.4.** (Left to the reader).

In Sect. 4.9 we give a program for finding the complete solution of the Riemann problem for ideal gases, excluding the cases in which vacuum is present.

## 4.6 The Riemann Problem in the Presence of Vacuum

The admission of flowing material adjacent to *vacuum* plays an important role in a number of practical applications. Loosely, vacuum is characterised by the condition  $\rho = 0$ . It follows that the total energy per unit mass also vanishes,  $E = 0$ . Values of pressure and particle velocity in vacuum are discussed later. Naturally, in vacuum regions the Euler equations, or any other mathematical model based on the continuum assumption, are no longer a valid description of the physics. Here we discuss solutions of the Euler equations in domains adjacent to regions of vacuum. As for the non–vacuum case described previously, the simplest problem involving the vacuum state is furnished by the Riemann problem. There are two obvious cases to consider. One is that in which the left non–vacuum state is adjacent to a right vacuum state at the initial time  $t = 0$ . The second case is simply the previous case reversed, the right non–vacuum state is adjacent to a left vacuum state. There is a third case, in which both left and right data states are non–vacuum states, but the vacuum state is generated in the interaction of the data states via the Riemann problem. The solution of the Riemann problem in the presence of vacuum involves the computation of free boundaries separating vacuum regions from those in which material exists.

In the presence of vacuum the structure of the solution of the Riemann problem is different from that of the conventional case shown in Fig. 4.1; the *Star Region* does no longer exist. Attempts at using the pressure equation (4.5) and an iterative scheme to solve it will fail, simply because the scheme would be assuming a solution structure that does not exist. The temptation to use *small values* of density and pressure to *simulate* vacuum with the Riemann solver for the non–vacuum case will also prove frustrating. If this is done in approximate Riemann solvers, then one is effectively changing the local wave structure of the solution.

Concerning the admissible elementary waves present in the structure of the solution of the Riemann problem including the vacuum state, an important observation is that a shock wave cannot be adjacent to a vacuum region. This is stated in the following proposition

**Proposition 4.5.** *A shock wave cannot be adjacent to a vacuum region.*

*Proof.* Let us consider a left non–vacuum constant state  $\mathbf{W}_L = (\rho_L, u_L, p_L)^T$  adjacent to a right vacuum state  $\mathbf{W}_0 \equiv (\rho_0, u_0, p_0)^T$  at the initial time  $t = 0$ , where  $\rho_0 = 0$ . Assume these states are connected by a discontinuity of speed  $S$ . Application of the Rankine–Hugoniot conditions, see Sect. 3.1.3 of Chap. 3, gives

$$\rho_L u_L - \rho_0 u_0 = S(\rho_L - \rho_0), \quad (4.68)$$

$$\rho_L u_L^2 + p_L - (\rho_0 u_0^2 + p_0) = S(\rho_L u_L - \rho_0 u_0), \quad (4.69)$$

$$u_L(E_L + p_L) - u_0(E_0 + p_0) = S(E_L - E_0). \quad (4.70)$$

As  $E_0 = 0$  and assuming  $u_0$  to be *finite*, manipulation of the equations gives

$$u_L = u_0 = S, \quad p_L = p_0. \quad (4.71)$$

It follows that a shock wave cannot be adjacent to a region of vacuum,  $p_L = p_0$ . The proposition is thus proved.

From the result (4.71) it also follows that a contact discontinuity can be adjacent to a region of vacuum,  $u_L = u_0 = S$ , which makes perfect physical sense. This wave separates a region of material from a region of no material and is therefore a boundary. The velocity  $u_0$  of the front is also the maximum particle velocity across the wave system connecting a non-vacuum state with the vacuum state and is called the *escape velocity*. It turns out that  $u_0$  is completely determined by the data on the non-vacuum state. As to the pressure  $p_0$  we note that the previous result is independent of the specification of  $p_0$ . However, in order to determine the solution of the Riemann problem admitting regions of vacuum it is necessary to make some statements regarding pressure  $p$  and sound speed  $a$ . This is most conveniently done by specifying an equation of state. As the only possible waves are contacts and rarefactions, it is reasonable to adopt an isentropic-type equation of state and assume that this is valid all the way up to the boundary separating material from vacuum. We take the EOS

$$p = p(\rho), \quad (4.72)$$

subject to the following conditions, see Liu and Smoller [373],

$$p(0) = 0, \quad p'(0) = 0, \quad p'(\rho) > 0, \quad p''(\rho) > 0. \quad (4.73)$$

Thus from now on we denote the vacuum state by  $\mathbf{W}_0 = (\rho_0, u_0, p_0)^T \equiv (0, u_0, 0)^T$ . From equation (4.71) the velocity  $u_0$  is the speed of the boundary between the region of material and the region of vacuum. As pointed out in solving the Riemann problem including  $\mathbf{W}_0$  there are three cases to consider. The structure of the solution of the Riemann problem for each of the three cases is depicted in Figs. 4.15, 4.16 and 4.17 respectively. We now study each case in detail.

#### 4.6.1 Case 1: Vacuum Right State

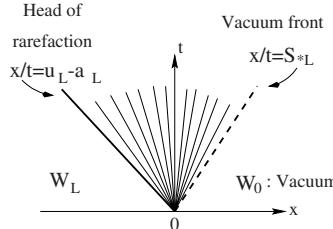
In this case the Riemann problem has data of the type

$$\mathbf{W}(x, 0) = \begin{cases} \mathbf{W}_L \neq \mathbf{W}_0 & \text{if } x < 0, \\ \mathbf{W}_0 \text{ (vacuum)} & \text{if } x > 0, \end{cases} \quad (4.74)$$

with  $\mathbf{W}_L = (\rho_L, u_L, p_L)^T$  and  $\mathbf{W}_0 = (0, u_0, 0)^T$ .

This problem may be seen as a mathematical model for the following physical situation: a shock tube of constant cross sectional area is filled with

a gas at uniform conditions. Assume the right-hand boundary is a piston of speed  $u_P = 0$  at time  $t = 0$ . Suppose the piston is *instantaneously accelerated* to a speed  $u_P > 0$  (to the right). Clearly if  $u_P > u_0$  vacuum will take place. The incipient cavitation case is  $u_P = u_0$ . The structure of the solution of the



**Fig. 4.15.** Riemann problem solution for a right vacuum state

Riemann problem with initial data (4.74) is shown in Fig. 4.15 and consists of a left rarefaction wave and a contact wave that coalesces with the tail of the rarefaction. Obviously across the contact,  $\Delta u = \Delta p = 0$ . Note that the right non-linear wave of Fig. 4.1 is absent in Fig. 4.15. The physical interpretation of this is that there is no medium for this wave to propagate through. The exact solution follows directly from the methods of previous sections if the speed of the *contact front* is known. Application of the Generalised Riemann Invariant  $I_L(u, a)$  to connect a point on the left data state with a point along the contact gives

$$u_0 + \frac{2a_0}{\gamma - 1} = u_L + \frac{2a_L}{\gamma - 1}. \quad (4.75)$$

From the assumptions (4.73) on the equation of state and the definition of sound speed we find that the sound speed vanishes along the contact, that is,  $a_0 = 0$ . Use of (4.73) with  $a_0 = 0$  gives the speed of the *front* as

$$S_{*L} \equiv u_0 = u_L + \frac{2a_L}{\gamma - 1}. \quad (4.76)$$

It is worth remarking that the conditions  $\rho_0 = 0$ ,  $p_0 = 0$  do not, in general, imply  $a_0 = 0$ . The value of  $a_0$  depends on the particular equation of state. The complete solution  $\mathbf{W}(x, t)$  can now be written as

$$\mathbf{W}_{L0}(x, t) = \begin{cases} \mathbf{W}_L & \text{if } \frac{x}{t} \leq u_L - a_L, \\ \mathbf{W}_{Lf\text{an}} & \text{if } u_L - a_L < \frac{x}{t} < S_{*L}, \\ \mathbf{W}_0 & \text{if } \frac{x}{t} \geq S_{*L}, \end{cases} \quad (4.77)$$

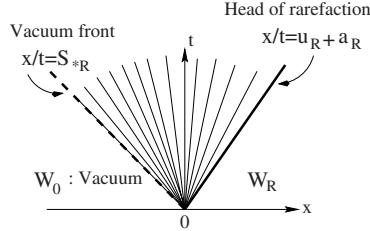
where  $\mathbf{W}_{Lf\text{an}}$  is the solution inside the left rarefaction fan given by (4.56).

### 4.6.2 Case 2: Vacuum Left State

In this case the Riemann problem has data of the type

$$\mathbf{W}(x, 0) = \begin{cases} \mathbf{W}_0 \text{ (vacuum) if } x < 0, \\ \mathbf{W}_R \neq \mathbf{W}_0 \text{ if } x > 0. \end{cases} \quad (4.78)$$

The structure of the solution in the  $x-t$  plane is illustrated in Fig. 4.16.



**Fig. 4.16.** Riemann problem solution for a left vacuum state

Compared with the regular Riemann problem case of Fig. 4.1, the left non-linear wave is missing and the contact wave separating the vacuum state from the non-vacuum state coalesces with the tail of the right rarefaction wave. The speed of the contact is found to be

$$S_{*R} = u_R - \frac{2a_R}{\gamma - 1}, \quad (4.79)$$

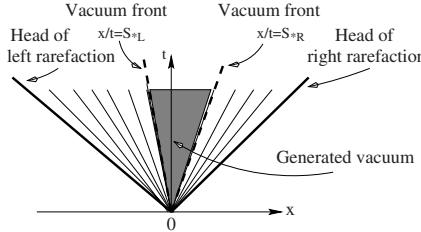
and thus the complete solution is given by

$$\mathbf{W}_{R0}(x, t) = \begin{cases} \mathbf{W}_0 \text{ if } \frac{x}{t} \leq S_{*R}, \\ \mathbf{W}_{R\text{fan}} \text{ if } S_{*R} < \frac{x}{t} < u_R + a_R, \\ \mathbf{W}_R \text{ if } \frac{x}{t} \geq u_R + a_R, \end{cases} \quad (4.80)$$

where  $\mathbf{W}_{R\text{fan}}$  is the solution inside the right fan and is given by (4.63).

### 4.6.3 Case 3: Generation of Vacuum

This case has general data  $\mathbf{W}_L = (\rho_L, u_L, p_L)^T \neq \mathbf{W}_0$ ,  $\mathbf{W}_R = (\rho_R, u_R, p_R)^T \neq \mathbf{W}_0$  but combinations of particle and sound speeds are such that *vacuum is generated* as part of the interaction between  $\mathbf{W}_L$  and  $\mathbf{W}_R$ . The structure of the solution for this case is depicted in Fig. 4.17. There are two rarefaction waves and two contact waves of speeds  $S_{*L}$  and  $S_{*R}$  that enclose the generated



**Fig. 4.17.** Riemann problem solution for non-vacuum data states that do not satisfy the pressure positivity condition. Vacuum is generated in the middle of two rarefaction waves

vacuum state. The speeds  $S_{*L}$  and  $S_{*R}$  are as given by equations (4.76) and (4.79) respectively. The full solution reads

$$\mathbf{W}(x, t) = \begin{cases} \mathbf{W}_{L0}(x, t) & \text{if } \frac{x}{t} \leq S_{*L}, \\ \mathbf{W}_0 \text{ (vacuum)} & \text{if } S_{*L} < \frac{x}{t} < S_{*R}, \\ \mathbf{W}_{R0}(x, t) & \text{if } \frac{x}{t} \geq S_{*R}, \end{cases} \quad (4.81)$$

where  $\mathbf{W}_{L0}(x, t)$  and  $\mathbf{W}_{R0}(x, t)$  are the solutions for the two previous cases given by (4.77) and (4.80) respectively. Note that for solution (4.81) to apply the condition  $S_{*L} \leq S_{*R}$  must be valid. This implies

$$(\Delta u)_{\text{crit}} \equiv \frac{2a_L}{\gamma - 1} + \frac{2a_R}{\gamma - 1} \leq u_R - u_L, \quad (4.82)$$

which is consistent with the *pressure positivity condition* (4.40). The *vacuum condition* (4.82), which is stated purely in terms of the data values for particle velocity and sound speed, can be very useful in practical applications of the Riemann problem to numerical methods of the Godunov type. A note of caution is in order: it is also possible to determine a *vacuum generating condition* for approximate Riemann solvers, but this will in general be different from the exact condition (4.82). Two useful references concerning vacuum are the papers by Munz [358] and Munz et. al. [359].

## 4.7 The Riemann Problem for Covolume Gases

A small perturbation of the ideal gas equation of state (4.3) is the covolume EOS (4.4). For details on the Thermodynamics of covolume gases see Sect. 1.2.2 of Chap. 1. We now solve the Riemann problem for the Euler equations (4.1)-(4.2) for covolume gases. The solution methodology [498] applied is the same as for the ideal gas case and we shall therefore omit much of the detail.

The sound speed for covolume gases is

$$a = \sqrt{(p/\rho^2 - e_\rho)/e_p} = \sqrt{\frac{\gamma p}{\rho(1-b\rho)}} . \quad (4.83)$$

The structure of the solution in the  $x-t$  plane is identical to that for the Riemann problem with the ideal gas EOS (4.3), in which  $b = 0$ . The solution procedure presented for the ideal gas case applies directly to the covolume case, except for finding the solution inside rarefaction waves. No closed-form solution can be found in this case and an extra iterative step is required.

#### 4.7.1 Solution for Pressure and Particle Velocity.

In solving the Riemann problem we need to determine the flow quantities in the *Star Region*, see Fig. 4.1. The first step is to find the pressure and particle velocity.

**Proposition 4.6 (Solution for  $p_*$  and  $u_*$ ).** *The solution for pressure  $p_*$  and particle velocity  $u_*$  of the Riemann problem (4.1), (4.2) with the covolume gas Equation of State (4.4) in the unknown Star Region is given by the root of the algebraic equation*

$$f(p, \mathbf{W}_L, \mathbf{W}_R) \equiv f_L(p, \mathbf{W}_L) + f_R(p, \mathbf{W}_R) + \Delta u = 0 , \quad \Delta u \equiv u_R - u_L , \quad (4.84)$$

where the function  $f_L$  is given by

$$f_L(p, \mathbf{W}_L) = \begin{cases} (p - p_L) \left[ \frac{A_L}{p + B_L} \right]^{\frac{1}{2}} & \text{if } p > p_L \text{ (shock)} , \\ \frac{2a_L(1-b\rho_L)}{(\gamma-1)} \left[ \left( \frac{p}{p_L} \right)^{\frac{\gamma-1}{2\gamma}} - 1 \right] & \text{if } p \leq p_L \text{ (rarefaction)} , \end{cases} \quad (4.85)$$

the function  $f_R$  is given by

$$f_R(p, \mathbf{W}_R) = \begin{cases} (p - p_R) \left[ \frac{A_R}{p + B_R} \right]^{\frac{1}{2}} & \text{if } p > p_R \text{ (shock)} , \\ \frac{2a_R(1-b\rho_R)}{(\gamma-1)} \left[ \left( \frac{p}{p_R} \right)^{\frac{\gamma-1}{2\gamma}} - 1 \right] & \text{if } p \leq p_R \text{ (rarefaction)} , \end{cases} \quad (4.86)$$

and the data-dependent constants  $A_L, B_L, A_R, B_R$  are given by

$$\left. \begin{aligned} A_L &= \frac{2(1-b\rho_L)}{(\gamma+1)\rho_L} , & B_L &= \frac{(\gamma-1)}{(\gamma+1)}p_L , \\ A_R &= \frac{2(1-b\rho_R)}{(\gamma+1)\rho_R} , & B_R &= \frac{(\gamma-1)}{(\gamma+1)}p_R . \end{aligned} \right\} \quad (4.87)$$

The particle velocity in the Star Region is given by

$$u_* = \frac{1}{2}(u_L + u_R) + \frac{1}{2} [f_R(p_*) - f_L(p_*)] . \quad (4.88)$$

*Proof.* As for the case of ideal gases the derivation of the function  $f_L$  requires the analysis of the left non-linear wave, which can either be a shock wave or a rarefaction wave. Similarly for the function  $f_R$ .

We first consider  $f_L$  and assume the **left wave is a shock wave** of speed  $S_L$ . Application of the Rankine–Hugoniot Conditions yields the following relations of interest: The density  $\rho_{*L}$  is related to the pressure  $p_*$  behind the left shock via

$$\rho_{*L} = \rho_L \left[ \frac{\frac{p_*}{p_L} + \frac{\gamma-1}{\gamma+1}}{\left( \frac{\gamma-1+2b\rho_L}{\gamma+1} \right) \frac{p_*}{p_L} + \frac{\gamma+1-2b\rho_L}{\gamma+1}} \right] \quad (4.89)$$

The *mass flux*

$$Q_L = -\frac{p_* - p_L}{u_* - u_L} \quad (4.90)$$

is worked out to be

$$Q_L = \left[ \frac{p_* + B_L}{A_L} \right]^{\frac{1}{2}}, \quad (4.91)$$

with  $A_L$  and  $B_L$  as given by (4.87). From (4.90) we obtain

$$u_* = u_L - f_L(p_*, \mathbf{W}_L), \quad (4.92)$$

with

$$f_L(p_*, \mathbf{W}_L) = (p_* - p_L) \left[ \frac{A_L}{p_* + B_L} \right]^{\frac{1}{2}}$$

and

$$A_L = \frac{2(1 - b\rho_L)}{(\gamma + 1)\rho_L}, \quad B_L = \frac{(\gamma - 1)}{(\gamma + 1)}p_L.$$

This is the sought expression for  $f_L$  for the case in which the left wave is a shock wave.

When the **left wave is a rarefaction** we apply the  $I_L(u, a)$  Generalised Riemann Invariant to obtain

$$u_L + \frac{2a_L}{\gamma - 1}(1 - b\rho_L) = u_* + \frac{2a_{*L}}{\gamma - 1}(1 - b\rho_{*L}), \quad (4.93)$$

where  $a_L$  and  $a_{*L}$  denote the sound speed on the left and right states bounding the left rarefaction wave. Now the isentropic law gives

$$\rho_{*L} = \rho_L \left( \frac{1 - b\rho_{*L}}{1 - b\rho_L} \right) \left( \frac{p_*}{p_L} \right)^{\frac{1}{\gamma}}. \quad (4.94)$$

Substitution of  $\rho_{*L}$  from (4.94) into the definition of  $a_{*L}$  reduces equation (4.93) to

$$u_* = u_L - f_L(p_*, \mathbf{W}_L), \quad (4.95)$$

with

$$f_L(p_*, \mathbf{W}_L) = \frac{2a_L(1 - b\rho_L)}{(\gamma - 1)} \left[ \left( \frac{p_*}{\rho_L} \right)^{\frac{\gamma-1}{2\gamma}} - 1 \right] ,$$

which is the required expression for the function  $f_L$  for the case in which the left wave is a rarefaction wave.

We now determine the function  $f_R$ . For a **right shock wave** of speed  $S_R$  we find

$$\rho_{*R} = \rho_R \left[ \frac{\frac{p_*}{\rho_R} + \frac{\gamma-1}{\gamma+1}}{\left( \frac{\gamma-1+2b\rho_R}{\gamma+1} \right) \frac{p_*}{\rho_R} + \frac{\gamma+1-2b\rho_R}{\gamma+1}} \right] . \quad (4.96)$$

The *mass flux*

$$Q_R = \frac{p_* - p_R}{u_* - u_R} \quad (4.97)$$

becomes

$$Q_R = \left[ \frac{p_* + B_R}{A_R} \right]^{\frac{1}{2}} \quad (4.98)$$

and leads to

$$u_* = u_R + f_R(p_*, \mathbf{W}_R) , \quad (4.99)$$

where

$$f_R(p_*, \mathbf{W}_R) = (p_* - p_R) \left[ \frac{A_R}{p_* + B_R} \right]^{\frac{1}{2}} ,$$

with  $A_R$  and  $B_R$  as given by (4.87). This is the sought expression for  $f_R$  for the case in which the right wave is a shock wave.

When the **right wave is a rarefaction** we apply the  $I_R(u, a)$  Generalised Riemann Invariant to obtain

$$u_R - \frac{2a_R}{\gamma - 1} (1 - b\rho_R) = u_* - \frac{2a_{*R}}{\gamma - 1} (1 - b\rho_{*R}) , \quad (4.100)$$

where  $a_{*R}$  and  $a_R$  denote the sound speed on the left and right states bounding the right rarefaction wave. Now the isentropic law gives

$$\rho_{*R} = \rho_R \left( \frac{1 - b\rho_{*R}}{1 - b\rho_R} \right) \left( \frac{p_*}{p_R} \right)^{\frac{1}{\gamma}} . \quad (4.101)$$

Substitution of  $\rho_{*R}$  from (4.101) into the definition of  $a_{*L}$  reduces equation (4.100) to

$$u_* = u_R + f_R(p_*, \mathbf{W}_R) , \quad (4.102)$$

with

$$f_R(p_*, \mathbf{W}_R) = \frac{2a_R(1 - b\rho_R)}{(\gamma - 1)} \left[ \left( \frac{p_*}{p_R} \right)^{\frac{\gamma-1}{2\gamma}} - 1 \right] ,$$

which is the required expression for the function  $f_R$  for the case in which the right wave is a rarefaction wave.

The numerical solution of the algebraic equation  $f(p) = 0$  given by (4.84) yields the pressure  $p_*$ . The particle velocity  $u_*$  may be computed from (4.92) if  $p_* > p_L$  or from (4.95) if  $p_* \leq p_L$  or from (4.99) if  $p_* > p_R$  or from (4.102) if  $p_* \leq p_R$ . It can also be found from a mean value as

$$u_* = \frac{1}{2}(u_L + u_R) + \frac{1}{2}(f_R(p_*) - f_L(p_*))$$

and the proposition is thus proved.

### 4.7.2 Numerical Solution for Pressure

The numerical solution of the pressure equation is carried out by a Newton–Raphson iteration scheme, as done for the ideal gas case in Sect. 4.3. This requires the calculation of the derivative of the function  $f(p)$ . Details of this are found in [498]. A possible guess value  $p_0$  for the iteration scheme is given by a *Two–Rarefaction* approximation to  $p_*$ , namely

$$p_{\text{TR}} = \left[ \frac{(1 - b\rho_L)a_L + (1 - b\rho_R)a_R - \frac{1}{2}(\gamma - 1)(u_R - u_L)}{(1 - b\rho_L)a_L/p_L^{\frac{\gamma-1}{2\gamma}} + (1 - b\rho_R)a_R/p_R^{\frac{\gamma-1}{2\gamma}}} \right]^{\frac{2\gamma}{\gamma-1}}, \quad (4.103)$$

and results from the exact function (4.84) for pressure under the assumption that the two non–linear waves are rarefaction waves. If the solution actually consists of two rarefactions then  $p_{\text{TR}}$  is exact and no iteration is required. A second guess value results from a linearised solution based on primitive variables. This is

$$\left. \begin{aligned} p_0 &= \max(TOL, p_{\text{PV}}), \\ p_{\text{PV}} &= \frac{1}{2}(p_L + p_R) - \frac{1}{8}(u_R - u_L)(\rho_L + \rho_R)(a_L + a_R). \end{aligned} \right\} \quad (4.104)$$

It is worth remarking here that the form of the  $p_{\text{PV}}$  approximation for covolume remains identical to that for the ideal gas case. The equation of state makes its input through the sound speeds  $a_L$  and  $a_R$ . The *Two–Shock* approximation is applied in exactly the same way as in the ideal gas case with the appropriate definitions for the quantities involved. See Sect. 9.4 of Chapt. 9.

### 4.7.3 The Complete Solution

So far, we know how to find  $p_*$  and  $u_*$ . The density values  $\rho_{*L}$  and  $\rho_{*R}$  follow from the appropriate (determined by the value of  $p_*$ ) left and right wave relations. If the left wave is a shock wave,  $p_* > p_L$ , then  $\rho_{*L}$  follows from (4.89). As in the ideal gas case the shock speed  $S_L$  is

$$S_L = u_L - Q_L/\rho_L,$$

where the mass flux  $Q_L$  is given by (4.91). If the left wave is a rarefaction wave,  $p_* \leq p_L$ ,  $\rho_{*L}$  follows from (4.94). The rarefaction fan is enclosed by the characteristics of speeds

$$S_{HL} = u_L - a_L, \quad S_{TL} = u_* - a_{*L}.$$

For a right shock wave,  $p_* > p_R$ ,  $\rho_{*L}$  is given by (4.96) and the shock speed is

$$S_R = u_R + Q_R / \rho_R,$$

where the mass flux  $Q_R$  is given by (4.98). For a right rarefaction,  $p_* \leq p_R$ ,  $\rho_{*L}$  follows from (4.101) and the wave is enclosed by the characteristics of speeds

$$S_{TR} = u_* + a_{*R}, \quad S_{HR} = u_R + a_R.$$

Solution values inside rarefaction fans cannot be obtained directly, as in the ideal gas case. For covolume gases an extra iterative procedure is needed.

#### 4.7.4 Solution Inside Rarefactions

Unlike the ideal gas case the solution inside rarefaction waves for covolume gases is not direct. An extra iterative procedure is needed. Here we give the details for the case of a left rarefaction. Consider a general point  $(x, t)$  inside a left rarefaction fan and a characteristic ray through the origin and the point  $(x, t)$ . The slope of the characteristic is

$$\frac{x}{t} = u - a. \quad (4.105)$$

Use of the  $I_L(u, a)$  Generalised Riemann Invariant allows us to write

$$a \left[ 1 + \frac{2}{\gamma - 1} (1 - b\rho) \right] = I_L(u_L, a_L) - \frac{x}{t}, \quad (4.106)$$

where  $\rho$ ,  $u$  and  $a$  are the sought solution values at  $(x, t)$  inside the left rarefaction wave. Use of the isentropic relation for the covolume EOS gives

$$p = p_L \left( \frac{1 - b\rho_L}{\rho_L} \right)^\gamma \left( \frac{\rho}{1 - b\rho} \right)^\gamma. \quad (4.107)$$

Further algebraic manipulations lead to a non-linear algebraic equation for the density  $\rho$  at the point  $(x, t)$  inside the left rarefaction fan, namely

$$Z_L(\rho) \equiv \rho^{\gamma-1} (\gamma + 1 - 2b\rho)^2 - \beta_L (1 - b\rho)^{\gamma+1} = 0, \quad (4.108)$$

where the constant  $\beta_L$  is given by

$$\beta_L = \frac{\{(\gamma - 1)[I_L(u_L, a_L) - \frac{x}{t}]\}^2}{\gamma p_L [(1 - b\rho_L)/\rho_L]^\gamma}, \quad (4.109)$$

where  $I_L(u_L, a_L)$  is the left Riemann invariant evaluated on the left data state.

Equation (4.108) is solved numerically for  $\rho$  using a Newton–Raphson iteration, for which one needs the derivative

$$Z'_L(\rho) = (\gamma + 1)[b\beta_L(1 - b\rho)^\gamma + (\gamma + 1 - 2b\rho)(\gamma - 1 - 2b\rho)\rho^{\gamma-2}] . \quad (4.110)$$

Once  $\rho$  has been found to a given accuracy the pressure  $p$  follows immediately from equation (4.107). The sound speed  $a$  follows from definition (4.83) and the velocity  $u$  follows directly from (4.105).

The solution at a point  $(x, t)$  inside the right rarefaction fan is found in an entirely analogous way. In this case the density function  $Z_R(\rho)$  is

$$Z_R(\rho) \equiv \rho^{\gamma-1}(\gamma + 1 - 2b\rho)^2 - \beta_R(1 - b\rho)^{\gamma+1} = 0 , \quad (4.111)$$

where the constant  $\beta_R$  is given by

$$\beta_R = \frac{\{(\gamma - 1)[I_R(u_R, a_R) - \frac{x}{t}]\}^2}{\gamma p_R[(1 - b\rho_R)/\rho_R]^\gamma} , \quad (4.112)$$

where  $I_R(u_R, a_R)$  is the right Riemann invariant evaluated on the right data state.

The solution  $\mathbf{W} = (\rho, u, p)^T$  of the Riemann problem for the Euler equations with the covolume equation of state is now known at any point  $(x, t)$  in the relevant domain. The solution sampling procedure is entirely analogous to the ideal gas case of Sect. 4.5 and is omitted.

**Exercise 4.7.** Write a flow chart for sampling the solution of the Riemann problem for covolume gases at any point  $(x, t)$  in a domain of interest  $x_L \leq x \leq x_R$ ,  $t > 0$ , with  $x_L < 0$  and  $x_R > 0$ .

**Solution 4.8.** (Left to the reader).

## 4.8 The Split Multi-Dimensional Case

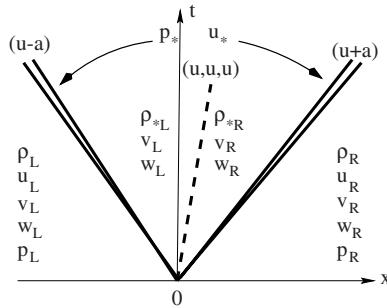
For the purpose of using the Riemann problem solution in conjunction with numerical methods of the Godunov type, see Chaps. 5, 6, 14 and 16, it is useful to solve the split multi-dimensional Riemann problem. See Chap. 3. The  $x$ -split Riemann problem is the IVP for

$$\left. \begin{aligned} \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x &= \mathbf{0} , \\ \mathbf{U} &= \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ E \end{bmatrix} , \quad \mathbf{F}(\mathbf{U}) = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ \rho uw \\ u(E + p) \end{bmatrix} , \end{aligned} \right\} \quad (4.113)$$

with initial conditions

$$\mathbf{U}(x, 0) = \begin{cases} \mathbf{U}_L & \text{if } x < 0, \\ \mathbf{U}_R & \text{if } x > 0. \end{cases} \quad (4.114)$$

We note here that the  $x$ -direction will in general be understood as the direction *normal* to the boundary of a domain in two or three-dimensional space, see Sect. 3.2.4 of Chap. 3 and Sect. 16.7.3 of Chap. 16.



**Fig. 4.18.** Structure of the solution of the Riemann problem for the split three-dimensional case

As seen in Sect. 3.2.4 of Chap. 3 the exact solution to this problem is virtually identical to that of the genuine one-dimensional problem discussed previously. Fig. 4.18 shows the structure of the solution in terms of primitive variables  $\mathbf{W} = (\rho, u, v, w, p)^T$ . The outer non-linear waves are exactly the same as in the one-dimensional case. The multiplicity 3 of the eigenvalue  $\lambda = u$  generates three, *coincident middle waves*. So effectively there are three waves that separate four constant states  $\mathbf{W}_L$ ,  $\mathbf{W}_{*L}$ ,  $\mathbf{W}_{*R}$  and  $\mathbf{W}_R$ . In the *Star Region* between the left and right waves the solution for pressure, *normal velocity* and density is exactly the same as in the one-dimensional case. The two extra quantities  $v$  and  $w$  (tangential velocity components) only jump across the middle wave from their left data values to their right data values.

In summary, the solution to the  $x$ -split three-dimensional Riemann problem (4.113), (4.114) is exactly the same as that for the one-dimensional Riemann problem for the quantities  $\rho$ ,  $u$  and  $p$ ; for  $q = v$  and  $q = w$  we have, in addition,

$$q(x, t) = \begin{cases} q_L & \text{if } \frac{x}{t} < u_*, \\ q_R & \text{if } \frac{x}{t} > u_*. \end{cases} \quad (4.115)$$

As a matter of fact, for any passively advected quantity  $q(x, y, z, t)$  by the fluid, that is

$$q_t + uq_x + vq_y + wq_z = 0, \quad (4.116)$$

one can, by making use of the continuity equation, derive a new conservation law

$$(\rho q)_t + (u\rho q)_x + (v\rho q)_y + (w\rho q)_z = 0. \quad (4.117)$$

When this conservation law is added to the set of one-dimensional Euler equations the  $x$ -split Riemann problem has solution for  $\rho$ ,  $u$  and  $p$  as described earlier and the solution for the advected quantity  $q$  is as given by (4.115). One may have several advected quantities such as  $q$ . In chemically reactive compressible flows the quantity  $q$  may stand for the concentration of a chemical species, the progress variable of a chemical reaction or a fluid interface parameter.

Finally, we remark that although the exact solution to the split three-dimensional Riemann problem for any passively advected quantity, such as the tangential velocity components or concentration of chemical species, is so simple, approximate Riemann solvers may produce solutions for these quantities that are completely incorrect.

A highly relevant reference on the general theme of this chapter is the paper by Zhang [601]. Exact and approximate Riemann solvers for compressible liquids with various equations of state are given in [263]. For exact Riemann solvers for real gases see Colella and Glaz [135] and Menikoff and Plohr [349]. Approximate Riemann solvers for the ideal Euler equations are presented in Chaps. 9 to 12.

The methodology can be applied quite directly to solve the Riemann problem for other hyperbolic systems. For the steady supersonic Euler equations, exact Riemann solvers have been given by, amongst others, Marshall and Plohr [340]; Honma, Wada and Inomata [257]; Dawes [151]; Toro and Chou [533] and Toro and Chakravarthy [532]. Approximate Riemann solvers for the steady supersonic Euler equations have been given by Roe [407], Pandolfi [374], Toro and Chou [533] and Toro and Chakravarthy [532].

## 4.9 FORTRAN Program for Exact Riemann Solver

A listing of a FORTRAN 77 program to compute the exact solution to the Riemann problem for the one-dimensional time-dependent Euler equations for ideal gases is now given. The data file is called `exact.ini`. The main program calls two routines: `STARPU` and `SAMPLE`. The first routine solves iteratively for the pressure  $p_*$  and then computes the particle speed  $u_*$ ; the guessed value for the iteration is provided by the subroutine `GUESSP`. The subroutine `SAMPLE` finds, for given  $p_*$ ,  $u_*$  and  $S = x/t$ , the solution of the Riemann problem at the point  $(x, t)$ . This routine can then be utilised in numerical methods to solve the general initial boundary value problem for the Euler equations. For the Godunov first-order method, see Chap. 6, one calls `SAMPLE` with  $S = 0$ . For Glimm's method, see Chap. 7, one calls `SAMPLE` with  $S = \theta x/t$ , where  $\theta$  is a random number. The source code is also part of the library `NUMERICA` [519].

```
*-----*
*  
C      EXACT RIEMANN SOLVER  
C      FOR THE EULER EQUATIONS  
*  
C      Name of program: HE-E1RPEXACT  
*  
C      Purpose: to solve the Riemann problem exactly,  
C                  for the time dependent one dimensional  
C                  Euler equations for an ideal gas  
*  
C      Input file: exact.ini  
C      Output file: exact.out (exact solution)  
*  
C      Programer: E. F. Toro  
*  
C      Last revision: February 1st 1999  
*  
C      Theory is found in Chapter 4 of Reference 1  
*  
C      1. Toro, E. F., "Riemann Solvers and Numerical  
C         Methods for Fluid Dynamics"  
C         Springer-Verlag,  
C         Second Edition, 1999  
*  
C      This program is part of  
*  
C      NUMERICA  
C      A Library of Source Codes for Teaching,  
C      Research and Applications,  
C      by E. F. Toro  
C      Published by NUMERITEK LTD,  
C      Website: www.numeritek.com  
*-----*
```

IMPLICIT NONE

\*  
C Declaration of variables:  
\*  
 INTEGER I, CELLS  
\*  
 REAL GAMMA, G1, G2, G3, G4, G5, G6, G7, G8,  
 & DL, UL, PL, CL, DR, UR, PR, CR,

```

&      DIAPH, DOMLEN, DS, DX, PM, MPA, PS, S,
&      TIMEOUT, UM, US, XPOS
*
COMMON /GAMMAS/ GAMMA, G1, G2, G3, G4, G5, G6, G7, G8
COMMON /STATES/ DL, UL, PL, CL, DR, UR, PR, CR
*
OPEN(UNIT = 1,FILE = 'exact.ini',STATUS = 'UNKNOWN')
*
C Initial data and parameters are read in
*
READ(1,*)DOMLEN      ! Domain length
READ(1,*)DIAPH        ! Initial discontinuity position
READ(1,*)CELLS         ! Number of computing cells
READ(1,*)GAMMA         ! Ratio of specific heats
READ(1,*)TIMEOUT       ! Output time
READ(1,*)DL            ! Initial density on left state
READ(1,*)UL            ! Initial velocity on left state
READ(1,*)PL            ! Initial pressure on left state
READ(1,*)DR            ! Initial density on right state
READ(1,*)UR            ! Initial velocity on right state
READ(1,*)PR            ! Initial pressure on right state
READ(1,*)MPA           ! Normalising constant
*
CLOSE(1)
*
C Compute gamma related constants
*
G1 = (GAMMA - 1.0)/(2.0*GAMMA)
G2 = (GAMMA + 1.0)/(2.0*GAMMA)
G3 = 2.0*GAMMA/(GAMMA - 1.0)
G4 = 2.0/(GAMMA - 1.0)
G5 = 2.0/(GAMMA + 1.0)
G6 = (GAMMA - 1.0)/(GAMMA + 1.0)
G7 = (GAMMA - 1.0)/2.0
G8 = GAMMA - 1.0
*
C Compute sound speeds
*
CL = SQRT(GAMMA*PL/DL)
CR = SQRT(GAMMA*PR/DR)
*
C The pressure positivity condition is tested for
*
IF(G4*(CL+CR).LE.(UR-UL))THEN
*
```

```
C      The initial data is such that vacuum is generated.
C      Program stopped.
*
      WRITE(6,*)
      WRITE(6,*)'***Vacuum is generated by data***'
      WRITE(6,*)'***Program stopped***'
      WRITE(6,*)
*
      STOP
ENDIF
*
C      Exact solution for pressure and velocity in star
C      region is found
*
      CALL STARPU(PM, UM, MPA)
*
      DX = DOMLEN/REAL(CELLS)
*
C      Complete solution at time TIMEOUT is found
*
      OPEN(UNIT = 2,FILE = 'exact.out',STATUS = 'UNKNOWN')
*
      DO 10 I = 1, CELLS
*
         XPOS = (REAL(I) - 0.5)*DX
         S     = (XPOS - DIAPH)/TIMEOUT
*
C      Solution at point (X,T) = ( XPOS - DIAPH,TIMEOUT)
C      is found
*
         CALL SAMPLE(PM, UM, S, DS, US, PS)
*
C      Exact solution profiles are written to exact.out.
*
         WRITE(2, 20)XPOS, DS, US, PS/MPA, PS/DS/G8/MPA
*
10      CONTINUE
*
         CLOSE(2)
*
20      FORMAT(5(F14.6, 2X))
*
         END
*
-----*
```

```

*
      SUBROUTINE STARPU(P, U, MPA)
*
      IMPLICIT NONE
*
C      Purpose: to compute the solution for pressure and
C              velocity in the Star Region
*
C      Declaration of variables
*
      INTEGER I, NRITER
*
      REAL      DL, UL, PL, CL, DR, UR, PR, CR,
&           CHANGE, FL, FLD, FR, FRD, P, POLD, PSTART,
&           TOLPRE, U, UDIFF, MPA
*
      COMMON /STATES/ DL, UL, PL, CL, DR, UR, PR, CR
      DATA TOLPRE, NRITER/1.0E-06, 20/
*
C      Guessed value PSTART is computed
*
      CALL GUESSP(PSTART)
*
      POLD = PSTART
      UDIFF = UR - UL
*
      WRITE(6,*)'-----',
      WRITE(6,*)' Iteration number      Change   ',
      WRITE(6,*)'-----',
*
      DO 10 I = 1, NRITER
*
          CALL PREFUN(FL, FLD, POLD, DL, PL, CL)
          CALL PREFUN(FR, FRD, POLD, DR, PR, CR)
          P      = POLD - (FL + FR + UDIFF)/(FLD + FRD)
          CHANGE = 2.0*ABS((P - POLD)/(P + POLD))
          WRITE(6, 30)I, CHANGE
          IF(CHANGE.LE.TOLPRE)GOTO 20
          IF(P.LT.0.0)P = TOLPRE
          POLD = P
*
10    CONTINUE
*
      WRITE(6,*)'Divergence in Newton-Raphson iteration'
*
```

```

20    CONTINUE
*
C      Compute velocity in Star Region
*
U = 0.5*(UL + UR + FR - FL)
*
WRITE(6,*),-----
WRITE(6,*),      Pressure          Velocity',
WRITE(6,*),-----
WRITE(6,40)P/MPA, U
WRITE(6,*),-----
*
30    FORMAT(5X, I5,15X, F12.7)
40    FORMAT(2(F14.6, 5X))
*
      RETURN
      END
*
*-----*
*
SUBROUTINE GUESSP(PM)
*
C      Purpose: to provide a guess value for pressure
C              PM in the Star Region. The choice is made
C              according to adaptive Riemann solver using
C              the PVRS, TRRS and TSRS approximate
C              Riemann solvers. See Sect. 9.5 of Chapt. 9
C              of Ref. 1
*
      IMPLICIT NONE
*
C      Declaration of variables
*
      REAL      DL, UL, PL, CL, DR, UR, PR, CR,
&           GAMMA, G1, G2, G3, G4, G5, G6, G7, G8,
&           CUP, GEL, GER, PM, PMAX, PMIN, PPV, PQ,
&           PTL, PTR, QMAX, QUSER, UM
*
COMMON /GAMMAS/ GAMMA, G1, G2, G3, G4, G5, G6, G7, G8
COMMON /STATES/ DL, UL, PL, CL, DR, UR, PR, CR
*
      QUSER = 2.0
*
C      Compute guess pressure from PVRS Riemann solver
*
```

```

CUP  = 0.25*(DL + DR)*(CL + CR)
PPV  = 0.5*(PL + PR) + 0.5*(UL - UR)*CUP
PPV  = AMAX1(0.0, PPV)
PMIN = AMIN1(PL, PR)
PMAX = AMAX1(PL, PR)
QMAX = PMAX/PMIN
*
* IF(QMAX.LE.QUSER.AND.
& (PMIN.LE.PPV.AND.PPV.LE.PMAX))THEN
*
C      Select PVRS Riemann solver
*
      PM = PPV
ELSE
    IF(PPV.LT.PMIN)THEN
*
C      Select Two-Rarefaction Riemann solver
*
      PQ  = (PL/PR)**G1
      UM  = (PQ*UL/CL + UR/CR +
&          G4*(PQ - 1.0))/(PQ/CL + 1.0/CR)
      PTL = 1.0 + G7*(UL - UM)/CL
      PTR = 1.0 + G7*(UM - UR)/CR
      PM  = 0.5*(PL*PTL**G3 + PR*PTR**G3)
ELSE
*
C      Select Two-Shock Riemann solver with
C      PVRS as estimate
*
      GEL = SQRT((G5/DL)/(G6*PL + PPV))
      GER = SQRT((G5/DR)/(G6*PR + PPV))
      PM  = (GEL*PL + GER*PR - (UR - UL))/(GEL + GER)
ENDIF
ENDIF
*
      RETURN
END
*
*-----*
*
      SUBROUTINE PREFUN(F, FD, P, DK, PK, CK)
*
C      Purpose: to evaluate the pressure functions
C              FL and FR in exact Riemann solver
*
```

```

IMPLICIT NONE
*
C Declaration of variables
*
REAL AK, BK, CK, DK, F, FD, P, PK, PRAT, QRT,
& GAMMA, G1, G2, G3, G4, G5, G6, G7, G8
*
COMMON /GAMMAS/ GAMMA, G1, G2, G3, G4, G5, G6, G7, G8
*
IF(P.LE.PK)THEN
*
C Rarefaction wave
*
PRAT = P/PK
F = G4*CK*(PRAT**G1 - 1.0)
FD = (1.0/(DK*CK))*PRAT**(-G2)
ELSE
*
C Shock wave
*
AK = G5/DK
BK = G6*PK
QRT = SQRT(AK/(BK + P))
F = (P - PK)*QRT
FD = (1.0 - 0.5*(P - PK)/(BK + P))*QRT
ENDIF
*
RETURN
END
*
*-----*
*
SUBROUTINE SAMPLE(PM, UM, S, D, U, P)
*
C Purpose: to sample the solution throughout the wave
C pattern. Pressure PM and velocity UM in the
C Star Region are known. Sampling is performed
C in terms of the 'speed' S = X/T. Sampled
C values are D, U, P
*
C Input variables : PM, UM, S, /GAMMAS/, /STATES/
C Output variables: D, U, P
*
IMPLICIT NONE
*
```

```

C      Declaration of variables
*
REAL      DL, UL, PL, CL, DR, UR, PR, CR,
&        GAMMA, G1, G2, G3, G4, G5, G6, G7, G8,
&        C, CML, CMR, D, P, PM, PML, PMR, S,
&        SHL, SHR, SL, SR, STL, STR, U, UM
*
COMMON /GAMMAS/ GAMMA, G1, G2, G3, G4, G5, G6, G7, G8
COMMON /STATES/ DL, UL, PL, CL, DR, UR, PR, CR
*
IF(S.LE.UM)THEN
*
C      Sampling point lies to the left of the contact
C      discontinuity
*
IF(PM.LE.PL)THEN
*
C      Left rarefaction
*
SHL = UL - CL
*
IF(S.LE.SHL)THEN
*
C      Sampled point is left data state
*
D = DL
U = UL
P = PL
ELSE
CML = CL*(PM/PL)**G1
STL = UM - CML
*
IF(S.GT.STL)THEN
*
C      Sampled point is Star Left state
*
D = DL*(PM/PL)**(1.0/GAMMA)
U = UM
P = PM
ELSE
*
C      Sampled point is inside left fan
*
U = G5*(CL + G7*UL + S)
C = G5*(CL + G7*(UL - S))

```

```

D = DL*(C/CL)**G4
P = PL*(C/CL)**G3
ENDIF
ENDIF
ELSE
*
C      Left shock
*
PML = PM/PL
SL = UL - CL*SQRT(G2*PML + G1)
*
IF(S.LE.SL)THEN
*
C      Sampled point is left data state
*
D = DL
U = UL
P = PL
*
ELSE
*
C      Sampled point is Star Left state
*
D = DL*(PML + G6)/(PML*G6 + 1.0)
U = UM
P = PM
ENDIF
ENDIF
ELSE
*
C      Sampling point lies to the right of the contact
C      discontinuity
*
IF(PM.GT.PR)THEN
*
C      Right shock
*
PMR = PM/PR
SR = UR + CR*SQRT(G2*PMR + G1)
*
IF(S.GE.SR)THEN
*
C      Sampled point is right data state
*
D = DR

```

```

U = UR
P = PR
ELSE
*
C      Sampled point is Star Right state
*
D = DR*(PMR + G6)/(PMR*G6 + 1.0)
U = UM
P = PM
ENDIF
ELSE
*
C      Right rarefaction
*
SHR = UR + CR
*
IF(S.GE.SHR)THEN
*
C      Sampled point is right data state
*
D = DR
U = UR
P = PR
ELSE
CMR = CR*(PM/PR)**G1
STR = UM + CMR
*
IF(S.LE.STR)THEN
*
C      Sampled point is Star Right state
*
D = DR*(PM/PR)**(1.0/GAMMA)
U = UM
P = PM
ELSE
*
C      Sampled point is inside left fan
*
U = G5*(-CR + G7*UR + S)
C = G5*(CR - G7*(UR - S))
D = DR*(C/CR)**G4
P = PR*(C/CR)**G3
ENDIF
ENDIF
ENDIF

```

```
ENDIF
*
RETURN
END
*
*-----*
```

## Notions on Numerical Methods

We assume the reader to be familiar with some basic concepts on numerical methods for partial differential equations in general. In particular, we shall assume the concepts of truncation error, order of accuracy, consistency, modified equation, stability and convergence. For background on these concepts the reader may consult virtually any standard book on numerical methods for differential equations. As general references, useful textbooks are those of Smith [450], Anderson et. al. [7], Mitchell and Griffiths [352], Roache [405], Richtmyer and Morton [402], Hoffmann [253] and Fletcher [192]. Very relevant textbooks to the main themes of this book are Sod [454], Holt [254], Hirsch Volumes I [251] and II [252], LeVeque [308], Godlewski and Raviart [215], Kröner [291] and Thomas [484].

The contents of this chapter are designed specifically to provide the necessary background for the application of *high-resolution upwind and centred numerical methods* to hyperbolic conservation laws. Our prime objective is to present the basic Godunov method [216] in a simple setting so as to make the application of upwind methods to non-linear systems of conservation laws an easier task in the forthcoming chapters. Essential background material is given in Chap. 2. For those who are absolute beginners in the field I would recommend the following self study programme as a way of obtaining more benefit from this text book: (a) read chapters 2 to 4 of the book by Hoffmann [253] and do exercises, (b) read chapters 7 to 10 of the book by Hirsch Vol. I [251] and do exercises.

### 5.1 Discretisation: Introductory Concepts

Our concern is the utilisation of numerical methods for solving partial differential equations (PDEs). Numerical methods replace the *continuous* problem represented by the PDEs by a *finite* set of *discrete* values. These are obtained by first discretising the domain of the PDEs into a finite set of points or volumes via a *mesh* or *grid*. The corresponding discretisation of the PDEs

on the grid results in discrete values. In the Finite Difference approach one regards these values as *point values* defined at grid points. The Finite Volume approach regards these discrete values as *averages over finite volumes*. We are mostly interested in the second approach but for the purpose of introducing some of the basic concepts in numerical methods we also consider the Finite Difference approach. For most of this chapter we restrict the discussion to model problems, such as the model PDE

$$u_t + au_x = 0 ,$$

with  $u = u(x, t)$  and  $a \neq 0$  a constant wave propagation speed. See Sect. 2.1 of Chap. 2. In this equation there are two partial derivatives, namely a time derivative  $u_t$  and a spatial derivative  $u_x$ .

### 5.1.1 Approximation to Derivatives

Given a sufficiently smooth function  $g(x)$ , by using Taylor's theorem, we can always find the value of  $g(x)$  at any neighbouring point  $x_0 + \Delta x$  of  $x = x_0$  if we know  $g(x)$  and all its derivatives  $g^{(k)}(x)$  at  $x = x_0$ , that is

$$g(x_0 + \Delta x) = g(x_0) + \sum_k \frac{(\Delta x)^k}{k!} g^{(k)}(x_0) . \quad (5.1)$$

By truncating the Taylor series one can obtain approximations to  $g(x_0 + \Delta x)$ . Also one can obtain approximations to derivatives of  $g(x)$ . Consider a function  $g(x)$  and three equally-spaced points  $x_0 - \Delta x$ ,  $x_0$  and  $x_0 + \Delta x$ , as shown in Fig. 5.1. As an illustrative example we shall derive three approximations to the first derivative of  $g(x)$  at the point  $x_0$ . By neglecting terms of third order and higher,  $O(\Delta x^3)$ , we can write

$$g(x_0 + \Delta x) = g(x_0) + \Delta x g^{(1)}(x_0) + \frac{(\Delta x)^2}{2} g^{(2)}(x_0) + O(\Delta x^3) \quad (5.2)$$

and

$$g(x_0 - \Delta x) = g(x_0) - \Delta x g^{(1)}(x_0) + \frac{(\Delta x)^2}{2} g^{(2)}(x_0) + O(\Delta x^3) . \quad (5.3)$$

Neglecting second order terms in (5.2) leads immediately to an approximation to the first derivative  $g^{(1)}(x)$  of  $g(x)$ , that is

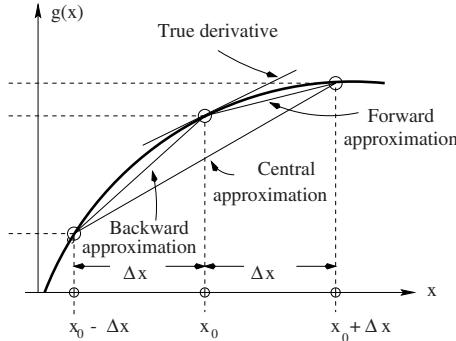
$$g^{(1)}(x_0) = \frac{g(x_0 + \Delta x) - g(x_0)}{\Delta x} + O(\Delta x) . \quad (5.4)$$

This is a first-order approximation to  $g^{(1)}(x)$  at  $x = x_0$ , the leading term in the remaining part contains  $\Delta x$  to the power unity. It is a *one-sided* approximation to the first derivative of  $g(x)$ , usually called a *forward* finite difference

approximation; the points used by the approximation are  $x_0$  (the point at which the derivative is sought) and its neighbour  $x_0 + \Delta x$  on the right hand side. From (5.3) we can also obtain a *backward* first-order approximation

$$g^{(1)}(x_0) = \frac{g(x_0) - g(x_0 - \Delta x)}{\Delta x} + O(\Delta x). \quad (5.5)$$

Note that by subtracting (5.3) from (5.2) we obtain



**Fig. 5.1.** Finite difference approximations to the first derivative of a function  $g(x)$  at the point  $x_0$ : backward, centred and forward

$$g^{(1)}(x_0) = \frac{g(x_0 + \Delta x) - g(x_0 - \Delta x)}{2\Delta x} + O(\Delta x^2). \quad (5.6)$$

This is a *central* finite-difference approximation to  $g^{(1)}(x_0)$  and is second-order accurate; it uses the left and right hand side neighbours of  $x = x_0$ . It can also be obtained by taking a mean value of the forward and backward approximations (5.4) and (5.5). For the problems of our concern here the distinction between one-sided and central approximations is significant. *Upwind Methods* may be viewed as one-sided differencing schemes. The question of which side is also of paramount importance.

### 5.1.2 Finite Difference Approximation to a PDE

Consider the Initial Boundary Value Problem (IBVP) for the linear advection equation in the domain  $[0, L] \times [0, T]$  on the  $x-t$  plane. This consists of a PDE together with initial condition (IC) and boundary conditions (BCs), namely

$$\left. \begin{aligned} \text{PDE : } & u_t + au_x = 0, \\ \text{IC : } & u(x, 0) = u_0(x), \\ \text{BCs : } & u(0, t) = u_l(t), \quad u(L, t) = u_r(t). \end{aligned} \right\} \quad (5.7)$$

Solving this IBVP means evolving the solution  $u(x, t)$  in time starting from the initial condition  $u_0(x)$  at time  $t = 0$  and subject to boundary conditions. For the moment we assume that the boundary constraints take the form of prescribed boundary functions  $u_l(t)$  and  $u_r(t)$ . The application of boundary conditions is intimately linked to the physics of the problem at hand. A more meaningful discussion is presented in Chap. 6 in the context of physically more meaningful systems of PDEs.

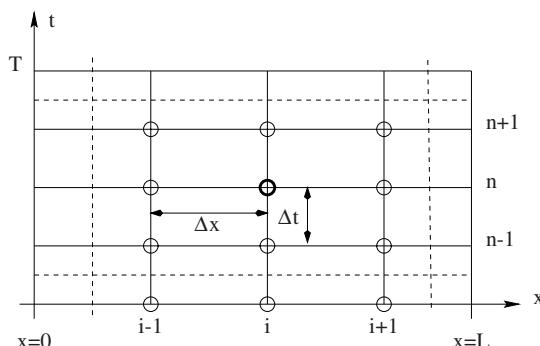
A possible finite difference mesh for discretising the domain is depicted in Fig. 5.2. This is a regular grid of dimensions  $\Delta x$  (spacing of grid points in the  $x$ -direction) by  $\Delta t$  (spacing in the  $t$ -direction). In general if  $[0, L]$  is discretised by  $M$  equally spaced grid points then

$$\Delta x = \frac{L}{M-1}. \quad (5.8)$$

The points of the mesh are positioned at  $(i\Delta x, n\Delta t)$  on the  $x-t$  plane, with  $i = 0, \dots, M$  and  $n = 0, \dots$ . Often we shall use the notation  $x_i = i\Delta x$ ,  $t^n = n\Delta t$ . The discrete values of the function  $u(x, t)$  at  $(i\Delta x, n\Delta t)$  will be denoted by

$$u_i^n \equiv u(i\Delta x, n\Delta t) \equiv u(x_i, t^n).$$

The superscript  $n$  refers to the time discretisation and is called the *time level*. The subscript  $i$  refers to the space discretisation and designates the spatial position in the mesh. We shall also use the symbol  $u_i^n$  to denote an *approximation* to the exact mesh value  $u(i\Delta x, n\Delta t)$ . The distinction will be made at appropriate places. If the IBVP (5.7) has given data at a time level



**Fig. 5.2.** Finite difference discretisation of domain on  $x-t$  plane. Regular mesh of dimensions  $\Delta x \times \Delta t$  is assumed

$n$ , say, this can be represented by a set of discrete values  $u_i^n, i = 0, \dots, M$ . Solving (5.7) requires finding the solution at the next time level  $n + 1$ , that is, we want to find the set  $u_i^{n+1}, i = 0, \dots, M$ . The extreme boundary points  $u_0^n, u_M^n$  are determined by the particular *boundary conditions* to be applied. For the moment we simply assume that these are prescribed for all time levels.

Consider now approximations to the time derivative  $u_t$  at the grid point  $(x_i, t^n)$ . A first-order forward approximation gives

$$u_t = \frac{u_i^{n+1} - u_i^n}{\Delta t}. \quad (5.9)$$

For the spatial derivatives one could use the second-order central approximation

$$u_x = \frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x}. \quad (5.10)$$

By replacing  $u_t$  and  $u_x$  in the PDE in (5.7) by their respective approximations (5.9) and (5.10) we obtain

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} + a \left[ \frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x} \right] = 0. \quad (5.11)$$

This is the discrete analogue of the PDE in (5.7). The differential equation has been replaced by a finite difference equation. As all values at the time level  $n$  in (5.11) are prescribed data values at the initial time, or have already been computed, we can solve for the *single unknown*  $u_i^{n+1}$  at the new time level as

$$u_i^{n+1} = u_i^n - \frac{1}{2}c [u_{i+1}^n - u_{i-1}^n], \quad (5.12)$$

where

$$c = \frac{\Delta t a}{\Delta x} = \frac{a}{\Delta x / \Delta t} \quad (5.13)$$

is a *dimensionless* quantity known as the *Courant number*; it is also known as the Courant–Friedrichs–Lowy number, or CFL number. This quantity can be regarded as the ratio of two speeds, namely the wave propagation speed  $a$  in the partial differential equation in (5.7) and the *grid speed*  $\Delta x / \Delta t$  defined by the discretisation of the domain. Formula (5.12) provides an *explicit* scheme for evolving the solution in time and has resulted from approximating the time and space derivatives of the PDE in (5.7) by first and second order finite differences respectively. This appears to be a reasonable step.

It is disappointing however, and perhaps surprising, to realise that the resulting scheme (5.12) is totally useless. It is *unconditionally unstable*. This can be seen by performing a von Neumann stability analysis. Consider the *trial solution*  $u_i^n = A^n e^{Ii\theta}$ .  $A$  is the amplitude,  $\theta = P\Delta x$  is the phase angle,  $P$  is the wave number in the  $x$ -direction,  $\lambda = 2\pi/P$  is the wave length and  $I = \sqrt{-1}$  is the unit complex number. Substitution of the trial solution into the scheme (5.12) gives  $A = 1 - Ic \sin \theta$ . For stability one requires  $\|A\| \leq 1$ . But note that  $\|A\| = 1 + c^2 \sin^2 \theta \geq 1$  and thus the scheme is unstable under all circumstances, unconditionally unstable.

## 5.2 Selected Difference Schemes

We study some of the most well-known finite difference schemes, of first and second order of accuracy.

### 5.2.1 The First Order Upwind Scheme

One way of remedying our failed attempt at devising a useful numerical method for the PDE in (5.7) is to replace the central finite difference approximation to the spatial derivative  $u_x$  by a first-order *one-sided approximation*. Two choices are

$$u_x = \frac{u_i^n - u_{i-1}^n}{\Delta x}, \quad (5.14)$$

$$u_x = \frac{u_{i+1}^n - u_i^n}{\Delta x}. \quad (5.15)$$

It turns out that only one of these yields a useful numerical scheme. The correct choice of either (5.14) or (5.15) will depend on the *sign of the wave propagation speed*  $a$  of the differential equation in (5.7). Suppose  $a$  is positive, then (5.14) together with (5.9) give the scheme

$$u_i^{n+1} = u_i^n - c(u_i^n - u_{i-1}^n). \quad (5.16)$$

A von-Neumann stability analysis of (5.16) gives

$$\|A\|^2 = (1 - c)^2 + 2c(1 - c) \cos \theta + c^2,$$

from which it follows that the scheme is stable if the CFL number  $c$  lies between zero and unity; it is *conditionally stable* with *stability condition*

$$0 \leq c \leq 1. \quad (5.17)$$

Recall that the CFL number depends on the speed  $a$ , the mesh spacing  $\Delta x$  and the time step size  $\Delta t$ . Of these,  $a$  is prescribed at the outset,  $\Delta x$  is chosen on desired accuracy or on computing resources available. We are left with some freedom to choose  $\Delta t$ , the time step size. The stability restriction (5.17) on  $c$  means a restriction on  $\Delta t$  and thus we are not free to choose the time step at our will, at least for the schemes under discussion.

Scheme (5.16) is called the *first-order upwind method* and is due to Courant, Isaacson and Rees [144]; we shall also call it the *CIR scheme*. The key feature of this numerical method is the fact that the discretisation has been performed according to the sign of the wave propagation speed  $a$  in the differential equation. The physics and mathematics embodied in the PDE are intimately linked to the discretisation procedure. The term *upwind*, or upstream, refers to the fact that spatial differencing is performed using mesh points on the side from which information (wind) flows; see Fig. 5.3. For positive  $a$  the upwind side is the left side and for negative  $a$  the upwind side is the right side.

Suppose that for positive  $a$  we use the *downwind* information to perform the spatial differencing. This gives the *downwind scheme*

$$u_i^{n+1} = u_i^n - c(u_{i+1}^n - u_i^n), \quad (5.18)$$

which can easily be checked to be *unconditionally unstable*. Thus, in order to obtain a useful one-sided scheme one must perform the spatial differencing according to the sign of the speed  $a$  of the PDE. For negative speed  $a$  the upwind scheme is (5.18).

In order to formulate the upwind scheme, in a unified manner, for both positive and negative wave speeds  $a$  we introduce the following notation

$$a^+ = \max(a, 0) = \frac{1}{2}(a + |a|), \quad a^- = \min(a, 0) = \frac{1}{2}(a - |a|), \quad (5.19)$$

where  $|a|$  denotes the absolute value of  $a$ , that is,

$$|a| = a \text{ if } a \geq 0, \quad |a| = -a \text{ if } a \leq 0.$$

It can easily be verified that for any value of the speed  $a$  the speeds  $a^+$  and  $a^-$  satisfy

$$a^+ \geq 0, \quad a^- \leq 0.$$

Clearly, for  $a \geq 0$  one has  $a^+ = a$  and  $a^- = 0$ ; for  $a \leq 0$  one has  $a^+ = 0$  and  $a^- = a$ . Based on the speeds  $a^+$  and  $a^-$  we define Courant numbers

$$c^+ = \Delta t a^+ / \Delta x, \quad c^- = \Delta t a^- / \Delta x. \quad (5.20)$$

Using the above notation the *first-order upwind scheme* can be expressed in general form as

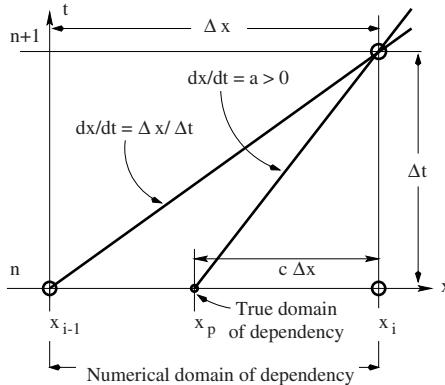
$$u_i^{n+1} = u_i^n - c^+ (u_i^n - u_{i-1}^n) - c^- (u_{i+1}^n - u_i^n). \quad (5.21)$$

For  $a \geq 0$  the second difference term vanishes leading to (5.16); if  $a \leq 0$  we obtain (5.18). The idea of *splitting the difference* into positive and negative components can be generalised to systems of conservation laws. Now the stability condition for the upwind scheme (5.21) is

$$0 \leq |c| \leq 1. \quad (5.22)$$

The *stencil* of the first-order upwind scheme for the case  $a > 0$  is shown in Fig. 5.3. It has a triangular shape, the three mesh points involved define a triangle. The base of the triangle defines the *numerical domain of dependence* of the scheme. This is generally different from the (true) domain of dependence of the PDE, see Sect. 2.3.5 of Chap. 2. One can relate the meaning of the scheme and its stability condition to the exact behaviour of the differential equation in (5.7).

Consider the characteristic of speed  $a$  through the mesh point  $(x_i, t^{n+1})$  in Fig. 5.3, at which a numerical solution  $u_i^{n+1}$  is sought. Since the exact solution



**Fig. 5.3.** Stencil of first-order upwind scheme (CIR) for positive speed  $a$  of wave propagation. Upwind direction lies on the left hand side.

of the PDE in (5.7) is constant along characteristics of speed  $dx/dt = a$ , see Sect. 2.2 of Chap. 2, the true solution at  $(x_i, t^{n+1})$  is

$$u(x_i, t^{n+1}) = u(x_p, t^n), \quad (5.23)$$

with  $x_p$  between  $x_{i-1}$  and  $x_i$ . Unfortunately, the only values at time level  $n$  available to us are those at the grid points and thus we could not set  $u_i^{n+1} = u(x_p, t^n)$ , unless  $c = 1$  of course.

One may however utilise the information at the mesh points  $x_{i-1}$  and  $x_i$  at the time level  $n$  to produce an estimate of the data at the point  $x_p$ . For instance we can construct a linear interpolation function  $\tilde{u}(x)$  based on the two points  $(x_{i-1}, u_{i-1}^n)$  and  $(x_i, u_i^n)$ . From Fig. 5.3, the distance  $d$  between  $x_p$  and  $x_i$  is

$$d = \Delta t a = \frac{\Delta t a}{\Delta x} \Delta x = c \Delta x,$$

and thus

$$x_p = (i - 1)\Delta x + (1 - c)\Delta x.$$

The linear interpolant is then

$$\tilde{u}(x) = u_{i-1}^n + \frac{(u_i^n - u_{i-1}^n)}{\Delta x} (x - x_{i-1}),$$

which if evaluated at  $x = x_p$  gives

$$\tilde{u}(x_p) = u_i^n - c(u_i^n - u_{i-1}^n).$$

This is precisely the first-order upwind scheme (5.16), which can then be seen as a *linear interpolation* scheme. Fig. 5.3 is also useful for interpreting the stability condition (5.17), which states

$$0 \leq \frac{a}{\Delta x / \Delta t} \leq 1.$$

Thus for stability, the grid speed  $\Delta x / \Delta t$  must be larger than the speed  $a$  in the PDE. In the context of Fig. 5.3 this means that the characteristics

$$\frac{dx}{dt} = \Delta x / \Delta t, \quad \frac{dx}{dt} = a,$$

must be as shown. That is *the numerical domain of dependence must contain the true domain of dependence of the PDE*, which is in fact the single point  $x_p$ .

A truncation error analysis reveals that the CIR scheme is first-order accurate in space and time. Moreover, the corresponding *modified equation* is

$$q_t + aq_x = \alpha_{\text{cir}} q_{xx}, \quad (5.24)$$

where  $\alpha_{\text{cir}}$  is the numerical viscosity coefficient of the CIR scheme and is given by

$$\alpha_{\text{cir}} = \frac{1}{2} \Delta x a (1 - |c|). \quad (5.25)$$

The advection-diffusion equation (5.24) is the actual equation solved by the numerical scheme, to second-order accuracy in fact; see LeVeque [308]. The viscous term  $\alpha_{\text{cir}} q_{xx}$  is responsible for the production of the *artificial or numerical viscosity* of the scheme. This vanishes when  $\Delta x = 0$ , which is impossible in practice. It also vanishes when  $|c| = 1$ , which is only of the academic importance, as for non-linear systems it is impossible to enforce this CFL number unity condition. The case  $a = 0$  is interesting, as it allows for the perfect resolution of *stationary discontinuities*. In general  $\alpha_{\text{cir}} > 0$  and one of the consequences is that discontinuities in the solution tend to be heavily *spread or smeared* and extreme values tend to be *clipped*. This is a disadvantage of the CIR scheme, which is in fact common to all first-order methods, with the exception of the Random Choice Method, see Chap. 7.

Scheme (5.21) is part of a wider class of methods that can be written as

$$\left. \begin{aligned} u_i^{n+1} &= H(u_{i-l_L}^n, \dots, u_{i+l_R}^n), \\ &= \sum_{k=-l_L}^{l_R} b_k u_{i+k}^n, \end{aligned} \right\} \quad (5.26)$$

with  $l_L$  and  $l_R$  two non-negative integers;  $b_k$ ,  $k = -l_L, \dots, l_R$  are the coefficients of the scheme and  $u_{i+k}^n$  are data values at time level  $n$ . For the CIR scheme (5.21) the coefficients are given by

$$b_{-1} = c^+, \quad b_0 = (1 - |c|), \quad b_1 = -c^-. \quad (5.27)$$

Under the CFL stability condition (5.22) we see that

$$b_k \geq 0, \quad \forall k, \quad (5.28)$$

i.e. all coefficients (5.27) are *positive or zero*.

**Definition 5.1 (Monotone Schemes).** A scheme of the form (5.26) is said to be monotone if all coefficients  $b_k$  are positive or zero. An alternative definition is given in terms of the function  $H$  in (5.26); this is required to satisfy

$$\frac{\partial H}{\partial u_k^n} \geq 0 , \quad \forall k . \quad (5.29)$$

The class of monotone schemes form the basis of modern schemes for conservation laws. Monotone schemes are, however, at most first-order accurate; high-order extensions are studied in Chaps. 13, 14 and 16.

### 5.2.2 Other Well-Known Schemes

Another first-order scheme is that of **Lax and Friedrichs**. The scheme is sometimes also called the Lax Method [299], or the scheme of Keller and Lax. This does not require the differencing to be performed according to upwind directions and can be seen as a way of stabilising the unstable scheme (5.12) obtained from forward in time and central in space approximations to the partial derivatives. The Lax–Friedrichs scheme results if  $u_i^n$  in the time derivative (5.9) is replaced by

$$\frac{1}{2} (u_{i-1}^n + u_{i+1}^n) ,$$

a mean value of the two neighbours at time level  $n$ . Then the modified scheme becomes

$$u_i^{n+1} = \frac{1}{2} (u_{i-1}^n + u_{i+1}^n) - \frac{1}{2} c (u_{i+1}^n - u_{i-1}^n)$$

or

$$u_i^{n+1} = \frac{1}{2} (1+c) u_{i-1}^n + \frac{1}{2} (1-c) u_{i+1}^n . \quad (5.30)$$

A von Neumann stability analysis reveals that scheme (5.30) is stable under the stability condition (5.22) and a truncation error analysis says that the scheme is first-order accurate. The modified equation is like (5.24) with numerical viscosity coefficient given by

$$\alpha_{lf} = \frac{\Delta x a}{2c} (1 - c^2) . \quad (5.31)$$

By comparing  $\alpha_{lf}$  with  $\alpha_{cir}$  we see that the Lax–Friedrichs scheme is considerably more diffusive than the CIR scheme; in fact for  $0 \leq c \leq 1$  we have

$$2 \leq \alpha_{lf}/\alpha_{cir} = \frac{1+c}{c} < \infty .$$

When written in the form (5.26) the coefficients of the Lax–Friedrichs scheme are found to be

$$b_{-1} = \frac{1}{2}(1+c), \quad b_0 = 0, \quad b_1 = \frac{1}{2}(1-c).$$

Under the stability condition (5.22) all coefficients  $b_k$  in the Lax–Friedrich scheme (5.30) are positive or zero. Therefore the scheme is *monotone*.

A scheme of historic as well as practical importance is that of **Lax and Wendroff** [302]. For a comprehensive treatment of the family of Lax–Wendroff schemes see Hirsch [252], Chap. 17. The basic Lax–Wendroff scheme is second–order accurate in both space and time. There are several ways of deriving the scheme for the model equation in (5.7). A rather unconventional derivation is this: for the time derivative  $u_t$  insist on the first–order forward approximation (5.9); for the space derivative  $u_x$  take an average of the upwind (stable if  $a > 0$ ) and downwind (unstable if  $a > 0$ ) approximations (5.14) and (5.15) respectively, that is

$$u_x = \beta_1 \frac{u_i^n - u_{i-1}^n}{\Delta x} + \beta_2 \frac{u_{i+1}^n - u_i^n}{\Delta x}. \quad (5.32)$$

If the coefficients  $\beta_1, \beta_2$  are chosen as

$$\beta_1 = \frac{1}{2}(1+c), \quad \beta_2 = \frac{1}{2}(1-c), \quad (5.33)$$

the resulting scheme is the Lax–Wendroff method

$$u_i^{n+1} = \frac{1}{2}c(1+c)u_{i-1}^n + (1-c^2)u_i^n - \frac{1}{2}c(1-c)u_{i+1}^n. \quad (5.34)$$

This scheme is second–order in space and time although all finite difference approximations used to generate it are first–order accurate. Moreover, one of the terms in the spatial derivative originates from an unconditionally unstable scheme and yet the Lax–Wendroff scheme is stable with stability condition (5.22). This scheme is a good example to show that the order of accuracy of the scheme cannot in general be inferred from the order of accuracy of the finite difference approximations to the partial derivatives involved.

When written in the form (5.26) the Lax–Wendroff scheme (5.34) has coefficients

$$b_{-1} = \frac{(1+c)c}{2}, \quad b_0 = 1 - c^2, \quad b_1 = -\frac{(1-c)c}{2}.$$

Therefore this scheme is *not monotone*. Not all coefficients in (5.34) are positive or zero. The fact that a scheme is not monotone is associated with the phenomenon of *spurious oscillations* in the numerical solution in the vicinity of sharp gradients, such as at discontinuities; see Chap. 13.

Another second–order accurate scheme for (5.7) is the upwind method of **Warming and Beam** [574]. For positive speed  $a$  it reads

$$u_i^{n+1} = \frac{1}{2}c(c-1)u_{i-2}^n + c(2-c)u_{i-1}^n + \frac{1}{2}(c-1)(c-2)u_i^n. \quad (5.35)$$

Note that the scheme is fully one-sided in the sense that all the mesh points involved, other than the centre of the stencil, are on the left hand side of the centre of the stencil. There is an equivalent scheme for negative speed  $a$ . Clearly the Warming–Beam scheme is not monotone. The stability restriction for this scheme is

$$0 \leq |c| \leq 2 . \quad (5.36)$$

The enlarged stability range means that one may advance in time with a larger time step  $\Delta t$ , which has a bearing on the efficiency of the scheme.

Yet another second order scheme is the **Fromm** scheme [195]. For the linear advection equation in (5.7), for  $a > 0$ , the scheme reads

$$\left. \begin{aligned} u_i^{n+1} = & -\frac{1}{4}(1-c)cu_{i-2}^n + \frac{1}{4}(5-c)cu_{i-1}^n \\ & + \frac{1}{4}(1-c)(4+c)u_i^n - \frac{1}{4}(1-c)cu_{i+1}^n , \end{aligned} \right\} \quad (5.37)$$

which has stability restriction (5.22). Also, it can be easily verified that the Fromm scheme is not monotone.

Second-order schemes such as the Lax–Wendroff, Warming–Beam and Fromm schemes have modified equation of the form

$$q_t + aq_x = \beta q_{xxx} , \quad (5.38)$$

which is a *dispersive* equation. See LeVeque [308] for details.

### 5.3 Conservative Methods

Computing solutions containing discontinuities, such as shock waves, poses stringent requirements on (i) the mathematical formulation of the governing equations and (ii) the numerical schemes to solve the equations. As seen in Chaps. 2 and 3 the formulation of the equations can be differential or integral. Also, there are various choices for the set of variables to be used. One obvious choice is the set of conserved variables. In Sect. 3.3 of Chap. 3, through an example, we highlighted the fact that formulations based on variables other than the conserved variables (non-conservative variables) fail at shock waves. They give the wrong jump conditions; consequently they give the wrong shock strength, the wrong shock speed and thus the wrong shock position. Recent work by Hou and Le Floch [259] has shown that non-conservative schemes do not converge to the correct solution if a shock wave is present in the solution. The classical result of Lax and Wendroff [302], on the other hand, says that conservative numerical methods, *if convergent*, do converge to the weak solution of the conservation law. Therefore, it appears as if there is no choice but to work with conservative methods if shock waves are part of the solution. There are alternative special procedures involving shock fitting [355], [353] and adaptive primitive-conservative schemes [508], [279]. Some primitive-variable schemes are presented in Sect. 14.6 of Chap. 14.

### 5.3.1 Basic Definitions

Here we shall study conservative shock capturing methods. Of the class of conservative methods we are particularly interested in *upwind methods*, but not exclusively. This section is designed to introduce some basic concepts in the simple setting of model problems. Consider a scalar conservation law written in differential form

$$u_t + f(u)_x = 0, \quad (5.39)$$

where  $f = f(u)$  is the flux function. The choice of flux  $f(u) = au$  reproduces the linear advection equation in (5.7). In order to include weak solutions of (5.39) we must use the integral form of the equations. Two possibilities are

$$\oint (u \, dx - f \, dt) = 0 \quad (5.40)$$

and

$$\int_{x_1}^{x_2} u(x, t_2) \, dx = \int_{x_1}^{x_2} u(x, t_1) \, dx + \int_{t_1}^{t_2} f(u(x_1, t)) \, dt - \int_{t_1}^{t_2} f(u(x_2, t)) \, dt \quad (5.41)$$

for any rectangular control volume  $[x_1, x_2] \times [t_1, t_2]$ . See Sect. 2.4 of Chap. 2.

**Definition 5.2 (Conservative Method).** A conservative scheme for the scalar conservation law (5.39) is a numerical method of the form

$$u_i^{n+1} = u_i^n + \frac{\Delta t}{\Delta x} \left[ f_{i-\frac{1}{2}} - f_{i+\frac{1}{2}} \right], \quad (5.42)$$

where

$$f_{i+\frac{1}{2}} = f_{i+\frac{1}{2}}(u_{i-l_L}^n, \dots, u_{i+l_R}^n), \quad (5.43)$$

with  $l_L, l_R$  two non-negative integers;  $f_{i+\frac{1}{2}}$  is called the numerical flux, an approximation to the physical flux  $f(u)$  in (5.39).

For any particular choice of numerical flux  $f_{i+\frac{1}{2}}$  a corresponding conservative scheme results. A fundamental requirement on the numerical flux is the *consistency condition*

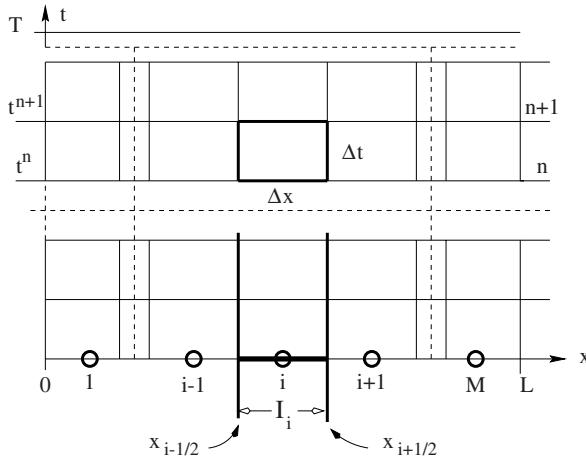
$$f_{i+\frac{1}{2}}(v, \dots, v) = f(v). \quad (5.44)$$

This means that if all arguments in (5.43) are equal to  $v$  then the numerical flux is identical to the physical flux at  $u = v$ . See LeVeque [308].

**Exercise 5.3.** Verify that the choice of numerical flux

$$f_{i+\frac{1}{2}} = f_{i+\frac{1}{2}}(u_i^n, u_{i+1}^n) = \frac{1}{2}(f_i^n + f_{i+1}^n),$$

with  $f_i^n = f(u_i^n)$ ,  $f_{i+1}^n = f(u_{i+1}^n)$ , when substituted into the conservative formula (5.42), reproduces the unconditionally unstable scheme (5.12), when applied to the linear advection equation, in which  $f(u) = au$ .



**Fig. 5.4.** Discretisation of domain  $[0, L]$  into  $M$  finite volumes  $I_i$  (computing cells)

**Solution 5.4.** (Left to the reader).

The conservative scheme (5.42) requires a redefinition of the discretisation of the domain. Now one is concerned with cell averages defined over *finite volumes*. A domain  $[0, L] \times [0, T]$  in the  $x$ - $t$  plane is discretised as shown in Fig. 5.4. The spatial domain of length  $L$  is subdivided into  $M$  *finite volumes*, called *computing cells* or simply *cells*, given as

$$x_{i-\frac{1}{2}} = (i - 1)\Delta x \leq x \leq i\Delta x = x_{i+\frac{1}{2}} . \quad (5.45)$$

The extreme values  $x_{i-\frac{1}{2}}$  and  $x_{i+\frac{1}{2}}$  of cell  $I_i$  define the position of the inter-cell boundaries at which the corresponding intercell numerical fluxes must be specified. The size of the cell is

$$\Delta x = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}} = \frac{L}{M} . \quad (5.46)$$

Obviously one may discretise the domain into cells of irregular size. For simplicity we assume regular meshes in this chapter. The average value of  $u(x, t)$  in cell  $i$ , the cell average, at a fixed time  $t = t^n = n\Delta t$  is defined as

$$u_i^n = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} u(x, t^n) dx . \quad (5.47)$$

Note here that although within cell  $i$  one may have spatial variations of  $u(x, t)$  at time  $t = t^n$ , the integral average value  $u_i^n$  given above is constant. We shall assign that constant value at the centre of the cell, which gives rise to *cell-centred conservative methods*. Computationally, we shall deal with approximations to the cell averages  $u_i^n$ , which for simplicity we shall still denote as  $u_i^n$ . The set of cell averages (5.47) defines a piece-wise constant distribution of the solution at time  $t^n$ ; see Fig. 5.5.

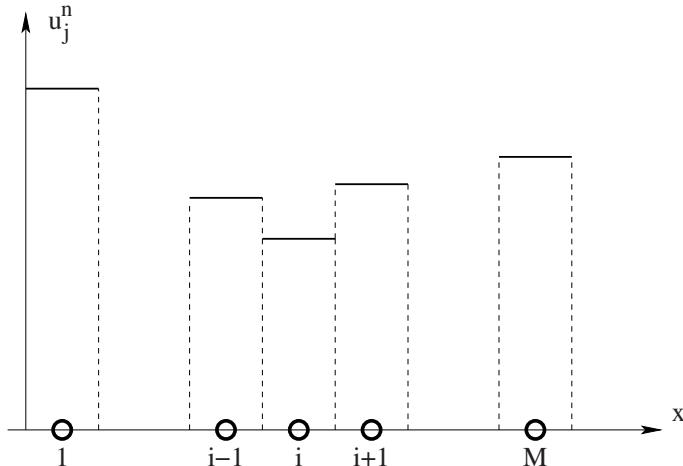


Fig. 5.5. Piece–wise constant distribution of data at time level  $n$

### 5.3.2 Godunov’s First–Order Upwind Method

Godunov [216] is credited with the first successful conservative extension of the CIR scheme (5.21) to non–linear systems of conservation laws. When applied to the scalar conservation law (5.39) with  $f(u) = au$ , Godunov’s scheme reduces to the CIR scheme, allowing for appropriate interpretation of the values  $\{u_i^n\}$ .

Godunov’s first–order upwind method is a conservative method of the form (5.42), where the intercell numerical fluxes  $f_{i+\frac{1}{2}}$  are computed by using solutions of local Riemann problems. A basic assumption of the method is that at a given time level  $n$  the data has a piece–wise constant distribution of the form (5.47), as depicted in Fig. 5.5. The data at time level  $n$  may be seen as pairs of constant states  $(u_i^n, u_{i+1}^n)$  separated by a discontinuity at the intercell boundary  $x_{i+\frac{1}{2}}$ . Then, locally, one can define a Riemann problem

$$\text{PDE : } u_t + f(u)_x = 0 .$$

$$\text{IC : } u(x, 0) = u_0(x) = \begin{cases} u_i^n & \text{if } x < 0 , \\ u_{i+1}^n & \text{if } x > 0 , \end{cases}$$

This *local* Riemann problem may be solved analytically, if desired. Thus, at a given time level  $n$ , at each intercell boundary  $x_{i+\frac{1}{2}}$  we have the *local* Riemann problem  $RP(u_i^n, u_{i+1}^n)$  with initial data  $(u_i^n, u_{i+1}^n)$ . What is then needed is a way of finding the solution of the *global* problem at a later time level  $n + 1$ .

#### First Version of Godunov’s Method

Godunov proposed the following scheme to update a cell value  $u_i^n$  to a new value  $u_i^{n+1}$ : solve the two Riemann problems  $RP(u_{i-1}^n, u_i^n)$  and  $RP(u_i^n, u_{i+1}^n)$

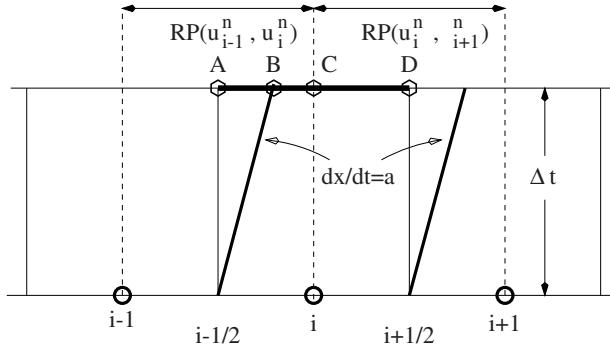
for the conservation law (5.39), take an integral average in cell  $i$  of the combined solutions of these two local problems and assign the value to  $u_i^{n+1}$ . Fig. 5.6 illustrates the situation for the special case  $f(u) = au$ ,  $a > 0$ . The exact solution of  $RP(u_{i-1}^n, u_i^n)$  for  $a > 0$ , see Sect. 2.2.2 of Chap. 2, is

$$u_{i-\frac{1}{2}}(x/t) = \begin{cases} u_{i-1}^n & \text{if } x/t < a, \\ u_i^n & \text{if } x/t > a, \end{cases} \quad (5.48)$$

where the local origin of the Riemann problem is  $(0, 0)$ . Likewise the solution  $u_{i+\frac{1}{2}}(x/t)$  of  $RP(u_i^n, u_{i+1}^n)$  is given by

$$u_{i+\frac{1}{2}}(x/t) = \begin{cases} u_i^n & \text{if } x/t < a, \\ u_{i+1}^n & \text{if } x/t > a. \end{cases} \quad (5.49)$$

The Godunov scheme defines the updated solution as



**Fig. 5.6.** Illustration of Godunov's method for the linear advection equation for positive speed  $a$ . Riemann problem solutions are averaged at time  $t = \Delta t$  inside cell  $I_i$

$$u_i^{n+1} = \frac{1}{\Delta x} \left[ \int_0^{\frac{1}{2}\Delta x} u_{i-\frac{1}{2}}(x/\Delta t) dx + \int_{-\frac{1}{2}\Delta x}^0 u_{i+\frac{1}{2}}(x/\Delta t) dx \right]. \quad (5.50)$$

This integral is evaluated at time  $\Delta t$  (local time) between the points  $A$  and  $D$  in Fig. 5.6. Note that we only use the right half of the solution  $u_{i-\frac{1}{2}}(x/t)$  and the left half of the solution  $u_{i+\frac{1}{2}}(x/t)$ . The reader should realise that *each solution has its own local frame of reference* with origin  $(0, 0)$  corresponding with the intercell boundaries at  $x_{i-\frac{1}{2}}$  and  $x_{i+\frac{1}{2}}$ . In order to evaluate the integral we impose a restriction on the size of the time step  $\Delta t$ . We set

$$c = \frac{a\Delta t}{\Delta x} \leq \frac{1}{2}. \quad (5.51)$$

The first term in (5.50) involves the intervals  $AB$  and  $BC$ ; these have lengths respectively given by

$$l_{AB} = c\Delta x, \quad l_{BC} = \left(\frac{1}{2} - c\right)\Delta x.$$

The integrand is give by (5.48) and (5.49); hence we have

$$\frac{1}{\Delta x} \int_0^{\frac{1}{2}\Delta x} u_{i-\frac{1}{2}}(x/\Delta t) dx = cu_{i-1}^n + \left(\frac{1}{2} - c\right)u_i^n.$$

The second term gives

$$\frac{1}{\Delta x} \int_{-\frac{1}{2}\Delta x}^0 u_{i+\frac{1}{2}}(x/\Delta t) dx = \frac{1}{2}u_i^n.$$

Hence (5.50) becomes

$$u_i^{n+1} = u_i^n + c(u_{i-1}^n - u_i^n), \quad (5.52)$$

which is identical to the CIR first order upwind method (5.16) for positive speed  $a$ . The conservative character of the Godunov method is self evident: the updated solution is obtained by integral averaging, a conservative process, of local exact solutions of the conservation law.

**Exercise 5.5.** Using geometric arguments, show that the Godunov approach described above also reproduces the CIR scheme for negative  $a$ .

**Solution 5.6.** (Left to the reader).

### Second Version of Godunov's Method

A second interpretation of the Godunov method leads directly to the conservative formula (5.42), which is easier to apply in practice and avoids the over-restrictive CFL-like condition (5.51). The integral average (5.50) of the solution of the Riemann problems  $RP(u_{i-1}^n, u_i^n)$  and  $RP(u_i^n, u_{i+1}^n)$  can also be written as

$$u_i^{n+1} = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \tilde{u}(x, \Delta t) dx, \quad (5.53)$$

where  $\tilde{u}(x, t)$  is understood as the combined solution of  $RP(u_{i-1}^n, u_i^n)$  and  $RP(u_i^n, u_{i+1}^n)$ . Since  $\tilde{u}(x, t)$  is an exact solution to the original conservation law (5.39) we can make use of it in its integral form (5.41), say, in the control volume  $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [0, \Delta t]$  to write

$$\left. \begin{aligned} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \tilde{u}(x, \Delta t) dx &= \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \tilde{u}(x, 0) dx + \int_0^{\Delta t} f(\tilde{u}(x_{i-\frac{1}{2}}, t)) dt \\ &\quad - \int_0^{\Delta t} f(\tilde{u}(x_{i+\frac{1}{2}}, t)) dt. \end{aligned} \right\} \quad (5.54)$$

Using the definition of cell averages (5.47) into equation (5.54) followed by division through by  $\Delta x$  we reproduce the conservative formula

$$u_i^{n+1} = u_i^n + \frac{\Delta t}{\Delta x} \left[ f_{i-\frac{1}{2}} - f_{i+\frac{1}{2}} \right], \quad (5.55)$$

with the intercell fluxes defined as time integral averages, namely

$$f_{i-\frac{1}{2}} = \frac{1}{\Delta t} \int_0^{\Delta t} f(\tilde{u}(x_{i-\frac{1}{2}}, t)) dt, \quad f_{i+\frac{1}{2}} = \frac{1}{\Delta t} \int_0^{\Delta t} f(\tilde{u}(x_{i+\frac{1}{2}}, t)) dt. \quad (5.56)$$

Thus, by invoking the integral form of the conservation laws on a control or finite volume  $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [0, \Delta t]$  in  $x-t$  space we have arrived at the conservative formula (5.55) with intercell fluxes (5.56); these are time integral averages of the physical flux  $f(u)$  evaluated at the intercell boundaries. The integrand  $f(\tilde{u}(x, t))$  at each cell interface depends on the exact solution  $\tilde{u}(x, t)$  of the Riemann problem along the  $t$ -axis (local coordinates); this is given by

$$\tilde{u}(x_{i-\frac{1}{2}}, t) = u_{i-\frac{1}{2}}(0), \quad \tilde{u}(x_{i+\frac{1}{2}}, t) = u_{i+\frac{1}{2}}(0), \quad (5.57)$$

and the intercell fluxes  $f_{i-\frac{1}{2}}$  and  $f_{i+\frac{1}{2}}$  become

$$f_{i-\frac{1}{2}} = f(u_{i-\frac{1}{2}}(0)), \quad f_{i+\frac{1}{2}} = f(u_{i+\frac{1}{2}}(0)). \quad (5.58)$$

In general, one expresses the Godunov intercell numerical flux as

$$f_{i+\frac{1}{2}}^{\text{god}} = f(u_{i+\frac{1}{2}}(0)), \quad (5.59)$$

where  $u_{i+\frac{1}{2}}(0)$  denotes the exact solution  $u_{i+\frac{1}{2}}(x/t)$  of the Riemann problem  $RP(u_i^n, u_{i+1}^n)$  evaluated at  $x/t = 0$ , i.e. the solution is evaluated along the intercell boundary, which coincides with the  $t$ -axis in the local frame of the Riemann problem solution. We have thus defined the second version of Godunov's method for a general scalar conservation law (5.39), as the conservative formula (5.55) together with the intercell numerical flux (5.59). For the special conservation law with flux  $f(u) = au$ ,  $a > 0$ , we have

$$f_{i-\frac{1}{2}} = au_{i-1}^n, \quad f_{i+\frac{1}{2}} = au_i^n, \quad (5.60)$$

which if substituted in the conservative formula (5.55) reproduce the CIR scheme. The second version (5.55), (5.59) of Godunov's method is the one that is mostly used in practice.

**Exercise 5.7.** Verify that the second version of the Godunov method based on the conservative formula (5.55) and the Godunov intercell flux (5.59) also reproduces the CIR scheme when applied to (5.39) with flux  $f(u) = au$  and  $a < 0$ .

**Solution 5.8.** (Left to the reader).

### 5.3.3 Godunov's Method for Burgers's Equation

As a way of illustrating Godunov's method in the context of *non-linear* PDEs we apply the scheme to the inviscid Burgers equation

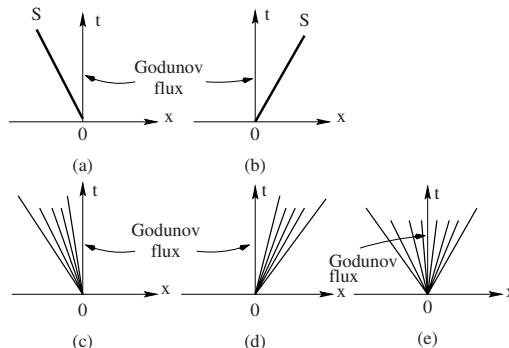
$$u_t + f(u)_x = 0, \quad f(u) = \frac{1}{2}u^2. \quad (5.61)$$

We adopt the second version (5.55) with numerical flux given by (5.59). We need the solution  $u_{i+\frac{1}{2}}(x/t)$  of the Riemann problem  $RP(u_i^n, u_{i+1}^n)$ . As seen in Sect. 2.4.2 of Chap. 2, the solution is a shock wave, when  $u_i^n > u_{i+1}^n$ , and a rarefaction wave when  $u_i^n \leq u_{i+1}^n$ . The complete solution is

$$\left. \begin{aligned} u_{i+\frac{1}{2}}(x/t) &= \begin{cases} u_i^n & \text{if } S \geq x/t \\ u_{i+1}^n & \text{if } S \leq x/t \end{cases} & \text{if } u_i^n > u_{i+1}^n, \\ S &= \frac{1}{2}(u_i^n + u_{i+1}^n) \end{aligned} \right\} \quad (5.62)$$

$$\left. \begin{aligned} u_{i+\frac{1}{2}}(x/t) &= \begin{cases} u_i^n & \text{if } x/t \leq u_i^n \\ x/t & \text{if } u_i^n < x/t < u_{i+1}^n \\ u_{i+1}^n & \text{if } x/t \geq u_{i+1}^n \end{cases} & \text{if } u_i^n \leq u_{i+1}^n. \end{aligned} \right\}$$

The Godunov's flux requires  $u_{i+\frac{1}{2}}(0)$ ; this is the solution  $u_{i+\frac{1}{2}}(x/t)$  evaluated along the intercell boundary  $x_{i+\frac{1}{2}}$ , that corresponds to  $x/t = 0$  in the frame of the local Riemann problem. The second stage is *to identify all possible wave patterns in the solution* of the Riemann problem. For Burgers's equation there are five possibilities. These are illustrated in Fig. 5.7. If the solution is a shock



**Fig. 5.7.** Five possible wave patterns in the solution of the Riemann problem for the inviscid Burgers equations, when evaluating the Godunov flux

wave then cases (a) and (b) can occur. The sought value  $u_{i+\frac{1}{2}}(0)$ , on the  $t$ -axis, depends on the sign of the shock speed  $S$ . If the solution is a rarefaction wave then the three possible cases are illustrated in Figs. 5.7c, 5.7d and 5.7e. Applying terminology from Gas Dynamics to the rarefaction cases, Fig. 5.7c

is called *supersonic to the left* and that of Fig. 5.7d *supersonic to the right*. The case of Fig. 5.7e is that of a *transonic rarefaction* or *sonic rarefaction*; as the wave is crossed, there is a sign change in the characteristic speed  $u$  and thus there is one point at which  $u = 0$ , a *sonic point*. The complete sought solution is summarised as follows

$$\left. \begin{aligned} u_{i+\frac{1}{2}}(0) &= \begin{cases} u_i^n & \text{if } S \geq 0 \\ u_{i+1}^n & \text{if } S \leq 0 \end{cases} & \text{if } u_i^n > u_{i+1}^n, \\ S &= \frac{1}{2}(u_i^n + u_{i+1}^n) \end{aligned} \right\} \quad (5.63)$$

$$u_{i+\frac{1}{2}}(0) = \begin{cases} u_i^n & \text{if } 0 \leq u_i^n \\ 0 & \text{if } u_i^n < 0 < u_{i+1}^n \\ u_{i+1}^n & \text{if } 0 \geq u_{i+1}^n \end{cases} \quad \text{if } u_i^n \leq u_{i+1}^n.$$

Naturally, Godunov's method can also be implemented using approximate solutions to the Riemann problem. For a review on the Godunov scheme in the context of two well-known approximations to the Riemann problem solution for Burgers's equation, the reader is referred to the paper by van Leer [562].

In applying Godunov's scheme to solve Burgers's equation there are two more issues to consider. One concerns the application of boundary conditions and the other is to do with the choice of the time step  $\Delta t$ .

## Boundary Conditions

The conservative formula (5.55) can be applied directly to all cells  $i$ , for  $i = 2, \dots, M-1$ . The two required intercell fluxes at  $x_{i-\frac{1}{2}}$  and  $x_{i+\frac{1}{2}}$  are defined in terms of the corresponding Riemann problems. For each of the cells 1 and  $M$ , which are adjacent to the left and right boundaries respectively, we only have one intercell flux. Some special procedure needs to be implemented. Let us consider the left boundary  $x = 0$ . One possibility is to assume a boundary function  $u_l(t)$  prescribed there. Then we could define an intercell flux at the boundary by setting  $f_{\frac{1}{2}} = f(u_l(t))$ .

A more attractive alternative is to specify a *fictitious cell 0* to the left of the boundary  $x = 0$  together with a cell average  $u_0^n$ , at each time level  $n$ , so that a Riemann problem  $RP(u_0^n, u_1^n)$  can be posed and solved to find the *missing* intercell flux  $f_{\frac{1}{2}}$ . For the right boundary we prescribe a fictitious cell  $M+1$  and a cell average  $u_{M+1}^n$  to find the intercell flux  $f_{M+\frac{1}{2}}$ . The prescription of the fictitious states depends entirely on the physics of the particular problem at hand. Provisionally, for the inviscid Burgers's equation we apply the boundary conditions

$$u_0^n = u_1^n, \quad u_{M+1}^n = u_M^n. \quad (5.64)$$

Note that the fictitious states here are given in terms of the data at the states within the computational domain adjacent to the boundaries. This particular type of boundary conditions will cause no disturbance at the boundaries; waves will hopefully go through the boundaries as if the boundaries were not

there. Often one speaks of *transparent* or *transmissive* boundary conditions, in this case.

## Choosing the Time Step

As seen for the linear case, the choice of the size of the time step  $\Delta t$  in the conservative formula is related to the stability condition of the particular scheme. For Godunov's method (the CIR scheme in this case) the choice of  $\Delta t$  depends on the restriction on the Courant number  $c$ . For non-linear problems, at each time level, there are multiple wave speeds and thus multiple associated Courant numbers. In deriving the second version of Godunov's method we made the implicit assumption that the value of the local Riemann problem solution along the intercell boundary is constant. This means that the fastest wave at a given time travels for at most one cell length  $\Delta x$  in the sought time step  $\Delta t$ . Denoting by  $S_{\max}^n$  the maximum wave speed throughout the domain at time level  $n$  we define the *maximum Courant number*  $C_{\text{cfl}}$

$$C_{\text{cfl}} = \Delta t S_{\max}^n / \Delta x , \quad (5.65)$$

where  $\Delta t$  is such that

$$0 < C_{\text{cfl}} \leq 1 .$$

We shall often call  $C_{\text{cfl}}$  the *CFL coefficient* or the *Courant number coefficient*. The time step  $\Delta t$  follows as

$$\Delta t = C_{\text{cfl}} \Delta x / S_{\max}^n . \quad (5.66)$$

A matter of crucial importance is the estimation of the maximum speed  $S_{\max}^n$ . In realistic applications this can be difficult and frustrating, as inappropriate choices can cause the scheme to *crash*, no matter how sophisticated this is. For Burgers's equation one can identify wave speeds, such as those emerging from solutions of Riemann problems at the intercell boundaries, and characteristic speeds  $u$ . At any time level  $n$ , there are  $M + 2$  characteristic speeds  $u_i^n$  (including the fictitious cells) and hence one possibility is to take  $S_{\max}^n$  as the maximum of these, in absolute value. Another possibility is to select, at each time level, a speed  $S_{i+\frac{1}{2}}^n$  from the solution Riemann problem at each cell interface; this information is available as part of the flux computation process. For the case of a shock, one obviously takes the shock speed. For the case of a rarefaction there are two characteristic speeds of significance, namely those of the head and tail of the expansion. Thus we define an intercell speed

$$S_{i+\frac{1}{2}}^n = \begin{cases} |\frac{1}{2}(u_i^n + u_{i+1}^n)| & (\text{shock}), \\ \max(|u_i^n|, |u_{i+1}^n|) & (\text{rarefaction}). \end{cases} \quad (5.67)$$

Finally, we define a maximum wave speed at time level  $n$  as follows

$$S_{\max}^n = \max_i \left\{ S_{i+\frac{1}{2}}^n \right\} \text{ for } i = 0, \dots, M+1 . \quad (5.68)$$

Having chosen  $C_{\text{cfl}}$ , the time step  $\Delta t$  to march the solution to the next time level is given by (5.66). For a scheme with linearised stability condition  $|c| \leq 1$ , one usually takes the empirical value  $C_{\text{cfl}} = 0.9$ .

In Sect. 5.6 we give a listing of a FORTRAN program for Godunov's method in conjunction with the exact Riemann solver, to solve numerically the inviscid Burgers equation.

### 5.3.4 Conservative Form of Difference Schemes

Conventional finite difference schemes can often be expressed in conservation form (5.42). It is a question of finding an intercell flux (5.43).

#### The Lax–Friedrichs Scheme

Recall that the Lax–Friedrichs scheme as applied to the linear advection equation is

$$u_i^{n+1} = \frac{(1+c)}{2} u_{i-1}^n + \frac{(1-c)}{2} u_{i+1}^n . \quad (5.69)$$

It is easy to verify that the conservative formula (5.42) together with the intercell flux

$$f_{i+\frac{1}{2}} = \frac{(1+c)}{2c} f(u_i^n) + \frac{(c-1)}{2c} f(u_{i+1}^n) \quad (5.70)$$

reproduces the Lax–Friedrichs scheme. Therefore, at least for the linear advection equation, one can recast the Lax–Friedrichs scheme in conservative form.

An interesting way of viewing the Lax–Friedrichs scheme (5.69) is as an integral average *within cell  $i$* , namely

$$u_i^{n+1} = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \tilde{u}(x, \frac{1}{2} \Delta t) dx , \quad (5.71)$$

in which  $\tilde{u}(x, t)$  is the solution of the Riemann problem  $RP(u_{i-1}^n, u_{i+1}^n)$  (note subscripts), that is

$$\tilde{u}(x/t) = \begin{cases} u_{i-1}^n & \text{if } x/t < a , \\ u_{i+1}^n & \text{if } x/t > a . \end{cases} \quad (5.72)$$

**Exercise 5.9.** Verify that exact evaluation of the integral (5.71) reproduces the Lax–Friedrichs scheme (5.69).

**Solution 5.10.** (Left to the reader).

*Remark 5.11.* We note that the Lax–Friedrichs solution at cell  $i$  is a weighted average of the solution of the Riemann problem with the *left and right neighbour states* as data, at time  $t = \frac{1}{2}\Delta t$ . One could state that the Lax–Friedrichs scheme is *upwind biased*, as the upwind term has always the larger weight.

Let us now attempt to generalise interpretation (5.71) of the Lax–Friedrichs scheme to non–linear systems of conservation laws

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0}. \quad (5.73)$$

Now (5.71) reads

$$\mathbf{U}_i^{n+1} = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \tilde{\mathbf{U}}(x, \frac{1}{2}\Delta t) dx, \quad (5.74)$$

where  $\tilde{\mathbf{U}}(x, t)$  is the solution of the Riemann problem  $RP(\mathbf{U}_{i-1}^n, \mathbf{U}_{i+1}^n)$ . Given an exact Riemann solver, e.g. see Chap. 4, one could then implement this Riemann–problem based version of the Lax–Friedrichs scheme. The author has implemented this version of the scheme for non–linear systems. The numerical results obtained are indistinguishable from those obtained from the conventional form of the scheme. Version (5.74) offers no obvious advantages over the conventional form; in fact it is more expensive and complex and hence is of no practical use. A stochastic evaluation of the integral leads to a random choice type method; see Sect. 7.3 of Chap. 7.

If the space integral (5.74) at time  $\frac{1}{2}\Delta t$  is replaced by invoking the integral form of conservation law (Chap. 2, equation (2.67)) in the control volume  $[-\frac{1}{2}\Delta x, \frac{1}{2}\Delta x] \times [0, \frac{1}{2}\Delta t]$  we obtain

$$\left. \begin{aligned} \int_{-\frac{1}{2}\Delta x}^{\frac{1}{2}\Delta x} \tilde{\mathbf{U}}(x, \frac{1}{2}\Delta t) dx &= \int_{-\frac{1}{2}\Delta x}^{\frac{1}{2}\Delta x} \tilde{\mathbf{U}}(x, 0) dx + \int_0^{\frac{1}{2}\Delta t} \mathbf{F}(\tilde{\mathbf{U}}(-\frac{1}{2}\Delta x, t)) dt \\ &\quad - \int_0^{\frac{1}{2}\Delta t} \mathbf{F}(\tilde{\mathbf{U}}(\frac{1}{2}\Delta x, t)) dt. \end{aligned} \right\} \quad (5.75)$$

Direct evaluation of the integrals and use of the definition of integral averages (5.74) in cell  $i$ , as applied to systems, yield

$$\mathbf{U}_i^{n+1} = \frac{1}{2}(\mathbf{U}_{i-1}^n + \mathbf{U}_{i+1}^n) + \frac{1}{2} \frac{\Delta t}{\Delta x} (\mathbf{F}_{i-1}^n - \mathbf{F}_{i+1}^n).$$

Simple algebraic manipulations of this expression lead to a conservative version of the Lax–Friedrichs scheme for systems

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n + \frac{\Delta t}{\Delta x} \left[ \mathbf{F}_{i-\frac{1}{2}} - \mathbf{F}_{i+\frac{1}{2}} \right], \quad (5.76)$$

with the Lax–Friedrichs intercell flux given by

$$\mathbf{F}_{i+\frac{1}{2}}^{\text{LF}} = \frac{1}{2}(\mathbf{F}_i^n + \mathbf{F}_{i+1}^n) + \frac{1}{2} \frac{\Delta x}{\Delta t} (\mathbf{U}_i^n - \mathbf{U}_{i+1}^n). \quad (5.77)$$

This is the conventional numerical flux for the Lax–Friedrichs scheme when applied to systems of conservation laws (5.73). No mention of the Riemann problem is needed in this formulation. Compare the conservative formula (5.76) for systems with the conservative formula (5.55) for scalar conservation laws.

### The Lax–Wendroff Scheme

The Lax–Wendroff scheme (5.34) as applied to the linear advection equation may also be written in conservation form (5.42). The intercell numerical flux is

$$f_{i+\frac{1}{2}} = \frac{(1+c)}{2}(au_i^n) + \frac{(1-c)}{2}(au_{i+1}^n),$$

which is a weighted average of fluxes on the left and right of the interface. For the linear advection equation, it is easy to check that this can also be obtained from

$$f_{i+\frac{1}{2}} = f\left(u_{i+\frac{1}{2}}^{n+\frac{1}{2}}\right), \quad u_{i+\frac{1}{2}}^{n+\frac{1}{2}} = \frac{(1+c)}{2}u_i^n + \frac{(1-c)}{2}u_{i+1}^n.$$

For non-linear scalar conservation laws, such as Burgers's equations (5.61), this generalises to

$$f_{i+\frac{1}{2}} = f\left(u_{i+\frac{1}{2}}^{n+\frac{1}{2}}\right); \quad u_{i+\frac{1}{2}}^{n+\frac{1}{2}} = \frac{1}{\Delta x} \int_{-\frac{1}{2}\Delta x}^{\frac{1}{2}\Delta x} u_{i+\frac{1}{2}}(x, \frac{1}{2}\Delta t) dx, \quad (5.78)$$

where  $u_{i+\frac{1}{2}}(x, t)$  is the solution of the Riemann problem  $RP(u_i^n, u_{i+1}^n)$ . A straightforward Riemann–problem based generalisation of the Lax–Wendroff scheme to non–linear hyperbolic systems (5.73) ([499], [506]) reads

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}^{n+\frac{1}{2}}); \quad \mathbf{U}_{i+\frac{1}{2}}^{n+\frac{1}{2}} = \frac{1}{\Delta x} \int_{-\frac{1}{2}\Delta x}^{\frac{1}{2}\Delta x} \mathbf{U}_{i+\frac{1}{2}}(x, \frac{1}{2}\Delta t) dx,$$

where  $\mathbf{U}_{i+\frac{1}{2}}(x, t)$  is the solution of the Riemann problem  $RP(\mathbf{U}_i^n, \mathbf{U}_{i+1}^n)$ . This scheme is called the Weighted Average Flux (WAF) method and is studied in Chaps. 13, 14 and 16.

As done for the Lax–Friedrichs scheme one may replace the integral involving the solution of the Riemann problem by invoking the integral form of the conservation laws, see Sect. 2.4.1 of Chap. 2, in the control volume  $[-\frac{1}{2}\Delta x, \frac{1}{2}\Delta x] \times [0, \frac{1}{2}\Delta t]$  to obtain

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}\left(\mathbf{U}_{i+\frac{1}{2}}^{n+\frac{1}{2}}\right); \quad \mathbf{U}_{i+\frac{1}{2}}^{n+\frac{1}{2}} = \frac{1}{2}(\mathbf{U}_i^n + \mathbf{U}_{i+1}^n) + \frac{1}{2} \frac{\Delta t}{\Delta x} (\mathbf{F}_i^n - \mathbf{F}_{i+1}^n). \quad (5.79)$$

This scheme is known as the two–step Richtmyer version of the Lax–Wendroff method, as applied to non–linear systems of conservation laws (5.73). No mention of the Riemann problem solution is necessary here.

*Remark 5.12.* Note the similarities between the reinterpretations and generalisations of the Lax–Friedrichs and Lax–Wendroff schemes. For both schemes one ends up with two formulations. For the Lax–Friedrichs scheme the weighted–average character leads to an integral formulation involving the solution of the Riemann problem. The second version of the scheme eliminates the role of the Riemann problem by utilising the integral form of the conservation law and leads to the conventional form of the Lax–Friedrichs scheme for non–linear systems. For the Lax–Wendroff method the procedure is entirely analogous. An integral average interpretation leads to a Riemann–problem based extension to non–linear systems [499]. Utilisation of the integral form of the conservation law eliminates the role of the Riemann problem and leads to the conventional Richtmyer version of the scheme. Both versions of the Lax–Wendroff method have actually been applied in practice.

**Exercise 5.13.** Verify that the Fromm scheme (5.37) as applied to the linear advection equation in (5.7), for positive  $a$ , can be written in conservation form (5.42) with numerical flux

$$f_{i+\frac{1}{2}}^{\text{FR}} = -\frac{1}{4}(1-c)f_{i-1} + f_i + \frac{1}{4}(1-c)f_{i+1},$$

where  $c$  is the Courant number.

**Solution 5.14.** (Left to the reader).

**Exercise 5.15.** Verify that the numerical flux of the scheme of Warming and Beam (5.35), as applied to (5.7) with  $a > 0$ , is

$$f_{i+\frac{1}{2}}^{\text{WB}} = \frac{1}{2}(c-1)f_{i-1} + \frac{1}{2}(3-c)f_i.$$

**Solution 5.16.** (Left to the reader).

*Remark 5.17.* The Warming–Beam numerical flux can be derived from (5.78) in terms of integral averages of solutions of Riemann problems under the assumption  $1 \leq c \leq 2$ . An extension of this interpretation to non–linear systems was proposed by Toro and Billett [527].

**Exercise 5.18.** Apply (5.78) to the linear advection equation and derive the Warming–Beam flux for negative speed  $a$ . Assume  $-2 \leq c \leq -1$ .

**Solution 5.19.** (Left to the reader).

## 5.4 Upwind Schemes for Linear Systems

Here we apply the first–order upwind scheme to hyperbolic systems with constant coefficients

$$\mathbf{U}_t + \mathbf{A}\mathbf{U}_x = \mathbf{0}. \quad (5.80)$$

For background on mathematical properties of these PDEs see Sect. 2.3 of Chap. 2. We denote the real eigenvalues of the  $m \times m$  constant coefficient matrix  $\mathbf{A}$  by  $\lambda_j$  with  $j = 1, \dots, m$  and assume they are ordered as

$$\lambda_1 < \lambda_2 < \lambda_3 \dots < \lambda_m.$$

The corresponding right eigenvectors are denoted by  $\mathbf{K}^{(1)}, \mathbf{K}^{(2)}, \dots, \mathbf{K}^{(m)}$ . Note here that in general the eigenvalues  $\lambda_j$  can be of any sign and thus a one-sided differencing scheme applied to (5.80) directly will only work if all the eigenvalues are of the same sign. In the general case with eigenvalues of mixed sign the particular chosen side for the differencing will be *upwind* for only some of the eigenvalues and *downwind* for the rest. The difficulty can be resolved by splitting the matrix into two matrices, one of them having positive or zero eigenvalues and the other having negative or zero eigenvalues. From the assumption of hyperbolicity,  $\mathbf{A}$  may be diagonalised as

$$\mathbf{A} = \mathbf{K}\Lambda\mathbf{K}^{-1}, \quad (5.81)$$

where  $\mathbf{K}$  is the matrix whose columns are the right eigenvectors  $\mathbf{K}^{(j)}$ ,  $\mathbf{K}^{-1}$  is the inverse of  $\mathbf{K}$  and  $\Lambda$  is the diagonal matrix formed by the eigenvalues of  $\mathbf{A}$ , namely

$$\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix}. \quad (5.82)$$

See Sect. 2.3.1 of Chap. 2. In terms of the *characteristic variables*

$$\mathbf{V} = \mathbf{K}^{-1}\mathbf{U}, \quad (5.83)$$

system (5.80) becomes the *decoupled* system

$$\mathbf{V}_t + \Lambda\mathbf{V}_x = \mathbf{0}, \quad (5.84)$$

where the  $j$ -th equation

$$\frac{\partial}{\partial t}v_j + \lambda_j \frac{\partial}{\partial x}v_j = 0, \text{ for } j = 1, \dots, m, \quad (5.85)$$

involves only the variable  $v_j(x, t)$ .

### 5.4.1 The CIR Scheme

From a numerical point of view the decomposition of (5.80) into the decoupled set (5.84) with component equations (5.85) is very convenient. Each component equation (5.85) is a linear advection equation with characteristic

speed  $\lambda_j = \text{constant}$ , just as the PDE in (5.7), in which the speed is  $a$ , a constant. The CIR first-order upwind scheme (5.21) can directly be applied to each component equation (5.85). As for the scalar case we introduce the following definitions

$$\left. \begin{aligned} \lambda_j^+ &\equiv \max(\lambda_j, 0) = \frac{1}{2}(\lambda_j + |\lambda_j|), \\ \lambda_j^- &\equiv \min(\lambda_j, 0) = \frac{1}{2}(\lambda_j - |\lambda_j|), \end{aligned} \right\} \quad (5.86)$$

where  $|\lambda_j|$  is the absolute value of  $\lambda_j$ . The following relations can be easily verified

$$\lambda_j = \lambda_j^+ + \lambda_j^-, \quad |\lambda_j| = \lambda_j^+ - \lambda_j^-. \quad (5.87)$$

Then, the CIR scheme (5.21) applied to each PDE in (5.85) for the characteristic variables reads

$$\begin{aligned} (v_j)_i^{n+1} &= (v_j)_i^n - \frac{\Delta t}{\Delta x} \lambda_j^+ [(v_j)_i^n - (v_j)_{i-1}^n] \\ &\quad - \frac{\Delta t}{\Delta x} \lambda_j^- [(v_j)_{i+1}^n - (v_j)_i^n]. \end{aligned} \quad (5.88)$$

This is a straight generalisation of the first-order upwind scheme (5.21) to the decoupled linear hyperbolic system (5.84).

Based on definitions (5.86) we can form the positive  $\Lambda^+$  and negative  $\Lambda^-$  components of the diagonal matrix  $\Lambda$ , namely

$$\Lambda^\pm \equiv \begin{pmatrix} \lambda_1^\pm & & 0 \\ & \ddots & \\ 0 & & \lambda_m^\pm \end{pmatrix}. \quad (5.89)$$

Property (5.81) allows us to introduce the positive and negative components of the coefficient matrix  $\mathbf{A}$  as

$$\mathbf{A}^- = \mathbf{K} \Lambda^- \mathbf{K}^{-1}, \quad \mathbf{A}^+ = \mathbf{K} \Lambda^+ \mathbf{K}^{-1}. \quad (5.90)$$

Based on properties (5.87), the matrices (5.89), (5.90) can be shown to satisfy the following

$$\left. \begin{aligned} \Lambda &= \Lambda^+ + \Lambda^-, \quad |\Lambda| = \Lambda^+ - \Lambda^-, \\ \mathbf{A} &= \mathbf{A}^+ + \mathbf{A}^-, \quad |\mathbf{A}| = \mathbf{A}^+ - \mathbf{A}^- \end{aligned} \right\} \quad (5.91)$$

The CIR scheme (5.88) can be written as

$$\left. \begin{aligned} \mathbf{V}_i^{n+1} &= \mathbf{V}_i^n - \frac{\Delta t}{\Delta x} \Lambda^+ (\mathbf{V}_i^n - \mathbf{V}_{i-1}^n) \\ &\quad - \frac{\Delta t}{\Delta x} \Lambda^- (\mathbf{V}_{i+1}^n - \mathbf{V}_i^n) \end{aligned} \right\} \quad (5.92)$$

In terms of the original variables  $\mathbf{U} = \mathbf{K} \mathbf{V}$  this scheme may be expressed as

$$\left. \begin{aligned} \mathbf{U}_i^{n+1} &= \mathbf{U}_i^n - \frac{\Delta t}{\Delta x} \mathbf{A}^+ [\mathbf{U}_i^n - \mathbf{U}_{i-1}^n] \\ &\quad - \frac{\Delta t}{\Delta x} \mathbf{A}^- [\mathbf{U}_{i+1}^n - \mathbf{U}_i^n] . \end{aligned} \right\} \quad (5.93)$$

This is the generalisation of scheme (5.21) to linear hyperbolic systems (5.80) with constant coefficients. It is left to the reader to verify that the above result can be obtained by multiplying (5.92) from the left by the matrix  $\mathbf{K}$  and using (5.90). Note that by splitting the coefficient matrix  $\mathbf{A}$  into a positive and a negative component we have been able to retain the basic principle of performing the one-sided spatial differencing according to the sign of the characteristic speeds. The differencing  $\mathbf{U}_i^n - \mathbf{U}_{i-1}^n$  is upwind for the coefficient matrix  $\mathbf{A}^+$  and  $\mathbf{U}_{i+1}^n - \mathbf{U}_i^n$  is upwind for the coefficient matrix  $\mathbf{A}^-$ .

### 5.4.2 Godunov's Method

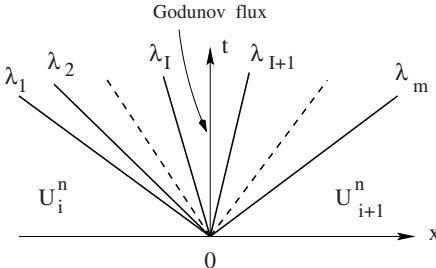
Consider the constant coefficient, linear hyperbolic system (5.80) written in conservation-law form

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0}, \quad \mathbf{F}(\mathbf{U}) \equiv \mathbf{A}\mathbf{U}. \quad (5.94)$$

The Godunov first-order upwind method utilises the conservative formula (5.76) and requires the solution  $\mathbf{U}_{i+\frac{1}{2}}(x/t)$  of the local Riemann problem  $RP(\mathbf{U}_i^n, \mathbf{U}_{i+1}^n)$  for (5.94) to compute the intercell numerical flux

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}(0)). \quad (5.95)$$

See Sect. 5.3.2. Here  $\mathbf{U}_{i+\frac{1}{2}}(0)$  is the value of the solution  $\mathbf{U}_{i+\frac{1}{2}}(x/t)$  at  $x/t = 0$  along the intercell boundary. As seen in Sect. 2.3.3 of Chap. 2, the solution



**Fig. 5.8.** Evaluation of the Godunov intercell flux for linear hyperbolic systems with constant coefficients

$\mathbf{U}_{i+\frac{1}{2}}(x/t)$  can be easily found by first expanding the initial data  $\mathbf{U}_i^n, \mathbf{U}_{i+1}^n$  in terms of the right eigenvectors as

$$\mathbf{U}_i^n = \sum_{j=1}^m \alpha_j \mathbf{K}^{(j)}, \quad \mathbf{U}_{i+1}^n = \sum_{j=1}^m \beta_j \mathbf{K}^{(j)}. \quad (5.96)$$

The general solution at any point  $(x, t)$  is given by

$$\mathbf{U}_{i+\frac{1}{2}}(x/t) = \sum_{j=1}^I \beta_j \mathbf{K}^{(j)} + \sum_{j=I+1}^m \alpha_j \mathbf{K}^{(j)}, \quad (5.97)$$

where  $I$  is the largest integer with  $1 \leq I \leq m$  such that  $x/t \geq \lambda_I$ . The Godunov flux (5.95) requires the solution at  $x/t = 0$  in (5.97). See Fig. 5.8. For  $x/t = 0$   $I$  is such that  $\lambda_I \leq 0$  and  $\lambda_{I+1} \geq 0$ , then  $\mathbf{U}_{i+\frac{1}{2}}(0)$  is obtained by manipulating (5.97), namely

$$\mathbf{U}_{i+\frac{1}{2}}(0) = \mathbf{U}_i^n + \sum_{j=1}^I (\beta_j - \alpha_j) \mathbf{K}^{(j)} \quad (5.98)$$

or

$$\mathbf{U}_{i+\frac{1}{2}}(0) = \mathbf{U}_{i+1}^n - \sum_{j=I+1}^m (\beta_j - \alpha_j) \mathbf{K}^{(j)}. \quad (5.99)$$

Recall that the jump across wave  $j$  with eigenvalue  $\lambda_j$  and eigenvector  $\mathbf{K}^{(j)}$  is given by  $(\beta_j - \alpha_j)\mathbf{K}^{(j)}$ . Note that the solution of the Riemann problem, at  $x/t = 0$ , as given by (5.98), can be interpreted as being the left data state  $\mathbf{U}_i^n$  plus all wave jumps across waves of negative or zero speed. Similarly, the form (5.99) gives the solution as the right data state  $\mathbf{U}_{i+1}^n$  minus the wave jumps across all waves of positive or zero speeds. By combining (5.98) and (5.99) we obtain

$$\mathbf{U}_{i+\frac{1}{2}}(0) = \frac{1}{2}(\mathbf{U}_i^n + \mathbf{U}_{i+1}^n) - \frac{1}{2} \sum_{j=1}^m \text{sign}(\lambda_j)(\beta_j - \alpha_j) \mathbf{K}^{(j)}. \quad (5.100)$$

The Godunov intercell numerical flux (5.95) can now be obtained by evaluating  $\mathbf{F}(\mathbf{U})$  at any of the expressions (5.98)–(5.100) for the solution of the Riemann problem. Use of (5.98) gives

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}_i^n + \sum_{j=1}^I \mathbf{A}(\beta_j - \alpha_j) \mathbf{K}^{(j)}, \quad (5.101)$$

and since  $\mathbf{A}\mathbf{K}^{(j)} = \lambda_j \mathbf{K}^{(j)}$ ,

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}_i^n + \sum_{j=1}^I (\beta_j - \alpha_j) \lambda_j \mathbf{K}^{(j)}. \quad (5.102)$$

Similarly, (5.99) gives

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}_{i+1}^n - \sum_{j=I+1}^m (\beta_j - \alpha_j) \lambda_j \mathbf{K}^{(j)}, \quad (5.103)$$

or combining (5.102) and (5.103) we obtain

$$\mathbf{F}_{i+\frac{1}{2}} = \frac{1}{2} (\mathbf{F}_i^n + \mathbf{F}_{i+1}^n) - \frac{1}{2} \sum_{j=1}^m (\beta_j - \alpha_j) |\lambda_j| \mathbf{K}^{(j)}. \quad (5.104)$$

Next we show that the Godunov flux can also be expressed in two more alternative forms.

**Proposition 5.20.** *The Godunov flux (5.95) to solve (5.94) via (5.76) can be written as*

$$\mathbf{F}_{i+\frac{1}{2}} = \frac{1}{2} (\mathbf{F}_i^n + \mathbf{F}_{i+1}^n) - \frac{1}{2} |\mathbf{A}| (\mathbf{U}_{i+1}^n - \mathbf{U}_i^n). \quad (5.105)$$

*Proof.* Starting from (5.104) and using the properties (5.87) and (5.91) one writes

$$\begin{aligned} \mathbf{F}_{i+\frac{1}{2}} &= \frac{1}{2} (\mathbf{F}_i^n + \mathbf{F}_{i+1}^n) - \frac{1}{2} \sum_{j=1}^m (\beta_j - \alpha_j) (\lambda_j^+ - \lambda_j^-) \mathbf{K}^{(j)} \\ &= \frac{1}{2} (\mathbf{F}_i^n + \mathbf{F}_{i+1}^n) - \frac{1}{2} \sum_{j=1}^m (\beta_j - \alpha_j) [\lambda_j^+ \mathbf{K}^{(j)} - \lambda_j^- \mathbf{K}^{(j)}] \\ &= \frac{1}{2} (\mathbf{F}_i^n + \mathbf{F}_{i+1}^n) - \frac{1}{2} \sum_{j=1}^m (\beta_j - \alpha_j) [\mathbf{A}^+ \mathbf{K}^{(j)} - \mathbf{A}^- \mathbf{K}^{(j)}] \\ &= \frac{1}{2} (\mathbf{F}_i^n + \mathbf{F}_{i+1}^n) - \frac{1}{2} \sum_{j=1}^m (\beta_j - \alpha_j) [\mathbf{A}^+ - \mathbf{A}^-] \mathbf{K}^{(j)} \\ &= \frac{1}{2} (\mathbf{F}_i^n + \mathbf{F}_{i+1}^n) - \frac{1}{2} |\mathbf{A}| \sum_{j=1}^m (\beta_j - \alpha_j) \mathbf{K}^{(j)}. \end{aligned}$$

Hence

$$\mathbf{F}_{i+\frac{1}{2}} = \frac{1}{2} (\mathbf{F}_i^n + \mathbf{F}_{i+1}^n) - \frac{1}{2} |\mathbf{A}| (\mathbf{U}_{i+1}^n - \mathbf{U}_i^n)$$

and the proposition is proved.

**Proposition 5.21.** *The Godunov flux (5.95) for (5.94) can be written in flux-split form as*

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{A}^+ \mathbf{U}_i^n + \mathbf{A}^- \mathbf{U}_{i+1}^n. \quad (5.106)$$

*Proof.* The result follows directly from manipulating (5.105) and using appropriate definitions. Alternatively we have

$$\begin{aligned} \mathbf{F}_{i+\frac{1}{2}} &= \mathbf{A} \mathbf{U}_{i+\frac{1}{2}}(0) \\ &= \sum_{j=1}^I \beta_j \mathbf{A} \mathbf{K}^{(j)} + \sum_{j=I+1}^m \alpha_j \mathbf{A} \mathbf{K}^{(j)} \end{aligned}$$

$$\begin{aligned}
&= \sum_{j=1}^I \beta_j \lambda_j \mathbf{K}^{(j)} + \sum_{j=I+1}^m \alpha_j \lambda_j \mathbf{K}^{(j)} \\
&= \sum_{j=1}^m \beta_j \lambda_j^- \mathbf{K}^{(j)} + \sum_{j=1}^m \alpha_j \lambda_j^+ \mathbf{K}^{(j)} \\
&= \sum_{j=1}^m \beta_j \mathbf{A}^- \mathbf{K}^{(j)} + \sum_{j=1}^m \alpha_j \mathbf{A}^+ \mathbf{K}^{(j)} \\
&= \mathbf{A}^+ \sum_{j=1}^m \alpha_j \mathbf{K}^{(j)} + \mathbf{A}^- \sum_{j=1}^m \beta_j \mathbf{K}^{(j)} \\
&= \mathbf{A}^+ \mathbf{U}_i^n + \mathbf{A}^- \mathbf{U}_{i+1}^n ,
\end{aligned}$$

and the proposition is proved.

*Remark 5.22.* The intercell flux has been *split* as

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}_{i+\frac{1}{2}}^+ + \mathbf{F}_{i+\frac{1}{2}}^- , \quad (5.107)$$

where the positive  $\mathbf{F}_{i+\frac{1}{2}}^+$  and negative  $\mathbf{F}_{i+\frac{1}{2}}^-$  flux components are

$$\mathbf{F}_{i+\frac{1}{2}}^+ = \mathbf{A}^+ \mathbf{U}_i^n , \quad \mathbf{F}_{i+\frac{1}{2}}^- = \mathbf{A}^- \mathbf{U}_i^n . \quad (5.108)$$

Note that, trivially, the respective Jacobian matrices have eigenvalues that are all positive (or zero) and all negative (or zero).

**Exercise 5.23.** Consider the linearised equations of Gas Dynamics

$$\mathbf{U}_t + \mathbf{A} \mathbf{U}_x = \mathbf{0} ,$$

with

$$\mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \equiv \begin{bmatrix} \rho \\ u \end{bmatrix} , \quad \mathbf{A} = \begin{bmatrix} 0 & \rho_0 \\ a^2/\rho_0 & 0 \end{bmatrix} .$$

Using the results of Sect. 2.3.4 of Chap. 2

- Find the matrices  $\Lambda^-$ ,  $\Lambda^+$ ,  $\mathbf{A}^-$ ,  $\mathbf{A}^+$ .
- Write the scheme (5.93) in full, that is for the two components of the vector of unknowns.
- Compute the Godunov intercell flux directly by using the explicit solution of the Riemann problem in the *Star Region*

$$\left. \begin{aligned} \rho_* &= \frac{1}{2}(\rho_L + \rho_R) - \frac{1}{2}(u_R - u_L)\rho_0/a , \\ u_* &= \frac{1}{2}(u_L + u_R) - \frac{1}{2}(\rho_R - \rho_L)a/\rho_0 . \end{aligned} \right\}$$

How many possible wave patterns do you need to consider here?

- Write a computer program to solve the linearised equations of Gas Dynamics using the method of Godunov.

**Solution 5.24.** (Left to the reader).

## 5.5 Sample Numerical Results

To complete this chapter, we present some numerical results obtained by some of the most well known schemes as applied to two model PDEs.

### 5.5.1 Linear Advection

We apply four schemes to solve

$$u_t + f(u)_x = 0, \quad f(u) = au, \quad a = \text{constant} \quad (5.109)$$

with two types of initial conditions.

#### Test 1 for linear advection (smooth data)

Here the initial condition is the smooth profile

$$u(x, 0) = \alpha e^{-\beta x^2}. \quad (5.110)$$

In the computations we take  $a = 1.0$ ,  $\alpha = 1.0$ ,  $\beta = 8.0$  and a CFL coefficient  $C_{\text{cfl}} = 0.8$ ; the initial profile  $u(x, 0)$  is evaluated in the interval  $-1 \leq x \leq 1$ . Computed results are shown in Figs. 5.9 to 5.11; these correspond respectively to the output times  $t = 1.0$  unit (125 time steps),  $t = 10.0$  units (1250 time steps),  $t = 100.0$  units (12499 time steps). In each figure we compare the exact solution (shown by full lines) with the numerical solution (symbols) for the Godunov method, the Lax–Friedrichs method, the Lax–Wendroff method and the Warming–Beam method.

The results of Fig. 5.9 are in many ways representative of the quality of each scheme. Collectively these results are also representative of most of the current successes and limitations of numerical methods for PDEs governing wave propagation. The first-order method of Godunov (CIR scheme) has modified equation of the form (5.24), where  $\alpha_{\text{cir}}$  is a numerical viscosity coefficient. This is responsible for the *clipping* of the peak values. As seen earlier  $\alpha_{\text{cir}} < \alpha_{\text{lf}}$ , which explains the fact that the Lax–Friedrichs scheme gives even more diffused results. For the computational parameters used  $\alpha_{\text{cir}} = 0.1\Delta x$  and  $\alpha_{\text{lf}} = 0.225\Delta x$ .

The results from the Lax–Wendroff method and the Warming–Beam method, both second-order accurate, are much more accurate than those of the first-order schemes. There are however, slight signs of error in the position of the wave. For the Lax–Wendroff scheme the computed wave *lags behind* the true wave (lagging phase error), while for the Warming–Beam method the computed wave *is ahead* of the true wave (leading phase error). The phase errors of second-order accurate schemes are explained by the dispersive term of the modified equation (5.38).

The limitations of the schemes are more clearly exposed if the solution is evolved for longer times. Fig. 5.10 shows results at the output time  $t =$

10.0 units (1250 time steps). Compare with Fig. 5.9. The numerical diffusion inherent in first-order methods has ruined the solution of the Godunov and Lax–Friedrichs schemes. Computed peak values are only of the order of 30 to 40% of the true peak values. The second-order methods are still giving more satisfactory results than their first-order counterparts, but now the numerical dispersion errors are clearly visible. Numerical diffusion is beginning to show its effects too.

Fig. 5.11 shows results at the output time  $t = 100.0$  units (12499 time steps). Compare with Figs. 5.9. and 5.10. These results are truly disappointing and clearly expose the limitations of numerical methods for computing solutions to problems involving long time evolution of wave phenomena. In acoustics one may require the computation of (i) very weak signals (ii) over long distances. The combination of these two requirements rules out automatically a wide range of otherwise acceptable numerical methods for PDEs. See Tam and Webb [480]. The numerical diffusion of the first-order schemes has virtually flattened the wave, while the numerical dispersion of the second-order methods has resulted in unacceptable position errors, in addition to clipping by numerical diffusion.

## Test 2 for linear advection (discontinuous data)

Now the initial data for (5.109) consists of a *square wave*, namely

$$u(x, 0) = \begin{cases} 0 & \text{if } x \leq 0.3, \\ 1 & \text{if } 0.3 \leq x \leq 0.7, \\ 0 & \text{if } x \geq 0.7. \end{cases} \quad (5.111)$$

The computed results for the three output times are shown in Figs. 5.12 to 5.14. As for Test 1 the effects of numerical diffusion in the first-order methods and the effects of dispersion in the second-order methods lead to visible errors in the numerical solution (symbols), as compared with the exact solution (full line). First-order methods *smear* discontinuities over many computing cells; as expected this error is more apparent in the Lax–Friedrichs scheme. Note also the *pairing* of neighbouring values in the Lax–Friedrichs scheme. Second-order methods reduce the smearing of discontinuities, but at the cost of *overshoots* and *undershoots* in the vicinity of the discontinuities. These *spurious oscillations* are highly undesirable features of second and higher-order methods. We shall return to this theme in Chaps. 13 and 14, where improved methods for dealing with discontinuities will be presented.

Fig. 5.13 shows results for Test 2 at time  $t = 10.0$  units (1250 time steps). The errors observed in Fig. 5.12 are now exaggerated. Fig. 5.14 shows results at time  $t = 100.0$  units (12499 time steps). Once again first-order methods have *lost* the solution while second-order methods exhibit unacceptable position errors, in addition to spurious oscillations produced near discontinuities.

### 5.5.2 The Inviscid Burgers Equation

Our Test 3 consists of the inviscid Burgers equation

$$u_t + f(u)_x = 0 , \quad f(u) = \frac{1}{2}u^2 \quad (5.112)$$

in the domain  $[0, \frac{3}{2}]$  with initial conditions

$$u(x, 0) = \begin{cases} -\frac{1}{2} & \text{if } x \leq \frac{1}{2}, \\ 1 & \text{if } \frac{1}{2} \leq x \leq 1, \\ 0 & \text{if } x \geq 1. \end{cases} \quad (5.113)$$

We solve this problem numerically on a domain of length  $L = 1.5$  discretised by  $M = 75$  equally spaced cells of width  $\Delta x = 0.02$ ; the CFL coefficient used is 0.8. Fig. 5.15 shows computed results (symbol) along with the exact (line) solution, for the Godunov and Lax–Friedrichs schemes at time  $t = 0.5$  units (32 time steps). Two new features are now present in solving non–linear PDEs. First the discontinuity on the right is a shock wave. This satisfies the entropy condition, see Sect. 2.4.2 of Chap. 2, and characteristics on either side of the discontinuity converge into the discontinuity. This compression mechanism helps the more accurate resolution of shock waves. Compare with Fig. 5.12. The Godunov method resolves the shock much more sharply (3 cells) than the Lax–Friedrichs scheme (10 cells). The second new feature to note in this non–linear example is the *entropy glitch* at  $x = \frac{1}{2}$ . This corresponds to a sonic point, see Sect. 2.4.2 of Chap. 2. The entropy glitch affects the Godunov method and not the Lax–Friedrichs method. A question of crucial importance is the construction of entropy satisfying schemes [369].

More advanced concepts on numerical methods are presented in Chap. 13 for scalar problems. Chaps. 14, 15 and 16 deal with numerical methods for non–linear systems.

## 5.6 FORTRAN Program for Godunov’s Method

A listing of a FORTRAN program to compute the numerical solution to the inviscid Burgers equation is included.

```
*-----*
*-----*
*-----*
C      Godunov's method for the inviscid Burgers's      *
C              equation                                     *
*-----*
C      Name of program: HL-B1GOD                         *
*-----*
C      Purpose: to solve the inviscid Burgers equation   *
C                  using the Godunov first order upwind    *
C                  scheme in conjunction with the exact    *
C                  Riemann solver                           *
*-----*
C      Input  file: b1god.ini                            *
C      Output file: numer.out (numerical)               *
*-----*
C      Programer: E. F. Toro                            *
*-----*
C      Last revision: February 7th 1999                 *
*-----*
C      Theory is found in Section 5.3.3, Chapter 5 of     *
C      Reference 1.                                      *
*-----*
C      1. Toro, E. F., "Riemann Solvers and Numerical      *
C                  Methods for Fluid Dynamics"             *
C                  Springer-Verlag,                      *
C                  Second Edition, 1999                     *
*-----*
C      This program is part of                          *
*-----*
C      NUMERICA                                         *
C      A Library of Source Codes for Teaching,          *
C      Research and Applications,                      *
C      by E. F. Toro                                    *
C      Published by NUMERITEK LTD,                    *
C      Website: www.numeritek.com                   *
*-----*
*-----*
*-----*
C      Driver program
*-----*
IMPLICIT NONE
*-----*
C      Declaration of variables:
*-----*
```

```
INTEGER ITEST, CELLS, N, NFREQ, NTMAXI
*
REAL      CFLCOE, DOMLEN, DT, TIME, TIMEOUT, TIMETO
*
COMMON /DATAIN/ CFLCOE, DOMLEN, ITEST, CELLS,
&           NFREQ, NTMAXI, TIMEOUT
COMMON /DELTAT/ DT
*
DATA TIMETO /1.0E-07/
*
C Parameters of problem are read in from
C file "b1god.ini"
*
CALL READER
*
C Initial conditions are set up
*
CALL INITIA(DOMLEN, ITEST, CELLS)
*
WRITE(6,*),'-----',
WRITE(6,*),'    Time step N          TIME          TIMEOUT',
WRITE(6,*),'-----',
*
C Time marching procedure
*
TIME = 0.0
*
DO 10 N = 1, NTMAXI
*
C Boundary conditions are set
*
CALL BCONDI(CELLS)
*
C Courant-Friedrichs-Lowy (CFL) condition imposed
*
CALL CFLCON(CFLCOE, CELLS, TIME, TIMEOUT)
*
TIME = TIME + DT
*
C Intercell numerical fluxes are computed
*
CALL FLUXES(CELLS)
*
C Solution is updated according to
C conservative formula
```

```
*          CALL UPDATE(CELLS)
*
*          IF(MOD(N,NFREQ).EQ.0)WRITE(6,20)N, TIME
*
C          Check output time
*
*          IF(ABS(TIME - TIMEOUT).LE.TIMETO)THEN
*
C              Solution is written to "numer.out" at
C              specified time TIMEOUT
*
*              CALL OUTPUT(CELLS)
*
*              WRITE(6,*),-----
*              WRITE(6,*),    Number of time steps = ',N
*
*              STOP
        ENDIF
*
10     CONTINUE
*
20     FORMAT(I12,6X, F12.7)
*
      END
*
*-----*
*
*          SUBROUTINE READER
*
C          Purpose: to read initial parameters of the problem
*
          IMPLICIT NONE
*
C          Declaration of variables
*
          INTEGER   ITEST, CELLS, NFREQ, NTMAXI
*
          REAL      CFLCOE, DOMLEN, TIMEOUT
*
          COMMON /DATAIN/ CFLCOE, DOMLEN, ITEST, CELLS, NFREQ,
&                           NTMAXI, TIMEOUT
*
          OPEN(UNIT = 1,FILE = 'b1god.ini',STATUS = 'UNKNOWN')
*
```

```

READ(1,*)CFLCOE ! Courant number coefficient
READ(1,*)DOMLEN ! Domain length
READ(1,*)ITEST ! Test problem
READ(1,*)CELLS ! Number of cells in domain
READ(1,*)NFREQ ! Output frequency to screen
READ(1,*)NTMAXI ! Maximum number of time steps
READ(1,*)TIMEOUT ! Output time
*
CLOSE(1)
*
WRITE(6,*)'-----',
WRITE(6,*)'Data read in is echoed to screen',
WRITE(6,*)'-----',
WRITE(6,*)'CFLCOE = ',CFLCOE
WRITE(6,*)'DOMLEN = ',DOMLEN
WRITE(6,*)'ITEST = ',ITEST
WRITE(6,*)'CELLS = ',CELLS
WRITE(6,*)'NFREQ = ',NFREQ
WRITE(6,*)'NTMAXI = ',NTMAXI
WRITE(6,*)'TIMEOUT = ',TIMEOUT
WRITE(6,*)'-----',
*
RETURN
END
*
*-----*
*
SUBROUTINE INITIA(DOMLEN, ITEST, CELLS)
*
C Purpose: to set initial conditions for solution U
C           and initialise other variables. There are
C           two choices of initial conditions,
C           determined by ITEST
*
C Local variables:
*
C   Name          Description
*   ===          =====
*
C   DX            Spatial mesh size
C   I              Variable in do loop
C   ITEST         Defines test problem
C   FLUX          Array for intercell fluxes
C   U              Array for numerical solution
C   XPOS          Position along x-axis

```

```

C      XRIGHT        Left diaphragm
C      XMIDL         Middle diaphragm
C      XRIGHT        Right diaphragm
*
IMPLICIT NONE
*
C      Declaration of variables
*
INTEGER I, ITEST, CELLS, IDIM
*
REAL      DOMLEN, DX, FLUX, U, XLEFT, XPOS, XMIDL,
&          XRIGHT
*
PARAMETER (IDIM = 1000)
*
DIMENSION FLUX(0:IDIM + 1), U(0:IDIM + 1)
*
COMMON /DELTAX/ DX
COMMON /FLUXFS/ FLUX
COMMON /SOLUTI/ U
*
C      Calculate mesh size DX
*
DX = DOMLEN/REAL(CELLS)
*
C      Initialise arrays
*
DO 10 I = 0, IDIM + 1
    FLUX(I) = 0.0
    U(I)    = 0.0
10  CONTINUE
*
IF(ITEST.EQ.1)THEN
*
C      Test 1: smooth profile
*
XPOS     = -1.0
*
DO 20 I = 1, CELLS
    XPOS = XPOS + 2.0/REAL(CELLS)
    U(I) = EXP(-8.0*XPOS*XPOS)
20  CONTINUE
*
ELSE
*
```

```

C      Test 2: square waves
*
      XLEFT   = 0.1*DOMLEN
      XMIDL  = 0.5*DOMLEN
      XRIGHT = 0.9*DOMLEN
*
      DO 30 I = 1, CELLS
*
      XPOS = (REAL(I)-1.0)*DX
*
      IF(XPOS.LT.XLEFT)THEN
          U(I) = -1.0
      ENDIF
*
      IF(XPOS.GE.XLEFT.AND.XPOS.LE.XMIDL)THEN
          U(I) = 1.0
      ENDIF
*
      IF(XPOS.GT.XMIDL.AND.XPOS.LE.XRIGHT)THEN
          U(I) = 0.0
      ENDIF
*
      IF(XPOS.GT.XRIGHT)THEN
          U(I) = -1.0
      ENDIF
*
      30      CONTINUE
*
      ENDIF
*
      RETURN
END
*
*-----*
*
      SUBROUTINE BCOND1(CELLS)
*
C      Purpose: to apply boundary conditions
*
      IMPLICIT NONE
*
C      Declaration of variables
*
      INTEGER CELLS, IDIM
*
```

```
      REAL      U
*
*      PARAMETER (IDIM = 1000)
*
*      DIMENSION U(0:IDIM + 1)
*
*      COMMON /SOLUTI/ U
*
C      Left boundary, periodic boundary condition
*
*      U(0) = U(CELLS)
*
C      Right boundary, periodic boundary condition
*
*      U(CELLS + 1) = U(1)
*
*      RETURN
*      END
*
*-----*
*
*      SUBROUTINE CFLCON(CFLCOE, CELLS, TIME, TIMEOUT)
*
C      Purpose: to apply the CFL condition to compute a
C              stable time step DT
*
*      IMPLICIT NONE
*
C      Declaration of variables
*
*      INTEGER I, CELLS, IDIM
*
*      REAL      CFLCOE, DT, DX, SMAX, TIME, TIMEOUT, U
*
*      PARAMETER (IDIM = 1000)
*
*      DIMENSION U(0:IDIM + 1)
*
*      COMMON /SOLUTI/ U
*      COMMON /DELTAT/ DT
*      COMMON /DELTAX/ DX
*
*      SMAX = -1.0E+06
*
C      Find maximum characteristic speed
```

```

*
      DO 10 I = 0, CELLS + 1
          IF(ABS(U(I)).GT.SMAX)SMAX = ABS(U(I))
10    CONTINUE
*
      DT = CFLCOE*DX/SMAX
*
C      Check size of DT to avoid exceeding output time
*
      IF((TIME + DT).GT.TIMEOUT)THEN
*
C          Recompute DT
*
          DT = TIMEOUT - TIME
ENDIF
*
      RETURN
END
*
*-----*
*
      SUBROUTINE UPDATE(CELLS)
*
C      Purpose: to update the solution to a new time level
C              using the explicit conservative formula
*
      IMPLICIT NONE
*
C      Declaration of variables
*
      INTEGER I, CELLS, IDIM
*
      REAL      DT, DX, DTODX, FLUX, U
*
      PARAMETER (IDIM = 1000)
*
      DIMENSION U(0:IDIM + 1), FLUX(0:IDIM + 1)
*
      COMMON /DELTAT/ DT
      COMMON /DELTAX/ DX
      COMMON /FLUXFS/ FLUX
      COMMON /SOLUTI/ U
*
      DTODX = DT/DX
*
```

```
DO 10 I = 1, CELLS
      U(I) = U(I) + DTODX*(FLUX(I-1) - FLUX(I))
10   CONTINUE
*
      RETURN
      END
*
*-----*
*
SUBROUTINE OUTPUT(CELLS)
*
C Purpose: to output the solution at a specified time
C           TIMEOUT
*
IMPLICIT NONE
*
C Declaration of variables
*
INTEGER I, CELLS, IDIM
*
REAL      DX, U, XPOS
*
PARAMETER (IDIM = 1000)
*
DIMENSION U(0:IDIM + 1)
*
COMMON /DELTAX/ DX
COMMON /SOLUTI/ U
*
OPEN(UNIT = 1,FILE = 'numer.out',STATUS = 'UNKNOWN')
*
DO 10 I = 1, CELLS
      XPOS = REAL(I)*DX
      WRITE(1,20)XPOS, U(I)
10   CONTINUE
*
      CLOSE(1)
*
20   FORMAT(2(4X, F10.5))
*
      RETURN
      END
*
*-----*
```

```

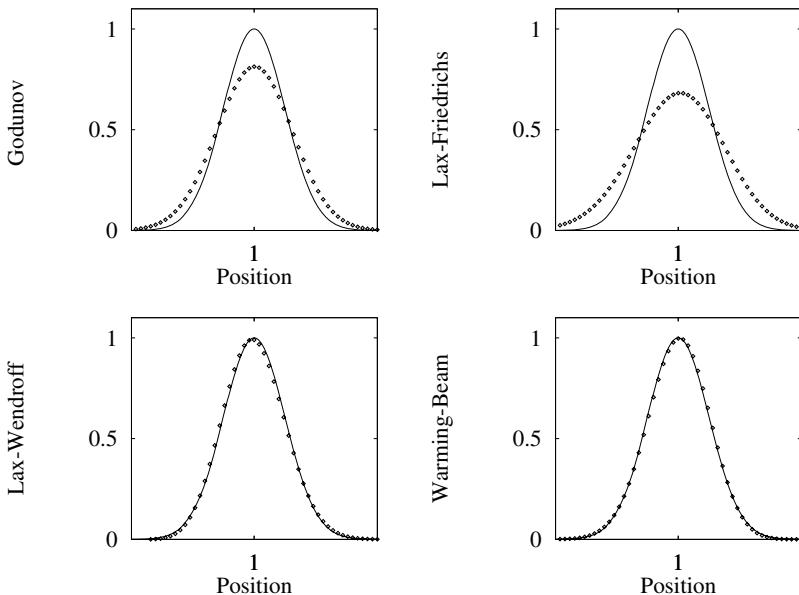
SUBROUTINE FLUXES(CELLS)
*
C Purpose: to compute intercell fluxes according to
C           the Godunov first-order upwind method,
C           in conjunction with the exact Riemann
C           solver
*
IMPLICIT NONE
*
C Declaration of variables
*
INTEGER I, CELLS, IDIM
*
REAL     FLUX, U, UL, UR, USTAR
*
PARAMETER (IDIM = 1000)
*
DIMENSION FLUX(0:IDIM + 1), U(0:IDIM + 1)
*
COMMON /FLUXFS/ FLUX
COMMON /SOLUTI/ U
*
C Compute intercell flux FLUX(I), I = 0, CELLS
C Solution of Riemann problem RP(I, I+1) is stored
C in FLUX(I)
*
DO 10 I = 0, CELLS
*
C Define states UL (Left) and UR (Right) for local
C Riemann problem RP(UL, UR)
*
UL = U(I)
UR = U(I+1)
*
C Solve the Riemann problem RP(UL, UR) exactly
*
CALL RIEMANN(UL, UR, USTAR)
*
C Compute Godunov intercell flux
*
FLUX(I) = 0.5*USTAR*USTAR
*
10    CONTINUE
*
RETURN

```

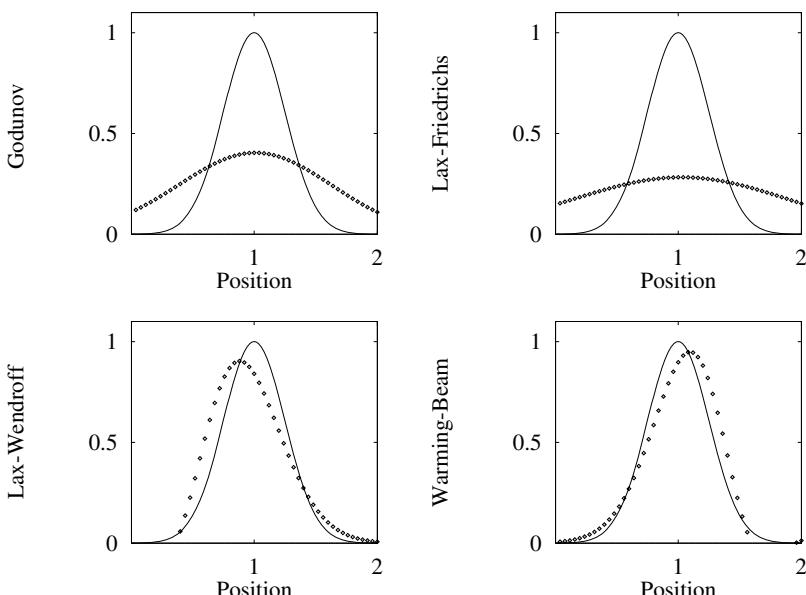
```

      END
*
*-----*
*
      SUBROUTINE RIEMANN(UL, UR, USTAR)
*
C      Purpose: to solve the Riemann problem for the inviscid
C              Burgers equation exactly.
*
C      Local variables:
*
C      Name      Description
*      ===       =====
*
C      UL        Left data state
C      UR        Right data state
C      S         Shock speed
C      USTAR     Sampled state
*
      IMPLICIT NONE
*
      REAL    S, UL, UR, USTAR
*
      IF(UL.GT.UR)THEN
*
C          Solution is a shock wave
C          Compute shock speed S
*
          S = 0.5*(UL + UR)
*
C          Sample the state along the t-axis
*
          IF(S.GE.0.0)THEN
              USTAR = UL
          ELSE
              USTAR = UR
          ENDIF
*
          ELSE
*
C          Solution is a rarefaction wave.
C          There are 3 cases:
*
          IF(UL.GE.0.0)THEN
*
```

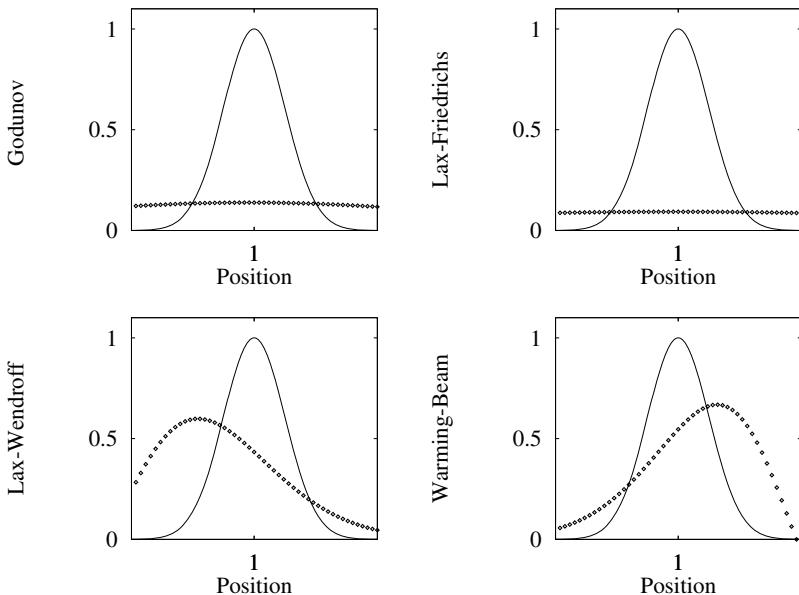
```
C          Right supersonic rarefaction
*
USTAR = UL
ENDIF
*
IF(UR.LE.0.0)THEN
*
C          Left supersonic rarefaction
*
USTAR = UR
ENDIF
*
IF(UL.LE.0.0.AND.UR.GE.0.0)THEN
*
C          Transonic rarefaction
*
USTAR = 0.0
ENDIF
*
ENDIF
*
RETURN
END
*
*-----*
```



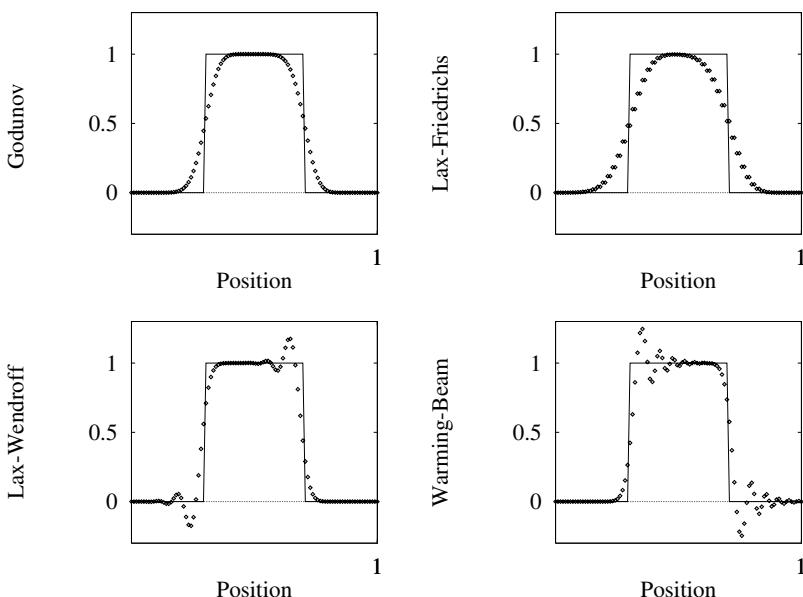
**Fig. 5.9.** Test 1: Comparison of numerical results for four numerical schemes (symbols) with the exact solution (line) at output time of 1 unit (125 time steps)



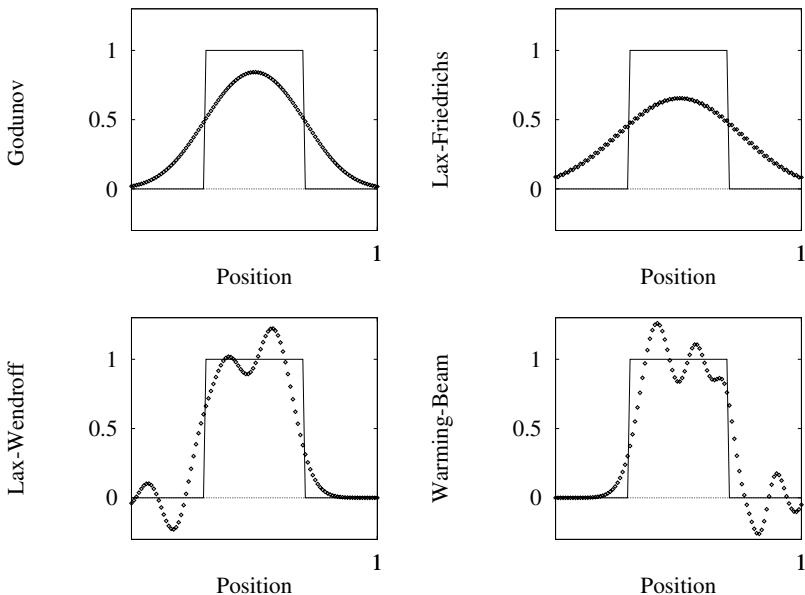
**Fig. 5.10.** Test 1: Comparison of numerical results for four numerical schemes (symbols) with the exact solution (line) at time 10 units (1250 time steps)



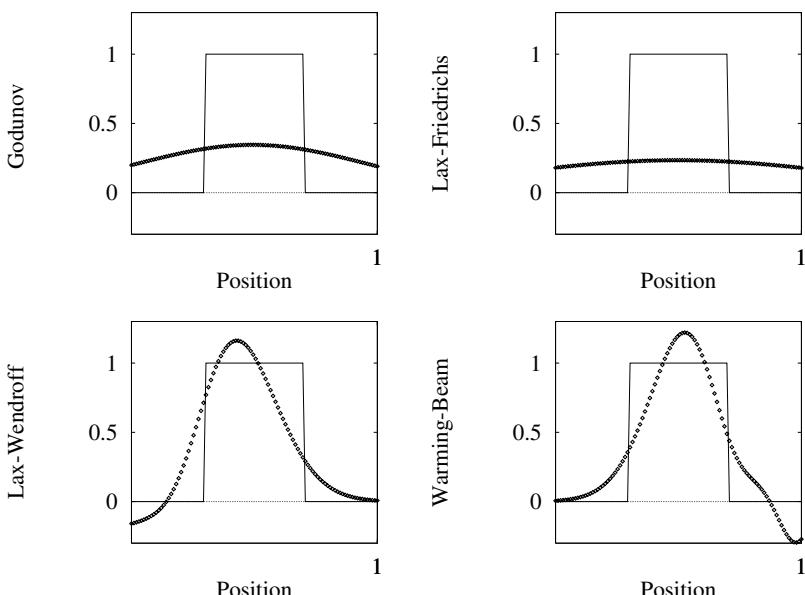
**Fig. 5.11.** Test 1: Comparison of numerical results for four numerical schemes (symbols) with the exact solution (line) at time 100 units (12499 time steps)



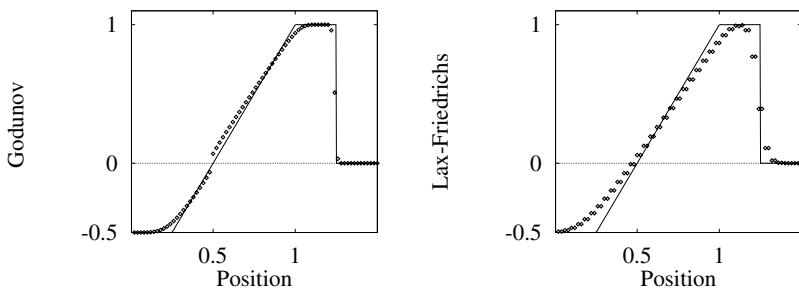
**Fig. 5.12.** Test 2: Comparison of numerical results for four numerical schemes (symbols) with the exact solution (line) at output time of 1 unit (125 time steps)



**Fig. 5.13.** Test 2: Comparison of numerical results for four numerical schemes (symbols) with the exact solution (line) at time 10 units (1250 time steps)



**Fig. 5.14.** Test 2: Comparison of numerical results for four numerical schemes (symbols) with the exact solution (line) at time 100 units (12499 time steps)



**Fig. 5.15.** Test 3: Inviscid Burgers's equation. Comparison of numerical results for two schemes (symbols) with the exact solution (line) at time 0.5 units (32 time steps)

---

# The Method of Godunov for Non-linear Systems

It was almost 40 years ago when Godunov [216] produced a conservative extension of the first-order upwind scheme of Courant, Isaacson and Rees [144] to non-linear systems of hyperbolic conservation laws. In Chap. 5 we advanced a description of Godunov's method in terms of scalar equations and linear systems with constant coefficients. In this chapter, we describe the scheme for general non-linear hyperbolic systems; in particular, we give a detailed description of the technique as applied to the time-dependent, one dimensional Euler equations. The essential ingredient of Godunov's method is the solution of the Riemann problem, which may be the exact solution or some suitable approximation to it. Here, we present the scheme in terms of the exact solution. In Chaps. 9 to 12 we shall present versions of Godunov's scheme that utilise approximate Riemann solvers; these, if used cautiously, will provide an improvement to the efficiency of the scheme. As seen in Chap. 5 the method is only first-order accurate, which makes it unsuitable for application to practical problems; well-resolved solutions will require the use of very fine meshes, with the associated computing expense. Second and third order extensions of the basic Godunov method will be studied in Chap. 13 for scalar conservation laws; some of these high-order methods are extended to non-linear systems in Chaps. 14 and 16.

Relevant background for studying the Godunov's method is found in all preceding chapters, but detailed review of Chaps. 4 and 5 might be found particularly helpful.

## 6.1 Bases of Godunov's Method

Consider the general Initial-Boundary Value Problem (IBVP) for non-linear systems of hyperbolic conservation laws

$$\left. \begin{array}{l} \text{PDEs : } \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0} , \\ \text{ICs : } \mathbf{U}(x, 0) = \mathbf{U}^{(0)}(x) , \\ \text{BCs : } \mathbf{U}(0, t) = \mathbf{U}_l(t) , \quad \mathbf{U}(L, t) = \mathbf{U}_r(t) . \end{array} \right\} \quad (6.1)$$

Here,  $\mathbf{U}(x, t)$  is the vector of conserved variables;  $\mathbf{F}(\mathbf{U})$  is the vector of fluxes;  $\mathbf{U}^{(0)}(x)$  is the initial data at time  $t = 0$ ;  $[0, L]$  is the spatial domain and boundary conditions are, for the moment, assumed to be represented by the boundary functions  $\mathbf{U}_l(t)$  and  $\mathbf{U}_r(t)$ . We make the assumption that the solution of IVBP (6.1) does exist.

In order to admit discontinuous solutions we must use one of the integral forms of the conservation laws in (6.1). Here we adopt

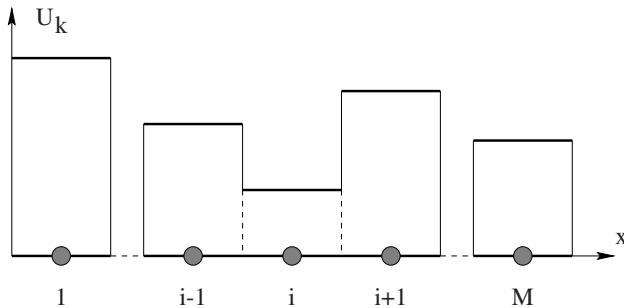
$$\left. \begin{aligned} \int_{x_1}^{x_2} \mathbf{U}(x, t_2) dx &= \int_{x_1}^{x_2} \mathbf{U}(x, t_1) dx + \int_{t_1}^{t_2} \mathbf{F}(\mathbf{U}(x_1, t)) dt \\ &\quad - \int_{t_1}^{t_2} \mathbf{F}(\mathbf{U}(x_2, t)) dt , \end{aligned} \right\} \quad (6.2)$$

for any control volume  $[x_1, x_2] \times [t_1, t_2]$  in the domain of interest; see Sect. 2.4.1 of Chap. 2.

We discretise the spatial domain  $[0, L]$  into  $M$  computing cells or finite volumes  $I_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$  of regular size  $\Delta x = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}} = L/M$ , with  $i = 1, \dots, M$ . For a given cell  $I_i$  the location of the cell centre  $x_i$  and the cell boundaries  $x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}$  are given by

$$x_{i-\frac{1}{2}} = (i-1)\Delta x, \quad x_i = (i - \frac{1}{2})\Delta x, \quad x_{i+\frac{1}{2}} = i\Delta x. \quad (6.3)$$

See Fig. 5.4 of Chap. 5. We denote the temporal domain by  $[0, T]$ , where  $T$  is some output time, not a boundary. The discretisation of the time interval  $[0, T]$  is generally done in time steps  $\Delta t$  of variable size; recall that for non-linear systems wave speeds vary in space and time, and thus the choice of  $\Delta t$  is carried out as marching in time proceeds. Given general initial data



**Fig. 6.1.** Piece-wise constant distribution of data at time level  $n$ , for a single component of the vector  $\mathbf{U}$

$\tilde{\mathbf{U}}(x, t^n)$  for (6.1) at time  $t = t^n$  say, in order to evolve the solution to a time  $t^{n+1} = t^n + \Delta t$ , the Godunov method first assumes a *piece-wise constant* distribution of the data. Formally, this is realised by defining *cell averages*

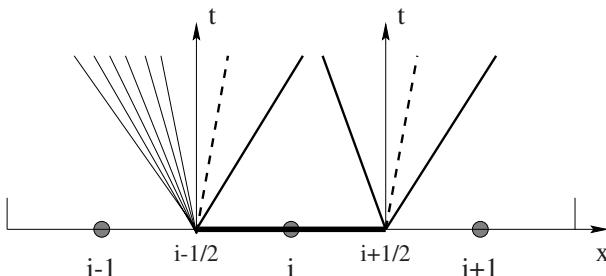
$$\mathbf{U}_i^n = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \tilde{\mathbf{U}}(x, t^n) dx , \quad (6.4)$$

which produces the desired piecewise constant distribution  $\mathbf{U}(x, t^n)$ , with

$$\mathbf{U}(x, t^n) = \mathbf{U}_i^n, \text{ for } x \text{ in each cell } I_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] , \quad (6.5)$$

as illustrated in Fig. 6.1 for a single component  $\mathbf{U}_k$  of the vector of conserved variables. Data now consists of a set  $\{\mathbf{U}_i^n\}$  of constant states. Naturally these are in terms of conserved variables, but other variables may be derived to proceed with the implementation of numerical methods. In particular, for the Godunov method we use the solution of the Riemann problem in terms of primitive variables, which for the Euler equations are  $\mathbf{W} = (\rho, u, p)^T$ ;  $\rho$  is density,  $u$  is velocity and  $p$  is pressure.

Once the piece-wise constant distribution of data has been established the next step in the Godunov method is to solve the Initial Value Problem (IVP) for the original conservation laws but with the modified initial data (6.5). Effectively, this generates local Riemann problems  $RP(\mathbf{U}_i^n, \mathbf{U}_{i+1}^n)$  with data  $\mathbf{U}_i$  (left side) and  $\mathbf{U}_{i+1}^n$  (right side), centred at the intercell boundary positions  $x_{i+\frac{1}{2}}$ . As seen for the Euler equations in Chap. 4, the solution of  $RP(\mathbf{U}_i^n, \mathbf{U}_{i+1}^n)$  is a similarity solution and depends on the ratio  $\bar{x}/\bar{t}$ , see (6.7), and the data states  $\mathbf{U}_i^n, \mathbf{U}_{i+1}^n$ ; the solution is denoted by  $\mathbf{U}_{i+\frac{1}{2}}(\bar{x}/\bar{t})$ , where  $(\bar{x}, \bar{t})$  are the local coordinates for the local Riemann problem. Fig. 6.2 shows typical wave patterns emerging from intercell boundaries  $x_{i-\frac{1}{2}}$  and  $x_{i+\frac{1}{2}}$  when solving the two Riemann problems  $RP(\mathbf{U}_{i-1}^n, \mathbf{U}_i^n)$  and  $RP(\mathbf{U}_i^n, \mathbf{U}_{i+1}^n)$ . For a



**Fig. 6.2.** Typical wave patterns emerging from solutions of local Riemann problems at intercell boundaries  $i - \frac{1}{2}$  and  $i + \frac{1}{2}$

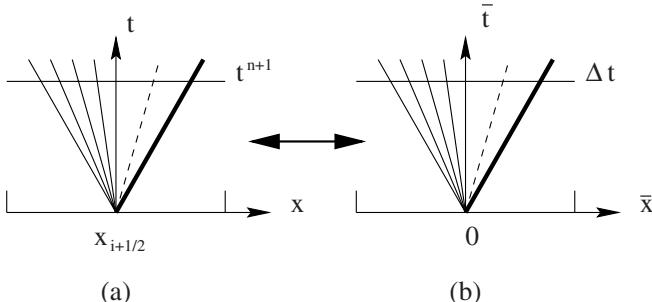
time step  $\Delta t$  that is sufficiently small, to avoid wave interaction, one can define a global solution  $\tilde{\mathbf{U}}(x, t)$  in the strip  $0 \leq x \leq L$ ,  $t^n \leq t \leq t^{n+1}$  in terms of the local solutions as follows

$$\tilde{\mathbf{U}}(x, t) = \mathbf{U}_{i+\frac{1}{2}}(\bar{x}/\bar{t}) , \quad x \in [x_i, x_{i+1}] , \quad (6.6)$$

where the correspondence between the global  $(x, t)$  and local  $(\bar{x}, \bar{t})$  coordinates is given by

$$\left. \begin{aligned} \bar{x} &= x - x_{i+\frac{1}{2}} & \bar{t} &= t - t^n, \\ x &\in [x_i, x_{i+1}] & t &\in [t^n, t^{n+1}], \\ \bar{x} &\in [-\frac{\Delta x}{2}, \frac{\Delta x}{2}] & \bar{t} &\in [0, \Delta t], \end{aligned} \right\} \quad (6.7)$$

and is illustrated in Fig. 6.3. Having found a solution  $\tilde{\mathbf{U}}(x, t)$  in terms of solu-



**Fig. 6.3.** Correspondence between the global (a) and local (b) frames of reference for the solution of the Riemann problem

tions  $\mathbf{U}_{i+\frac{1}{2}}(\bar{x}/\bar{t})$  to local Riemann problems, the Godunov method advances the solution to a time  $t^{n+1} = t^n + \Delta t$  by defining a new set of average values  $\{\mathbf{U}_i^{n+1}\}$ , in a way to be described. We shall often use  $(x, t)$  to mean the local frame of reference  $(\bar{x}, \bar{t})$ .

## 6.2 The Godunov Scheme

The first version of Godunov's method defines new average values  $\mathbf{U}_i^{n+1}$  at time  $t^{n+1} = t^n + \Delta t$  via the integrals

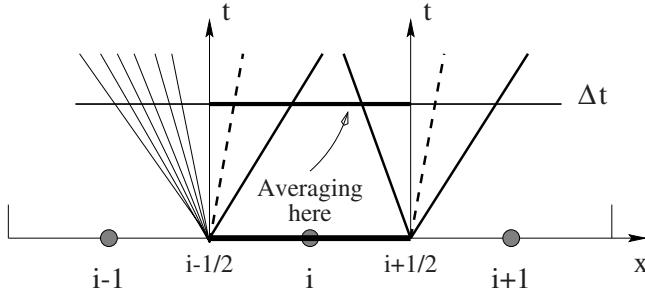
$$\mathbf{U}_i^{n+1} = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \tilde{\mathbf{U}}(x, t^{n+1}) dx \quad (6.8)$$

within each cell  $I_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ . This averaging process is illustrated in Fig. 6.4.

Note first that in order to perform the averaging, we need to make the assumption that no wave interaction takes place within cell  $I_i$ , in the chosen time  $\Delta t$ . This is satisfied by imposing the following restriction on the size of  $\Delta t$ , namely

$$\Delta t \leq \frac{\frac{1}{2} \Delta x}{S_{\max}^n}, \quad (6.9)$$

where  $S_{\max}^n$  denotes the maximum *wave velocity* present throughout the domain at time  $t^n$ . A consequence of this restriction is that only two Riemann



**Fig. 6.4.** Godunov averaging of local solutions to Riemann problems within cell  $I_i$  at a fixed time  $\Delta t$

problem solutions affect cell  $I_i$ , namely the right travelling waves of  $\mathbf{U}_{i-\frac{1}{2}}(x/t)$  and the left travelling waves of  $\mathbf{U}_{i+\frac{1}{2}}(x/t)$ , see Fig. 6.4. Thus  $\mathbf{U}_i^{n+1}$ , given by (6.8), can be expressed as

$$\mathbf{U}_i^{n+1} = \frac{1}{\Delta x} \int_0^{\frac{1}{2}\Delta x} \mathbf{U}_{i-\frac{1}{2}}\left(\frac{x}{\Delta t}\right) dx + \frac{1}{\Delta x} \int_{-\frac{1}{2}\Delta x}^0 \mathbf{U}_{i+\frac{1}{2}}\left(\frac{x}{\Delta t}\right) dx , \quad (6.10)$$

after using (6.6) and (6.8). This version of Godunov's method can obviously be implemented as a practical computational scheme. We note however that it has two main drawbacks. First, the CFL-like condition (6.9) is computationally somewhat restrictive on  $\Delta t$ . Second, the evaluation of the integrals in (6.10), although possible, could be involved. Rarefaction waves are bound to add to the complexity of the scheme. The *second version* of Godunov's method is more attractive and is given by the following statement.

**Proposition 6.1.** *The Godunov method can be written in conservative form*

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n + \frac{\Delta t}{\Delta x} [\mathbf{F}_{i-\frac{1}{2}} - \mathbf{F}_{i+\frac{1}{2}}] , \quad (6.11)$$

with intercell numerical flux given by

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}(0)) , \quad (6.12)$$

if the time step  $\Delta t$  satisfies the condition

$$\Delta t \leq \frac{\Delta x}{S_{\max}^n} . \quad (6.13)$$

*Proof.* The integrand  $\tilde{\mathbf{U}}(x, t)$  in (6.8) is an exact solution of the conservation laws, see equation (6.6). We can therefore apply the integral form (6.2) of the conservation laws to any control volume  $[x_1, x_2] \times [t_1, t_2]$ . In particular, we can apply it to the case in which  $x_1 = x_{i-\frac{1}{2}}$ ,  $x_2 = x_{i+\frac{1}{2}}$ ,  $t_1 = t^n$ ,  $t_2 = t^{n+1}$ . From (6.4) we then have

$$\left. \begin{aligned} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \tilde{\mathbf{U}}(x, t^{n+1}) dx &= \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \tilde{\mathbf{U}}(x, t^n) dx \\ &+ \int_0^{\Delta t} \mathbf{F}[\tilde{\mathbf{U}}(x_{i-\frac{1}{2}}, t)] dt - \int_0^{\Delta t} \mathbf{F}[\tilde{\mathbf{U}}(x_{i+\frac{1}{2}}, t)] dt . \end{aligned} \right\} \quad (6.14)$$

In terms of local solutions, as in (6.6), and assuming condition (6.13) we have

$$\left. \begin{aligned} \tilde{\mathbf{U}}(x_{i-\frac{1}{2}}, t) &= \mathbf{U}_{i-\frac{1}{2}}(0) = \text{constant} , \\ \tilde{\mathbf{U}}(x_{i+\frac{1}{2}}, t) &= \mathbf{U}_{i+\frac{1}{2}}(0) = \text{constant} , \end{aligned} \right\} \quad (6.15)$$

where  $\mathbf{U}_{i+\frac{1}{2}}(0)$  is the solution of the Riemann problem  $RP(\mathbf{U}_i^n, \mathbf{U}_{i+1}^n)$  along the ray  $x/t = 0$ , which is the t-axis in the local frame. Similarly,  $\mathbf{U}_{i-\frac{1}{2}}(0)$  is the solution of  $RP(\mathbf{U}_{i-1}^n, \mathbf{U}_i^n)$  along the t-axis. Division of (6.14) through by  $\Delta x$  gives

$$\left. \begin{aligned} \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \tilde{\mathbf{U}}(x, t^{n+1}) dx &= \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \tilde{\mathbf{U}}(x, t^n) dx \\ &+ \frac{\Delta t}{\Delta x} [\mathbf{F}(\mathbf{U}_{i-\frac{1}{2}}(0)) - \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}(0))] , \end{aligned} \right\} \quad (6.16)$$

which by virtue of (6.4) and (6.15) leads to the desired result (6.11)–(6.12), and thus the proposition has been proved.

The following remarks are in order:

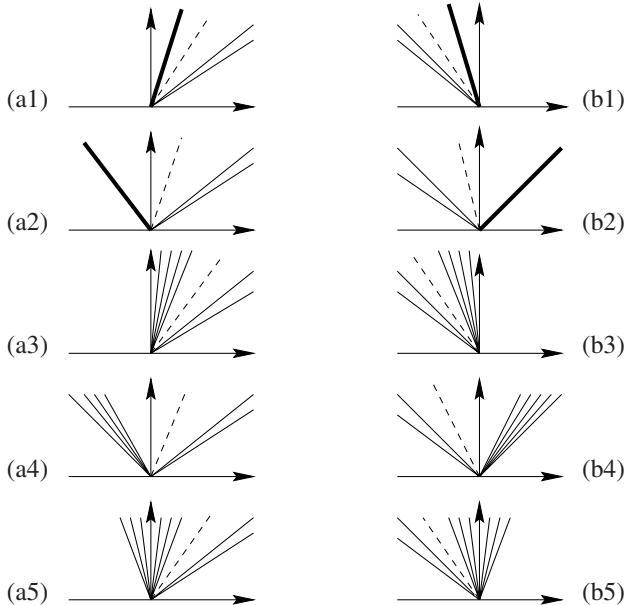
- The CFL condition (6.13) for the second version (6.11)–(6.12) of the Godunov method is more generous than (6.9), thus allowing a larger time step. This in turn results in a more efficient time-marching scheme. Here a wave is allowed to travel, at most, a complete cell length  $\Delta x$  in a time  $\Delta t$ .
- Condition (6.13) remains valid even if wave interaction takes place in time  $\Delta t$  within cell  $I_i$ , under the assumption that no wave acceleration takes place as a consequence of wave interaction; this is a kind of linearity assumption. Condition (6.13) is necessary in (6.16) when computing the fluxes along the left and right intercell boundaries.
- The second version (6.11)–(6.12) of the Godunov method is the one that is used for practical computations.

### 6.3 Godunov's Method for the Euler Equations

Here we describe Godunov's method for the specific case of the time-dependent, one-dimensional Euler equations. As data  $\{\mathbf{U}_i^n\}$  at time level  $n$  is assumed, in order to march the solution to time level  $n+1$  via the conservative formula (6.11) we need to compute the intercell fluxes  $\mathbf{F}_{i-\frac{1}{2}}$  and  $\mathbf{F}_{i+\frac{1}{2}}$ .

### 6.3.1 Evaluation of the Intercell Fluxes

For a generic cell interface at  $x_{i+\frac{1}{2}}$  we compute the Godunov flux  $\mathbf{F}_{i+\frac{1}{2}}$  according to (6.12). We therefore require the solution  $\mathbf{U}_{i+\frac{1}{2}}(x/t)$  of the Riemann problem  $RP(\mathbf{U}_i^n, \mathbf{U}_{i+1}^n)$  evaluated at the point  $S = x/t = 0$ .



**Fig. 6.5.** Possible wave patterns in evaluating the Godunov flux for the Euler equations: (a) positive particle speed in the Star Region (b) negative particle speed in the Star Region

In Chap. 4 we presented the complete exact solution to a general Riemann problem  $RP(\mathbf{U}_i^n, \mathbf{U}_{i+1}^n)$  for the Euler equations. In practice we use the solution in terms of the primitive variables, which we denote by  $\mathbf{W}_{i+\frac{1}{2}}(x/t)$ . Having found  $\mathbf{W}_{i+\frac{1}{2}}(x/t)$  its evaluation at any point  $S = x/t$  is carried out by the subroutine `SAMPLE` in the FORTRAN program given in Sect. 4.9 of Chap. 4. Sampling requires the identification of ten possible wave patterns; these are illustrated in Fig. 6.5. The flow chart of Fig. 4.14 in Chap. 4 relates to the five sub-cases arising from the case in which the sampling point  $S$  lies to the left of the contact discontinuity. There is an analogous flow chart for the five sub-cases arising from the case in which the sampling point  $S$  lies to the right of the contact discontinuity. For the Godunov method the sampling is performed for the special value  $S = x/t = 0$ . Unfortunately, this does not simplify the sampling procedure and all ten possible wave patterns must be taken into account; these are shown in Fig. 6.5. Recall that in our convention

a shock is a single thick ray, a contact is a dashed line and a rarefaction wave is obviously a fan. A wave of unknown character is represented by a pair of rays emanating from the origin. There are two situations, each of which has five cases, namely (a)  $0 \leq u_*$  (positive particle speed in the *Star Region*) and (b)  $0 \geq u_*$  (negative particle speed in the *Star Region*). The sampled value  $\mathbf{W}_{i+\frac{1}{2}}(0)$  needed for evaluating the Godunov flux is given in Table 6.1 for all ten possible wave patterns; see Fig. 6.5. Consider for example the situation in which  $u_*$  is positive. In order to compute correctly the value  $\mathbf{W}_{i+\frac{1}{2}}(0)$  along the  $t$ -axis (left of contact) we must identify the character of the left wave. This can be a shock, cases (a1) and (a2), or a rarefaction wave, cases (a3), (a4) and (a5). If the left wave is a shock wave we compute the state  $\mathbf{W}_{*L}$  between the left shock and the contact using shock relations, see Sect. 4.5.1 of Chap. 4. Then the speed  $S_L$  of the left shock is computed. This then allows us to test whether the shock speed is positive (supersonic flow) or negative (subsonic flow). If  $S_L \geq 0$  then

$$\mathbf{W}_{i+\frac{1}{2}}(0) = \mathbf{W}_L .$$

If  $S_L \leq 0$  then

$$\mathbf{W}_{i+\frac{1}{2}}(0) = \mathbf{W}_{*L} .$$

The analysis for the remaining cases (a3) to (a5) is analogous, as is for the set of cases (b1) to (b5). Details are omitted.

Sub-case	Case (a): positive speed $u_*$	case (b): negative speed $u_*$
1	$\mathbf{W}_L$	$\mathbf{W}_R$
2	$\mathbf{W}_{*L}$	$\mathbf{W}_{*R}$
3	$\mathbf{W}_L$	$\mathbf{W}_R$
4	$\mathbf{W}_{*L}$	$\mathbf{W}_{*R}$
5	$\mathbf{W}_{Lf}$	$\mathbf{W}_{Rf}$

Table 6.1. Value of  $\mathbf{W}_{i+\frac{1}{2}}(0)$  required for evaluating the Godunov flux, for all ten possible wave patterns in the solution of the Riemann problem

Having identified the desired value  $\mathbf{W}_{i+\frac{1}{2}}(0)$  the intercell (6.12) becomes

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}(\mathbf{W}_{i+\frac{1}{2}}(0)) .$$

**Exercise 6.2.** Construct a flow chart for computing the Godunov flux for the time-dependent, one-dimensional Euler equations.

**Solution 6.3.** (Left to the reader).

**Exercise 6.4.** Draw all possible wave patterns required for evaluating the Godunov flux for the isentropic equations of Gas Dynamics; see Sect. 2.4.4 of Chap. 4. Construct a flow chart for computing the Godunov flux.

**Solution 6.5.** (Left to the reader).

### 6.3.2 Time Step Size

So far we know how to compute the intercell flux (6.12) to be used in the conservative formula (6.11). The spatial discretisation length  $\Delta x$  is chosen on desired accuracy or available computing resources. What remains to be determined in (6.11) is the size of the time step  $\Delta t$ . This is based on the condition (6.13). The time step is then given by

$$\Delta t = \frac{C_{\text{cfl}} \Delta x}{S_{\max}^n}. \quad (6.17)$$

Here  $C_{\text{cfl}}$  is a Courant or CFL coefficient satisfying

$$0 < C_{\text{cfl}} \leq 1. \quad (6.18)$$

The closer the coefficient  $C_{\text{cfl}}$  is to 1, the more efficient the time marching scheme is.  $S_{\max}^n$  is the largest wave speed present throughout the domain at time level  $n$ . This means that no wave present in the solution of all Riemann problems travels more than a distance  $\Delta x$  in time  $\Delta t$ . As discussed in Chap. 5 in the context of simple problems, there are various ways of estimating  $S_{\max}^n$ . For the time-dependent, one dimensional Euler equations a reliable choice is

$$S_{\max}^n = \max_i \left\{ |S_{i+\frac{1}{2}}^{\text{L}}|, |S_{i+\frac{1}{2}}^{\text{R}}| \right\}, \quad (6.19)$$

for  $i = 0, \dots, M$ , where  $S_{i+\frac{1}{2}}^{\text{L}}, S_{i+\frac{1}{2}}^{\text{R}}$  are the wave speeds of the left and right non-linear waves present in the solution of the Riemann problem  $RP(\mathbf{U}_i^n, \mathbf{U}_{i+1}^n)$ . Recall that this Riemann problem generates three waves; the fastest are the non-linear waves, which can be shocks or rarefactions. For rarefaction waves one selects the speed of the head. For shock waves one selects the shock speed, naturally. Note that in sampling the wave speeds one must include the boundaries, as these might generate large wave speeds. Using (6.19) to find  $S_{\max}^n$  and thus  $\Delta t$  according to (6.17), is a simple and very reliable procedure. As the local solutions of Riemann problems are available for flux evaluation, it is just a question of using this information to find  $\Delta t$  as well. For multi-dimensional problems however, this scheme for estimating the maximum wave speed is unsuitable; see Sect. 16.3.2 of Chap. 16.

A popular alternative for estimating  $S_{\max}^n$ , which extends to multi-dimensional problems, is

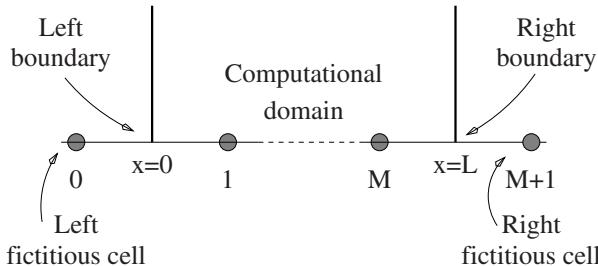
$$S_{\max}^n = \max_i \{|u_i^n| + a_i^n\}. \quad (6.20)$$

Only data values for the particle velocity  $u_i^n$  and sound speed  $a_i^n$  are used here. It is not difficult to see however that (6.20) can lead to an underestimate of  $S_{\max}^n$ . For instance, assume shock-tube data in which the flow is stationary at time  $t = 0$ . Then  $u_i^n = 0$  and the sound speed is the only contribution to  $S_{\max}^n$ . Underestimating  $S_{\max}^n$  results in a choice of  $\Delta t$  that is too large and instabilities may be developed from the beginning of the computations.

A possible way of remedying this, is by choosing the CFL coefficient  $C_{\text{cfl}}$  in (6.17) cautiously. If  $S_{\max}^n$  is known reliably then the choice  $C_{\text{cfl}} = 1$  is probably adequate, although this implies that waves pass through each other without acceleration, which is a kind of linearity assumption. A practical choice is  $C_{\text{cfl}} = 0.9$ . If there are uncertainties in the estimate for  $S_{\max}^n$ , such as when (6.20) is used, a more conservative choice for  $C_{\text{cfl}}$  is advised. In spite of the alluded disadvantages of choice (6.20), it provides a practical approach when computing solutions to multi-dimensional problems. See Chap. 16.

### 6.3.3 Boundary Conditions

For a domain  $[0, L]$  discretised into  $M$  computing cells of length  $\Delta x$  we need boundary conditions at the boundaries  $x = 0$  and  $x = L$  as illustrated in Fig. 6.6. Numerically, such boundary conditions are expected to provide numerical fluxes  $\mathbf{F}_{\frac{1}{2}}$ , and  $\mathbf{F}_{M+\frac{1}{2}}$ . These are required in order to apply the conservative formula (6.11) to update the extreme cells  $I_1$  and  $I_M$  to the next time level  $n + 1$ . The boundary conditions may result in direct prescription of  $\mathbf{F}_{\frac{1}{2}}$  and  $\mathbf{F}_{M+\frac{1}{2}}$ . Alternatively, we may prescribe *fictitious data values* in the fictitious cells  $I_0$  and  $I_{M+1}$ , adjacent to  $I_1$  and  $I_M$  respectively; see Fig. 6.6. In this way, boundary Riemann problems  $RP(\mathbf{U}_0^n, \mathbf{U}_1^n)$  and  $RP(\mathbf{U}_M^n, \mathbf{U}_{M+1}^n)$  are solved and the corresponding Godunov fluxes  $\mathbf{F}_{\frac{1}{2}}$  and  $\mathbf{F}_{M+\frac{1}{2}}$  are computed, as done for the interior cells. The imposition of boundary conditions



**Fig. 6.6.** Boundary conditions. Fictitious cells outside the computational domain are created

is, fundamentally, a physical problem. Great care is required in their numerical implementation. For the Godunov method this task tends to be facilitated by the fact that local Riemann problem solutions are used. Here we consider only two types of boundaries: *reflective* and *transparent* or *transmissive*.

#### Reflective Boundaries

Consider the boundary at  $x = L$  and suppose it physically consists of a fixed, reflective impermeable wall. Then the physical situation is correctly

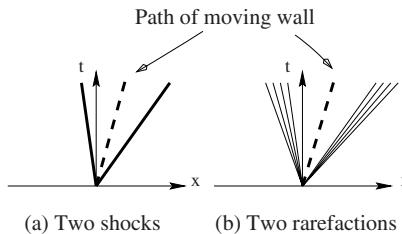
modelled by creating a fictitious state  $\mathbf{W}_{M+1}^n$  on the right hand side of the boundary and defining the boundary Riemann problem  $RP(\mathbf{W}_M^n, \mathbf{W}_{M+1}^n)$ . The fictitious state  $\mathbf{W}_{M+1}^n$  is defined from the known state  $\mathbf{W}_M^n$  inside the computational domain, namely

$$\rho_{M+1}^n = \rho_M^n, \quad u_{M+1}^n = -u_M^n, \quad p_{M+1}^n = p_M^n. \quad (6.21)$$

The exact solution of this boundary Riemann problem consists of either (i) two shock waves if  $u_M^n > 0$  or (ii) two rarefaction waves if  $u_M^n \leq 0$ . In both cases  $u_* = 0$  along the boundary; this is the desired condition at the solid, fixed impermeable boundary. Consequently, the only non-zero contribution to the flux vector at the boundary is in the momentum component and is due to the pressure  $p_*$  corresponding to  $u_* = 0$ . In both cases the solution can be obtained in closed form, no iteration is required. As a matter of fact, a closed-form solution exists for the more general case in which the fluid under consideration obeys the covolume equation of state and the impermeable wall moves with a prescribed speed  $u_{\text{wall}}$  [498]. The boundary conditions are

$$\rho_{M+1}^n = \rho_M^n, \quad u_{M+1}^n = -u_M^n + 2u_{\text{wall}}, \quad p_{M+1}^n = p_M^n. \quad (6.22)$$

The exact solution of the Riemann problem  $RP(\mathbf{W}_M^n, \mathbf{W}_{M+1}^n)$  containing a moving boundary is symmetric about the path of the moving wall and consists of either (a) two shocks or (b) two rarefactions, with the contact wave coinciding with the moving wall, as desired. See Fig. 6.7.



**Fig. 6.7.** Boundary Riemann problem for moving wall. Contact surface coincides with moving solid boundary: (a) solution consists of two shocks and the contact (b) solution consists of two rarefactions and the contact

We now find the exact solution for  $p_*$  and  $u_*$  in the moving-wall Riemann problem. From the analysis of the exact function for pressure, see Sect. 4.3 of Chap 4, it is seen that if  $\Delta u = -2(u_M - u_{\text{wall}}) = 0$ , that is  $u_M = u_{\text{wall}}$ , then the solution  $p_*$  for pressure at the boundary is  $p_* = p_M = p_{M+1}$ . For  $\Delta u > 0$  we have  $p_* < p_M = p_{M+1}$ , that is, the solution consists of two rarefaction waves, see Fig. 6.7b. For  $\Delta u < 0$  we have  $p_* > p_M = p_{M+1}$  and the solution consists of two shocks, see Fig. 6.7a. For the case of two rarefaction waves,  $u_M \leq u_{\text{wall}}$ , direct utilisation of the data in the pressure function  $f(p) = 0$ , see Sect. 4.3 of Chap 4, gives

$$p_* = p_M \left[ 1 + \frac{1}{2}(\gamma - 1) \left( \frac{C_M}{a_M} \right) \right]^{\frac{2\gamma}{\gamma-1}}. \quad (6.23)$$

For the case of two shocks,  $u_M > u_{\text{wall}}$ , we have

$$p_* = p_M + \frac{C_M}{2A_M} \left\{ C_M + [C_M^2 + 4A_M(B_M + p_M)]^{\frac{1}{2}} \right\}, \quad (6.24)$$

where

$$A_M = \frac{2}{(\gamma + 1)\rho_M}, \quad B_M = \frac{(\gamma - 1)}{(\gamma + 1)}p_M, \quad C_M = u_M - u_{\text{wall}}. \quad (6.25)$$

As anticipated, the solution for the velocity  $u_*$  in both cases is found to be

$$u_* = u_{\text{wall}}. \quad (6.26)$$

These closed-form solutions for the pressure and velocity at the boundary (fixed or moving) can also be utilised in the Godunov method even when this is used in conjunction with approximate Riemann solvers, particularly if these are thought to be inaccurate for boundary data Riemann problems. A useful discussion on solid-body boundary conditions for the Euler equations in multi-dimensional domains is given by Rizzi [404]. A recommended paper on boundary conditions for hyperbolic problems is that of Thompson [489].

## Transmissive Boundaries

Transmissive, or transparent boundaries arise from the need to define finite, or sufficiently small, computational domains. The corresponding boundary conditions are a numerical attempt to produce boundaries that allow the passage of waves without any effect on them. For one-dimensional problems the objective is reasonably well attained. For multi-dimensional problems this is a substantial area of current research, usually referred to as *open-end* boundary conditions, *transparent* boundary conditions, *far-field* boundary conditions, *radiation* boundary conditions or *non-reflecting* boundary conditions. For a transmissive right boundary we suggest the boundary conditions

$$\rho_{M+1}^n = \rho_M^n, \quad u_{M+1}^n = u_M^n, \quad p_{M+1}^n = p_M^n. \quad (6.27)$$

This data produces a trivial Riemann problem. No wave of finite strength is produced at the boundary that may affect the flow inside the domain. Useful publications dealing with transparent boundary conditions are those of Giles [205], Bayliss and Turkel [34], Roe [413] and Karni [277].

For an assumed mesh of size  $\Delta x$ , we have defined all details for the practical implementation of the Godunov method, see (6.11)–(6.13). These are

- intercell fluxes
- the maximum wave speed  $S_{max}^n$  to compute the time step size  $\Delta t$ , and

- boundary conditions.

*Remark 6.6.* The wave speeds generated at the boundaries, after applying boundary conditions, must be taken into account when selecting the time step  $\Delta t$ .

**Exercise 6.7.** Write a flow chart to implement the Godunov method to solve the one-dimensional, time dependent Euler equations in a tube of constant cross sectional area. Assume the left wall is impermeable and fixed and the right wall is transparent.

**Solution 6.8.** (Left to the reader).

## 6.4 Numerical Results and Discussion

Here we illustrate the performance of Godunov's first-order upwind method for the Euler equations on test problems with exact solution. For comparison we also show numerical results obtained by the Lax–Friedrichs and Richtmyer (or two-step Lax–Wendroff) methods, discussed in Chap. 5. In all chosen tests, data consists of two constant states  $\mathbf{W}_L = (\rho_L, u_L, p_L)^T$  and  $\mathbf{W}_R = (\rho_R, u_R, p_R)^T$ , separated by a discontinuity at a position  $x = x_0$ . The states  $\mathbf{W}_L$  and  $\mathbf{W}_R$  are given in Table 6.2. The ratio of specific heats is chosen to be  $\gamma = 1.4$ . For all test problems the spatial domain is the interval  $[0, 1]$  which is discretised with  $M = 100$  computing cells. The Courant number coefficient is  $C_{\text{cfl}} = 0.9$ ; boundary conditions are transmissive and  $S_{\max}^n$  is found using the simplified formula (6.20).

*Remark 6.9.* Given that formula (6.20) is not reliable, see discussion in Sect. 6.3.2, in all computations presented here we take, for the the first 5 time steps, a Courant number coefficient  $C_{\text{cfl}}$  reduced by a factor of 0.2. This allows for waves to begin to form, after which formula (6.20) becomes more reliable.

The exact solutions were found by running the code HE-E1RPEXACT of the library *NUMERICA* [519] and the numerical solutions were obtained by running the code HE-E1GODSTATE of *NUMERICA*.

Test	$\rho_L$	$u_L$	$p_L$	$\rho_R$	$u_R$	$p_R$
1	1.0	0.75	1.0	0.125	0.0	0.1
2	1.0	-2.0	0.4	1.0	2.0	0.4
3	1.0	0.0	1000.0	1.0	0.0	0.01
4	5.99924	19.5975	460.894	5.99242	-6.19633	46.0950
5	1.0	-19.59745	1000.0	1.0	-19.59745	0.01

Table 6.2. Data for five test problems with exact solution. Test 5 is like Test 3 with negative uniform background speed

Test 1 is a *modified version* of the popular Sod's test [453]; the solution consists of a right shock wave, a right travelling contact wave and a left *sonic* rarefaction wave; this test is very useful in assessing the *entropy satisfaction* property of numerical methods. Test 2 has solution consisting of two symmetric rarefaction waves and a trivial contact wave of zero speed; the *Star Region* between the non-linear waves is close to vacuum, which makes this problem a suitable test for assessing the performance of numerical methods for low-density flows; this is the so called *123 problem* introduced in chapter Chap. 4. Test 3 is designed to assess the robustness and accuracy of numerical methods; its solution consists of a strong shock wave, a contact surface and a left rarefaction wave. Test 4 is also designed to test robustness of numerical methods; the solution consists of three strong discontinuities travelling to the right. See Sect. 4.3.3 of Chap. 4 for more details on the exact solution of these test problems. Test 5 is also designed to test the robustness of numerical methods but the main reason for devising this test is to assess the ability of the numerical methods to resolve *slowly-moving contact discontinuities*. The exact solution of Test 5 consists of a left rarefaction wave, a right-travelling shock wave and a *stationary* contact discontinuity. For each test we select a convenient position  $x_0$  of the initial discontinuity and an output time. These are stated in the legend of each figure displaying computational results.

Figs. 6.8 to 6.12 show comparisons between exact solutions (line) and numerical solutions (symbol) at a given output time obtained by the Godunov method, for all five test problems. The quantities shown are density  $\rho$ , particle speed  $u$ , pressure  $p$  and specific internal energy  $e$ . For comparison, we also solved these test problems using the Lax–Friedrichs method, see Figs. 6.13 to 6.17, and the Richtmyer (or two-step Lax–Wendroff) method, which failed to produce solutions to Tests 2 to 5. For Test 1 the solution of the Richtmyer scheme is shown in Fig. 6.18.

#### 6.4.1 Numerical Results for Godunov's Method

The results for Test 1, shown in Fig. 6.8, are typical of the Godunov's first-order accurate method described in this chapter.

The numerical approximation of the shock wave, of zero-width transition in the exact solution, has a transition region of width approximately  $4\Delta x$ ; that is, the shock has been *smeared* over 4 computing cells. This spreading of shock waves may seem unsatisfactory, but it is quite typical of numerical solutions; in fact most first-order methods will spread a shock wave even more. A satisfactory feature of the numerical shock wave of Fig. 6.8 is that it is monotone, there are no spurious oscillations in the vicinity of the shock, at least for this example. Monotonicity of shock waves computed by the Godunov method depends on the speed of the shock and it holds in most cases except when the shock speed is very close to zero. The contact discontinuity, seen in the density and internal energy plots, is smeared over 20 cells; generally contact waves are much more difficult to resolve accurately than shock waves.

This is due to the linear character of contacts; characteristics on either side of the wave run parallel to the wave. In shock waves, characteristics on either side of the wave run into the shock, a compression mechanism that helps the numerical resolution of shock waves. As for the shock case, the solution for the contact is perfectly monotone.

Another positive feature of the numerical approximation of the discontinuities is that their speed of propagation is correct and thus their average positions are correct. This is a consequence of the conservative character of Godunov's method. The rarefaction wave is a smooth flow feature and is reasonably well approximated by the method except near the head and the tail, where a discontinuity in derivative exists. Another visible error in the rarefaction is the small discontinuous jump within the rarefaction. This is sometimes referred to as the *entropy glitch* and arises only in the presence of *sonic* rarefaction waves, as in the present case. Godunov's method is theoretically entropy satisfying [244] and we therefore expect the size of the jump in the entropy glitch to tend to zero as the mesh size  $\Delta x$  tends to zero. Fig. 6.19 shows the result obtained by refining the mesh by a factor of 5. It appears as if the numerical solution does converge to the exact solution.

The performance of Godunov's method on Test 2, see Fig. 6.9, is generally quite satisfactory as regards the physical variables  $p$ ,  $u$  and  $\rho$  but not so much for the specific internal energy, which is computed from  $\rho$  and  $p$  as  $e = p/((\gamma - 1)\rho)$ . In this low density example both pressure and density are close to zero and thus small errors will be exaggerated by their ratio. In any case, it is generally accepted that plots of the internal energy  $e$  can be quite revealing of the quality of the numerical solution. On the other hand pressure is probably the easiest quantity to get right. The main point of Test 2 is to make the reader aware that this class of low density flows can easily cause numerical methods to fail; even the robust Godunov method fails if used in conjunction with certain approximate Riemann solvers [182]. The Richtmyer (or two-step Lax-Wendroff) method fails to give a solution to this problem.

Test 3 is a very severe problem and is designed to test the robustness of the Godunov method, the results of which are shown in Fig. 6.10. The emerging right travelling shock wave has pressure ratio  $p_*/p_R = 46000$  and a corresponding shock Mach number of 198. For flows involving such strong shock waves as this, one would seriously question the validity of the ideal gas equations of state. However, from the point of view of assessing the robustness of numerical schemes, the validity of the test problem as a mathematical/numerical problem still holds. As for Test 1, the resolution of discontinuities is worst for the contact wave; as a consequence of this, post shock values are not attained, as is clearly seen in the density plot. The velocity plot shows a kind of overshoot near the tail of the rarefaction. The Richtmyer (or two-step Lax-Wendroff) scheme failed for this test.

As seen in Fig. 6.11 the solution of Test 4 consists of three discontinuities: two shock waves and a contact. They all travel to the right; the left shock has a small positive speed. The complete wave system has resulted from the

interaction of two strong shock waves propagating in opposite directions. The right shock is the fastest wave and is smeared over 5 cells, as seen in the pressure plot; the left, slowly moving shock is sharply resolved (two cells) but is not monotone; there are some *low frequency* spurious oscillations in its vicinity, as seen in the internal energy plot. The contact discontinuity is heavily smeared. Slowly moving shocks are sharply resolved by Godunov's method. In fact, *isolated* shocks and contacts of zero speed are perfectly resolved, if non-defective Riemann solvers are used, see Chap. 10. The phenomenon of spurious oscillations in slowly moving shocks has been studied by Roberts [406] and is so far, to the author's knowledge, an unresolved difficulty. Billett and Toro [60] investigated some possible cures of the problem for the Euler equations. See also the recent papers by Arora and Roe [19] and by Karni and Čanić [280].

Test 5 is like Test 3 but with a uniform, negative background speed so as to obtain a virtually stationary contact discontinuity. In addition to testing the robustness of numerical methods, Test 5 is mainly designed to test the ability of numerical methods to resolve slowly-moving or stationary contact discontinuities. Fig. 6.12 shows the result obtained from the Godunov method as compared with the exact solution. For this test problem the contact discontinuity is *virtually stationary*; the Godunov method resolves this discontinuity very sharply indeed. This result should be compared with that obtained by the Lax–Friedrichs method, Fig. 6.17. Test 5 is a very challenging test problem, as we shall illustrate in subsequent chapters dealing with other numerical methods.

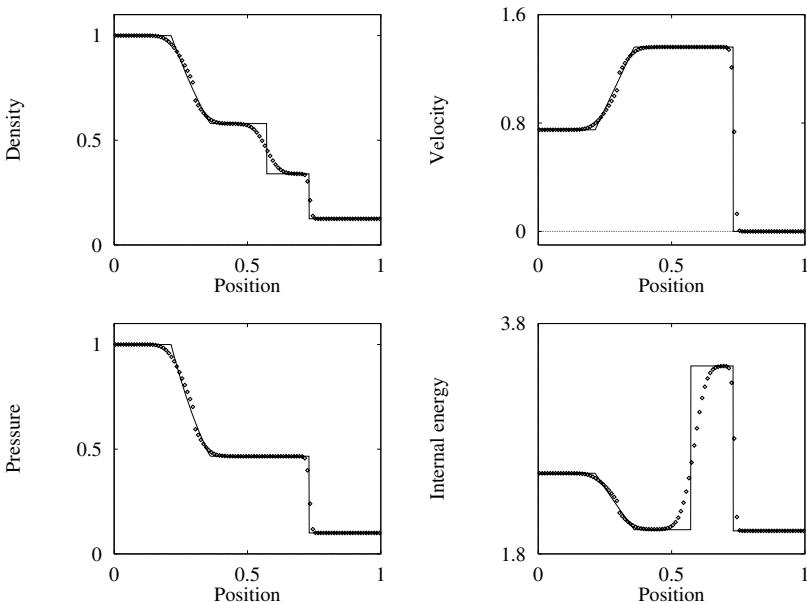
#### 6.4.2 Numerical Results from Other Methods

First we apply the Lax–Friedrichs method, see Sect. 5.3.4 of Chap. 5, to Tests 1 to 5. The numerical results are shown in Figs. 6.13 to 6.17. The results for Test 1 are shown in Fig. 6.13 and are to be compared with those of Godunov's method, Fig. 6.8. The Lax–Friedrichs scheme has the peculiar property of *pairing* cell values, which enhances smearing. The shock wave is resolved with about 8 cells and looks acceptable. The resolution of the rarefaction wave and the contact discontinuity is very poor. The solution for Test 2 is shown in Fig. 6.14. Note how inaccurate the solution for internal energy is; compare with Fig. 6.9 and with the exact solution. A large class of methods are known to have difficulties with this kind of *symmetric* Riemann problems [512]. Fig. 6.15 shows results for Test 3; the scheme is unable to attain the post-shock density values; compare with Fig. 6.10. The results for Test 4 are shown in Fig. 6.16, which are to be compared with those of Fig. 6.11. The Lax–Friedrichs method, although simpler and cheaper, is significantly less accurate than the Godunov method. Fig. 6.17 shows the result from the Lax–Friedrichs scheme for Test 5; this result and that of Fig. 6.12 show the crucial difference between two major classes of numerical methods, namely *centred methods* and

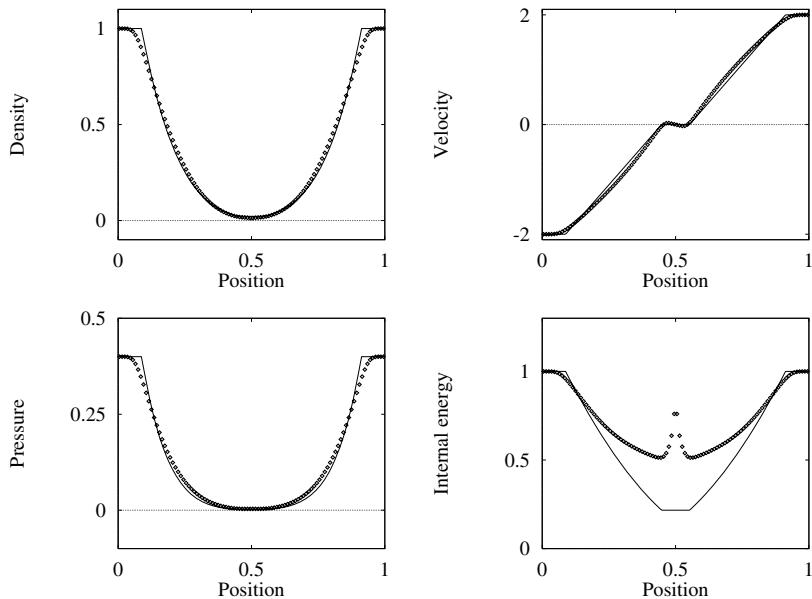
*Godunov-type methods.* Test 5 highlights the fact that resolving *linear waves* is perhaps one of the most challenging tasks for numerical methods today.

The second-order Richtmyer (or two-step Lax-Wendroff) method was applied to Test 1 and the results are shown in Fig. 6.18; compare with Figs. 6.8 and 6.13. The solution is generally more accurate in the smooth regions of the flow, as is to be expected from a second-order accurate method; discontinuities are also more sharply resolved but *spurious oscillations* near discontinuities appear; see Chap. 5. This method failed to give a solution at all, for Tests 2 to 5.

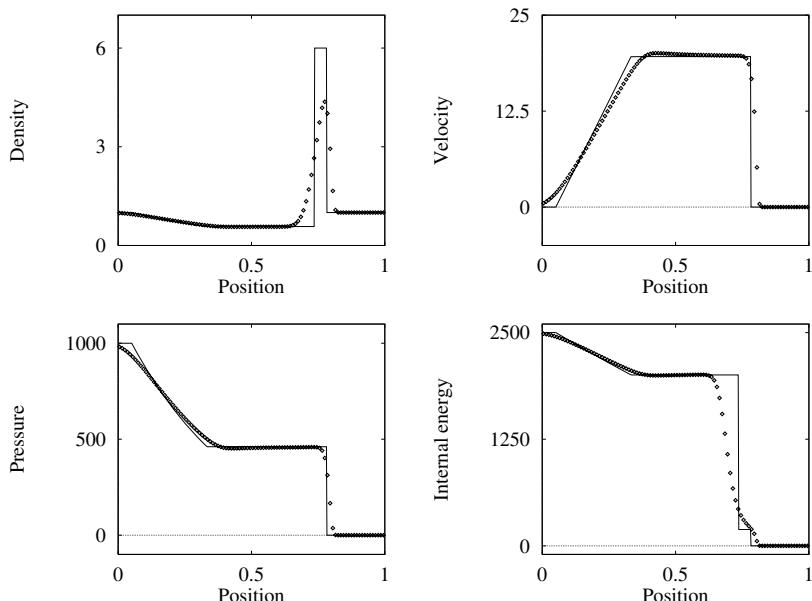
In this chapter we have presented the Godunov method as used in conjunction with the exact Riemann solver. Godunov's method can also be used with approximate Riemann solvers. In Chaps. 9 to 12 we present several approximate Riemann solvers for the Euler equations. Second and third order extensions of Godunov's method are presented in Chap. 13 for scalar problems. In Chaps. 14 and 16 we present second-order TVD schemes for non-linear systems.



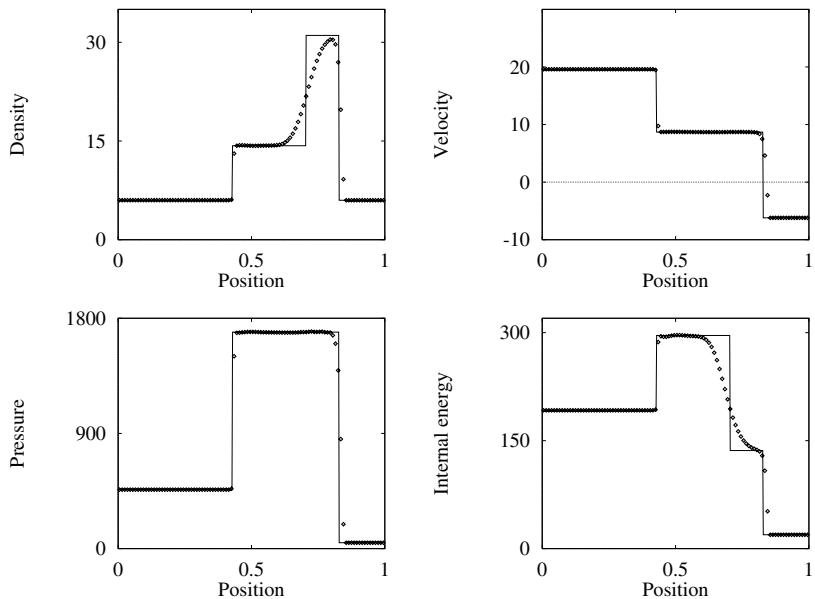
**Fig. 6.8.** Godunov's method applied to Test 1, with  $x_0 = 0.3$ . Numerical (symbol) and exact (line) solutions are compared at the output time 0.2 units



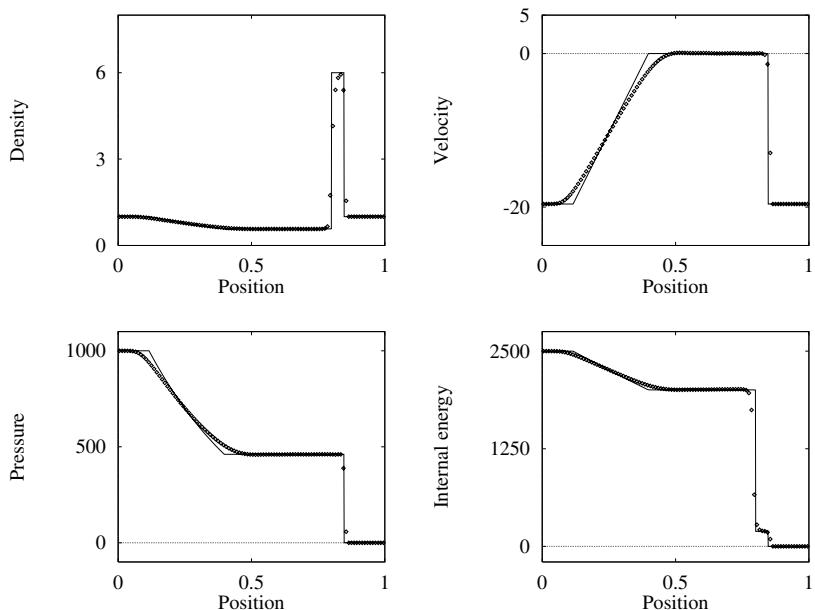
**Fig. 6.9.** Godunov's method applied to Test 2, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at the output time 0.15 units



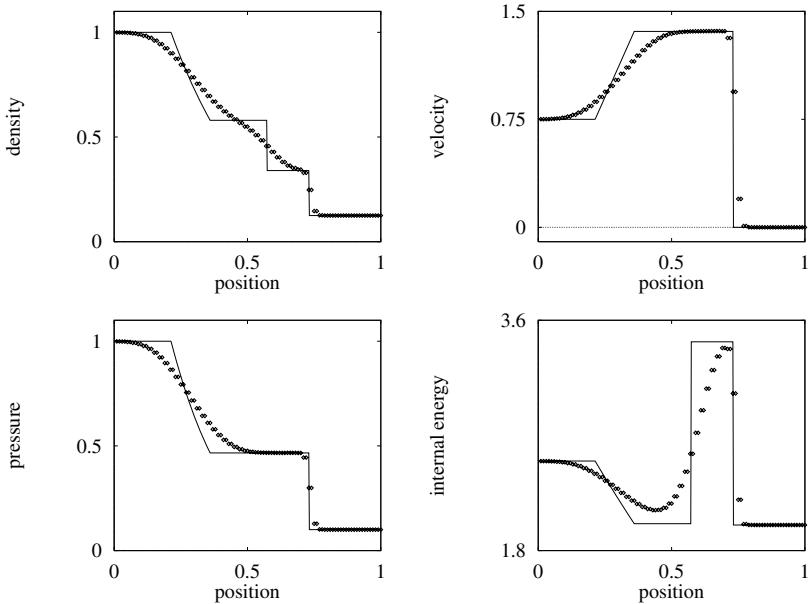
**Fig. 6.10.** Godunov's method applied to Test 3, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at the output time 0.012 units



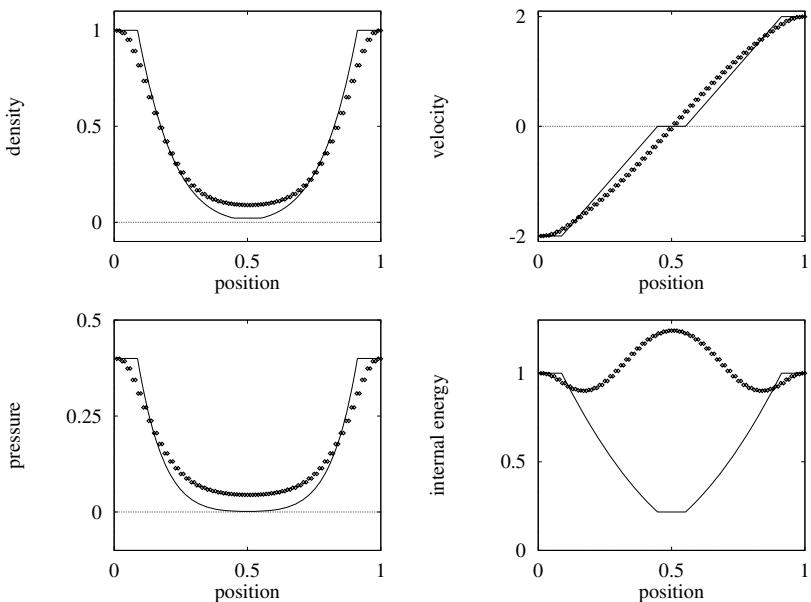
**Fig. 6.11.** Godunov's method applied to Test 4, with  $x_0 = 0.4$ . Numerical (symbol) and exact (line) solutions are compared at the output time 0.035 units



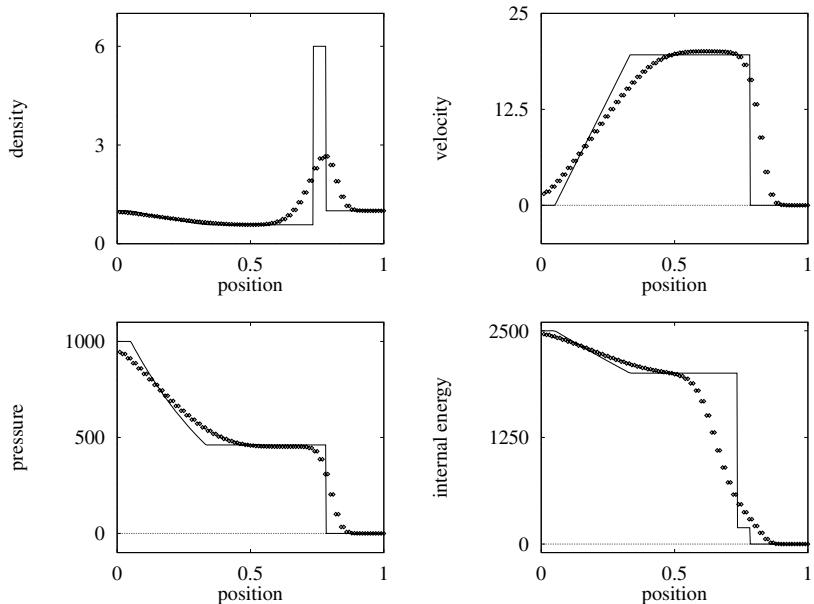
**Fig. 6.12.** Godunov's method applied to Test 5, with  $x_0 = 0.8$ . Numerical (symbol) and exact (line) solutions are compared at the output time 0.012



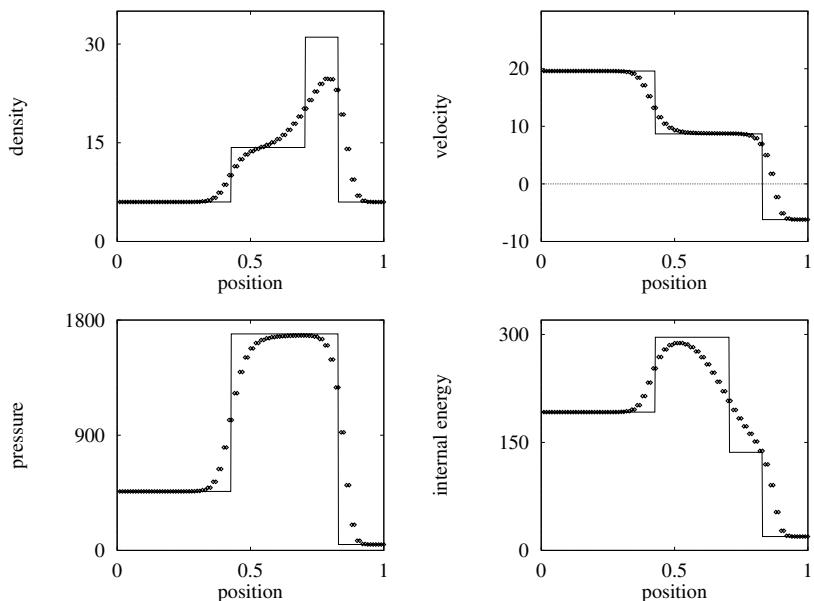
**Fig. 6.13.** The Lax–Friedrichs method applied to Test 1, with  $x_0 = 0.3$ . Numerical (symbol) and exact (line) solutions are compared at the output time 0.2 units



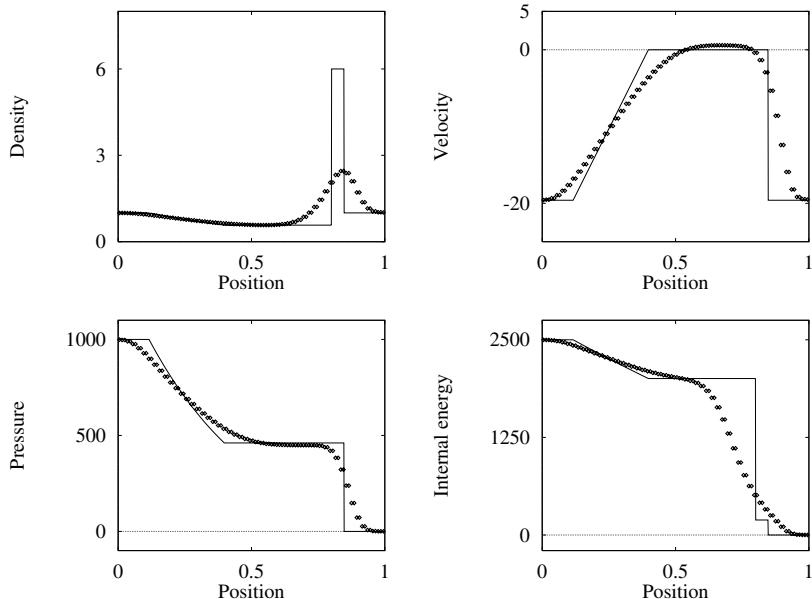
**Fig. 6.14.** The Lax–Friedrichs method applied to Test 2, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at the output time 0.15 units



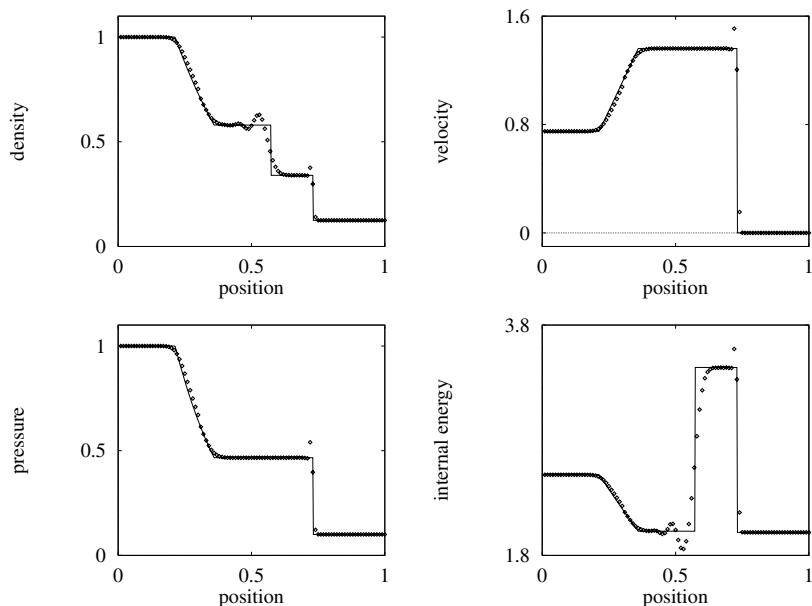
**Fig. 6.15.** The Lax–Friedrichs method applied to Test 3, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at the output time 0.012 units



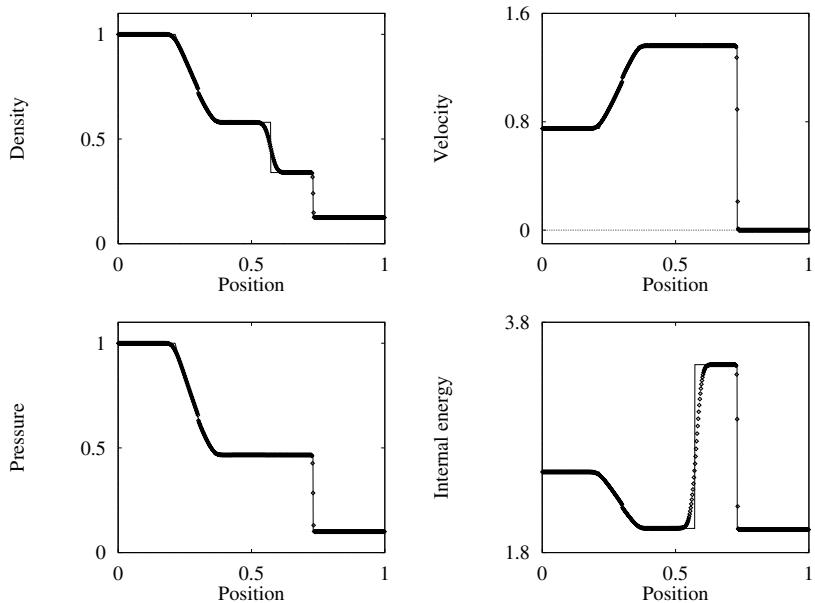
**Fig. 6.16.** The Lax–Friedrichs method applied to Test 4, with  $x_0 = 0.4$ . Numerical (symbol) and exact (line) solutions are compared at the output time 0.035 units



**Fig. 6.17.** The Lax–Friedrichs method applied to Test 5, with  $x_0 = 0.8$ . Numerical (symbol) and exact (line) solutions are compared at the output time 0.012



**Fig. 6.18.** The Richtmyer method applied to Test 1, with  $x_0 = 0.3$ . Numerical (symbol) and exact (line) solutions are compared at the output time 0.2 units



**Fig. 6.19.** Godunov's method applied to Test 1 with fine mesh,  $M = 500$ . Numerical (symbol) and exact (line) solutions are compared at time 0.2 units

## Random Choice and Related Methods

### 7.1 Introduction

In 1965, Glimm [212] introduced the Random Choice Method (RCM) as part of a constructive proof of existence of solutions to a class of non-linear systems of hyperbolic conservation laws. In 1976, Chorin [110] successfully implemented a modified version of the method, as a computational technique, to solve the Euler equations of Gas Dynamics. In essence, to implement the RCM one requires (i) exact solutions of local Riemann problems and (ii) a random sampling procedure to pick up states to be assigned to the next time level. As we shall see, there is a great deal of commonality between the RCM and the Godunov method presented in Chap. 6. Both schemes use the exact solution of the Riemann problem, although Godunov's method can also be implemented using approximate Riemann solvers, as we shall see in Chaps. 9 to 12. The two methods differ in the way the local Riemann problem solutions are utilised to march to the next time level: the Godunov method takes an integral *average* of local solutions of Riemann problems, while the RCM picks a single state, contained in the local solutions, at *random*. The random sampling procedure is carried out by employing a sequence of random numbers. The statistical properties of these random numbers have a significant effect on the accuracy of the Random Choice Method.

Since the introduction of the RCM as a computational scheme by Chorin, there have been many contributions to the development of the method. Chorin himself [111] extended the RCM to combustion problems; Sod [452] applied the RCM to the one-dimensional Euler equations for cylindrically and spherically symmetric flows, thereby introducing a way of dealing with algebraic source terms. Concus [138] applied the RCM to a non-linear scalar equation governing the two-phase flow of petroleum in underground reservoirs. Major contributions to the method were presented by Colella [131], [132]; these include a better understanding of the method, its strengths and limitations, and improved random sampling techniques. Marshall and Mendez [339] applied the RCM to the one-dimensional shallow water equations. Li and Holt [317]

applied the RCM to the study of underwater explosions. Marshall and Plohr [340] applied the RCM to solve the steady supersonic Euler equations, see also Shi and Gottlieb [443], and to the study of shock wave diffraction phenomena. Gottlieb [221] compared the implementation of the RCM on staggered and non-staggered grids and introduced an effective way of using irregular meshes. Toro [498] applied the RCM to covolume gases with moving boundaries. Applications of the RCM to the study of reactive flows were performed by Saito and Glass [422], Takano [476], Singh and Clarke [449] and Dawes [151]. Olivier and Grönig [364] applied the RCM to solve the two-dimensional time dependent Euler equations to study shock focussing and diffraction phenomena in water and air.

Essentially, the RCM is applicable to scalar problems in any number of dimensions and to non-linear systems in two independent variables. Examples of these systems are the one-dimensional, time dependent Euler equations, the two-dimensional, steady supersonic Euler equations and the one-dimensional shallow water equations. By using splitting schemes, see Chap. 15, one can also solve extensions of these systems to include algebraic source terms or even terms to model viscous diffusion; see Sod [455] and Honma and Glass [256]. A fundamental limitation of the RCM is its inability to solve multi-dimensional non-linear systems via splitting schemes, which usually work well when extending other schemes to multi-dimensional problems; see Chap. 16. An attraction of the RCM is its ability to handle complex wave interaction involving discontinuities such as shock waves and material interfaces; these are resolved as true discontinuities. Most other methods will smear discontinuities over several computing cells, a problem that is particularly serious for contact surfaces. Although computed discontinuities in the RCM have infinite resolution, the position of these waves at any given time has an error, which is random in character. The randomness of the RCM also shows in resolving smooth waves, such as rarefactions. Such randomness is tolerable when solving homogeneous systems, i.e. no source terms. In the presence of source terms however, the randomness tends to be enhanced.

This chapter is primarily devoted to the conventional Random Choice Method, but we also present what appears to be a new random choice method [513] that is analogous to the Lax–Friedrichs (deterministic) scheme. In addition we present a, deterministic, first-order centred (FORCE) scheme based on a reinterpretation of the conventional RCM on a staggered grid. The presentation of the schemes is given in terms of the time-dependent, one dimensional Euler equations. The reader is advised to review Chap. 4 before proceeding with the study of the present chapter.

## 7.2 RCM on a Non-Staggered Grid

We consider the general Initial Boundary Value Problem (IBVP) for non-linear systems of hyperbolic conservation laws, namely

$$\left. \begin{array}{l} \text{PDEs : } \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0}, \\ \text{ICs : } \mathbf{U}(x, 0) = \mathbf{U}^{(0)}(x), \\ \text{BCs : } \mathbf{U}(0, t) = \mathbf{U}_l(t), \quad \mathbf{U}(L, t) = \mathbf{U}_r(t). \end{array} \right\} \quad (7.1)$$

We assume a solution to this IBVP exists. Here  $\mathbf{U}(x, t)$  is the vector of conserved variables,  $\mathbf{F}(\mathbf{U})$  is the vector of fluxes,  $\mathbf{U}^{(0)}(x)$  is the initial data at time  $t = 0$ ,  $[0, L]$  is the spatial domain and boundary conditions are, for the moment, assumed to be represented by the boundary functions  $\mathbf{U}_l(t)$  and  $\mathbf{U}_r(t)$ .

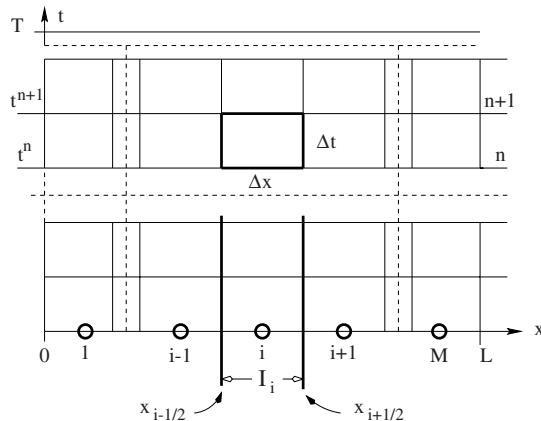
In the RCM the only step in which one is required to work with the vector of conserved variables is at the level of the Riemann problem, when enforcing the Rankine–Hugoniot Conditions at shocks. All other steps of the method are more conveniently performed in terms of the vector of primitive variables, which for the Euler equations are  $\mathbf{W} = (\rho, u, p)^T$ ;  $\rho$  is density,  $u$  is velocity and  $p$  is pressure.

### 7.2.1 The Scheme for Non–Linear Systems

As for the Godunov method studied in Chap. 6, we discretise the spatial domain  $[0, L]$  into  $M$  computing cells  $I_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$  of size  $\Delta x = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}} = L/M$ , with  $i = 1, \dots, M$ . For a given cell  $I_i$ , the location of the cell centre  $x_i$  and the cell boundaries  $x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}$  are given by

$$x_{i-\frac{1}{2}} = (i-1)\Delta x, \quad x_i = (i - \frac{1}{2})\Delta x, \quad x_{i+\frac{1}{2}} = i\Delta x. \quad (7.2)$$

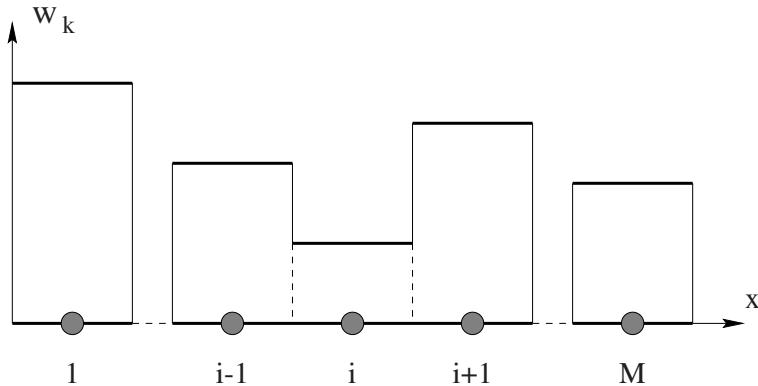
For convenience we choose cells of regular size  $\Delta x$ , but this is not a necessary



**Fig. 7.1.** Discretisation of domain for the Random Choice Method on a non–staggered grid

requirement for implementing the RCM. Fig. 7.1 illustrates the non–staggered

grid arrangement for this version of the Random Choice Method. The solution is updated at the cell centres  $x_i$ , every time step. Obviously, the time step is, in general, of different size for every time step. Given general initial

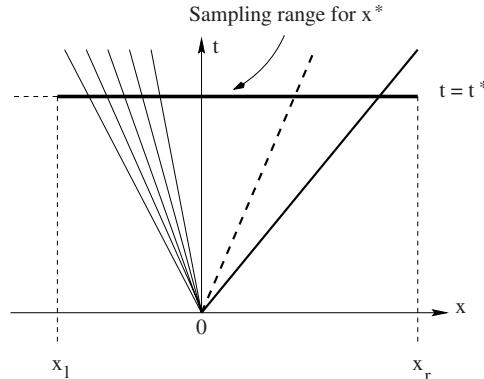


**Fig. 7.2.** Piecewise constant distribution of data at time level  $n$ . Pairs of neighbouring states define data for local Riemann problems

data  $\mathbf{W}(x, t^n)$  at time  $t = t^n$  say, in order to evolve the solution to a time  $t^{n+1} = t^n + \Delta t$ , the Random Choice Method first assumes a *piecewise constant* distribution of the data. Formally, this may be realised by defining *cell averages* as in the Godunov method, see Sect. 6.1 of Chap. 6. For the RCM this is not necessary and we assume that the given data at the cell centres  $x_i$  is constant throughout the respective cell  $I_i$ . We then have  $\mathbf{W}(x, t^n) = \mathbf{W}_i^n$  in each cell  $I_i$ . Fig. 7.2 shows the distribution of a typical variable  $w_k$  at a given time level  $n$ .

The pairs of neighbouring, constant, states  $\mathbf{W}_i^n, \mathbf{W}_{i+1}^n$  define local Riemann problems  $RP(\mathbf{W}_i^n, \mathbf{W}_{i+1}^n)$ , which have similarity solutions  $\mathbf{W}_{i+\frac{1}{2}}(x/t)$ . In Chap. 4 we provided the complete exact solution to the Riemann problem for the Euler equations along with a deterministic sampling procedure contained in the FORTRAN 77 program of Sect. 4.9. Given a time  $t^*$ , the sampling routine SAMPLE evaluates  $\mathbf{W}_{i+\frac{1}{2}}(x/t^*)$  at any point  $x$  in an interval  $[x_l, x_r]$  with  $x_l \leq 0 \leq x_r$ . A detailed understanding of the complete exact solution of the Riemann problem is essential for understanding and implementing the RCM. Fig. 7.3 illustrates the structure of a typical Riemann problem solution and a typical sampling range of the solution at a given time  $t^*$ , across the wave structure. In the Random Choice Method the particular point  $x = x^*$  is picked up at random within the sampling range  $[x_l, x_r]$ . The sampling routine SAMPLE evaluates  $\mathbf{W}_{i+\frac{1}{2}}(x^*/t^*)$  automatically. See Sect. 4.9 of Chap. 4.

The Random Choice Method updates the solution from the data value  $\mathbf{W}_i^n$  in cell  $I_i$  at time level  $n$ , to the value  $\mathbf{W}_i^{n+1}$  at time level  $n+1$ , in two steps as follows:

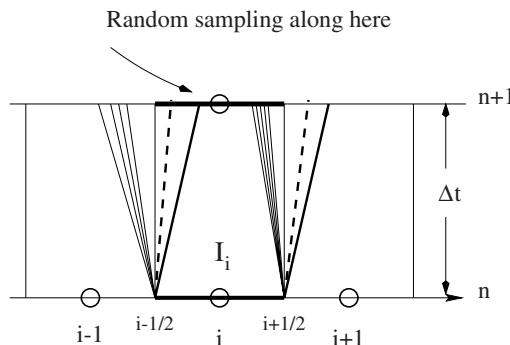


**Fig. 7.3.** Spacial sampling range in the solution of the Riemann problem at a given time  $t = t^*$

Step I: Solve the Riemann problems  $RP(\mathbf{W}_{i-1}^n, \mathbf{W}_i^n)$  and  $RP(\mathbf{W}_i^n, \mathbf{W}_{i+1}^n)$  to find their respective solutions  $\mathbf{W}_{i-\frac{1}{2}}(x/t)$  and  $\mathbf{W}_{i+\frac{1}{2}}(x/t)$ . Fig. 7.4 shows typical wave patterns emerging from the intercell boundaries  $x_{i-\frac{1}{2}}$  and  $x_{i+\frac{1}{2}}$ .

Step II: *Random sample* these solutions at time  $\Delta t$  within cell  $I_i$  to pick up a state and assign it to cell  $I_i$ . The random sampling range is shown in Fig. 7.4 by a thick horizontal line. The picked up state depends on a random, or quasi-random, number  $\theta^n$  in the interval  $[0, 1]$ . The updated solution is then

$$\mathbf{W}_i^{n+1} = \begin{cases} \mathbf{W}_{i-\frac{1}{2}}(\theta^n \Delta x / \Delta t), & \text{if } 0 \leq \theta^n \leq \frac{1}{2}, \\ \mathbf{W}_{i+\frac{1}{2}}((\theta^n - 1) \Delta x / \Delta t), & \text{if } \frac{1}{2} < \theta^n \leq 1. \end{cases} \quad (7.3)$$



**Fig. 7.4.** The RCM on non-staggered grid. Solution is updated to time level  $n$  by random sampling solutions of Riemann problems within cell  $I_i$  at time  $\Delta t$

The case in which  $0 \leq \theta^n \leq \frac{1}{2}$  is illustrated in Fig. 7.5. Here the updated solution depends on the random sampling procedure applied to the right side of the left Riemann problem solution  $\mathbf{W}_{i-\frac{1}{2}}(x/t)$ . The particular randomly chosen state is returned by the sampling routine **SAMPLE**, called with the argument

$$S \equiv x/t = \frac{\theta^n \Delta x}{\Delta t}.$$

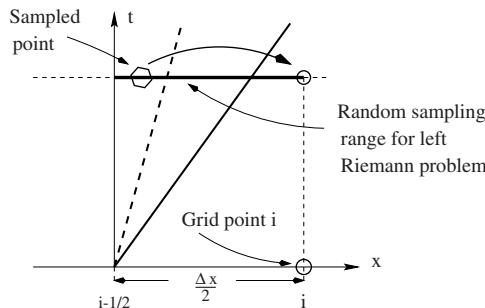
The resulting state is then assigned to the grid point  $i$ , which is regarded as the solution in cell  $I_i$  for the next time level. A similar procedure is applied if  $\frac{1}{2} \leq \theta^n \leq 1$ . In this case one samples the left side of the right Riemann problem solution  $\mathbf{W}_{i+\frac{1}{2}}(x/t)$ . When programming the non-staggered version of the Random Choice Method there are many ways of organising the tasks of (i) solving of Riemann problems and (ii) random sampling their solutions.

Concerning the use of random numbers in the scheme, Chorin [110] established that one only requires a single random number  $\theta^n$  for a complete time level  $n$ . In Glimm's proof [212] one may take one random number per time step per cell. In Sect. 7.5 we discuss generation of random numbers and their properties.

Finally, we note the relationship between the RCM scheme to obtain the updated value  $\mathbf{U}_i^{n+1}$  and the Godunov method, see Chap. 6. The Godunov scheme, instead of random sampling the solution of the relevant Riemann problems, will take the integral average of these local Riemann problem solutions

$$\mathbf{U}_i^{n+1} = \frac{1}{\Delta x} \int_0^{\frac{1}{2} \Delta x} \mathbf{U}_{i-\frac{1}{2}} \left( \frac{x}{\Delta t} \right) dx + \frac{1}{\Delta x} \int_{-\frac{1}{2} \Delta x}^0 \mathbf{U}_{i+\frac{1}{2}} \left( \frac{x}{\Delta t} \right) dx. \quad (7.4)$$

In order to preserve the conservative character of the Godunov method, the



**Fig. 7.5.** RCM sampling of the right-hand side of the left Riemann problem solution, when  $0 \leq \theta^n \leq \frac{1}{2}$ . Sampled state is assigned to the centre of cell  $I_i$

averaging is performed in terms of the conserved variables. The RCM, on the other hand, is not strictly conservative, although one may regard the scheme as being conservative in a statistical sense.

### 7.2.2 Boundary Conditions and the Time Step Size

The solution updating procedure just described is completely defined for all cells  $I_i$ , except for those next to the boundaries, namely  $I_1$  and  $I_M$ . In order to update these two cells we apply boundary conditions. This is carried out in exactly the same way as for the Godunov method, see Sects. 6.3.2 and 6.3.3 of Chap. 6. Fictitious states  $\mathbf{W}_0$  and  $\mathbf{W}_{M+1}$  adjacent to states  $\mathbf{W}_1$  and  $\mathbf{W}_M$  are defined. The otherwise missing Riemann problem solutions  $\mathbf{W}_{\frac{1}{2}}(x/t)$  and  $\mathbf{W}_{M+\frac{1}{2}}(x/t)$  at the boundaries are now defined and the random sampling procedure can now be extended to the full computational domain. We consider two types of boundary conditions, as for the one-dimensional time dependent Euler equations.

- (I) *Transmissive Boundary Conditions.* Here the fictitious states are given as

$$\left. \begin{aligned} \rho_0^n &= \rho_1^n, & u_0^n &= u_1^n, & p_0^n &= p_1^n, \\ \rho_{M+1}^n &= \rho_M^n, & u_{M+1}^n &= u_M^n, & p_{M+1}^n &= p_M^n. \end{aligned} \right\} \quad (7.5)$$

- (II) *Reflective Boundary Conditions.* Here we state the boundary conditions that apply to reflective left and right boundaries moving with respective speeds  $u_{\text{wl}}$  and  $u_{\text{wr}}$ . The fictitious states are given by

$$\left. \begin{aligned} \rho_0^n &= \rho_1^n, & u_0^n &= -u_1^n + 2u_{\text{wl}}, & p_0^n &= p_1^n, \\ \rho_{M+1}^n &= \rho_M^n, & u_{M+1}^n &= -u_M^n + 2u_{\text{wr}}, & p_{M+1}^n &= p_M^n. \end{aligned} \right\} \quad (7.6)$$

For a more complete discussion on boundary conditions see Sect. 6.3.3 of Chap. 6.

The choice of the time step  $\Delta t$  is determined by a CFL condition. Note first that in order to perform the random sampling described previously, we need to make the assumption that no wave interaction takes place within cell  $I_i$ , in the chosen time  $\Delta t$ , see Fig. 7.4. This is satisfied by imposing the following restriction on the size of  $\Delta t$ , namely

$$\Delta t \leq \frac{\frac{1}{2}\Delta x}{S_{\max}^n}, \quad (7.7)$$

where  $S_{\max}^n$  denotes the maximum *wave velocity* present throughout the domain at time  $t^n$ . A consequence of this restriction is that only two Riemann problem solutions affect cell  $I_i$ , namely the right travelling waves of  $\mathbf{W}_{i-\frac{1}{2}}(x/t)$  and the left travelling waves of  $\mathbf{W}_{i+\frac{1}{2}}(x/t)$ . Condition (7.7) may be expressed in the standard form

$$\Delta t = \frac{C_{\text{cfl}} \Delta x}{S_{\max}^n}, \quad (7.8)$$

where the CFL coefficient  $C_{\text{cfl}}$  satisfies

$$0 < C_{\text{cfl}} \leq \frac{1}{2}. \quad (7.9)$$

Hence the Random Choice Method has a stability limit of  $\frac{1}{2}$ , which is only half that of Godunov's method.

Concerning the choice of the maximum wave speed  $S_{\max}^n$  the reader is referred to Sect. 6.3.2 of Chap. 6. Virtually all relevant remarks made there in the context of the Godunov method apply. For the RCM we recommend the use of the true waves arising from the solutions of local Riemann problems. The collective experience in applying the RCM is that the scheme is not too sensitive to underestimating  $S_{\max}^n$ . If  $S_{\max}^n$  is underestimated then the chosen time step  $\Delta t$  will be too large and the likely consequence will not be signs of instabilities, as one would expect, but computed waves will propagate at *obviously* the wrong speed.

We have presented the Random Choice Method on a non-staggered grid as applied to any time-dependent one dimensional non-linear systems of hyperbolic conservation laws. Details of boundary conditions for the time-dependent one dimensional Euler equations have been given, for which numerical results are presented in Sect. 7.6.

## 7.3 A Random Choice Scheme of the Lax–Friedrichs Type

Here we present a Random Choice Method that arises from interpreting the Lax–Friedrichs scheme as an integral average of solutions of Riemann problems; see [513], [515]. If these averages are transformed by use of the integral form of the conservation laws one recovers the usual Lax–Friedrichs scheme for non-linear systems, thus eliminating the role of the Riemann problem. If the role of the Riemann problem solution is preserved and the integral averages are interpreted in a stochastic sense one obtains a Random Choice Method of the Lax–Friedrichs type.

### 7.3.1 Review of the Lax–Friedrichs Scheme

As seen in Sect. 5.3.4 of Chap. 5, the Lax–Friedrichs scheme as applied to the linear advection equation

$$u_t + f(u)_x = 0, \quad f(u) = au \quad (7.10)$$

reads

$$u_i^{n+1} = \frac{(1+c)}{2} u_{i-1}^n + \frac{(1-c)}{2} u_{i+1}^n. \quad (7.11)$$

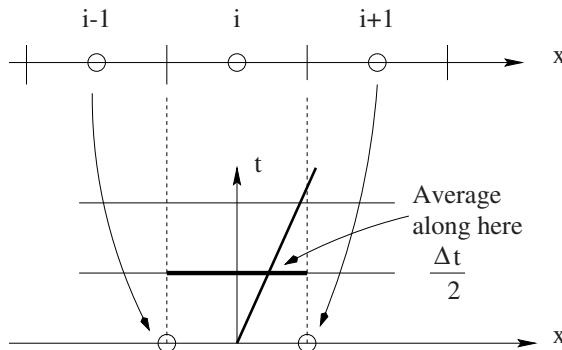
This is obviously identical to the integral average

$$u_i^{n+1} = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \hat{u}_i(x, \frac{1}{2}\Delta t) dx \quad (7.12)$$

within cell  $i$ , in which  $\hat{u}_i(x, t)$  is the solution of the Riemann problem  $RP(u_{i-1}^n, u_{i+1}^n)$ , that is

$$\hat{u}_i(x/t) = \begin{cases} u_{i-1}^n & \text{if } x/t < a, \\ u_{i+1}^n & \text{if } x/t > a. \end{cases} \quad (7.13)$$

See Fig. 7.6. The Lax–Friedrichs solution in cell  $i$  at time  $t^{n+1} = t^n + \Delta t$  is a *weighted average* of the solution of the Riemann problem with the left  $u_{i-1}^n$  and right  $u_{i+1}^n$  neighbour states as data, at time  $t = \frac{1}{2}\Delta t$ . Note the two peculiarities of the scheme, (i) the data states do not include  $u_i^n$  and (ii) the time for the averaging is half the full time step  $\Delta t$ .



**Fig. 7.6.** Reinterpretation of the Lax–Friedrichs scheme for the linear advection equation.

### 7.3.2 The Scheme

We first generalise interpretation (7.12) of the Lax–Friedrichs scheme to non–linear systems of conservation laws

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0}. \quad (7.14)$$

The scheme reads

$$\mathbf{U}_i^{n+1} = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \hat{\mathbf{U}}_i(x, \frac{1}{2}\Delta t) dx, \quad (7.15)$$

where  $\hat{\mathbf{U}}_i(x, t)$  is the solution of the Riemann problem  $RP(\mathbf{U}_{i-1}^n, \mathbf{U}_{i+1}^n)$ . There are now three possible routes to follow. These are

- (i) Solve Riemann problems and find the updated solution by evaluating (7.15) directly. Numerical results of this scheme are indistinguishable from those obtained from the conventional, much simpler, Lax–Friedrichs scheme. We therefore discard this as a useful scheme.

- (ii) Apply the integral form of the conservation laws to (7.15) to reproduce the conventional Lax–Friedrichs scheme for non–linear systems, and thus eliminate the role of the Riemann problem, see Sect. 5.3.4 of Chap. 5.
- (iii) Keep the role of the Riemann problem and reinterpret (7.15) in a stochastic sense. We obtain

$$\mathbf{U}_i^{n+1} = \frac{1}{\Delta x} \int_{-\frac{1}{2}\Delta x}^{\frac{1}{2}\Delta x} \hat{\mathbf{U}}_i(\theta^n \Delta x, \frac{1}{2}\Delta t) dx , \quad (7.16)$$

where  $\theta^n$  is a random number satisfying

$$-\frac{1}{2} \leq \theta^n \leq \frac{1}{2} . \quad (7.17)$$

A Random Choice Scheme for updating the solution to the new time level is thus obtained, namely

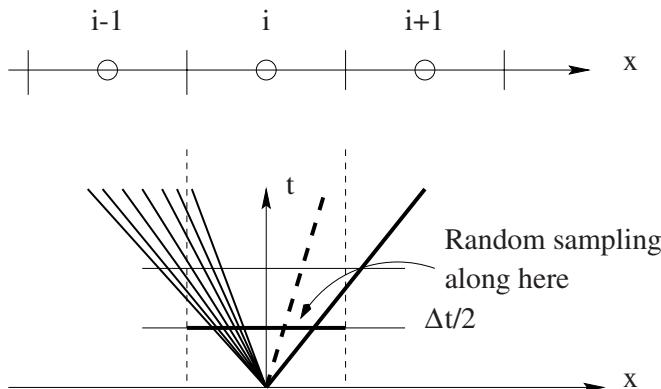
$$\mathbf{U}_i^{n+1} = \hat{\mathbf{U}}_i(\theta^n \Delta x, \frac{1}{2}\Delta t) . \quad (7.18)$$

The conventional RCM has stability restriction (7.9), while the random choice scheme (7.18) has stability condition

$$0 < C_{cfl} \leq 1 , \quad (7.19)$$

which represents an improvement by a factor of 2.

Fig. 7.7 illustrates the Random Choice Method of the Lax–Friedrichs type as applied to non–linear systems. The programming of the scheme is straightforward. Numerical results will be presented in Sect. 7.6.



**Fig. 7.7.** Random choice scheme of the Lax–Friedrichs type for non–linear systems. Updated solution in cell  $I_i$  at time  $\Delta t$  is obtained from random sampling solution of Riemann problem  $RP(\mathbf{U}_{i-1}^n, \mathbf{U}_{i+1}^n)$  at time  $\frac{1}{2}\Delta t$

## 7.4 The RCM on a Staggered Grid

As stated earlier the RCM can be implemented on a non-staggered grid and on a staggered grid. The former version was described in Sect. 7.2. Here we describe the RCM on a staggered grid and derive an associated deterministic scheme that is conservative, first-order accurate and centred.

### 7.4.1 The Scheme for Non-Linear Systems

The staggered grid version of the RCM to solve (7.14) updates  $\mathbf{U}_i^n$  to a new value  $\mathbf{U}_i^{n+1}$  in two steps, as illustrated in Fig. 7.8. The steps are:

Step (I) Solve the Riemann problems  $RP(\mathbf{U}_{i-1}^n, \mathbf{U}_i^n)$  and  $RP(\mathbf{U}_i^n, \mathbf{U}_{i+1}^n)$  to find respective solutions

$$\hat{\mathbf{U}}_{i-\frac{1}{2}}^{n+\frac{1}{2}}(x, t), \quad \hat{\mathbf{U}}_{i+\frac{1}{2}}^{n+\frac{1}{2}}(x, t). \quad (7.20)$$

*Random sample* these solutions at a stable time  $\Delta t^{n+\frac{1}{2}}$ , that is

$$\mathbf{U}_{i-\frac{1}{2}}^{n+\frac{1}{2}} = \hat{\mathbf{U}}_{i-\frac{1}{2}}^{n+\frac{1}{2}}(\theta^n \Delta x, \Delta t^{n+\frac{1}{2}}), \quad \mathbf{U}_{i+\frac{1}{2}}^{n+\frac{1}{2}} = \hat{\mathbf{U}}_{i+\frac{1}{2}}^{n+\frac{1}{2}}(\theta^n \Delta x, \Delta t^{n+\frac{1}{2}}). \quad (7.21)$$

Step (II) Solve  $RP(\mathbf{U}_{i-\frac{1}{2}}^{n+\frac{1}{2}}, \mathbf{U}_{i+\frac{1}{2}}^{n+\frac{1}{2}})$  to find solution  $\hat{\mathbf{U}}_i^{n+1}(x, t)$  and random sample it, at a stable time  $\Delta t^{n+1}$ , to obtain  $\mathbf{U}_i^{n+1}$ , that is

$$\mathbf{U}_i^{n+1} = \hat{\mathbf{U}}_i^{n+1}(\theta^{n+1} \Delta x, \Delta t^{n+1}). \quad (7.22)$$

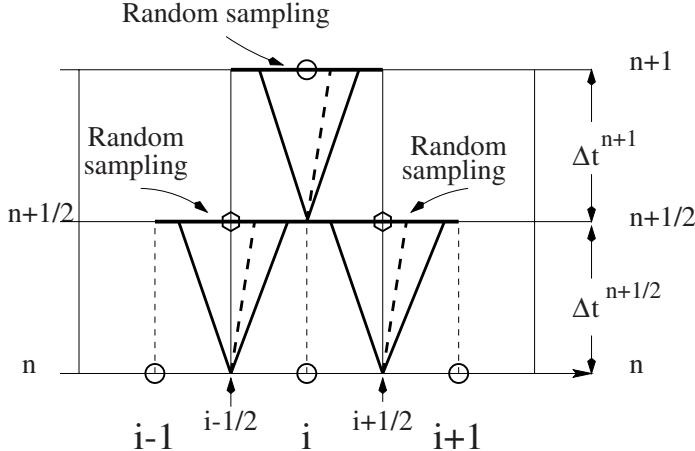
The time steps  $\Delta t^{n+\frac{1}{2}}$  and  $\Delta t^{n+1}$  need not be the same but must be chosen according to the usual stability restriction (7.8)–(7.9) for the RCM. As for the case of the non-staggered RCM, one may use the primitive variables to describe the staggered grid RCM. However, for the theme of the next section we assume the vector  $\mathbf{U}$  in (7.20)–(7.22) to be the vector of conserved variables.

### 7.4.2 A Deterministic First-Order Centred Scheme (FORCE)

Here we present a First-Order Centred deterministic scheme (FORCE) [513], [515], that is obtained by replacing the stochastic steps (7.21)–(7.22) by deterministic versions, via integral averages of Riemann problem solutions. We preserve the previous notation and assume

$$\Delta t^{n+\frac{1}{2}} = \Delta t^{n+1} = \frac{1}{2} \Delta t.$$

The stochastic integrals (7.21) are replaced by the deterministic integrals



**Fig. 7.8.** Illustration of the Random Choice Method, on a staggered grid

$$\mathbf{U}_{i-\frac{1}{2}}^{n+\frac{1}{2}} = \frac{1}{\Delta x} \int_{-\frac{1}{2}\Delta x}^{\frac{1}{2}\Delta x} \hat{\mathbf{U}}_{i-\frac{1}{2}}^{n+\frac{1}{2}}(x, \frac{\Delta t}{2}) dx \quad (7.23)$$

and

$$\mathbf{U}_{i+\frac{1}{2}}^{n+\frac{1}{2}} = \frac{1}{\Delta x} \int_{-\frac{1}{2}\Delta x}^{\frac{1}{2}\Delta x} \hat{\mathbf{U}}_{i+\frac{1}{2}}^{n+\frac{1}{2}}(x, \frac{\Delta t}{2}) dx . \quad (7.24)$$

Then we apply the integral form of the conservation laws, see Sect. 2.4.1 of Chap. 2, to expressions (7.23) and (7.24). The result is

$$\mathbf{U}_{i-\frac{1}{2}}^{n+\frac{1}{2}} = \frac{1}{2} (\mathbf{U}_{i-1}^n + \mathbf{U}_i^n) + \frac{\Delta t}{2\Delta x} (\mathbf{F}_{i-1}^n - \mathbf{F}_i^n) , \quad (7.25)$$

$$\mathbf{U}_{i+\frac{1}{2}}^{n+\frac{1}{2}} = \frac{1}{2} (\mathbf{U}_i^n + \mathbf{U}_{i+1}^n) + \frac{\Delta t}{2\Delta x} (\mathbf{F}_i^n - \mathbf{F}_{i+1}^n) . \quad (7.26)$$

We denote by  $\hat{\mathbf{U}}_i(x, t)$  the solution of the Riemann problem  $RP(\mathbf{U}_{i-\frac{1}{2}}^{n+\frac{1}{2}}, \mathbf{U}_{i+\frac{1}{2}}^{n+\frac{1}{2}})$  and define an average  $\mathbf{U}_i^{n+1}$  at the complete time step  $\Delta t$  in terms of an integral average of  $\hat{\mathbf{U}}_i(x, t)$  at the (local) time  $t = \frac{1}{2}\Delta t$ , namely

$$\mathbf{U}_i^{n+1} = \frac{1}{\Delta x} \int_{-\frac{1}{2}\Delta x}^{\frac{1}{2}\Delta x} \hat{\mathbf{U}}_i(x, \frac{1}{2}\Delta t) dx . \quad (7.27)$$

This is the deterministic version of (7.22). Applying the integral form of the conservation laws to the right-hand side of (7.27) gives

$$\mathbf{U}_i^{n+1} = \frac{1}{2} \left[ \mathbf{U}_{i-\frac{1}{2}}^{n+\frac{1}{2}} + \mathbf{U}_{i+\frac{1}{2}}^{n+\frac{1}{2}} \right] + \frac{\Delta t}{2\Delta x} \left[ \mathbf{F}_{i-\frac{1}{2}}^{n+\frac{1}{2}} - \mathbf{F}_{i+\frac{1}{2}}^{n+\frac{1}{2}} \right] , \quad (7.28)$$

where

$$\mathbf{F}_{i+\frac{1}{2}}^{n+\frac{1}{2}} = \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}^{n+\frac{1}{2}}) \equiv \mathbf{F}_{i+\frac{1}{2}}^{\text{RI}} \quad (7.29)$$

and  $\mathbf{F}_{i+\frac{1}{2}}^{\text{RI}}$  is the intercell numerical flux for the Richtmyer scheme; see Sect. 5.3.4 of Chap. 5.

Thus the deterministic version of the staggered-grid RCM scheme (7.21)–(7.22) becomes (7.28). The scheme is obviously conservative and when written in conservation form we have

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n + \frac{\Delta t}{\Delta x} \left( \mathbf{F}_{i-\frac{1}{2}} - \mathbf{F}_{i+\frac{1}{2}} \right) \quad (7.30)$$

with intercell numerical flux

$$\mathbf{F}_{i+\frac{1}{2}}^{\text{force}} = \frac{1}{2} \left[ \mathbf{F}_{i+\frac{1}{2}}^{n+\frac{1}{2}} + \frac{1}{2} (\mathbf{F}_i^n + \mathbf{F}_{i+1}^n) \right] + \frac{1}{4} \frac{\Delta x}{\Delta t} (\mathbf{U}_i^n - \mathbf{U}_{i+1}^n) . \quad (7.31)$$

A surprising outcome is that the intercell flux (7.31) is in fact the arithmetic mean of the fluxes for the Richtmyer and Lax–Friedrichs schemes, namely

$$\mathbf{F}_{i+\frac{1}{2}}^{\text{force}} = \frac{1}{2} \left( \mathbf{F}_{i+\frac{1}{2}}^{\text{RI}} + \mathbf{F}_{i+\frac{1}{2}}^{\text{LF}} \right) . \quad (7.32)$$

### 7.4.3 Analysis of the FORCE Scheme

For the linear advection equation (7.10) the conservative scheme (7.30), (7.32) yields

$$u_i^{n+1} = b_{-1} u_{i-1}^n + b_0 u_i^n + b_1 u_{i+1}^n , \quad (7.33)$$

with coefficients given as

$$b_{-1} = \frac{1}{4}(1+c)^2 , \quad b_0 = \frac{1}{2}(1-c^2) , \quad b_1 = \frac{1}{4}(1-c)^2 . \quad (7.34)$$

**Proposition 7.1.** *The scheme (7.33)–(7.34) is*

- *stable, with stability condition*

$$0 \leq |c| \leq 1 , \quad (7.35)$$

*where*

$$c = \frac{\Delta t a}{\Delta x} : \text{Courant Number} \quad (7.36)$$

- *monotone, and*
- *has modified equation*

$$q_t + a q_x = \alpha_{\text{fo}} q_{xx} , \quad \alpha_{\text{fo}} = \frac{1}{4} a \Delta x \left( \frac{1-c^2}{c} \right) = \frac{1}{2} \alpha_{\text{lf}} , \quad (7.37)$$

*where  $\alpha_{\text{lf}}$  is the coefficient of artificial viscosity for the Lax–Friedrichs scheme.*

*Proof.* The scheme can be shown to be stable with stability condition (7.35) by using standard von Neumann analysis, or more directly by utilising Billett's result, which says [58] that a scheme of the form (7.33)–(7.34) is stable if and only if

1.  $b_0(b_{-1} + b_1) \geq 0$
2.  $b_0(b_{-1} + b_1) + 4b_{-1}b_1 \equiv B \geq 0$ .

The first condition leads to (7.35) directly, while the second condition produces

$$B = \frac{1}{4} \left[ (1 - c^2)(1 + c^2) + (1 + c)^2(1 - c)^2 \right].$$

A sufficient condition for  $B \geq 0$  is again  $b_0 \geq 0$ , which confirms the sought stability restriction (7.35). Concerning monotonicity, by inspection, the scheme is monotone for Courant numbers satisfying the stability condition, i.e. all coefficients are non-negative, see Sect. 5.2.1 of Chap. 5. The result (7.37) is obtained by using standard analysis.

Numerical results of the FORCE scheme for the Euler equations are presented in Sect. 7.6. High-order extensions are presented in Chaps. 13 and 14; these high-order schemes also extend to multi-dimensional problems following the splitting techniques presented in Chap. 16.

## 7.5 Random Numbers

The quality of the computed RCM solution depends crucially on the random numbers  $\{\theta^n\}$ . Research in this area has produced some very effective guidelines. For example, it has been established that the more random the generation of  $\{\theta^n\}$ , the worse the computed RCM results. Colella [131], [132] introduced the use of pseudo-random numbers of the van der Corput type.

### 7.5.1 Van der Corput Pseudo-Random Numbers

A general van der Corput sequence  $\{\theta^n\}$  depends on two parameters  $k_1$ ,  $k_2$  with  $k_1 > k_2 > 0$  both integer and relatively prime. The  $(k_1, k_2)$  van der Corput sequence  $\{\theta^n\}$  is formally defined [233] as

$$\theta^n = \sum_{i=0}^m A_i k_1^{-(i+1)}, \quad (7.38)$$

$$A_i = k_2 a_i (\text{mod } k_1), \quad (7.39)$$

$$n = \sum_{i=0}^m a_i k_1^i. \quad (7.40)$$

To explain the definition of the pseudo-random number  $\theta^n$  in (7.38) we start from equation (7.40), which gives the non-negative integer  $n$  in scale of notation with radix  $k_1$ . If  $k_1 = 2$ , (7.40) gives the binary expansion of  $n$ . For example, the binary expansion of the integer 3 is

$$3 = 1 \times 2^0 + 1 \times 2^1$$

and  $m = 1$ .

The next stage is to find the coefficients  $A_i$  in (7.38) according to equation (7.39); this says that  $A_i$  is the remainder when dividing the product  $k_2 a_i$  by  $k_1$  ( $A_i < k_1$ ). The simplest case is given by  $k_2 = 1$ , for which  $A_i \equiv a_i$ ,  $\forall i$ . Having found the number  $m$  and the modified coefficients  $A_i$ , the random number  $\theta^n$  corresponding to the integer  $n$  is completely determined by the summation (7.38).

**Exercise 7.2 (Van der Corput sequences).** Find  $m$ , the coefficients  $a_i$  and  $A_i$  and the corresponding first 10 random numbers  $\theta^n$  of the  $(2, 1)$  and  $(5, 3)$  van der Corput sequences ( $n = 1, \dots, 10$ ).

**Solution 7.3.** Results are shown in Table 7.1 for the  $(2, 1)$  (binary) van der Corput sequence and in Table 7.2 for the  $(5, 3)$  van der Corput sequence.

$n$	$m$	$a_0$	$a_1$	$a_2$	$a_3$	$A_0$	$A_1$	$A_2$	$A_3$	$\theta^n$
1	0	1				1				0.5000
2	1	0	1			0	1			0.2500
3	1	1	1			1	1			0.7500
4	2	0	0	1		0	0	1		0.1250
5	2	1	0	1		1	0	1		0.6250
6	2	0	1	1		0	1	1		0.3750
7	2	1	1	1		1	1	1		0.8750
8	3	0	0	0	1	0	0	0	1	0.0625
9	3	1	0	0	1	1	0	0	1	0.5625
10	3	0	1	0	1	0	1	0	1	0.3125

Table 7.1: Number  $m$ , coefficients  $a_i$ ,  $A_i$  and random numbers  $\theta^n$  for the  $(2, 1)$  van der Corput sequence, for  $n = 1, \dots, 10$

### 7.5.2 Statistical Properties

A desirable statistical property of the sequence of numbers  $\{\theta^n\}$  is that  $\{\theta^n\}$  be uniformly distributed over  $[0, 1]$ . Following Olivier and Grönig [364] we study three statistical quantities that help to characterise the sequence of random numbers  $\{\theta^n\}$ . These are: the arithmetic mean, the standard deviation and the so called chi-square statistics.

$n$	$m$	$a_0$	$a_1$	$A_0$	$A_1$	$\theta^n$
1	0	1		3		0.6000
2	0	2		1		0.2000
3	0	3		4		0.8000
4	0	4		2		0.4000
5	1	0	1	0	3	0.1200
6	1	1	1	3	3	0.7200
7	1	2	1	1	3	0.3200
8	1	3	1	4	3	0.9200
9	1	4	1	2	3	0.5200
10	1	0	2	0	1	0.0400

Table 7.2: Number  $m$ , coefficients  $a_i$ ,  $A_i$  and random numbers  $\theta^n$  for the  $(5, 3)$  van der Corput sequence, for  $n = 1, \dots, 10$ .

The *arithmetic mean*  $x_{\text{ar}}$  of the set  $\{\theta^n\}_{n=1}^N$  is defined as

$$x_{\text{ar}} = \frac{1}{N} \sum_{n=1}^N \theta^n. \quad (7.41)$$

For an optimally equidistributed sequence  $\{\theta^n\}$  in  $[0, 1]$  we expect  $x_{\text{ar}}$  to be close to  $\frac{1}{2}$ .

The *standard deviation* is

$$x_{\text{sd}} = \left[ \frac{1}{(N-1)} \sum_{n=1}^N (\theta^n - x_{\text{ar}})^2 \right]^{\frac{1}{2}}. \quad (7.42)$$

The *chi-square statistic*  $x_{\text{sq}}$  is computed as follows: the interval  $[0, 1]$  is subdivided into  $D$  equally spaced subintervals  $R_i$ . Then we consider a total of  $N$  random numbers  $\theta^n$  and count the number  $c(R_i)$  of random numbers  $\theta^n$  that fall inside  $R_i$ , for  $i = 1, \dots, D$ . We denote by  $p_i = p_i(\theta^n)$  the probability that the number  $\theta^n$  falls inside the subinterval  $R_i$ . The expected number of random numbers in the interval  $R_i$  is thus  $Np_i$ . As an illustrative example we subdivide  $[0, 1]$  into  $D = 2$  subintervals and choose  $N = 4$  random numbers  $\{\theta^1, \theta^2, \theta^3, \theta^4\}$ . For a uniformly distributed sequence one would expect  $p_1(\theta^n) = \frac{1}{2}$  and so the expected number of random numbers  $\theta^n$  falling into subinterval  $R_1 \equiv [0, \frac{1}{2})$  is  $Np_1 = 4 \times \frac{1}{2} = 2$ . The same holds for subinterval  $R_2 \equiv [\frac{1}{2}, 1]$ . Then the *chi-square statistic* is defined as

$$x_{\text{sq}}^2 = \sum_{i=1}^N \frac{(c(R_i) - Np_i)^2}{Np_i}. \quad (7.43)$$

For a uniformly distributed sequence  $\{\theta^n\}$  we expect  $x_{\text{sq}}$  to be small. Better RCM numerical results are obtained with pseudo-random sequences  $\{\theta^n\}$  for which  $x_{\text{sq}}^2$  is small.

As in [364], we analyse the quantities  $x_{\text{ar}}$ ,  $x_{\text{sd}}$  and  $x_{\text{sq}}$  for 6 van der Corput sequences. Table 7.3 shows the statistical properties of  $N = 200$  random numbers. For the *chi-square test* the interval  $[0, 1]$  was subdivided into  $D = 20$  equally spaced subintervals.

$(k_1, k_2)$	(2,1)	(3,2)	(3,1)	(5,1)	(7,3)	(5,3)
$x_{\text{ar}}$	0.49459	0.49673	0.49457	0.49373	0.49965	0.49598
$x_{\text{sd}}$	0.28885	0.28780	0.28799	0.28914	0.28842	0.28907
$x_{\text{sq}}^2$	1.00000	1.00000	1.00000	1.40000	1.00000	0.20000

Table 7.3: Statistical properties of 6 van de Corput sequences  $(k_1, k_2)$

Our results agree well with those of Olivier and Grönig for the arithmetic mean  $x_{\text{ar}}$ ; for  $x_{\text{sq}}^2$  our results are similar but not identical. Our values for the standard deviation  $x_{\text{sd}}$  are much smaller than those quoted in [364]. Olivier and Grönig also analysed the quantities  $x_{\text{ar}}$ ,  $x_{\text{sd}}$  and  $x_{\text{sq}}$  for other sequences of random numbers, including the modified random numbers suggested by Chorin [110]. They observed that in all cases  $x_{\text{ar}}$  was close to  $\frac{1}{2}$ ; they also observed that all sequences tested had similar values for the standard deviation  $x_{\text{sd}}$ . The quantity that was different was  $x_{\text{sq}}$ , which led them to conclude that this was the statistical property of significance. Van der Corput sequences have a small value for  $x_{\text{sq}}$ , as compared with other sequences, and are known to produce very good computational results when used in the Random Choice Method. Anderson and Gottlieb [8] suggested a sequence of pseudo-random numbers of similar qualities to the van der Corput sequences. See [8] for details. Our tests for the Anderson and Gottlieb numbers give  $x_{\text{ar}} = 0.4981$ ,  $x_{\text{sd}} = 0.2891$  and  $x_{\text{sq}}^2 = 0.6$ . Compare with results of Table 7.3.

### 7.5.3 Propagation of a Single Shock

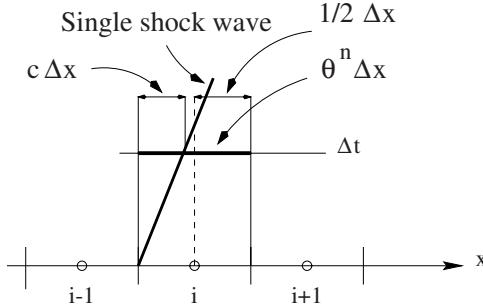
Here we consider a single shock wave of positive speed  $S$  connecting two constant states  $\mathbf{U}_L$  and  $\mathbf{U}_R$ . We solve the time-dependent Euler equations with initial data

$$\mathbf{U}(x, 0) = \begin{cases} \mathbf{U}_L & \text{if } x < 0, \\ \mathbf{U}_R & \text{if } x > 0. \end{cases} \quad (7.44)$$

The exact solution is

$$\mathbf{U}(x, t) = \begin{cases} \mathbf{U}_L & \text{if } x/t < S, \\ \mathbf{U}_R & \text{if } x/t > S. \end{cases} \quad (7.45)$$

On applying the Random Choice Method on a non-staggered grid with Courant number  $c \leq \frac{1}{2}$ , based on the shock speed  $S$ , the shock wave propagates



**Fig. 7.9.** Propagation of a single shock by the Random Choice Method, on a non-staggered grid. Shock propagates by comparing  $\theta^n$  with  $c$

to the right as a true discontinuity. The situation at any time level is illustrated in Fig. 7.9, where the shock wave is assumed located at the interface  $x_{i-\frac{1}{2}}$ . The shock wave crosses the line  $t = \Delta t$  at  $c\Delta x$  and by virtue of the CFL condition this point lies to the left of the middle of cell  $I_i$ , namely  $\frac{1}{2}\Delta x$ . A random position inside cell  $I_i$  is given by  $\theta^n\Delta x$ , with  $\theta^n$  a random number in the interval  $[0, 1]$ . If  $\theta^n\Delta x \leq c\Delta x$  then the randomly selected state is the post shock state  $\mathbf{U}_L$ , which is then assigned to the whole of the cell  $I_i$  for the next time level. This means that the shock moves to the right by a complete distance  $\Delta x$ . If  $\theta^n\Delta x > c\Delta x$  then the shock does not move at all. The position  $X_s$  of the shock after  $n$  time steps is given by

$$X_s = \sum_{i=1}^n \Delta x P_i , \quad (7.46)$$

where

$$P_i = \begin{cases} 1 , & \text{if } \theta^i \leq c , \\ 0 , & \text{if } \theta^i > c . \end{cases} \quad (7.47)$$

Table 7.4 shows calculations by hand of the shock position error, normalised by the mesh size  $\Delta x$ , of the RCM solution for 10 time steps. We use two Courant numbers for each of the van der Corput sequences (2, 1) and (5, 3). For instance, if the RCM is used with the (2, 1) van der Corput sequence and a CFL number  $c = \frac{1}{4}$ , then at the time step  $n = 5$  the error in the shock position is  $\frac{3}{4}\Delta x$ . In the first 10 time steps the largest error takes place at  $n = 8$  and is equal to  $\Delta x$ . For the (2, 1) van der Corput sequence and a CFL number  $c = \frac{1}{2}$  the maximum position error is also  $\Delta x$ . The results for the (5, 3) van der Corput sequence are more accurate than those for the (2, 1) sequence. For  $c = \frac{1}{4}$  the maximum error observed is also  $\frac{3}{4}\Delta x$  but there are two time levels at which the solution is exact. For the case  $c = \frac{1}{2}$  the maximum position error is  $\frac{1}{2}\Delta x$  and the solution is exact every other time step.

$n$	$(2, 1), c = \frac{1}{4}$	$(2, 1), c = \frac{1}{2}$	$(5, 3), c = \frac{1}{4}$	$(5, 3), c = \frac{1}{2}$
1	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{2}$
2	$\frac{1}{2}$	1	$\frac{1}{2}$	0
3	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{2}$
4	1	1	0	0
5	$\frac{3}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{1}{2}$
6	$\frac{1}{2}$	1	$\frac{1}{2}$	0
7	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{2}$
8	1	1	0	0
9	$\frac{3}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{2}$
10	$\frac{1}{2}$	1	$\frac{1}{2}$	0

Table 7.4: Position error of shock computed by the RCM. Two van der Corput sequences and two values of the Courant number  $c$  are used

**Exercise 7.4.** Verify the results of Table 4 and find the shock position error for the van der Corput sequences  $(3, 2)$  and  $(7, 3)$  for Courant numbers  $\frac{1}{4}$  and  $\frac{1}{3}$ .

**Solution 7.5.** (Left to the reader).

## 7.6 Numerical Results

In this section we show some numerical results for three methods, namely the conventional Random Choice Method on a non-staggered grid associated with the Godunov Method (denoted by RCMG), the Lax–Friedrichs type Random Choice Method (7.18) (denoted by RCMLF) and the First–Order Centred (FORCE) scheme (7.30), (7.32). We solve five test problems with exact solution for the one-dimensional time dependent Euler equations. The initial data consists of two constant states  $\mathbf{W}_L = [\rho_L, u_L, p_L]^T$  and  $\mathbf{W}_R = [\rho_R, u_R, p_R]^T$ . These are given in Table 7.5. For a discussion on the exact solution of these test problems see Sect. 6.4, Chapt. 6. Codes of the library *NUMERICA* [519] were used to obtain the displayed results. The exact solutions were found by running the code *HE-E1RPEXACT*, the numerical solutions using RCMG were obtained by running the code *HE-E1RCM* and the numerical solutions using FORCE were obtained by running the code *HE-E1FOCENT*.

Test	$\rho_L$	$u_L$	$p_L$	$\rho_R$	$u_R$	$p_R$
1	1.0	0.75	1.0	0.125	0.0	0.1
2	1.0	-2.0	0.4	1.0	2.0	0.4
3	1.0	0.0	1000.0	1.0	0.0	0.01
4	5.99924	19.5975	460.894	5.99242	-6.19633	46.0950
5	1.0	-19.59745	1000.0	1.0	-19.59745	0.01

Table 7.5. Data for five test problems with exact solution.

We discretise the domain  $[0, 1]$  into  $M = 100$  computing cells. For the random choice methods we use a Courant number coefficient  $C_{\text{cfl}} = 0.45$  and for the FORCE method we use  $C_{\text{cfl}} = 0.9$ .

Figs. 7.10 to 7.13 show results for RCMG applied to Tests 1 to 4. The corresponding results for the RCMLF are shown in Figs. 7.14 to 7.17 and those for the FORCE scheme are shown in Figs. 7.18 to 7.22. All these results are to be compared with those of Chap. 6.

The RCMG results of Fig. 7.10 are, by any standards, very accurate. Jump discontinuities such as shock waves and contacts are resolved as true discontinuities. Also, discontinuities in derivative, such as those along the head and tail of rarefaction waves, are also very well resolved. The complexity of the RCM is comparable to that of the Godunov method, and thus it is fair to compare Fig. 7.10 with Fig. 6.8 of Chap. 6. The entropy glitch inside the sonic rarefaction for the RCM result is smaller than that for the Godunov method. The RCM results for Test 2 are shown in Fig. 7.11; this problem does not contain jump discontinuities and exposes a weakness of the RCM, namely the random noise in smooth parts of the flow. The results of Figs. 7.12 and 7.13 exhibit the true merits of the RCM for computing solutions containing multiple jump discontinuities; compare with the Godunov results of Figs. 6.10 and 6.11.

The computational results for the Lax–Friedrichs type random choice method (RCMLF) are shown in Figs. 7.14 to 7.17. These are somewhat inferior to those of the conventional RCM. However RCMLF has the advantage of having twice the stability range of the conventional RCM. Notice the pairing of neighbouring states, which is also a feature of the conventional Lax–Friedrichs scheme; compare Fig. 7.16 with Fig. 6.12.

Figs. 7.18 to 7.22 show the results for the First–Order Centred scheme FORCE as applied to Tests 1 to 5. Compare with the corresponding results for the Godunov and Lax–Friedrichs methods in Chap. 6. In general, the results for the FORCE scheme look inferior to those of the Godunov method, particularly for Test 5; the former scheme is however significantly simpler and more efficient than the latter. The numerical results of the FORCE scheme are superior to those of the Lax–Friedrichs scheme (see results of Chap. 6), with both schemes being comparable in complexity and efficiency.

## 7.7 Concluding Remarks

We have presented random choice and related methods to solve time-dependent one dimensional hyperbolic conservation laws. Details have been given for the Euler equations, for which numerical results have been presented. It is well known that the conventional RCM, on staggered and non-staggered grids, is only directly applicable to hyperbolic systems in *two independent variables*. For the time-dependent one dimensional Euler equations these are

time  $t$  and space  $x$ . For the steady supersonic Euler equations the two independent variables are the flow direction  $x$ , which is a time-like variable, and space  $y$ .

The great merit of the RCM is its ability to resolve discontinuities with infinite resolution, as true discontinuities. The main disadvantage of the RCM is the randomness in the smooth parts of the flow. One way of eliminating this unwanted randomness is by resorting to hybrid approaches, whereby the RCM is used at discontinuities only; elsewhere in the flow one may use some other scheme, see Toro and Roe [539], [540] for instance. A crucial question is this: *can the RCM be extended to in-homogeneous (sources) systems or to systems with more than two independent variables?*

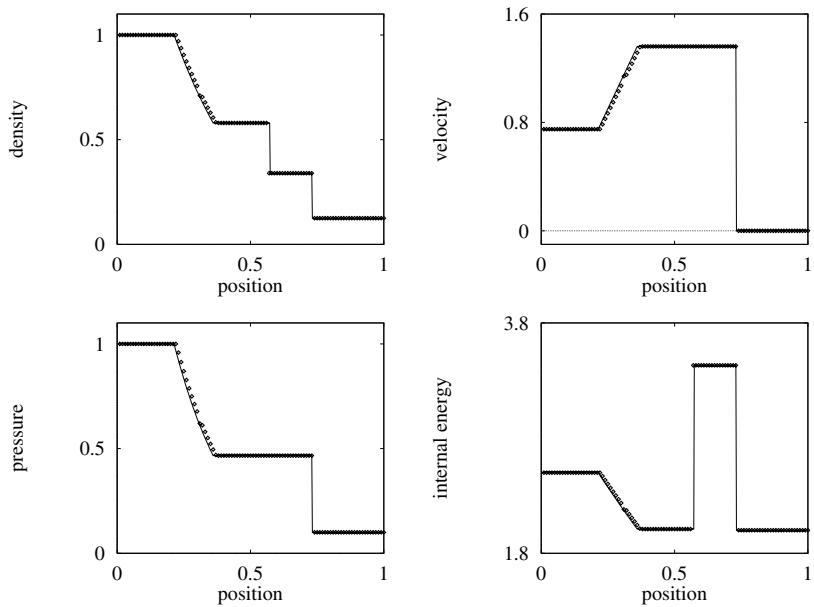
By applying splitting schemes, see Chap. 15, random choice methods can be extended to solve in-homogeneous systems

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{S}(\mathbf{U}) . \quad (7.48)$$

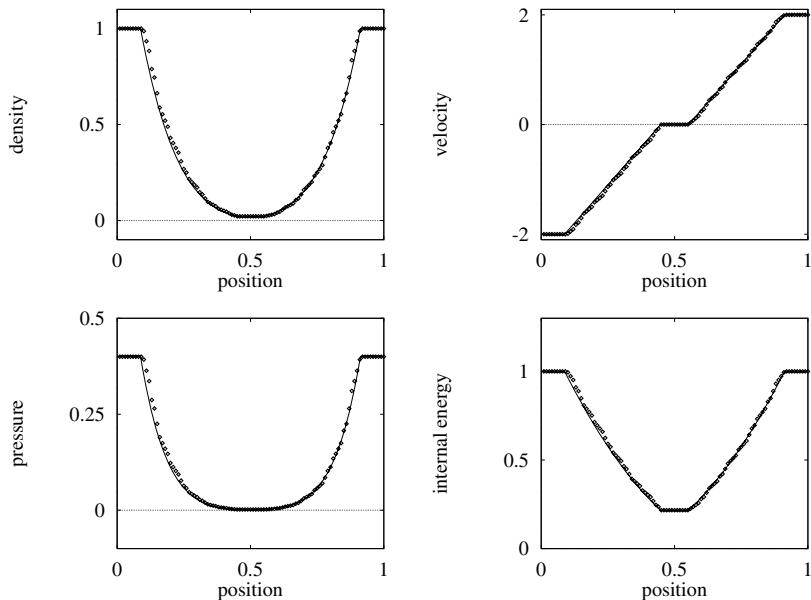
The *source term vector*  $\mathbf{S}(\mathbf{U})$  may be an algebraic function of the flow variables, such as in cylindrically or spherically symmetric flow [453].  $\mathbf{S}(\mathbf{U})$  may also involve higher order spatial derivatives such as in viscous terms; this means that the RCM can be applied to parabolic equations, e.g. the time-dependent, one dimensional Navier-Stokes equations [256]. The inclusion of algebraic source terms retains the infinite resolution of discontinuities but may enhance the randomness in smooth parts of the flow, present in the homogeneous part of the problem. In some special cases, see Glimm et. al. [214], the source term vector  $\mathbf{S}(\mathbf{U})$  may be incorporated into the solution of the Riemann problem. This significantly alleviates the problem of enhanced randomness.

Efforts to extend the RCM, retaining its distinctive feature, to solve problems in more than two independent variables, such as the time-dependent, two dimensional Euler equations and the three dimensional steady supersonic Euler equations, have so far proved unsuccessful. For details on how the use of splitting techniques to carry out the extensions fail, were reported by Colella [131]. If resolving shocks with infinite resolution is abandoned then splitting techniques applied to the RCM work reasonably well. Colella [131] introduced artificial viscosity, as in conventional finite difference methods. Toro [500] proposed a hybrid approach to extend the RCM to solve the time-dependent, two dimensional Euler equations (three independent variables); here the RCM is used at contacts and shear waves and a high-resolution shock capturing method is used elsewhere. This approach gives up the infinite resolution of shocks but retains infinite resolution of shear waves and material interfaces, features that are exceedingly difficult to resolve by most methods. Recently, Loh et.al. [332] have reported their work to extend the RCM to solve the three-dimensional steady supersonic Euler equations (three independent variables). They also give up infinite resolution of shocks but retain infinite resolution of slip surfaces by virtue of a Lagrangian approach.

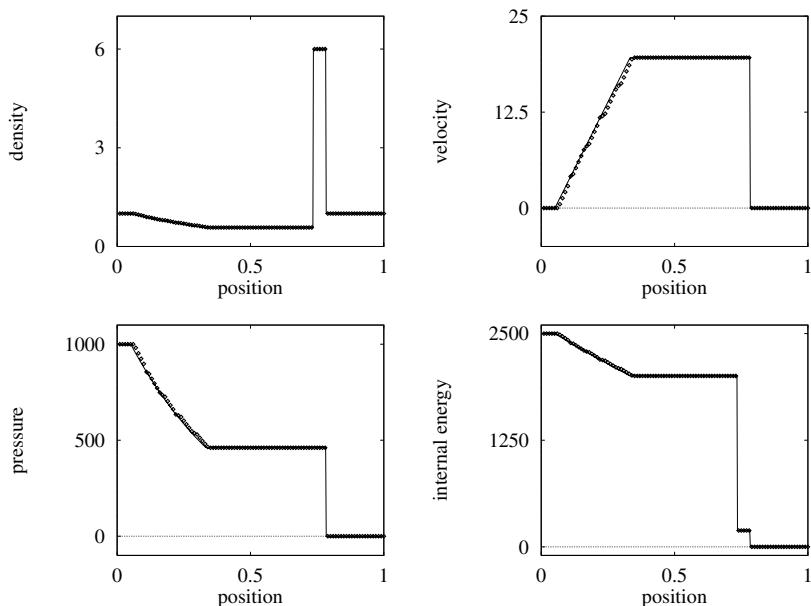
At the present time the RCM offers an accurate numerical method for solving non-linear systems of the form (7.48) in conjunction with operator splitting techniques, as presented in Chap. 15.



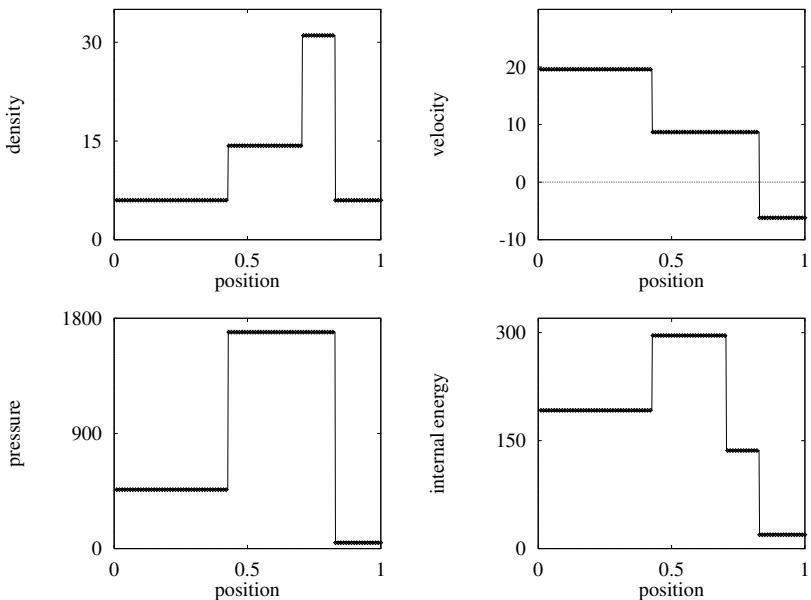
**Fig. 7.10.** Random Choice Method applied to Test 1, with  $x_0 = 0.3$ . Numerical (symbol) and exact (line) solutions are compared at the output time 0.2 units



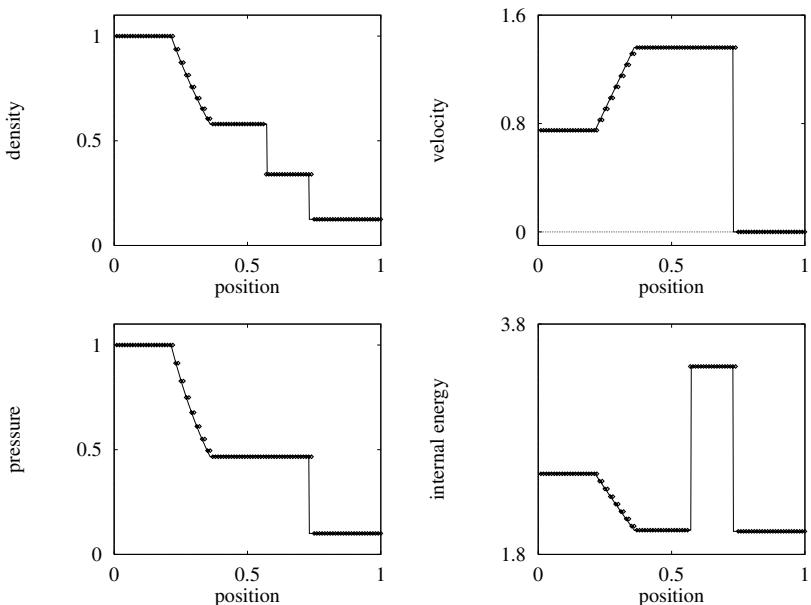
**Fig. 7.11.** Random Choice Method applied to Test 2, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at the output time 0.15 units



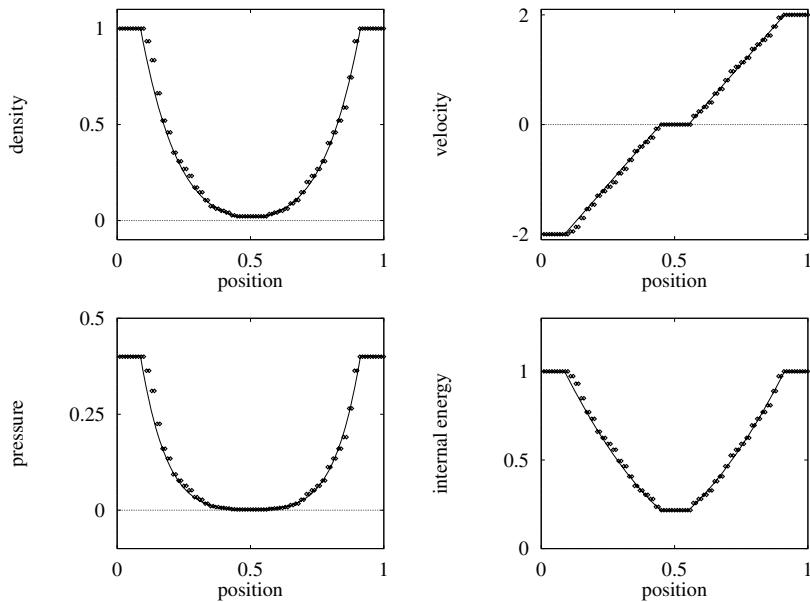
**Fig. 7.12.** Random Choice Method applied to Test 3, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at the output time 0.012 units



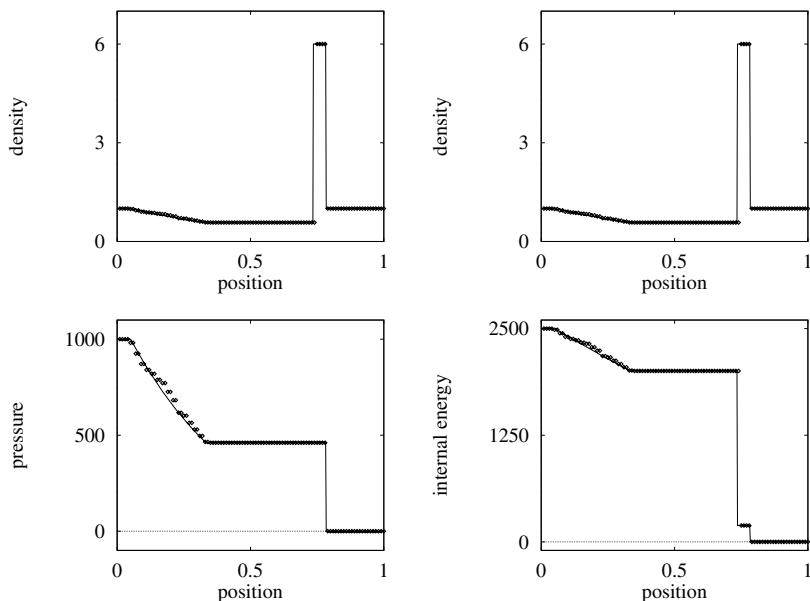
**Fig. 7.13.** Random Choice Method applied to Test 4, with  $x_0 = 0.4$ . Numerical (symbol) and exact (line) solutions are compared at the output time 0.035 units



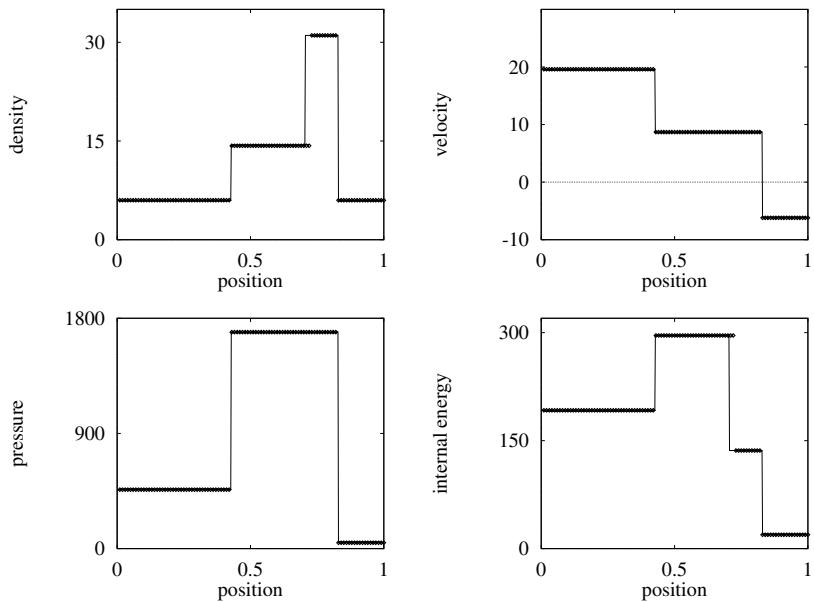
**Fig. 7.14.** Lax-Friedrichs Random Choice scheme applied to Test 1, with  $x_0 = 0.3$ . Numerical (symbol) and exact (line) solutions are compared at time 0.2 units



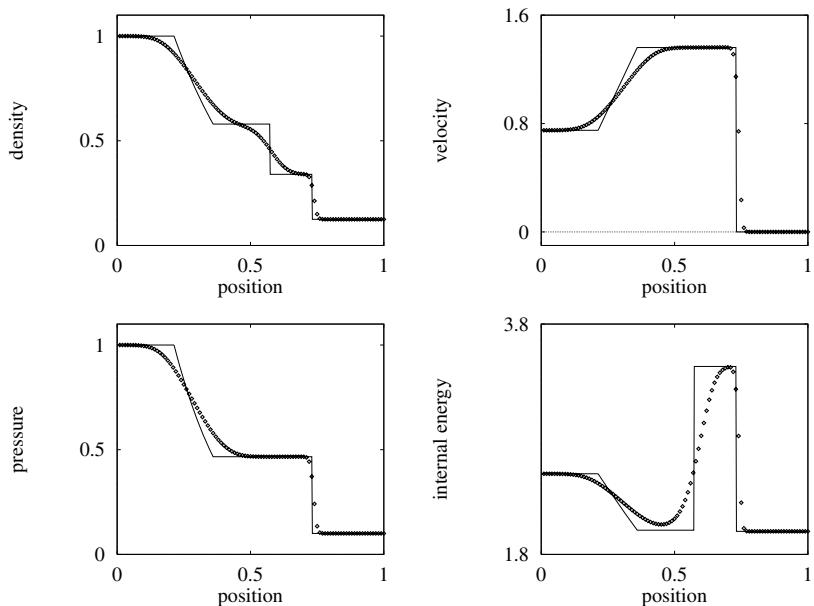
**Fig. 7.15.** Lax–Friedrichs Random Choice scheme applied to Test 2, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at time 0.15 units



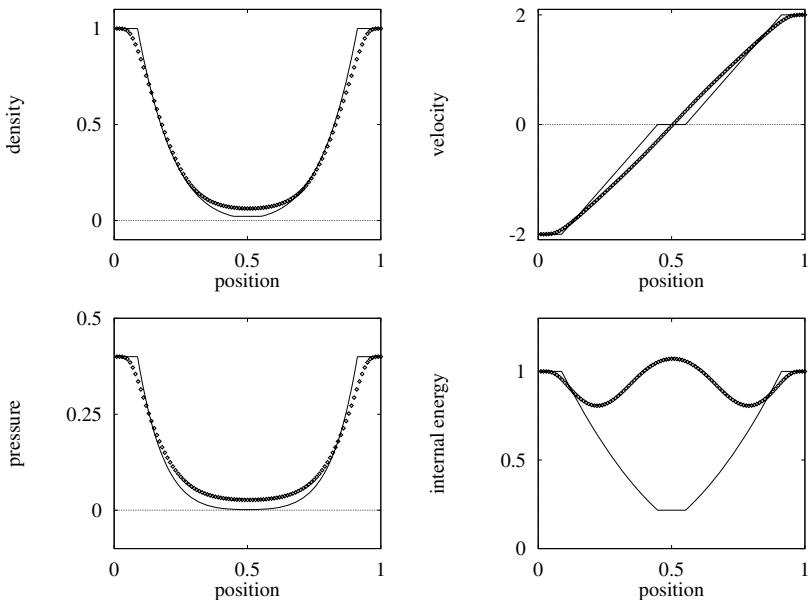
**Fig. 7.16.** Lax–Friedrichs Random Choice scheme applied to Test 3, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012 units



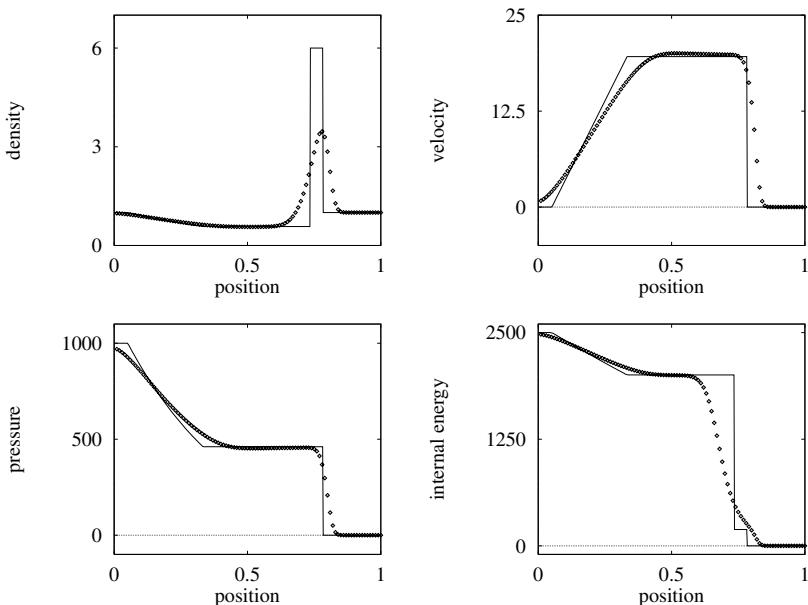
**Fig. 7.17.** Lax-Friedrichs Random Choice scheme applied to Test 4, with  $x_0 = 0.4$ . Numerical (symbol) and exact (line) solutions are compared at time 0.035 units



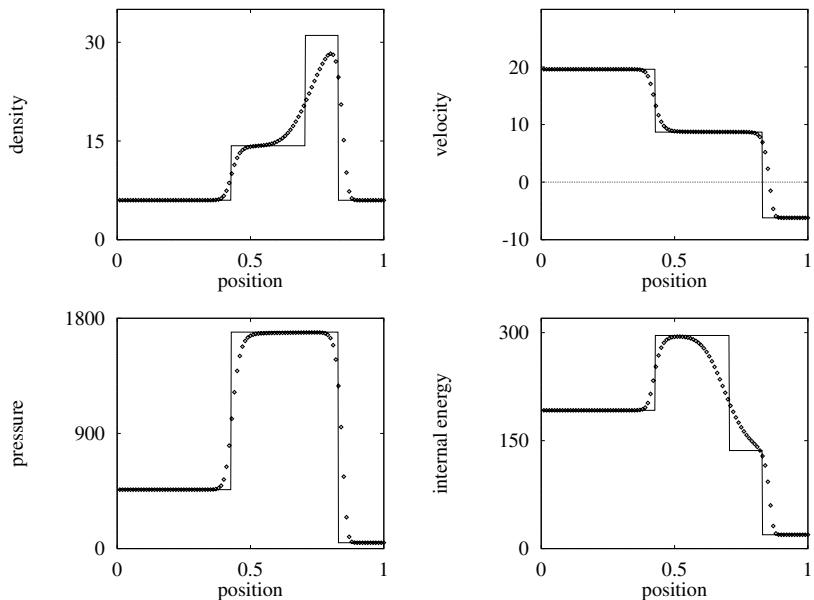
**Fig. 7.18.** FORCE scheme applied to Test 1, with  $x_0 = 0.3$ . Numerical (symbol) and exact (line) solutions are compared at the output time 0.2 units



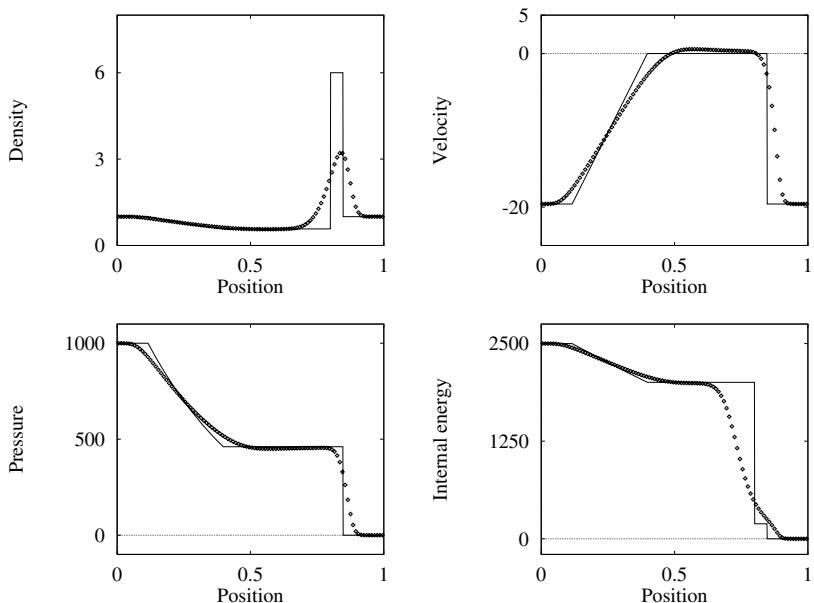
**Fig. 7.19.** FORCE scheme applied to Test 2, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at the output time 0.15 units



**Fig. 7.20.** FORCE scheme applied to Test 3, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at the output time 0.012 units



**Fig. 7.21.** FORCE scheme applied to Test 4, with  $x_0 = 0.4$ . Numerical (symbol) and exact (line) solutions are compared at the output time 0.035 units



**Fig. 7.22.** FORCE scheme applied to Test 5, with  $x_0 = 0.8$ . Numerical (symbol) and exact (line) solutions are compared at the output time 0.012 units

---

## Flux Vector Splitting Methods

### 8.1 Introduction

A distinguishing feature of upwind numerical methods is this: *the discretisation of the equations on a mesh is performed according to the direction of propagation of information on that mesh*. In this way, salient features of the physical phenomena modelled by the equations are incorporated into the discretisation schemes. There are essentially two approaches for identifying upwind directions, namely the *Godunov approach* [216] studied in Chap. 6, and the *Flux Vector Splitting* (FVS) approach [424], [463], [560], [561] to be studied in this chapter. These two approaches are often referred to as the *Riemann approach* and the *Boltzmann approach* [244]. The respective numerical methods derived from these two approaches are often referred to as *Flux Difference Splitting Methods* and *Flux Vector Splitting Methods*. For a review on both of these approaches the paper by Harten, Lax and van Leer [244] is highly recommended. Closely related schemes to FVS, not studied here, are the KFVS or kinetic schemes, see for example Pullin [389], Perthame [380], [381], Mandal and Desphande [336], Xu and Prendergast [587], Xu et. al. [586], Xu [585] and Yang *et. al* [591].

The identification of upwind directions in Flux Vector Splitting Methods is achieved with less effort than in Godunov-type methods, leading to simpler and somewhat more efficient schemes. These two features are very attractive and have made FVS schemes very popular within a large community of practitioners. The Flux Vector Splitting approach is particularly well suited for *implicit methods*; these are popular in Aerodynamics, where the computation of steady solutions is of great practical value. The reduced sophistication of FVS schemes however, as compared with Godunov-type schemes, results in poorer resolution of discontinuities, particularly stationary contact and shear waves. In applications to the Navier–Stokes equations, it is reported by van Leer, Thomas and Roe [565] that their FVS scheme is considerably less accurate than Godunov’s method with Roe’s approximate Riemann solver [407]. A key feature of the FVS approach is its reliance on a special property of the

equations, namely the *homogeneity property*. As seen in Sect. 3.1.1 of Chap. 3, the Euler equations satisfy this property but there are important examples, such as the shallow water equations, that do not. The homogeneity property may however be circumvented so as to be able to apply the FVS approach, see Vázquez–Cendón [568].

The pioneering works of Sanders and Prendergast [424], Steger and Warming [463] and van Leer [560], [561] has been followed by numerous applications as well as by increased research efforts to improve further the technique. See for example the papers [12], [13], [166], [328], [578] and [387], amongst many others.

The purpose of this chapter is to give an elementary introduction to Flux Vector Splitting methods. Sects. 8.2 and 8.3 are devoted to a simple introduction to the FVS approach. In Sect. 8.4 we derive FVS methods for the time-dependent Euler equations following the methodologies of Steger and Warming [463], that of van Leer [560], [561] and the recently proposed approach of Liou and Steffen [328]. Numerical results are presented in Sect. 8.5. Techniques to construct high-order schemes based on FVS are found in Chaps. 13 and 14. In Chap. 15 we show how to solve systems with source terms and in Chap. 16 we deal with approaches to construct multidimensional schemes. Essential background material for reading this chapter is found in Chaps. 2, 3 and 5.

## 8.2 The Flux Vector Splitting Approach

In this section we introduce the flux vector splitting approach in the simple setting of model hyperbolic systems, namely the small perturbation steady supersonic equations and the isothermal equations of Gas Dynamics; see Sect. 1.6.2 of Chap. 1 and Sects. 2.1 and 2.4.1 of Chap. 2 for details on these systems.

### 8.2.1 Upwind Differencing

Consider the small perturbation steady supersonic equations

$$u_x - a^2 v_y = 0 , \quad v_x - u_y = 0 , \quad (8.1)$$

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

$$a = \sqrt{\frac{1}{M_\infty^2 - 1}} \quad (8.2)$$

is the sound speed and  $M_\infty$  is the free-stream Mach number, assumed to be greater than unity. Equations (8.1) may be rewritten as

$$\mathbf{U}_x + \mathbf{A} \mathbf{U}_y = \mathbf{0} , \quad (8.3)$$

with

$$\mathbf{U} = \begin{bmatrix} u \\ v \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & -a^2 \\ -1 & 0 \end{bmatrix}. \quad (8.4)$$

The eigenvalues of the coefficient matrix  $\mathbf{A}$  are

$$\lambda_1 = -a, \quad \lambda_2 = +a, \quad (8.5)$$

with corresponding right eigenvectors

$$\mathbf{K}^{(1)} = \begin{bmatrix} a \\ 1 \end{bmatrix}, \quad \mathbf{K}^{(2)} = \begin{bmatrix} a \\ -1 \end{bmatrix}. \quad (8.6)$$

Given the mixed character of the eigenvalues ( $\lambda_1 = -a$  is negative and  $\lambda_2 = +a$  is positive), a finite difference discretisation of (8.3) has limited choices for the spatial derivative, if upwind bias is to be applied. Consider, for instance, the *one-sided difference schemes*

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \frac{\Delta x}{\Delta y} \mathbf{A} [\mathbf{U}_i^n - \mathbf{U}_{i-1}^n], \quad (8.7)$$

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \frac{\Delta x}{\Delta y} \mathbf{A} [\mathbf{U}_{i+1}^n - \mathbf{U}_i^n]. \quad (8.8)$$

Clearly scheme (8.7) is upwind relative to the eigenvalue  $\lambda_2 = a > 0$  but is downwind, and thus unstable, relative to the eigenvalue  $\lambda_1 = -a < 0$ . A similar observation applies to scheme (8.8). For the case in which all eigenvalues have the same sign the difficulty of choosing the upwind direction does not arise.

As seen in Sect. 5.4 of Chap. 5, general linear hyperbolic systems with constant coefficients may be solved by the CIR first-order upwind method by decomposing the coefficient matrix  $\mathbf{A}$  into a positive component  $\mathbf{A}^+$  and a negative component  $\mathbf{A}^-$ , such that

$$\mathbf{A} = \mathbf{A}^+ + \mathbf{A}^-, \quad (8.9)$$

where  $\mathbf{A}^+$  has positive or zero eigenvalues and  $\mathbf{A}^-$  has negative or zero eigenvalues. One then has the upwind scheme

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \frac{\Delta x}{\Delta y} \mathbf{A}^+ [\mathbf{U}_i^n - \mathbf{U}_{i-1}^n] - \frac{\Delta x}{\Delta y} \mathbf{A}^- [\mathbf{U}_{i+1}^n - \mathbf{U}_i^n]. \quad (8.10)$$

The *Split-Coefficient Matrix Scheme* of Chakravarthy et. al. [97], [251] is an extension of this procedure to non-linear systems, in non-conservative form.

The CIR upwind scheme, when applied to general linear hyperbolic systems with constant coefficients, may be written in conservative form by defining the flux vector

$$\mathbf{F} = \mathbf{A} \mathbf{U}. \quad (8.11)$$

Then the splitting (8.9) of the coefficient matrix  $\mathbf{A}$  results in a natural splitting of the flux vector  $\mathbf{F}$ , namely

$$\mathbf{F} = \mathbf{F}^+ + \mathbf{F}^- . \quad (8.12)$$

In this way the CIR upwind scheme can be written in conservative form

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \frac{\Delta t}{\Delta x} [\mathbf{F}_{i+\frac{1}{2}} - \mathbf{F}_{i-\frac{1}{2}}] , \quad (8.13)$$

where the intercell numerical flux

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}_i^+(\mathbf{U}_i^n) + \mathbf{F}_i^-(\mathbf{U}_{i+1}^n) \quad (8.14)$$

is identical to the Godunov intercell flux. See Sect. 5.4 of Chap. 5 for details. The Flux Vector Splitting Method is a generalisation of this to non-linear systems in conservation form.

### 8.2.2 The FVS Approach

Here we consider a general system of  $m$  non-linear hyperbolic conservation laws

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0} . \quad (8.15)$$

From the assumption of hyperbolicity the Jacobian matrix

$$\mathbf{A}(\mathbf{U}) = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} \quad (8.16)$$

may be expressed as

$$\mathbf{A} = \mathbf{K} \boldsymbol{\Lambda} \mathbf{K}^{-1} , \quad (8.17)$$

where  $\boldsymbol{\Lambda}$  is the diagonal matrix formed by the eigenvalues of  $\mathbf{A}$ , namely

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{bmatrix} . \quad (8.18)$$

The matrix  $\mathbf{K}$  is

$$\mathbf{K} = [\mathbf{K}^{(1)}, \mathbf{K}^{(2)}, \dots, \mathbf{K}^{(m)}] , \quad (8.19)$$

where the column  $\mathbf{K}^{(i)}$  is the right eigenvector of  $\mathbf{A}$  corresponding to  $\lambda_i$  and  $\mathbf{K}^{-1}$  is the inverse of  $\mathbf{K}$ . Recall our usual convention of ordering the eigenvalues in *increasing* order.

As anticipated in the previous section, the Flux Vector Splitting method aims at generalising (8.14) to non-linear systems (8.15). That is, FVS requires a splitting of the flux vector  $\mathbf{F}$  into two component  $\mathbf{F}^+$  and  $\mathbf{F}^-$  such that

$$\mathbf{F}(\mathbf{U}) = \mathbf{F}^+(\mathbf{U}) + \mathbf{F}^-(\mathbf{U}) , \quad (8.20)$$

under the restriction that the eigenvalues  $\hat{\lambda}_i^+$  and  $\hat{\lambda}_i^-$  of the Jacobian matrices

$$\hat{\mathbf{A}}^+ = \frac{\partial \mathbf{F}^+}{\partial \mathbf{U}}, \quad \hat{\mathbf{A}}^- = \frac{\partial \mathbf{F}^-}{\partial \mathbf{U}} \quad (8.21)$$

satisfy the condition

$$\hat{\lambda}_i^+ \geq 0, \quad \hat{\lambda}_i^- \leq 0. \quad (8.22)$$

The splitting is also required to reproduce *regular upwinding* when all eigenvalues  $\lambda_i$  of the coefficient matrix  $\mathbf{A}$  are one-sided, that is, all positive or zero, or all negative or zero. That is to say

$$\left. \begin{array}{l} \mathbf{F}^+ = \mathbf{F}, \quad \mathbf{F}^- = \mathbf{0} \quad \text{if } \lambda_i \geq 0 \quad \text{for } i = 1, \dots, m, \\ \mathbf{F}^+ = \mathbf{0}, \quad \mathbf{F}^- = \mathbf{F} \quad \text{if } \lambda_i \leq 0 \quad \text{for } i = 1, \dots, m. \end{array} \right\} \quad (8.23)$$

If in addition to hyperbolicity, the system (8.15) satisfies the *homogeneity property*

$$\mathbf{F}(\mathbf{U}) = \mathbf{A}(\mathbf{U})\mathbf{U}, \quad (8.24)$$

just as in the linear constant coefficient case, then the sought splitting is easily accomplished by identifying a suitable splitting of the Jacobian matrix  $\mathbf{A}$ . As seen in Sect. 3.1.1 of Chap. 3, the time-dependent Euler equations satisfy the *homogeneity property*.

From the diagonalisation of  $\mathbf{A}$  given by (8.17), a splitting of  $\mathbf{A}$  may be accomplished by an appropriate splitting of the diagonal matrix  $\mathbf{\Lambda}$ . This in turn, may be split by identifying a splitting of the eigenvalues  $\lambda_i$ ,  $i = 1, \dots, m$  of  $\mathbf{A}$ . Suppose we may split the eigenvalues  $\lambda_i$  as

$$\lambda_i = \lambda_i^+ + \lambda_i^-, \quad (8.25)$$

such that  $\lambda_i^+ \geq 0$  and  $\lambda_i^- \leq 0$ . Then  $\mathbf{\Lambda}$  may be split as

$$\mathbf{\Lambda} = \mathbf{\Lambda}^+ + \mathbf{\Lambda}^-, \quad (8.26)$$

where

$$\mathbf{\Lambda}^+ = \begin{bmatrix} \lambda_1^+ & 0 \\ \ddots & \ddots \\ 0 & \lambda_m^+ \end{bmatrix}, \quad \mathbf{\Lambda}^- = \begin{bmatrix} \lambda_1^- & 0 \\ \ddots & \ddots \\ 0 & \lambda_m^- \end{bmatrix}. \quad (8.27)$$

A natural splitting of  $\mathbf{A}$  results, namely

$$\mathbf{A} = \mathbf{A}^+ + \mathbf{A}^-, \quad (8.28)$$

with

$$\mathbf{A}^+ = \mathbf{K}\mathbf{\Lambda}^+\mathbf{K}^{-1}, \quad \mathbf{A}^- = \mathbf{K}\mathbf{\Lambda}^-\mathbf{K}^{-1}. \quad (8.29)$$

Then, if (8.24) is satisfied, we can split  $\mathbf{F}(\mathbf{U})$  as

$$\mathbf{F} = \mathbf{F}^+ + \mathbf{F}^-, \quad (8.30)$$

where

$$\mathbf{F}^+ = \mathbf{A}^+ \mathbf{U}, \quad \mathbf{F}^- = \mathbf{A}^- \mathbf{U}. \quad (8.31)$$

Steger and Warming [463] proposed a splitting of the eigenvalues  $\lambda_i$  as in (8.25) with definitions

$$\lambda_i^+ = \frac{1}{2}(\lambda_i + |\lambda_i|), \quad \lambda_i^- = \frac{1}{2}(\lambda_i - |\lambda_i|), \quad (8.32)$$

where  $|\lambda_i|$  is the absolute value of  $\lambda_i$  namely,

$$|\lambda_i| = \begin{cases} \lambda_i & \text{if } \lambda_i \geq 0, \\ -\lambda_i & \text{if } \lambda_i \leq 0. \end{cases} \quad (8.33)$$

Clearly

$$\lambda_i^+ \geq 0, \quad \lambda_i^- \leq 0, \quad \text{for } i = 1, \dots, m. \quad (8.34)$$

**Exercise 8.1.** Verify that the following properties are satisfied

$$\left. \begin{aligned} \lambda_i &= \lambda_i^+ + \lambda_i^- & ; \quad |\lambda_i| &= \lambda_i^+ - \lambda_i^-, \\ \mathbf{\Lambda} &= \mathbf{\Lambda}^+ + \mathbf{\Lambda}^- & ; \quad |\mathbf{\Lambda}| &= \mathbf{\Lambda}^+ - \mathbf{\Lambda}^-, \\ \mathbf{A} &= \mathbf{A}^+ + \mathbf{A}^- & ; \quad |\mathbf{A}| &= \mathbf{A}^+ - \mathbf{A}^-. \end{aligned} \right\} \quad (8.35)$$

**Solution 8.2.** (Left to the reader).

### 8.3 FVS for the Isothermal Equations

In order to illustrate the FVS approach we consider the isothermal equations of Gas Dynamics

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0}, \quad (8.36)$$

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \end{bmatrix}, \quad \mathbf{F}(\mathbf{U}) = \begin{bmatrix} \rho u \\ \rho u^2 + \rho a^2 \end{bmatrix}, \quad (8.37)$$

where the sound speed  $a$  is a positive constant. For details on the eigenstructure of this system see Sect. 2.4 of Chap. 2. The Jacobian matrix is

$$\mathbf{A} = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} = \begin{bmatrix} 0 & 1 \\ a^2 - u^2 & 2u \end{bmatrix}. \quad (8.38)$$

The eigenvalues of  $\mathbf{A}$  are

$$\lambda_1 = u - a, \quad \lambda_2 = u + a \quad (8.39)$$

and the matrix  $\mathbf{K}$  of corresponding right eigenvectors is

$$\mathbf{K} = \begin{bmatrix} 1 & 1 \\ u - a & u + a \end{bmatrix}. \quad (8.40)$$

**Exercise 8.3.** Verify that system (8.36)–(8.37) satisfy the *homogeneity property* (8.24).

**Solution 8.4.** (Left to the reader).

### 8.3.1 Split Fluxes

Given any splitting (8.25) with

$$\mathbf{\Lambda}^+ = \begin{bmatrix} \lambda_1^+ & 0 \\ 0 & \lambda_2^+ \end{bmatrix}, \quad \mathbf{\Lambda}^- = \begin{bmatrix} \lambda_1^- & 0 \\ 0 & \lambda_2^- \end{bmatrix}, \quad (8.41)$$

we require the computation of the matrices  $\mathbf{A}^+$  and  $\mathbf{A}^-$  as given by (8.29). One then requires the determination of the inverse  $\mathbf{K}^{-1}$  of the matrix  $\mathbf{K}$ , the products of three matrices as in (8.29) and finally the products (8.31) to find the flux components. For large systems this may be a rather tedious algebraic task. For the isothermal equations we have

$$\mathbf{K}^{-1} = \frac{1}{2a} \begin{bmatrix} u + a & -1 \\ a - u & 1 \end{bmatrix}. \quad (8.42)$$

Now, given any of the two components (8.27) of  $\mathbf{\Lambda}$ ,  $\mathbf{A}^\alpha$ , say, we compute

$$\mathbf{A}^\alpha = \mathbf{K} \mathbf{\Lambda}^\alpha \mathbf{K}^{-1}.$$

The result is

$$\mathbf{A}^\alpha = \frac{1}{2a} \begin{bmatrix} \lambda_1^\alpha(u+a) - \lambda_2^\alpha(u-a) & \lambda_2^\alpha - \lambda_1^\alpha \\ (u^2 - a^2)(\lambda_1^\alpha - \lambda_2^\alpha) & \lambda_2^\alpha(u+a) - \lambda_1^\alpha(u-a) \end{bmatrix}. \quad (8.43)$$

Application of (8.31) gives the flux vector component

$$\mathbf{F}^\alpha = \mathbf{A}^\alpha \mathbf{U},$$

that is

$$\mathbf{F}^\alpha = \frac{\rho}{2} \begin{bmatrix} \lambda_1^\alpha + \lambda_2^\alpha \\ \lambda_1^\alpha(u-a) + \lambda_2^\alpha(u+a) \end{bmatrix}. \quad (8.44)$$

Note that the expression for the component  $\mathbf{F}^\alpha$  given by (8.44) is general. For  $\alpha = +$  and  $\alpha = -$  the flux components  $\mathbf{F}^+$  and  $\mathbf{F}^-$  are

$$\mathbf{F}^+ = \frac{\rho}{2} \begin{bmatrix} \lambda_1^+ + \lambda_2^+ \\ \lambda_1^+(u-a) + \lambda_2^+(u+a) \end{bmatrix}, \quad (8.45)$$

and

$$\mathbf{F}^- = \frac{\rho}{2} \begin{bmatrix} \lambda_1^- + \lambda_2^- \\ \lambda_1^-(u-a) + \lambda_2^-(u+a) \end{bmatrix}. \quad (8.46)$$

**Exercise 8.5.** For the split fluxes (8.45)–(8.46), for the case of subsonic flow,

- (i) Find the Jacobian matrices

$$\hat{\mathbf{A}}^+ = \frac{\partial \mathbf{F}^+}{\partial \mathbf{U}}, \quad \hat{\mathbf{A}}^- = \frac{\partial \mathbf{F}^-}{\partial \mathbf{U}}.$$

- (ii) Find the eigenvalues  $\hat{\lambda}_i^+$  and  $\hat{\lambda}_i^-$ .

**Solution 8.6.** For the positive flux component  $\mathbf{F}^+$  the Jacobian matrix is

$$\hat{\mathbf{A}}^+ = \frac{\partial \mathbf{F}^+}{\partial \mathbf{U}} = \begin{bmatrix} \frac{1}{2}a & \frac{1}{2} \\ \frac{1}{2}(a^2 - u^2) & u + a \end{bmatrix}.$$

The eigenvalues are the roots of the characteristic polynomial

$$\lambda^2 - \left(\frac{3}{2}a + u\right)\lambda + \frac{1}{4}(u + a)^2 = 0,$$

namely,

$$\hat{\lambda}_1^+ = \frac{1}{4}a \left[ 2M + 3 - \sqrt{4M + 5} \right], \quad \hat{\lambda}_2^+ = \frac{1}{4}a \left[ 2M + 3 + \sqrt{4M + 5} \right].$$

*Remark 8.7.* Note that

$$\hat{\mathbf{A}}^+ \neq \mathbf{A}^+$$

and that

$$\hat{\lambda}_i^+ \neq \lambda_i^+.$$

Note also that  $\hat{\lambda}_i^+ > 0$ , that is, none of the eigenvalues vanish. Numerically, this particular property is not desirable, and which unfortunately also carries over to the Euler equations. As we shall see in the next section, there are other splitting schemes that remove this difficulty.

### 8.3.2 FVS Numerical Schemes

The FVS approach can be used to solve (8.36) using the *explicit* conservative scheme

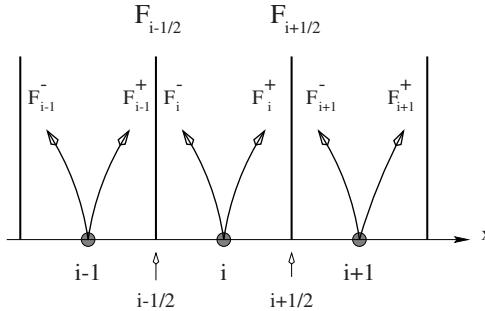
$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \frac{\Delta t}{\Delta x} [\mathbf{F}_{i+\frac{1}{2}} - \mathbf{F}_{i-\frac{1}{2}}], \quad (8.47)$$

where the FVS numerical flux is given by

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}_i^+(\mathbf{U}_i^n) + \mathbf{F}_{i+1}^-(\mathbf{U}_{i+1}^n). \quad (8.48)$$

Fig. 8.1 provides a physical interpretation of (8.48). The intercell numerical flux  $\mathbf{F}_{i+\frac{1}{2}}$  is made out from two contributions; one comes from the *forward* component  $\mathbf{F}_i^+$  in the *left cell*  $I_i$  and the other comes from the *backward* component  $\mathbf{F}_{i+1}^-$  in the *right cell*  $I_{i+1}$ .

The Steger and Warming [463] splitting (8.32) in a computational set up is as follows: we consider a computing cell  $I_i$  at time level  $n$ , where  $\mathbf{U}_i^n$  is the vector of conserved variables and  $\mathbf{F}_i^n \equiv \mathbf{F}(\mathbf{U}_i^n)$  is the vector of fluxes. The three cases to consider are illustrated in Fig. 8.2 and are



**Fig. 8.1.** Splitting of the flux function *within* each computing cell  $I_i$  at time level  $n$

- Case (a) Left supersonic flow:  $\lambda_2 = u_i^n + a_i^n \leq 0$ . Fig. 8.2a illustrates the situation in a cell  $I_i$  at time level  $n$ . Clearly

$$\left. \begin{aligned} \lambda_1^+ &= 0, \quad \lambda_1^- = \lambda_1 = u_i^n - a_i^n, \\ \lambda_2^+ &= 0, \quad \lambda_2^- = \lambda_2 = u_i^n + a_i^n, \\ \mathbf{F}_i^+ &= \mathbf{0}, \quad \mathbf{F}_i^- = \mathbf{F}_i^n. \end{aligned} \right\} \quad (8.49)$$

- Case (b) Right supersonic flow:  $\lambda_1 = u_i^n - a_i^n \geq 0$ . See Fig. 8.2b. Obviously

$$\left. \begin{aligned} \lambda_1^+ &= \lambda_1 = u_i^n - a_i^n, \quad \lambda_1^- = 0, \\ \lambda_2^+ &= \lambda_2 = u_i^n + a_i^n, \quad \lambda_2^- = 0, \\ \mathbf{F}_i^+ &= \mathbf{F}_i^n, \quad \mathbf{F}_i^- = \mathbf{0}. \end{aligned} \right\} \quad (8.50)$$

- Case (c) Subsonic flow:  $\lambda_1 = u_i^n - a_i^n \leq 0 \leq \lambda_2 = u_i^n + a_i^n$ . See Fig. 8.2c. Evidently

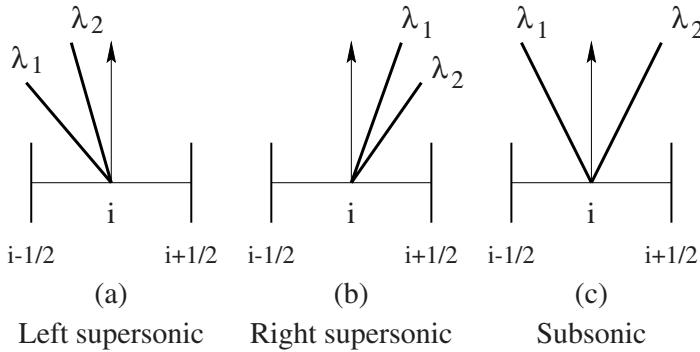
$$\left. \begin{aligned} \lambda_1^+ &= 0, \quad \lambda_1^- = \lambda_1 = u_i^n - a_i^n, \\ \lambda_2^+ &= \lambda_2 = u_i^n + a_i^n, \quad \lambda_2^- = 0. \end{aligned} \right\} \quad (8.51)$$

According to (8.45)–(8.46) the fluxes  $\mathbf{F}_i^+$  and  $\mathbf{F}_i^-$  for the subsonic case are given by

$$\mathbf{F}_i^+ = \frac{\rho_i^n}{2} \left[ \frac{u_i^n + a_i^n}{(u_i^n + a_i^n)^2} \right], \quad \mathbf{F}_i^- = \frac{\rho_i^n}{2} \left[ \frac{u_i^n - a_i^n}{(u_i^n - a_i^n)^2} \right]. \quad (8.52)$$

## 8.4 FVS Applied to the Euler Equations

Here we present three Flux Vector Splitting schemes applied to the time dependent Euler equations.



**Fig. 8.2.** Possible flow patterns in cell  $I_i$  at time  $n$ : (a) supersonic flow to the left  
(b) supersonic flow to the right (c) subsonic

#### 8.4.1 Recalling the Equations

The one-dimensional Euler Equations in conservation-law form are given by

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0}, \quad (8.53)$$

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ E \end{bmatrix}, \quad \mathbf{F}(\mathbf{U}) = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ u(E + p) \end{bmatrix}. \quad (8.54)$$

As seen in Sect. 3.1.1 of Chap. 3, the Jacobian matrix  $\mathbf{A}$  is given by

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2}(\gamma - 3)u^2 & (3 - \gamma)u & \gamma - 1 \\ \frac{1}{2}(\gamma - 2)u^3 - \frac{a^2 u}{\gamma - 1} & \frac{3 - 2\gamma}{2}u^2 + \frac{a^2}{\gamma - 1} & \gamma u \end{bmatrix} \quad (8.55)$$

and the system is hyperbolic with real eigenvalues

$$\lambda_1 = u - a, \quad \lambda_2 = u, \quad \lambda_3 = u + a. \quad (8.56)$$

The matrix  $\mathbf{K}$  of corresponding right eigenvectors is

$$\mathbf{K} = \begin{bmatrix} 0 & 1 & 0 \\ u - a & u & u + a \\ H - ua & \frac{1}{2}u^2 & H + ua \end{bmatrix}. \quad (8.57)$$

Here  $H$  is the enthalpy

$$H = (E + p)/\rho = \frac{1}{2}u^2 + \frac{a^2}{(\gamma - 1)}. \quad (8.58)$$

As explained in Sect. 3.2.4 of Chap. 3, the three-dimensional Euler equations may be dealt with by only considering the flux component *normal* to

the computing cell interface, see also Sect. 16.2 of Chap. 16. In constructing numerical methods for Cartesian geometries it is sufficient to consider the flux in any of the coordinate directions. For general geometries this is modified by use of rotation matrices; see Sect. 3.2 of Chap. 3. We thus state the schemes for the  $x$ -split three dimensional Euler equations

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0}, \quad (8.59)$$

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ E \end{bmatrix}, \quad \mathbf{F}(\mathbf{U}) = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ \rho uw \\ u(E + p) \end{bmatrix}. \quad (8.60)$$

The Jacobian matrix  $\mathbf{A}$ , see Sect. 3.2.2 of Chap. 3, is given by

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ \hat{\gamma}H - u^2 - a^2 & (3 - \gamma)u & -\hat{\gamma}v & -\hat{\gamma}w & \hat{\gamma} \\ -uv & v & u & 0 & 0 \\ -uw & w & 0 & u & 0 \\ \frac{1}{2}u[(\gamma - 3)H - a^2] & H - \hat{\gamma}u^2 & -\hat{\gamma}uv & -\hat{\gamma}uw & \gamma u \end{bmatrix}, \quad (8.61)$$

where

$$H = (E + p)/\rho = \frac{1}{2}\mathbf{V}^2 + \frac{a^2}{(\gamma - 1)}, \quad \mathbf{V}^2 = u^2 + v^2 + w^2, \quad \hat{\gamma} = \gamma - 1. \quad (8.62)$$

This system is hyperbolic with real eigenvalues

$$\lambda_1 = u - a, \quad \lambda_2 = \lambda_3 = \lambda_4 = u, \quad \lambda_5 = u + a. \quad (8.63)$$

The matrix of corresponding right eigenvectors is

$$\mathbf{K} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 \\ u - a & u & 0 & 0 & u + a \\ v & v & 1 & 0 & v \\ w & w & 0 & 1 & w \\ H - ua & \frac{1}{2}\mathbf{V}^2 & v & w & H + ua \end{bmatrix} \quad (8.64)$$

As seen in Chap. 3 the one-dimensional Euler equations satisfy the *homogeneity property*

$$\mathbf{F}(\mathbf{U}) = \mathbf{A}(\mathbf{U})\mathbf{U}. \quad (8.65)$$

**Exercise 8.8.** Verify that the split three-dimensional Euler equations (8.59)–(8.60) also satisfy the homogeneity property (8.65).

**Solution 8.9.** (Left to the reader).

### 8.4.2 The Steger–Warming Splitting

For a splitting (8.25)–(8.27) we require an expression for the inverse  $\mathbf{K}^{-1}$  of the matrix  $\mathbf{K}$ , in order to find the split Jacobians (8.29).

#### The One-Dimensional Case

For the one-dimensional Euler equations we have

$$\mathbf{K}^{-1} = \frac{(\gamma - 1)}{2a^2} \begin{bmatrix} \frac{1}{2}u^2 + \frac{ua}{\gamma-1} & -u - \frac{a}{\gamma-1} & 1 \\ \frac{2a^2}{\gamma-1} - u^2 & 2u & -2 \\ \frac{1}{2}u^2 - \frac{ua}{\gamma-1} & \frac{a}{\gamma-1} - u & 1 \end{bmatrix}. \quad (8.66)$$

Then, for any component  $\Lambda^\alpha$  of the two components of  $\Lambda$  in (8.26) the corresponding Jacobian component is

$$\mathbf{A}^\alpha = \mathbf{K}\Lambda^\alpha\mathbf{K}^{-1}.$$

The associated split flux component  $\mathbf{F}^\alpha = \mathbf{A}^\alpha\mathbf{U}$  is

$$\mathbf{F}^\alpha = \frac{\rho}{2\gamma} \begin{bmatrix} \lambda_1^\alpha + 2(\gamma - 1)\lambda_2^\alpha + \lambda_3^\alpha \\ (u - a)\lambda_1^\alpha + 2(\gamma - 1)u\lambda_2^\alpha + (u + a)\lambda_3^\alpha \\ (H - ua)\lambda_1^\alpha + (\gamma - 1)u^2\lambda_2^\alpha + (H + ua)\lambda_3^\alpha \end{bmatrix}, \quad (8.67)$$

where the eigenvalues  $\lambda_k^\alpha$  are given by (8.32), for  $\alpha = +, -$ .

#### The Three-Dimensional Case

For the three-dimensional case we have

$$\mathbf{K}^{-1} = \frac{(\gamma - 1)}{2a^2} \begin{bmatrix} H + \frac{a}{\hat{\gamma}}(u - a) - (u + \frac{a}{\hat{\gamma}}) & -v & -w & 1 \\ -2H + \frac{4}{\hat{\gamma}}a^2 & 2u & 2v & 2w & -2 \\ -\frac{2va^2}{\hat{\gamma}} & 0 & \frac{2a^2}{\hat{\gamma}} & 0 & 0 \\ -\frac{2wa^2}{\hat{\gamma}} & 0 & 0 & \frac{2a^2}{\hat{\gamma}} & 0 \\ H - \frac{a}{\hat{\gamma}}(u + a) & -u + \frac{a}{\hat{\gamma}} & -v & -w & 1 \end{bmatrix} \quad (8.68)$$

and the resulting split flux component  $\mathbf{F}^\alpha = \mathbf{A}^\alpha\mathbf{U}$  is found to be

$$\mathbf{F}^\alpha = \frac{\rho}{2\gamma} \begin{bmatrix} \lambda_1^\alpha + 2(\gamma - 1)\lambda_2^\alpha + \lambda_5^\alpha \\ (u - a)\lambda_1^\alpha + 2(\gamma - 1)u\lambda_2^\alpha + (u + a)\lambda_5^\alpha \\ v\lambda_1^\alpha + 2(\gamma - 1)v\lambda_2^\alpha + v\lambda_5^\alpha \\ w\lambda_1^\alpha + 2(\gamma - 1)w\lambda_2^\alpha + w\lambda_5^\alpha \\ (H - ua)\lambda_1^\alpha + (\gamma - 1)\mathbf{V}^2\lambda_2^\alpha + (H + ua)\lambda_5^\alpha \end{bmatrix}. \quad (8.69)$$

**Exercise 8.10.** Verify expressions (8.68) and (8.69) above.

**Solution 8.11.** (Left to the reader).

### 8.4.3 The van Leer Splitting

Van Leer [560], [561] constructed a splitting for the Euler equations that has some extra desirable properties, namely

- (I) The split Jacobian matrices

$$\hat{\mathbf{A}}^+ = \frac{\partial \mathbf{F}^+}{\partial \mathbf{U}}, \quad \hat{\mathbf{A}}^- = \frac{\partial \mathbf{F}^-}{\partial \mathbf{U}}$$

are required to be continuous.

- (II) The split fluxes are degenerate *for subsonic flow*, that is  $\hat{\mathbf{A}}^+$ ,  $\hat{\mathbf{A}}^-$  have a zero eigenvalue.

Van Leer expresses the flux vector  $\mathbf{F}$  as a function of density, sound speed and Mach number  $M = \frac{u}{a}$ , that is

$$\mathbf{F} = \mathbf{F}(\rho, a, M) = \begin{bmatrix} \rho a M \\ \rho a^2 (M^2 + \frac{1}{\gamma}) \\ \rho a^3 M (\frac{1}{2} M^2 + \frac{1}{\gamma-1}) \end{bmatrix} \equiv \begin{bmatrix} f_{\text{mas}} \\ f_{\text{mom}} \\ f_{\text{ene}} \end{bmatrix}. \quad (8.70)$$

For the mass flux

$$f_{\text{mas}} = \rho a M$$

one requires quadratics in  $M$  and the split mass fluxes are

$$f_{\text{mas}}^+ = \frac{1}{4} \rho a (1 + M)^2, \quad f_{\text{mas}}^- = -\frac{1}{4} \rho a (1 - M)^2. \quad (8.71)$$

The momentum split fluxes are

$$f_{\text{mom}}^+ = f_{\text{mas}}^+ \frac{2a}{\gamma} \left[ \frac{(\gamma-1)}{2} M + 1 \right], \quad f_{\text{mom}}^- = f_{\text{mas}}^- \frac{2a}{\gamma} \left[ \frac{(\gamma-1)}{2} M - 1 \right] \quad (8.72)$$

and the energy split fluxes are

$$f_{\text{ene}}^+ = \frac{\gamma^2}{2(\gamma^2 - 1)} \frac{[f_{\text{mom}}^+]^2}{f_{\text{mas}}^+}, \quad f_{\text{ene}}^- = \frac{\gamma^2}{2(\gamma^2 - 1)} \frac{[f_{\text{mom}}^-]^2}{f_{\text{mas}}^-}. \quad (8.73)$$

In vector form we have

$$\mathbf{F}^+ = \frac{1}{4} \rho a (1 + M)^2 \begin{bmatrix} 1 \\ \frac{2a}{\gamma} \left( \frac{\gamma-1}{2} M + 1 \right) \\ \frac{2a^2}{\gamma^2-1} \left( \frac{\gamma-1}{2} M + 1 \right)^2 \end{bmatrix}, \quad (8.74)$$

$$\mathbf{F}^- = -\frac{1}{4} \rho a (1 - M)^2 \begin{bmatrix} 1 \\ \frac{2a}{\gamma} \left( \frac{\gamma-1}{2} M - 1 \right) \\ \frac{2a^2}{\gamma^2-1} \left( \frac{\gamma-1}{2} M - 1 \right)^2 \end{bmatrix}. \quad (8.75)$$

For the  $x$ -split three dimensional Euler equations the split flux formulae are

$$\mathbf{F}^+ = \frac{1}{4}\rho a(1+M)^2 \begin{bmatrix} 1 \\ \frac{2a}{\gamma}(\frac{\gamma-1}{2}M+1) \\ v \\ w \\ \frac{2a^2}{\gamma^2-1}(\frac{\gamma-1}{2}M+1)^2 + \frac{1}{2}(v^2+w^2) \end{bmatrix}, \quad (8.76)$$

and

$$\mathbf{F}^- = -\frac{1}{4}\rho a(1-M)^2 \begin{bmatrix} 1 \\ \frac{2a}{\gamma}(\frac{\gamma-1}{2}M-1) \\ v \\ w \\ \frac{2a^2}{\gamma^2-1}(\frac{\gamma-1}{2}M-1)^2 + \frac{1}{2}(v^2+w^2) \end{bmatrix}, \quad (8.77)$$

where the Mach number is still  $M = \frac{u}{a}$ .

Concerning stability, van Leer [560] gives the following practical stability condition

$$C_{\text{cfl}} \equiv \frac{\Delta t}{\Delta x}(|u|+a) \leq \frac{2\gamma + |M|(3-\gamma)}{\gamma + 3}. \quad (8.78)$$

Note that  $C_{\text{cfl}} = C_{\text{cfl}}(M)$  and that when  $\gamma = 1.4$  we have

$$C_{\text{cfl}}^{\max} = 1 \text{ for } |M| = 1, \quad C_{\text{cfl}}^{\min} = \frac{2\gamma}{\gamma + 3} \approx 0.636\dots, \text{ for } |M| = 0. \quad (8.79)$$

*Remark 8.12.* The CFL condition for the explicit FVS scheme is more restrictive than that for the Godunov method, for which  $C_{\text{cfl}}$  is close to unity. See Sect. 6.3.2 of Chap. 6 for a discussion on the CFL condition.

#### 8.4.4 The Liou–Steffen Scheme

A recent scheme that attempts to combine features from the Flux Vector Splitting and Godunov approaches is due to Liou and Steffen [328]. The scheme has been formulated in terms of the time-dependent Euler equations and relies on splitting the flux vector  $\mathbf{F}$  into a *convective* component  $\mathbf{F}^{(c)}$  and a *pressure* component  $\mathbf{F}^{(p)}$ . For the  $x$ -split three dimensional flux we have

$$\mathbf{F}(\mathbf{U}) = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ \rho uw \\ u(E+p) \end{bmatrix} = \begin{bmatrix} \rho u \\ \rho u^2 \\ \rho uv \\ \rho uw \\ \rho uH \end{bmatrix} + \begin{bmatrix} 0 \\ p \\ 0 \\ 0 \\ 0 \end{bmatrix} \equiv \mathbf{F}^{(c)} + \mathbf{F}^{(p)}, \quad (8.80)$$

with the obvious definitions for the convective component  $\mathbf{F}^{(c)}$  and the pressure component  $\mathbf{F}^{(p)}$ . By introducing the Mach number and enthalpy

$$M = \frac{u}{a}, \quad H = \frac{E + p}{\rho}$$

we write

$$\mathbf{F}^{(c)} = M \begin{bmatrix} \rho a \\ \rho au \\ \rho av \\ \rho aw \\ \rho aH \end{bmatrix} \equiv M \hat{\mathbf{F}}^{(c)}, \quad (8.81)$$

with the obvious notation for the vector  $\hat{\mathbf{F}}^{(c)}$ . In defining the intercell numerical flux  $\mathbf{F}_{i+\frac{1}{2}}$ , Liou and Steffen take

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}_{i+\frac{1}{2}}^{(c)} + \mathbf{F}_{i+\frac{1}{2}}^{(p)}, \quad (8.82)$$

where the *convective flux component* is given by

$$\mathbf{F}_{i+\frac{1}{2}}^{(c)} = M_{i+\frac{1}{2}} \left[ \hat{\mathbf{F}}^{(c)} \right]_{i+\frac{1}{2}}, \quad (8.83)$$

with definition

$$[\bullet]_{i+\frac{1}{2}} = \begin{cases} [\bullet]_i & \text{if } M_{i+\frac{1}{2}} \geq 0, \\ [\bullet]_{i+1} & \text{if } M_{i+\frac{1}{2}} \leq 0. \end{cases} \quad (8.84)$$

Note that the flux vector in (8.83) is upwinded according to the sign of the convection, or advection, speed *implied* in the intercell Mach number  $M_{i+\frac{1}{2}}$ . For this reason Liou and Steffen call their scheme AUSM, which stands for Advection Upstream Splitting Method.

The cell-interface Mach number is given by the splitting

$$M_{i+\frac{1}{2}} = M_i^+ + M_{i+1}^- \quad (8.85)$$

with the positive and negative components yet to be defined. The splitting of the pressure flux component depends on the splitting of the pressure itself, namely

$$p_{i+\frac{1}{2}} = p_i^+ + p_{i+1}^- . \quad (8.86)$$

For the splitting of the Mach number Liou and Steffen follow van Leer and set

$$M^\pm = \begin{cases} \pm \frac{1}{4}(M \pm 1)^2 & \text{if } |M| \leq 1, \\ \frac{1}{2}(M \pm |M|) & \text{if } |M| > 1. \end{cases} \quad (8.87)$$

For splitting the pressure they suggest two choices, namely

$$p^\pm = \begin{cases} \frac{1}{2}p(1 \pm M) & \text{if } |M| \leq 1 \\ \frac{1}{2}p \frac{(M \pm |M|)}{|M|} & \text{if } |M| > 1 \end{cases} \quad (8.88)$$

and

$$p^\pm = \begin{cases} \frac{1}{4}p(M \pm 1)^2(2 \mp M) & \text{if } |M| \leq 1, \\ \frac{1}{2}p \frac{(M \pm |M|)}{|M|} & \text{if } |M| > 1. \end{cases} \quad (8.89)$$

For more details see the original paper by Liou and Steffen [328] and the more recent publication of Liou [327].

## 8.5 Numerical Results

Here we illustrate the performance of three FVS-type schemes on the one-dimensional, time dependent Euler equations for ideal gases with  $\gamma = 1.4$ , namely the Steger–Warming FVS scheme, the van Leer FVS scheme and the AUSM scheme of Liou and Steffen. Numerical results are compared with the exact solution. The respective results are obtained from running two codes of *NUMERICA* [519], namely HE-E1FVS (FVS schemes) and HE-E1RPEXACT (exact Riemann solver).

### 8.5.1 Tests

We use five test problems with exact solution. Data consists of two constant states  $\mathbf{W}_L = [\rho_L, u_L, p_L]^T$  and  $\mathbf{W}_R = [\rho_R, u_R, p_R]^T$ , separated by a discontinuity at a position  $x = x_0$ , and are given in Table 8.1. The exact and numerical solutions are found in the spatial domain  $0 \leq x \leq 1$ . The numerical solution is computed with  $M = 100$  cells. The Courant number coefficient is taken as  $C_{\text{cfl}} = 0.9$ , except for the van Leer scheme, for which we took  $C_{\text{cfl}} = 0.6$ . In implementing the CFL condition we use the simple formula given by equation 6.20 of Chap. 6 to estimate the maximum wave speed. Therefore, for all methods, we reduce the CFL number further to 0.2 of that given by the CFL condition, for the first 5 time steps. Boundary conditions are transmissive. For each test problem we select a convenient position  $x_0$  of the initial discontinuity and the output time; these are stated in the legend of each figure displaying computational results. All numerical results should be compared with those from Godunov’s method, Figs. 6.8 to 6.12 of Chap. 6. For more details on the exact solutions of the test problems see Sect. 4.3.3 of Chap. 4.

Test	$\rho_L$	$u_L$	$p_L$	$\rho_R$	$u_R$	$p_R$
1	1.0	0.75	1.0	0.125	0.0	0.1
2	1.0	-2.0	0.4	1.0	2.0	0.4
3	1.0	0.0	1000.0	1.0	0.0	0.01
4	5.99924	19.5975	460.894	5.99242	-6.19633	46.0950
5	1.0	-19.59745	1000.0	1.0	-19.59745	0.01

Table 8.1. Data for five test problems with exact solution. Test 5 is like Test 3 with negative uniform background speed

### 8.5.2 Results for Test 1

Test 1 is a *modified* version of the popular Sod’s test [453]; the solution consists of a right shock wave, a right travelling contact wave and a left *sonic* rarefaction wave; this test is very useful in assessing the *entropy satisfaction*

property of numerical methods. Figs. 8.3 to 8.5 show the results for the three FVS schemes.

In the results from the Steger–Warming scheme, shown in Fig. 8.3, the resolution of the shock and the right travelling contact is comparable with that of Godunov’s method, Fig. 6.8 of Chap. 6. The resolution of the left rarefaction is less satisfactory; the head and tail are visibly smeared and the sonic point, as expected, is not handled correctly. The results from the van Leer scheme, shown in Fig. 8.4, are virtually identical to those of Godunov’s method of Fig. 6.8 for the rarefaction and contact, but the shock is broader. The performance at the sonic point is comparable with that of Godunov’s method and better than that of the Steger–Warming scheme. The results from the Liou and Steffen scheme are shown in Fig. 8.5. In comparison with Godunov’s method, the shock wave is more sharply resolved and the contact wave is similar but the resolution of the rarefaction is not as good, particularly near the sonic point.

### 8.5.3 Results for Test 2

The exact solution of Test 2 consists of two symmetric rarefaction waves and a trivial contact wave of zero speed; the *Star Region* between the non-linear waves is close to vacuum, which makes this problem a suitable test for assessing the performance of numerical methods for low-density flows [182]; this is the so called *123 problem* introduced in Sect. 4.3.3 of Chap. 4. Figs. 8.6 to 8.8 show the results for the three FVS schemes.

The results from the Steger–Warming scheme, shown in Fig. 8.6, are virtually identical to those of the Godunov method, Fig. 6.9 of Chap. 6. The results from the van Leer scheme, shown in Fig. 8.7, are also comparable with those from the Godunov method. The heads of the rarefactions are slightly more diffused. The Liou and Steffen scheme, Fig. 8.8, gives results that are comparable with those of Godunov’s method and slightly more accurate than those from van Leer’s scheme; in the vicinity of the trivial contact, where both density and pressure are close to zero, the results are somewhat erratic, see velocity and internal energy plots.

In view of the fact that Godunov-type methods with linearised Riemann solvers will fail for this test problem [182], it is quite remarkable to note that all three FVS-type schemes described in this chapter actually run and give, overall, good results.

### 8.5.4 Results for Test 3

Test 3 is designed to assess the robustness and accuracy of numerical methods; its solution consists of a strong right travelling shock wave of shock Mach number 198, a contact surface and a left rarefaction wave. Figs. 8.9 and 8.10 show the results for two FVS schemes.

The Steger–Warming result, shown in Fig. 8.9, is seen to be overall less accurate than the corresponding result from the Godunov method, shown in Fig. 6.10 of Chap. 6; the numerical solution has an unphysical dip behind the shock wave, which is more clearly seen in the velocity and pressure plots. The results from the van Leer scheme, shown in Fig. 8.10, are also less accurate than those from the Godunov method, but they are more accurate than the results from the Steger–Warming scheme. The Liou and Steffen scheme, as coded by the author, failed to give a solution at all for this very severe test problem, even when reducing the CFL number to a value as low as 0.1.

### 8.5.5 Results for Test 4

Test 4, as Test 3, is also designed to assess the robustness of numerical methods; data originates from two very strong shock waves travelling towards each other and the solution consists of three strong discontinuities travelling to the right; the left shock wave moves to the right very slowly, which adds another difficulty [406] to numerical methods. Figs. 8.11 to 8.13 show the results for the three FVS schemes.

The results from the Steger and Warming scheme, shown in Fig. 8.11, are overall comparable with those of Godunov’s method shown in Fig. 6.11 of Chapter 6. The only visible difference is seen near the left slowly moving shock, and as expected, this is more diffused in the Steger–Warming result; however, it appears as if the low frequency oscillations seen in the Godunov results are significantly reduced in the Steger–Warming scheme. The results from the van Leer scheme, shown in Fig. 8.12, are comparable with those of Godunov’s method and are more accurate than those from the Steger–Warming scheme. The slowly-moving shock is resolved with two interior cells, instead of one in the Godunov’s method, but low-frequency spurious oscillations are just about visible. The results from the Liou and Steffen scheme, shown in Fig. 8.13, are comparable with the Godunov and van Leer results for this test; the fast right shock is more sharply resolved than with the other methods, but at the cost of an overshoot; the slowly moving left shock is slightly more smeared than in the van Leer result.

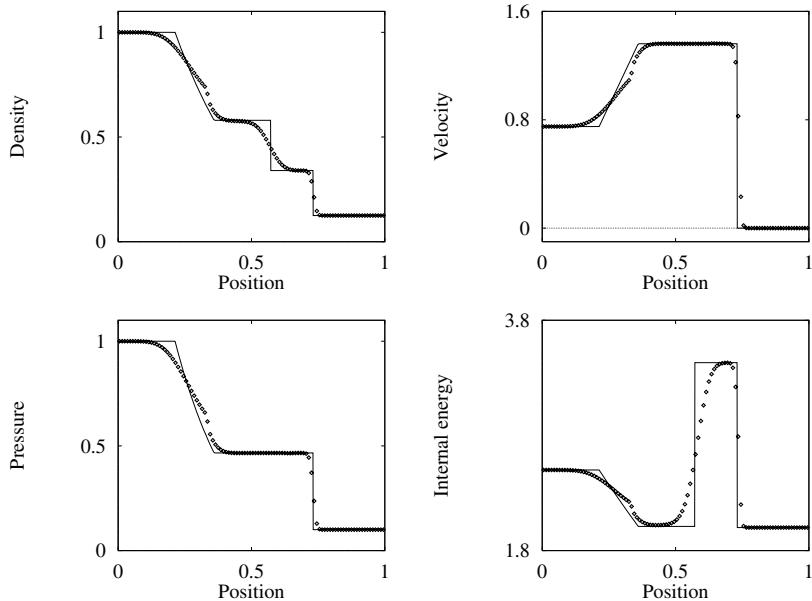
### 8.5.6 Results for Test 5

Test 5 is effectively Test 3, with a negative uniform background speed so as to obtain a stationary contact discontinuity. Test 5 is also designed to test the robustness of numerical methods but the main reason for devising this test is for assessing the ability of numerical methods to resolve *slowly-moving contact discontinuities*. The exact solution of Test 5 consists of a left rarefaction wave, a right-travelling shock wave (slow) and a *stationary* contact discontinuity. Figs. 8.14 to 8.16 show the results for the three FVS schemes and Fig. 8.17 shows the respective results obtained from the Godunov method used in conjunction with the exact Riemann solver, code HE-E1GODSTATE

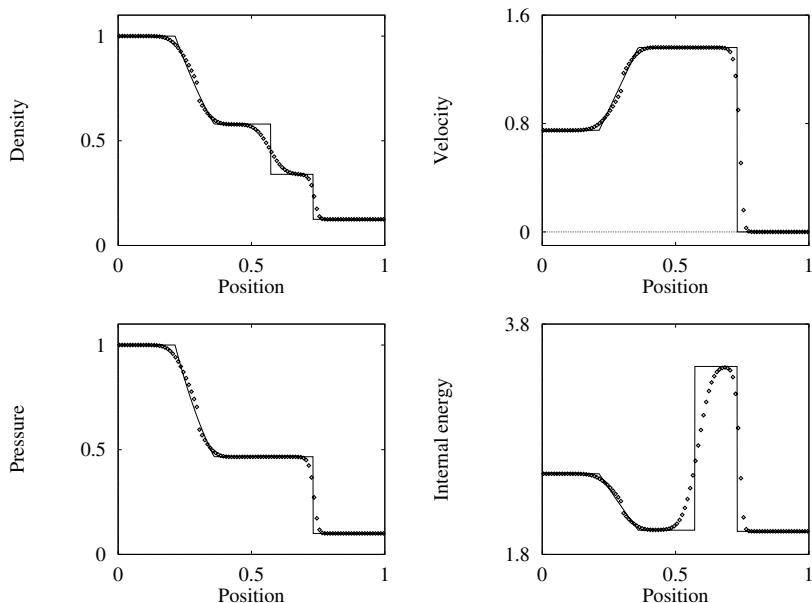
of *NUMERICA* [519]. We note that at the chosen output time, the right travelling shock wave has propagated only 5 cells, in 81 time steps. For this test problem the results from the Steger/Warming and van Leer FVS schemes are similar, in that the contact discontinuity is heavily smeared, even for a relatively short evolution time. The Liou and Steffen FVS scheme, Fig. 8.16, differs from the other two FVS schemes in that it resolves the contact discontinuity more sharply; note however the unphysical oscillations in the vicinity of the shock wave, the contact discontinuity and even near the tail of the rarefaction. For comparison, the results from the Godunov method used in conjunction with the exact Riemann solver are displayed in Fig. 8.17. These are obviously superior to any of the FVS schemes, for this test problem.

The numerical experiments presented suggest that Flux Vector Splitting Schemes give, generally, results of similar quality to those obtained by the Godunov method. The difference between these two upwind approaches is evident when *slowly or stationary* contact waves are present. For multidimensional problems this has important implications for the accurate resolution of shear layers, material interfaces and vortical flows. The Liou and Steffen FVS-type scheme is an exception, as it does resolve contacts more accurately than the Warming–Beam and van Leer schemes, although there are questions marks about its robustness. For Test 3 the Liou and Steffen scheme *crashed* and for Test 5 produced large unphysical oscillations. It is worth remarking that the Godunov method was used in conjunction with the exact Riemann solver, to obtain the numerical results of Fig. 8.17. If the Godunov scheme is used with linearised Riemann solvers, then it would fail for low-density flows, such as Test 2 for example, whereas the FVS-type schemes appear to be much less sensitive; they all produced acceptable results for Test 2. In addition, if the Godunov method is used in conjunction with *incomplete* Riemann solvers, such as those that ignore the presence of linear waves in the structure of the solution of the Riemann problem, then the resolution of contacts will be as poor as that of FVS-type schemes, such as the Warming–Beam and van Leer schemes. The selection of the Riemann solver is crucial to the performance of the Godunov method. See Chaps. 9 to 12.

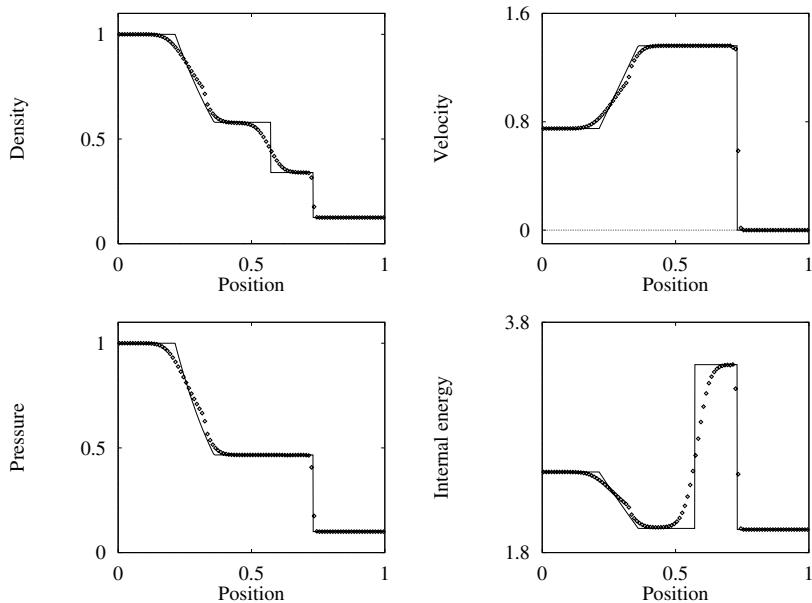
For details on how to extend FVS-type schemes to higher order of accuracy for one-dimensional homogeneous problems the reader is referred to Chaps. 13 and 14. Methods for treating source terms are given in Chapt. 15 and techniques to extend the methods to solve multidimensional problems are given in Chapt. 16. For multidimensional, steady state, applications of Flux Vector Splitting methods, readers are encouraged to consult, amongst many others, the following references: [12], [13], [600], [166], [578].



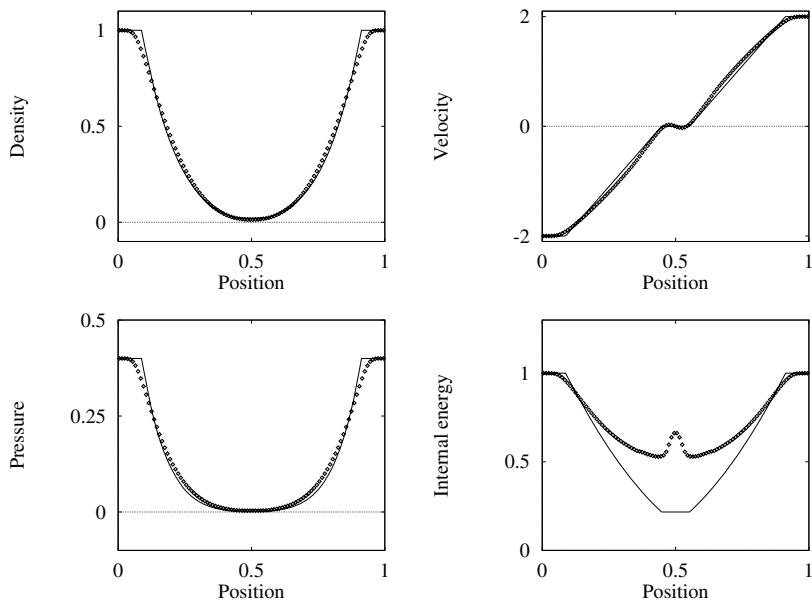
**Fig. 8.3.** Steger and Warming FVS scheme applied to Test 1, with  $x_0 = 0.3$ . Numerical (symbol) and exact (line) solutions are compared at time 0.2 units



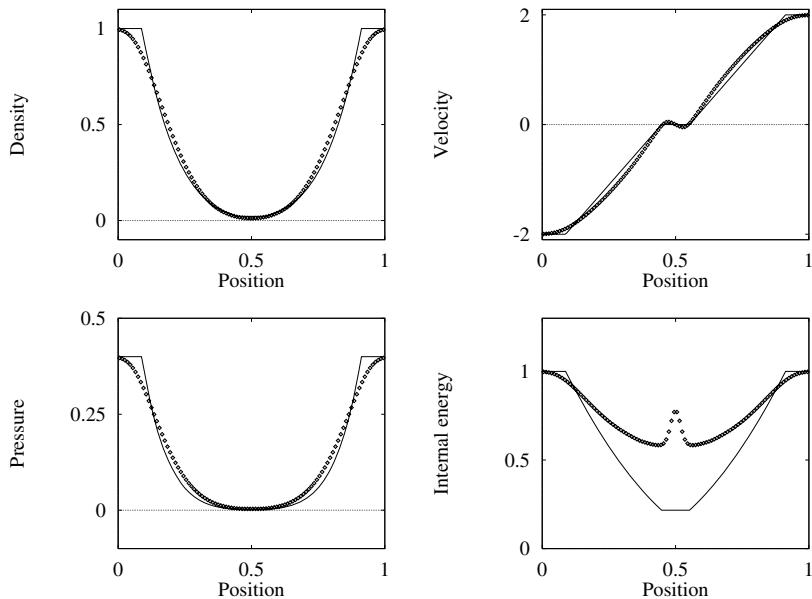
**Fig. 8.4.** Van Leer FVS scheme applied to Test 1, with  $x_0 = 0.3$ . Numerical (symbol) and exact (line) solutions are compared at time 0.2 units



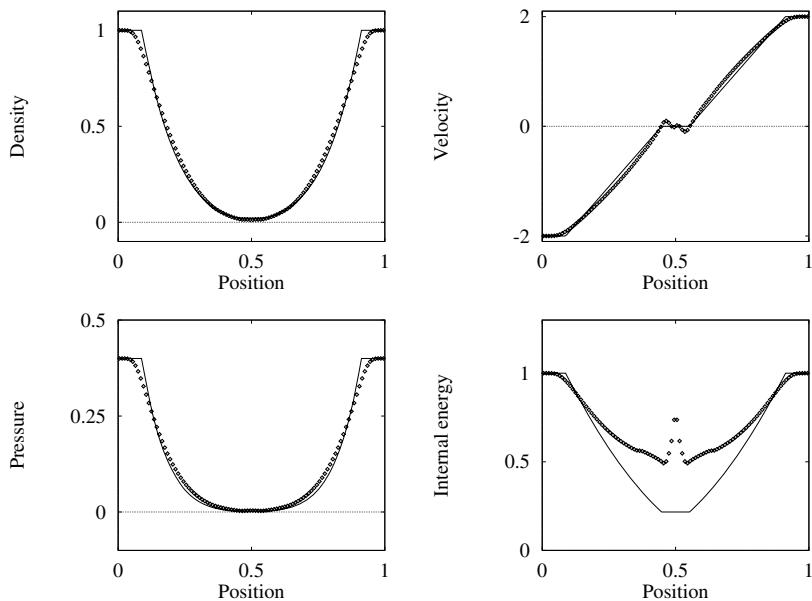
**Fig. 8.5.** Liou and Steffen scheme applied to Test 1, with  $x_0 = 0.3$ . Numerical (symbol) and exact (line) solutions are compared at time 0.2 units



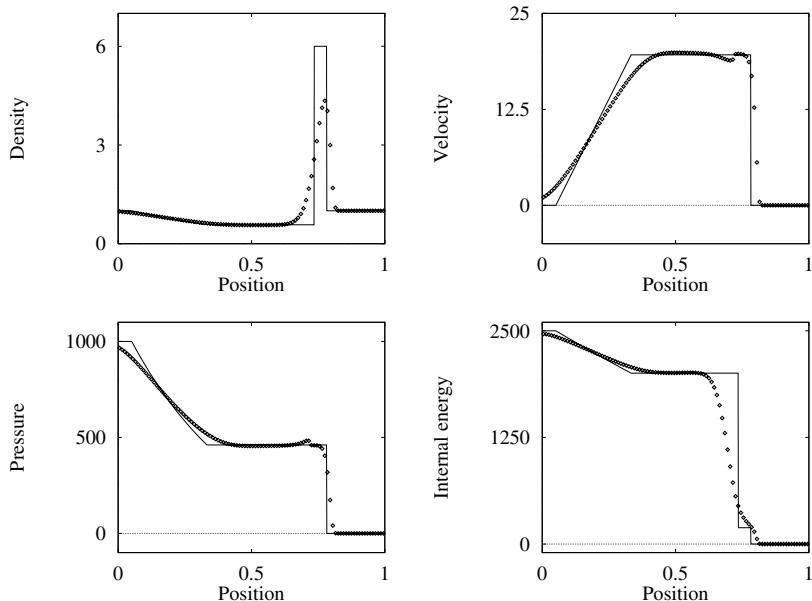
**Fig. 8.6.** Steger and Warming FVS scheme applied to Test 2, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at time 0.15 units



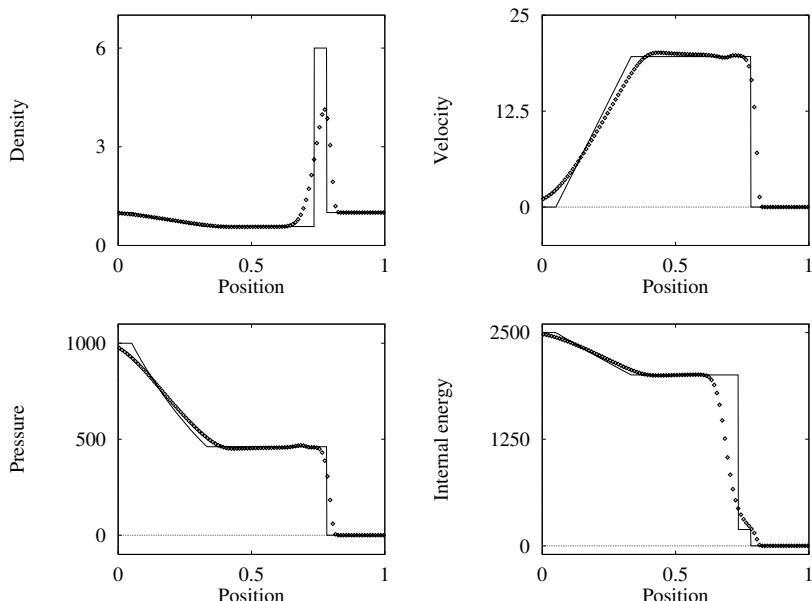
**Fig. 8.7.** Van Leer FVS scheme applied to Test 2, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at time 0.15 units



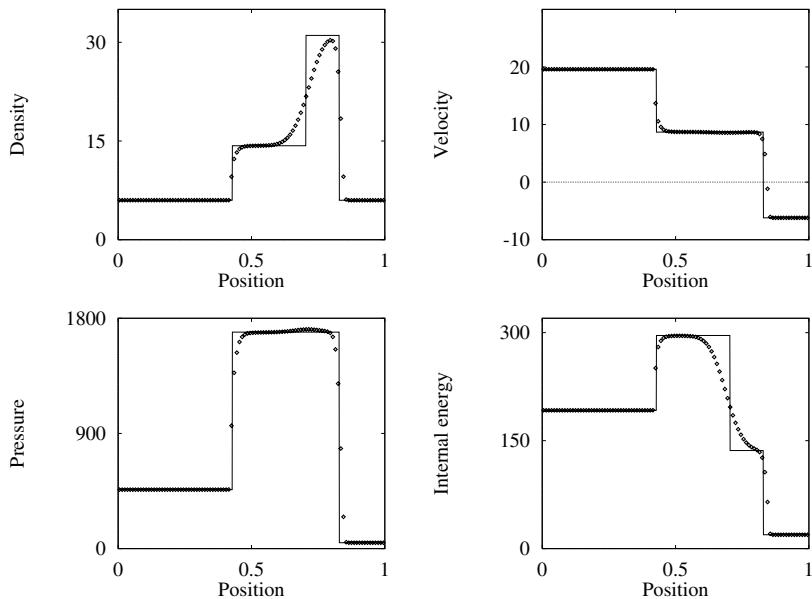
**Fig. 8.8.** Liou and Steffen scheme applied to Test 2, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at time 0.15 units



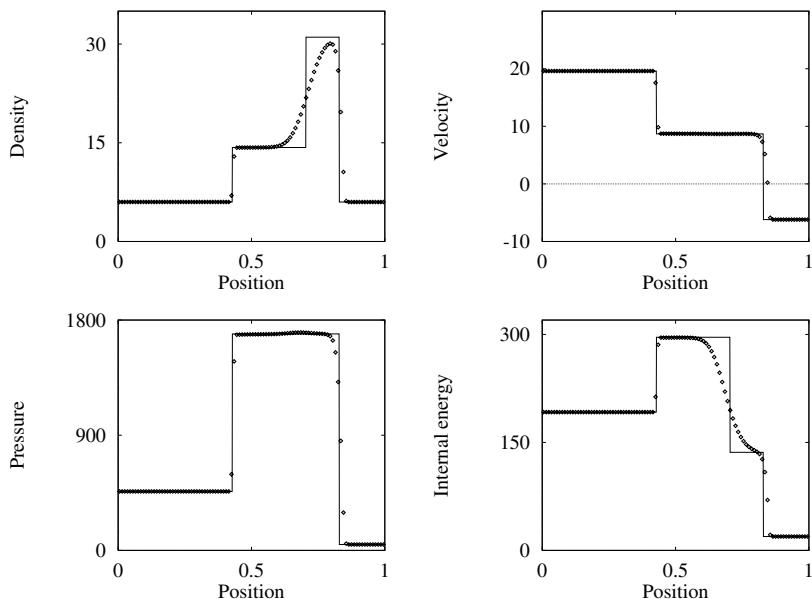
**Fig. 8.9.** Steger and Warming FVS scheme applied to Test 3, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012 units



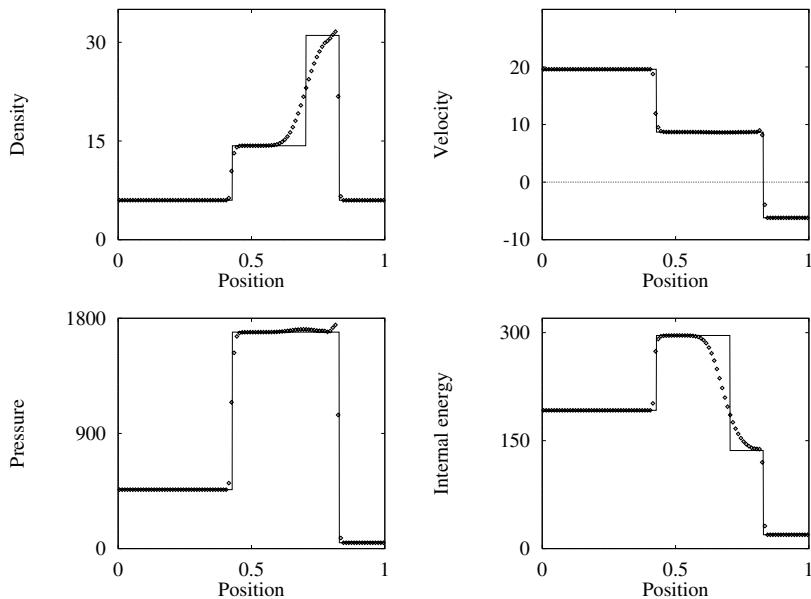
**Fig. 8.10.** Van Leer FVS scheme applied to Test 3, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012 units



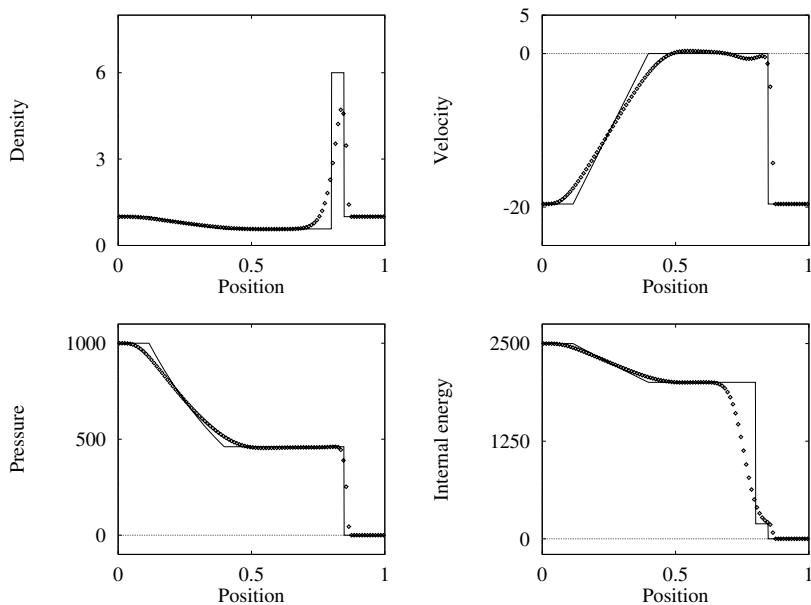
**Fig. 8.11.** Steger and Warming FVS scheme applied to Test 4, with  $x_0 = 0.4$ . Numerical (symbol) and exact (line) solutions are compared at time 0.035 units



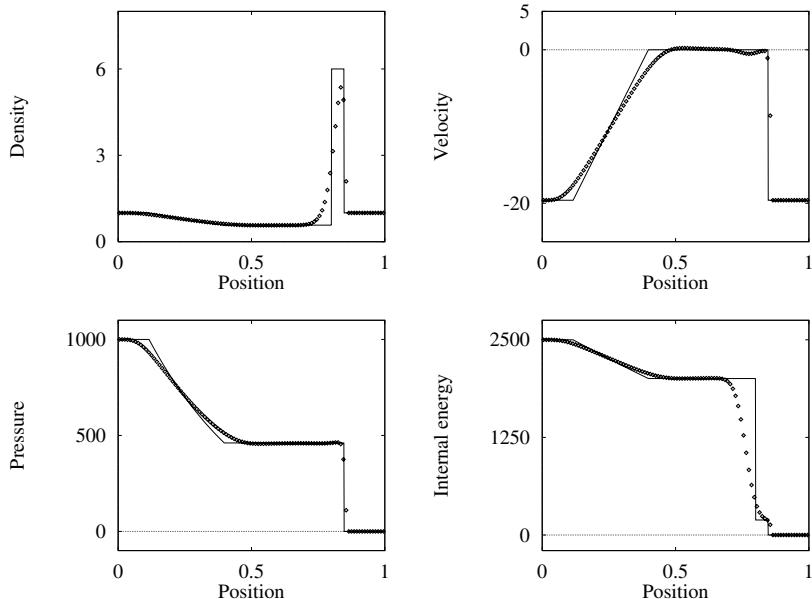
**Fig. 8.12.** Van Leer FVS scheme applied to Test 4, with  $x_0 = 0.4$ . Numerical (symbol) and exact (line) solutions are compared at time 0.035 units



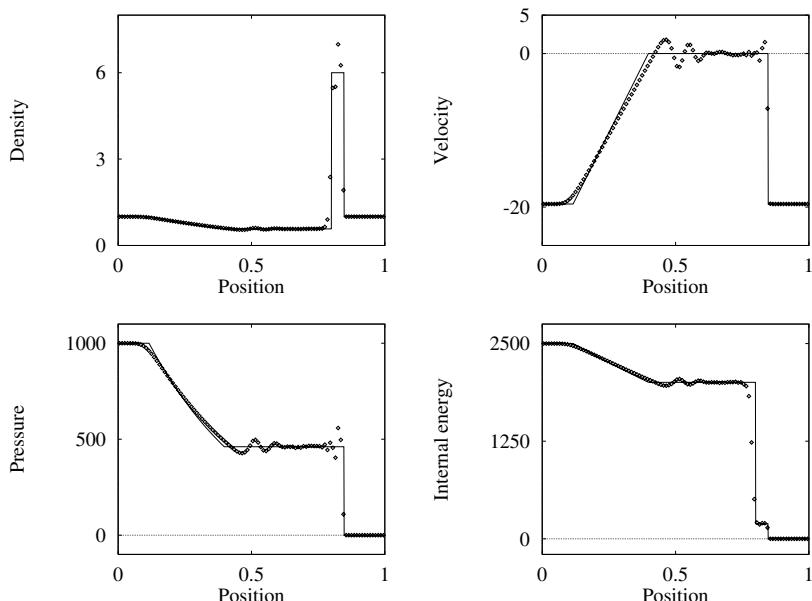
**Fig. 8.13.** Liou and Steffen scheme applied to Test 4, with  $x_0 = 0.4$ . Numerical (symbol) and exact (line) solutions are compared at time 0.035 units



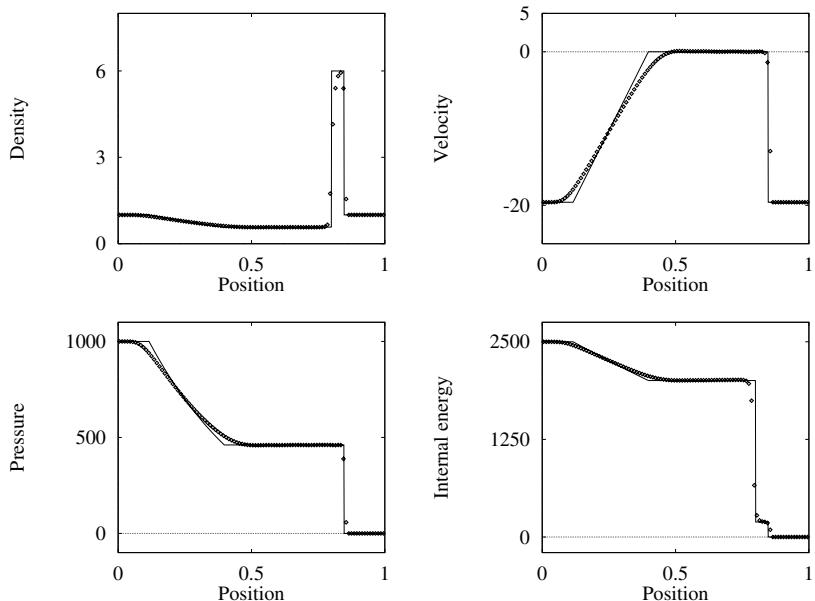
**Fig. 8.14.** Steger and Warming FVS scheme applied to Test 5, with  $x_0 = 0.8$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012 units



**Fig. 8.15.** Van Leer FVS scheme applied to Test 5, with  $x_0 = 0.8$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012 units



**Fig. 8.16.** Liou and Steffen scheme applied to Test 5, with  $x_0 = 0.8$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012 units



**Fig. 8.17.** Godunov scheme applied to Test 5, with  $x_0 = 0.8$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012 units

---

## Approximate–State Riemann Solvers

### 9.1 Introduction

The method of Godunov [216] and its high–order extensions require the solution of the Riemann problem. In a practical computation this is solved billions of times, making the Riemann problem solution process the single most demanding task in the numerical method. In Chap. 4 we provided exact Riemann solvers for the Euler equations for ideal and covolume gases. An iterative procedure is always involved and the associated computational effort may not always be justified. This effort may increase dramatically by equations of state of complicated algebraic form or by the complexity of the particular system of equations being solved, or both. Approximate, non–iterative solutions have the potential to provide the necessary items of information for numerical purposes. There are essentially two ways of extracting approximate information from the solution of the Riemann problem to be used in Godunov–type methods: one approach is to find an *approximation to the numerical flux* employed in the numerical method, directly, see Chaps. 10, 11 and 12; the other approach is to find an *approximation to a state* and then evaluate the physical flux function at this state. It is the latter route the one we follow in this chapter.

We present, approximate, Riemann solvers that do not need an iteration process. We provide an approximate *solution for the state* required to evaluate the Godunov flux. The approximations can be used directly in the first–oder Godunov method and its high–order extensions. Some of the approximations presented are exceedingly simple but not accurate enough to produce *robust numerical methods*. This difficulty is resolved by designing *hybrid schemes* that combine various approximate solvers in and adaptive fashion. There are other uses of the explicit approximate solutions presented here. For instance, the simplest solutions can be used in the *characteristic limiting* of high–order Godunov type methods based on the MUSCL approach; see Sect. 13.4 of Chap. 13. They also provide valuable information of use in other well known approximate Riemann solvers. For instance, Roe’s approximate Riemann solver, [407]

to be studied in Chap. 11, requires an entropy fix; the results of this chapter may be used to provide the state values in the Harten–Hyman entropy fix [243]. The approximate Riemann solver of Osher [372], to be studied in Chap. 12, requires intersection points for the integration paths; the approximations of this chapter can be used directly. The HLL approach of Harten, Lax and van Leer [244] for deriving approximate solutions to the Riemann problem, to be studied in Chap. 10, requires estimates for the smallest and largest signal velocities in the Riemann problem; again, the pressure–velocity approximation of this chapter can directly lead to estimates for wave speeds. The approximate solutions presented in this chapter may also be of use in other computational approaches, such as in *front tracking schemes* [468], [403]. The techniques discussed here can easily be extended to other systems, such as the shallow water equations, the steady supersonic Euler equations, the artificial compressibility equations (see Sect. 1.6.3 of Chap. 1) and the Euler equations with general equation of state.

Useful background for studying this chapter is found in Chaps. 2, 3, 4, and 6. The rest of this chapter is organised as follows: in Sect. 9.2 we recall the Godunov flux and the Riemann problem solution, in Sect. 9.3 we present very simple Riemann solvers based on primitive variable formulations of the Euler equations. In Sect. 9.4 we study approximations based on the exact function for pressure, namely the two–rarefaction approximation and the two–shock approximation. Hybrid schemes are dealt with in Sect. 9.5 and numerical results are presented in Sect. 9.6.

## 9.2 The Riemann Problem and the Godunov Flux

We want to solve numerically the general Initial Boundary Value Problem (IBVP)

$$\left. \begin{array}{l} \text{PDEs : } \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0} , \\ \text{ICs : } \mathbf{U}(x, 0) = \mathbf{U}^{(0)}(x) , \\ \text{BCs : } \mathbf{U}(0, t) = \mathbf{U}_l(t) , \quad \mathbf{U}(L, t) = \mathbf{U}_r(t) , \end{array} \right\} \quad (9.1)$$

utilising the explicit conservative formula

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n + \frac{\Delta t}{\Delta x} [\mathbf{F}_{i-\frac{1}{2}} - \mathbf{F}_{i+\frac{1}{2}}] , \quad (9.2)$$

along with the Godunov intercell numerical flux

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}(0)) . \quad (9.3)$$

We assume that the solution of IBVP (9.1) exists. Here  $\mathbf{U}_{i+\frac{1}{2}}(0)$  is the similarity solution  $\mathbf{U}_{i+\frac{1}{2}}(x/t)$  of the Riemann problem

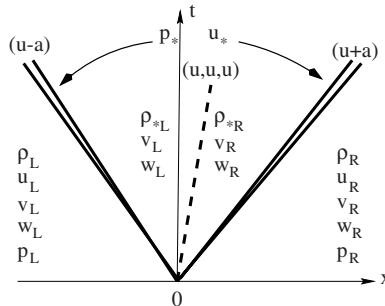
$$\left. \begin{array}{l} \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0} , \\ \mathbf{U}(x, 0) = \begin{cases} \mathbf{U}_L & \text{if } x < 0 , \\ \mathbf{U}_R & \text{if } x > 0 , \end{cases} \end{array} \right\} \quad (9.4)$$

evaluated at  $x/t = 0$ . Fig. 9.1 shows the structure of the exact solution of the Riemann problem for the  $x$ -split three-dimensional Euler equations, for which the vectors of conserved variables and fluxes are

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ E \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ \rho uw \\ u(E + p) \end{bmatrix}. \quad (9.5)$$

The value  $x/t = 0$  for the Godunov flux corresponds to the  $t$ -axis. See Chap. 6 for details. The piecewise constant initial data, in terms of primitive variables, is

$$\mathbf{W}_L = \begin{bmatrix} \rho_L \\ u_L \\ v_L \\ w_L \\ p_L \end{bmatrix}, \quad \mathbf{W}_R = \begin{bmatrix} \rho_R \\ u_R \\ v_R \\ w_R \\ p_R \end{bmatrix}. \quad (9.6)$$



**Fig. 9.1.** Structure of the solution of the Riemann problem for the  $x$ -split, three dimensional Euler equations. Data and solution are given in terms of primitive variables

The purpose of this chapter is to find *approximate* solutions to the Riemann problem in order to evaluate the Godunov flux. As seen in Chap. 6, the evaluation of the flux requires the identification of the appropriate wave pattern in the Riemann problem solution; as depicted in Fig. 9.2, there are ten possibilities to be considered.

In our solution procedure we split the task of solving the complete Riemann problem into three subproblems, namely

(I) The star values

$$p_*, u_*, \rho_{*L}, \rho_{*R} \quad (9.7)$$

in the *Star Region* between the non-linear waves, see Fig. 9.1.

- (II) The solution for the tangential velocity components  $v$  and  $w$  throughout the wave structure, and
- (III) The solution for  $\rho$ ,  $u$  and  $p$  inside sonic rarefactions.

Cases (II) and (III) are dealt with in the rest of this section, while case (I) is the subject of the rest of the chapter.

### 9.2.1 Tangential Velocity Components

Recall that in the exact solution, the values of the tangential velocity components  $v$  and  $w$  do not change across the non-linear waves but do change, discontinuously, across the *middle* wave. Thus, given an approximate solution  $u_*$  for the normal velocity component in the *Star Region*, the solution for the tangential velocity components  $v$  and  $w$  is

$$v(x, t), w(x, t) = \begin{cases} v_L, w_L & \text{if } \frac{x}{t} \leq u_* , \\ v_R, w_R & \text{if } \frac{x}{t} > u_* . \end{cases} \quad (9.8)$$

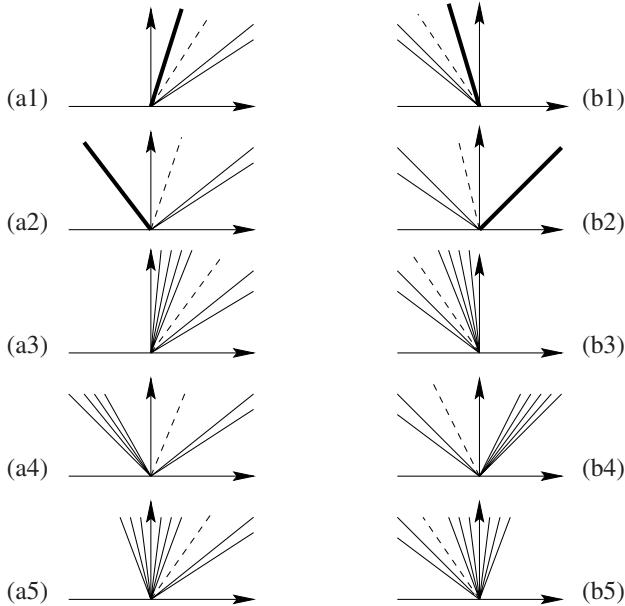
In this way, the solution for the tangential velocity components is, in a sense, *exact*; the only approximation being that for  $u_*$ . As a matter of fact, any *passive* scalar quantity  $q(x, y, z, t)$  advected with the fluid will have this property. In the study of multi-component flow, this quantity could be a species concentration; in practical applications there can be many of such quantities. Hence, it is very important that the approximate solution of the Riemann problem preserves the correct behaviour, as in (9.8).

### 9.2.2 Sonic Rarefactions

Assuming the solution for the star values (9.7) is available, we then need to identify the correct values along the  $t$ -axis, in order to evaluate the Godunov flux. The cases (a1) to (a4) and (b1) to (b4) of Fig. 9.2 can be dealt with once solutions for (9.7) and (9.8) are available. The *sonic flow* cases (a5) and (b5) must be treated separately. For these two cases we recommend the use of the exact solution, which, as seen in Sect. 4.4 of Chap. 4 for ideal gases, is non-iterative.

The solution along the  $t$ -axis inside a *left sonic rarefaction* is obtained by setting  $x/t = 0$  in

$$\mathbf{W}_{\text{Lfan}} = \begin{cases} \rho = \rho_L \left[ \frac{2}{(\gamma+1)} + \frac{(\gamma-1)}{(\gamma+1)a_L} (u_L - \frac{x}{t}) \right]^{\frac{2}{\gamma-1}}, \\ u = \frac{2}{(\gamma+1)} \left[ a_L + \frac{(\gamma-1)}{2} u_L + \frac{x}{t} \right], \\ p = p_L \left[ \frac{2}{(\gamma+1)} + \frac{(\gamma-1)}{(\gamma+1)a_L} (u_L - \frac{x}{t}) \right]^{\frac{2\gamma}{\gamma-1}}. \end{cases} \quad (9.9)$$



**Fig. 9.2.** Possible wave patterns in evaluating the Godunov flux for the Euler equations: (a) positive particle speed in the Star Region (b) negative particle speed in the Star Region

The solution along the  $t$ -axis inside a *right sonic rarefaction* is obtained by setting  $x/t = 0$  in

$$\mathbf{W}_{\text{Rfan}} = \begin{cases} \rho = \rho_R \left[ \frac{2}{(\gamma+1)} - \frac{(\gamma-1)}{(\gamma+1)a_R} (u_R - \frac{x}{t}) \right]^{\frac{2}{\gamma-1}}, \\ u = \frac{2}{(\gamma+1)} \left[ -a_R + \frac{(\gamma-1)}{2} u_R + \frac{x}{t} \right], \\ p = p_R \left[ \frac{2}{(\gamma+1)} - \frac{(\gamma-1)}{(\gamma+1)a_R} (u_R - \frac{x}{t}) \right]^{\frac{2\gamma}{\gamma-1}}. \end{cases} \quad (9.10)$$

The rest of this chapter is devoted to providing approximate solutions for the *star values* (9.7). We study four approaches as well as two adaptive schemes that combine various approximations.

### 9.3 Primitive Variable Riemann Solvers (PVRS)

A very simple linearised solution to the Riemann problem [502] for the  $x$ -split, three dimensional time dependent Euler equations (9.4)–(9.5) can be obtained in terms of the primitive variables  $\rho, u, v, w, p$ . The corresponding governing equations, see Sect. 3.2.3 of Chap. 3, are

$$\mathbf{W}_t + \mathbf{A}(\mathbf{W})\mathbf{W}_x = \mathbf{0}, \quad (9.11)$$

where the coefficient matrix  $\mathbf{A}(\mathbf{W})$  is given by

$$\mathbf{A} = \begin{bmatrix} u & \rho & 0 & 0 & 0 \\ 0 & u & 0 & 0 & 1/\rho \\ 0 & 0 & u & 0 & 0 \\ 0 & 0 & 0 & u & 0 \\ 0 & \rho a^2 & 0 & 0 & u \end{bmatrix}. \quad (9.12)$$

The eigenvalues of  $\mathbf{A}(\mathbf{W})$  are

$$\lambda_1 = u - a, \lambda_2 = \lambda_3 = \lambda_4 = u, \lambda_5 = u + a \quad (9.13)$$

and the matrix of corresponding right eigenvectors is

$$\mathbf{K} = \begin{bmatrix} \rho & 1 & \rho & \rho & \rho \\ -a & 0 & 0 & 0 & a \\ 0 & v & 1 & v & 0 \\ 0 & w & w & 1 & 0 \\ \rho a^2 & 0 & 0 & 0 & \rho a^2 \end{bmatrix}. \quad (9.14)$$

The difficulty in solving (9.11) is due to the fact that the coefficient matrix  $\mathbf{A}(\mathbf{W})$  depends on the solution vector  $\mathbf{W}$  itself. If  $\mathbf{A}(\mathbf{W})$  were to be *constant*, then we could apply, directly, the various techniques studied in Sect. 2.3.3 of Chap. 2 for solving linear hyperbolic systems with constant coefficients.

Assume that the initial data  $\mathbf{W}_L, \mathbf{W}_R$  and the solution  $\mathbf{W}(x/t)$  are *close* to a constant state  $\bar{\mathbf{W}}$ . Then, by setting

$$\bar{\mathbf{A}} \equiv \mathbf{A}(\bar{\mathbf{W}}) \quad (9.15)$$

we approximate the Riemann problem for (9.11) by the Riemann problem for the linear hyperbolic systems with constant coefficients

$$\mathbf{W}_t + \bar{\mathbf{A}}\mathbf{W}_x = \mathbf{0}. \quad (9.16)$$

We now solve this *approximate problem*, with initial data (9.6), *exactly*. In Sect. 2.3.3 of Chap. 2 we studied various techniques that are directly applicable to this problem. One possibility is to apply Rankine–Hugoniot Conditions across each wave of speed  $\bar{\lambda}_i$ . Thus we treat (9.16) as the system in *conservative form*

$$\mathbf{W}_t + \mathbf{F}(\mathbf{W})_x = \mathbf{0}, \quad \mathbf{F}(\mathbf{W}) \equiv \bar{\mathbf{A}}\mathbf{W}. \quad (9.17)$$

Then, across a wave of speed  $\bar{\lambda}_i$  we have

$$\Delta \mathbf{F} \equiv \bar{\mathbf{A}} \Delta \mathbf{W} = \bar{\lambda}_i \Delta \mathbf{W}. \quad (9.18)$$

Direct application of (9.18) to the  $\bar{\lambda}_1$  and  $\bar{\lambda}_5$  waves gives four useful relations, namely

$$\left. \begin{aligned} (u_* - u_L)\bar{\rho} + \bar{a}(\rho_{*L} - \rho_L) &= 0, \\ (p_* - p_L)/\bar{\rho} + \bar{a}(u_* - u_L) &= 0, \\ (u_R - u_*)\bar{\rho} - \bar{a}(\rho_R - \rho_{*R}) &= 0, \\ (p_R - p_*)/\bar{\rho} - \bar{a}(u_R - u_*) &= 0. \end{aligned} \right\} \quad (9.19)$$

The complete solution for the unknowns (9.7) is then given by

$$\left. \begin{aligned} p_* &= \frac{1}{2}(p_L + p_R) + \frac{1}{2}(u_L - u_R)(\bar{\rho}\bar{a}), \\ u_* &= \frac{1}{2}(u_L + u_R) + \frac{1}{2}(p_L - p_R)/(\bar{\rho}\bar{a}), \\ \rho_{*L} &= \rho_L + (u_L - u_*)(\bar{\rho}/\bar{a}), \\ \rho_{*R} &= \rho_R + (u_* - u_R)(\bar{\rho}/\bar{a}). \end{aligned} \right\} \quad (9.20)$$

Notice that in this linearised solution we only need to specify constant values for  $\bar{\rho}$  and  $\bar{a}$ . There is some freedom in making the choice. Selecting some average of the data values  $\rho_L$ ,  $\rho_R$ ,  $a_L$ ,  $a_R$  appears sensible. The choice may be constrained to satisfy some desirable properties of the Riemann problem solution, such as exact recognition of particular flow features. Here we suggest to select the simple *arithmetic means*

$$\bar{\rho} = \frac{1}{2}(\rho_L + \rho_R), \quad \bar{a} = \frac{1}{2}(a_L + a_R). \quad (9.21)$$

Note that if the data states  $\mathbf{W}_L$  and  $\mathbf{W}_R$  are connected by a *single isolated* contact discontinuity or shear wave, then the solution is actually *exact*, regardless of the particular choice for the averages  $\bar{\rho}$  and  $\bar{a}$ . This is in fact a very important property; contacts and shear waves turn out to be some of the most challenging flow features to resolve correctly by any numerical method.

Another way of obtaining approximate solutions for the star values is to use the *characteristic equations*, see Sect. 3.1.2 of Chap. 3,

$$dp - \rho a du = 0 \text{ along } dx/dt = u - a, \quad (9.22)$$

$$dp - a^2 d\rho = 0 \text{ along } dx/dt = u, \quad (9.23)$$

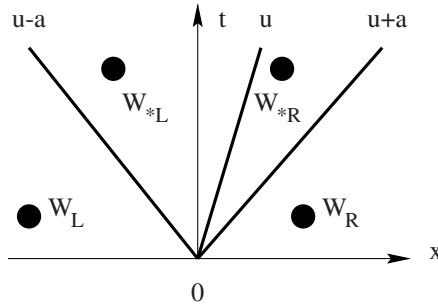
$$dp + \rho a du = 0 \text{ along } dx/dt = u + a. \quad (9.24)$$

These differential relations hold true *along characteristic* directions. First we set

$$C = \rho a. \quad (9.25)$$

Then, in order to find the star values we *connect* the state  $\mathbf{W}_{*L}$  to the data state  $\mathbf{W}_L$ , see Fig. 9.1, by integrating (9.24) along the characteristic of speed  $u + a$ , where  $C$  is evaluated at the foot of the characteristic. See Fig. 9.3. The results is

$$p_* + C_L u_* = p_L + C_L u_L. \quad (9.26)$$



**Fig. 9.3.** Solution for star values using characteristic equations

Similarly, we connect  $\mathbf{W}_{*R}$  to the data state  $\mathbf{W}_R$  by integrating (9.22) along the characteristic of speed  $u - a$ , with  $C$  is evaluated at the foot of the characteristic. We obtain

$$p_* - C_R u_* = p_R - C_R u_R . \quad (9.27)$$

The values  $\rho_{*L}$  and  $\rho_{*R}$  are obtained by connecting  $\mathbf{W}_{*L}$  to  $\mathbf{W}_L$  and  $\mathbf{W}_{*R}$  to  $\mathbf{W}_R$  via (9.23). The complete solution is

$$\left. \begin{aligned} p_* &= \frac{1}{C_L + C_R} [C_R p_L + C_L p_R + C_L C_R (u_L - u_R)] , \\ u_* &= \frac{1}{C_L + C_R} [C_L u_L + C_R u_R + (p_L - p_R)] , \\ \rho_{*L} &= \rho_L + (p_* - p_L)/a_L^2 , \\ \rho_{*R} &= \rho_R + (p_* - p_R)/a_R^2 . \end{aligned} \right\} \quad (9.28)$$

In this approximation we do not need to make a choice for the averages  $\bar{\rho}$  and  $\bar{a}$ ; their values are replaced by data values at the foot of the corresponding characteristics. If  $C_L = C_R = \bar{\rho}\bar{a}$ , then the two approximations (9.20) and (9.28) are identical.

The two linearised approximations (9.20) and (9.28) for the star values are exceedingly simple and may be useful in a variety of ways. The approaches might prove very useful in solving the Riemann problem for complicated sets of equations.

We have now given the complete approximate solution to the sub-problems (9.7)–(9.10). In order to evaluate the Godunov flux (9.3) we need to *sample* the solution to find the value  $\mathbf{W}_{i+\frac{1}{2}}(0)$  along the  $t$ -axis. This sampling procedure is virtually identical, although simpler, to the sampling procedure for the exact Riemann problem solution presented in Chap. 4. The reader is made aware that the numerical schemes associated with the simple linearised solutions just derived may not be robust enough to be used with absolute confidence under *all flow conditions*. In Sect. 9.5 we study hybrid Riemann solvers, which combine simple and sophisticated solvers to provide schemes that have effectively

the computational cost of the simplest Riemann solvers and the robustness of the sophisticated Riemann solvers.

## 9.4 Approximations Based on the Exact Solver

In Chap. 4 we presented an exact Riemann solver based on the pressure equation

$$f(p) \equiv f_L(p, \mathbf{W}_L) + f_R(p, \mathbf{W}_R) + \Delta u = 0 , \quad \Delta u = u_R - u_L , \quad (9.29)$$

with

$$f_K(p) = \begin{cases} (p - p_K) \left[ \frac{A_K}{p+B_K} \right]^{\frac{1}{2}} & \text{if } p > p_K \text{ (shock)}, \\ \frac{2a_K}{(\gamma-1)} \left[ \left( \frac{p}{p_K} \right)^z - 1 \right] & \text{if } p \leq p_K \text{ (rarefaction)}, \end{cases} \quad (9.30)$$

$$z = \frac{\gamma-1}{2\gamma} , \quad A_K = \frac{2}{(\gamma+1)\rho_K} , \quad B_K = \left( \frac{\gamma-1}{\gamma+1} \right) p_K , \quad K = L, R . \quad (9.31)$$

Various approximations based on  $f(p) = 0$  can be obtained, including curve-fitting procedures [509]. Here we give approximations based on the rarefaction and shock branches (9.30) of  $f(p)$ .

### 9.4.1 A Two-Rarefaction Riemann Solver (TRRS)

Recall that the non-linear waves in the Riemann problem solution are either shock or rarefaction waves and finding their type is part of the solution procedure. If one assumes *a-priori* that both non-linear waves are rarefactions then (9.29), with the appropriate choice of  $f_L$  and  $f_R$  in (9.30), becomes

$$\frac{2a_L}{(\gamma-1)} \left[ \left( \frac{p}{p_L} \right)^z - 1 \right] + \frac{2a_R}{(\gamma-1)} \left[ \left( \frac{p}{p_R} \right)^z - 1 \right] + u_R - u_L = 0 .$$

Solving this equation for pressure  $p_*$  gives the approximation

$$p_* = \left[ \frac{a_L + a_R - \frac{\gamma-1}{2}(u_R - u_L)}{a_L/p_L^z + a_R/p_R^z} \right]^{\frac{1}{z}} . \quad (9.32)$$

Having found  $p_*$  one can obtain the particle velocity  $u_*$  from any of the rarefaction wave relations

$$u_* = u_L - \frac{2a_L}{(\gamma-1)} \left[ \left( \frac{p_*}{p_L} \right)^z - 1 \right] \quad (9.33)$$

or

$$u_* = u_R + \frac{2a_R}{(\gamma - 1)} \left[ \left( \frac{p_*}{p_R} \right)^z - 1 \right]. \quad (9.34)$$

Alternatively, one can eliminate  $p_*$  from (9.33) and (9.34) to obtain a closed-form solution for the particle velocity

$$u_* = \frac{P_{LR}u_L/a_L + u_R/a_R + 2(P_{LR} - 1)/(\gamma - 1)}{P_{LR}/a_L + 1/a_R}, \quad P_{LR} = \left( \frac{p_L}{p_R} \right)^z. \quad (9.35)$$

Computing  $p_*$  from (9.32) requires the evaluation of 3 fractional powers. A more efficient approach is to calculate  $u_*$  from (9.35), which only requires one fractional power, and then evaluate  $p_*$  from (9.33) or (9.34), or from a mean value as

$$p_* = \frac{1}{2} \left\{ p_L \left[ 1 + \frac{(\gamma - 1)}{2a_L} (u_L - u_*) \right]^{\frac{1}{z}} + p_R \left[ 1 + \frac{(\gamma - 1)}{2a_R} (u_* - u_R) \right]^{\frac{1}{z}} \right\}. \quad (9.36)$$

Being consistent with the assumption that the two nonlinear waves are rarefaction waves, the computation of the densities  $\rho_{*L}$  and  $\rho_{*R}$  on either side of the contact discontinuity is obtained from the isentropic law, see Sect. 3.1.2 of Chap. 3. The result is

$$\rho_{*L} = \rho_L \left( \frac{p_*}{p_L} \right)^{\frac{1}{\gamma}}, \quad \rho_{*R} = \rho_R \left( \frac{p_*}{p_R} \right)^{\frac{1}{\gamma}}. \quad (9.37)$$

An improved version of the two-rarefaction solution is obtained by using exact relations, for given  $p_*$  or  $u_*$ . For instance, suppose  $p_*$  is given by (9.32) say, then  $u_*$  can be found from

$$u_* = \frac{1}{2}(u_L + u_R) + \frac{1}{2} [f_R(p_*) - f_L(p_*)], \quad (9.38)$$

where the functions  $f_L$  and  $f_R$  are evaluated according to the exact relations (9.30) by comparing  $p_*$  with  $p_L$  and  $p_R$ . The densities  $\rho_{*L}$  and  $\rho_{*R}$  can be found from the isentropic law if the  $K$  wave is a rarefaction ( $p_* \leq p_K$ ) or from the shock relation if the  $K$  wave is a shock wave ( $p_* > p_K$ ).

The two-rarefaction approximation is generally quite robust; it is more accurate, although more expensive, than the simple PVRS solutions (9.20) or (9.28) of the previous section. The TRRS is in fact exact when both nonlinear waves are actually rarefaction waves. This can be detected a-priori by the condition

$$f(p_{min}) \geq 0 \text{ with } p_{min} = \min(p_L, p_R). \quad (9.39)$$

See Sect. 4.3 of Chap. 4 for details on the behaviour of the pressure function.

We have now given another approximate solution to the problem (9.7). The solution for (9.9)–(9.10) follows. The evaluation of the Godunov flux (9.3) requires sampling the solution to find the value  $\mathbf{W}_{i+\frac{1}{2}}(0)$  along the  $t$ -axis, in the usual way. See Sect. 4.5 of Chap. 4.

### 9.4.2 A Two-Shock Riemann Solver (TSRS)

By assuming that *both non-linear waves are shock waves* in (9.29)–(9.30) one can derive the two-shock approximation

$$f(p) = (p - p_L)g_L(p) + (p - p_R)g_R(p) + u_R - u_L = 0 , \quad (9.40)$$

with

$$g_K(p) = \left[ \frac{A_K}{p + B_K} \right]^{\frac{1}{2}} \quad (9.41)$$

and  $A_K, B_K$  given by (9.31). Unfortunately, this approximation does not lead to a closed-form solution. Further approximations must be constructed [168], [384], [509]. Obvious approximations to the two-shock approximation involve quadratic equations. These do not generally lead to robust schemes. One difficulty is the non-uniqueness of solutions and making the correct choice; the exact solution, as seen in Chap. 4, is unique. The other problem is the case of complex roots (non-existence) even for data for which the exact problem has a solution; in our experience these can occur very often and is therefore the most serious difficulty of the two-shock approach.

An alternative approach [509] is as follows. First we assume an estimate  $p_0$  for the solution for pressure. Then we insert this estimate in the functions (9.41), which in turn are substituted into equation (9.40). We obtain

$$(p - p_L)g_L(p_0) + (p - p_R)g_R(p_0) + u_R - u_L = 0 .$$

The solution of this equation is immediate:

$$p_* = \frac{g_L(p_0)p_L + g_R(p_0)p_R - (u_R - u_L)}{g_L(p_0) + g_R(p_0)} . \quad (9.42)$$

Being consistent with the two-shock assumption we derive a solution for the velocity  $u_*$  as

$$u_* = \frac{1}{2}(u_L + u_R) + \frac{1}{2}[(p_* - p_R)g_R(p_0) - (p_* - p_L)g_L(p_0)] . \quad (9.43)$$

Solution values for  $\rho_{*L}$  and  $\rho_{*R}$  obtained from shock relations, see Sect. 3.1.3 of Chap. 3, namely

$$\rho_{*L} = \rho_L \left[ \frac{\frac{p_*}{p_L} + \frac{(\gamma-1)}{(\gamma+1)}}{\frac{(\gamma-1)}{(\gamma+1)} \frac{p_*}{p_L} + 1} \right] , \quad \rho_{*R} = \rho_R \left[ \frac{\frac{p_*}{p_R} + \frac{(\gamma-1)}{(\gamma+1)}}{\frac{(\gamma-1)}{(\gamma+1)} \frac{p_*}{p_R} + 1} \right] . \quad (9.44)$$

As to the choice for the pressure estimate  $p_0$  we propose

$$p_0 = \max(0, p_{pvrs}) , \quad (9.45)$$

where  $p_{pvrs}$  is the solution (9.20) for pressure given by the PVRS solver of Sect. 9.3.

We have just presented another approximate solution to the problem (9.7). As before, the solution for (9.8)–(9.10) follows. The evaluation of the Godunov flux (9.3) requires sampling the solution to find the value  $\mathbf{W}_{i+\frac{1}{2}}(0)$  along the  $t-$  axis, in the usual way. See Sect. 4.5 of Chap. 4.

The approximation (9.42)–(9.44) to the star values (9.7) is more efficient than the TRRS and only slightly more expensive than the PVRS approximations. Also TSRS is more accurate than TRRS and PVRS for a wider range of flow conditions, except for *near vacuum conditions*, where TRRS is very accurate or indeed exact. As for the case of the TRRS approximation, we can improve the TSRS by using the true wave relations whenever possible. For instance, for given  $p_*$  as computed from (9.42), one can obtain  $u_*$ ,  $\rho_{*L}$  and  $\rho_{*R}$  from exact wave relations. This is bound to improve the accuracy of the derived quantities.

## 9.5 Adaptive Riemann Solvers

In a typical flow field the overwhelming majority of local Riemann problems are a representation for smooth flow. Large gradients occur only near shock waves, contact surfaces, shear waves or some other sharp flow features. Large gradients generate Riemann problems with *widely different data states*  $\mathbf{W}_L, \mathbf{W}_R$ . Generally, it is in this kind of situations where approximate Riemann solvers can be fatally inaccurate, leading to failure of the numerical method being used. The rationale behind the use of hybrid schemes is *the use of simple Riemann solvers in regions of smooth flow and near isolated contacts and shear waves, and more sophisticated Riemann solvers elsewhere, in an adaptive fashion*.

Successful implementations of adaptive schemes involving the PVRS and the exact Riemann solvers were presented in [502] for the two-dimensional, time dependent Euler equations. Toro and Chou [533] extended the idea to the case of the steady supersonic Euler equations. Quirk [400] implemented this Riemann–solver adaptation approach in a MUSCL–type scheme, used in conjunction with adaptive mesh refinement techniques.

Here we present two hybrid schemes to compute the star values (9.7). Problems (9.8)–(9.10) are solved as before and the sampling is handled as described in Sect. 4.5 of Chap. 4.

### 9.5.1 An Adaptive Iterative Riemann Solver (AIRS)

This adaptive scheme makes use of two Riemann solvers: any of the primitive–variable Riemann solvers PVRS of Sect. 9.3 and the exact Riemann solver of Chap. 4. The PVRS scheme is used if the following two conditions are met:

$$Q = p_{\max}/p_{\min} < Q_{\text{user}} \quad (9.46)$$

and

$$p_{\min} < p_* < p_{\max}, \quad (9.47)$$

where

$$p_{\min} \equiv \min(p_L, p_R), \quad p_{\max} \equiv \max(p_L, p_R), \quad p_* \equiv p_{pvr} . \quad (9.48)$$

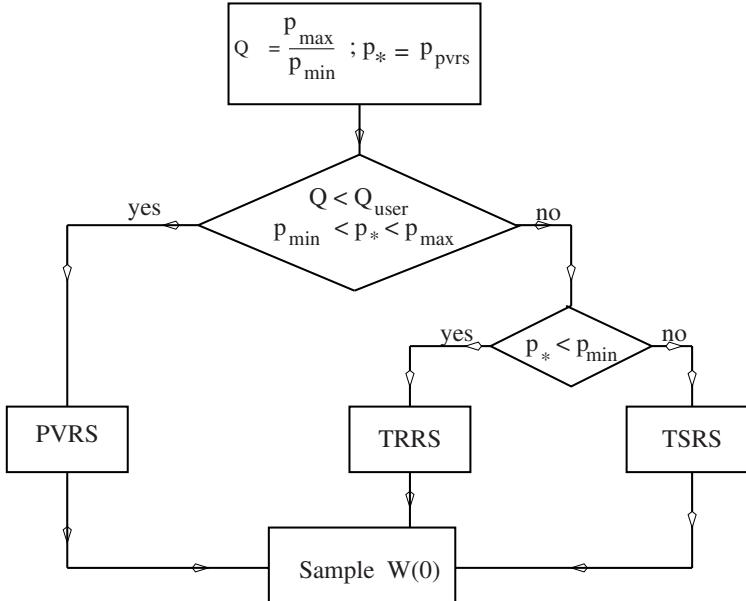
Otherwise, the exact Riemann solver is used.

Some remarks on the switching conditions (9.46)–(9.47) are in order. Condition (9.46) ensures that the pressure data values  $p_L, p_R$  are not widely different. Condition (9.47) imposes an extra restriction on the use of PVRS. The pressure restriction (9.46) is not sufficient; in fact for  $Q \approx 1, (p_L \approx p_R)$  and  $\Delta u = u_R - u_L$  negative and large in absolute value, strong shock waves are present in the solution of the Riemann problem, that is  $p_* > p_{\max}$ . For  $\Delta u$  large and positive  $p_* < p_{\min}$  and strong rarefactions are present in the exact solution of the Riemann problem. Condition (9.47) is effectively a condition on  $\Delta u$  and excludes the two–rarefaction and the two–shock cases; both of these cases occur naturally at reflected boundaries, where it would be unwise to employ unreliable approximations. Also, these two cases are inconsistent with condition (9.46) on pressure ratios.

A choice of the switching parameter  $Q_{\text{user}}$  is to be made. Extensive testing suggests that the value  $Q_{\text{user}} = 2$  is perfectly adequate to give both very robust and efficient schemes. Even much larger values of  $Q_{\text{user}}$  can give accurate solutions, but the gains are not significant and thus the cautious choice of  $Q_{\text{user}} = 2$  is recommended. For typical flow conditions and meshes, over 90% of all Riemann problems are handled by the cheap linearised Riemann solver. Effectively, the resulting schemes have the efficiency of the cheapest Riemann solvers and the robustness of the exact Riemann solver. A disadvantage of this hybrid PVRS–EXACT scheme is the iterative character of the robust component of the scheme, namely the exact Riemann solver. This may be inconvenient for some computer architectures. One possibility here is to fix the number of iterations in the exact Riemann solver. In our experience, *one iteration* leads to very accurate values for pressure and subsequent quantities derived. This is due in part to the provision of a sophisticated starting value for the iteration procedure.

### 9.5.2 An Adaptive Noniterative Riemann Solver (ANRS)

Here we propose to combine a PVRS scheme, as the cheap component, together with the non–iterative TRRS and TSRS solvers of Sects. 9.4.1 and 9.4.2 to provide the robust component of the adaptive scheme. The use of PVRS is again restricted by conditions (9.46)–(9.47) of the previous scheme. If any of conditions (9.46) or (9.47) are not met we use either TRRS or TSRS. The switching between TRRS and TSRS is motivated by the behaviour of the exact function for pressure, see Sect. 4.3.1 of Chap. 4, and is as follows. If



**Fig. 9.4.** Flow chart for Adaptive Noniterative Riemann Solver (ANRS) involving PVRS, TRRS and TSRS

$$p_{\text{pvrs}} \leq p_{\min}, \quad (9.49)$$

then we use TRRS, otherwise we use TSRS. The flow chart of Fig. 9.4 illustrates the implementation of this adaptive scheme. The problems of computing the tangential velocity components, handling sonic flow and the sampling procedure to find the Godunov flux are dealt with as described in the previous sections. This adaptive noniterative Riemann solver is recommended for practical applications.

## 9.6 Numerical Results

Here we assess the performance of Godunov's first-order upwind method used in conjunction with the approximate Riemann solvers presented in this chapter. We select five test problems for the one-dimensional, time dependent Euler equations for ideal gases with  $\gamma = 1.4$ ; these have exact solutions, which are evaluated by running the code HE-E1RPEXACT of NUMERICA [519].

In all chosen tests, data consists of two constant states  $\mathbf{W}_L = [\rho_L, u_L, p_L]^T$  and  $\mathbf{W}_R = [\rho_R, u_R, p_R]^T$ , separated by a discontinuity at a position  $x = x_0$ . The states  $\mathbf{W}_L$  and  $\mathbf{W}_R$  are given in Table 9.1. The ratio of specific heats is chosen to be  $\gamma = 1.4$ . For all test problems the spatial domain is the interval  $[0, 1]$  which is discretised with  $M = 100$  computing cells. The Courant number

coefficient is  $C_{\text{cfl}} = 0.9$ ; boundary conditions are transmissive and  $S_{\max}^n$  is found using the simplified formula (6.20) of Chapt. 6. But given that this formula is somewhat unreliable, see discussion of Sect. 6.3.2 of Chapter 6, in all computations presented here we take, for the the first 5 time steps, a Courant number coefficient  $C_{\text{cfl}}$  reduced by a factor of 0.2.

Test	$\rho_L$	$u_L$	$p_L$	$\rho_R$	$u_R$	$p_R$
1	1.0	0.75	1.0	0.125	0.0	0.1
2	1.0	-2.0	0.4	1.0	2.0	0.4
3	1.0	0.0	1000.0	1.0	0.0	0.01
4	5.99924	19.5975	460.894	5.99242	-6.19633	46.0950
5	1.0	-19.59745	1000.0	1.0	-19.59745	0.01

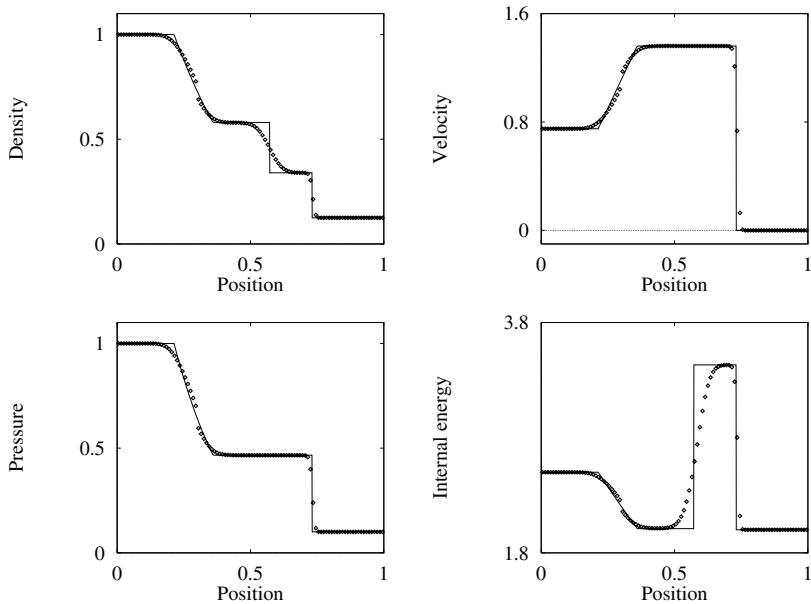
Table 9.1. Data for five test problems with exact solution. Test 5 is like Test 3 with negative uniform background speed

Test 1 is a *modified version* of the popular Sod's test [453]; the solution consists of a right shock wave, a right travelling contact wave and a left *sonic* rarefaction wave; this test is very useful in assessing the *entropy satisfaction* property of numerical methods. Test 2 has solution consisting of two symmetric rarefaction waves and a trivial contact wave of zero speed; the *Star Region* between the non-linear waves is close to vacuum, which makes this problem a suitable test for assessing the performance of numerical methods for low-density flows; this is the so called *123 problem* introduced in chapter Chap. 4. Test 3 is designed to assess the robustness and accuracy of numerical methods; its solution consists of a strong shock wave, a contact surface and a left rarefaction wave. Test 4 is also designed to test robustness of numerical methods; the solution consists of three strong discontinuities travelling to the right. See Sect. 4.3.3 of Chap. 4 for more details on the exact solution of these test problems. Test 5 is also designed to test the robustness of numerical methods but the main reason for devising this test is to assess the ability of the numerical methods to resolve *slowly-moving contact discontinuities*. The exact solution of Test 5 consists of a left rarefaction wave, a right-travelling shock wave and a *stationary* contact discontinuity. For each test we select a convenient position  $x_0$  of the initial discontinuity and an output time. These are stated in the legend of each figure displaying computational results.

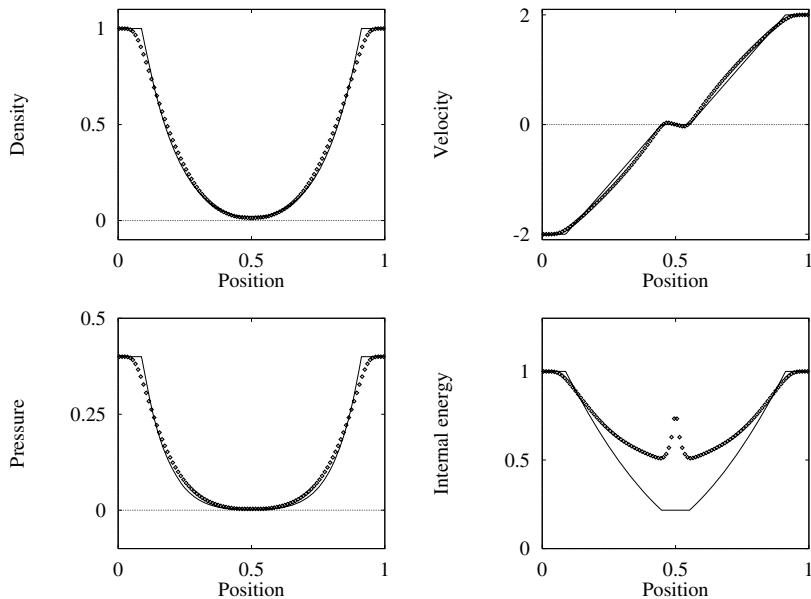
We present numerical results for two of the approximate Riemann solvers presented in this chapter, namely the Two-Shock Riemann solver (TSRS) and the Adaptive Noniterative Riemann Solver (ANRS). The numerical solutions are obtained by running the code HE-E1GODSTATE of *NUMERICA* [519]. The results from TSRS are shown in Figs. 9.5 to 9.9 and those of ANRS are shown in Figs. 9.10 to 9.14. All of these results are to be compared with those obtained from the Godunov scheme used in conjunction with the exact Riemann solver, see Figs. 6.8 to 6.12, Chapt. 6; to plotting accuracy, there is

no difference in the computed results. The two approximate Riemann solvers TSRS and ANRS are recommended for practical applications.

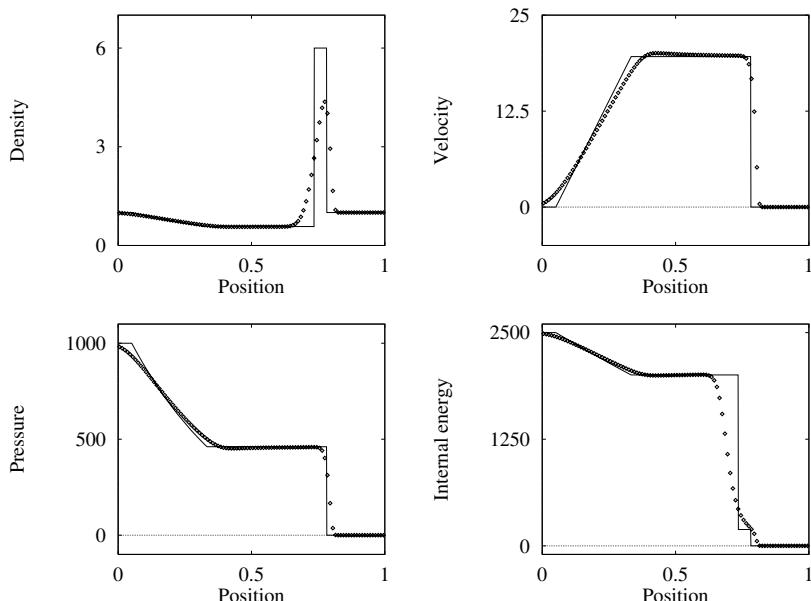
The Godunov-type methods based on the approximate-state Riemann solvers of this chapter are extended to second-order of accuracy using the techniques of Chaps. 13 and 14, for one-dimensional problems. Approaches for including source terms are given in Chapt. 15 and for solving multidimensional problems in Chap. 16.



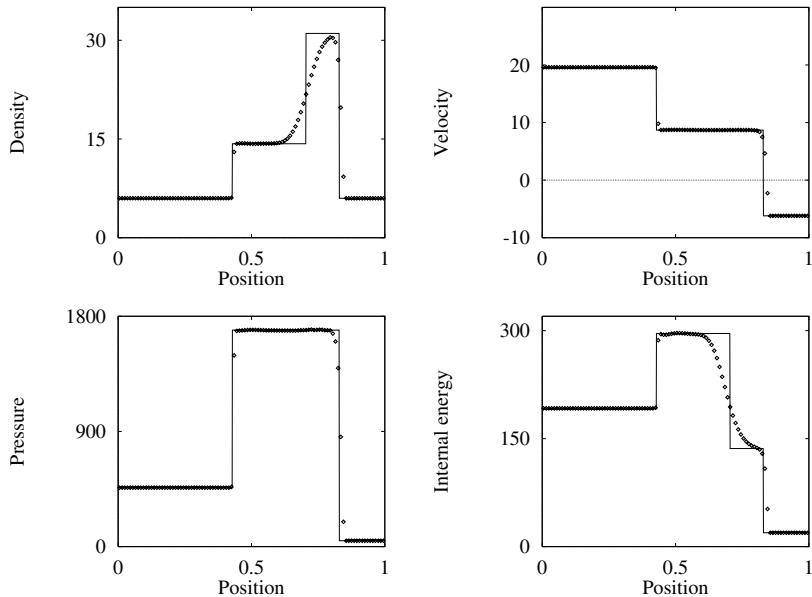
**Fig. 9.5.** Two-Shock Riemann Solver applied to Test 1, with  $x_0 = 0.3$ . Numerical (symbol) and exact (line) solutions are compared at time 0.2 units



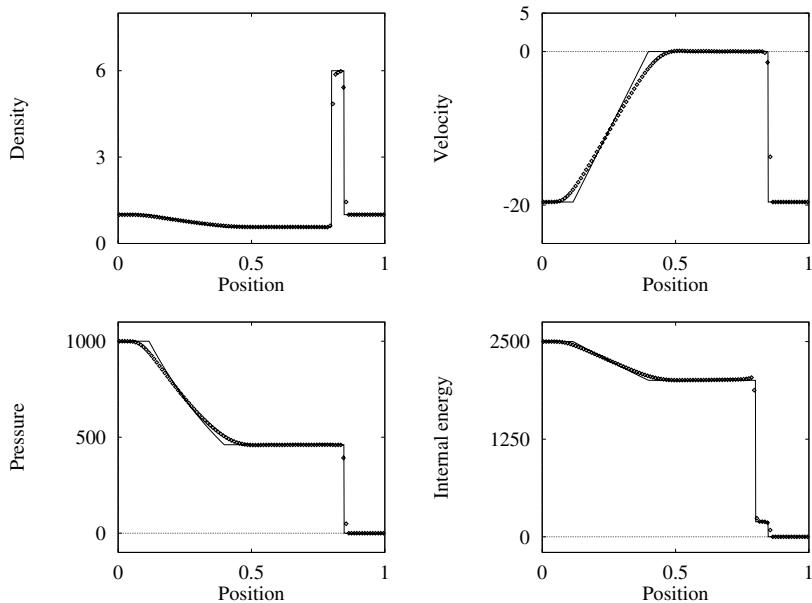
**Fig. 9.6.** Two-Shock Riemann Solver applied to Test 2, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at time 0.15 units



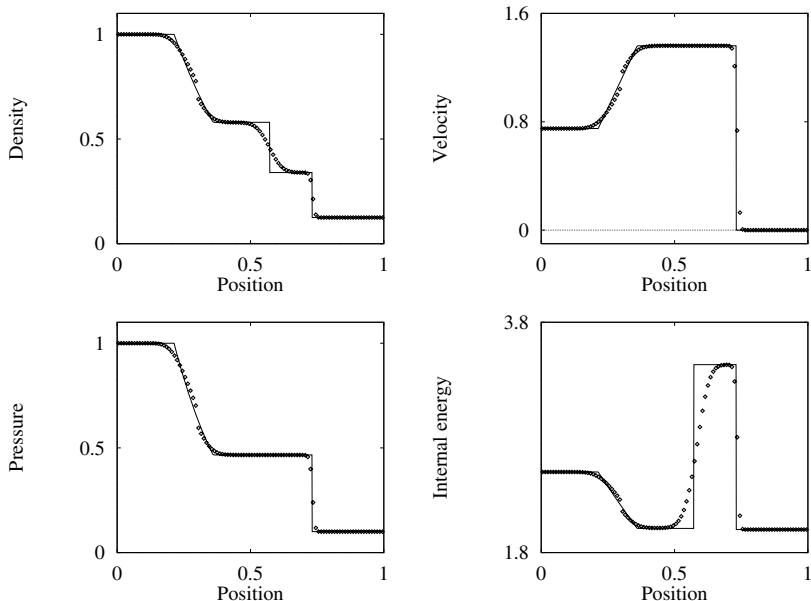
**Fig. 9.7.** Two-Shock Riemann Solver applied to Test 3, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012 units



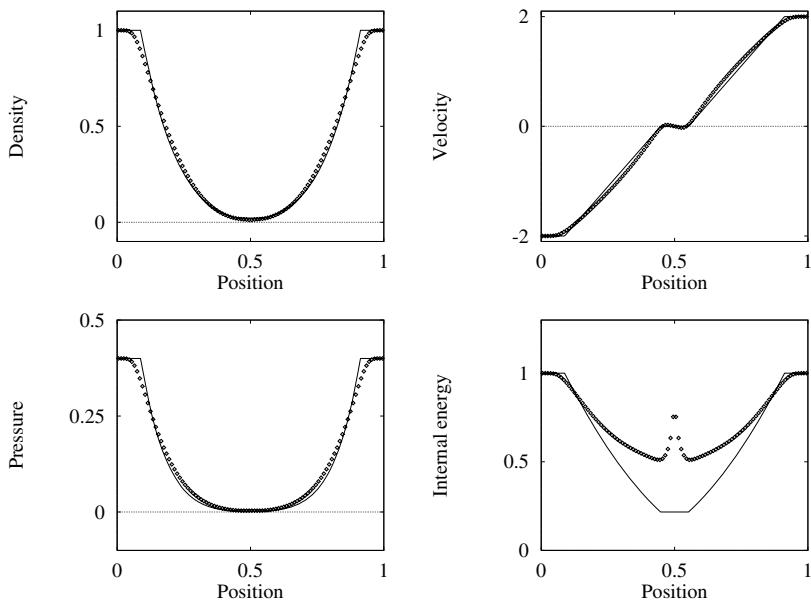
**Fig. 9.8.** Two-Shock Riemann Solver applied to Test 4, with  $x_0 = 0.4$ . Numerical (symbol) and exact (line) solutions are compared at time 0.035 units



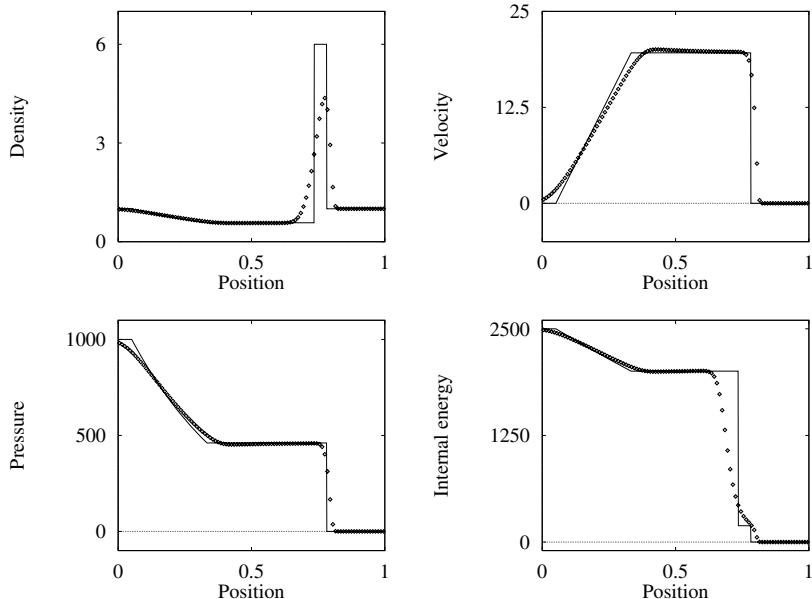
**Fig. 9.9.** Two-Shock Riemann Solver applied to Test 5, with  $x_0 = 0.8$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012 units



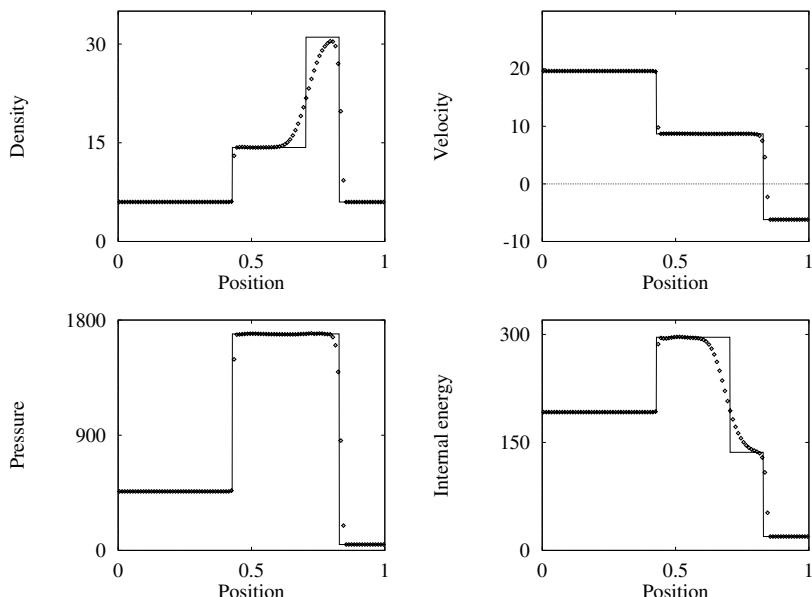
**Fig. 9.10.** Adaptive Noniterative Riemann Solver applied to Test 1, with  $x_0 = 0.3$ . Numerical (symbol) and exact (line) solutions are compared at time 0.2 units



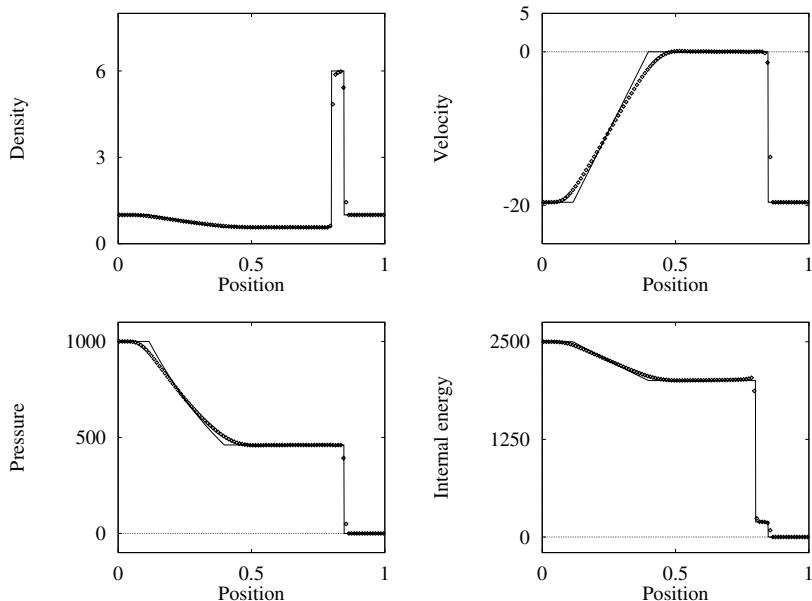
**Fig. 9.11.** Adaptive Noniterative Riemann Solver applied to Test 2, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at time 0.15 units



**Fig. 9.12.** Adaptive Noniterative Riemann Solver applied to Test 3, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012 units



**Fig. 9.13.** Adaptive Noniterative Riemann Solver applied to Test 4, with  $x_0 = 0.4$ . Numerical (symbol) and exact (line) solutions are compared at time 0.035 units



**Fig. 9.14.** Adaptive Noniterative Riemann Solver applied to Test 5, with  $x_0 = 0.8$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012 units

---

## The HLL and HLLC Riemann Solvers

The approximate Riemann solver proposed by Harten Lax and van Leer (HLL) in 1983 requires estimates for the fastest signal velocities emerging from the initial discontinuity at the interface, resulting in a *two-wave model* for the structure of the exact solution. A more accurate method is the HLLC, introduced by Toro and collaborators in 1992. This method assumes a *three-wave model*, resulting in better resolution of intermediate waves.

### 10.1 Introduction

For the purpose of computing a Godunov flux, Harten, Lax and van Leer [244] presented a novel approach for solving the Riemann problem approximately. The resulting Riemann solvers have become known as HLL Riemann solvers. In this approach an *approximation for the intercell numerical flux is obtained directly*, unlike the Riemann solvers presented previously in Chaps. 4 and 9. The central idea is to assume, for the solution, a wave configuration that consists of *two waves* separating three constant states. Assuming that the wave speeds are given by some algorithm, application of the integral form of the conservation laws gives a closed-form, approximate expression for the flux. The approach produced practical schemes after the contributions of Davis [150] and Einfeldt [181], who independently proposed various ways of computing the wave speeds required to completely determine the intercell flux. The two-wave HLL approach, along with the wave speed estimates proposed by Einfeldt [181] is known as the HLLE solver. The resulting HLL-type Riemann solvers form the bases of very efficient and robust approximate Godunov-type methods.

One difficulty with these schemes, however, is the assumption of a *two-wave* configuration. This is correct only for hyperbolic systems of two equations, such as the one-dimensional shallow water equations. For larger systems, such as the Euler equations or the split two-dimensional shallow water equations for example, the two-wave assumption is incorrect. As a con-

sequence the resolution of physical features such as contact surfaces, shear waves and material interfaces, can be very inaccurate. For the limiting case in which these features are stationary relative to the mesh, the resulting numerical smearing is unacceptable. In view of this situation Einfeldt proposed [181] a modification to the HLLE scheme, called HLLM, in which the single intermediate state in the HLL approach is modified by means of a linear distribution. The modification involves some parameters that control the amount of excessive dissipation for intermediate waves. Particular choices of these parameters and of the wave speed estimates reduce the HLLM scheme to a modified version of the Roe solver. See [182] for further details on both HLLE and HLLEM.

A different approach to remedy the problem of intermediate waves in the HLL approach waves was taken by Toro, Spruce and Speares in 1992 [541], [542]. They proposed the HLLC Riemann solver (C standing for Contact), as applied to the time-dependent Euler equations. HLLC is a three-wave model, resulting two-star states for the intermediate region of the Riemann-problem solution fan. A precursor to HLLC was also anticipated in [505]. Early applications of HLLC include the steady supersonic two-dimensional Euler equations [532] and the time-dependent two dimensional shallow water equations [193], [194]. Batten and collaborators [32] analyzed the HLLC scheme and proposed new ways of estimating the wave speeds. See also the work of Batten, Leschziner and Goldberg [33], in which they proposed implicit versions of the HLLC Riemann solver, with application to turbulent flows. In later work by Linde and others [324], [325] modifications to the HLL two-wave approach were also explored, in order to reduce numerical dissipation of contact waves.

In the last decade or so we have seen further developments of the HLLC method as well as ambitious applications. A quick electronic search, by typing for example *HLLC solver*, will give hundreds of useful entries on the subject. Recall that the Euler equations have *three* distinct characteristic fields in one, two and three space dimensions, see Chapter 3, section 3.2. This is why HLLC as proposed in [541], [542], [96] is a *complete* Riemann solver, for the Euler equations; that is the approximate wave structure of HLLC contains all the characteristic fields of the exact problem. However, for systems with eigen-structure containing more than three distinct characteristic fields, the HLLC becomes *incomplete*, tending to behave like HLL for the one-dimensional Euler equations. The incomplete character of a Riemann solver affects the resolution of intermediate waves, particularly when these move slowly relative to the mesh. Therefore, the obvious way of improving the HLLC approach is to admit the correct number of characteristic fields for the system of interest. Works along these lines include [230], [474] and [75]. Other interesting developments and ambitious applications are found in the following works, to name but a few, [24], [553], [54], [54], [74], [360], [580], [318], [397], [351], [334], [572], [6], [53], [285], [199], [255], [237], [382], [361], [86], [425], [602].

In this Chapter we present the HLL and HLLC Riemann solvers as applied to the three-dimensional, time dependent Euler equations. The principles can

easily be extended to solve other hyperbolic systems. Useful background reading is found in Chaps. 3, 4, 6 and 9. The rest of this chapter is organised as follows: Sect. 10.1 recalls the Riemann problem. In Sect. 10.3 we present the original approach of Harten, Lax and van Leer. In Sect. 10.4 we present the HLLC Riemann solver and in Sect. 10.5 we give various algorithms for computing the required wave speeds. A summary of the HLLC schemes is presented in Sect. 10.6. In Sect. 10.7 we analyse the behaviour of the approximate Riemann solvers in the presence of contacts and passive scalars. Numerical results are shown in Sect. 10.8 and in Sect. 10.9 contains some concluding remarks.

## 10.2 The Riemann Problem

We are concerned with solving numerically the general Initial Boundary Value Problem (IBVP)

$$\left. \begin{array}{l} \text{PDEs : } \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0} , \\ \text{ICs : } \mathbf{U}(x, 0) = \mathbf{U}^{(0)}(x) , \\ \text{BCs : } \mathbf{U}(0, t) = \mathbf{U}_l(t) , \quad \mathbf{U}(L, t) = \mathbf{U}_r(t) , \end{array} \right\} \quad (10.1)$$

in a domain  $0 \leq x \leq L$ , with appropriate boundary conditions. We use the explicit conservative formula

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \frac{\Delta t}{\Delta x} [\mathbf{F}_{i+\frac{1}{2}} - \mathbf{F}_{i-\frac{1}{2}}] , \quad (10.2)$$

with the numerical flux  $\mathbf{F}_{i+\frac{1}{2}}$  yet to be defined.

### 10.2.1 The Godunov Flux

In Chap. 6 we defined the Godunov intercell numerical flux as

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}(0)) , \quad (10.3)$$

in which  $\mathbf{U}_{i+\frac{1}{2}}(0)$  is the exact similarity solution  $\mathbf{U}_{i+\frac{1}{2}}(x/t)$  of the Riemann problem

$$\left. \begin{array}{l} \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0} , \\ \mathbf{U}(x, 0) = \begin{cases} \mathbf{U}_L & \text{if } x < 0 , \\ \mathbf{U}_R & \text{if } x > 0 , \end{cases} \end{array} \right\} \quad (10.4)$$

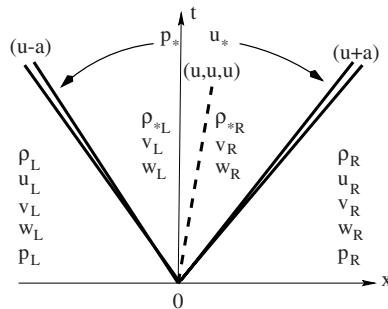
evaluated at  $x/t = 0$ . Fig. 10.1 shows the structure of the exact solution of the Riemann problem for the  $x$ -split, three dimensional Euler equations, for which the vectors of conserved variables and fluxes are

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ E \end{bmatrix} , \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ \rho uw \\ u(E + p) \end{bmatrix} . \quad (10.5)$$

The value  $x/t = 0$  for the Godunov flux corresponds to the  $t$ -axis. See Chaps. 4 and 6 for details. The piece-wise constant initial data, in terms of primitive variables, is

$$\mathbf{W}_L = \begin{bmatrix} \rho_L \\ u_L \\ v_L \\ w_L \\ p_L \end{bmatrix}, \quad \mathbf{W}_R = \begin{bmatrix} \rho_R \\ u_R \\ v_R \\ w_R \\ p_R \end{bmatrix}. \quad (10.6)$$

In Chap. 9 we provided approximations to the state  $\mathbf{U}_{i+\frac{1}{2}}(x/t)$  and obtained



**Fig. 10.1.** Structure of the exact solution of the Riemann problem for the  $x$ -split three dimensional Euler equations. There are five wave families associated with the eigenvalues  $u - a$ ,  $u$  (of multiplicity 3) and  $u + a$ .

a corresponding approximate Godunov method by evaluating the physical flux function  $\mathbf{F}$  at this approximate state; see (10.3). The purpose of this chapter is to find *direct approximations to the flux function*  $\mathbf{F}_{i+\frac{1}{2}}$  following the novel approach proposed by Harten, Lax and van Leer [238].

### 10.2.2 Integral Relations

Consider Fig. 10.2, in which the whole of the wave structure arising from the exact solution of the Riemann problem is contained in the control volume  $[x_L, x_R] \times [0, T]$ , that is

$$x_L \leq TS_L, \quad x_R \geq TS_R, \quad (10.7)$$

where  $S_L$  and  $S_R$  are the *fastest signal velocities* perturbing the initial data states  $\mathbf{U}_L$  and  $\mathbf{U}_R$  respectively, and  $T$  is a chosen time. The integral form of the conservation laws in (10.4), in the control volume  $[x_L, x_R] \times [0, T]$  reads

$$\int_{x_L}^{x_R} \mathbf{U}(x, T) dx = \int_{x_L}^{x_R} \mathbf{U}(x, 0) dx + \int_0^T \mathbf{F}(\mathbf{U}(x_L, t)) dt - \int_0^T \mathbf{F}(\mathbf{U}(x_R, t)) dt. \quad (10.8)$$

See Sect. 2.4.1 of Chap. 2 for details on integral forms of conservation laws. Evaluation of the right-hand side of this expression gives

$$\int_{x_L}^{x_R} \mathbf{U}(x, T) dx = x_R \mathbf{U}_R - x_L \mathbf{U}_L + T(\mathbf{F}_L - \mathbf{F}_R), \quad (10.9)$$

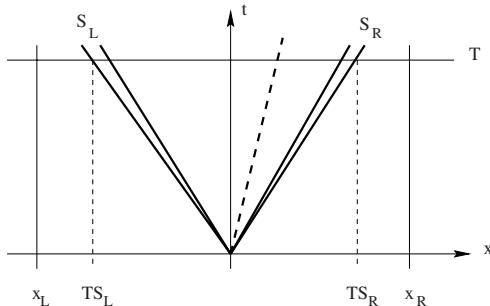
where  $\mathbf{F}_L = \mathbf{F}(\mathbf{U}_L)$  and  $\mathbf{F}_R = \mathbf{F}(\mathbf{U}_R)$ . We call the integral relation (10.9) the *consistency condition*. Now we split the integral on the left-hand side of (10.8) into three integrals, namely

$$\int_{x_L}^{x_R} \mathbf{U}(x, T) dx = \int_{x_L}^{TS_L} \mathbf{U}(x, T) dx + \int_{TS_L}^{TS_R} \mathbf{U}(x, T) dx + \int_{TS_R}^{x_R} \mathbf{U}(x, T) dx$$

and evaluate the first and third terms on the right-hand side. We obtain

$$\int_{x_L}^{x_R} \mathbf{U}(x, T) dx = \int_{TS_L}^{TS_R} \mathbf{U}(x, T) dx + (TS_L - x_L) \mathbf{U}_L + (x_R - TS_R) \mathbf{U}_R. \quad (10.10)$$

Comparing (10.10) with (10.9) gives



**Fig. 10.2.** Control volume  $[x_L, x_R] \times [0, T]$  on  $x$ - $t$  plane.  $S_L$  and  $S_R$  are the fastest signal velocities arising from the solution of the Riemann problem.

$$\int_{TS_L}^{TS_R} \mathbf{U}(x, T) dx = T(S_R \mathbf{U}_R - S_L \mathbf{U}_L + \mathbf{F}_L - \mathbf{F}_R). \quad (10.11)$$

On division through by the length  $T(S_R - S_L)$ , which is the width of the wave system of the solution of the Riemann problem between the slowest and fastest signals at time  $T$ , we have

$$\frac{1}{T(S_R - S_L)} \int_{TS_L}^{TS_R} \mathbf{U}(x, T) dx = \frac{S_R \mathbf{U}_R - S_L \mathbf{U}_L + \mathbf{F}_L - \mathbf{F}_R}{S_R - S_L}. \quad (10.12)$$

Thus, the integral average of the exact solution of the Riemann problem between the slowest and fastest signals at time  $T$  is a known constant, provided

that the signal speeds  $S_L$  and  $S_R$  are known; such constant is the right-hand side of (10.12) and we denote it by

$$\mathbf{U}^{hll} = \frac{S_R \mathbf{U}_R - S_L \mathbf{U}_L + F_L - F_R}{S_R - S_L}. \quad (10.13)$$

We now apply the integral form of the conservation laws to the left portion of Fig. 10.2, that is the control volume  $[x_L, 0] \times [0, T]$ . We obtain

$$\int_{TS_L}^0 \mathbf{U}(x, T) dx = -TS_L \mathbf{U}_L + T(\mathbf{F}_L - \mathbf{F}_{0L}), \quad (10.14)$$

where  $\mathbf{F}_{0L}$  is the flux  $\mathbf{F}(\mathbf{U})$  along the  $t$ -axis. Solving for  $\mathbf{F}_{0L}$  we find

$$\mathbf{F}_{0L} = \mathbf{F}_L - S_L \mathbf{U}_L - \frac{1}{T} \int_{TS_L}^0 \mathbf{U}(x, T) dx. \quad (10.15)$$

Evaluation of the integral form of the conservation laws on the control volume  $[0, x_R] \times [0, T]$  yields

$$\mathbf{F}_{0R} = \mathbf{F}_R - S_R \mathbf{U}_R + \frac{1}{T} \int_0^{TS_R} \mathbf{U}(x, T) dx. \quad (10.16)$$

The reader can easily verify that the equality

$$\mathbf{F}_{0L} = \mathbf{F}_{0R}$$

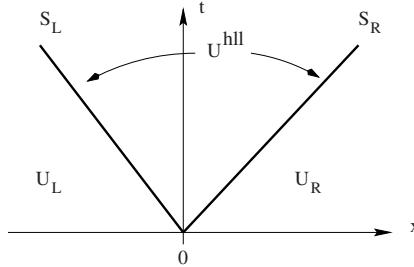
results in the *consistency condition* (10.9). All relations so far are exact, as we are assuming the exact solution of the Riemann problem.

### 10.3 The HLL Approximate Riemann Solver

Harten, Lax and van Leer [244] put forward the following approximate Riemann solver

$$\tilde{\mathbf{U}}(x, t) = \begin{cases} \mathbf{U}_L & \text{if } \frac{x}{t} \leq S_L, \\ \mathbf{U}^{hll} & \text{if } S_L \leq \frac{x}{t} \leq S_R, \\ \mathbf{U}_R & \text{if } \frac{x}{t} \geq S_R, \end{cases} \quad (10.17)$$

where  $\mathbf{U}^{hll}$  is the constant state vector given by (10.13) and the speeds  $S_L$  and  $S_R$  are assumed to be known. Fig. 10.3 shows the structure of this approximate solution of the Riemann problem, called the HLL Riemann solver. Note that this approximation consists of just three constant states separated by two waves. The *Star Region* consists of a *single* constant state; all intermediate states separated by intermediate waves are *lumped* into the single state  $\mathbf{U}^{hll}$ . The corresponding flux  $\mathbf{F}^{hll}$  along the  $t$ -axis is found from the relations (10.15) or (10.16), with the exact integrand replaced by the approximate solution (10.17). Note that we *do not take*  $\mathbf{F}^{hll} = \mathbf{F}(\mathbf{U}^{hll})$ . The non-trivial case of



**Fig. 10.3.** Approximate HLL Riemann solver. Solution in the *Star Region* consists of a single state  $\mathbf{U}^{hll}$  separated from data states by two waves of speeds  $S_L$  and  $S_R$ .

interest is the subsonic case  $S_L \leq 0 \leq S_R$ . Substitution of the integrand in (10.15) or (10.16) by  $\mathbf{U}^{hll}$  in (10.13) gives

$$\mathbf{F}^{hll} = \mathbf{F}_L + S_L(\mathbf{U}^{hll} - \mathbf{U}_L), \quad (10.18)$$

or

$$\mathbf{F}^{hll} = \mathbf{F}_R + S_R(\mathbf{U}^{hll} - \mathbf{U}_R). \quad (10.19)$$

Note that relations (10.18) and (10.19) are also obtained from applying Rankine–Hugoniot conditions across the left and right waves respectively; see Sect. 2.4.2 of Chap. 2 and Sect. 3.1.3 of Chap. 3 for details on the Rankine–Hugoniot conditions. Use of (10.13) in (10.18) or (10.19) gives the HLL flux

$$\mathbf{F}^{hll} = \frac{S_R \mathbf{F}_L - S_L \mathbf{F}_R + S_L S_R (\mathbf{U}_R - \mathbf{U}_L)}{S_R - S_L}. \quad (10.20)$$

The corresponding HLL intercell flux for the approximate Godunov method is then given by

$$\mathbf{F}_{i+\frac{1}{2}}^{hll} = \begin{cases} \mathbf{F}_L & \text{if } 0 \leq S_L, \\ \frac{S_R \mathbf{F}_L - S_L \mathbf{F}_R + S_L S_R (\mathbf{U}_R - \mathbf{U}_L)}{S_R - S_L}, & \text{if } S_L \leq 0 \leq S_R, \\ \mathbf{F}_R & \text{if } 0 \geq S_R. \end{cases} \quad (10.21)$$

Given an algorithm to compute the speeds  $S_L$  and  $S_R$  we have an approximate intercell flux (10.21) to be used in the conservative formula (10.2) to produce an approximate Godunov method. Procedures to estimate the wave speeds  $S_L$  and  $S_R$  are given in Sect. 10.5. Harten, Lax and van Leer [244] showed that the Godunov scheme (10.2), (10.21), if convergent, converges to the weak solution of the conservation laws. In fact they proved that the converged solution is also the physical, entropy satisfying, solution of the conservation laws. Their results actually apply to a larger class of approximate Riemann solvers. One of the requirements is *consistency with the integral form of the conservation laws*.

That is, an approximate solution  $\tilde{\mathbf{U}}(x, t)$  is consistent with the integral form of the conservation laws if, when substituted for the exact solution  $\mathbf{U}(x, t)$  in the consistency condition (10.9), the right-hand side remains unaltered.

A shortcoming of the HLL scheme is exposed by contact discontinuities, shear waves and material interfaces, or any type of *intermediate waves*. For the Euler equations these waves are associated with the multiple eigenvalue  $\lambda_2 = \lambda_3 = \lambda_4 = u$ . See Fig. 10.1. Note that in the integral (10.12), all that matters is the average across the wave structure, without regard for the spatial variations of the solution of the Riemann problem in the *Star Region*. As pointed out by Harten, Lax and van Leer themselves [244], this defect of the HLL scheme may be corrected by restoring the missing waves. Accordingly, Toro, Spruce and Speares [541], [542] proposed the so called *HLLC scheme*, where C stands for *Contact*. In this scheme the missing middle waves are put back into the structure of the approximate Riemann solver.

## 10.4 The HLLC Approximate Riemann Solver

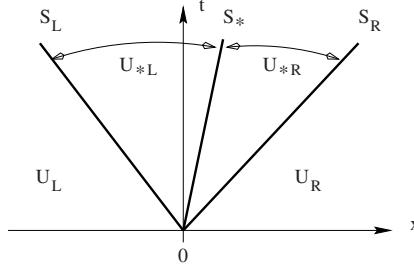
The HLLC scheme is a modification of the HLL scheme described in the previous section, whereby the missing contact and shear waves in the Euler equations are restored. The scheme was first presented in terms of the time-dependent, two dimensional Euler equations [541], [542]. Early applications include the steady supersonic two-dimensional Euler equations [532] and the time-dependent two dimensional shallow water equations [193], [194].

### 10.4.1 Useful Relations

Consider Fig. 10.2, in which the complete structure of the solution of the Riemann problem is contained in a sufficiently large control volume  $[x_L, x_R] \times [0, T]$ . Now, in addition to the slowest and fastest signal speeds  $S_L$  and  $S_R$  we include a middle wave of speed  $S_*$ ; for the Euler equations this corresponds to the multiple eigenvalue  $\lambda_2 = \lambda_3 = \lambda_4 = u$ . See Fig. 10.4. Evaluation of the integral form of the conservation laws in the control volume reproduces the result of equation (10.12), even if variations of the integrand across the wave of speed  $S_*$  are allowed. Note that the consistency condition (10.9) effectively becomes the condition (10.12). By splitting the left-hand side of integral (10.12) into two terms we obtain

$$\left. \begin{aligned} \frac{1}{T(S_R - S_L)} \int_{TS_L}^{TS_R} \mathbf{U}(x, T) dx &= \frac{1}{T(S_R - S_L)} \int_{TS_L}^{TS_*} \mathbf{U}(x, T) dx \\ &\quad + \frac{1}{T(S_R - S_L)} \int_{TS_*}^{TS_R} \mathbf{U}(x, T) dx . \end{aligned} \right\} \quad (10.22)$$

We define the integral averages



**Fig. 10.4.** HLLC approximate Riemann solver. Solution in the *Star Region* consists of two constant states separated from each other by a middle wave of speed  $S_*$ .

$$\left. \begin{aligned} \mathbf{U}_{*L} &= \frac{1}{T(S_* - S_L)} \int_{TS_L}^{TS_*} \mathbf{U}(x, T) dx , \\ \mathbf{U}_{*R} &= \frac{1}{T(S_R - S_*)} \int_{TS_*}^{TS_R} \mathbf{U}(x, T) dx . \end{aligned} \right\} \quad (10.23)$$

By substitution of (10.23) into (10.22) and use of (10.12), the consistency condition (10.9) becomes

$$\left( \frac{S_* - S_L}{S_R - S_L} \right) \mathbf{U}_{*L} + \left( \frac{S_R - S_*}{S_R - S_L} \right) \mathbf{U}_{*R} = \mathbf{U}^{hll} , \quad (10.24)$$

where  $\mathbf{U}^{hll}$  is given by (10.12)–(10.13). The HLLC approximate Riemann solver is given as follows

$$\tilde{\mathbf{U}}(x, t) = \begin{cases} \mathbf{U}_L & , \text{if } \frac{x}{t} \leq S_L , \\ \mathbf{U}_{*L} & , \text{if } S_L \leq \frac{x}{t} \leq S_* , \\ \mathbf{U}_{*R} & , \text{if } S_* \leq \frac{x}{t} \leq S_R , \\ \mathbf{U}_R & , \text{if } \frac{x}{t} \geq S_R . \end{cases} \quad (10.25)$$

We seek a corresponding HLLC numerical flux defined as

$$\mathbf{F}_{i+\frac{1}{2}}^{hllc} = \begin{cases} \mathbf{F}_L & , \text{if } 0 \leq S_L , \\ \mathbf{F}_{*L} & , \text{if } S_L \leq 0 \leq S_* , \\ \mathbf{F}_{*R} & , \text{if } S_* \leq 0 \leq S_R , \\ \mathbf{F}_R & , \text{if } 0 \geq S_R , \end{cases} \quad (10.26)$$

with the intermediate fluxes  $\mathbf{F}_{*L}$  and  $\mathbf{F}_{*R}$  still to be determined. Fig. 10.4 shows the structure of the HLLC approximate Riemann solver.

By integrating over appropriate control volumes, or more directly, by applying Rankine–Hugoniot Conditions across each of the waves of speeds  $S_L$ ,  $S_*$ ,  $S_R$ , we obtain

$$\mathbf{F}_{*L} = \mathbf{F}_L + S_L(\mathbf{U}_{*L} - \mathbf{U}_L) , \quad (10.27)$$

$$\mathbf{F}_{*R} = \mathbf{F}_{*L} + S_*(\mathbf{U}_{*R} - \mathbf{U}_{*L}) , \quad (10.28)$$

$$\mathbf{F}_{*R} = \mathbf{F}_R + S_R(\mathbf{U}_{*R} - \mathbf{U}_R) . \quad (10.29)$$

Compare relations (10.27) and (10.29) for the HLLC scheme with (10.18) and (10.19) for the HLL scheme. Substitution of  $\mathbf{F}_{*L}$  from (10.27) and  $\mathbf{F}_{*R}$  from (10.29) into (10.28) gives identically the consistency condition (10.24). Hence conditions (10.27)–(10.29) are sufficient for ensuring consistency; these are three equations for the four unknowns vectors  $\mathbf{U}_{*L}$ ,  $\mathbf{F}_{*L}$ ,  $\mathbf{U}_{*R}$ ,  $\mathbf{F}_{*R}$ .

#### 10.4.2 The HLLC Flux for the Euler Equations

We seek the solution for the two unknown intermediate fluxes  $\mathbf{F}_{*L}$  and  $\mathbf{F}_{*R}$ . From (10.27)–(10.29) we see that it is sufficient to find solutions for the two intermediate state vectors  $\mathbf{U}_{*L}$  and  $\mathbf{U}_{*R}$ . There are more unknowns than equations and some extra conditions need to be imposed, in order to solve the algebraic problem. Obvious conditions to impose are those satisfied by the exact solution; for pressure and normal component of velocity we have

$$\left. \begin{aligned} p_{*L} &= p_{*R} = p_* , \\ u_{*L} &= u_{*R} = u_* , \end{aligned} \right\} \quad (10.30)$$

and for tangential velocity components we have

$$\left. \begin{aligned} v_{*L} &= v_L , \quad v_{*R} = v_R , \\ w_{*L} &= w_L , \quad w_{*R} = w_R . \end{aligned} \right\} \quad (10.31)$$

See Chap. 4. In addition, it is entirely justified, and convenient, to set

$$S_* = u_* \quad (10.32)$$

and thus if an estimate for  $S_*$  is known, the normal velocity component  $u_*$  in the *Star Region* is known. Now equations (10.27) and (10.29) can be rearranged as

$$S_L \mathbf{U}_{*L} - \mathbf{F}_{*L} = S_L \mathbf{U}_L - \mathbf{F}_L , \quad (10.33)$$

and

$$S_R \mathbf{U}_{*R} - \mathbf{F}_{*R} = S_R \mathbf{U}_R - \mathbf{F}_R , \quad (10.34)$$

where the right-hand sides of (10.33) and (10.34) are known constant vectors. We also note the useful relation between  $\mathbf{U}$  and  $\mathbf{F}$ , namely

$$\mathbf{F}(\mathbf{U}) = u \mathbf{U} + p \mathbf{D} , \quad \mathbf{D} = [0, 1, 0, 0, u]^T . \quad (10.35)$$

Assuming that the wave speeds  $S_L$  and  $S_R$  are known and performing algebraic manipulations of the first and second components of equations (10.33)–(10.34) one obtains the following solutions for pressure in the two *Star Regions*

$$p_{*L} = p_L + \rho_L (S_L - u_L)(S_* - u_L) , \quad p_{*R} = p_R + \rho_R (S_R - u_R)(S_* - u_R) . \quad (10.36)$$

From (10.30)  $p_{*L} = p_{*R}$ , which from (10.36) allows us to obtain an expression for the speed  $S_*$  purely in terms of the assumed speeds  $S_L$  and  $S_R$ , namely

$$S_* = \frac{p_R - p_L + \rho_L u_L (S_L - u_L) - \rho_R u_R (S_R - u_R)}{\rho_L (S_L - u_L) - \rho_R (S_R - u_R)} . \quad (10.37)$$

Thus, we only need to provide estimates for  $S_L$  and  $S_R$ , just as for the simpler HLL solver.

Algebraic manipulation of (10.33) and (10.34) and using the corresponding values  $p_{*L}$  and  $p_{*R}$  from (10.36) gives the intermediate fluxes  $\mathbf{F}_{*L}$  and  $\mathbf{F}_{*R}$  as

$$\mathbf{F}_{*K} = \mathbf{F}_K + S_K (\mathbf{U}_{*K} - \mathbf{U}_K) , \quad (10.38)$$

for K=L and K=R, with the intermediate states given as

$$\mathbf{U}_{*K} = \rho_K \left( \frac{S_K - u_K}{S_K - S_*} \right) \begin{bmatrix} 1 \\ S_* \\ v_K \\ w_K \\ \frac{E_K}{\rho_K} + (S_* - u_K) \left[ S_* + \frac{p_K}{\rho_K (S_K - u_K)} \right] \end{bmatrix} . \quad (10.39)$$

The final choice of the HLLC flux is made according to (10.26).

A variation in the formulation of the HLLC solver (10.38)–(10.39) is the following. From equations (10.33) and (10.34) we may write the following solutions for the state vectors  $\mathbf{U}_{*L}$  and  $\mathbf{U}_{*R}$

$$\mathbf{U}_{*K} = \frac{S_K \mathbf{U}_K - \mathbf{F}_K + p_{*K} \mathbf{D}_*}{S_L - S_*} , \quad \mathbf{D}_* = [0, 1, 0, 0, S_*] , \quad (10.40)$$

with  $p_{*L}$  and  $p_{*R}$  as given by (10.36). Substitution of  $p_{*K}$  from (10.36) into (10.40) followed by use of (10.27) and (10.29) gives direct expressions for the intermediate fluxes as

$$\mathbf{F}_{*K} = \frac{S_* (S_K \mathbf{U}_K - \mathbf{F}_K) + S_K (p_K + \rho_L (S_K - u_K) (S_* - u_K)) D_*}{S_K - S_*} , \quad (10.41)$$

with the final choice of the HLLC flux made again according to (10.26).

We remark here that the HLLC formulation (10.38)–(10.39) enforces the condition  $p_{*L} = p_{*R}$ , which is satisfied by the exact solution. In the alternative HLLC formulation (10.41) we relax such condition, being more consistent with the pressure approximations (10.36).

A different HLLC flux is obtained by assuming a single mean pressure value in the *Star Region*, and given by the arithmetic average of the pressures in (10.36), namely

$$P_{LR} = \frac{1}{2} [p_L + p_R + \rho_L(S_L - u_L)(S_* - u_L) + \rho_R(S_R - u_R)(S_* - u_R)] . \quad (10.42)$$

Then the intermediate state vectors are given by

$$\mathbf{U}_{*K} = \frac{S_K \mathbf{U}_K - \mathbf{F}_K + P_{LR} \mathbf{D}_*}{S_K - S_*} . \quad (10.43)$$

Substitution of these into (10.27) and (10.29) gives the fluxes  $\mathbf{F}_{*L}$  and  $\mathbf{F}_{*R}$  as

$$\mathbf{F}_{*K} = \frac{S_*(S_K \mathbf{U}_K - \mathbf{F}_K) + S_K P_{LR} \mathbf{D}_*}{S_K - S_*} . \quad (10.44)$$

Again the final choice of HLLC flux is made according to (10.26).

*Remark: general equation of state.* All manipulations so far, assuming that wave speed estimates for  $S_L$  and  $S_R$  are available, are valid for any equation of state. The equation of state only enters in prescribing estimates for  $S_L$  and  $S_R$ .

### 10.4.3 Multidimensional and Multicomponent Flow

Here we consider extensions of the HLLC solver to two areas of application, namely multidimensional flow and multicomponent flow.

The presentation of the HLLC scheme has been made for the  $x$ -split three-dimensional Euler equations, for which the corresponding eigenvalues are denoted here as  $\lambda_1 = u - a$ ,  $\lambda_2 = u$  (multiplicity 3),  $\lambda_3 = u + a$ , where  $u$  is the *normal* velocity component and  $a$  is the speed of sound. In a general multidimensional situation, see Chapt. 16, one usually requires the flux in the direction *normal* to a volume (or element) interface, which is not necessarily aligned with any of the Cartesian directions. In this case the form of the governing equations remains identical to the  $x$ -split system (10.4), (10.5). There will be a normal and two tangential components of velocity as before, and all the results obtained so far will be applicable.

In the study of multicomponent flow, one considers the advection of *chemical species* by the flow, the carrier fluid. For example, let us consider  $m$  species of concentrations  $q_l$ , for  $l = 1, \dots, m$ , advected with the normal fluid speed  $u$ . This means that for each species we can write the following advection equation

$$\partial_t q_l + u \partial_x q_l = 0 ,$$

for  $l = 1, \dots, m$ . Note that these equations are written in non-conservative form. However, by combining these with the continuity equation we obtain a conservative form of these equations, namely

$$(\rho q_l)_t + (\rho u q_l)_x = 0 , \text{ for } l = 1, \dots, m .$$

The eigenvalues of the enlarged system are as before, with the exception of  $\lambda_2 = u$ , which now, in three space dimensions, has multiplicity  $m + 3$ . These conservation equations can then be added as new components to the conservation equations in (10.1) or (10.4), with the enlarged vectors of conserved variables and fluxes given as

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ E \\ \rho q_1 \\ \dots \\ \rho q_l \\ \dots \\ \rho q_m \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ \rho uw \\ u(E + p) \\ \rho u q_1 \\ \dots \\ \rho u q_l \\ \dots \\ \rho u q_m \end{bmatrix}. \quad (10.45)$$

The HLLC flux accommodates these new equations in a very natural way, and nothing special needs to be done. If the HLLC flux (10.38) is used, with  $\mathbf{F}$  as in (10.45), then the intermediate state vectors are given by

$$\mathbf{U}_{*K} = \rho_K \left( \frac{S_K - u_K}{S_K - S_*} \right) \begin{bmatrix} 1 \\ S_* \\ v_K \\ w_K \\ \frac{E_K}{\rho_K} + (S_* - u_K) \left[ S_* + \frac{p_K}{\rho_K(S_K - u_K)} \right] \\ (q_1)_K \\ \dots \\ (q_l)_K \\ \dots \\ (q_m)_K \end{bmatrix}. \quad (10.46)$$

for  $K = L$  and  $K = R$ . In this manner the HLLC flux will resolve the additional intermediate fields as the exact Riemann solver.

Note that the tangential velocity components  $v$  and  $w$  are special cases of passive scalars; compare (10.46) with (10.39) for  $q = v$  and  $q = w$ .

## 10.5 Wave-Speed Estimates

In order to determine completely the numerical fluxes in both the HLL and HLLC Riemann solvers we need to provide an algorithm for computing the wave speeds  $S_L$  and  $S_R$ . For the HLLC scheme one requires in addition an estimate for the speed of the middle wave  $S_*$ , but as seen in (10.37), this can in fact be computed once  $S_L$  and  $S_R$  are known. Thus the pending task is to

determine estimates for  $S_L$  and  $S_R$ . One approach is to estimate the speeds directly; another approach relies on pressure estimates in the *Star Region*, which are then utilised to obtain  $S_L$  and  $S_R$  using exact wave relations.

### 10.5.1 Direct Wave Speed Estimates

The most well known approach for estimating bounds for the minimum and maximum signal velocities present in the solution of the Riemann problem is to provide, directly, wave speeds  $S_L$  and  $S_R$ . Davis [150] suggested the simple estimates

$$S_L = u_L - a_L, \quad S_R = u_R + a_R \quad (10.47)$$

and

$$S_L = \min \{u_L - a_L, u_R - a_R\}, \quad S_R = \max \{u_L + a_L, u_R + a_R\}. \quad (10.48)$$

These estimates make use of data values only, are exceedingly simple but are not recommended for practical computations. Both Davis [150] and Einfeldt [181], proposed to use the Roe [407] average eigenvalues for the left and right non-linear waves, that is

$$S_L = \tilde{u} - \tilde{a}, \quad S_R = \tilde{u} + \tilde{a}, \quad (10.49)$$

where  $\tilde{u}$  and  $\tilde{a}$  are the Roe-average particle and sound speeds respectively, given as follows

$$\tilde{u} = \frac{\sqrt{\rho_L} u_L + \sqrt{\rho_R} u_R}{\sqrt{\rho_L} + \sqrt{\rho_R}}, \quad \tilde{a} = \left[ (\gamma - 1) \left( \tilde{H} - \frac{1}{2} \tilde{u}^2 \right) \right]^{1/2}, \quad (10.50)$$

with the enthalpy  $H = (E + p)/\rho$  approximated as

$$\tilde{H} = \frac{\sqrt{\rho_L} H_L + \sqrt{\rho_R} H_R}{\sqrt{\rho_L} + \sqrt{\rho_R}}. \quad (10.51)$$

Complete details of the Roe Riemann solver are given in Chap. 11.

Motivated by the Roe eigenvalues Einfeldt [181] proposed the estimates

$$S_L = \bar{u} - \bar{d}, \quad S_R = \bar{u} + \bar{d}, \quad (10.52)$$

for his HLLE solver, where

$$\bar{d}^2 = \frac{\sqrt{\rho_L} a_L^2 + \sqrt{\rho_R} a_R^2}{\sqrt{\rho_L} + \sqrt{\rho_R}} + \eta_2 (u_R - u_L)^2 \quad (10.53)$$

and

$$\eta_2 = \frac{1}{2} \frac{\sqrt{\rho_L} \sqrt{\rho_R}}{(\sqrt{\rho_L} + \sqrt{\rho_R})^2}. \quad (10.54)$$

These wave speed estimates are reported to lead to effective and robust Godunov-type schemes. More details on the HLL solver and its modification HLLEM, are found in [182]. In this paper the authors also analyze the effect of the choice of wave speed estimates on the Riemann solver and introduce the concept of *positively conservative* Riemann solvers, for the Euler equations. These are solvers for which, for physically admissible data, density and internal energy remain positive during the calculations. See Batten et al. [32] for further discussion on direct wave speed estimates.

Davis made some observations regarding the relationship between the chosen wave speeds and some well-known numerical methods. Suppose that for a given Riemann problem we can identify a positive speed  $S^+$ . Then by choosing  $S_L = -S^+$  and  $S_R = S^+$  in the HLL flux (10.20) one obtains a Rusanov flux [418]

$$\mathbf{F}_{i+1/2} = \frac{1}{2}(\mathbf{F}_L + \mathbf{F}_R) - \frac{1}{2}S^+(\mathbf{U}_R - \mathbf{U}_L). \quad (10.55)$$

As to the choice of the speed  $S^+$ , Davis [150] considered

$$S^+ = \max \{|u_L - a_L|, |u_R - a_R|, |u_L + a_L|, |u_R + a_R|\}.$$

Actually, the above speed is bounded by

$$S^+ = \max \{|u_L| + a_L, |u_R| + a_R\}. \quad (10.56)$$

This choice is likely to produce a more robust scheme and is also simpler than Davis's choice.

Another possible choice is  $S^+ = S_{max}^n$ , the maximum wave speed present at the appropriate time found by imposing the Courant stability condition; see Sect. 6.3.2 of Chap. 6. This speed is related to the time step  $\Delta t$  and the grid spacing  $\Delta x$  via

$$S_{max}^n = \frac{C_{cfl}\Delta x}{\Delta t}, \quad (10.57)$$

where  $C_{cfl}$  is the Courant number coefficient, usually chosen (empirically) to be  $C_{cfl} \approx 0.9$ , for a scheme of linear stability limit of unity. For  $C_{cfl} = 1$  one has  $S^+ = \frac{\Delta x}{\Delta t}$ , which results in the Lax–Friedrichs numerical flux

$$F_{i+1/2} = \frac{1}{2}(F_L + F_R) - \frac{1}{2}\frac{\Delta x}{\Delta t}(U_R - U_L). \quad (10.58)$$

See Sect. 5.3.4 of Chap. 5 and Sect. 7.3.1 of Chap. 7.

In the next section we propose a different way of finding wave-speed estimates.

### 10.5.2 Pressure-Based Wave Speed Estimates

A different approach for finding wave speed estimates was proposed by Toro et. al. [542], whereby one first finds an estimate for the pressure  $p_*$  in

the *Star Region*. Then, estimates for  $S_L$  and  $S_R$  are derived. This is a simple task and several reliable choices are available. Suppose we have an estimate  $p_*$  for the pressure in the *Star Region*. Then we choose the following wave speeds

$$S_L = u_L - a_L q_L, \quad S_R = u_R + a_R q_R, \quad (10.59)$$

where

$$q_K = \begin{cases} 1 & \text{if } p_* \leq p_K \\ \left[ 1 + \frac{\gamma+1}{2\gamma} (p_*/p_K - 1) \right]^{1/2} & \text{if } p_* > p_K. \end{cases} \quad (10.60)$$

This choice of wave speeds discriminates between shock and rarefaction waves. If the  $K$  wave ( $K = L$  or  $K = R$ ) is a rarefaction then the speed  $S_K$  corresponds to the characteristic speed of the head of the rarefaction, which carries the fastest signal. If the wave is a shock wave then the speed corresponds to an approximation of the true shock speed; the wave relations used are exact but the pressure ratio across the shock is approximated, because  $p_*$  is an approximation to the pressure behind the shock wave. We propose to use the state approximations of Chap. 9 to find  $p_*$ .

The PVRS approximate Riemann solver [502] presented in Sect. 9.3 of Chap. 9 gives

$$p_{pvrs} = \frac{1}{2}(p_L + p_R) - \frac{1}{2}(u_R - u_L)\bar{\rho}\bar{a}, \quad (10.61)$$

where

$$\bar{\rho} = \frac{1}{2}(\rho_L + \rho_R), \quad \bar{a} = \frac{1}{2}(a_L + a_R). \quad (10.62)$$

This approximation for pressure can be used directly into (10.59)–(10.60) to obtain wave speed estimates for the HLL and HLLC schemes. See also Eq. (9.28) of Chapt. 9 for alternative estimates for  $p_*$ .

Another choice is furnished by the Two-Rarefaction Riemann solver TRRS of Sect. 9.4.1 of Chap. 9, namely

$$p_{tr} = \left[ \frac{a_L + a_R - \frac{\gamma-1}{2}(u_R - u_L)}{a_L/p_L^z + a_R/p_R^z} \right]^{1/z}, \quad (10.63)$$

where

$$P_{LR} = \left( \frac{p_L}{p_R} \right)^z; \quad z = \frac{\gamma-1}{2\gamma}. \quad (10.64)$$

The Two-Shock Riemann solver TSRS of Sect. 9.4.2 of Chap. 9 gives

$$p_{ts} = \frac{g_L(p_0)p_L + g_R(p_0)p_R - \Delta u}{g_L(p_0) + g_R(p_0)}, \quad (10.65)$$

where

$$g_K(p) = \left[ \frac{A_K}{p + B_K} \right]^{1/2}, \quad p_0 = \max(0, p_{pvrs}), \quad (10.66)$$

for  $K = L$  and  $K = R$ .

In computational practice we could use the hybrid scheme of Sect. 9.5.2 of Chap. 9 to determine  $p_*$ . See Chap. 9 for full details. The HLL approximate Riemann solver with the hybrid pressure-based wave speed estimates has been implemented in the NAG routine D03PXF [319] for Godunov-type methods to solve the time-dependent, one dimensional Euler equations for ideal gases. For ideal gases we find that the simplified PVRS scheme, with  $p_* = \max(0, p_{pvrs})$  is very simple and also is found to be sufficiently robust.

## 10.6 Summary of HLLC Fluxes

Here we summarize the HLLC scheme, based on a particular choice of wave speeds. To compute the HLLC flux one performs the following steps:

- *Step I: pressure estimate.* Compute estimate for the pressure  $p_*$  in the *Star Region* as

$$\left. \begin{aligned} p_* &= \max(0, p_{pvrs}), & p_{pvrs} &= \frac{1}{2}(p_L + p_R) - \frac{1}{2}(u_R - u_L)\bar{\rho}\bar{a}, \\ \bar{\rho} &= \frac{1}{2}(\rho_L + \rho_R), & \bar{a} &= \frac{1}{2}(a_L + a_R). \end{aligned} \right\} \quad (10.67)$$

There are other possible choices for estimating the pressure  $p_*$ . See (10.63) and (10.65).

- *Step II: wave speed estimates.* Compute the wave speed estimates for  $S_L$  and  $S_R$  as

$$S_L = u_L - a_L q_L, \quad S_R = u_R + a_R q_R, \quad (10.68)$$

with

$$q_K = \begin{cases} 1 & \text{if } p_* \leq p_K \\ \left[ 1 + \frac{\gamma+1}{2\gamma} \left( \frac{p_*}{p_K} - 1 \right) \right]^{1/2} & \text{if } p_* > p_K. \end{cases} \quad (10.69)$$

Then compute the intermediate speed  $S_*$  in terms of  $S_L$  and  $S_R$  as

$$S_* = \frac{p_R - p_L + \rho_L u_L (S_L - u_L) - \rho_R u_R (S_R - u_R)}{\rho_L (S_L - u_L) - \rho_R (S_R - u_R)}. \quad (10.70)$$

Other choices of  $S_L$  and  $S_R$  are possible. See for example (10.49) and (10.52)

- *Step III: HLLC flux.* Compute the HLLC flux, according to

$$\mathbf{F}_{i+\frac{1}{2}}^{hllc} = \begin{cases} \mathbf{F}_L & \text{if } 0 \leq S_L, \\ \mathbf{F}_{*L} & \text{if } S_L \leq 0 \leq S_*, \\ \mathbf{F}_{*R} & \text{if } S_* \leq 0 \leq S_R, \\ \mathbf{F}_R & \text{if } 0 \geq S_R, \end{cases} \quad (10.71)$$

with

$$\mathbf{F}_{*K} = \mathbf{F}_K + S_K(\mathbf{U}_{*K} - \mathbf{U}_K) \quad (10.72)$$

and

$$\mathbf{U}_{*K} = \rho_K \left( \frac{S_K - u_K}{S_K - S_*} \right) \begin{bmatrix} 1 \\ S_* \\ v_K \\ w_K \\ \frac{E_K}{\rho_K} + (S_* - u_K) \left[ S_* + \frac{p_K}{\rho_K(S_K - u_K)} \right] \end{bmatrix}. \quad (10.73)$$

There are two variants of the HLLC flux in the third step, as seen below.

- *Step III: HLLC flux, Variant 1.* Compute the numerical fluxes as

$$\mathbf{F}_{*K} = \frac{S_*(S_K \mathbf{U}_K - \mathbf{F}_K) + S_K(p_K + \rho_L(S_K - u_K)(S_* - u_K))\mathbf{D}_*}{S_K - S_*}, \quad \left. \begin{aligned} \mathbf{D}_* &= [0, 1, 0, 0, S_*]^T, \end{aligned} \right\} \quad (10.74)$$

and the final HLLC flux chosen according to (10.71).

- *Step III: HLLC flux, Variant 2.* Compute the numerical fluxes as

$$\mathbf{F}_{*K} = \frac{S_*(S_K \mathbf{U}_K - \mathbf{F}_K) + S_K P_{LR} \mathbf{D}_*}{S_K - S_*}, \quad (10.75)$$

with  $\mathbf{D}_*$  as in (10.74) and

$$P_{LR} = \frac{1}{2} [p_L + p_R + \rho_L(S_L - u_L)(S_* - u_L) + \rho_R(S_R - u_R)(S_* - u_R)]. \quad (10.76)$$

The final HLLC flux is chosen according to (10.71).

## 10.7 Contact Waves and Passive Scalars

Here we study the special case of a passive scalar  $q(x, t)$  transported with the fluid speed  $u(x, t)$ . The time-dependent, one dimensional Euler equations are augmented by the extra conservation law

$$(\rho q)_t + (\rho qu)_x = 0. \quad (10.77)$$

Consider the special IVP in which  $p = \text{constant}$ ,  $\rho = \text{constant}$ ,  $u = \text{constant}$  and

$$q(x, 0) = q_0(x) = \begin{cases} q_L & \text{if } x \leq 0 \\ q_R & \text{if } x > 0. \end{cases} \quad (10.78)$$

Clearly, the non-trivial part of the exact solution is

$$q(x, t) = q_0(x - ut). \quad (10.79)$$

Application of the HLL Riemann solver to this problem gives the following expression for the numerical flux

$$f_{i+\frac{1}{2}}^{\text{hll}} = \frac{1}{2} \left( 1 + \frac{1}{M} \right) f_i + \frac{1}{2} \left( 1 - \frac{1}{M} \right) f_{i+1}, \quad (10.80)$$

where  $M = \frac{u}{a}$  is the Mach number and the wave speeds have been taken to be

$$S_L = u - a, \quad S_R = u + a.$$

Obviously, this flux applies only in the subsonic regime  $u - a \leq 0 \leq u + a$ . For sonic flow, the flux (10.80) reduces identically to the Godunov flux computed from the exact Riemann solver. For subsonic flow  $1/M > 1$  and the resulting scheme is *more diffusive than the Godunov method* when used in conjunction with the exact Riemann solver. For the special case

$$M = \frac{u\Delta t}{\Delta x}$$

the HLL scheme reproduces the Lax–Friedrichs method, which is exceedingly diffusive, see Chaps. 5 and 6. The limiting case of a stationary passive scalar is the worst. Note that the analysis includes the important cases  $q = v$  and  $q = w$ , the tangential velocity components in three-dimensional flow.

The analysis for an isolated contact can be carried out in a similar manner; by using an appropriate choice of the wave speeds the resulting HLL flux is identical to (10.80), and thus the same observations as for a passive scalar apply. The HLLC solver, on the other hand, behaves as the exact Riemann solver; for the limiting case in which the wave is stationary, the HLLC numerical scheme gives infinite resolution; the reader can verify this algebraically. In the next section on numerical results we compare the HLL and HLLC schemes for this type of problems; see Fig. 10.9. The relevance of these observations is that the HLL scheme, unlike the HLLC scheme, will add excessive numerical dissipation to the resolution of special but important flow features such as material interfaces, shear waves and vortices.

## 10.8 Numerical Results

Here we assess the performance of Godunov's first-order method used in conjunction with the HLL and HLLC approximate Riemann solvers presented in this chapter. The HLLC results shown correspond to the version (10.38)–(10.39). For both HLL and HLLC, the wave speed estimates for  $S_L$  and  $S_R$  are based on a pressure estimate obtained from the adaptive scheme of section 9.5.2 of Chapter 9. For HLLC we note that, for the tests considered, all three versions of HLLC give identical results when using the simple algorithm (10.67)–(10.70).

We select seven test problems for the one-dimensional, time dependent Euler equations for ideal gases with  $\gamma = 1.4$ ; these have exact solutions. In all chosen tests, data consists of two constant states  $\mathbf{W}_L = [\rho_L, u_L, p_L]^T$  and  $\mathbf{W}_R = [\rho_R, u_R, p_R]^T$ , separated by a discontinuity at a position  $x = x_0$ . The states  $\mathbf{W}_L$  and  $\mathbf{W}_R$  are given in Table 10.1. The exact and numerical solutions are found in the spatial domain  $0 \leq x \leq 1$ . The numerical solution is computed with  $M = 100$  cells and the CFL condition is as for all previous computations, see Chap. 6; the chosen Courant number coefficient is  $C_{\text{cfl}} = 0.9$ ; boundary conditions are transmissive.

The exact solutions were found by running the code HE-E1RPEXACT of the library *NUMERICA* [518] and the numerical solutions were obtained by running the code HE-E1GODFLUX of *NUMERICA*.

Test	$\rho_L$	$u_L$	$p_L$	$\rho_R$	$u_R$	$p_R$
1	1.0	0.75	1.0	0.125	0.0	0.1
2	1.0	-2.0	0.4	1.0	2.0	0.4
3	1.0	0.0	1000.0	1.0	0.0	0.01
4	5.99924	19.59775	460.894	5.99242	-6.19633	46.0950
5	1.0	-19.59745	1000.0	1.0	-19.59745	0.01
6	1.4	0.0	1.0	1.0	0.0	1.0
7	1.4	0.1	1.0	1.0	0.1	1.0

Table 10.1. Data for seven test problems with exact solution

Test 1 is a *modified* version of Sod's problem [453]; the solution has a right shock wave, a right travelling contact wave and a left *sonic* rarefaction wave; this test is useful for assessing the *entropy satisfaction* property of numerical methods. The solution of Test 2 consists of two symmetric rarefaction waves and a trivial contact wave; the *Star Region* between the non-linear waves is close to vacuum, which makes this problem a suitable test for assessing the performance of numerical methods for low-density flows. Test 3 is designed to assess the robustness and accuracy of numerical methods; its solution consists of a strong shock wave of shock Mach number 198, a contact surface and a left rarefaction wave. Test 4 is also a very severe test, its solution consists of three strong discontinuities travelling to the right. A detailed discussion on the

exact solution of Tests 1 to 4 is found in Sect. 4.3.3 of Chap. 4. Test 5 is also designed to test the robustness of numerical methods but the main reason for devising this test is to assess the ability of the numerical methods to resolve *slowly-moving contact discontinuities*. The exact solution of Test 5 consists of a left rarefaction wave, a right-travelling shock wave and a *stationary* contact discontinuity. Test 6 corresponds to an isolated stationary contact wave and Test 7 corresponds to an isolated contact moving slowly to the right. The purpose of Tests 6 and 7 is to illustrate the likely performance of HLL and HLLC for contacts, shear waves and vortices. For each test problem we select a convenient position  $x_0$  of the initial discontinuity and the output time. These are stated in the legend of each figure displaying computational results.

We compare computed results with the exact solution for three first-order methods, namely the Godunov method used in conjunction with the HLL and HLLC approximate Riemann solvers, and the Rusanov scheme. In all three schemes we compute wave speed estimates by using the adaptive noniterative scheme of Sect. 9.5.2 of Chapt. 9. Figs. 10.5 to 10.9 show results for Godunov's method with the HLLC Riemann solver. Figs. 10.10 to 10.14 show results for the Godunov method with the HLL Riemann solver and Figs. 10.15 to 10.19 show results for Rusanov's method. Fig. 10.20 shows results aimed at comparing the performance of HLL and HLLC for isolated, stationary and slowly moving contact discontinuities.

The numerical results obtained from the Godunov method in conjunction with the HLL and HLLC approximate Riemann solvers are broadly similar to those obtained from Godunov's method in conjunction with the exact Riemann solver. See results of Chapt. 6. Some points to note are the following: the *sonic rarefaction* of Test 1 is better resolved by the HLL and HLLC approximate Riemann solvers than by the exact Riemann solver. The resolution of the stationary contact (non-isolated) of Test 5 by the HLLC Riemann solver is comparable with that of the exact Riemann solver. The HLL Riemann solver however, as anticipated by the analysis of Sect. 10.7, diffuses the contact wave to similar levels seen in the Flux Vector Splitting methods of Steger–Warming and van Leer, see results of Chap. 8. The advantage of HLLC over HLL is the resolution of slowly-moving contact discontinuities; this point is further emphasised by the results of Tests 6 and 7 for an *isolated* contact wave. The HLLC Riemann solver preserves the excellent entropy-satisfaction property of the HLL Riemann solver. The Rusanov scheme is broadly similar to the HLL Riemann solver in that it also diffuses slowly moving contacts. For Test 1 containing a sonic rarefaction however, the Rusanov scheme is clearly inferior to the HLL scheme, compare Fig. 10.15 with Fig. 10.10.

The results of Tests 6 and 7 using both the HLL and the HLLC schemes are shown in Fig. 10.20. As anticipated by the analysis of Sect. 10.7, the HLL scheme will give unacceptably smeared results for stationary and slowly moving contact waves. The HLLC behaves like the exact Riemann solver for this type of problem; it has much less numerical dissipation for slowly moving contacts and it gives infinite resolution for stationary contact waves. The

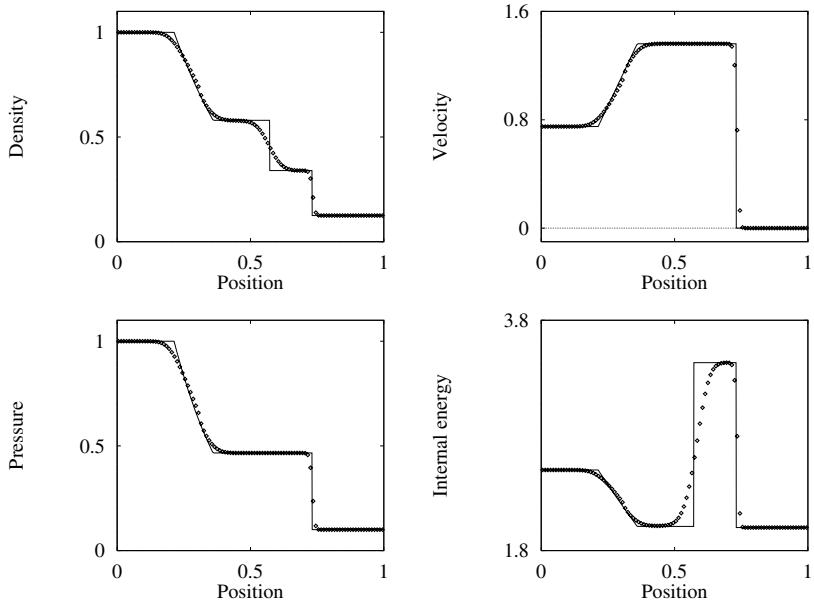
same observations apply to augmented systems of equations containing species equations, and to shear waves and vortices in multiple space dimensions.

## 10.9 Closing Remarks

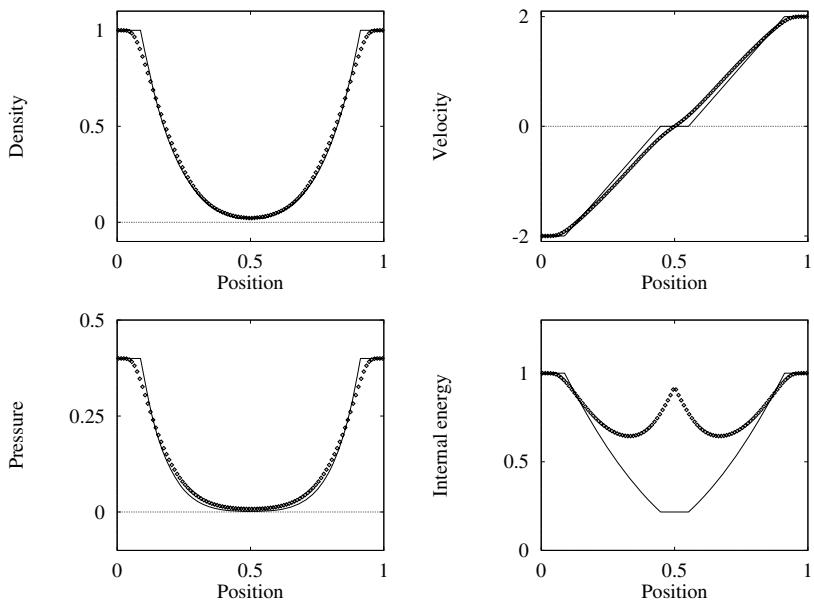
We have first studied HLL and HLLC approximate Riemann solvers for the split three-dimensional Euler equations. Then we have indicated the manner in which these solvers can be extended to three-dimensional flow and to multicomponent flow, noting that HLLC will perform as the exact Riemann solver in these more general situations. This is due to the fact that tangential velocity components and species concentrations are all represented by the intermediate characteristic field  $\lambda_2 = u$ , where  $u$  is understood as the normal velocity component. HLLC, unlike HLL, captures correctly this characteristic field, which is enough to correctly capture contact discontinuities, shear waves and contact discontinuities associated with all the species equations.

The approximate Riemann solvers of this chapter may be applied in conjunction with the Godunov first-order upwind method presented in Chap. 6. Second-order Total Variation Diminishing (TVD) extensions of the schemes are presented in Chap. 13 for scalar problems and in Chap. 14 for non-linear one dimensional systems. In Chap. 15 we present techniques that allow the extension of these schemes to solve problems with source terms. In Chap. 16 we study techniques to extend the methods of this chapter to three-dimensional problems. Implicit versions of the HLL and HLLC Riemann solvers have been developed by Batten, Leschziner and Goldberg [33], who have also applied the schemes to turbulent flows. The HLLC scheme can be used as the building block for high-order methods, semi discrete, fully discrete, finite volume and discontinuous Galerkin finite element methods, on structured and unstructured grids. See for example [6], [86], [237], [334], [361], [382], [553] and [572]. At this stage, two useful remarks on the HLLC flux are worth mentioning. The first concerns the positivity/negativity of the momentum flux Safranov [420]. The second (Dr V. A. Titarev, personal communication) concerns the question of robustness of the choice of wave speeds in the HLL and HLLC solvers for the case of very high speed flow impinging on solid stationary walls; some of the well known wave speed estimates may fail.

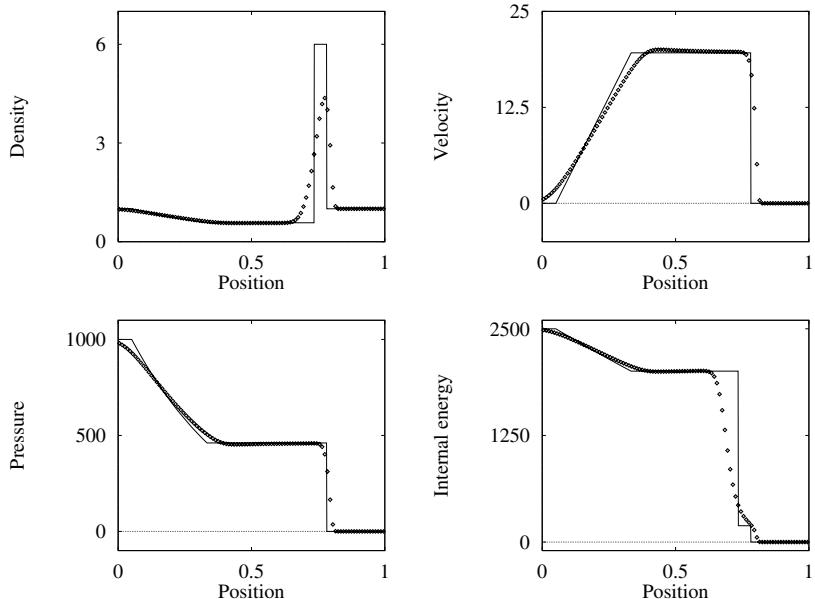
Perhaps the most significant advance of the HLLC approach concerns systems with more than three distinct characteristic fields, such as systems for multiphase flow and the MHD equations, for example. A proper treatment of these, following the HLLC approach, requires the construction of an appropriate wave model that includes, ideally, all the characteristic fields of the relevant system. Developments in this direction are found, for example, in [230], [474] and [75].



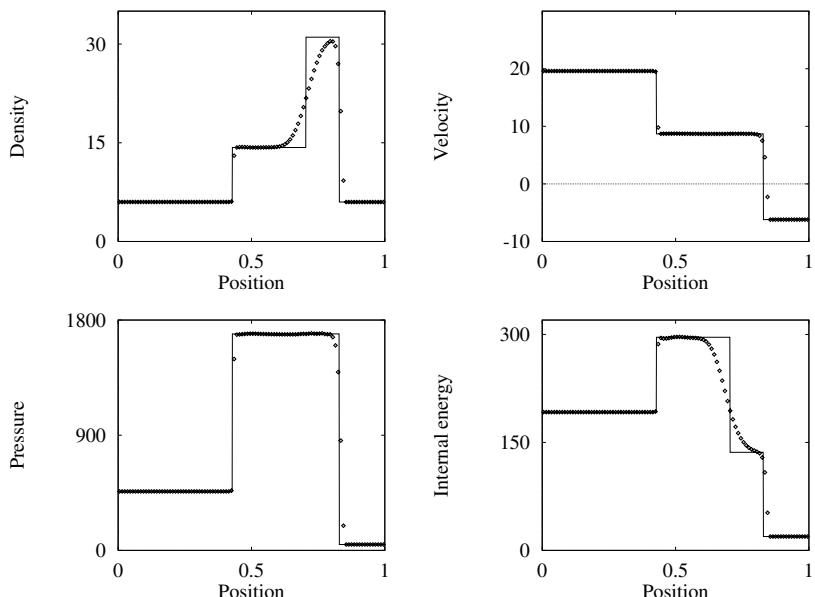
**Fig. 10.5.** Godunov's method with HLLC Riemann solver applied to Test 1, with  $x_0 = 0.3$ . Numerical (symbol) and exact (line) solutions are compared at time 0.2.



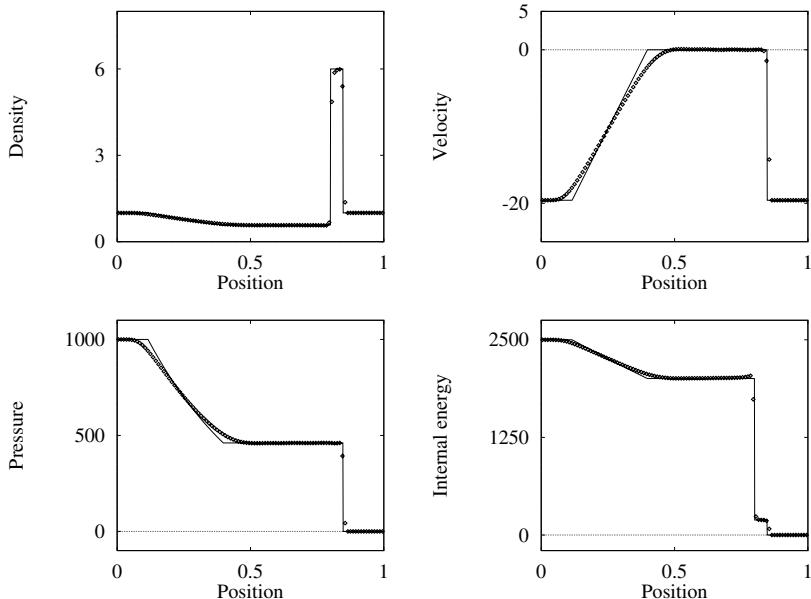
**Fig. 10.6.** Godunov's method with HLLC Riemann solver applied to Test 2, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at time 0.15.



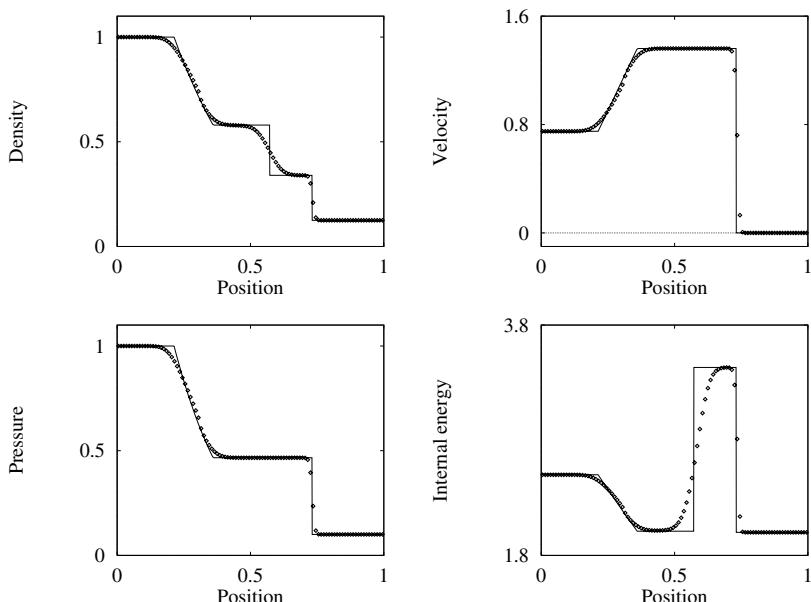
**Fig. 10.7.** Godunov's method with HLLC Riemann solver applied to Test 3, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012.



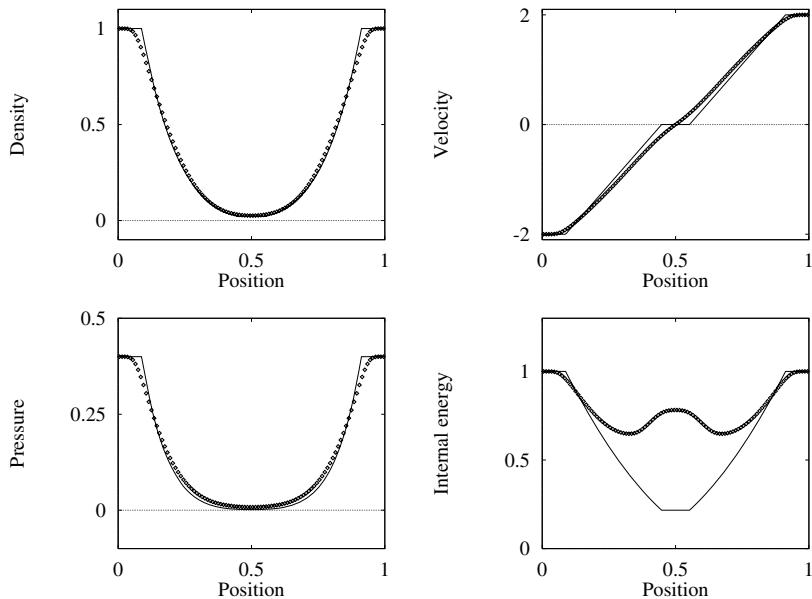
**Fig. 10.8.** Godunov's method with HLLC Riemann solver applied to Test 4, with  $x_0 = 0.4$ . Numerical (symbol) and exact (line) solutions are compared at time 0.035.



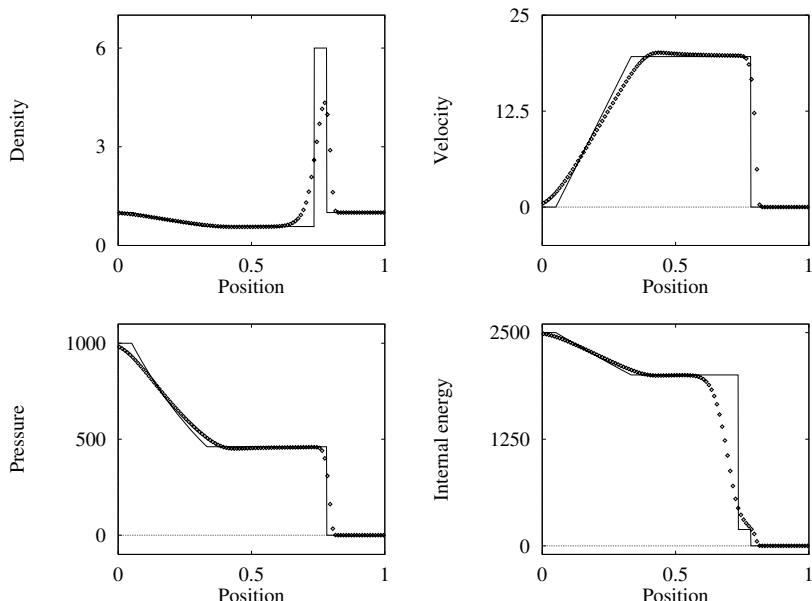
**Fig. 10.9.** Godunov's method with HLLC Riemann solver applied to Test 5, with  $x_0 = 0.8$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012.



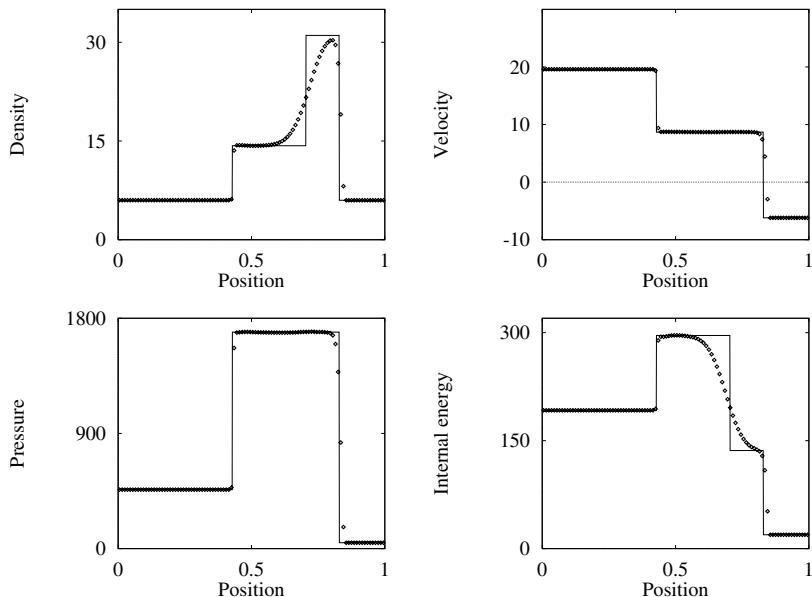
**Fig. 10.10.** Godunov's method with HLL Riemann solver applied to Test 1, with  $x_0 = 0.3$ . Numerical (symbol) and exact (line) solutions are compared at time 0.2.



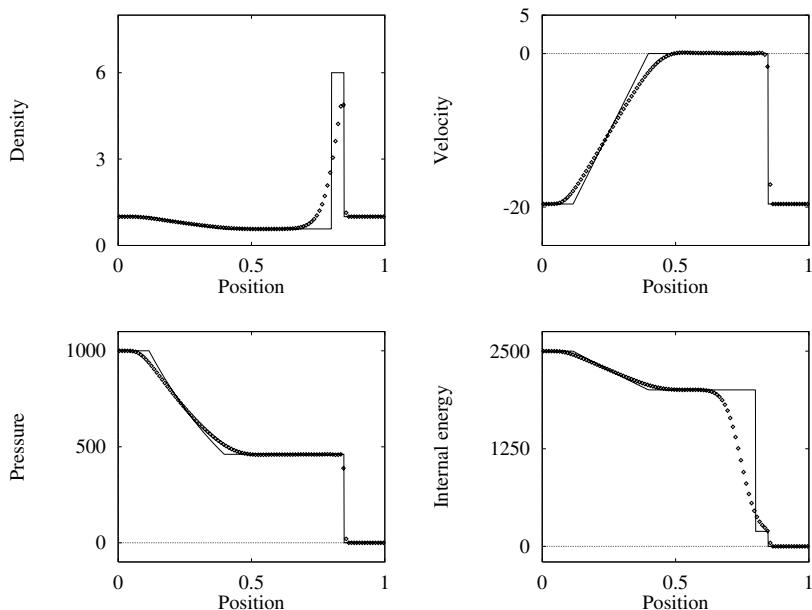
**Fig. 10.11.** Godunov's method with HLL Riemann solver applied to Test 2, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at time 0.15.



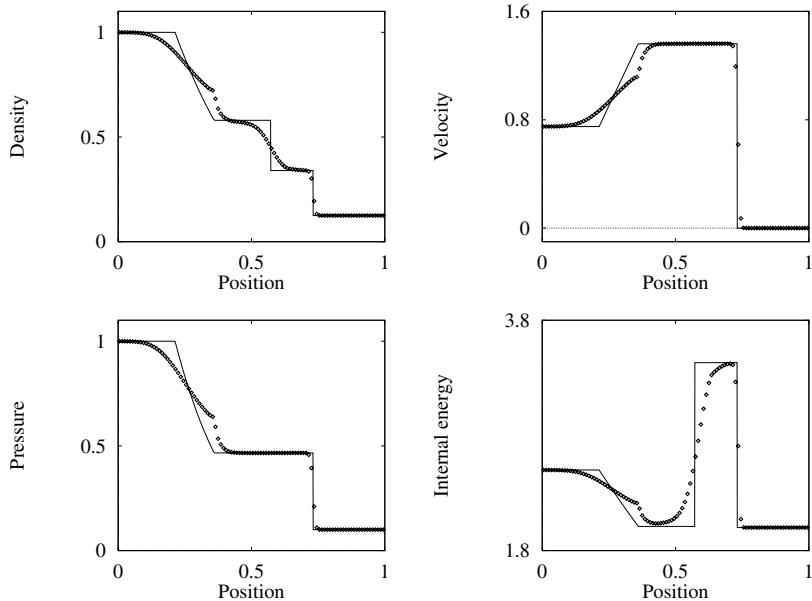
**Fig. 10.12.** Godunov's method with HLL Riemann solver applied to Test 3, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012.



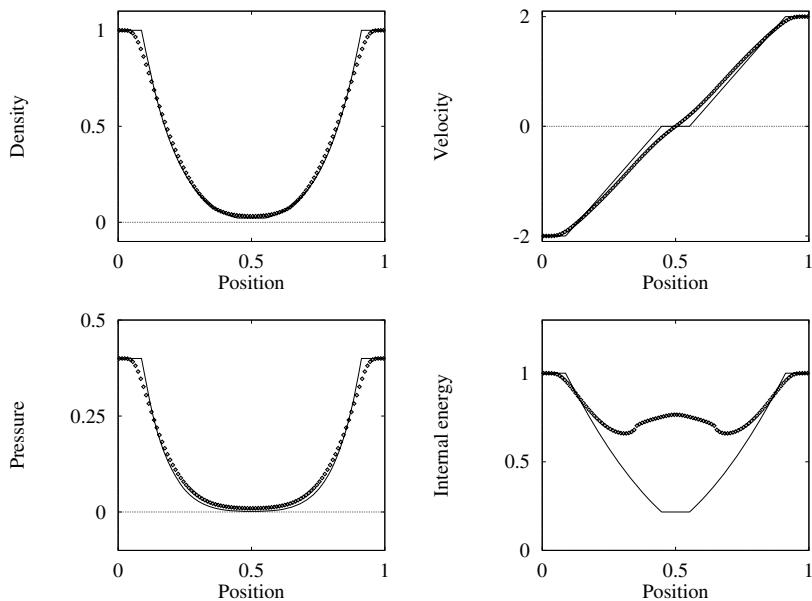
**Fig. 10.13.** Godunov's method with HLL Riemann solver applied to Test 4, with  $x_0 = 0.4$ . Numerical (symbol) and exact (line) solutions are compared at time 0.035.



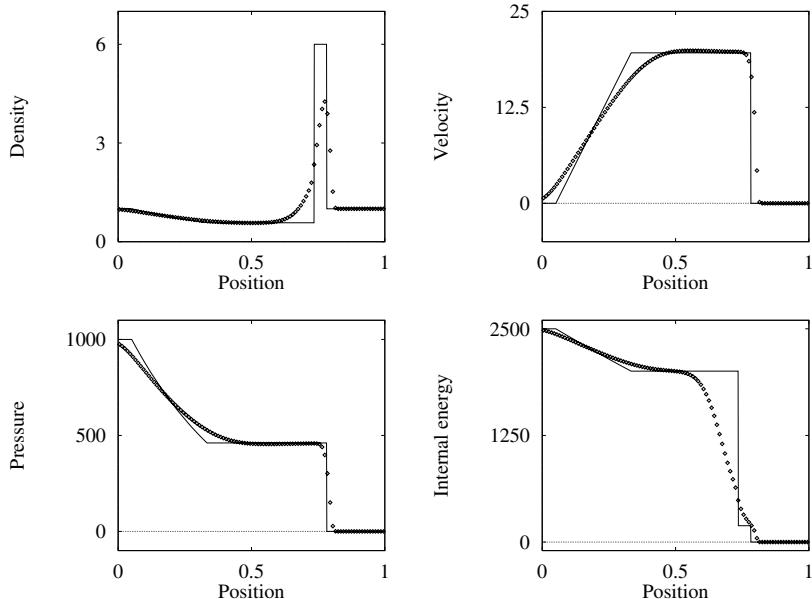
**Fig. 10.14.** Godunov's method with HLL Riemann solver applied to Test 5, with  $x_0 = 0.8$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012.



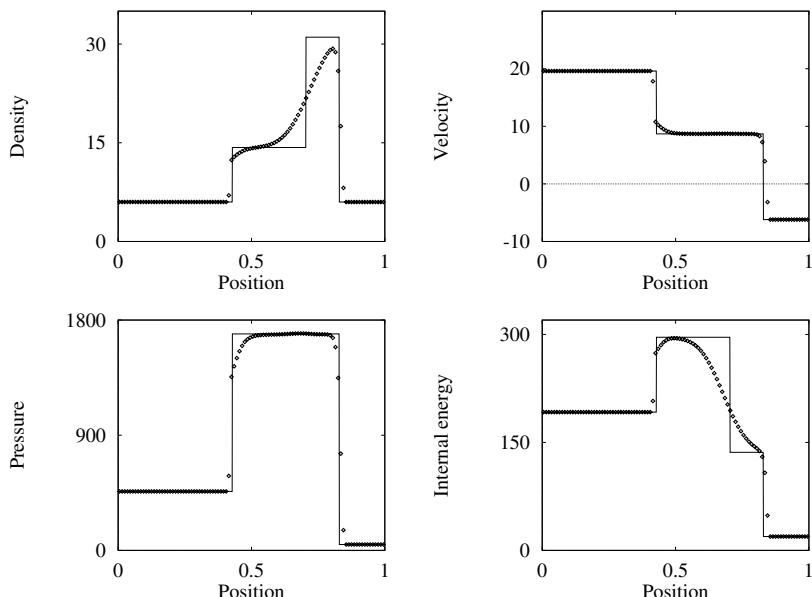
**Fig. 10.15.** Rusanov's method applied to Test 1, with  $x_0 = 0.3$ . Numerical (symbol) and exact (line) solutions are compared at time 0.2.



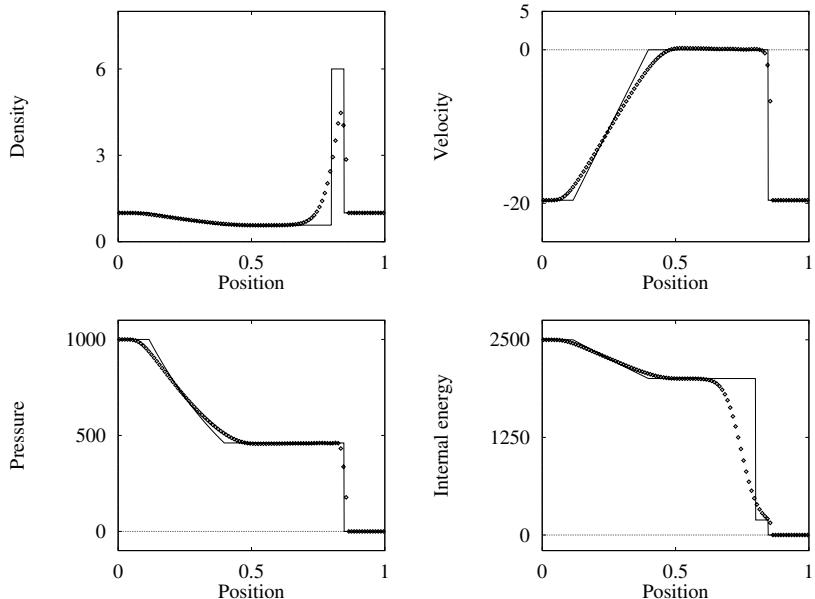
**Fig. 10.16.** Rusanov's method applied to Test 2, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at time 0.15.



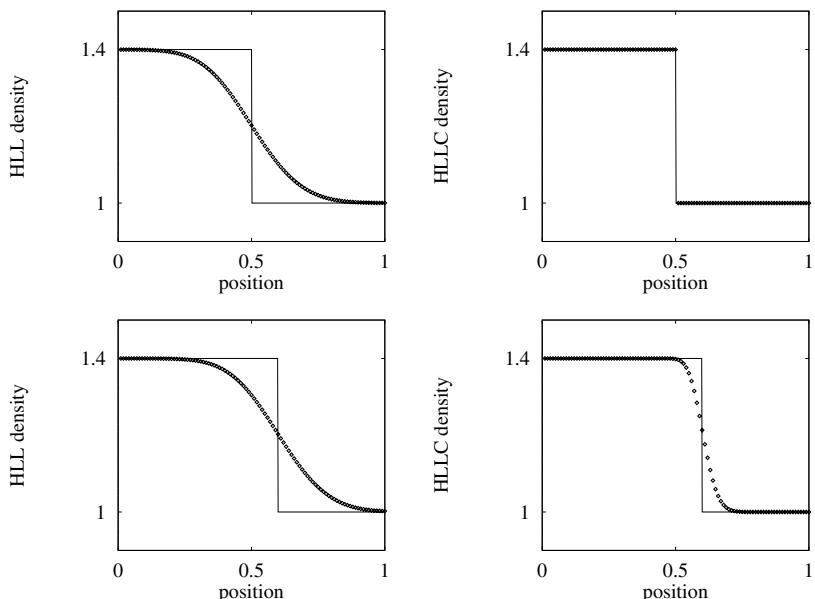
**Fig. 10.17.** Rusanov's method applied to Test 3, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012.



**Fig. 10.18.** Rusanov's method applied to Test 4, with  $x_0 = 0.4$ . Numerical (symbol) and exact (line) solutions are compared at time 0.035.



**Fig. 10.19.** Rusanov's method applied to Test 5, with  $x_0 = 0.8$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012.



**Fig. 10.20.** Godunov's method with HLL (left) and HLLC (right) Riemann solvers applied to Tests 6 and 7. Numerical (symbol) and exact (line) solutions are compared at time 2.0.

---

## The Riemann Solver of Roe

Perhaps, the most well-known of all approximate Riemann solvers today, is the one due to Roe, which was first presented in the open literature in 1981 [407]. Since then, the method has not only been refined, but it has also been applied to a very large variety of physical problems. Refinements to the Roe approach were introduced by Roe and Pike [416], whereby the computation of the necessary items of information does not explicitly require the Roe averaged Jacobian matrix. This second methodology appears to be simpler and is thus useful in solving the Riemann problem for new, complicated sets of hyperbolic conservations laws, or for conventional systems but for complex media. Glaister exploited the Roe–Pike approach to extend Roe’s method to the time-dependent Euler equations with a general equation of state [208], [209]. The large body of experience accumulated by many workers over a considerable period of time has led to various improvements of the scheme. As originally presented the Roe scheme computes rarefaction shocks, thus violating the entropy condition. Harten and Hyman [243], Roe and Pike [416], Roe [414], Dubois and Mehlman [167] and others, have produced appropriate modifications to the scheme. Einfeldt et. al. [182] produced corrections to the basic Roe scheme to avoid the so-called *vacuum problem* near low-density flows; they also showed that in fact this anomaly afflicts *all linearised Riemann solvers*.

Ambitious applications of the Roe scheme were presented by Brio and Wu [80], who utilised Roe’s method to solve the Magneto-Hydrodynamic equations (MHD). Clarke et. al. [118] applied the method in conjunction with adaptive gridding to the computation of two-dimensional unsteady detonation waves in solid materials. Giraud and Manzini [206] produced parallel implementations of the Roe scheme for two-dimensional Gas Dynamics. LeVeque and Shyue [313] have applied the Roe scheme in the context of *front tracking* in two space dimensions. Marx has applied the Roe scheme to solve the incompressible Navier–Stokes equations [345], [346] and the compressible Navier–Stokes equations [344] using *implicit* versions of the scheme; see also McNeil [348]. The method has also been applied to multiphase flows; Toro

[501] solved reactive multi-phase problems in the context of propulsion systems via a phase-splitting procedure; recently, Sainsaulieu [421] has extended the Roe scheme to a class of multiphase flow problems without phase splitting.

The purpose of this chapter is to present the approximate Riemann solver of Roe as applied to the three-dimensional time dependent Euler equations. For the numerical methods considered here, we only need to derive the Riemann solver for the split three-dimensional equations. After a general introduction to the method, we present both the methodology of Roe and that of Roe and Pike. Both methodologies are suitably illustrated via the simpler isothermal equations. Useful background reading is found in Chaps. 2 to 6.

## 11.1 Bases of the Roe Approach

In this section we describe the Roe approach for a general system of  $m$  hyperbolic conservation laws. Detailed application of the scheme to the isothermal and Euler equations are given in subsequent sections.

### 11.1.1 The Exact Riemann Problem and the Godunov Flux

We are concerned with solving numerically the general Initial Boundary Value Problem (IBVP)

$$\left. \begin{array}{l} \text{PDEs : } \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0}, \\ \text{ICs : } \mathbf{U}(x, 0) = \mathbf{U}^{(0)}(x), \\ \text{BCs : } \mathbf{U}(0, t) = \mathbf{U}_l(t), \quad \mathbf{U}(L, t) = \mathbf{U}_r(t), \end{array} \right\} \quad (11.1)$$

in a domain  $x_l \leq x \leq x_r$ , utilising the explicit conservative formula

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n + \frac{\Delta t}{\Delta x} [\mathbf{F}_{i-\frac{1}{2}} - \mathbf{F}_{i+\frac{1}{2}}]. \quad (11.2)$$

We assume the solution of IBVP (11.1) exists. In Chap. 6 we defined the Godunov intercell numerical flux

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}(0)), \quad (11.3)$$

in which  $\mathbf{U}_{i+\frac{1}{2}}(0)$  is the exact similarity solution  $\mathbf{U}_{i+\frac{1}{2}}(x/t)$  of the Riemann problem

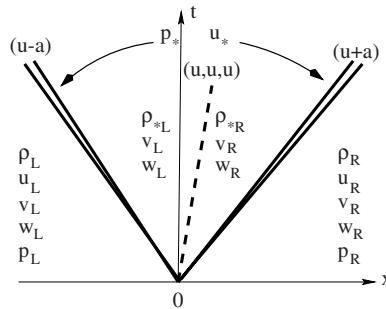
$$\left. \begin{array}{l} \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0}, \\ \mathbf{U}(x, 0) = \begin{cases} \mathbf{U}_L & \text{if } x < 0, \\ \mathbf{U}_R & \text{if } x > 0 \end{cases} \end{array} \right\} \quad (11.4)$$

evaluated at  $x/t = 0$ . Fig. 11.1 shows the structure of the exact solution of the Riemann problem for the  $x$ -split three dimensional Euler equations, for which the vectors of conserved variables and fluxes are

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ E \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ \rho uw \\ u(E + p) \end{bmatrix}. \quad (11.5)$$

The *Star Region* between the left and right waves contains the unknowns of the problem. The particular value at  $x/t = 0$  corresponds to the  $t$ -axis and is the value required by the Godunov flux. See Chaps. 4 and 6 for details. The piece-wise constant initial data, in terms of primitive variables, is

$$\mathbf{W}_L = \begin{bmatrix} \rho_L \\ u_L \\ v_L \\ w_L \\ p_L \end{bmatrix}, \quad \mathbf{W}_R = \begin{bmatrix} \rho_R \\ u_R \\ v_R \\ w_R \\ p_R \end{bmatrix}. \quad (11.6)$$



**Fig. 11.1.** Structure of the solution of the Riemann problem for the  $x$ -split three dimensional Euler equations

In Chap. 4 we provided an algorithm to compute the exact solution  $\mathbf{U}_{i+\frac{1}{2}}(x/t)$  and in Chap. 6 we utilised this solution in the Godunov method. In Chap. 9 we provided approximations to the state  $\mathbf{U}_{i+\frac{1}{2}}(x/t)$  and obtained a corresponding approximate Godunov method by evaluating the physical flux function  $\mathbf{F}$  at this approximate state. The purpose of this chapter is to find *direct approximations to the flux function*  $\mathbf{F}_{i+\frac{1}{2}}$  following the approach proposed by Roe [407] and Roe and Pike [416].

### 11.1.2 Approximate Conservation Laws

Roe [407] solved the Riemann problem (11.4) approximately. By introducing the Jacobian matrix

$$\mathbf{A}(\mathbf{U}) = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} \quad (11.7)$$

and using the chain rule the conservation laws

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0}$$

in (11.4) may be written as

$$\mathbf{U}_t + \mathbf{A}(\mathbf{U})\mathbf{U}_x = \mathbf{0}. \quad (11.8)$$

Roe's approach replaces the Jacobian matrix  $\mathbf{A}(\mathbf{U})$  in (11.8) by a *constant* Jacobian matrix

$$\tilde{\mathbf{A}} = \tilde{\mathbf{A}}(\mathbf{U}_L, \mathbf{U}_R), \quad (11.9)$$

which is a function of the data states  $\mathbf{U}_L, \mathbf{U}_R$ . In this way the original PDEs in (11.4) are replaced by

$$\mathbf{U}_t + \tilde{\mathbf{A}}\mathbf{U}_x = \mathbf{0}. \quad (11.10)$$

This is a linear system with constant coefficients. The original, Riemann problem (11.4) is then replaced by the *approximate Riemann problem*

$$\left. \begin{aligned} \mathbf{U}_t + \tilde{\mathbf{A}}\mathbf{U}_x &= \mathbf{0} \\ \mathbf{U}(x, 0) &= \begin{cases} \mathbf{U}_L, & x < 0 \\ \mathbf{U}_R, & x > 0 \end{cases} \end{aligned} \right\}, \quad (11.11)$$

which is then *solved exactly*. The approximate problem results from replacing the original non-linear conservation laws by a linearised system with constant coefficients but the initial data of the exact problem is retained.

For a general hyperbolic system of  $m$  conservation laws, the Roe Jacobian matrix  $\tilde{\mathbf{A}}$  is required to satisfy the following properties:

Property (A): Hyperbolicity of the system.  $\tilde{\mathbf{A}}$  is required to have real eigenvalues  $\tilde{\lambda}_i = \tilde{\lambda}_i(\mathbf{U}_L, \mathbf{U}_R)$ , which we choose to order as

$$\tilde{\lambda}_1 \leq \tilde{\lambda}_2 \leq \cdots \leq \tilde{\lambda}_m, \quad (11.12)$$

and a complete set of linearly independent right eigenvectors

$$\tilde{\mathbf{K}}^{(1)}, \tilde{\mathbf{K}}^{(2)}, \dots, \tilde{\mathbf{K}}^{(m)}. \quad (11.13)$$

Property (B): Consistency with the exact Jacobian

$$\tilde{\mathbf{A}}(\mathbf{U}, \mathbf{U}) = \mathbf{A}(\mathbf{U}). \quad (11.14)$$

Property (C): Conservation across discontinuities

$$\mathbf{F}(\mathbf{U}_R) - \mathbf{F}(\mathbf{U}_L) = \tilde{\mathbf{A}}(\mathbf{U}_R - \mathbf{U}_L). \quad (11.15)$$

Property (A) on hyperbolicity is an obvious requirement; the approximate problem should at the very least preserve the mathematical character of the

original non-linear system. Property (B) ensures consistency with the conservation laws. Property (C) ensures conservation. It also ensures exact recognition of isolated discontinuities; that is, if the data  $\mathbf{U}_L, \mathbf{U}_R$  are connected by a *single, isolated* discontinuity, then the approximate Riemann solver recognises this wave exactly. Note however that this does not mean that the corresponding, approximate, Godunov method with the Roe approximate numerical flux will in general give exact solutions for isolated discontinuities.

The construction of matrices satisfying properties (A)–(C) for *general hyperbolic systems* can be very complicated and thus computationally unattractive. For the *specific* case of the Euler equations of Gas Dynamics Roe [407] proposed a relatively simple way of constructing a matrix  $\tilde{\mathbf{A}}$ . Later, Roe and Pike [416] proposed a simpler approach, where the explicit construction of  $\tilde{\mathbf{A}}$  is actually avoided.

### 11.1.3 The Approximate Riemann Problem and the Intercell Flux

Once the matrix  $\tilde{\mathbf{A}}(\mathbf{U}_L, \mathbf{U}_R)$ , its eigenvalues  $\tilde{\lambda}_i(\mathbf{U}_L, \mathbf{U}_R)$  and right eigenvectors  $\tilde{\mathbf{K}}^{(i)}(\mathbf{U}_L, \mathbf{U}_R)$  are available, one solves the Riemann problem (11.11) by direct application of methods discussed in Sect. 2.3 of Chap. 2 and Sect. 5.4 of Chap. 5, for linear hyperbolic systems with constant coefficients. By projecting the data difference

$$\Delta \mathbf{U} = \mathbf{U}_R - \mathbf{U}_L$$

onto the right eigenvectors we write

$$\Delta \mathbf{U} = \mathbf{U}_R - \mathbf{U}_L = \sum_{i=1}^m \tilde{\alpha}_i \tilde{\mathbf{K}}^{(i)}, \quad (11.16)$$

from which one finds the *wave strengths*  $\tilde{\alpha}_i = \tilde{\alpha}_i(\mathbf{U}_L, \mathbf{U}_R)$ . The solution  $\mathbf{U}_{i+\frac{1}{2}}(x/t)$  evaluated along the  $t$ -axis,  $x/t = 0$ , is given by

$$\mathbf{U}_{i+\frac{1}{2}}(0) = \mathbf{U}_L + \sum_{\tilde{\lambda}_i \leq 0} \tilde{\alpha}_i \tilde{\mathbf{K}}^{(i)}, \quad (11.17)$$

or

$$\mathbf{U}_{i+\frac{1}{2}}(0) = \mathbf{U}_R - \sum_{\tilde{\lambda}_i \geq 0} \tilde{\alpha}_i \tilde{\mathbf{K}}^{(i)}. \quad (11.18)$$

We now find the corresponding numerical flux. Recall that we have replaced the original set of conservation laws in (11.4) by the constant coefficient linear system (11.10); this can be viewed as a modified system of conservation laws

$$\bar{\mathbf{U}}_t + \bar{\mathbf{F}}(\bar{\mathbf{U}})_x = \mathbf{0}, \quad (11.19)$$

with flux function

$$\bar{\mathbf{F}}(\bar{\mathbf{U}}) = \tilde{\mathbf{A}}\bar{\mathbf{U}}. \quad (11.20)$$

The corresponding numerical flux, see (11.3), *is not* the obvious choice

$$\mathbf{F}_{i+\frac{1}{2}} = \tilde{\mathbf{A}}\bar{\mathbf{U}}_{i+\frac{1}{2}}(0),$$

where  $\bar{\mathbf{U}}_{i+\frac{1}{2}}(0)$  is given by any of (11.17)–(11.18). That this would be incorrect becomes obvious when, for instance, assuming right supersonic flow in (11.17) one would compute an intercell flux  $\mathbf{F}_{i+\frac{1}{2}} \neq \mathbf{F}_L$ . Instead, the correct expression for the corresponding numerical flux is obtained from any of the integral relations

$$\mathbf{F}_{0L} = \mathbf{F}_L - S_L \mathbf{U}_L - \frac{1}{T} \int_{TS_L}^0 \mathbf{U}(x, T) dx, \quad (11.21)$$

$$\mathbf{F}_{0R} = \mathbf{F}_R - S_R \mathbf{U}_R + \frac{1}{T} \int_0^{TS_R} \mathbf{U}(x, T) dx, \quad (11.22)$$

derived in Sect. 10.2 of Chap. 10. Here  $S_L, S_R$  are the smallest and largest signal speeds in the exact solution of the Riemann problem with data  $\mathbf{U}_L, \mathbf{U}_R$  and  $T$  is a positive time. If the integrand  $\mathbf{U}(x, t)$  in (11.21) or (11.22) is replaced by some approximate solution, then equality of the fluxes  $\mathbf{F}_{0L}$  and  $\mathbf{F}_{0R}$  requires the approximate solution to satisfy a *Consistency Condition*, see Sect. 10.2 of Chap. 10.

If  $\bar{\mathbf{U}}_{i+\frac{1}{2}}(x, t)$  is the solution of the Riemann problem for the modified conservation laws (11.19) with data  $\mathbf{U}_L, \mathbf{U}_R$ , then the integrals in (11.21) and (11.22) respectively, are

$$\int_{TS_L}^0 \bar{\mathbf{U}}_{i+\frac{1}{2}}(x, T) dx = T[\bar{\mathbf{F}}(\mathbf{U}_L) - \bar{\mathbf{F}}(\bar{\mathbf{U}}_{i+\frac{1}{2}}(0))] - TS_L \mathbf{U}_L \quad (11.23)$$

and

$$\int_0^{TS_R} \bar{\mathbf{U}}_{i+\frac{1}{2}}(x, T) dx = T[\bar{\mathbf{F}}(\bar{\mathbf{U}}_{i+\frac{1}{2}}(0)) - \bar{\mathbf{F}}(\mathbf{U}_R)] + TS_R \mathbf{U}_R. \quad (11.24)$$

Substitution of (11.23) and (11.24) into (11.21) and (11.22) gives

$$\mathbf{F}_{0L} = \bar{\mathbf{F}}(\bar{\mathbf{U}}_{i+\frac{1}{2}}(0)) + \mathbf{F}(\mathbf{U}_L) - \bar{\mathbf{F}}(\mathbf{U}_L) \quad (11.25)$$

and

$$\mathbf{F}_{0R} = \bar{\mathbf{F}}(\bar{\mathbf{U}}_{i+\frac{1}{2}}(0)) + \mathbf{F}(\mathbf{U}_R) - \bar{\mathbf{F}}(\mathbf{U}_R). \quad (11.26)$$

Finally, by using  $\bar{\mathbf{U}}_{i+\frac{1}{2}}(0)$  as given by (11.17) or (11.18) and the definition of the flux  $\bar{\mathbf{F}} = \tilde{\mathbf{A}}\bar{\mathbf{U}}$  we obtain the numerical flux as

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}_L + \sum_{\tilde{\lambda}_i \leq 0} \tilde{\alpha}_i \tilde{\lambda}_i \tilde{\mathbf{K}}^{(i)}, \quad (11.27)$$

or

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}_R - \sum_{\tilde{\lambda}_i \geq 0} \tilde{\alpha}_i \tilde{\lambda}_i \tilde{\mathbf{K}}^{(i)}. \quad (11.28)$$

Alternatively, we may also write

$$\mathbf{F}_{i+\frac{1}{2}} = \frac{1}{2}(\mathbf{F}_L + \mathbf{F}_R) - \frac{1}{2} \sum_{i=1}^m \tilde{\alpha}_i |\tilde{\lambda}_i| \tilde{\mathbf{K}}^{(i)}. \quad (11.29)$$

We remark that all previous relations (11.19)–(11.29) are valid for any hyperbolic system and any linearisation of it. In order to compute Roe's numerical flux for a particular system of hyperbolic conservation laws, one requires expressions for the wave strengths  $\tilde{\alpha}_i$ , the eigenvalues  $\tilde{\lambda}_i$  and the right eigenvectors  $\tilde{\mathbf{K}}^{(i)}$  in any of the flux expressions (11.27)–(11.29). Note that the Jacobian matrix  $\tilde{\mathbf{A}}(\mathbf{U}_L, \mathbf{U}_R)$  is not explicitly required by the numerical flux. In the next two sections we give details on methodologies to find  $\tilde{\alpha}_i$ ,  $\tilde{\lambda}_i$  and  $\tilde{\mathbf{K}}^{(i)}$ . There are two approaches, namely the original approach presented by Roe in 1981 [407] and the Roe–Pike approach [416].

## 11.2 The Original Roe Method

In order for the approximate Godunov method based on (11.2) with the Roe-type numerical flux (11.27)–(11.29) to be completely determined, we need to find the average eigenvalues  $\tilde{\lambda}_i$ , the corresponding averaged right eigenvectors  $\tilde{\mathbf{K}}^{(i)}$  and averaged wave strengths  $\tilde{\alpha}_i$ . In his original paper [407] Roe finds an averaged Jacobian matrix  $\tilde{\mathbf{A}}$ , the Roe matrix, from which  $\tilde{\lambda}_i$ ,  $\tilde{\mathbf{K}}^{(i)}$  and  $\tilde{\alpha}_i$  follow directly. In constructing a matrix  $\tilde{\mathbf{A}}$  the properties (A)–(C), equations (11.12)–(11.15), are enforced. It is not difficult to think of candidates  $\tilde{\mathbf{A}}$  that satisfy the first two properties. Property C is crucial and is the one that narrows down the choices. Roe showed that the existence of a matrix  $\tilde{\mathbf{A}}$  satisfying Property C is assured by the mean value theorem. An early line of attack in constructing a matrix  $\tilde{\mathbf{A}}$  satisfying all desirable properties is reported by Sells [440]. Roe identifies some disadvantages of this approach; it is argued, for instance, that the construction is far from unique and that the resulting schemes are too complicated.

A breakthrough in constructing  $\tilde{\mathbf{A}}$  resulted from Roe's ingenious idea of introducing a *parameter vector*  $\mathbf{Q}$ , such that both the vector of conserved variables  $\mathbf{U}$  and the flux vector  $\mathbf{F}(\mathbf{U})$  could be expressed in terms of  $\mathbf{Q}$ . That is

$$\mathbf{U} = \mathbf{U}(\mathbf{Q}), \quad \mathbf{F} = \mathbf{F}(\mathbf{Q}). \quad (11.30)$$

Two important steps then follow. First, the changes

$$\Delta \mathbf{U} = \mathbf{U}_R - \mathbf{U}_L, \quad \Delta \mathbf{F} = \mathbf{F}(\mathbf{U}_R) - \mathbf{F}(\mathbf{U}_L) \quad (11.31)$$

can be expressed in terms of the change  $\Delta \mathbf{Q} = \mathbf{Q}_R - \mathbf{Q}_L$ . Then, averages are obtained in terms of *simple arithmetic means* of  $\mathbf{Q}$ . Next, we illustrate the technique as applied to a simple set of conservation laws.

### 11.2.1 The Isothermal Equations

Consider the isothermal equations

$$\left. \begin{aligned} \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x &= \mathbf{0}, \\ \mathbf{U} \equiv \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} &\equiv \begin{bmatrix} \rho \\ \rho u \end{bmatrix}; \quad \mathbf{F} \equiv \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \equiv \begin{bmatrix} \rho u \\ \rho u^2 + a^2 \rho \end{bmatrix}, \end{aligned} \right\} \quad (11.32)$$

where  $a$  is a constant sound speed. See Sect. 1.6.2 of Chap. 1. See also Sect. 2.4.1 of Chap. 2, where the eigenstructure of the equations is given. The exact Jacobian, eigenvalues and corresponding right eigenvectors are

$$\left. \begin{aligned} \mathbf{A}(\mathbf{U}) &= \begin{bmatrix} 0 & 1 \\ a^2 - u^2 & 2u \end{bmatrix}, \\ \lambda_1 = u - a, \quad \lambda_2 = u + a, \\ \mathbf{K}^{(1)} = \begin{bmatrix} 1 \\ u - a \end{bmatrix}, \quad \mathbf{K}^{(2)} = \begin{bmatrix} 1 \\ u + a \end{bmatrix}. \end{aligned} \right\} \quad (11.33)$$

Choose the *parameter vector*

$$\mathbf{Q} \equiv \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \equiv \frac{\mathbf{U}}{\sqrt{\rho}} = \begin{bmatrix} \sqrt{\rho} \\ \sqrt{\rho}u \end{bmatrix}. \quad (11.34)$$

Then  $\mathbf{U}$  and  $\mathbf{F}$  can be expressed in terms of the components  $q_1, q_2$  of  $\mathbf{Q}$ , namely

$$\mathbf{U} \equiv \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \equiv q_1 \mathbf{Q} = \begin{bmatrix} q_1^2 \\ q_1 q_2 \end{bmatrix} \quad (11.35)$$

and

$$\mathbf{F} \equiv \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \equiv \begin{bmatrix} q_1 q_2 \\ q_2^2 + a^2 q_1^2 \end{bmatrix}. \quad (11.36)$$

One now looks for an averaged vector  $\tilde{\mathbf{Q}} = (\tilde{q}_1, \tilde{q}_2)^T$ . This is found by simple *arithmetic averaging*

$$\tilde{\mathbf{Q}} = \begin{bmatrix} \tilde{q}_1 \\ \tilde{q}_2 \end{bmatrix} = \frac{1}{2}(\mathbf{Q}_L + \mathbf{Q}_R) = \frac{1}{2} \begin{bmatrix} \sqrt{\rho_L} + \sqrt{\rho_R} \\ \sqrt{\rho_L}u_L + \sqrt{\rho_R}u_R \end{bmatrix}. \quad (11.37)$$

Then two matrices  $\tilde{\mathbf{B}} = \tilde{\mathbf{B}}(\tilde{\mathbf{Q}})$  and  $\tilde{\mathbf{C}} = \tilde{\mathbf{C}}(\tilde{\mathbf{Q}})$  are found, such that the jumps  $\Delta \mathbf{U}$  and  $\Delta \mathbf{F}$  in (11.31) can be expressed in terms of the jump  $\Delta \mathbf{Q}$ , namely

$$\Delta \mathbf{U} = \tilde{\mathbf{B}} \Delta \mathbf{Q}; \quad \Delta \mathbf{F} = \tilde{\mathbf{C}} \Delta \mathbf{Q}. \quad (11.38)$$

Use of these two expressions produces

$$\Delta \mathbf{F} = (\tilde{\mathbf{C}} \tilde{\mathbf{B}}^{-1}) \Delta \mathbf{U}, \quad (11.39)$$

which if compared with condition (C), equation (11.15), produces the Roe averaged matrix

$$\tilde{\mathbf{A}} = \tilde{\mathbf{C}} \tilde{\mathbf{B}}^{-1}. \quad (11.40)$$

Matrices  $\tilde{\mathbf{B}}$  and  $\tilde{\mathbf{C}}$  satisfying (11.38) are

$$\tilde{\mathbf{B}} = \begin{bmatrix} 2\tilde{q}_1 & 0 \\ \tilde{q}_2 & \tilde{q}_1 \end{bmatrix}; \quad \tilde{\mathbf{C}} = \begin{bmatrix} \tilde{q}_2 & \tilde{q}_1 \\ 2a^2\tilde{q}_1 & 2\tilde{q}_2^2 \end{bmatrix}, \quad (11.41)$$

which the reader can easily verify. The sought Roe matrix is then

$$\tilde{\mathbf{A}} = \begin{bmatrix} 0 & 1 \\ a^2 - \tilde{u}^2 & 2\tilde{u} \end{bmatrix}, \quad (11.42)$$

where  $\tilde{u}$  is the *Roe averaged velocity* and is given by

$$\tilde{u} = \frac{\sqrt{\rho_L} u_L + \sqrt{\rho_R} u_R}{\sqrt{\rho_L} + \sqrt{\rho_R}}. \quad (11.43)$$

Compare (11.42) with the matrix in (11.33). As the sound speed  $a$  is constant, no averaged  $\tilde{\rho}$  is required.

Having found  $\tilde{\mathbf{A}}$  one computes the averaged eigenvalues, eigenvectors and wave strengths. The eigenvalues of  $\tilde{\mathbf{A}}$  are

$$\tilde{\lambda}_1 = \tilde{u} - a; \quad \tilde{\lambda}_2 = \tilde{u} + a \quad (11.44)$$

and are all real. The corresponding averaged right eigenvectors are

$$\tilde{\mathbf{K}}^{(1)} = \begin{bmatrix} 1 \\ \tilde{u} - a \end{bmatrix}; \quad \tilde{\mathbf{K}}^{(2)} = \begin{bmatrix} 1 \\ \tilde{u} + a \end{bmatrix} \quad (11.45)$$

and are easily seen to be linearly independent. Thus condition (A) is satisfied. To find the wave strengths  $\tilde{\alpha}_i$  we solve the  $2 \times 2$  linear system, see (11.16),

$$\Delta \mathbf{U} \equiv \begin{bmatrix} \Delta u_1 \\ \Delta u_2 \end{bmatrix} = \sum_{i=1}^2 \tilde{\alpha}_i \tilde{\mathbf{K}}^{(i)}.$$

The solution is easily verified to be

$$\left. \begin{aligned} \tilde{\alpha}_1 &= \frac{\Delta u_1(\tilde{u} + a) - \Delta u_2}{2a}, \\ \tilde{\alpha}_2 &= \frac{-\Delta u_1(\tilde{u} - a) + \Delta u_2}{2a}, \end{aligned} \right\} \quad (11.46)$$

with the obvious definitions  $\Delta u_1 \equiv \rho_R - \rho_L$ ,  $\Delta u_2 \equiv \rho_R u_R - \rho_L u_L$ . The corresponding Roe numerical flux  $\mathbf{F}_{i+\frac{1}{2}}$  now follows from using (11.43)–(11.46) into any of the expressions (11.27)–(11.29).

### 11.2.2 The Euler Equations

Here we present the Roe Riemann solver as applied to the Riemann problem (11.4)–(11.5) for the  $x$ -split three dimensional time dependent Euler equations for ideal gases. Details of the Euler equations are found in Sect. 1.1 and Sect. 1.2 of Chap. 1; mathematical properties of the Euler equations are studied in Chap. 3.

The exact,  $x$ -direction Jacobian matrix  $\mathbf{A}(\mathbf{U})$  is

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ \hat{\gamma}H - u^2 - a^2 & (3 - \gamma)u & -\hat{\gamma}v & -\hat{\gamma}w & \hat{\gamma} \\ -uv & v & u & 0 & 0 \\ -uw & w & 0 & u & 0 \\ \frac{1}{2}u[(\gamma - 3)H - a^2] & H - \hat{\gamma}u^2 & -\hat{\gamma}uv & -\hat{\gamma}uw & \gamma u \end{bmatrix}, \quad (11.47)$$

where  $\hat{\gamma} = \gamma - 1$ . The eigenvalues are

$$\lambda_1 = u - a, \quad \lambda_2 = \lambda_3 = \lambda_4 = u, \quad \lambda_5 = u + a, \quad (11.48)$$

where  $a = \sqrt{\gamma p/\rho}$  is the sound speed. The corresponding right eigenvectors are

$$\left. \begin{aligned} \mathbf{K}^{(1)} &= \begin{bmatrix} 1 \\ u - a \\ v \\ w \\ H - ua \end{bmatrix}; \quad \mathbf{K}^{(2)} = \begin{bmatrix} 1 \\ u \\ v \\ w \\ \frac{1}{2}V^2 \end{bmatrix}; \quad \mathbf{K}^{(3)} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ v \end{bmatrix} \\ \mathbf{K}^{(4)} &= \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ w \end{bmatrix}; \quad \mathbf{K}^{(5)} = \begin{bmatrix} 1 \\ u + a \\ v \\ w \\ H + ua \end{bmatrix}. \end{aligned} \right\} \quad (11.49)$$

Here  $H$  is the total enthalpy

$$H = \frac{E + p}{\rho} \quad (11.50)$$

and  $E$  is the total energy per unit volume

$$E = \frac{1}{2}\rho\mathbf{V}^2 + \rho e, \quad (11.51)$$

with

$$\mathbf{V}^2 = u^2 + v^2 + w^2 \quad (11.52)$$

and  $e$  denoting the specific internal energy, which for ideal gases, see Sect. 1.2 of Chap. 1, is

$$e = \frac{p}{(\gamma - 1)\rho} . \quad (11.53)$$

Roe chooses the *parameter vector*

$$\mathbf{Q} \equiv \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \end{bmatrix} \equiv \sqrt{\rho} \begin{bmatrix} 1 \\ u \\ v \\ w \\ H \end{bmatrix} , \quad (11.54)$$

which has the property that every component  $u_i$  of  $\mathbf{U}$  and every component  $f_i$  of  $\mathbf{F}(\mathbf{U})$  in (11.4)–(11.5) is a quadratic in the components  $q_i$  of  $\mathbf{Q}$ . For instance  $u_1 = q_1^2$  and  $f_1 = q_1 q_2$ , etc. Actually, the property is also valid for the components of the  $\mathbf{G}$  and  $\mathbf{H}$  fluxes for the full three-dimensional Euler equations.

As done for the isothermal equations, see equations (11.38), one can express the jumps  $\Delta\mathbf{U}$  and  $\Delta\mathbf{F}$  in terms of the jump  $\Delta\mathbf{Q}$  via two matrices  $\tilde{\mathbf{B}}$  and  $\tilde{\mathbf{C}}$ . Roe [407] gives the following expressions

$$\tilde{\mathbf{B}} = \begin{pmatrix} 2\tilde{q}_1 & 0 & 0 & 0 & 0 \\ \tilde{q}_2 & \tilde{q}_1 & 0 & 0 & 0 \\ \tilde{q}_3 & 0 & \tilde{q}_1 & 0 & 0 \\ \tilde{q}_4 & 0 & 0 & \tilde{q}_1 & 0 \\ \frac{\tilde{q}_5}{\gamma} & \frac{\gamma-1}{\gamma}\tilde{q}_2 & \frac{\gamma-1}{\gamma}\tilde{q}_3 & \frac{\gamma-1}{\gamma}\tilde{q}_4 & \frac{\tilde{q}_1}{\gamma} \end{pmatrix} \quad (11.55)$$

and

$$\tilde{\mathbf{C}} = \begin{pmatrix} \tilde{q}_2 & \tilde{q}_1 & 0 & 0 & 0 \\ \frac{\gamma-1}{\gamma}\tilde{q}_5 & \frac{\gamma+1}{\gamma}\tilde{q}_2 & -\frac{\gamma-1}{\gamma}\tilde{q}_3 & -\frac{\gamma-1}{\gamma}\tilde{q}_4 & \frac{\gamma-1}{\gamma}\tilde{q}_1 \\ 0 & \tilde{q}_3 & \tilde{q}_2 & 0 & 0 \\ 0 & \tilde{q}_4 & 0 & \tilde{q}_2 & 0 \\ 0 & \tilde{q}_5 & 0 & 0 & \tilde{q}_2 \end{pmatrix} . \quad (11.56)$$

The sought Roe matrix is then given by

$$\tilde{\mathbf{A}} = \tilde{\mathbf{B}} \tilde{\mathbf{C}}^{-1} . \quad (11.57)$$

The eigenvalues of  $\tilde{\mathbf{A}}$  are

$$\tilde{\lambda}_1 = \tilde{u} - \tilde{a} , \quad \tilde{\lambda}_2 = \tilde{\lambda}_3 = \tilde{\lambda}_4 = \tilde{u} , \quad \tilde{\lambda}_5 = \tilde{u} + \tilde{a} \quad (11.58)$$

and the corresponding right eigenvectors are

$$\left. \begin{aligned} \tilde{\mathbf{K}}^{(1)} &= \begin{bmatrix} 1 \\ \tilde{u} - \tilde{a} \\ \tilde{v} \\ \tilde{w} \\ \tilde{H} - \tilde{u}\tilde{a} \end{bmatrix}; \quad \tilde{\mathbf{K}}^{(2)} = \begin{bmatrix} 1 \\ \tilde{u} \\ \tilde{v} \\ \tilde{w} \\ \frac{1}{2}\tilde{\mathbf{V}}^2 \end{bmatrix}; \quad \tilde{\mathbf{K}}^{(3)} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \tilde{v} \end{bmatrix} \\ \tilde{\mathbf{K}}^{(4)} &= \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ \tilde{w} \end{bmatrix}; \quad \tilde{\mathbf{K}}^{(5)} = \begin{bmatrix} 1 \\ \tilde{u} + \tilde{a} \\ \tilde{v} \\ \tilde{w} \\ \tilde{H} + \tilde{u}\tilde{a} \end{bmatrix}. \end{aligned} \right\} \quad (11.59)$$

The symbol  $\tilde{r}$  in (11.58), (11.59) denotes a Roe average for a variable  $r$ . The relevant averages are given as follows

$$\left. \begin{aligned} \tilde{u} &= \frac{\sqrt{\rho_L}u_L + \sqrt{\rho_R}u_R}{\sqrt{\rho_L} + \sqrt{\rho_R}}, \\ \tilde{v} &= \frac{\sqrt{\rho_L}v_L + \sqrt{\rho_R}v_R}{\sqrt{\rho_L} + \sqrt{\rho_R}}, \\ \tilde{w} &= \frac{\sqrt{\rho_L}w_L + \sqrt{\rho_R}w_R}{\sqrt{\rho_L} + \sqrt{\rho_R}}, \\ \tilde{H} &= \frac{\sqrt{\rho_L}H_L + \sqrt{\rho_R}H_R}{\sqrt{\rho_L} + \sqrt{\rho_R}}, \\ \tilde{a} &= \left( (\gamma - 1)(\tilde{H} - \frac{1}{2}\tilde{\mathbf{V}}^2) \right)^{\frac{1}{2}}, \end{aligned} \right\} \quad (11.60)$$

where  $\tilde{\mathbf{V}}^2 = \tilde{u}^2 + \tilde{v}^2 + \tilde{w}^2$ .

In order to determine completely the Roe numerical flux  $\mathbf{F}_{i+\frac{1}{2}}$  we need, in addition, the wave strengths  $\tilde{\alpha}_i$ . These are obtained by projecting the jump  $\Delta\mathbf{U}$  onto the right, averaged eigenvectors (11.59), namely

$$\Delta\mathbf{U} = \sum_{i=1}^5 \tilde{\alpha}_i \tilde{\mathbf{K}}^{(i)}. \quad (11.61)$$

When written in full these equations read

$$\tilde{\alpha}_1 + \tilde{\alpha}_2 + \tilde{\alpha}_5 = \Delta u_1, \quad (11.62)$$

$$\tilde{\alpha}_1(\tilde{u} - \tilde{a}) + \tilde{\alpha}_2\tilde{u} + \tilde{\alpha}_5(\tilde{u} + \tilde{a}) = \Delta u_2, \quad (11.63)$$

$$\tilde{\alpha}_1\tilde{v} + \tilde{\alpha}_2\tilde{v} + \tilde{\alpha}_3 + \tilde{\alpha}_5\tilde{v} = \Delta u_3, \quad (11.64)$$

$$\tilde{\alpha}_1\tilde{w} + \tilde{\alpha}_2\tilde{w} + \tilde{\alpha}_4 + \tilde{\alpha}_5\tilde{w} = \Delta u_4, \quad (11.65)$$

$$\tilde{\alpha}_1(\tilde{H} - \tilde{u}\tilde{a}) + \frac{1}{2}\tilde{\mathbf{V}}^2\tilde{\alpha}_2 + \tilde{\alpha}_3\tilde{v} + \tilde{\alpha}_4\tilde{w} + \tilde{\alpha}_5(\tilde{H} + \tilde{u}\tilde{a}) = \Delta u_5 . \quad (11.66)$$

Here the right-hand side terms of equations (11.62)–(11.66) are known: they are jumps  $\Delta u_i$  in the conserved quantity  $u_i$ , namely

$$\Delta u_i = (u_i)_R - (u_i)_L .$$

Before solving these equations we note that in the purely one-dimensional case

$$\tilde{v} = \tilde{w} = 0 , \quad \tilde{\alpha}_3 = \tilde{\alpha}_4 = 0 , \quad \tilde{\mathbf{K}}^{(3)} = \tilde{\mathbf{K}}^{(4)} = \mathbf{0} \quad (11.67)$$

and the problem reduces to solving (11.62), (11.63) and (11.66) for  $\tilde{\alpha}_1$ ,  $\tilde{\alpha}_2$  and  $\tilde{\alpha}_5$ , with terms involving  $\tilde{\alpha}_3$  and  $\tilde{\alpha}_4$  being absent.

For the  $x$ -split three dimensional problem the system (11.62)–(11.66) may be viewed in exactly the same manner as for the one-dimensional case. Use of equation (11.62) into (11.64) and (11.65) gives directly

$$\tilde{\alpha}_3 = \Delta u_3 - \tilde{v}\Delta u_1 ; \quad \tilde{\alpha}_4 = \Delta u_4 - \tilde{w}\Delta u_1 . \quad (11.68)$$

Then one solves (11.62), (11.63) and (11.66) for  $\tilde{\alpha}_1$ ,  $\tilde{\alpha}_2$ ,  $\tilde{\alpha}_5$ . Computationally, it is convenient to arrange the solution as follows

$$\left. \begin{aligned} \tilde{\alpha}_2 &= \frac{\gamma - 1}{\tilde{a}^2} \left[ \Delta u_1(\tilde{H} - \tilde{u}^2) + \tilde{u}\Delta u_2 - \overline{\Delta u}_5 \right] , \\ \tilde{\alpha}_1 &= \frac{1}{2\tilde{a}} [\Delta u_1(\tilde{u} + \tilde{a}) - \Delta u_2 - \tilde{a}\tilde{\alpha}_2] , \\ \tilde{\alpha}_5 &= \Delta u_1 - (\tilde{\alpha}_1 + \tilde{\alpha}_2) , \end{aligned} \right\} \quad (11.69)$$

where

$$\overline{\Delta u}_5 = \Delta u_5 - (\Delta u_3 - \tilde{v}\Delta u_1)\tilde{v} - (\Delta u_4 - \tilde{w}\Delta u_1)\tilde{w} . \quad (11.70)$$

## An Algorithm

To compute the Roe numerical flux  $\mathbf{F}_{i+\frac{1}{2}}$  according to any of the formulae (11.27)–(11.29) we do the following:

- (1) Compute the Roe average values for  $\tilde{u}$ ,  $\tilde{v}$ ,  $\tilde{w}$ ,  $\tilde{H}$  and  $\tilde{a}$  according to (11.60).
- (2) Compute the averaged eigenvalues  $\tilde{\lambda}_i$  according to (11.58).
- (3) Compute the averaged right eigenvectors  $\tilde{\mathbf{K}}^{(i)}$  according to (11.59).
- (4) Compute the wave strengths  $\tilde{\alpha}_i$  according to (11.68)–(11.70).
- (5) Use all of the above quantities to compute  $\mathbf{F}_{i+\frac{1}{2}}$ , according to any of the formulae (11.27)–(11.29).

For the pure one-dimensional case, virtually all the required information for the application of the above algorithm is contained in this Chapter. An entropy fix is given in Sect. 11.4. The remaining items such as choosing the time step size and boundary conditions are found in Chap. 6. For two and three dimensional applications the reader requires the additional information provided in Chap. 16.

## 11.3 The Roe–Pike Method

Recall that solving the Riemann problem (11.4) approximately using Roe’s method means finding *averaged* eigenvalues  $\tilde{\lambda}_i$ , right eigenvectors  $\tilde{\mathbf{K}}^{(i)}$  and wave strengths  $\tilde{\alpha}_i$ , so that the Roe numerical flux may be evaluated by any of the formulae (11.27)–(11.29). In the previous section this task was carried out by following the original Roe approach, where the averaged Jacobian matrix  $\tilde{\mathbf{A}}$  is first sought. In this section we present a different approach, due to Roe and Pike [416], whereby the construction of  $\tilde{\mathbf{A}}$  is avoided; instead, one seeks directly averages of a set of scalar quantities that can then be used to evaluate the eigenvalues, right eigenvectors and wave strengths needed in formulae (11.27)–(11.29).

### 11.3.1 The Approach

The approach assumes, of course, that the appropriate original system is hyperbolic and that analytical expressions for the eigenvalues  $\lambda_i$  and the set of linearly independent right eigenvectors  $\mathbf{K}^{(i)}$  are available. Analytical expressions  $\hat{\alpha}_i$  for the wave strengths require extra work via an *extra linearisation*. One then selects a suitable vector of scalar quantities, typically the vector  $\mathbf{W}$  of primitive variables in (11.6) or variations of it, for which an average  $\tilde{\mathbf{W}}$  is to be found. The values of  $\tilde{\lambda}_i$ ,  $\tilde{\mathbf{K}}^{(i)}$  and  $\tilde{\alpha}_i$  are then found by direct evaluation of the analytical expressions for  $\lambda_i$ ,  $\mathbf{K}^{(i)}$  and  $\hat{\alpha}_i$  at the state  $\tilde{\mathbf{W}}$ . There are two distinct steps in the Roe–Pike approach.

#### Linearisation about a Reference State

To find analytical expressions for the wave strengths  $\alpha_i$  Roe and Pike assume a linearised form of the governing equations based on the assumption that the data states  $\mathbf{U}_L$  and  $\mathbf{U}_R$  are close to a reference state  $\hat{\mathbf{U}}$ , to order  $O(\Delta^2)$ . Linearisation of the conservation laws in (11.4) about this state  $\hat{\mathbf{U}}$  gives

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x \equiv \mathbf{U}_t + \left( \frac{\partial \mathbf{F}}{\partial \mathbf{U}} \right) \mathbf{U}_x \approx \mathbf{U}_t + \hat{\mathbf{A}} \mathbf{U}_x ,$$

where

$$\mathbf{U}_t + \hat{\mathbf{A}} \mathbf{U}_x = \mathbf{0} \quad (11.71)$$

is an approximation to the original conservation laws. Here  $\hat{\mathbf{A}}$  is the Jacobian matrix, assumed available, computed at the reference state  $\hat{\mathbf{U}}$ . Eigenvalues and right eigenvectors follow. Analytical expressions for the wave strengths  $\hat{\alpha}_i$  in the solution of the linear Riemann problem

$$\begin{aligned} \mathbf{U}_t + \hat{\mathbf{A}} \mathbf{U}_x &= \mathbf{0} , \\ \mathbf{U}(x, 0) &= \left\{ \begin{array}{l} \mathbf{U}_L \text{ if } x < 0 , \\ \mathbf{U}_R \text{ if } x > 0 , \end{array} \right\} \end{aligned} \quad (11.72)$$

are found by decomposing the data jump  $\Delta\mathbf{U}$  onto the right eigenvectors, in the usual way; see Sect. 2.3 of Chap. 2 and Sect. 5.4 of Chap. 5. That is we solve

$$\Delta\mathbf{U} = \mathbf{U}_R - \mathbf{U}_L = \sum_{k=1}^m \hat{\alpha}_k \hat{\mathbf{K}}^{(k)}. \quad (11.73)$$

Before proceeding, we note that this linearisation *is not* the Roe linearisation resulting from the Roe matrix  $\tilde{\mathbf{A}}$ ; it is merely a step to find some sufficiently simple analytical expressions for the wave strengths, which can then be evaluated at the unknown *Roe–Pike average state*  $\tilde{\mathbf{W}}$ , yet to be found.

### The Algebraic Problem for the Average State

The sought Roe–Pike average vector  $\tilde{\mathbf{W}}$  is then found by first setting

$$\tilde{\alpha}_i = \hat{\alpha}_i(\tilde{\mathbf{W}}), \quad \tilde{\lambda}_i = \lambda_i(\tilde{\mathbf{W}}), \quad \tilde{\mathbf{K}}^{(i)} = \mathbf{K}^{(i)}(\tilde{\mathbf{W}}); \quad (11.74)$$

the analytical expressions for  $\lambda_i$ ,  $\mathbf{K}^{(i)}$  and  $\hat{\alpha}_i$  are evaluated at the unknown average state  $\tilde{\mathbf{W}}$ . Then  $\tilde{\mathbf{W}}$  is found by solving the algebraic problem posed by the following two sets of equations

$$\Delta\mathbf{U} = \mathbf{U}_R - \mathbf{U}_L = \sum_{k=1}^m \tilde{\alpha}_k \tilde{\mathbf{K}}^{(k)} \quad (11.75)$$

and

$$\Delta\mathbf{F} = \mathbf{F}_R - \mathbf{F}_L = \sum_{k=1}^m \tilde{\alpha}_k \tilde{\lambda}_k \tilde{\mathbf{K}}^{(k)}. \quad (11.76)$$

In the following section we illustrate the Roe–Pike approach in terms of a simple system of conservation laws.

#### 11.3.2 The Isothermal Equations

We solve the Riemann problem

$$\left. \begin{aligned} \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x &= \mathbf{0}, \\ \mathbf{U}(x, 0) &= \begin{cases} \mathbf{U}_L & \text{if } x < 0, \\ \mathbf{U}_R & \text{if } x > 0, \end{cases} \end{aligned} \right\} \quad (11.77)$$

for the isothermal equations using the Roe–Pike approach; the vectors  $\mathbf{U}$  and  $\mathbf{F}$  are given in (11.32). The exact Jacobian matrix, eigenvalues and right eigenvectors are

$$\left. \begin{aligned} \mathbf{A}(\mathbf{U}) &= \begin{bmatrix} 0 & 1 \\ a^2 - u^2 & 2u \end{bmatrix}, \\ \lambda_1 &= u - a, \quad \lambda_2 = u + a, \\ \mathbf{K}^{(1)} &= \begin{bmatrix} 1 \\ u - a \end{bmatrix}, \quad \mathbf{K}^{(2)} = \begin{bmatrix} 1 \\ u + a \end{bmatrix}. \end{aligned} \right\} \quad (11.78)$$

## Linearisation about a Reference State

Assume that the data states  $\mathbf{U}_L$  and  $\mathbf{U}_R$  are close to a state  $\hat{\mathbf{U}}$  to order  $O(\Delta^2)$ . Linearisation of the conservation laws in (11.77) about this state  $\hat{\mathbf{U}}$  gives linear Riemann problem

$$\begin{aligned} \mathbf{U}_t + \hat{\mathbf{A}}\mathbf{U}_x &= \mathbf{0}, \\ \mathbf{U}(x, 0) &= \begin{cases} \mathbf{U}_L & \text{if } x < 0, \\ \mathbf{U}_R & \text{if } x > 0. \end{cases} \end{aligned} \quad (11.79)$$

Here  $\hat{\mathbf{A}}$  is the Jacobian  $\mathbf{A}$  evaluated at the reference state  $\hat{\mathbf{U}}$ , which in terms of primitive variables is denoted by  $\hat{\mathbf{W}} = (\hat{\rho}, \hat{u})^T$ . The complete eigenstructure is

$$\left. \begin{aligned} \mathbf{A}(\mathbf{U}) &= \begin{bmatrix} 0 & 1 \\ a^2 - \hat{u}^2 & 2\hat{u} \end{bmatrix}, \\ \hat{\lambda}_1 &= \hat{u} - a, \quad \hat{\lambda}_2 = \hat{u} + a, \\ \hat{\mathbf{K}}^{(1)} &= \begin{bmatrix} 1 \\ \hat{u} - a \end{bmatrix}, \quad \hat{\mathbf{K}}^{(2)} = \begin{bmatrix} 1 \\ \hat{u} + a \end{bmatrix}. \end{aligned} \right\} \quad (11.80)$$

Recall that the sound speed  $a$  is constant. We look for solutions of (11.79). The system is linear with constant coefficients. One can therefore deploy appropriate techniques studied in Sect. 2.3 of Chap. 2 and Sect. 5.4 of Chap. 5. We decompose the data jump  $\Delta\mathbf{U}$  onto the right eigenvectors as follows

$$\Delta\mathbf{U} = \mathbf{U}_R - \mathbf{U}_L = \sum_{k=1}^2 \hat{\alpha}_k \hat{\mathbf{K}}^{(k)} = \hat{\alpha}_1 \hat{\mathbf{K}}^{(1)} + \hat{\alpha}_2 \hat{\mathbf{K}}^{(2)}, \quad (11.81)$$

where analytical expressions for the coefficients  $\hat{\alpha}_1, \hat{\alpha}_2$  are to be found. Writing (11.81) in full gives

$$\Delta\rho = \rho_R - \rho_L = \hat{\alpha}_1 + \hat{\alpha}_2, \quad (11.82)$$

$$\Delta(\rho u) = (\rho u)_R - (\rho u)_L = \hat{\alpha}_1(\hat{u} - a) + \hat{\alpha}_2(\hat{u} + a). \quad (11.83)$$

It can easily be shown that

$$\Delta(\rho u) = \hat{\rho}\Delta u + \hat{u}\Delta\rho + O(\Delta^2), \quad (11.84)$$

where the leading term in  $O(\Delta^2)$  is

$$(\rho_R - \hat{\rho})(u_R - \hat{u}) - (\rho_L - \hat{\rho})(u_L - \hat{u}).$$

By neglecting  $O(\Delta^2)$ , (11.83) becomes

$$\hat{\rho}\Delta u + \hat{u}\Delta\rho = \hat{\alpha}_1(\hat{u} - a) + \hat{\alpha}_2(\hat{u} + a). \quad (11.85)$$

Solving equations (11.82) and (11.85) gives the sought analytical expressions for  $\hat{\alpha}_1$  and  $\hat{\alpha}_2$ , namely

$$\hat{\alpha}_1 = \frac{1}{2} \left[ \Delta\rho - \hat{\rho} \frac{\Delta u}{a} \right], \quad \hat{\alpha}_2 = \frac{1}{2} \left[ \Delta\rho + \hat{\rho} \frac{\Delta u}{a} \right]. \quad (11.86)$$

Compare these with expressions (11.46). The reader may easily verify that, to within  $O(\Delta^2)$ , the following two sets of equations are identically satisfied

$$\Delta\mathbf{U} = \mathbf{U}_R - \mathbf{U}_L = \sum_{k=1}^2 \hat{\alpha}_k \hat{\mathbf{K}}^{(k)}, \quad \Delta\mathbf{F} = \mathbf{F}_R - \mathbf{F}_L = \sum_{k=1}^2 \hat{\alpha}_k \hat{\lambda}_k \hat{\mathbf{K}}^{(k)}. \quad (11.87)$$

Here we give details for the second set. In full, these equations read

$$\Delta(\rho u) = \hat{\alpha}_1 \hat{\lambda}_1 + \hat{\alpha}_2 \hat{\lambda}_2, \quad (11.88)$$

$$\Delta(\rho u^2 + \rho a^2) = \hat{\alpha}_1 \hat{\lambda}_1 (\hat{u} - a) + \hat{\alpha}_2 \hat{\lambda}_2 (\hat{u} + a). \quad (11.89)$$

Equation (11.88) may be written as

$$\hat{\rho} \Delta u + \hat{u} \Delta \rho = \hat{u} (\hat{\alpha}_1 + \hat{\alpha}_2) + a (\hat{\alpha}_2 - \hat{\alpha}_1),$$

which after using (11.86) becomes an identity. To prove (11.89) we first expand its left-hand side

$$\Delta(\rho u^2 + \rho a^2) = 2\hat{\rho}\hat{u}\Delta u + \hat{u}^2\Delta\rho + a^2\Delta\rho.$$

The right-hand side of (11.89) can be expressed as

$$(\hat{\alpha}_1 + \hat{\alpha}_2)(\hat{u}^2 + a^2) + 2\hat{u}a(\hat{\alpha}_2 - \hat{\alpha}_1).$$

Therefore, after use of (11.86), equation (11.89) becomes an identity and thus the second set of equations in (11.87), to order  $O(\Delta^2)$ , is identically satisfied.

### The Algebraic Problem for the Average State

For the general case in which the data states  $\mathbf{U}_L$  and  $\mathbf{U}_R$  are not necessarily close, the Roe–Pike approach proposes the algebraic problem of finding the Roe–Pike averages  $\tilde{\rho}$  and  $\tilde{u}$  such that the two conditions (11.75) and (11.76) are valid, namely

$$\Delta\mathbf{U} = \sum_{k=1}^2 \tilde{\alpha}_k \tilde{\mathbf{K}}^{(k)}, \quad \Delta\mathbf{F} = \sum_{k=1}^2 \tilde{\alpha}_k \tilde{\lambda}_k \tilde{\mathbf{K}}^{(k)}. \quad (11.90)$$

Here, according to (11.74),  $\tilde{\alpha}_k$ ,  $\tilde{\lambda}_k$  and  $\tilde{\mathbf{K}}^{(k)}$  are obtained by evaluating the available analytical expressions at the sought averages  $\tilde{\rho}, \tilde{u}$ . For the wave strengths these are given by (11.86). For the eigenvalues and right eigenvectors they are given by (11.78). We then set

$$\tilde{\alpha}_1 = \frac{1}{2} \left[ \Delta\rho - \tilde{\rho} \frac{\Delta u}{a} \right], \quad \tilde{\alpha}_2 = \frac{1}{2} \left[ \Delta\rho + \tilde{\rho} \frac{\Delta u}{a} \right], \quad (11.91)$$

$$\tilde{\lambda}_1 = \tilde{u} - a, \quad \tilde{\lambda}_2 = \tilde{u} + a, \quad (11.92)$$

$$\tilde{\mathbf{K}}^{(1)} = \begin{bmatrix} 1 \\ \tilde{u} - a \end{bmatrix}, \quad \tilde{\mathbf{K}}^{(2)} = \begin{bmatrix} 1 \\ \tilde{u} + a \end{bmatrix}. \quad (11.93)$$

Writing conditions (11.90) in full produces

$$\Delta\rho = \tilde{\alpha}_1 + \tilde{\alpha}_2, \quad (11.94)$$

$$\Delta(\rho u) = \tilde{\alpha}_1(\tilde{u} - a) + \tilde{\alpha}_2(\tilde{u} + a), \quad (11.95)$$

$$\Delta(\rho u) = \tilde{\lambda}_1\tilde{\alpha}_1 + \tilde{\lambda}_2\tilde{\alpha}_2, \quad (11.96)$$

$$\Delta(\rho u^2 + a^2\rho) = \tilde{\lambda}_1\tilde{\alpha}_1(\tilde{u} - a) + \tilde{\lambda}_2\tilde{\alpha}_2(\tilde{u} + a). \quad (11.97)$$

These are a set of four non-linear algebraic equations for the two unknowns  $\tilde{\rho}$  and  $\tilde{u}$ . Note however that, by virtue of (11.91), (11.94) is an identity, for any average value  $\tilde{\rho}$ . Also, (11.95) is identical to (11.96) and thus we work with (11.96) and (11.97) only. From equation (11.96) one obtains

$$\Delta(\rho u) = \tilde{u}(\tilde{\alpha}_1 + \tilde{\alpha}_2) + a(\tilde{\alpha}_2 - \tilde{\alpha}_1).$$

Use of (11.91) here leads to

$$\Delta(\rho u) = \tilde{\rho}\Delta u + \tilde{u}\Delta\rho. \quad (11.98)$$

From (11.97) we write

$$\Delta(\rho u^2 + \rho a^2) = (\tilde{\alpha}_1 + \tilde{\alpha}_2)(\tilde{u}^2 + a^2) + 2a\tilde{u}(\tilde{\alpha}_2 - \tilde{\alpha}_1),$$

which after using (11.91) and the exact relation

$$\Delta(\rho u^2 + \rho a^2) = \Delta(\rho u^2) + a^2\Delta\rho$$

leads to the result

$$\Delta(\rho u^2) = 2\tilde{u}\tilde{\rho}\Delta u + \tilde{u}^2\Delta\rho. \quad (11.99)$$

Elimination of  $\tilde{\rho}$  from (11.98) and (11.99) leads to a quadratic equation for  $\tilde{u}$ , namely

$$\Delta\rho\tilde{u}^2 - 2\Delta(\rho u)\tilde{u} + \Delta(\rho u^2) = 0. \quad (11.100)$$

This equation has two solutions, namely

$$\tilde{u} = \frac{\Delta(\rho u) \pm \sqrt{[\Delta(\rho u)]^2 - \Delta\rho\Delta(\rho u^2)}}{\Delta\rho}. \quad (11.101)$$

After using the definition  $\Delta r = r_R - r_L$  the discriminant is found to be

$$\rho_L\rho_R(\Delta u)^2,$$

which simplifies (11.101) to

$$\tilde{u} = \frac{\Delta(\rho u) \pm \Delta u \sqrt{\rho_L \rho_R}}{\Delta \rho}. \quad (11.102)$$

The root obtained by taking the *negative* sign in (11.102) produces the Roe–averaged velocity

$$\tilde{u} = \frac{\sqrt{\rho_L} u_L + \sqrt{\rho_R} u_R}{\sqrt{\rho_L} + \sqrt{\rho_R}}. \quad (11.103)$$

Compare (11.103) with (11.43). From (11.98) we obtain

$$\tilde{\rho} = \sqrt{\rho_L \rho_R}. \quad (11.104)$$

We have thus found algebraic expressions for the sought Roe–Pike averages  $\tilde{\rho}$  and  $\tilde{u}$ . We observe that the second root obtained by taking the *positive* sign in (11.102) leads to the *spurious* solution

$$\tilde{u} = \frac{\sqrt{\rho_R} u_R - \sqrt{\rho_L} u_L}{\sqrt{\rho_R} - \sqrt{\rho_L}}. \quad (11.105)$$

There is a very good reason for rejecting this as a *useful solution*; in the trivial case  $\rho_L = \rho_R$ ,  $u_L \neq u_R$  the solution  $\tilde{u}$  is not even defined.

Having found the Roe–Pike averages  $\tilde{\rho}$  and  $\tilde{u}$  we can then compute the wave strengths  $\tilde{\alpha}_k$ , the eigenvalues  $\tilde{\lambda}_k$  and the right eigenvectors  $\tilde{\mathbf{K}}^{(k)}$  according to expressions (11.91)–(11.93). The Roe numerical flux  $\mathbf{F}_{i+\frac{1}{2}}$  to be used in the conservative formula (11.2) can now be obtained from any of the relations (11.27)–(11.29).

### 11.3.3 The Euler Equations

We solve the Riemann problem (11.4) for the  $x$ –split, three dimensional Euler equations using the Roe–Pike method. Assuming the analytical expressions (11.48)–(11.49) for the eigenvalues and eigenvectors, one then linearises the equations about a state  $\hat{\mathbf{U}}$  to find analytical expressions for the wave strengths; this is done under the assumption that both data states  $\mathbf{U}_L, \mathbf{U}_R$  are close to  $\hat{\mathbf{U}}$  to  $O(\Delta^2)$ . This leads to the linear system

$$\left. \begin{aligned} \mathbf{U}_t + \hat{\mathbf{A}} \mathbf{U}_x &= \mathbf{0}, \\ \mathbf{U}(x, t) &= \begin{cases} \mathbf{U}_L, & x < 0, \\ \mathbf{U}_R, & x > 0. \end{cases} \end{aligned} \right\} \quad (11.106)$$

The Jacobian matrix  $\hat{\mathbf{A}}$  is obtained by evaluating the exact Jacobian matrix (11.47) at the state  $\hat{\mathbf{U}}$ ; the eigenvalues  $\hat{\lambda}_i$  are

$$\hat{\lambda}_1 = \hat{u} - \hat{a}, \quad \hat{\lambda}_2 = \hat{\lambda}_3 = \hat{\lambda}_4 = \hat{u}, \quad \hat{\lambda}_5 = \hat{u} + \hat{a} \quad (11.107)$$

and the right eigenvectors  $\hat{\mathbf{K}}^{(i)}$  are

$$\left. \begin{aligned} \hat{\mathbf{K}}^{(1)} &= \begin{bmatrix} 1 \\ \hat{u} - \hat{a} \\ \hat{v} \\ \hat{w} \\ \hat{H} - \hat{u}\hat{a} \end{bmatrix}; \quad \hat{\mathbf{K}}^{(2)} = \begin{bmatrix} 1 \\ \hat{u} \\ \hat{v} \\ \hat{w} \\ \frac{1}{2}\hat{V}^2 \end{bmatrix}; \quad \hat{\mathbf{K}}^{(3)} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \hat{v} \end{bmatrix} \\ \hat{\mathbf{K}}^{(4)} &= \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ \hat{w} \end{bmatrix}; \quad \hat{\mathbf{K}}^{(5)} = \begin{bmatrix} 1 \\ \hat{u} + \hat{a} \\ \hat{v} \\ \hat{w} \\ \hat{H} + \hat{u}\hat{a} \end{bmatrix}. \end{aligned} \right\} \quad (11.108)$$

By expanding the data jump  $\Delta\mathbf{U}$  onto the right eigenvectors we write

$$\Delta\mathbf{U} = \sum_{i=1}^5 \hat{\alpha}_i \hat{\mathbf{K}}^{(i)} \quad (11.109)$$

The solution of this  $5 \times 5$  linear system will provide analytical expressions for the wave strengths  $\hat{\alpha}_i$ . As a matter of fact we can use the solution for the wave strengths obtained in the Roe original method, (11.68)–(11.70), and reinterpret the solution appropriately. These are

$$\left. \begin{aligned} \hat{\alpha}_3 &= \Delta u_3 - \hat{v} \Delta u_1, \\ \hat{\alpha}_4 &= \Delta u_4 - \hat{w} \Delta u_1, \\ \hat{\alpha}_2 &= \frac{\gamma-1}{\hat{a}^2} [\Delta u_1 (\hat{H} - \hat{u}^2) + \hat{u} \Delta u_2 - \overline{\Delta u_5}], \\ \hat{\alpha}_1 &= \frac{1}{2\hat{a}} [\Delta u_1 (\hat{u} + \hat{a}) - \Delta u_2 - \hat{a} \hat{\alpha}_2], \\ \hat{\alpha}_5 &= \Delta u_1 - (\hat{\alpha}_1 + \hat{\alpha}_2), \end{aligned} \right\} \quad (11.110)$$

where

$$\overline{\Delta u_5} = \Delta u_5 - (\Delta u_3 - \hat{v} \Delta u_1) \hat{v} - (\Delta u_4 - \hat{w} \Delta u_1) \hat{w}. \quad (11.111)$$

By applying the operator

$$\Delta(rs) = \hat{r} \Delta s + \hat{s} \Delta r + O(\Delta^2) \quad (11.112)$$

and neglecting  $O(\Delta^2)$  we arrive at the following solution:

$$\left. \begin{aligned} \hat{\alpha}_1 &= \frac{1}{2\hat{a}^2} [\Delta p - \hat{\rho}\hat{a}\Delta u], \\ \hat{\alpha}_2 &= \Delta\rho - \Delta p/\hat{a}^2, \\ \hat{\alpha}_3 &= \hat{\rho}\Delta v, \\ \hat{\alpha}_4 &= \hat{\rho}\Delta w, \\ \hat{\alpha}_5 &= \frac{1}{2\hat{a}^2} [\Delta p + \hat{\rho}\hat{a}\Delta u] \end{aligned} \right\} \quad (11.113)$$

The second step in the Roe–Pike method is to find an average state

$$\tilde{\mathbf{W}} = (\tilde{\rho}, \tilde{u}, \tilde{v}, \tilde{w}, \tilde{a})^T, \quad (11.114)$$

such that the algebraic problem posed by the following two sets of equations

$$\Delta \mathbf{U} = \sum_{i=1}^5 \tilde{\alpha}_i \tilde{\mathbf{K}}^{(i)}, \quad (11.115)$$

$$\Delta \mathbf{F} = \sum_{i=1}^5 \tilde{\alpha}_i \tilde{\lambda}_i \tilde{\mathbf{K}}^{(i)}, \quad (11.116)$$

is satisfied, where

$$\tilde{\alpha}_i = \hat{\alpha}_i(\tilde{\mathbf{W}}), \quad \tilde{\lambda}_i = \lambda_i(\tilde{\mathbf{W}}), \quad \tilde{\mathbf{K}}^{(i)} = \mathbf{K}^{(i)}(\tilde{\mathbf{W}}), \quad (11.117)$$

with  $\lambda_i$  and  $\mathbf{K}^{(i)}$  given by (11.48)–(11.49) and  $\hat{\alpha}_i$  given by (11.113). Details of the algebra for the one-dimensional case are given by Roe and Pike [416]. For the  $x$ -split three dimensional case the solution for the average vector  $\tilde{\mathbf{W}}$  is

$$\left. \begin{aligned} \tilde{\rho} &= \sqrt{\rho_L \rho_R}, \\ \tilde{u} &= \frac{\sqrt{\rho_L} u_L + \sqrt{\rho_R} u_R}{\sqrt{\rho_L} + \sqrt{\rho_R}}, \\ \tilde{v} &= \frac{\sqrt{\rho_L} v_L + \sqrt{\rho_R} v_R}{\sqrt{\rho_L} + \sqrt{\rho_R}}, \\ \tilde{w} &= \frac{\sqrt{\rho_L} w_L + \sqrt{\rho_R} w_R}{\sqrt{\rho_L} + \sqrt{\rho_R}}, \\ \tilde{H} &= \frac{\sqrt{\rho_L} H_L + \sqrt{\rho_R} H_R}{\sqrt{\rho_L} + \sqrt{\rho_R}}, \\ \tilde{a} &= \left( (\gamma - 1)(\tilde{H} - \frac{1}{2}\tilde{\mathbf{V}}^2) \right)^{\frac{1}{2}}, \end{aligned} \right\} \quad (11.118)$$

where  $\tilde{\mathbf{V}}^2 = \tilde{u}^2 + \tilde{v}^2 + \tilde{w}^2$ . These are identical to the Roe averages obtained by the original Roe method, see (11.60). Now  $\tilde{\alpha}_i$ ,  $\tilde{\lambda}_i$  and  $\tilde{\mathbf{K}}^{(i)}$  are computed according to (11.117) and then the Roe intercell flux  $\mathbf{F}_{i+\frac{1}{2}}$  follows from any of the formulae (11.27)–(11.29).

## An Algorithm

To compute the Roe numerical flux  $\mathbf{F}_{i+\frac{1}{2}}$  according to any of the formulae (11.27)–(11.29) we do the following:

- (1) Compute the Roe average values according to (11.118).
- (2) Compute the eigenvalues  $\tilde{\lambda}_i$  using the analytical expressions (11.107) evaluated on the averages (11.118).
- (3) Compute the right eigenvectors using the analytical expressions (11.108) evaluated on the averages (11.118).
- (4) Compute the wave strengths using the analytical expressions (11.113) evaluated on the averages (11.118).
- (5) Use all of the above quantities to compute  $\mathbf{F}_{i+\frac{1}{2}}$ , according to any of the formulae (11.27)–(11.29).

Before applying the scheme as described to practical problems, a modification to handle *sonic flow* correctly is required. This is the subject of the next section.

## 11.4 An Entropy Fix

Linearised Riemann problem solutions consist of discontinuous jumps only. See Sect. 2.3 of Chap. 2. This can be a good approximation for contacts and shocks, in that the discontinuous character of the wave is correct, although the size of the jump may not be correctly approximated by the linearised solution. Rarefaction waves, on the other hand, carry a continuous change in flow variables, and as time increases, they tend to spread; that is spatial gradients tend to decay. Quite clearly then, the linearised approximation via discontinuous jumps is grossly incorrect. In a practical computational set up however, it is only in the case in which the rarefaction wave is *transonic*, or *sonic*, where linearised approximations encounter difficulties; these show up in the form of unphysical, entropy violating discontinuous waves, sometimes called *rarefaction shocks*.

### 11.4.1 The Entropy Problem

Consider the Riemann problem whose initial data is that of Test 1 in Table 11.1. The structure of the exact solution of this problem, depicted in Fig. 11.2, contains a left *sonic rarefaction*, a contact discontinuity of speed  $u_*$  and a right shock wave. As the left rarefaction is *sonic* the eigenvalue  $\lambda_1 = u - a$  changes from negative to positive, as the wave is crossed from left to right. There is a point at which  $\lambda_1 = u - a = 0$ , giving the sonic flow condition  $u = a$ .

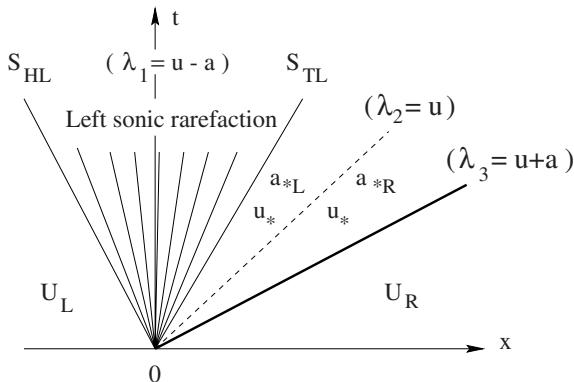
$$\lambda_1(\mathbf{U}_L) = S_{HL} = u_L - a_L < 0$$

is the speed of the *head* of the rarefaction and

$$\lambda_1(\mathbf{U}_{*L}) = S_{TL} = u_* - a_{*L} > 0$$

is the speed of the *tail*. Fig. 11.4 shows the numerical (symbols) and exact (line) solutions of this problem, where the numerical solution is obtained by

Roe's method as described so far. The numerical solution within the rarefaction exhibits a discontinuity within the wave; this discontinuity is unphysical, it violates the entropy condition. See Sect. 2.4.2 of Chap. 2. Recall that a physically admissible discontinuity of speed  $S$  requires  $S_b \geq S \geq S_a$  where  $S_b$  and  $S_a$  are characteristic speeds behind and ahead of the wave respectively. That is, characteristics move into the discontinuity; the limiting case of parallel characteristic speeds is that of a contact discontinuity. For the example above, the opposite happens. See Sect. 2.4.2 of Chap. 2, for a discussion on entropy-violating solutions.



**Fig. 11.2.** Left transonic rarefaction wave. Left eigenvalue  $\lambda_1 = u - a$  changes sign as the wave is crossed from left to right

Roe's solver can be modified so as to avoid entropy violating solutions. This is usually referred to as an *entropy fix*. Harten and Hyman [243] suggested an entropy fix for Roe's method, which has widespread use. Other ways of correcting the scheme have been discussed by Roe and Pike [416], Roe [414], Sweby [469] and Dubois and Mehlman [167], amongst others. Here we present the details of the Harten–Hyman approach.

### 11.4.2 The Harten–Hyman Entropy Fix

The general approach is presented in the original paper of Harten and Hyman of 1983 [243]. A description can also be found in [308]. The presentation here is tailored specifically to the time-dependent Euler equations, for which we only need to consider the left and right non-linear waves associated with the eigenvalues  $\lambda_1 = u - a$  and  $\lambda_5 = u + a$  respectively. Our version of the Harten–Hyman entropy fix relies on estimates for particle velocity  $u_*$  and sound speeds  $a_{*L}$ ,  $a_{*R}$  in the *Star Region*; see Figs. 11.1 and 11.2. Various ways of finding these are given in Sect. 11.4.3.

### Left Transonic Rarefaction

Consider the situation depicted in Fig. 11.2. Assuming  $u_*$  and  $a_{*L}$  are available, we compute the speeds

$$\lambda_1^L = u_L - a_L ; \quad \lambda_1^R = u_* - a_{*L} . \quad (11.119)$$

If

$$\lambda_1^L < 0 < \lambda_1^R , \quad (11.120)$$

then the left wave is a *transonic, or sonic, rarefaction wave*. In these circumstances the entropy fix is required and is enforced as follows. The single jump

$$\mathbf{U}_{*L} - \mathbf{U}_L = \tilde{\alpha}_1 \tilde{\mathbf{K}}^{(1)} \quad (11.121)$$

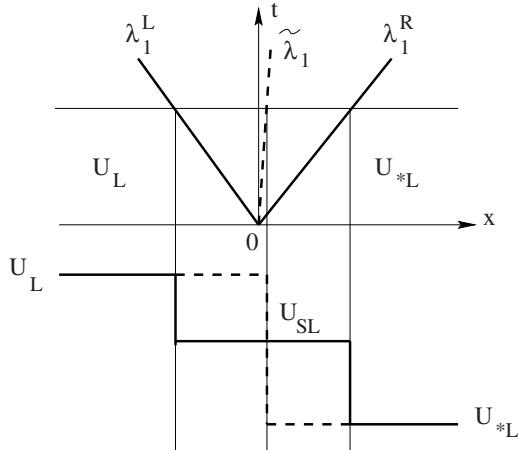
travelling with speed  $\tilde{\lambda}_1$  is split into two smaller jumps  $\mathbf{U}_{SL} - \mathbf{U}_L$  and  $\mathbf{U}_{*L} - \mathbf{U}_{SL}$  travelling respectively at speeds  $\lambda_1^L$  and  $\lambda_1^R$ , where  $\mathbf{U}_{SL}$  is a *transonic state* yet to be found; see Fig. 11.3. Application of the integral form of the conservation laws, see Chaps. 3 and 10, gives

$$\lambda_1^R (\mathbf{U}_{SL} - \mathbf{U}_{*L}) + \lambda_1^L (\mathbf{U}_L - \mathbf{U}_{SL}) = \tilde{\lambda}_1 (\mathbf{U}_L - \mathbf{U}_{*L}) , \quad (11.122)$$

from which we obtain

$$\mathbf{U}_{SL} = \frac{(\tilde{\lambda}_1 - \lambda_1^L) \mathbf{U}_L + (\lambda_1^R - \tilde{\lambda}_1) \mathbf{U}_{*L}}{\lambda_1^R - \lambda_1^L} . \quad (11.123)$$

To compute the Roe intercell flux we adopt the one-sided formulae (11.27), namely



**Fig. 11.3.** Entropy fix for left transonic rarefaction wave. Single jump  $\mathbf{U}_{*L} - \mathbf{U}_L$  travelling with speed  $\tilde{\lambda}_1$  is split into the two jumps  $\mathbf{U}_{SL} - \mathbf{U}_L$  and  $\mathbf{U}_{*L} - \mathbf{U}_{SL}$  travelling with speeds  $\lambda_1^L$  and  $\lambda_1^R$ . Profile shown is a representation a single variable

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}_L + \sum_{\tilde{\lambda}_k \leq 0} \tilde{\lambda}_k \tilde{\alpha}_k \tilde{\mathbf{K}}^{(k)}, \quad (11.124)$$

where in the present case the summation applies to the single jump  $\mathbf{U}_{SL} - \mathbf{U}_L$  travelling with speed  $\lambda_1^L < 0$ ; in view of (11.122) the jump is

$$\mathbf{U}_{SL} - \mathbf{U}_L = \frac{(\lambda_1^R - \tilde{\lambda}_1)}{(\lambda_1^R - \lambda_1^L)} (\mathbf{U}_{*L} - \mathbf{U}_L). \quad (11.125)$$

But the Roe approximation gives

$$\mathbf{U}_{*L} - \mathbf{U}_L = \tilde{\alpha}_1 \tilde{\mathbf{K}}^{(1)} \quad (11.126)$$

and thus the flux jump  $(\Delta \mathbf{F})_1^L$  across the wave of speed  $\lambda_1^L$  is

$$(\Delta \mathbf{F})_1^L = \lambda_1^L \left( \frac{\lambda_1^R - \tilde{\lambda}_1}{\lambda_1^R - \lambda_1^L} \right) \tilde{\alpha}_1 \tilde{\mathbf{K}}^{(1)}. \quad (11.127)$$

By defining the new wave speed

$$\bar{\lambda}_1 = \lambda_1^L \left( \frac{\lambda_1^R - \tilde{\lambda}_1}{\lambda_1^R - \lambda_1^L} \right), \quad (11.128)$$

the intercell flux (11.27) becomes

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}_L + \bar{\lambda}_1 \tilde{\alpha}_1 \tilde{\mathbf{K}}^{(1)}. \quad (11.129)$$

### Right Transonic Rarefaction

For a right transonic rarefaction, the entropy fix procedure is entirely analogous to the left rarefaction case. Assuming the speeds  $u_*$  and  $a_{*R}$  are available, we compute the two wave speeds

$$\lambda_5^L = u_* + a_{*R}, \quad \lambda_5^R = u_R + a_R. \quad (11.130)$$

If

$$\lambda_5^L < 0 < \lambda_5^R \quad (11.131)$$

then the right wave is a *transonic rarefaction wave*. The transonic state  $\mathbf{U}_{SR}$  is defined between the waves of speeds  $\lambda_5^L$  and  $\lambda_5^R$  and is given by

$$\mathbf{U}_{SR} = \frac{(\lambda_5^R - \tilde{\lambda}_5) \mathbf{U}_R + (\tilde{\lambda}_5 - \lambda_5^L) \mathbf{U}_{*R}}{\lambda_5^R - \lambda_5^L}. \quad (11.132)$$

Next we define the new wave speed

$$\bar{\lambda}_5 = \lambda_5^R \left( \frac{\tilde{\lambda}_5 - \lambda_5^L}{\lambda_5^R - \lambda_5^L} \right) \quad (11.133)$$

and then use the one-sided flux formula (11.28) to compute the numerical flux. The resulting Roe numerical flux is

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}_R - \bar{\lambda}_5 \tilde{\alpha}_5 \tilde{\mathbf{K}}^{(5)}. \quad (11.134)$$

In the present version of the Harten–Hyman entropy fix we have used the one-sided flux formulae (11.27) and (11.28). The procedure can be easily adapted for use in conjunction with the centred formulae (11.29), if desired.

Next we discuss ways of finding the speeds  $u_*$ ,  $a_{*L}$  and  $a_{*R}$  needed to implement the entropy fix.

### 11.4.3 The Speeds $u_*$ , $a_{*L}$ , $a_{*R}$

The *star* states  $\mathbf{U}_{*L}$ ,  $\mathbf{U}_{*R}$  are required in order to obtain the speeds  $u_*$ ,  $a_{*L}$ ,  $a_{*R}$  and thus the characteristic speeds in (11.119) and (11.130). We present various possible alternatives.

#### The Roe–Averaged States

Given the Roe–averaged  $\tilde{\alpha}_i$  and  $\tilde{\mathbf{K}}^{(i)}$  one can find the state  $\mathbf{U}_{*L}$  as

$$\mathbf{U}_{*L} = \mathbf{U}_L + \tilde{\alpha}_1 \tilde{\mathbf{K}}^{(1)}, \quad (11.135)$$

which leads to

$$\left. \begin{aligned} \rho_{*L} &= \rho_L + \tilde{\alpha}_1, & u_* &= \frac{\rho_L u_L + \tilde{\alpha}_1 (\tilde{u} - \tilde{a})}{\rho_L + \tilde{\alpha}_1}, \\ p_* &= (\gamma - 1) \left[ E_L + \tilde{\alpha}_1 (\tilde{H} - \tilde{u}\tilde{a}) - \frac{1}{2} \rho_{*L} u_*^2 \right] \end{aligned} \right\} \quad (11.136)$$

Then we compute the sound speed

$$a_{*L} = \sqrt{\frac{\gamma p_*}{\rho_{*L}}} \quad (11.137)$$

and thus the speeds  $\lambda_1^L$  and  $\lambda_1^R$  in Eq. (11.119) follow. For the right wave one has

$$\mathbf{U}_{*R} = \mathbf{U}_R - \tilde{\alpha}_5 \tilde{\mathbf{K}}^{(5)}, \quad (11.138)$$

which produces

$$\left. \begin{aligned} \rho_{*R} &= \rho_R - \tilde{\alpha}_5, & u_* &= \frac{\rho_R u_R - \tilde{\alpha}_5 (\tilde{u} + \tilde{a})}{\rho_R - \tilde{\alpha}_5}, \\ p_* &= (\gamma - 1) \left[ E_R - \tilde{\alpha}_5 (\tilde{H} + \tilde{u}\tilde{a}) - \frac{1}{2} \rho_{*R} u_*^2 \right] \end{aligned} \right\} \quad (11.139)$$

The sound speed follows as  $a_{*R} = \sqrt{\frac{\gamma p_*}{\rho_{*R}}}$  and thus the wave speeds  $\lambda_5^L$  and  $\lambda_5^R$  in Eq. (11.130) are determined.

## The PVRS Approximation

Another way of estimating the required wave speeds is by using the Primitive–Variable Riemann Solver (PVRS) of Toro [502] presented in Sect. 9.3 of Chap. 9. The relevant solution values are

$$\left. \begin{aligned} p_* &= \frac{1}{2}(p_L + p_R) + \frac{1}{2}(u_L - u_R)\bar{\rho}\bar{a}, \\ u_* &= \frac{1}{2}(u_L + u_R) + \frac{1}{2}(p_L - p_R)/(\bar{\rho}\bar{a}), \\ \rho_{*L} &= \rho_L + (u_L - u_*)\bar{\rho}/\bar{a}, \\ \rho_{*R} &= \rho_R + (u_* - u_R)\bar{\rho}/\bar{a}, \end{aligned} \right\} \quad (11.140)$$

with

$$\bar{\rho} = \frac{1}{2}(\rho_L + \rho_R), \quad \bar{a} = \frac{1}{2}(a_L + a_R). \quad (11.141)$$

In order to avoid negative pressures we recommend replacing the linearised solution  $p_*$  by  $\max\{0, p_*\}$ . The sound speeds  $a_{*L}, a_{*R}$  are then computed in the usual way.

## TRRS Approximation

Another possibility is to use the Two–Rarefaction Riemann Solver (TRRS) discussed in Chap. 9, Sect. 9.4.1. The pressure  $p_*$  is given by

$$p_* = \left[ \frac{a_L + a_R - \frac{\gamma-1}{2}(u_R - u_L)}{a_L/p_L^z + a_R/p_R^z} \right]^{\frac{1}{z}}, \quad (11.142)$$

with  $z = \frac{\gamma-1}{2\gamma}$ . For the left non–linear wave the sound speed and particle velocity follow directly as

$$a_{*L} = a_L(p_*/p_L)^z, \quad u_* = u_L + \frac{2}{(\gamma-1)}(a_L - a_{*L}). \quad (11.143)$$

For the right non–linear wave we have

$$a_{*R} = a_R(p_*/p_R)^z, \quad u_* = u_R + \frac{2}{(\gamma-1)}(a_{*R} - a_R). \quad (11.144)$$

Hence speeds (11.119) and (11.130) are determined.

## Other Alternatives

Both the PVRS and the Roe linearised solutions for the speeds  $u_*, a_{*L}, a_{*R}$  may fail in the vicinity of low density flow [182]. The TRRS approximation presented above would not suffer from such difficulties; in fact, in the case in which both non–linear waves are rarefactions such an approximation would be exact. But as seen in equations (11.142)–(11.144) there are four fractional powers to be computed in each case, which makes this approximation rather expensive to use. A robust and yet more efficient scheme is the Two–Shock Riemann Solver (TSRS) [509] of Sect. 9.4.2, Chap. 9. Even better is the adaptive Riemann solver scheme of Sect. 9.5.2, Chap. 9.

## 11.5 Numerical Results and Discussion

Here we illustrate the performance of Godunov's first-order upwind method used in conjunction with the Roe approximate Riemann solver, discuss the results and point directions for extending the method.

### 11.5.1 The Tests

We select five test problems for the one-dimensional, time dependent Euler equations for ideal gases with  $\gamma = 1.4$ ; these have exact solutions. In all chosen tests, data consists of two constant states  $\mathbf{W}_L = [\rho_L, u_L, p_L]^T$  and  $\mathbf{W}_R = [\rho_R, u_R, p_R]^T$ , separated by a discontinuity at a position  $x = x_0$ . The states  $\mathbf{W}_L$  and  $\mathbf{W}_R$  are given in Table 11.1. The exact and numerical solutions are found in the spatial domain  $0 \leq x \leq 1$ . The numerical solution is computed with  $M = 100$  cells and the CFL condition is as for all previous computations, see Chap. 6; the chosen Courant number coefficient is  $C_{\text{cfl}} = 0.9$ ; boundary conditions are transmissive.

The exact solutions were found by running the code HE-E1RPEXACT of the library *NUMERICA* [519] and the numerical solutions were obtained by running the code HE-E1GODFLUX of *NUMERICA*.

Test	$\rho_L$	$u_L$	$p_L$	$\rho_R$	$u_R$	$p_R$
1	1.0	0.75	1.0	0.125	0.0	0.1
2	1.0	-2.0	0.4	1.0	2.0	0.4
3	1.0	0.0	1000.0	1.0	0.0	0.01
4	5.99924	19.5975	460.894	5.99242	-6.19633	46.0950
5	1.0	-19.59745	1000.0	1.0	-19.59745	0.01

Table 11.1. Data for five test problems with exact solution, for the time-dependent, one dimensional Euler equations

Test 1 is a *modified* version of Sod's problem [453]; the solution has a right shock wave, a right travelling contact wave and a left *sonic* rarefaction wave; this test is useful in assessing the *entropy satisfaction* property of numerical methods. The solution of Test 2 consists of two symmetric rarefaction waves and a trivial contact wave; the *Star Region* between the non-linear waves is close to vacuum, which makes this problem a suitable test for assessing the performance of numerical methods for low-density flows. Test 3 is designed to assess the robustness and accuracy of numerical methods; its solution consists of a strong shock wave of shock Mach number 198, a contact surface and a left rarefaction wave. Test 4 is also a very severe test, its solution consists of three strong discontinuities travelling to the right. A detailed discussion on the exact solution of the test problems is found in Sect. 4.3.3 of Chap. 4. Test 5 is also designed to test the robustness of numerical methods but the main reason for devising this test is to assess the ability of numerical methods to resolve

*slowly-moving contact discontinuities.* The exact solution of Test 5 consists of a left rarefaction wave, a right-travelling shock wave and a *stationary* contact discontinuity. For each test problem we select a convenient position  $x_0$  of the initial discontinuity and the output time. These are stated in the legend of each figure displaying computational results.

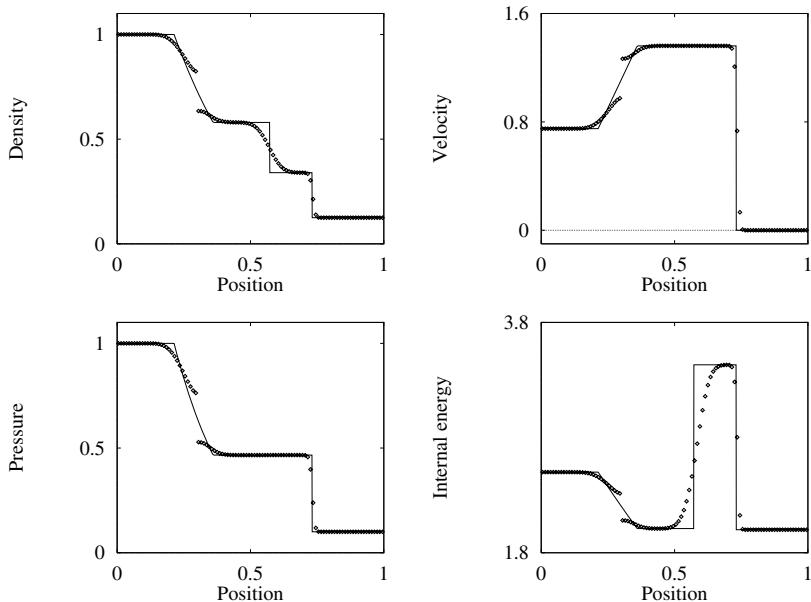
### 11.5.2 The Results

The computed results for Tests 1 to 5 using the Godunov first-order method in conjunction with the Roe approximate Riemann solver are shown in Figs. 11.4–11.8, where the numerical solution is shown by the symbols and the full line denotes the exact solution. As discussed earlier, Fig. 11.4 shows the results obtained from the Roe Riemann solver without the entropy fix and, as expected, the computed solution is obviously incorrect. Fig. 11.5 shows the corresponding results from the modified scheme using the Harten–Hyman entropy fix presented in the previous section. These results are, to plotting accuracy, almost indistinguishable from those obtained by the Godunov method in conjunction with the exact Riemann solver; see Fig. 6.8, Chap. 6. As a matter of fact, near the sonic point, the modified Roe solution looks slightly better; it also looks better than the Flux Vector Splitting solution, with the van Leer splitting, see Fig. 8.4 of Chap. 8. The HLL and HLLC solutions of Chap. 10, still seem to be the most accurate near sonic points. Compare also with the Osher results of Chap. 12. As anticipated, the Roe solver will fail near low-density flows; Test 2 contains two strong rarefactions with a low density and low pressure region in the middle and the Roe method, as described, does actually fail on this test. To compute successfully this kind of flows one must modify the Roe solver following the methodology of Einfeldt et. al. [182]. The results for Tests 3 and 4 are virtually identical to those of Godunov’s method with the exact Riemann solver, as the reader can verify by comparing Figs. 11.6 and 11.7 with Figs. 6.10 and 6.11 of Chap. 6. The results for Test 5 are also very similar to those obtained from the Godunov method with the exact Riemann solver; note however that the (non-isolated) stationary contact is not as sharply resolved as with the approximate HLLC Riemann solver of Chapt. 10, see Fig. 10.9. As expected of course, the resolution of the stationary contact is better than that of the Flux Vector Splitting Method with the Steger–Warming splitting and that with the van Leer splitting, see Figs. 8.14 and 8.15 of Chap. 8.

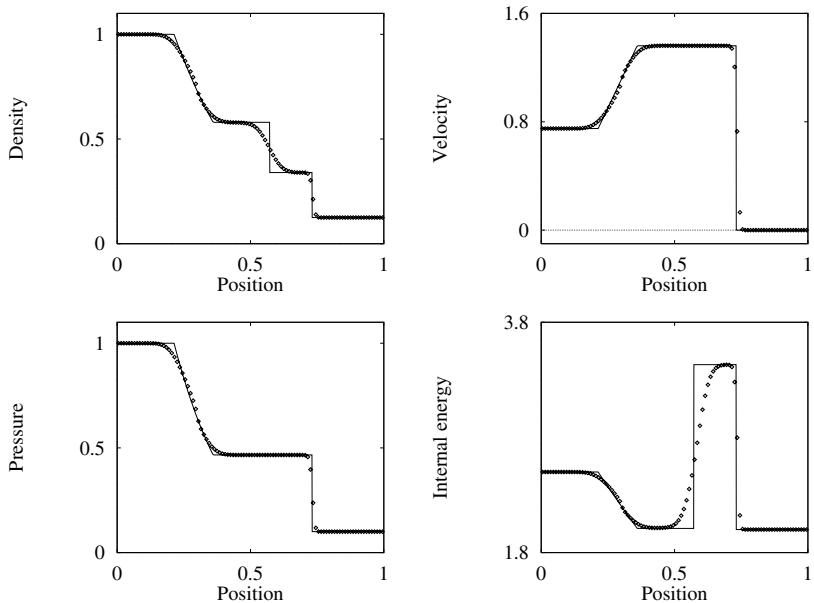
## 11.6 Extensions

The Roe approximate Riemann solver, following the original method of Roe and that of Roe and Pike, has been presented and illustrated via the isothermal equations of gas dynamics and the split three-dimensional, time dependent Euler equations. Details of the Roe solver for the three-dimensional

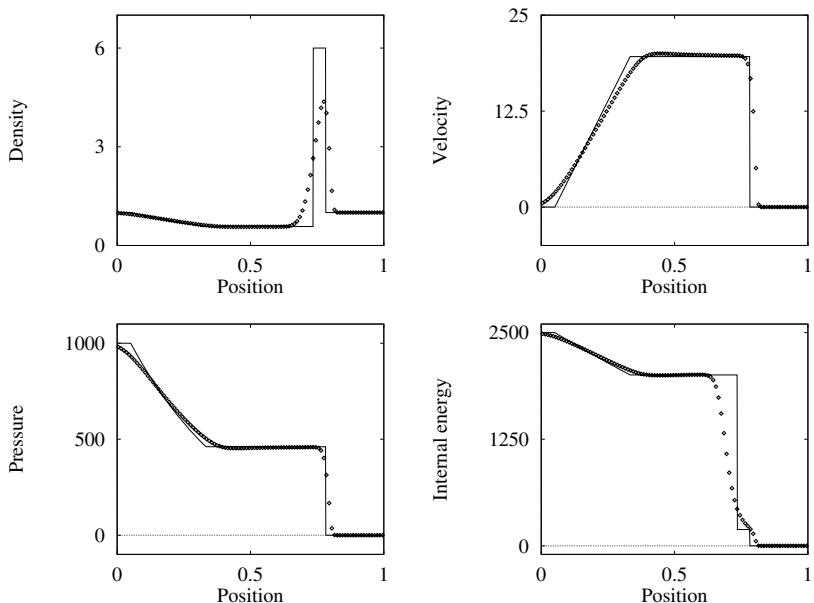
steady supersonic Euler equations are found in the original paper of Roe [407]. For one-dimensional applications all the required information is contained in this chapter and Chap. 6. Second-order Total Variation Diminishing (TVD) extensions of the schemes are presented in Chap. 13 for scalar problems and in Chap. 14 for non-linear one dimensional systems. In chap. 15 we present techniques that allow the extension of these schemes to solve problems with source terms. In Chap. 16 we study techniques to extend the methods of this chapter to three-dimensional problems.



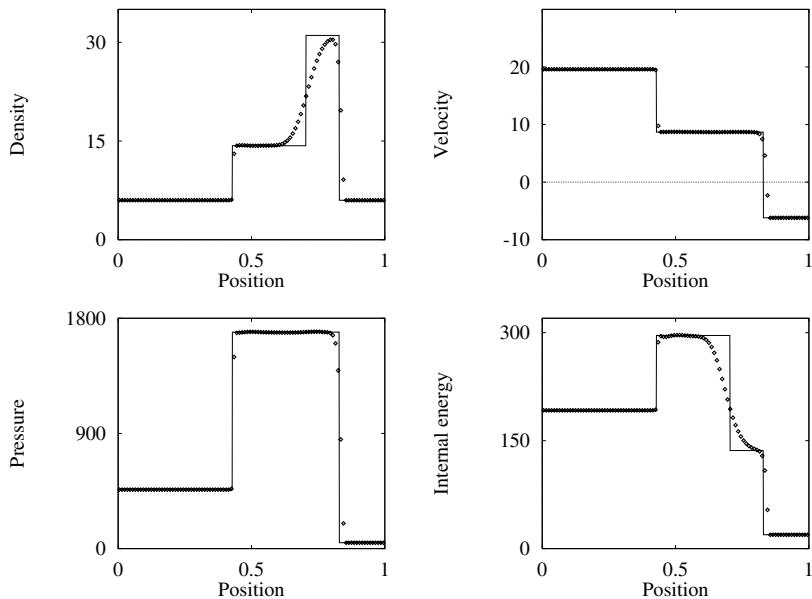
**Fig. 11.4.** Godunov's method with Roe's Riemann solver (no entropy fix) for Test 1,  $x_0 = 0.3$ . Numerical (symbol) and exact (line) solutions compared at time 0.2



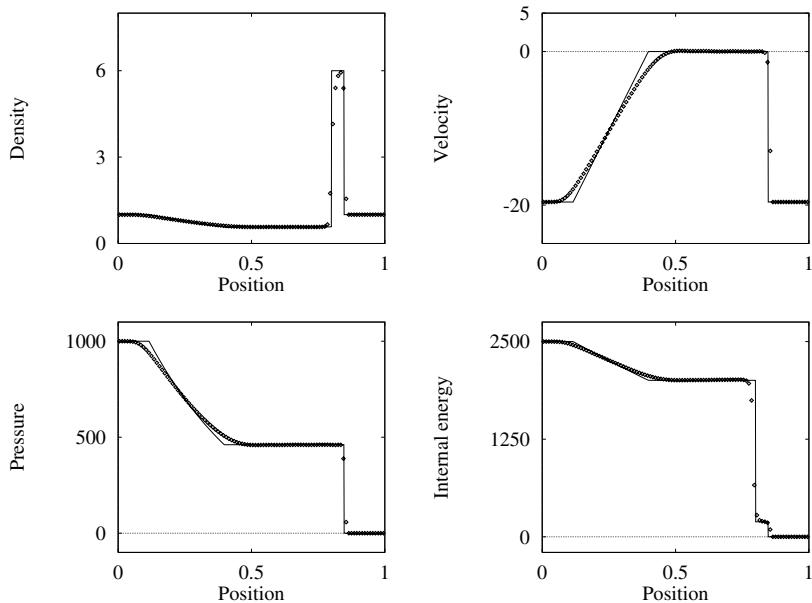
**Fig. 11.5.** Godunov's method with Roe's Riemann solver applied to Test 1, with  $x_0 = 0.3$ . Numerical (symbol) and exact (line) solutions are compared at time 0.2



**Fig. 11.6.** Godunov's method with Roe's Riemann solver applied to Test 3, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012



**Fig. 11.7.** Godunov's method with Roe's Riemann solver applied to Test 4, with  $x_0 = 0.4$ . Numerical (symbol) and exact (line) solutions are compared at time 0.035



**Fig. 11.8.** Godunov's method with Roe's Riemann solver applied to Test 5, with  $x_0 = 0.8$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012

---

## The Riemann Solver of Osher

Osher's approximate Riemann solver is one of the earliest in the literature. The bases of the approach were communicated in the papers by Engquist and Osher in 1981 [185] and Osher and Solomon the following year [372]. Applications to the Euler equations were published later in a paper by Osher and Chakravarthy [370]. Since then the scheme has gained increasing popularity, particularly within the CFD community concerned with Steady Aerodynamics; see for example the works of Spekreijse [458], [459], Hemker and Spekreijse [247], Koren and Spekreijse [290], Qin et. al. [393], [394], [395], [396], [390], [391], [392]. One of the attractions of Osher's scheme is the smoothness of the numerical flux; the scheme has also been proved to be entropy satisfying and in practical computations it is seen to handle sonic flow well. A distinguishing feature of the Osher scheme is its performance near slowly-moving shock waves; see Roberts [406], Billett and Toro [60] and Arora and Roe [19]. The scheme is closely related to the Flux Vector Splitting approach described in Chap. 8 and, as Godunov's method of Chap. 6, it is a generalisation of the CIR scheme described in Chap. 5 for linear hyperbolic systems with constant coefficients. For a scalar conservation law, van Leer [562] studied in detail the relationship between the Osher scheme and some other Riemann solvers available at the time. Useful background material for reading this chapter is found in the previous Chaps. 2, 3, 5, 6, 8 and 9.

The derivation of the Osher intercell numerical flux depends on integration in phase space. Such operation involves the choice of *integration paths*, *intersection points* and *sonic points*. The integration paths are taken to be *integral curves* associated with the set of right eigenvectors and to date there are essentially two ways of *ordering* these integration paths. The most recent approach orders the integration paths such that these correspond to physically meaningful relations across wave families in *physical space*. In the current literature this is called *physical ordering* or *P-ordering*. Osher's original scheme utilises the ordering of paths that is precisely the inverse of the P-ordering; this is usually called *O-ordering*. Intersection and sonic points are computed via *Generalised Riemann Invariants*.

The purpose of this chapter is to present the Osher scheme in a way that can be directly applied to practical problems involving hyperbolic conservation laws. We first describe the principles behind Osher's method as applied to any non-linear system of hyperbolic conservation laws. We then give detailed applications of the approach to a single scalar equation, to the isentropic equations of Gas Dynamics and to the split three-dimensional time dependent Euler equations for ideal gases. For one-dimensional applications, all the required information is found in this chapter and in Chap. 6.

## 12.1 Osher's Scheme for a General System

Here we give, in a self-contained manner, some of the basic aspects of the Osher scheme for a general non-linear system of hyperbolic conservation laws.

### 12.1.1 Mathematical Bases

Osher's approach to upwind differencing provides an approximation to the Godunov numerical flux of Chap. 6 and results from evaluating the physical flux  $\mathbf{F}(\mathbf{U})$  at various states  $\mathbf{U}_k$ ; these include the data states  $\mathbf{U}_L$ ,  $\mathbf{U}_R$ , intersection points and sonic points. Consider a system of  $m$  hyperbolic conservation laws

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0} \quad (12.1)$$

and the conservative scheme

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n + \frac{\Delta t}{\Delta x} \left[ \mathbf{F}_{i-\frac{1}{2}} - \mathbf{F}_{i+\frac{1}{2}} \right] \quad (12.2)$$

to solve it numerically. The objective of this chapter is to provide an expression for the numerical flux  $\mathbf{F}_{i+\frac{1}{2}}$  following Osher's approach.

We assume (12.1) to be *strictly hyperbolic* with eigenvalues

$$\lambda_1(\mathbf{U}) < \lambda_2(\mathbf{U}) < \cdots < \lambda_m(\mathbf{U}) \quad (12.3)$$

and corresponding right eigenvectors

$$\mathbf{K}^{(1)}(\mathbf{U}), \mathbf{K}^{(2)}(\mathbf{U}), \dots, \mathbf{K}^{(m)}(\mathbf{U}).$$

From hyperbolicity, the Jacobian matrix

$$\mathbf{A}(\mathbf{U}) = \partial \mathbf{F} / \partial \mathbf{U} \quad (12.4)$$

is diagonalisable, that is

$$\mathbf{A}(\mathbf{U}) = \mathbf{K}(\mathbf{U}) \mathbf{\Lambda}(\mathbf{U}) \mathbf{K}^{-1}(\mathbf{U}), \quad (12.5)$$

where  $\mathbf{K}(\mathbf{U})$  is the non-singular matrix whose columns are the right eigenvectors of  $\mathbf{A}(\mathbf{U})$ , that is

$$\mathbf{K}(\mathbf{U}) = [\mathbf{K}^{(1)}(\mathbf{U}); \mathbf{K}^{(2)}(\mathbf{U}); \dots; \mathbf{K}^{(m)}(\mathbf{U})] \quad (12.6)$$

and  $\Lambda(\mathbf{U})$  is the diagonal matrix formed by the eigenvalues  $\lambda_i(\mathbf{U})$

$$\Lambda(\mathbf{U}) = \begin{pmatrix} \lambda_1(\mathbf{U}) & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \lambda_m(\mathbf{U}) \end{pmatrix}. \quad (12.7)$$

See Sect. 2.3 of Chap. 2 and Sect. 3.2 of Chap. 3. As done in Sect. 2.3 of Chap. 2 for linear systems with constant coefficients, we introduce the following notation

$$\lambda_i^+(\mathbf{U}) = \max(\lambda_i(\mathbf{U}), 0); \quad \lambda_i^-(\mathbf{U}) = \min(\lambda_i(\mathbf{U}), 0) \quad (12.8)$$

to define diagonal matrices

$$\Lambda^+(\mathbf{U}) = \begin{pmatrix} \lambda_1^+(\mathbf{U}) & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \lambda_m^+(\mathbf{U}) \end{pmatrix}, \quad (12.9)$$

$$\Lambda^-(\mathbf{U}) = \begin{pmatrix} \lambda_1^-(\mathbf{U}) & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \lambda_m^-(\mathbf{U}) \end{pmatrix}, \quad (12.10)$$

and

$$|\Lambda(\mathbf{U})| = \begin{pmatrix} |\lambda_1(\mathbf{U})| & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & |\lambda_m(\mathbf{U})| \end{pmatrix}. \quad (12.11)$$

We also introduce

$$|\mathbf{A}(\mathbf{U})| = \mathbf{K}(\mathbf{U})|\Lambda(\mathbf{U})|\mathbf{K}^{-1}(\mathbf{U}). \quad (12.12)$$

But

$$|\lambda_i(\mathbf{U})| = \lambda_i^+(\mathbf{U}) - \lambda_i^-(\mathbf{U})$$

and hence

$$|\Lambda(\mathbf{U})| = \Lambda^+(\mathbf{U}) - \Lambda^-(\mathbf{U}),$$

which if substituted in (12.12) gives

$$|\mathbf{A}(\mathbf{U})| = \mathbf{A}^+(\mathbf{U}) - \mathbf{A}^-(\mathbf{U}), \quad (12.13)$$

with

$$\mathbf{A}^+(\mathbf{U}) = \mathbf{K}(\mathbf{U})\Lambda^+(\mathbf{U})\mathbf{K}^{-1}(\mathbf{U}) \quad (12.14)$$

and

$$\mathbf{A}^-(\mathbf{U}) = \mathbf{K}(\mathbf{U})\Lambda^-(\mathbf{U})\mathbf{K}^{-1}(\mathbf{U}). \quad (12.15)$$

These matrices produce a splitting of the Jacobian matrix as

$$\mathbf{A}(\mathbf{U}) = \mathbf{A}^+(\mathbf{U}) + \mathbf{A}^-(\mathbf{U}), \quad (12.16)$$

where  $\mathbf{A}^+(\mathbf{U})$  has positive or zero eigenvalues and  $\mathbf{A}^-(\mathbf{U})$  has negative or zero eigenvalues. Note that this is a direct generalisation to non-linear systems of the Jacobian splitting for linear systems with constant coefficients performed in Sect. 5.4 of Chap. 5. See also Sect. 8.2.2 of Chap. 8 on the Flux Vector Splitting approach.

### 12.1.2 Osher's Numerical Flux

Osher's approach assumes that there exist vector-valued functions  $\mathbf{F}^+(\mathbf{U})$  and  $\mathbf{F}^-(\mathbf{U})$  that satisfy

$$\mathbf{F}(\mathbf{U}) = \mathbf{F}^+(\mathbf{U}) + \mathbf{F}^-(\mathbf{U}) \quad (12.17)$$

and

$$\frac{\partial \mathbf{F}^+}{\partial \mathbf{U}} = \mathbf{A}^+(\mathbf{U}); \quad \frac{\partial \mathbf{F}^-}{\partial \mathbf{U}} = \mathbf{A}^-(\mathbf{U}). \quad (12.18)$$

If the initial data  $\mathbf{U}_L$ ,  $\mathbf{U}_R$  of the Riemann problem for the conservation laws (12.1) is denoted by

$$\mathbf{U}_0 \equiv \mathbf{U}_L = \mathbf{U}_i^n; \quad \mathbf{U}_1 \equiv \mathbf{U}_R = \mathbf{U}_{i+1}^n, \quad (12.19)$$

then the corresponding numerical flux to be used in (12.2) is

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}^+(\mathbf{U}_0) + \mathbf{F}^-(\mathbf{U}_1). \quad (12.20)$$

Sect. 8.2.2 of Chap. 8. Using the integral relations

$$\int_{\mathbf{U}_0}^{\mathbf{U}_1} \mathbf{A}^-(\mathbf{U}) d\mathbf{U} = \mathbf{F}^-(\mathbf{U}_1) - \mathbf{F}^-(\mathbf{U}_0)$$

and

$$\int_{\mathbf{U}_0}^{\mathbf{U}_1} \mathbf{A}^+(\mathbf{U}) d\mathbf{U} = \mathbf{F}^+(\mathbf{U}_1) - \mathbf{F}^+(\mathbf{U}_0),$$

we can express (12.20) in three different forms, namely

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}(\mathbf{U}_0) + \int_{\mathbf{U}_0}^{\mathbf{U}_1} \mathbf{A}^-(\mathbf{U}) d\mathbf{U}, \quad (12.21)$$

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}(\mathbf{U}_1) - \int_{\mathbf{U}_0}^{\mathbf{U}_1} \mathbf{A}^+(\mathbf{U}) d\mathbf{U} \quad (12.22)$$

and

$$\mathbf{F}_{i+\frac{1}{2}} = \frac{1}{2} [\mathbf{F}(\mathbf{U}_0) + \mathbf{F}(\mathbf{U}_1)] - \frac{1}{2} \int_{\mathbf{U}_0}^{\mathbf{U}_1} |\mathbf{A}(\mathbf{U})| d\mathbf{U}. \quad (12.23)$$

Compare these flux formulae with those of Sect. 5.4.2 of Chap. 5 for linear systems with constant coefficients and with those for the Roe flux in Sect. 11.1.3 of Chap. 11.

The integration with respect to  $\mathbf{U}$  in (12.21)–(12.23) is carried out in phase space  $\mathbf{R}^m$ . Elements of this vector space are vectors

$$\mathbf{U} = [u_1, u_2, \dots, u_m]^T, \quad (12.24)$$

whose components  $u_i$  are real numbers. In general, the integrals (12.21)–(12.23) depend on the integration path chosen. Osher's approach is to select particular integration paths so as to make the actual integration tractable.

### 12.1.3 Osher's Flux for the Single-Wave Case

The solution of the Riemann problem for (12.1) with data  $\mathbf{U}_0, \mathbf{U}_1$  has  $m$  waves, in general. Osher's scheme utilises partial information on the solution to provide integration paths to evaluate the integrals (12.21)–(12.23), which in turn produce an expression for the numerical flux. Consider first the simplest case in which all waves in the solution of the Riemann problem, except for that associated with the eigenvalue  $\lambda_j(\mathbf{U})$  and eigenvector  $\mathbf{K}^{(j)}(\mathbf{U})$ , are trivial. That is, the states  $\mathbf{U}_0$  and  $\mathbf{U}_1$  are connected by a single  $j$ -wave. Associated with any vector field  $\mathbf{K}^{(k)}(\mathbf{U})$ , there are *integral curves*. These have the property that their tangent lies in the direction of the eigenvector  $\mathbf{K}^{(k)}(\mathbf{U})$ , at any point  $\mathbf{U}$  in phase space. For background on integral curves see [308] and [596]. An integration path  $I_k(\mathbf{U})$  is now taken to be an integral curve of  $\mathbf{K}^{(k)}(\mathbf{U})$  connecting  $\mathbf{U}_0$  and  $\mathbf{U}_1$ . It is important to note here that the eigenvector  $\mathbf{K}^{(k)}(\mathbf{U})$  and the eigenvalue  $\lambda_k(\mathbf{U})$  associated with the relevant integration path are not necessarily those corresponding to the non-trivial  $j$ -wave family in question.

Suppose  $I_k(\mathbf{U})$  is parameterised by  $\mathbf{U}(\xi)$ ,  $0 \leq \xi \leq \xi_1$ , and

$$\mathbf{U}_0 = \mathbf{U}(0), \quad \mathbf{U}_1 = \mathbf{U}(\xi_1); \quad (12.25)$$

then

$$\frac{d\mathbf{U}(\xi)}{d\xi} = \mathbf{K}^{(k)}(\mathbf{U}(\xi)). \quad (12.26)$$

By performing a change of variables, utilising (12.25)–(12.26) and the fact that  $\lambda_k^-(\mathbf{U})$  is an eigenvalue of  $\mathbf{A}^-(\mathbf{U})$  with eigenvector  $\mathbf{K}^{(k)}(\mathbf{U})$ , we have

$$\int_{\mathbf{U}_0}^{\mathbf{U}_1} \mathbf{A}^-(\mathbf{U}) d\mathbf{U} = \int_0^{\xi_1} \lambda_k^-[\mathbf{U}(\xi)] \mathbf{K}^{(k)}[\mathbf{U}(\xi)] d\xi. \quad (12.27)$$

We first consider the case in which the  $k$ -th field is *linearly degenerate*, that is  $\lambda_k(\mathbf{U})$  is constant along  $I_k(\mathbf{U})$ ; see Sect. 2.4.3 of Chap. 2 and Sect. 3.1.3 of Chap. 3. If  $\lambda_k(\mathbf{U}) \geq 0 \quad \forall \mathbf{U}$  along  $I_k(\mathbf{U})$ , then  $\lambda_k^-(\mathbf{U}) = 0$ , see (12.8), and thus from (12.27) we have

$$\int_{\mathbf{U}_0}^{\mathbf{U}_1} \mathbf{A}^-(\mathbf{U}) d\mathbf{U} = \mathbf{0}. \quad (12.28)$$

If  $\lambda_k(\mathbf{U}) \leq 0 \quad \forall \mathbf{U}$  along  $I_k(\mathbf{U})$ , then manipulations of (12.27) give

$$\int_{\mathbf{U}_0}^{\mathbf{U}_1} \mathbf{A}^-(\mathbf{U}) d\mathbf{U} = \int_0^{\xi_1} \frac{d\mathbf{F}}{d\mathbf{U}}[\mathbf{U}(\xi)] \frac{d\mathbf{U}}{d\xi} d\xi$$

and thus

$$\int_{\mathbf{U}_0}^{\mathbf{U}_1} \mathbf{A}^-(\mathbf{U}) d\mathbf{U} = \mathbf{F}(\mathbf{U}_1) - \mathbf{F}(\mathbf{U}_0). \quad (12.29)$$

Hence, if the  $k$ -field is linearly degenerate, use of (12.28) and (12.29) in the one-sided flux formula (12.21) gives the Osher's intercell flux as

$$\mathbf{F}_{i+\frac{1}{2}} = \begin{cases} \mathbf{F}(\mathbf{U}_0) & \text{if } \lambda_k \geq 0 \\ \mathbf{F}(\mathbf{U}_1) & \text{if } \lambda_k < 0. \end{cases} \quad (12.30)$$

An entirely equivalent derivation of the Osher flux results from using the flux formulae (12.22) or (12.23). Note that in the case of a single wave associated with a linearly degenerate field, the Osher flux is identical to the Godunov flux if  $j = k$ .

We next consider the case in which the  $k$ -th field is *genuinely non-linear*, that is the eigenvalue  $\lambda_k(\mathbf{U})$  is *monotone* along  $I_k(\mathbf{U})$ . This means that  $\lambda_k(\mathbf{U})$  changes sign *at most once*, along  $I_k(\mathbf{U})$ . If  $\lambda_k(\mathbf{U})$  does not change sign along  $I_k(\mathbf{U})$  then this is simply like the linearly degenerate case above and the flux is given by (12.30). If  $\lambda_k(\mathbf{U})$  changes sign at  $\xi = \xi_S$ , then there are two cases to consider. First assume

$$\lambda_k[\mathbf{U}(\xi)] \geq 0, \quad \forall \xi \in [0, \xi_S]; \quad \lambda_k[\mathbf{U}(\xi)] \leq 0, \quad \forall \xi \in [\xi_S, \xi_1].$$

Then (12.27) can be split into an integral between 0 and  $\xi_S$  and an integral between  $\xi_S$  and  $\xi_1$ . The first integral is zero, as  $\lambda_k^-(\mathbf{U}) = \min(\lambda_k(\mathbf{U}), 0)$  and  $\lambda_k(\mathbf{U}) \geq 0$ , see (12.8). The second integral gives

$$\int_{\xi_S}^{\xi_1} \frac{d\mathbf{F}}{d\mathbf{U}}[\mathbf{U}(\xi)] \frac{d\mathbf{U}}{d\xi} d\xi = \mathbf{F}(\mathbf{U}_1) - \mathbf{F}(\mathbf{U}_S),$$

where  $\mathbf{U}_S = \mathbf{U}(\xi_S)$  and is called a *sonic point*. The case

$$\lambda_k[\mathbf{U}(\xi)] \leq 0, \quad \forall \xi \in [0, \xi_S]; \quad \lambda_k[\mathbf{U}(\xi)] \geq 0, \quad \forall \xi \in [\xi_S, \xi_1]$$

can be treated in a similar way to give

$$\int_{\mathbf{U}_0}^{\mathbf{U}_1} \mathbf{A}^-(\mathbf{U}) d\mathbf{U} = \mathbf{F}(\mathbf{U}_S) - \mathbf{F}(\mathbf{U}_0) .$$

Collecting results, for a single wave, we have

$$\int_{\mathbf{U}_0}^{\mathbf{U}_1} \mathbf{A}^-(\mathbf{U}) d\mathbf{U} = \begin{cases} \mathbf{0} & \text{if } \lambda_k(\mathbf{U}) \geq 0 , \\ \mathbf{F}(\mathbf{U}_1) - \mathbf{F}(\mathbf{U}_0) & \text{if } \lambda_k(\mathbf{U}) \leq 0 , \\ \mathbf{F}(\mathbf{U}_1) - \mathbf{F}(\mathbf{U}_S) & \text{if } \lambda_k(\mathbf{U}_0) \geq 0 , \lambda_k(\mathbf{U}_1) \leq 0 , \\ & \quad \lambda_k(\mathbf{U}_S) = 0 , \\ \mathbf{F}(\mathbf{U}_S) - \mathbf{F}(\mathbf{U}_0) & \text{if } \lambda_k(\mathbf{U}_0) \leq 0 , \lambda_k(\mathbf{U}_1) \geq 0 , \\ & \quad \lambda_k(\mathbf{U}_S) = 0 . \end{cases}$$

By substituting these expressions into the one-sided flux formula (12.21) we obtain the Osher intercell flux, for the case in which the states  $\mathbf{U}_0$  and  $\mathbf{U}_1$  are connected by the single  $j$ -wave.

$$\mathbf{F}_{i+\frac{1}{2}} = \begin{cases} \mathbf{F}(\mathbf{U}_0) , & \lambda_k(\mathbf{U}) \geq 0 , \\ \mathbf{F}(\mathbf{U}_1) , & \lambda_k(\mathbf{U}) \leq 0 , \\ \mathbf{F}(\mathbf{U}_0) + \mathbf{F}(\mathbf{U}_1) - \mathbf{F}(\mathbf{U}_S) , & \lambda_k(\mathbf{U}_0) \geq 0 , \lambda_k(\mathbf{U}_1) \leq 0 , \\ & \quad \lambda_k(\mathbf{U}_S) = 0 , \\ \mathbf{F}(\mathbf{U}_S) , & \lambda_k(\mathbf{U}_0) \leq 0 , \lambda_k(\mathbf{U}_1) \geq 0 , \\ & \quad \lambda_k(\mathbf{U}_S) = 0 . \end{cases} \quad (12.31)$$

These results can be applied directly to any scalar, non-linear conservation law

$$u_t + f(u)_x = 0 . \quad (12.32)$$

#### 12.1.4 Osher's Flux for the Inviscid Burgers Equation

Consider (12.32) with flux function

$$f(u) = \frac{1}{2}u^2 . \quad (12.33)$$

This gives the inviscid Burgers equation; see Sect. 2.4.2 of Chap. 2. Recall that the exact solution of the Riemann problem for (12.32)–(12.33) with data  $u_L \equiv u_0$ ,  $u_R \equiv u_1$  is

$$u(x/t) = \begin{cases} u_0 & \text{if } x/t \leq S = \frac{1}{2}(u_0 + u_1) \\ u_1 & \text{if } x/t > S \end{cases} \quad (12.34)$$

when  $u_0 > u_1$  (shock case), and

$$u(x/t) = \begin{cases} u_0 & \text{if } x/t \leq u_0 , \\ x/t & \text{if } u_0 \leq x/t \leq u_1 , \\ u_1 & \text{if } x/t \geq u_1 , \end{cases} \quad (12.35)$$

when  $u_0 \leq u_1$  (rarefaction case). As, trivially, the only eigenvalue is the characteristic speed  $\lambda = u$ , direct application of (12.31) gives

$$\mathbf{F}_{i+\frac{1}{2}} = \begin{cases} f(u_0) & \text{if } u_0, u_1 > 0 , \\ f(u_1) & \text{if } u_0, u_1 < 0 , \\ f(u_0) + f(u_1) & \text{if } u_0 \geq 0 \geq u_1 , \\ 0 & \text{if } u_0 \leq 0 \leq u_1 . \end{cases} \quad (12.36)$$

At the sonic point  $u_S = 0$  and thus  $f(u_S) = 0$ . Note that the Godunov flux  $f_{i+\frac{1}{2}}^{god}$ , obtained from the exact solution to the Riemann problem, is identical to Osher's flux (12.36), except in the case of a *transonic shock* ( $u_0 \geq 0 \geq u_1$ ), where

$$f_{i+\frac{1}{2}}^{god} = \begin{cases} f(u_0) & \text{if } S \geq 0 , \\ f(u_1) & \text{otherwise .} \end{cases}$$

For a full discussion on the relationship between the Godunov scheme, with the exact Riemann solver, and approximate Riemann solvers, including Osher's scheme, see the paper by van Leer [562].

### 12.1.5 Osher's Flux for the General Case

In the previous section we analysed Osher's numerical flux for the case of two states  $\mathbf{U}_0, \mathbf{U}_1$  connected by a single wave, where an integration path was chosen to be tangential to a right eigenvector. For the general case of  $m$  wave families Osher chooses a set  $\{I_k(\mathbf{U})\}$ ,  $k = 1, \dots, m$ , of *partial* integration paths such that  $I_k(\mathbf{U})$  is tangential to  $\mathbf{K}^{(k)}(\mathbf{U})$  and two successive paths  $I_k(\mathbf{U})$  and  $I_{k+1}(\mathbf{U})$  intersect at a *single* point

$$\mathbf{U}_{k/m} = I_k(\mathbf{U}) \cap I_{k+1}(\mathbf{U}) \quad (12.37)$$

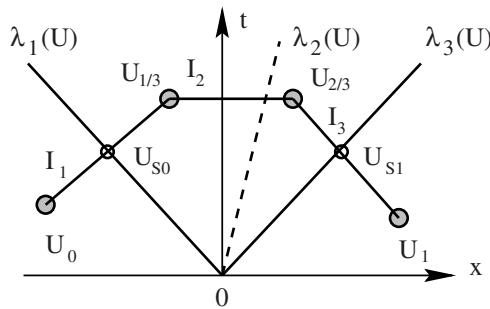
in phase space, called an *intersection point*. The data points  $\mathbf{U}_0$  and  $\mathbf{U}_1$  are now to be interpreted as

$$\mathbf{U}_0 \equiv \mathbf{U}_{(k-1)/m} , \quad \mathbf{U}_1 \equiv \mathbf{U}_{k/m} .$$

The total integration path is the union of all partial paths, namely

$$I(\mathbf{U}) = \cup I_k(\mathbf{U}) . \quad (12.38)$$

Fig. 12.1 illustrates a choice of integration paths for the case of a  $3 \times 3$  hyperbolic system such as the one-dimensional, time dependent Euler equations. The vectors  $\mathbf{U}_0, \mathbf{U}_{1/3}, \mathbf{U}_{2/3}$  and  $\mathbf{U}_1$  are vectors in phase space  $\mathbf{R}^3$  and can be thought of as being the four constant states arising in the exact solution to the corresponding Riemann problem represented in physical space  $x-t$  in Fig. 12.1. See Chap. 4 for details of the exact solution of the Riemann problem for the Euler equations. The points  $\mathbf{U}_{\frac{1}{3}}$  and  $\mathbf{U}_{\frac{2}{3}}$  are the *intersection points* in phase space, and the points  $\mathbf{U}_{S0}$  and  $\mathbf{U}_{S1}$  are a representation of the potential *sonic points* that may arise from the non-linear fields associated with  $\lambda_1(\mathbf{U})$  and  $\lambda_3(\mathbf{U})$ .



**Fig. 12.1.** Possible configuration of integration paths  $I_k(\mathbf{U})$ , intersection points  $\mathbf{U}_{\frac{1}{3}}, \mathbf{U}_{\frac{2}{3}}$  and sonic points  $\mathbf{U}_{S0}, \mathbf{U}_{S1}$  in physical space  $x-t$  for a  $3 \times 3$  system

For the general case, and assuming for the moment that the intersection points (12.37) and sonic points are known, the integration along  $I(\mathbf{U})$  to evaluate (12.21), say, can now be performed by integrating along each partial integration path  $I_k(\mathbf{U})$ . But since these have been chosen to be tangential to the corresponding eigenvector  $\mathbf{K}^{(k)}(\mathbf{U})$ , the results of the previous section, see (12.31), can now be applied directly. The determination of the intersection points  $\mathbf{U}_{k/m}$  and sonic points requires extra information about the solution of the Riemann problem. Traditionally, these have been determined by the use of Generalised Riemann Invariants, in at least two different ways, as we shall see.

## 12.2 Osher's Flux for the Isothermal Equations

The isothermal equations

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0}, \quad (12.39)$$

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ \rho u^2 + a^2 \rho \end{bmatrix}, \quad \mathbf{F}(\mathbf{U}) = \begin{bmatrix} \rho u \\ \rho u^2 + a^2 \rho \\ \rho u^3 + 2a\rho u \end{bmatrix}, \quad (12.40)$$

have Jacobian matrix

$$\mathbf{A}(\mathbf{U}) = \begin{bmatrix} 0 & 1 \\ a^2 - u^2 & 2u \end{bmatrix} \quad (12.41)$$

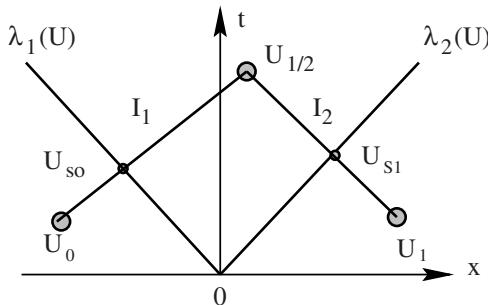
with eigenvalues

$$\lambda_1(\mathbf{U}) = u - a, \quad \lambda_2(\mathbf{U}) = u + a, \quad (12.42)$$

and corresponding right eigenvectors

$$\mathbf{K}^{(1)}(\mathbf{U}) = \begin{bmatrix} 1 \\ u - a \end{bmatrix}, \quad \mathbf{K}^{(2)}(\mathbf{U}) = \begin{bmatrix} 1 \\ u + a \end{bmatrix}. \quad (12.43)$$

Recall that the sound speed  $a$  is constant here; for details on the isothermal equations see Sect. 1.6.2 of Chap. 1 and Sect. 2.4.1 of Chap. 2. The structure of the exact solution to the Riemann problem for (12.39)–(12.40) with initial data  $\mathbf{U}_0, \mathbf{U}_1$  is depicted in Fig. 12.2 in the  $x-t$  plane, where the intersection point  $\mathbf{U}_{\frac{1}{2}}$  is identified with the solution of the Riemann problem between the non-linear waves; potential sonic points are also shown.



**Fig. 12.2.** Structure of the solution of the Riemann problem for the isothermal equations in physical space  $x-t$ . Integration paths are  $I_1$  and  $I_2$ , intersection point is  $\mathbf{U}_{\frac{1}{2}}$ , potential sonic points are  $\mathbf{U}_{S0}$  and  $\mathbf{U}_{S1}$ ; P-ordering

### 12.2.1 Osher's Flux with P-Ordering

There are two ways of ordering the integration paths in Osher's scheme, namely the *original* Osher ordering, or *O-ordering*, and the *physical* ordering following valid relations across waves in physical space, or *P-ordering*. First we apply *P-ordering*. It is easy to check, see Sect. 2.4.3 of Chap. 2, that the Riemann Invariants along the path  $I_1(\mathbf{U})$ , across the left wave in physical space, give

$$u_{\frac{1}{2}} + a \ln(\rho_{\frac{1}{2}}) = u_0 + a \ln(\rho_0). \quad (12.44)$$

Similarly, across the right wave

$$u_{\frac{1}{2}} - a \ln(\rho_{\frac{1}{2}}) = u_0 - a \ln(\rho_1) . \quad (12.45)$$

The simultaneous solution for  $\rho_{\frac{1}{2}}$  and  $u_{\frac{1}{2}}$  is

$$u_{\frac{1}{2}} = \frac{1}{2}(u_0 + u_1) + \frac{1}{2}a \ln(\rho_0/\rho_1) , \quad (12.46)$$

$$\rho_{\frac{1}{2}} = \sqrt{\rho_0 \rho_1} \exp \left[ -\frac{u_1 - u_0}{2a} \right] \quad (12.47)$$

and thus the intersection point

$$\mathbf{U}_{\frac{1}{2}} = \begin{bmatrix} \rho_{\frac{1}{2}} \\ \rho_{\frac{1}{2}} u_{\frac{1}{2}} \end{bmatrix} \quad (12.48)$$

is determined.

	Behaviour of $\lambda_1(\mathbf{U}) = u - a$	$\int_{\mathbf{U}_0}^{\mathbf{U}_{\frac{1}{2}}} \mathbf{A}^-(\mathbf{U}) d\mathbf{U}$
1	$u_0 - a \geq 0 , u_{\frac{1}{2}} - a \geq 0$	$\mathbf{0}$
2	$u_0 - a \leq 0 , u_{\frac{1}{2}} - a \leq 0$	$\mathbf{F}_{\frac{1}{2}} - \mathbf{F}_0$
3	$u_0 - a \geq 0 , u_{\frac{1}{2}} - a \leq 0$	$\mathbf{F}_{\frac{1}{2}} - \mathbf{F}_{S0}$
4	$u_0 - a \leq 0 , u_{\frac{1}{2}} - a \geq 0$	$\mathbf{F}_{S0} - \mathbf{F}_0$

**Table 12.1.** Evaluation of integral along integration path  $I_1(\mathbf{U})$  with P-ordering,  $\mathbf{F}_k \equiv \mathbf{F}(\mathbf{U}_k)$

The left and right sonic points  $\mathbf{U}_{S0}$  and  $\mathbf{U}_{S1}$  can easily be found by using the sonic conditions  $u_{S0} = a$  from  $\lambda_1 = u - a = 0$ ,  $u_{S1} = -a$  from  $\lambda_2 = u + a = 0$  and the Riemann Invariants. The result is

$$\rho_{S0} = \rho_0 \exp \left[ \frac{u_0 - a}{a} \right] , \quad u_{S0} = a , \quad (12.49)$$

$$\rho_{S1} = \rho_1 \exp \left[ -\frac{u_1 + a}{a} \right] , \quad u_{S1} = -a . \quad (12.50)$$

Behaviour of $\lambda_2(\mathbf{U}) = u + a$		$\int_{\mathbf{U}_{\frac{1}{2}}}^{\mathbf{U}_1} \mathbf{A}^-(\mathbf{U}) d\mathbf{U}$
1	$u_{\frac{1}{2}} + a \geq 0, u_1 + a \geq 0$	$\mathbf{0}$
2	$u_{\frac{1}{2}} + a \leq 0, u_1 + a \leq 0$	$\mathbf{F}_1 - \mathbf{F}_{\frac{1}{2}}$
3	$u_{\frac{1}{2}} + a \geq 0, u_1 + a \leq 0$	$\mathbf{F}_1 - \mathbf{F}_{S1}$
4	$u_{\frac{1}{2}} + a \leq 0, u_1 + a \geq 0$	$\mathbf{F}_{S1} - \mathbf{F}_{\frac{1}{2}}$

**Table 12.2.** Evaluation of integral along integration path  $I_2(\mathbf{U})$  with P-ordering,  $\mathbf{F}_k \equiv \mathbf{F}(\mathbf{U}_k)$

It is worth remarking at this stage, that the solution (12.46)–(12.47) for the intersection point is exact when the left and right waves are both rarefaction waves; in the general case it is an approximation. It is in fact the Two-Rarefaction approximation TRRS presented in Sect. 9.4.1 of Chap. 9 for the Euler equations.

### Integration Along Partial Paths

In order to compute the Osher flux we use the one-sided flux formula

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}_0 + \int_{\mathbf{U}_0}^{\mathbf{U}_1} \mathbf{A}^-(\mathbf{U}) d\mathbf{U}, \quad (12.51)$$

where the integral is evaluated along each of the partial integration paths  $I_1(\mathbf{U})$  and  $I_2(\mathbf{U})$  shown in Fig. 12.2. For each case the integration is performed according to the local characteristic configuration. Tables 12.1 and 12.2 show the results for  $I_1(\mathbf{U})$  and  $I_2(\mathbf{U})$  respectively. As

$$\int_{\mathbf{U}_0}^{\mathbf{U}_1} \mathbf{A}^-(\mathbf{U}) d\mathbf{U} = \int_{\mathbf{U}_0}^{\mathbf{U}_{\frac{1}{2}}} \mathbf{A}^-(\mathbf{U}) d\mathbf{U} + \int_{\mathbf{U}_{\frac{1}{2}}}^{\mathbf{U}_1} \mathbf{A}^-(\mathbf{U}) d\mathbf{U}, \quad (12.52)$$

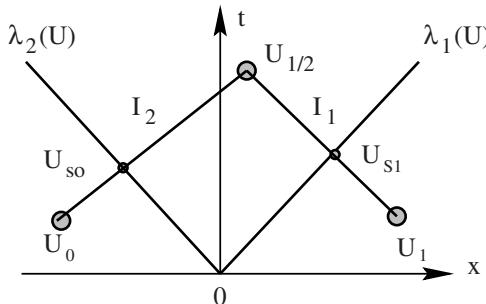
strictly speaking, one should consider all 16 possible characteristic configurations  $(i, j)$  that result from Tables 12.1 and 12.2. Closer examination of all cases reveals that 4 possibilities are unrealisable. These are (1,2), (1,4), (4,2) and (4,4). For instance case (1,2) contains the requirements

$$u_{\frac{1}{2}} - a \geq 0 ; \quad u_{\frac{1}{2}} + a \leq 0 . \quad (12.53)$$

But  $a > 0$  and thus these two conditions are contradictory. The remaining 12 cases can be tabulated as in the  $3 \times 4$  Table 12.3, which gives the Osher intercell numerical flux  $\mathbf{F}_{i+\frac{1}{2}}$ . We separate conditions on the intersection point  $\mathbf{U}_{\frac{1}{2}}$  (first column) from conditions on the data points  $\mathbf{U}_0$  and  $\mathbf{U}_1$  (top row). Note that in general  $\mathbf{F}_{i+\frac{1}{2}}$  is a combination of physical flux values at several points in phase space. In contrast, the Godunov flux obtained from the exact Riemann solver of Chap. 4 and from the approximate Riemann solvers of Chaps. 9, 10 and 11, consists of a *single value*. This has a bearing on the simplicity and computational efficiency of the schemes.

### 12.2.2 Osher's Flux with O-Ordering

In the original Osher scheme the ordering of the partial integration paths  $I_k(\mathbf{U})$  is inverted. Fig. 12.3 illustrates the corresponding path configuration for the isothermal equations. The integration path connecting  $\mathbf{U}_0$  to  $\mathbf{U}_{\frac{1}{2}}$  is tangential to the eigenvector  $\mathbf{K}^{(2)}(\mathbf{U})$  and that connecting  $\mathbf{U}_{\frac{1}{2}}$  to  $\mathbf{U}_1$  is tangential to  $\mathbf{K}^{(1)}(\mathbf{U})$ . Compare with Fig. 12.2. The O-ordering of the Osher scheme is used both for the integration paths as well as for the determination of the intersection point  $\mathbf{U}_{\frac{1}{2}}$  and the sonic points  $\mathbf{U}_{S0}, \mathbf{U}_{S1}$ . The data state



**Fig. 12.3.** Configuration of integration paths  $I_k(\mathbf{U})$ , intersection point  $\mathbf{U}_{\frac{1}{2}}$  and potential sonic points  $\mathbf{U}_{S0}, \mathbf{U}_{S1}$  in physical space  $x-t$  when using the O-ordering

$\mathbf{U}_0$  is connected to  $\mathbf{U}_{\frac{1}{2}}$  via the *right* (relation valid across the right wave) Riemann Invariant to give

$$u_{\frac{1}{2}} - a \ln(\rho_{\frac{1}{2}}) = u_0 - a \ln(\rho_0) .$$

Compare with (12.44). The data point  $\mathbf{U}_1$  is connected to  $\mathbf{U}_{\frac{1}{2}}$  via the *left* (relation valid across the left wave) Riemann Invariant to produce

$$u_{\frac{1}{2}} + a \ln(\rho_{\frac{1}{2}}) = u_1 + a \ln(\rho_1) .$$

	$u_0 - a \geq 0$ $u_1 + a \geq 0$	$u_0 - a \geq 0$ $u_1 + a \leq 0$	$u_0 - a \leq 0$ $u_1 + a \leq 0$	$u_0 - a \leq 0$ $u_1 + a \geq 0$
$u_{\frac{1}{2}} - a \geq 0$	$\mathbf{F}_0$	$\mathbf{F}_0 + \mathbf{F}_1$ $-\mathbf{F}_{S1}$	$\mathbf{F}_{S0} + \mathbf{F}_1$ $-\mathbf{F}_{S1}$	$\mathbf{F}_{S0}$
$u_{\frac{1}{2}} - a \leq 0$ $u_{\frac{1}{2}} + a \geq 0$	$\mathbf{F}_0 + \mathbf{F}_{\frac{1}{2}}$ $-\mathbf{F}_{S0}$	$\mathbf{F}_0 + \mathbf{F}_{\frac{1}{2}}$ $-\mathbf{F}_{S0} + \mathbf{F}_1$ $-\mathbf{F}_{S1}$	$\mathbf{F}_{\frac{1}{2}} + \mathbf{F}_1$ $-\mathbf{F}_{S1}$	$\mathbf{F}_{\frac{1}{2}}$
$u_{\frac{1}{2}} + a \leq 0$	$\mathbf{F}_0 + \mathbf{F}_{S1}$ $-\mathbf{F}_{S0}$	$\mathbf{F}_0 + \mathbf{F}_1$ $-\mathbf{F}_{S0}$	$\mathbf{F}_1$	$\mathbf{F}_{S1}$

**Table 12.3.** Osher's intercell flux for the isothermal equations using P-ordering,  $\mathbf{F}_k \equiv \mathbf{F}(\mathbf{U}_k)$

Compare with (12.45). The resulting solution is

$$u_{\frac{1}{2}} = \frac{1}{2}(u_0 + u_1) + \frac{1}{2}a \ln(\rho_1/\rho_0), \quad (12.54)$$

$$\rho_{\frac{1}{2}} = \sqrt{\rho_0 \rho_1} \exp \left[ \frac{u_1 - u_0}{2a} \right] \quad (12.55)$$

and thus the intersection point  $\mathbf{U}_{\frac{1}{2}}$  in (12.48) is determined. The left and right sonic points are evaluated using the sonic conditions  $\lambda_2 = u + a = 0$  and  $\lambda_1 = u - a = 0$  and the right and left Riemann Invariants, respectively. The result is

$$u_{S0} = -a, \quad \rho_{S0} = \rho_0 \exp \left[ -\frac{u_0 + a}{a} \right], \quad (12.56)$$

$$u_{S1} = a, \quad \rho_{S1} = \rho_1 \exp \left[ \frac{u_1 - a}{a} \right]. \quad (12.57)$$

Compare solutions (12.54)–(12.57) with (12.46)–(12.50) obtained with the P-ordering of the Osher scheme. Note that the solution for the intersection point given by the O-ordering has no physical meaning as a solution for the *Star Region* between the left and right wave families. If the exact solution of the Riemann problem consists of two rarefactions, then solution (12.47) is exact, in which case the density is expected to decrease; solution (12.55) gives an increase in density in the *Star Region*, which is obviously incorrect. We expand on this point when dealing with the Euler equations later in Sect. 12.3.3.

Integration along the path  $I_2(\mathbf{U})$  connecting  $\mathbf{U}_0$  to  $\mathbf{U}_{\frac{1}{2}}$  produces the results of Table 12.4. The results of integrating along the path  $I_1(\mathbf{U})$  connecting  $\mathbf{U}_{\frac{1}{2}}$  to  $\mathbf{U}_1$  are given in Table 12.5.

	Behaviour of $\lambda_2(\mathbf{U}) = u + a$	$\int_{\mathbf{U}_0}^{\mathbf{U}_{\frac{1}{2}}} \mathbf{A}^-(\mathbf{U}) d\mathbf{U}$
1	$u_0 + a \geq 0, u_{\frac{1}{2}} + a \geq 0$	$\mathbf{0}$
2	$u_0 + a \leq 0, u_{\frac{1}{2}} + a \leq 0$	$\mathbf{F}_{\frac{1}{2}} - \mathbf{F}_0$
3	$u_0 + a \geq 0, u_{\frac{1}{2}} + a \leq 0$	$\mathbf{F}_{\frac{1}{2}} - \mathbf{F}_{S0}$
4	$u_0 + a \leq 0, u_{\frac{1}{2}} + a \geq 0$	$\mathbf{F}_{S0} - \mathbf{F}_0$

**Table 12.4.** Integration along path  $I_2(\mathbf{U})$  connecting  $\mathbf{U}_0$  to  $\mathbf{U}_{\frac{1}{2}}$  (O-ordering),  $\mathbf{F}_k \equiv \mathbf{F}(\mathbf{U}_k)$

	Behaviour of $\lambda_1(\mathbf{U}) = u - a$	$\int_{\mathbf{U}_{\frac{1}{2}}}^{\mathbf{U}_1} \mathbf{A}^-(\mathbf{U}) d\mathbf{U}$
1	$u_1 - a \geq 0, u_{\frac{1}{2}} - a \geq 0$	$\mathbf{0}$
2	$u_1 - a \leq 0, u_{\frac{1}{2}} - a \leq 0$	$\mathbf{F}_1 - \mathbf{F}_{\frac{1}{2}}$
3	$u_1 - a \leq 0, u_{\frac{1}{2}} - a \geq 0$	$\mathbf{F}_1 - \mathbf{F}_{S1}$
4	$u_1 - a \geq 0, u_{\frac{1}{2}} - a \leq 0$	$\mathbf{F}_{S1} - \mathbf{F}_{\frac{1}{2}}$

**Table 12.5.** Integration along path  $I_1(\mathbf{U})$  connecting  $\mathbf{U}_{\frac{1}{2}}$  to  $\mathbf{U}_1$  (O-ordering),  $\mathbf{F}_k \equiv \mathbf{F}(\mathbf{U}_k)$

Out of the 16 possible combinations 12 cases are realisable. Table 12.6 gives the Osher numerical flux  $\mathbf{F}_{i+\frac{1}{2}}$  for all 12 cases. Compare with Table 12.3 in which the flux is computed using the P-ordering.

	$u_0 + a \geq 0$ $u_1 - a \geq 0$	$u_0 + a \geq 0$ $u_1 - a \leq 0$	$u_0 + a \leq 0$ $u_1 - a \leq 0$	$u_0 + a \leq 0$ $u_1 - a \geq 0$
$u_{\frac{1}{2}} - a \geq 0$	$\mathbf{F}_0$	$\mathbf{F}_0 + \mathbf{F}_1$ $-\mathbf{F}_{S1}$	$\mathbf{F}_{S0} + \mathbf{F}_1$ $-\mathbf{F}_{S1}$	$\mathbf{F}_{S0}$
$u_{\frac{1}{2}} - a \leq 0$ $u_{\frac{1}{2}} + a \geq 0$	$\mathbf{F}_0 + \mathbf{F}_{S1}$ $-\mathbf{F}_{\frac{1}{2}}$	$\mathbf{F}_0 + \mathbf{F}_1$ $-\mathbf{F}_{\frac{1}{2}}$	$\mathbf{F}_{S0} + \mathbf{F}_1$ $-\mathbf{F}_{\frac{1}{2}}$	$\mathbf{F}_{S0} + \mathbf{F}_{S1}$ $-\mathbf{F}_{\frac{1}{2}}$
$u_{\frac{1}{2}} + a \leq 0$	$\mathbf{F}_0 + \mathbf{F}_{S1}$ $-\mathbf{F}_{S0}$	$\mathbf{F}_0 + \mathbf{F}_1$ $-\mathbf{F}_{S0}$	$\mathbf{F}_1$	$\mathbf{F}_{S1}$

**Table 12.6.** Osher's intercell flux for the isothermal equations using O-ordering of integration paths,  $\mathbf{F}_k = \mathbf{F}(\mathbf{U}_k)$

### 12.3 Osher's Scheme for the Euler Equations

Here we develop in detail the Osher scheme, with both P and O orderings, for the time-dependent Euler equations. We first consider the one-dimensional case

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0}, \quad (12.58)$$

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ E \end{bmatrix}, \quad \mathbf{F}(\mathbf{U}) = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ u(E + p) \end{bmatrix}. \quad (12.59)$$

Details of the Euler equations are found in Sect. 1.1 of Chap. 1 and Chap. 3. We require an expression for the intercell flux  $\mathbf{F}_{i+\frac{1}{2}}$  in the explicit conservative formula

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n + \frac{\Delta t}{\Delta x} [\mathbf{F}_{i-\frac{1}{2}} - \mathbf{F}_{i+\frac{1}{2}}]. \quad (12.60)$$

Recall that the Jacobian matrix  $\mathbf{A}(\mathbf{U})$ , see Sect. 3.1.2 of Chap. 3, has eigenvalues

$$\lambda_1 = u - a, \quad \lambda_2 = u, \quad \lambda_3 = u + a \quad (12.61)$$

and right eigenvectors

$$\mathbf{K}^{(1)} = \begin{bmatrix} 1 \\ u-a \\ H-ua \end{bmatrix}, \quad \mathbf{K}^{(2)} = \begin{bmatrix} 1 \\ u \\ \frac{1}{2}u^2 \end{bmatrix}, \quad \mathbf{K}^{(3)} = \begin{bmatrix} 1 \\ u+a \\ H+ua \end{bmatrix}. \quad (12.62)$$

It is instructive to relate Osher's scheme to the solution of the Riemann problem in the classical sense, see Chap. 4 for details. Fig. 12.1 shows the structure of the solution of the Riemann problem with data  $\mathbf{U}_0, \mathbf{U}_1$  in the  $x-t$  plane. Also shown there are the partial integration paths  $I_1(\mathbf{U}), I_2(\mathbf{U}), I_3(\mathbf{U})$ , the intersection points  $\mathbf{U}_{\frac{1}{3}}, \mathbf{U}_{\frac{2}{3}}$  and the sonic points  $\mathbf{U}_{S0}, \mathbf{U}_{S1}$  in phase space; the illustrated paths follow the *P-ordering*. There are essentially two steps in obtaining the Osher flux formulae. First the intersection points  $\mathbf{U}_{\frac{1}{3}}$  and  $\mathbf{U}_{\frac{2}{3}}$  are obtained. We identify these points with the states  $\mathbf{U}_{*L}$  and  $\mathbf{U}_{*R}$  in the solution of the Riemann problem, see Chaps. 4 and 9. The second step consists of evaluating the integral in (12.21), for instance, along the integration paths  $I_k(\mathbf{U})$  to obtain the intercell flux.

### 12.3.1 Osher's Flux with P-Ordering

The physical or *P-ordering* of integration paths for the Euler equations is illustrated in Fig. 12.1. States  $\mathbf{U}_0$  and  $\mathbf{U}_{\frac{1}{3}}$  are connected by the partial integration path  $I_1(\mathbf{U})$ , which is taken to be tangential to the right eigenvector  $\mathbf{K}^{(1)}(\mathbf{U})$  in (12.62). Similarly,  $I_2(\mathbf{U})$  connects  $\mathbf{U}_{\frac{1}{3}}$  to  $\mathbf{U}_{\frac{2}{3}}$  and  $I_3(\mathbf{U})$  connects  $\mathbf{U}_{\frac{2}{3}}$  to  $\mathbf{U}_1$ . In the P-ordering the intersection points  $\mathbf{U}_{\frac{1}{3}}, \mathbf{U}_{\frac{2}{3}}$  and the sonic points  $\mathbf{U}_{S0}, \mathbf{U}_{S1}$  are obtained by using the physically correct Generalised Riemann Invariants; see Sect. 3.1.3 of Chap. 3 for details.

### Intersection Points and Sonic Points

Effectively, the intersection points  $\mathbf{U}_{\frac{1}{3}}, \mathbf{U}_{\frac{2}{3}}$  can be taken to be the solution of the Riemann problem with data  $\mathbf{U}_0, \mathbf{U}_1$  in the conventional sense. In the spirit of Osher's scheme we obtain  $\mathbf{U}_{\frac{1}{3}}$  and  $\mathbf{U}_{\frac{2}{3}}$  utilising the Generalised Riemann Invariants

$$I_L \equiv u + \frac{2a}{\gamma - 1} = \text{constant} \quad (12.63)$$

and

$$I_R \equiv u - \frac{2a}{\gamma - 1} = \text{constant} \quad (12.64)$$

to relate  $\mathbf{U}_0$  to  $\mathbf{U}_{\frac{1}{3}}$  and  $\mathbf{U}_{\frac{2}{3}}$  to  $\mathbf{U}_1$  respectively. Recall that if the left and right non-linear waves are rarefaction waves then relations (12.63)–(12.64) are exact. These waves can be either shock or rarefactions and thus the derived intersection points  $\mathbf{U}_{\frac{1}{3}}$  and  $\mathbf{U}_{\frac{2}{3}}$  are, in general, approximations. The underlying assumption is that in the solution of the Riemann problem with data  $\mathbf{U}_0, \mathbf{U}_1$ , both non-linear waves are rarefaction waves. This corresponds to the

Two-Rarefaction approximation TRRS presented in Sect. 9.4.1 of Chap. 9. Using (12.63) across the left wave gives

$$u_* + \frac{2a_{\frac{1}{3}}}{\gamma - 1} = u_0 + \frac{2a_0}{\gamma - 1} \quad (12.65)$$

and use of (12.64) across the right wave gives

$$u_* - \frac{2a_{\frac{2}{3}}}{\gamma - 1} = u_1 - \frac{2a_1}{\gamma - 1}. \quad (12.66)$$

Here  $u_*$  is the common particle velocity for  $\mathbf{U}_{\frac{1}{3}}$  and  $\mathbf{U}_{\frac{2}{3}}$ . Recall that, see Chaps. 3 and 4, the pressure  $p_*$  is also common, that is

$$u_{\frac{1}{3}} = u_{\frac{2}{3}} = u_* = \text{constant}, \quad p_{\frac{1}{3}} = p_{\frac{2}{3}} = p_* = \text{constant}. \quad (12.67)$$

In addition to (12.65)–(12.66) the isentropic law applied to the left and right waves gives

$$a_{\frac{1}{3}} = a_0(p_*/p_0)^z, \quad a_{\frac{2}{3}} = a_1(p_*/p_1)^z, \quad (12.68)$$

with

$$z = \frac{\gamma - 1}{2\gamma}. \quad (12.69)$$

From (12.65) and (12.68) we obtain

$$u_* = u_0 - \frac{2a_0}{\gamma - 1} \left[ \left( \frac{p_*}{p_0} \right)^z - 1 \right]. \quad (12.70)$$

Similarly, use of (12.66) and (12.68) gives

$$u_* = u_1 + \frac{2a_1}{\gamma - 1} \left[ \left( \frac{p_*}{p_1} \right)^z - 1 \right]. \quad (12.71)$$

Solving for  $p_*$  and  $u_*$  gives

$$p_* = \left[ \frac{a_0 + a_1 - (u_1 - u_0)(\gamma - 1)/2}{a_0/p_0^z + a_1/p_1^z} \right]^{\frac{1}{z}}, \quad (12.72)$$

$$u_* = \frac{Hu_0/a_0 + u_1/a_1 + 2(H - 1)/(\gamma - 1)}{H/a_0 + 1/a_1}, \quad (12.73)$$

with

$$H = (p_0/p_1)^z.$$

The density values  $\rho_{\frac{1}{3}}, \rho_{\frac{2}{3}}$  could be obtained from (12.68)–(12.69) or more directly as

$$\rho_{\frac{1}{3}} = \rho_0 \left( \frac{p_*}{p_0} \right)^{\frac{1}{\gamma}}, \quad \rho_{\frac{2}{3}} = \rho_1 \left( \frac{p_*}{p_1} \right)^{\frac{1}{\gamma}}. \quad (12.74)$$

The complete solution for  $\mathbf{U}_{\frac{1}{3}}, \mathbf{U}_{\frac{2}{3}}$  is given by (12.72)–(12.74). This is identical to the two-rarefaction approximation of Sect. 9.4.1 of Chap. 9. The computation of the sonic points  $\mathbf{U}_{S0}$  (left wave) and  $\mathbf{U}_{S1}$  (right wave) is performed by first enforcing the sonic conditions  $\lambda_1 = u - a = 0$  and  $\lambda_3 = u + a = 0$ , respectively and then applying the *corresponding* Generalised Riemann Invariants. The solution for the left sonic point is

$$\left. \begin{aligned} u_{S0} &= \frac{\gamma-1}{\gamma+1} u_0 + \frac{2a_0}{\gamma+1}, & a_{S0} &= u_{S0}, \\ \rho_{S0} &= \rho_0 \left( \frac{a_{S0}}{a_0} \right)^{\frac{2}{\gamma-1}}, & p_{S0} &= p_0 \left( \frac{\rho_{S0}}{\rho_0} \right)^\gamma. \end{aligned} \right\} \quad (12.75)$$

For the right sonic point the solution is

$$\left. \begin{aligned} u_{S1} &= \frac{\gamma-1}{\gamma+1} u_1 - \frac{2a_1}{\gamma+1}, & a_{S1} &= -u_{S1}, \\ \rho_{S1} &= \rho_1 \left( \frac{a_{S1}}{a_1} \right)^{\frac{2}{\gamma-1}}, & p_{S1} &= p_1 \left( \frac{\rho_{S1}}{\rho_1} \right)^\gamma. \end{aligned} \right\} \quad (12.76)$$

### Integration Along Partial Paths

We adopt expression (12.21) for the Osher intercell flux

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}_0 + \int_{\mathbf{U}_0}^{\mathbf{U}_1} \mathbf{A}^-(\mathbf{U}) d\mathbf{U},$$

where the integral in phase space along the path

$$I(\mathbf{U}) = I_1(\mathbf{U}) \cup I_2(\mathbf{U}) \cup I_3(\mathbf{U})$$

gives

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}_0 + \int_{\mathbf{U}_0}^{\mathbf{U}_{\frac{1}{3}}} \mathbf{A}^-(\mathbf{U}) d\mathbf{U} + \int_{\mathbf{U}_{\frac{1}{3}}}^{\mathbf{U}_{\frac{2}{3}}} \mathbf{A}^-(\mathbf{U}) d\mathbf{U} + \int_{\mathbf{U}_{\frac{2}{3}}}^{\mathbf{U}_1} \mathbf{A}^-(\mathbf{U}) d\mathbf{U}. \quad (12.77)$$

The integration along each partial path  $I_k(\mathbf{U})$  is performed individually and the results are added to produce  $\mathbf{F}_{i+\frac{1}{2}}$ . The partial integrations are easily performed following the methodology presented in previous sections. For any given path there is only one wave involved. The left and right waves define genuinely non-linear fields, see Sect. 3.1.3 of Chap. 3, and the corresponding eigenvalue changes sign at most once, generating the sonic-point values. The second field (middle wave) is linearly degenerate, see Sect. 3.1.3 of Chap. 3, and thus the eigenvalue  $\lambda_2(\mathbf{U})$  is constant along the path  $I_2(\mathbf{U})$ .

The integration results for each path are given in Table 12.7 and are labelled  $A$ ,  $B$  and  $C$  respectively. To obtain the intercell flux  $\mathbf{F}_{i+\frac{1}{2}}$  the integral terms in (12.77) must be selected according to the behaviour of the eigenvalue

$A :$	Behaviour of $\lambda_1(\mathbf{U}) = u - a$	$\int_{\mathbf{U}_0}^{\mathbf{U}_{\frac{1}{3}}} \mathbf{A}^-(\mathbf{U}) d\mathbf{U}$
1	$u_0 - a_0 \geq 0, \quad u_* - a_{\frac{1}{3}} \geq 0$	$\mathbf{0}$
2	$u_0 - a_0 \leq 0, \quad u_* - a_{\frac{1}{3}} \leq 0$	$\mathbf{F}_{\frac{1}{3}} - \mathbf{F}_0$
3	$u_0 - a_0 \geq 0, \quad u_* - a_{\frac{1}{3}} \leq 0$	$\mathbf{F}_{\frac{1}{3}} - \mathbf{F}_{S0}$
4	$u_0 - a_0 \leq 0, \quad u_* - a_{\frac{1}{3}} \geq 0$	$\mathbf{F}_{S0} - \mathbf{F}_0$
$B :$	Behaviour of $\lambda_2(\mathbf{U}) = u$	$\int_{\mathbf{U}_{\frac{1}{3}}}^{\mathbf{U}_{\frac{2}{3}}} \mathbf{A}^-(\mathbf{U}) d\mathbf{U}$
1	$u_* \geq 0$	$\mathbf{0}$
2	$u_* < 0$	$\mathbf{F}_{\frac{2}{3}} - \mathbf{F}_{\frac{1}{3}}$
$C :$	Behaviour of $\lambda_3(\mathbf{U}) = u + a$	$\int_{\mathbf{U}_{\frac{2}{3}}}^{\mathbf{U}_1} \mathbf{A}^-(\mathbf{U}) d\mathbf{U}$
1	$u_1 + a_1 \geq 0, \quad u_* + a_{\frac{2}{3}} \geq 0$	$\mathbf{0}$
2	$u_1 + a_1 \leq 0, \quad u_* + a_{\frac{2}{3}} \leq 0$	$\mathbf{F}_1 - \mathbf{F}_{\frac{2}{3}}$
3	$u_* + a_{\frac{2}{3}} \geq 0, \quad u_1 + a_1 \leq 0$	$\mathbf{F}_1 - \mathbf{F}_{S1}$
4	$u_* + a_{\frac{2}{3}} \leq 0, \quad u_1 + a_1 \geq 0$	$\mathbf{F}_{S1} - \mathbf{F}_{\frac{2}{3}}$

**Table 12.7.** Integration along partial paths for the Euler equations following the P-ordering,  $\mathbf{F}_k \equiv \mathbf{F}(\mathbf{U}_k)$

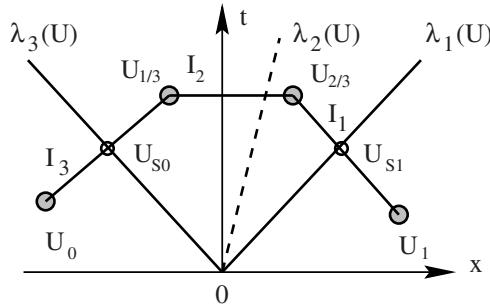
$\lambda_k(\mathbf{U})$  along the path  $I_k(\mathbf{U})$ . As seen in Table 12.7 there are 32 possible combinations  $i, j, l$ ; of these only 16 are realisable, which are tabulated as in Table 12.8. Here, in order to identify the correct expression for the resulting intercell flux  $\mathbf{F}_{i+\frac{1}{2}}$ , we split conditions on the intersection points, first column, from conditions on the data, top row.

	$u_0 - a_0 \geq 0$ $u_1 + a_1 \geq 0$	$u_0 - a_0 \geq 0$ $u_1 + a_1 \leq 0$	$u_0 - a_0 \leq 0$ $u_1 + a_1 \geq 0$	$u_0 - a_0 \leq 0$ $u_1 + a_1 \leq 0$
$u^* \geq 0$ $u^* - a_{\frac{1}{3}} \geq 0$	$\mathbf{F}_0$	$\mathbf{F}_0 + \mathbf{F}_1$ $- \mathbf{F}_{S1}$	$\mathbf{F}_{S0}$	$\mathbf{F}_{S0} - \mathbf{F}_{S1}$ $+ \mathbf{F}_1$
$u^* \geq 0$ $u^* - a_{\frac{1}{3}} \leq 0$	$\mathbf{F}_0 - \mathbf{F}_{S0}$ $+ \mathbf{F}_{\frac{1}{3}}$	$\mathbf{F}_0 - \mathbf{F}_{S0}$ $+ \mathbf{F}_{\frac{1}{3}} - \mathbf{F}_{S1}$ $+ \mathbf{F}_1$	$\mathbf{F}_{\frac{1}{3}}$	$\mathbf{F}_1 + \mathbf{F}_{\frac{1}{3}}$ $- \mathbf{F}_{S1}$
$u^* \leq 0$ $u^* + a_{\frac{2}{3}} \geq 0$	$\mathbf{F}_0 - \mathbf{F}_{S0}$ $+ \mathbf{F}_{\frac{2}{3}}$	$\mathbf{F}_0 - \mathbf{F}_{S0}$ $+ \mathbf{F}_{\frac{2}{3}} - \mathbf{F}_{S1}$ $+ \mathbf{F}_1$	$\mathbf{F}_{\frac{2}{3}}$	$\mathbf{F}_{\frac{2}{3}} - \mathbf{F}_{S1}$ $+ \mathbf{F}_1$
$u^* \leq 0$ $u^* + a_{\frac{2}{3}} \leq 0$	$\mathbf{F}_0 - \mathbf{F}_{S0}$ $+ \mathbf{F}_{S1}$	$\mathbf{F}_0 - \mathbf{F}_{S0}$ $+ \mathbf{F}_1$	$\mathbf{F}_{S1}$	$\mathbf{F}_1$

**Table 12.8.** Osher's flux formulae for the Euler equations using P-ordering of integration paths,  $\mathbf{F}_k \equiv \mathbf{F}(\mathbf{U}_k)$

### 12.3.2 Osher's Flux with O-Ordering

As originally presented, Osher's scheme uses the O-ordering of integration paths; this is precisely the opposite of the P-ordering described previously. The approach is used consistently to determine the intersection points, the sonic points and for performing the integration in (12.77). Fig. 12.4 illustrates the O-ordering of Osher's scheme as applied to a  $3 \times 3$  system, such as the one-dimensional Euler equations. We combine the configuration for the integration paths, the intersection points and the sonic points in phase space with the Riemann problem solution with data  $\mathbf{U}_0, \mathbf{U}_1$  in physical space  $x-t$ . The O-ordering can now be interpreted as assigning the eigenvalue  $\lambda_3(\mathbf{U})$  and



**Fig. 12.4.** Configuration of integration paths  $I_k(\mathbf{U})$ , intersection points  $\mathbf{U}_{\frac{1}{3}}$ ,  $\mathbf{U}_{\frac{2}{3}}$  and sonic points  $\mathbf{U}_{S0}$ ,  $\mathbf{U}_{S1}$  in physical space  $x-t$  for a  $3 \times 3$  system, following the O-ordering

eigenvector  $\mathbf{K}^{(3)}(\mathbf{U})$  to the wave family with eigenvalue  $\lambda_1(\mathbf{U})$  and eigenvector  $\mathbf{K}^{(1)}(\mathbf{U})$  (left wave), and vice-versa.

### Intersection Points and Sonic Points

First, the intersection points  $\mathbf{U}_{\frac{1}{3}}$  and  $\mathbf{U}_{\frac{2}{3}}$  are determined by using Generalised Riemann Invariants, as done in Sect. 12.1.3 with P-ordering. The difference is that  $\mathbf{U}_0$  and  $\mathbf{U}_{\frac{1}{3}}$  are connected using the *right* Riemann Invariant and  $\mathbf{U}_{\frac{2}{3}}$  is connected to  $\mathbf{U}_1$  using the *left* Riemann Invariant. Thus we write

$$I_R \equiv u_* - \frac{2}{\gamma - 1} a_{\frac{1}{3}} = u_0 - \frac{2}{\gamma - 1} a_0 , \quad (12.78)$$

$$I_L \equiv u_* + \frac{2}{\gamma - 1} a_{\frac{2}{3}} = u_1 + \frac{2}{\gamma - 1} a_1 . \quad (12.79)$$

See (12.63)–(12.66). Again, we assume  $u_*$  and  $p_*$  are the common particle velocity and pressure at points  $\mathbf{U}_{\frac{1}{3}}$  and  $\mathbf{U}_{\frac{2}{3}}$ . Use of (12.68) into (12.78) gives

$$u_* = u_0 + \frac{2}{(\gamma - 1)} a_0 \left[ \left( \frac{p_*}{p_0} \right)^z - 1 \right] , \quad (12.80)$$

and use of (12.68) into (12.79) produces

$$u_* = u_1 - \frac{2}{(\gamma - 1)} a_1 \left[ \left( \frac{p_*}{p_1} \right)^z - 1 \right] . \quad (12.81)$$

Solving for  $p_*$  gives

$$p_* = \left[ \frac{a_0 + a_1 + (u_1 - u_0)(\gamma - 1)/2}{a_0/p_0^z + a_1/p_1^z} \right]^{\frac{1}{z}} . \quad (12.82)$$

Compare with solution (12.72) using P-ordering. Equations (12.80) and (12.81) can be rearranged as

$$p_* = p_0 \left[ \frac{\gamma - 1}{2a_0} (u_* - u_0) + 1 \right]^{\frac{1}{z}}, \quad (12.83)$$

$$p_* = p_1 \left[ \frac{\gamma - 1}{2a_1} (u_1 - u_*) + 1 \right]^{\frac{1}{z}}, \quad (12.84)$$

whose solution for  $u_*$  is

$$u_* = \frac{Hu_0/a_0 + u_1/a_1 - 2(H-1)/(\gamma-1)}{H/a_0 + 1/a_1}, \quad (12.85)$$

with

$$H = (p_0/p_1)^z, \quad z = \frac{\gamma - 1}{2\gamma}.$$

Compare with solution (12.73) using the P-ordering. The solution for  $\rho_{\frac{1}{3}}$  and  $\rho_{\frac{2}{3}}$  is, using the isentropic law,

$$\rho_{\frac{1}{3}} = \rho_0(p_*/p_0)^{\frac{1}{\gamma}}, \quad \rho_{\frac{2}{3}} = \rho_1(p_*/p_1)^{\frac{1}{\gamma}}. \quad (12.86)$$

To find the sonic points  $\mathbf{U}_{S0}$  and  $\mathbf{U}_{S1}$  we first connect  $\mathbf{U}_0$  to  $\mathbf{U}_{\frac{1}{3}}$  via the *right* Riemann Invariant to obtain

$$u_{S0} = \frac{2}{\gamma-1}a_{S0} + u_0 - \frac{2}{\gamma-1}a_0.$$

Then by enforcing the sonic condition

$$\lambda_3(\mathbf{U}) = u_{S0} + a_{S0} = 0$$

along  $I_3(\mathbf{U})$  and applying the isentropic law one obtains the solution

$$\left. \begin{aligned} u_{S0} &= \frac{\gamma-1}{\gamma+1}u_0 - \frac{2a_0}{\gamma+1}, & a_{S0} &= -u_{S0}, \\ \rho_{S0} &= \rho_0 \left( \frac{a_{S0}}{a_0} \right)^{\frac{2}{\gamma-1}}, & p_{S0} &= p_0 \left( \frac{\rho_{S0}}{\rho_0} \right)^\gamma. \end{aligned} \right\} \quad (12.87)$$

The solution for the right sonic point  $\mathbf{U}_{S1}$  is

$$\left. \begin{aligned} u_{S1} &= \frac{\gamma-1}{\gamma+1}u_1 + \frac{2}{\gamma+1}a_1, & a_{S1} &= u_{S1}, \\ \rho_{S1} &= \rho_1 \left( \frac{a_{S1}}{a_1} \right)^{\frac{2}{\gamma-1}}, & p_{S1} &= p_1 \left( \frac{\rho_{S1}}{\rho_1} \right)^\gamma. \end{aligned} \right\} \quad (12.88)$$

### Integration Along Partial Paths

To compute Osher's intercell flux

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}_0 + \int_{\mathbf{U}_0}^{\mathbf{U}_1} \mathbf{A}^-(\mathbf{U}) d\mathbf{U}, \quad (12.89)$$

we need to find the integral

$$\int_{\mathbf{U}_0}^{\mathbf{U}_1} \mathbf{A}^-(\mathbf{U}) d\mathbf{U} = \int_{\mathbf{U}_0}^{\mathbf{U}_{\frac{1}{3}}} \mathbf{A}^-(\mathbf{U}) d\mathbf{U} + \int_{\mathbf{U}_{\frac{1}{3}}}^{\mathbf{U}_{\frac{2}{3}}} \mathbf{A}^-(\mathbf{U}) d\mathbf{U} + \int_{\mathbf{U}_{\frac{2}{3}}}^{\mathbf{U}_1} \mathbf{A}^-(\mathbf{U}) d\mathbf{U}. \quad (12.90)$$

Using the O-ordering one has

$$\int_{\mathbf{U}_0}^{\mathbf{U}_{\frac{1}{3}}} \mathbf{A}^-(\mathbf{U}) d\mathbf{U} = \int_{I_3(\mathbf{U})} \mathbf{A}^-(\mathbf{U}) d\mathbf{U}, \quad (12.91)$$

$$\int_{\mathbf{U}_{\frac{1}{3}}}^{\mathbf{U}_{\frac{2}{3}}} \mathbf{A}^-(\mathbf{U}) d\mathbf{U} = \int_{I_2(\mathbf{U})} \mathbf{A}^-(\mathbf{U}) d\mathbf{U}, \quad (12.92)$$

$$\int_{\mathbf{U}_{\frac{2}{3}}}^{\mathbf{U}_1} \mathbf{A}^-(\mathbf{U}) d\mathbf{U} = \int_{I_1(\mathbf{U})} \mathbf{A}^-(\mathbf{U}) d\mathbf{U}. \quad (12.93)$$

The evaluation of the three terms in (12.90) according to (12.91)–(12.93) is given in Table 12.9, by cases labelled A, B and C respectively. The first column contains the sub-case numbers. The second column contains the behaviour of the eigenvalue along the corresponding partial path. For instance in case A one considers the sign of the eigenvalue  $\lambda_3(\mathbf{U})$  along the path  $I_3(\mathbf{U})$  joining the points  $\mathbf{U}_0$  and  $\mathbf{U}_{\frac{1}{3}}$ . The third column contains the resulting integral in (12.91)–(12.93). There are 32 combinations, of which only 16 are realisable.

Just to illustrate the method of analysis, consider first the combination  $(A_1, B_1, C_1)$ , that is

$$A_1 : u_0 + a_0 \geq 0, u_* + a_{\frac{1}{3}} \geq 0,$$

$$B_1 : \quad u_* \geq 0,$$

$$C_1 : u_1 - a_1 \geq 0, u_* - a_{\frac{2}{3}} \geq 0.$$

As all eigenvalues are positive throughout, the integrals are all zero. The combination  $(A_1, B_2, C_1)$  gives

$$A_1 : u_0 + a_0 \geq 0, u_* + a_{\frac{1}{3}} \geq 0,$$

$$B_2 : \quad u_* \leq 0,$$

$$C_1 : u_1 - a_1 \geq 0, u_* - a_{\frac{2}{3}} \geq 0.$$

Clearly these conditions cannot be satisfied simultaneously,  $u_* \leq 0$  contradicts  $u_* - a_{\frac{2}{3}} \geq 0$ , as  $a_{\frac{2}{3}} > 0$ .

The Osher intercell flux with O-ordering is given in table 12.10. The logic involved can be organised so as to test conditions on the data points (top row) and conditions on the intersection points (first column). The Osher's flux

$A :$	Behaviour of $\lambda_3(\mathbf{U}) = u - a$	$\int_{\mathbf{U}_0}^{\mathbf{U}_{\frac{1}{3}}} \mathbf{A}^-(\mathbf{U}) d\mathbf{U}$
1	$u_0 + a_0 \geq 0, \quad u_* + a_{\frac{1}{3}} \geq 0$	$\mathbf{0}$
2	$u_0 + a_0 \leq 0, \quad u_* + a_{\frac{1}{3}} \leq 0$	$\mathbf{F}_{\frac{1}{3}} - \mathbf{F}_0$
3	$u_0 + a_0 \geq 0, \quad u_* + a_{\frac{1}{3}} \leq 0$	$\mathbf{F}_{\frac{1}{3}} - \mathbf{F}_{S0}$
4	$u_0 + a_0 \leq 0, \quad u_* + a_{\frac{1}{3}} \geq 0$	$\mathbf{F}_{S0} - \mathbf{F}_0$
$B :$	Behaviour of $\lambda_2(\mathbf{U}) = u$	
1	$u_* \geq 0$	$\mathbf{0}$
2	$u_* \leq 0$	$\mathbf{F}_{\frac{2}{3}} - \mathbf{F}_{\frac{1}{3}}$
$C :$	Behaviour of $\lambda_1(\mathbf{U}) = u - a$	
1	$u_1 - a_1 \geq 0, \quad u_* - a_{\frac{2}{3}} \geq 0$	$\mathbf{0}$
2	$u_1 - a_1 \leq 0, \quad u_* - a_{\frac{2}{3}} \leq 0$	$\mathbf{F}_1 - \mathbf{F}_{\frac{2}{3}}$
3	$u_1 - a_1 \geq 0, \quad u_* - a_{\frac{2}{3}} \leq 0$	$\mathbf{F}_{S1} - \mathbf{F}_{\frac{2}{3}}$
4	$u_1 - a_1 \leq 0, \quad u_* - a_{\frac{2}{3}} \geq 0$	$\mathbf{F}_1 - \mathbf{F}_{S1}$

**Table 12.9.** Osher's flux for the Euler equations. Evaluation of integral following O-ordering,  $\mathbf{F}_k \equiv \mathbf{F}(\mathbf{U}_k)$

formulae with P and O ordering given by Tables 12.8 and 12.10 respectively, are valid for any  $3 \times 3$  non-linear hyperbolic system. The specific properties of a particular system enter in the determination of the sonic and intersection points only. The formulae can be directly applied, for instance, to the split two-dimensional shallow water equations.

	$u_0 + a_0 \geq 0$ $u_1 - a_1 \geq 0$	$u_0 + a_0 \geq 0$ $u_1 - a_1 \leq 0$	$u_0 + a_0 \leq 0$ $u_1 - a_1 \geq 0$	$u_0 + a_0 \leq 0$ $u_1 - a_1 \leq 0$
$u_* + a_{\frac{1}{3}} \leq 0$	$\mathbf{F}_0 - \mathbf{F}_{S0} + \mathbf{F}_{S1}$	$\mathbf{F}_0 - \mathbf{F}_{S0} + \mathbf{F}_1$	$\mathbf{F}_{S1}$	$\mathbf{F}_1$
$u_* + a_{\frac{1}{3}} \geq 0$ $u_* \leq 0$	$\mathbf{F}_0 - \mathbf{F}_{\frac{1}{3}} + \mathbf{F}_{S1}$	$\mathbf{F}_0 - \mathbf{F}_{\frac{1}{3}} + \mathbf{F}_1$	$\mathbf{F}_{S0} - \mathbf{F}_{\frac{1}{3}} + \mathbf{F}_{S1}$	$\mathbf{F}_{S0} - \mathbf{F}_{\frac{1}{3}} + \mathbf{F}_1$
$u_* - a_{\frac{2}{3}} \geq 0$	$\mathbf{F}_0$	$\mathbf{F}_0 - \mathbf{F}_{S1} + \mathbf{F}_1$	$\mathbf{F}_{S0}$	$\mathbf{F}_{S0} - \mathbf{F}_{S1} + \mathbf{F}_1$
$u_* - a_{\frac{2}{3}} \leq 0$ $u_* \geq 0$	$\mathbf{F}_0 - \mathbf{F}_{\frac{2}{3}} + \mathbf{F}_{S1}$	$\mathbf{F}_0 - \mathbf{F}_{\frac{2}{3}} + \mathbf{F}_1$	$\mathbf{F}_{S0} - \mathbf{F}_{\frac{2}{3}} + \mathbf{F}_{S1}$	$\mathbf{F}_{S0} - \mathbf{F}_{\frac{2}{3}} + \mathbf{F}_1$

**Table 12.10.** Osher's flux formulae for the Euler equations using O-ordering of integration paths,  $\mathbf{F}_k = \mathbf{F}(\mathbf{U}_k)$

### 12.3.3 Remarks on Path Orderings

It is useful to compare the pressure solutions (12.72) and (12.82) when computing the intersection points using the P and O orderings for the integration paths. Let us redefine the respective solutions as

$$p_*^{(P)} = \left[ \frac{a_0 + a_1 - \Delta u(\gamma - 1)/2}{a_0/p_0^z + a_1/p_1^z} \right]^{1/z} \quad (12.94)$$

and

$$p_*^{(O)} = \left[ \frac{a_0 + a_1 + \Delta u(\gamma - 1)/2}{a_0/p_0^z + a_1/p_1^z} \right]^{1/z}, \quad (12.95)$$

where  $\Delta u = u_1 - u_0$  is the velocity difference in the data. The reader is encouraged to review Sect. 4.3.1 of Chap. 4, in which a detailed discussion is

given on the influence of  $\Delta u$  on the solution for pressure in the exact solution. First we note that  $p_*^{(P)} = p_*^{(O)}$  when  $\Delta u = 0$  and that for  $\Delta u \neq 0$  the two solutions are not only different but more importantly, they have *very different behaviour*. This does not seem to matter too much in practical computations, except for special but important situations. There are at least two cases that deserve special attention.

The case  $\Delta u \gg 0$ , see Sect. 4.3.1 of Chap. 4, is associated with strong rarefaction waves. In fact there is a limit for which the pressure  $p_*$  becomes negative and is associate with the *pressure positivity condition* stated in Sect. 4.3.1 of Chap. 4. In the *incipient cavitation case* the pressure is 0. The solution  $p_*^{(P)}$  will correctly reflect this physical situation of low-density flow, including the detection of *vacuum*. The author is not aware of this having been exploited in the context of Osher's Riemann solver with P-ordering. On the other hand, the solution  $p_*^{(O)}$  for  $\Delta u \gg 0$  obtained with the O-ordering, will give unrealistically large values for the pressure at the intersection points, which is more consistent with the presence of strong shock waves, rather than strong rarefaction waves. Such large pressure values will lead the Osher scheme to be very inaccurate or simply to fail for low-density flows, just as linearised Riemann solvers do; see Einfeldt et. al. [182] for a discussion on numerical difficulties for low-density flows. We illustrate this point through Test 2 of Table 12.11, for which the scheme actually fails.

The case  $\Delta u \ll 0$ , see Sect. 4.3.1 of Chap. 4, is associated with strong shock waves. Again the P-ordering solution (12.94) will correctly reflect this. However the O-ordering solution (12.95) will not. More importantly, there will be a limiting strong shock situation for which the O-ordering pressure is *undefined* and the scheme will again fail; see Test 5 in Table 12.11. The failure condition in this case is analytic, namely

$$a_0 + a_1 + \Delta u(\gamma - 1)/2 \leq 0. \quad (12.96)$$

It is paradoxical that the scheme fails in the presence of strong shocks, which is consistent with large pressure, through a pressure solution  $p_*^{(O)}$  that is undefined for being so close to vacuum conditions. One could devise some sort of fix to remedy this situation. One possibility is to abandon the O-ordering altogether or switch to the P-ordering locally, in a kind of adaptive ordering.

#### 12.3.4 The Split Three-Dimensional Case

The extension of the Osher scheme to two and three-dimensional problems is straightforward. All methods considered here require expressions for the split fluxes. In the  $x$ -split, three dimensional case, for instance, we require in addition the  $y$  and  $z$  momentum flux components  $\rho uv$  and  $\rho uw$ . In turn, this requires the extra components  $\rho v$  and  $\rho w$  in the vector of conserved variables  $\mathbf{U}$ . These new components will be needed in the intersection points and sonic

points. But, as seen in Sect. 4.8 of Chap. 4, the exact solution for  $v$  and  $w$  in the Riemann problem is given by

$$q(x, t) = \begin{cases} q_L & \text{if } x/t < u_* \\ q_R & \text{if } x/t > u_* \end{cases}, \quad (12.97)$$

where  $q = v$  and  $q = w$ . The approximate solution for the tangential components of velocity obtained from using Generalised Riemann Invariants in the Osher scheme preserves the form of the exact solution (12.97), the only approximation being that of the normal velocity component  $u_*$ . Therefore, for computing  $\mathbf{U}_{S0}$  and  $\mathbf{U}_{\frac{1}{3}}$  we take  $v_L$  and  $w_L$ ; for computing  $\mathbf{U}_{S1}$  and  $\mathbf{U}_{\frac{2}{3}}$  we take  $v_R$  and  $w_R$ . Solution (12.97) ensures that contact waves and shear waves (and shear layers in Navier–Stokes applications) are well resolved by Osher’s scheme, a property that is common to the exact Riemann solver of Chap. 6, the approximate-state Riemann solvers of Chap. 9, the HLLC Riemann solver of Chap. 10 and Roe’s solver of Chap. 11.

## 12.4 Numerical Results and Discussion

Here we illustrate the performance of Godunov’s first-order method used in conjunction with Osher-type approximate Riemann solvers. We select six test problems for the one-dimensional time dependent Euler equations for ideal gases with  $\gamma = 1.4$ ; these have exact solutions. In all chosen tests, data consists of two constant states  $\mathbf{W}_L = [\rho_L, u_L, p_L]^T$  and  $\mathbf{W}_R = [\rho_R, u_R, p_R]^T$ , separated by a discontinuity at a position  $x = x_0$ . The states  $\mathbf{W}_L$  and  $\mathbf{W}_R$  are given in Table 12.11. The exact and numerical solutions are found in the spatial domain  $0 \leq x \leq 1$ . The numerical solution is computed with  $M = 100$  cells and the CFL condition is as for all previous computations, see Chap. 6; the chosen Courant number coefficient is  $C_{\text{cfl}} = 0.9$ ; boundary conditions are transmissive. The exact solutions were found by running the code *HE-E1RPEXACT* of the library *NUMERICA* [519] and the numerical solutions were obtained by running the code *HE-E1GODOSHER* of *NUMERICA*.

Test	$\rho_L$	$u_L$	$p_L$	$\rho_R$	$u_R$	$p_R$
1	1.0	0.75	1.0	0.125	0.0	0.1
2	1.0	-2.0	0.4	1.0	2.0	0.4
3	1.0	0.0	1000.0	1.0	0.0	0.01
4	5.99924	19.5975	460.894	5.99242	-6.19633	46.0950
5	1.0	-19.59745	1000.0	1.0	-19.59745	0.01
6	1.0	2.0	0.1	1.0	-2.0	0.1

Table 12.11. Data for six test problems with exact solution, for the time-dependent one dimensional Euler equations

Test 1 is a *modified* version of Sod's problem [453]; the solution has a right shock wave, a right travelling contact wave and a left *sonic* rarefaction wave; this test is useful in assessing the *entropy satisfaction* property of numerical methods. The solution of Test 2 consists of two symmetric rarefaction waves and a trivial contact wave; the *Star Region* between the non-linear waves is close to vacuum, which makes this problem a suitable test for assessing the performance of numerical methods for low-density flows. Test 3 is designed to assess the robustness and accuracy of numerical methods; its solution consists of a strong right shock wave of shock Mach number 198, a contact surface and a left rarefaction wave. Test 4 is also a very severe test, its solution consists of three strong discontinuities travelling to the right. A detailed discussion on the exact solution of the test problems is found in Chap. 4. Test 5 is also designed to test the robustness of numerical methods but the main reason for devising this test is to assess the ability of numerical methods to resolve *slowly-moving contact discontinuities*. The exact solution of Test 5 consists of a left rarefaction wave, a right-travelling shock wave and a *stationary* contact discontinuity. Test 6 simulates the collision of two uniform streams; the exact solution consists of two strong symmetric shock waves and a trivial contact discontinuity. The purpose of Test 6 is to illustrate the fact that Osher's scheme with O-ordering will fail for a range of problems of this kind. For each test problem we select a convenient position  $x_0$  of the initial discontinuity and the output time. These are stated in the legend of each figure displaying computational results.

The computed results for Tests 1, 3, 4 and 5 using the Godunov first-order method in conjunction with the Osher approximate Riemann solver with O-ordering are shown in Figs. 12.5–12.8. The scheme failed for Tests 2 and 6. The result of Fig. 12.5 is virtually identical to that of the Godunov method with an exact Riemann solver, Fig. 6.8 of Chap. 6; note also that both schemes give comparable performance in the vicinity of the sonic point. The results for Tests 3 and 4, shown in Figs. 12.6 and 12.7, obtained with O-ordering compare well with those of the exact Riemann solver, except for the fact that the slowly moving shock in Fig. 12.7 does not have the spurious oscillations that other Riemann solvers produce [406], [60], [19], [280]; this is a distinguishing property of the Osher approach.

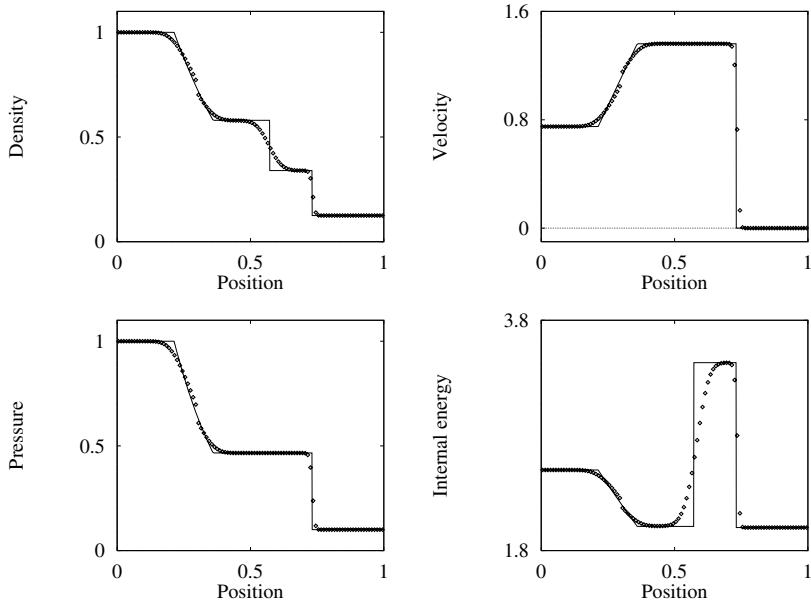
Results for Tests 1 to 6 from the Osher scheme with P-ordering are shown in Figs. 12.9–12.14. The results for Tests 1 to 4 are, overall, as accurate as those obtained from an exact Riemann solver; the P-ordering scheme actually works for Tests 2 and 6, whereas O-ordering scheme does not; note however the spurious oscillations in Test 6 near the shocks, a feature of P-ordering schemes already noted by Osher. For Test 5 the P-ordering scheme gives the incorrect solution; compare Fig. 12.13 with Fig. 12.8 and note the scales. First, there are very large unphysical overshoots in density, velocity and pressure. More intriguing is the fact that the expected right-travelling shock does not propagate at all; no signal propagates to the right of the initial discontinuity at  $x = 0.8$ . As a matter of fact, it looks as if the shock wave does not actually

form and the state behind the contact discontinuity has the incorrect values, see pressure and velocity plots. My numerical results have been confirmed by Dr N. Qin (private communication), who used an independently written code to solve Test 5.

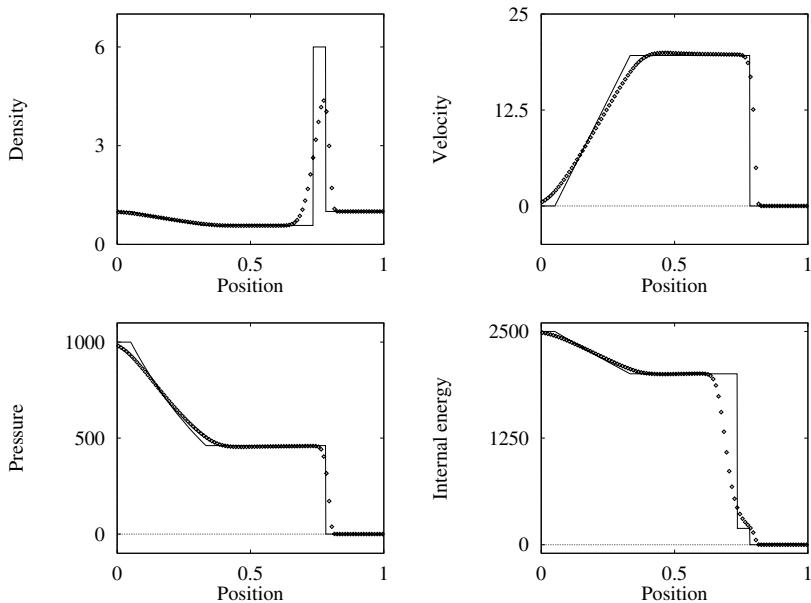
Earlier we made some remarks concerning a modified Osher scheme based on an adaptive ordering. In view of the numerical results it seems as if O-ordering could be used in all cases except for the two situations in which the solution of the Riemann problem contains either two shocks or two rarefactions. Such situations could be identified reliably and cheaply by using any of the approximate state Riemann solvers of Chap. 9.

## 12.5 Extensions

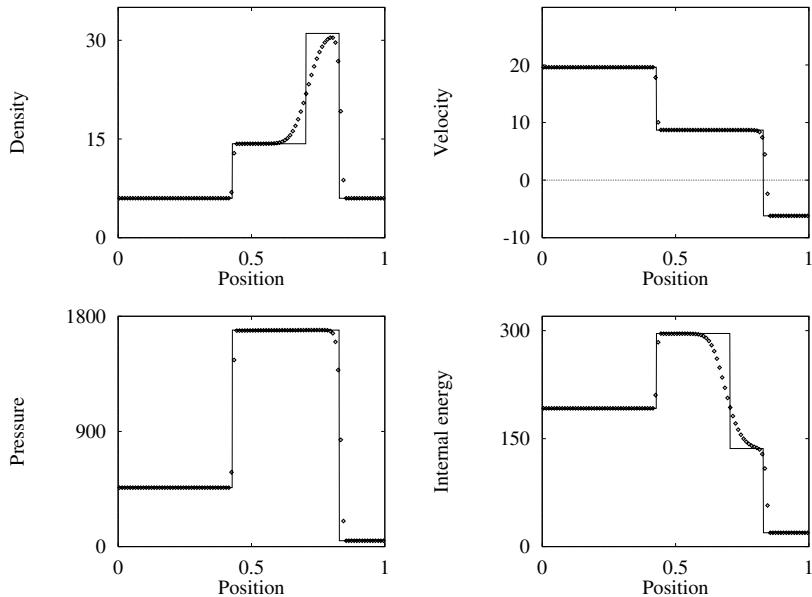
The approximate Riemann solvers of this chapter may be applied in conjunction with the Godunov first-order upwind method presented in Chap. 6. Second-order Total Variation Diminishing (TVD) extensions of the schemes are presented in Chap. 13 for scalar problems and in Chap. 14 for non-linear one dimensional systems. In chap. 15 we present techniques that allow the extension of these schemes to solve problems with source terms. In Chap. 16 we study techniques to extend the methods of this chapter to three-dimensional problems.



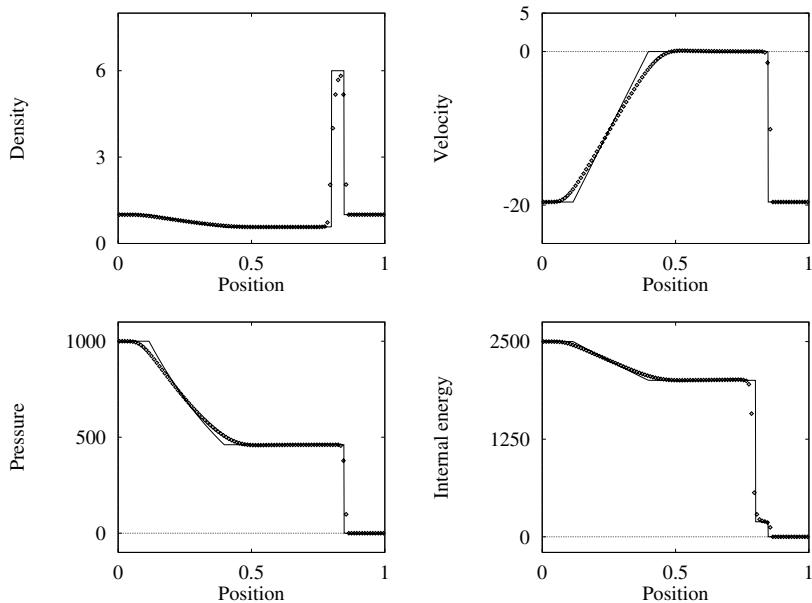
**Fig. 12.5.** Osher scheme with O-ordering applied to Test 1, with  $x_0 = 0.3$ . Numerical (symbol) and exact (line) solutions are compared at time 0.2 units



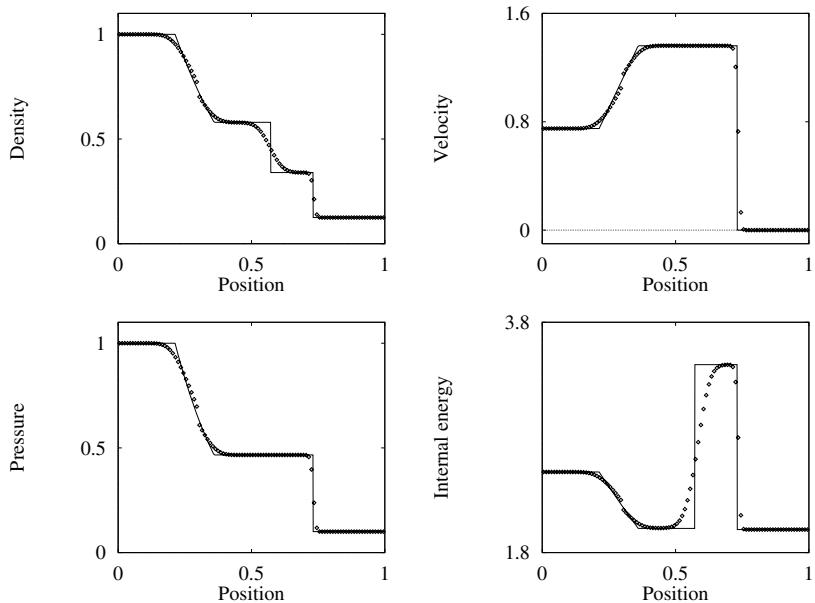
**Fig. 12.6.** Osher scheme with O-ordering applied to Test 3, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012 units



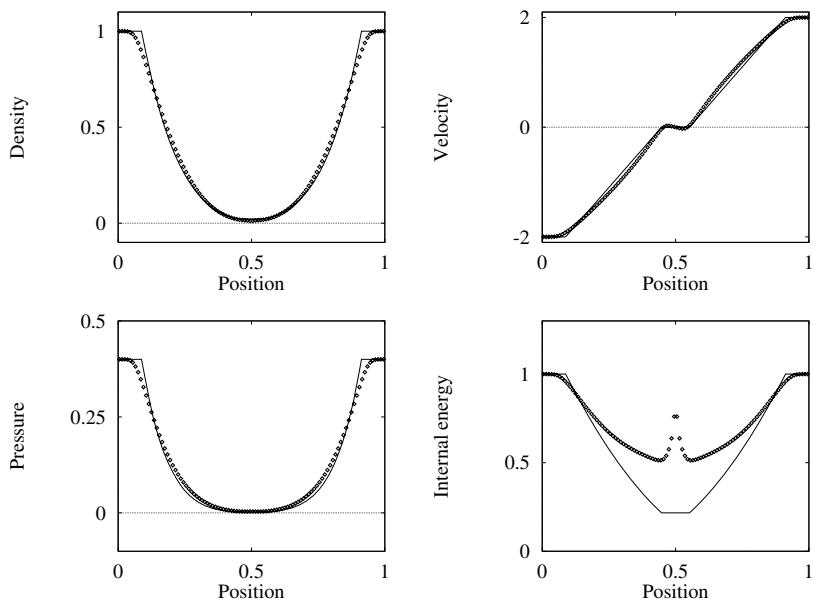
**Fig. 12.7.** Osher scheme with O-ordering applied to Test 4, with  $x_0 = 0.4$ . Numerical (symbol) and exact (line) solutions are compared at time 0.035 units



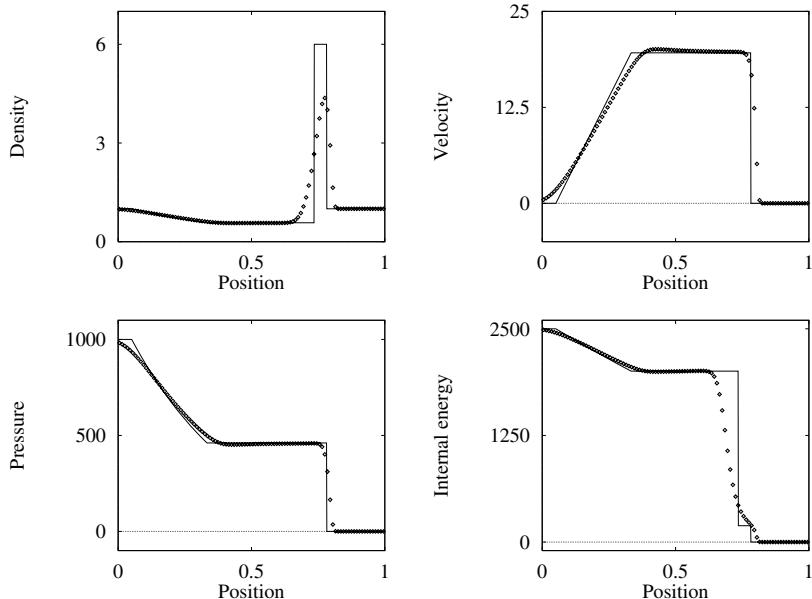
**Fig. 12.8.** Osher scheme with O-ordering applied to Test 5, with  $x_0 = 0.8$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012 units



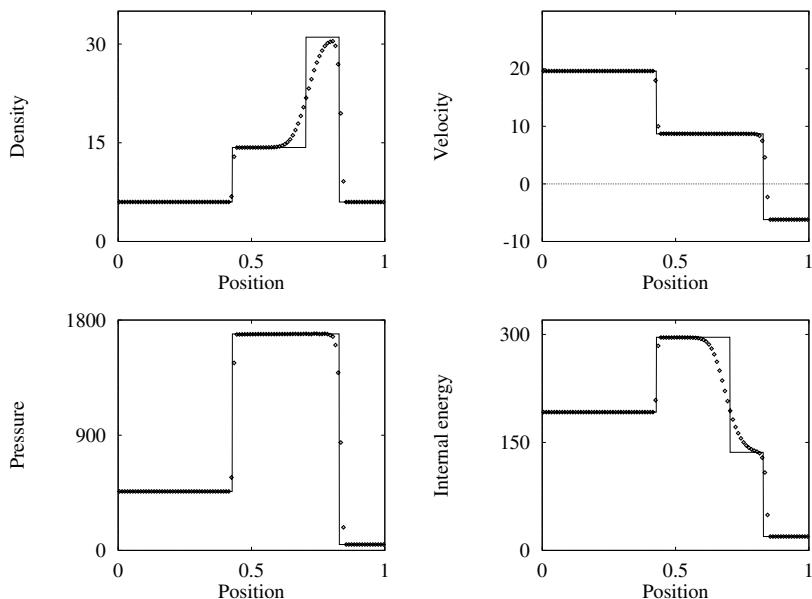
**Fig. 12.9.** Osher scheme with P-ordering applied to Test 1, with  $x_0 = 0.3$ . Numerical (symbol) and exact (line) solutions are compared at time 0.2 units



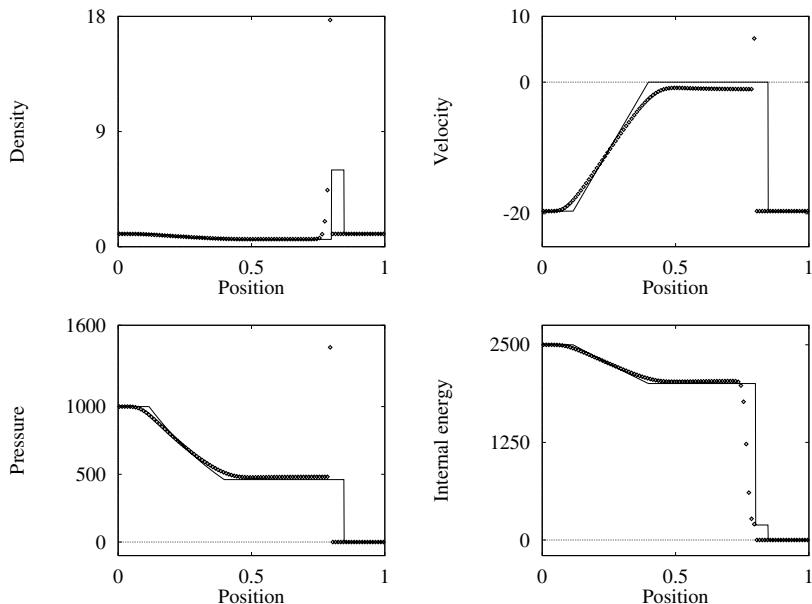
**Fig. 12.10.** Osher scheme with P-ordering applied to Test 2, with  $x_0 = 0.3$ . Numerical (symbol) and exact (line) solutions are compared at time 0.2 units



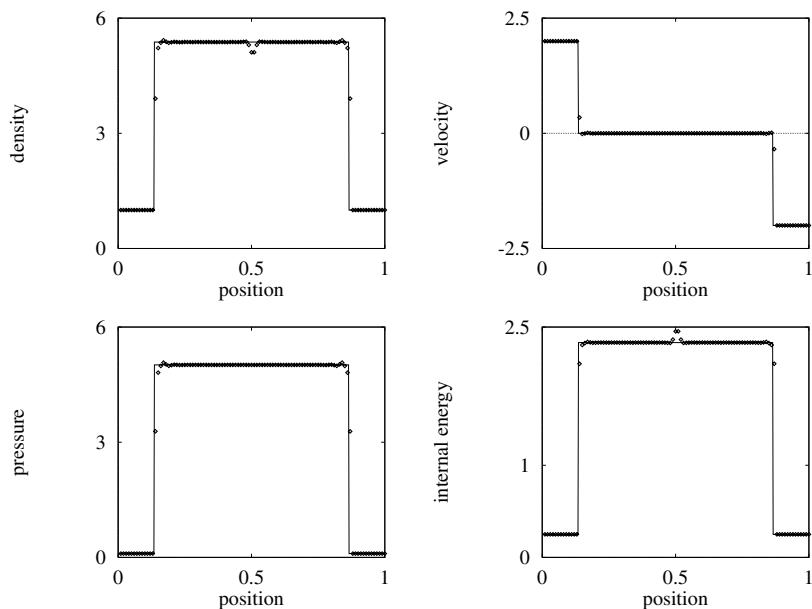
**Fig. 12.11.** Osher scheme with P-ordering applied to Test 3, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012 units



**Fig. 12.12.** Osher scheme with P-ordering applied to Test 4, with  $x_0 = 0.4$ . Numerical (symbol) and exact (line) solutions are compared at time 0.035 units



**Fig. 12.13.** Osher scheme with P-ordering applied to Test 5, with  $x_0 = 0.8$ . Numerical (symbol) and exact (line) solutions are compared at time 0.012 units



**Fig. 12.14.** Osher scheme with P-ordering applied to Test 6, with  $x_0 = 0.5$ . Numerical (symbol) and exact (line) solutions are compared at time 0.8 units

---

## High-Order and TVD Methods for Scalar Equations

Central to this chapter is the resolution of two contradictory requirements on numerical methods, namely high-order of accuracy and absence of *spurious* (unphysical) oscillations in the vicinity of large gradients. It is well-known that high-order linear (constant coefficients) schemes produce unphysical oscillations in the vicinity of large gradients. This was illustrated by some numerical results shown in Chap. 5. On the other hand, the class of *monotone methods*, defined in Sect. 5.2 of Chap. 5, do not produce unphysical oscillations. However, *monotone methods are at most first order accurate* and are therefore of limited use. These difficulties are embodied in the statement of Godunov's theorem [216] to be studied in Sect. 13.5.3. One way of resolving the contradiction between linear schemes of high-order of accuracy and absence of spurious oscillations is by constructing non-linear methods. *Total Variation Diminishing Methods*, or TVD Methods for short, are a prominent class of non-linear methods, which we shall study in detail in this chapter.

### 13.1 Introduction

The purpose of this chapter is to introduce and develop the concept of TVD methods in a simple setting, so as to aid the reader to extend the techniques to more general problems. TVD methods are one of the most significant achievements in the development of numerical methods for partial differential equations (PDEs) in the last 20 years or so, although the preliminary ideas can be traced as far back as 1959, with the pioneering work of Godunov and later by van Leer [554], the Flux Corrected Transport (FCT) approach of Boris and Book [70], [69], [71] and many others. The theoretical bases of TVD methods are sound for scalar one-dimensional problems only. In practical, non-linear, multidimensional problems, the accumulated experience of numerous applications has demonstrated that the one-dimensional scalar theory serves well as a guideline for extending the ideas, on a more or less empirical basis. TVD schemes have their roots on the fundamental question of *convergence* of

schemes for *non-linear* problems. Total-Variation Stable schemes have been proved to be convergent. TVD methods are a class of Total-Variation Stable schemes and they are based on the requirement that the total variation of the numerical solution be non-increasing in time. As a matter of fact the *exact* solution of a scalar conservation law does possess this property and therefore TVD numerical methods just mimic a property of the exact solution. Apart from being a class of methods proved to be convergent, for scalar problems, the TVD concept is also extremely useful in *designing schemes*, as we shall see in this chapter.

TVD schemes are intimately linked to traditional *Artificial Viscosity Methods* [402]. Both TVD and Artificial Viscosity Methods attempt to circumvent Godunov's theorem by constructing schemes of accuracy  $p > 1$ , such that *spurious oscillations near high gradients* are eliminated or controlled. Both classes of schemes resort effectively to the same mechanism, namely the addition of *artificial viscosity*. Artificial Viscosity Methods do this *explicitly*, by adding extra terms to the PDEs. In TVD methods, on the other hand, the artificial viscosity is inherent in the scheme itself, but the way this is activated is rather sophisticated.

The TVD approach has managed to circumvent Godunov's theorem by resorting to *non-linear schemes*, even when applied to linear problems. The inherent artificial viscosity of the schemes is controlled in a very precise manner so that spurious oscillations are eliminated and high-order accuracy in smooth parts of the solution is retained. Historically, TVD methods have almost exclusively been associated with high-order *upwind* methods. Over the last few years, however, the TVD concept has been extended to *centred* schemes, as distinct from *upwind* schemes. This development has effectively provided Artificial Viscosity Methods with a more rational basis and has tended to unify the developments on numerical methods for conservation laws.

Relevant contributions to the general area of TVD and related methods can be found in the papers by Godunov [216]; van Leer [554]; Boris and Book [70]; Book, Boris and Hain [69]; Boris and Book [71]; Majda and Osher [335]; Zalesak [597]; Crandall and Majda [145]; Roe [408], [409], [410]; van Albada, van Leer and Roberts [552]; Harten [238], [239]; Sweby [470], [471]; Baines and Sweby [23]; Davis [149]; Osher [369]; Osher and Chakravarthy [371]; Jameson and Lax [267]; Yee, Warming and Harten [595]; Shu [446]; Yee [592], [593]; Tadmor [473]; Leonard [307]; Gaskell and Lau [198]; Swanson and Turkel [467]; Jorgenson and Turkel [274]; Cockburn and Shu [124]; Liu and Lax [330]; Billett and Toro [65]; Toro and Billett [538]; and many others.

In this chapter, after a brief review of some basic properties of schemes in Sect. 13.2, we present several approaches for constructing high-order methods. The emphasis is on approaches that extend to non-linear systems and can be suitably modified according to some TVD criteria; see Sects. 13.3 and 13.4. General theoretical aspects relevant to TVD methods are discussed in Sect. 13.5. In Sect. 13.6 we introduce TVD methods. We then present two approaches for constructing TVD methods, namely the *Flux Limiter Approach*

of Sect. 13.7 and the *Slope Limiter Approach* of Sect. 13.8. For each of these two classes of TVD methods we present *upwind*-based schemes and *centred* schemes. Numerical results on model problems are included.

The reader is encouraged to review Chaps. 2 and 5 before proceeding with this chapter. Application of the TVD concepts to non-linear systems is found in Chaps. 14 and 16.

## 13.2 Basic Properties of Selected Schemes

The purpose of this section is to review a selected number of schemes and discuss some of their main properties. In the main, the discussion is centred on the linear advection equation

$$u_t + f(u)_x = 0 , \quad f(u) = au , \quad (13.1)$$

where  $u = u(x, t)$  and  $a$  is a constant wave propagation speed.

### 13.2.1 Selected Schemes

In Chap. 5 we introduced a number of well known schemes to solve (13.1) numerically. Prominent examples are the Godunov first-order upwind method (the CIR scheme for linear systems)

$$\left. \begin{aligned} u_i^{n+1} &= u_i^n - c(u_i^n - u_{i-1}^n) , & a > 0 , \\ u_i^{n+1} &= u_i^n - c(u_{i+1}^n - u_i^n) , & a < 0 , \end{aligned} \right\} \quad (13.2)$$

and the Lax–Wendroff method

$$u_i^{n+1} = \frac{1}{2}c(1+c)u_{i-1}^n + (1-c^2)u_i^n - \frac{1}{2}c(1-c)u_{i+1}^n . \quad (13.3)$$

Here  $c$  is the CFL or Courant number

$$c = \frac{a\Delta t}{\Delta x} . \quad (13.4)$$

For background on these well-known schemes see Sect. 5.2 of Chap. 5. A general form of these schemes is

$$u_i^{n+1} = \sum_{k=-k_L}^{k_R} b_k u_{i+k}^n , \quad (13.5)$$

where  $k_L$  and  $k_R$  are two non-negative integers. We assume that these schemes (13.5) are *linear*, that is all the coefficients  $b_k$  are constant. Any scheme (13.5) is entirely determined by the coefficients  $b_k$ . Table 13.1 lists the coefficients for several schemes. The first on the list is the Lax–Friedrichs method (LF);

the second scheme is the First-Order Centred scheme (FORCE) introduced in Sect. 7.4.2 of Chap. 7; the third and fourth schemes are the two versions of the Godunov first-order upwind method (GODu). The fifth scheme is the Godunov first-order centred method (GODc) [217]. LW is the Lax-Wendroff method; WB is the Warming-Beam scheme and FR is the Fromm scheme.

Properties of the various schemes are stated in terms of the coefficients  $b_k$ . An important feature of a given scheme is its *support*; this is defined as the number  $k_L + k_R + 1$  of data mesh point values  $u_i^n$  involved in the summation (13.5). One speaks of a *compact scheme* when its support is small. For example the support of the GODu scheme has only two points, whereas the support of the Fromm scheme has four points.

	$b_{-2}$	$b_{-1}$	$b_0$	$b_1$	$b_2$
LF	0	$\frac{1}{2}(1+c)$	0	$\frac{1}{2}(1-c)$	0
FORCE	0	$\frac{1}{4}(1+c)^2$	$\frac{1}{2}(1-c^2)$	$\frac{1}{4}(1-c)^2$	0
GODu $a > 0$	0	$c$	$1-c$	0	0
GODu $a < 0$	0	0	$1+c$	$-c$	0
GODc	0	$\frac{1}{2}c(1+2c)$	$1-2c^2$	$-\frac{1}{2}c(1-2c)$	0
LW	0	$\frac{1}{2}c(1+c)$	$1-c^2$	$-\frac{1}{2}c(1-c)$	0
WB, $a > 0$	$\frac{1}{2}c(c-1)$	$(2-c)c$	$\frac{1}{2}(c-2)(c-1)$	0	0
WB, $a < 0$	0	0	$\frac{1}{2}(c+2)(c+1)$	$-(c+2)c$	$\frac{1}{2}c(c+1)$
FR, $a > 0$	$-\frac{1}{4}c(1-c)$	$\frac{1}{4}c(5-c)$	$\frac{1}{4}(1-c)(4+c)$	$-\frac{1}{4}c(1-c)$	0
FR, $a < 0$	0	$\frac{1}{4}c(1+c)$	$\frac{1}{4}(1+c)(4-c)$	$-\frac{1}{4}c(5+c)$	$\frac{1}{4}c(1+c)$

**Table 13.1.** Coefficients  $b_k$  for several first and second order linear schemes

Another feature of a scheme is the *bias* of its support. Notice for instance that the support of the GODu scheme is biased towards the left of the centre point  $u_i^n$  or to the right, depending on the sign of the wave speed  $a$  in the conservation law (13.1). For  $a > 0$  the support is biased to the left; if  $a < 0$  the support is biased to the right. For this reason we say that the GODu scheme is *upwind*. By contrast the Lax-Friedrichs scheme has a *centred support*; there is always one point to the left and one point to the right of the centre of the stencil, regardless of the sign of  $a$ . Note however that the upwind coefficient of the Lax-Friedrichs scheme is always larger than the downwind coefficient; schemes whose support configuration does not depend on the sign of characteristic speeds are called *centred schemes*. Examples of centred schemes are Lax-Friedrichs, FORCE, Godunov's first-order centred method (GODc) and the Lax-Wendroff method.

Given a numerical method, there are four fundamental properties associated with it, namely consistency, stability, convergence and accuracy. A most

useful reference in this regard is Hirsch, [251]. See also Smith [450], Mitchell and Griffiths [352] and Hoffmann [253].

### 13.2.2 Accuracy

The first five schemes listed in Table 13.1 are first-order accurate. The remaining schemes are second-order accurate. Here we quantify the accuracy of a scheme by the form of the leading term in its *local truncation error*. Appropriate definitions and techniques to analyse the accuracy of a scheme are found in any textbook on numerical methods for PDEs. See for example the references listed above.

Next we state a theorem due to Roe [408] that facilitates the verification of the accuracy of any scheme of the form (13.5).

**Theorem 13.1 (Roe [408]).** *A scheme of the form (13.5) is  $p$ -th ( $p \geq 0$ ) order accurate in space and time if and only if*

$$\sum_{k=-k_L}^{k_R} k^q b_k = (-c)^q, \quad 0 \leq q \leq p. \quad (13.6)$$

*Proof.* (Exercise)

*Remark 13.2.* For each integer value of  $q$ , with  $0 \leq q \leq p$ , we verify that the sum of terms  $k^q b_k$ , for all integers  $k$  with  $-k_L \leq k \leq k_R$ , reproduces identically the power  $(-c)^q$ , where  $c$  is the Courant number (13.4). The condition for  $q = 0$  is simply the *Consistency Condition* of a scheme.

*Example 13.3 (Godunov's upwind method).*

The GODu scheme for  $a \geq 0$  has coefficients  $b_{-1} = c$ ,  $b_0 = 1 - c$ . For  $q = 0$  we have the summation

$$\sum_{k=-1}^0 k^0 b_k = (-1)^0 \times c + 0^0 \times (1 - c) = 1 = (-c)^0.$$

For  $q = 1$

$$\sum_{k=-1}^0 k^1 b_k = (-1)^1 \times c + 0^1 \times (1 - c) = -c = (-c)^1.$$

Thus all conditions for first order accuracy are satisfied. Note that the GODu scheme is *not* second-order accurate, as for  $q = 2$

$$\sum_{k=-1}^0 k^2 b_k = (-1)^2 \times c + 0^2 \times (1 - c) = c \neq (-c)^2.$$

*Example 13.4 (The Lax–Wendroff method).* Application of (13.6) with  $q = 0, 1, 2$  gives

$$\sum_{k=-1}^1 k^0 b_k = 1 = (-c)^0, \quad \sum_{k=-1}^1 k^1 b_k = -c = (-c)^1, \quad \sum_{k=-1}^1 k^2 b_k = (-c)^2.$$

Therefore the Lax–Wendroff method is second–order accurate in space and time. This scheme is *not* third–order accurate, as

$$\sum_{k=-1}^1 k^3 b_k = -c \neq (-c)^3.$$

**Exercise 13.5.** Apply (13.6) to prove that the Lax–Friedrichs, FORCE and Godunov’s first order centred (GODc) methods are first order accurate in space and time. Show also that the Warming–Beam and Fromm schemes are second order accurate in space and time.

**Solution 13.6.** (Left to the reader).

**Exercise 13.7 (The Lax–Wendroff Scheme).** Use (13.6) to show that any centred three–point support scheme of the form (13.5) to solve (13.1) that is second–order accurate in space and time must be identical to the Lax–Wendroff method.

**Solution 13.8.** (Left to the reader).

### 13.2.3 Stability

Background reading on stability analysis of numerical methods is found in any text book on numerical methods for differential equations. We particularly recommend Hirsch [251] Chaps. 7–10, Smith [450] Chaps. 3 and 4, Mitchell and Griffiths [352] Chap. 4 and Hoffmann [253] Chap. 4.

Naturally, the only useful numerical methods are those that are *convergent*. Unfortunately, it is difficult or impossible to prove, theoretically, the convergence of a particular numerical method. For linear problems a most useful result is the *Lax Equivalence Theorem*. This states that *the only convergent schemes are those that are both consistent and stable*. There are various techniques available for studying stability. However, even for simple problems and schemes, the algebra involved may be cumbersome and simply not lead to analytical conditions on the parameters of the scheme to assume stability or otherwise. For non–linear problems there is a severe lack of theorems, and in practice one relies heavily on linear analysis and on numerical experimentation as guidance. A useful practice when dealing with non–linear problems is to compute the numerical solution on a sequence of meshes of decreasing

mesh size and observe the behaviour of the numerical solutions as the mesh is refined. Exact solutions, even for simplified initial conditions, can also be very helpful here in assessing the performance of numerical methods.

For any centred three-point scheme of the form (13.5) we quote a useful result on stability derived for the model advection-diffusion equation [58]. Here we state the special case applicable to the model equation (13.1).

**Proposition 13.9.** *Any centred, three-point scheme*

$$u_i^{n+1} = b_{-1}u_{i-1}^n + b_0u_i^n + b_1u_{i+1}^n$$

for the model equation (13.1) is stable if the following two conditions on the coefficients  $b_k$  are met, namely

$$\left. \begin{array}{l} (i) \quad b_0(b_{-1} + b_1) \geq 0, \\ (ii) \quad b_0(b_{-1} + b_1) + 4b_{-1}b_1 \geq 0. \end{array} \right\} \quad (13.7)$$

*Proof.* (Exercise).

*Example 13.10 (The Godunov first-order upwind method GODu).* For  $a > 0$ ,  $b_{-1} = c$ ,  $b_0 = 1 - c$ ,  $b_1 = 0$ . Thus (i) and (ii) become identical; they both give  $c(1 - c) \geq 0$ . But  $a > 0$  and so  $c > 0$ . Thus  $1 - c > 0$  and therefore  $0 \leq c \leq 1$  is the stability condition for the GODu scheme when  $a \geq 0$ . For  $a < 0$  we obtain the condition  $-1 \leq c \leq 0$  and thus for both positive and negative speed  $a$  the stability condition for the GODu scheme becomes

$$0 \leq |c| \leq 1. \quad (13.8)$$

**Exercise 13.11 (Stability Conditions).** Apply (13.7) to prove that the stability condition of the Lax-Friedrichs and FORCE schemes is as given by (13.8) and that the stability condition for the Godunov first order centred method GODc is

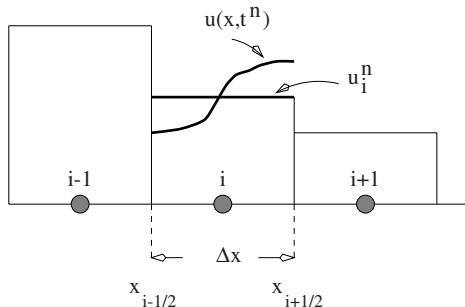
$$0 \leq |c| \leq \frac{1}{2}\sqrt{2}. \quad (13.9)$$

**Solution 13.12.** Left to the reader.

In the next two sections we study various ways of constructing high order numerical methods. The emphasis is on explicit, fully discrete schemes that are based on the solution of the Riemann problem and that can be extended to non-linear systems of conservation laws. We shall consider two approaches. One is the Weighted Average Flux approach and the other is the MUSCL or Variable Extrapolation approach.

### 13.3 WAF-Type High Order Schemes

The Weighted Average Flux (or WAF for short) approach is a generalisation of the Lax–Wendroff and the Godunov first-order upwind schemes to non-linear systems of conservation laws. It is also a generalisation of the Warming–Beam method. The WAF approach has its origins in the random flux scheme [497], which is second-order accurate in space and time in a statistical sense [539]. The WAF approach is *deterministic* and leads to fully discrete, explicit second order accurate schemes. See Refs. [499], [506], [504] [527], [63].



**Fig. 13.1.** Integral averages produce piece-wise constant data  $u_i^n$  at time level  $n$

In this section we present the approach as applied to a general scalar conservation law

$$u_t + f(u)_x = 0 , \quad (13.10)$$

where  $u = u(x, t)$  and  $f(u)$  is a convex flux function, see Sect. 2.4 of Chap. 2. The scheme is based on the explicit conservative formula

$$u_i^{n+1} = u_i^n + \frac{\Delta t}{\Delta x} \left[ f_{i-\frac{1}{2}} - f_{i+\frac{1}{2}} \right] . \quad (13.11)$$

Full details for the linear advection case (13.1) will be developed. Useful background reading is found in Sect. 5.3 of Chap. 5.

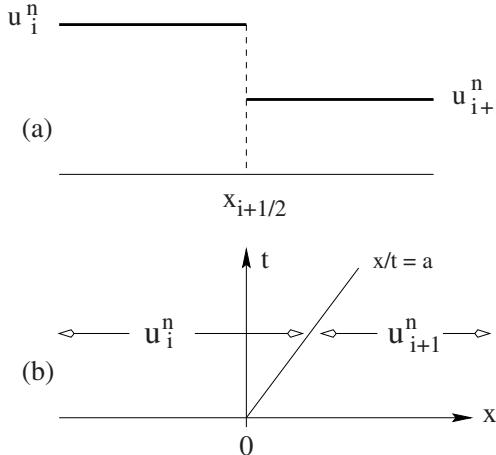
#### 13.3.1 The Basic WAF Scheme

The WAF method assumes piece-wise constant data  $\{u_i^n\}$ , as in the Godunov first-order upwind method; that is,  $u_i^n$  is an integral average of the solution  $u(x, t^n)$  within a cell  $I_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$  at time  $t = t^n$ , namely

$$u_i^n = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} u(x, t^n) dx . \quad (13.12)$$

Fig. 13.1 illustrates the cell average  $u_i^n$  in cell  $I_i$  of length  $x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}} = \Delta x$ . In practice, the formality of defining  $u_i^n$  is not necessary. The initial data  $\{u_i^0\}$

at time  $t = 0$  is regarded as cell averages and for all subsequent time levels these averages are produced automatically by formula (13.11). Thus, all we require is a definition of the numerical flux. In the original presentation of the



**Fig. 13.2.** The Riemann problem for the linear advection equation: (a) piece-wise constant initial data, (b) exact solution on  $x$ - $t$  plane for  $a > 0$

WAF method, the intercell flux was defined as an integral average of the flux function  $f(u)$ , namely

$$f_{i+\frac{1}{2}}^{waf} = \frac{1}{\Delta x} \int_{-\frac{1}{2}\Delta x}^{\frac{1}{2}\Delta x} f[u_{i+\frac{1}{2}}(x, \frac{1}{2}\Delta t)] dx. \quad (13.13)$$

Here, the integration range goes from the middle of cell  $I_i$  to the middle of cell  $I_{i+1}$ . The integrand is the physical flux function  $f(u)$  in the conservation law (13.10) (linear or non-linear) evaluated at  $u_{i+\frac{1}{2}}(x, \frac{1}{2}\Delta t)$ , where  $u_{i+\frac{1}{2}}(x, t)$  is the solution of the Riemann problem with *piece-wise constant data*  $(u_i^n, u_{i+1}^n)$ . In defining (13.13) we have assumed a regular mesh of size  $\Delta x$ , but this definition can easily be altered to include the case of irregular meshes.

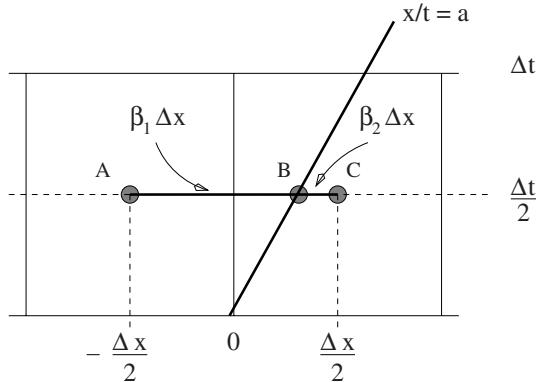
Now we develop, in full detail, the WAF scheme as applied to the linear advection equation (13.1). As seen in Sect. 2.2 of Chap. 2, the solution  $u_{i+\frac{1}{2}}(x, t)$  of the Riemann problem for (13.1) is trivial and consists of the two data states  $u_i^n, u_{i+1}^n$  separated by a wave of speed  $a$ . Fig. 13.2(a) shows the initial data and Fig. 13.2(b) shows the solution  $u_{i+\frac{1}{2}}(x, t)$  for  $a > 0$ . Algebraically

$$u_{i+\frac{1}{2}}(x, t) = \begin{cases} u_i^n, & \frac{x}{t} < a, \\ u_{i+1}^n, & \frac{x}{t} > a. \end{cases} \quad (13.14)$$

Obviously the two states present in the solution  $u_{i+\frac{1}{2}}(x, t)$  are constant and thus the exact evaluation of the integral (13.13) is trivial, see Fig. 13.3. The integration domain is the line AC, which is subdivided into the segments AB and BC. Note that the lengths  $|AB|$  and  $|BC|$ , when normalised by  $\Delta x$ , are respectively given by

$$\beta_1 = \frac{1}{2}(1 + c), \quad \beta_2 = \frac{1}{2}(1 - c). \quad (13.15)$$

The integral (13.13) becomes the summation



**Fig. 13.3.** Evaluation of the WAF flux for the linear advection equation, for  $a > 0$ . Weights normalised by distance  $\Delta x$  are  $\beta_1, \beta_2$

$$f_{i+\frac{1}{2}}^{waf} = \frac{1}{\Delta x} \left[ \frac{1}{2}(1 + c)\Delta x(a u_i^n) \right] + \frac{1}{\Delta x} \left[ \frac{1}{2}(1 - c)\Delta x(a u_{i+1}^n) \right],$$

or

$$f_{i+\frac{1}{2}}^{waf} = \frac{1}{2}(1 + c)(a u_i^n) + \frac{1}{2}(1 - c)(a u_{i+1}^n). \quad (13.16)$$

This numerical flux is identical to that of the Lax-Wendroff method in conservation form. For non-linear conservation laws (13.10)  $f_{i+\frac{1}{2}}^{waf}$ , as given by (13.13), becomes a Riemann-problem based generalisation of the Lax-Wendroff method, see Sect 5.3.4 of Chap. 5.

Note that for  $a > 0$  the WAF flux as given by (13.16) is a *weighted average* of the upwind flux  $f_i = a u_i^n$ , with weight  $\beta_1 = \frac{1}{2}(1 + c)$  and the downwind flux  $f_{i+1} = a u_{i+1}^n$  with weight  $\beta_2 = \frac{1}{2}(1 - c)$ . For  $a < 0$  the derivation is analogous and the result is in fact identical to (13.16), but now  $\beta_1 = \frac{1}{2}(1 + c)$  is the downwind weight and  $\beta_2 = \frac{1}{2}(1 - c)$  is the upwind weight. The upwind weight is always larger than the downwind weight and thus the WAF method is, in this sense, upwind biased. The WAF flux contains the Godunov flux in one of its terms, and it may thus be also regarded as an extension of the Godunov first order upwind method.

**Exercise 13.13.** Evaluate the WAF numerical flux (13.13) for the linear advection equation (13.1) for the case  $a < 0$ .

**Solution 13.14.** Left to the reader.

It is convenient to express intercell numerical fluxes for the linear advection equation (13.1) as a linear combination of data flux values  $f_i = au_i^n$ , namely

$$f_{i+\frac{1}{2}} = \sum_{k=-1}^2 \beta_k (au_{i+k}^n). \quad (13.17)$$

Table 13.2 lists the coefficients  $\beta_k$  for a range of schemes.

	$\beta_{-1}$	$\beta_0$	$\beta_1$	$\beta_2$
LF	0	$\frac{1}{2c}(1+c)$	$-\frac{1}{2c}(1-c)$	0
FORCE	0	$\frac{1}{4c}(1+c)^2$	$-\frac{1}{4c}(1-c)^2$	0
GODu	0	$\frac{1}{2}(1 + \text{sign}(c))$	$\frac{1}{2}(1 - \text{sign}(c))$	0
GODc	0	$\frac{1}{2}(1+2c)$	$\frac{1}{2}(1-2c)$	0
LW	0	$\frac{1}{2}(1+c)$	$\frac{1}{2}(1-c)$	0
WB, $a > 0$	$-\frac{1}{2}(1-c)$	$\frac{1}{2}(3-c)$	0	0
WB, $a < 0$	0	0	$\frac{1}{2}(3+c)$	$-\frac{1}{2}(1+c)$
FR, $a > 0$	$-\frac{1}{4}(1-c)$	1	$\frac{1}{4}(1-c)$	0
FR, $a < 0$	0	$\frac{1}{4}(1+c)$	1	$-\frac{1}{4}(1+c)$
SLIC	$-\frac{1}{16}(1-c)$ $(1+c)^2$	$\frac{1}{4c}(1+c)^2$ $+\frac{1}{16c}(1+c)$ $(1-c)^2$	$\frac{1}{16c}(1-c)$ $(1+c)^2$ $+\frac{1}{4c}(1-c)^2$	$-\frac{1}{16c}(1+c)$ $(1-c)^2$

**Table 13.2.** Coefficients  $\beta_k$  in the expression for the intercell flux for various schemes. See also Table 13.1

**Exercise 13.15.** Evaluate the WAF flux, according to (13.13), for the inviscid Burgers equation,  $f(u) = \frac{1}{2}u^2$  in (13.10). Recall that the Riemann problem at the interface  $i + \frac{1}{2}$ , see Sect. 2.4.2 of Chap. 2, has two cases, namely the shock case  $u_i^n > u_{i+1}^n$  and the rarefaction case  $u_i^n \leq u_{i+1}^n$ .

**Solution 13.16.** Left to the reader.

### 13.3.2 Generalisations of the WAF Scheme

Generalisations of the basic WAF flux (13.13) may be carried out in a number of ways. The definition of the flux to include irregular meshes of size  $(\Delta x)_i$  is obvious,

$$f_{i+\frac{1}{2}}^{waf} = \frac{1}{\frac{1}{2}[(\Delta x)_i + (\Delta x)_{i+1}]} \int_{-\frac{1}{2}(\Delta x)_i}^{\frac{1}{2}(\Delta x)_{i+1}} f[u_{i+\frac{1}{2}}(x, \frac{\Delta t}{2})] dx \quad (13.18)$$

and definitions (13.13) and (13.18) are actually special cases of the more general formula

$$f_{i+\frac{1}{2}}^{waf} = \frac{1}{t_2 - t_1} \frac{1}{x_2 - x_1} \int_{t_1}^{t_2} \int_{x_1}^{x_2} f[\hat{u}_{i+\frac{1}{2}}(x, t)] dx dt. \quad (13.19)$$

Here we integrate  $f(\hat{u}_{i+\frac{1}{2}})$  in a box  $[x_1, x_2] \times [t_1, t_2]$  in the  $x-t$  plane, where  $\hat{u}_{i+\frac{1}{2}}(x, t)$  is the solution of some relevant initial value problem.

Let us consider the following conditions on the integral (13.19) as applied to the linear advection equation (13.1):

- (i)  $x_1 = -\frac{1}{2}\Delta x$ ,  $x_2 = \frac{1}{2}\Delta x$ ,
- (ii)  $t_1 = 0$ ,  $t_2 = \Delta t$ ,
- (iii) the time integral in (13.19) is approximated by the mid-point rule in time,
- (iv)  $\hat{u}_{i+\frac{1}{2}} \equiv u_{i+\frac{1}{2}}(x, t)$  is the solution of the Riemann problem with data  $u_i^n, u_{i+1}^n$ ,
- (v)  $\Delta t$  is such that  $|c| = \frac{\Delta t}{\Delta x} |a| < 1$ .

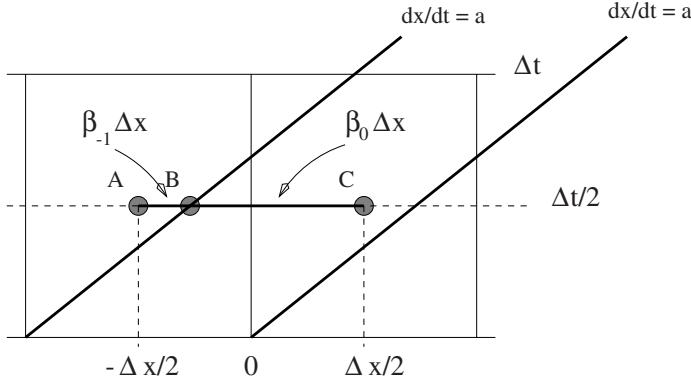
If all these conditions are fulfilled, then (13.19) reduces identically to (13.13). If in condition (iii) above we replace the approximate integration by exact integration, then two second-order accurate methods result; for  $c \leq \frac{1}{2}$  one obtains the Lax-Wendroff method; for  $\frac{1}{2} \leq c \leq 1$  one obtains a second-order method which appears to be new.

Consider the case in which conditions (i)-(iv) are imposed but condition (v) is weakened so that the CFL number  $|c|$  may be larger than unity. A consequence of relaxing (v) in this way is that  $\hat{u}_{i+\frac{1}{2}}$  may become the solution of more than one piece-wise constant data Riemann problem. For the linear advection equation this results in an explicit method with arbitrarily large time step  $\Delta t$ , see Casulli and Toro [95]. Extensions of this large-time step scheme to non-linear problems is, however, very difficult. The particular case in which  $|c| \leq 2$  is of interest and we call the resulting scheme WAF CFL2 [527]; this scheme is effectively a combined, Riemann-problem based extension of the Lax-Wendroff scheme and the Warming-Beam scheme [574], namely

$$f_{i+\frac{1}{2}}^{wafcfl2} = \begin{cases} f_{i+\frac{1}{2}}^{lw} = \frac{1}{2}(1+c)(au_i^n) + \frac{1}{2}(1-c)(au_{i+1}^n), & |c| < 1, \\ f_{i+\frac{1}{2}}^{wb}, & 1 \leq |c| \leq 2, \end{cases} \quad (13.20)$$

where  $f_{i+\frac{1}{2}}^{wb}$  is the Warming-Beam numerical flux

$$f_{i+\frac{1}{2}}^{wb} = \begin{cases} -\frac{1}{2}(1-c)(au_{i-1}^n) + \frac{1}{2}(3-c)(au_i^n), & \text{if } a > 0, \\ \frac{1}{2}(3+c)(au_{i+1}^n) - \frac{1}{2}(1+c)(au_{i+2}^n), & \text{if } a < 0. \end{cases} \quad (13.21)$$



**Fig. 13.4.** Evaluation of the WAF flux for the linear advection equation, for  $a > 0$  and  $1 \leq c \leq 2$

Fig. 13.4 illustrates the flux evaluation for  $a > 0$  and  $1 \leq c \leq 2$ . Now the normalised lengths  $|AB|$  and  $|BC|$  for  $a > 0$  are  $\beta_{-1} = -\frac{1}{2}(1-c)$ ,  $\beta_0 = \frac{1}{2}(3-c)$ . For  $a < 0$  the weights are  $\beta_1 = \frac{1}{2}(3+c)$ ,  $\beta_2 = -\frac{1}{2}(1+c)$ . See (13.17) and Table 13.2.

The WAF CFL2 scheme of Toro and Billett has been extended to non-linear systems, with a suitable TVD condition in [527], where it is applied to the time-dependent two dimensional Euler equations. A final remark on the WAF CFL2 scheme is this. If the integral (13.19), under the conditions following it, is evaluated by the mid-point rule in space and exact integration in time, then a first-order upwind method of *Courant number two* stability restriction is produced. This has numerical flux

$$f_{i+\frac{1}{2}}^{tb} = \begin{cases} f_{i+\frac{1}{2}}^{GODu} = \frac{1}{2}(1 + \text{sign}(c))(au_i^n) \\ \quad + \frac{1}{2}(1 - \text{sign}(c))(au_{i+1}^n), & \text{if } |c| \leq 1, \\ \frac{(c-1)}{c}(au_{i-1}^n) + \frac{1}{c}(au_i^n), & \text{if } a > 0, 1 \leq c \leq 2, \\ -\frac{1}{c}(au_{i+1}^n) + \frac{(c+1)}{c}(au_{i+2}^n), & \text{if } a < 0, 1 \leq |c| \leq 2. \end{cases} \quad (13.22)$$

**Exercise 13.17.** Substitute the above numerical fluxes (13.22) into the conservative formula (13.11) to derive the scheme in the form (13.5). Check the accuracy of the scheme using formula (13.6) and verify that the scheme produces the exact solution for  $|c| = 1$  and  $|c| = 2$ .

**Solution 13.18.** Left to the reader.

A Total Variation Diminishing (TVD) modification of the WAF method is constructed in Sect. 13.7.1. In Chap. 14 we extend the WAF method to non-linear systems of conservation laws; in Chap. 15 we extend the scheme

further to deal with source terms and in Chap. 16 we extend the scheme to deal with multi-dimensional problems.

In the following section we present another approach for deriving high-order numerical methods.

## 13.4 MUSCL-Type High-Order Methods

Van Leer, see [557], [558], [559], introduced the idea of *modifying* the piecewise constant data (13.12) in the first-order Godunov method, as a first step to achieve higher order of accuracy. This approach has become known as the MUSCL or Variable Extrapolation approach, where MUSCL stands for Monotone Upstream-centred Scheme for Conservation Laws. The MUSCL approach is routinely used in practice today; it allows the construction of very high order methods, fully discrete, semi-discrete and also implicit methods. As we shall see later in this chapter, high-order (linear) schemes produce spurious oscillations. See numerical results of Chap. 5. The MUSCL approach implies (i) high-order of accuracy obtained by data reconstruction and (ii) the reconstruction is constrained so as to avoid spurious oscillations, and thus the justification of the word *monotone* in the name of the scheme. In this section we shall only be concerned with the first aspect of MUSCL, that is as a method for increasing the accuracy of first-order schemes.

### 13.4.1 Data Reconstruction

The simplest way of modifying the piecewise constant data  $\{u_i^n\}$  is to replace the constant states  $u_i^n$  by piecewise linear functions  $u_i(x)$ . As for the first-order Godunov method, one assumes that  $u_i^n$  represents an integral average in cell  $I_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$  as given by (13.12).

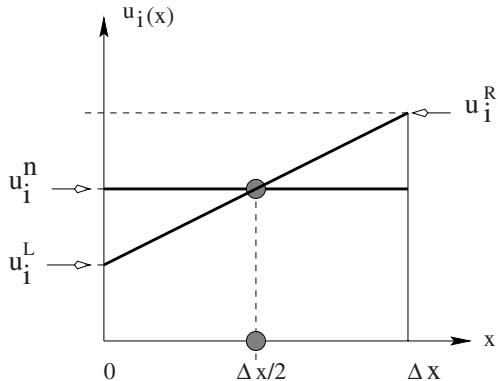
A piece-wise linear, *local* reconstruction of  $u_i^n$  is

$$u_i(x) = u_i^n + \frac{(x - x_i)}{\Delta x} \Delta_i, \quad x \in [0, \Delta x], \quad (13.23)$$

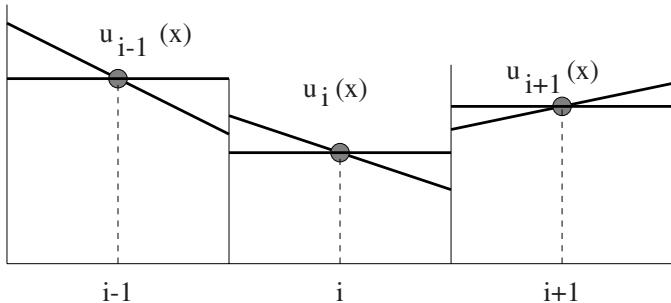
where  $\frac{\Delta_i}{\Delta x}$  is a suitably chosen *slope* of  $u_i(x)$  in cell  $I_i$ . In what follows we call  $\Delta_i$  a slope. Note that the linear function  $u_i(x)$  is defined *locally* in cell  $I_i$ , that is  $x \in [0, \Delta x]$ . Fig. 13.5 illustrates the situation. The centre of the cell  $x_i$  in local co-ordinates is  $x = \frac{1}{2}\Delta x$  and  $u_i(x_i) = u_i^n$ . The values of  $u_i(x)$  at the extreme points play a fundamental role; they are given by

$$u_i^L = u_i(0) = u_i^n - \frac{1}{2}\Delta_i; \quad u_i^R = u_i(\Delta x) = u_i^n + \frac{1}{2}\Delta_i \quad (13.24)$$

and are usually called *boundary extrapolated values*; hence the alternative name of *Variable Extrapolation Method*. Note that the integral of  $u_i(x)$  in cell  $I_i$  is identical to that of  $u_i^n$  and thus the reconstruction process retains



**Fig. 13.5.** Piece–wise linear MUSCL reconstruction of data in a single computing cell  $I_i$ ; boundary extrapolated values are  $u_i^L, u_i^R$



**Fig. 13.6.** Piece–wise linear MUSCL reconstruction for three successive computing cells  $I_{i-1}, I_i, I_{i+1}$

conservation. Fig. 13.6 illustrates the piece–wise linear reconstruction process applied to three successive cells. As a consequence of having modified the data, at each interface  $i + \frac{1}{2}$  one now may consider the so called *Generalised Riemann Problem* (or GRP)

$$\left. \begin{aligned} u_t + f(u)_x &= 0, \\ u(x, 0) &= \begin{cases} u_i(x), & x < 0, \\ u_{i+1}(x), & x > 0, \end{cases} \end{aligned} \right\} \quad (13.25)$$

to compute an intercell Godunov-type flux  $f_{i+\frac{1}{2}}$ . Fig. 13.7 illustrates the initial data and the structure of the solution of the Generalised Riemann problem. The solution no longer contains uniform regions as in the conventional Riemann problem in which the data is piece–wise constant; wave paths are now curved in  $x$ - $t$  space.

Ben–Artzi and Falcovitz [37] were the first to develop a method based on the GRP. They applied their scheme to the Euler equations. Naturally, for

non-linear systems the exact solution of the generalised Riemann problem is exceedingly complicated, but for the purpose of flux evaluation, approximate information may be obtained. Most approaches do in fact give up the solution of the generalised Riemann problem and rely instead on judicious use of the boundary extrapolated values  $u_i^L, u_i^R$  in (13.24), for each function  $u_i(x)$  (13.23). In this way, one may instead consider the *piece-wise constant data* Riemann problem

$$\left. \begin{aligned} u_t + f(u)_x &= 0, \\ u(x, 0) &= \begin{cases} u_i^R, & x < 0, \\ u_{i+1}^L, & x > 0, \end{cases} \end{aligned} \right\} \quad (13.26)$$

which if used in conjunction with some other procedures can produce useful high-order methods.

As to the choice of slopes  $\Delta_i$  in (13.23) we define

$$\Delta_i = \frac{1}{2}(1 + \omega)\Delta u_{i-\frac{1}{2}} + \frac{1}{2}(1 - \omega)\Delta u_{i+\frac{1}{2}}, \quad (13.27)$$

where as usual

$$\Delta u_{i-\frac{1}{2}} \equiv u_i^n - u_{i-1}^n, \quad \Delta u_{i+\frac{1}{2}} \equiv u_{i+1}^n - u_i^n \quad (13.28)$$

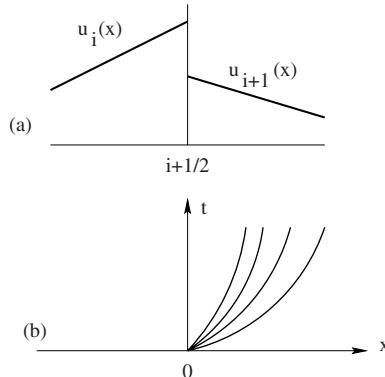
and  $\omega$  is a free parameter in the real interval  $[-1, 1]$ . For  $\omega = 0$ ,  $\Delta_i$  is a central-difference approximation, multiplied by  $\Delta x$ , to the first spatial derivative of the numerical solution at time level  $n$ .

MUSCL data reconstructions that are more accurate than the piece-wise linear reconstruction (13.23) are possible. A piece-wise quadratic reconstruction [251] is

$$u_i(x) = u_i^n + \frac{(x - x_i)}{\Delta x}\Delta_i^{(1)} + \frac{3\kappa}{2(\Delta x)^2} \left[ (x - x_i)^2 - \frac{(\Delta x)^2}{12} \right] \Delta_i^{(2)}, \quad (13.29)$$

where  $\Delta_i^{(1)} = \Delta_i$  as in (13.27),  $\Delta_i^{(2)}$  is associated with an estimate for the second space derivative of  $u_i(x)$  in cell  $I_i$  and  $\kappa$  is a parameter. For  $\kappa = \frac{1}{3}$  this piece-wise quadratic reconstruction leads to third-order accurate *space* discretisations. The Piece-wise Parabolic Method (PPM) of Colella and Woodward [137], [584] is based on a piece-wise parabolic reconstruction of the data.

The next sections present various methodologies based on MUSCL data reconstruction to derive high-order methods. Three of the approaches use, in addition, solutions of associated Riemann problems. We also present a high-order scheme that uses the MUSCL reconstruction step, but instead of using the Riemann problem solution it uses the First-Order Centred Scheme FORCE introduced in Sect. 7.4.2 Chap. 7, giving rise to a centred high-order method.



**Fig. 13.7.** The Generalised Riemann Problem (GRP): (a) piece–wise linear data (b) structure of the solution in  $x$ – $t$  plane

### 13.4.2 The MUSCL–Hancock Method (MHM)

Van Leer [562] attributes this method to S. Hancock, and we therefore adopt the name MUSCL–Hancock Method, or MHM, for this approach. MHM has three distinct steps to construct fully discrete second–order accurate schemes based on (13.11) to solve (13.10). These are:

Step (I) Data reconstruction as in (13.23) with boundary extrapolated values as in (13.24), namely

$$u_i^L = u_i^n - \frac{1}{2}\Delta_i ; \quad u_i^R = u_i^n + \frac{1}{2}\Delta_i .$$

Step (II) Evolution of  $u_i^L, u_i^R$  by a time  $\frac{1}{2}\Delta t$  according to

$$\left. \begin{aligned} \bar{u}_i^L &= u_i^L + \frac{1}{2} \frac{\Delta t}{\Delta x} [f(u_i^L) - f(u_i^R)] , \\ \bar{u}_i^R &= u_i^R + \frac{1}{2} \frac{\Delta t}{\Delta x} [f(u_i^L) - f(u_i^R)] . \end{aligned} \right\} \quad (13.30)$$

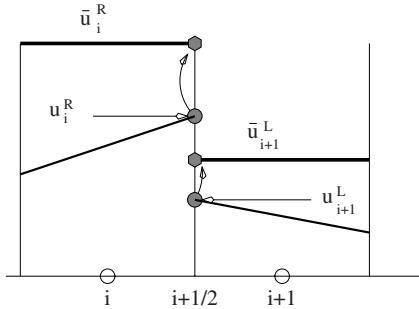
Step (III) Solution of the piece–wise constant data Riemann problem

$$\left. \begin{aligned} u_t + f(u)_x &= 0 , \\ u(x, 0) &= \begin{cases} \bar{u}_i^R , & x < 0 , \\ \bar{u}_{i+1}^L , & x > 0 , \end{cases} \end{aligned} \right\} \quad (13.31)$$

to find the *similarity* solution  $u_{i+\frac{1}{2}}(x/t)$ . Fig. 13.8 illustrates steps (I) and (II) at the intercell boundary position  $i+\frac{1}{2}$ ; the boundary extrapolated values  $u_i^R, u_{i+1}^L$  are evolved to  $\bar{u}_i^R, \bar{u}_{i+1}^L$ . These form the *piece–wise constant data* for

a conventional Riemann problem at the cell interface  $i + \frac{1}{2}$  with solution  $u_{i+\frac{1}{2}}(x/t)$ . The intercell numerical flux  $f_{i+\frac{1}{2}}$  is then obtained in exactly the same way as in the Godunov first-order upwind method, see Sect. 5.3.2 of Chaps. 5 and Chap. 6, namely

$$f_{i+\frac{1}{2}} = f(u_{i+\frac{1}{2}}(0)) . \quad (13.32)$$



**Fig. 13.8.** Boundary extrapolated values. At each interface  $i + \frac{1}{2}$  boundary extrapolated values  $u_i^R, u_{i+1}^L$  are evolved to  $\bar{u}_i^R, \bar{u}_{i+1}^L$ , to form the piece-wise constant data for a conventional Riemann problem at the intercell boundary

Now we develop the details of the scheme for the linear advection equation (13.1). Step (I) gives the boundary extrapolated values (13.24). In step (II) we need the fluxes  $f(u_i^L), f(u_i^R)$ . As for the linear advection equation  $f(u) = au$ , then we have

$$f(u_i^L) = a[u_i^n - \frac{1}{2}\Delta_i] ; \quad f(u_i^R) = a[u_i^n + \frac{1}{2}\Delta_i] .$$

Substitution of these into (13.30) gives the evolved boundary extrapolated values in cell  $I_i$ , namely

$$\bar{u}_i^L = u_i^n - \frac{1}{2}(1+c)\Delta_i ; \quad \bar{u}_i^R = u_i^n + \frac{1}{2}(1-c)\Delta_i . \quad (13.33)$$

In step (III) one solves the conventional Riemann problem at the interface  $i + \frac{1}{2}$  with data

$$(\bar{u}_i^R, \bar{u}_{i+1}^L)$$

the solution of which is

$$u_{i+\frac{1}{2}}(x/t) = \begin{cases} \bar{u}_i^R = u_i^n + \frac{1}{2}(1-c)\Delta_i , & \text{if } x/t < a , \\ \bar{u}_{i+1}^L = u_{i+1}^n - \frac{1}{2}(1+c)\Delta_{i+1} , & \text{if } x/t > a . \end{cases} \quad (13.34)$$

The intercell flux is now as in the first-order Godunov method (13.32). The result is

$$f_{i+\frac{1}{2}}^{mh} = \frac{1}{2}(1 + sign(c))f(\bar{u}_i^R) + \frac{1}{2}(1 - sign(c))f(\bar{u}_{i+1}^L), \quad (13.35)$$

where

$$f(\bar{u}_i^R) = a[u_i^n + \frac{1}{2}(1 - c)\Delta_i]; \quad f(\bar{u}_{i+1}^L) = a[u_{i+1}^n - \frac{1}{2}(1 + c)\Delta_{i+1}]. \quad (13.36)$$

Recall that  $sign(x) = 1$  if  $x > 0$  and  $sign(x) = -1$  if  $x < 0$ , for any real number  $x$ .

When expressed as in (13.17) the MHM flux for the linear advection equation (13.1) has coefficients  $\beta_k$  that depend on the Courant number  $c$  and the parameter  $\omega$ .

**Exercise 13.19.** Using the notation (13.17) verify that the coefficients  $\beta_k$  in (13.35), for  $a > 0$ , are

$$\beta_{-1} = -\frac{1}{4}(1 - c)(1 + \omega); \beta_0 = 1 + \frac{1}{2}(1 - c)\omega; \beta_1 = \frac{1}{4}(1 - c)(1 - \omega); \beta_2 = 0 \quad (13.37)$$

and that for  $a < 0$  they are

$$\beta_{-1} = 0; \beta_0 = \frac{1}{4}(1 + c)(1 + \omega); \beta_1 = 1 - \frac{1}{2}(1 + c)\omega; \beta_2 = -\frac{1}{4}(1 + c)(1 - \omega). \quad (13.38)$$

**Solution 13.20.** Left to the reader.

**Exercise 13.21.** Verify that if the parameter  $\omega$  in the slopes (13.27) is  $\omega = 0$ , then MHM reproduces the Fromm scheme, see Tables 13.1 and 13.2. Assume  $a > 0$  and verify that for  $\omega = 1$  the scheme is identical to the Warming–Beam scheme and that for  $\omega = -1$  the scheme is identical to the Lax–Wendroff method; see Tables 13.1 and 13.2.

**Solution 13.22.** Left to the reader.

**Exercise 13.23.** Apply (13.6) to show that for any value of  $\omega$  in the slopes  $\Delta_i$  in (13.27) MHM is second–order accurate in space and time, for both  $a < 0$  and  $a > 0$ .

**Solution 13.24.** Left to the reader.

**Exercise 13.25.** Apply (13.6) to show that MHM is third–order accurate in space and time if the parameter  $\omega$  in (13.27) is chosen to be

$$\omega = \frac{1}{3}(2c - sign(c)). \quad (13.39)$$

**Solution 13.26.** Left to the reader.

A Total Variation Diminishing ( TVD) version of the MHM is constructed in Sect. 13.8. Extensions to non–linear systems in one space dimension are given in Chap. 14. Application of the method to systems involving source terms is discussed in Chap. 15 and extension to multidimensional problems is given in Chap. 16.

### 13.4.3 The Piece-Wise Linear Method (PLM)

Colella [133] proposed the Piece-wise Linear Method, or PLM, to obtain second-order accurate generalisations of the Godunov first-order upwind method. PLM is also based on the MUSCL data reconstruction step discussed in Sec. 13.4.1. For any conservation law (13.10) as solved by (13.11), PLM defines the numerical flux  $f_{i+\frac{1}{2}}$  as

$$f_{i+\frac{1}{2}}^{PLM} = \frac{1}{2}[f(u_{i+\frac{1}{2}}^{grp}(0, 0)) + f(u_{i+\frac{1}{2}}^{grp}(0, \Delta t))], \quad (13.40)$$

where  $u_{i+\frac{1}{2}}^{grp}(x, t)$  is the solution of the Generalised Riemann problem (13.25). See Fig. 13.7. PLM does not solve the Generalised Riemann problem directly; instead, it uses a trapezium rule approximation in time to the integral

$$f_{i+\frac{1}{2}} = \frac{1}{\Delta t} \int_0^{\Delta t} f[u_{i+\frac{1}{2}}^{grp}(0, t)] dt, \quad (13.41)$$

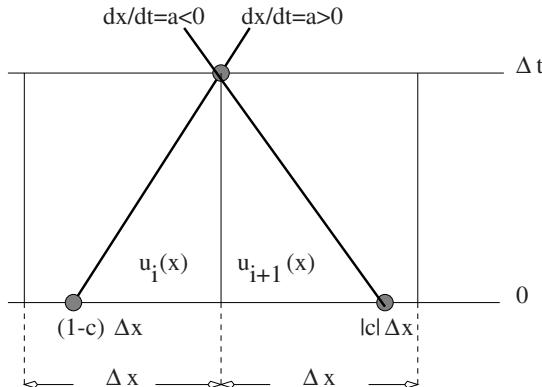
where  $u_{i+\frac{1}{2}}^{grp}(0, t)$  is the solution of the Generalised Riemann Problem along the cell interface position ( $t$ -axis). In order to determine the two terms in the PLM flux (13.40), one proceeds as follows

- Step (I) Data reconstruction in each cell  $I_i$  as in (13.23) with slopes as in (13.27) and computation of the boundary extrapolated values  $u_i^L, u_i^R$ , for each cell  $I_i$ .
- Step (II) The first term is found by solving the conventional Riemann problem

$$\left. \begin{aligned} u_t + f(u)_x &= 0, \\ u(x, 0) &= \begin{cases} u_i^R, & x < 0, \\ u_{i+1}^L, & x > 0, \end{cases} \end{aligned} \right\} \quad (13.42)$$

with piece-wise constant data  $u_i^R, u_{i+1}^L$ . As  $u_{i+\frac{1}{2}}^{grp}(0, 0)$  is solely determined by the interaction of the extreme data values  $u_i^R, u_{i+1}^L$  in the Generalised Riemann Problem (13.25), the solution  $u_{i+\frac{1}{2}}(x/t)$  of (13.42) produces the sought result, namely  $u_{i+\frac{1}{2}}^{grp}(0, 0) = u_{i+\frac{1}{2}}(0)$ .

- Step (III) The second term in (13.40) depends on  $u_{i+\frac{1}{2}}^{grp}(0, \Delta t)$ , which is found by tracing characteristics from the point  $(0, \Delta t)$  back to the *reconstructed piece-wise linear data* (13.23); Fig. 13.9 illustrates this step for the linear advection equation (13.1). Then  $u_{i+\frac{1}{2}}^{grp}(0, \Delta t) = u_k(\hat{x})$ , where  $k$  is either  $i$  or  $i + 1$  and  $\hat{x}$  is the point of intersection of the characteristic with the  $x$ -axis.



**Fig. 13.9.** Tracing characteristic  $dx/dt = a$  back to the reconstructed data  $u_i$  to find the constant value at the interface position  $i + \frac{1}{2}$  at time  $\Delta t$

We now work out the details for the linear advection equation (13.1). Step (I) produces the extrapolated values (13.24). Step (II) gives the solution of (13.42), which, when evaluated along the  $t$ -axis gives

$$u_{i+\frac{1}{2}}(0) = \begin{cases} u_i^R &= u_i^n + \frac{1}{2}\Delta_i, & \text{if } a > 0, \\ u_{i+1}^L &= u_{i+1}^n - \frac{1}{2}\Delta_{i+1}, & \text{if } a < 0, \end{cases}$$

which in turn leads to the first term in (13.40). In step (III), one traces the characteristic  $dx/dt = a$  from the point  $(0, \Delta t)$  back to cell  $I_i$  if  $a > 0$  or to cell  $I_{i+1}$  if  $a < 0$ , to find the intersection point  $\hat{x}$ ; this is

$$\hat{x} = \begin{cases} c\Delta x &, \text{ if } a \leq 0, \\ (1 - c)\Delta x &, \text{ if } a \geq 0. \end{cases}$$

See Fig. 13.9. For  $a > 0$ , evaluation of  $u_i(x)$  at  $\hat{x} = (1 - c)\Delta x$  in (13.23) gives

$$u_i[(1 - c)\Delta x] = u_i^n + \left(\frac{1}{2} - c\right)\Delta_i.$$

As this value is constant along the characteristic we have  $u_{i+\frac{1}{2}}^{grp}(0, \Delta t) = u_i^n + (\frac{1}{2} - c)\Delta_i$ , which gives the second term for the PLM flux (13.40).

**Exercise 13.27.** Verify that for both  $a > 0$  and  $a < 0$  the PLM flux is

$$f_{i+\frac{1}{2}}^{plm} = \begin{cases} a[u_i^n + \frac{1}{2}(1 - c)\Delta_i], & a > 0, \\ a[u_{i+1}^n - \frac{1}{2}(1 + c)\Delta_{i+1}], & a < 0. \end{cases} \quad (13.43)$$

**Solution 13.28.** Left to the reader.

The PLM intercell flux is identical to the MHM flux (13.35)–(13.36), for the linear advection equation and leads to the same numerical fluxes (13.17)

with coefficients (13.37)–(13.38), with third-order accuracy for the parameter value given by (13.39). See Colella [133] for details on the extension of PLM to non-linear systems. A simplified version of the PLM to construct conservative and non-conservative methods is found in [517].

### 13.4.4 The Generalised Riemann Problem (GRP) Method

The Generalised Riemann Problem, or GRP, method was first put forward by Ben-Artzi and Falcovitz [37], [38], [36], who applied it to the Euler equations. Here we present the scheme as applied to the scalar conservation law (13.10), as solved by (13.11). The basic ingredient of the GRP method is the solution of the Generalised Riemann Problem (13.25) to obtain a Godunov-type numerical flux that yields a second-order accurate scheme. The GRP method defines a numerical flux as

$$f_{i+\frac{1}{2}}^{grp} = f(u_{i+\frac{1}{2}}^{grp}(0, \frac{1}{2}\Delta t)) , \quad (13.44)$$

where  $u_{i+\frac{1}{2}}^{grp}(x, t)$  is the solution of the Generalised Riemann Problem (13.25) and  $u_{i+\frac{1}{2}}^{grp}(0, \frac{1}{2}\Delta t)$  is the mid-point rule approximation in time to the integral (13.41). As a matter of fact, a further approximation to  $u_{i+\frac{1}{2}}^{grp}(0, \frac{1}{2}\Delta t)$  is required to obtain the numerical flux (13.44). The GRP scheme then has the following steps

- Step (I) Data reconstruction in each cell  $I_i$  as in (13.23) with slopes as in (13.27).
- Step (II) Taylor expansion of  $u_{i+\frac{1}{2}}^{grp}(x, t)$  about  $(0, 0)$  to obtain an approximation to  $u_{i+\frac{1}{2}}^{grp}(0, \frac{1}{2}\Delta t)$ , namely

$$u_{i+\frac{1}{2}}^{grp}(0, \frac{1}{2}\Delta t) = u_{i+\frac{1}{2}}^{grp}(0, 0) + \frac{1}{2}\Delta t \frac{\partial}{\partial t} u_{i+\frac{1}{2}}^{grp}(0, 0) , \quad (13.45)$$

with error terms  $O(\Delta t^2)$ .

- Step (III) Determination of the two terms on the right hand side of (13.45).

The first term  $u_{i+\frac{1}{2}}^{grp}(0, 0)$  in the expansion (13.45) is the value of  $u_{i+\frac{1}{2}}^{grp}(x, t)$  immediately after the interaction of the piece-wise linear states  $u_i(x), u_{i+1}(x)$  in the GRP (13.25). This value is solely determined by the extrapolated values (13.24). Thus  $u_{i+\frac{1}{2}}^{grp}(0, 0)$  is the solution of the piece-wise constant data Riemann problem (13.42) evaluated along the  $t$ -axis, namely

$$u_{i+\frac{1}{2}}^{grp}(0) = \begin{cases} u_i^R &= u_i^n + \frac{1}{2}\Delta_i , & \text{if } a > 0 , \\ u_{i+1}^L &= u_{i+1}^n - \frac{1}{2}\Delta_{i+1} , & \text{if } a < 0 . \end{cases} \quad (13.46)$$

This is identical to the step leading to the first term in the PLM flux (13.40). To determine the second term in (13.45) there are various possibilities.

Here we suggest a modification of the original GRP method whereby the time derivative in (13.45) is replaced by a space derivative. The required value along the cell interface results from solving an extra Riemann problem for gradients. For the linear advection equation (13.1) the modification is based on the following result

**Proposition 13.29.** *For any  $p$ -th order spatial derivative  $v = \partial^p u / \partial x^p$ , where  $u(x, t)$  is a solution of the linear advection equation (13.1), we have*

$$v_t + av_x = 0 .$$

*Proof.* (Left as an exercise).

The proposition states that any spatial gradient of  $u(x, t)$ , if defined, obeys the original PDE. Hence one may pose Riemann problems for gradients  $\partial^p u / \partial x^p$ , which if assumed piecewise constant, lead to conventional Riemann problems.

For the linear advection equation (13.1) we have  $u_t = -au_x$  and thus the time derivative in the second term of (13.45) can be replaced by a space derivative. This in turn, by virtue of the above proposition, can be found by solving the gradient ( $p = 1$ ) Riemann problem

$$\left. \begin{aligned} v_t + av_x &= 0 , \\ v(x, 0) \equiv u_x(x, 0) &= \begin{cases} \frac{\Delta_i}{\Delta x} , & x < 0 , \\ \frac{\Delta_{i+1}}{\Delta x} , & x > 0 . \end{cases} \end{aligned} \right\} \quad (13.47)$$

The solution is

$$v(x, t) \equiv u_x(x, t) = \begin{cases} \frac{\Delta_i}{\Delta x} , & \frac{x}{t} < a , \\ \frac{\Delta_{i+1}}{\Delta x} , & \frac{x}{t} > a . \end{cases} \quad (13.48)$$

Therefore, the GRP flux for the linear advection equation (13.1) is

$$f_{i+\frac{1}{2}}^{GRP} = \begin{cases} a[u_i^n + \frac{1}{2}(1-c)\Delta_i] , & a > 0 , \\ a[u_{i+1}^n - \frac{1}{2}(1+c)\Delta_{i+1}] , & a < 0 . \end{cases} \quad (13.49)$$

The GRP flux is identical to those of the MUSCL Hancock Method (13.35) and of Colella’s PLM scheme (13.43). Therefore all the schemes of the form (13.5) with coefficients (13.37)–(13.38) are reproduced for any value of the slope parameter  $\omega$  in (13.27). Third order accuracy is achieved by the special value (13.39).

*Remark 13.30.* The modification of the GRP method to include solutions of Riemann problems for spatial derivatives has potential for constructing higher-order GRP-type schemes. Some preliminary results on a third-order scheme were reported by former MSc student Cáceres [84]. The approach also has potential for constructing schemes for convection-diffusion PDEs, where a viscous flux component involves space derivatives. Some preliminary results were reported by former MSc student Cheney [104], who derived schemes with enlarged stability regions.

One extension of the GRP scheme to the Euler equations is found in the original paper of Ben-Artzi and Falcovitz [37]. Hillier [250] has successfully applied the GRP scheme to study shock diffraction problems; see also [38] and [36]. The modified GRP scheme is applied in [517] to construct non-conservative and conservative schemes for non-linear systems.

### 13.4.5 Slope-Limiter Centred (SLIC) Schemes

Recall that one aspect of the MUSCL approach allows the construction of high-order extensions of the Godunov first-order upwind method via reconstruction of the data. Here we extend the approach for constructing high-order versions of any low-order scheme with numerical flux

$$f_{i+\frac{1}{2}}^{LO} = f_{i+\frac{1}{2}}^{LO}(u_L, u_R)$$

that depends on the two states  $u_L$  and  $u_R$  on the left and right of the interface respectively. For the linear advection equation (13.1), Table 13.2 shows four possible choices for the low-order flux, including the Godunov first-order upwind method. In this section we are interested in low-order schemes that avoid the solution of the Riemann problem, and thus possible choices for the low-order scheme are the Godunov first-order centred, the Lax-Friedrichs and the FORCE schemes, all of which extend to non-linear systems. As we are fundamentally interested in oscillation-free versions of these high-order schemes, we construct Slope-Limiter versions of these centred schemes; we therefore adopt the name of SLIC methods [538]. We describe the SLIC schemes in terms of the scalar conservation law (13.10) as solved by the explicit conservative formula (13.11). The schemes have three steps, the first two being identical to those of the MUSCL-Hancock method of Sect. 13.4.2. The steps are

Step (I) Data reconstruction as in (13.23) with boundary extrapolated values as in (13.24), namely

$$u_i^L = u_i^n - \frac{1}{2}\Delta_i ; \quad u_i^R = u_i^n + \frac{1}{2}\Delta_i . \quad (13.50)$$

Step (II) Evolution of  $u_i^L, u_i^R$  by a time  $\frac{1}{2}\Delta t$  according to

$$\left. \begin{aligned} \bar{u}_i^L &= u_i^L + \frac{1}{2} \frac{\Delta t}{\Delta x} [f(u_i^L) - f(u_i^R)] , \\ \bar{u}_i^R &= u_i^R + \frac{1}{2} \frac{\Delta t}{\Delta x} [f(u_i^L) - f(u_i^R)] . \end{aligned} \right\} \quad (13.51)$$

Step (III) Now, instead of solving the Riemann problem with data  $(\bar{u}_i^R, \bar{u}_{i+1}^L)$  to find the Godunov first–order upwind flux, we compute a low–order flux with data arguments  $(\bar{u}_i^R, \bar{u}_{i+1}^L)$  and thus we have

$$f_{i+\frac{1}{2}}^{slice} = f_{i+\frac{1}{2}}^{LO}(\bar{u}_i^R, \bar{u}_{i+1}^L) . \quad (13.52)$$

One possible choice for the low–order flux is the FORCE flux. Recall that for data  $(u_L, u_R)$  at the cell interface  $i + \frac{1}{2}$  the FORCE flux, see Sect. 7.4.2 of Chap. 7, is

$$f_{i+\frac{1}{2}}^{force}(u_L, u_R) = \frac{1}{2} [f_{i+\frac{1}{2}}^{RI}(u_L, u_R) + f_{i+\frac{1}{2}}^{LF}(u_L, u_R)] , \quad (13.53)$$

where  $f_{i+\frac{1}{2}}^{RI}$  is the Richtmyer flux and  $f_{i+\frac{1}{2}}^{LF}$  is the Lax–Friedrichs flux.

We consider a general low–order flux

$$f_{i+\frac{1}{2}}^{LO}(u_L, u_R) = \beta_0(au_L) + \beta_1(au_R) \quad (13.54)$$

for the linear advection equation in which the flux function is  $f(u) = au$ . See Table 13.2 for examples of low–order fluxes. Application of (13.54) in (13.52) gives the high–order flux

$$f_{i+\frac{1}{2}}^{slice} = \beta_0 a[u_i^n + \frac{1}{2}(1-c)\Delta_i] + \beta_1 a[u_{i+1}^n - \frac{1}{2}(1+c)\Delta_{i+1}] . \quad (13.55)$$

Compare this with the MUSCL–Hancock flux (13.35). Direct substitution of these fluxes into (13.11) gives a five–point support *centred* scheme based on the MUSCL approach, namely

$$u_i^{n+1} = \sum_{k=-2}^2 b_k u_{i+k}^n \quad (13.56)$$

with coefficients

$$\left. \begin{aligned} b_{-2} &= -\frac{1}{4}\beta_0 c(1-c)(1+\omega) , \\ b_{-1} &= \beta_0 c(1 + \frac{1}{4}(1-c)(1+3\omega)) + \frac{1}{4}\beta_1 c(1+c)(1+\omega) , \\ b_0 &= 1 - \beta_0 c(1 - \frac{1}{4}(1-c)(1-3\omega)) + \beta_1 c(1 - \frac{1}{4}(1+c)(1+3\omega)) , \\ b_1 &= -\beta_1 c(1 + \frac{1}{4}(1+c)(1-3\omega)) - \frac{1}{4}\beta_0 c(1-c)(1-\omega) , \\ b_2 &= \frac{1}{4}\beta_1 c(1+c)(1-\omega) . \end{aligned} \right\} \quad (13.57)$$

**Theorem 13.31.** *Schemes (13.56)–(13.57) are second order accurate in space and time for any value  $\omega$  and for any coefficients  $\beta_0, \beta_1$  that give a consistent low-order flux (13.54). Provided  $\beta_0$  and  $\beta_1$  also give a first order flux, schemes (13.56)–(13.57) are third order accurate in space and time when*

$$\omega = \frac{3c(\beta_0 - \beta_1) - 2c^2 - 1}{3(\beta_0 - \beta_1 - c)}. \quad (13.58)$$

*Proof.* Conventional truncation error analysis gives the desired result. A more direct alternative is to apply Roe's theorem (13.1), [408], formula (13.16).

*Remark 13.32.* Note that (13.58) always gives a well-defined  $\omega$  under the conditions of the theorem because the denominator is only zero when

$$\beta_0 = \frac{1}{2}(1+c) \quad \text{and} \quad \beta_1 = \frac{1}{2}(1-c),$$

which are the coefficients for the second order Lax–Wendroff flux.

**Exercise 13.33.** Consider schemes (13.56)–(13.57). Verify that the following statements are true.

1. If the first order scheme is the Godunov upwind scheme, third order accuracy is achieved for

$$\omega = \frac{1}{3}(2c - \text{sign}(c)).$$

2. If the first order scheme is the Godunov centred scheme, third order accuracy is achieved for

$$\omega = \frac{4c^2 - 1}{3c}.$$

3. If the first order scheme is the FORCE scheme, third order accuracy is achieved for

$$\omega = \frac{1}{3}c.$$

4. If the first order scheme is the Lax–Friedrichs scheme, third order accuracy is achieved for

$$\omega = \frac{2c}{3}.$$

5. If the coefficients  $\beta_0, \beta_1$  are chosen as

$$\beta_0 = \frac{1}{6} \left( \frac{c+1}{c-\omega} \right) (1 + 2c - 3\omega); \quad \beta_1 = \frac{1}{6} \left( \frac{c-1}{c-\omega} \right) (1 - 2c + 3\omega),$$

then the scheme (13.56)–(13.57) is third order accurate  $\forall \omega$ .

**Solution 13.34.** (Left to the reader).

A Total Variation Diminishing (TVD) version of these centred schemes will be constructed in Sect. 13.8. In Chap. 14 we extend the schemes to non-linear systems of conservation laws.

### 13.4.6 Other Approaches

There are two more approaches for constructing numerical methods, which are closely related to the explicit, fully discrete schemes studied previously; these are the semi–discrete methods and implicit methods. Both families of methods admit first–order schemes as special cases. Here we give a brief description.

### 13.4.7 Semi–Discrete Schemes

Consider the scalar conservation law (13.10). In the semi–discrete approach, or *method of lines*, one separates the space and time discretisation processes. First one assumes some discretisation in space, while leaving the problem continuous in time. This results in an Ordinary Differential Equation (ODE) in time, namely

$$\frac{du}{dt} = \frac{1}{\Delta x} (f_{i-\frac{1}{2}} - f_{i+\frac{1}{2}}). \quad (13.59)$$

Here

$$f_{i+\frac{1}{2}} = f_{i+\frac{1}{2}}(\{u_i(t)\}, t) \quad (13.60)$$

is an intercell numerical flux, an approximation to the true flux  $f(u(x_{i+\frac{1}{2}}, t))$  at the interface  $i + \frac{1}{2}$  at time  $t$ .

A possible choice for  $f_{i+\frac{1}{2}}$  in (13.59) is the Godunov first–order upwind flux. A second–order space discretisation may be obtained by first applying the MUSCL reconstruction step (13.23) followed by the solution of the Riemann problem (13.42), where the initial data are the right and left boundary extrapolated values (13.24). Higher–order reconstructions lead to corresponding high–order space discretisations.

The time discretisation in (13.10) results from solving the ODE (13.59). There are many techniques for solving ODEs numerically; see Refs. [296] and [251]. The simplest method for solving Ordinary Differential Equations is the first–order Euler method, which if applied to (13.59) leads to the explicit conservative formula (13.11). Runge–Kutta methods are particularly popular for solving ODEs, in the context of the semi–discrete approach. See Jameson et. al. [268] and Gottlieb and Shu [223]. Related works based on the method of lines are those of Berzins [55], [56].

The separation of the time and space discretisation processes in the semi–discrete approach allows enormous flexibility and is well suited for deriving very high–order schemes, such as the families of UNO and ENO schemes. See [245], [242], [240], [241].

### 13.4.8 Implicit Methods

Natural extensions of the second–order fully discrete schemes based on (13.11) to solve (13.10) are conservative schemes

$$u_i^{n+1} = u_i^n + \frac{\Delta t}{\Delta x} (f_{i-\frac{1}{2}} - f_{i+\frac{1}{2}}), \quad (13.61)$$

where the intercell numerical flux  $f_{i+\frac{1}{2}}$  depends on both data values  $\{u_i^n\}$  and unknown values  $\{u_i^{n+1}\}$ , that is

$$f_{i+\frac{1}{2}} = f_{i+\frac{1}{2}}(u_{i-r+1}^n, \dots, u_{i+r}^n; u_{i-s+1}^{n+1}, \dots, u_{i+s}^{n+1}), \quad (13.62)$$

with  $r$  and  $s$  two non-negative integers.

See Harten [239] for approaches for constructing Total Variation Diminishing (TVD) implicit schemes. A most useful reference on implicit TVD methods is the VKI Lecture Notes by Helen Yee [593]; see also [595].

Implicit methods involve the solution of systems of algebraic equations at each time step, which is an expensive process on both CPU time and storage. The advantage of implicit schemes over explicit schemes is that the choice of the time step  $\Delta t$  is, at least in theory, not restricted by stability considerations. Implicit schemes are particularly well suited for solving steady state problems by marching in time. Steady Aerodynamics is perhaps the most prominent area of application of implicit schemes [268], [595].

Several ways of constructing high-order methods have been presented. While high order of accuracy is a very desirable feature of numerical methods, we shall see that, if unmodified, these schemes produce *spurious oscillations* in the vicinity of high gradients of the solution, see Fig. 13.10. The next three sections are devoted to study ways of eliminating the unwanted spurious oscillations while preserving high order of accuracy in the smooth parts of the solution.

## 13.5 Monotone Schemes and Accuracy

Here we study the class of monotone methods and their relation to accuracy. We prove the important theorem due to Godunov that asserts that *monotone, linear schemes are at most first-order accurate*.

### 13.5.1 Monotone Schemes

A useful class of methods for non-linear scalar conservation laws (13.10) are those which are *monotone*. These were defined in Chap. 5 but we recall the definition here.

**Definition 13.35 (Monotone Schemes).** *A scheme*

$$u_i^{n+1} = H(u_{i-k_L+1}^n, \dots, u_{i+k_R}^n), \quad (13.63)$$

*with  $k_L$  and  $k_R$  two non-negative integers, is said to be monotone if*

$$\frac{\partial H}{\partial u_j^n} \geq 0 \quad \forall j. \quad (13.64)$$

*That is,  $H$  is a non-decreasing function of each of its arguments.*

This definition of a monotone scheme is actually equivalent to the following property:

$$\text{if } v_i^n \geq u_i^n \forall i \text{ then } v_i^{n+1} \geq u_i^{n+1}. \quad (13.65)$$

This property in turn is the discrete version of the following property of the exact solution of the conservation law (13.10): if two initial data functions  $v_0(x)$  and  $u_0(x)$  for (13.10) satisfy  $v_0(x) \geq u_0(x) \forall x$ , then their corresponding solutions  $v(x, t)$  and  $u(x, t)$  satisfy  $v(x, t) \geq u(x, t)$ ,  $t > 0$ . Hence *monotone schemes mimic a basic property of exact solutions of conservation laws* (13.10). See [238] and [308].

Next we state a useful property of monotone schemes when applied to the scalar non-linear conservation law (13.10).

**Theorem 13.36.** *Given the data set  $\{u_i^n\}$ , if the solution set  $\{u_i^{n+1}\}$  is obtained with a monotone method (13.63) then*

$$\max_i\{u_i^{n+1}\} \leq \max_i\{u_i^n\}; \quad \min_i\{u_i^{n+1}\} \geq \min_i\{u_i^n\}. \quad (13.66)$$

*Proof.* Define  $v_i^n = \max_j\{u_j^n\} = \text{constant } \forall i$ . Then application of (13.63) gives  $v_i^{n+1} = v_i^n$ . As  $v_i^n \geq u_i^n$ , application of (13.65) gives  $v_i^{n+1} = v_i^n \geq u_i^{n+1}$  and therefore  $\max_i\{u_i^{n+1}\} \leq \max_i\{u_i^n\}$ . The second inequality in (13.66) follows analogously.

An obvious consequence of (13.66) is

$$\left. \begin{aligned} \max_i\{u^n\} &\leq \max_i\{u_i^{n-1}\} \leq \cdots \leq \max_i\{u_i^0\}, \\ \min_i\{u^n\} &\geq \min_i\{u_i^{n-1}\} \geq \cdots \geq \min_i\{u_i^0\}. \end{aligned} \right\} \quad (13.67)$$

This result says that *no new extrema are created* and thus spurious oscillations do not appear; see Fig. 13.10.

In numerical solutions computed with monotone methods, minima increase and maxima decrease as time evolves. This results in *clipping of extrema*, which is in fact a disadvantage of monotone methods. See Figs. 5.10 and 5.11 of Chap. 5, which show results for two monotone methods, namely the Lax–Friedrichs method and the Godunov first-order upwind method.

Another very useful consequence of (13.66) is that solutions of monotone schemes satisfy

$$\min_j\{u_j^n\} \leq u_i^{n+1} \leq \max_j\{u_j^n\}. \quad (13.68)$$

This follows from noting that

$$\min_j\{u_j^{n+1}\} \leq u_i^{n+1} \leq \max_j\{u_j^{n+1}\}$$

and direct application of (13.66). Condition (13.68) says that the solution at any point  $i$  is bounded by the minimum and maximum of the data.

The following result applies to all three-point schemes for non-linear scalar conservation laws (13.10).

**Theorem 13.37 (Monotonicity and the Flux).** *If a three-point scheme of the form*

$$u_i^{n+1} = u_i^n + \frac{\Delta t}{\Delta x} [f_{i-\frac{1}{2}} - f_{i+\frac{1}{2}}] \quad (13.69)$$

for the non-linear conservation law (13.10) is monotone then

$$\frac{\partial}{\partial u_i^n} f_{i+\frac{1}{2}}(u_i^n, u_{i+1}^n) \geq 0 \text{ and } \frac{\partial}{\partial u_{i+1}^n} f_{i+\frac{1}{2}}(u_i^n, u_{i+1}^n) \leq 0. \quad (13.70)$$

That is, in a monotone scheme the numerical flux  $f_{i+\frac{1}{2}}(u_i^n, u_{i+1}^n)$  is an increasing (meaning non-decreasing) function of its first argument and a decreasing (meaning non-increasing) function of its second argument.

*Proof.* In (13.69) we define

$$H(u_{i-1}^n, u_i^n, u_{i+1}^n) \equiv u_i^n + \frac{\Delta t}{\Delta x} [f_{i-\frac{1}{2}}(u_{i-1}^n, u_i^n) - f_{i+\frac{1}{2}}(u_i^n, u_{i+1}^n)].$$

The requirement  $\frac{\partial H}{\partial u_{i-1}^n} \geq 0$  implies  $\frac{\partial f_{i-\frac{1}{2}}}{\partial u_{i-1}^n}(u_{i-1}^n, u_i^n) \geq 0$ , while the requirement  $\frac{\partial H}{\partial u_{i+1}^n} \geq 0$  implies  $\frac{\partial f_{i+\frac{1}{2}}}{\partial u_{i+1}^n}(u_i^n, u_{i+1}^n) \leq 0$ , and the result follows.

**Example 13.38 (The Lax–Friedrichs Scheme).** For a general non-linear conservation law (13.10) the Lax–Friedrichs flux is

$$f_{i+\frac{1}{2}}^{LF}(u_i^n, u_{i+1}^n) = \frac{1}{2}[f(u_i^n) + f(u_{i+1}^n)] + \frac{1}{2} \frac{\Delta x}{\Delta t} (u_i^n - u_{i+1}^n).$$

Application of conditions (13.70) to the Lax–Friedrichs flux shows that monotonicity is ensured provided

$$-1 \leq \frac{\Delta t \lambda_{max}(u)}{\Delta x} \leq 1,$$

where  $\lambda(u) = \partial f / \partial u$  is the characteristic speed and  $\lambda_{max}(u)$  is the maximum. That is, provided the CFL stability condition is enforced properly, the Lax–Friedrichs method is monotone, when applied to non-linear scalar conservation laws (13.10).

**Exercise 13.39.** Derive conditions for the FORCE flux

$$f_{i+\frac{1}{2}}^{force} = \frac{1}{2}[f_{i+\frac{1}{2}}^{LF} + f_{i+\frac{1}{2}}^{RI}],$$

where  $f_{i+\frac{1}{2}}^{RI}$  is the Richtmyer flux, see Sect. 7.4.2 of Chap. 7, to give a monotone scheme, when applied to the inviscid Burgers equation

$$u_t + f(u)_x = 0, \quad f(u) = \frac{1}{2}u^2.$$

**Solution 13.40.** Left to the reader.

Next we state a result for linear schemes (13.5) to solve the linear advection equation (13.1).

**Theorem 13.41 (Positivity of Coefficients).** *A scheme*

$$u_i^{n+1} = \sum_{k=-k_L}^{k_R} b_k u_{i+k}^n \quad (13.71)$$

*for the linear advection equation (13.1) is monotone if and only if all coefficients  $b_k$  are non-negative, that is  $b_k \geq 0 \forall k$ .*

*Proof.* By defining

$$H \equiv \sum_{k=-k_L}^{k_R} b_k u_{i+k}^n$$

and using (13.64) the result  $b_k \geq 0$  follows immediately.

*Example 13.42.* Note that the first-order upwind scheme (see Table 13.1) is always monotone, whereas the Lax–Wendroff method is not monotone. As a matter of fact, none of the second-order accurate schemes, whose coefficients  $b_k$  are listed in Table 13.1, is monotone.

**Exercise 13.43.** Show that the Godunov first-order centred method [217], see Table 13.1, is non-monotone for the range of Courant numbers  $0 \leq |c| \leq \frac{1}{2}$  and is monotone for the range  $\frac{1}{2} \leq |c| \leq \frac{1}{2}\sqrt{2}$ .

**Solution 13.44.** Left to the reader.

### 13.5.2 A Motivating Example

Suppose we want to solve the linear advection equation (13.1) with a conservative scheme (13.11) in conjunction with a WAF-type flux

$$f_{i+\frac{1}{2}}^\alpha = \frac{1}{\Delta x} \int_{-\frac{1}{2}\Delta x}^{\frac{1}{2}\Delta x} f[u_{i+\frac{1}{2}}(x, \alpha\Delta t)] dx, \quad (13.72)$$

where  $\alpha \geq 0$  is a dimensionless parameter and  $u_{i+\frac{1}{2}}(x, t)$  is the solution of the Riemann problem with data  $(u_i^n, u_{i+1}^n)$ . Compare (13.72) with (13.13) and (13.19). In fact  $\alpha = \frac{1}{2}$  reproduces the WAF flux (13.16) that leads to the Lax–Wendroff method; see Sect. 13.3.1. Evaluation of (13.72), see Fig. 13.3, for the linear advection equation gives

$$f_{i+\frac{1}{2}}^\alpha = \frac{1}{2}(1 + 2\alpha c)(au_i^n) + \frac{1}{2}(1 - 2\alpha c)(au_{i+1}^n). \quad (13.73)$$

Appropriate choices for  $\alpha$  give familiar schemes. For instance,

$$\alpha = \begin{cases} \frac{1}{2} & : \text{Lax-Wendroff method;} \\ \frac{1}{2|c|} & : \text{Godunov first-order upwind method;} \\ 1 & : \text{Godunov first-order centred method;} \\ \frac{1}{4c^2}(1+c^2) & : \text{FORCE scheme;} \\ \frac{1}{2c^2} & : \text{Lax-Friedrichs method.} \end{cases} \quad (13.74)$$

Substitution of (13.73) into (13.11) gives

$$u_i^{n+1} = b_{-1}u_{i-1}^n + b_0u_i^n + b_1u_{i+1}^n, \quad (13.75)$$

with coefficients

$$b_{-1} = \frac{1}{2}(1+2\alpha c)c; \quad b_0 = 1 - 2\alpha c^2; \quad b_1 = -\frac{1}{2}(1-2\alpha c)c. \quad (13.76)$$

A stability analysis of the scheme (13.75)–(13.76) using (13.7) shows that the scheme is stable for the following values of the parameter  $\alpha$ :

$$\frac{1}{2} \leq \alpha \leq \frac{1}{2c^2}. \quad (13.77)$$

The flux (13.73) is a weighted average of the upwind and downwind fluxes, with  $\alpha$  controlling the relative size of their contributions. For  $\alpha > \frac{1}{2}$  the upwind flux contribution is increased and the downwind contribution is decreased. The minimum upwind contribution consistent with a stable scheme is  $\alpha = \frac{1}{2}$ , the Lax-Wendroff method. Also note that the upper limit for stability is  $\alpha = \frac{1}{2c^2}$ ; that is, the maximum upwind contribution results in the Lax-Friedrichs method. Any attempt at increasing the upwind contribution further results in an unstable method.

An accuracy analysis of (13.75)–(13.76) using (13.6) says that the scheme is second-order accurate in space and time for the single value  $\alpha = \frac{1}{2}$  (the Lax-Wendroff method), and for any value of  $\alpha$  with  $\frac{1}{2} < \alpha \leq \frac{1}{2c^2}$  the scheme is first-order accurate. As soon as the upwind contribution in the Lax-Wendroff scheme is slightly increased, the accuracy of the scheme drops to first order.

According to Theorem (13.41), monotonicity of the scheme (13.75)–(13.76) holds if the coefficients (13.76) are non-negative.

**Exercise 13.45 (Monotonicity).** Verify that scheme (13.75)–(13.76) is monotone for values of  $\alpha$  in the range

$$\frac{1}{2|c|} \leq \alpha \leq \frac{1}{2c^2}. \quad (13.78)$$

**Solution 13.46.** Left to the reader.

An important observation follows from (13.77) and (13.78): for the range of values

$$\frac{1}{2} < \alpha \leq \frac{1}{2|c|} \quad (13.79)$$

of the parameter  $\alpha$  the scheme (13.75) is stable, first-order accurate, but *not monotone*.

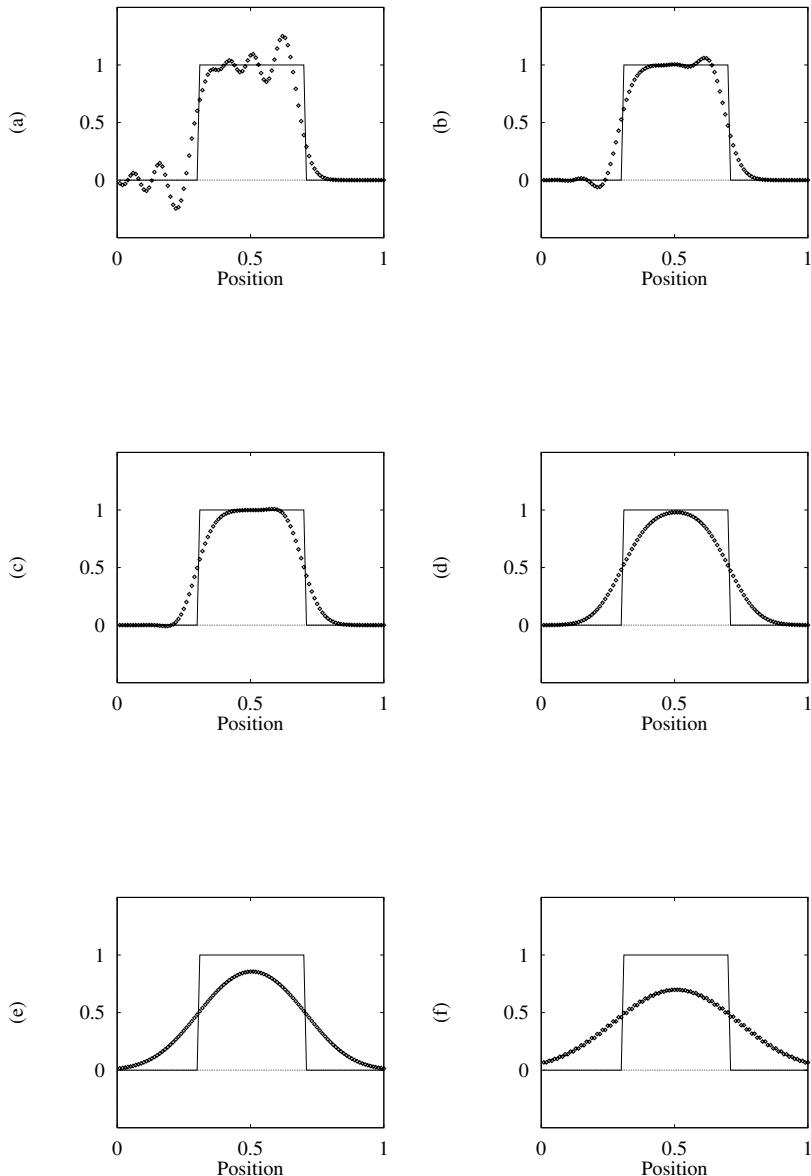
*Remark 13.47.* There are infinitely many first-order accurate schemes that are stable but not monotone. In (13.75) we have one example for each value of  $\alpha$  in the range given by (13.79).

**Exercise 13.48.** On the  $c$ - $\alpha$  plane, identify the regions of stability, instability and monotonicity of scheme (13.75)–(13.76). Note that the Godunov first-order centred scheme is stable for  $0 \leq |c| \leq \frac{1}{2}\sqrt{2}$  and is monotone only in the range  $\frac{1}{2} \leq |c| \leq \frac{1}{2}\sqrt{2}$ .

**Solution 13.49.** Left to the reader.

Fig. 13.10 shows numerical results (symbols) compared with the exact solution (line), for the linear advection equation (13.1) with a square wave as initial condition; see Test 2, Sect. 5.5.1 of Chap. 5. The computational domain is  $[0, 10]$  which is discretised with  $M = 1000$  cells; the CFL number used is  $c = \frac{1}{4}$  and the output time is  $t = 1.0$ . The schemes used result from the conservative formula (13.11) and the numerical flux (13.73), with six fixed values for the parameter  $\alpha$ ; see legend to figure. A relevant observation to make is that the second-order Lax–Wendroff method (a) produces high resolution of the discontinuities (narrow transition zone) but also produces spurious oscillations in their vicinity; the result labelled (b) is some first-order accurate method resulting from the choice  $\alpha = \frac{3}{4}$  and the result confirms the fact that the scheme is not monotone, as predicted by (13.79). This first-order accurate scheme is not monotone, as remarked earlier. The result labelled (c) corresponds to the Godunov first-order centred scheme, which for the CFL number used is not monotone, in agreement with condition (13.79). The result (d) is oscillation-free and corresponds to a scheme with a larger value of  $\alpha$  ( $\alpha = 2$ ) so that the Godunov first-order upwind method is reproduced. Larger values of  $\alpha$  produce other schemes; (e) corresponds to the First–Order Centred (FORCE) scheme ( $\alpha = 4.25$ ) and (f) corresponds to the Lax–Friedrichs method ( $\alpha = 8$ ), both giving oscillation-free results.

*Remark 13.50. Monotonicity and Accuracy.* Our empirical observation based on the numerical results of Chap. 5 and those of Fig. 13.10 is that schemes that are second order accurate produce spurious oscillations, and schemes that are monotone are inaccurate. The observation is in fact true in general, for fixed coefficient (linear) schemes, as we shall prove in the next section. A



**Fig. 13.10.** Numerical results at time  $t = 1$  for six numerical schemes contained in the motivating example of Sect. 13.5.2; symbols denote numerical solution and lines denote the exact solution; CFL number used is  $c = \frac{1}{4}$ . (a):  $\alpha = \frac{1}{2}$  (Lax–Wendroff); (b):  $\alpha = \frac{3}{4}$  (Some first–order non–monotone method); (c):  $\alpha = 1$  (Godunov first–order centred, non–monotone); (d):  $\alpha = 2$  (Godunov first–order upwind); (e):  $\alpha = 4.25$  (First–Order Centred, FORCE); (f):  $\alpha = 8$  (Lax–Friedrichs)

possible way forward is to vary the parameter  $\alpha$  so as to maintain accuracy in smooth parts of the solution and monotonicity near high gradients. This would require  $\alpha$  to be a function of the data set  $\{u_i^n\}$ , which would then lead to non-linear schemes; see Sect. 13.7.1.

### 13.5.3 Monotone Schemes and Godunov's Theorem

Our empirical observations so far reveal that *linear* second-order accurate schemes (13.5) to solve the linear advection equation (13.1) are distinctly better than first-order methods for problems with smooth solutions; see computational results of Figs. 5.9 to 5.11 of Chap. 5. However, for solutions involving high gradients, such as near a discontinuity, these methods produce spurious oscillations; see Figs. 5.12 to 5.14 of Chap. 5 and Fig. 13.10. On the other hand, monotone methods avoid spurious oscillations near high gradients; see results of Figs. 5.12 to 5.14 of Chap. 5. Their disadvantage though, seems to be their limited accuracy, which is clearly exposed in the results of Figs. 5.9 to 5.11 of Chap. 5.

Godunov's theorem establishes, theoretically, that the desirable properties of accuracy and monotonicity are, for linear schemes, contradictory requirements. The following result applies to the linear advection equation (13.1) and linear schemes (13.5).

**Theorem 13.51 (Godunov's Theorem).** *There are no monotone, linear schemes (13.5) for (13.1) of second or higher order of accuracy.*

*Proof.* The proof given here is based on the accuracy relation (13.6). Denote by  $s_q$  the summation

$$s_q = \sum_{k=-k_L}^{k_R} k^q b_k , \quad (13.80)$$

where  $b_k$  are the coefficients (constant) of the (linear) scheme (13.5). For second-order accuracy one requires

$$s_0 = 1 , \quad s_1 = -c , \quad s_2 = c^2 . \quad (13.81)$$

From definition (13.80)

$$\begin{aligned} s_2 &= \sum_{k=-k_L}^{k_R} k^2 b_k \\ &= \sum_{k=-k_L}^{k_R} (k + c)^2 b_k - 2c \sum_{k=-k_L}^{k_R} k b_k - c^2 \sum_{k=-k_L}^{k_R} b_k \\ &= \left[ \sum_{k=-k_L}^{k_R} (k + c)^2 b_k \right] - 2cs_1 - c^2 s_0 . \end{aligned} \quad (13.82)$$

Use of (13.81) into (13.82) gives

$$\left[ \sum_{k=-k_L}^{k_R} (k+c)^2 b_k \right] + c^2 \geq c^2 . \quad (13.83)$$

The above inequality holds, as a monotone linear scheme satisfies  $b_k \geq 0$ . Equality in (13.83), and thus second order accuracy, is only possible if  $b_k = 0 \forall k$  or when  $c = -k_0$ , that is for integer Courant numbers, and  $b_k = 0 \forall k \neq k_0$ . Thus we have proved the theorem for schemes satisfying the condition  $0 \leq |c| \leq 1$ .

The case of integer Courant numbers larger than unity is only of theoretical interest, as for non-linear systems this is an impossible requirement to impose on numerical methods.

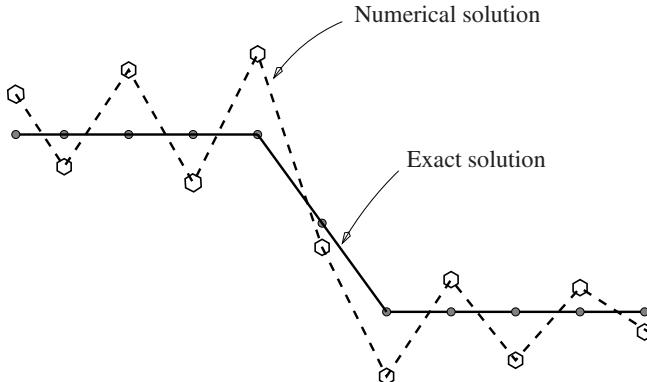
*Remark 13.52. Consequences of Godunov's Theorem.* Another way to express Godunov's theorem is that monotone schemes are at most first-order accurate. First-order methods are too inaccurate to be of practical interest. We therefore must search for other classes of schemes that, ideally, allow for both the oscillation-free property of monotone schemes and the accuracy of high order methods to coexist. In other words we must look for ways of circumventing Godunov's theorem. The key to this lies on the assumption made in the theorem that the schemes have fixed coefficients (linear schemes).

### 13.5.4 Spurious Oscillations and High Resolution

The problem of spurious oscillations in the vicinity of high gradients is depicted in the sketch of Fig. 13.11, where the full line denotes the exact solution and the dotted line denotes the numerical solution obtained by some linear method of second or higher order of accuracy. Different methods will produce different patterns for the oscillatory profile. The Lax-Wendroff method will produce spurious oscillations *behind* the wave, whereas the Warming-Beam method will produce spurious oscillations *ahead* of the wave; see Fig. 5.12 of Chap. 5. This is related to the form of the leading term in the local truncation error of the method.

The data  $u_i^n$  and the exact solution are monotone decreasing functions but the numerical solution is *not*; new local extrema have been introduced, thus violating a fundamental property of solutions of scalar conservation laws. See Sect. 13.6.1 and the statements preceding (13.94). See also conditions (13.94).

The aim of the rest of this chapter is to provide the necessary theoretical background and construct numerical methods that avoid the problem of spurious oscillations, sometimes called the Gibb phenomena, while retaining accuracy. In fact we aim at constructing numerical methods that have the following properties



**Fig. 13.11.** Illustration of the numerical phenomenon of spurious oscillations near high gradients.

- (i) The schemes have second or higher order of accuracy in smooth parts of the solution
- (ii) The schemes produce numerical solutions free from spurious oscillations
- (iii) The schemes produce *high-resolution* of discontinuities, that is the number of mesh points in the transition zone containing the numerical wave is narrow in comparison with that of first-order monotone methods.

Schemes satisfying the above properties are called *high-resolution methods*, after Harten [238].

### 13.5.5 Data Compatibility

One of the first reported attempts at providing a rational approach to circumventing Godunov's theorem is due to Roe [408]. The central idea is to construct adaptive algorithms that would adjust themselves to the local nature of the solution. This results, obviously, in variable coefficient (non-linear) schemes, even when applied to linear PDEs. The coefficients must be functions of the data.

The following definition relates an algorithm for the linear equation (13.1) with particular classes of data.

**Definition 13.53 (Data Compatible Algorithm).** A scheme is compatible with a given data set  $\{u_i^n\}$  if the solution  $u_i^{n+1}$  at each point  $i$ , as given by the algorithm, is bounded by the upwind pair  $(u_{i-s}^n, u_i^n)$ , where  $s \equiv \text{sign}(c) = \text{sign}(a)$ .

*Remark 13.54.* For a given point  $i$  the upwind pair is  $(u_{i-1}^n, u_i^n)$  for  $c > 0$  and  $(u_i^n, u_{i+1}^n)$  for  $c < 0$ . The characteristic through  $(i, n + 1)$  of the PDE (13.1) intersects the real line at a point between the upwind pair.

We may therefore rephrase the above definition as follows: a scheme is compatible with the data if  $u_i^{n+1}$  lies between  $u_{i-s}^n$  and  $u_i^n$ . The data compatibility requirement of an algorithm is conveniently expressed as follows:

$$\min\{u_{i-s}^n, u_i^n\} \leq u_i^{n+1} \leq \max\{u_{i-s}^n, u_i^n\}. \quad (13.84)$$

Compare (13.84) with (13.68) derived from Theorem (13.36).

**Proposition 13.55 (Data Compatibility).** *The data compatibility condition (13.84) is equivalent to requiring*

$$0 \leq \frac{u_i^{n+1} - u_i^n}{u_{i-s}^n - u_i^n} \leq 1. \quad (13.85)$$

*Proof.* First assume  $u_{i-s}^n \leq u_i^n$ . Imposing (13.84) gives  $u_{i-s}^n \leq u_i^{n+1} \leq u_i^n$ , that is  $u_{i-s}^n - u_i^n \leq u_i^{n+1} - u_i^n \leq 0$ . Division through by  $u_{i-s}^n - u_i^n < 0$  gives (13.85). The case  $u_{i-s}^n \geq u_i^n$  follows similarly.

*Example 13.56.* The flux for the Godunov first-order upwind method to solve (13.1) is

$$f_{i+\frac{1}{2}} = \frac{1}{2}(1+s)(au_i^n) + \frac{1}{2}(1-s)(au_{i+1}^n).$$

Substitution of  $f_{i+\frac{1}{2}}$  and  $f_{i-\frac{1}{2}}$  into (13.11) gives

$$u_i^{n+1} = u_i^n + sc(u_{i-s}^n - u_i^n). \quad (13.86)$$

The data compatibility condition (13.85) for (13.86) is

$$0 \leq \frac{u_i^{n+1} - u_i^n}{u_{i-s}^n - u_i^n} = s|c| \leq 1,$$

which is *always* satisfied. That is, the Godunov first-order upwind method is compatible with all possible sets of data. This is not a property enjoyed by the Lax-Wendroff method, for instance.

Next we consider data compatibility conditions for any three point scheme

$$u_i^{n+1} = b_{-1}u_{i-1}^n + b_0u_i^n + b_1u_{i+1}^n \quad (13.87)$$

applied to the linear advection equation (13.1). We first note that (13.87) can be written in *incremental form* as

$$u_i^{n+1} = u_i^n - b_{-1}\Delta u_{i-\frac{1}{2}} + b_1\Delta u_{i+\frac{1}{2}}, \quad (13.88)$$

where as usual  $\Delta u_{i-\frac{1}{2}} = u_i^n - u_{i-1}^n$  and  $\Delta u_{i+\frac{1}{2}} = u_{i+1}^n - u_i^n$ .

It is easily verified that any three-point scheme (13.87) or (13.88) for (13.1) has data compatibility conditions

$$\begin{aligned} 0 \leq b_{-1} - b_1/r_i &\leq 1, \text{ if } c > 0, \\ 0 \leq b_1 - b_{-1}r_i &\leq 1, \text{ if } c < 0, \end{aligned} \quad (13.89)$$

where

$$r_i = \frac{\Delta u_{i-\frac{1}{2}}}{\Delta u_{i+\frac{1}{2}}} . \quad (13.90)$$

*Example 13.57 (The Lax–Wendroff Method).*

Manipulations of the previous result (13.89), (13.90) show that the Lax–Wendroff method is *not* compatible with all possible sets of data, it is compatible only with data satisfying

$$\left. \begin{aligned} -\frac{1+c}{1-c} &\leq 1/r_i \leq \frac{c+2}{c}, & \text{if } 0 \leq c \leq 1, \\ -\frac{1-c}{1+c} &\leq r_i \leq \frac{c-2}{c}, & \text{if } -1 \leq c \leq 0. \end{aligned} \right\}$$

*Example 13.58 (The Warming and Beam Method).* Roe [408] showed that the Warming–Beam Method is *not* compatible with all possible sets of data, it is compatible only with data satisfying

$$\left. \begin{aligned} \frac{c-2}{c} &\leq r_i \leq \frac{c-3}{c-1}, & \text{if } 0 \leq c \leq 1, \\ \frac{c}{c-2} &\leq 1/r_i \leq \frac{c+1}{c-1}, & \text{if } 1 \leq c \leq 2 \end{aligned} \right\}$$

and

$$\left. \begin{aligned} \frac{c+2}{c} &\leq 1/r_i \leq \frac{c+3}{c+1}, & \text{if } -1 \leq c \leq 0, \\ \frac{c+2}{c} &\leq r_i \leq \frac{c-1}{c}, & \text{if } -2 \leq c \leq -1. \end{aligned} \right\}$$

Roe [408] also observed that given the set  $S_{DC}$  of all possible data states one can select combinations of schemes such that the subsets for which the schemes satisfy the data compatibility requirement do cover the full set  $S_{DC}$ ; in this way he constructed *adaptive, non-linear schemes that are monotone and second order accurate*, thus circumventing Godunov’s Theorem.

More recent applications of the data compatibility concept impose directly the data compatibility requirement (13.85) on schemes that depend on some free parameter to allow for variable coefficients. In Sect. 13.7.1 we illustrate this approach to modify the scheme (13.11) with flux (13.73) so as to avoid spurious oscillations and maintain second-order accuracy in smooth parts of the solution.

## 13.6 Total Variation Diminishing ( TVD) Methods

The real issue concerning numerical methods is convergence. For non-linear systems convergence proofs rely on *non-linear stability*, the theory of

which relies on functional analysis concepts, such as *compactness*. Sets of functions whose total variation is bounded lead to compact sets. Total Variation Stable methods are then defined as those whose mesh-dependent approximations lie in compact function sets. It can then be proved that Total Variation Stable methods are convergent. See Harten [238] and LeVeque [308] for details.

A subclass of Total Variation Stable methods are those whose total variation does not increase in time; these are commonly referred to as *Total Variation Diminishing* (TVD) methods, or Total Variation Non-Increasing (TVNI) methods. See Harten [238], [239] and Sweby [470]. Here we are interested in studying and designing TVD methods.

### 13.6.1 The Total Variation

Given a function  $u = u(x)$ , the total variation of  $u$  is defined as

$$TV(u) = \lim_{\delta \rightarrow 0} \sup \frac{1}{\delta} \int_{-\infty}^{\infty} |u(x + \delta) - u(x)| dx . \quad (13.91)$$

See Apostol [15] for definitions of supremum ( $\sup$ ) and related concepts of Real Analysis. If  $u(x)$  is smooth then (13.91) is identical to

$$TV(u) = \int_{-\infty}^{\infty} |u'(x)| dx . \quad (13.92)$$

In fact, even for discontinuous  $u(x)$ , (13.92) is still correct in the sense of Distribution Theory, see Smoller [451], Chap. 7. If  $u = u(x, t)$  then one can generalise definitions (13.91) or (13.92), but in fact for convergence purposes, it suffices to define the total variation of  $u(x, t)$  at fixed times  $t = t^n$ , which we denote by  $TV(u(t))$ . Moreover, if  $u^n = \{u_i^n\}$  is a *mesh function*, then the total variation of  $u^n$  is defined as

$$TV(u^n) = \sum_{i=-\infty}^{\infty} |u_{i+1}^n - u_i^n| . \quad (13.93)$$

Obviously, in order for  $TV(u^n)$  to be finite, one must assume  $u_i^n = 0$  or  $u_i^n = \text{constant}$  as  $i \rightarrow \pm\infty$ . For a moment, let us just consider the results of Fig. 13.10 again, where computed results at time  $t = 1$  are shown. Clearly, for the Lax–Wendroff result (a) the total variation of the numerical solution is larger than that of the exact solution; highly oscillatory solutions have large total variation. Of the numerical solutions shown, that of the Lax–Friedrichs method (f) has the smallest total variation. It is easy to see that as time evolves the total variation of the Lax–Friedrichs solution will decrease; the total variation of the exact solution of the linear equation (13.1) will remain constant.

A fundamental property of the exact solution of the IVP for the non-linear scalar conservation law (13.10), when the initial data  $u(x, 0)$  has bounded total variation, is

- (i) No new local extrema in  $x$  may be created
- (ii) The value of a local minimum increases (it does not decrease) and the value of a local maximum decreases (it does not increase).

From this it follows that the Total Variation  $TV(u(t))$  is a decreasing function of time (non-increasing); that is,

$$TV(u(t_2)) \leq TV(u(t_1)) \quad \forall t_2 \geq t_1. \quad (13.94)$$

See Harten [238] and Lax [301]. This property of the exact solution is the one that we want to mimic when designing numerical methods.

### 13.6.2 TVD and Monotonicity Preserving Schemes

Consider a numerical scheme of the form

$$u_i^{n+1} = H(u_{i-k_L+1}^n, \dots, u_{i+k_R}^n), \quad (13.95)$$

with  $k_L$  and  $k_R$  two non-negative integers, to solve the scalar conservation law (13.10). Motivated by property (13.94) we introduce the following definition

**Definition 13.59 (TVD Schemes).** *Scheme (13.95) is said to be a Total Variation Diminishing (TVD) scheme, or Total Variation Non-Increasing (TVNI) scheme, if*

$$TV(u^{n+1}) \leq TV(u^n), \quad \forall n. \quad (13.96)$$

An obvious consequence of the above definition is that

$$TV(u^n) \leq TV(u^{n-1}) \leq \dots \leq TV(u^0), \quad (13.97)$$

where  $\{u_i^0\}$  is data at time  $t = 0$ . Next we define another class of numerical methods.

**Definition 13.60.** *Schemes of the form (13.95) for the scalar, nonlinear conservation law (13.10) are said to be Monotonicity Preserving Schemes if whenever the data  $\{u_i^n\}$  is monotone the solution set  $\{u_i^{n+1}\}$  is monotone in the same sense. That is, if  $\{u_i^n\}$  is monotone increasing so is  $\{u_i^{n+1}\}$  and if  $\{u_i^n\}$  is monotone decreasing so is  $\{u_i^{n+1}\}$ .*

An important result that relates monotone, TVD and monotonicity preserving schemes for non-linear scalar conservation laws (13.10) is given by the following theorem, the proof of which we omit here.

**Theorem 13.61.** *In general, the set  $S_{mon}$  of monotone schemes is contained in the set  $S_{tvd}$  of TVD schemes and this in turn is contained in the set  $S_{mpr}$  of monotonicity preserving schemes, that is*

$$S_{mon} \subseteq S_{tvd} \subseteq S_{mpr}. \quad (13.98)$$

*Proof.* See Harten [238]; see also [308].

For linear schemes (13.5) to solve the linear advection equation (13.1) one can prove that monotone schemes are equivalent to monotonicity preserving schemes, as the following theorem states.

**Theorem 13.62.** *A linear scheme (13.5) as applied to the linear advection equation (13.1) is Monotonicity Preserving if and only if the coefficients  $b_k$  are non-negative, i.e.  $b_k \geq 0 \forall k$ .*

*Proof.* In the first part of the proof, we assume the scheme is monotonicity preserving and show that  $b_k \geq 0$ . Applying the scheme (13.5) to two consecutive points leads to

$$u_{i+1}^{n+1} - u_i^{n+1} = \sum_{k=-k_L}^{k_R} b_k (u_{i+1+k}^n - u_{i+k}^n). \quad (13.99)$$

In proving this part by contradiction we assume  $b_K < 0$  for some  $K$ . The objective now is to contradict the hypothesis that the scheme is Monotonicity Preserving. This is achieved by considering the special case

$$u_j^n = \begin{cases} 0 & j < i + 1 + K, \\ 1 & j \geq i + 1 + K. \end{cases} \quad (13.100)$$

Thus

$$u_{i+1}^{n+1} - u_i^{n+1} = b_K < 0.$$

But if  $\{u_i^n\}$  is monotone increasing for instance, then by hypothesis, we must have  $u_{i+1}^{n+1} - u_i^{n+1} \geq 0$ . Therefore  $b_k \geq 0, \forall k$ .

In the second part of the proof we assume  $b_k \geq 0$  in (13.5) and prove that the scheme is Monotonicity Preserving. If  $\{u_i^n\}$  is monotone then all terms in the summation (13.99) are of the same sign, as  $b_k \geq 0$ ; consequently the left hand side of (13.99) is of the same sign and thus  $\{u_i^{n+1}\}$  is monotone in the same sense as  $\{u_i^n\}$ .

*Remark 13.63.* Recall that a monotone linear scheme (13.5) to solve (13.1) is monotone if and only if  $b_k \geq 0$ . Hence, in this case, monotone schemes are identical to monotonicity preserving schemes and by (13.98) they are also equivalent to TVD schemes.

*Remark 13.64.* The most useful class is that of TVD methods, which have some precise mathematical properties that allow proofs of convergence. For an example of convergence proofs see for instance the work of Sweby and Baines [472] and relevant references cited there. The TVD conditions also allow the practical construction of numerical methods having the TVD property.

Harten [238] considered the class of non-linear schemes

$$u_i^{n+1} = u_i^n - C_{i-\frac{1}{2}} \Delta u_{i-\frac{1}{2}} + D_{i+\frac{1}{2}} \Delta u_{i+\frac{1}{2}}, \quad (13.101)$$

where  $\Delta u_{i+\frac{1}{2}} = u_{i+1}^n - u_i^n$  and the coefficients  $C_{i-\frac{1}{2}}, D_{i+\frac{1}{2}}$  are in general assumed to be functions of the data. Harten [238] proved the following important result

**Theorem 13.65 (Harten).** *For any scheme of the form (13.101) to solve (13.10), a sufficient condition for the scheme to be TVD is that the coefficients satisfy*

$$\left. \begin{aligned} C_{i+\frac{1}{2}} &\geq 0; & D_{i+\frac{1}{2}} &\geq 0, \\ 0 &\leq C_{i+\frac{1}{2}} + D_{i+\frac{1}{2}} && \leq 1. \end{aligned} \right\} \quad (13.102)$$

*Proof.* Apply the scheme (13.101) to two consecutive cells  $i$  and  $i+1$ . We obtain

$$u_i^{n+1} = u_i^n - C_{i-\frac{1}{2}}(u_i^n - u_{i-1}^n) + D_{i+\frac{1}{2}}(u_{i+1}^n - u_i^n) \quad (13.103)$$

and

$$u_{i+1}^{n+1} = u_{i+1}^n - C_{i+\frac{1}{2}}(u_{i+1}^n - u_i^n) + D_{i+\frac{3}{2}}(u_{i+2}^n - u_{i+1}^n). \quad (13.104)$$

Subtracting (13.103) from (13.104) gives

$$u_{i+1}^{n+1} - u_i^{n+1} = (u_{i+1}^n - u_i^n)(1 - C_{i+\frac{1}{2}} - D_{i+\frac{1}{2}})$$

$$+ C_{i-\frac{1}{2}}(u_i^n - u_{i-1}^n) + D_{i+\frac{3}{2}}(u_{i+2}^n - u_{i+1}^n).$$

Taking absolute values on both sides of the above inequality leads to

$$|u_{i+1}^{n+1} - u_i^{n+1}| \leq |1 - C_{i+\frac{1}{2}} - D_{i+\frac{1}{2}}| |u_{i+1}^n - u_i^n|$$

$$+ |C_{i-\frac{1}{2}}| |u_i^n - u_{i-1}^n| + |D_{i+\frac{3}{2}}| |u_{i+2}^n - u_{i+1}^n|.$$

As  $|r| = r$  for any non-negative real number  $r$ , application of the sufficient conditions (13.102) gives

$$|u_{i+1}^{n+1} - u_i^{n+1}| \leq |u_{i+1}^n - u_i^n| + C_{i-\frac{1}{2}} |u_i^n - u_{i-1}^n| - C_{i+\frac{1}{2}} |u_{i+1}^n - u_i^n|$$

$$+ D_{i+\frac{3}{2}} |u_{i+2}^n - u_{i+1}^n| - D_{i+\frac{1}{2}} |u_{i+1}^n - u_i^n|.$$

Summing over  $i$  gives

$$\begin{aligned} \sum_i |u_{i+1}^{n+1} - u_i^{n+1}| &\leq \sum_i |u_{i+1}^n - u_i^n| + \sum_i C_{i-\frac{1}{2}} |u_i^n - u_{i-1}^n| \\ &\quad - \sum_i C_{i+\frac{1}{2}} |u_{i+1}^n - u_i^n| + \sum_i D_{i+\frac{3}{2}} |u_{i+2}^n - u_{i+1}^n| \\ &\quad - \sum_i D_{i+\frac{1}{2}} |u_{i+1}^n - u_i^n|. \end{aligned}$$

Clearly the second and third summations on the right-hand side of the above inequality cancel and so do the fourth and fifth summations, leading to the sought result

$$TV(u^{n+1}) \leq TV(u^n).$$

*Remark 13.66.* During the proof of the above theorem we assumed

$$C_{i-\frac{1}{2}} \geq 0, \quad 1 - C_{i+\frac{1}{2}} - D_{i+\frac{1}{2}} \geq 0, \quad D_{i+\frac{3}{2}} \geq 0.$$

By appropriate shifting of the position at which expressions (13.103), (13.104) are applied, the above conditions are equivalent to the conditions (13.102) of the theorem.

*Remark 13.67.* The coefficients  $C_{i+\frac{1}{2}}$ ,  $D_{i+\frac{1}{2}}$  in Harten's theorem, may in general, be data dependent. The theorem therefore applies to nonlinear schemes. This fact can be then used to circumvent Godunov's theorem (13.51), which applies to linear schemes only. Harten's theorem (13.65) offers a very useful tool for constructing high resolution schemes.

*Remark 13.68.* Schemes that allow a controlled increase in the total variation have also been constructed. They are usually referred to as Total Variation Bounded Schemes, or TVB schemes. See for instance the work of Shu [446], [445].

## 13.7 Flux Limiter Methods

In this section we construct Total Variation Diminishing (TVD) methods following the *flux limiter approach*, see Sweby [470]. In Sect. 13.7.1 we first consider a particular method, namely a TVD version of the Weighted Average Flux (WAF) method described in Sect. 13.3. We then present the general flux limiter approach as applied to families of numerical methods; in Sect. 13.7.2 we present *upwind* based flux limiter methods and in Sect. 13.7.4 we present  *(non-upwind) flux limiter methods.*

### 13.7.1 TVD Version of the WAF Method

Recall the motivating example of Sect. 13.5.2 in which the conservative scheme (13.11) has numerical flux (13.73), which depends on a free parameter  $\alpha$ . Various, fixed values of  $\alpha$  reproduce familiar schemes. One may exploit  $\alpha$  by regarding it as a function of data  $u_i^n$  and constrained by some TVD requirement to construct non-linear oscillation-free schemes of second order accuracy in smooth parts of the solution. This is in fact the approach we follow here to construct a TVD version of the basic WAF scheme presented in Sect. 13.3.1 to solve the model linear advection equation (13.1).

For convenience we introduce a new parameter  $\phi$  and re-write the WAF flux (13.16) as

$$f_{i+\frac{1}{2}} = \frac{1}{2}(1 + \phi)(au_i^n) + \frac{1}{2}(1 - \phi)(au_{i+1}^n). \quad (13.105)$$

Compare with (13.73). Three particular choices of  $\phi$  are  $\phi = |c|$  (which leads to the Lax–Wendroff method),  $\phi = s \equiv \text{sign}(a) = \text{sign}(c)$  (which produces the Godunov first-order upwind method) and  $\phi = -s$  (which gives the downwind method, unconditionally unstable).

The purpose is to find appropriate ranges for  $\phi$  as a function of some data-dependent variables that produce a TVD version of the WAF scheme. Based on the extreme cases  $\phi = s$  and  $\phi = -s$  above we restrict  $\phi$  to satisfy the following requirement

$$-1 \leq \phi \leq 1. \quad (13.106)$$

Other choices for the upper and lower bounds of  $\phi$  are of course also possible. By introducing appropriate notation for  $\phi$  at the intercell boundaries the relevant intercell fluxes become

$$\left. \begin{aligned} f_{i+\frac{1}{2}} &= \frac{1}{2}(1 + \phi_{i+\frac{1}{2}})(au_i^n) + \frac{1}{2}(1 - \phi_{i+\frac{1}{2}})(au_{i+1}^n), \\ f_{i-\frac{1}{2}} &= \frac{1}{2}(1 + \phi_{i-\frac{1}{2}})(au_{i-1}^n) + \frac{1}{2}(1 - \phi_{i-\frac{1}{2}})(au_i^n). \end{aligned} \right\} \quad (13.107)$$

Thus, for each intercell boundary position  $i + \frac{1}{2}$  we seek a value  $\phi_{i+\frac{1}{2}}$  that will produce a modified WAF flux leading to a TVD scheme.

## Application of Data Compatibility

Instead of using Harten’s Theorem we impose the slightly stronger constraint embodied in the data compatibility condition (13.85). First we assume that  $a > 0$  in the PDE (13.1). Substitution of the fluxes (13.107) into the conservative formula (13.11) followed by some obvious algebraic manipulations, to reproduce the ratio in (13.85), give

$$\frac{u_i^{n+1} - u_i^n}{u_{i-1}^n - u_i^n} = \frac{1}{2}c \left[ \frac{1}{r_{i+\frac{1}{2}}} \left( 1 - \phi_{i+\frac{1}{2}} \right) + \phi_{i-\frac{1}{2}} + 1 \right], \quad (13.108)$$

where

$$r_{i+\frac{1}{2}} = \frac{u_i^n - u_{i-1}^n}{u_{i+1}^n - u_i^n}. \quad (13.109)$$

Compare the left-hand side of (13.108) with the ratio of consecutive data changes in the data compatibility condition (13.85). Direct application of condition (13.85) to the right-hand side of (13.108) leads to

$$-1 \leq \frac{1}{r_{i+\frac{1}{2}}} \left( 1 - \phi_{i+\frac{1}{2}} \right) + \phi_{i-\frac{1}{2}} \leq \frac{2 - c}{c}. \quad (13.110)$$

For negative speed  $a$  in the PDE (13.1) ( $c < 0$ ) one arrives at the same result (13.110), with the CFL number  $c$  replaced by its absolute value and  $r_{i+\frac{1}{2}}$  defined as

$$r_{i+\frac{1}{2}} = \frac{u_{i+2}^n - u_{i+1}^n}{u_{i+1}^n - u_i^n}. \quad (13.111)$$

Thus, for both positive and negative waves speeds in (13.1) we obtain

$$-1 \leq \frac{1}{r_{i+\frac{1}{2}}} (1 - \phi_{i+\frac{1}{2}}) + \phi_{i-\frac{1}{2}} \leq \frac{2 - |c|}{|c|}, \quad (13.112)$$

with

$$r_{i+\frac{1}{2}} = \begin{cases} \frac{u_i^n - u_{i-1}^n}{u_{i+1}^n - u_i^n} & , \quad a > 0, \\ \frac{u_{i+2}^n - u_{i+1}^n}{u_{i+1}^n - u_i^n} & , \quad a < 0. \end{cases} \quad (13.113)$$

Thus  $r_{i+\frac{1}{2}}$  is always the ratio of the *upwind change*,  $\Delta u_{i-\frac{1}{2}} \equiv u_i^n - u_{i-1}^n$  for  $a > 0$  or  $\Delta u_{i+\frac{3}{2}} \equiv u_{i+2}^n - u_{i+1}^n$  for  $a < 0$ , to the *local change*  $\Delta u_{i+\frac{1}{2}} = u_{i+1}^n - u_i^n$ . That is

$$r_{i+\frac{1}{2}} = \frac{\Delta_{upw}}{\Delta_{loc}},$$

with the obvious definitions for  $\Delta_{upw}$  and  $\Delta_{loc}$ . Fig. 13.12. illustrates the identification of local and upwind changes across appropriate waves. The ratio  $r_{i+\frac{1}{2}}$  of local changes in the data is now regarded as a *flow parameter* that will cause the value of the sought function  $\phi$  to adjust to local conditions on the data. So far we have identified two quantities that may be regarded as the independent variables of the function  $\phi$ , namely  $|c|$  and the flow parameter  $r_{i+\frac{1}{2}}$ . Thus we define

$$\phi_{i+\frac{1}{2}} = \phi_{i+\frac{1}{2}}(r_{i+\frac{1}{2}}, |c|). \quad (13.114)$$

Occasionally, we shall omit the argument  $|c|$ .

### Construction of the TVD Region

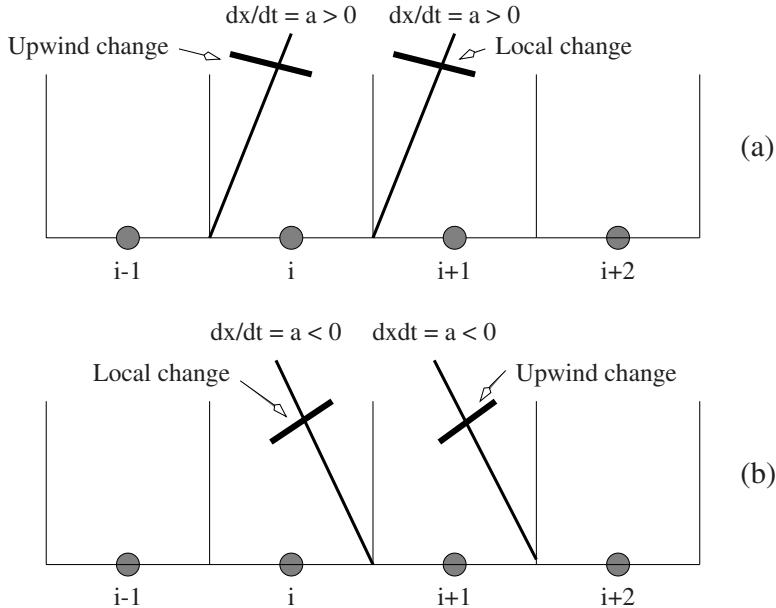
The purpose here is to find a suitable range of values for  $\phi_{i+\frac{1}{2}}$  as a function of  $r_{i+\frac{1}{2}}$  and  $|c|$ . To this end we select two inequalities, one for  $\phi_{i+\frac{1}{2}}$  and one for  $\phi_{i-\frac{1}{2}}$  so that both inequalities in (13.112) are automatically satisfied. We take

$$-1 - L \leq \frac{1}{r_{i+\frac{1}{2}}} (1 - \phi_{i+\frac{1}{2}}) \leq \frac{2(1 - |c|)}{|c|} \quad (13.115)$$

and

$$L \leq \phi_{i-\frac{1}{2}} \leq 1, \quad (13.116)$$

with the lower bound  $L$  in the interval  $[-1, |c|]$ . This will allow some freedom in selecting how much *downwinding* is to be allowed. For  $L = -1$ , see (13.112),



**Fig. 13.12.** Identification of upwind direction to form the ratio  $r_{i+\frac{1}{2}}$  of upwind change to local change: (a) positive wave speed, upwind change lies on the left (b) negative wave speed, upwind change lies on the right

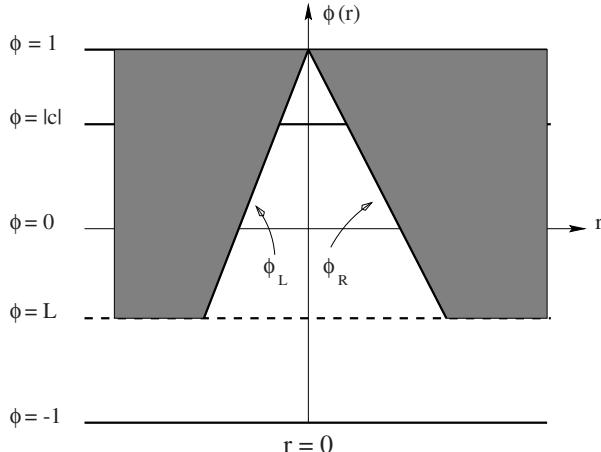
full downwinding is allowed; for  $L = 0$  no downwinding is allowed. Note that by adding (13.115) and (13.116) we reproduce (13.112). We study (13.115) in detail. Subscripts are omitted, for convenience.

The *left inequality* in (13.115):  $-(1 + L) \leq \frac{1}{r}(1 - \phi)$  contains two cases, namely

- Positive  $r$ : if  $r > 0$  then  $-(1 + L)r \leq 1 - \phi$ , which implies
 
$$\phi \leq 1 + (1 + L)r \equiv \phi_L(r, |c|) . \quad (13.117)$$
- Negative  $r$ : if  $r < 0$  then  $-(1 + L)r \geq 1 - \phi$ , which implies
 
$$\phi \geq 1 + (1 + L)r \equiv \phi_L(r, |c|) . \quad (13.118)$$

The *right inequality* in (13.115):  $\frac{1}{r}(1 - \phi) \leq \frac{2(1 - |c|)}{|c|}$  has two cases, namely

- Positive  $r$ : if  $r > 0$  we obtain
 
$$\phi \geq 1 - \frac{2(1 - |c|)}{|c|}r \equiv \phi_R(r, |c|) . \quad (13.119)$$
- Negative  $r$ : if  $r < 0$  we obtain
 
$$\phi \leq 1 - \frac{2(1 - |c|)}{|c|}r \equiv \phi_R(r, |c|) . \quad (13.120)$$



**Fig. 13.13.** TVD regions for the WAF method are shown by shaded area in the  $r-\phi$  plane, for fixed values of the CFL number  $c$

Fig. 13.13 illustrates the derived TVD regions (dark zone) in the plane  $r - \phi$  that satisfies conditions (13.115)–(13.116). There is a TVD region for  $r < 0$  and a TVD region for  $r > 0$ . The choice of the lower bound  $L$  determines the left TVD region; for  $L = -1$  this region coalesces to the single line  $\phi_R(r, |c|) = 1$  (Godunov first-order upwind). Any fixed value of the limiter function  $\phi(r, |c|)$  results in a scheme with a numerical viscosity  $\alpha_\phi$ , see Sect. 5.2 of Chap. 5. For  $|c| \leq \phi(r, |c|) \leq 1$  this numerical viscosity is positive and results in *spreading of discontinuities* and *clipping of extrema*. For  $L \leq \phi(r, |c|) \leq |c|$  the numerical viscosity is negative and results in steepening of discontinuities. Obviously the case  $\phi(r, |c|) = |c|$  does not add any extra numerical viscosity and corresponds to the Lax–Wendroff method.

A *limiter function*  $\phi = \phi(r, |c|)$  can be constructed within the TVD region and any choice will produce an oscillation-free scheme (13.11) with flux (13.105); the only restriction is

$$\phi(1, |c|) = |c| . \quad (13.121)$$

This ensures second-order accuracy for values of  $r$  close to 1, that is when the upwind change is comparable to the local change (smooth part of the solution), see (13.113).

### Construction of WAF Limiter Functions

There is scope for the imagination in constructing functions  $\phi(r, |c|)$ . Here we give five WAF limiter functions, namely

$$\phi_{ua}(r, |c|) = \begin{cases} 1 & \text{if } r \leq 0 , \\ 1 - \frac{2(1-|c|)r}{|c|} & \text{if } 0 \leq r \leq \frac{|c|}{1-|c|} , \\ -1 & \text{if } r \geq \frac{|c|}{1-|c|} , \end{cases} \quad (13.122)$$

$$\phi_{sa}(r, |c|) = \begin{cases} 1 & \text{if } r \leq 0 , \\ 1 - 2(1 - |c|)r & \text{if } 0 \leq r \leq \frac{1}{2} , \\ |c| & \text{if } \frac{1}{2} \leq r \leq 1 , \\ 1 - (1 - |c|)r & \text{if } 1 \leq r \leq 2 , \\ 2|c| - 1 & \text{if } r \geq 2 \end{cases} \quad (13.123)$$

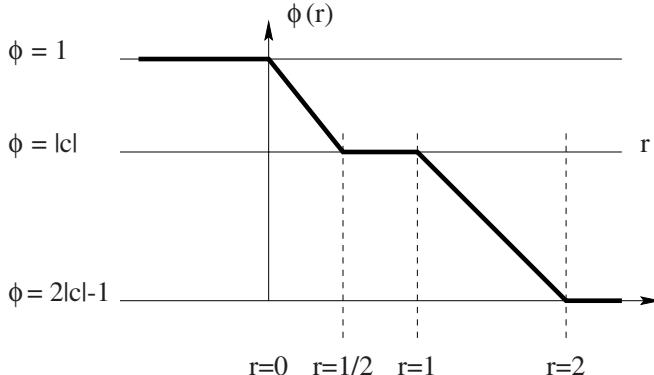
$$\phi_{vl}(r, |c|) = \begin{cases} 1 & \text{if } r \leq 0 , \\ 1 - \frac{2(1-|c|)r}{1+r} & \text{if } r \geq 0 , \end{cases} \quad (13.124)$$

$$\phi_{va}(r, |c|) = \begin{cases} 1 & \text{if } r \leq 0 , \\ 1 - \frac{(1-|c|)r(1+r)}{1+r^2} & \text{if } r \geq 0 , \end{cases} \quad (13.125)$$

and

$$\phi_{ma}(r, |c|) = \begin{cases} 1 & \text{if } r \leq 0 , \\ 1 - (1 - |c|)r & \text{if } 0 \leq r \leq 1 , \\ |c| & \text{if } r \geq 1 . \end{cases} \quad (13.126)$$

Each of these WAF limiter functions above corresponds to the well known flux limiters ULTRABEE, SUPERBEE, VANLEER, VANALBADA and MINBEE, to be studied in Sect. 13.7.2. We therefore give the WAF limiters analogous names. The function  $\phi_{ua}(r, |c|)$  (ULTRAA) corresponds to the flux limiter ULTRABEE of Roe [409];  $\phi_{sa}(r, |c|)$  (SUPERA) corresponds to the flux limiter SUPERBEE of Roe [409];  $\phi_{vl}(r, |c|)$  (VANLEER) corresponds to the flux limiter of van Leer [554], [555];  $\phi_{va}(r, |c|)$  (VANALBADA) corresponds to the limiter proposed by van Albada [552];  $\phi_{ma}(r, |c|)$  (MINA) corresponds to the flux limiter MINBEE of Roe [409]. Fig. 13.14 illustrates the limiter function SUPERA. Note that  $\phi_{ua}(r, |c|)$  does not satisfy the second-order requirement (13.121).



**Fig. 13.14.** WAF limiter function SUPER-A (corresponding to flux limiter SUPER-BEE) is a piece-wise linear function shown by thick line

### Numerical Experiments

We consider the same two test problems for the linear advection (13.1) as in Chap. 5. In the computations we take  $a = 1.0$  and a CFL coefficient  $C_{\text{cfl}} = 0.8$ . Computed results are shown at the output times  $t = 1.0$  unit (125 time steps) and  $t = 10.0$  unit (1250 time steps). In each figure we compare the exact solution (shown by full lines) with the numerical solution (symbols).

The initial conditions for Test 1 (smooth data) and Test 2 (discontinuous data, square wave) are respectively given by

$$u(x, 0) = \alpha e^{-\beta x^2} ; \quad u(x, 0) = \begin{cases} 0 & \text{if } x \leq 0.3 , \\ 1 & \text{if } 0.3 \leq x \leq 0.7 , \\ 0 & \text{if } x \geq 0.7 . \end{cases} \quad (13.127)$$

In the computations for Test 1 we take  $\alpha = 1.0$ ,  $\beta = 8.0$ ; the initial profile  $u(x, 0)$  is evaluated in the interval  $-1 \leq x \leq 1$ . The mesh used is  $\Delta x = 0.02$ . For Test 2 the mesh used is  $\Delta x = 0.01$ . In both test problems the initial profile and that at the output time is resolved by 100 computing cells.

Results for Test 1 are shown in Figs. 13.15 ( $t=1$ ) and 13.16 ( $t=10$ ). Results for Test 2 are shown in Figs. 13.17 ( $t=1$ ) and 13.18 ( $t=10$ ). In each figure we first show the linear schemes corresponding to  $\phi(r, |c|) = |c| = \text{constant}$  (Lax-Wendroff) and to  $\phi(r, |c|) = 1 = \text{constant}$  (Godunov first order upwind). The other solutions correspond to TVD versions of the WAF method for four choices of the limiter function  $\phi(r, |c|)$ ; these are: ULTRA, SUPER-A, VANLEER and MINA, given respectively by (13.122)–(13.124) and (13.126). In analysing the results of Fig. 13.15 one must bear in mind that these are for a test with smooth initial data and that the solution has been evolved for a short time (125 time steps). The results from the Lax-Wendroff (or WAF without the TVD property) are excellent and no apparent improvements are offered by the TVD version of WAF using any of the limiter functions tested.

The result of the Godunov first order upwind method shows the effects of the *numerical viscosity* inherent in this method; as a consequence extrema are unduly smoothed out. Of the four TVD results shown, that of the limiter ULTRA is the worst of the group. As noted earlier this function does not pass through the point of second-order accuracy; it adds *negative numerical viscosity* (locally) and results in the wrong *steepening* and *squaring* of the solution profiles. Clearly this limiter is unsuitable, at least for smooth solutions. All other TVD results are very satisfactory. SUPER shows a tendency to *squaring* of the solution profile near extrema, similar to ULTRA. This is in fact a well known feature of SUPER. The limiters VANLEER and MINA give similar results, at least for this output time.

Compare the results for Fig. 13.15 with those of Fig. 13.16 ( $t=10$ ). The results of Fig. 13.16 (1250 time steps) begin to show the effect of the dispersive errors in the Lax–Wendroff method, the computed solution lags behind (the speed  $a$  is positive) the true solution. For very long evolution times this error becomes unacceptable. The Godunov first-order upwind method now clearly shows the effects of its numerical viscosity, which increases as the number of time steps increases. As for  $t = 1$ , the TVD results that are acceptable are those produced by SUPER, VANLEER AND MINA. The result of ULTRA is completely unacceptable. The tendency of the limiter SUPER to *square* smooth profiles unduly is now visible. The limiters VANLEER and MINA are beginning to show the effect of their inherent numerical viscosity acting near extrema.

The main conclusion from the results of Figs. 13.15 and 13.16 is that the TVD results of the limiter functions SUPER and VANLEER are the best, a conclusion that is also valid for non-linear systems (see Chap. 14).

In discussing the results of Fig. 13.17 one must bear in mind that the solution contains two discontinuities and that the output time is short (125 time steps). Now, unlike Test 1, the schemes (linear schemes) Lax–Wendroff and Godunov first-order upwind give the worse results. As noted earlier, the Lax–Wendroff scheme resolves steep fronts more sharply but at the cost of introducing *spurious oscillations*. Godunov's first-order upwind method on the other hand, while avoiding the production of spurious oscillations, *smears* sharp fronts excessively. It is in fact the limitations of these (linear) methods that motivates the introduction of TVD methods. These have the best features of both classes of schemes.

Of the four limiter functions tested in Fig. 13.17, ULTRA produces the best results; the corresponding results of Figs. 13.15 and 13.16 are the worst of their group. It is worth noting that at even time steps the discontinuities in the ULTRA profile have one intermediate point and none at odd time steps. The results from the other limiter functions are also satisfactory. Recall that the larger the limiter function is (within the TVD region), the more numerical viscosity is added; thus, of the four TVD results the one given by MINA is the one with the largest numerical viscosity, which explains the smearing of the discontinuities. Compare results with those of Fig. 13.18, which correspond to

1250 time steps; these results clearly expose the limitations of Lax-Wendroff and Godunov first-order upwind method. All TVD results are superior, with ULTRA giving the best result followed by SUPERA, VANLEER and MINA. In practical computations the solutions will contain smooth parts as well as discontinuities. It is found that SUPERA and VANLEER give the best results. Extension of TVD methods to non-linear systems is carried out in Chap. 14.

### 13.7.2 The General Flux-Limiter Approach

A well established approach for constructing high-order Total Variation Diminishing (TVD) schemes is the *flux limiter approach* [70], [69], [71], [470], [409]. This requires a high-order flux  $f_{i+\frac{1}{2}}^{HI}$  associated with a scheme of accuracy greater than or equal to two and a low-order flux  $f_{i+\frac{1}{2}}^{LO}$  associated with a *monotone*, first-order scheme. We present the approach in terms of the model conservation law

$$u_t + f(u)_x = 0 ; \quad f(u) = au \quad (13.128)$$

as solved by

$$u_i^{n+1} = u_i^n + \frac{\Delta t}{\Delta x} [f_{i-\frac{1}{2}} - f_{i+\frac{1}{2}}] . \quad (13.129)$$

One then defines a high-order TVD flux as

$$f_{i+\frac{1}{2}}^{TVD} = f_{i+\frac{1}{2}}^{LO} + \phi_{i+\frac{1}{2}} [f_{i+\frac{1}{2}}^{HI} - f_{i+\frac{1}{2}}^{LO}] , \quad (13.130)$$

where  $\phi_{i+\frac{1}{2}}$  is a flux limiter function yet to be determined. To preserve some generality we assume that  $f_{i+\frac{1}{2}}^{HI}$  and  $f_{i+\frac{1}{2}}^{LO}$  are respectively of the form

$$\left. \begin{aligned} f_{i+\frac{1}{2}}^{LO} &= \alpha_0 au_i^n + \alpha_1 au_{i+1}^n , \\ f_{i+\frac{1}{2}}^{HI} &= \beta_0 au_i^n + \beta_1 au_{i+1}^n . \end{aligned} \right\} \quad (13.131)$$

The choice

$$\alpha_0 = \frac{1}{2}(1+s) , \quad \alpha_1 = \frac{1}{2}(1-s) , \quad (13.132)$$

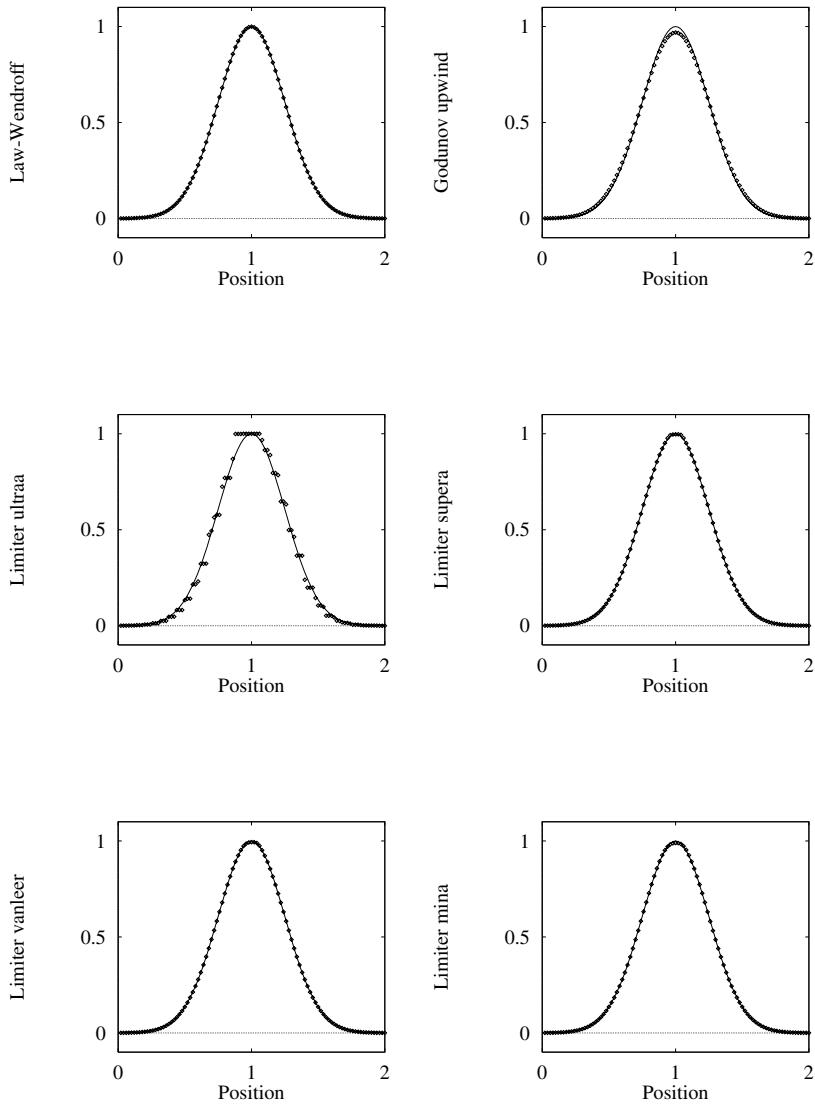
with  $s = \text{sign}(a)$ , reduces  $f_{i+\frac{1}{2}}^{LO}$  to the Godunov first order upwind flux. For the choice

$$\alpha_0 = \frac{1}{4c}(1+c)^2 ; \quad \alpha_1 = -\frac{1}{4c}(1-c)^2 \quad (13.133)$$

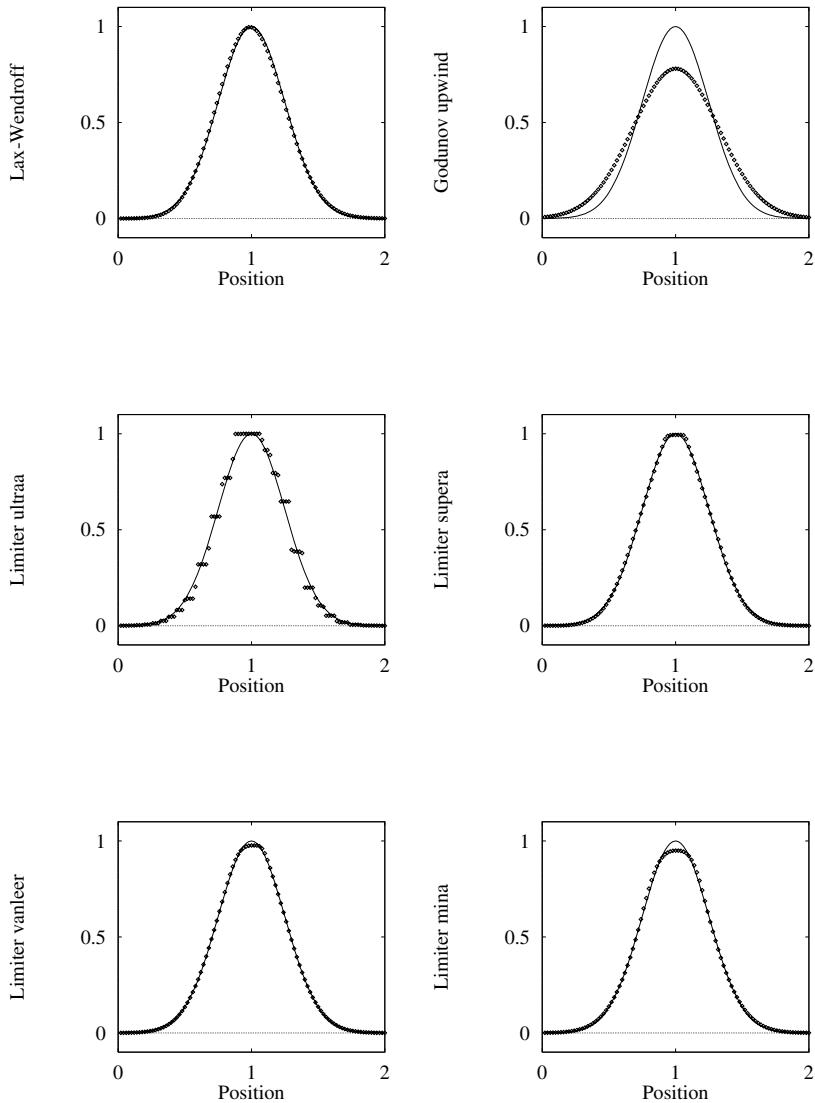
$f_{i+\frac{1}{2}}^{LO}$  becomes the FORCE flux and the choice

$$\alpha_0 = \frac{1}{2c}(1+c) ; \quad \alpha_1 = -\frac{1}{2c}(1-c) \quad (13.134)$$

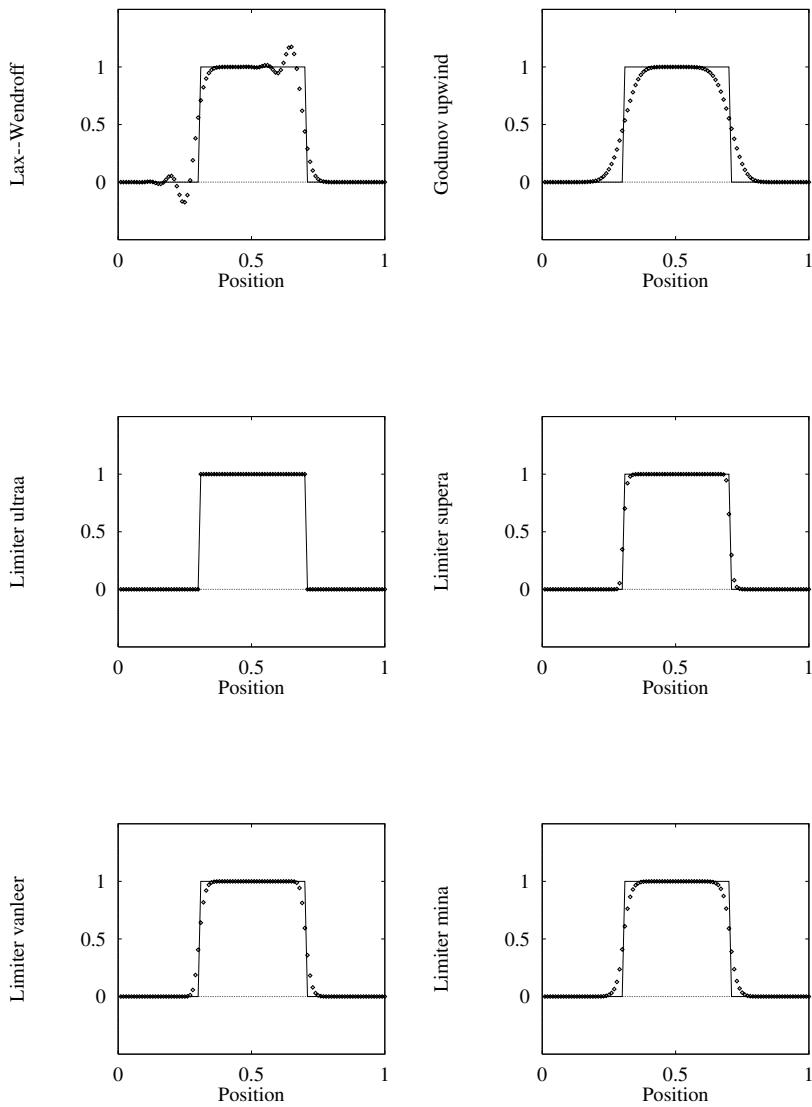
reproduces the Lax-Friedrichs flux; see (13.17) and Table 13.2. The two-point support for the flux  $f_{i+\frac{1}{2}}^{HI}$  means that this flux must be the Lax-Wendroff flux, for which



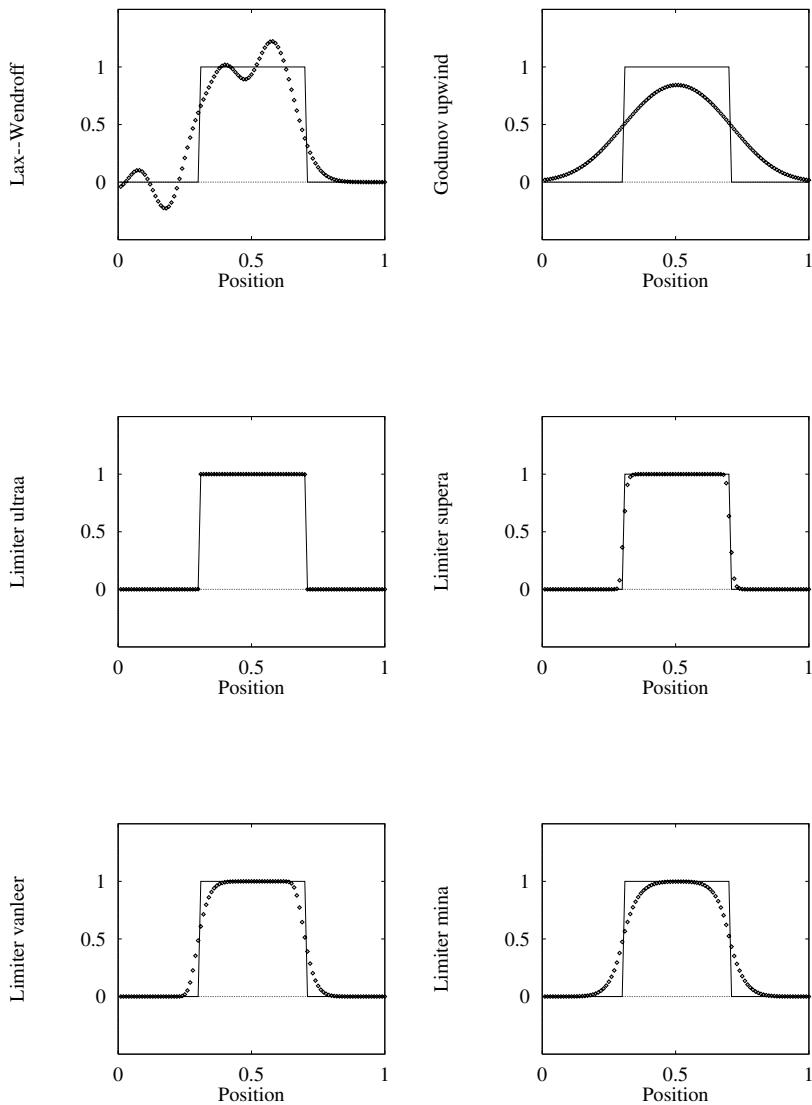
**Fig. 13.15.** Numerical (symbols) and exact (line) solutions for Test 1 at time  $t = 1$ . The results of the linear schemes Lax–Wendroff and Godunov first order upwind are compared with results from TVD version of the WAF method using four limiter functions, namely: ULTRAa, SUPERA, VANLEER and MINA. Compare the results with those of Fig. 13.16.



**Fig. 13.16.** Numerical (symbols) and exact (line) solutions for Test 1 at time  $t = 10$ . The results of the linear schemes Lax–Wendroff and Godunov first order upwind are compared with results from TVD version of the WAF method using four limiter functions, namely: ULTRA, SUPERA, VANLEER and MINA. Compare the results with those of Fig. 13.15.



**Fig. 13.17.** Numerical (symbols) and exact (line) solutions for Test 2 at time  $t = 1$ . The results of the linear schemes Lax–Wendroff and Godunov first order upwind are compared with results from TVD version of the WAF method using four limiter functions, namely: ULTRAa, SUPERA, VANLEER and MINA. Compare the results with those of Fig. 13.18.



**Fig. 13.18.** Numerical (symbols) and exact (line) solutions for Test 2 at time  $t = 10$ . The results of the linear schemes Lax–Wendroff and Godunov first order upwind are compared with results from TVD version of the WAF method using four limiter functions, namely: ULTRA, SUPERA, VANLEER and MINA. Compare the results with those of Fig. 13.17.

$$\beta_0 = \frac{1}{2}(1+c) ; \quad \beta_1 = \frac{1}{2}(1-c) . \quad (13.135)$$

Therefore the coefficients for  $\beta_0$  and  $\beta_1$  for the high-order flux are fixed but the coefficients  $\alpha_0$  and  $\alpha_1$  for the low-order flux are general. Substitution of (13.131) into (13.130) gives

$$f_{i+\frac{1}{2}}^{TVD} = [\alpha_0 + (\beta_0 - \alpha_0)\phi_{i+\frac{1}{2}}](au_i^n) + [\alpha_1 + (\beta_1 - \alpha_1)\phi_{i+\frac{1}{2}}](au_{i+1}^n) , \quad (13.136)$$

which if substituted into (13.129) produces

$$\left. \begin{aligned} u_i^{n+1} &= u_i^n - C\Delta u_{i-\frac{1}{2}} + D\Delta u_{i+\frac{1}{2}} , \\ C &= c[\alpha_0 + (\beta_0 - \alpha_0)\phi_{i-\frac{1}{2}}] , \\ D &= -c[\alpha_1 + (\beta_1 - \alpha_1)\phi_{i+\frac{1}{2}}] , \\ \Delta u_{i-\frac{1}{2}} &= u_i^n - u_{i-1}^n ; \quad \Delta u_{i+\frac{1}{2}} = u_{i+1}^n - u_i^n . \end{aligned} \right\} \quad (13.137)$$

In the next section we specialise the flux limiter approach to particular choices of low and high order schemes.

### 13.7.3 TVD Upwind Flux Limiter Schemes

Here we assume that the low-order flux is that of the Godunov first-order upwind scheme and the high order flux is that of the Lax–Wendroff scheme. The derivation of the TVD scheme relies on identifying *upwind directions*. Flux limiters are now denoted by  $\psi_{i+\frac{1}{2}}(r)$ .

First assume  $a > 0$  in the model conservation law (13.128). Then

$$\left. \begin{aligned} \alpha_0 &= 1 ; \quad \alpha_1 = 0 ; \quad c > 0 , \\ C &= c[1 + (\beta_0 - 1)\psi_{i-\frac{1}{2}}] ; \quad D = -c\beta_1\psi_{i+\frac{1}{2}} , \\ \beta_0 &= \frac{1}{2}(1+c) ; \quad \beta_1 = \frac{1}{2}(1-c) . \end{aligned} \right\} \quad (13.138)$$

Now we rewrite (13.137) as

$$\left. \begin{aligned} u_i^{n+1} &= u_i^n - \hat{C}\Delta u_{i-\frac{1}{2}} , \\ \hat{C} &= C - D/r ; \quad r = \Delta u_{i-\frac{1}{2}}/\Delta u_{i+\frac{1}{2}} . \end{aligned} \right\} \quad (13.139)$$

Application of the TVD condition (13.102) of Harten's theorem gives the inequalities

$$0 \leq c \left[ 1 + (\beta_0 - 1)\psi_{i-\frac{1}{2}} + \beta_1\psi_{i+\frac{1}{2}} \frac{1}{r} \right] \leq 1 . \quad (13.140)$$

We now impose a global constraint, *independent* of  $r$ , on the sought limiter functions, namely

$$\psi_B \leq \psi_{i-\frac{1}{2}} \leq \psi_T , \quad \forall i , \forall r . \quad (13.141)$$

This constraint may be re-written as

$$c[1 + (\beta_0 - 1)\psi_T] \leq c[1 + (\beta_0 - 1)\psi_{i-\frac{1}{2}}] \leq c[1 + (\beta_0 - 1)\psi_B] . \quad (13.142)$$

Now we consider the following two inequalities

$$-c[1 + (\beta_0 - 1)\psi_T] \leq c\beta_1\psi_{i+\frac{1}{2}} \frac{1}{r} \leq 1 - c[1 + (\beta_0 - 1)\psi_B] . \quad (13.143)$$

Note that (13.142) and (13.143) reproduce the TVD condition (13.140) identically. As (13.142) is only a re-statement of (13.141) we only work with inequalities (13.143) to find the limiter function  $\psi_{i+\frac{1}{2}}$  at the intercell boundary  $i + \frac{1}{2}$ .

Analysis of the left inequality in (13.143) leads to

$$\psi_{i+\frac{1}{2}}(r) \begin{cases} \geq \psi_L(r) , & \text{if } r > 0 , \\ \leq \psi_L(r) , & \text{if } r < 0 , \end{cases} \quad (13.144)$$

where

$$\psi_L(r) = (\psi_T - 1/\beta_1)r . \quad (13.145)$$

Analysis of the right inequality in (13.143) produces the constraints

$$\psi_{i+\frac{1}{2}}(r) \begin{cases} \leq \psi_R(r) , & r > 0 , \\ \geq \psi_R(r) , & r < 0 , \end{cases} \quad (13.146)$$

where

$$\psi_R(r) = \left[ \psi_B + \frac{(1-c)}{c\beta_1} \right] r . \quad (13.147)$$

Using constraints (13.141) together with (13.144)–(13.147), one can draw the TVD region for the flux limiter  $\psi_{i+\frac{1}{2}}(r)$  for the case  $a > 0$ .

Before doing that we consider the case  $a < 0$ . Now the upwind direction is on the right hand side of the relevant intercell boundary. The scheme may now be written as

$$\left. \begin{aligned} u_i^{n+1} &= u_i^n + \hat{D}\Delta u_{i+\frac{1}{2}} ; \\ \hat{D} &= D - C/r ; \quad r = \frac{\Delta u_{i+\frac{1}{2}}}{\Delta u_{i-\frac{1}{2}}} . \end{aligned} \right\} \quad (13.148)$$

Now we regard the position  $i - \frac{1}{2}$  as the local position so that  $r$  remains the ratio of upwind to local changes. Harten's TVD conditions (13.102) lead to

$$0 \leq -c \left[ 1 + (\beta_1 - 1)\psi_{i+\frac{1}{2}} + \beta_0\psi_{i-\frac{1}{2}} \frac{1}{r} \right] \leq 1 . \quad (13.149)$$

Now we impose the global constraint

$$\psi_B \leq \psi_{i+\frac{1}{2}} \leq \psi_T . \quad (13.150)$$

Following the same steps as for the case  $a > 0$  we end up with the conditions

$$c[1 + (\beta_1 - 1)\psi_T] \leq -c\beta_0\psi_{i-\frac{1}{2}} \leq 1 + c[1 + (\beta_1 - 1)\psi_B] \quad (13.151)$$

for the flux limiter function  $\psi_{i-\frac{1}{2}}$  at the intercell boundary  $i - \frac{1}{2}$ . Note that by replacing  $c$  with its absolute value  $|c|$  conditions (13.151) are identical to conditions (13.143), with the appropriate interpretation for the *local* intercell position.

Thus conditions (13.141) together with (13.144)–(13.147) apply to both  $a > 0$  and  $a < 0$ , provided we replace  $c$  by  $|c|$  and correctly interpret the ratio  $r$ . Hence, the limiter function at a general intercell position  $i + \frac{1}{2}$  satisfies

$$\psi_B \leq \psi(r) \leq \psi_T , \quad (13.152)$$

$$\psi(r) \begin{cases} \geq \psi_L(r) , & \text{if } r > 0 , \\ \leq \psi_L(r) , & \text{if } r < 0 , \end{cases} \quad (13.153)$$

$$\psi(r) \begin{cases} \leq \psi_R(r) , & \text{if } r > 0 , \\ \geq \psi_R(r) , & \text{if } r < 0 , \end{cases} \quad (13.154)$$

$$\psi_L(r) = \left( \psi_T - \frac{2}{1 - |c|} \right) r , \quad (13.155)$$

$$\psi_R(r) = \left( \psi_B + \frac{2}{|c|} \right) r , \quad (13.156)$$

$$r = \frac{\Delta_{upw}}{\Delta_{loc}} = \begin{cases} \frac{u_i^n - u_{i-1}^n}{u_{i+1}^n - u_i^n} , & a > 0 , \\ \frac{u_{i+2}^n - u_{i+1}^n}{u_{i+1}^n - u_i^n} , & a < 0 . \end{cases} \quad (13.157)$$

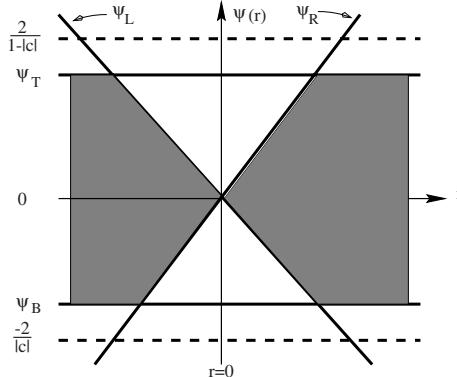
The choice of the bottom and top bounds  $\psi_B$  and  $\psi_T$  in (13.152) determines the functions  $\psi_L$  and  $\psi_R$  in (13.155)–(13.156) and thus the TVD region. Fig. 13.19 shows the resulting TVD region for

$$\psi_B \geq -\frac{2}{|c|} ; \quad \psi_T \leq \frac{2}{1 - |c|} . \quad (13.158)$$

As for the derivation of the TVD regions for the WAF method in Sect. 13.7.1, the choice of the bottom and top boundaries  $\psi_B$  and  $\psi_T$  depends on what schemes one wants to reproduce as *boundary schemes*. The choice

$$\psi_B = 0 ; \quad \psi_T = \frac{2}{1 - |c|} . \quad (13.159)$$

allows for all schemes between the Godunov first-order upwind scheme and the downwind scheme. In this case the TVD region for  $r < 0$  coalesces to the single line  $\psi(r) = 0$  (Godunov first-order upwind for negative  $r$ ).

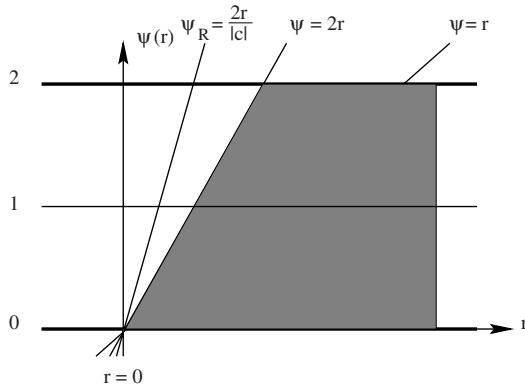


**Fig. 13.19.** General TVD region (dark zone) for flux limiter schemes based on the Godunov first-order upwind method and the Lax–Wendroff scheme. Compare with Fig. 13.13 showing the TVD region for the WAF method

The Sweby TVD region [470] is reproduced if we take

$$\psi_T = 2 \leq \frac{2}{1 - |c|}, \quad \forall c; \quad \psi_B = 0 \quad (13.160)$$

and replace  $2/|c|$  by 2 in  $\psi_R$ . See Fig. 13.20. Five flux limiter functions are



**Fig. 13.20.** Sweby's TVD Region, a special case of that shown in Fig. 13.19. Here the bottom and top boundaries are 0 and 2 respectively, so that the TVD region and limiters are independent of CFL number

the following: ULTRABEE is given by

$$\psi_{ub}(r) = \begin{cases} 0, & r \leq 0, \\ \frac{2}{|c|}r, & 0 \leq r \leq \frac{|c|}{1-|c|}, \\ \frac{2}{1-|c|}, & r \geq \frac{|c|}{1-|c|}, \end{cases} \quad (13.161)$$

SUPERBEE is given by

$$\psi_{sb}(r) = \begin{cases} 0, & r \leq 0, \\ 2r, & 0 \leq r \leq \frac{1}{2}, \\ 1, & \frac{1}{2} \leq r \leq 1, \\ r, & 1 \leq r \leq 2, \\ 2, & r \geq 2, \end{cases} \quad (13.162)$$

VANLEER is given by

$$\psi_{vl}(r) = \begin{cases} 0, & r \leq 0, \\ \frac{2r}{1+r}, & r \geq 0, \end{cases} \quad (13.163)$$

VANALBADA is given by

$$\psi_{va}(r) = \begin{cases} 0, & r \leq 0, \\ \frac{r(1+r)}{1+r^2}, & r \geq 0, \end{cases} \quad (13.164)$$

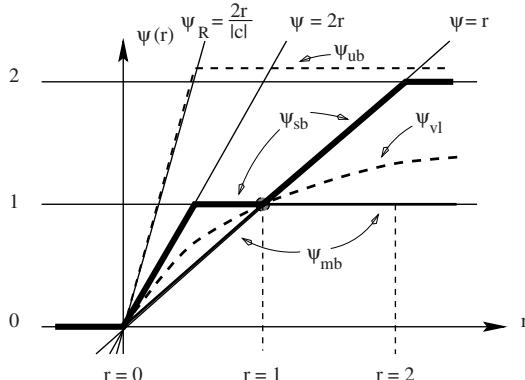
and MINBEE (or MINMOD) is given by

$$\psi_{mb}(r) = \begin{cases} 0, & r \leq 0, \\ r, & 0 \leq r \leq 1, \\ 1, & r \geq 1. \end{cases} \quad (13.165)$$

Fig. 13.21 illustrates four of these flux limiters constructed from Sweby's TVD region, namely: ULTRABEE, SUPERBEE, VANLEER and MINBEE. These are related to the WAF limiter functions given by (13.122)–(13.126). There is in fact a direct correspondence between the WAF limiter functions  $\phi(r)$  and the flux limiter functions  $\psi(r)$  of this section. Such correspondence is given by

$$\phi(r) = 1 - (1 - |c|)\psi(r), \quad (13.166)$$

so that for any given conventional flux limiter function  $\psi(r)$  there is a corresponding WAF limiter function  $\phi(r)$ , and vice-versa.



**Fig. 13.21.** Four flux limiter functions constructed from TVD region of Fig. 13.20: ULTRABEE, SUPERBEE, VANLEER and MINBEE. Compare with WAF limiter of Fig. 13.14

### 13.7.4 TVD Centred Flux Limiter Schemes

We follow the general flux limiter approach discussed in Sect. 13.7.2. Now, the low-order flux  $f_{i+\frac{1}{2}}^{LO}$  is assumed to be a *centred flux*, with general coefficients  $\alpha_0$  and  $\alpha_1$  in (13.131). The high-order flux is still assumed to be that of the Lax–Wendroff method; see equations (13.128) to (13.137).

#### TVD Criteria for Centred Schemes

Convenient TVD conditions for constructing centred TVD schemes are first presented. Such conditions are a generalisation of the classical data compatibility conditions of Sect. 13.5.5 and the necessary conditions given by Harten's Theorem (13.65). We assume scheme (13.129) to be expressed as in (13.137), namely

$$u_i^{n+1} = u_i^n - C \Delta u_{i-\frac{1}{2}} + D \Delta u_{i+\frac{1}{2}}, \quad (13.167)$$

where the coefficients  $C$  and  $D$  are, in general, data-dependent. By defining

$$r_i \equiv \frac{\Delta u_{upw}}{\Delta u_{dow}} \equiv \frac{\Delta u_{i-\frac{s}{2}}}{\Delta u_{i+\frac{s}{2}}} = \begin{cases} \frac{\Delta u_{i-\frac{1}{2}}}{\Delta u_{i+\frac{1}{2}}}, & \text{if } a > 0, \\ \frac{\Delta u_{i+\frac{1}{2}}}{\Delta u_{i-\frac{1}{2}}}, & \text{if } a < 0, \end{cases} \quad (13.168)$$

scheme (13.167) produces

$$R_i^{(s)} \equiv \frac{u_i^{n+1} - u_i^n}{u_{i-s}^n - u_i^n} = \begin{cases} C - D/r_i, & \text{if } a > 0, \\ D - C/r_i, & \text{if } a < 0, \end{cases} \quad (13.169)$$

where  $s = \text{sign}(a)$ .

Next we state a result giving sufficient conditions for scheme (13.167) to be TVD [538].

**Theorem 13.69.** Scheme (13.167) for (13.128) is TVD if

$$-\frac{\epsilon_R}{r_i} \leq R_i^{(s)} \leq \epsilon_L , \quad \text{if } r_i > 0 , \quad (13.170)$$

$$0 \leq R_i^{(s)} \leq 1 - \epsilon_0 - \epsilon_0/r_i , \quad \text{if } r_i < 0 , \quad (13.171)$$

where  $\epsilon_L, \epsilon_0, \epsilon_R$  are real numbers satisfying

$$0 \leq \epsilon_L, \epsilon_0, \epsilon_R \leq 1 . \quad (13.172)$$

*Proof.* We first prove result (13.170), which refers to monotone increasing or monotone decreasing data. Monotonicity is ensured, and thus the TVD property, if conditions (13.68) are enforced, namely

$$\min_k \{u_k^n\} \leq u_i^{n+1} \leq \max_k \{u_k^n\} . \quad (13.173)$$

First consider the case of monotone decreasing data

$$u_{i-1}^n > u_i^n \geq u_{i+1}^n$$

and assume

$$(1 - \epsilon_R)u_i^n + \epsilon_R u_{i+1}^n \leq u_i^{n+1} \leq (1 - \epsilon_L)u_i^n + \epsilon_L u_{i-1}^n , \quad (13.174)$$

which, in a sense, is a special case of (13.173), as it is more restrictive than (13.173). Assuming  $a > 0$ , subtracting  $u_i^n$  and dividing through by  $u_{i-1}^n - u_i^n > 0$  leads to the sought result (13.170). For  $a < 0$  the proof is similar. The case of monotone increasing data follows in an analogous manner. Condition (13.171) refers to extrema, local maxima or local minima. First assume a local minimum, i.e.

$$u_i^n \leq u_{i-1}^n ; \quad u_i^n \leq u_{i+1}^n .$$

We impose the following monotonicity condition

$$u_i^n \leq u_i^{n+1} \leq (1 - \epsilon_0)u_{i-1}^n + \epsilon_0 u_{i+1}^n ,$$

which is a special case of (13.173). Simple manipulations lead to the sought result (13.171). The case of a local maximum leads to the same result and thus the TVD conditions (13.170), (13.171) are proved.

*Remark 13.70.* The TVD condition (13.170) relates to points away from extrema and is more relaxed than the Data Compatibility condition (13.85), namely

$$0 \leq R_i^{(s)} \leq 1 , \quad (13.175)$$

which, as discussed previously, is perfectly adequate for deriving limiter functions for the case in which the underlying first-order scheme is the Godunov first order *upwind* method [216]. However, for the case in which the underlying

first order scheme is *centred*, direct application of (13.175) leads to over restrictive TVD regions that may actually exclude the underlying monotone first-order scheme! The TVD condition (13.171) is more restrictive than (13.75) but since this condition relates to extrema, for which the scheme is locally first-order accurate, the TVD conditions (13.170)–(13.171) are overall more relaxed than (13.175).

## TVD Regions

We first prove a result that concerns the construction of the TVD region for centred flux limiter schemes.

**Theorem 13.71.** *In order to ensure a TVD scheme, flux limiters  $\phi(r)$  must lie in a region (the TVD region) satisfying the following constraints*

$$\phi_B \leq \phi(r) \leq \phi_T , \quad (13.176)$$

$$\phi_L^+(r) \leq \phi(r) \leq \phi_R^+(r) , \quad r > 0 , \quad (13.177)$$

$$\phi_R^-(r) \leq \phi(r) \leq \phi_L^-(r) , \quad r \leq 0 , \quad (13.178)$$

$$\phi_L^+(r) = (S_L + \phi_T)r , \quad (13.179)$$

$$\phi_R^+(r) = \phi_g + (S_R^+ + \phi_B)r , \quad (13.180)$$

$$\phi_L^-(r) = \phi_g + (S_L + \phi_T)r , \quad (13.181)$$

$$\phi_R^-(r) = (S_R^- + \phi_B)r . \quad (13.182)$$

Here  $\phi_B, \phi_T$  are global lower and upper bounds and definitions for the parameters  $S_L, \phi_g, S_R^+, S_R^-$  are given in Table 13.3.

	$a > 0$	$a < 0$	FORCE	Lax–Friedrichs	Godunov
$S_L$	$\frac{\alpha_0}{\alpha_1 - \beta_1}$	$\frac{\alpha_1}{\alpha_0 - \beta_0}$	$\frac{1 +  c }{1 -  c }$	$\frac{1}{1 -  c }$	$\frac{2}{1 -  c }$
$\phi_g$	$\frac{\alpha_1}{\alpha_1 - \beta_1}$	$\frac{\alpha_0}{\alpha_0 - \beta_0}$	$\frac{1 -  c }{1 +  c }$	$\frac{1}{1 +  c }$	0
$S_R^+$	$\frac{\alpha_0 - 1/ c }{\alpha_1 - \beta_1}$	$\frac{\alpha_1 - 1/ c }{\alpha_0 - \beta_0}$	$\frac{3 +  c }{1 +  c }$	$\frac{1}{1 +  c }$	$\frac{2}{ c }$
$S_R^-$	$\frac{\alpha_0 - 1/ c  - \alpha_1}{\alpha_1 - \beta_1}$	$\frac{\alpha_1 - 1/ c  - \alpha_0}{\alpha_0 - \beta_0}$	2	0	$\frac{2}{ c }$

**Table 13.3.** Values of useful quantities for the FORCE scheme, the Lax–Friedrichs scheme and the Godunov first order upwind method.

*Proof.* (For details of proof see Toro and Billett [538]).

A few remarks are in order. The TVD region is determined by inequalities involving six straight lines, namely the horizontal bounds  $\phi_B, \phi_T$  and the functions defined by (13.179)–(13.182). Fig. 13.22 depicts the general TVD region defined by the theorem. We have chosen  $\phi_B = 0$  and  $\phi_T$  consistent with a non-positive slope for  $\phi_L^-(r)$  and  $\phi_R^+(r)$ . In fact it suffices to take these functions to be constant and we may therefore set

$$\phi_T = -\frac{\alpha_0}{\alpha_1 - \beta_1} ; \quad a > 0 ; \quad \phi_T = -\frac{\alpha_1}{\alpha_0 - \beta_0} ; \quad a < 0 . \quad (13.183)$$

and

$$\phi_L^+(r) = 0 ; \quad \phi_L^-(r) \equiv \phi_g = \begin{cases} \frac{\alpha_1}{\alpha_1 - \beta_1}, & a > 0, \\ \frac{\alpha_0}{\alpha_0 - \beta_0}, & a < 0 . \end{cases} \quad (13.184)$$

Table 13.3 lists expressions for various quantities involved, for three first-order schemes, including the Godunov first order upwind method.

*Remark 13.72.* In the course of the proof of the above theorem [538] it is found that the convenient values for the parameters  $\epsilon_L, \epsilon_0, \epsilon_R$  in (13.170), (13.171) are

$$\epsilon_L = 1 , \quad \epsilon_0 = \epsilon_R = -c\alpha_1 . \quad (13.185)$$

If we enforced the usual Data Compatibility Conditions (13.175), we would

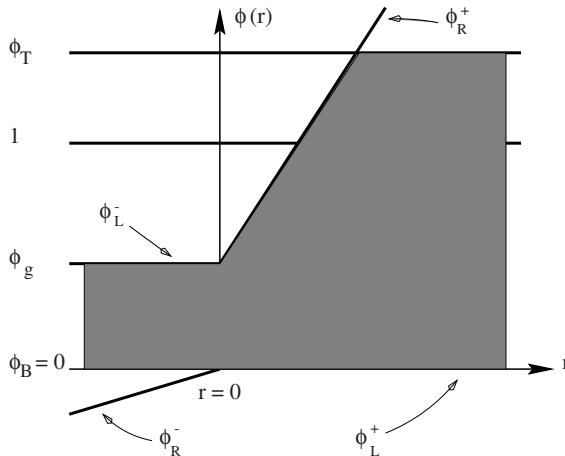


Fig. 13.22. TVD region for flux limiter centred schemes

have had  $\epsilon_0 = \epsilon_R = 0$  and thus the TVD region in Fig. 13.22 would have excluded limiter functions  $\phi(r)$  with  $\phi(r) \leq \phi_g > 0$  except for the special case  $|c| = 1$ , for which  $\phi_g = 0$ . Even the basic case  $\phi(r) = 0$  that reproduces the low-order monotone scheme in (13.130) would be excluded. For upwind methods, this difficulty does not arise, as the downwind coefficient ( $\alpha_1$  for  $a > 0$  and  $\alpha_0$  for  $a < 0$ ) is zero, and thus  $\phi_g = 0 \forall c$ .

## Construction of Limiters

The task ahead is to construct flux limiters  $\phi(r)$  to be used in the flux limiter scheme (13.130). In this connection the following result applies.

**Theorem 13.73.** *The centred TVD flux*

$$f_{i+\frac{1}{2}}^{(c)} = f_{i+\frac{1}{2}}^{LO} + \phi_{i+\frac{1}{2}}(r)(f_{i+\frac{1}{2}}^{LW} - f_{i+\frac{1}{2}}^{LO}) , \quad (13.186)$$

where  $f_{i+\frac{1}{2}}^{LO}$  is the flux for some first-order centred monotone scheme and  $f_{i+\frac{1}{2}}^{LW}$  is the Lax-Wendroff flux, reduces to the flux for the Godunov first-order upwind flux

$$f_{i+\frac{1}{2}}^{(g)} = \frac{1}{2}(1+s)(au_i^n) + \frac{1}{2}(1-s)(au_{i+1}^n) , \quad (13.187)$$

with  $s = \text{sign}(a)$ , when

$$\phi(r) = \phi_g , \quad (13.188)$$

and upwind flux limiters  $\psi(r)$  are related to centred flux limiters  $\phi(r)$  by the equation

$$\phi(r) = \phi_g + (1 - \phi_g)\psi(r) . \quad (13.189)$$

*Proof.* (Left as an exercise).

*Remark 13.74.* Based on the above result, we call  $\phi_g$  the Godunov point, see Table 13.3, and note that the conventional TVD conditions (13.175), while suitable for upwind TVD schemes, do not admit flux limiters below  $\phi_g$ . That is, schemes that have larger numerical viscosity such as FORCE and the Lax-Friedrichs method, are not included in the TVD region.

*Remark 13.75.* Based on the first result of Theorem (13.73) one can generalise the Godunov first-order upwind method to non-linear scalar conservation laws by simply defining a flux as in (13.186) with  $\phi_{i+\frac{1}{2}}(r) = \phi_g$ . Such extension depends on the particular centred low order flux  $f_{i+\frac{1}{2}}^{LO}$  used and on a Courant number  $c$ ; this may be obtained from the characteristic speed (local) or the CFL coefficient (global). A local choice is

$$c_{i+\frac{1}{2}} = \frac{\Delta t s_{i+\frac{1}{2}}}{\Delta x} , \quad s_{i+\frac{1}{2}} \equiv |a_{i+\frac{1}{2}}| , \quad (13.190)$$

where  $a_{i+\frac{1}{2}}$  is a characteristic speed at the cell interface  $i + \frac{1}{2}$ . We note that such extension *does not require the solution of a local Riemann problem*, as does the Godunov first-order upwind method.

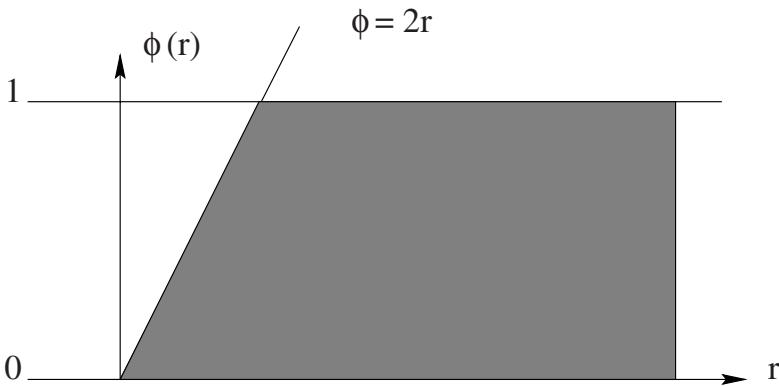
Before constructing flux limiters  $\phi(r)$  we note that the parameters  $S_L$ ,  $\phi_g$ ,  $S_R^+$ ,  $S_R^-$  in (13.179)–(13.182) define the TVD region, and they depend on the Courant number  $|c|$ ; see Table 13.3. By identifying possible minimum and maximum values of  $|c|$ , i.e.

$$c_{min} = \min |c|, \quad c_{max} = \max |c|, \quad (13.191)$$

we construct a TVD region based on the most restrictive conditions. The corresponding values for the relevant parameters are

$$\phi_T = \phi_T(c_{min}), \quad \phi_g = \phi_g(c_{max}), \quad S_R^+ = S_R^+(c_{max}). \quad (13.192)$$

Recall that we set  $\phi_L^+(r) = 0$ ;  $\phi_L^-(r) = \phi_g$  and  $S_L$  in (13.179), (13.181) is defined once  $\phi_T$  is chosen. For the model equation (13.128) we interpret (13.191) as meaning  $c_{min} = 0$  and  $c_{max} = 1$ , leading to a Courant–number independent TVD region depicted in Fig. 13.23. It follows that the only possible flux



**Fig. 13.23.** TVD region for flux limiter centred schemes, that is independent of Courant number

limiters are of the MINBEE type [470], [409], the most prominent example being

$$\phi_{mb}(r) = \begin{cases} 0, & r \leq 0, \\ r, & 0 \leq r \leq 1, \\ 1, & r \geq 1. \end{cases} \quad (13.193)$$

Note that  $\phi_{mb}(r)$  is the conventional upwind flux limiter MINBEE (13.165) [470], [409]. This is perfectly adequate for most realistic applications but has the disadvantage of being over diffusive. In general, larger values of  $\phi(r)$  add less numerical viscosity than smaller ones. It is therefore desirable to preserve at least part of the TVD region above  $\phi = 1$ . Using the relation (13.189) between conventional upwind flux limiters  $\psi(r)$  and the sought centred flux limiters  $\phi(r)$  we construct these as follows

$$\phi(r) = \hat{\phi}_g + (1 - \hat{\phi}_g)\psi(r), \quad (13.194)$$

with

$$\hat{\phi}_g = \begin{cases} 0, & r \leq 1, \\ \phi_g \equiv \phi_g(c_{max}), & r \geq 1. \end{cases} \quad (13.195)$$

By choosing the upwind flux limiters ULTRABEE, SUPERBEE, VANLEER and VANALBADA we obtain *corresponding* centred flux limiters

$$\phi_{ub}(r) = \begin{cases} 0, & r \leq 0, \\ \min\{\phi_T, \phi_g + S_R^+ r\}, & r > 0, \end{cases} \quad (13.196)$$

$$\phi_{sb}(r) = \begin{cases} 0, & r \leq 0, \\ 2r, & 0 \leq r \leq \frac{1}{2}, \\ 1, & \frac{1}{2} \leq r \leq 1, \\ \min\{2, \phi_g + (1 - \phi_g)r\}, & r > 1, \end{cases} \quad (13.197)$$

$$\phi_{vl}(r) = \begin{cases} 0, & r \leq 0, \\ \frac{2r}{1+r}, & 0 \leq r \leq 1, \\ \phi_g + \frac{2(1-\phi_g)r}{1+r}, & r \geq 1, \end{cases} \quad (13.198)$$

$$\phi_{va}(r) = \begin{cases} 0, & r \leq 0, \\ \frac{r(1+r)}{1+r^2}, & 0 \leq r \leq 1, \\ \phi_g + \frac{(1-\phi_g)r(1+r)}{1+r^2}, & r \geq 1, \end{cases} \quad (13.199)$$

We stressed that the corresponding centred limiters (13.196)–(13.199) are only *analogous* to the upwind limiters, they are *not equivalent*.

## 13.8 Slope Limiter Methods

The MUSCL approach introduced in Sect. 13.4 allows the construction of high order methods. In Sects. 13.4.2 to 13.4.4 we constructed second and third order accurate extensions of the Godunov first-order upwind method using this approach. In Sect. 13.4.5 we constructed second and third order accurate extensions of *centred* first order schemes following the MUSCL–Hancock methodology. All of these high order schemes, as stated, will produce spurious oscillations in the vicinity of high gradients, see Godunov’s theorem in Sect. 13.5.3. In this section we construct non-linear versions of these schemes by replacing the slopes  $\Delta_i$  in the data reconstruction step (13.23) by limited slopes  $\bar{\Delta}_i$ , according to some TVD constraints.

### 13.8.1 TVD Conditions

First we note that in constructing fully discrete second-order explicit TVD methods, restricting the boundary extrapolated values  $u_i^L, u_i^R$  in (13.24) to satisfy

$$\left. \begin{aligned} \min\{u_{i-1}^n, u_i^n\} &\leq u_i^L \leq \max\{u_{i-1}^n, u_i^n\}, \\ \min\{u_{i+1}^n, u_i^n\} &\leq u_i^R \leq \max\{u_i^n, u_{i+1}^n\}. \end{aligned} \right\} \quad (13.200)$$

does not lead to useful results. Instead, we impose a restriction on the *evolved* boundary extrapolated values  $\bar{u}_i^L, \bar{u}_i^R$ , see (13.30) or (13.33), and prove the following result.

**Theorem 13.76.** If the evolved boundary extrapolated values  $\{\bar{u}_k^{L,R}\}$  satisfy

$$\left. \begin{aligned} \min\{u_{i-1}^n, u_i^n\} &\leq \bar{u}_i^L \leq \max\{u_{i-1}^n, u_i^n\}, \quad \forall i, \\ \min\{u_i^n, u_{i+1}^n\} &\leq \bar{u}_i^R \leq \max\{u_i^n, u_{i+1}^n\}, \quad \forall i, \end{aligned} \right\} \quad (13.201)$$

then, for any monotone scheme of the form

$$v_i^{n+1} = \sum_{k=-1}^1 b_k v_{i+k}^n \quad (13.202)$$

applied to  $\{\bar{u}_k^{L,R}\}$  one has

$$\min_k \{u_k^n\}_{k=i-2}^{i+2} \leq u_i^{n+1} \leq \max_k \{u_k^n\}_{k=i-1}^{i+2}. \quad (13.203)$$

Hence the corresponding MUSCL scheme is TVD.

*Proof.* From conditions (13.201) we have

$$\min_k \{u_k^n\}_{k=i-1}^{i+1} \leq \bar{u}_i^{L,R} \leq \max_k \{u_k^n\}_{k=i-1}^{i+1}, \quad \forall i. \quad (13.204)$$

For any monotone scheme (13.202) applied to  $\{u_i^n; \bar{u}_k^{L,R}\}_{k=i-1}^{i+1}$  we will have

$$\min_k \{u_i^n; \bar{u}_k^{L,R}\}_{k=i-1}^{i+1} \leq \bar{u}_i^{n+1} \leq \max_k \{u_i^n; \bar{u}_k^{L,R}\}_{k=i-1}^{i+1}. \quad (13.205)$$

But from (13.201)

$$\left. \begin{aligned} \min\{u_k^n\}_{k=i-2}^{i+2} &\leq \min\{u_i^n; \bar{u}_k^{L,R}\}_{k=i-1}^{i+1}, \\ \max\{u_i^n; \bar{u}_k^{L,R}\}_{k=i-1}^{i+1} &\leq \max\{u_k^n\}_{k=i-2}^{i+2}. \end{aligned} \right\} \quad (13.206)$$

From (13.205) and (13.206) inequalities (13.203) follow and the result is thus proved.

### 13.8.2 Construction of TVD Slopes

Now we construct *limited* slopes  $\bar{\Delta}_i$  to replace  $\Delta_i$  in (13.23), or (13.33), according to the TVD constraints (13.201) and prove the following result.

**Theorem 13.77.** If the limited slopes  $\bar{\Delta}_i$  are chosen according to

$$\left. \begin{aligned} \bar{\Delta}_i &= \frac{1}{2} [\text{sign}(\Delta_{i-\frac{1}{2}}) + \text{sign}(\Delta_{i+\frac{1}{2}})] \times \min[\beta_{i-\frac{1}{2}} |\Delta_{i-\frac{1}{2}}|, \beta_{i+\frac{1}{2}} |\Delta_{i+\frac{1}{2}}|] \\ \beta_{i-\frac{1}{2}} &= \frac{2}{1+c}, \quad \beta_{i+\frac{1}{2}} = \frac{2}{1-c}, \end{aligned} \right\} \quad (13.207)$$

then the resulting MUSCL scheme is TVD.

*Proof.* First assume  $u_{i-1}^n \leq u_i^n$ . Application of (13.201) leads to

$$u_{i-1}^n \leq u_i^n - \frac{1}{2}(1+c)\bar{\Delta}_i \leq u_i^n .$$

Use of these two inequalities produces

$$0 \leq \bar{\Delta}_i \leq \frac{2}{1+c}\Delta_{i-\frac{1}{2}} ; \quad \Delta_{i-\frac{1}{2}} \geq 0 . \quad (13.208)$$

The case in which  $u_{i-1}^n \geq u_i^n$  leads to

$$\frac{2}{1+c}\Delta_{i-\frac{1}{2}} \leq \bar{\Delta}_i \leq 0 ; \quad \Delta_{i-\frac{1}{2}} \leq 0 . \quad (13.209)$$

Similarly, for  $u_i^n \leq u_{i+1}^n$  one obtains

$$0 \leq \bar{\Delta}_i \leq \frac{2}{1-c}\Delta_{i+\frac{1}{2}} ; \quad \Delta_{i+\frac{1}{2}} \geq 0 \quad (13.210)$$

and for  $u_i^n \geq u_{i+1}^n$  one obtains

$$\frac{2}{1-c}\Delta_{i+\frac{1}{2}} \leq \bar{\Delta}_i \leq 0 ; \quad \Delta_{i+\frac{1}{2}} \leq 0 . \quad (13.211)$$

From inspection of conditions (13.208) and (13.210) when  $\Delta_{i-\frac{1}{2}} \geq 0$ ,  $\Delta_{i+\frac{1}{2}} \geq 0$ , one requires  $\bar{\Delta}_i$  to be bounded by zero and the minimum of  $\frac{2}{1+c}\Delta_{i-\frac{1}{2}}$  and  $\frac{2}{1-c}\Delta_{i+\frac{1}{2}}$ . Conditions (13.209) and (13.211), when  $\Delta_{i-\frac{1}{2}} \leq 0$ ,  $\Delta_{i+\frac{1}{2}} \leq 0$ , require  $\bar{\Delta}_i$  to lie between

$$\text{sign}(\Delta_{i-\frac{1}{2}}) \times \min\left\{\frac{2}{1+c}|\Delta_{i-\frac{1}{2}}|, \frac{2}{1-c}|\Delta_{i+\frac{1}{2}}|\right\}$$

and zero. When  $\Delta_{i-\frac{1}{2}}$  and  $\Delta_{i+\frac{1}{2}}$  are of opposite sign, the only possible choice is  $\bar{\Delta}_i = 0$ . Hence all the above constraints may then be written as in (13.207).

The reader may easily verify that (13.207) may also be written as

$$\bar{\Delta}_i = \text{sign}(\Delta_{i+\frac{1}{2}}) \times \max\{0, \min[\beta_{i-\frac{1}{2}}\Delta_{i-\frac{1}{2}}\text{sign}(\Delta_{i+\frac{1}{2}}), \beta_{i+\frac{1}{2}}\Delta_{i+\frac{1}{2}}]\} . \quad (13.212)$$

### 13.8.3 Slope Limiters

The TVD analysis of the previous section is sufficient to construct *upwind* and *centred* slope limiter methods. A re-interpretation of this analysis in terms of a *slope limiter*  $\xi_i$  such that

$$\bar{\Delta}_i = \xi_i \Delta_i , \quad (13.213)$$

with  $\Delta_i$  as given by

$$\Delta_i = \frac{1}{2}(1 + \omega)\Delta u_{i-\frac{1}{2}} + \frac{1}{2}(1 - \omega)\Delta u_{i+\frac{1}{2}}, \quad (13.214)$$

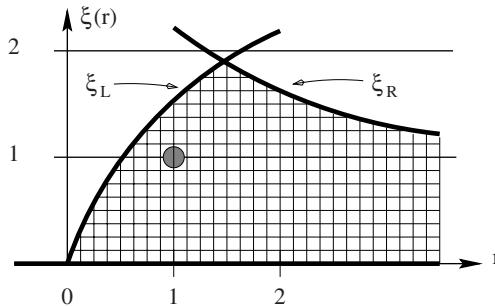
is useful in a number of ways. It leads to a TVD region for  $\xi(r)$  given as follows

$$\xi(r) = 0 \text{ for } r \leq 0, \quad 0 \leq \xi(r) \leq \min\{\xi_L(r), \xi_R(r)\} \text{ for } r > 0, \quad (13.215)$$

where

$$\left. \begin{aligned} \xi_L(r) &= \frac{2\beta_{i-\frac{1}{2}}r}{1 - \omega + (1 + \omega)r}, \\ \xi_R(r) &= \frac{2\beta_{i+\frac{1}{2}}}{1 - \omega + (1 + \omega)r}, \\ r &= \frac{\Delta_{i-\frac{1}{2}}}{\Delta_{i+\frac{1}{2}}}. \end{aligned} \right\} \quad (13.216)$$

The corresponding TVD region is depicted in Fig. 13.24, see also [23]. A



**Fig. 13.24.** TVD region for slope limiters. For negative  $r$  TVD region is single line  $\xi = 0$  and for positive  $r$  TVD region lies between 0 and  $\min\{\xi_L(r), \xi_R(r)\}$ .

possible slope limiter function  $\xi(r)$  is

$$\xi(r) = \begin{cases} 0, & r \leq 0 \\ \min\{\xi_L(r), \xi_R(r)\}, & r > 0 \end{cases} \quad (13.217)$$

This slope limiter is analogous to the flux limiter ULTRABEE, in that one follows the boundaries of the TVD region. It is expected to be over compressive for smooth solutions but to give satisfactory results for discontinuities. A slope limiter that is *analogous* (not equivalent) to the SUPERBEE flux limiter is

$$\xi(r) = \begin{cases} 0, & \text{if } r \leq 0, \\ 2r, & \text{if } 0 \leq r \leq \frac{1}{2}, \\ 1, & \text{if } \frac{1}{2} \leq r \leq 1, \\ \min\{r, \xi_R(r), 2\}, & \text{if } r \geq 1. \end{cases} \quad (13.218)$$

A van Leer-type slope limiter is

$$\xi(r) = \begin{cases} 0, & r \leq 0, \\ \min\left\{\frac{2r}{1+r}, \xi_R(r)\right\}, & r \geq 0. \end{cases} \quad (13.219)$$

A van Albada-type slope limiter is

$$\xi(r) = \begin{cases} 0, & r \leq 0, \\ \min\left\{\frac{r(1+r)}{1+r^2}, \xi_R(r)\right\}, & r \geq 0. \end{cases} \quad (13.220)$$

A MINBEE-type slope limiter is

$$\xi(r) = \begin{cases} 0, & r \leq 0, \\ r, & 0 \leq r \leq 1, \\ \min\{1, \xi_R(r)\}, & r \geq 1. \end{cases} \quad (13.221)$$

One issue is the choice of the coefficients  $\beta_{i-\frac{1}{2}}$ ,  $\beta_{i+\frac{1}{2}}$  in (13.207) and (13.216). There are two cases to consider, namely the case of positive Courant number  $c$  and the case of negative Courant number  $c$ . For schemes that are stable in the range  $|c| \leq 1$  one has

$$\begin{cases} \beta_{i-\frac{1}{2}} \in [1, 2] \text{ and } \beta_{i+\frac{1}{2}} \in [2, \infty) & \text{when } a > 0, \\ \beta_{i-\frac{1}{2}} \in [2, \infty) \text{ and } \beta_{i+\frac{1}{2}} \in [1, 2] & \text{when } a < 0. \end{cases} \quad (13.222)$$

For upwind methods one may exploit the upwind information contained in the Courant number  $c$ . For *centred* (non-upwind) methods such information is unavailable in general and we therefore recommend the choice

$$\beta_{i-\frac{1}{2}} = \beta_{i+\frac{1}{2}} = 1, \quad (13.223)$$

which gives a *smaller* TVD region, valid for all Courant numbers in the range  $|c| \leq 1$ . One may improve this by utilising a limited amount of wave information via the CFL coefficient, which must be available for any explicit scheme to solve any hyperbolic system. See Sect. 6.3.2 of Chap. 6 on the CFL condition.

### 13.8.4 Limited Slopes Obtained from Flux Limiters

By establishing a relationship between upwind-based flux limiter schemes and MUSCL-type schemes for the model linear advection equation (13.1), one may select the *limited slopes*  $\bar{\Delta}_i$  in the reconstruction step so as to reproduce conventional (upwind) flux limiters  $\psi_{i+\frac{1}{2}}$ . From (13.130), the intercell flux for the flux limiter method resulting from the Godunov first order upwind and the Lax-Wendroff schemes is

$$f_{i+\frac{1}{2}} = \begin{cases} au_i^n + \psi_{i+\frac{1}{2}}[\frac{1}{2}(1-c)\Delta_{i+\frac{1}{2}}]a, & a > 0, \\ au_{i+1}^n - \psi_{i+\frac{1}{2}}[\frac{1}{2}(1+c)\Delta_{i+\frac{1}{2}}]a, & a < 0, \end{cases} \quad (13.224)$$

where  $\psi_{i+\frac{1}{2}}$  is an upwind flux limiter, see Sect. 13.7.3. An upwind-based slope limiter method has intercell flux

$$f_{i+\frac{1}{2}} = \begin{cases} au_i^n + \frac{1}{2}(1-c)a\bar{\Delta}_i, & a > 0, \\ au_{i+1} - \frac{1}{2}(1+c)a\bar{\Delta}_{i+1}, & a < 0, \end{cases} \quad (13.225)$$

where  $\bar{\Delta}_i$  and  $\bar{\Delta}_{i+1}$  are *limited slopes* in cells  $i$  and  $i+1$  respectively. See (13.35), (13.36), (13.43) and (13.49). By comparing (13.224) and (13.225) one obtains

$$\psi_{i+\frac{1}{2}} \Delta_{i+\frac{1}{2}} = \begin{cases} \bar{\Delta}_i, & a > 0, \\ \bar{\Delta}_{i+1}, & a < 0. \end{cases} \quad (13.226)$$

As  $\psi_{i+\frac{1}{2}} = 1 \forall i$  reproduces the Lax–Wendroff method in the flux–limiter scheme, the choice of slopes

$$\Delta_i = \begin{cases} \Delta_{i+\frac{1}{2}}, & a > 0, \\ \Delta_{i-\frac{1}{2}}, & a < 0 \end{cases} \quad (13.227)$$

in the MUSCL schemes also reproduces the Lax–Wendroff method. Note that this would be obtained by an appropriate value of the parameter  $\omega$  in (13.27). Using this relation one can construct *limited slopes*

$$\bar{\Delta}_i = \bar{\Delta}_i(\Delta_{i-\frac{1}{2}}, \Delta_{i+\frac{1}{2}}) \quad (13.228)$$

so as to reproduce conventional (upwind) flux limiters. The reader can easily verify that two methods for obtaining *limited slopes* result from

$$\bar{\Delta}_i = \begin{cases} \max[0, \min(\beta\Delta_{i-\frac{1}{2}}, \Delta_{i+\frac{1}{2}}), \min(\Delta_{i-\frac{1}{2}}, \beta\Delta_{i+\frac{1}{2}})], & \Delta_{i+\frac{1}{2}} > 0, \\ \min[0, \max(\beta\Delta_{i-\frac{1}{2}}, \Delta_{i+\frac{1}{2}}), \max(\Delta_{i-\frac{1}{2}}, \beta\Delta_{i+\frac{1}{2}})], & \Delta_{i+\frac{1}{2}} < 0 \end{cases} \quad (13.229)$$

for particular values of the parameter  $\beta$ . The value  $\beta = 1$  reproduces the MINBEE flux limiter (13.165), which may also be written as

$$\psi_{mi}(r) = \max[0, \min(1, r)]. \quad (13.230)$$

$\beta = 2$  reproduces the SUPERBEE flux limiter (13.162), which may also be written as

$$\psi_{sb}(r) = \max[0, \min(2r, 1), \min(r, 2)]. \quad (13.231)$$

The parameter  $r$  is defined in (13.157).

*Remark 13.78.* Note that upwind MUSCL type schemes based on the limited slopes (13.229) have the Lax–Wendroff as the base second–order scheme, when applied to the linear advection equation (13.1), and *not* the Fromm scheme; see Sect. 13.4.

## 13.9 Extensions of TVD Methods

In this chapter we have studied a variety of TVD schemes for the scalar, homogeneous linear partial differential equation (13.1). Crucial questions concern the extension of these methods to scalar inhomogeneous (source terms) equations, convection-diffusion equations and non-linear systems. The extension of the schemes to homogeneous non-linear systems is carried out in Chap. 14. Here we make some remarks concerning model PDEs with source terms and diffusion terms.

### 13.9.1 TVD Schemes in the Presence of Source Terms

Sweby [471] considered the problem of devising TVD schemes to solve the inhomogeneous PDE

$$u_t + au_x = s(u) , \quad (13.232)$$

where  $s(u)$  is an algebraic function of the unknown  $u = u(x, t)$ , usually called a *source term*, or forcing term. In fact an even more basic problem is that of just devising numerical methods to solve (13.232) after some fashion. In Chap. 15 we present splitting methods, whereby the full inhomogeneous problem is *split* into a homogenous problem, for which a TVD scheme can be applied directly, and an Ordinary Differential Equation (ODE) that can be solved by some appropriate ODE solver. This approach carries over to non-linear inhomogeneous systems of PDEs. For special cases in which a change of variables allows us to re-write the PDEs as a, new, homogeneous problem, one applies TVD methods directly to the homogeneous problem and thus to the full original inhomogeneous problem. See Sweby [471] for a discussion on this approach and Watson et. al. [576] for an application to the non-linear shallow water equations with variable bed elevation.

On the question of whether one can construct TVD schemes for general inhomogeneous PDEs of the form (13.232), it should first be realised that the TVD concept itself is inappropriate here. The effect of the source term may cause the Total Variation of the exact solution to increase and it would therefore be absurd to attempt to produce numerical methods to compute approximate solutions with properties not enjoyed by the exact solution.

For general problems one resorts to the splitting schemes presented in Chap. 15.

### 13.9.2 TVD Schemes in the Presence of Diffusion Terms

Consider the question of devising TVD schemes to solve the convection-diffusion equation

$$u_t + au_x = \alpha_{phy} u_{xx} , \quad (13.233)$$

where  $\alpha_{phy}$  is a physical or natural viscosity coefficient, assumed constant here. A very simple way of solving this PDE is to deploy the splitting schemes

studied in Chap. 15. These splitting schemes apply a TVD method to the convection part  $u_t + au_x = 0$  and the parabolic PDE  $u_t = \alpha_{phy} u_{xx}$  is solved by some other appropriate method. In general these methods are found to be quite successful for solving the Navier–Stokes equations [82], [529].

Recall however that when a TVD method is applied to  $u_t + au_x = 0$ , one effectively adds a numerical viscosity term to this pure convection PDE, so that one solves

$$u_t + au_x = \alpha_{num} u_{xx}, \quad (13.234)$$

where  $\alpha_{num}$  is a numerical or artificial viscosity coefficient; see Sect. 5.2 of Chap. 5. The value of  $\alpha_{num}$  depends on the flux or slope limiter used and its purpose is to avoid spurious oscillations in the vicinity of high gradients in the numerical solution. As the complete PDE (13.233) to be solved contains a physical viscosity term, an obvious question is this: is the physical viscosity by itself sufficient to guarantee oscillation free solutions? The answer is no, in general. Toro [507] considered this problem and found that for certain TVD discretisation schemes one does not require artificial viscosity only if

$$R_{cell} \leq \frac{2}{1 - |c|}, \quad (13.235)$$

where  $R_{cell}$  is the cell Reynolds number defined as  $R_{cell} = \frac{|c|}{d}$ ;  $c$  is the Courant number and  $d = \frac{\alpha_{phy} \Delta t}{\Delta x^2}$  is the diffusion number; in this case the physical viscosity is sufficient to guarantee oscillation-free solutions. Otherwise, numerical viscosity is also needed. Viscous flux limiters were constructed in [507], whereby the absolute minimum of *extra numerical viscosity* is added to the physical viscosity to guarantee oscillation-free solutions. These viscous limiters use fully the physical viscosity and minimise the amount of numerical viscosity; in fact this is turned off when condition (13.235) is satisfied. See [507] for details on preliminary results for model problems.

In the next section we present some numerical results for the linear advection equation.

## 13.10 Numerical Results for Linear Advection

We consider the same two test problems for the linear advection (13.1) as in Sect. 13.7.2. We take  $a = 1$  and a CFL coefficient  $C_{cfl} = 0.8$ . Computed results are shown at the output time  $t = 10$  units (1250 time steps). In each figure we compare the exact solution (shown by full lines) with the numerical solution (symbols).

The initial conditions for Test 1 (smooth data) and Test 2 (discontinuous data, square wave) are respectively given by

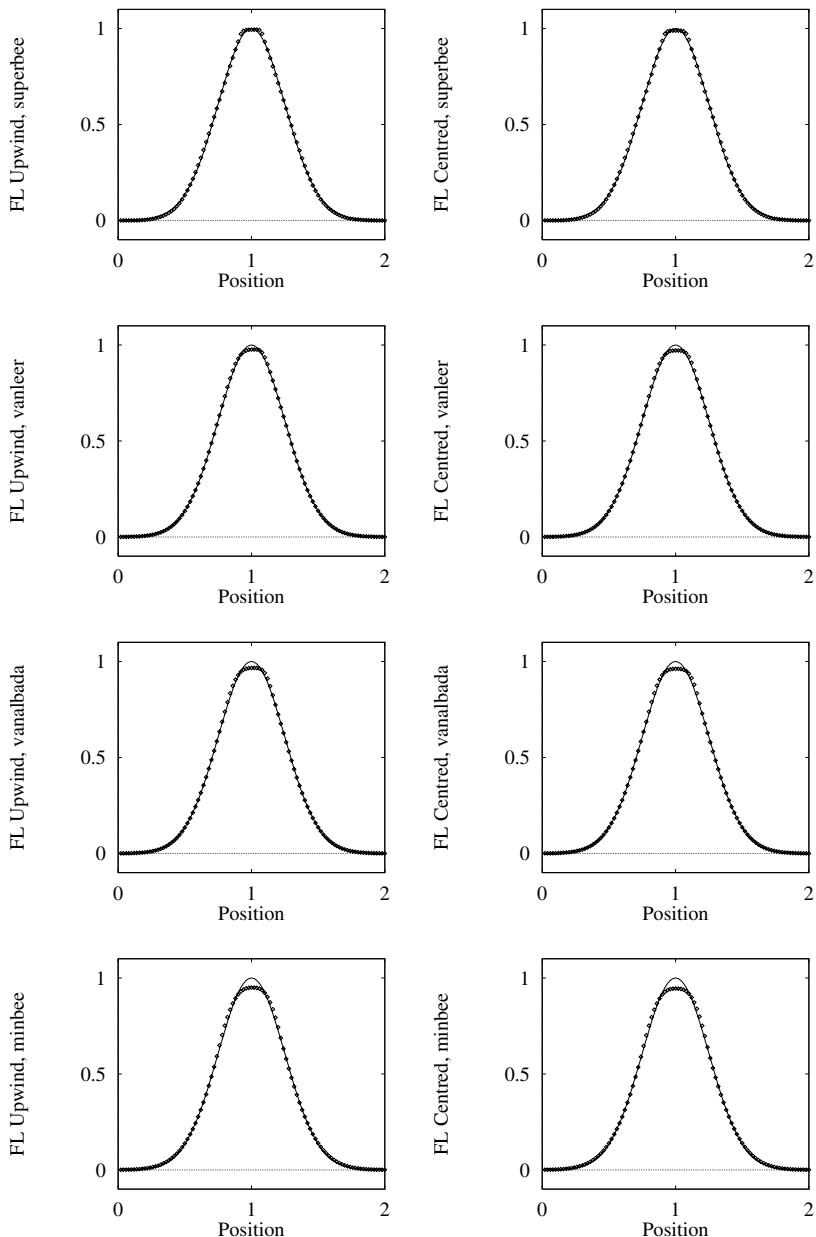
$$u(x, 0) = \alpha e^{-\beta x^2}; \quad u(x, 0) = \begin{cases} 0 & \text{if } x \leq 0.3, \\ 1 & \text{if } 0.3 \leq x \leq 0.7, \\ 0 & \text{if } x \geq 0.7. \end{cases} \quad (13.236)$$

For Test 1 we take  $\alpha = 1.0$  and  $\beta = 8.0$ ; the initial profile  $u(x, 0)$  is evaluated in the interval  $-1 \leq x \leq 1$ ; the mesh used is  $\Delta x = 0.02$ . For Test 2 the mesh used is  $\Delta x = 0.01$ . In both test problems the initial profile and that at the output time is resolved by 100 computing cells.

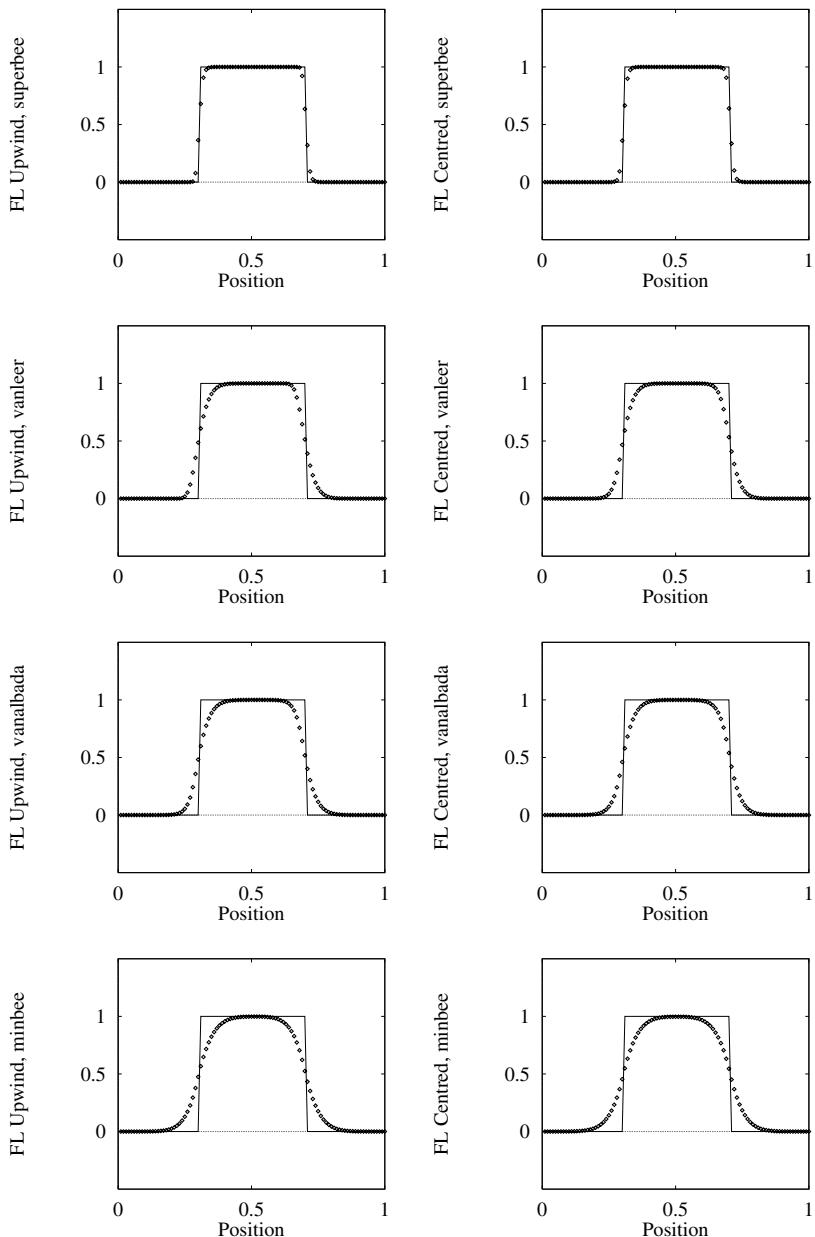
Results are presented for two classes of methods, namely *flux limiter methods* and *slope limiter methods*. For each of these two classes we consider one *upwind*-based method and one *centred* method. The upwind schemes are high-order, TVD extensions of the Godunov first-order upwind scheme. The centred schemes used are second-order, TVD extensions of the FORCE scheme (with  $\omega = 0$ ), see Sect. 13.2.1. The centred schemes are the **Flux LImiter Centred**, or **FLIC**, scheme (Sect. 13.7.4) and the **Slope LImiter Centred**, or **SLIC**, scheme (Sect. 13.8). In the centred TVD schemes we use the FORCE flux as the low-order flux. All of these schemes are extended to non-linear systems in Chap. 14.

Results from the flux limiter methods are shown in Figs. 13.25 and 13.26 for Tests 1 and 2 respectively. Results from the slope limiter methods are shown in Figs. 13.27 and 13.28 for the same tests. The results of the left column in each of the figures are those of the upwind based scheme and those on the right column are obtained with the centred TVD scheme. Four *types* of limiter functions have been used, namely SUPERBEE, VANLEER, VANALBADA and MINBEE; *note that the limiters of the upwind scheme are not equivalent to those of the centred scheme*.

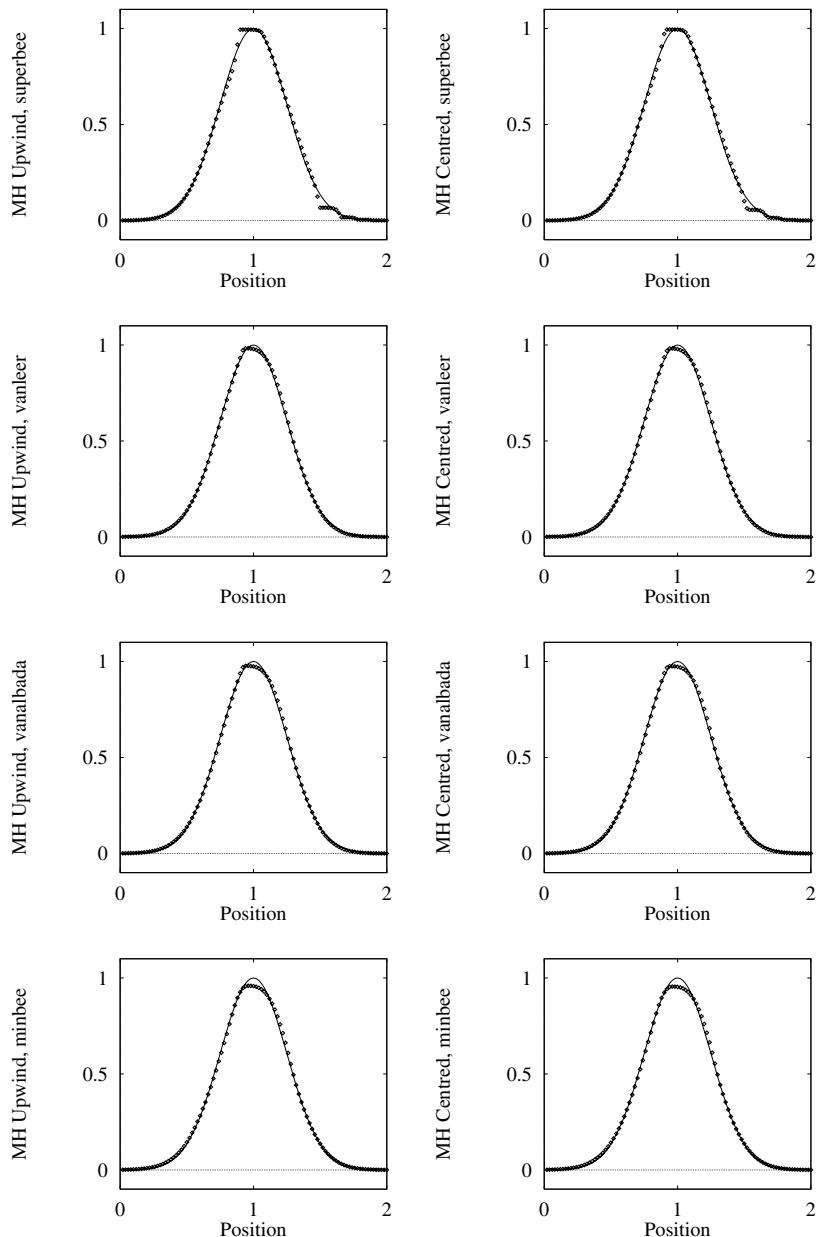
The main conclusion that may be drawn from the results is that there appears to be no obvious advantages, at least for this scalar problem, in using the upwind based TVD scheme. We note however, as will be seen in Chap. 14, that for non-linear systems, upwind methods give more accurate solutions.



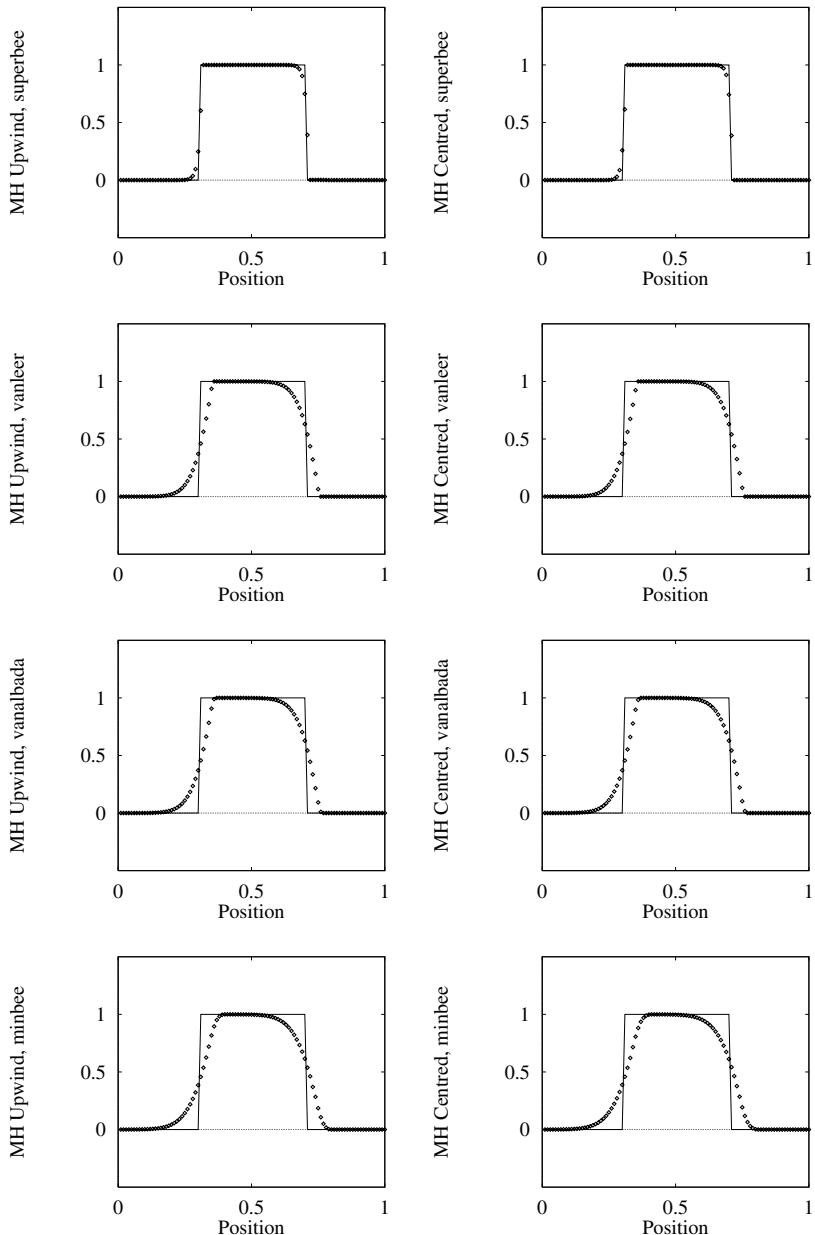
**Fig. 13.25.** Numerical (symbols) and exact (line) solutions for Tests 1 at time  $t = 10$ . Results from Flux-Limiter Upwind and Flux-Limiter Centred (FORCE) are compared using various flux limiters.



**Fig. 13.26.** Numerical (symbols) and exact (line) solutions for Tests 2 at time  $t = 10$ . Results from Flux-Limiter Upwind and Flux-Limiter Centred (FORCE) are compared using various flux limiters.



**Fig. 13.27.** Numerical (symbols) and exact (line) solutions for Tests 2 at time  $t = 10$ . Results from MUSCL–Hancock Upwind and MUSCL–Hancock Centred (FORCE) are compared using various slope limiters.



**Fig. 13.28.** Numerical (symbols) and exact (line) solutions for Tests 2 at time  $t = 10$ . Results from MUSCL–Hancock Upwind and MUSCL–Hancock Centred (FORCE) are compared using various slope limiters.

---

## High-Order and TVD Schemes for Non-Linear Systems

This chapter is concerned with TVD *upwind* and *centred* schemes for non-linear systems of conservation laws that depend on time  $t$ , or a time-like variable  $t$ , and one space dimension  $x$ . The upwind schemes are extensions of the Godunov first order upwind method of Chap. 6 and can be applied with any of the Riemann solvers presented in Chap. 4 (exact) and Chaps. 9 to 12 (approximate); they can also be used with the Flux Vector Splitting flux of Chap. 8. The centred schemes are extensions of the First Order Centred (FORCE) method presented in Chap. 7. All the TVD schemes are in effect the culmination of work carried out in all previous chapters, particularly Chap. 13, where the TVD concept was developed in the context of simple scalar problems. The schemes are presented in terms of the time-dependent one dimensional Euler equations for ideal gases, which are introduced in Chap. 1 and studied in detail in Chap. 3. Applications to other systems may be easily accomplished. Techniques for extending the methods to systems with source terms, as for reactive flows for instance, are given in Chap. 15 and to multidimensional systems in Chap. 16.

### 14.1 Introduction

We study fully discrete, explicit methods that are suitable for time-dependent problems. *Upwind* TVD schemes are more accurate than their *centred* TVD counterparts; a disadvantage of upwind schemes is their complexity and the computing cost due to the solution of the Riemann problem, see Chaps. 4, 6, 8, 9, 10, 11 and 12. For ideal gases the extra expense of upwind schemes is not significant, if efficient Riemann solvers are used. The TVD *centred* schemes presented are considerably simpler to implement and somewhat more efficient to apply. They are particularly recommended to the reader who does not want to get involved with the details of the Riemann problem. Also, for complicated sets of conservation laws for which the solution of the Riemann problem is not available, or is complex, or is prohibitively

expensive to compute, then TVD centred schemes are definitely an option to consider.

The methods are applied to the Initial-Boundary Value Problem

$$\left. \begin{array}{l} PDEs : \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0}, \\ ICs : \mathbf{U}(x, 0) = \mathbf{U}^{(0)}(x), \\ BCs : \mathbf{U}(0, t) = \mathbf{U}_l(t); \quad \mathbf{U}(L, t) = \mathbf{U}_r(t). \end{array} \right\} \quad (14.1)$$

We assume the spatial domain  $[0, L]$  to be discretised into  $M$  computing cells  $I_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ , and for simplicity we assume the mesh size  $\Delta x = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$  to be constant. The time-dependent Euler equations are typical conservation laws in (14.1), for which the vectors  $\mathbf{U}$  of conserved variables and  $\mathbf{F} = \mathbf{F}(\mathbf{U})$  of fluxes are

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ E \end{bmatrix}; \quad \mathbf{F}(\mathbf{U}) = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ u(E + p) \end{bmatrix} \quad (14.2)$$

respectively. Here  $\rho$  is density,  $u$  is velocity,  $p$  is pressure and  $E$  is total energy per unit volume. See Chaps. 1 and 3 for background on the Euler equations.

We solve the general IBVP (14.1) using the explicit, conservative fully discrete scheme

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n + \frac{\Delta t}{\Delta x} [\mathbf{F}_{i-\frac{1}{2}} - \mathbf{F}_{i+\frac{1}{2}}], \quad (14.3)$$

where  $\mathbf{F}_{i+\frac{1}{2}}$  is the numerical flux at the cell interface position  $x_{i+\frac{1}{2}}$ . For background on the meaning of the conservative formula (14.3) the reader is referred to Sect. 5.3 of Chap. 5 and Chap. 6. The choice of the time step  $\Delta t$  in (14.3) and the implementation of boundary conditions is addressed in Sect. 14.2. The choice of the numerical flux  $\mathbf{F}_{i+\frac{1}{2}}$  in (14.3) determines the scheme. Here we study two upwind-based second order TVD schemes and two centred second-order TVD schemes. These are extensions of schemes presented in detail for scalar conservation laws in Chap. 13. In addition, we present two upwind second order TVD schemes based on *primitive variables* and adaptive primitive-conservative variations of these.

The rest of this chapter is organised as follows: Sect. 14.2 discusses the application of the CFL condition to compute the time step  $\Delta t$  and the implementation of boundary conditions. Sects. 14.3 and 14.4 present two upwind TVD schemes, namely the WAF method and the MUSCL-Hancock scheme respectively. Sect. 14.5 presents two centred TVD schemes; the first method, called FLIC, is a flux limiter extension of the First Order Centred (FORCE) scheme; the other, called SLIC, is a slope limiter extension of FORCE. Sect. 14.6 presents primitive variable schemes and an adaptive primitive-conservative scheme. Numerical results are shown in Sect. 14.7.

## 14.2 CFL and Boundary Conditions

Before describing particular schemes contained in (14.3) by specifying the numerical flux, we briefly address two basic problems, namely the choice of the time step  $\Delta t$  and application of boundary conditions. See Sects. 6.3.2 and 6.3.3 of Chap. 6 for full discussion on these topics, in the context of the Godunov first order upwind method.

All methods considered in this chapter have linearised stability constraint  $|c| \leq 1$ , where  $c$  is the Courant number. For non-linear systems we implement this condition as

$$\Delta t = C_{cfl} \frac{\Delta x}{S_{max}^{(n)}}, \quad (14.4)$$

where  $\Delta x$  is the mesh spacing,  $S_{max}^{(n)}$  is the maximum wave speed present at time level  $n$  and  $C_{cfl}$  is the CFL coefficient, with  $C_{cfl} \in (0, 1]$ . See Sect. 6.3.2 of Chap. 6 for details on possible choices for  $S_{max}^{(n)}$ . A practical choice, which must be used with caution, is

$$S_{max}^{(n)} = \max_i \{|u_i^n| + a_i^n\}, \quad (14.5)$$

where the range for  $i$  must include data arising from boundary conditions. As remarked in Sect. 6.3.2 of Chap. 6, inappropriate choices for  $S_{max}^{(n)}$  in (14.4) can result in the scheme becoming unstable, even for *small* values of  $C_{cfl}$ .

A detailed discussion on boundary conditions for the first-order Godunov method is given in Sect. 6.3.3 of Chap. 6, which the reader is encouraged to consult before proceeding. For the second order methods discussed in this chapter the application of boundary conditions is fundamentally the same as for the Godunov method. We assume the computational domain  $[0, L]$  to be discretised by  $M$  cells  $I_i$ , so that cells  $i = 1, \dots, M$  lie within the computational domain. In applying boundary conditions we now require *two* fictitious cells next to each boundary. For the left boundary at  $x = 0$  the fictitious cells are denoted by  $i = -1$  and  $i = 0$  and for the right boundary they are denoted by  $i = M + 1$  and  $i = M + 2$ . We discuss two types of boundary conditions.

*Transmissive boundary conditions* are given by

$$\left. \begin{array}{l} \mathbf{W}_0^n = \mathbf{W}_1^n; \quad \mathbf{W}_{-1}^n = \mathbf{W}_2^n, \\ \mathbf{W}_{M+1}^n = \mathbf{W}_M^n; \quad \mathbf{W}_{M+2}^n = \mathbf{W}_{M-1}^n, \end{array} \right\} \quad (14.6)$$

where  $\mathbf{W}$  may be the vector of conserved variables or some other variables, such as the primitive variables.

*Reflective boundary conditions* are applied at reflective, moving or stationary, boundaries. Suppose the speed of a reflective solid boundary at  $x = L$  is  $u_{wall}$ , then the reflective boundary conditions for the Euler equations are applied as

$$\left. \begin{aligned} \rho_{M+1}^n &= \rho_M^n ; & u_{M+1}^n &= -u_M^n + 2u_{wall} ; & p_{M+1}^n &= p_M^n , \\ \rho_{M+2}^n &= \rho_{M-1}^n ; & u_{M+2}^n &= -u_{M-1}^n + 2u_{wall} ; & p_{M+2}^n &= p_{M-1}^n . \end{aligned} \right\} \quad (14.7)$$

For a moving reflective left boundary the situation is entirely analagous to the case just given.

Having decided on the application of boundary conditions and on a strategy to compute the time step  $\Delta t$ , all that remains to be discussed is the selection of the intercell flux  $\mathbf{F}_{i+\frac{1}{2}}$ , for a scheme (14.3) to be completely determined. Sects. 14.3 to 14.6 present various ways of choosing the intercell flux.

### 14.3 Weighted Average Flux (WAF) Schemes

The Weighted Average Flux (WAF) approach for obtaining second-order extensions of the Godunov first-order upwind method has its origins in the Random Flux Method [497], [539]. WAF is a deterministic approach and has two versions. In the first the intercell flux is an integral average of the physical flux across the full structure of the solution of a local Riemann problem. The second version computes an averaged state and then  $\mathbf{F}_{i+\frac{1}{2}}$  results from evaluating the physical flux  $\mathbf{F}$  at this state.

Useful background is found in Sect. 13.3 of Chap. 13, where the WAF approach is presented in detail, when applied to scalar equations.

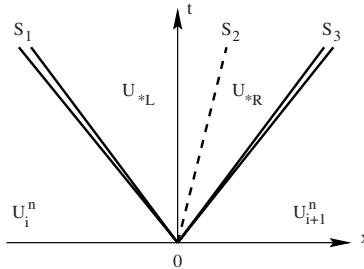
#### 14.3.1 The Original Version of WAF

The simplest WAF flux is given as

$$\mathbf{F}_{i+\frac{1}{2}} = \frac{1}{\Delta x} \int_{-\frac{1}{2}\Delta x}^{\frac{1}{2}\Delta x} \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}(x, \frac{\Delta t}{2})) dx , \quad (14.8)$$

where  $\mathbf{U}_{i+\frac{1}{2}}(x, t)$  is the solution of the Riemann problem with *piece-wise constant data*  $\mathbf{U}_i^n, \mathbf{U}_{i+1}^n$  at the interface position  $i + \frac{1}{2}$ .  $\mathbf{U}_{i+\frac{1}{2}}(x, t)$  is exactly the same solution as for the Godunov first-order upwind method, see Chap. 6. Chap. 4 gives the detailed, exact solution of the Riemann problem for the Euler equations. The approximate Riemann solvers of Chaps. 9 to 12 may also be used here. Fig. 14.1 depicts the structure of the solution of the Riemann problem for the Euler equations. There are three waves of speeds  $S_1, S_2, S_3$  that separate four constant states:  $\mathbf{U}_i^n, \mathbf{U}_{*L}, \mathbf{U}_{*R}, \mathbf{U}_{i+1}^n$ . The left and right waves may be shock or rarefaction waves. In the latter case there will be additional, non-uniform states present, namely those inside rarefaction fans. Fig 14.2 illustrates the evaluation of the integral average (14.8), for a wave structure assumed to contain no rarefaction waves. By setting

$$\mathbf{U}^{(1)} = \mathbf{U}_i^n ; \quad \mathbf{U}^{(2)} = \mathbf{U}_{*L} ; \quad \mathbf{U}^{(3)} = \mathbf{U}_{*R} ; \quad \mathbf{U}^{(4)} = \mathbf{U}_{i+1}^n \quad (14.9)$$



**Fig. 14.1.** Structure of the solution of the Riemann problem with data  $\mathbf{U}_i^n, \mathbf{U}_{i+1}^n$

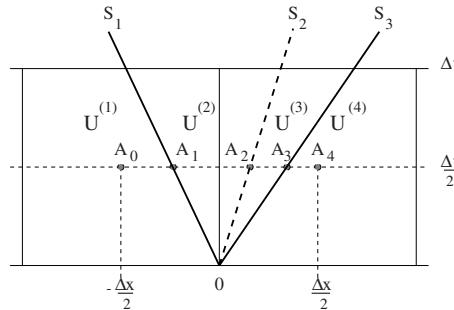
the integral (14.8) gives

$$\mathbf{F}_{i+\frac{1}{2}} = \sum_{k=1}^{N+1} \beta_k \mathbf{F}_{i+\frac{1}{2}}^{(k)}, \quad (14.10)$$

where  $\mathbf{F}_{i+\frac{1}{2}}^{(k)} = \mathbf{F}(\mathbf{U}^{(k)})$ ,  $N$  is the number of waves in the solution of the Riemann problem and  $\beta_k$ ,  $k = 1, \dots, 4$ , are the normalised lengths of the segments  $A_{k-1}A_k$ ,

$$\beta_k = \frac{|A_{k-1}A_k|}{\Delta x}.$$

It can be seen easily that, in terms of the wave speeds  $S_k$ , the weights are



**Fig. 14.2.** Evaluation of the WAF intercell flux. Integral average becomes summation involving flux terms in regions 1 to 4 with weights obtained by normalising lengths  $|A_0A_1|$  to  $|A_3A_4|$

$$\left. \begin{aligned} \beta_k &= \frac{1}{2}(c_k - c_{k-1}), \\ c_k &= \frac{\Delta t S_k}{\Delta x}, \quad c_0 = -1, \quad c_{N+1} = 1. \end{aligned} \right\} \quad (14.11)$$

Here  $c_k$  is the Courant number for wave  $k$  of speed  $S_k$ . Formula (14.10) is a *weighted average flux*, and thus the name of the scheme.

Substitution of  $\beta_k$  in (14.11) into (14.10) gives an alternative form for the WAF flux

$$\mathbf{F}_{i+\frac{1}{2}} = \frac{1}{2}(\mathbf{F}_i + \mathbf{F}_{i+1}) - \frac{1}{2} \sum_{k=1}^N c_k \Delta \mathbf{F}_{i+\frac{1}{2}}^{(k)}, \quad (14.12)$$

where

$$\Delta \mathbf{F}_{i+\frac{1}{2}}^{(k)} = \mathbf{F}_{i+\frac{1}{2}}^{(k+1)} - \mathbf{F}_{i+\frac{1}{2}}^{(k)} \quad (14.13)$$

is the flux jump across wave  $k$  of CFL number  $c_k$ .

### 14.3.2 A Weighted Average State Version

A possible variant of the scheme is obtained by first defining a *weighted average state* as

$$\bar{\mathbf{W}}_{i+\frac{1}{2}} = \frac{1}{\Delta x} \int_{-\frac{1}{2}\Delta x}^{\frac{1}{2}\Delta x} \mathbf{W}_{i+\frac{1}{2}}(x, \frac{1}{2}\Delta t) dx, \quad (14.14)$$

where  $\mathbf{W}$  is a suitable vector of variables and  $\mathbf{W}_{i+\frac{1}{2}}$  is the solution of the Riemann problem with data  $\mathbf{W}_i^n, \mathbf{W}_{i+1}^n$ . For the assumed wave structure of Fig. 14.2 this integral becomes

$$\bar{\mathbf{W}}_{i+\frac{1}{2}} = \sum_{k=1}^N \beta_k \mathbf{W}_{i+\frac{1}{2}}^{(k)}, \quad (14.15)$$

where the weights  $\beta_k$  are as given by (14.11) and  $\mathbf{W}_{i+\frac{1}{2}}^{(k)}$  is the value of  $\mathbf{W}_{i+\frac{1}{2}}$  in region  $k$ . Expanding (14.15) gives

$$\bar{\mathbf{W}}_{i+\frac{1}{2}} = \frac{1}{2}(\mathbf{W}_i^n + \mathbf{W}_{i+1}^n) - \frac{1}{2} \sum_{k=1}^N c_k [\mathbf{W}_{i+\frac{1}{2}}^{(k+1)} - \mathbf{W}_{i+\frac{1}{2}}^{(k)}]. \quad (14.16)$$

Compare (14.15), (14.16) with (14.10), (14.12). An intercell flux is now defined as

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}(\bar{\mathbf{W}}_{i+\frac{1}{2}}). \quad (14.17)$$

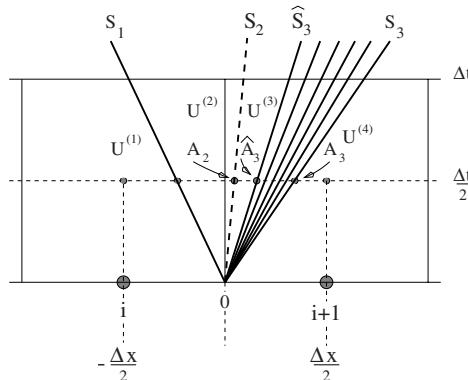
As to the choice of the vector  $\mathbf{W}$  in (14.17) we recommend the set of primitive variables, i.e.  $\mathbf{W} = (\rho, u, p)^T$ .

Computationally, formulation (14.17), (14.16) is more efficient than formulation (14.12); it involves only one flux evaluation per component of the system.

### 14.3.3 Rarefactions in State Riemann Solvers

When the Riemann solver provides a state at which the physical flux is to be evaluated, such as in the exact Riemann solver of Chap. 4 and the approximate state Riemann solvers of Chap. 9, one requires a special treatment of rarefaction waves in the WAF method. For other Riemann solvers such as HLL and HLLC of Chap. 10 such special treatment is not needed.

Let us consider the case of the exact Riemann solver. In the presence of rarefaction waves in the solution  $\mathbf{U}_{i+\frac{1}{2}}(x, t)$  of the Riemann problem, definitions (14.8) and (14.14) still apply. As the exact solution across rarefactions is available, see Chap. 4, the exact evaluation of (14.8) or (14.14) is possible and the result is again a summation like (14.10) or (14.15), with extra terms to account for regions occupied by rarefactions.



**Fig. 14.3.** Treatment of right non-sonic rarefaction when evaluating the WAF flux

For the Euler equations, waves 1 (left) and 3 (right) may be shocks or rarefactions; wave 2 (middle) is always a contact discontinuity of speed  $S_2 = u_*$ . See Sect. 3.1.3 of Chap. 3. Fig. 14.3 depicts the case in which the right wave is a rarefaction, whose bounding characteristics have speeds

$$\hat{S}_3 = S_{TR} = u_* + a_{*R} \quad (\text{Tail}) ; \quad S_3 = S_{HR} = u_{i+1}^n + a_{i+1}^n \quad (\text{Head}) .$$

Evaluation of the integral (14.8) gives one term for the zone across the rarefaction, namely

$$\frac{1}{\Delta x} \int_{\hat{\mathbf{A}}_3}^{\mathbf{A}_3} \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}(x, \frac{\Delta t}{2})) dx , \quad (14.18)$$

which can be found exactly, if desired. This results in one extra state, one extra weight  $\beta_k$  and one extra wave speed  $\hat{S}_3$ .

Computational experience suggests that for the purpose of evaluating the integral (14.8) or (14.14) one may ‘lump’ the rarefaction state together with

the closest *constant* state in the direction of the  $t$ -axis. The integral in the resulting enlarged zone,  $\mathbf{A}_2$  to  $\mathbf{A}_3$  in Fig. 14.3, is then approximated as

$$\frac{1}{\Delta x} \int_{\mathbf{A}_2}^{\mathbf{A}_3} \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}(x, \frac{\Delta t}{2})) dx \approx \frac{1}{2}(c_3 - c_2)\mathbf{F}_{i+\frac{1}{2}}^{(3)}, \quad (14.19)$$

where

$$\left. \begin{aligned} S_3 &= S_{HR} = u_{i+1}^n + a_{i+1}^n, \\ \mathbf{F}_{i+\frac{1}{2}}^{(3)} &= \begin{cases} \mathbf{F}(\mathbf{U}^{(3)}) & : \text{non-sonic rarefaction}, \\ \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}(0)) & : \text{sonic rarefaction}. \end{cases} \end{aligned} \right\} \quad (14.20)$$

Fig. 14.4 illustrates the case in which the right rarefaction is sonic. The required state lies inside the rarefaction along the  $t$ -axis. Recall that for ideal gases the solution inside rarefactions is given in closed form, see Sect. 4.4 of Chap. 4.

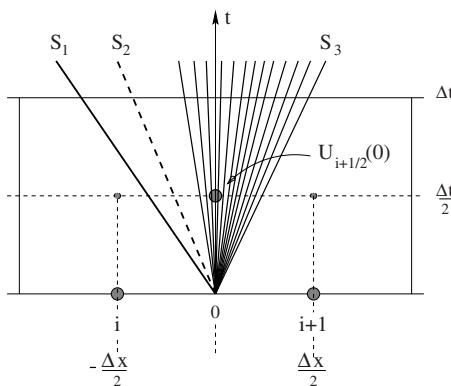
For a left rarefaction we have

$$\frac{1}{\Delta x} \int_{\mathbf{A}_1}^{\mathbf{A}_2} \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}(x, \frac{\Delta t}{2})) dx \approx \frac{1}{2}(c_2 - c_1)\mathbf{F}_{i+\frac{1}{2}}^{(2)}, \quad (14.21)$$

where

$$\left. \begin{aligned} S_1 &= S_{HL} = u_i^n - a_i^n, \\ \mathbf{F}_{i+\frac{1}{2}}^{(2)} &= \begin{cases} \mathbf{F}(\mathbf{U}^{(2)}) & : \text{non-sonic rarefaction}, \\ \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}(0)) & : \text{sonic rarefaction}. \end{cases} \end{aligned} \right\} \quad (14.22)$$

Note that the flux term evaluated at  $\mathbf{U}_{i+\frac{1}{2}}(0)$  is the Godunov flux, which is



**Fig. 14.4.** Treatment of right sonic rarefaction when evaluating the WAF flux

always part of the WAF flux.

The special treatment of rarefactions, just described, also applies when using the Approximate–State Riemann Solvers of Chap. 9. When using the HLL and HLLC Riemann solvers of Chap. 10, no need for this special treatment arises; this is also the case if one was to use the Flux Vector Splitting intercell flux of Chap. 8.

#### 14.3.4 TVD Version of WAF Schemes

As seen in Sect. 13.3 of Chap. 13, the WAF schemes are second order accurate in space and time. According to Godunov’s Theorem, Sect. 13.5.3 of Chap. 13, spurious oscillations in the vicinity of high gradients are to be expected. Here we extend the Total Variation Diminishing (TVD) version of the WAF method derived in Sect. 13.7.1 to non–linear systems.

Strictly speaking, the extension to non–linear systems is somewhat empirical but is found to work well in practice. In the scalar case, for which the TVD property was enforced rigourously, there is one conservation law and one wave. In the case of the Euler equations in (14.1), (14.2) there are three conservation laws and three wave families. For the purpose of applying a TVD constraint to non–linear systems, a valuable, though empirical, observation is that the solution to the complete system may be characterised by *jumps in a single quantity*  $q$ , across each of the three waves. We therefore require the computation of three limiter functions  $\phi_{i+\frac{1}{2}}(r)$  per intercell boundary. The reader is encouraged to review carefully Sect 13.7.1 before proceeding with the rest of this section.

The TVD modification of the WAF flux (14.12) is

$$\mathbf{F}_{i+\frac{1}{2}} = \frac{1}{2}(\mathbf{F}_i + \mathbf{F}_{i+1}) - \frac{1}{2} \sum_{k=1}^N \text{sign}(c_k) \phi_{i+\frac{1}{2}}^{(k)} \Delta \mathbf{F}_{i+\frac{1}{2}}^{(k)}, \quad (14.23)$$

where

$$\phi_{i+\frac{1}{2}}^{(k)} = \phi_{i+\frac{1}{2}}(r^{(k)}) \quad (14.24)$$

is a WAF limiter function, as derived in Sect. 13.7.1 of Chap. 13. The flow parameter  $r^{(k)}$  refers to wave  $k$  in the solution  $\mathbf{U}_{i+\frac{1}{2}}(x, t)$  of the Riemann problem and is the ratio

$$r^{(k)} = \begin{cases} \frac{\Delta q_{i-\frac{1}{2}}^{(k)}}{\Delta q_{i+\frac{1}{2}}^{(k)}}, & \text{if } c_k > 0, \\ \frac{\Delta q_{i+\frac{1}{2}}^{(k)}}{\Delta q_{i+\frac{3}{2}}^{(k)}}, & \text{if } c_k < 0. \end{cases} \quad (14.25)$$

As has already been indicated, one selects a single quantity  $q$  which is known to change across each wave family in the solution of the Riemann problem.

For the Euler equations the choices  $q \equiv \rho$  (density) or  $q \equiv e$  (specific internal energy) give very satisfactory results. For multidimensional problems, see Chap. 16, jumps in tangential velocity components must also be used.  $\Delta q_{i-\frac{1}{2}}^{(k)}$  denotes the jump in  $q$  across wave  $k$  in the solution  $\mathbf{U}_{i-\frac{1}{2}}(x, t)$  of the Riemann problem with data  $(\mathbf{U}_{i-1}^n, \mathbf{U}_i^n)$ ,  $\Delta q_{i+\frac{3}{2}}^{(k)}$  is the jump in  $q$  across wave  $k$  in the solution  $\mathbf{U}_{i+\frac{3}{2}}(x, t)$  of the Riemann problem with data  $(\mathbf{U}_{i+1}^n, \mathbf{U}_{i+2}^n)$ ;  $\Delta q_{i+\frac{1}{2}}^{(k)}$  is the corresponding jump across wave  $k$  in the *local* Riemann problem solution  $\mathbf{U}_{i+\frac{1}{2}}(x, t)$  with data  $(\mathbf{U}_i^n, \mathbf{U}_{i+1}^n)$ .

The TVD version of the weighted average state version of the scheme has

$$\overline{\mathbf{W}}_{i+\frac{1}{2}} = \frac{1}{2}(\mathbf{W}_i^n + \mathbf{W}_{i+1}^n) - \frac{1}{2} \sum_{k=1}^N \text{sign}(c_k) \phi_k (\mathbf{W}_{i+\frac{1}{2}}^{(k+1)} - \mathbf{W}_{i+\frac{1}{2}}^{(k)}) \quad (14.26)$$

and the TVD flux is

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}(\overline{\mathbf{W}}_{i+\frac{1}{2}}).$$

As to the WAF limiters  $\phi$  one may use any of the following

$$\phi_{sa}(r, |c|) = \begin{cases} 1 & \text{if } r \leq 0, \\ 1 - 2(1 - |c|)r & \text{if } 0 \leq r \leq \frac{1}{2}, \\ |c| & \text{if } \frac{1}{2} \leq r \leq 1, \\ 1 - (1 - |c|)r & \text{if } 1 \leq r \leq 2, \\ 2|c| - 1 & \text{if } r \geq 2, \end{cases} \quad (14.27)$$

$$\phi_{vl}(r, |c|) = \begin{cases} 1 & \text{if } r \leq 0, \\ 1 - \frac{(1-|c|)2r}{1+r} & \text{if } r \geq 0, \end{cases} \quad (14.28)$$

$$\phi_{va}(r, |c|) = \begin{cases} 1 & \text{if } r \leq 0, \\ 1 - \frac{(1-|c|)r(1+r)}{(1+r^2)} & \text{if } r \geq 0, \end{cases} \quad (14.29)$$

$$\phi_{ma}(r, |c|) = \begin{cases} 1 & \text{if } r \leq 0, \\ 1 - (1 - |c|)r & \text{if } 0 \leq r \leq 1, \\ |c| & \text{if } r \geq 1. \end{cases} \quad (14.30)$$

For convenience, subscripts and superscripts have been omitted. The WAF limiter functions are related, and are entirely equivalent, to conventional flux limiters  $\psi(r)$  via

$$\phi(r) = 1 - (1 - |c|)\psi(r).$$

$\phi_{sa}$  is related to SUPERBEE,  $\phi_{vl}$  is related to van Leer's limiter,  $\phi_{va}$  is related to van Albada's limiter and  $\phi_{ma}$  is related to MINBEE. See Sect. 13.7.1 of Chap. 13, for details on the derivation of limiter functions for the WAF method.

### 14.3.5 Riemann Solvers

The WAF method may be used with the exact Riemann solver or approximate Riemann solvers. An exact Riemann solver is given in Chap. 4 for the Euler equations. Approximate Riemann solvers for the Euler equations are given in Chaps. 9 to 12; one may also use the Flux–Vector splitting intercell flux given in Chap. 8. We particularly recommend the adaptive, approximate-state Riemann solvers of Chap. 9 and the HLLC approximate Riemann solver presented in Chap. 10 for the Euler equations. The WAF approach may be generalised to irregular grids. A Courant–number 2 extension of the WAF method is reported in [527]. Numerical results are presented in Sect. 14.7

### 14.3.6 Summary of the WAF Method

For a given domain  $[0, L]$  and a mesh size  $\Delta x$ , determined by the choice of the number  $M$  of computing cells, one first sets the initial conditions  $\mathbf{U}^0 \equiv \mathbf{U}(x, 0)$  at time  $t = 0$ . Then, for each time step  $n$  one performs the following operations

- Operation (I): **Boundary conditions.** This is carried out according to (14.6) or (14.7), for example.
- Operation (II): **Computation of time step.** This is carried out according to (14.4) and (14.5). A choice of  $C_{cfl}$  must be made at the beginning of the computations. One usually takes  $C_{cfl} = 0.9$ . Recall however that the choice of  $S_{max}^{(n)}$  is crucial and given that (14.5) produces somewhat unreliable estimates for the true speeds we recommend that when solving problems with shock–tube like data, the CFL coefficient  $C_{cfl}$  be set to small number, e.g. 0.2, for a few time steps, e.g. 5.
- Operation (III): **Solution of Riemann problem.** For each pair of data states  $(\mathbf{U}_i^n, \mathbf{U}_{i+1}^n)$ , compute the solution of Riemann problem and compute: states  $\mathbf{U}^k$ , fluxes  $\mathbf{F}_{i+\frac{1}{2}}^k$ , wave speeds  $S_k$  and wave jumps  $\Delta q_{i+\frac{1}{2}}$ , where  $q$  may be chosen to be density, when solving the Euler equations.

**Operation (IV): TVD intercell flux.** At each cell interface  $i + \frac{1}{2}$  compute the ratios  $r^k$  according to (14.25), compute Courant numbers  $c_k$  according to (14.11), compute limiter functions according to any of (14.27) to (14.30), and, finally compute the intercell flux at position  $i + \frac{1}{2}$  according to (14.23); alternatively one computes the intercell flux according to (14.17), (14.26).

**Operation (V): Updating of solution.** Proceed to update the conserved variables according to the conservative formula (14.3)

**Operation (VI): Next time level.** Go to (I).

## 14.4 The MUSCL–Hancock Scheme

Here we present a scheme that achieves second order accuracy by applying the MUSCL approach [554], [555], [556], [557], [558], [559], [562]. MUSCL stands for Monotonic Upstream-Centred Scheme for Conservation Laws. An introduction to this general approach was given in Sect. 13.4 of Chap. 13, in the context of scalar conservation laws. Three MUSCL type schemes were presented in detail, namely the MUSCL–Hancock method [562], [400], the Piece-Wise Linear method (PLM) of Colella [133] and the Generalised Riemann Problem (GRP) approach of Ben-Artzi and Falcovitz [37], [38], [36]. The reader is encouraged to review Sects. 13.4.1 and 13.4.2 of Chap. 13 before proceeding.

### 14.4.1 The Basic Scheme

In this section we present the extension of the MUSCL–Hancock scheme to any system of non-linear hyperbolic conservation laws

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0} \quad (14.31)$$

and discuss the details when (14.31) are the Euler equations, with the vector of conserved variables  $\mathbf{U}$  and fluxes  $\mathbf{F}(\mathbf{U})$  as given by (14.2). The objective is to construct a flux  $\mathbf{F}_{i+\frac{1}{2}}$  in (14.3) that is a second order extension of the Godunov first-order upwind method, Chap. 6. Given that the bases of the scheme were discussed in great detail in Sects. 13.4.1 and 13.4.2 of Chap. 13, we restrict ourselves here to a succinct, algorithmic presentation of the method.

The MUSCL–Hancock approach achieves a second order extension of the Godunov first order upwind method if the intercell flux  $\mathbf{F}_{i+\frac{1}{2}}$  is computed according to the following steps:

**Step (I): Data Reconstruction.** In the following, we assume a choice of variables  $\mathbf{W}$  has been made. One possibility is the conserved variables  $\mathbf{W} = \mathbf{U}$  of course; another choice is offered by the *primitive*

or physical variables. For the Euler equations (14.1), (14.2) these are  $\mathbf{W} = (\rho, u, p)^T$ . Here we present the scheme in terms of the conserved variables  $\mathbf{U}$ . In the data reconstruction step, data cell average values  $\mathbf{U}_i^n$  are locally replaced by piecewise linear functions in each cell  $I_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ , namely

$$\mathbf{U}_i(x) = \mathbf{U}_i^n + \frac{(x - x_i)}{\Delta x} \Delta_i, \quad x \in [0, \Delta x]. \quad (14.32)$$

$\Delta_i$  is a suitably chosen slope vector (actually a difference) of  $\mathbf{U}_i(x)$  in cell  $I_i$ . The extreme points  $x = 0$  and  $x = \Delta x$ , in local coordinates, correspond to the intercell boundaries  $x_{i-\frac{1}{2}}$  and  $x_{i+\frac{1}{2}}$ , in global coordinates, respectively. The values of  $\mathbf{U}_i(x)$  at the extreme points are

$$\mathbf{U}_i^L = \mathbf{U}_i^n - \frac{1}{2} \Delta_i; \quad \mathbf{U}_i^R = \mathbf{U}_i^n + \frac{1}{2} \Delta_i \quad (14.33)$$

and are usually called boundary extrapolated values. Note that  $\mathbf{U}$  and  $\Delta_i$  are vectors of three components for the Euler equations and thus there are six scalar extrapolated values in (14.33).

- Step (II): **Evolution.** For each cell  $I_i$ , the boundary extrapolated values  $\mathbf{U}_i^L, \mathbf{U}_i^R$  in (14.33) are evolved by a time  $\frac{1}{2}\Delta t$ , according to

$$\left. \begin{aligned} \bar{\mathbf{U}}_i^L &= \mathbf{U}_i^L + \frac{1}{2} \frac{\Delta t}{\Delta x} [\mathbf{F}(\mathbf{U}_i^L) - \mathbf{F}(\mathbf{U}_i^R)], \\ \bar{\mathbf{U}}_i^R &= \mathbf{U}_i^R + \frac{1}{2} \frac{\Delta t}{\Delta x} [\mathbf{F}(\mathbf{U}_i^L) - \mathbf{F}(\mathbf{U}_i^R)]. \end{aligned} \right\} \quad (14.34)$$

Note that this evolution step is entirely contained in each cell  $I_i$ , as the intercell fluxes are evaluated at the boundary extrapolated values of each cell. At each intercell position  $i + \frac{1}{2}$  there are two fluxes, namely  $\mathbf{F}(\mathbf{U}_i^R)$  and  $\mathbf{F}(\mathbf{U}_{i+1}^L)$ , which are in general distinct. This does not really affect the conservative character of the overall method, as this step is only an intermediate step; the intercell flux  $\mathbf{F}_{i+\frac{1}{2}}$  to be used in (14.3) is yet to be evaluated.

- Step (III): **The Riemann Problem.** To compute the intercell flux  $\mathbf{F}_{i+\frac{1}{2}}$  one now solves the conventional Riemann problem with data

$$\mathbf{U}_L \equiv \bar{\mathbf{U}}_i^R; \quad \mathbf{U}_R \equiv \bar{\mathbf{U}}_{i+1}^L \quad (14.35)$$

to obtain the similarity solution  $\mathbf{U}_{i+\frac{1}{2}}(x/t)$ . The intercell flux  $\mathbf{F}_{i+\frac{1}{2}}$  is now computed in exactly the same way as in the Godunov first order upwind method, namely

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}(0)). \quad (14.36)$$

Here  $\mathbf{U}_{i+\frac{1}{2}}(0)$  denotes the value of  $\mathbf{U}_{i+\frac{1}{2}}(x/t)$  at  $x/t = 0$ , i.e. the value of  $\mathbf{U}_{i+\frac{1}{2}}(x/t)$  along the  $t$ -axis. Fig. 14.5 illustrates only one of the ten possible wave patterns in the solution of the Riemann problem that must be taken into account when computing the Godunov flux. See Chaps. 4 and 6 for details.

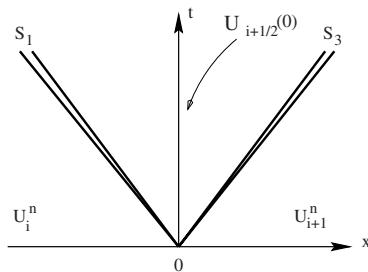
A possible choice for the *slope* vector  $\Delta_i$  in (14.32), (14.33) is

$$\Delta_i = \frac{1}{2}(1 + \omega)\Delta_{i-\frac{1}{2}} + \frac{1}{2}(1 - \omega)\Delta_{i+\frac{1}{2}}, \quad (14.37)$$

where

$$\Delta_{i-\frac{1}{2}} \equiv \mathbf{U}_i^n - \mathbf{U}_{i-1}^n; \quad \Delta_{i+\frac{1}{2}} \equiv \mathbf{U}_{i+1}^n - \mathbf{U}_i^n. \quad (14.38)$$

and  $\omega \in [-1, 1]$ .



**Fig. 14.5.** Evaluation of the Godunov flux for a particular wave pattern occurring in the solution of the local Riemann problem

#### 14.4.2 A Variant of the Scheme

It is possible to modify steps (I) and (II) above by formulating the reconstruction process in terms of a vector  $\mathbf{W}$  of primitive variables. For the Euler equations we may choose  $\mathbf{W} = (\rho, u, p)^T$  and we have

$$\mathbf{W}_t + \mathbf{A}(\mathbf{W})\mathbf{W}_x = \mathbf{0}. \quad (14.39)$$

Then steps (I) and (II) can be reduced to a single step. As the evolution step does not need to be performed in terms of the conserved variables, we replace (14.34) by the scheme

$$\bar{\mathbf{W}}_i^{L,R} = \mathbf{W}_i^{L,R} + \frac{1}{2} \frac{\Delta t}{\Delta x} \mathbf{A}(\mathbf{W}_i^n)[\mathbf{W}_i^L - \mathbf{W}_i^R]. \quad (14.40)$$

Here  $\mathbf{A}(\mathbf{W}_i^n)$  denotes the matrix  $\mathbf{A}$  in (14.39) evaluated at the data cell average  $\mathbf{W}_i^n$ . Substitution of  $\mathbf{W}_i^L$ ,  $\mathbf{W}_i^R$ , as obtained from (14.33), into (14.40) gives

$$\left. \begin{aligned} \bar{\mathbf{W}}_i^L &= \mathbf{W}_i^n - \frac{1}{2} [\mathbf{I} + \frac{\Delta t}{\Delta x} \mathbf{A}(\mathbf{W}_i^n)] \Delta_i , \\ \bar{\mathbf{W}}_i^R &= \mathbf{W}_i^n + \frac{1}{2} [\mathbf{I} - \frac{\Delta t}{\Delta x} \mathbf{A}(\mathbf{W}_i^n)] \Delta_i , \end{aligned} \right\} \quad (14.41)$$

where  $\mathbf{I}$  is the identity matrix. Thus the complete MUSCL–Hancock scheme reduces to solving the Riemann problem with data

$$\left. \begin{aligned} \mathbf{W}_L &= \mathbf{W}_i^n + \frac{1}{2} [\mathbf{I} - \frac{\Delta t}{\Delta x} \mathbf{A}(\mathbf{W}_i^n)] \Delta_i , \\ \mathbf{W}_R &= \mathbf{W}_{i+1}^n - \frac{1}{2} [\mathbf{I} + \frac{\Delta t}{\Delta x} \mathbf{A}(\mathbf{W}_{i+1}^n)] \Delta_{i+1} , \end{aligned} \right\} \quad (14.42)$$

to find  $\mathbf{W}_{i+\frac{1}{2}}(x/t)$  and evaluate the intercell flux as

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}(\mathbf{W}_{i+\frac{1}{2}}(0)) ,$$

in which the vector  $\mathbf{U}_{i+\frac{1}{2}}(0)$  in (14.36) is replaced by  $\mathbf{W}_{i+\frac{1}{2}}(0)$ . See also Sect. 14.6.3.

For primitive variables, in the one-dimensional time dependent Euler equations, we have

$$\mathbf{A} = \begin{pmatrix} u & \rho & 0 \\ 0 & u & \frac{1}{\rho} \\ 0 & \rho a^2 & u \end{pmatrix} , \quad (14.43)$$

where  $a$  is the speed of sound. See Chap. 3 for matrices for the Euler equations in three space dimensions. See also Chap. 8.

### 14.4.3 TVD Version of the Scheme

As seen in Sect. 13.4.2 of Chap. 13, the MUSCL–Hancock scheme is second-order accurate in space and time for any value of the parameter  $\omega$  in the slopes (14.37), when applied to the linear advection equation  $u_t + au_x = 0$ . For  $\omega = 0$  the scheme reduces to the Fromm method. It follows from Godunov’s Theorem, see Sect. 13.5.3 of Chap. 13, that spurious oscillations will be produced in the vicinity of strong gradients. The purpose here is to remedy this problem using some Total Variation Diminishing (TVD) constraint.

We present TVD versions of the scheme. These are extensions of the analysis of Sect. 13.8 of Chap. 13, for the linear advection equation, to non-linear systems of conservation laws (14.31). The extension is somewhat empirical; the rigour of the scalar case can not be preserved for non-linear systems. In spite of this, the TVD versions of the schemes work well and computed solutions are, in the main, oscillation free. The TVD version of the scheme is obtained by replacing the slopes  $\Delta_i$  in (14.33) by *limited* slopes  $\bar{\Delta}_i$ . As discussed in Sect. 13.8, there are essentially two ways of achieving this.

## Limited Slopes

A well established approach [12], [400] is to select *limited slopes* directly from forcing equivalence of the schemes with conventional flux limiter methods, for the model scalar equation. Limited slopes are thus obtained from

$$\bar{\Delta}_i = \begin{cases} \max[0, \min(\beta\Delta_{i-\frac{1}{2}}, \Delta_{i+\frac{1}{2}}), \min(\Delta_{i-\frac{1}{2}}, \beta\Delta_{i+\frac{1}{2}})] , & \Delta_{i+\frac{1}{2}} > 0 , \\ \min[0, \max(\beta\Delta_{i-\frac{1}{2}}, \Delta_{i+\frac{1}{2}}), \max(\Delta_{i-\frac{1}{2}}, \beta\Delta_{i+\frac{1}{2}})] , & \Delta_{i+\frac{1}{2}} < 0 \end{cases} \quad (14.44)$$

for particular values of the parameter  $\beta$ . The value  $\beta = 1$  reproduces the MINBEE (or MINMOD) flux limiter and  $\beta = 2$  reproduces the SUPERBEE flux limiter. Recall that  $\bar{\Delta}_i$  is a vector and thus the limiting process is applied to each component of the system being solved.

As noted in Sect. 13.4 of Chap. 13, upwind MUSCL type schemes based on the limited slopes (14.44) have the Lax–Wendroff as the base second-order scheme, when applied to the linear advection equation, and *not* the Fromm scheme.

A refinement of the above approach is the so called *characteristic limiting* or wave-by-wave limiting; see for instance [400] and [510]. First note that the slope  $\Delta_i$  in cell  $i$  is a function of the jumps  $\Delta_{i-\frac{1}{2}}, \Delta_{i+\frac{1}{2}}$  given in (14.38), namely

$$\bar{\Delta}_i = \bar{\Delta}_i(\Delta_{i-\frac{1}{2}}, \Delta_{i+\frac{1}{2}}) . \quad (14.45)$$

In this approach one decomposes the arguments  $\Delta_{i-\frac{1}{2}}, \Delta_{i+\frac{1}{2}}$  above into jumps across characteristic fields emerging respectively from the Riemann problem solutions with data  $(\mathbf{U}_{i-1}^n, \mathbf{U}_i^n)$  and  $(\mathbf{U}_i^n, \mathbf{U}_{i+1}^n)$ , namely

$$\Delta_{i-\frac{1}{2}} = \sum_{k=1}^N \Delta_{i-\frac{1}{2}}^{(k)} ; \quad \Delta_{i+\frac{1}{2}} = \sum_{k=1}^N \Delta_{i+\frac{1}{2}}^{(k)} \quad (14.46)$$

where  $\Delta_{i+\frac{1}{2}}^{(k)}$  is the jump across wave  $k$  at interface  $i + \frac{1}{2}$  and  $N$  is the number of waves. See Fig. 14.6. One then replaces (14.45) by the summation

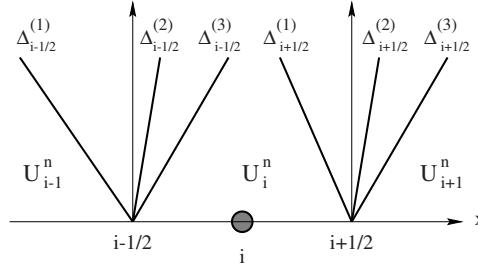
$$\bar{\Delta}_i = \sum_{k=1}^N \bar{\Delta}_i^{(k)}(\Delta_{i-\frac{1}{2}}^{(k)}, \Delta_{i+\frac{1}{2}}^{(k)}) , \quad (14.47)$$

where

$$\bar{\Delta}_i^{(k)} = \bar{\Delta}_i^{(k)}(\Delta_{i-\frac{1}{2}}^{(k)}, \Delta_{i+\frac{1}{2}}^{(k)}) \quad (14.48)$$

is simply a *limited slope* as applied to a scalar problem, obtained from (14.44), for instance.

The application of characteristic limiting requires the solution of the Riemann problem with data  $(\mathbf{U}_i^n, \mathbf{U}_{i+1}^n)$  at each cell interface  $i + \frac{1}{2}$ . Note that this means that the MUSCL–Hancock scheme, or some other MUSCL-type



**Fig. 14.6.** Characteristic limiting. Wave–by–wave decomposition of gradients obtained from local solutions of Riemann problems, with cell averages as initial data

schemes, requires the solution of *two* Riemann problems per cell interface per time step. As for the purpose of characteristic limiting, accuracy and robustness of the Riemann solver are not of paramount importance, one may use the simplest Riemann solver available. Quirk [400] reports satisfactory results when using the Primitive–Variable Linearised Riemann solver presented in Sect. 9.3 of Chap. 9; see [502]. The collective experience of using characteristic limiting is generally a positive one and is well worth the extra cost, particularly when high resolution of contact and shear waves is required.

### Use of Slope Limiters

Another approach is to find a *slope limiter*  $\xi_i$  such that

$$\overline{\Delta}_i = \xi_i \Delta_i , \quad (14.49)$$

with  $\Delta_i$  as given by (14.37), for example. This approach leads to a TVD region for  $\xi(r)$  given as follows:

$$\xi(r) = 0 \text{ for } r \leq 0 , \quad 0 \leq \xi(r) \leq \min\{\xi_L(r), \xi_R(r)\} \text{ for } r > 0 , \quad (14.50)$$

where

$$\left. \begin{aligned} \xi_L(r) &= \frac{2\beta_{i-\frac{1}{2}}r}{1 - \omega + (1 + \omega)r} , \\ \xi_R(r) &= \frac{2\beta_{i+\frac{1}{2}}}{1 - \omega + (1 + \omega)r} , \\ r &= \frac{\Delta_{i-\frac{1}{2}}}{\Delta_{i+\frac{1}{2}}} , \end{aligned} \right\} \quad (14.51)$$

and

$$\beta_{i-\frac{1}{2}} = \frac{2}{1 + c} , \quad \beta_{i+\frac{1}{2}} = \frac{2}{1 - c} . \quad (14.52)$$

The coefficients  $\beta_{i-\frac{1}{2}}$  and  $\beta_{i+\frac{1}{2}}$  above are derived for the scalar case in Sect. 13.8.3 of Chap. 13 and  $c$  is the Courant number for the single wave present.

By considering the limit values of  $\beta_{i-\frac{1}{2}}$  and  $\beta_{i+\frac{1}{2}}$  one may eliminate the dependence on  $c$  and simply set  $\beta_{i-\frac{1}{2}} = \beta_{i+\frac{1}{2}} = 1$  in (14.52). A more refined approach would be to adopt the characteristic limiting approach described earlier in this section. This would require solving extra Riemann problems to find jumps across each wave and their respective Courant numbers, so that one may define coefficients  $\beta_{i+\frac{1}{2}}$  as functions of Courant numbers  $c_{i+\frac{1}{2}}^{(k)}$ .

In Chap. 13 we constructed slope limiters that are analogous to conventional flux limiters, such as SUPERBEE and MINBEE. We stress however that they are only analogous, not equivalent. A slope limiter that is analogous to the SUPERBEE flux limiter is

$$\xi_{sb}(r) = \begin{cases} 0, & \text{if } r \leq 0, \\ 2r, & \text{if } 0 \leq r \leq \frac{1}{2}, \\ 1, & \text{if } \frac{1}{2} \leq r \leq 1, \\ \min\{r, \xi_R(r), 2\}, & \text{if } r \geq 1. \end{cases} \quad (14.53)$$

A van Leer-type slope limiter is

$$\xi_{vl}(r) = \begin{cases} 0, & r \leq 0, \\ \min\left\{\frac{2r}{1+r}, \xi_R(r)\right\}, & r \geq 0. \end{cases} \quad (14.54)$$

A van Albada-type slope limiter is

$$\xi_{va}(r) = \begin{cases} 0, & r \leq 0, \\ \min\left\{\frac{r(1+r)}{1+r^2}, \xi_R(r)\right\}, & r \geq 0. \end{cases} \quad (14.55)$$

A MINBEE-type slope limiter is

$$\xi_{mb}(r) = \begin{cases} 0, & r \leq 0, \\ r, & 0 \leq r \leq 1, \\ \min\{1, \xi_R(r)\}, & r \geq 1. \end{cases} \quad (14.56)$$

As to Riemann solvers to be used the reader is referred to Sect. 14.3.5.

#### 14.4.4 Summary of the MUSCL–Hancock Method

For a given domain  $[0, L]$  and a mesh size  $\Delta x$ , determined by the choice of the number  $M$  of computing cells, one first sets the initial conditions  $\mathbf{U}^0 \equiv \mathbf{U}(x, 0)$  at time  $t = 0$ . Then, for each time step  $n$  one performs the following operations

Operation (I): **Boundary conditions.** This is carried out according to (14.6) or (14.7), for example.

- Operation (II): **Computation of time step.** This is carried out according to (14.4) and (14.5). A choice of  $C_{cfl}$  must be made at the beginning of the computations. One usually takes  $C_{cfl} = 0.9$ . Recall however that the choice of  $S_{max}^{(n)}$  is crucial and given that (14.5) produces somewhat unreliable estimates for the true speeds we recommend that when solving problems with shock–tube like data, the CFL coefficient  $C_{cfl}$  be set to small number, e.g. 0.2, for a few time steps, e.g. 5.
- Operation (III): **Boundary extrapolated values.** This involves Steps (I) and (II) of Sect. 14.4.1, where the slopes  $\Delta_i$  in (14.32), (14.33) are replaced by the limited slopes  $\overline{\Delta}_i$ , which are to be obtained by applying the methods of Sect. 14.4.3. The operation results in TVD, evolved boundary extrapolated values  $\overline{\mathbf{U}}_i^{L,R}$ , for each cell  $i$ .
- Operation (IV): **Solution of Riemann problem.** At each intercell position  $i + \frac{1}{2}$  one finds the solution of Riemann problem with data  $(\overline{\mathbf{U}}_i^R, \overline{\mathbf{U}}_{i+1}^L)$  and computes the intercell flux according to (14.36).
- Operation (V): **Updating of solution.** Proceed to update the conserved variables according to conservative formula (14.3)
- Operation (VI): **Next time level.** Go to (I).

In the next section we present *centred* (non upwind) TVD schemes for non–linear systems of conservation laws.

## 14.5 Centred TVD Schemes

In this section we present two second–order TVD schemes that are extensions of the First–Order Centred (FORCE) scheme [513], [515]. The first scheme is a flux limiter blending of the FORCE scheme and the Richtmyer scheme [402]. The second scheme is of the slope limiter type and results from replacing the Godunov flux by the FORCE flux in the MUSCL–Hancock scheme of Sect. 14.4. Details on the construction and theory of these TVD centred schemes are found in Toro and Billett [526], [538].

A key issue in constructing TVD *centred* schemes is the fact that the dependence of the scheme, via the limiter functions, on the direction of wave propagation is impossible to implement or is simply undesirable. Upwind based TVD schemes have plenty of local information on directionality supplied by the Riemann problem or some other type of characteristic information. This upwind information is usually put to good use in constructing limiter functions with a minimum of artificial viscosity.

### 14.5.1 Review of the FORCE Flux

We solve systems of hyperbolic conservation laws

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0} \quad (14.57)$$

such as the Euler equations, for which  $\mathbf{U}$  and  $\mathbf{F}(\mathbf{U})$  are given by (14.2). A classical scheme of first order accuracy to solve (14.57) is that of Lax–Friedrichs, whose numerical flux at the interface of two states  $\mathbf{U}_L, \mathbf{U}_R$  is

$$\mathbf{F}_{i+\frac{1}{2}}^{LF} = \mathbf{F}_{i+\frac{1}{2}}^{LF}(\mathbf{U}_L, \mathbf{U}_R) = \frac{1}{2}[\mathbf{F}(\mathbf{U}_L) + \mathbf{F}(\mathbf{U}_R)] + \frac{1}{2} \frac{\Delta x}{\Delta t} [\mathbf{U}_L - \mathbf{U}_R]. \quad (14.58)$$

A second-order accurate scheme of interest to us here is the Richtmyer scheme [402], which computes a numerical flux by first defining an intermediate state

$$\mathbf{U}_{i+\frac{1}{2}}^{RI} \equiv \mathbf{U}_{i+\frac{1}{2}}^{RI}(\mathbf{U}_L, \mathbf{U}_R) = \frac{1}{2}(\mathbf{U}_L + \mathbf{U}_R) + \frac{1}{2} \frac{\Delta t}{\Delta x} [\mathbf{F}(\mathbf{U}_L) - \mathbf{F}(\mathbf{U}_R)] \quad (14.59)$$

and then setting

$$\mathbf{F}_{i+\frac{1}{2}}^{RI} = \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}^{RI}). \quad (14.60)$$

The First Order Centred (FORCE) scheme [513] was derived in Chap. 7 for non-linear systems (14.57) and has numerical flux

$$\mathbf{F}_{i+\frac{1}{2}}^{force} = \mathbf{F}_{i+\frac{1}{2}}^{force}(\mathbf{U}_L, \mathbf{U}_R) = \frac{1}{2}[\mathbf{F}_{i+\frac{1}{2}}^{LF}(\mathbf{U}_L, \mathbf{U}_R) + \mathbf{F}_{i+\frac{1}{2}}^{RI}(\mathbf{U}_L, \mathbf{U}_R)]. \quad (14.61)$$

*Remark 14.1.* We note that the FORCE flux results from replacing random values in the staggered-grid version of the Random Choice Method by integral averages; an unexpected outcome is that it turns out to be a simple mean value of the Lax–Friedrichs and Richtmyer fluxes. Motivated by the averaging outcome, one may find the *optimal* averaging of the Lax–Friedrichs and Richtmyer fluxes to reproduce the Godunov first order upwind method for model problems; however, this depends on *upwind* information.

In the next two sections we construct second-order TVD extensions of the FORCE scheme for non-linear systems.

### 14.5.2 A Flux Limiter Centred (FLIC) Scheme

The general flux limiter approach combines a low order monotone flux  $\mathbf{F}_{i+\frac{1}{2}}^{LO}$  and a high order flux  $\mathbf{F}_{i+\frac{1}{2}}^{HI}$  as

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}_{i+\frac{1}{2}}^{LO} + \phi_{i+\frac{1}{2}} [\mathbf{F}_{i+\frac{1}{2}}^{HI} - \mathbf{F}_{i+\frac{1}{2}}^{LO}], \quad (14.62)$$

where  $\phi_{i+\frac{1}{2}}$  is a flux limiter. For details on the flux limiter approach as applied to scalar problems see Sect. 13.7.2 of Chap. 13.

Here we construct a Flux LImiter Centred (FLIC) scheme by taking

$$\mathbf{F}_{i+\frac{1}{2}}^{LO} \equiv \mathbf{F}_{i+\frac{1}{2}}^{force}; \quad \mathbf{F}_{i+\frac{1}{2}}^{HI} \equiv \mathbf{F}_{i+\frac{1}{2}}^{RI} \quad (14.63)$$

where  $\mathbf{F}_{i+\frac{1}{2}}^{force}$  and  $\mathbf{F}_{i+\frac{1}{2}}^{RI}$  are the fluxes for the FORCE and Richtmyer schemes given respectively by (14.61) and (14.60). The scalar version of this scheme was studied in detail in Sect. 13.7.4, where a Total Variation Diminishing (TVD) version of the scheme was also constructed. Here we extend that analysis to non-linear systems in a somewhat empirical manner.

The analysis for the scalar case contains wave propagation information via the Courant number  $c$ . For non-linear systems it will be actually possible to retain, at no cost, some of this information by making use of the CFL coefficient  $C_{cfl}$  in (14.4) required to compute a stable time step  $\Delta t$  in (14.3).

In Sect. 13.7.4 of Chap. 13 we established a relationship between conventional upwind flux limiters  $\psi(r)$  and centred flux limiters  $\phi(r)$ , so that these may be found as follows

$$\phi(r) = \hat{\phi}_g + (1 - \hat{\phi}_g)\psi(r), \quad (14.64)$$

with

$$\hat{\phi}_g = \begin{cases} 0, & r \leq 1, \\ \phi_g \equiv (1 - c_{max})/(1 + c_{max}), & r \geq 1. \end{cases} \quad (14.65)$$

The function  $\phi_g$  above retains some upwinding at no extra cost, as one may set for instance  $c_{max} = C_{cfl}$ . First order centred schemes other than FORCE may also be used, in which case the expression for  $\phi_g$  is different. See Sect. 13.7.4 of Chap. 13 for details. By choosing the upwind flux limiters SUPERBEE, VANLEER, VANALBADA and MINBEE we obtain corresponding *centred flux limiters*

$$\phi_{sb}(r) = \begin{cases} 0, & r \leq 0, \\ \frac{2r}{1+r}, & 0 \leq r \leq \frac{1}{2}, \\ 1, & \frac{1}{2} \leq r \leq 1, \\ \min\{2, \phi_g + (1 - \phi_g)r\}, & r > 1, \end{cases} \quad (14.66)$$

$$\phi_{vl}(r) = \begin{cases} 0, & r \leq 0, \\ \frac{2r}{1+r}, & 0 \leq r \leq 1, \\ \phi_g + \frac{2(1-\phi_g)r}{1+r}, & r \geq 1, \end{cases} \quad (14.67)$$

$$\phi_{va}(r) = \begin{cases} 0, & r \leq 0, \\ \frac{r(1+r)}{1+r^2}, & 0 \leq r \leq 1, \\ \phi_g + \frac{(1-\phi_g)r(1+r)}{1+r^2}, & r \geq 1, \end{cases} \quad (14.68)$$

$$\phi_{mb}(r) = \begin{cases} 0, & r \leq 0, \\ r, & 0 \leq r \leq 1, \\ 1, & r \geq 1. \end{cases} \quad (14.69)$$

For the Euler equations we recommend the following procedure: we first define  $q \equiv E$  (total energy) and set

$$r_{i+\frac{1}{2}}^L = \frac{\Delta q_{i-\frac{1}{2}}}{\Delta q_{i+\frac{1}{2}}} ; \quad r_{i+\frac{1}{2}}^R = \frac{\Delta q_{i+\frac{3}{2}}}{\Delta q_{i+\frac{1}{2}}} . \quad (14.70)$$

Then we compute a single flux limiter

$$\phi^{LR} = \min\{\phi(r_{i+\frac{1}{2}}^L), \phi(r_{i+\frac{1}{2}}^R)\} , \quad (14.71)$$

where  $\phi$  is any of the limiter functions (14.66)–(14.69). Then the single limiter  $\phi^{LR}$  is applied to all three flux components in (14.62). For other systems, some experimentation might be necessary before arriving at some satisfactory procedure. A fairly general approach for a system whose vector of conserved variables has components  $u_k$ ,  $k = 1, \dots, m$ , is to first apply (14.70)–(14.71) to every component  $u_k$  and obtain a limiter  $\phi_k^{LR}$ . Then, one can select the final limiter as

$$\phi^{LR} = \min_k (\phi_k^{LR}) , \quad k = 1, \dots, m .$$

Methods for extending the FLIC scheme to systems with source terms are given in Chap. 15 and techniques for solving multidimensional problems are given in Chap. 16. Numerical results are presented in Sect. 14.7.

In the next section we present another centred TVD scheme.

### 14.5.3 A Slope Limiter Centred (SLIC) Scheme

Here we construct another second order extension of the FORCE scheme [513]. This results from replacing the Godunov upwind flux in the MUSCL–Hancock scheme [562] by the FORCE flux. The scheme as applied to the scalar case was presented in Sect. 13.8 of Chap 13. The TVD version, derived in Sect. 13.8.2 and 13.8.3, results from limiting the slopes in the data reconstruction step.

The **Slope LImiter Centred (SLIC)** scheme has three steps. Steps (I) and (II) are exactly the same as in the MUSCL–Hancock scheme, see Sect. 14.3. The first step results in boundary extrapolated values

$$\mathbf{U}_i^L = \mathbf{U}_i^n - \frac{1}{2} \Delta_i ; \quad \mathbf{U}_i^R = \mathbf{U}_i^n + \frac{1}{2} \Delta_i \quad (14.72)$$

in each cell  $I_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ , where  $\Delta_i$  is a slope vector (a difference). The second step evolves  $\mathbf{U}_i^L, \mathbf{U}_i^R$  by a time  $\frac{1}{2} \Delta t$  according to

$$\left. \begin{aligned} \overline{\mathbf{U}}_i^L &= \mathbf{U}_i^L + \frac{1}{2} \frac{\Delta t}{\Delta x} [\mathbf{F}(\mathbf{U}_i^L) - \mathbf{F}(\mathbf{U}_i^R)] , \\ \overline{\mathbf{U}}_i^R &= \mathbf{U}_i^R + \frac{1}{2} \frac{\Delta t}{\Delta x} [\mathbf{F}(\mathbf{U}_i^L) - \mathbf{F}(\mathbf{U}_i^R)] . \end{aligned} \right\} \quad (14.73)$$

Now, instead of solving the Riemann problem with data

$$\mathbf{W}_L \equiv \overline{\mathbf{U}}_i^R ; \quad \mathbf{W}_R \equiv \overline{\mathbf{U}}_{i+1}^L \quad (14.74)$$

to find the Godunov flux at the intercell position  $i + \frac{1}{2}$ , we evaluate the FORCE flux

$$\mathbf{F}_{i+\frac{1}{2}}^{force} = \mathbf{F}_{i+\frac{1}{2}}^{force}(\overline{\mathbf{U}}_i^R, \overline{\mathbf{U}}_{i+1}^L) , \quad (14.75)$$

as given by (14.61). For the slopes  $\Delta_i$  in (14.72) we take

$$\Delta_i = \frac{1}{2}(1 + \omega)\Delta_{i-\frac{1}{2}} + \frac{1}{2}(1 - \omega)\Delta_{i+\frac{1}{2}} , \quad (14.76)$$

where  $\omega \in [-1, 1]$  is a parameter and  $\Delta_{i-\frac{1}{2}}, \Delta_{i+\frac{1}{2}}$  denote jumps across interfaces  $i - \frac{1}{2}$  and  $i + \frac{1}{2}$  respectively, see (14.38).

As seen in section 13.4.5 of Chap. 13, this scheme is second-order accurate in space and time and stable with Courant number  $c$  satisfying  $|c| \leq 1$ , for any  $\omega \in [-1, 1]$ , when applied to the scalar, linear advection equation. According to Godunov's theorem, see Sect. 13.5.3, spurious oscillations will be produced in the vicinity of strong gradients. To avoid this difficulty, a Total Variation Diminishing (TVD) version of the scheme for the scalar case was constructed in Sect. 13.8. This was achieved by replacing the slopes  $\Delta_i$  by *limited* slopes  $\overline{\Delta}_i$ . The extension to non-linear systems is somewhat empirical. Any of the slope limiters given by (14.53) to (14.56), with  $\beta_{i-\frac{1}{2}} = \beta_{i+\frac{1}{2}} = 1$  in (14.51) will lead to satisfactory results.

*Remark 14.2.* The choice of limited slopes given by (14.44) will not in general give a TVD method when the underlying first order scheme is centred, and therefore (14.44) is not a choice of limiting to be considered here.

In the next section we present upwind TVD schemes based on non-conservative formulations of the equations.

## 14.6 Primitive–Variable Schemes

Here we construct two second-order TVD *upwind* schemes for hyperbolic systems expressed in *non-conservative form*. The schemes are respectively based on the WAF and MUSCL–Hancock approaches presented in Sects. 14.3 and 14.4 for constructing conservative methods. Modifications of the PLM [133] and the GRP [37] approaches have also been used to construct primitive and conservative schemes [510], [262]. Full details on the construction of primitive, conservative and adaptive schemes are given in [517].

### 14.6.1 Formulation of the Equations and Primitive Schemes

We shall call *primitive–variable method* or simply primitive method, any scheme that is based on the *non-conservative* form of the equations, namely

$$\mathbf{W}_t + \mathbf{A}(\mathbf{W})\mathbf{W}_x = \mathbf{0}, \quad (14.77)$$

where  $\mathbf{A}$  is a coefficient matrix that depends on the particular choice of variables  $\mathbf{W}$ . See Sect. 3.2.2 of Chap. 3, for various possible formulations of the governing equations. Note that (14.77) includes the case in which  $\mathbf{W}$  is the vector of conserved variables.

It is well known that *primitive-variable* schemes will compute shock waves with the wrong strength and thus the wrong speed of propagation. At the numerical level there is some recent work by Hou and Le Floch [259] which is also relevant here. However, if the problem is shock free then primitive-variable schemes are perfectly adequate. Conservative methods, on the other hand, have been reported to give erroneous solutions when computing material interfaces by Clarke et. al. [118] and by Karni [278]. As a matter of fact, there are classes of initial-value problems (IVPs) for which conservative methods give erroneous solutions [508], [512], while primitive-variable schemes successfully avoid some of these difficulties. These IVPs include material interfaces, shear waves and *symmetric data* Riemann problems. Other relevant examples in which primitive methods would be adequate are (i) the artificial compressibility equations associated with the incompressible Navier–Stokes equations, [514] (ii) linear constant coefficient hyperbolic systems, see Chap. 5, (iii) systems admitting very weak shock waves such as in Acoustics and (iv) hyperbolic systems with unknown conservative formulations [549]; we note however that if such systems contain shocks, then their primitive form will not give the correct solutions. There is a large body of work concerned with primitive-variable schemes and their combination with shock-fitting techniques; see for instance the works of Moretti and co-workers [355], [353], [354], [157]. Upwind TVD primitive-variable schemes have been constructed by Karni [278], Toro [508], [511], [510] and more recently by Abgrall [1]. Obviously, if shock waves are present then primitive schemes must be ruled out. However, by making use of adaptive schemes, where shock waves are treated by a conservative method, one can preserve the advantages of both approaches. Adaptive primitive-conservative schemes have been put forward by Toro [508], [511], Karni [279] and Abgrall [1]; see also Ivings et. al. [262].

We describe the schemes for the *augmented*, one dimensional time dependent Euler equations, see Chaps. 1 and 3. We express the equations in terms of the primitive variables  $\mathbf{W} = (\rho, u, p, q)^T$ , where  $q$  is any *passive scalar*. The governing equations are of the form (14.77), where the coefficient matrix  $\mathbf{A}(\mathbf{W})$  is

$$\mathbf{A}(\mathbf{W}) = \begin{pmatrix} u & \rho & 0 & 0 \\ 0 & u & \frac{1}{\rho} & 0 \\ 0 & \rho a^2 & u & 0 \\ 0 & 0 & 0 & u \end{pmatrix}. \quad (14.78)$$

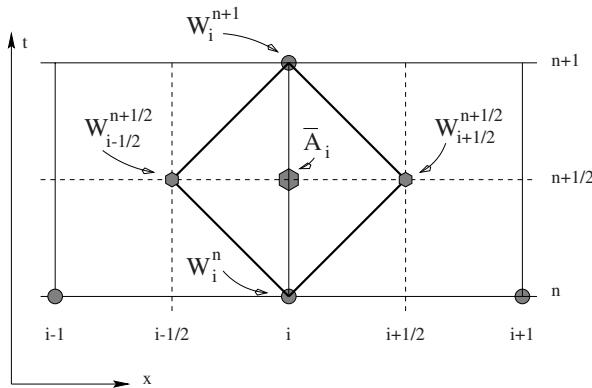
The variable  $q = q(x, t)$  may represent a reaction progress variable when modelling chemically active flows, a variable associated with a material interface,

the concentration of some chemical species, or simply the tangential velocity component in two-dimensional gas dynamics.

Now we propose a two-step finite-difference scheme to solve (14.77), namely

$$\mathbf{W}_i^{n+1} = \mathbf{W}_i^n + \frac{\Delta t}{\Delta x} \bar{\mathbf{A}}_i \left[ \mathbf{W}_{i-\frac{1}{2}}^{n+\frac{1}{2}} - \mathbf{W}_{i+\frac{1}{2}}^{n+\frac{1}{2}} \right]. \quad (14.79)$$

Fig. 14.7 shows the stencil of the scheme. At the half-time level  $n + \frac{1}{2}$ , the matrix  $\bar{\mathbf{A}}_i$  is defined at the grid point  $i$  and the *intermediate* states  $\mathbf{W}_{i+\frac{1}{2}}^{n+\frac{1}{2}}$  are defined at the intermediate positions  $i + \frac{1}{2}$ . The scheme is completely determined once  $\bar{\mathbf{A}}_i$  and  $\mathbf{W}_{i+\frac{1}{2}}^{n+\frac{1}{2}}$  are defined. We note that the problem of finding  $\mathbf{W}_{i+\frac{1}{2}}^{n+\frac{1}{2}}$  is similar to that of finding an intercell flux in conservative methods, see equation (14.3). We shall study two approaches.



**Fig. 14.7.** Stencil for primitive–variable schemes on a staggered grid. Point values  $\mathbf{W}_i^n$  are updated to  $\mathbf{W}_i^{n+1}$ , intermediate values are  $\mathbf{W}_{i-\frac{1}{2}}^{n+\frac{1}{2}}$  and  $\mathbf{W}_{i+\frac{1}{2}}^{n+\frac{1}{2}}$

### 14.6.2 A WAF–Type Primitive Variable Scheme

Here we follow the WAF approach of Sect. 14.3 for constructing the intermediate vectors  $\mathbf{W}_{i+\frac{1}{2}}^{n+\frac{1}{2}}$  and the coefficient matrix  $\bar{\mathbf{A}}_i$ . First we define

$$\mathbf{W}_{i+\frac{1}{2}}^{n+\frac{1}{2}} = \frac{1}{\Delta x} \int_{-\frac{1}{2}\Delta x}^{\frac{1}{2}\Delta x} \mathbf{W}_{i+\frac{1}{2}}(x, \frac{1}{2}\Delta t) dx. \quad (14.80)$$

See equation (14.14). So far this is identical to the Weighted Average State version of WAF, see Sect. 14.3.2. Here  $\mathbf{W}_{i+\frac{1}{2}}(x, t)$  is the solution of the conventional Riemann problem with data  $(\mathbf{W}_i^n, \mathbf{W}_{i+1}^n)$ . An approximation to the integral is

$$\mathbf{W}_{i+\frac{1}{2}}^{n+\frac{1}{2}} = \sum_{k=1}^N \beta_k \mathbf{W}_{i+\frac{1}{2}}^{(k)}, \quad (14.81)$$

where  $N$  is the total number of waves in the solution of the Riemann problem ( $N = 4$  in the present case). The weights or coefficients are given by

$$\left. \begin{aligned} \beta_k &= \frac{1}{2}(c_k - c_{k-1}), \\ c_k &= \frac{\Delta t S_k}{\Delta x}, \quad c_0 = -1, \quad c_{N+1} = 1, \end{aligned} \right\} \quad (14.82)$$

where  $c_k$  is the Courant number for wave  $k$  of speed  $S_k$ . Manipulation of (14.81)–(14.82) leads to

$$\mathbf{W}_{i+\frac{1}{2}}^{n+\frac{1}{2}} = \frac{1}{2}(\mathbf{W}_i^n + \mathbf{W}_{i+1}^n) - \frac{1}{2} \sum_{k=1}^N c_k [\mathbf{W}_{i+\frac{1}{2}}^{(k+1)} - \mathbf{W}_{i+\frac{1}{2}}^{(k)}]. \quad (14.83)$$

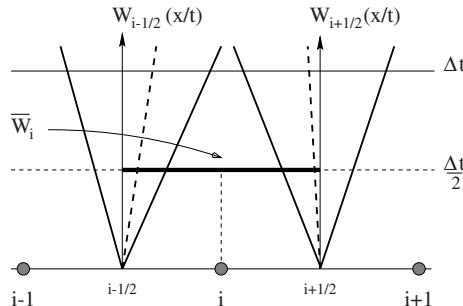
To compute the coefficient matrix  $\bar{\mathbf{A}}_i$  we first define a state  $\bar{\mathbf{W}}_i$  and then set

$$\bar{\mathbf{A}}_i = \mathbf{A}(\bar{\mathbf{W}}_i). \quad (14.84)$$

The simplest choice is obviously  $\bar{\mathbf{W}}_i = \mathbf{W}_i^n$ , which we do not believe to be sufficiently robust to be used with confidence. A sophisticated choice for  $\bar{\mathbf{A}}_i$  is

$$\bar{\mathbf{W}}_i = \frac{1}{\Delta x} \int_0^{\frac{1}{2}\Delta x} \mathbf{W}_{i-\frac{1}{2}}(x, \frac{1}{2}\Delta t) dx + \frac{1}{\Delta x} \int_{-\frac{1}{2}\Delta x}^0 \mathbf{W}_{i+\frac{1}{2}}(x, \frac{1}{2}\Delta t) dx, \quad (14.85)$$

which is an integral average of the solutions  $\mathbf{W}_{i-\frac{1}{2}}(x/t)$  and  $\mathbf{W}_{i+\frac{1}{2}}(x/t)$  of the two Riemann problems *affecting* mesh point  $i$ , at the half-time level. See Fig. 14.8. A suitable approximation to this integral is



**Fig. 14.8.** State  $\bar{\mathbf{W}}_i$  to compute the coefficient matrix is an integral average of solutions of Riemann problems affecting mesh point  $i$

$$\bar{\mathbf{W}}_i = \sum_{j=j_L}^{N+1} \beta_{i-\frac{1}{2}}^{(j)} \mathbf{W}_{i-\frac{1}{2}}^{(j)} + \sum_{j=1}^{j_R} \beta_{i+\frac{1}{2}}^{(j)} \mathbf{W}_{i+\frac{1}{2}}^{(j)}, \quad (14.86)$$

with the coefficients defined as

$$\left. \begin{aligned} \beta_{i-\frac{1}{2}}^{(j)} &= \frac{1}{2}(c_j - c_{j-1}) \text{ for } j > j_L ; \beta_{i-\frac{1}{2}}^{j_L} = \frac{1}{2}c_{j_L}, \\ \beta_{i+\frac{1}{2}}^{(j)} &= \frac{1}{2}(c_j - c_{j-1}) \text{ for } j < j_R ; \beta_{i+\frac{1}{2}}^{j_R} = -\frac{1}{2}c_{j_R}. \end{aligned} \right\} \quad (14.87)$$

The superscript  $j_L$  refers to the value of  $k$  in the solution of the Riemann problem with data  $(\mathbf{W}_{i-1}^n, \mathbf{W}_i^n)$  such that the speed  $S_{k-1}$  is negative and  $S_k$  is positive. A similar interpretation holds for  $j_R$ . Another approximation to (14.85) that is reported to work well [262] is

$$\bar{\mathbf{W}}_i = \frac{1}{2}(\mathbf{W}_{i-\frac{1}{2}}^{n+\frac{1}{2}} + \mathbf{W}_{i+\frac{1}{2}}^{n+\frac{1}{2}}), \quad (14.88)$$

which makes use of information made available by (14.83).

In approximating the integrals (14.80) and (14.85) by the summations (14.81) and (14.86) we have assumed the simplified wave structure of Figs. 14.2 and 14.8. The treatment of rarefactions is entirely analogous to the conservative version of the scheme and is discussed in Sect. 14.3.3.

**Exercise 14.3.** Apply the primitive scheme (14.79) to the model equation

$$u_t + au_x = 0, \quad a = \text{constant}. \quad (14.89)$$

Verify that for  $a > 0$  and  $a < 0$  the resulting scheme is the Lax–Wendroff method, see Chap. 13, which is second–order accurate in space and time.

Show also that if the intermediate state (14.80) is computed by a mid–point rule, then the scheme reduces to the Godunov first–order upwind scheme, see Chap. 6, which is monotone.

*Remark 14.4.* If the above observation is generalised and we compute the intermediate state vectors  $\bar{\mathbf{W}}_{i+\frac{1}{2}}^{n+\frac{1}{2}}$  in (14.80) by a mid–point rule approximation, then one obtains a first–order accurate primitive scheme, that is the non–conservative counterpart of Godunov’s conservative method studied in Chap. 6.

Based on the first–order primitive scheme above one can construct a Total Variation Diminishing (TVD) version of the second–order primitive method. This is most easily accomplished by replacing (14.83) by

$$\mathbf{W}_{i+\frac{1}{2}}^{n+\frac{1}{2}} = \frac{1}{2}(\mathbf{W}_i^n + \mathbf{W}_{i+1}^n) - \frac{1}{2} \sum_{k=1}^N \text{sign}(c_k) \phi_k [\mathbf{W}_{i+\frac{1}{2}}^{(k+1)} - \mathbf{W}_{i+\frac{1}{2}}^{(k)}], \quad (14.90)$$

where  $\phi_k$  is any of the WAF limiter functions (14.27)–(14.30). See Chap. 13 for general background on TVD schemes and limiter functions.

We summarise the application of the primitive WAF scheme as applied to (14.77): One first solves the Riemann problems with data  $(\mathbf{W}_i^n, \mathbf{W}_{i+1}^n)$  to find  $\mathbf{W}_{i+\frac{1}{2}}(x/t)$ ; one then computes the intermediate state vectors  $\mathbf{W}_{i+\frac{1}{2}}^{n+\frac{1}{2}}$  according to (14.90) and the coefficient matrix  $\bar{\mathbf{A}}_i$  according to (14.84), with  $\bar{\mathbf{W}}$  obtained from (14.85) or (14.88); the primitive scheme (14.79) is thus completely determined. The linearised stability condition of the scheme is  $|c| \leq 1$ .

*Remark 14.5.* The primitive WAF scheme is most easily converted into the conservative WAF scheme of Sect. 14.3.2 by simply computing an intercell flux as  $\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}(\mathbf{W}_{i+\frac{1}{2}}^{n+\frac{1}{2}})$  and advancing the solution via the conservative formula (14.3). See Sect. 14.3.2.

Concerning Riemann solvers the reader is referred to Sect. 14.3.5.

### 14.6.3 A MUSCL–Hancock Primitive Scheme

A different primitive scheme results if the computation of  $\mathbf{W}_{i+\frac{1}{2}}^{n+\frac{1}{2}}$  and the coefficient matrix  $\bar{\mathbf{A}}_i$  in (14.79) is carried out via the MUSCL–Hancock approach, first put forward for constructing conservative methods [562]. See Sect. 14.4 for details on this approach.

First we define a procedure for computing the intermediate values  $\mathbf{W}_{i+\frac{1}{2}}^{n+\frac{1}{2}}$ . This is done in essentially the same way as for conservative schemes. After the reconstruction step (14.32) in terms of  $\mathbf{W}$ , there are three distinct steps, namely

Step (I): **Extrapolation.** Using the reconstructed solution in (14.32) in terms of the vector of primitive variables  $\mathbf{W}$  we obtain the boundary extrapolated values

$$\mathbf{W}_i^L = \mathbf{W}_i^n - \frac{1}{2}\Delta_i; \quad \mathbf{W}_i^R = \mathbf{W}_i^n + \frac{1}{2}\Delta_i. \quad (14.91)$$

Step (II): **Evolution.** We evolve the values  $\mathbf{W}_i^{L,R}$  by a time  $\frac{1}{2}\Delta t$  according to

$$\bar{\mathbf{W}}_i^{L,R} = \mathbf{W}_i^{L,R} + \frac{1}{2} \frac{\Delta t}{\Delta x} \mathbf{A}(\mathbf{W}_i^n)(\mathbf{W}_i^L - \mathbf{W}_i^R). \quad (14.92)$$

Step (III): **The Riemann problem.** In order to compute  $\mathbf{W}_{i+\frac{1}{2}}^{n+\frac{1}{2}}$  we find the solution  $\mathbf{W}_{i+\frac{1}{2}}(x/t)$  of the Riemann problem with piece-wise constant data  $(\bar{\mathbf{W}}_i^R, \bar{\mathbf{W}}_{i+1}^L)$  and set

$$\mathbf{W}_{i+\frac{1}{2}}^{n+\frac{1}{2}} = \mathbf{W}_{i+\frac{1}{2}}(0), \quad (14.93)$$

which requires a solution sampling procedure entirely analogous to that of the Godunov method described in Chap. 6. See also Chap. 9.

There are several possible choices for the matrix  $\overline{\mathbf{A}}_i$  in (14.79). By using (14.84) one only requires the computation of a state  $\overline{\mathbf{W}}_i$ . A possible choice is

$$\overline{\mathbf{W}}_i = \frac{1}{2}(\overline{\mathbf{W}}_i^L + \overline{\mathbf{W}}_i^R). \quad (14.94)$$

Some algebra gives the simple expression

$$\overline{\mathbf{W}}_i = \mathbf{W}_i^n - \frac{1}{2} \frac{\Delta t}{\Delta x} \mathbf{A}(\mathbf{W}_i^n) \Delta_i. \quad (14.95)$$

Another choice is given by (14.85), where now  $\mathbf{W}_{i+\frac{1}{2}}(x/t)$  is the solution of the Riemann problem in Step (III) above. The integral may now be approximated as in (14.86), (14.87). A trapezium–rule approximation gives another choice,

$$\overline{\mathbf{W}}_i = \frac{1}{2}(\mathbf{W}_{i-\frac{1}{2}}^{n+\frac{1}{2}} + \mathbf{W}_{i+\frac{1}{2}}^{n+\frac{1}{2}}). \quad (14.96)$$

**Exercise 14.6.** Apply the primitive MUSCL–Hancock scheme described to the model equation (14.89). Verify that for  $a > 0$  and  $a < 0$  the resulting scheme is the Fromm scheme, see Chap. 13, which is second–order accurate in space and time.

Verify also that if the slopes  $\Delta_i$  in (14.91) are identically zero, then the scheme reduces to the Godunov first–order upwind scheme (the CIR scheme), see Chap. 13, which is monotone.

**Solution 14.7.** Left to the reader.

*Remark 14.8.* Based on the remark above one can construct a first–order primitive scheme for non–linear systems by setting the slopes to zero. The resulting scheme is the primitive counterpart of Godunov’s conservative method studied in Chap. 6 and which is identical to the primitive first order scheme obtained via the WAF approach in Sect. 14.6.2.

A Total Variation Diminishing (TVD) version of the second–order primitive MUSCL–Hancock method is easily constructed. To this end one replaces the slopes  $\Delta_i$  by *limited slopes*  $\overline{\Delta}_i$  in the same way as in the conservative MUSCL–Hancock scheme described in Sect. 14.4.3.

*Remark 14.9.* The primitive MUSCL–Hancock scheme is most easily converted into the conservative MUSCL–Hancock scheme by simply computing an intercell flux as  $\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}(\mathbf{W}_{i+\frac{1}{2}}^{n+\frac{1}{2}})$ , where  $\mathbf{W}_{i+\frac{1}{2}}^{n+\frac{1}{2}}$  is given by (14.93), and advancing the solution via the conservative formula (14.3). See Sects. 14.4.1 and 14.4.2.

Concerning Riemann solvers the reader is referred to Sect. 14.3.5.

#### 14.6.4 Adaptive Primitive–Conservative Schemes

Here we present an adaptive procedure whereby a conservative scheme is used at shocks only and a primitive scheme is used elsewhere. Details of the approach are found in [508], [511], [517]; see also Ivings et. al. [262]. The experience reported in [539], [540] in constructing adaptive schemes is useful here.

Consider a mesh point  $i$  as depicted in Fig. 14.8. This mesh point is updated by a conservative method *only if* any of the two neighbouring Riemann problem solutions  $\mathbf{W}_{i-\frac{1}{2}}(x/t)$ ,  $\mathbf{W}_{i+\frac{1}{2}}(x/t)$  contains a shock travelling in the direction of the mesh point  $i$ . The local solutions  $\mathbf{W}_{i-\frac{1}{2}}(x/t)$ ,  $\mathbf{W}_{i+\frac{1}{2}}(x/t)$  contain all the required information and thus the adaptive scheme is easily implemented. Denote by  $p_{i-\frac{1}{2}}^*$  and  $p_{i+\frac{1}{2}}^*$  the solutions for pressure in the star regions of  $\mathbf{W}_{i-\frac{1}{2}}(x/t)$  and  $\mathbf{W}_{i+\frac{1}{2}}(x/t)$  respectively. See Chaps. 4 and 9 for details on the structure of the solution of the Riemann problem. Define

$$S_{str} = 1 + PTOL, \quad (14.97)$$

where  $PTOL$  is a small positive quantity yet to be defined. If

$$\left. \begin{array}{l} p_{i-\frac{1}{2}}^*/p_i^n > S_{str} \text{ and } S_{i-\frac{1}{2}} > 0, \\ \text{or if} \\ p_{i+\frac{1}{2}}^*/p_i^n > S_{str} \text{ and } S_{i+\frac{1}{2}} < 0 \end{array} \right\} \quad (14.98)$$

then mesh point  $i$  is advanced via a conservative method (14.3); otherwise one advances the solution via a primitive scheme (14.79) of Sect. 14.6.1.

In (14.98)  $S_{i-\frac{1}{2}}$  denotes the speed of a shock wave contained in the solution  $\mathbf{W}_{i-\frac{1}{2}}(x/t)$  of the Riemann problem and  $S_{i+\frac{1}{2}}$  denotes the speed of a shock wave contained in the solution  $\mathbf{W}_{i+\frac{1}{2}}(x/t)$ . Experience reported in [508] and [262] suggests that the choice of  $PTOL$  in (14.97) is not too critical and any value in the range  $(0, 0.1)$  gives satisfactory results. If  $PTOL$  is too large one risks computing shocks of *moderate* strength with the primitive scheme, which will lead to shock waves being in the wrong position. Small values of  $PTOL$  simply mean that the conservative method is used almost everywhere.

Note that these adaptive schemes are *essentially primitive*, whereby conservative methods are only used at isolated shocked points. Given that any of the primitive schemes presented can most easily be converted into their conservative counterparts, the adaptive primitive–conservative approach is easily implemented. Alternative hybrid schemes have been presented by Karni [279] and independently by Abgrall [1].

In the next section we present some numerical results for a test problem with exact solution.

## 14.7 Some Numerical Results

Here we present numerical results for some of the schemes studied in this chapter, as applied to the Euler equations. As a test problem with exact solution we use Test 1 of Sect. 6.4 of Chap. 6, which has initial data consisting of two constant states  $\mathbf{W}_L = [\rho_L, u_L, p_L]^T$  and  $\mathbf{W}_R = [\rho_R, u_R, p_R]^T$ , separated by a discontinuity at a position  $x = 0.3$ . The data states are  $\mathbf{W}_L = [1.0, 0.0, 1.0]^T$  and  $\mathbf{W}_R = [0.125, 0.0, 0.1]^T$ . The ratio of specific heats is chosen to be  $\gamma = 1.4$ . The exact and numerical solutions are found in the spatial domain  $0 \leq x \leq 1$ . The numerical solution is computed with  $M = 100$  cells and the Courant number coefficient is  $C_{\text{cfl}} = 0.9$ ; boundary conditions are transmissive and  $S_{\max}^{(n)}$  is found using the simplified formula (14.5). The test problem is a modified version of the popular Sod's test [453]; the solution consists of a right shock wave, a right travelling contact wave and a left *sonic* rarefaction wave, a feature that is very useful for assessing the *entropy satisfaction* property of numerical methods. The exact solution was found by running the code *HE-E1RPEXACT* of the library *NUMERICA* [519]. The numerical solutions were obtained by running the codes *HE-E1WAF*, *HE-E1MUSHAN* and *HE-E1SLIC* of *NUMERICA*.

We present four sets of results. The first two sets correspond to *upwind* methods and include the WAF method of Sect. 14.3, see Figs. 14.9 and 14.10, and the MUSCL–Hancock method of Sect. 14.4, see Figs. 14.11 to 14.14. All upwind results have been obtained with the HLLC Riemann solver of Sect. 10.4 of Chap. 10. The third and fourth sets of results correspond to *centred* methods and include the Flux Limiter Centred (FLIC) scheme of Sect. 14.5.2, see Figs. 14.15 and 14.16, and the Slope Limiter Centred (SLIC) scheme of Sect. 14.5.3, see Figs. 14.17 and 14.18. For each method we show results for at least two limiter functions.

### 14.7.1 Upwind TVD Methods

Figs. 14.9 and 14.10 show results for the WAF method using the limiter functions that are *equivalent* to the flux limiters SUPERBEE (14.27) and MINBEE (14.30); the first result gives very sharp resolution of discontinuities but small spurious oscillations are just visible on the plots; the second result using MINBEE has no signs of spurious oscillations but discontinuities are somewhat smeared, especially the contact discontinuity. In practice one tends to use some limiter in between MINBEE and SUPERBEE, such as VANLEER or VANALBADA, given respectively by (14.28) and (14.29).

Figs. 14.11 to 14.14 show results from the MUSCL–Hancock method. The results of Figs. 14.11 and 14.12 were obtained from the *limited slopes* (14.44), the first one corresponds to the flux limiter SUPERBEE and the second one corresponds to the flux limiter MINBEE. Compare with Figs. 14.9 and 14.10 respectively. The result from SUPERBEE gives sharp discontinuities but has visible spurious oscillations, specially in the internal energy plot. The details

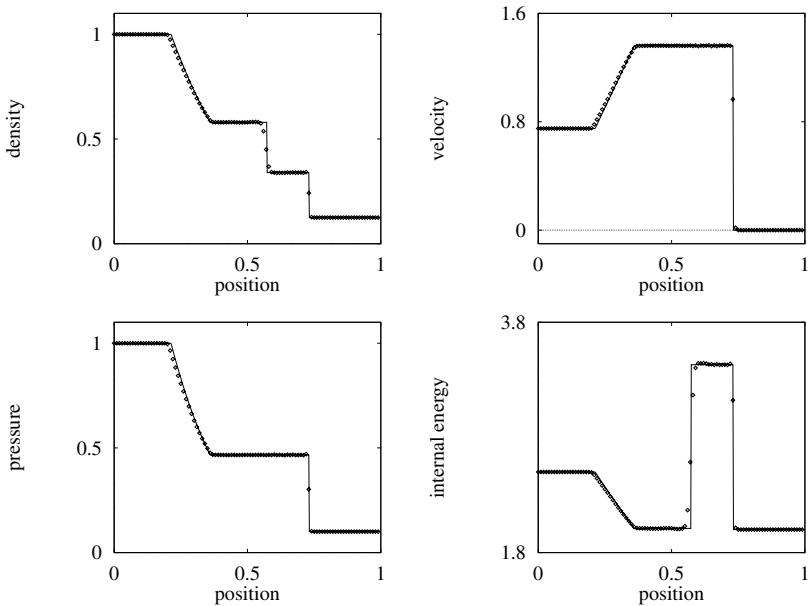
of the particular Riemann solver being used may also have an effect. The result from MINBEE has virtually eliminated the spurious oscillations but discontinuities are somewhat smeared. The results of Fig. 14.13 are obtained from the slope limiter (14.53), which is *analogous* to SUPERBEE, and the results of Fig. 14.14 are obtained from the slope limiter (14.56), which is analogous to MINBEE. In both slope limiters we have set  $c = C_{cfl}$  in (14.51), (14.52). Overall, the WAF method appears to give better results than the MUSCL–Hancock scheme, although the use of *characteristic limiting*, see Sect. 14.4.3, would improve the quality of the MUSCL scheme, but at the cost of solving an extra, though cheaper, Riemann problem per cell interface per time step. All results should be compared with those obtained from the Godunov first order upwind method of Chap. 6; see Fig. 6.8.

### 14.7.2 Centred TVD Methods

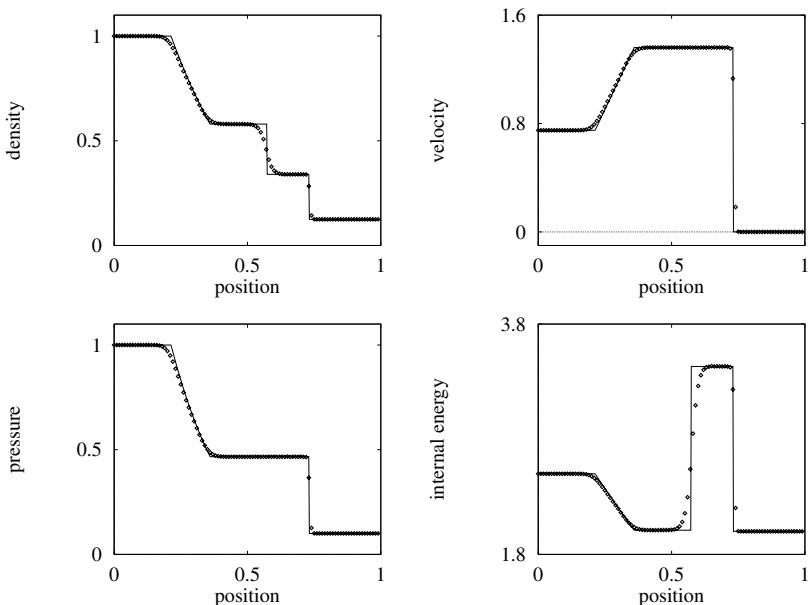
Figs. 14.15 and 14.16 show results from the Flux Limiter *Centred* method, FLIC, for two *centred* flux limiter functions, namely (14.66) and (14.69); these are *analogous* to SUPERBEE and MINBEE respectively. Overall these results are less accurate than those from the *upwind* methods. Note however that the results of Fig. 14.15 are comparable with those of Figs. 14.10, 14.12 and 14.14, obtained from the most diffusive of upwind limiters used. Also, the most compressive limiter for upwind methods tends to overshoot, which does not happen with the most compressive limiter for the centred method.

Figs. 14.17 and 14.18 show results from the Slope Limiter *Centred* method, SLIC, for two slope limiter functions, namely (14.53) and (14.56). In fact these limiters are the same as those used for the upwind MUSCL–Hancock scheme to produce the results of Figs. 14.13 and 14.14. Overall these results are very similar to those obtained from the other centred method, FLIC, and thus the same remarks apply.

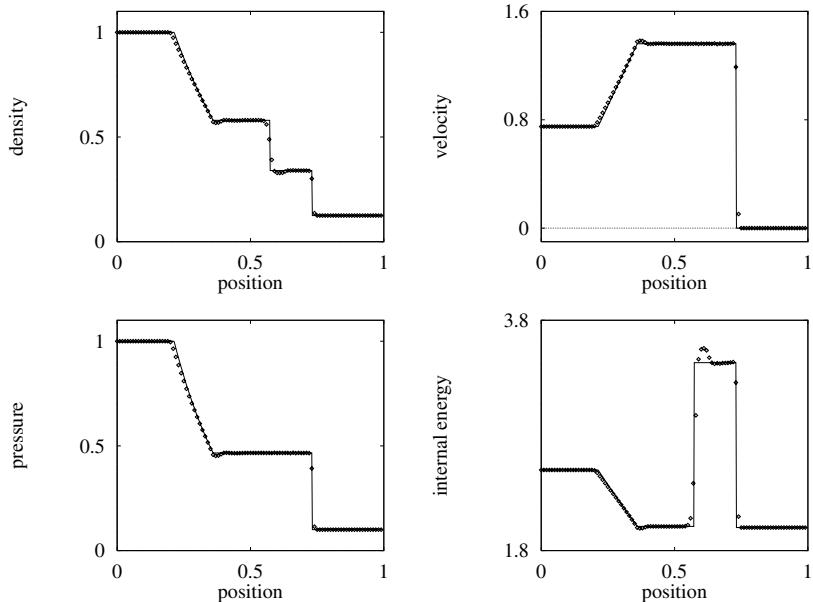
In order to appreciate the gains in going from first-order schemes to second-order TVD schemes, the reader should compare all the results of this chapter with those obtained from (i) the Godunov method of Chap. 6 (Fig. 6.8), (ii) the Lax–Friedrichs scheme of Chap. 5 (Fig. 6.13) and (iii) the FORCE scheme of Chap. 7 (Fig. 7.18). They should also be compared with those obtained from a non-TVD second order method, such as the Richtmyer scheme of Chap. 5 (Fig. 6.18).



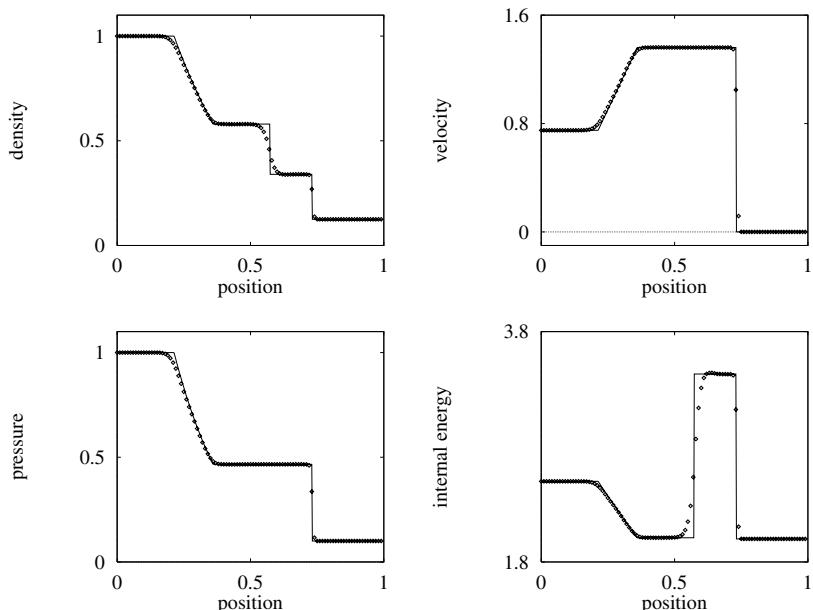
**Fig. 14.9.** WAF Scheme with HLLC Riemann solver and SUPERBEE applied to Test 1. Numerical (symbol) and exact (line) solutions at time 0.2 units



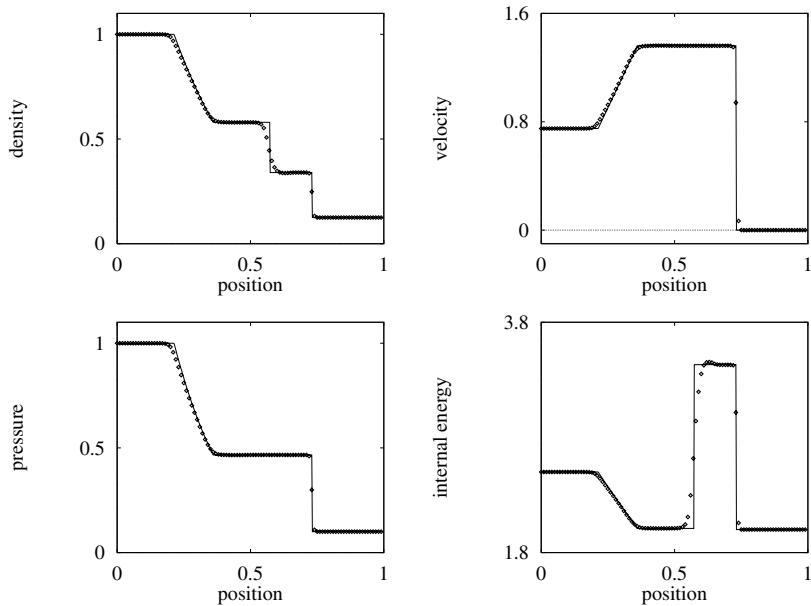
**Fig. 14.10.** WAF Scheme with HLLC Riemann solver and MINBEE applied to Test 1. Numerical (symbol) and exact (line) solutions at time 0.2 units



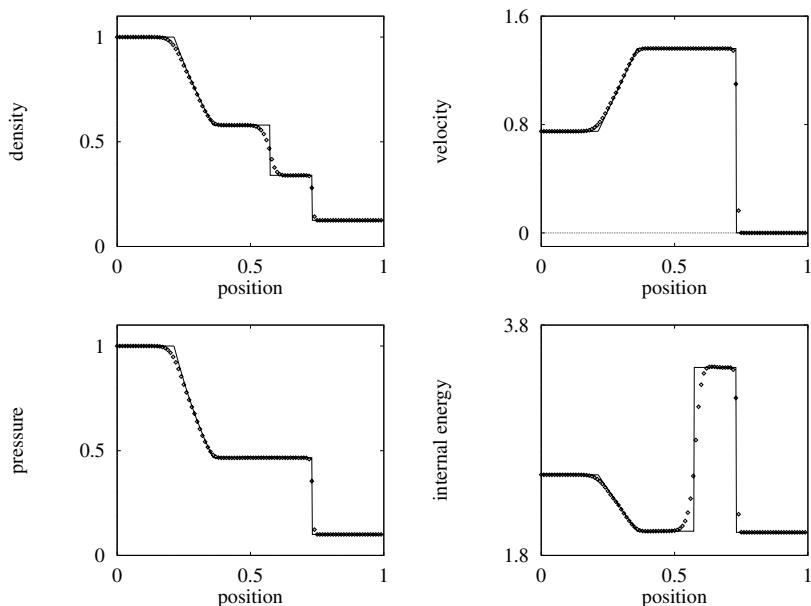
**Fig. 14.11.** MUSCL–Hancock scheme, HLLC Riemann solver and SUPERBEE for to Test 1. Numerical (symbol) and exact (line) solutions at time 0.2 units



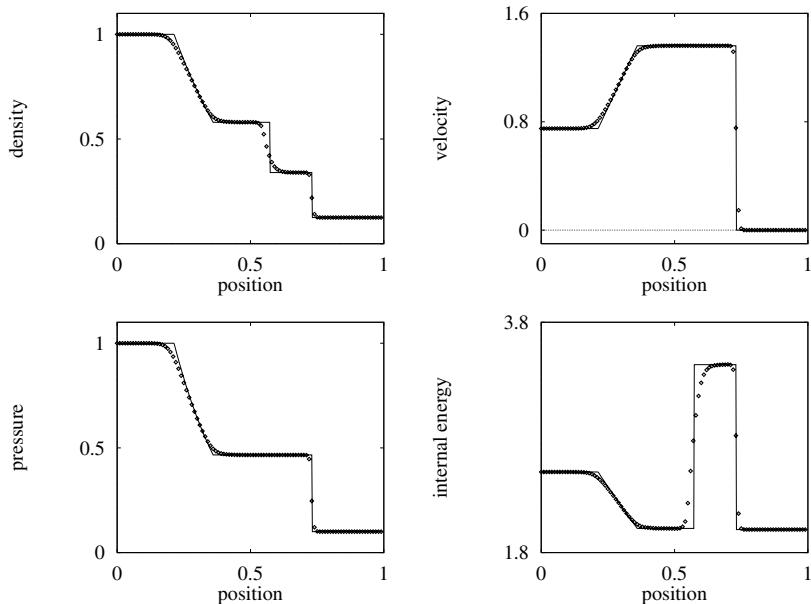
**Fig. 14.12.** MUSCL–Hancock Scheme with HLLC Riemann solver and MINBEE applied to Test 1. Numerical (symbol) and exact (line) solutions at time 0.2 units



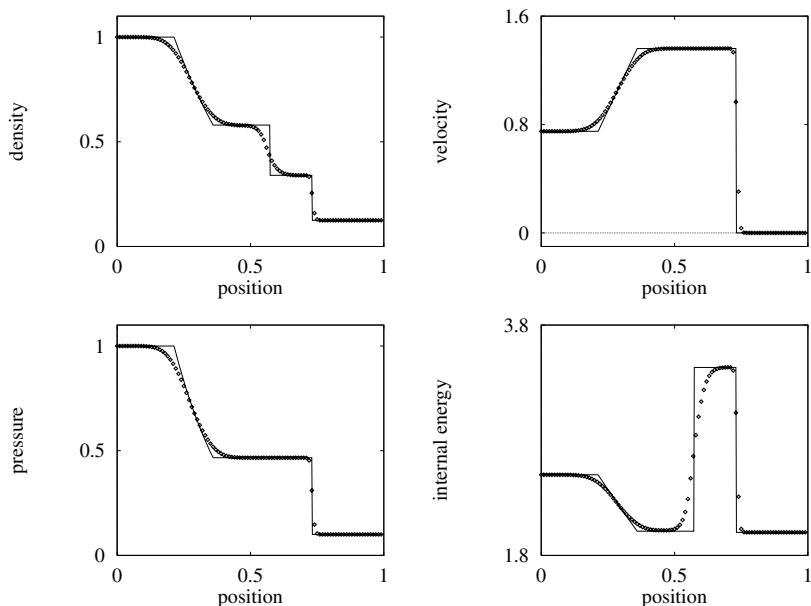
**Fig. 14.13.** MUSCL–Hancock Scheme, HLLC Riemann solver and SUPERBEEs for Test 1. Numerical (symbol) and exact (line) solutions at time 0.2 units



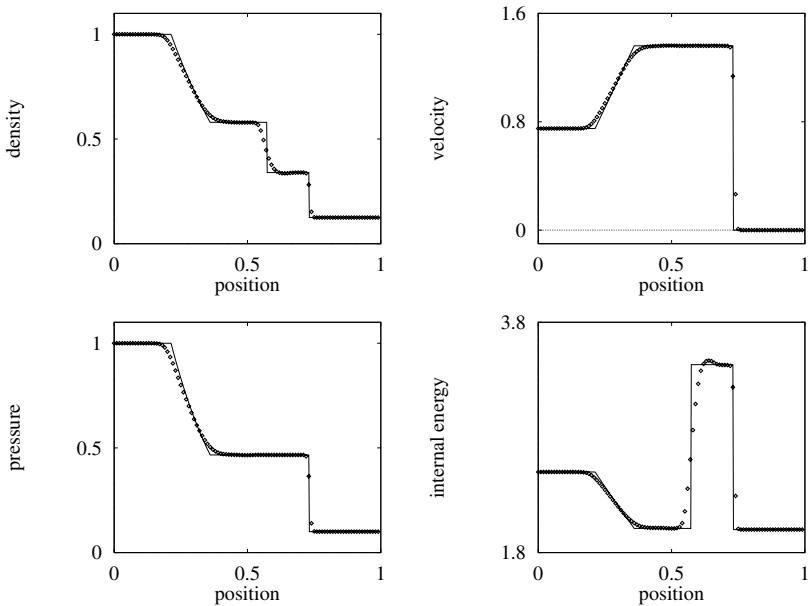
**Fig. 14.14.** MUSCL–Hancock Scheme with HLLC Riemann solver and MINBEEs applied to Test 1. Numerical (symbol) and exact (line) solutions at time 0.2 units



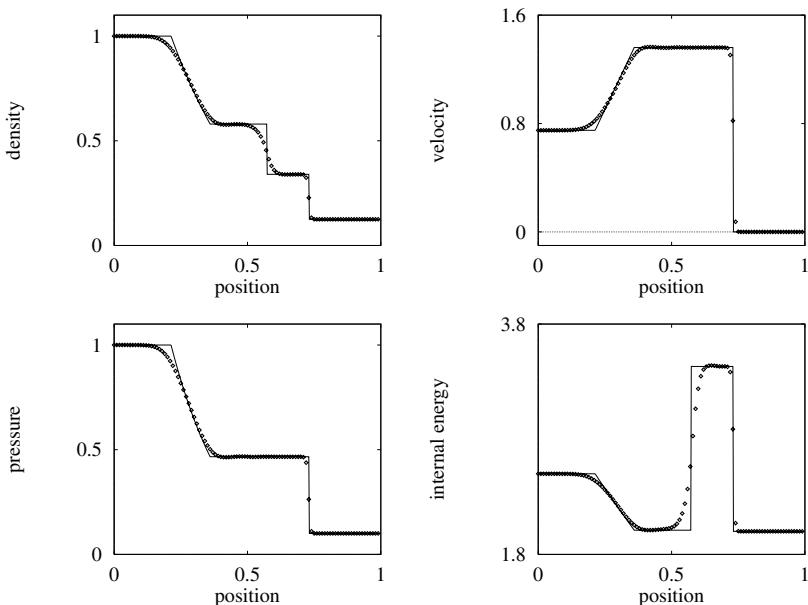
**Fig. 14.15.** Flux Limiter Centred (FLIC) scheme with Superbee applied to Test 1. Numerical (symbol) and exact (line) solutions at time 0.2 units



**Fig. 14.16.** Flux Limiter Centred (FLIC) scheme with Minbee applied to Test 1. Numerical (symbol) and exact (line) solutions at time 0.2 units



**Fig. 14.17.** Slope Limiter Centred (SLIC) scheme with Superbee applied to Test 1. Numerical (symbol) and exact (line) solutions at time 0.2 units



**Fig. 14.18.** Slope Limiter Centred (SLIC) scheme with Minbee applied to Test 1. Numerical (symbol) and exact (line) solutions at time 0.2 units

---

## Splitting Schemes for PDEs with Source Terms

### 15.1 Introduction

This chapter is concerned with numerical methods for solving non-linear systems of hyperbolic conservation laws with source terms

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{S}(\mathbf{U}) . \quad (15.1)$$

$\mathbf{U}$  is the vector of unknowns,  $\mathbf{F}(\mathbf{U})$  is the vector of fluxes and  $\mathbf{S}(\mathbf{U})$  is a vector of *sources*, which in general is an algebraic function of  $\mathbf{U}$  or other physical parameters of the problem at hand. In Chap. 2 we studied some properties for the *pure advection* hyperbolic problem

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0} . \quad (15.2)$$

This *homogeneous system*, in which  $\mathbf{S}(\mathbf{U}) \equiv \mathbf{0}$  (no sources) is a simplified version of (15.1). The time-dependent one dimensional Euler equations of Chap. 3 are one example of a homogeneous system of this kind. Another simplification of (15.1) results from the assumption of no spatial variations,  $\mathbf{F}(\mathbf{U})_x = \mathbf{0}$ , in which case one obtains

$$\frac{d}{dt}\mathbf{U} = \mathbf{S}(\mathbf{U}) , \quad (15.3)$$

which is a system of Ordinary Differential Equations (ODEs).

Inhomogeneous systems of the form (15.1),  $\mathbf{S}(\mathbf{U}) \neq \mathbf{0}$ , arise naturally in many problems of practical interest. A whole class of inhomogeneous systems are derived when reducing the spatial dimensionality of multidimensional problems. For example, under the assumption of spherically or cylindrically symmetric flow, the three or two dimensional Euler equations become a one-dimensional system of the form (15.1); see Sect. 1.6.2 of Chap. 1. In this case the source terms are geometric in character. Sources of similar type are present in the shallow water equations for flow on non-horizontal channels; see Sect. 1.6.3 of Chap. 1.

Important examples of inhomogeneous systems of the form (15.1) arise in the study of the fluid dynamics of reactive gaseous mixtures, where in addition to the fluid dynamics governed by a system like (15.2), there are chemical reactions between the constituent gases, which in the absence of fluid flow may be modelled by systems of the form (15.3). Examples of problems of this kind arise in the study of hypersonic flows [119], [9] and detonation waves, see for instance [191], [72], [78], [118] and [363].

Chemically active flows contain a range of widely varying time scales, which leads to *stiff* ODEs of the form (15.3), [296]. The problem of *stiffness* in ODEs may be resolved by resorting to *implicit methods*. For chemically active flow models (15.1), stiffness may not be resolved by simply using implicit methods. If the mesh is not sufficiently fine in both space and time, then spurious solutions travelling at unphysical speeds may be computed [136], [314], [49], [78], [227]. See also the recent review paper by Yee and Sweby [594]. There are still a number of unresolved problems in solving systems like (15.2), which are the subject of current research.

There are essentially two ways of constructing methods to solve inhomogeneous systems of the form (15.1). One approach attempts to preserve some coupling between the two processes in (15.1). These two processes might be represented by the systems (15.2) (advection) and (15.3) (reaction-like). LeVeque and Yee [314] report on a predictor–corrector scheme of the MacCormack type with a TVD constraint. The idea of *upwinding the source terms* [412], may be seen as an attempt to couple the two processes involved, although the eigenstructure used in projecting the source terms is oblivious to the influence of these.

Another approach is to *split* (15.1), for a time  $\Delta t$ , into the advection problem (15.2) and the source problem (15.3). At first sight this might appear unreasonable. However, for the case of a model inhomogeneous PDE, splitting is actually exact. For more general problems, the fact that one can construct high-order splitting schemes following this approach is also somewhat reassuring. In addition, computational experience suggests that splitting is a viable approach, *if used with caution*. The main attraction of splitting schemes is in the fact that one can deploy the optimal, existing schemes for each subproblem. For instance, to solve the homogeneous subproblem (15.2) one may use directly any of the schemes presented in Chaps. 6 to 14, or any other appropriate method. To solve the subproblem (15.3) one may use directly any of the ODE solvers available. If the system is known to be stiff, then *stiff* solvers must be used.

This chapter is concerned with splitting schemes to solve (15.1). In Sect. 15.2 we show that for a model equation, splitting is exact; in Sect. 15.3 we present numerical schemes based on the splitting approach; in Sect. 15.4 we briefly review some basic aspects of numerical methods for Ordinary Differential Equations (ODEs). Concluding remarks are given in Sect. 15.5.

## 15.2 Splitting for a Model Equation

The simplest model hyperbolic equation of the form (15.1) is given by

$$u_t + au_x = \lambda u , \quad (15.4)$$

where  $a$  is constant wave propagation speed and  $\lambda$  is a constant parameter. This simple model equation will prove very useful in discussing possible strategies for solving (15.1) numerically.

Consider the initial value problem (IVP) for (15.4), namely

$$\left. \begin{array}{l} \text{PDE : } u_t + au_x = \lambda u , \\ \text{IC : } u(x, 0) = u_0(x) . \end{array} \right\} \quad (15.5)$$

Here  $u = u(x, t)$ ,  $-\infty < x < \infty$ ,  $t > 0$  and  $u_0(x)$  is the initial data for the problem at  $t = 0$ . It is easy to verify that the exact solution of IVP (15.5) is

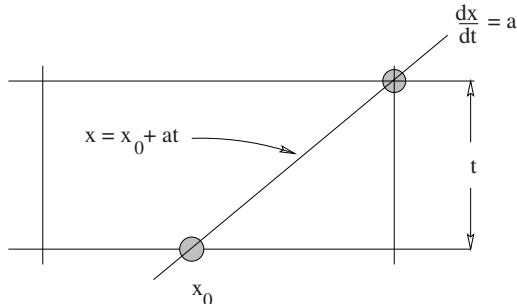
$$u(x, t) = u_0(x - at) e^{\lambda t} . \quad (15.6)$$

Note in particular that if  $\lambda = 0$  we recover the exact solution for the homogeneous equation  $u_t + au_x = 0$ , namely  $u_0(x - at)$ ; see Sect. 2.2 of Chap. 2.

A geometric interpretation of the original IVP (15.5) results if we view (15.5) as an IVP involving an ODE along characteristics, namely

$$\left. \begin{array}{l} \frac{d}{dt} u = \lambda u ; u(0) = u(x_0) , \\ \frac{d}{dt} x = a ; x(0) = x_0 . \end{array} \right\} \quad (15.7)$$

Fig. 15. 1 illustrates the situation. The ODE in (15.7) requires initial data at



**Fig. 15.1.** Illustration of operator splitting scheme for model PDE with source term

the foot of the characteristic curve  $x = x_0 + at$ , namely the point  $x_0$ . Actually the initial data is then  $u(0) = u(x_0) = u(x - at)$ , which is the solution of the homogeneous problem in (15.5),  $\lambda = 0$ .

Next we show that the exact solution (15.6) can also be obtained by solving exactly a pair of initial value problems in succession.

**Theorem 15.1 (Splitting of Source Term).**

The exact solution (15.6) of the inhomogeneous IVP (15.5) can be found by solving exactly the following pair of IVP's.

$$\left. \begin{array}{l} PDE : r_t + ar_x = 0 , \\ IC \quad : r(x, 0) = u_0(x) , \end{array} \right\} \implies r(x, t) \quad (15.8)$$

$$\left. \begin{array}{l} ODE : \frac{d}{dt}s = \lambda s , \\ IC \quad : s(0) = r(x, t) , \end{array} \right\} \implies u(x, t) \quad (15.9)$$

Note here that the initial condition of IVP (15.8) is the actual initial condition for the original IVP (15.5) and the initial condition for IVP (15.9) is the solution  $r(x, t)$  of IVP (15.8).

*Proof.* Clearly the solution of IVP (15.8) is  $r(x, t) = u_0(x - at)$ , while the exact solution of IVP (15.9) is  $s(x(t), t) = s(0)e^{\lambda t}$ . But  $s(0) = r(x, t) = u_0(x - at)$  and thus the resulting solution of IVPs (15.8) and (15.9) is

$$s(x, t) = u_0(x - at)e^{\lambda t} ,$$

which is the exact solution (15.6) of the original inhomogeneous IVP (15.5), and the theorem is thus proved.

The result on the splitting scheme obtained by solving in succession (15.8), (15.9) can be expressed in the succinct form

$$u(x, t) = S^{(t)}C^{(t)}[u_0(x)] . \quad (15.10)$$

We interpret  $C^{(t)}$  as the *solution operator* for the advection problem (15.8) applied over a time  $t$  and  $S^{(t)}$  as the *solution operator* for the ODE (15.9) applied for a time  $t$ .

**Exercise 15.2.** Show that the exact solution  $u(x, t)$  of IVP (15.5) can be obtained by solving

$$\left. \begin{array}{l} PDE : u_t + f(u)_x = 0 ; f(u) = au , \\ IC \quad : u(x, 0) = u_0(x) \end{array} \right\} \implies \bar{u}^{n+1} \quad (15.11)$$

and

$$\left. \begin{array}{l} ODE : \frac{d}{dx}f(u) = \lambda u , \\ IC \quad : \bar{u}^{n+1} \end{array} \right\} \implies u^{n+1} \quad (15.12)$$

in succession.

This result says that the splitting scheme (15.8), (15.9) modified by replacing the ODE in time by an ODE in space also gives the exact solution. This splitting scheme can be expressed in the succinct form

$$u(x, t) = S^{(x)}C^{(t)}[u_0(x)] , \quad (15.13)$$

where  $S^{(x)}$  denotes the solution operator for the ODE in  $x$  in (15.12).

**Solution 15.3.** Left to the reader.

Glimm, Marshall and Plohr [214] constructed numerical splitting schemes of the form (15.11), (15.12) to solve numerically one-dimensional flows with area variation. In this case the source term involves a spatial derivative and does not depend on time.

In the next section we construct numerical methods based on the splitting approach, or fractional step approach [466], [588].

## 15.3 Numerical Methods Based on Splitting

We have shown that for the model inhomogeneous PDE (15.4) the splitting approach, as described in the previous section, is *exact*. For non-linear systems (15.1) this result is no longer valid. However, approximate, numerical schemes based on the splitting approach can be constructed. We first consider the scalar case.

### 15.3.1 Model Equations

The splitting scheme (15.8)–(15.9), represented by (15.10), is exact if the operators  $C$  and  $S$  are exact. Here we are interested in constructing numerical methods for the scalar IVP

$$\left. \begin{array}{l} \text{PDE : } u_t + f(u)_x = s(u) : 0 \leq x \leq L , \\ \text{IC : } u(x, t^n) = u^n , \end{array} \right\} \quad (15.14)$$

To this end we replace the exact operators  $C^{(t)}$  and  $S^{(t)}$  in (15.10) by approximate operators and re-state the problem in a numerical context. Given the IVP (15.14), we want to evolve the solution from its initial value  $u^n$  at a time  $t^n$ , by one time step of size  $\Delta t$ , to a value  $u^{n+1}$  at time  $t^{n+1} = t^n + \Delta t$ . We assume the spatial domain  $[0, L]$  has been discretised into a finite number  $M$  of cells  $i$  (finite volume approach) or grid points  $i$  (finite difference approach). Here  $u^n$  is a set of discrete values  $u_i^n$  at time  $t^n$ . The discrete analogue of the splitting scheme (15.8)–(15.9) is now

$$\left. \begin{array}{l} \text{PDE : } u_t + f(u)_x = 0 , \\ \text{IC : } u(x, t^n) = u^n , \end{array} \right\} \implies \bar{u}^{n+1} \quad (15.15)$$

$$\left. \begin{array}{l} \text{ODE : } \frac{d}{dt} u = s(u) , \\ \text{IC : } \bar{u}^{n+1} . \end{array} \right\} \implies u^{n+1} \quad (15.16)$$

The initial condition for the advection problem (15.15) is the initial condition for the complete problem (15.14). The solution of (15.15) after a time  $\Delta t$  is denoted by  $\bar{u}^{n+1}$  and is used as the initial condition for the second IVP (15.16). This second IVP accounts for the presence of the source term  $s(u)$

and is also solved for a complete time step  $\Delta t$ ; this solution is then regarded as an approximation to the solution  $u^{n+1}$  of the full problem (15.14) at a time  $t^{n+1} = t^n + \Delta t$ . If the numerical analogues of  $S^{(t)}$  and  $C^{(t)}$  in (15.10) are still denoted by  $S$  and  $C$ , then we can write the splitting of (15.14) into (15.15)–(15.16) as

$$u^{n+1} = S^{(\Delta t)} C^{(\Delta t)}(u^n). \quad (15.17)$$

Each numerical sub–problem (15.15), (15.16) is dealt with separately, for a time step  $\Delta t$ . One requires a numerical method to solve the homogeneous advection problem (15.15) and another numerical method to solve the ordinary differential equation in (15.16), with the initial data taken from the solution of (15.15). This procedure for solving inhomogeneous systems is exceedingly simple but is only first–order accurate in time, when  $S$  and  $C$  are at least first–order accurate solution operators. A second–order accurate scheme is

$$u^{n+1} = S^{(\frac{1}{2}\Delta t)} C^{(\Delta t)} S^{(\frac{1}{2}\Delta t)}(u^n), \quad (15.18)$$

where  $S$  and  $C$  are at least second–order accurate solution operators.

### 15.3.2 Schemes for Systems

Here we extend the application of the splitting scheme of the previous section to non–linear systems of the form (15.1). The generalisation of (15.15), (15.16) to solve (15.1) is straightforward. Given the IVP

$$\left. \begin{array}{l} \text{PDE's : } \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{S}(\mathbf{U}) ; 0 \leq x \leq L , \\ \text{IC : } \mathbf{U}(x, t^n) = \mathbf{U}^n , \end{array} \right\} \quad (15.19)$$

we want to evolve  $\mathbf{U}^n$  from time  $t = t^n$  to the new value  $\mathbf{U}^{n+1}$  at  $t = t^{n+1}$  in a time step  $\Delta t = t^{n+1} - t^n$ . The splitting (15.15), (15.16) becomes

$$\left. \begin{array}{l} \text{PDE's : } \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0} , \\ \text{IC : } \mathbf{U}(x, t^n) = \mathbf{U}^n \end{array} \right\} \Rightarrow \overline{\mathbf{U}}^{n+1}, \quad (15.20)$$

$$\left. \begin{array}{l} \text{PDE's : } \frac{d}{dt} \mathbf{U} = \mathbf{S}(\mathbf{U}) , \\ \text{IC's : } \overline{\mathbf{U}}^{n+1} \end{array} \right\} \Rightarrow \mathbf{U}^{n+1}. \quad (15.21)$$

The analogue of the first–order scheme (15.17) is

$$\mathbf{U}^{n+1} = S^{(\Delta t)} C^{(\Delta t)}(\mathbf{U}^n). \quad (15.22)$$

A second–order accurate scheme for systems is

$$\mathbf{U}^{n+1} = S^{(\frac{1}{2}\Delta t)} C^{(\Delta t)} S^{(\frac{1}{2}\Delta t)}(\mathbf{U}^n). \quad (15.23)$$

A splitting scheme based on (15.11), (15.12) is

$$\left. \begin{array}{l} \text{PDEs : } \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0}, \\ \text{IC : } \mathbf{U}(x, t^n) = \mathbf{U}^n, \end{array} \right\} \Rightarrow \bar{\mathbf{U}}^{n+1} \quad (15.24)$$

$$\left. \begin{array}{l} \text{ODEs : } \frac{d}{dx} \mathbf{F}(\mathbf{U}) = \mathbf{S}(\mathbf{U}), \\ \text{IC : } \bar{\mathbf{U}}^{n+1}. \end{array} \right\} \Rightarrow \mathbf{U}^{n+1} \quad (15.25)$$

There appears to be little experience in using this approach. For source terms that are independent of time or involve spatial derivatives, this approach may be advantageous. See [214].

The attraction of splitting schemes is in the freedom available in choosing the numerical operators  $S$  and  $C$ . In general, one may choose the best scheme for each type of problems. For solving the advection (homogeneous) IVP (15.20) one can, for instance, use any of the schemes studied in Chaps. 6 to 14, or some other method. For solving the ODEs in (15.21), (15.25) one may choose some appropriate ODE solver, see next section.

## 15.4 Remarks on ODE Solvers

There is a vast literature on ODEs and on numerical methods for solving ODEs. For theoretical properties of ODEs see for example Brown [81], Ince and Sneddon [260], Sánchez [423] and Coddington and Levinson [130]. Almost any textbook on Numerical Analysis will contain some chapter on schemes for ODEs. See for example Hildebrand [249]; Mathews [347]; Conte and de Boor [139]; Maron and Lopez [337]; Johnson and Riess [273]; Kahaner, Moler and Nash [275]. Advanced textbooks are those of Gear [200], Lambert [296] and Shampine [441].

### 15.4.1 First-Order Systems of ODEs

Here we recall some very basic facts about first-order systems of ODEs

$$\frac{d}{dt} \mathbf{U}(t) \equiv \mathbf{U}' = \mathbf{S}(t, \mathbf{U}(t)). \quad (15.26)$$

Here  $\mathbf{U} = \mathbf{U}(t)$  and  $\mathbf{S}(t, \mathbf{U}(t))$  are vector-valued functions of  $m$  components

$$\mathbf{U} = [u_1, u_2, \dots, u_m]^T; \quad \mathbf{S} = [s_1, s_2, \dots, s_m]^T \quad (15.27)$$

and the independent variable  $t$  is a time-like variable. The Jacobian  $\mathbf{A}(\mathbf{U})$  is defined as the matrix

$$\mathbf{A}(\mathbf{U}) = \partial \mathbf{S} / \partial \mathbf{U} = \begin{bmatrix} \partial s_1 / \partial u_1 & \dots & \partial s_1 / \partial u_m \\ \partial s_2 / \partial u_1 & \dots & \partial s_2 / \partial u_m \\ \vdots & \vdots & \vdots \\ \partial s_m / \partial u_1 & \dots & \partial s_m / \partial u_m \end{bmatrix}. \quad (15.28)$$

The entries  $a_{ij}$  of  $\mathbf{A}(\mathbf{U})$  are partial derivatives of the components  $s_i$  of the vector  $\mathbf{S}$  with respect to the components  $u_j$  of the vector  $\mathbf{U}$ , that is  $a_{ij} = \partial s_i / \partial u_j$ . The eigenvalues  $\lambda_i$  of  $\mathbf{A}$  are the solutions of the *characteristic polynomial*

$$|\mathbf{A} - \lambda \mathbf{I}| = \det(\mathbf{A} - \lambda \mathbf{I}) = 0 , \quad (15.29)$$

where  $\mathbf{I}$  is the identity matrix. Generally, the eigenvalues are complex numbers. Trivially, the eigenvalue of the model ODE

$$u'(t) = \lambda u(t) \quad (15.30)$$

is  $\lambda$ . The behaviour of a system of ODEs is, in the main, determined by the behaviour of its eigenvalues. For instance, the exact solution of (15.30) with initial condition  $u(0) = 1$  is  $u(t) = e^{\lambda t}$ . For  $t$  close to 0 the solution varies rapidly if the eigenvalue is negative and large in absolute value. For  $t$  away from zero the solution is almost indistinguishable from 0.

An important property of ODEs is that of *stability*. Generally speaking stable solutions are bounded. Note that the solution of the linear ODE (15.30) is bounded only if  $\lambda < 0$ . Geometrically, a solution  $\mathbf{U}(t)$  is stable if any other solution of the ODE whose initial condition is sufficiently close to that of  $\mathbf{U}(t)$  remains in a *tube* enclosing  $\mathbf{U}(t)$ . If the diameter of the tube tends to 0 as  $t$  tends to  $\infty$ , the solution is said to be *asymptotically stable*. Stability of solutions  $\mathbf{U}(t)$  is characterised in terms of the eigenvalues  $\lambda_j$  of the Jacobian matrix. In particular if the real part of every eigenvalue is negative the solution is asymptotically stable.

Another feature of ODEs is that of *stiffness*. Stiff ODEs are usually associated with processes operating on disparate time scales. Chemical kinetics is a classical source of stiff ODEs. The stiffness of a system is generally determined by the behaviour of the eigenvalues of the system. In addition, the time interval over which the solution is sought is also a consideration in determining the stiffness of the system. There will be intervals of *rapid variations* (transient) of the solution and intervals of *slow variation*. The single ODE (15.30) is stiff for  $\lambda \ll 0$  and for time  $t$  in the vicinity of 0.

Following Lambert [296], a nonlinear system of the form (15.26) is said to be stiff if

- (i)  $Re(\lambda_j) < 0$ ,  $j = 1, 2, \dots, m$  and
- (ii)  $\lambda_{\max} \equiv \max_j |Re(\lambda_j)| \gg \lambda_{\min} \equiv \min_j |Re(\lambda_j)|$ .

Here  $Re(\lambda_j)$  denotes the real part of the complex number  $\lambda_j$ . The stiffness ratio is defined as  $R_{stif} = \lambda_{\max}/\lambda_{\min}$ . Modest values of  $R_{stif}$ , e.g. 20, are sufficient to cause serious numerical difficulties to explicit methods. In real applications  $R_{stif}$  may be as large as  $10^6$ .

Before thinking of numerical methods to solve ODEs, a fundamental question is to investigate whether the ODEs are stiff or not; this will determine the appropriate numerical methods to be used for solving the equations. See Kahaner, Moler and Nash [275], Gear [200] and Lambert [296].

### 15.4.2 Numerical Methods

We are interested in solving the Initial Value Problem (IVP) for (15.26) with initial condition

$$\mathbf{U}(t^0) = \mathbf{U}_0. \quad (15.31)$$

Discretise the domain of integration  $[t^0, t^f]$  through the partition  $t^0 < t^1 < t^2 \dots < t^n < t^{n+1} \dots < t^f$ . One way of constructing numerical methods to solve the IVP (15.26), (15.31) is by using Taylor series expansions. Another way is to integrate (15.26) between  $t^n$  and  $t^{n+1}$  to obtain

$$\mathbf{U}(t^{n+1}) = \mathbf{U}(t^n) + \int_{t^n}^{t^{n+1}} \mathbf{S}(t, \mathbf{U}(t)) dt. \quad (15.32)$$

Various numerical methods are obtained depending on the way the integral is evaluated. The *Euler Method* results from evaluating the integral at the *old* time,

$$\mathbf{U}^{n+1} = \mathbf{U}^n + \Delta t \mathbf{S}(t^n, \mathbf{U}^n). \quad (15.33)$$

where  $\Delta t = t^{n+1} - t^n$  is the time step and  $\mathbf{U}^n \approx \mathbf{U}(t^n)$ . The Euler method is *explicit* and first-order accurate. The *Backward Euler Method*, also first order accurate but *implicit*, results from evaluating the integral at the *new* time  $t^{n+1}$ , namely

$$\mathbf{U}^{n+1} = \mathbf{U}^n + \Delta t \mathbf{S}(t^{n+1}, \mathbf{U}^{n+1}). \quad (15.34)$$

A second-order implicit method results from a *trapezium rule* approximation to the integral, giving the *Trapezoidal Method*

$$\mathbf{U}^{n+1} = \mathbf{U}^n + \frac{1}{2} \Delta t [\mathbf{S}(t^n, \mathbf{U}^n) + \mathbf{S}(t^{n+1}, \mathbf{U}^{n+1})]. \quad (15.35)$$

A second-order, two stage *Runge–Kutta method* (explicit) is

$$\left. \begin{aligned} \mathbf{K}_1 &= \Delta t \mathbf{S}(t^n, \mathbf{U}^n), \\ \mathbf{K}_2 &= \Delta t \mathbf{S}(t^n + \Delta t, \mathbf{U}^n + \mathbf{K}_1), \\ \mathbf{U}^{n+1} &= \mathbf{U}^n + \frac{1}{2} [\mathbf{K}_1 + \mathbf{K}_2]. \end{aligned} \right\} \quad (15.36)$$

A fourth-order, four stage Runge–Kutta method (explicit) is

$$\left. \begin{aligned} \mathbf{K}_1 &= \Delta t \mathbf{S}(t^n, \mathbf{U}^n), \\ \mathbf{K}_2 &= \Delta t \mathbf{S}(t^n + \frac{1}{2} \Delta t, \mathbf{U}^n + \frac{1}{2} \mathbf{K}_1), \\ \mathbf{K}_3 &= \Delta t \mathbf{S}(t^n + \frac{1}{2} \Delta t, \mathbf{U}^n + \frac{1}{2} \mathbf{K}_2), \\ \mathbf{K}_4 &= \Delta t \mathbf{S}(t^n + \Delta t, \mathbf{U}^n + \mathbf{K}_3), \\ \mathbf{U}^{n+1} &= \mathbf{U}^n + \frac{1}{6} [\mathbf{K}_1 + 2\mathbf{K}_2 + 2\mathbf{K}_3 + \mathbf{K}_4]. \end{aligned} \right\} \quad (15.37)$$

Stability of numerical methods is a most important issue. To illustrate this point consider the model ODE (15.30) as solved by the explicit Euler method (15.33). The scheme reads

$$u^{n+1} = (1 + \Delta t \lambda) u^n . \quad (15.38)$$

Clearly, for stability one requires that the ODE itself be stable,  $\lambda < 0$ , and that the *amplification factor* satisfy  $|1 + \lambda \Delta t| \leq 1$ . Therefore the time step  $\Delta t$  must satisfy the stability restriction

$$\Delta t \leq \frac{2}{|\lambda|} . \quad (15.39)$$

For large  $|\lambda|$  (stiff ODE)  $\Delta t$  can be extremely small, which means that the method becomes very inefficient or even useless in practice.

On the other hand, the Trapezoidal method (15.35), which is implicit, gives

$$u^{n+1} = \frac{\left(1 + \frac{1}{2}\Delta t \lambda\right)}{\left(1 - \frac{1}{2}\Delta t \lambda\right)} u^n . \quad (15.40)$$

This is stable for any  $\Delta t$ , provided  $\lambda \leq 0$ , that is whenever the ODE itself is stable.

Explicit methods are much simpler to use than implicit methods. The latter require the solution of non-linear algebraic equations at each time step and are therefore much more expensive. However, as illustrated, for stiff problems implicit methods are the only methods to use in any practical situation.

### 15.4.3 Implementation Details for Split Schemes

There are two facts that need to be emphasised when solving ODEs in the context of the splitting schemes described in Sect. 15.3. First, at every time  $t^n$ , at each mesh point  $i$  one has a system of ODEs to solve; second, the time evolution of the ODEs is generally short and is dictated by the time step in the overall splitting scheme. This second point is relevant when choosing ODE solvers.

Before selecting a method, an analysis of the ODEs must be performed. If the problem is non-stiff then a high-order explicit method is recommended. A stability analysis of the method must be carried out and enforced when selecting the size of the time step  $\Delta t$ . For simplicity, let us assume we want to implement the first-order splitting scheme (15.22). A practical problem is to determine the time step  $\Delta t$ . One first determines the time step  $\Delta t_c$  for the advection problem (15.20). If this problem is solved by some explicit method, e.g. Godunov's first order upwind method (see Chap. 6), then  $\Delta t_c$  is found from some stability constraint, i.e. the Courant condition, see Chap. 6. The solution of the advection problem is found at every mesh point  $i$  and this gives  $\bar{\mathbf{U}}^{n+1}$ , which is then used as the initial data for the ODE step (15.21). If the ODEs are solved by some implicit method, then there will be no stability restriction on the time step, and therefore one can advance the solution via the ODE solver by a time  $\Delta t_s = \Delta t_c$ , in *one go*. However, if an explicit method is used to solve the ODEs, then a stable time step  $\Delta t_s$  must first be found.

If  $\Delta t_s \geq \Delta t_c$ , then one may again advance the solution via the ODEs by a time  $\Delta t_c$  in *one go*. Hence the final solution at time  $t^{n+1}$  has been advanced by a time  $\Delta t = \Delta t_c$ . If  $\Delta t_s < \Delta t_c$ , then one possibility is to update via the ODEs in  $k$  steps of size  $\Delta t_k = \Delta t_c/k$ , where  $k$  is a positive integer such such  $\Delta t_k$  is a stable time step for the ODE solver. The previous observations apply directly when implementing the second-order splitting scheme (15.23). A useful reference is Chiang and Hoffmann [108].

## 15.5 Concluding Remarks

We have only presented one approach for treating source terms. There are other approaches, but at the present time there appears to be no clear, and sufficiently general, alternative to splitting. The idea of *upwinding* the source terms proposed by Roe [412] appears to work well for certain problems. See for instance the work of Vázquez [568] and that of Bermúdez and Vázquez [51]. See also the recent paper of Vázquez [569], which addresses the issue of geometric and friction source terms in shallow water models. For steady-state problems computed by time-marching schemes the reader should consult the recent paper by LeVeque [310].

The reader is strongly encouraged to utilise problems with exact solutions, whenever available, to carefully assess the numerical methods before applying them to *the real problem*. The simplest test problem is the IVP (15.5) with exact solution (15.6). More scalar test problems are found in [314], [136], [227]. A test problem with exact solution for a  $2 \times 2$  non-linear system is the so called Fickett detonation analogue [191], see also example 2.4.3 of Sect. 2.4.2 in Chap. 2. This problem is exploited in [118] for testing numerical methods for detonation waves in high-energy solids. For the Euler equations, a test problem with exact solution is reported in [120]. Details of the solution are given in [115] and applications are also shown in [412] and [471].

For certain types of problems, such as detonation waves, there are serious difficulties in designing numerical methods to properly account for the fluid dynamics and the chemistry. For sufficiently fine meshes such difficulties may be overcome but at a cost that is impossible to meet with current computing resources, if realistic problems in multidimensions are to be solved. Since the early papers by Colella, Majda and Roytburd [136] and that of LeVeque and Yee [314], there has been a noticeable increase in the interest for hyperbolic systems with source terms, both numerically and theoretically. See, amongst others: Griffiths, Stuart and Yee [227]; Berkenbosch, Kaasschieter and Boonkamp [49]; Bourlioux, Majda and Roytburd [78]; Fey, Jeltsch and Müller [190]; Chalabi [98], [99], [100], [101]; Benkhaldoun and Chalabi [44]; Pember [377], [378]; Schroll and Tveito [432]; Schroll and Winther [434]; Schroll, Tveito and Winther [433]; Corberán and Gascón [140]; Lorenz and Schroll [333]; Lafon and Yee [294], [295]; and especially the review paper of Yee and Sweby [594].

The splitting approach may also be applied to treat diffusion like terms [82], [529] in exactly the same manner as for algebraic source terms. The splitting approach also offers one way of solving multidimensional problems; this topic is dealt with in Chap. 16.

---

## Methods for Multi–Dimensional PDEs

### 16.1 Introduction

This chapter is concerned with numerical methods for solving non–linear systems of hyperbolic conservation laws in multidimensions. For Cartesian geometries one may write the equations of our interest here as

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_y + \mathbf{H}(\mathbf{U})_z = \mathbf{0}, \quad (16.1)$$

where  $t$  denotes time or a time–like variable and  $x, y, z$  are Cartesian coordinate directions.  $\mathbf{U}$  is the vector of conserved variables and  $\mathbf{F}(\mathbf{U}), \mathbf{G}(\mathbf{U}), \mathbf{H}(\mathbf{U})$  are vectors of fluxes in the  $x, y, z$  directions respectively. A prominent example are the time–dependent three dimensional Euler equations studied in Sect. 3.2 of Chap. 3. Other examples are the time–dependent two dimensional shallow water equations and the artificial compressibility equations. See Sect. 1.6.3 of Chap. 1.

We shall present two ways of solving (16.1) numerically. The first approach is *dimensional splitting* or *method of fractional steps* [466], [588]. This method is also known as time–operator splitting. In this approach one applies one–dimensional methods in each coordinate direction. A simple extension of the one–dimensional methods is required to solve the one–dimensional conservation laws *augmented* by the extra components of velocity present in multidimensional problems. Any of the methods studied in Chaps. 6 to 14 can be applied to solve (16.1), in conjunction with dimensional splitting. The other approach studied here is the *finite volume method*, whereby in updating the solution within some control volume one includes all the intercell flux contributions in a single step. Only some of the one–dimensional approaches discussed in Chaps. 6 to 14 can be extended to multidimensional problems following the (unsplit) finite volume approach. There is a close relationship between dimensional splitting and the finite volume method. If the schemes used in the dimensional splitting approach are of the finite–volume type in one space dimension, see Chaps. 5 to 12 and 14, then the resulting splitting schemes may

be viewed as being predictor–corrector type finite volume schemes in multi-dimensions. Both the dimensional splitting and the finite volume approaches may be extended to deal with non–Cartesian geometries.

## 16.2 Dimensional Splitting

Before introducing splitting schemes for multidimensional systems of conservation laws, we study the splitting approach as applied to a model conservation law in three space dimensions, for which the dimensional splitting method can be shown to be exact.

### 16.2.1 Splitting for a Model Problem

Consider the Initial–Value Problem (IVP) for the linear advection equation in three space dimensions

$$\left. \begin{array}{l} \text{PDE : } u_t + a_1 u_x + a_2 u_y + a_3 u_z = 0 , \\ \text{IC : } u(x, y, z, t^n) = u_0(x, y, z) \equiv u^n . \end{array} \right\} \quad (16.2)$$

Here  $a_1$ ,  $a_2$  and  $a_3$  are the three velocity components of a constant velocity vector  $\mathbf{V} = (a_1, a_2, a_3)$ . Let us consider three one-dimensional IVPs in the  $x$ ,  $y$  and  $z$  directions respectively. The IVP in the  $x$ –direction, or  $x$  sweep, is

$$\left. \begin{array}{l} \text{PDE : } u_t + a_1 u_x = 0 \\ \text{IC : } u^n \end{array} \right\} \xrightarrow{\Delta t} u^{n+\frac{1}{3}} . \quad (16.3)$$

The initial data for this IVP is the initial data for the original full IVP (16.2) and its solution after a time  $\Delta t$  is denoted by  $u^{n+\frac{1}{3}}$ . The  $y$ –direction IVP, or  $y$  sweep, is

$$\left. \begin{array}{l} \text{PDE : } u_t + a_2 u_y = 0 \\ \text{IC : } u^{n+\frac{1}{3}} \end{array} \right\} \xrightarrow{\Delta t} u^{n+\frac{2}{3}} . \quad (16.4)$$

The initial data for this IVP is the solution  $u^{n+\frac{1}{3}}$  of IVP (16.3) and its solution after a time  $\Delta t$  is denoted by  $u^{n+\frac{2}{3}}$ . The  $z$ –direction IVP, or  $z$  sweep, is

$$\left. \begin{array}{l} \text{PDE : } u_t + a_3 u_z = 0 \\ \text{IC : } u^{n+\frac{2}{3}} \end{array} \right\} \xrightarrow{\Delta t} u^{n+1} . \quad (16.5)$$

The initial data for this IVP is the solution  $u^{n+\frac{2}{3}}$  of IVP (16.4) and its solution after a time  $\Delta t$  is denoted by  $u^{n+1}$ , which is then regarded as the solution to the full problem (16.2) after a time  $\Delta t$ .

Let us denote by  $\mathcal{X}^{(t)}$ ,  $\mathcal{Y}^{(t)}$  and  $\mathcal{Z}^{(t)}$  the solution operators for IVPs (16.3), (16.4) and (16.5) respectively, when solved *exactly* for a time  $t$ . Then the following result can be proved.

**Proposition 16.1 (Dimensional Splitting).** *The exact solution  $u^{n+1}$  of the three-dimensional IVP (16.2) after a time  $\Delta t$  can be obtained by solving exactly the sequence of three one-dimensional IVPs (16.3)–(16.5), each for a time  $\Delta t$ , that is*

$$u^{n+1} = \mathcal{Z}^{(\Delta t)} \mathcal{Y}^{(\Delta t)} \mathcal{X}^{(\Delta t)}(u^n). \quad (16.6)$$

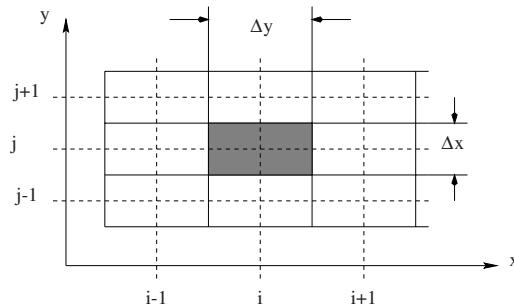
*Proof.* (Left as an exercise. See [308]).

### 16.2.2 Splitting Schemes for Two-Dimensional Systems

For non-linear systems, dimensional splitting is not exact but one may construct approximate splitting schemes. Consider the two-dimensional initial value problem

$$\left. \begin{array}{l} \text{PDE : } \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_y = \mathbf{0}, \\ \text{IC : } \mathbf{U}(x, y, t^n) = \mathbf{U}^n. \end{array} \right\} \quad (16.7)$$

The initial data at a time  $t^n$  is given by the set  $\mathbf{U}^n$  of discrete cell average values  $\mathbf{U}_{i,j}^n$ ; index  $i$  refers to the  $x$ -coordinate direction and index  $j$  refers to the  $y$ -coordinate direction. Fig. 16.1 illustrates a finite volume discretisation of a two-dimensional domain in the  $x$ - $y$  plane.



**Fig. 16.1.** Discretisation of two-dimensional Cartesian domain into finite volumes  $I_{ij}$  of area  $\Delta x \times \Delta y$

The dimensional splitting approach replaces (16.7) by a pair of one-dimensional IVPs. The simplest version of the approach replaces (16.7) by the sequence of IVPs

$$\left. \begin{array}{l} \text{PDEs : } \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0} \\ \text{ICs : } \mathbf{U}^n \end{array} \right\} \xrightarrow{\Delta t} \mathbf{U}^{n+\frac{1}{2}} \quad (16.8)$$

and

$$\left. \begin{array}{l} \text{PDEs : } \mathbf{U}_t + \mathbf{G}(\mathbf{U})_y = \mathbf{0} \\ \text{ICs : } \mathbf{U}^{n+\frac{1}{2}} \end{array} \right\} \xrightarrow{\Delta t} \mathbf{U}^{n+1}. \quad (16.9)$$

In the first IVP (16.8) one solves a one-dimensional problem in the  $x$ -direction for a time step  $\Delta t$ . This is called the  $x$  *sweep* and its solution is denoted by  $\mathbf{U}^{n+\frac{1}{2}}$ . Note that *for each strip labelled  $j$* , see Fig. 16.1, one solves the one-dimensional problem (16.8). In the next IVP (16.9) one solves a one-dimensional problem in the  $y$ -direction, also for a time step  $\Delta t$ . This is called the  $y$  *sweep*. The initial condition for the second IVP (16.9) is the solution  $\mathbf{U}^{n+\frac{1}{2}}$  of IVP (16.8). Note here that *for each strip labelled  $i$* , see Fig. 16.1, one solves the one-dimensional problem (16.9). Both sweeps have a common time step  $\Delta t$ ; in Sec. 16.3.2 we discuss ways of determining  $\Delta t$ .

If  $\mathcal{X}^{(t)}$  and  $\mathcal{Y}^{(t)}$  are approximate solution operators for IVPs (16.8) and (16.9), then the splitting (16.8), (16.9) of the original two-dimensional IVP (16.7) can be written thus

$$\mathbf{U}^{n+1} = \mathcal{Y}^{(\Delta t)} \mathcal{X}^{(\Delta t)} (\mathbf{U}^n). \quad (16.10)$$

There is no particular reason for applying the operators in the order just described. An equivalent scheme is

$$\mathbf{U}^{n+1} = \mathcal{X}^{(\Delta t)} \mathcal{Y}^{(\Delta t)} (\mathbf{U}^n). \quad (16.11)$$

It can be shown, see Strang [466], that for general systems both splitting schemes (16.10) and (16.11) are first-order accurate in time if the individual operators  $\mathcal{X}$  and  $\mathcal{Y}$  are at least first-order accurate accurate in time. A second-order accurate splitting [466] is

$$\mathbf{U}^{n+1} = \frac{1}{2} \left[ \mathcal{X}^{(\Delta t)} \mathcal{Y}^{(\Delta t)} + \mathcal{Y}^{(\Delta t)} \mathcal{X}^{(\Delta t)} \right] (\mathbf{U}^n), \quad (16.12)$$

provided each of the solution operators are at least second-order accurate in time. Note however that this scheme requires double the amount of work of schemes (16.10) and (16.11). More attractive second-order schemes are

$$\mathbf{U}^{n+1} = \mathcal{X}^{(\frac{1}{2}\Delta t)} \mathcal{Y}^{(\Delta t)} \mathcal{X}^{(\frac{1}{2}\Delta t)} (\mathbf{U}^n) \quad (16.13)$$

and

$$\mathbf{U}^{n+1} = \mathcal{Y}^{(\frac{1}{2}\Delta t)} \mathcal{X}^{(\Delta t)} \mathcal{Y}^{(\frac{1}{2}\Delta t)} (\mathbf{U}^n). \quad (16.14)$$

Schemes (16.13) and (16.14) require about 50% more work than the first order schemes (16.10) and (16.11). Strang [466] suggested a modification to schemes (16.13) and (16.14) so that they become as efficient as (16.10) or (16.11), while still preserving formal second-order accuracy. Suppose (16.13) is applied over  $m$  time steps of size  $\Delta t$

$$\mathbf{U}^{n+m} = \left( \mathcal{X}^{(\frac{1}{2}\Delta t)} \mathcal{Y}^{(\Delta t)} \mathcal{X}^{(\frac{1}{2}\Delta t)} \right)^m (\mathbf{U}^n). \quad (16.15)$$

For example, two successive applications of (16.13) give

$$\mathcal{X}^{(\frac{1}{2}\Delta t)} \mathcal{Y}^{(\Delta t)} \mathcal{X}^{(\frac{1}{2}\Delta t)} \mathcal{X}^{(\frac{1}{2}\Delta t)} \mathcal{Y}^{(\Delta t)} \mathcal{X}^{(\frac{1}{2}\Delta t)} (\mathbf{U}^n).$$

The two successive applications of  $\mathcal{X}^{(\frac{1}{2}\Delta t)}$  can be combined into a single application of  $\mathcal{X}^{(\Delta t)}$ . It is then easy to see that (16.15) can be expressed as

$$\mathbf{U}^{n+m} = \mathcal{X}^{(\frac{1}{2}\Delta t)} \left[ \mathcal{Y}^{(\Delta t)} \mathcal{X}^{(\Delta t)} \right]^{m-1} \mathcal{Y}^{(\Delta t)} \mathcal{X}^{(\frac{1}{2}\Delta t)} (\mathbf{U}^n) \quad (16.16)$$

or

$$\mathbf{U}^{n+m} = \mathcal{X}^{(\frac{1}{2}\Delta t)} \mathcal{Y}^{(\Delta t)} \left[ \mathcal{X}^{(\Delta t)} \mathcal{Y}^{(\Delta t)} \right]^{m-1} \mathcal{X}^{(\frac{1}{2}\Delta t)} (\mathbf{U}^n). \quad (16.17)$$

Note that in the manipulations above we have assumed that the size of the time step  $\Delta t$  is constant for at least  $m$  time steps. For linear systems this can certainly be imposed but not for non-linear systems, as  $\Delta t$  depends on the wave speeds. In practice one still applies the above schemes (16.16), (16.17), where the choice of  $m$  depends on the required output times. In implementing these schemes it is reasonable to ensure that the size of the time step  $\Delta t$  in the operators outside the squared brackets is constant. These splitting schemes are almost as efficient as the first-order splittings (16.10), (16.11).

Two more second-order accurate schemes are

$$\mathbf{U}^{n+2} = \mathcal{X}^{(\Delta t)} \mathcal{Y}^{(\Delta t)} \mathcal{Y}^{(\Delta t)} \mathcal{X}^{(\Delta t)} (\mathbf{U}^n) \quad (16.18)$$

and

$$\mathbf{U}^{n+2} = \mathcal{Y}^{(\Delta t)} \mathcal{X}^{(\Delta t)} \mathcal{X}^{(\Delta t)} \mathcal{Y}^{(\Delta t)} (\mathbf{U}^n). \quad (16.19)$$

These can be shown to be second-order accurate every other time step, see Warming and Beam [574], and are as efficient as the first-order schemes (16.10) and (16.11).

The principle of dimensional splitting applies directly to three-dimensional problems, which is the subject of the next section.

### 16.2.3 Splitting Schemes for Three-Dimensional Systems

Consider the three-dimensional IVP

$$\left. \begin{array}{l} \text{PDE : } \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_y + \mathbf{H}(\mathbf{U})_z = \mathbf{0}, \\ \text{IC : } \mathbf{U}(x, y, z, t^n) = \mathbf{U}^n \end{array} \right\} \quad (16.20)$$

The initial data  $\mathbf{U}^n$  at a time  $t^n$  is a set of discrete cell averages  $\mathbf{U}_{i,j,k}^n$  on a three-dimensional Cartesian computational domain. The simplest version of the dimensional splitting approach replaces (16.20) by three one-dimensional IVPs, namely

$$\left. \begin{array}{l} \text{PDEs : } \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0} \\ \text{ICs : } \mathbf{U}^n \end{array} \right\} \xrightarrow{\Delta t} \mathbf{U}^{n+\frac{1}{3}}, \quad (16.21)$$

$$\left. \begin{array}{l} \text{PDEs : } \mathbf{U}_t + \mathbf{G}(\mathbf{U})_y = \mathbf{0} \\ \text{ICs : } \mathbf{U}^{n+\frac{1}{3}} \end{array} \right\} \xrightarrow{\Delta t} \mathbf{U}^{n+\frac{2}{3}}, \quad (16.22)$$

and

$$\left. \begin{array}{l} \text{PDEs : } \mathbf{U}_t + \mathbf{H}(\mathbf{U})_z = \mathbf{0} \\ \text{ICs : } \mathbf{U}^{n+\frac{2}{3}} \end{array} \right\} \xrightarrow{\Delta t} \mathbf{U}^{n+1}. \quad (16.23)$$

The first IVP (16.21) solves an *augmented* one-dimensional problem in the  $x$ -direction, the  $x$  *sweep*, for a time step  $\Delta t$  and the solution is denoted by  $\mathbf{U}^{n+\frac{1}{3}}$ . The next IVP (16.22) solves a one-dimensional (augmented) problem in the  $y$ -direction, the  $y$  *sweep*, for a time step  $\Delta t$ ; the initial condition for the second IVP (16.22) is the solution of IVP (16.21). The next IVP (16.23) solves an augmented one-dimensional problem in the  $z$ -direction, the  $z$  *sweep*, for a time step  $\Delta t$ . The initial condition for the third IVP (16.23) is the solution of IVP (16.22). It is important to realise that the size of the time step  $\Delta t$  in each one-dimensional IVP is constant during one time step.  $\mathbf{U}^{n+1}$  is regarded as the solution of IVP (16.20) after a time  $\Delta t$ .

If  $\mathcal{X}^{(t)}$ ,  $\mathcal{Y}^{(t)}$  and  $\mathcal{Z}^{(t)}$  are approximate solution operators for IVPs (16.21), (16.22) and (16.23) respectively, then the simplest, first-order accurate, splitting based on (16.21)–(16.23) can be written thus

$$\mathbf{U}^{n+1} = \mathcal{Z}^{(\Delta t)} \mathcal{Y}^{(\Delta t)} \mathcal{X}^{(\Delta t)} (\mathbf{U}^n). \quad (16.24)$$

Naturally, other orderings of the operators are possible to produce other first-order splitting schemes. A three-dimensional splitting, see Shang [442], that is second-order accurate every other time step is

$$\mathbf{U}^{n+2} = \mathcal{X}^{(\Delta t)} \mathcal{Y}^{(\Delta t)} \mathcal{Z}^{(\Delta t)} \mathcal{Z}^{(\Delta t)} \mathcal{Y}^{(\Delta t)} \mathcal{X}^{(\Delta t)} (\mathbf{U}^n). \quad (16.25)$$

A three-dimensional splitting scheme that is second-order accurate every time step [61] is

$$\mathbf{U}^{n+1} = \mathcal{X}^{(\frac{1}{2}\Delta t)} \mathcal{Y}^{(\frac{1}{2}\Delta t)} \mathcal{Z}^{(\Delta t)} \mathcal{Y}^{(\frac{1}{2}\Delta t)} \mathcal{X}^{(\frac{1}{2}\Delta t)} (\mathbf{U}^n). \quad (16.26)$$

Other combinations of one-dimensional operators will produce other second order accurate splitting schemes. The tools for constructing splitting schemes and analysing their accuracy, see Strang [466] and Shang [442], rely on Taylor series expansions and therefore the derived splitting schemes are strictly valid only for problems with smooth solutions. For discontinuous solutions splitting schemes may produce erroneous results. Ironically, the failure of dimensional splitting procedures is closely related to how successful the one-dimensional schemes used are for computing discontinuous solutions. For instance, the Random Choice Method (RCM), see Chap. 7, has the unique property of computing contact surfaces and shock waves as true discontinuities. However, dimensional splitting schemes based on the RCM have been shown to be a complete failure, see Colella [132]. An exception to these difficulties are contact surfaces; these waves can be correctly preserved by splitting schemes based on the RCM, see Toro [500].

In general, schemes that smear discontinuities in one space dimension work well in dimensional splitting procedures. We particularly recommend the use

of the one-dimensional schemes presented in Chaps. 6 to 14, except for the Random Choice method of Chap. 7.

In the next section we illustrate via an example, some of the practical details involved when implementing dimensional splitting schemes to solve three-dimensional time dependent problems.

## 16.3 Practical Implementation of Splitting Schemes in Three Dimensions

As an example, we consider the time-dependent, three dimensional Euler equations

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_y + \mathbf{H}(\mathbf{U})_z = \mathbf{0}, \quad (16.27)$$

with

$$\left. \begin{aligned} \mathbf{U} &= \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ E \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ \rho uw \\ u(E + p) \end{bmatrix}, \\ \mathbf{G} &= \begin{bmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ \rho vw \\ v(E + p) \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} \rho w \\ \rho uw \\ \rho vw \\ \rho w^2 + p \\ w(E + p) \end{bmatrix}. \end{aligned} \right\} \quad (16.28)$$

See Chaps. 1 and 3 for details on the Euler equations.

### 16.3.1 Handling the Sweeps by a Single Subroutine

For simplicity let us assume we apply the first-order splitting scheme (16.24). For convenience we re-order the equations so that every sweep is handled by a *single one-dimensional subroutine*. In the  $x$  sweep we solve

$$\begin{bmatrix} \rho \\ \rho u \\ E \\ \rho v \\ \rho w \end{bmatrix}_t + \begin{bmatrix} \rho u \\ \rho u^2 + p \\ u(E + p) \\ \rho uv \\ \rho uw \end{bmatrix}_x = \mathbf{0}. \quad (16.29)$$

In the the  $y$  sweep one solves

$$\begin{bmatrix} \rho \\ \rho v \\ E \\ \rho u \\ \rho w \end{bmatrix}_t + \begin{bmatrix} \rho v \\ \rho v^2 + p \\ v(E + p) \\ \rho vu \\ \rho vw \end{bmatrix}_y = \mathbf{0}. \quad (16.30)$$

In the  $z$  sweep one solves

$$\begin{bmatrix} \rho \\ \rho w \\ E \\ \rho u \\ \rho v \end{bmatrix}_t + \begin{bmatrix} \rho w \\ \rho w^2 + p \\ w(E + p) \\ \rho w u \\ \rho w v \end{bmatrix}_z = \mathbf{0}. \quad (16.31)$$

In the  $x$  sweep (16.29)  $u$  is the *normal* velocity components and the first three equations look identical to the pure one-dimensional problem in  $x$ . Note however that the full problem (16.29) differs from the pure one-dimensional problem in two respects, namely (i) there are two extra equations for momentum in the  $y$  and  $z$  directions and (ii) the total energy  $E$  involves contributions from the tangential velocity components  $v$  and  $w$  via the kinetic energy.

As seen in Sect. 3.1.2 of Chap. 3, use of the continuity equation into the expanded form of the  $y$  and  $z$  momentum equations gives the two advection equations

$$v_t + uv_x = 0, \quad w_t + uw_x = 0. \quad (16.32)$$

These say that the *tangential velocity components*  $v$  and  $w$  are passively advected with the normal velocity component  $u$ . As a matter of fact, the extra equations for momentum in the  $y$  and  $z$  directions have exactly the same form as species equations present in the study of reactive mixtures.

For the  $y$  sweep (16.30) the normal velocity component is  $v$  and the tangential velocity components are  $u$  and  $w$ . For the  $z$  sweep (16.31) the normal velocity component is  $w$  and the tangential velocity components are  $u$  and  $v$ .

It is now obvious that a single subroutine can be used to deal with all three sweeps in a dimensional splitting scheme. Suppose we are solving the  $x$  sweep. Then an explicit conservative method, see Chaps. 5 to 14, reads

$$\mathbf{U}_{i,j,k}^{n+\frac{1}{3}} = \mathbf{U}_{i,j,k}^n + \frac{\Delta t}{\Delta x} [\mathbf{F}_{i-\frac{1}{2},j,k}^n - \mathbf{F}_{i+\frac{1}{2},j,k}^n], \quad \forall j, \forall k, \quad (16.33)$$

where  $\mathbf{F}_{i+\frac{1}{2},j,k}^n$  is the numerical flux at the cell interface position  $x_{i+\frac{1}{2}}$ . For background on the meaning of the conservative formula (16.33) the reader is referred to Chap. 5, Sect. 5.3 and Chap. 6. For the  $y$  sweep (16.30) the updating conservative formula reads

$$\mathbf{U}_{i,j,k}^{n+\frac{2}{3}} = \mathbf{U}_{i,j,k}^{n+\frac{1}{3}} + \frac{\Delta t}{\Delta y} [\mathbf{G}_{i,j-\frac{1}{2},k}^{n+\frac{1}{3}} - \mathbf{G}_{i,j+\frac{1}{2},k}^{n+\frac{1}{3}}], \quad \forall i, \forall k, \quad (16.34)$$

and for the  $z$  sweep (16.31) the updating conservative formula reads

$$\mathbf{U}_{i,j,k}^{n+1} = \mathbf{U}_{i,j,k}^{n+\frac{2}{3}} + \frac{\Delta t}{\Delta z} [\mathbf{H}_{i,j,k-\frac{1}{2}}^{n+\frac{2}{3}} - \mathbf{H}_{i,j,k+\frac{1}{2}}^{n+\frac{2}{3}}], \quad \forall i, \forall j. \quad (16.35)$$

Obviously, given the way the equations have been re-ordered, all three sweeps (16.33)–(16.35) can be handled by a single subroutine. Let us denote

the subroutine by ONED(...). Then, sweeps (16.33)–(16.35) are dealt with by doing

$$\left. \begin{array}{ll} \text{for the } x \text{ sweep} & \text{CALL ONED}(\rho, u, v, w, p, \Delta t, \Delta x, \dots), \\ \text{for the } y \text{ sweep} & \text{CALL ONED}(\rho, v, u, w, p, \Delta t, \Delta y, \dots), \\ \text{for the } z \text{ sweep} & \text{CALL ONED}(\rho, w, u, v, p, \Delta t, \Delta z, \dots). \end{array} \right\} \quad (16.36)$$

Note that we simply interchange the order of the velocity components in the single subroutine to solve for all the  $x$ ,  $y$  and  $z$  sweeps. To implement some of the other splitting schemes such as (16.25) or (16.26), the order in (16.36), as well as the size of the time step, are appropriately changed.

### 16.3.2 Choice of Time Step Size

The explicit conservative schemes considered here require the computation of a time step  $\Delta t$  to be used in (16.33)–(16.35), such that *stability* of the numerical method is ensured. The choice of  $\Delta t$  depends on (i) the intercell flux used, i.e. the particular method used, (ii) the wave speeds present and (iii) the dimensions of the mesh used. See Sect. 6.3.2 for a full discussion of this topic for one-dimensional problems. To preserve some generality let us assume that the dimensions of cell  $I_{i,j,k}$  are  $\Delta x_{i,j,k}$ ,  $\Delta y_{i,j,k}$  and  $\Delta z_{i,j,k}$ . Then, one way of choosing  $\Delta t$  is

$$\Delta t = C_{cfl} \times \min_{i,j,k} \left[ \frac{\Delta x_{i,j,k}}{S_{i,j,k}^{n,x}}, \frac{\Delta y_{i,j,k}}{S_{i,j,k}^{n,y}}, \frac{\Delta z_{i,j,k}}{S_{i,j,k}^{n,z}} \right]. \quad (16.37)$$

Here  $S_{i,j,k}^{n,d}$  is the speed of the fastest wave present at time level  $n$  travelling in the  $d$  direction, with  $d = x, y, z$ .  $C_{cfl}$  is the CFL coefficient and is chosen according to the linear stability condition of the particular numerical method in use, which depends on the numerical flux. For the Godunov first-order upwind method  $0 < C_{cfl} \leq 1$  and in practice one usually takes  $C_{cfl} \approx 0.9$ . Perhaps the most critical point is finding reliable estimates for the speeds  $S_{i,j,k}^{n,d}$ . The simplest choice is

$$S_{i,j,k}^{n,d} = |d_{i,j,k}^n| + a_{i,j,k}^n \quad (16.38)$$

where  $d_{i,j,k}^n$  is the particle velocity component in the  $d$ -direction and  $a_{i,j,k}^n$  is the sound speed at time level  $n$  in cell  $I_{i,j,k}$ . As remarked in Sect. 6.3.2 of Chap. 6 for one-dimensional problems, choice (16.38) is inadequate for problems with initial conditions involving severe gradients. One may underestimate the wave speeds, for instance, when the initial conditions are of the shock-tube problem type. In such cases  $d_{i,j,k}^n$  might be zero and the only contribution to  $S_{i,j,k}^{n,d}$  is the sound speed  $a_{i,j,k}^n$ , which then results in a gross underestimate for  $S_{i,j,k}^{n,d}$ . As a consequence a gross overestimate for  $\Delta t$  is computed which leads the method outside its stability range at the beginning of the computations.

### 16.3.3 The Intercell Flux and the TVD Condition

In discussing the choice of the intercell fluxes in (16.33)–(16.35) we shall consider the case of a single sweep, the  $x$  sweep say. The intercell flux  $\mathbf{F}_{i+\frac{1}{2},j,k}^n$  in (16.33) may be computed by any of the methods discussed in Chaps. 6 to 12 or the high-order TVD schemes of Chaps. 13 and 14.

The computation of  $\mathbf{F}_{i+\frac{1}{2},j,k}^n$  by the Godunov first-order upwind method of Chap. 6 is straightforward. What is new in multidimensional problems is the presence of tangential velocity components. Locally, these are related to shear waves. The particular details of the flux can have a profound effect in the accuracy with which shear waves and vortices are resolved by the method. The Godunov method with the exact Riemann solver of Chap. 4, see Sect. 4.8, gives accurate resolution of shear waves. The approximate Riemann solvers of Chaps. 9, 11, 12 and the HLLC Riemann solver of Chap. 10 will also resolve shear waves accurately. The Flux Vector Splitting schemes of Chap. 8, except for the version of Sect. 8.4.4, and the HLL Riemann solver of Sect. 10.3 of Chap. 10 will smear contact surfaces and shear waves. The FORCE flux of Chap. 7, see also Chap. 14, will smear contacts and shear waves too.

High-order TVD extensions of the previous methods will generally improve the overall accuracy, but if shear waves are unduly smeared by the underlying first-order scheme then their resolution by the high-order extensions will also be unsatisfactory. The implementation of the MUSCL–Hancock scheme to compute  $\mathbf{F}_{i+\frac{1}{2},j,k}^n$  is a straightforward application of the one-dimensional approach, see Sect. 14.4 of Chap. 14. The same is true for of the centred schemes FLIC and SLIC presented of Sects. 14.5.2 and 14.5.3 of Chap. 14.

The *limiting* process in enforcing the TVD constraints will now involve the tangential velocity components. For MUSCL-type schemes this is carried out automatically for each equation. The details of the particular Riemann solver used will become important in attempting to resolve contacts and shear waves. Moreover, when applying characteristic limiting, see Sect. 14.4.3, the particular characteristic decomposition used must properly account for the presence of contact and shear waves. For flux-limiter type schemes, such as the WAF and FLIC methods the limiting procedures are different.

The Weighted Average Flux (WAF) method of Sect. 14.3 has intercell flux

$$\mathbf{F}_{i+\frac{1}{2},j,k} = \frac{1}{2}(\mathbf{F}_{i,j,k}^n + \mathbf{F}_{i+1,j,k}^n) - \frac{1}{2} \sum_{k=1}^5 \text{sign}(c_k) \phi_{i+\frac{1}{2}}^{(k)} \Delta \mathbf{F}_{i+\frac{1}{2}}^{(k)}, \quad (16.39)$$

where

$$\phi_{i+\frac{1}{2}}^{(k)} = \phi_{i+\frac{1}{2}}^{(k)}(r^{(k)}) \quad (16.40)$$

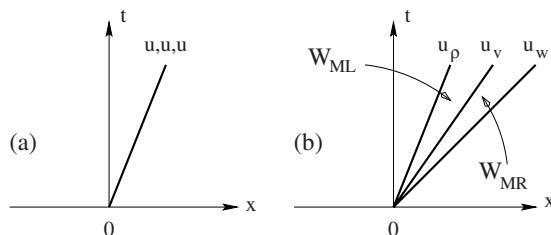
is a limiter function for wave  $k$  and  $r^{(k)} = r^{(k)}(q)$  is the *flow parameter*, computed from ratios of wave jumps of a quantity  $q$ , see Sect. 14.3.4. Recall that in the eigenstructure of  $x$ -split three-dimensional Euler equations one has one eigenvalue of multiplicity three, that is  $\lambda_2 = \lambda_4 = \lambda_5 = u$ . Note that we

have re-ordered the eigenvalues, consistently with the order of the equations in (16.29). These are associated with three *linear* waves:  $\lambda_2$  is associated with a contact wave of speed  $u$ ,  $\lambda_4$  is associated with a  $v$ -shear wave of speed  $u$  and  $\lambda_5$  is associated with a  $w$ -shear wave of speed  $u$ ;  $\lambda_1 = u - a$  and  $\lambda_3 = u + a$  correspond to the left and right non-linear waves in the solution of the Riemann problem. See Sect. 4.8 for details. For the *limiting* process one needs to compute five limiter functions, one for each wave; they are

$$\left. \begin{array}{l} \phi_{i+\frac{1}{2}}^{(1)} \text{ is the limiter function for the left non-linear wave ,} \\ \phi_{i+\frac{1}{2}}^{(2)} \text{ is the limiter function for the contact wave ,} \\ \phi_{i+\frac{1}{2}}^{(3)} \text{ is the limiter function for the right non-linear wave ,} \\ \phi_{i+\frac{1}{2}}^{(4)} \text{ is the limiter function for the } v\text{-shear wave ,} \\ \phi_{i+\frac{1}{2}}^{(5)} \text{ is the limiter function for the } w\text{-shear wave .} \end{array} \right\} \quad (16.41)$$

These limiters are then applied to each flux component in (16.39). Recall that the limiter function depends on a flow parameter  $r$ , which is computed from a selected quantity  $q$ . For the first three limiters one takes  $q = \rho$ , or the specific internal energy  $e$ ; for limiter 4 one takes  $q = v$  and for limiter 5 one takes  $q = w$ .

The application of the limiter functions in (16.39) may be viewed as applying the *un-limited* WAF flux to a *dissipated wave structure*, where the original wave paths are displaced. Fig. 16.2 (a) shows the *un-limited* wave structure



**Fig. 16.2.** Wave structure for the linear waves (contact and shear waves). (a) un-limited structure before applying the TVD condition, (b) limited structure after applying the TVD condition

for the contact and shear waves; Fig. 16.2. (b) illustrates one possible wave pattern in the dissipated wave structure, where  $u_p$ ,  $u_v$  and  $u_w$  denote the *new* speeds of the contact,  $v$ -shear and  $w$ -shear waves respectively. Then the wave pattern of Fig. 14.2. (b) corresponds to

$$\phi_{i+\frac{1}{2}}^{(2)} \leq \phi_{i+\frac{1}{2}}^{(4)} \leq \phi_{i+\frac{1}{2}}^{(5)}$$

and we call this the 2–4–5 order. The dissipated wave structure results in *two new intermediate states*  $\mathbf{W}_{ML}$  and  $\mathbf{W}_{MR}$ , with corresponding fluxes  $\mathbf{F}_{ML}$

and  $\mathbf{F}_{MR}$ ; here  $\mathbf{W} = (\rho, u, v, w, p)^T$  denotes the vector of primitive variables. For the two-dimensional Euler equations there are only two possible orders; for the three-dimensional case there are six possible orders. Table 16.1 lists all six possible orders for  $\rho$ ,  $v$  and  $w$ . The values for  $p$  and  $u$  are simply  $p_*$  and  $u_*$  respectively.

Order	Middle left			Middle right		
	$\rho$	$v$	$w$	$\rho$	$v$	$w$
2–4–5	$\rho_{*R}$	$v_L$	$w_L$	$\rho_{*R}$	$v_R$	$w_L$
2–5–4	$\rho_{*R}$	$v_L$	$w_L$	$\rho_{*R}$	$v_R$	$w_R$
4–2–5	$\rho_{*L}$	$v_R$	$w_L$	$\rho_{*R}$	$v_R$	$w_L$
4–5–2	$\rho_{*L}$	$v_R$	$w_L$	$\rho_{*L}$	$v_R$	$w_R$
5–2–4	$\rho_{*L}$	$v_L$	$w_R$	$\rho_{*R}$	$v_L$	$w_R$
5–4–2	$\rho_{*L}$	$v_L$	$w_R$	$\rho_{*L}$	$v_R$	$w_R$

**Table 16.1.** Six possible cases in dissipated wave structure occurring in limiting the WAF flux for the three dimensional Euler equations

A simple procedure is to take the maximum dissipation as the single dissipation to be applied to all three linear waves, that is

$$\phi_{max}^{(2)} \equiv \phi_{i+\frac{1}{2}}^{(2)} = \phi_{i+\frac{1}{2}}^{(4)} = \phi_{i+\frac{1}{2}}^{(5)} = \max[\phi_{i+\frac{1}{2}}^{(2)}, \phi_{i+\frac{1}{2}}^{(4)}, \phi_{i+\frac{1}{2}}^{(5)}]. \quad (16.42)$$

This will obviously result in excessive dissipation for some of the waves and is not recommended.

Another approach is this. First apply the limiting procedure for waves 1, 2 and 3 as for the one-dimensional case, the first three equations in (16.29). Waves 4 and 5 are treated separately by regarding the equations for the  $y$  and  $z$  momentum components as scalar equations of the form

$$(\rho\eta)_t + (f_{i+\frac{1}{2}}^1\eta)_x = 0, \quad (16.43)$$

where  $\eta = v, w$  and  $f_{i+\frac{1}{2}}^1 = \rho u$  is the TVD flux for the mass equation. Then, the TVD flux component for the  $\eta$  momentum equation is

$$(\rho\eta)^{tvd} = [\frac{1}{2}(1 + \phi^n)\eta_{i,j,k}^n + \frac{1}{2}(1 - \phi^n)\eta_{i+1,j,k}^n] \times f_{i+\frac{1}{2}}^1, \quad (16.44)$$

where  $\phi^n$  is a limiter function based on jumps in  $\eta$  across the  $\eta$ -shear wave.

*Remark 16.2.* The attraction of the approach just described is its simplicity as well as its ability to treat, in an entirely similar manner, any passive scalar  $\eta$ . More importantly, one can apply the same procedure for any number of passive scalars, as, for example, in the study of reactive multicomponent flows.

*Remark 16.3.* The equation for total energy per unit mass  $E$ , unlike the pure one dimensional case, contains the tangential velocity components and therefore the limiting for the first three equations in (16.29) based on  $\rho$  alone is somewhat unsatisfactory.

In the next section we study a methodology to solve multidimensional problems that is distinct from the splitting approach discussed so far in this chapter, namely unsplit finite volume methods. Some of the previous remarks on intercell fluxes and applying the TVD condition will carry over to finite volume methods.

## 16.4 Unsplit Finite Volume Methods

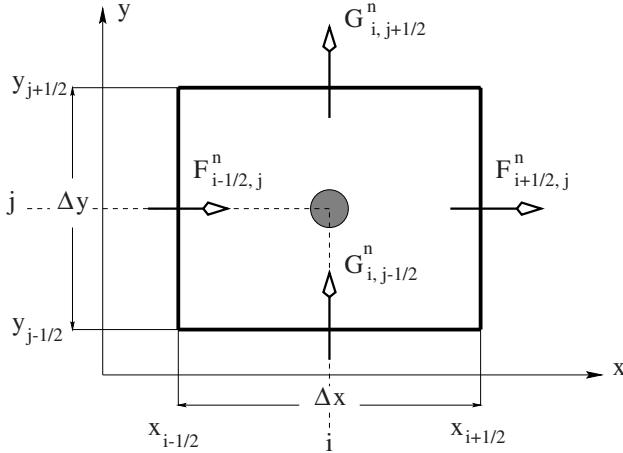
In this section we present the *unsplit*, finite volume approach for solving multidimensional problems. This is an alternative to dimensional splitting, presented in Sect. 16.2. When finite volume one-dimensional schemes are applied in conjunction with dimensional splitting, there exists a close relationship between both approaches. Both approaches are based on enforcing the integral form of the conservation laws on discrete or finite volumes, as explained in Sect. 16.5 for general non-Cartesian geometries. The difference stems from the fact that in the so-called unsplit finite volume method the solution is advanced by accounting for *all flux contributions in a single step*. In what follows, a finite volume scheme will be understood as an unsplit finite volume scheme. We present two approaches for constructing second-order upwind finite volumes schemes in two and three space dimensions, namely the MUSCL–Hancock and the WAF approaches. These were introduced in Chap. 14 for constructing one-dimensional schemes; in Sects. 16.2 and 16.3 of this chapter these schemes were utilised for solving multidimensional problems via dimensional splitting. There are other approaches for constructing upwind finite volume schemes for multidimensional hyperbolic conservations laws. The reader is referred to the works of Colella [134], LeVeque [309] and Casper and Atkins [88]. The review paper by Vinokur [570] is highly recommended.

### 16.4.1 Introductory Concepts

Consider a time-dependent two dimensional system of conservation laws

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_y = \mathbf{0}. \quad (16.45)$$

For simplicity assume that the boundaries of the computational domain are aligned with the coordinate directions  $x$  and  $y$ . Consider a typical finite volume or computational cell  $I_{i,j}$  of dimensions  $\Delta x \times \Delta y$ , as depicted in Fig. 16.3. The cell average  $\mathbf{U}_{i,j}^n$  is assigned to the *centre of the cell*, which gives rise to *cell centred methods*. To each intercell boundary there corresponds a numerical flux.



**Fig. 16.3.** Finite Volume discretisation of Cartesian domain. Typical computing cell  $I_{i,j}$  has four intercell boundaries with corresponding intercell fluxes

An explicit finite volume scheme to solve (16.45) reads

$$\mathbf{U}_{i,j}^{n+1} = \mathbf{U}_{i,j}^n + \frac{\Delta t}{\Delta x} [\mathbf{F}_{i-\frac{1}{2},j} - \mathbf{F}_{i+\frac{1}{2},j}] + \frac{\Delta t}{\Delta y} [\mathbf{G}_{i,j-\frac{1}{2}} - \mathbf{G}_{i,j+\frac{1}{2}}]. \quad (16.46)$$

The cell average  $\mathbf{U}_{i,j}^n$  in cell  $I_{i,j}$  at time level  $n$  is updated to time level  $n+1$  via (16.46) in a *single step*, involving flux contributions *from all intercell boundaries*. This conservative formula is the natural extension of one-dimensional conservative formulae, such as (16.33)–(16.35) for example, and is completely determined once the intercell numerical fluxes are specified and a choice of mesh has been made.

The simplest, upwind finite volume scheme results from applying the Godunov first-order upwind fluxes across each intercell boundary, in exactly the same way as done for one-dimensional problems, that is

$$\mathbf{F}_{i+\frac{1}{2},j} = \mathbf{F}(\mathbf{U}_{i+\frac{1}{2},j}(0)); \quad \mathbf{G}_{i,j+\frac{1}{2}} = \mathbf{G}(\mathbf{U}_{i,j+\frac{1}{2}}(0)), \quad (16.47)$$

where  $\mathbf{U}_{i+\frac{1}{2},j}(x/t)$  is the solution of the Riemann problem

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0}, \\ \mathbf{U}(x, 0) = \begin{cases} \mathbf{U}_{i,j}^n & \text{if } x < 0, \\ \mathbf{U}_{i+1,j}^n & \text{if } x > 0 \end{cases} \quad (16.48)$$

and  $\mathbf{U}_{i,j+\frac{1}{2}}(y/t)$  is the solution of the Riemann problem

$$\mathbf{U}_t + \mathbf{G}(\mathbf{U})_y = \mathbf{0}, \\ \mathbf{U}(y, 0) = \begin{cases} \mathbf{U}_{i,j}^n & \text{if } y < 0, \\ \mathbf{U}_{i,j+1}^n & \text{if } y > 0. \end{cases} \quad (16.49)$$

For details on one-dimensional fluxes see Chaps. 5 to 14.

*Example 16.4 (Godunov Finite Volume Method).* Consider the linear advection equation

$$u_t + f(u)_x + g(u)_y = 0 ; \quad f(u) = a_1 u , \quad g(u) = a_2 u , \quad (16.50)$$

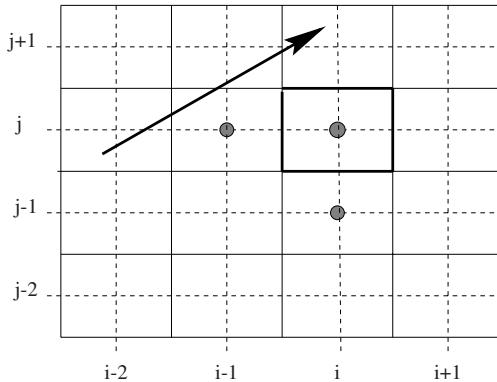
where  $a_1$  and  $a_2$  are constant velocity components in the  $x$  and  $y$  directions respectively. By assuming that  $a_1 > 0$ ,  $a_2 > 0$ , it is easy to verify that the Godunov finite volume scheme (16.46)–(16.47) as applied to (16.50) reads

$$u_{i,j}^{n+1} = u_{i,j}^n + c_1(u_{i-1,j}^n - u_{i,j}^n) + c_2(u_{i,j-1}^n - u_{i,j}^n) , \quad (16.51)$$

where

$$c_1 = \frac{a_1 \Delta t}{\Delta x} ; \quad c_2 = \frac{a_2 \Delta t}{\Delta y} \quad (16.52)$$

are the Courant numbers in the  $x$  and  $y$  directions respectively. The stencil of the scheme is shown in Fig. 16.4. The centre of the stencil is the point  $(i, j)$ . In addition to the cell average  $u_{i,j}^n$  at the centre of the stencil there are only two more data values contributing to the updating. These are  $u_{i-1,j}^n$ , which is the upwind value in the  $x$ -direction, and  $u_{i,j-1}^n$ , which is the upwind value in the  $y$ -direction. Note the scheme reduces identically to the Godunov first-order upwind method, see Sect. 5.3.2 of Chap. 5, if  $a_1 = 0$  or  $a_2 = 0$ . In two dimensions the *upwinding* is not complete; for instance the most obvious upwind value for  $a_1 > 0$  and  $a_2 > 0$  in a two-dimensional sense is  $u_{i-1,j-1}^n$ , and this does not contribute to the updating of  $u_{i,j}^n$  via (16.51).



**Fig. 16.4.** Stencil of the Godunov finite volume scheme for the linear advection equation with positive velocity components  $a_1$  and  $a_2$ . The arrow represents the direction of the velocity vector

For three-dimensional time dependent systems of conservation laws

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_y + \mathbf{H}(\mathbf{U})_z = \mathbf{0} , \quad (16.53)$$

finite volume schemes read

$$\mathbf{U}_{i,j,k}^{n+1} = \mathbf{U}_{i,j,k}^n + \frac{\Delta t}{\Delta x} [\mathbf{F}_{i-\frac{1}{2},j,k} - \mathbf{F}_{i+\frac{1}{2},j,k}] + \frac{\Delta t}{\Delta y} [\mathbf{G}_{i,j-\frac{1}{2},k} - \mathbf{G}_{i,j+\frac{1}{2},k}] + \frac{\Delta t}{\Delta z} [\mathbf{H}_{i,j,k-\frac{1}{2}} - \mathbf{H}_{i,j,k+\frac{1}{2}}]. \quad (16.54)$$

Consider the three-dimensional linear advection equation

$$u_t + f(u)_x + g(u)_y + h(u)_z = 0; \quad f(u) = a_1 u, \quad g(u) = a_2 u, \quad h(u) = a_3 u, \quad (16.55)$$

where  $a_1$ ,  $a_2$  and  $a_3$  are constant velocity components in the  $x$ ,  $y$  and  $z$  directions respectively.

**Exercise 16.5.** Assume that  $a_1 > 0$ ,  $a_2 > 0$ ,  $a_3 > 0$  and show that the three-dimensional finite volume scheme obtained by straight application of the one-dimensional Godunov flux is

$$u_{i,j,k}^{n+1} = u_{i,j,k}^n + c_1(u_{i-1,j,k}^n - u_{i,j,k}^n) + c_2(u_{i,j-1,k}^n - u_{i,j,k}^n) + c_3(u_{i,j,k-1}^n - u_{i,j,k}^n), \quad (16.56)$$

where

$$c_1 = \frac{a_1 \Delta t}{\Delta x}; \quad c_2 = \frac{a_2 \Delta t}{\Delta y}; \quad c_3 = \frac{a_3 \Delta t}{\Delta z} \quad (16.57)$$

are the Courant numbers in the  $x$ ,  $y$  and  $z$  directions respectively.

### 16.4.2 Accuracy and Stability of Multidimensional Schemes

#### Accuracy Theorems

A useful theorem due to Roe [408] concerning the accuracy of general constant coefficient schemes for the one-dimensional linear advection equation was stated in Sect. 13.2.1 of Chap. 13. We now state the generalisation of this result to two and three dimensional schemes [62].

Two-dimensional schemes such as (16.51) can be written in the general form

$$u_{i,j}^{n+1} = \sum_{\alpha,\beta} b_{\alpha,\beta} u_{i+\alpha,j+\beta}^n, \quad (16.58)$$

where  $b_{\alpha,\beta}$  are the, constant, coefficients of the scheme. Schemes of the form (16.58) are a straight generalisation of the one-dimensional schemes studied in Sect. 13.2.1 of Chap. 13. We now state the following result.

**Theorem 16.6.** *The two-dimensional scheme (16.58) to solve (16.50) is  $m$ th-order accurate in space and time if and only if the coefficients satisfy*

$$\sum_{\alpha,\beta} \alpha^q \beta^r b_{\alpha,\beta} = (-c_1)^q (-c_2)^r \quad (16.59)$$

for all integer pairs  $(q, r)$  such that  $q \geq 0$ ,  $r \geq 0$  and  $q + r \leq m$ . Here  $c_1$  and  $c_2$  are the directional Courant numbers (16.52).

*Proof.* For details of the proof see Billett and Toro [62].

Three-dimensional schemes, such as (16.56), to solve the three-dimensional linear advection equation (16.55) can be written as

$$u_{i,j,k}^{n+1} = \sum_{\alpha,\beta,\gamma} b_{\alpha,\beta,\gamma} u_{i+\alpha,j+\beta,k+\gamma}^n, \quad (16.60)$$

where  $b_{\alpha,\beta,\gamma}$  are the coefficients of the scheme, assumed to be constant. The following result is a generalisation of the previous accuracy theorem to three-dimensional schemes.

**Theorem 16.7.** *The three-dimensional scheme (16.60) to solve (16.55) is  $m$ th-order accurate in space and time if and only if the coefficients  $b_{\alpha,\beta,\gamma}$  satisfy*

$$\sum_{\alpha,\beta,\gamma} \alpha^q \beta^r \gamma^s b_{\alpha,\beta,\gamma} = (-c_1)^q (-c_2)^r (-c_3)^s \quad (16.61)$$

for all integer triples  $(q, r, s)$  such that  $q \geq 0, r \geq 0, s \geq 0$  and  $q + r + s \leq m$ . Here  $c_1, c_2$  and  $c_3$  are the directional Courant numbers (16.57).

*Proof.* For details of the proof see Billett and Toro [62].

**Exercise 16.8.** Apply the two previous accuracy theorems (16.6) and (16.7) to show that the Godunov finite volume schemes (16.51) and (16.56) to solve the linear advection equations (16.50) and (16.55) are first-order accurate in space and time.

**Solution 16.9.** (Left to the reader).

### Computational Test for Linear Stability

There are various methods for analysing the stability of schemes for the linear advection equation; see Hirsch [251], Chaps. 7 to 10, for a comprehensive presentation of techniques. Here we assume the reader to be familiar with the popular von Neumann method.

For a one-dimensional scheme whose coefficients depend only on the Courant number  $c_1$  the von Neumann method derives an algebraic expression

$$A = A(c_1, \theta_1)$$

for the *amplification factor*, where  $c_1$  is the Courant number and  $\theta_1$  is the *phase angle* in the range  $0 \leq \theta_1 \leq \pi$ . For sufficiently simple schemes the stability condition  $\|A\| \leq 1$  can be used to derive, algebraically, conditions on the Courant number for the scheme to be linearly stable. There are instances, however, where the algebraic task becomes intractable. In this case, drawing a contour plot of  $\|A\|$ , the modulus of the complex number  $A$ , for a *large number* of values  $(c_1, \theta_1)$  in a rectangle  $[0, c_1^{max}] \times [0, \pi]$ , will give an *indication* as to

the linear stability region of the scheme. Here  $c_1^{max}$  is chosen to be larger than the *expected* stability limit.

For two-dimensional schemes the amplification factor reads

$$A = A(c_1, c_2, \theta_1, \theta_2) .$$

Here  $c_1, c_2$  are the Courant numbers in the  $x$  and  $y$  directions and  $\theta_1, \theta_2$  are phase angles in the  $x$  and  $y$  directions. Deriving algebraic conditions for the scheme to be stable is now generally much harder than in the one-dimensional case. A *numerical test* can be performed as follows [62]. For a given pair  $c_1, c_2$  of Courant numbers in the rectangle  $[0, c_1^{max}] \times [0, c_2^{max}]$  compute  $\|A\|$  for a *large number*  $M_{ang}$  of values  $(\theta_1, \theta_2)$  of phase angles in the rectangle  $[0, \pi] \times [0, \pi]$ . Record the number  $M_{cfl}$  for which  $\|A\| \leq 1$  and define the ratio

$$S = S(c_1, c_2) = \frac{M_{cfl}}{M_{ang}} .$$

If  $S = 1$  the scheme is regarded as stable for the pair  $c_1, c_2$ . A contour plot of  $S = S(c_1, c_2)$  for a large number of pairs  $(c_1, c_2)$  in the range  $[0, c_1^{max}] \times [0, c_2^{max}]$  will give an *indication* of the linear stability of the two-dimensional scheme.

In three space dimensions the amplification factor reads

$$A = A(c_1, c_2, c_3, \theta_1, \theta_2, \theta_3)$$

and one draws contour plots on planes within  $[0, c_1^{max}] \times [0, c_2^{max}] \times [0, c_3^{max}]$ .

*Remark 16.10.* The above stability tests, useful in practice as they may be, do not constitute proofs of linear stability, and caution is needed when quoting the conclusions of the tests.

*Example 16.11.* Application of the stability test just described to the Godunov finite volume scheme (16.51) *indicates* that the scheme has linear stability condition

$$c_1 + c_2 \leq 1 . \quad (16.62)$$

For  $c_1 = c_2$  one has  $c_1 = c_2 \leq \frac{1}{2}$ , which is half the stability range of the one-dimensional Godunov scheme. For the the three-dimensional Godunov finite volume scheme (16.56), application of the above test suggests that the scheme is linearly stable under the condition

$$c_1 + c_2 + c_3 \leq 1 . \quad (16.63)$$

For  $c_1 = c_2 = c_3$  one has  $c_1 = c_2 = c_3 \leq \frac{1}{3}$ , which is one third of the stability limit of the one-dimensional, Godunov first-order upwind scheme.

The previous example illustrates the fact that in constructing finite volume schemes by straightforward application of one-dimensional fluxes one,

at best, ends up with schemes with a *reduced stability range*. There are cases in which this approach leads to unconditionally unstable finite volume methods, even though the one-dimensional operators are stable for the respective one-dimensional method.

In the next section we construct a finite volume method that is second-order accurate in space and time.

## 16.5 A MUSCL–Hancock Finite Volume Scheme

Here we extend the second-order MUSCL–Hancock approach, presented in Chap. 14 for one-dimensional systems, to construct unsplit finite volume schemes for multi-dimensional conservation laws. We present the details of the approach for the two-dimensional case and construct a scheme of the form (16.46). Background on the MUSCL–Hancock approach for scalar one-dimensional problems is found in Sect. 13.4.2 of Chap. 13; Sect. 14.4 of Chap. 14 gives details of the approach for one-dimensional non-linear systems. The reader is encouraged to review the details of the one dimensional MUSCL–Hancock scheme before proceeding with this section. As in the one-dimensional case the scheme has the following three steps

- (I) **Data Reconstruction and Boundary Extrapolated Values.** The cell averages  $\mathbf{U}_{i,j}^n$  are reconstructed, independently, in the  $x$  and  $y$  directions by selecting respective slope vectors (differences)  $\Delta_i$  and  $\Delta_j$ . Boundary extrapolated values are

$$\left. \begin{aligned} \mathbf{U}_{i,j}^{-x} &= \mathbf{U}_{i,j}^n - \frac{1}{2}\Delta_i ; & \mathbf{U}_{i,j}^{+x} &= \mathbf{U}_{i,j}^n + \frac{1}{2}\Delta_i ; \\ \mathbf{U}_{i,j}^{-y} &= \mathbf{U}_{i,j}^n - \frac{1}{2}\Delta_j ; & \mathbf{U}_{i,j}^{+y} &= \mathbf{U}_{i,j}^n + \frac{1}{2}\Delta_j . \end{aligned} \right\} \quad (16.64)$$

- (II) **Evolution of Boundary Extrapolated Values.**

$$\hat{\mathbf{U}}_{i,j}^l = \mathbf{U}_{i,j}^l + \frac{\Delta t}{2\Delta x} [\mathbf{F}(\mathbf{U}_{i,j}^{-x}) - \mathbf{F}(\mathbf{U}_{i,j}^{+x})] + \frac{\Delta t}{2\Delta y} [\mathbf{G}(\mathbf{U}_{i,j}^{-y}) - \mathbf{G}(\mathbf{U}_{i,j}^{+y})] \quad (16.65)$$

for  $l = -x, +x, -y, +y$ .

- (III) **Solution of Riemann problems.** At each intercell position  $(i + \frac{1}{2}, j)$  one solves the  $x$ -split one-dimensional Riemann problem with data  $\hat{\mathbf{U}}_{i,j}^{+x}$ ,  $\hat{\mathbf{U}}_{i+1,j}^{-x}$  to find a solution  $\mathbf{U}_{i+1/2,j}(x/t)$ . Similarly, at each interface position  $(i, j + \frac{1}{2})$  one solves the  $y$ -split one-dimensional Riemann problem with data  $\hat{\mathbf{U}}_{i,j}^{+y}$ ,  $\hat{\mathbf{U}}_{i,j+1}^{-y}$  to find a solution  $\mathbf{U}_{i,j+1/2}(y/t)$ . The corresponding intercell fluxes are found as in the one-dimensional Godunov method, namely

$$\mathbf{F}_{i+\frac{1}{2},j} = \mathbf{F}(\mathbf{U}_{i+\frac{1}{2},j}(0)) ; \quad \mathbf{G}_{i,j+\frac{1}{2}} = \mathbf{G}(\mathbf{U}_{i,j+\frac{1}{2}}(0)) . \quad (16.66)$$

*Example 16.12.* Here we apply the MUSCL–Hancock finite volume scheme to the linear advection equation (16.50). It is easily shown that the intercell fluxes are

$$\left. \begin{aligned} f_{i+\frac{1}{2},j} &= a_1[u_{i,j}^n + \frac{1}{2}(1-c_1)\Delta_i - \frac{1}{2}c_2\Delta_j], \\ g_{i,j+\frac{1}{2}} &= a_2[u_{i,j}^n + \frac{1}{2}(1-c_2)\Delta_j - \frac{1}{2}c_1\Delta_i]. \end{aligned} \right\} \quad (16.67)$$

By taking central difference approximations for the slopes  $\Delta_i$ ,  $\Delta_j$  and substituting the resulting fluxes into the finite volume formula (16.46) one obtains the scheme

$$\left. \begin{aligned} u_{i,j}^{n+1} &= u_{i,j}^n \\ -c_1 \{ (u_{i,j}^n - u_{i-1,j}^n) &+ \frac{1}{4}(1-c_1)[(u_{i+1,j}^n - u_{i,j}^n) - (u_{i-1,j}^n - u_{i-2,j}^n)] \} \\ + \frac{1}{4}c_1c_2[(u_{i,j+1}^n - u_{i,j-1}^n) &- (u_{i-1,j+1}^n - u_{i-1,j-1}^n)] \\ -c_2 \{ (u_{i,j}^n - u_{i,j-1}^n) &+ \frac{1}{4}(1-c_2)[(u_{i,j+1}^n - u_{i,j}^n) - (u_{i,j-1}^n - u_{i,j-2}^n)] \} \\ + \frac{1}{4}c_1c_2[(u_{i+1,j}^n - u_{i-1,j}^n) &- (u_{i+1,j-1}^n - u_{i-1,j-1}^n)]. \end{aligned} \right\} \quad (16.68)$$

The stencil of the scheme as applied to (16.50) with  $a_1 > 0$  and  $a_2 > 0$  is shown in Fig. 16.5 and has 10 points.

**Exercise 16.13.** Apply the accuracy theorem (16.6), see equation (16.59), to show that scheme (16.68) is second-order accurate in space and time.

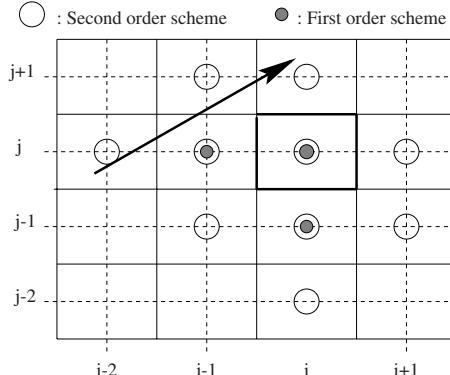
**Solution 16.14.** Left to the reader.

*Remark 16.15.* Stability tests according to the method of Sect. 16.4.2 suggest that the MUSCL–Hancock scheme (16.68) has the same stability condition (16.62), as the first-order Godunov finite volume scheme (16.51). The unsplit finite volume MUSCL–Hancock scheme has increased the accuracy of the Godunov finite volume scheme but not its stability range.

**Exercise 16.16.** Attempt to extend the MUSCL–Hancock approach to construct a three-dimensional finite volume scheme for non-linear systems of conservation laws. Apply the scheme to the three-dimensional linear advection equation (16.55) and show that it is second-order accurate in space and time.

**Solution 16.17.** Left to the reader.

As to the computation of the numerical fluxes, these can be found by using the exact Riemann solver, given in Chap. 6 for the Euler equations, or any of the approximate Riemann solvers presented in Chaps. 9 to 12. One can also use the Flux Vector Splitting schemes of Chap. 8.



**Fig. 16.5.** Stencil of the MUSCL–Hancock Finite Volume scheme for the linear advection equation with positive velocity components  $a_1$  and  $a_2$ . The arrow represents the direction of the velocity vector. Also shown is the stencil of the first–order version of the scheme

In order to *control* spurious oscillations in the two–dimensional schemes, one applies one–dimensional TVD constraints in exactly the same manner as for one–dimensional problems, by replacing the slopes by *limited slopes*; see Chaps. 13 and 14. The resulting schemes are usually referred to, incorrectly, as *two–dimensional TVD schemes*. The same remark applies to three–dimensional schemes. We first note that the schemes are *not strictly* TVD as understood for one–dimensional model equations, see Chap. 13; the second order slope limited scheme will still produce spurious oscillations near high gradients. Also note that if the TVD conditions were to be enforced in a two–dimensional sense, then, as proved by Goodman and LeVeque [219], the scheme would be oscillation free but at most first–order accurate.

## 16.6 WAF–Type Finite Volume Schemes

The Weighted Average Flux (WAF) approach was first introduced [499], [506] to construct schemes to solve one–dimensional hyperbolic systems of conservation laws. The WAF schemes are of the Godunov type and of second order accuracy in space and time. Detailed descriptions of the schemes along with their TVD constraints are presented in Sect. 13.7 of Chap. 13 for scalar problems and in Sects. 14.2–14.3 of Chap. 14 for non–linear systems. By means of dimensional splitting these WAF schemes may be applied to solve multidimensional problems. See Sects. 16.2–16.3 of this chapter about details for the two and three dimensional Euler equations.

The WAF approach may be extended to derive unsplit finite volume schemes for multi–dimensional hyperbolic systems [59], [63], [61]. In the next section we first present some details of the approach as applied to the two

and three dimensional linear advection equations. Then we study two of the possible extensions to multidimensional non-linear systems in two and three space dimensions.

### 16.6.1 Two-Dimensional Linear Advection

Consider the two-dimensional linear advection equation with constant coefficients

$$u_t + f(u)_x + g(u)_y = 0 ; \quad f(u) = a_1 u ; \quad g(u) = a_2 u \quad (16.69)$$

and finite volume schemes of the form

$$u_{i,j}^{n+1} = u_{i,j}^n + \frac{\Delta t}{\Delta x} [f_{i-\frac{1}{2},j} - f_{i+\frac{1}{2},j}] + \frac{\Delta t}{\Delta y} [g_{i,j-\frac{1}{2}} - g_{i,j+\frac{1}{2}}] \quad (16.70)$$

to solve (16.69). The objective is to prescribe the intercell fluxes  $f_{i+\frac{1}{2},j}$  and  $g_{i,j+\frac{1}{2}}$ . The WAF approach specifies a flux in the  $x$ -direction as follows

$$f_{i+\frac{1}{2},j} = \frac{1}{V(B)} \int_{t_1}^{t_2} \int_{x_1}^{x_2} \int_{y_1}^{y_2} f(u_{i+\frac{1}{2}}(x, y, t)) dx dy dt , \quad (16.71)$$

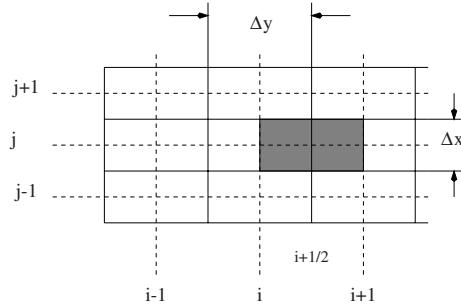
where  $B \equiv [t_1, t_2] \times [x_1, x_2] \times [y_1, y_2]$  is the domain of integration in  $t$ - $x$ - $y$  space, that includes the relevant intercell boundary at  $x_{i+\frac{1}{2}}$ ;  $V(B)$  denotes the volume of  $B$ . In general  $f(u)$  is the flux component perpendicular to the intercell boundary. An analogous definition holds for  $g_{i,j+\frac{1}{2}}$ . Particular schemes are obtained by specifying (i) the spatial and temporal domain of integration, (ii) the integrand and (iii) integration schemes to evaluate (16.71). Here we make the following choices

$$\left. \begin{aligned} t_1 &= 0 , & t_2 &= \Delta t , & c_1 &\leq 1 , & c_2 &\leq 1 , \\ x_1 &= -\frac{1}{2}\Delta x , & x_2 &= \frac{1}{2}\Delta x , & y_1 &= 0 , & y_2 &= \Delta y . \end{aligned} \right\} \quad (16.72)$$

Fig. 16.6 depicts the spatial integration range to compute the intercell flux. The choice of  $\Delta t$  and the conditions on the Courant numbers  $c_1$  and  $c_2$  are obviously related. The integrand is determined by specifying  $u_{i+\frac{1}{2}}(x, y, t)$ . Here we assume  $u_{i+\frac{1}{2}}(x, y, t)$  to be the solution of *relevant* two-dimensional Riemann problems for (16.69), with initial data  $u(x, y, 0) = u_{i,j}^n$  for  $(x, y)$  in the computing cell  $I_{i,j}$ , overlapping the spatial range of integration in a time  $\Delta t$ .

For the model equation (16.69) the solution of the two-dimensional Riemann problem is trivial. Exact integration in (16.71) leads to a second-order accurate scheme in space and time, see [59], [63]. Here we only consider schemes arising from approximate evaluation of the integral (16.71).

Assume that  $a_1 > 0$  and  $a_2 > 0$  in (16.69) and restrict the Courant numbers to satisfy  $c_1 \leq \frac{1}{2}$  and  $c_2 \leq \frac{1}{2}$ ; this in turn restricts the number of states



**Fig. 16.6.** Spacial integration range for evaluating the WAF intercell flux at  $i + \frac{1}{2}$

that influence  $f_{i+\frac{1}{2},j}$  to four, namely  $u_{i,j}^n$ ,  $u_{i+1,j}^n$ ,  $u_{i,j-1}^n$  and  $u_{i+1,j-1}^n$ . These states are laid out in the four quadrants of a rectangle centred at  $(x_{i+\frac{1}{2}}, y_{j-\frac{1}{2}})$ , and form the initial conditions of a two-dimensional Riemann problem. For different signs of  $a_1$  and  $a_2$  the relevant two-dimensional Riemann problem has different initial conditions.

Having specified the integration domain and the integrand in (16.71), we are left with the task of choosing integration schemes to evaluate the integral.

### First–Order Schemes

Apply exact integration in time, the midpoint rule perpendicular to the boundary, and exact integration parallel to the boundary. The integral form of the intercell flux (16.71) simplifies in this case to

$$f_{i+\frac{1}{2},j} = \frac{1}{\Delta t} \frac{1}{\Delta y} \int_0^{\Delta t} \int_{B \cap \{x=0\}} f(u_{i+\frac{1}{2}}(0, y, t)) dy dt , \quad (16.73)$$

where we integrate over the plane  $x = 0$ , in local coordinates. This gives

$$f_{i+\frac{1}{2},j} = \frac{1}{2}(2 - c_2)f_{i,j} + \frac{1}{2}c_2f_{i,j-1} . \quad (16.74)$$

The flux  $g_{i,j+\frac{1}{2}}$  follows similarly. Substitution of all four intercell fluxes into the finite volume formula (16.70) produces the scheme

$$\left. \begin{aligned} u_{i,j}^{n+1} &= (1 - c_1 - c_2 + c_1c_2)u_{i,j}^n + c_2(1 - c_1)u_{i,j-1}^n \\ &\quad + c_1(1 - c_2)u_{i-1,j}^n + c_1c_2u_{i-1,j-1}^n . \end{aligned} \right\} \quad (16.75)$$

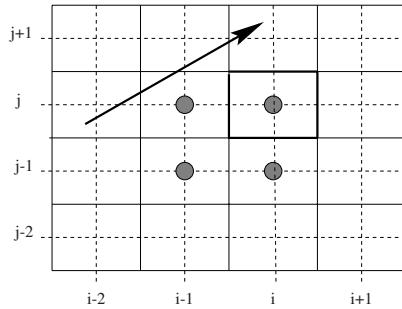
This WAF two-dimensional finite volume scheme turns out to be identical to the *Corner Transport Upwind* (CTU) scheme of Colella [134].

**Exercise 16.18.** Using the accuracy theorem (16.6), see (16.59), show that the CTU scheme (16.75) is first–order accurate in space and time.

Fig. 16.7 shows the stencil of two-dimensional first-order upwind scheme (16.75). Compare with the stencil for the Godunov finite volume method shown in Fig. 16.4. The CTU (WAF) finite volume scheme does contain the upwind point  $u_{i-1,j-1}^n$  and is the natural two-dimensional extension of the one-dimensional Godunov first-order upwind scheme. Colella proved the scheme (16.75) to be stable under the condition

$$\max\{c_1, c_2\} \leq 1. \quad (16.76)$$

Thus, the WAF approach is capable of producing first-order two dimensional finite volume schemes with twice the stability limit of the Godunov finite volume scheme (16.51), which is constructed by straightforward application of the Godunov one-dimensional fluxes.



**Fig. 16.7.** Stencil of the first-order WAF finite volume scheme (identical to the CTU scheme of Colella) for the two-dimensional linear advection equation with positive  $a_1$  and  $a_2$ . The arrow indicates the direction of the velocity vector

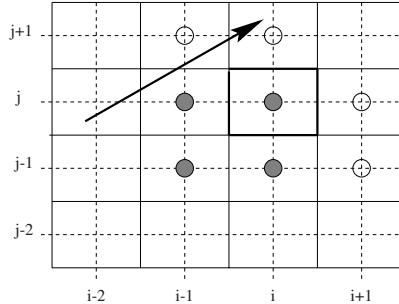
## A Second-Order Scheme

We now consider the midpoint rule in time combined with exact integration in both space directions in (16.71) and assume that  $c_1 \leq 1$ ,  $c_2 \leq 1$ . This gives the intercell flux

$$f_{i+\frac{1}{2},j} = \left. \begin{aligned} & \frac{1}{4}(1+c_1)(2-c_2)f_{i,j} + \frac{1}{4}(1-c_1)(2-c_2)f_{i+1,j} \\ & + \frac{1}{4}(1+c_1)c_2f_{i,j-1} + \frac{1}{4}(1-c_1)c_2f_{i+1,j-1} \end{aligned} \right\} \quad (16.77)$$

The flux  $g_{i,j+\frac{1}{2}}$  follows similarly. Substitution of all fluxes into the finite volume formula (16.70) gives the scheme in full as

$$u_{i,j}^{n+1} = \left. \begin{aligned} & [1 - \frac{1}{2}c_1^2(2-c_2) - \frac{1}{2}c_2^2(2-c_1)] u_{i,j}^n \\ & - \frac{1}{4}c_1(1-c_1)(2-c_2)u_{i+1,j}^n - \frac{1}{4}(2-c_1)c_2(1-c_2)u_{i,j+1}^n \\ & - \frac{1}{4}c_1c_2(1-c_1)u_{i+1,j-1}^n - \frac{1}{4}c_1c_2(1-c_2)u_{i-1,j+1}^n \\ & + \frac{1}{4}[c_1(1+c_1)(2-c_2) - 2c_1c_2^2]u_{i-1,j}^n \\ & + \frac{1}{4}c_1c_2(2+c_1+c_2)u_{i-1,j-1}^n. \end{aligned} \right\} \quad (16.78)$$



**Fig. 16.8.** Stencil of the second-order WAF finite volume scheme for the two-dimensional linear advection equation with positive  $a_1$  and  $a_2$ . Compare with stencil of first-order scheme of Fig. 16.7

When  $c_1 = 0$  or  $c_2 = 0$ , the scheme reduces to the one dimensional Lax–Wendroff scheme, and when  $c_1 = c_2 = 1$ , it reduces to  $u_{i,j}^{n+1} = u_{i-1,j-1}^n$ , which reproduces the exact solution under the given initial conditions.

**Exercise 16.19.** Apply the accuracy theorem (16.6) to show that scheme (16.78) is second-order accurate in space and time.

**Solution 16.20.** Left to the reader.

The eight-point stencil for the second-order WAF finite volume scheme (16.78) is shown in Fig. 16.8; compare with the ten-point stencil for the second-order MUSCL–Hancock scheme (16.68) shown in Fig. 16.5. Numerical stability tests, as described in Sect. 16.4.2, suggest that scheme (16.78) is stable provided the condition (16.76) holds. The MUSCL–Hancock scheme (16.68) has stability restriction  $c_1 + c_2 \leq 1$ , which is half that of this WAF finite volume scheme. Both WAF finite volume schemes (16.75) (CTU) and (16.78) are much more stable than the two-dimensional Lax–Wendroff scheme [303], [252].

### 16.6.2 Three-Dimensional Linear Advection

The WAF approach is now extended to generate finite volume schemes for the linear advection equation in three space dimensions

$$u_t + f(u)_x + g(u)_y + h(u)_z = 0 ; \quad f(u) = a_1 u , \quad g(u) = a_2 u , \quad h(u) = a_3 u . \quad (16.79)$$

The finite volume schemes have the form

$$\left. \begin{aligned} u_{i,j,k}^{n+1} = u_{i,j,k}^n &+ \frac{\Delta t}{\Delta x} [f_{i-\frac{1}{2},j,k} - f_{i+\frac{1}{2},j,k}] + \frac{\Delta t}{\Delta y} [g_{i,j-\frac{1}{2},k} - g_{i,j+\frac{1}{2},k}] \\ &+ \frac{\Delta t}{\Delta z} [h_{i,j,k-\frac{1}{2}} - h_{i,j,k+\frac{1}{2}}] , \end{aligned} \right\} \quad (16.80)$$

where  $f_{i+\frac{1}{2},j,k}$ ,  $g_{i,j+\frac{1}{2},k}$  and  $h_{i,j,k+\frac{1}{2}}$  are the intercell fluxes in the  $x$ ,  $y$  and  $z$  directions respectively. A general WAF intercell flux in three space dimensions reads

$$f_{i+\frac{1}{2},j,k} = \frac{1}{V(B)} \int_{t_1}^{t_2} \int_{x_1}^{x_2} \int_{y_1}^{y_2} \int_{z_1}^{z_2} f(u_{i+\frac{1}{2}}(x, y, z, t)) dx dy dz dt, \quad (16.81)$$

where  $B = [t_1, t_2] \times [x_1, x_2] \times [y_1, y_2] \times [z_1, z_2]$  is the domain of integration in  $t-x-y-z$  space;  $V(B) = (t_2 - t_1) \times (x_2 - x_1) \times (y_2 - y_1) \times (z_2 - z_1)$  is the volume of  $B$ . The fluxes  $g_{i,j+\frac{1}{2},k}$  and  $h_{i,j,k+\frac{1}{2}}$  are defined similarly.

## A First-Order Scheme

A three-dimensional version of the first-order WAF scheme for the two-dimensional linear advection equation is now derived. A choice of parameter values consistent with (16.72) is assumed. By using the midpoint rule in time, the midpoint rule in space perpendicular to the boundary, and exact integration in space in both directions parallel to the boundary, the resulting intercell flux is

$$f_{i+\frac{1}{2},j,k} = \frac{1}{\Delta y \Delta z} \int_B \bigcap_{\{x=0\}} f(u_{i+\frac{1}{2}}(0, y, z, \frac{1}{2}\Delta t)) dy dz. \quad (16.82)$$

This can be written explicitly as

$$\left. \begin{aligned} f_{i+\frac{1}{2},j,k} &= \frac{1}{4}(2 - c_2)(2 - c_3)f_{i,j,k} + \frac{1}{4}c_2(2 - c_3)f_{i,j-1,k} \\ &\quad + \frac{1}{4}(2 - c_2)c_3f_{i,j,k-1} + \frac{1}{4}c_2c_3f_{i,j-1,k-1}. \end{aligned} \right\} \quad (16.83)$$

The derivation of the fluxes  $g_{i,j+\frac{1}{2},k}$  and  $h_{i,j,k+\frac{1}{2}}$  follows by symmetry. Substituting these into the finite volume formula (16.80) gives the scheme as

$$\left. \begin{aligned} u_{i,j,k}^{n+1} &= \left[ 1 - \frac{1}{4}c_1(2 - c_2)(2 - c_3) - \frac{1}{4}(2 - c_1)(2 - c_2)c_3 \right] u_{i,j,k}^n \\ &\quad - \left[ \frac{1}{4}(2 - c_1)c_2(2 - c_3) - \frac{1}{4}(2 - c_1)(2 - c_2)c_3 \right] u_{i,j,k}^n \\ &\quad + c_1 \left[ (1 - c_2)(1 - c_3) - \frac{1}{4}c_2c_3 \right] u_{i-1,j,k}^n \\ &\quad + \frac{1}{4}(4 - 3c_1)c_2c_3u_{i,j-1,k-1}^n \\ &\quad + c_2 \left[ (1 - c_1)(1 - c_3) - \frac{1}{4}c_1c_3 \right] u_{i,j-1,k}^n \\ &\quad + \frac{1}{4}c_1(4 - 3c_2)c_3u_{i-1,j,k-1}^n \\ &\quad + c_3 \left[ (1 - c_1)(1 - c_2) - \frac{1}{4}c_1c_2 \right] u_{i,j,k-1}^n \\ &\quad + \frac{1}{4}c_1c_2(4 - 3c_3)u_{i-1,j-1,k}^n \\ &\quad + \frac{3}{4}c_1c_2c_3u_{i-1,j-1,k-1}^n. \end{aligned} \right\} \quad (16.84)$$

By setting  $c_1 = c_2 = c_3 = c$  in (16.84) it can be seen that not all the coefficients are positive if  $c \geq \frac{2}{3}$ ; see for example the coefficient of  $u_{i-1,j,k}$ . A sufficient condition for the scheme to remain *monotone* is

$$\max\{c_1, c_2, c_3\} \leq \frac{2}{3}. \quad (16.85)$$

**Exercise 16.21.** Apply the accuracy theorem (16.7) to show that scheme (16.84) is first-order accurate in space and time. This scheme is a natural three-dimensional extension of the one-dimensional Godunov first-order upwind scheme.

### A Second Order Scheme

Working from the experience of the two dimensional case, the midpoint rule in time and exact integration in all three space directions will now be used to generate a second-order scheme. The flux in the  $x$ -direction becomes

$$f_{i+\frac{1}{2},j,k} = \frac{1}{V(B)} \int_B f(u_{i+\frac{1}{2}}(x, y, z, \frac{1}{2}\Delta t)) dx dy dz, \quad (16.86)$$

where  $u_{i+\frac{1}{2}}(x, y, z, t)$  is the exact solution to the three-dimensional Riemann problem for (16.79) with initial conditions  $\{u_{i,j,k}^n\}$ . By assuming  $a_1 > 0$ ,  $a_2 > 0$ ,  $a_3 > 0$  in (16.79) and performing the integration we obtain the intercell flux

$$\left. \begin{aligned} f_{i+\frac{1}{2},j,k} = & \frac{1}{8}(1+c_1)(2-c_2)(2-c_3)f_{i,j,k} \\ & + \frac{1}{8}(1-c_1)(2-c_2)(2-c_3)f_{i+1,j,k} \\ & + \frac{1}{8}(1+c_1)c_2(2-c_3)f_{i,j-1,k} \\ & + \frac{1}{8}(1-c_1)c_2(2-c_3)f_{i+1,j-1,k} \\ & + \frac{1}{8}(1+c_1)(2-c_2)c_3f_{i,j,k-1} \\ & + \frac{1}{8}(1-c_1)(2-c_2)c_3f_{i+1,j,k-1} \\ & + \frac{1}{8}(1+c_1)c_2c_3f_{i,j-1,k-1} \\ & + \frac{1}{8}(1-c_1)c_2c_3f_{i+1,j-1,k-1} \end{aligned} \right\} \quad (16.87)$$

The fluxes  $g_{i,j+\frac{1}{2},k}$  and  $h_{i,j,k+\frac{1}{2}}$  follow by symmetry. Substituting all the fluxes into the finite volume formula (16.80) gives the scheme

$$\left. \begin{aligned} u_{i,j,k}^{n+1} = & u_{i,j,k}^n \\ & - \frac{1}{4} [c_1^2(2-c_2)(2-c_3) + (2-c_1)c_2^2(2-c_3) + (2-c_1)(2-c_2)c_3^2] u_{i,j,k}^n \\ & + \frac{1}{8} c_1 [(1+c_1)(2-c_2)(2-c_3) - 2c_2^2(2-c_3) - 2(2-c_2)c_3^2] u_{i-1,j,k}^n \\ & + \frac{1}{8} c_2 [(2-c_1)(1+c_2)(2-c_3) - 2c_1^2(2-c_3) - 2(2-c_1)c_3^2] u_{i,j-1,k}^n \\ & + \frac{1}{8} c_3 [(2-c_1)(2-c_2)(1+c_3) - 2(2-c_1)c_2^2 - 2c_1^2(2-c_2)] u_{i,j,k-1}^n \\ & + \frac{1}{8} c_1 c_2 [(2-c_3)(2+c_1+c_2) - 2c_3^2] u_{i-1,j-1,k}^n \\ & + \frac{1}{8} c_1 c_3 [(2-c_2)(2+c_1+c_3) - 2c_2^2] u_{i-1,j,k-1}^n \\ & + \frac{1}{8} c_2 c_3 [(2-c_1)(2+c_2+c_3) - 2c_1^2] u_{i,j-1,k-1}^n \\ & + \frac{1}{8} c_1 c_2 c_3 (3+c_1+c_2+c_3) u_{i-1,j-1,k-1}^n \\ & - \frac{1}{8} c_1 (1-c_1)(2-c_2)(2-c_3) u_{i+1,j,k}^n - \frac{1}{8} c_1 (1-c_1)c_2(2-c_3) u_{i+1,j-1,k}^n \\ & - \frac{1}{8} c_1 (1-c_1)(2-c_2)c_3 u_{i+1,j,k-1}^n - \frac{1}{8} c_1 (1-c_1)c_2c_3 u_{i+1,j-1,k-1}^n \\ & - \frac{1}{8} (2-c_1)c_2(1-c_2)(2-c_3) u_{i,j+1,k}^n - \frac{1}{8} c_1 c_2(1-c_2)(2-c_3) u_{i-1,j+1,k}^n \\ & - \frac{1}{8} (2-c_1)c_2(1-c_2)c_3 u_{i,j+1,k-1}^n - \frac{1}{8} c_1 c_2(1-c_2)c_3 u_{i-1,j+1,k-1}^n \\ & - \frac{1}{8} (2-c_1)(2-c_2)c_3(1-c_3) u_{i,j,k+1}^n - \frac{1}{8} (2-c_1)c_2c_3(1-c_3) u_{i,j-1,k+1}^n \\ & - \frac{1}{8} c_1 (2-c_2)c_3(1-c_3) u_{i-1,j,k+1}^n - \frac{1}{8} c_1 c_2c_3(1-c_3) u_{i-1,j-1,k+1}^n \end{aligned} \right\} \quad (16.88)$$

**Exercise 16.22.** Apply the accuracy theorem (16.7) to show that scheme (16.88) is second-order accurate in space and time.

**Solution 16.23.** Left to the reader.

Application of the numerical tests for stability according to the method of Sect. 16.4.2 suggests that the three-dimensional scheme (16.88) is stable under the condition

$$\max\{c_1, c_2, c_3\} \leq \frac{2}{3}, \quad (16.89)$$

which is also the condition (16.85) for monotonicity for the first-order three dimensional scheme (16.84). The three-dimensional Godunov finite volume scheme (16.56) has stability restriction

$$\max\{c_1, c_2, c_3\} \leq \frac{1}{3}, \quad (16.90)$$

which is half that of the WAF second-order finite volume scheme (16.88).

### 16.6.3 Schemes for Two-Dimensional Nonlinear Systems

The construction of multi-dimensional WAF-type schemes for non-linear systems is based on *re-interpreting* the flux formulae derived for the linear advection equation in multidimensions.

#### An Average Flux Second-Order Scheme

First we re-write the intercell flux (16.77) as

$$f_{i+\frac{1}{2},j} = \frac{1}{2}(1+c_1)f(u_{i,j}^{god}) + \frac{1}{2}(1-c_1)f(u_{i+1,j}^{god}), \quad (16.91)$$

where

$$\left. \begin{aligned} u_{i,j}^{god} &= u_{i,j}^n + \frac{\frac{1}{2}\Delta t}{\Delta y} \left[ g_{i,j-\frac{1}{2}}^{god} - g_{i,j+\frac{1}{2}}^{god} \right], \\ u_{i+1,j}^{god} &= u_{i+1,j}^n + \frac{\frac{1}{2}\Delta t}{\Delta y} \left[ g_{i+1,j-\frac{1}{2}}^{god} - g_{i+1,j+\frac{1}{2}}^{god} \right] \end{aligned} \right\} \quad (16.92)$$

and  $g_{i,j+\frac{1}{2}}^{god} = a_2 u_{i,j}^n$  is the Godunov first-order upwind flux,  $a_2 > 0$ .

Formulae (16.91) and (16.92) are now *re-interpreted* in terms of two well known *one-dimensional* operators that are valid for non-linear systems. Formulae (16.91) is the Weighted Average Flux scheme that results from solving the one-dimensional Riemann problem for the linear advection equation  $u_t + a_1 u_x = 0$ , at the intercell position  $i + \frac{1}{2}$ , with *modified* initial data  $(u_{i,j}^{god}, u_{i+1,j}^{god})$ . See Sect. 13.3 of Chap. 13 for details on the one dimensional WAF flux for the linear advection equation. The modified data results from applying the Godunov first-order upwind method to the one dimensional advection equation  $u_t + a_2 u_y = 0$  for a time  $\frac{1}{2}\Delta t$ .

We now construct intercell fluxes  $\mathbf{F}_{i+\frac{1}{2},j}$  and  $\mathbf{G}_{i,j+\frac{1}{2}}$  for the finite volume scheme (16.46) to solve the two-dimensional nonlinear system (16.45) following the WAF approach. This is done by *imitating* (16.91)–(16.92). The intercell fluxes result from applying two one-dimensional operators in succession, namely

$$\mathbf{F}_{i+\frac{1}{2},j} = \mathbf{L}_{x,\frac{1}{2}\Delta t}^{waf} \left( \mathbf{L}_{y,\frac{1}{2}\Delta t}^{god}(\mathbf{U}_{i,j}^n), \mathbf{L}_{y,\frac{1}{2}\Delta t}^{god}(\mathbf{U}_{i+1,j}^n) \right) \quad (16.93)$$

and

$$\mathbf{G}_{i,j+\frac{1}{2}} = \mathbf{L}_{y,\frac{1}{2}\Delta t}^{waf} \left( \mathbf{L}_{x,\frac{1}{2}\Delta t}^{god}(\mathbf{U}_{i,j}^n), \mathbf{L}_{x,\frac{1}{2}\Delta t}^{god}(\mathbf{U}_{i,j+1}^n) \right). \quad (16.94)$$

Here the operator  $\mathbf{L}_{s,dt}^{waf}(\mathbf{Z}_L, \mathbf{Z}_R)$  applies to a pair of states  $(\mathbf{Z}_L, \mathbf{Z}_R)$  and is defined as

$$\mathbf{L}_{s,dt}^{waf}(\mathbf{Z}_L, \mathbf{Z}_R) = \frac{1}{\Delta s} \int_{-\frac{1}{2}\Delta s}^{\frac{1}{2}\Delta s} \mathbf{E} \left( \mathbf{Z}_{LR} \left( \frac{s}{dt} \right) \right) ds. \quad (16.95)$$

This is an integral average of the flux component  $\mathbf{E}$  in the  $s$ -direction normal to the interface between the states  $\mathbf{Z}_L$  and  $\mathbf{Z}_R$  and gives a WAF-type flux.  $\mathbf{Z}_{LR}$  denotes the solution of the Riemann problem with data  $(\mathbf{Z}_L, \mathbf{Z}_R)$ . In practice one approximates the WAF flux by a summation, namely

$$\mathbf{F}_{LR} = \frac{1}{2} [\mathbf{F}(\mathbf{Z}_L) + \mathbf{F}(\mathbf{Z}_R)] - \frac{1}{2} \sum_{k=1}^N \text{sign}(c_{LR}^{(k)}) \phi_{LR}^{(k)} \Delta \mathbf{F}_{LR}^{(k)}, \quad (16.96)$$

where  $c_{LR}^{(k)}$  is the Courant number for wave  $k$  in the solution of the Riemann problem,  $\Delta \mathbf{F}_{LR}^{(k)}$  is the flux jump across wave  $k$  and  $\phi_{LR}^{(k)} = \phi_{LR}^{(k)}(r_{LR}^{(k)})$  is a one-dimensional limiter function for wave  $k$ . For details on the WAF flux see Sect. 14.3 of Chap. 14 and Sect. 16.3.3 of this chapter.

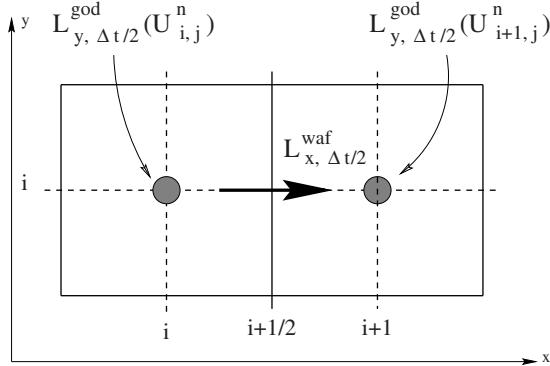
The second operator in (16.93) and (16.94) is the Godunov one dimensional operator

$$\mathbf{L}_{s,dt}^{god}(\mathbf{Z}^n) = \mathbf{Z}^n + \frac{dt}{\Delta s} \left[ \mathbf{K}_{l-\frac{1}{2}}^{god} - \mathbf{K}_{l+\frac{1}{2}}^{god} \right], \quad (16.97)$$

as applied to a state  $\mathbf{Z}^n$  over a time  $dt$  in the direction  $s$  and  $\mathbf{K}_{l+\frac{1}{2}}^{god}$  is the one-dimensional Godunov first-order upwind flux.

Fig. 16.9 depicts the evaluation of the intercell flux  $\mathbf{F}_{i+\frac{1}{2},j}$  in (16.93), for which one performs the following two steps:

- (I) **Modification of Data:** The initial data on the left and right side of the interface is modified by applying the Godunov first-order upwind scheme in the  $y$ -direction (cross direction) for a time  $\frac{1}{2}\Delta t$  to it. One obtains the two states  $\mathbf{L}_{y,\frac{1}{2}\Delta t}^{god}(\mathbf{U}_{i,j}^n)$  and  $\mathbf{L}_{y,\frac{1}{2}\Delta t}^{god}(\mathbf{U}_{i+1,j}^n)$ .
- (II) **The Riemann Problem:** Solve the Riemann problem in the  $x$ -direction with the modified data  $(\mathbf{L}_{y,\frac{1}{2}\Delta t}^{god}(\mathbf{U}_{i,j}^n), \mathbf{L}_{y,\frac{1}{2}\Delta t}^{god}(\mathbf{U}_{i+1,j}^n))$  and compute the WAF flux in the usual way; see (16.96). See also (16.39).



**Fig. 16.9.** Intercell flux  $\mathbf{F}_{i+\frac{1}{2},j}$  is obtained by applying the WAF operator  $\mathbf{L}_{x,\frac{1}{2}\Delta t}^{waf}$  in the  $x$ -direction to modified data in cells  $(i, j)$  and  $(i+1, j)$ . Modified data is obtained by applying the Godunov operator  $\mathbf{L}_{y,\frac{1}{2}\Delta t}^{god}$  in the  $y$ -direction to initial states  $\mathbf{U}_{i,j}^n$ ,  $\mathbf{U}_{i+1,j}^n$

The evaluation of the intercell flux  $\mathbf{G}_{i,j+\frac{1}{2}}$  in (16.94) is entirely analogous. One first modifies the initial data by applying the Godunov operator in the  $x$ -direction for a time  $\frac{1}{2}\Delta t$  to it and then solves the Riemann problem in the  $y$ -direction to compute a WAF flux in the  $y$ -direction.

### An Average State Scheme

Another extension to non-linear systems is based on the definition of a Weighted Average State (WAS) operator

$$\mathbf{L}_{s,\Delta t}^{was}(\mathbf{Z}_L, \mathbf{Z}_R) = \frac{1}{\Delta s} \int_{-\frac{1}{2}\Delta s}^{\frac{1}{2}\Delta s} \mathbf{Z}_{LR} \left( \frac{s}{\Delta t} \right) ds \quad (16.98)$$

This is identical to the Weighted Average State version of WAF in one dimension, see Sect. 14.3.2 of Chap. 14. The flux  $\mathbf{F}_{i+\frac{1}{2},j}$  is now given by

$$\mathbf{F}_{i+\frac{1}{2},j} = \mathbf{F} \left( \mathbf{L}_{y,\frac{1}{2}\Delta t}^{god} \left( \mathbf{L}_{x,\frac{1}{2}\Delta t}^{was} (\mathbf{U}_{i,j}^n, \mathbf{U}_{i+1,j}^n) \right) \right) \quad (16.99)$$

and the flux  $\mathbf{G}_{i,j+\frac{1}{2}}$  is given by

$$\mathbf{G}_{i,j+\frac{1}{2}} = \mathbf{G} \left( \mathbf{L}_{x,\frac{1}{2}\Delta t}^{god} \left( \mathbf{L}_{y,\frac{1}{2}\Delta t}^{was} (\mathbf{U}_{i,j}^n, \mathbf{U}_{i,j+1}) \right) \right) . \quad (16.100)$$

### First-Order Scheme and the TVD Condition

Recall that the WAF one-dimensional flux with the one dimensional TVD condition includes the Godunov first-order upwind scheme by simply setting the one-dimensional limiter function to an appropriate value, namely

$\phi_{LR}^{(k)} = \text{sign}(c_{LR}^{(k)}) \times |c_{LR}^{(k)}|$ . Therefore, by utilising the WAF operator in its first–order mode in (16.91) and applying the scheme to the two–dimensional linear advection equation (16.69) one obtains the first–order WAF finite volume scheme, which is equivalent to the CTU scheme of Colella for the linear advection equation. As seen earlier this first–order scheme is monotone and has linear stability limit 1. We therefore obtain an extension of the CTU scheme to two–dimensional nonlinear systems of hyperbolic conservation laws by simply applying the WAF operator in its first–order mode (Godunov’s first order upwind scheme) in (16.93)–(16.94), or (16.99)–(16.100).

The second–order scheme results from the second–order WAF operator in (16.93)–(16.94) or (16.99)–(16.100). In order to control the expected *spurious oscillations* we simply apply the WAF flux together with its TVD constraint. Note that this TVD constraint is applied only in a one–dimensional sense. The resulting scheme is *not TVD in a two–dimensional sense*. As a matter of fact spurious oscillations are still present, even for the model two–dimensional equation. In applications to two and three–dimensional nonlinear systems we find that spurious oscillations are small. Billett and Toro [65] have produced some preliminary results on a Total Variation Bounded (TVB) version of the multidimensional WAF schemes.

#### 16.6.4 Schemes for Three–Dimensional Nonlinear Systems

Following the same approach as for two–dimensional non–linear systems, we now construct WAF finite volume schemes of the form (16.54) for solving three–dimensional non–linear systems of hyperbolic conservation laws (16.53). We *re–interpret* the schemes (16.84) and (16.88) derived for the three–dimensional linear advection equation (16.79).

A *second–order Weighted Average Flux Scheme* has intercell flux in the  $x$ –direction given by

$$\left. \begin{aligned} \mathbf{F}_{i+\frac{1}{2},j,k} &= \mathbf{L}_{x,\frac{1}{2}\Delta t}^{\text{waf}}(\mathbf{U}_{i,j,k}^{\text{god}}, \mathbf{U}_{i+1,j,k}^{\text{god}}), \\ \mathbf{U}_{i,j,k}^{\text{god}} &= \mathbf{L}_{y,\frac{1}{2}\Delta t}^{\text{god}}(\mathbf{L}_{z,\frac{1}{2}\Delta t}^{\text{god}}(\mathbf{U}_{i,j,k}^n)), \\ \mathbf{U}_{i+1,j,k}^{\text{god}} &= \mathbf{L}_{y,\frac{1}{2}\Delta t}^{\text{god}}(\mathbf{L}_{z,\frac{1}{2}\Delta t}^{\text{god}}(\mathbf{U}_{i+1,j,k}^n)), \end{aligned} \right\} \quad (16.101)$$

numerical flux in the  $y$ –direction given by

$$\left. \begin{aligned} \mathbf{G}_{i,j+\frac{1}{2},k} &= \mathbf{L}_{y,\frac{1}{2}\Delta t}^{\text{waf}}(\mathbf{U}_{i,j,k}^{\text{god}}, \mathbf{U}_{i,j+1,k}^{\text{god}}), \\ \mathbf{U}_{i,j,k}^{\text{god}} &= \mathbf{L}_{z,\frac{1}{2}\Delta t}^{\text{god}}(\mathbf{L}_{x,\frac{1}{2}\Delta t}^{\text{god}}(\mathbf{U}_{i,j,k}^n)), \\ \mathbf{U}_{i,j+1,k}^{\text{god}} &= \mathbf{L}_{z,\frac{1}{2}\Delta t}^{\text{god}}(\mathbf{L}_{x,\frac{1}{2}\Delta t}^{\text{god}}(\mathbf{U}_{i,j+1,k}^n)) \end{aligned} \right\} \quad (16.102)$$

and numerical flux in the  $z$ -direction given by

$$\left. \begin{aligned} \mathbf{H}_{i,j,k+\frac{1}{2}} &= \mathbf{L}_{z, \frac{1}{2}\Delta t}^{waf}(\mathbf{U}_{i,j,k}^{god}, \mathbf{U}_{i,j,k+1}^{god}), \\ \mathbf{U}_{i,j,k}^{god} &= \mathbf{L}_{x, \frac{1}{2}\Delta t}^{god}(\mathbf{L}_{y, \frac{1}{2}\Delta t}^{god}(\mathbf{U}_{i,j,k}^n)), \\ \mathbf{U}_{i,j,k+1}^{god} &= \mathbf{L}_{x, \frac{1}{2}\Delta t}^{god}(\mathbf{L}_{y, \frac{1}{2}\Delta t}^{god}(\mathbf{U}_{i,j,k+1}^n)). \end{aligned} \right\} \quad (16.103)$$

**Exercise 16.24.** Show that the intercell fluxes (16.101) reproduces the intercell flux (16.87) for the linear advection equation (16.79) with  $a_1 > 0$ ,  $a_2 > 0$ ,  $a_3 > 0$ .

**Solution 16.25.** Left to the reader.

*Remark 16.26.* A first-order scheme for three dimensional nonlinear systems is obtained by replacing the WAF flux by the Godunov first order upwind flux in (16.101)–(16.103).

**Exercise 16.27.** Extend the Weighted Average State scheme with fluxes (16.99)–(16.100) to construct a finite volume scheme for three-dimensional non-linear systems. Verify that the corresponding fluxes reproduce (16.87) for the linear advection equation (16.79) with  $a_1 > 0$ ,  $a_2 > 0$ ,  $a_3 > 0$ .

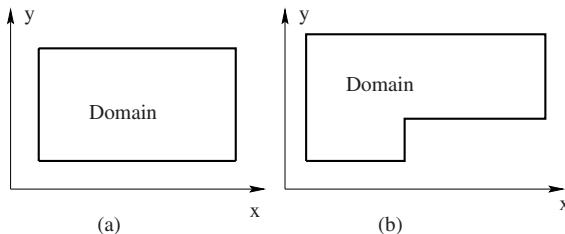
**Solution 16.28.** Left to the reader.

## 16.7 Non-Cartesian Geometries

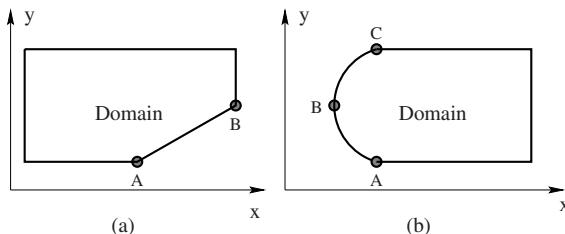
### 16.7.1 Introduction

A well defined fluid dynamic problem will include (a) a set of partial differential equations (PDEs) or a set of Integral Equations (b) the domain in which the governing equations are to be solved and (c) initial and boundary conditions. The numerical solution of this mathematical problem requires the discretisation of the domain via the generation of a grid or mesh and the discretisation of the equations on the mesh. The grid generation process discretises the continuous domain into a finite number of mesh points (finite difference interpretation) or into a finite number of volumes (finite volume interpretation). So far we have assumed that all domains are Cartesian, that is their boundaries are perfectly aligned with the Cartesian coordinate directions  $x, y, z$ . Fig. 16.10 shows two examples of Cartesian domains in two space dimensions. In practice, most domains *are not* Cartesian; Fig. 16.11 shows two simple examples of non-Cartesian domains in two dimensions. Having specified the domain and its boundaries, the process of determining the interior points or finite volumes is known as *grid generation*, or *mesh generation*. The generation of grids for Cartesian domains is exceedingly simple, particularly

for a rectangular domain such as that of Fig. 16.10 (a). For non-Cartesian domains such as those of Fig. 16.11, both the generation of a grid and the discretisation of the PDEs is much more involved. The topic of grid generation is a large area of study in its own right and is not pursued here. We assume that a mesh is available. Introductory material on grid generation is found in the textbook by Hoffmann [253]. Comprehensive treatments of the subject are found in the textbook edited by Thompson [487], the book by Thompson, Warsi and Mastin [486] and in the VKI Lectures by Weatherhill [577]; see also [485]. The purpose of this section is to introduce the reader to approaches



**Fig. 16.10.** Examples of two-dimensional Cartesian domains in  $x$ - $y$  space



**Fig. 16.11.** Examples of two-dimensional non-Cartesian domains in  $x$ - $y$  space (a) straight segment  $AB$  is not aligned with any of the Cartesian coordinates directions (b) arc  $ABC$  is not aligned with any of the Cartesian coordinates directions

for dealing with non-Cartesian domains. In particular, we give details of the finite volume approach, whereby one discretises the PDEs directly in physical space.

### 16.7.2 General Domains and Coordinate Transformation

The simplest approach for dealing with non-Cartesian domains such as those of Fig. 16.11 is to insist on using a Cartesian representation of the boundaries of the domain. In this case those portions of the boundary which are not aligned with any of the coordinate directions will end up with a *staircase* like representation. This simple procedure allows once more the use of

Cartesian meshes but the errors introduced at the boundaries are too large and therefore this approach is not recommended. An elegant method that retains the Cartesian mesh in the interior of the domain and treats in a special manner those computing cells that are *cut* by the true boundary is the so-called *Cartesian cut cell* method. Details on this approach are found in [114], [598], [47], [400], [589] and [590].

The most well-known approach for dealing with general domains is based on transforming the domain in *physical space* ( $t, x, y, z$ ) to a domain in *computational space* ( $\tau, \xi, \eta, \zeta$ ). The coordinate transformation maps the, irregular, domain in physical space to a perfectly regular computational domain, which is then discretised by a perfectly regular mesh as done for Cartesian domains. One defines the transformation via

$$\tau = t, \quad \xi = \xi(t, x, y, z), \quad \eta = \eta(t, x, y, z), \quad \zeta = \zeta(t, x, y, z). \quad (16.104)$$

By use of the chain rule, partial derivatives of any quantity  $\phi = \phi(t, x, y, z)$  read

$$\left. \begin{aligned} \phi_t &= \phi_\tau + \phi_\xi \xi_t + \phi_\eta \eta_t + \phi_\zeta \zeta_t, \\ \phi_x &= \phi_\xi \xi_x + \phi_\eta \eta_x + \phi_\zeta \zeta_x, \\ \phi_y &= \phi_\xi \xi_y + \phi_\eta \eta_y + \phi_\zeta \zeta_y, \\ \phi_z &= \phi_\xi \xi_z + \phi_\eta \eta_z + \phi_\zeta \zeta_z, \end{aligned} \right\} \quad (16.105)$$

so that PDEs defined in physical space in terms of the  $(t, x, y, z)$  coordinates are transformed in computational space in terms of the generalised coordinates  $(\tau, \xi, \eta, \zeta)$ .

*Example 16.29.* Suppose we want to solve the linear advection equation

$$u_t + a_1 u_x + a_2 u_y = 0, \quad (16.106)$$

where  $a_1$  and  $a_2$  are constants, in a non-Cartesian domain. Use of relations (16.105) reduces (16.106) in physical space to

$$u_\tau + \hat{a}_1 u_\xi + \hat{a}_2 u_\eta = 0 \quad (16.107)$$

in computational space. Note that this PDE has variable coefficients

$$\hat{a}_1 = \xi_t + a_1 \xi_x + a_2 \xi_y, \quad \hat{a}_2 = \eta_t + a_1 \eta_x + a_2 \eta_y \quad (16.108)$$

and is more complicated than the original PDE (16.106), which has constant coefficients. The variable coefficients  $\hat{a}_1$  and  $\hat{a}_2$  include the transformation derivatives  $\xi_x$ ,  $\xi_y$ ,  $\eta_x$  and  $\eta_y$ . These partial derivatives are called *metrics* and may be computed analytically from (16.104), or numerically, as we shall explain.

Next we transform the three-dimensional time dependent Euler equations, or any three-dimensional non-linear system of conservation laws

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_y + \mathbf{H}(\mathbf{U})_z = \mathbf{0}. \quad (16.109)$$

By using relations (16.105) one obtains the transformed system in generalised coordinates

$$\hat{\mathbf{U}}_\tau + \hat{\mathbf{F}}_\xi + \hat{\mathbf{G}}_\eta + \hat{\mathbf{H}}_\zeta = \mathbf{0} , \quad (16.110)$$

where

$$\left. \begin{aligned} \hat{\mathbf{U}} &= \frac{1}{J} \mathbf{U} , \\ \hat{\mathbf{F}} &= \frac{1}{J} [\xi_t \mathbf{U} + \xi_x \mathbf{F} + \xi_y \mathbf{G} + \xi_z \mathbf{H}] , \\ \hat{\mathbf{G}} &= \frac{1}{J} [\eta_t \mathbf{U} + \eta_x \mathbf{F} + \eta_y \mathbf{G} + \eta_z \mathbf{H}] , \\ \hat{\mathbf{H}} &= \frac{1}{J} [\zeta_t \mathbf{U} + \zeta_x \mathbf{F} + \zeta_y \mathbf{G} + \zeta_z \mathbf{H}] . \end{aligned} \right\} \quad (16.111)$$

with  $J$  to be specified.

*Remark 16.30.* Equations (16.110) are written in conservation form, just as were the original equations (16.109). For details see [253], [7] and [11].

## Computation of Metrics

One can relate differentials in the  $(t, x, y, z)$  system to differentials in the  $(\tau, \xi, \eta, \zeta)$  system so that the metrics can be found in terms of derivative of  $(t, x, y, z)$  with respect to  $(\tau, \xi, \eta, \zeta)$  in computational space, namely

$$\mathbf{A} = \mathbf{B}^{-1} , \quad (16.112)$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \xi_t & \xi_x & \xi_y & \xi_z \\ \eta_t & \eta_x & \eta_y & \eta_z \\ \zeta_t & \zeta_x & \zeta_y & \zeta_z \end{bmatrix} ; \quad \mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ x_\tau & x_\xi & x_\eta & x_\zeta \\ y_\tau & y_\xi & y_\eta & y_\zeta \\ z_\tau & z_\xi & z_\eta & z_\zeta \end{bmatrix} . \quad (16.113)$$

See [253] for details. The computation of derivatives in computational space is easily carried out numerically, as this is discretised by perfectly regular meshes. In solving for the components of  $\mathbf{A}$  via (16.112) one requires the computation of the inverse matrix of  $\mathbf{B}$ , which involves the expression

$$J = \frac{1}{\det(\mathbf{B})} , \quad (16.114)$$

where  $\det(\mathbf{B})$  is the determinant of  $\mathbf{B}$  and  $J$  is called the *Jacobian of the transformation*. Note that this is different from the definition of the Jacobian matrix associated with systems of hyperbolic conservation laws introduced in Chap. 2. Here the Jacobian is a scalar quantity given by

$$J = \frac{1}{x_\xi(y_\eta z_\zeta - y_\zeta z_\eta) - x_\eta(y_\xi z_\zeta - y_\zeta z_\xi) + x_\zeta(y_\xi z_\eta - y_\eta z_\xi)} . \quad (16.115)$$

In the next section we study the finite volume approach, whereby the equations are directly discretised in physical space.

### 16.7.3 The Finite Volume Method for Non-Cartesian Domains

The Finite Volume Method is another approach for dealing with general non-Cartesian domains. In this approach the governing conservation laws, expressed in integral form, are discretised *directly* in physical space and one still works with reference to a *Cartesian frame*. Background reading on the finite volume method is found in [251], [570]. Here we present a brief description of the method for two and three dimensional problems. First recall that the integral form of the three-dimensional conservation laws (16.109) is

$$\frac{d}{dt} \int \int \int_V \mathbf{U} dV + \int \int_A \mathcal{H} \cdot \mathbf{n} dA = \mathbf{0}, \quad (16.116)$$

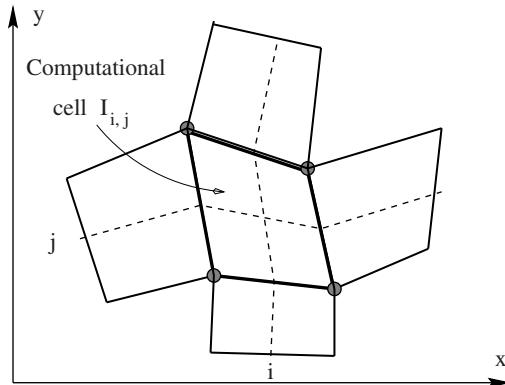
where  $V$  is a *control volume*,  $A$  is the boundary of  $V$ ,  $\mathcal{H} = (\mathbf{F}, \mathbf{G}, \mathbf{H})$  is the tensor of fluxes,  $\mathbf{n}$  is the outward unit vector normal to the surface  $A$ ,  $dA$  is an area element and  $\mathcal{H} \cdot \mathbf{n} dA$  is the flux component normal to the boundary  $A$ . See Sect. 1.5 of Chap. 1 for details on the derivation of the integral form of the equations. The conservation laws (16.116) state that the time-rate of change of  $\mathbf{U}$  inside the volume  $V$  depends only on the *total flux* through the surface  $A$ , the boundary of the control volume  $V$ . In the finite volume approach, the integral form of the conservation laws is enforced on each control volume that results from discretising the domain in physical space into a finite number of finite volumes or computing cells. The conservation equations are thus satisfied at the discrete level, which is a distinguishing property of the Finite Volume Method.

#### The Two-Dimensional Case

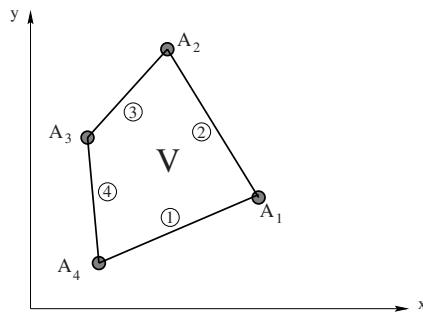
Consider some two-dimensional domain in  $x-y$  space and assume this has been discretised into a finite number of computing cells by some grid generation technique. Part of a simple structured mesh is shown in Fig. 16.12, where all computing cells may be identified by the indices  $i$  and  $j$ , just as in Cartesian meshes assumed previously. Consider now any computing cell from the discretised domain and regard it as the control volume  $V$  in the integral form of the equations (16.116). For generality assume that the boundary  $A$  of  $V$  is the union of  $N$  straight segments  $A_s A_{s+1}$ , where  $A_{N+1} \equiv A_1$ . Fig. 16.13 shows the special case of a quadrilateral finite volume,  $N = 4$ . We have assumed that a grid has been generated and that the  $(x_s, y_s)$  coordinates of the vertices  $A_s$  are known. The *total flux* through the boundaries in (16.116) may now be written as

$$\int \int_A \mathcal{H} \cdot \mathbf{n} dA = \sum_{s=1}^N \int_{A_s}^{A_{s+1}} \mathcal{H} \cdot \mathbf{n}_s dA, \quad (16.117)$$

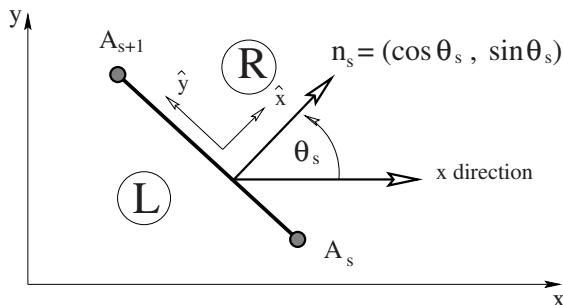
where  $s$  denotes the intercell boundary  $A_s A_{s+1}$ . We consistently adhere to the convention of transversing the boundary so that the interior of the volume



**Fig. 16.12.** Possible grid configuration for two-dimensional domain in  $x$ - $y$  space



**Fig. 16.13.** Example of quadrilateral finite volume  $V$  in two-dimensional domain in  $x$ - $y$  space



**Fig. 16.14.** General intercell boundary  $s$  of finite volume  $V$  in two-dimensional domain in  $x$ - $y$  space.

always lies on the left hand side of the boundary. Let us now consider any intercell boundary  $s$  as shown in Fig. 16.14, where  $L$  (left) denotes the interior side of the control volume  $V$  and  $R$  (right) denotes a state exterior to  $V$  and adjacent to side  $s$ . The first problem is to determine the direction of the outward unit vector  $\mathbf{n}_s$  normal to side  $s$ ; this is needed to evaluate the integral terms in (16.117). It is convenient to choose the  $x$ -direction as the *reference direction* and to define the angle  $\theta_s$  as the angle formed by the  $x$ -direction and the normal vector  $\mathbf{n}_s$ ; see Fig. 16.14. The components of  $\mathbf{n}_s$  can be found in terms of the angle  $\theta_s$  and it is easy to see that

$$\mathbf{n}_s = (\cos\theta_s, \sin\theta_s) . \quad (16.118)$$

Then, the total flux in (16.117) is

$$\sum_{s=1}^N \int_{A_s}^{A_{s+1}} \mathcal{H} \cdot \mathbf{n}_s \, dA = \sum_{s=1}^N \int_{A_s}^{A_{s+1}} [\cos\theta_s \mathbf{F}(\mathbf{U}) + \sin\theta_s \mathbf{G}(\mathbf{U})] \, dA . \quad (16.119)$$

By interpreting the first integral term in equations (16.116) as the time-rate of change of the average of  $\mathbf{U}$  inside volume  $V$  we may write

$$\frac{d}{dt} \int \int \int_V \mathbf{U} \, dV = |V| \frac{d}{dt} \mathbf{U} , \quad (16.120)$$

where  $|V|$  denotes the volume of  $V$ , the area of  $V$  in the two-dimensional case. Then, by substituting (16.119) and (16.120) into (16.116) we obtain

$$\frac{d}{dt} \mathbf{U} = -\frac{1}{|V|} \sum_{s=1}^N \int_{A_s}^{A_{s+1}} [\cos\theta_s \mathbf{F}(\mathbf{U}) + \sin\theta_s \mathbf{G}(\mathbf{U})] \, dA . \quad (16.121)$$

This finite volume formula leads to numerical schemes, once approximations to the intercell fluxes for each side  $s$  have been made and a time discretisation scheme has been chosen. In devising intercell flux approximations one requires to be more specific as to the locations of the discrete values of  $\mathbf{U}$ . Various flux approximations are discussed in Chap. 8 of [251]. Here we assume that the discrete values of the variable  $\mathbf{U}$  are cell averages within the volumes and assign them to the centres of the cells. This is known as the *cell centred* finite volume method.

One way of constructing intercell flux approximations arises from exploiting the *rotational invariance* of the governing equations. For the Euler equations see Sect. 3.2.1 of Chap. 3. For the two-dimensional case we have

$$\cos\theta_s \mathbf{F}(\mathbf{U}) + \sin\theta_s \mathbf{G}(\mathbf{U}) = \mathbf{T}_s^{-1} \mathbf{F}(\mathbf{T}_s \mathbf{U}) , \quad (16.122)$$

where  $\mathbf{T}_s \equiv \mathbf{T}(\theta_s)$  is the rotation matrix and  $\mathbf{T}_s^{-1}$  is its inverse, namely

$$\mathbf{T}(\theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} , \quad \mathbf{T}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} . \quad (16.123)$$

We note that  $\hat{\mathbf{U}} \equiv \mathbf{T}_s \mathbf{U}$  is the vector of *rotated* conserved variables obtained by applying the rotation matrix  $\mathbf{T}_s$  to the original vector of conserved variables  $\mathbf{U}$  and is aligned with the new rotated Cartesian frame  $(\hat{x}, \hat{y})$  shown in Fig. 16.14. Here the coordinate  $\hat{x}$  is the normal direction (normal to the intercell boundary) and  $\hat{y}$  is the tangential direction (parallel to the intercell boundary). Then equations (16.121) become

$$\frac{d}{dt} \mathbf{U} = -\frac{1}{|V|} \sum_{s=1}^N \int_{A_s}^{A_{s+1}} \mathbf{T}_s^{-1} \mathbf{F}(\mathbf{T}_s \mathbf{U}) dA . \quad (16.124)$$

So far, apart from interpretation (16.120), no numerical approximations have been made. In order to define numerical fluxes across the intercell boundaries  $s$  we use the equations in the rotated frame  $(\hat{x}, \hat{y})$ . They become an *augmented* one-dimensional system

$$\hat{\mathbf{U}}_t + \hat{\mathbf{F}}_{\hat{x}} = \mathbf{0} , \quad (16.125)$$

where  $\hat{\mathbf{F}} = \mathbf{F}(\hat{\mathbf{U}})$ . This *augmented* one-dimensional system is identical to the split multidimensional systems in Cartesian coordinates considered in Chap. 3, for which numerical fluxes were constructed in Chaps. 6 to 12 and 14.

Each term in the summation (16.124) is now approximated as

$$\int_{A_s}^{A_{s+1}} \mathbf{T}_s^{-1} \mathbf{F}(\mathbf{T}_s \mathbf{U}) dA \approx \mathcal{L}_s \mathbf{T}_s^{-1} \hat{\mathbf{F}}_s , \quad (16.126)$$

where  $\mathcal{L}_s$  is the length of segment  $A_s A_{s+1}$  and  $\hat{\mathbf{F}}_s$  is the intercell flux corresponding to the augmented one-dimensional system (16.125). Assuming numerical fluxes  $\hat{\mathbf{F}}_s$  have been constructed we then have the *semidiscrete* conservative scheme

$$\frac{d}{dt} \mathbf{U} = -\frac{1}{|V|} \sum_{s=1}^N \mathcal{L}_s \mathbf{T}_s^{-1} \hat{\mathbf{F}}_s , \quad (16.127)$$

where the time discretisation is still open to choice. See Sec. 13.4.7 of Chap. 13.

In a computational setup as depicted in Fig. 16.12 the volume  $V$  is a general computing cell labelled  $I_{i,j}$ . Let us now replace the time derivative by a forward in time approximation; we then obtain the fully discrete scheme

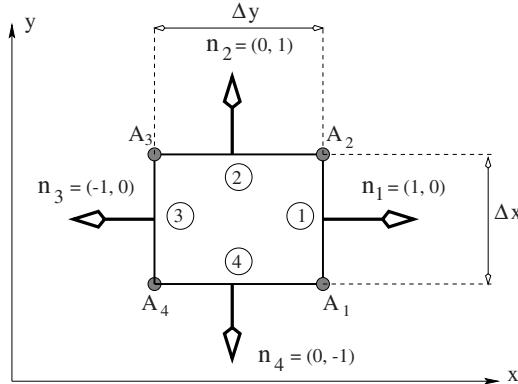
$$\mathbf{U}_{i,j}^{n+1} = \mathbf{U}_{i,j}^n - \frac{\Delta t}{|I_{i,j}|} \sum_{s=1}^N \mathcal{L}_s \mathbf{T}_s^{-1} \hat{\mathbf{F}}_s , \quad (16.128)$$

where  $|I_{i,j}|$  is the area of cell  $I_{i,j}$ .

**Exercise 16.31.** Show that if  $I_{i,j}$ ,  $\forall i, j$ , is a Cartesian cell of area  $\Delta x \times \Delta y$  as shown in Fig. 16.15, then the finite volume formula (16.128) reproduces the finite volume formula

$$\mathbf{U}_{i,j}^{n+1} = \mathbf{U}_{i,j}^n + \frac{\Delta t}{\Delta x} [\mathbf{F}_{i-\frac{1}{2},j} - \mathbf{F}_{i+\frac{1}{2},j}] + \frac{\Delta t}{\Delta y} [\mathbf{G}_{i,j-\frac{1}{2}} - \mathbf{G}_{i,j+\frac{1}{2}}] \quad (16.129)$$

studied in Sect. 16.4.



**Fig. 16.15.** Cartesian cell  $I_{i,j}$  in two-dimensional domain in  $x-y$  space.

**Solution 16.32.** The outward unit normals for sides 1 to 4 are  $\mathbf{n}_1 = (1, 0)$ ,  $\mathbf{n}_2 = (0, 1)$ ,  $\mathbf{n}_3 = (-1, 0)$ , and  $\mathbf{n}_4 = (0, -1)$ . The corresponding angles are  $\theta_1 = 0$ ,  $\theta_2 = \frac{1}{2}\pi$ ,  $\theta_3 = \pi$ ,  $\theta_4 = \frac{3}{2}\pi$ . The fluxes may be computed from (16.126), or more directly from (16.117); they are

$$\left. \begin{aligned} \int_{A_1}^{A_2} (\mathbf{F}, \mathbf{G}) \cdot \mathbf{n}_1 dA &= \Delta y \times \mathbf{F}_{12} , \\ \int_{A_2}^{A_3} (\mathbf{F}, \mathbf{G}) \cdot \mathbf{n}_2 dA &= \Delta x \times \mathbf{G}_{23} , \\ \int_{A_3}^{A_4} (\mathbf{F}, \mathbf{G}) \cdot \mathbf{n}_3 dA &= -\Delta y \times \mathbf{F}_{34} , \\ \int_{A_4}^{A_1} (\mathbf{F}, \mathbf{G}) \cdot \mathbf{n}_4 dA &= -\Delta x \times \mathbf{G}_{41} . \end{aligned} \right\} \quad (16.130)$$

By defining  $\mathbf{F}_{12} \equiv \mathbf{F}_{i+\frac{1}{2},j}$ ,  $\mathbf{F}_{34} \equiv \mathbf{F}_{i-\frac{1}{2},j}$ ,  $\mathbf{G}_{23} \equiv \mathbf{F}_{i,j+\frac{1}{2}}$ ,  $\mathbf{G}_{41} \equiv \mathbf{G}_{i,j-\frac{1}{2}}$ , we obtain the the desired result.

A possible way of constructing intercell numerical fluxes for (16.128) is following the Godunov approach, studied in Chap. 6 as used in conjunction with the exact Riemann solver, and in Chaps. 8 to 12 using approximate Riemann solvers. To obtain a Godunov first-order upwind flux we solve the Riemann problem for (16.125) with initial data

$$\hat{\mathbf{U}}(\hat{x}, 0) = \begin{cases} \hat{\mathbf{U}}_L = \mathbf{T}_s(\mathbf{U}_L) & \text{if } \hat{x} < 0 , \\ \hat{\mathbf{U}}_R = \mathbf{T}_s(\mathbf{U}_R) & \text{if } \hat{x} > 0 . \end{cases} \quad (16.131)$$

Here  $\mathbf{U}_L$  is the state in the interior of the finite volume  $V$  and  $\mathbf{U}_R$  is the state outside  $V$  separated from  $\mathbf{U}_L$  by the intercell boundary  $A_s A_{s+1}$ . See Fig. 16.14.

### Computation of Lengths, Normals and Areas

Here we give formulae for the lengths  $\mathcal{L}_s$ , the components of the outward unit normals  $\mathbf{n}_s$  and the areas of co-planar quadrilateral finite volumes. These quantities are needed in the conservative formula (16.128). Consider the quadrilateral  $V$  of Fig. 16.13 and assume the coordinates  $(x_s, y_s)$  of the points  $A_s$  are known. We first define

$$\Delta x_s \equiv x_{s+1} - x_s, \quad \Delta y_s \equiv y_{s+1} - y_s. \quad (16.132)$$

Then, the length  $\mathcal{L}_s$  of side  $s$  is given by

$$\Delta s = \sqrt{\Delta x_s^2 + \Delta y_s^2}. \quad (16.133)$$

The components  $\cos\theta_s$  and  $\sin\theta_s$  of the outward unit normal  $\mathbf{n}_s$  are

$$\cos\theta_s = \frac{\Delta y_s}{\Delta s}, \quad \sin\theta_s = -\frac{\Delta x_s}{\Delta s}. \quad (16.134)$$

Finally, the area of the co-planar quadrilateral  $V$  is given by

$$|V| = \frac{1}{2}|(x_3 - x_1) \times (y_4 - y_2) - (y_3 - y_1) \times (x_4 - x_2)|. \quad (16.135)$$

### Examples of Finite Volume Schemes

Consider sufficiently smooth structured grids formed by quadrilateral volumes as shown in Fig. 16.12. Then formula (16.128) contains four flux contributions. The most natural way of updating the solution is by applying (16.128) with all flux contributions included in a single step. We term this the *unsplit finite volume method* and is entirely consistent with the unsplit finite volume schemes for Cartesian meshes studied in Sect. 16.4. One may also reinterpret (16.128) as a *split finite volume scheme*, whereby for instance, flux terms associated with the  $i$  index are included in a predictor step and flux terms associated with the  $j$  index are included in the corrector step. The practical implementation of the split finite volume schemes is entirely analogous to the split schemes studied in Sects. 16.2 and 16.3.

First-order methods of the Godunov type are easily constructed. The intercell numerical flux depends only on the two rotated states  $(\hat{\mathbf{U}}_L, \hat{\mathbf{U}}_R)$ , the length  $\mathcal{L}_s$  and the normal  $\mathbf{n}_s$ . The following is a possible algorithm

- (I) For each intercell boundary  $s$  rotate the data  $(\mathbf{U}_L, \mathbf{U}_R)$  to obtain  $\hat{\mathbf{U}}_L = \mathbf{T}_s \mathbf{U}_L$ ,  $\hat{\mathbf{U}}_R = \mathbf{T}_s \mathbf{U}_R$ .

- (II) Solve the local Riemann problem with rotated data  $(\hat{\mathbf{U}}_L, \hat{\mathbf{U}}_R)$  to obtain the intercell flux  $\hat{\mathbf{F}}_s$ . Here one uses the exact Riemann solver, see Chaps. 6, or approximate Riemann solvers, see Chaps. 8 to 12.
- (III) Rotate  $\hat{\mathbf{F}}_s$  back by multiplying it by  $\mathbf{T}_s^{-1}$  to obtain the intercell flux  $\mathbf{T}_s^{-1}\hat{\mathbf{F}}_s$  for side  $s$ .
- (IV) Compute lengths  $\mathcal{L}_s$ , normals  $\mathbf{n}_s$  and areas  $|I_{i,j}|$  to complete the intercell flux evaluation.
- (V) Update the solution according to a split or unsplit scheme.

Recall that for Cartesian meshes an unsplit Godunov method has linearised stability limit  $\frac{1}{2}$ , whereas a split version has limit 1. The WAF-type first order unsplit finite volume scheme of Sect. 16.6.1 also has stability limit 1; this scheme can easily be implemented in non-Cartesian quadrilaterals.

The second-order TVD schemes presented in Chaps. 14, see also Sect. 16.2 and 16.3 of this chapter, can also be implemented for non-Cartesian quadrilateral meshes following the methodology for the Godunov first-order upwind method outlined above. Details on WAF-type methods are found in [456], [457] and [64].

### The Three-Dimensional Case

Three-dimensional finite volume schemes are easily constructed. Here one requires the evaluation of the area of of non-coplanar intercell boundaries and as well as volumes. For details see [251]. The rotational invariance property of the three-dimensional Euler equations, proved in Sect. 3.2.2 of Chap. 3, can also be exploited to derive finite volume formulae analogous to (16.128), with corresponding Godunov-type numerical fluxes, see [64] for example.

## Multidimensional Test Problems

This chapter is concerned with tests for assessing numerical solutions to multidimensional problems. The assessment of the numerical methods to be used in practical computations, prior to their actual application, is of considerable importance and cannot be emphasised enough. There are four classes of test problems that can be used, namely (A) tests with exact solution, (B) tests with reliable numerical solution to equivalent one-dimensional equations obtained under the assumption of symmetry for instance, (C) tests for which other numerical solutions are available and (D) tests for which experimental results are available. In the first three categories of test problems one solves the same or equivalent governing partial differential equations and thus one seeks *complete agreement* in the comparisons. Care is required in class (D) when experimental results are used. The governing PDEs might themselves not be an accurate description of the physical problem being solved. Typical questions to be asked are: will viscosity and heat conduction be important in the problem, will the equation of state be a correct description of the thermodynamics, will turbulence be important, etc. If one can isolate these effects, or account for their influence, then comparison between numerical and experimental results is useful in assessing the performance of the numerical methods. If the numerical solution is reliable, then one would expect the comparison with experimental results to be a way of verifying the validity of the governing equations as a suitable model of the physics. A useful reference here is the report by Albone [3].

Exact solutions for one-dimensional problems can be used to check two and three dimensional programs. For the two-dimensional case one can, for instance, use the exact Riemann solver of Chap. 4 with the initial data states separated by a straight line at an angle to the mesh; this includes the trivial cases of one-dimensional flow along coordinates directions. The same can be done for three-dimensional programs. In this chapter we consider test problems in categories (B), (C) and (D) above.

## 17.1 Explosions and Implosions

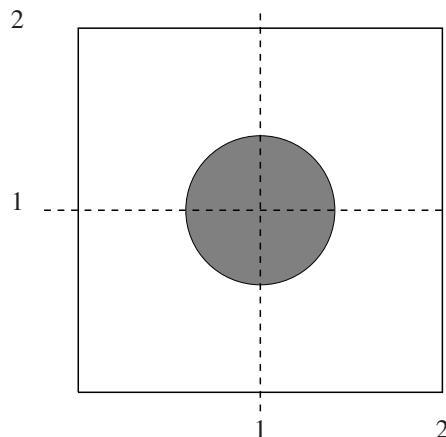
Here we consider test problems for the two and three dimensional Euler equations for ideal gases with  $\gamma = 1.4$ . The geometry and initial data for the problems are such that cylindrical and spherical symmetry can be enforced. The multidimensional Euler equations may then be simplified to the one-dimensional inhomogeneous system

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_r = \mathbf{S}(\mathbf{U}), \quad (17.1)$$

where

$$\mathbf{U} = \begin{bmatrix} \rho \\ \rho u \\ E \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ u(E + p) \end{bmatrix}, \quad \mathbf{S} = -\frac{\alpha}{r} \begin{bmatrix} \rho u \\ \rho u^2 \\ u(E + p) \end{bmatrix}. \quad (17.2)$$

See Sect. 1.6.2 of Chap. 1 for details. Here  $r$  is the radial direction,  $u$  is the radial velocity and  $\alpha$  is a parameter. For  $\alpha = 0$  we reproduce plain one-dimensional flow. For  $\alpha = 1$  we have cylindrical symmetry and equations (17.1)–(17.2) are equivalent to the two-dimensional Euler equations. For  $\alpha = 2$  we have spherical symmetry and (17.1)–(17.2) are equivalent to the three-dimensional Euler equations. The one-dimensional equations (17.1)–(17.2) with a geometric source term may be solved with a reliable one-dimensional method on a very fine mesh to provide very accurate numerical solutions to compare with the numerical solution of the full two and three dimensional Euler equations. These test problems belong to category (B) above.



**Fig. 17.1.** Initial conditions for cylindrical explosion. Dark circular zone has high pressure and high density

### 17.1.1 Explosion Test in Two-Space Dimensions

The two-dimensional Euler equations, see Chap. 3, are solved on the square domain  $2.0 \times 2.0$  in the  $x-y$  plane. The initial conditions consist of the region inside of a circle of radius  $R = 0.4$  centred at  $(1, 1)$  and the region outside the circle, see Fig. 17.1. The flow variables take constant values in each of these regions and are joined by a circular discontinuity at time  $t = 0$ . The two constant states for the two-dimensional Euler equations are chosen to be

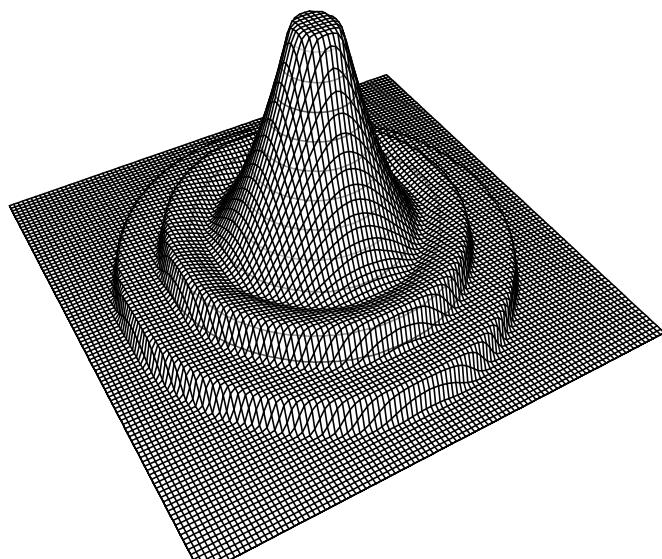
$$\left. \begin{array}{l} \rho_{ins} = 1.0, \quad \rho_{out} = 0.125, \\ u_{ins} = 0.0, \quad u_{out} = 0.0, \\ v_{ins} = 0.0, \quad v_{out} = 0.0, \\ p_{ins} = 1.0, \quad p_{out} = 0.1. \end{array} \right\} \quad (17.3)$$

Subscripts *ins* and *out* denote values inside and outside the circle respectively. This test problem is like a two-dimensional extension of the shock-tube problem used for assessing one-dimensional solutions in previous chapters.

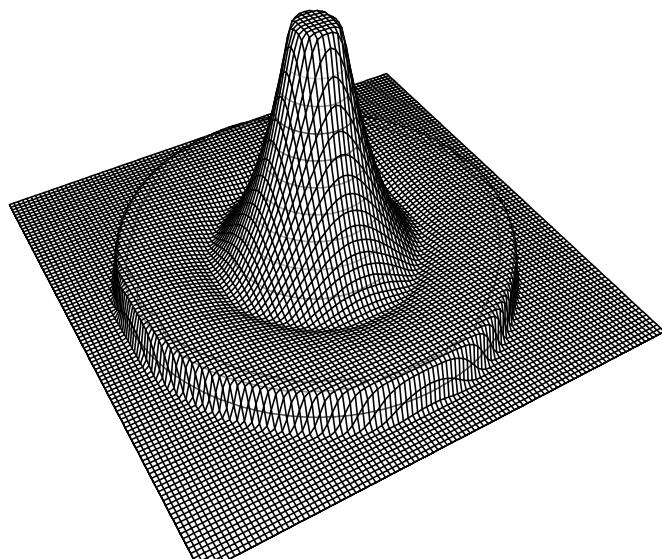
We solve the full two-dimensional Euler equations by the WAF method presented in Chap. 14 in conjunction with space splitting, see Sect. 16.2 of Chap. 16; the Riemann solver used is HLLC presented in Chap. 10. The limiter function used is VANLEER, see 13.7 and 13.8 of Chap. 13. The CFL coefficient used is  $C_{cfl} = 0.9$  and the mesh is of  $101 \times 101$  computing cells. We remark that in initialising the implosion/explosion problems, we modify the initial data on quadrilateral cells cutting the initial discontinuity, by assigning modified area-weighted values to the appropriate cells at the initial time  $t = 0$ . This procedure avoids the formation of small amplitude waves created at early times by the staircase configuration of the data; such waves are not spurious, they result from the data as given and the numerical method does its best to resolve them.

Fig. 17.2 shows the density distribution as a function of  $x$  and  $y$  at the output time  $t = 0.25$ . The solution exhibits a circular shock wave travelling away from the centre, a circular contact surface travelling in the same direction and a circular rarefaction travelling towards the origin  $(1, 1)$ . Fig. 17.3 shows the corresponding pressure distribution at the same time. As expected, the pressure is continuous across the contact surface. We remark that as time evolves, a complex wave pattern emerges. Some salient features of the solution are these. The circular shock wave travels outwards, becoming weaker as time evolves. The contact surface follows the shock becoming weaker also; at some point in time the contact comes to rest and then travels inwards. The rarefaction travelling towards the centre reflects, as a rarefaction, and over expands the flow so as to create an inward travelling shock wave; this circular shock wave implodes into the origin, reflects and travels outwards colliding with the contact surface. And so on.

The one dimensional inhomogeneous equations (17.1), (17.2) for the equivalent problem are solved numerically on a domain  $[0, 1]$  on the radial line  $r$ ,

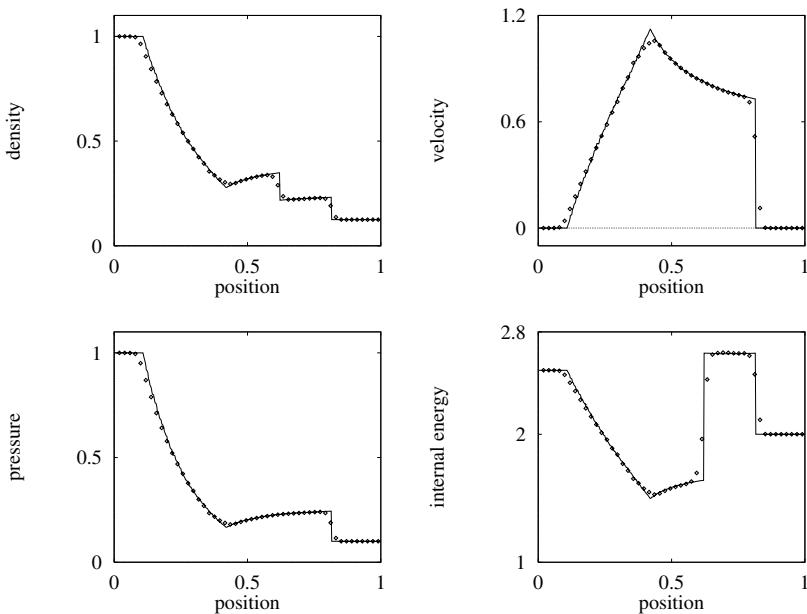


**Fig. 17.2.** Cylindrical explosion. Density distribution at time  $t = 0.25$ . Solution exhibits circular shock, circular contact and circular rarefaction



**Fig. 17.3.** Cylindrical explosion. Pressure distribution at time  $t = 0.25$ . Solution exhibits circular shock and circular rarefaction. Compare with Fig. 17.2

with equivalent shock–tube like data. The numerical method used is the Random Choice Method (RCM) described in Chap. 7, together with the splitting techniques of Chap. 15 to account for the geometric source term. We used  $M = 1000$  computing cells and a CFL coefficient  $C_{cfl} = 0.4$ . Recall that the RCM resolves discontinuities as true discontinuities, the only errors being those of the position of the waves. Under these circumstances we may regard the RCM solution as the *exact solution*.



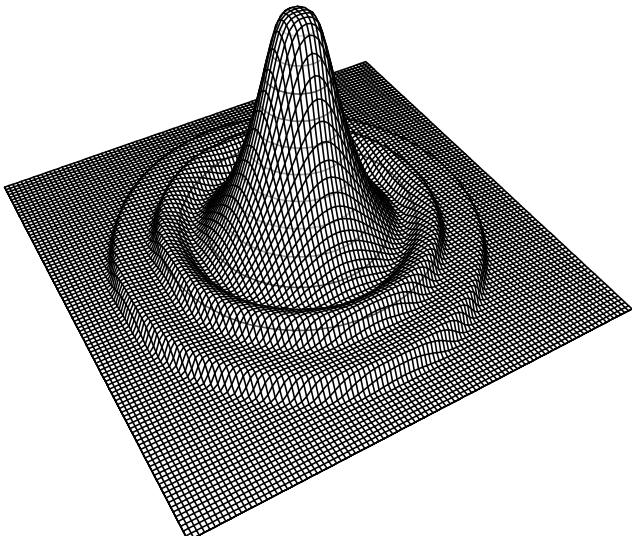
**Fig. 17.4.** Comparison between the two-dimensional WAF solution (symbol) and the one-dimensional radial RCM solution (line) at time=0.25

Fig. 17.4 shows a comparison between the one-dimensional radial solution (line) and the two-dimensional solution (symbols) along the radial line that is coincident with the  $x$ -axis. Comparisons along other radial directions give virtually identical results. This is confirmed by the very symmetric character of the numerical solution shown in Figs. 17.2 and 17.3. The shock wave is resolved with two mesh points and the contact surface is resolved with two to three mesh points. It is worth remarking that, at least for this test problem, the resolution of discontinuities that travel in *all directions* is essentially the same as that achieved in one-dimensional problems; see results of Chap. 14.

The initial data of the type (17.3) may be reversed so as to create an implosion problem, with *shock focussing* taking place as part of the solution. This can also be used to test programs and methods.

### 17.1.2 Explosion Test in Three Space Dimensions

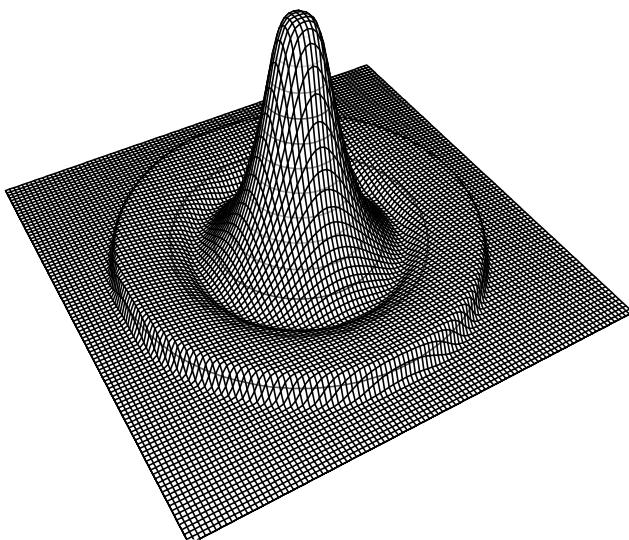
The three-dimensional Euler equations, see Chap. 3, are solved on a cube  $2.0 \times 2.0 \times 2.0$  in  $x-y-z$  space. The initial conditions consist of the region inside of a sphere with radius  $R = 0.4$  centred at  $(1, 1, 1)$  and a region outside the sphere. The flow variables take constant values in each of these regions and are joined by a spherical discontinuity at time  $t = 0$ . The data values are the same as those for the two-dimensional case in (17.3), with the  $w$  velocity component initialised as  $w_{ins} = w_{out} = 0$ . This test problem is like a three-dimensional extension of the shock-tube problem.



**Fig. 17.5.** Spherical explosion. Density distribution on plane  $z = 0$  at time  $t = 0.25$ . Solution exhibits a spherical shock, a spherical contact and a spherical rarefaction

Fig. 17.5 shows the density distribution as a function of  $x$  and  $y$  on the plane  $z = 0$  at the output time  $t = 0.25$ . The solution exhibits a spherical shock wave, a spherical contact surface travelling in the same direction and a spherical rarefaction travelling towards the origin  $(1, 1, 1)$ . Fig. 17.6 shows the corresponding pressure distribution; the pressure is continuous across the contact surface.

The one-dimensional equations (17.1)–(17.2) for the equivalent problem are solved on a domain  $[0, 1]$  on the radial line  $r$  with equivalent shock-tube like data by the Random Choice Method, with  $M = 1000$  computing cells and a CFL coefficient of  $C_{cfl} = 0.4$ .



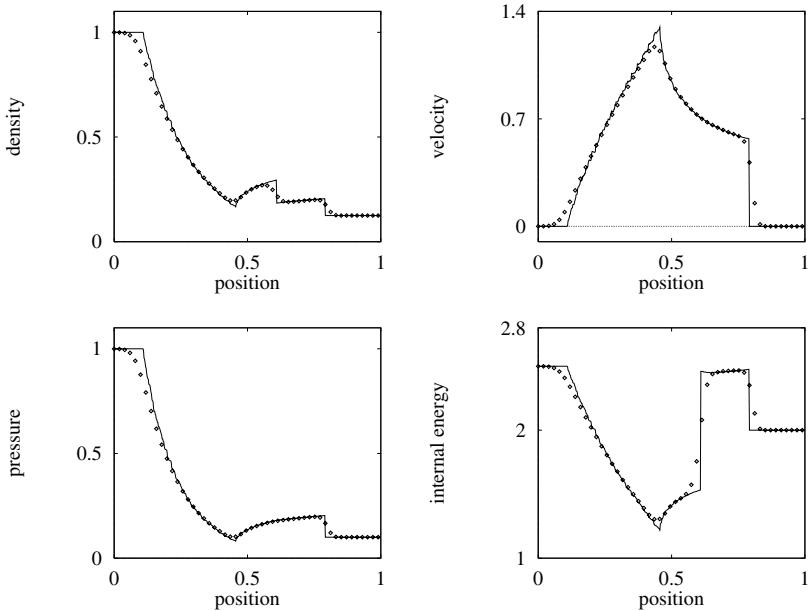
**Fig. 17.6.** Spherical explosion. Pressure distribution on plane  $z = 0$  at time  $t = 0.25$ . Solution exhibits a spherical shock and a spherical rarefaction. Compare with Fig. 17.5

Fig. 17.7 shows a comparison between the one-dimensional radial solution (line) and the three-dimensional solution (symbols) along the radial line that is coincident with the  $x$ -axis. Agreement is good and, as for the two-dimensional equations, the resolution of discontinuities in three space dimensions is as good as in one-dimensional problems.

## 17.2 Shock Wave Reflection from a Wedge

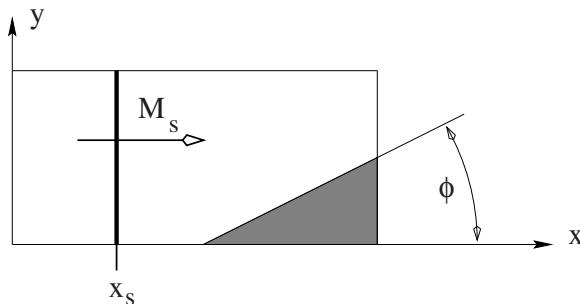
The subject of shock waves is large and embodies a substantial international community of scientists. *The International Symposium on Shock Waves* is a series of bi-annual meetings devoted exclusively to shock waves. A good source of information are the proceedings of these meetings. See for example Bershader and Hanson [52]; Grönig [228]; Takayama [478]; Brun and Dumitrescu [83]; and Jagadeesh, Arunan and Reddy [266].

Shock wave phenomena of considerable physical interest are those arising from the reflection of a shock wave from a wedge placed at an angle to the incident shock wave direction; see the sketch of Fig. 17.8. At the initial time a shock wave of shock Mach number  $M_S$ , perfectly perpendicular to the  $x$ -direction is placed at a position  $x_S$ . The shock travels in the  $x$ -direction and encounters a wedge that makes an angle  $\phi$  with the shock propagation



**Fig. 17.7.** Spherical explosion. Comparison between the three-dimensional WAF solution (symbol) and the one-dimensional radial RCM solution (line) at time=0.25

direction. The shock wave reflection pattern that emerges may fall into one of four possible categories. These are called: (a) regular reflection, (b) single Mach reflection, (c) complex Mach reflection and (d) double Mach reflection. The occurrence of a particular type of reflection depends on the shock Mach number  $M_S$  and the wedge angle  $\phi$ . A comprehensive study of these wave patterns is found in the book by Ben-Dor [42]. Informative references, amongst many others, are those of Heilig [246], Heilig and Reichenbach [246], Dewey [158] and Takayama [477].



**Fig. 17.8.** Sketch of initial conditions for shock wave reflection from a wedge.

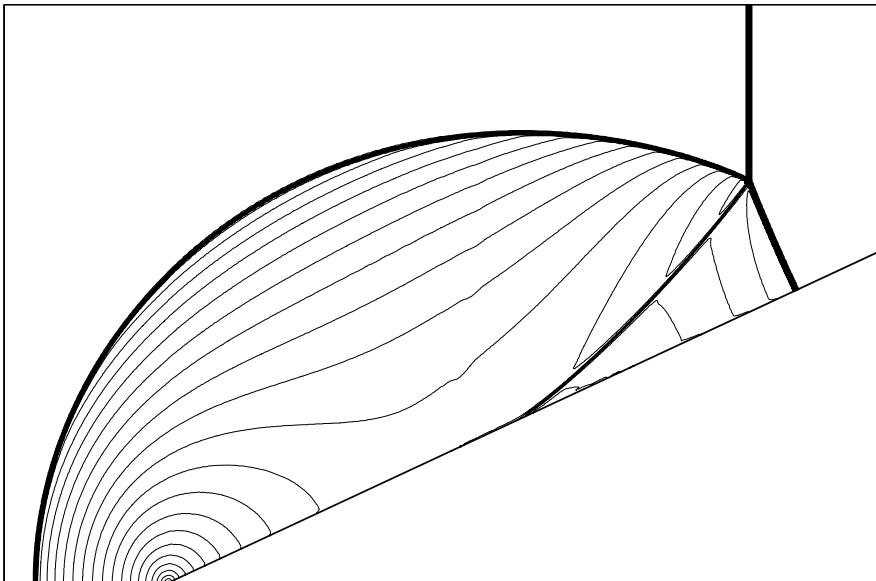
Two very useful reference are the UTIAS report by Glaz, Colella, Glass and Deschambault [210] and the paper by the same authors [211]. Many experimental and numerical results for this type of problems are presented and discussed in these works.

Here we consider the case of Mach number  $M_S = 1.7$  and  $\phi = 25$  degrees for which Fig. 17.9 shows experimental and computational results. The top frame shows the experimental result (Courtesy of Professor K. Takayama, Shock Wave Research Center, Sendai, Japan) and the bottom frame shows the computed result for density. The computational results were obtained on a rectangular domain of  $25.0 \times 16.5$  on the  $x-y$  plane, with the apex of the wedge placed at  $x = 4.69$  and the initial shock placed at  $x = 4.0$ . The conditions ahead of the shock are ambient conditions, with  $\rho_a = 1.225 \text{ kg/m}^3$ ,  $p_a = 1.01325 \times 10^5 \text{ Pa}$ ,  $u_a = 0$ ,  $v_a = 0$ . Given conditions ahead of the shock and the shock Mach number  $M_S$ , the state behind the shock is obtained from the *Rankine–Hugoniot conditions*, see Sect. 3.1.3 of Chap. 3. *Start up errors* appear in the numerical solution, which may affect the fine details of the solution. A common practice [211], [250] is to run the problem until the *numerical shock* has been established and then reset the initial conditions connecting the states ahead and behind the shock, not through the Rankine–Hugoniot conditions, but through the numerical shock. For this purpose one may simply use a one-dimensional code to find the numerical shock profile and give this as the initial conditions for the problem, or one may run the two-dimensional code and propagate the shock for some distance and then reset the state behind the shock. In the computations shown, we use the latter option and allow the shock to travel a distance of 90 % of that between its initial position and that of the apex of the wedge.

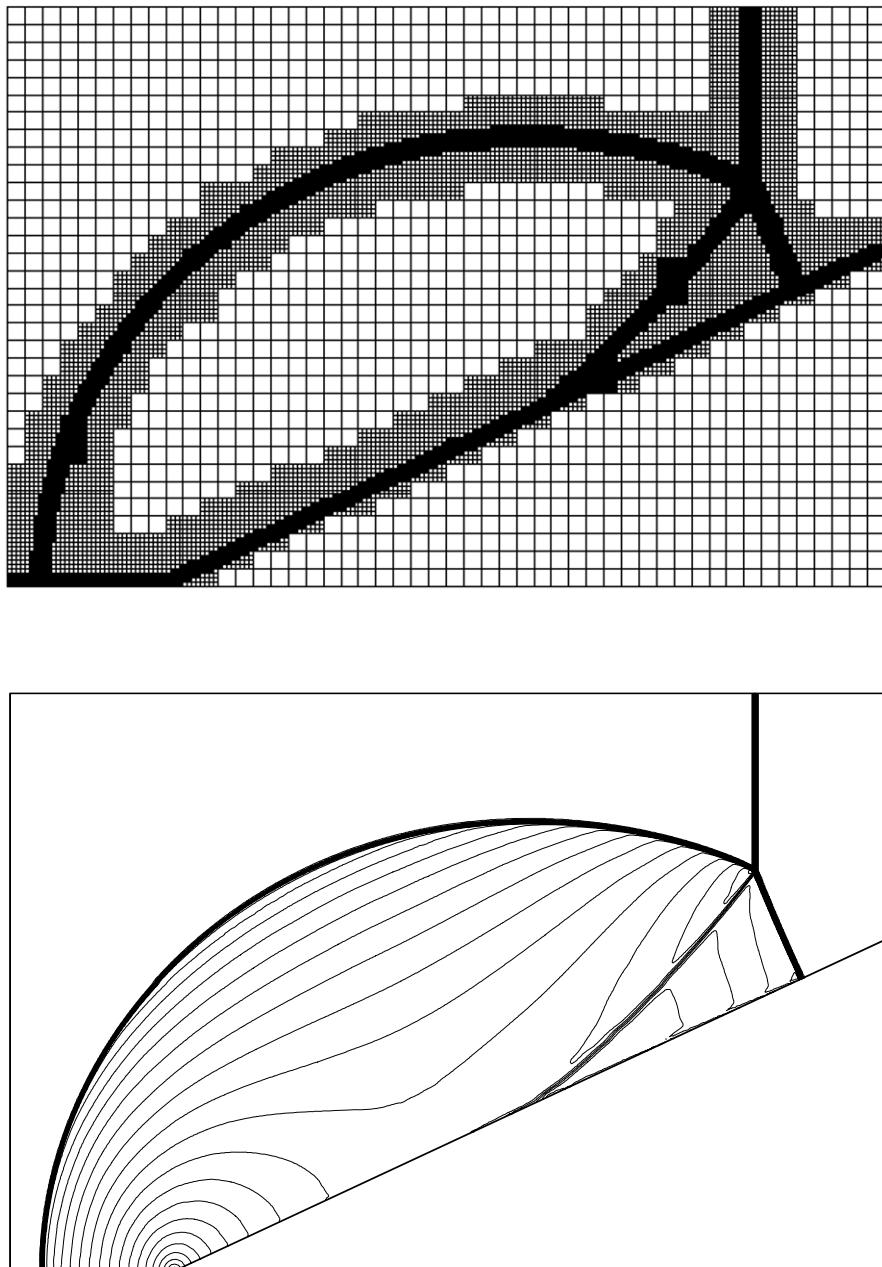
The numerical method used here is the WAF method with dimensional splitting, see Sect. 16.2 and 16.3 of Chap. 16. The Riemann solver used is HLLC of Sect. 10.4 of Chap. 10. The limiter function used is the VANLEER limiter; see Sects. 13.7 and 13.8 of Chap. 13, see also Chap. 14. The CFL coefficient is  $C_{cfl} = 0.9$  and the maximum wave speed is estimated as in Sect. 16.3.2 of Chap. 16. The bottom frame of Fig. 17.9 shows the numerical solution obtained with a regular curvilinear mesh of  $800 \times 528$ .

Fig. 17.10 shows the numerical results obtained with a scheme that combines the CHIMERA approach [20], [462], [43], [105], [106], an Adaptive Mesh Refinement (AMR) approach [45], [48], [46], [399], [118], [400] and a Cartesian mesh approach. The base grid used is  $50 \times 33$  with two further levels of refinement, each with a refinement factor of four. The top frame of Fig. 17.10 shows the structure of the mesh at the output time. Details of the CHIMERA–AMR–Cartesian approach used are found in Boden [66] and Boden and Toro [67]. The bottom frame of Fig. 17.10 shows the corresponding computed density, compare with bottom frame of Fig. 17.9.

The wave pattern of the solution corresponds to *single Mach reflection*. There are three shocks meeting at the triple point, namely, part of the incident shock, the reflected shock and the Mach stem, which joins the triple point



**Fig. 17.9.** Shock reflection problem for  $M_S = 1.7$  and wedge angle of  $\phi = 25$  degrees. Top frame: experimental result (Courtesy of Professor K. Takayama). Bottom frame: density contours from fine regular mesh solution obtained with WAF method and HLLC Riemann solver.



**Fig. 17.10.** Shock reflection problem for  $M_S = 1.7$  and wedge angle of  $\phi = 25$  degrees. Top frame: AMR-CHIMERA grid structure at output time. Bottom frame: density contours from AMR-CHIMERA solution obtained with WAF method with HLLC Riemann solver (Courtesy of Dr. E. P. Boden).

with the wedge surface. From the triple point there emerges a slip surface that joins the wedge surface at a sharp angle. The qualitative agreement between the experimental result and the computations is excellent. The reader is warned however, that for high shock Mach numbers, when both real gas and dissipative effects are important, the agreement deteriorates, unless of course the Euler equations are replaced with the Navier–Stokes equations and the ideal gas equation of state is replaced with a suitable real-gas equation of state. See the paper by Glaz et. al. [210], [211] on these issues.

Generally, the overall wave structure of the solution of this problem can be computed well with almost any modern shock-capturing method. A delicate feature is the *slip surface*, across which discontinuities in density and velocity occur. The particular Riemann solver used in Godunov-type methods can, however, have a profound influence on the solution. For instance, the HLL Riemann solver of Chap. 10, the structure of which does not account for contact and shear waves, will smear the slip surface to unacceptable levels. Other Riemann solvers may be sensitive to extreme conditions. Godunov-type upwind methods with complete Riemann solvers will generally be more accurate than other methods, such as incomplete Riemann solvers, centred methods and flux vector splitting methods.

---

## FORCE Fluxes in Multiple Space Dimensions

This chapter contains a multidimensional extension of the centred FORCE numerical flux introduced in chapter 7. The extension applies to structured and unstructured meshes in two and three spatial dimensions. The resulting one-step schemes in conservative form are shown to be monotone and linearly stable for a generous range of Courant numbers. Sample numerical results are shown for the basic first-order FORCE schemes as well as for higher-order extensions in space and time using the ADER approach.

### 18.1 Introduction

The pioneering work of Lax and Wendroff [302], and more recently that of Hou and LeFloch [259], have established, theoretically, that numerical methods for systems of non-linear hyperbolic conservation laws must be conservative. Then, a key task is the prescription of monotone intercell numerical fluxes. These will then constitute the building block for a wide range of high-order numerical methods constructed in the frameworks of finite volume and discontinuous Galerkin finite element methods, in either fully discrete or semi-discrete form, on structured or unstructured meshes. The construction of numerical fluxes has been a central research issue for over five decades, of which two very prominent and representative examples are the Lax–Friedrichs flux, see chapter 5, and the Godunov flux, see chapter 6. These two methods contain key ideas that remain the pillars of current research and are representative of two distinct approaches for prescribing numerical fluxes, respectively termed *centred* (or symmetric) and *upwind* (or Riemann–problem based, or characteristic–based). Upwind schemes explicitly use wave propagation information contained in the differential equations for the construction of the numerical flux, which is usually accomplished by solving a local one-dimensional Riemann problem, in the direction normal to a cell interface. See Chapter 4 for details of the exact solution of the *classical* Riemann problem for the Euler equations. Approximate Riemann solvers are discussed in Chapters 7 to

12. Centred schemes, on the other hand, do not explicitly use wave propagation information contained in the corresponding Riemann problem. That is, centred schemes do not solve the Riemann problem, in the conventional manner. However it is fair to point out that centred schemes are not independent of the Riemann problem; they use both the differential equations and the initial conditions of the relevant Riemann problem. Examples of classical centred schemes are introduced in Chapter 5. In general, it is accepted that Godunov-type, upwind methods are *potentially* more accurate than centred-type methods; but the latter are simpler and more general than the former. The potentially enhanced accuracy of upwind methods is mainly realized for intermediate waves, especially if they move slowly relative to the mesh. In particular, contact and shear waves are more accurately resolved by upwind methods than by centred methods. However at this point it is necessary to state that the potentially superior accuracy of upwind methods depends crucially on the particular Riemann solver used. In addition to being linear or non-linear, a Riemann solver may also be classified as *complete* or *incomplete*. A *complete* Riemann solver is that whose wave model includes all characteristic fields present in the exact Riemann problem. Otherwise we speak of *incomplete* Riemann solvers. The simplest incomplete Riemann solver for a system is the Rusanov scheme, a one-wave scheme, see Chapter 10. Another example is the HLL Riemann solver, a two-wave scheme, see Chapter 10; this solver is incomplete for any system of more than two characteristic fields. Obviously, the exact Riemann solver is complete. The approximate HLLC solver, Chapter 10, is complete for the Euler equations in one, two and three space dimensions. The Roe Riemann solver and the Osher–Solomon Riemann solver are complete, for the systems they have been constructed. Regarding resolution of intermediate waves, flux vector splitting solvers, Chapter 8, perform like incomplete Riemann solvers; results are similar to those of the incomplete HLL solver. Incomplete Riemann solvers tend to smear intermediate waves very heavily; these waves are not contained in their approximate wave models. The overall performance of incomplete Riemann solvers and flux vector splitting methods is comparable, at least for the larger range of Courant numbers, to that of centred-type methods, which tend to be much simpler. This means that unless one has a *complete* Riemann solver available, it is more convenient to use a centred-type method, provided such method can be used for general meshes in conservative form and can be extended to high order of accuracy in space and time.

The FORCE scheme is a centred-type approach first introduced for one-dimensional hyperbolic problems in [513] and [528]. FORCE is related to the centred schemes of Nessyahu and Tadmor [362] and to the works of Arminjon and collaborators [16]. Both the Nessyahu–Tadmor and the FORCE approaches have the common feature of applying an integral averaging operator on staggered control volumes. The Nessyahu–Tadmor scheme consists of a two-step procedure applied on staggered grids; the scheme does not have a conservative form and is subject to a CFL restriction of  $1/2$ , in one space

dimension. The FORCE scheme is a one-step procedure in conservative form on a non-staggered grid, with a corresponding numerical flux, the FORCE flux; it is subject to a CFL restriction of unity, in one space dimension. The Nessyahu–Tadmor schemes have been applied to multidimensional problems on regular Cartesian grids by Jiang and Tadmor [271] and to unstructured triangular and tetrahedral meshes by Arminjon and collaborators, see for example [16]. Available theoretical results for the Nessyahu–Tadmor schemes include proofs of convergence. Haasdonk et al. [231] proved convergence of the first-order version of the Nessyahu–Tadmor scheme for non-linear scalar conservation laws on unstructured triangular meshes in two space dimensions. For the FORCE scheme convergence was proved by Chen and Toro [103] for the case of two non-linear systems of conservation laws in one space dimension, namely the isentropic equation of gas dynamics and the non-linear shallow water equations with a source term due to bed elevation. There is a significant body of literature on the related theme generally referred to as *central schemes*. The reader is encouraged to consult the following works and references therein [57], [315], [316], [155], [293], [286], [17], [16], [18], to name but a few.

There is also an interesting analogy between the FORCE scheme [513] and the composite schemes of Liska and Wendroff [329]. In the composite schemes the solution is advanced in time by alternating between the Lax–Friedrichs scheme, used by  $L_a$  steps, and the Lax–Wendroff scheme, used by  $L_b$  steps, with empirical choices for the numbers  $L_a$ ,  $L_b$ , trying to reconcile accuracy with monotonicity. In the FORCE scheme the *composite aspect* of the scheme is found at the level of the flux, which is precisely the arithmetic mean, weight of 1/2, of the Lax–Friedrichs and the Lax–Wendroff fluxes. The weight 1/2 is significant, as it is precisely the value that gives a monotone scheme with the maximum region of monotonicity, without resorting to explicit wave propagation information.

The material of this chapter is based on recent work reported in [536] and is a multidimensional extension of the, centred, FORCE numerical flux introduced in Chapter 7 for the case of one space dimension. The extension applies to structured and unstructured meshes in two and three space dimensions. The resulting one-step schemes in conservative form are shown to be monotone and linearly stable for a generous range of Courant numbers. The schemes can be extended to arbitrary order of accuracy in both space and time following the ADER approach in the frameworks of finite volume and discontinuous Galerkin finite elements methods, as will be introduced in Chapter 20.

The rest of this chapter is structured as follows. In section 2 we review the FORCE flux for the one-dimensional case. In Section 3 we construct multidimensional versions of FORCE for regular Cartesian meshes in two and three space dimensions. In Section 4 we analyze the monotonicity and linear stability of the multidimensional FORCE schemes on Cartesian meshes. In Section 5 we extend the FORCE approach and construct numerical fluxes on

general meshes, in two and three space dimensions. In Section 6 we show some numerical results. Conclusions are drawn in section 7.

## 18.2 Review of FORCE in One Space Dimension

Here we first review the FORCE flux in one space dimension. To this end we consider  $m \times m$  one-dimensional hyperbolic systems in conservation-law form

$$\partial_t \mathbf{Q} + \partial_x \mathbf{F}(\mathbf{Q}) = \mathbf{0}, \quad (18.1)$$

where  $\mathbf{Q}$  is a vector of  $m$  components, the conserved variables, and  $\mathbf{F}(\mathbf{Q})$  is the corresponding vector of fluxes. The finite volume method to solve (18.1) reads

$$\mathbf{Q}_i^{n+1} = \mathbf{Q}_i^n - \frac{\Delta t}{\Delta x} [\mathbf{F}_{i+\frac{1}{2}} - \mathbf{F}_{i-\frac{1}{2}}], \quad (18.2)$$

where  $\mathbf{F}_{i+\frac{1}{2}}$  is the numerical flux,  $\Delta x$  is the length of the *control volume* and  $\Delta t$  is the time step.

### 18.2.1 FORCE and Related Fluxes

We first review the reference first-order method, Godunov's upwind method [216], which defines the intercell numerical flux  $\mathbf{F}_{i+\frac{1}{2}}$  in terms of the solution, if available, of the corresponding Riemann problem

$$\left. \begin{aligned} \partial_t \mathbf{Q} + \partial_x \mathbf{F}(\mathbf{Q}) &= \mathbf{0}, \quad x \in (-\infty, \infty), \quad t > 0, \\ \mathbf{Q}(x, 0) &= \begin{cases} \mathbf{Q}_i^n & \text{if } x < 0, \\ \mathbf{Q}_{i+1}^n & \text{if } x > 0. \end{cases} \end{aligned} \right\} \quad (18.3)$$

See Chaps. 4 and 6. The so-called *Riemann fan* in the  $x$ - $t$  plane consists of  $m+1$  constant states separated by  $m$  wave families, each one associated with a real eigenvalue  $\lambda^{(k)}$ . The similarity solution of (18.3) depends on the ratio  $x/t$  and is denoted by  $\mathbf{Q}_{i+\frac{1}{2}}(x/t)$ . The Godunov intercell numerical flux is found by first evaluating  $\mathbf{Q}_{i+\frac{1}{2}}(x/t)$  at  $x/t = 0$ , that is along the  $t$ -axis, and then evaluating the physical flux vector  $\mathbf{F}(\mathbf{Q})$  in (18.3) at  $\mathbf{Q}_{i+\frac{1}{2}}(0)$ , namely

$$\mathbf{F}_{i+\frac{1}{2}}^{God} = \mathbf{F}(\mathbf{Q}_{i+\frac{1}{2}}(0)). \quad (18.4)$$

The exact solution of (18.3) for complicated systems, if possible, will generally involve the iterative solution of a non-linear system and thus in practice, whenever available, one uses approximate Riemann solvers. See Chaps. 7 to 12.

Non-upwind (or centred, or symmetric) schemes do not explicitly utilize wave propagation information and are thus simpler and more generally applicable. Commonly, the numerical fluxes can be computed explicitly as simple algebraic functions of the initial condition in (18.3), namely

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{H}(\mathbf{Q}_i^n, \mathbf{Q}_{i+1}^n), \quad (18.5)$$

where  $\mathbf{H}(\mathbf{V}_L, \mathbf{V}_R)$  is a two-point flux function, required to satisfy the consistency condition  $\mathbf{H}(\mathbf{V}, \mathbf{V}) = \mathbf{F}(\mathbf{V})$ . A classical example of a simple numerical flux that leads to a monotone scheme (18.2) is the Lax–Friedrichs flux

$$\mathbf{F}_{i+\frac{1}{2}}^{lf} = \frac{1}{2}[\mathbf{F}(\mathbf{Q}_i^n) + \mathbf{F}(\mathbf{Q}_{i+1}^n)] - \frac{1}{2} \frac{\Delta x}{\Delta t} (\mathbf{Q}_{i+1}^n - \mathbf{Q}_i^n). \quad (18.6)$$

One may interpret simple fluxes, such as the Lax–Friedrichs flux, as resulting from a *low-level approximation* to the solution of the Riemann problem (18.3). The Lax–Friedrichs flux, when inserted into (18.2), leads to a scheme that is linearly stable up to Courant number unity for the model hyperbolic equation. In addition, the scheme is monotone for the scalar version of (18.1).

Another classical, and slightly more elaborate, flux is the two-step Lax–Wendroff flux

$$\mathbf{F}_{i+\frac{1}{2}}^{lw} = \mathbf{F}(\mathbf{Q}_{i+\frac{1}{2}}^{lw}), \quad \mathbf{Q}_{i+\frac{1}{2}}^{lw} = \frac{1}{2} (\mathbf{Q}_i^n + \mathbf{Q}_{i+1}^n) - \frac{1}{2} \frac{\Delta t}{\Delta x} [\mathbf{F}(\mathbf{Q}_{i+1}^n) - \mathbf{F}(\mathbf{Q}_i^n)], \quad (18.7)$$

which leads to a scheme that is linearly stable up to Courant number unity for the model hyperbolic equation, non-monotone and second-order accurate in space and time.

A simple flux that leads to a monotone scheme is the FORCE flux, introduced in Chapt. 7. This flux was first derived [513] from a deterministic interpretation of the staggered-grid version of Glimm's method [212] and results in a non-staggered one-step conservative scheme of the form (18.2) with intercell numerical flux

$$\mathbf{F}_{i+\frac{1}{2}}^{\text{force}} = \frac{1}{4} \left[ \mathbf{F}(\mathbf{Q}_i^n) + 2\mathbf{F}(\mathbf{Q}_{i+\frac{1}{2}}^{lw}) + \mathbf{F}(\mathbf{Q}_{i+1}^n) - \frac{\Delta x}{\Delta t} (\mathbf{Q}_{i+1}^n - \mathbf{Q}_i^n) \right].$$

This can be rewritten as

$$\mathbf{F}_{i+\frac{1}{2}}^{\text{force}} = \frac{1}{2} (\mathbf{F}_{i+\frac{1}{2}}^{lw} + \mathbf{F}_{i+\frac{1}{2}}^{lf}), \quad (18.8)$$

the arithmetic average of the Lax–Friedrichs flux (18.6) and the two-step Lax–Wendroff flux (18.7).

For further details on the FORCE flux see [513], [528] and Chapt. 7 of this book, sections 7.4.2 and 7.4.3. See also [103], where convergence is proved for the case of two non-linear hyperbolic systems, namely the equations of isentropic gas dynamics and the shallow water equations with a bottom slope source term.

### 18.2.2 Monotonicity and Numerical Viscosity

Here we analyze monotonicity and numerical viscosity of the FORCE scheme for the model hyperbolic equation

$$\partial_t q + \lambda \partial_x q = 0, \quad \lambda: \text{constant}, \quad (18.9)$$

with physical flux function  $f(q) = \lambda q$ . We also discuss a possible generalization of FORCE, called GFORCE [496], [524]. The corresponding conservative scheme for (18.9) reads

$$q_i^{n+1} = q_i^n - \frac{\Delta t}{\Delta x} [f_{i+\frac{1}{2}} - f_{i-\frac{1}{2}}], \quad (18.10)$$

for which the Lax–Friedrichs and Lax–Wendroff, respectively, are

$$f_{i+\frac{1}{2}}^{lf} = \frac{(1+c)}{2c} (\lambda q_i^n) - \frac{(1-c)}{2c} (\lambda q_{i+1}^n) \quad (18.11)$$

and

$$f_{i+\frac{1}{2}}^{lw} = \frac{1}{2} (1+c) (\lambda q_i^n) + \frac{1}{2} (1-c) (\lambda q_{i+1}^n). \quad (18.12)$$

Here  $c$  is the CFL or Courant number  $c = \frac{\lambda \Delta t}{\Delta x}$ . Then the FORCE flux is easily verified to be

$$f_{i+\frac{1}{2}}^{force} = \frac{(1+c)^2}{4c} (\lambda q_i^n) - \frac{(1-c)^2}{4c} (\lambda q_{i+1}^n), \quad (18.13)$$

the arithmetic mean of the Lax–Friedrichs and Lax–Wendroff fluxes.

Now consider FORCE-type fluxes (18.8), see [496] and [524], as convex averages of the Lax–Friedrichs and the two-step Lax–Wendroff fluxes, as follows

$$f_{i+\frac{1}{2}}^{(\omega)} = \omega f_{i+\frac{1}{2}}^{lw} + (1-\omega) f_{i+\frac{1}{2}}^{lf}, \quad 0 \leq \omega \leq 1, \quad (18.14)$$

which written in full becomes

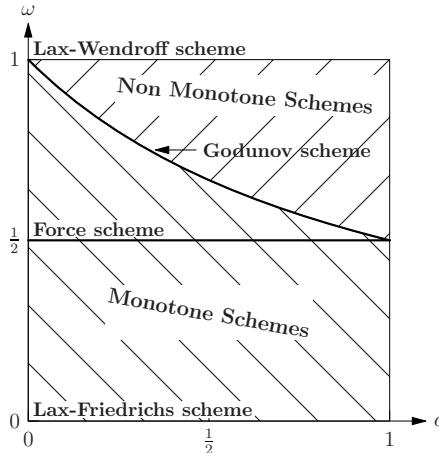
$$f_{i+\frac{1}{2}}^{(\omega)} = \frac{1}{2} (1+c) \left[ \frac{(1-c)\omega + 1}{c} \right] (\lambda q_i^n) + \frac{1}{2} (1-c) \left[ \frac{(1-c)\omega - 1}{c} \right] (\lambda q_{i+1}^n). \quad (18.15)$$

When inserted in (18.10), this flux yields a three-point scheme whose coefficient of numerical viscosity is

$$\alpha = \frac{1}{2} \left[ \frac{(1-c^2)(1-\omega)}{c} \right] \lambda \Delta x, \quad (18.16)$$

which varies linearly with  $\omega$ ; it is a maximum for the Lax–Friedrichs flux ( $\omega = 0$ ) and a minimum (zero) for the Lax–Wendroff flux ( $\omega = 1$ ).

It is instructive to analyze the weighted FORCE-type schemes in the  $c-\omega$  plane of Fig. 18.1. Particular values of the weight  $\omega$  in (18.14) reproduce well-known numerical fluxes. For example, the bottom horizontal line with constant



**Fig. 18.1.** Weighted FORCE-type fluxes defined as convex averages of the Lax–Friedrichs and Lax–Wendroff fluxes in the  $c - \omega$  plane. Here  $c$  is Courant number and  $\omega$  is a weight, whose constant values 0,  $\frac{1}{2}$  and 1 reproduce the Lax–Friedrichs, FORCE and Lax–Wendroff fluxes;  $\omega = 1/(1 + |c|)$  reproduces the Godunov upwind method.

weight  $\omega = 0$  gives the Lax–Friedrichs flux (18.11). The top horizontal line with constant weight  $\omega = 1$  gives the Lax–Wendroff flux (18.12). For the constant weight  $\omega = \frac{1}{2}$  we reproduce the FORCE flux (18.13).

We could also consider weights  $\omega(c)$  that depend on wave propagation information, such as the Courant number  $c$ . In particular, by comparing the coefficients of the flux (18.15) with those of the Godunov upwind flux, see Chapter 5, we obtain the Courant number dependent weight given by

$$\omega_g(c) = \frac{1}{1 + |c|}, \quad (18.17)$$

for which the average flux (18.14) reproduces identically Godunov's upwind flux.

The Godunov curve  $\omega_g(c)$  divides the unit square of Fig. 18.1 into two sub-regions. The region associated with weights lying above  $\omega_g(c)$  contains non-monotone schemes and the region below  $\omega_g(c)$  contains monotone schemes. The range of *constant* weights  $\omega$  (no dependence on  $c$ ), with  $0 \leq \omega \leq \frac{1}{2}$ , defines a sub-class of monotone schemes, with the FORCE flux (18.13) corresponding to the limiting case  $\omega = \frac{1}{2}$ , which is the scheme with the smallest numerical viscosity within the class, and thus the optimal scheme, for which no dependence on  $c$  is required, that is no explicit upwind information is needed.

In order to improve upon the centred FORCE flux, that is, reduce its numerical viscosity further, we need to consider entering the region  $\frac{1}{2} < \omega \leq 1$ , within which all schemes (18.14) with constant weight  $\omega$  are *non-monotone*; monotonicity is lost for the larger values of the CFL number  $c$ . Another way

to put it is by saying that for  $\frac{1}{2} < \omega \leq 1$  the scheme will remain monotone for increasingly smaller CFL numbers.

Let us consider the following subset of the square  $[0, 1] \times [0, 1]$  in the  $c - \omega$  plane of Fig. 18.1

$$T_g = \left\{ (c, \omega) : 0 \leq c \leq 1 \text{ and } \frac{1}{2} < \omega \leq \omega_g(c) = \frac{1}{1 + |c|} \right\}. \quad (18.18)$$

We call GFORCE (for Generalized FORCE, or Godunov FORCE) the weighted scheme (18.14) with weight  $\omega$  as given by (18.17). For the linear advection equation (18.9) GFORCE is identical to the Godunov upwind scheme. For scalar non-linear equations and for non-linear systems, GFORCE is a generalization of Godunov; such generalizations rely on the choice of a Courant number to be inserted in (18.17), and thus GFORCE requires a limited amount of upwind information. See [496] and [524] for details.

We return to the linear scalar case. The set of points  $T_g$  forms a kind of triangular sub region, with a curved *hypotenuse*. It is revealing to interpret the square  $[0, 1] \times [0, 1]$  of area unity in the  $c - \omega$  plane as representing the difference in total numerical viscosity between the Lax–Friedrichs scheme (maximum) and the Lax–Wendroff flux (minimum, zero). The FORCE scheme reduces the viscosity of the Lax–Friedrichs flux by 50% and is the optimal scheme in the class of methods (18.14) that makes no explicit use of wave propagation information. The upper boundary of  $T_g$ , given by (18.17), gives precisely the Godunov upwind scheme, the monotone scheme with the smallest coefficient of numerical viscosity given by

$$\alpha_{god} = \frac{1}{2}(1 - |c|)\lambda\Delta x. \quad (18.19)$$

The Godunov scheme reduces the numerical viscosity by a further 19% ( $\ln 2 - 1/2$ ), relative to the FORCE scheme; this is the maximum possible reduction within the class of monotone methods; the Godunov scheme is the optimal scheme within the class of monotone methods. This 19% reduction in numerical viscosity is due to the use of upwinding.

The GFORCE flux (18.14), (18.17) as applied to non-linear problems attempts to reduce the numerical viscosity by an amount that approaches the magic 19% of the Godunov scheme. For the linear scalar case this is fully achieved, as GFORCE becomes identical to the Godunov upwind scheme. For the non-linear scalar case the approach is also very satisfactory. Unfortunately, for non-linear systems the success is only partial. Further reductions in numerical viscosity, without solving the Riemann problem in the conventional manner, are possible by using the MUSTA (Multi–Stage) approach [496], [495], [524], [521].

As previously discussed, FORCE is the optimal centred flux in the class of schemes defined by (18.14), the one that has the smallest numerical viscosity without explicitly using upwind information. However, the crucial limitation

of the FORCE flux is linear stability in two and three space dimensions, as shown in [528]. This limitation also afflicts the Lax–Friedrichs and the Lax–Wendroff fluxes as stated here. One possible remedy for the FORCE scheme is to use it in the framework of dimensional splitting schemes. However, this would rule out unstructured meshes as well as accuracy greater than two, even on structured meshes. In the next section we introduce a generalization of the FORCE approach to two and three space dimensions. We start with the Cartesian mesh case.

## 18.3 FORCE Schemes on Cartesian Meshes

We consider general systems of  $m$  non-linear hyperbolic equations written in differential conservation–law form

$$\partial_t \mathbf{Q} + \partial_x \mathbf{F}(\mathbf{Q}) + \partial_y \mathbf{G}(\mathbf{Q}) + \partial_z \mathbf{H}(\mathbf{Q}) = \mathbf{0}, \quad (18.20)$$

where  $\mathbf{Q}$  is the vector of conserved variables,  $\mathbf{F}(\mathbf{Q})$ ,  $\mathbf{G}(\mathbf{Q})$  and  $\mathbf{H}(\mathbf{Q})$  are respectively the flux vectors in the  $x$ ,  $y$  and  $z$  coordinate directions.

### 18.3.1 The Two-Dimensional Case

We first consider the two-dimensional case,  $\mathbf{H}(\mathbf{Q}) = \mathbf{0}$ , and conservative schemes on Cartesian meshes, of the form

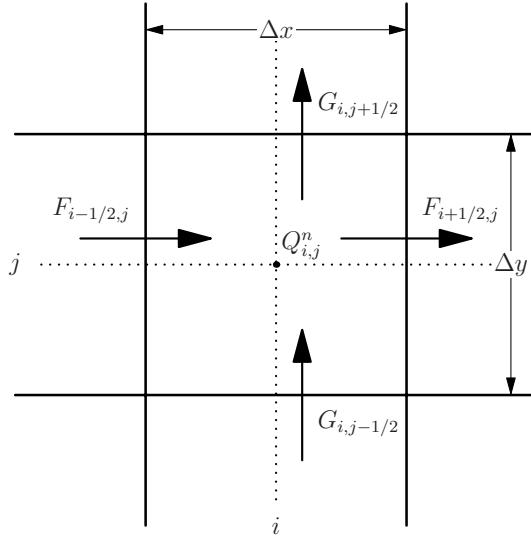
$$\mathbf{Q}_{i,j}^{n+1} = \mathbf{Q}_{i,j}^n - \frac{\Delta t}{\Delta x} [\mathbf{F}_{i+\frac{1}{2},j} - \mathbf{F}_{i-\frac{1}{2},j}] - \frac{\Delta t}{\Delta y} [\mathbf{G}_{i,j+\frac{1}{2}} - \mathbf{G}_{i,j-\frac{1}{2}}]. \quad (18.21)$$

The task at hand is to construct numerical fluxes for each of the four cell interfaces:  $\mathbf{F}_{i+\frac{1}{2},j}$ ,  $\mathbf{G}_{i,j+\frac{1}{2}}$ ,  $\mathbf{F}_{i-\frac{1}{2},j}$  and  $\mathbf{G}_{i,j-\frac{1}{2}}$ . Fig. 18.2 illustrates the situation.

**Meshes and notation.** The construction of FORCE schemes requires two meshes: the *primary mesh*, where the numerical scheme provides cell averages, and a *staggered mesh*, used to define numerical fluxes for the conservative form of the scheme.

The *primary mesh* is the mesh chosen for the discretization of the domain of interest. Having chosen the *primary mesh*, there are various possible choices for the *staggered mesh*. One option is to use *vertex-based* staggered meshes, as suggested by Jiang and Tadmor [271]. An alternative option is to use *edge-based* staggered meshes, as suggested in [536], which is the approach we consider here.

Fig. 18.3 depicts the situation for the two-dimensional case introduced in Fig. 18.2. The  $(i, j)$  numbering is typical of Cartesian meshes and structured non–Cartesian meshes. The chosen Cartesian mesh determines the computing cells denoted by  $C_{i,j}$ . There are four edges (or faces or inter-cell boundaries)



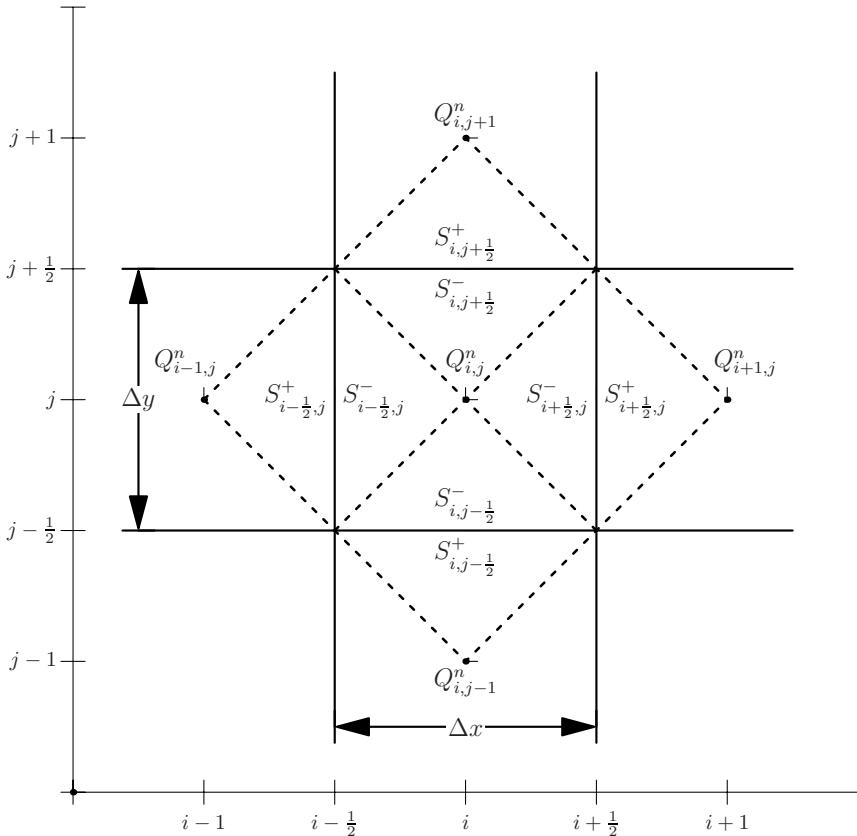
**Fig. 18.2.** Cartesian cell  $C_{i,j}$  in two-dimensional domain in  $x-y$  space.  $\mathbf{Q}_{i,j}^n$  is cell average in cell  $C_{i,j}$  at time  $t^n$ .

of  $C_{i,j}$ . For each edge  $k$  we adopt the conventional anticlock orientation so that the cell  $C_{i,j}$  lies on the left of edge  $k$  and the immediate neighbour of  $C_{i,j}$  across edge  $k$  lies on the right. To form an element of the staggered mesh, that corresponds to edge  $k$ , one joins the barycentre of  $C_{i,j}$  and that of its neighbour across edge  $k$ , with the vertices of edge  $k$ . This results in a quadrilateral straddling face  $k$ , as depicted in Fig. 18.3. The four control volumes and the corresponding edges of cell  $C_{i,j}$  are defined as follows

$$\left. \begin{aligned} S_{i+\frac{1}{2},j} &= S_{i+\frac{1}{2},j}^- \cup S_{i+\frac{1}{2},j}^+, \quad S_{i+\frac{1}{2},j}^0 = S_{i+\frac{1}{2},j}^- \cap S_{i+\frac{1}{2},j}^+, \\ S_{i,j+\frac{1}{2}} &= S_{i,j+\frac{1}{2}}^- \cup S_{i,j+\frac{1}{2}}^+, \quad S_{i,j+\frac{1}{2}}^0 = S_{i,j+\frac{1}{2}}^- \cap S_{i,j+\frac{1}{2}}^+, \\ S_{i-\frac{1}{2},j} &= S_{i-\frac{1}{2},j}^- \cup S_{i-\frac{1}{2},j}^+, \quad S_{i-\frac{1}{2},j}^0 = S_{i-\frac{1}{2},j}^- \cap S_{i-\frac{1}{2},j}^+, \\ S_{i,j-\frac{1}{2}} &= S_{i,j-\frac{1}{2}}^- \cup S_{i,j-\frac{1}{2}}^+, \quad S_{i,j-\frac{1}{2}}^0 = S_{i,j-\frac{1}{2}}^- \cap S_{i,j-\frac{1}{2}}^+. \end{aligned} \right\} \quad (18.22)$$

The portion of each volume of the staggered mesh that lies inside  $C_{i,j}$  (on the left) is labelled with superscript "minus", and that outside (on the right) is labelled with superscript "plus". See Fig. 18.3.

**Stages in the construction of the schemes.** The construction of the FORCE schemes is achieved in three stages:



**Fig. 18.3.** Primary mesh of Cartesian cells  $C_{i,j}$  and edge-based staggered mesh in two-dimensional domain in  $x-y$  space.

- *I: Averages on staggered mesh.* In the first stage we assume initial condition  $\mathbf{Q}_{i,j}^n$  for the conservation laws at time  $t^n$  in cells  $C_{i,j}$ . One then evolves the solution for a time  $\frac{1}{2}\Delta t$  in each control volume in (18.22) of the staggered mesh associated with an edge of  $C_{i,j}$ . The evolution is carried out by applying the integral form of the conservation laws (18.20). See Chapter 16 for background on the integral form of the conservation laws and related concepts.

For a generic control volume  $S_k$ , assumed to be a polygon of  $n_k$  edges, the integral form of the conservation laws (18.20) applied for a time  $\delta t$  reads

$$\mathbf{Q}_k^1 = \mathbf{Q}_k^0 - \frac{\delta t}{|S_k|} \sum_{l=1}^{n_k} \int_{v_l}^{v_{l+1}} [\cos\theta_l \mathbf{F}(\mathbf{Q}) + \sin\theta_l \mathbf{G}(\mathbf{Q})] ds . \quad (18.23)$$

Here  $\mathbf{Q}_k^0$  is the initial condition at time  $t = 0$ ,  $\mathbf{Q}_k^1$  is the integral average at time  $t = \delta t$ ,  $\theta_l$  is the angle between the  $x$ -axis and the normal vector

to face  $l$  joining vertices  $v_l$  and  $v_{l+1}$ , with  $v_{n_k+1} \equiv v_1$ ,  $|S_k|$  is the area of control volume  $S_k$  and  $\mathbf{n}_k = (\cos \theta_l, \sin \theta_l)$  is the outward unit normal vector to edge  $l$  of  $S_k$ . Application of (18.23) to each of the secondary volumes in (18.22) yields

$$\left. \begin{aligned} \mathbf{Q}_{i+\frac{1}{2},j}^{n+\frac{1}{2}} &= \frac{1}{2}(\mathbf{Q}_{i,j}^n + \mathbf{Q}_{i+1,j}^n) - \frac{\Delta t}{\Delta x}[\mathbf{F}(\mathbf{Q}_{i+1,j}^n) - \mathbf{F}(\mathbf{Q}_{i,j}^n)], \\ \mathbf{Q}_{i,j+\frac{1}{2}}^{n+\frac{1}{2}} &= \frac{1}{2}(\mathbf{Q}_{i,j}^n + \mathbf{Q}_{i,j+1}^n) - \frac{\Delta t}{\Delta y}[\mathbf{G}(\mathbf{Q}_{i,j+1}^n) - \mathbf{G}(\mathbf{Q}_{i,j}^n)], \\ \mathbf{Q}_{i-\frac{1}{2},j}^{n+\frac{1}{2}} &= \frac{1}{2}(\mathbf{Q}_{i-1,j}^n + \mathbf{Q}_{i,j}^n) - \frac{\Delta t}{\Delta x}[\mathbf{F}(\mathbf{Q}_{i,j}^n) - \mathbf{F}(\mathbf{Q}_{i-1,j}^n)], \\ \mathbf{Q}_{i,j-\frac{1}{2}}^{n+\frac{1}{2}} &= \frac{1}{2}(\mathbf{Q}_{i,j-1}^n + \mathbf{Q}_{i,j}^n) - \frac{\Delta t}{\Delta y}[\mathbf{G}(\mathbf{Q}_{i,j}^n) - \mathbf{G}(\mathbf{Q}_{i,j-1}^n)]. \end{aligned} \right\} \quad (18.24)$$

Note that the initial condition in cell  $C_{i,j}$  and all its immediate four neighbours determine initial conditions for each of the control volumes in (18.22) of the staggered mesh associated with the boundaries of  $C_{i,j}$ . For example, the initial condition at time  $t = t^n$  on  $S_{i+\frac{1}{2},j}^-$  is  $\mathbf{Q}_{i,j}^n$  and on  $S_{i+\frac{1}{2},j}^+$  is  $\mathbf{Q}_{i+1,j}^n$ , explaining the arithmetic means in (18.24).

- *II: Averages on primary mesh.* In the second stage of the construction of the schemes, we assume initial conditions (18.24) at time  $t = \frac{1}{2}\Delta t$  and evolve the solution by another time  $\frac{1}{2}\Delta t$  within each cell  $C_{i,j}$ , by applying again the integral form of the conservation laws (18.23), obtaining

$$\left. \begin{aligned} \mathbf{Q}_{i,j}^{n+1} &= \frac{1}{4}[\mathbf{Q}_{i+\frac{1}{2},j}^{n+\frac{1}{2}} + \mathbf{Q}_{i,j+\frac{1}{2}}^{n+\frac{1}{2}} + \mathbf{Q}_{i-\frac{1}{2},j}^{n+\frac{1}{2}} + \mathbf{Q}_{i,j-\frac{1}{2}}^{n+\frac{1}{2}}] \\ &\quad - \frac{1}{2}\frac{\Delta t}{\Delta x}[\mathbf{F}(\mathbf{Q}_{i+\frac{1}{2},j}^{n+\frac{1}{2}}) - \mathbf{F}(\mathbf{Q}_{i-\frac{1}{2},j}^{n+\frac{1}{2}})] \\ &\quad - \frac{1}{2}\frac{\Delta t}{\Delta y}[\mathbf{G}(\mathbf{Q}_{i,j+\frac{1}{2}}^{n+\frac{1}{2}}) - \mathbf{G}(\mathbf{Q}_{i,j-\frac{1}{2}}^{n+\frac{1}{2}})]. \end{aligned} \right\} \quad (18.25)$$

- *III: Conservative form and the FORCE flux.* In the final stage of the construction of the scheme we perform algebraic manipulations on (18.25) so as to reproduce the conservative formula (18.21). The resulting FORCE numerical flux is

$$\mathbf{F}_{i+\frac{1}{2},j} = \frac{1}{2}[\mathbf{F}(\mathbf{Q}_{i+\frac{1}{2},j}^{n+\frac{1}{2}}) + \frac{1}{2}(\mathbf{F}(\mathbf{Q}_{i,j}^n) + \mathbf{F}(\mathbf{Q}_{i+1,j}^n)) - \frac{1}{4}\frac{\Delta x}{\Delta t}(\mathbf{Q}_{i+1,j}^n - \mathbf{Q}_{i,j}^n)], \quad (18.26)$$

with analogous formulae for the other fluxes. As in the one-dimensional case, it is possible to rewrite the intercell flux (18.26) as the arithmetic average of two fluxes, namely

$$\mathbf{F}_{i+\frac{1}{2},j} = \frac{1}{2}(\mathbf{F}_{i+\frac{1}{2},j}^{lw2} + \mathbf{F}_{i+\frac{1}{2},j}^{lf2}) . \quad (18.27)$$

Here

$$\mathbf{F}_{i+\frac{1}{2},j}^{lw2} = \mathbf{F}(\mathbf{Q}_{i+\frac{1}{2},j}^{lw2}) , \quad (18.28)$$

with

$$\mathbf{Q}_{i+\frac{1}{2},j}^{lw2} = \frac{1}{2}(\mathbf{Q}_{i,j}^n + \mathbf{Q}_{i+1,j}^n) - \frac{1}{2} \left( \frac{2\Delta t}{\Delta x} \right) [\mathbf{F}(\mathbf{Q}_{i+1,j}^n) - \mathbf{F}(\mathbf{Q}_{i,j}^n)] . \quad (18.29)$$

The second flux term in (18.27) is

$$\mathbf{F}_{i+\frac{1}{2},j}^{lf2} = \frac{1}{2}[\mathbf{F}(\mathbf{Q}_{i,j}^n) + \mathbf{F}(\mathbf{Q}_{i+1,j}^n)] - \frac{1}{2} \left( \frac{\Delta x}{2\Delta t} \right) (\mathbf{Q}_{i+1,j}^n - \mathbf{Q}_{i,j}^n) . \quad (18.30)$$

Flux (18.27) may be regarded as a generalization of the Lax–Wendroff flux, or a generalization of the Godunov centred scheme, or a generalization of the WAF method. See the discussion of Section 13.5.2 on a motivating example including a parameter  $\alpha$ .

Flux (18.30) is a generalization of the Lax–Friedrichs flux. The factor  $\alpha = 2$  multiplying the time step  $\Delta t$  in (18.29) and (18.30) denotes the number of spatial dimensions. Compare (18.30) and (18.28) with the genuine one-dimensional case (18.6) and (18.7). The only difference is the factor  $\alpha = 2$ .

**Summary of the FORCE scheme:** The implementation of the numerical scheme to solve (18.20) in two space dimensions,  $\mathbf{H}(\mathbf{Q}) = \mathbf{0}$ , on Cartesian meshes is exceedingly simple. It consists of the conservative scheme (18.21) with numerical fluxes given by (18.27), with analogous definitions for the other faces. The extension to two-dimensional non–Cartesian but structured meshes is straightforward. The details are omitted.

### 18.3.2 The Three–Dimensional Case

We consider the conservative numerical scheme to solve (18.20) on Cartesian meshes, of the form

$$\left. \begin{aligned} \mathbf{Q}_{i,j,k}^{n+1} &= \mathbf{Q}_{i,j,k}^n - \frac{\Delta t}{\Delta x} [\mathbf{F}_{i+\frac{1}{2},j,k} - \mathbf{F}_{i-\frac{1}{2},j,k}] \\ &\quad - \frac{\Delta t}{\Delta y} [\mathbf{G}_{i,j+\frac{1}{2},k} - \mathbf{G}_{i,j-\frac{1}{2},k}] - \frac{\Delta t}{\Delta z} [\mathbf{H}_{i,j,k+\frac{1}{2}} - \mathbf{H}_{i,j,k-\frac{1}{2}}] , \end{aligned} \right\} \quad (18.31)$$

with the obvious notation for the numerical fluxes and the mesh parameters. The FORCE flux for  $\alpha = 3$  (three space dimensions), has the form

$$\mathbf{F}_{i+\frac{1}{2},j,k}^{force\alpha} = \frac{1}{2}(\mathbf{F}_{i+\frac{1}{2},j,k}^{lw\alpha} + \mathbf{F}_{i+\frac{1}{2},j,k}^{lf\alpha}) , \quad (18.32)$$

with

$$\mathbf{F}_{i+\frac{1}{2},j,k}^{lw\alpha} = \mathbf{F}(\mathbf{Q}_{i+\frac{1}{2},j,k}^{lw\alpha}) \quad (18.33)$$

and

$$\mathbf{Q}_{i+\frac{1}{2},j,k}^{lw\alpha} = \frac{1}{2}(\mathbf{Q}_{i,j,k}^n + \mathbf{Q}_{i+1,j,k}^n) - \frac{1}{2} \frac{\alpha \Delta t}{\Delta x} [\mathbf{F}(\mathbf{Q}_{i+1,j,k}^n) - \mathbf{F}(\mathbf{Q}_{i,j,k}^n)] . \quad (18.34)$$

The Lax–Friedrichs type flux is

$$\mathbf{F}_{i+\frac{1}{2},j,k}^{lf\alpha} = \frac{1}{2}[\mathbf{F}(\mathbf{Q}_{i,j,k}^n) + \mathbf{F}(\mathbf{Q}_{i+1,j,k}^n)] - \frac{1}{2} \frac{\Delta x}{\alpha \Delta t} (\mathbf{Q}_{i+1,j,k}^n - \mathbf{Q}_{i,j,k}^n) . \quad (18.35)$$

This concludes the derivation of FORCE fluxes for the case of Cartesian meshes in two and three space dimensions. Before proceeding to the construction of multidimensional FORCE fluxes for general meshes in Section 5, we study some properties of the multidimensional schemes on Cartesian meshes.

## 18.4 Properties of the FORCE Schemes

Here we study some basic properties of the derived multidimensional FORCE schemes for Cartesian meshes. But first, however, we re-interpret the multidimensional FORCE fluxes in a one-dimensional setting by keeping the number of space dimensions  $\alpha$  as a parameter in the one-dimensional flux. This results in a broad family of one-dimensional schemes with some intriguing properties. The results of some numerical experiments are also shown. Then we analyze monotonicity and linear stability of the two and three-dimensional schemes.

### 18.4.1 One-Dimensional Interpretation

Consider the one-dimensional, homogeneous hyperbolic system (18.1) and the conservative scheme (18.2). The one-dimensional version of the two and three-dimensional FORCE flux has the form

$$\mathbf{F}_{i+\frac{1}{2}}^{force\alpha} = \frac{1}{2} \left( \mathbf{F}_{i+\frac{1}{2}}^{lw\alpha} + \mathbf{F}_{i+\frac{1}{2}}^{lf\alpha} \right) , \quad (18.36)$$

where  $\mathbf{F}_{i+\frac{1}{2}}^{lw\alpha}$  and  $\mathbf{F}_{i+\frac{1}{2}}^{lf\alpha}$  are respectively the Lax–Wendroff and Lax–Friedrichs type fluxes. That is

$$\mathbf{F}_{i+\frac{1}{2}}^{lw\alpha} = \mathbf{F}(\mathbf{Q}_{i+\frac{1}{2}}^{lw\alpha}) , \quad \mathbf{Q}_{i+\frac{1}{2}}^{lw\alpha} = \frac{1}{2}(\mathbf{Q}_i^n + \mathbf{Q}_{i+1}^n) - \frac{1}{2} \frac{\alpha \Delta t}{\Delta x} (\mathbf{Q}_{i+1}^n - \mathbf{Q}_i^n) \quad (18.37)$$

and

$$\mathbf{F}_{i+\frac{1}{2}}^{lf\alpha} = \frac{1}{2}[\mathbf{F}(\mathbf{Q}_i^n) + \mathbf{F}(\mathbf{Q}_{i+1}^n)] - \frac{1}{2} \frac{\Delta x}{\alpha \Delta t} (\mathbf{Q}_{i+1}^n - \mathbf{Q}_i^n) . \quad (18.38)$$

The values  $\alpha = 1, 2, 3$  reproduce the one, two and three-dimensional cases. For the one-dimensional case,  $\alpha = 1$ ,  $\mathbf{F}_{i+\frac{1}{2}}^{lw\alpha}$  and  $\mathbf{F}_{i+\frac{1}{2}}^{lf\alpha}$  are identically to the

classical Lax–Wendroff and Lax–Friedrichs fluxes, and  $\mathbf{F}_{i+\frac{1}{2}}^{force\alpha}$  is just the one-dimensional FORCE flux (18.8). For the two-dimensional case,  $\alpha = 2$ ,  $\mathbf{F}_{i+\frac{1}{2}}^{lwx}$  is a known flux, namely, the Godunov centred scheme [217]; see exercise 13.5.3 in Chapter 13. Analysis of this scheme as applied to the model equation (18.9) says that the scheme is linearly stable in the range  $0 \leq |c| \leq \frac{1}{2}\sqrt{2}$ , where  $c$  is the Courant number. The scheme is monotone in the restricted range  $\frac{1}{2} \leq |c| \leq \frac{1}{2}\sqrt{2}$  of linear stability, being non-monotone elsewhere, which makes it useless in practice. The flux  $\mathbf{F}_{i+\frac{1}{2}}^{lfa\alpha}$ , with  $\alpha = 2$ , appears to be new and is analogous to the Lax–Friedrichs flux but has smaller numerical viscosity. The resulting scheme is monotone in its full range of linear stability, namely  $0 \leq |c| \leq \frac{1}{2}$  and is potentially a useful scheme for practical applications. For  $\alpha > 2$  all fluxes  $\mathbf{F}_{i+\frac{1}{2}}^{lwx\alpha}$  and  $\mathbf{F}_{i+\frac{1}{2}}^{lfa\alpha}$  are, to our knowledge, new. However, we note that the motivating example discussed in Section 13.5.2 is related to the Lax–Wendroff type schemes just considered.

For the model equation (18.9) the FORCE flux is

$$f_{i+\frac{1}{2}}^{force\alpha} = \frac{1}{2} \left( 1 + \frac{1 + \alpha^2 c^2}{2\alpha c} \right) (\lambda q_i^n) + \frac{1}{2} \left( 1 - \frac{1 + \alpha^2 c^2}{2\alpha c} \right) (\lambda q_{i+1}^n) \quad (18.39)$$

and the resulting three-point numerical scheme is

$$q_i^{n+1} = \frac{1}{2} c \left( 1 + \frac{1 + \alpha^2 c^2}{2\alpha c} \right) q_{i-1}^n + \left( 1 - \frac{1 + \alpha^2 c^2}{2\alpha c} \right) q_i^n - \frac{1}{2} c \left( 1 - \frac{1 + \alpha^2 c^2}{2\alpha c} \right) q_{i+1}^n. \quad (18.40)$$

It is easy to verify that the FORCE $\alpha$  scheme is monotone and linearly stable under the CFL-type restriction

$$c \leq \frac{\sqrt{2\alpha - 1}}{\alpha}. \quad (18.41)$$

The monotonicity and linear stability limit is a decreasing function of the dimension parameter  $\alpha$ .

In the next section we show some numerical results for one-dimensional problems obtained from the multidimensional FORCE schemes (18.36) regarding the dimension  $\alpha$  as a parameter in the scheme.

### 18.4.2 Some Numerical Experiments

Here we show some numerical results from applying the numerical flux  $\mathbf{F}_{i+\frac{1}{2}}^{force\alpha}$  given by (18.36) into the conservative scheme (18.2). We select two test problems for the one-dimensional Euler equations for ideal gases with  $\gamma = 1.4$ .

**Stationary Contact.** This test is the shock tube problem introduced as Test 6 in Chapter 10, whose exact solution is a stationary contact discontinuity.

This problem is the most demanding kind of tests for centred-type methods, for incomplete Riemann solvers and for conventional flux vector splitting schemes. The computational domain is the interval  $[0, 1]$ ; the initial conditions consist of two constant states separated by a discontinuity in density at  $x = 1/2$ ;  $\rho(x, 0) = \rho_L = 1.4$  for  $x < 1/2$ ,  $\rho(x, 0) = \rho_R = 1$  for  $x > 1/2$ . The other quantities are constant throughout, the particle velocity is  $u(x, 0) = 0$  and the pressure is  $p(x, 0) = 1$ . Transmissive boundary conditions are imposed. Computations are carried out till time  $t_{out} = 5$  units, which is much larger than that in Chapter 10. Boundary conditions are transmissive and the mesh has 100 cells. The Courant number used is 90% of the monotonicity and linear stability limit given by (18.41), with the exception of the van Leer scheme, for which we used a Courant number of 0.6.

Fig. 18.4 shows the numerical results as compared to the exact solution. The top frame shows the results of four FORCE fluxes (18.36), with the dimension parameter set to  $\alpha = 1, 2, 3, 15$ . It is seen that numerical viscosity is reduced visibly by increasing the dimension parameter  $\alpha$ . For simple curiosity we include the result for  $\alpha = 15$ . In the bottom frame of Fig. 18.4 we compare the above  $FORCE\alpha$  schemes with well-established methods. One surprising result is that two well-known upwind schemes, namely the Rusanov and the HLL schemes, have more numerical viscosity for this test than any of the centred  $FORCE\alpha$  schemes. It is also interesting to see that the Flux Vector Splitting scheme of van Leer, although better than the FORCE scheme with  $\alpha = 1$ , is more diffusive than the FORCE schemes with  $\alpha = 2$  and  $\alpha = 3$ . These results suggest that incomplete Riemann solvers and Flux Vector Splitting methods are comparable or even less accurate than the centred  $FORCE\alpha$  schemes of this Chapter for resolving slowly-moving intermediate waves, at least for the larger range of Courant numbers. Only complete Riemann solvers can perform better here.

**Woodward and Collela Test.** In this test [584] the computational domain is the interval  $[0, 1]$ , the initial condition in pressure consists of three constant states separated by two discontinuities, namely  $p(x, 0) = p_L = 1000$  for  $x < 1/10$ ,  $p(x, 0) = p_M = 0.01$  for  $1/10 < x < 9/10$  and  $p(x, 0) = p_R = 100$  for  $x > 9/10$ . Particle velocity and density are constant, with  $u(x, 0) = 0$  and  $\rho(x, 0) = 1$ . Reflective boundary conditions are imposed. The problem is solved on a mesh of 3000 cells, with a Courant number as for the previous test and results are displayed at the output time  $t = t_{out} = 0.038$  units. This problem does not have exact solution. We use the first-order Godunov upwind method with the exact Riemann solver as the reference scheme and note that this is the most accurate scheme that is monotone for the scalar case. The test problem is very severe and is suitable for testing robustness of numerical schemes.

Fig. 18.5 shows the results for the  $FORCE\alpha$  schemes compared with the reference solution of the Godunov method. As expected, the  $FORCE\alpha$  schemes are more diffusive than the reference first-order scheme. As the di-

mension parameter  $\alpha$  increases the solution tends to that of the reference scheme.

In the next section we analyze some properties of the multidimensional  $FORCE\alpha$  schemes.

### 18.4.3 Analysis in Multiple Space Dimensions

We study monotonicity and linear stability of the two and three-dimensional FORCE schemes in terms of the linear advection equation with constant coefficients

$$\partial_t q + \partial_x f(q) + \partial_y g(q) + \partial_z h(q) = 0, \quad (18.42)$$

with fluxes  $f(q) = \lambda_1 q$ ,  $g(q) = \lambda_2 q$  and  $h(q) = \lambda_3 q$ ; the velocity components  $\lambda_1$ ,  $\lambda_2$  and  $\lambda_3$  in the  $x$ ,  $y$  and  $z$  coordinate directions are assumed constant. The conservative numerical scheme for (18.42) is written as

$$\left. \begin{aligned} q_{i,j,k}^{n+1} &= q_{i,j,k}^n - \frac{\Delta t}{\Delta x} \left( f_{i+\frac{1}{2},j,k} - f_{i-\frac{1}{2},j,k} \right) \\ &\quad - \frac{\Delta t}{\Delta y} \left( g_{i,j+\frac{1}{2},k} - g_{i,j-\frac{1}{2},k} \right) \\ &\quad - \frac{\Delta t}{\Delta z} \left( h_{i,j,k+\frac{1}{2}} - h_{i,j,k-\frac{1}{2}} \right) \end{aligned} \right\} \quad (18.43)$$

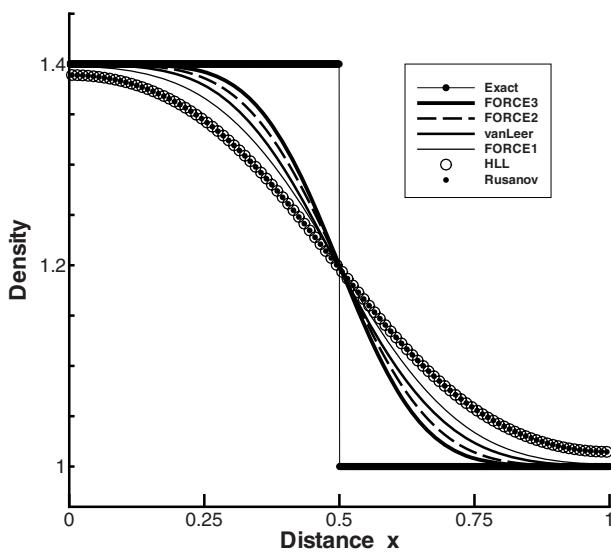
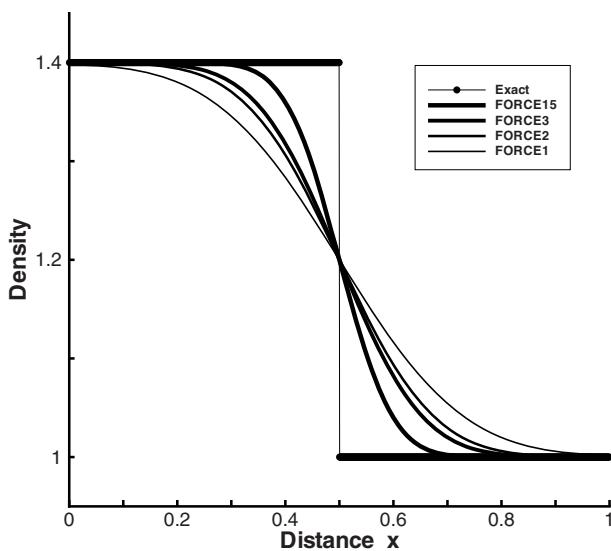
with FORCE numerical fluxes

$$\left. \begin{aligned} f_{i+\frac{1}{2},j,k} &= \frac{1}{2} \lambda_1 \left( q_{i,j,k}^n + q_{i+1,j,k}^n \right) - \frac{1}{2} \lambda_1 \left( \frac{\alpha^2 c_x^2 + 1}{2\alpha c_x} \right) \left( q_{i+1,j,k}^n - q_{i,j,k}^n \right), \\ g_{i,j+\frac{1}{2},k} &= \frac{1}{2} \lambda_2 \left( q_{i,j,k}^n + q_{i,j+1,k}^n \right) - \frac{1}{2} \lambda_2 \left( \frac{\alpha^2 c_y^2 + 1}{2\alpha c_y} \right) \left( q_{i,j+1,k}^n - q_{i,j,k}^n \right), \\ h_{i,j,k+\frac{1}{2}} &= \frac{1}{2} \lambda_3 \left( q_{i,j,k}^n + q_{i,j,k+1}^n \right) - \frac{1}{2} \lambda_3 \left( \frac{\alpha^2 c_z^2 + 1}{2\alpha c_z} \right) \left( q_{i,j,k+1}^n - q_{i,j,k}^n \right). \end{aligned} \right\} \quad (18.44)$$

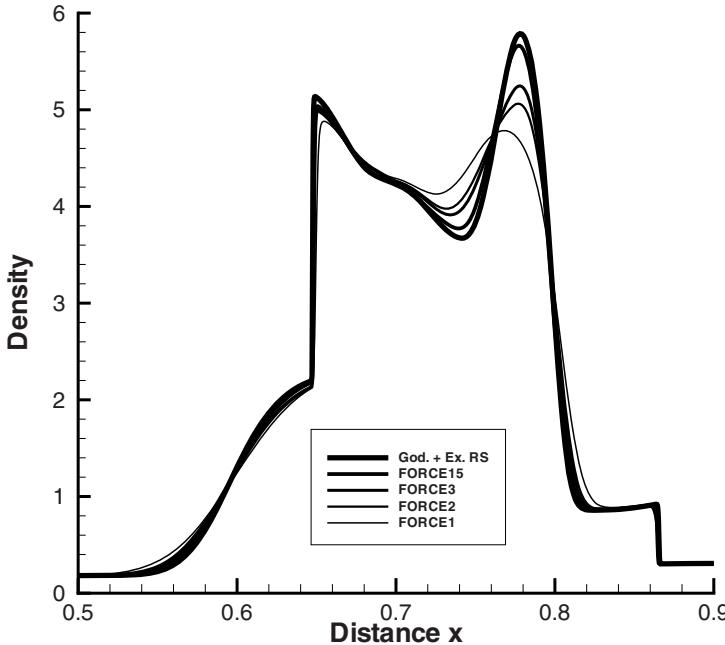
Here  $c_x = \frac{\lambda_1 \Delta t}{\Delta x}$ ,  $c_y = \frac{\lambda_2 \Delta t}{\Delta y}$  and  $c_z = \frac{\lambda_3 \Delta t}{\Delta z}$  are the directional Courant numbers, in the  $x$ ,  $y$  and  $z$  coordinate directions, respectively. Substitution of the numerical fluxes (18.44) into the conservative formula (18.43) yields the scheme

$$q_{i,j,k}^{n+1} = \sum_{l=-1}^1 \sum_{m=-1}^1 \sum_{r=-1}^1 \beta_{l,m,r} q_{i+l,j+m,k+r}^n, \quad (18.45)$$

with coefficients



**Fig. 18.4.** Stationary contact test. Comparison of numerical schemes with the exact solution. Top frame: results from the *FORCE $\alpha$*  schemes. Bottom frame: results from *FORCE $\alpha$*  schemes and from the two incomplete Riemann solvers Rusanov and HLL, and from the van Leer Flux Vector Splitting scheme.

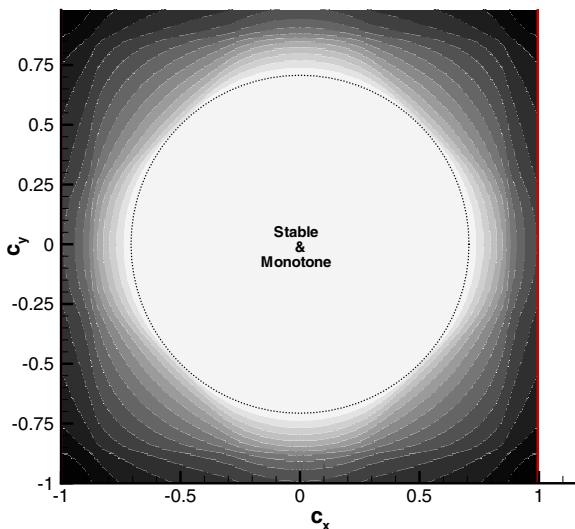


**Fig. 18.5.** Woodward and Colella test. Comparison of numerical results from the  $FORCE\alpha$  schemes with those from the first-order Godunov scheme with the exact Riemann solver.

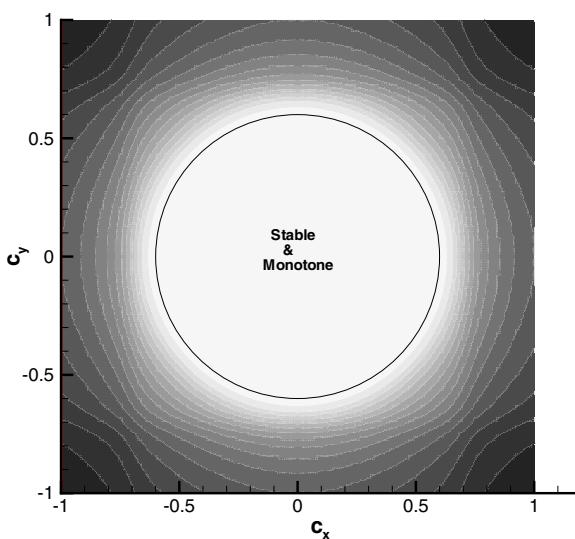
$$\left. \begin{aligned} \beta_{-1,0,0} &= \frac{1+\alpha c_x}{2\alpha} - \left( \frac{1-\alpha^2 c_x^2}{4\alpha} \right) , \quad \beta_{1,0,0} = \frac{1-\alpha c_x}{2\alpha} - \left( \frac{1-\alpha^2 c_x^2}{4\alpha} \right) , \\ \beta_{0,-1,0} &= \frac{1+\alpha c_y}{2\alpha} - \left( \frac{1-\alpha^2 c_y^2}{4\alpha} \right) , \quad \beta_{0,1,0} = \frac{1-\alpha c_y}{2\alpha} - \left( \frac{1-\alpha^2 c_y^2}{4\alpha} \right) , \\ \beta_{0,0,-1} &= \frac{1+\alpha c_z}{2\alpha} - \left( \frac{1-\alpha^2 c_z^2}{4\alpha} \right) , \quad \beta_{0,0,1} = \frac{1-\alpha c_z}{2\alpha} - \left( \frac{1-\alpha^2 c_z^2}{4\alpha} \right) , \\ \beta_{0,0,0} &= \frac{1}{2} (1 - \alpha c_x^2 - \alpha c_y^2 - \alpha c_z^2) . \end{aligned} \right\} \quad (18.46)$$

The monotonicity region of the scheme is found by requiring that the coefficients (18.46) be non-negative, which gives the following conditions

$$\left. \begin{aligned} c_x^2 + c_y^2 &\leq 1/2 && \text{in two space dimensions ,} \\ c_x^2 + c_y^2 + c_z^2 &\leq 1/3 && \text{in three space dimensions .} \end{aligned} \right\} \quad (18.47)$$



**Fig. 18.6.** Two-dimensional case. Linear stability and monotonicity region in the  $c_x - c_y$  plane.



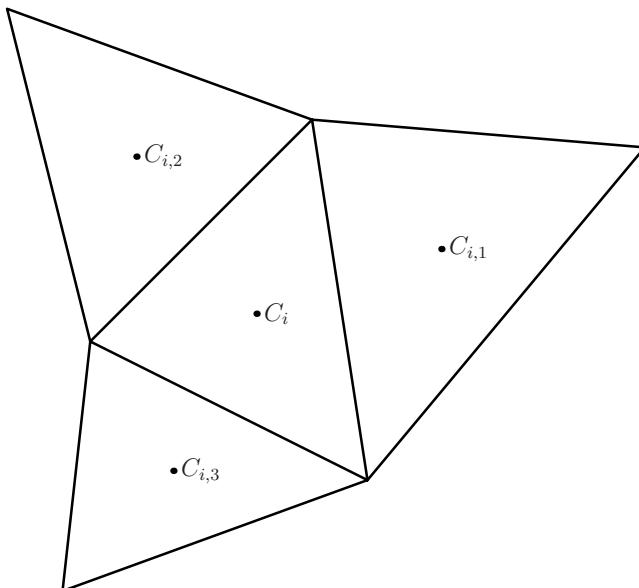
**Fig. 18.7.** Three-dimensional case. Linear stability and monotonicity region shown through a cut in the  $c_x - c_y$  plane.

Using the empirical method to study linear stability described in Section 16.4.2 of Chapter 16, see also [528] and [536], it is found that the linear stability region is also given by conditions (18.47) above. This is illustrated in Fig. 18.7. The top frame shows the two-dimensional case on the  $c_x$ - $c_y$  plane; the bottom frame illustrates the three-dimensional case, showing a cut through the  $c_x$ - $c_y$  plane.

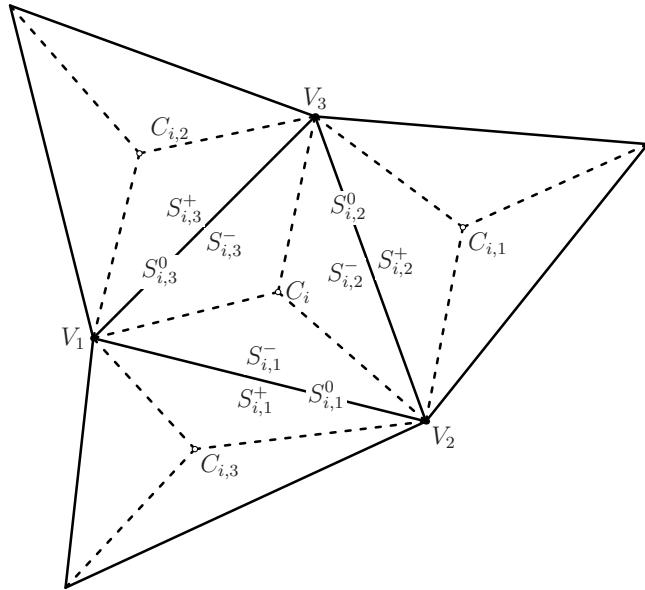
In the next section we construct FORCE schemes for general meshes in multiple space dimensions.

## 18.5 FORCE Schemes on General Meshes

Consider first the case of an unstructured mesh of triangular elements in two space dimensions. Fig. 18.8 depicts a general triangle  $C_i$  and its three neighbours having a common edge with  $C_i$ , denoted by  $C_{i,1}$ ,  $C_{i,2}$  and  $C_{i,3}$ . Recall that the construction of FORCE schemes requires the *primary mesh*, to compute cell averages via the conservative formula, and a *staggered mesh*, to define the numerical fluxes.



**Fig. 18.8.** Typical unstructured mesh of triangular elements in two space dimensions, called here the primary mesh. Cell  $C_i$  has three neighbours,  $C_{i,1}$ ,  $C_{i,2}$  and  $C_{i,3}$ .



**Fig. 18.9.** Primary mesh of triangular elements  $C_i$  and edge-based staggered mesh of elements  $S_{i,j} = S_{i,j}^- \cup S_{i,j}^+$ ,  $j = 1, 2, 3$ , and cell edges  $S_{i,j}^0 = S_{i,j}^- \cap S_{i,j}^+$ ,  $j = 1, 2, 3$ .

Fig. 18.9 shows the primary mesh represented by a cell  $C_i$ , its three neighbours  $C_{i,1}$ ,  $C_{i,2}$ ,  $C_{i,3}$ , and the edge-based staggered mesh of elements  $S_{i,j}$ ,  $j = 1, 2, 3$  shown in broken lines. The intercell boundaries of cell  $C_i$  are denoted by  $S_{i,j}^0$ ,  $j = 1, 2, 3$ .

For each intercell boundary  $S_{i,j}^0$ , visited in anticlockwise manner, there is an associated element of the staggered mesh

$$S_{i,j} = S_{i,j}^- \cup S_{i,j}^+, \quad j = 1, 2, 3, \quad (18.48)$$

where  $S_{i,j}^-$  is the portion inside  $C_i$  (left of edge  $j$ ) and  $S_{i,j}^+$  is the portion outside  $C_i$  (right of edge  $j$ ). See (18.22) and Fig. 18.3 for the two-dimensional Cartesian case. The intercell boundaries  $S_{i,j}^0$  are defined as

$$S_{i,j}^0 = C_i \cap C_{i,j}, \quad j = 1, 2, 3. \quad (18.49)$$

The elements of the staggered mesh are constructed as follows. For each edge  $S_{i,j}^0$ ,  $j = 1, 2, 3$ , we identify the barycentres of  $C_i$  (left) and of its right neighbour  $C_{i,j}$ . Then, a control volume is created around edge  $S_{i,j}^0$  by joining the two barycentres to the vertices of edge  $S_{i,j}^0$ . The procedure can be generalized to other types of meshes in two and three space dimensions.

In  $\alpha$  space dimensions we assume a conforming tessellation  $\mathcal{C}_\Omega$  of the computational domain  $\Omega \subset \mathbb{R}^\alpha$  by elements  $C_i$ , with

$$\mathcal{C}_\Omega = \cup_i C_i. \quad (18.50)$$

Each volume  $C_i$  has  $n_i$  plane faces  $S_{i,j}^0$  so that the outer boundary of  $C_i$  is

$$\partial C_i = \cup_{j=1}^{j=n_i} S_{i,j}^0 . \quad (18.51)$$

Each plane face  $S_{i,j}^0$  of cell  $C_i$  has a secondary volume as given by (18.48). The total volume of cell  $C_i$  is

$$|C_i| = \sum_{j=1}^{n_i} |S_{i,j}^-| . \quad (18.52)$$

Recall that for each (primary) control volume  $C_i$  one can define a cell average  $\mathbf{Q}_i^{n+1}$  at a time  $t = t^{n+1}$  by applying the integral form of the conservation laws, namely

$$\mathbf{Q}_i^{n+1} = \mathbf{Q}_i^n - \frac{\delta t}{|C_i|} \sum_{j=1}^{n_i} \int_{S_{i,j}^0} \bar{\mathbf{F}}(\mathbf{Q}) \cdot \mathbf{n}_j dA , \quad (18.53)$$

where  $\mathbf{Q}_i^n$  a cell average at time  $t = t^n$ ,  $\delta t = t^{n+1} - t^n$  is the evolution time,  $\mathbf{n}_j$  is the unit outward normal vector to face  $S_{i,j}^0$  and

$$\bar{\mathbf{F}}(\mathbf{Q}) \equiv [\mathbf{F}(\mathbf{Q}), \mathbf{G}(\mathbf{Q}), \mathbf{H}(\mathbf{Q})] .$$

As seen for the Cartesian case in Section 3, there are three stages in the construction of the FORCE schemes, as detailed below.

- *I: Averages on staggered mesh.* We assume the cell averages  $\mathbf{Q}_i^n$  in each cell  $C_i$  as initial conditions at time  $t = t^n$  and then apply the integral form of the conservation laws to each, edge-based, secondary volume  $S_{i,j}$  to obtain

$$\mathbf{Q}_{i,j}^{n+\frac{1}{2}} = \frac{\mathbf{Q}_i^n |S_{i,j}^-| + \mathbf{Q}_j^n |S_{i,j}^+|}{|S_{i,j}^-| + |S_{i,j}^+|} - \frac{1}{2} \frac{\Delta t |S_{i,j}^0|}{|S_{i,j}^-| + |S_{i,j}^+|} (\bar{\mathbf{F}}(\mathbf{Q}_j^n) - \bar{\mathbf{F}}(\mathbf{Q}_i^n)) \cdot \mathbf{n}_j . \quad (18.54)$$

- *II: Averages on primary mesh.* Assuming initial conditions  $\mathbf{Q}_{i,j}^{n+\frac{1}{2}}$  in each boundary-based secondary control volume we again apply the integral form of the conservation laws on the primary volume  $C_i$ , for another time  $\frac{1}{2}\Delta t$ , to obtain

$$\mathbf{Q}_i^{n+1} = \frac{1}{|C_i|} \sum_{j=1}^{n_i} \left( \mathbf{Q}_{i,j}^{n+\frac{1}{2}} S_{i,j}^- - \frac{1}{2} \Delta t |S_{i,j}^0| \bar{\mathbf{F}}(\mathbf{Q}_{i,j}^{n+\frac{1}{2}}) \cdot \mathbf{n}_j \right) . \quad (18.55)$$

Relations (18.54)–(18.55) are effectively a first-order accurate, explicit two-step method for solving (18.20) on a staggered mesh.

- *III: Conservative form and the FORCE flux.* The FORCE scheme can now be written as a one-step scheme in conservative form on a non-staggered mesh, with a corresponding numerical flux. After some algebraic manipulations of (18.55), involving Gauss' theorem and normalizing the face-normal vectors, we can recast the scheme (18.54)–(18.55) into the sought conservative form

$$\mathbf{Q}_i^{n+1} = \mathbf{Q}_i^n - \frac{\Delta t}{|\mathbf{C}_i|} \sum_{j=1}^{n_i} |S_{i,j}^0| \bar{\mathbf{F}}_{i,j}^{force\alpha} \cdot \mathbf{n}_j , \quad (18.56)$$

where the resulting FORCE flux is defined as

$$\bar{\mathbf{F}}_{i,j}^{force\alpha} = \frac{1}{2} \left( \bar{\mathbf{F}}_{i,j}^{lw\alpha}(\mathbf{Q}_i^n, \mathbf{Q}_j^n) + \bar{\mathbf{F}}_{i,j}^{lf\alpha}(\mathbf{Q}_i^n, \mathbf{Q}_j^n) \right) . \quad (18.57)$$

The FORCE flux on general meshes in multiple space dimensions is again the arithmetic average of two fluxes: a two-point flux of the Lax–Wendroff type and a two-point flux of the Lax–Friedrichs type. These two component fluxes appear to be new and are natural generalizations of the one-dimensional Lax–Wendroff and Lax–Friedrichs fluxes to general meshes in multiple space dimensions. The Lax–Wendroff type flux is given by the physical flux (tensor)  $\bar{\mathbf{F}}$  evaluated at the intermediate state (18.54) obtained from the first averaging procedure, namely

$$\bar{\mathbf{F}}_{i,j}^{lw\alpha}(\mathbf{Q}_i^n, \mathbf{Q}_j^n) = \bar{\mathbf{F}}(\mathbf{Q}_{i,j}^{n+\frac{1}{2}}) , \quad (18.58)$$

with

$$\mathbf{Q}_{i,j}^{n+\frac{1}{2}} = \frac{\mathbf{Q}_i^n |S_{i,j}^{(-)}| + \mathbf{Q}_j^n |S_{i,j}^{(+)}|}{|S_{i,j}^{(-)}| + |S_{i,j}^{(+)}|} - \frac{1}{2} \frac{\Delta t |S_{i,j}^{(0)}|}{|S_{i,j}^{(-)}| + |S_{i,j}^{(+)}|} (\bar{\mathbf{F}}(\mathbf{Q}_j^n) - \bar{\mathbf{F}}(\mathbf{Q}_i^n)) \cdot \mathbf{n}_j . \quad (18.59)$$

The Lax–Friedrichs type flux for general meshes in multiple space dimensions is defined as follows

$$\begin{aligned} \bar{\mathbf{F}}_{i,j}^{lf\alpha}(\mathbf{Q}_i^n, \mathbf{Q}_j^n) &= \frac{|S_{i,j}^{(+)}| \bar{\mathbf{F}}(\mathbf{Q}_i^n) + |S_{i,j}^{(-)}| \bar{\mathbf{F}}(\mathbf{Q}_j^n)}{|S_{i,j}^{(-)}| + |S_{i,j}^{(+)}|} \\ &\quad - \frac{2}{\Delta t |S_{i,j}^{(0)}|} \frac{|S_{i,j}^{(-)}| |S_{i,j}^{(+)}|}{|S_{i,j}^{(-)}| + |S_{i,j}^{(+)}|} (\mathbf{Q}_j^n - \mathbf{Q}_i^n) \cdot \mathbf{n}_j . \end{aligned} \quad \left. \right\} \quad (18.60)$$

**Summary of the FORCE scheme:** The implementation of the FORCE scheme to solve (18.20) in three space dimensions on unstructured meshes is straightforward. The scheme consists of the one-step explicit conservative formula (18.56) with numerical fluxes given by (18.57). Further details on the FORCE method just described are found in [536].

In the next section we illustrate the performance of the method, as applied to multidimensional test problems for the Euler equations on unstructured meshes.

## 18.6 Sample Numerical Results

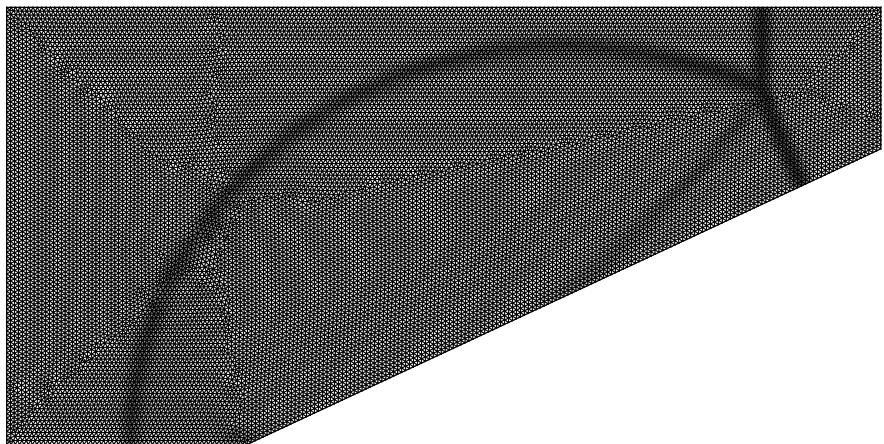
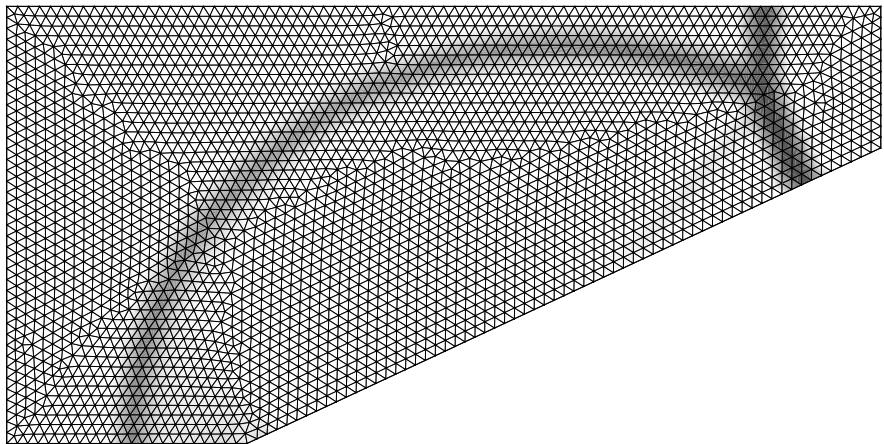
Here we apply the FORCE methods to the Mach reflection test problem described in Chapter 17. The problem consists of a single shock wave of shock Mach number 1.7, travelling along a two-dimensional channel and impinging on a wedge at an angle of  $\phi = 25$  degrees to the direction of the shock propagation. Fig. 17.9 of Chapter 17 shows reference solutions, including an experimental result for comparison. Here we solve the two-dimensional Euler equations for an ideal gas with  $\gamma = 1.4$ . We use unstructured triangular meshes and show results for the basic first-order FORCE scheme and for high-order extensions using the ADER approach in the framework of finite volumes with WENO non-linear reconstructions.

Fig. 18.10 shows the unstructured triangular meshes used along the numerical results for density. The mesh of the top frame has size 1/20 and that of the bottom frame has size 1/50. The numerical solutions are obtained with the ADER third-order extension of FORCE.

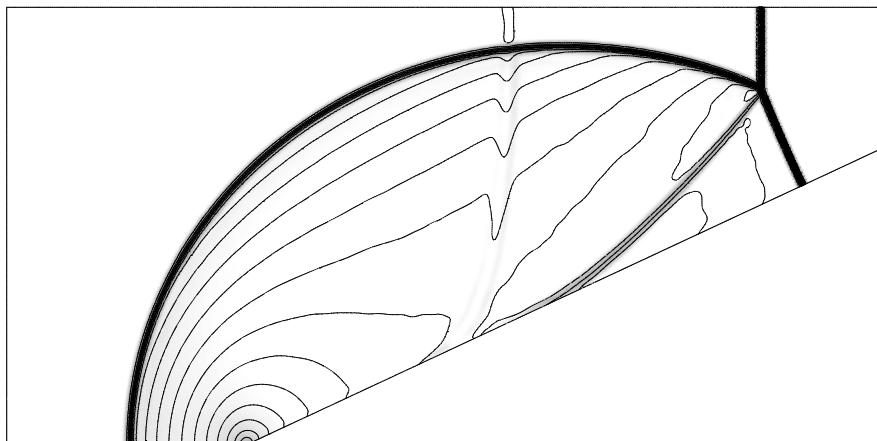
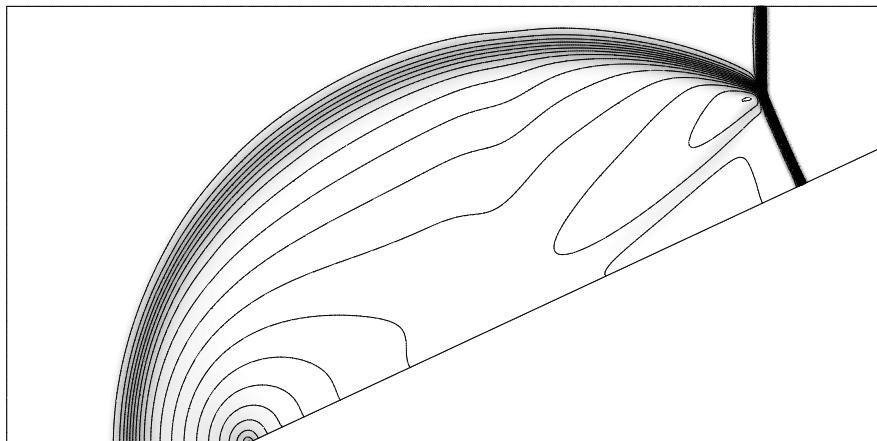
Fig. 18.11 shows the numerical results for density for a finer mesh of size 1/100. The top frame shows the numerical solution obtained with the first-order FORCE flux and the bottom frame shows the numerical solution are obtained with the third-order extension of the FORCE flux using ADER.

## 18.7 Concluding Remarks

In this chapter we have presented a generalization of the one-dimensional FORCE flux of Chapter 7 to multiple space dimensions, for solving hyperbolic equations in conservation-law form on general meshes. The schemes are linearly stable and monotone for a sufficiently generous range of Courant numbers. The FORCE schemes are first-order accurate and constitute a building block for constructing high-order schemes in the framework of existing methods, such as finite volume methods and discontinuous Galerkin finite element methods. Implementations can follow semi-discrete approaches or fully discrete approaches. In particular, in [536], the FORCE methods just described were extended to high-order of accuracy in space and time for two and three dimensional problems in the framework of ADER finite volume and DG methods; implementations on unstructured meshes of schemes of up to sixth order of accuracy in space and time were reported. The main feature of the proposed scheme is simplicity; it does not require explicit knowledge of the eigenstructure of the system to solve, nor the availability of a classical Riemann solver. This is attractive for very complex hyperbolic systems, for which a Riemann solver may not be available, or if available it might not be suitable.



**Fig. 18.10.** Shock reflection problem for shock Mach number 1.7 and wedge angle  $\phi = 25$  degrees. Results from third-order ADER extension of FORCE. Top frame: density contours and unstructured mesh of size 1/20. Bottom frame: density contours and unstructured mesh of size 1/50 (Courtesy of Dr. M. Dumbser).



**Fig. 18.11.** Shock reflection problem for shock Mach number 1.7 and wedge angle  $\phi = 25$  degrees. Solution for density on mesh of size  $1/100$ . Top frame: solution from first-order FORCE scheme. Bottom frame: solution from third-order ADER extension of FORCE (Courtesy of Dr. M. Dumbser).

---

# The Generalized Riemann Problem

In this chapter we study a technique for solving the generalized Riemann problem for systems of non-linear hyperbolic partial differential equations with source terms. The generalized Riemann problem studied here is a twofold generalization of the classical Riemann problem studied in previous chapters, namely: (i) the two vector fields that define the initial conditions are arbitrary but smooth away from the interface and (ii) the governing hyperbolic equations include source terms. We note that in the literature this Cauchy problem has also been termed Derivative Riemann Problem and High-Order Riemann Problem. In this book we shall adopt the terminology *Generalized Riemann Problem* and will be denoted by  $GRP_K$ . Here  $K$  is an arbitrary non-negative integer and stands for the order of the approximation. The solution of  $GRP_K$  can be used to construct one-step, fully discrete finite volume and discontinuous Galerkin finite element methods of  $(K + 1)$ -th order of accuracy in both space and time. An introduction to such generalized methods is given in the subsequent chapter.

## 19.1 Introduction

We consider a general non-linear system of hyperbolic balance laws in a given domain  $\Omega$ , along with general initial condition  $\mathbf{Q}_0(x)$

$$\left. \begin{array}{l} \text{PDEs: } \partial_t \mathbf{Q} + \partial_x \mathbf{F}(\mathbf{Q}) = \mathbf{S}(\mathbf{Q}), \quad x \in \Omega, \quad t > 0, \\ \text{IC: } \mathbf{Q}(x, 0) = \mathbf{Q}_0(x). \end{array} \right\} \quad (19.1)$$

$\mathbf{Q}$  is the vector of conserved variables,  $\mathbf{F}(\mathbf{Q})$  is the vector of fluxes and  $\mathbf{S}(\mathbf{Q})$  is the vector of sources, which does not contain differential terms of the unknown  $\mathbf{Q}$ . One class of numerical methods to solve (19.1) approximately is the finite volume method, which may be constructed by considering a control volume  $V = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [t^n, t^{n+1}]$  in the computational domain  $\Omega \times [0, T]$ , where

$T$  is an output time. Here  $x_{i-\frac{1}{2}}$  and  $x_{i+\frac{1}{2}}$  are the left and right intercell boundaries of the computational cell  $I_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ , with cell centre given by  $x_i = \frac{1}{2}(x_{i-\frac{1}{2}} + x_{i+\frac{1}{2}})$  and cell length given by  $\Delta x = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$ ;  $t^n$  and  $t^{n+1}$  are two successive times, with  $\Delta t = t^{n+1} - t^n$  being the time step. Integration of the governing equations in (19.1) in space and time in the control volume  $V$  yields the *exact* relation

$$\mathbf{Q}_i^{n+1} = \mathbf{Q}_i^n - \frac{\Delta t}{\Delta x} (\mathbf{F}_{i+\frac{1}{2}} - \mathbf{F}_{i-\frac{1}{2}}) + \Delta t \mathbf{S}_i , \quad (19.2)$$

where

$$\mathbf{Q}_i^n = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathbf{Q}(x, t^n) dx \quad (19.3)$$

is the spatial integral average of  $\mathbf{Q}(x, t)$  at time  $t = t^n$ , also called cell average; an analogous definition applies to  $\mathbf{Q}_i^{n+1}$ . Obviously the cell average is a constant but the integrand  $\mathbf{Q}(x, t^n)$  may admit a high-order distribution within the cell  $I_i$ , a polynomial distribution, for example. Computationally, such distribution is furnished by a reconstruction procedure using the set of cell averages  $\{\mathbf{Q}_i^n\}$ . For background on reconstruction procedures see, for example, [242], [270] and [25]; see also [175] and [176] for recent developments. The remaining terms in (19.2) are the integrals

$$\left. \begin{aligned} \mathbf{F}_{i+\frac{1}{2}} &= \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} \mathbf{F}(\mathbf{Q}(x_{i+\frac{1}{2}}, t)) dt , \\ \mathbf{S}_i &= \frac{1}{\Delta t \Delta x} \int_{t^n}^{t^{n+1}} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathbf{S}(\mathbf{Q}(x, t)) dx dt . \end{aligned} \right\} \quad (19.4)$$

$\mathbf{F}_{i+\frac{1}{2}}$  is the temporal integral average of  $\mathbf{F}(\mathbf{Q}(x, t))$  right at the interface  $x = x_{i+\frac{1}{2}}$  and  $\mathbf{S}_i$  is the space-time integral average of the source  $\mathbf{S}(\mathbf{Q}(x, t))$  within the control volume  $V$ .

A numerical method is obtained when in formula (19.2) one supplies suitable approximations for  $\mathbf{F}_{i+\frac{1}{2}}$  and  $\mathbf{S}_i$ . These approximations are called, respectively, the *numerical flux* and the *numerical source*, and are constructed by first finding an approximation  $\mathbf{Q}_{LR}(\tau) \approx \mathbf{Q}(x_{i+\frac{1}{2}}, t)$  at the interface, as a function of time, and an approximation  $\mathbf{Q}_i(x, t) \approx \mathbf{Q}(x, t)$  within the control volume, as a function of both space and time. Then, suitable numerical integration schemes are required to evaluate the integrals in (19.4) to finally obtain the numerical flux and the numerical source. Formula (19.2) can then be used numerically to evolve approximations to cell averages.

Most of the discussion regarding the construction of numerical methods centres on the definition of the numerical flux, which brings in the initial-value problem, or Cauchy problem, called the Riemann problem. The *classical* Riemann problem for a system of conservation laws is the Cauchy problem with

*piece-wise constant data*, that is the Cauchy problem with initial condition consisting of two *constant* states separated by a discontinuity. In chapter 4 we studied a method to solve this problem exactly, as applied to the Euler equations; in chapters 7 to 12 we presented techniques to solve the problem approximately. Godunov's first-order accurate method [216] is based on the exact or approximate solutions of this classical Riemann problem, see chapter 6. For most well-known hyperbolic systems one can compute the exact solution of this classical Riemann problem throughout the half plane  $t > 0$ ,  $-\infty < x < \infty$ , if desired, although for the purpose of computing an inter-cell flux for the Godunov method knowing the solution at the interface is sufficient.

In this chapter we study a generalization of the classical, piece-wise constant data, Riemann problem. The generalization considered here is twofold: (i) the two vector fields that define the initial conditions are arbitrary but smooth away from the interface and (ii) the governing hyperbolic equations include source terms. This generalized Riemann problem will be denoted by  $GRP_K$ , with  $K$  a non-negative integer representing the order of the approximation of the solution. Finding the full solution of this problem in the half-plane is probably an impossible task, in general. However for the purpose of constructing fully-discrete, one-step Godunov schemes of very high order of accuracy one only requires the solution of  $GRP_K$  right at the interface, as a function of time. Most available works on this subject refer to homogeneous systems and the simplest case is that obtained by replacing the piece-wise constant initial condition of the classical Riemann problem by a piece-wise linear initial condition. This particular Cauchy problem, which corresponds to  $GRP_1$  in the present notation, has also been termed the *generalized Riemann problem* or simply GRP, by some authors. Finding the solution of this problem in the half-plane is a difficult, or impossible task. However, for the Euler equations for ideal gases a complete solution was reported by Menshov [350]. Godunov-type numerical schemes of second order of accuracy can be constructed by solving approximately the  $GRP_1$  only at the interface. It seems as if the first researcher to construct one of such methods was Kolgan [287]; see also [289] and [288]. The well-known works of van Leer then followed, see for instance [554] and [555]; see also the recent article by van Leer [564] referring to the pioneering work of Kolgan. Subsequent results on second-order methods are also reported in [37], [76], [350], [133] and [499], to name but a few. Some of these methods are studied in chapters 14 and 16. Ben-Artzi and Falcovitz [37] solved the  $GRP_1$  approximately in a manner that is different to previous works. They also used the approximate solution to construct a second-order method. See also the related works [38], [36], [41], [40] and the recent textbook on the method of Ben-Artzi and Falcovitz [39].

The more general, still homogeneous, Cauchy problem with initial conditions consisting of two arbitrary but smooth vector fields separated by a discontinuity at the origin, has also been studied. Early results on the subject are found for example in [483], [236], [304], [76] and [305]. The high-order fully

discrete method first proposed by Harten and collaborators [242] also makes use of the solution of the generalized Riemann problem, although at first sight this might not be apparent. In the recent paper [91], a re-interpretation of the method of Harten et al. [242] is made in terms of the solution of the generalized Riemann problem at the interface and comparison is made with other recent solvers. The full generalized Riemann problem, including source terms, denoted here by  $GRP_K$ , was studied by Toro and Titarev [543], see also [523]. They used the solution of  $GRP_K$  to construct ADER methods; these are finite volume methods of the Godunov type, of arbitrary order of accuracy in space and time; see also [492]. A variant of the solution method of Harten and collaborators [242] for the full, inhomogeneous  $GRP_K$  problem has recently been reported in [173]; this solution method for  $GRP_K$  is able to deal with stiff source terms, while retaining high accuracy.

In this chapter we first study in complete detail the method proposed in [543] and [523] to solve the generalized Riemann problem  $GRP_K$  at the interface. The technique is a generalization of the method presented in section 13.4.4 of chapter 13 for the case  $K = 1$ , see [516] and [517]. The sought time-dependent solution at the interface is first expressed as a power series expansion of order  $K$ , as suggested, for example in [236], [304] and [305]. The leading term of the expansion is the solution of a classical, usually non-linear, Riemann problem, for which the techniques studied in chapters 6 and 9 apply. To evaluate the higher-order terms we utilize the Cauchy–Kowalewski procedure to express temporal derivatives in terms of functions of spatial derivatives. Then, in order to define the spatial derivatives at the interface we first construct new evolution equations for these and then solve additional classical Riemann problems whose solutions define all spatial derivatives at the interface. The coefficients of the expansion are then found by computing the functions of spatial derivatives and the complete series is then assembled to produce the final solution. In this manner this method of solution of the generalized Riemann problem of order  $K$  boils down to solving 1 classical non-linear Riemann problem for the leading term and  $K$  classical linear Riemann problems for spatial derivatives; see section 13.4.4 of chapter 13 for an introduction to the second-order case,  $K = 1$ . For smooth initial conditions the method becomes an application of the Cauchy–Kowalewski method. This case will rarely appear at cell interfaces but it will certainly be present in the interior of each finite volume or element and will be of use when evaluating the *numerical source* to high order of accuracy. With the solution available at the interface as a function of time, the straightforward generalization of Godunov’s method allows the construction of fully-discrete one-step schemes of arbitrary order of accuracy in space and time; an introduction to these methods is given in the following chapter.

The rest of this chapter is organized as follows. In section 19.2 we define the mathematical problem. In section 3 we give some mathematical preliminaries related to the Cauchy–Kowalewski theorem. In section 19.4 we study in detail the solution technique of [543] and [523] for the generalized Riemann problem

$GRP_K$  for a general hyperbolic system of balance laws. In section 19.5 we illustrate the details of the technique by solving in full detail the  $GRP_K$  for scalar equations with source terms. In section 6 we briefly discuss alternative methods of solution. Concluding remarks are found in section 19.7.

## 19.2 Statement of the Problem

We study the generalized Riemann problem  $GRP_K$  for hyperbolic systems of balance laws, defined as the initial-value problem

$$\left. \begin{aligned} \text{PDEs: } & \partial_t \mathbf{Q} + \partial_x \mathbf{F}(\mathbf{Q}) = \mathbf{S}(\mathbf{Q}), \quad x \in (-\infty, \infty), \quad t > 0, \\ \text{ICs: } & \mathbf{Q}(x, 0) = \begin{cases} \mathbf{Q}_L(x) & \text{if } x < 0, \\ \mathbf{Q}_R(x) & \text{if } x > 0. \end{cases} \end{aligned} \right\} \quad (19.5)$$

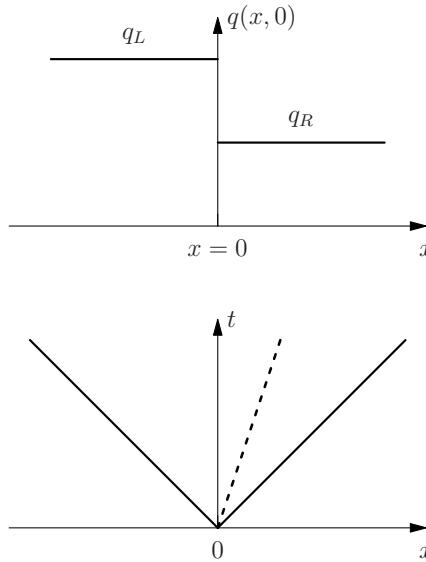
The initial conditions (ICs)  $\mathbf{Q}_L(x)$ ,  $\mathbf{Q}_R(x)$  are two vectors, the components of which are assumed to be smooth functions of  $x$ , with  $K$  continuous non-trivial spatial derivatives away from zero. That is, if the non-negative integers  $K_L$  and  $K_R$  are such that

$$\frac{d^k}{dx^k} \mathbf{Q}_L(x) \equiv 0 \quad \forall k > K_L, \quad \forall x < 0, \quad \frac{d^k}{dx^k} \mathbf{Q}_R(x) \equiv 0 \quad \forall k > K_R, \quad \forall x > 0,$$

then we call *generalized Riemann Problem* of order  $K$ , denoted by  $GRP_K$ , the Cauchy problem (19.5) with  $K = \max\{K_L, K_R\}$ . For example,  $GRP_0$  means that all first and higher-order spatial derivatives of the initial condition for the generalized Riemann problem away from the origin vanish identically; this case corresponds to the classical piece-wise constant data Riemann problem. Similarly,  $GRP_1$  means that all second and higher-order spatial derivatives of the initial condition for the generalized Riemann problem away from the origin vanish identically; this case corresponds to the piece-wise linear data Riemann problem, the problem studied, for example, in [37] and [39].

Fig. 19.1 illustrates the classical Riemann problem  $GRP_0$ ; the top frame shows the piece-wise constant initial condition for a single component  $q$  of the vector  $\mathbf{Q}(x, 0)$ , while the bottom frame shows the corresponding wave structure of the solution in the  $x-t$  plane, as for a  $3 \times 3$  system. The solution of this classical problem is a similarity solution that depends on the ratio  $x/t$  and characteristics are straight lines. The complete solution in the half-plane  $t > 0$ ,  $-\infty < x < \infty$  contains a set of constant wedge-like regions that are separated by discontinuities or smooth transitions.

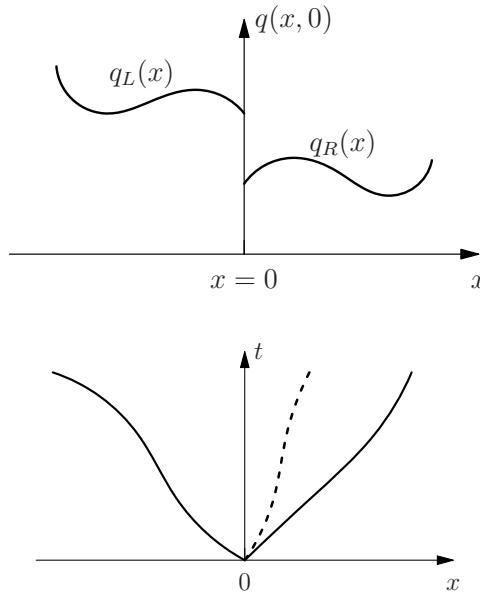
Fig. 19.2 illustrates the generalized Riemann problem  $GRP_K$ , for  $K > 0$ ; the top frame shows the piece-wise smooth initial condition for a single component  $q$  of the vector  $\mathbf{Q}(x, 0)$ , while the bottom frame shows the corresponding wave structure of the solution in the  $x-t$  plane. Now the solution no longer



**Fig. 19.1.** The classical Riemann problem for a typical  $3 \times 3$  hyperbolic system. Top frame shows piece-wise constant initial condition for a single component  $q$  of the vector  $\mathbf{Q}(x, 0)$ . Bottom frame shows the corresponding wave structure of the solution in the  $x$ - $t$  plane. Characteristics are straight lines.

contains regions of constant values and characteristics are curved lines. It is assumed however, that for a sufficiently short time, the solution of  $GRP_K$  (Fig. 19.2) has fundamentally the same structure as the solution of  $GRP_0$  (Fig. 19.1); the solution of  $GRP_K$  may be regarded as a perturbation of the solution of  $GRP_0$ .

The aim of this chapter is to study a semi-analytical method to find the solution of  $GRP_K$ , given by (19.5), at the origin  $x = 0$  and for  $t > 0$ , as a function of time. In order to solve numerically a system of hyperbolic balance laws using a high-order Godunov finite volume method of the form (19.2) one needs to compute (i) a *numerical flux* at each interface and (ii) a *numerical source* within each volume or element. To compute a numerical flux it is sufficient to find the solution  $\mathbf{Q}_{LR}(\tau)$  of the  $GRP_K$  at the interface  $x = 0$ , as a function of time  $\tau$  alone and then calculate an approximation to the temporal integral in (19.4).  $\mathbf{Q}_{LR}(\tau)$  will provide sufficient information to compute a numerical flux to construct a numerical scheme of  $(K+1)$ -th order of accuracy in both space and time. To compute a *numerical source* we need a high-order representation in space and time  $\mathbf{Q}_i(x, t)$  of the solution  $\mathbf{Q}(x, t)$  within the space-time volume to compute an approximation to the volume integral in (19.4) to the desired order of accuracy. This can be obtained in exactly the same manner as for the solution of the generalized Riemann problem, with the appropriate simplifications, resulting in the Cauchy-Kowalewski method, as



**Fig. 19.2.** The generalized Riemann problem  $GRP_K$  for a typical  $3 \times 3$  hyperbolic system. Top frame shows piece-wise smooth initial condition for a single component  $q$  of the vector  $\mathbf{Q}(x, 0)$ . Bottom frame shows the corresponding wave structure of the solution in the  $x$ - $t$  plane. Characteristics are curved lines.

pointed out earlier. We note that the classical case of piece-wise constant data reproduces the classical first-order upwind method of Godunov [216]. Before proceeding any further we provide some mathematical background, relevant to solving (19.5).

### 19.3 The Cauchy–Kowalewski Theorem

Here we provide some mathematical background centred on the classical Cauchy–Kowalewski theorem for the initial-value problem for differential equations. This theorem asserts the existence and uniqueness of an *analytic* solution of the problem provided that all functions involved are *analytic*. The solution is constructed as a power series expansion about the origin making use of the so-called Cauchy–Kowalewski procedure, whereby the governing equations and initial conditions are used repeatedly to determine the coefficients of the expansion. For detailed background on the Cauchy–Kowalewski theorem see for example [596], [186] and [383].

### 19.3.1 Series Expansions and Analytic Functions

Assume  $f(x)$  to be a  $C^\infty$  function of a single variable  $x \in I$ , where  $I$  is an open sub-interval of the set of real numbers  $\mathcal{R}^1$ . Consider a point  $x_0$  in  $I$ , then the Taylor series expansion of  $f(x)$  about  $x_0$  is

$$\sum_{k=0}^{\infty} [f^{(k)}(x_0)] \frac{(x - x_0)^k}{k!}. \quad (19.6)$$

If the above Taylor series expansion converges to  $f(x) \forall x$  in a neighbourhood of  $x_0$  then  $f(x)$  is said to be *analytic* at  $x_0$ . If  $f(x)$  is analytic  $\forall x \in I$ , then  $f(x)$  said to be *analytic* in  $I$ . We remark that for an arbitrary  $C^\infty$  function the Taylor series (19.6) may not converge, or if it does, it may not converge to  $f(x)$ . Thus, analytic functions are very special  $C^\infty$  functions. A familiar example of an analytic function is

$$f(x) = e^x = \sum_{k=0}^{\infty} \frac{1}{k!} x^k.$$

Simple examples of analytic functions are polynomial functions. As a matter of fact, when it comes to numerical applications of the solution of the generalized Riemann problem, we shall, in the main, be concerned with initial conditions in (19.5) consisting of polynomials.

For a function  $f$  of  $n$  variables  $x_1, x_2, \dots, x_n$  the Taylor series expansion about a point  $\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$  is

$$\sum_{(\alpha_1, \dots, \alpha_n)} \frac{\partial_{x_1}^{(\alpha_1)} \dots \partial_{x_n}^{(\alpha_n)} f(x_1^{(0)}, \dots, x_n^{(0)})}{\alpha_1! \dots \alpha_n!} (x_1 - x_1^{(0)})^{\alpha_1} \dots (x_n - x_n^{(0)})^{\alpha_n}, \quad (19.7)$$

where the summation is taken over all  $n$ -tuples  $(\alpha_1, \alpha_2, \dots, \alpha_n)$  of non-negative integers  $\alpha_i$ . For the  $\alpha_j$ -th order partial derivative of  $f$  with respect to  $x_i$  we use the notation

$$\partial_{x_i}^{(\alpha_j)} f(x_1, \dots, x_n) = \frac{\partial^{\alpha_j} f(x_1, \dots, x_n)}{\partial x_i^{\alpha_j}}. \quad (19.8)$$

As an example we consider a function  $f(x_1, x_2)$  of two variables. To second-order, we have the following expression for the Taylor series expansion

$$\left. \begin{aligned} f(x_1, x_2) &= f(x_1^{(0)}, x_2^{(0)}) \\ &+ \partial_{x_1} f(x_1^{(0)}, x_2^{(0)})(x_1 - x_1^{(0)}) \\ &+ \partial_{x_2} f(x_1^{(0)}, x_2^{(0)})(x_2 - x_2^{(0)}) \\ &+ \frac{1}{2} \partial_{x_1}^{(2)} f(x_1^{(0)}, x_2^{(0)})(x_1 - x_1^{(0)})^2 \\ &+ \partial_{x_1}^{(1)} \partial_{x_2}^{(1)} f(x_1^{(0)}, x_2^{(0)})(x_1 - x_1^{(0)})(x_2 - x_2^{(0)}) \\ &+ \frac{1}{2} \partial_{x_2}^{(2)} f(x_1^{(0)}, x_2^{(0)})(x_2 - x_2^{(0)})^2. \end{aligned} \right\} \quad (19.9)$$

In the next section we study a method to solve differential equations that is based on Taylor series expansions.

### 19.3.2 Illustration of the Cauchy–Kowalewski Theorem

Here we state the Cauchy–Kowalewski theorem for the special case of a first order partial differential equation

$$\text{PDE: } \partial_t q = G(x, t, q, \partial_x q) , \quad (19.10)$$

with initial condition

$$\text{IC: } q(x, 0) = q_0(x) . \quad (19.11)$$

The function  $G(v_1, v_2, v_3, v_4)$  in (19.10) is a function of four variables defined on a subset of  $\mathcal{R}^4$ . The initial condition (19.11) is assumed to be defined on an interval  $I$  of the real line containing the origin  $x = 0$ .

The Cauchy–Kowalewski theorem states that the solution  $q(x, t)$  of the initial–value problem (19.10)–(19.11) for  $(x, t)$  in a neighbourhood of  $(0, 0)$  is given by

$$q(x, t) = \sum_{(\alpha_1, \alpha_2)} \frac{\partial_x^{(\alpha_1)} \partial_t^{(\alpha_2)} q(0, 0)}{\alpha_1! \alpha_2!} x^{\alpha_1} t^{\alpha_2} , \quad (19.12)$$

if  $q_0(x)$  is analytic in a neighbourhood of the origin and  $G$  is analytic in a neighbourhood of  $(0, 0, q_0(0), q_0^{(1)}(0))$ , where  $q_0^{(1)}(x) \equiv \frac{d}{dx} q_0(x)$ . In addition, the theorem asserts that the solution is analytic and unique. See [596], [186] and [383].

As part of the proof of existence, the solution is constructed by computing all the coefficients of the expansion, for which one needs all the partial derivatives

$$\partial_x^{(\alpha_1)} \partial_t^{(\alpha_2)} q(x, t) \quad (19.13)$$

evaluated at the origin. In computing these partial derivatives one makes repeated use of both the governing PDE (19.10) and the initial condition (19.11) in what is known as the *Cauchy–Kowalewski procedure*, which is illustrated below via an example. In a numerical context, Cauchy–Kowalewski procedure is also known as the Lax–Wendroff procedure.

### 19.3.3 The Cauchy–Kowalewski Method

We first illustrate the application of the Cauchy–Kowalewski method to solve a simple initial–value problem and then briefly outline the application of the method to compute a high–order space–time distribution of the solution within a control volume to compute the numerical source in (19.2).

#### An Example

Consider the Cauchy problem

$$\left. \begin{aligned} \text{PDE: } & \partial_t q + \partial_x f(q) = 0, \quad x \in (-\infty, \infty), \quad t > 0, \\ \text{IC: } & q(x, 0) = q_0(x), \end{aligned} \right\} \quad (19.14)$$

where  $f(q)$  is the *flux function* and  $G(x, t, q, \partial_x q) = -\partial_x f(q)$  in (19.10). As an example, we compute the Cauchy–Kowalewski solution of the form (19.12) in a neighbourhood of the origin  $(0, 0)$  to second order of accuracy, for which we need to evaluate each term of the following expression

$$\left. \begin{aligned} q(x, t) = & q(0, 0) \\ & + \partial_t q(0, 0)t + \partial_x q(0, 0)x \\ & + \frac{1}{2}\partial_x^{(2)}q(0, 0)x^2 + \partial_t^{(1)}\partial_x^{(1)}q(0, 0)xt + \frac{1}{2}\partial_t^{(2)}q(0, 0)t^2. \end{aligned} \right\} \quad (19.15)$$

The leading term is simply

$$q(0, 0) = q_0(0). \quad (19.16)$$

The higher order terms will be computed in terms of functions of  $q_0(0)$  and its spatial derivatives. We first introduce the notation

$$q^{(k)}(0, 0) \equiv \partial_x^{(k)}q(0, 0) = \frac{d^k}{dx^k}q_0(0) \quad (19.17)$$

and

$$f^{(k)}(q) = \frac{d^k}{dq^k}f(q). \quad (19.18)$$

From the PDE, after using the chain rule, we have

$$\partial_t^{(1)}q(x, t) = -\partial_x^{(1)}f(q) = -f^{(1)}(q)\partial_x^{(1)}q. \quad (19.19)$$

We note that  $f^{(1)}(q)$  is the characteristic speed. For systems, this is the Jacobian matrix. The mixed derivative in (19.15) is obtained by differentiating (19.19) with respect to  $x$

$$\partial_x\partial_t q = -\partial_x \left[ f^{(1)}(q)\partial_x^{(1)}q \right],$$

which becomes

$$\partial_t\partial_x q = - \left[ f^{(1)}(q)\partial_x^{(2)}q + f^{(2)}(q)(\partial_x^{(1)}q)^2 \right]. \quad (19.20)$$

The second-order time derivative in (19.15) is obtained by differentiating (19.19) with respect to  $t$ , which after using (19.19) and (19.20) becomes

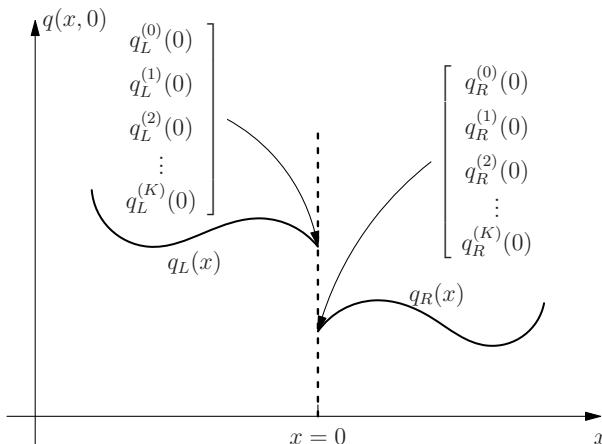
$$\partial_t^{(2)}q = f^{(1)}(q)f^{(2)}(q)\partial_x^{(1)}q + \left[ f^{(1)}(q) \right]^2(\partial_x^{(1)}q)^2 + f(q)f^{(1)}(q)\partial_x^{(2)}q. \quad (19.21)$$

**Remark:** each term in (19.15) can be expressed in terms of the initial value  $q_0(0)$  and space derivatives (19.17) of the initial condition (19.16). See (19.19)–(19.21).

## 19.4 A Method of Solution

We return to the Cauchy problem (19.5), the generalized Riemann problem  $GRP_K$ , in which the two initial states  $\mathbf{Q}_L(x)$  and  $\mathbf{Q}_R(x)$  are assumed to be smooth vector fields away from  $x = 0$ . For example, the components of the initial states may be polynomials of order at most  $K$ , defined respectively for  $x < 0$  and for  $x > 0$ , with discontinuities at  $x = 0$ . See Fig. 19.3. Away from  $x = 0$  we can use the Cauchy–Kowalewski method to construct a solution  $\mathbf{Q}(x, t)$  to (19.5), assuming, perhaps incorrectly, that all the requirements of the Cauchy–Kowalewski theorem are met.

Here we study a method to compute the solution of  $GRP_K$  given by (19.5), at  $x = 0$ , the interface, as a function of time. The method is based on the works [543] and [523] and is a generalization of the technique presented in section 13.4.4 of chapter 13 for the case  $K = 1$ , see [516] and [517]. Fig. 19.3. illustrates the initial conditions of  $GRP_K$  and the information available at  $t = 0$  at the interface  $x = 0$ , for a single component  $q(x, t)$  of the vector  $\mathbf{Q}(x, t)$ . The initial conditions  $\mathbf{Q}_L(x)$ ,  $\mathbf{Q}_R(x)$  are smooth away from  $x = 0$ , with one-sided function values and their spatial derivatives well defined and easily computed. At the interface  $x = 0$  however, all functions and their spatial derivatives are discontinuous, with left and right limits defining  $K + 1$  jumps  $\Delta^{(k)} q = q_R^{(k)} - q_L^{(k)}$ ,  $k = 0, 1, \dots, K$ , for each component  $q$  of the vector  $\mathbf{Q}$ . In the solution method the pairs  $(q_L^{(k)}, q_R^{(k)})$  will be the initial conditions for new (classical) Riemann problems, as we shall explain below.



**Fig. 19.3.** Illustration of the initial conditions for the generalized Riemann problem  $GRP_K$  for a single component  $q(x, t)$  of the vector of unknowns  $\mathbf{Q}(x, t)$ . The data  $q_L(x)$  and  $q_R(x)$  are smooth functions away from  $x = 0$  and have one-sided spatial derivatives at  $x = 0$ .

We seek a solution  $\mathbf{Q}_{LR}(\tau)$  of (19.5) at  $x = 0$  as the time power series expansion

$$\mathbf{Q}_{LR}(\tau) = \mathbf{Q}(0, 0_+) + \sum_{k=1}^K \left[ \partial_t^{(k)} \mathbf{Q}(0, 0_+) \right] \frac{\tau^k}{k!}, \quad (19.22)$$

where the leading term of the expansion is

$$\mathbf{Q}(0, 0_+) = \lim_{t \rightarrow 0_+} \mathbf{Q}(0, t).$$

The higher-order terms contain coefficients which are determined by the temporal partial derivatives  $\partial_t^{(k)} \mathbf{Q}(0, 0_+)$ . In what follows we describe a method to compute each of the terms of the series expansion.

#### 19.4.1 The Leading Term

The leading term  $\mathbf{Q}(0, 0_+)$  in the expansion (19.22) accounts for the *first-instant* interaction of the initial data via the governing PDEs in (19.5), which is realized solely by the boundary extrapolated values  $\mathbf{Q}_L(0)$  and  $\mathbf{Q}_R(0)$  of the ICs in (19.5). These two constant values constitute the initial data for a *classical Riemann problem*  $GRP_0$ , which is completely defined once appropriate evolution equations are specified. Trivially, the sought evolution equations are precisely those of the original hyperbolic system (19.5). Therefore, the leading term  $\mathbf{Q}(0, 0_+)$  is found from the *similarity solution* of the following homogeneous  $GRP_0$

$$\left. \begin{aligned} \text{PDEs: } & \partial_t \mathbf{Q} + \partial_x \mathbf{F}(\mathbf{Q}) = \mathbf{0}, \\ \text{ICs: } & \mathbf{Q}(x, 0) = \begin{cases} \mathbf{Q}_L(0) \equiv \lim_{x \rightarrow 0_-} \mathbf{Q}_L(x) & \text{if } x < 0, \\ \mathbf{Q}_R(0) \equiv \lim_{x \rightarrow 0_+} \mathbf{Q}_R(x) & \text{if } x > 0. \end{cases} \end{aligned} \right\} \quad (19.23)$$

We note that the source term  $\mathbf{S}(\mathbf{Q})$  in (19.5) is neglected here. However, the influence of the source terms is accounted for by the higher order terms in (19.22). Denoting the similarity solution of the  $GRP_0$  (19.23) by  $\mathbf{D}^{(0)}(x/t)$ , the sought leading term is given by evaluating this solution along the  $t$ -axis, that is along  $x/t = 0$ , namely

$$\mathbf{Q}(0, 0_+) = \mathbf{D}^{(0)}(0). \quad (19.24)$$

The value  $\mathbf{D}^{(0)}(0)$  is termed here the *Godunov state* and for the classical Riemann problem  $GRP_0$  corresponds to the numerical flux of the first-order upwind scheme of Godunov [216]. In what follows we shall extend the use of this terminology to mean the solution of classical Riemann problems for spatial derivatives evaluated at  $x/t = 0$ .

In practice, a *classical Riemann solver* is needed to obtain the leading term. See chapters 6 and 9. The determination of the higher-order terms in the expansion (19.22) is discussed next.

### 19.4.2 Higher-Order Terms

The coefficients of the higher order terms require the determination of the time derivatives. These in turn require the following steps: (i) expressing time derivatives in terms of spatial derivatives, (ii) construction of evolution equations for spatial derivatives and (iii) solution of classical Riemann problems for spatial derivatives. Below we treat each of these steps in detail.

#### Temporal and Spatial Derivatives

To compute the higher-order terms in (19.22) we need to obtain the partial derivatives

$$\partial_t^{(k)} \mathbf{Q}(0, 0_+) , \quad k = 1, 2, \dots, K . \quad (19.25)$$

If all time partial derivatives  $\partial_t^{(k)} \mathbf{Q}(x, t)$  were available on both sides of the initial discontinuity at  $x = 0$ , then one could implement a fairly direct approach to the evaluation of the higher order terms. The method presented in this section relies on the availability of all spatial derivatives away from the interface, since the initial data are given by two smooth functions of space, with all their one-sided spatial partial derivatives available. See Fig. 19.3.

We apply the Cauchy–Kowalewski procedure to express all *time* derivatives as *functions of space derivatives*, namely

$$\partial_t^{(k)} \mathbf{Q}(x, t) = \mathbf{P}^{(k)}(\partial_x^{(0)} \mathbf{Q}, \partial_x^{(1)} \mathbf{Q}, \dots, \partial_x^{(k)} \mathbf{Q}) . \quad (19.26)$$

We remark that the source term in (19.5) is also included in the functions  $\mathbf{P}^{(k)}$  in (19.26). For the simplest hyperbolic equation, the linear advection equation

$$\partial_t q(x, t) + \lambda \partial_x q(x, t) = 0 ,$$

we have

$$\left. \begin{aligned} \partial_t^{(1)} q(x, t) &= p^{(1)}(\partial_x^{(0)} q, \partial_x^{(1)} q) = -\lambda \partial_x^1 q(x, t) , \\ \partial_t^{(2)} q(x, t) &= p^{(2)}(\partial_x^{(0)} q, \partial_x^{(1)} q, \partial_x^{(2)} q) = \lambda^2 \partial_x^2 q(x, t) , \\ \partial_t^{(k)} q(x, t) &= p^{(k)}(\partial_x^{(0)} q, \partial_x^{(1)} q, \dots, \partial_x^{(k)} q) = (-\lambda)^k \partial_x^k q(x, t) . \end{aligned} \right\}$$

The temporal partial derivatives at  $x = 0$  for  $t > 0$  will only have a meaning if the functions  $\mathbf{P}^{(k)}$  can be given a meaning at  $x = 0$  for  $t > 0$ . This in turn depends on whether or not we can give a meaning to each of the arguments  $\partial_x^{(0)} \mathbf{Q}, \partial_x^{(1)} \mathbf{Q}, \dots, \partial_x^{(k)} \mathbf{Q}$  of  $\mathbf{P}^{(k)}$ , in other words, if we can give a meaning to all function values of  $\mathbf{Q}$  and its spatial derivatives at  $x = 0$  and  $t > 0$ .

Thus the basic problem is that of computing the spatial derivatives  $\partial_x^{(k)} \mathbf{Q}(0, 0_+)$ . At the initial time  $t = 0$ , for  $x < 0$  and for  $x > 0$  all spatial derivatives are defined, for which we introduce the notation

$$\mathbf{Q}_L^{(k)}(x) \equiv \frac{d^k}{dx^k} \mathbf{Q}_L(x), \quad \mathbf{Q}_R^{(k)}(x) \equiv \frac{d^k}{dx^k} \mathbf{Q}_R(x), \quad k = 1, 2, \dots, K. \quad (19.27)$$

At  $x = 0$  however, we have the *one-sided* derivatives

$$\left. \begin{array}{l} \mathbf{Q}_L^{(k)}(0) = \lim_{x \rightarrow 0^-} \mathbf{Q}_L^{(k)}(x) \\ \mathbf{Q}_R^{(k)}(0) = \lim_{x \rightarrow 0^+} \mathbf{Q}_R^{(k)}(x) \end{array} \right\} \quad k = 1, 2, \dots, K. \quad (19.28)$$

See Fig. 19.3. We thus have a set of  $K$  pairs  $(\mathbf{Q}_L^{(k)}(0), \mathbf{Q}_R^{(k)}(0))$  of constant vectors that could be used as the initial conditions for  $K$  classical Riemann problems, if in addition we had a set of corresponding evolution equations for the quantities  $\partial_x^{(k)} \mathbf{Q}(x, t)$ .

### Evolution equations for spatial derivatives

Fortunately, suitable evolution equations can easily be constructed. It can be verified that the quantity  $\partial_x^{(k)} \mathbf{Q}(x, t)$  obeys the following system of non-linear inhomogeneous evolution equations

$$\partial_t(\partial_x^{(k)} \mathbf{Q}(x, t)) + \mathbf{A}(\mathbf{Q}) \partial_x(\partial_x^{(k)} \mathbf{Q}(x, t)) = \mathbf{H}^{(k)}, \quad (19.29)$$

where the coefficient matrix  $\mathbf{A}(\mathbf{Q})$  is precisely the Jacobian matrix of system in (19.5). Equations (19.29) are obtained by manipulation of the PDEs in (19.5), with the right-hand side

$$\mathbf{H}^{(k)} = \mathbf{H}^{(k)}(\partial_x^{(0)} \mathbf{Q}, \partial_x^{(1)} \mathbf{Q}, \dots, \partial_x^{(k)} \mathbf{Q}) \quad (19.30)$$

being a function of  $\mathbf{Q}$  and all its spatial derivatives up to order  $k$ , including the source terms.

*Remark:* We note that for linear homogeneous systems with constant coefficient matrix  $\tilde{\mathbf{A}}$  the right-hand side  $\mathbf{H}^{(k)}$  in equations (19.29) vanishes identically and obviously  $\mathbf{A}(\mathbf{Q}) = \tilde{\mathbf{A}}$ . In other words, the spatial derivatives of all orders obey the original systems of PDEs identically.

To illustrate the situation we consider two examples. The simplest equation is the linear advection equation, for which it can easily be verified that the  $k$ -order spatial derivative  $\partial_x^{(k)} q(x, t)$  obeys the same differential equation, namely

$$\partial_t(\partial_x^{(k)} q(x, t)) + \lambda \partial_x(\partial_x^{(k)} q(x, t)) = 0.$$

For non-linear equations the above is not true. For example, for the Burgers equation, written in non-conservative form,

$$\partial_t q + q \partial_x q = 0,$$

the first order spatial derivative  $\partial_x^{(1)} q$  obeys the modified (inhomogeneous) Burgers-type equation

$$\partial_t(\partial_x^{(1)} q) + q\partial_x(\partial_x^{(1)} q) = -(\partial_x^{(1)} q)^2.$$

We return to the evolution equations (19.29). Clearly, in order to devise a practical method of solution we need to simplify these systems, hopefully, without destroying the high accuracy of the numerical schemes of the type (19.29) we are aiming to construct. Two simplifications are made. First we neglect the source terms  $\mathbf{H}^{(k)}$  and then we linearize the resulting homogeneous equations. Neglecting the effect of the source terms is justified, as we only need  $\partial_x^{(k)} \mathbf{Q}(x, 0_+)$  at the *first-instant* interaction of left and right states, leading to homogeneous non-linear systems for spatial derivatives. Then we perform a linearization of the homogeneous systems about the leading term  $\mathbf{Q}(0, 0_+)$  of the power series expansion (19.22), that is, the coefficient matrix  $\mathbf{A}(\mathbf{Q})$  in (19.29) is taken as the constant matrix

$$\mathbf{A}_{LR}^{(0)} = \mathbf{A}(\mathbf{Q}(0, 0_+)). \quad (19.31)$$

The leading term will generally be the solution of the non-linear Riemann problem (19.23), found by a non-linear Riemann solver, exact or approximate, and the proposed linearization based on the solution of this non-linear problem appears to be adequate to our purpose.

### Classical Riemann problems for spatial derivatives

In order to find the spatial derivatives  $\partial_x^{(k)} \mathbf{Q}(0, 0_+)$  in (19.26) we solve the following *homogeneous, linearized, classical* Riemann problems, for each  $k$

$$\left. \begin{aligned} \text{PDEs: } & \partial_t(\partial_x^{(k)} \mathbf{Q}(x, t)) + \mathbf{A}_{LR}^{(0)} \partial_x(\partial_x^{(k)} \mathbf{Q}(x, t)) = \mathbf{0}, \\ \text{ICs: } & \partial_x^{(k)} \mathbf{Q}(x, 0) = \begin{cases} \mathbf{Q}_L^{(k)}(0) & \text{if } x < 0, \\ \mathbf{Q}_R^{(k)}(0) & \text{if } x > 0. \end{cases} \end{aligned} \right\} \quad (19.32)$$

As already stated the (constant) Jacobian matrix  $\mathbf{A}_{LR}^{(0)}$  is the same coefficient matrix for all  $\partial_x^{(k)} \mathbf{Q}(x, t)$  and is evaluated only once, using the leading term (19.24) of the expansion (19.22).

We denote the similarity solution of (19.32) by  $\mathbf{D}^{(k)}(x/t)$ . See section 2.3.3 of chapter 2 for methods to solve Riemann problems for linear systems with constant coefficients, analytically. The relevant value at the interface is obtained by evaluating  $\mathbf{D}^{(k)}(x/t)$  at  $x/t = 0$ , namely

$$\partial_x^{(k)} \mathbf{Q}(0, 0_+) = \mathbf{D}^{(k)}(0). \quad (19.33)$$

We call this value the *Godunov state*, in analogy with the interface state (19.24), the leading term.

With the solution of  $\mathbf{Q}_{LR}(\tau)$  of the generalized Riemann problem  $GRP_K$  available at the interface, a *numerical flux* for scheme (19.2) to solve (19.1) is obtained by evaluating the time integral in (19.4) to the appropriate accuracy. In the presence of source terms in (19.1) the *numerical source* is computed as detailed below.

### 19.4.3 Source Terms

In order to obtain an approximation to the volume integral of the source term in (19.4) and thus obtain the numerical source of scheme (19.2), we propose to use the Cauchy–Kowalewski method described above. Given that within each cell  $I_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$  one assumes a smooth representation of the data using a reconstruction procedure, all spatial derivatives at the initial time are defined and readily computed. One possible procedure to define the function  $\mathbf{Q}(x, t)$  in (19.4) is to select a set of integration points  $\{x_j\}$  within the cell and for each  $j$  construct a time power series expansion analogous to (19.15), then substitute all time derivatives by functions of space derivatives, evaluate the space derivatives on the initial data and thus obtain a time-dependent function for each point  $j$ . Numerical evaluation of the volume integral in (19.4) to the appropriate order then follows. Alternatively, one can construct a space–time power series expansion, as in (19.15), then replace all time and mixed derivatives by space derivatives alone, evaluate these on the initial data, and in this manner determine the full series. Numerical evaluation of the volume integral in (19.4) to the appropriate order then follows, giving the numerical source of the scheme (19.2).

In the next section we summarize the complete solution procedure to obtain  $\mathbf{Q}_{LR}(\tau)$  for evaluating the numerical flux.

### 19.4.4 Summary: Numerical Flux and Numerical Source

The complete procedure for solving the generalized Riemann problem (19.5) and constructing a high-order numerical scheme is described in the following steps:

Step (I): **The leading term.** To compute the leading term in (19.22) one solves exactly or approximately the classical, homogeneous Riemann problem

$$\left. \begin{aligned} \text{PDEs: } & \partial_t \mathbf{Q} + \partial_x \mathbf{F}(\mathbf{Q}) = \mathbf{0}, \\ \text{ICs: } & \mathbf{Q}(x, 0) = \begin{cases} \mathbf{Q}_L(0) & \text{if } x < 0, \\ \mathbf{Q}_R(0) & \text{if } x > 0, \end{cases} \end{aligned} \right\} \quad (19.34)$$

to obtain the similarity solution  $\mathbf{D}^{(0)}(x/t)$ . The leading term is then given by the Godunov state

$$\mathbf{Q}(0, 0_+) = \mathbf{D}^{(0)}(0) . \quad (19.35)$$

Step (II): **Higher order terms.** There are three sub-steps here.

1. **Time derivatives in terms of spatial derivatives:** Use the Cauchy–Kowalewski procedure to express time derivatives in (19.22) in terms of functions of spatial derivatives

$$\partial_t^{(k)} \mathbf{Q}(x, t) = \mathbf{P}^{(k)}(\partial_x^{(0)} \mathbf{Q}, \partial_x^{(1)} \mathbf{Q}, \dots, \partial_x^{(k)} \mathbf{Q}) . \quad (19.36)$$

2. **Evolution equations for derivatives:** Construct evolution equations for spatial derivatives

$$\partial_t(\partial_x^{(k)} \mathbf{Q}(x, t)) + \mathbf{A}(\mathbf{Q}) \partial_x(\partial_x^{(k)} \mathbf{Q}(x, t)) = \mathbf{H}^{(k)} . \quad (19.37)$$

3. **Riemann problems for spatial derivatives:** Simplify (19.37) by neglecting source terms  $\mathbf{H}^{(k)}$  and linearizing the evolution equations. Then pose classical, homogeneous linearized Riemann problems for spatial derivatives

$$\left. \begin{aligned} \text{PDEs: } & \partial_t(\partial_x^{(k)} \mathbf{Q}(x, t)) + \mathbf{A}_{LR}^{(0)} \partial_x(\partial_x^{(k)} \mathbf{Q}(x, t)) = \mathbf{0} , \\ \text{ICs: } & \partial_x^{(k)} \mathbf{Q}(x, 0) = \begin{cases} \mathbf{Q}_L^{(k)}(0) & \text{if } x < 0 , \\ \mathbf{Q}_R^{(k)}(0) & \text{if } x > 0 . \end{cases} \end{aligned} \right\} \quad (19.38)$$

Solve these Riemann problems to obtain similarity solutions  $\mathbf{D}^{(k)}(x/t)$  and set

$$\partial_x^{(k)} \mathbf{Q}(0, 0_+) = \mathbf{D}^{(k)}(0) . \quad (19.39)$$

Step (III): **The solution.** Inserting  $\partial_x^{(k)} \mathbf{Q}(0, 0_+)$  into (19.36) one determines the coefficients in (19.22). Then the power series expansion reads

$$\mathbf{Q}_{LR}(\tau) = \mathbf{C}_0 + \mathbf{C}_1 \tau + \mathbf{C}_2 \tau^2 + \dots + \mathbf{C}_K \tau^K , \quad (19.40)$$

with  $\mathbf{C}_0$  as in (19.24) and

$$\mathbf{C}_k \equiv \frac{\partial_t^{(k)} \mathbf{Q}(0, 0_+)}{k!} = \frac{\mathbf{P}^{(k)}(\partial_x^{(0)} \mathbf{Q}, \partial_x^{(1)} \mathbf{Q}, \dots, \partial_x^{(k)} \mathbf{Q})}{k!} , \quad (19.41)$$

for  $k = 1, \dots, K$ . The solution technique for the generalized Riemann problem  $GRP_K$  reduces the problem to that of solving  $K + 1$  classical homogeneous Riemann problems, 1 (generally non-linear) Riemann problem to compute the leading term and  $K$  linearized Riemann problems to determine the higher order terms.

Step (IV): **Numerical flux.** The *numerical flux* for scheme (19.2) to solve the general problem (19.1) numerically is obtained by evaluating the time integral in (19.4), with  $\mathbf{Q}_{LR}(\tau)$  as the integrand, to the appropriate accuracy.

Step (V): **Numerical source.** To complete the numerical scheme (19.2) we need to compute the *numerical source*. This is achieved by evaluating the space–time volume integral in (19.4) to the appropriate accuracy, as described previously in section 19.4.3.

#### 19.4.5 Some remarks

The leading term requires the availability of a classical *Riemann solver*, exact or approximate. Chapter 4 gives the exact Riemann solver for the Euler equations and chapters 7 to 12 give a variety of approximate Riemann solvers. Recent works [496], [521], [523], [524], [495] provide alternative ways of solving the classical Riemann problem. These new *Riemann solvers* are simple and quite general. This means that also the generalized Riemann problem can be solved for more general hyperbolic systems [521].

The  $K$  linearized Riemann problems (19.38) for most well-known systems associated with the higher order terms can be solved analytically and no choice of a *Riemann solver* is necessary. Moreover, all of these linearized problems have the same eigenstructure, as the coefficient matrix is the same for all Riemann problems for derivatives.

For linear systems with constant coefficients the method of solution for  $GRP_K$  presented here is justified. For the non-linear case there appears to be no theoretical justification available so far, although the empirical evidence available is supportive, even for non-linear problems with source terms. See [543] and [521] for further details.

In principle, the technique can be applied to calculate the early–time solution of advection–reaction equations with piece–wise smooth initial conditions. One can set up a generalized Riemann problem at any desired position, taking care that at each point  $x = x_j$  of discontinuity in the initial condition one sets a corresponding generalized Riemann problem centred at  $x_j$ . The solution at each point  $x_j$ , for a small time  $\tau$ , can be used to check the results of numerical schemes.

### 19.5 Examples

Here we apply the technique studied in section 19.4 to solve the  $GRP_K$  for scalar advection–reaction equations

$$\partial_t q + \partial_x f(q) = s(q) , \quad (19.42)$$

where  $f(q)$  is the flux function and  $s(q)$  is a source term. We first deal with the simpler linear case without source term.

### 19.5.1 The Linear Advection Equation

We solve the generalized Riemann problem for the scalar linear advection equation with constant coefficient  $\lambda$  and no source term:

$$\left. \begin{aligned} \text{PDE: } & \partial_t q + \lambda \partial_x q = 0, \\ \text{IC: } & q(x, 0) = \begin{cases} q_L(x) & \text{if } x < 0, \\ q_R(x) & \text{if } x > 0, \end{cases} \end{aligned} \right\} \quad (19.43)$$

in which the initial conditions  $q_L(x)$  and  $q_R(x)$  are two smooth functions defined respectively for  $x \in (-\infty, 0)$  and  $x \in (0, \infty)$ . We may assume, for example, that  $q_L(x)$  and  $q_R(x)$  are polynomials of degree at most  $K$ .

According to the method studied in section 19.4, the solution of  $GRP_K$  (19.43) is given by the power series expansion

$$q_{LR}(\tau) = q(0, 0_+) + \sum_{k=1}^K \left[ \partial_t^{(k)} q(0, 0_+) \right] \frac{\tau^k}{k!}. \quad (19.44)$$

We determine each term of the expansion following the steps summarized in subsection 19.4.3.

**Step (I): the leading term.** The leading term  $q(0, 0_+)$  is computed from the similarity solution  $d^{(0)}(x/t)$  of the classical Riemann problem

$$\left. \begin{aligned} \text{PDE: } & \partial_t q + \lambda \partial_x q = 0, \\ \text{IC: } & q(x, 0) = \begin{cases} q_L(0) & \text{if } x < 0, \\ q_R(0) & \text{if } x > 0. \end{cases} \end{aligned} \right\} \quad (19.45)$$

Trivially, the solution in the half-plane  $t > 0$ ,  $-\infty < x < \infty$  is

$$d^{(0)}(x/t) = \begin{cases} q_L(0) & \text{if } \lambda > x/t, \\ q_R(0) & \text{if } \lambda < x/t, \end{cases} \quad (19.46)$$

and the sought leading term is the Godunov state

$$q(0, 0_+) = d^{(0)}(0) = \begin{cases} q_L(0) & \text{if } \lambda > 0, \\ q_R(0) & \text{if } \lambda < 0. \end{cases} \quad (19.47)$$

**Step (II): higher-order terms.** First we need the time derivatives expressed as functions of space derivatives. Using the Cauchy–Kowalewski procedure we obtain

$$\left. \begin{aligned} \partial_t q &= -\lambda \partial_x q, \\ \partial_t^{(2)} q &= \lambda^2 \partial_x^{(2)} q, \\ &\dots \dots \dots \\ \partial_t^{(k)} q &= (-1)^k \lambda^k \partial_x^{(k)} q, \end{aligned} \right\} \quad (19.48)$$

so that the functions  $p^{(k)}$  in (19.26) or (19.36) have a particularly simple form, namely  $p^{(k)} = (-1)^k \lambda^k \partial_x^{(k)} q$ ,  $\forall k > 0$ . Now we need to determine the space derivatives  $\partial_x^{(k)} q(0, 0_+)$ .

Following step (II)-2 of subsection 19.4.3 we formulate evolution equations for  $\partial_x^{(k)} q$ , which, as already seen obeys

$$\partial_t(\partial_x^{(k)} q) + \lambda \partial_x(\partial_x^{(k)} q) = 0, \quad \forall k > 0. \quad (19.49)$$

Note that the evolution equations for space derivatives are the *same* as the original (homogeneous) PDE for the unknown  $q(x, t)$ .

Following step (II)-3 of subsection 19.4.3 we solve the classical Riemann problems for space derivatives

$$\left. \begin{aligned} \text{PDE: } \partial_t(\partial_x^{(k)} q) + \lambda \partial_x(\partial_x^{(k)} q) &= 0, \\ \text{IC: } \partial_x^{(k)} q(x, 0) &= \begin{cases} q_L^{(k)}(0) & \text{if } x < 0, \\ q_R^{(k)}(0) & \text{if } x > 0, \end{cases} \end{aligned} \right\} \quad (19.50)$$

whose similarity solutions are denoted by  $d^{(k)}(x/t)$  and are given by

$$d^{(k)}(x/t) = \begin{cases} q_L^{(k)}(0) & \text{if } \lambda > x/t, \\ q_R^{(k)}(0) & \text{if } \lambda < x/t, \end{cases} \quad (19.51)$$

for  $k = 1, \dots, K$ . For each  $k$  the sought Godunov state is

$$\partial_x^{(k)} q(0, 0_+) = d^{(k)}(0) = \begin{cases} q_L^{(k)}(0) & \text{if } \lambda > 0, \\ q_R^{(k)}(0) & \text{if } \lambda < 0. \end{cases} \quad (19.52)$$

**Step (III): the complete solution.** The coefficients of the series expansion (19.40) are

$$c_k \equiv \frac{\partial_t^{(k)} q(0, 0_+)}{k!} = \frac{(-1)^k \lambda^k d^{(k)}(0)}{k!}, \quad (19.53)$$

for  $k = 1, \dots, K$ , with  $c_0$  the leading term given by (19.47). The complete solution is

$$q_{LR}(\tau) = c_0 + c_1 \tau + c_2 \tau^2 + \dots + c_K \tau^K, \quad (19.54)$$

which in full reads

$$q_{LR}(\tau) = \begin{cases} q_L(0) + \sum_{k=1}^K \left[ (-1)^k \lambda^k q_L^{(k)}(0) \right] \frac{\tau^k}{k!} & \text{if } \lambda > 0, \\ q_R(0) + \sum_{k=1}^K \left[ (-1)^k \lambda^k q_R^{(k)}(0) \right] \frac{\tau^k}{k!} & \text{if } \lambda < 0. \end{cases} \quad (19.55)$$

The numerical flux is now obtained by evaluating the time integral in (19.4), with  $q_{LR}(\tau)$  as the integrand. Note that in this case the integral can be evaluated exactly.

**Remark:** It is easy to verify that the above solution is correct. The exact solution of the initial value problem for the linear advection equation with general initial condition  $q_0(x)$  is  $q(x, t) = q_0(x - \lambda t)$ . For  $x = 0$  we have

$$q(0, t) = q_0(-\lambda t) = \begin{cases} q_L(-\lambda t) & \text{if } \lambda > 0, \\ q_R(-\lambda t) & \text{if } \lambda < 0. \end{cases} \quad (19.56)$$

By Taylor expanding  $q_L(-\lambda t)$  about  $x = 0_-$  for  $\lambda > 0$  or  $q_R(-\lambda t)$  about  $x = 0_+$  for  $\lambda < 0$  we obtain the desired result.

### 19.5.2 Linear Advection with a Source Term

Here we solve the  $GRP_K$  for the model advection–reaction equation

$$\partial_t q + \lambda \partial_x q = \beta q, \\ q(x, 0) \equiv q_0(x) = \begin{cases} q_L(x) & \text{if } x < 0, \\ q_R(x) & \text{if } x > 0. \end{cases} \quad (19.57)$$

The coefficient  $\lambda$  is a constant characteristic speed;  $\beta$  is also a constant, with dimensions reciprocal of time, and is normally taken as a negative real number in order to have a stable differential equation. The exact solution for a general initial condition  $q_0(x)$  is

$$q(x, t) = q_0(x - \lambda t) e^{\beta t}. \quad (19.58)$$

For the sake of clarity, at the cost of some repetitiveness, we solve the problem in complete detail. Applying the method studied, the solution of  $GRP_K$  (19.57) has the form

$$q_{LR}(\tau) = q(0, 0_+) + \sum_{k=1}^K \left[ \partial_t^{(k)} q(0, 0_+) \right] \frac{\tau^k}{k!}. \quad (19.59)$$

Note that the form of the solution is the same as that for the homogeneous case. The effect of the source term enters in the computation of the coefficients

for the terms of order greater than one. To determine each term in (19.59) we follow the steps of section 19.4.3.

According to Step (I) the leading term of the expansion, which is the same as for the homogeneous case, is given by (19.47). The higher-order terms are computed according to Step (II). We first express time derivatives as functions of spatial derivatives of  $q$  and the source term, using the Cauchy–Kowalewski procedure. For example

$$\left. \begin{aligned} \partial_t q &= -\lambda \partial_x q + \beta q , \\ \partial_t^{(2)} q &= \lambda^2 \partial_x^{(2)} q - 2\lambda\beta \partial_x^{(1)} q + \beta^2 q . \end{aligned} \right\} \quad (19.60)$$

For the general case it can be shown that

$$\partial_t^{(k)} q = \sum_{n=0}^k \left[ \frac{k!}{(k-n)!n!} \beta^n (-\lambda)^{k-n} \right] \partial_x^{(k-n)} q . \quad (19.61)$$

As seen in equations (19.61)–(19.62), in the present version of the technique, the source term in the PDE enters the solution via the time derivatives  $\partial_t^{(k)} q$ , for  $k = 1, \dots, K$ .

As for the homogeneous case, it is a question of computing space derivatives at the interface  $x = 0$  to determine the time derivatives (19.61). To this end we formulate evolution equations for space derivatives. It is immediately seen that the derivative  $\partial_x^{(k)} q$  obeys the same, original, advection–reaction evolution equation in (19.57), namely

$$\partial_t (\partial_x^{(k)} q) + \lambda \partial_x (\partial_x^{(k)} q) = \beta \partial_x^{(k)} q , \quad k = 1, \dots, K . \quad (19.62)$$

As in (19.37), we neglect the effect of the derivative source terms in the evolution of the derivatives. We thus solve homogeneous classical Riemann problems of the form (19.50) to find similarity solutions  $d^{(k)}(0)$  given by (19.51). Substitution of all space derivatives into (19.61) gives the time derivatives and thus the coefficients and the solution is given by (19.59).

We note that the solution  $q_{LR}(\tau)$  of the inhomogeneous  $GRP_K$  at the interface is again reduced to the solution of a sequence of  $K + 1$  classical, homogeneous Riemann problems. The effect of the source term comes into the coefficients of the higher-order terms of the expansion. At no point did we need to solve directly inhomogeneous problems to account for the presence of the source term.

The numerical flux is now obtained by evaluating the time integral in (19.4), with  $q_{LR}(\tau)$  as the integrand. The numerical source is evaluated as described in section 19.4.3.

### 19.5.3 Non–Linear Equation with a Source Term

We solve the  $GRP_K$

$$\left. \begin{aligned} \partial_t q + \partial_x f(q) &= s(q) , \\ q(x, 0) \equiv q_0(x) &= \begin{cases} q_L(x) & \text{if } x < 0 , \\ q_R(x) & \text{if } x > 0 , \end{cases} \end{aligned} \right\} \quad (19.63)$$

in which  $f(q)$  is flux function and  $s(q)$  is a source term. The solution has again the form

$$q_{LR}(\tau) = q(0, 0_+) + \sum_{k=1}^K \left[ \partial_t^{(k)} q(0, 0_+) \right] \frac{\tau^k}{k!} . \quad (19.64)$$

We determine each term of the expansion following the steps of section 19.4.3.

### Step (I): the leading term

The leading term results from the solution of the homogeneous classical Riemann problem

$$\left. \begin{aligned} \text{PDE: } \partial_t q + \partial_x f(q) &= 0 , \\ \text{IC: } q(x, 0) &= \begin{cases} q_L(0) & \text{if } x < 0 , \\ q_R(0) & \text{if } x > 0 \end{cases} \end{aligned} \right\} \quad (19.65)$$

denoted by  $d^{(0)}(x/t)$ , with the leading term given by

$$q(0, 0_+) = d^{(0)}(0) . \quad (19.66)$$

### Step (II): higher-order terms

We need to determine the time derivatives in (19.64), which is done in three sub-steps. First, using the Cauchy–Kowalewski procedure, we express time derivatives as functions of space derivatives. For example, the first-order time derivative, from the PDE, is

$$\partial_t^{(1)} q(x, t) = -f^{(1)}(q) \partial_x^{(1)} q + s(q) . \quad (19.67)$$

Then we compute the mixed derivative

$$\partial_x \partial_t q = \partial_x^{(1)} \left[ -f^{(1)}(q) \partial_x^{(1)} q + s(q) \right] , \quad (19.68)$$

which after some manipulations becomes

$$\partial_t \partial_x q(x, t) = -f^{(1)}(q) \partial_x^{(2)} q - f^{(2)}(\partial_x^{(1)} q)^2 + s^{(1)}(q) \partial_x^{(1)} q . \quad (19.69)$$

The second-order time derivative is obtained from differentiating (19.67) with respect to  $t$  and making appropriate substitutions, yielding

$$\left. \begin{aligned} \partial_t^{(2)} q &= [f^{(1)}]^2 \partial_x^{(2)} q + 2f^{(1)}(q)f^{(2)}(q)(\partial_x^{(1)} q)^2 \\ &\quad - [s(q)f^{(2)}(q) + 2f^{(1)}(q)s^{(1)}(q)] \partial_x^{(1)} q + s(q)s^{(1)}(q) . \end{aligned} \right\} \quad (19.70)$$

The higher-order time derivatives are constructed in a similar manner, ideally using a symbolic manipulator. In general, they are functions of  $q(x, t)$  and its spatial derivatives and have the form (19.36), namely

$$\partial_t^{(k)} q(x, t) = p^{(k)}(\partial_x^{(0)} q, \partial_x^{(1)} q, \dots, \partial_x^{(k)} q) . \quad (19.71)$$

We now construct evolution equations for the spatial derivatives, the arguments of  $p^{(k)}$ . These have the form

$$\partial_t(\partial_x^{(k)} q) + \lambda(q)\partial_x(\partial_x^{(k)} q) = h^{(k)} , \quad (19.72)$$

where

$$\lambda(q) = f^{(1)}(q) \quad (19.73)$$

is the characteristic speed. Neglecting the source term and linearizing we have the homogeneous linear problems

$$\left. \begin{aligned} \partial_t(\partial_x^{(k)} q) + \lambda_{LR}^{(0)}\partial_x(\partial_x^{(k)} q) &= 0 , \\ \partial_x^{(k)} q(x, 0) &= \begin{cases} q_L^{(k)}(0) & \text{if } x < 0 , \\ q_R^{(k)}(0) & \text{if } x > 0 , \end{cases} \end{aligned} \right\} \quad (19.74)$$

where  $\lambda_{LR}^{(0)} = f^{(1)}(q(0, 0_+))$  is a constant characteristic speed. The Godunov state is

$$d^{(k)}(0) = \begin{cases} q_L^{(k)}(0) & \text{if } \lambda_{LR}^{(0)} > 0 , \\ q_R^{(k)}(0) & \text{if } \lambda_{LR}^{(0)} < 0 . \end{cases} \quad (19.75)$$

Substitution of these into the right-hand side of (19.71) determines all the coefficients in (19.64) and hence the solution is complete.

The numerical flux is now obtained by evaluating the time integral in (19.4), with  $q_{LR}(\tau)$  as the integrand. The numerical source is evaluated as described in section 19.4.3.

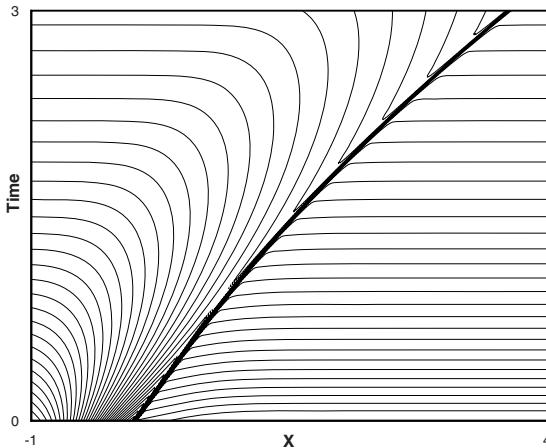
#### 19.5.4 The Burgers Equation with a Source Term

As an example we solve the generalized Riemann problem for the inviscid Burgers equation with a non-linear source term

$$\partial_t q + \partial_x \left( \frac{1}{2} q^2 \right) = e^{-q},$$

$$q(x, 0) = \begin{cases} q_L(x) = 2e^{-2(x-\frac{1}{5})^2} & \text{if } x < 0, \\ q_R(x) = \frac{1}{8}e^{-2(x+\frac{1}{5})^2} & \text{if } x > 0. \end{cases} \quad (19.76)$$

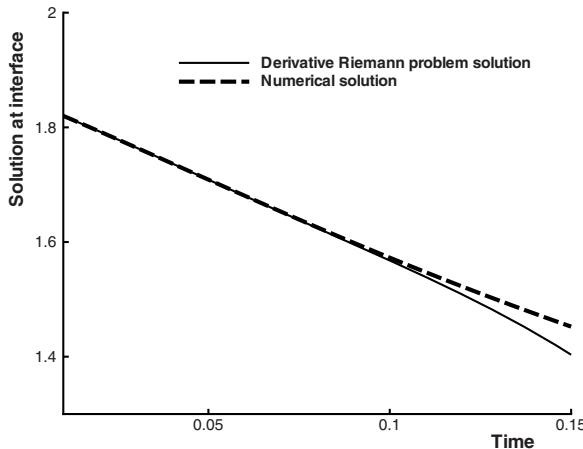
We wish to find the solution  $q_{LR}(\tau)$  right at the interface, as a function of time.



**Fig. 19.4.** Contour plot of the numerical solution of the generalized Riemann problem for the inviscid Burgers equation with a non-linear source term. The main feature of the solution is an accelerating shock wave that propagates to the right (Courtesy of Dr. V. A. Titarev).

Fig. 19.4 shows a contour plot of the *global* solution of (19.76) in the  $x-t$  plane. This solution was obtained numerically using a 5-th order WENO method [493] on a fine mesh of 1000 cells in the computational domain  $[-5, 5]$ . We regard this as the *reference* solution and we use it to compare with results from the semi-analytical solution procedure for the  $GRP_K$  studied in section 19.4. The dominant feature of the solution is an accelerating shock wave that emerges from the initial discontinuity in the initial condition at  $x = 0$ .

Fig. 19.5 depicts a comparison of the solution  $q_{LR}(\tau)$  given by (19.64) at the interface  $x = 0$  for  $0 < \tau < 0.15$ , with the reference numerical solution. We remark here that the reference (numerical) solution is not computed precisely at the interface  $x = 0$ . The nearest numerical values are those in the finite volumes immediately to the left and to the right of the interface. The solution  $q_{LR}(\tau)$  is valid precisely at the interface. Thus in order to carry out a comparison we use WENO high-order interpolation to obtain an interpo-



**Fig. 19.5.** Comparison between the reference solution (numerical solution interpolated to the interface position) and the  $GRP_K$  solution to sixth order of accuracy (Courtesy of Dr. V. A. Titarev).

lated numerical value at the interface. This is the value we compare with the solution  $q_{LR}(\tau)$  of  $GRP_K$ .

Fig. 19.5 shows that for relatively short times there is excellent agreement between the reference numerical solution and the solution  $q_{LR}(\tau)$  of  $GRP_K$ . It is also seen that, as expected, for larger times the accuracy of the solution  $q_{LR}(\tau)$  deteriorates visibly. We also note that for times very close to zero (not shown in Fig. 19.5) the reference solution is very inaccurate due to the smearing of the shock wave.

We can also compute an *error* by taking the difference between the reference numerical solution of Fig. 19.4 and the solution  $q_{LR}(\tau)$  of  $GRP_K$ . Table 19.1 shows errors at the four times 0.01, 0.025, 0.05 and 0.1 for solutions  $q_{LR}(\tau)$  of orders 1 to 6. It is seen that for small times, as expected, the solution is accurate, and becomes more accurate as the number of terms in the expansion increases. For *large* times the error is large. The error decreases as the number of terms in the expansion is increased, but this reduction in the error is less obvious for the larger times.

This example clearly illustrates the fact that the power series solution remains valid only for *small* times, as one would expect.

Order	$t = 0.010$	$t = 0.025$	$t = 0.050$	$t = 0.100$
1	0.2625656E-1	0.6724719E-1	0.1374329E0	0.2735556E0
2	0.5662654E-3	0.3021457E-2	0.8981471E-2	0.1665263E-1
3	0.5807203E-4	0.880652E-3	0.6626964E-2	0.4578111E-1
4	0.9508504E-6	0.4158045E-4	0.7508961E-3	0.1324177E-1
5	0.8990528E-7	0.7949777E-5	0.2128054E-3	0.463232E-2
6	0.5743972E-8	0.139097E-5	0.8609855E-4	0.4932606E-2

Table 19.1. Solution  $q_{LR}(\tau)$  of  $GRP_K$  for the Burgers equation with a non-linear source term. Errors of solution  $q_{LR}(\tau)$  of orders  $K = 1, \dots, 6$ , at four times.

In the next section we briefly discuss alternative ways of solving the generalized Riemann problem.

## 19.6 Other Solvers

In this chapter we have presented a semi-analytical method to solve the generalized Riemann problem for hyperbolic balance laws, based on a high-order Taylor expansion in time of the solution right at the interface. The technique is an arbitrary-order generalization of the second-order method proposed by Ben-Artzi and Falcovitz [37]. The material of this chapter closely follows the works [543] and [523]. To facilitate the discussion on various solvers we call this the Toro–Titarev  $GRP_K$  solver. Once the solution at the interface, as a function of time, is known then the numerical flux in (19.2) can be computed by appropriate integration, leading to a fully discrete, one-step method of arbitrary order of accuracy in both space and time.

Also the numerical method of Harten and collaborators [242] may be re-interpreted as a solver for the generalized Riemann problem  $GRP_K$ . See [91] for details on the re-interpretation and comparisons with other solvers. Harten and collaborators [242] first evolve in time the boundary extrapolated values of the data  $\mathbf{Q}_L(x)$  and  $\mathbf{Q}_R(x)$ , on the left and right of the interface; this is accomplished by means of a Taylor series expansion followed by use of the Cauchy–Kowalewski procedure. Then, at any time  $t_k$  one has two constant values of the evolved data, one to the left and one to the right of the interface. In order to find the solution *right at the interface* Harten and colleagues solve a classical Riemann problem at the interface and use the corresponding similarity solution to define the sought value of the solution of the generalized Riemann problem. We call this solver the Harten–Osher–Engquist–Chakravarthy solver.

A variant of the two solvers above was presented in [91], whereby the left and right boundary extrapolated data are evolved in time producing time-evolved temporal derivatives. These then form the initial conditions for a

classical Riemann problem for *time derivatives* directly, determining in this manner all the coefficients of the power series expansion right at the interface. We call this solver the Castro–Toro  $GRP_K$  solver. In [91], a comparison of the three solvers is made. It is established that for homogeneous linear hyperbolic systems with constant coefficients, all the three solves are algebraically equivalent. Otherwise, they are all different but are equivalent in terms of order of accuracy. In the evaluation of the  $GRP_K$  solvers of [91], the actual time-dependent solution of  $GRP_K$  was computed and compared to reference solutions for a range of test problems, for the Euler equations. In [530] such comparison was carried out for the Baer–Nunziato equations of two-phase compressible flow.

Numerical implementations pose two major tasks: (i) a non-linear reconstruction procedure (see, for example, [242], [270] and [25]), that provides the high-order data for the generalized Riemann problem and (ii) the practical implementation of the Cauchy–Kowalewski procedure to convert time derivatives to space derivatives. Reconstruction procedures in multiple space dimensions on unstructured meshes, although very complex, are by now well established. See for example [175] and [176]. However, the Cauchy–Kowalewski procedure remains the main challenge for implementing these three  $GRP_K$  solvers, specially for very high order of accuracy, in multiple space dimensions, on unstructured meshes and with source terms. We note, however, that computing time is not the main drawback of the Cauchy–Kowalewski procedure, but sheer complexity and thus the use of symbolic manipulators is a must in all three solvers mentioned above. For some systems, special algorithms have been developed to implement the Cauchy–Kowalewski procedure. See for example [176], for the three-dimensional Euler equations. An additional shortcoming of the solvers discussed so far is the ability to deal with stiff source terms.

In the recent paper [173], an alternative  $GRP_K$  solver has been presented, which we call the Dumbser–Enaux–Toro  $GRP_K$  solver. In very simple terms, this solver may be seen as a numerical variant of the Harten–Osher–Engquist–Chakravarthy solver. The initial data on the left and right of the interface is evolved in time numerically, avoiding in this manner the use of the Cauchy–Kowalewski procedure. Moreover, if the numerical evolution is carried out by an implicit method (locally implicit), then the generalized Riemann problem for hyperbolic systems with stiff source terms can be solved correctly. Dumbser and collaborators [173] use a locally implicit space–time discontinuous Galerkin method for the data evolution stage. Then to determine the solution of the generalized Riemann problem at the required position, the interface, one solves a classical Riemann problem at the required time  $t_k$ . The numerical flux to high-order is then determined in the usual way by appropriate quadrature. As a bonus of the Dumbser–Enaux–Toro method, one can straightforwardly evaluate the source term (stiff or non stiff) correctly; this involves a space–time volume integral in which one can use directly the evolved data for the predictor stage of the solver. Having determined the numerical

flux and the numerical source one has then a globally *explicit*, fully discrete, one-step scheme (19.2) of arbitrary order of accuracy in space and time. See [173] and [171] for recent applications of this  $GRP_K$  solver.

## 19.7 Concluding Remarks

We have studied in complete detail a semi-analytical method for solving the generalized Riemann problem  $GRP_K$  for systems of non-linear hyperbolic equations with source terms. The initial condition of  $GRP_K$  consists of two arbitrary but smooth vector fields joined by a jump discontinuity at the origin and the governing equations contain source terms, in general. The solution procedure studied reduces the full inhomogeneous non-linear problem to the solution of one non-linear homogeneous classical Riemann problem and a sequence of homogeneous classical linear Riemann problems for spatial derivatives. The solution technique has been illustrated through some simple scalar examples; further details of the method as applied to non-linear systems are given in [543] and [523]. Alternative methods have been briefly outlined. The main use of the generalized Riemann problem solution is in the construction of numerical fluxes and numerical sources that produce fully discrete, one-step numerical schemes of arbitrary order of accuracy in both space and time, for solving hyperbolic equations with source terms. This can be done in the frameworks of finite volume and discontinuous Galerkin finite element methods. In the following chapter we give an introduction to such methods in the framework of finite volumes.

## The ADER Approach

This chapter is an introduction to the ADER family of fully-discrete, one-step methods of arbitrary order of accuracy in space and time, for solving hyperbolic equations with source terms. These schemes are a generalization of the Godunov method. The numerical flux is computed as a time-integral average of the flux function evaluated at the solution of the generalized Riemann problem studied in chapter 19, and the numerical source is computed as a high-order space-time integral of the source term in the appropriate control volume. The ADER approach operates in the framework of finite volumes and of discontinuous Galerkin finite elements. Here we deal with the finite volume framework for one-dimensional model problems.

### 20.1 Introduction

The second-order TVD methods studied in previous chapters, have become a mature numerical technology that has been adopted by users not only in academia but also in industry, in research laboratories and in commercial software companies. However there is sufficient evidence that such methods are inadequate for a number of engineering and scientific problems. In particular, long-time evolution wave propagation problems require more accurate methods. It is also reasonable to assume that in the future, scientific computing and related disciplines will continue to drive errors towards machine zero. This means that numerical methods of very high order of accuracy will have an increasingly important role to play. However the obvious question to be asked is this: Given an error deemed acceptable, how can one best attain that error in terms of memory and processing time? Do we use a low-order method on a fine mesh or a high-order method on a coarse mesh? Available evidence [437], [436] suggests that for errors below a threshold it is distinctly better to use a method of high order of accuracy.

However, the construction of the kind of high-order methods that are suitable for wave propagation problems and other related fields, is subject

to stringent requirements. In particular, the methods must be conservative, in view of the Lax–Wendroff theorem [302], and non–linear, in view of Godunov’s theorem [216]. In addition, we speak of high–order of accuracy in both space and time, for problems involving source terms and to be solved in complex multidimensional geometries. We expect these methods to attain the theoretical convergence rates for smooth solutions and, at the same time, be able to compute discontinuous solutions without or much reduced spurious oscillations. The work of Harten and collaborators, see for example [242], introduced a family of fully–discrete, one–step methods that contained some of the basic desirable ingredients. The work of Shu and collaborators [447], [444] introduced semi–discrete versions of these methods, which are today the standard numerical methodology to use.

In this chapter we give an introduction to the ADER approach for constructing high–order methods. The approach was first put forward by Toro and collaborators in [537] for linear problems on Cartesian meshes. Schemes were formulated for one, two and three space dimensions; reported numerical implementations for linear schemes (fixed stencils) for one and two space dimensions included computations of up to 10–th order of accuracy in both space and time. The ADER schemes are one–step and fully discrete; they require a high–order spatial reconstruction (once per time step) and the solution of the generalized Riemann problem at the cell interface. The ADER schemes are akin to the methods proposed by Harten, Osher, Engquist and Chakravarthy [242]. As a matter of fact, for linear problems with constant coefficients the ADER schemes are identical to the schemes of [242]. For non–linear problems, however, they differ, even though they are equivalent in terms of order of accuracy. The ADER approach is an extension of the second–order GRP method of Ben–Artzi and Falcoviz [37] and the building block of the method is the solution of the generalized Riemann problem right at the interface. The scheme of Harten and collaborators, on the other hand, may be seen as an extension of the second–order MUSCL–Hancock scheme [562]; a distinctive feature of this method being the time evolution of the boundary extrapolated values of the reconstructed data, followed by solutions of classical Riemann problems at the interface. See [91] for a detailed comparison of both approaches.

Further developments of the ADER methodology for linear systems are due to Munz and collaborators [437], [436], [435]. The extension of ADER to non–linear problems for schemes of accuracy greater than two was first reported by Titarev and Toro in [492]; they used the solution of the generalized Riemann problem reported by Toro and Titarev in [543]. See also the work of Takakura and Toro [475] for the non–linear scalar case. Non–linear systems in two and three space dimensions using Cartesian meshes were dealt with in [494] and [490]; see also [544] and [524]. The case of non–Cartesian meshes was first tackled by Käser for the scalar non–linear case; see [281], [282] and [283]; see also the work of Castro [89]. The work of Dumbser and collaborators has taken the ADER approach a long way, by extending it to non–linear systems with source terms on unstructured meshes, in two and three space dimensions,

in the frameworks of finite volume and of discontinuous Galerkin finite element methods; see for instance [178], [169], [170], [175], [176] and [177]. The comparative study reported in [91] suggests that it is possible to widen the ADER framework so as to include the method of Harten and collaborators [242]. A significant step forward in this direction is the recent work of Dumbser and collaborators who have proposed a rather general solver for the generalized Riemann problem; see [173] and [171]. This solver allows, amongst other things, the proper treatment of stiff source terms, reconciling high accuracy and stiffness.

In this chapter we give an introduction to the ADER approach in the framework of finite volume methods. The rest of this chapter is structured as follows: in section 20.2 we describe the ADER approach in the finite volume framework; in section 20.3 we construct a second-order ADER method for a scalar equation with a source term; in section 20.4 we construct ADER schemes of arbitrary order of accuracy for the model scalar equation with a linear source term. Sample numerical results are shown in section 20.5 and conclusions are drawn in section 20.6.

## 20.2 The Finite Volume Method

The ADER methods studied in this chapter operate in both the frameworks of finite volumes and discontinuous Galerkin finite element methods. This introductory chapter deals exclusively with finite volume methods.

### 20.2.1 The Framework

Here we introduce the main features of the ADER approach as applied to a scalar conservation law with a source term

$$\partial_t q + \partial_x f(q) = s(q) . \quad (20.1)$$

A finite volume scheme to solve (20.1) approximately results from integrating equation (20.1) on a control or finite volume  $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [t^n, t^{n+1}]$  to obtain the formula

$$q_i^{n+1} = q_i^n - \frac{\Delta t}{\Delta x} [f_{i+\frac{1}{2}} - f_{i-\frac{1}{2}}] + \Delta t s_i , \quad (20.2)$$

where  $q_i^n$  is the spatial–integral average

$$q_i^n = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} q(x, t^n) dx , \quad (20.3)$$

$f_{i+\frac{1}{2}}$  is the time–integral average

$$f_{i+\frac{1}{2}} = \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} f(q(x_{i+\frac{1}{2}}, t)) dt , \quad (20.4)$$

and  $s_i$  is the volume–integral average

$$s_i = \frac{1}{\Delta t} \frac{1}{\Delta x} \int_{t^n}^{t^{n+1}} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} s(q(x,t)) dx dt , \quad (20.5)$$

with

$$\Delta x = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}} , \quad \Delta t = t^{n+1} - t^n . \quad (20.6)$$

In (20.3) the integrand is assumed to be the known initial condition at time  $t^n$  within the control volume. In (20.4)  $q(x_{i+\frac{1}{2}}, t)$  is the solution of (20.1) at the volume interface position  $x = x_{i+\frac{1}{2}}$ . In (20.5) the function  $q(x, t)$  is the solution of (20.1) within the volume  $x_{i-\frac{1}{2}} \leq x \leq x_{i+\frac{1}{2}}$ ,  $t^n \leq t \leq t^{n+1}$ . Under these assumptions relation (20.2) is *exact*.

Numerical methods of the finite volume type can be constructed by interpreting (20.2) as a numerical formula to update approximations to cell averages  $q_i^n$  within the cells  $I_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ , with cell or volume centre

$$x_i = \frac{1}{2}(x_{i-\frac{1}{2}} + x_{i+\frac{1}{2}}) , \quad (20.7)$$

cell interfaces  $x_{i-\frac{1}{2}}$  and  $x_{i+\frac{1}{2}}$ , mesh length  $\Delta x$  and time step  $\Delta t$ .

A particular finite volume method is constructed by specifying a particular approximation to the integrals (20.3)–(20.5). The integral (20.3) is obvious and results from taking  $q(x, t^n)$  as the initial condition inside the volume, or cell,  $I_i$ . Strictly speaking, at the initial time one should convert the given initial conditions into cells averages via (20.3), but for subsequent times the updating formula (20.2) takes care of this task automatically. What is crucial in a finite volume method is the design of a strategy to produce approximations to the integrals (20.4) and (20.5) that result in approximate expressions for  $f_{i+\frac{1}{2}}$ , called the *numerical flux*, and for  $s_i$ , called the *numerical source*. The numerical flux and numerical source will still be denoted respectively as  $f_{i+\frac{1}{2}}$  and  $s_i$ , so that (20.2) can be interpreted as a numerical formula.

In what follows we define the ADER approach for prescribing the numerical flux and the numerical source in the finite volume formula (20.2).

### 20.2.2 The Numerical Flux

The numerical flux  $f_{i+\frac{1}{2}}$  results from approximating the time–integral average (20.4) of the flux function at the interface. Where convenient we introduce local coordinates in which  $x$  is replaced by  $x - x_{i+\frac{1}{2}}$  and  $t$  is replaced by  $t - t^n$ . In the ADER approach the numerical flux is obtained by evaluating the integral

$$f_{i+\frac{1}{2}} = \frac{1}{\Delta t} \int_0^{\Delta t} f(q_{i+\frac{1}{2}}(\tau)) d\tau , \quad (20.8)$$

with a suitably high order numerical integration scheme, where the function of time  $q_{i+\frac{1}{2}}(\tau)$  at the interface position is the solution of a generalized Riemann problem. The generalized Riemann problem  $GRP_K$  for a given system with source terms has initial conditions with a piece-wise high-order spatial representation and whose solution at the interface has a high-order temporal representation. The  $GRP_K$  for (20.1) is the Cauchy problem

$$\left. \begin{aligned} \text{PDE: } & \partial_t q + \partial_x f(q) = s(q) , \\ \text{IC: } & q(x, 0) = \begin{cases} p_i(x) & \text{if } x < 0 , \\ p_{i+1}(x) & \text{if } x > 0 . \end{cases} \end{aligned} \right\} \quad (20.9)$$

Here  $p_i(x)$  and  $p_{i+1}(x)$  are assumed to be polynomials of degree  $K$  obtained from a conservative and non-oscillatory reconstruction procedure. The solution of (20.9) at the fixed interface position  $x_{i+\frac{1}{2}}$ , or  $x = 0$  in local coordinates, denoted by  $q_{i+\frac{1}{2}}(\tau)$ , is a function of time and has the form

$$q_{i+\frac{1}{2}}(\tau) = q(0, 0_+) + \sum_{k=1}^K \left[ \partial_t^{(k)} q(0, 0_+) \right] \frac{\tau^k}{k!} , \quad (20.10)$$

where

$$q(0, 0_+) = \lim_{t \rightarrow 0_+} q(0, t)$$

and  $K$  is the order of the approximation. See chapter 19 for details on solution methods for the generalized Riemann problem  $GRP_K$ . With the solution  $q_{i+\frac{1}{2}}(\tau)$  available one can compute the numerical flux according to (20.8). In the absence of source terms the resulting finite volume scheme (20.2) is of order  $K + 1$  in both space and time.

### 20.2.3 The Numerical Source

In the presence of source terms the ADER approach determines the numerical source  $s_i$  by approximating the space-time integral (20.5) within the space-time volume  $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [0, \Delta t]$  to obtain the average

$$s_i = \frac{1}{\Delta t} \frac{1}{\Delta x} \int_0^{\Delta t} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} s(q_i(x, t)) dx dt , \quad (20.11)$$

where  $q_i(x, t)$  denotes an approximation to the solution of (20.1) inside the volume  $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [0, \Delta t]$ , with  $q_i(x, 0) = p_i(x)$  in cell  $i$ , at the initial time  $t = 0$ . To obtain the required values of  $q_i(x, t)$  for  $x_{i-\frac{1}{2}} \leq x \leq x_{i+\frac{1}{2}}$ ,  $0 < t \leq \Delta t$  we use the Cauchy-Kowalewski procedure, introduced in chapter 19.

A possible version of the method considers fixed integration points  $\hat{x}$  within the cell  $i$  at which one constructs a Taylor series expansion in time. By means of the Cauchy–Kowalewski procedure one converts time derivatives to space derivatives using the PDE (20.1). The space derivatives in cell  $i$  are all readily available from the smooth reconstructed function  $p_i(x)$  and thus the solution at each integration point  $\hat{x}$  is known as an explicit function of time and therefore the integration can be carried out to obtain the numerical source  $s_i$  to the desired order of accuracy. The resulting finite volume scheme (20.2) for the advection equation with a source term (20.1) is of order  $K + 1$  in both space and time.

Two pending issues are the reconstruction procedure to obtain the smooth functions  $p_i(x)$  that form the initial condition of the generalized Riemann problem (20.9) and the application of the Cauchy–Kowalewski procedure to find  $q_i(x, t)$  in (20.11). In the following subsection we give a very succinct overview of the reconstruction problem.

#### 20.2.4 Reconstruction

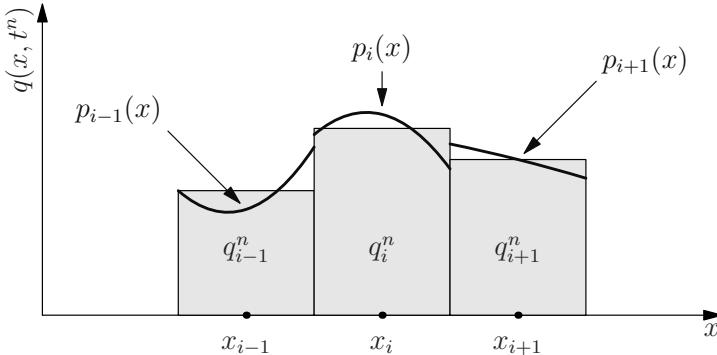
At any given time  $t^n$  one has a set of cell averages  $\{q_i^n\}$  that are approximations to (20.3). This set represents a piecewise constant distribution of the data. Note that the spatial variation of the solution within each cell has been lost in the averaging procedure. The task at hand is to *recover* the lost information by using the cells averages  $q_i^n$  to construct, or reconstruct, smooth functions  $p_i(x)$  that locally represent the spatial variation of the solution of (20.1). The reconstruction procedure to find the function  $p_i(x)$  must be conservative, that is, the integral average of  $p_i(x)$  inside cell  $i$  is identically  $q_i^n$ . In general, the functions  $p_i(x)$  have jump discontinuities at the cell interfaces  $x = x_{i+\frac{1}{2}}$ . Fig. 20.1 depicts typical reconstructed functions for three consecutive finite volumes  $i - 1$ ,  $i$  and  $i + 1$ . In this chapter we assume that the reconstructed functions are polynomials of a specified degree.

In principle, any of the classical interpolation techniques can be employed in the reconstruction procedure. In particular, one can construct polynomials on a *fixed* stencil formed by a set of averages

$$q_{i-k_L}^n, q_{i-k_L+1}^n, \dots, q_{i-1}^n, q_i^n, q_{i+1}^n, \dots, q_{i+k_R}^n, \quad (20.12)$$

where  $k_L$  and  $k_R$  are two non-negative integers.

A reconstruction based on a *fixed stencil* will be called a *linear reconstruction*. The resulting numerical schemes will be called *linear schemes*. Recall that from Godunov's theorem, see chapter 13, linear schemes of accuracy greater than one are oscillatory. If avoiding spurious oscillations is an important requirement on the numerical methods, then one must design *non-linear schemes*, even when applied to linear PDEs. A necessary condition for designing non-linear schemes is that the reconstruction be implemented on *variable* or *adaptive* stencils, rather than fixed stencils. Such reconstructions will be called *non-linear reconstructions*.



**Fig. 20.1.** Cell averages  $\{q_i^n\}$  define a piece-wise constant distribution  $\{p_i(x)\}$  on each cell  $i$ . Illustration of reconstructed polynomial functions  $p_{i-1}(x)$ ,  $p_i(x)$  and  $p_{i+1}(x)$  in cells  $i - 1$ ,  $i$  and  $i + 1$ , respectively.

The simplest non-linear reconstruction is based on the Total Variation Diminishing, or TVD, criterion studied in chapter 13. Higher-order non-linear ADER schemes can be constructed on the basis of other types of reconstructions, such as ENO and WENO interpolations.

The ENO (Essentially Non-Oscillatory) interpolation method [245] is suitable for approximating functions that contain discontinuities or exhibit large gradients. For a given interval  $[a, b]$  and a set of samples  $(x_i, h(x_i))$ , for  $i = 0, \dots, N$ , classical interpolation methods construct a *global polynomial* of order  $N$  that interpolates all points, using for example, the Newton form of the polynomial. Instead, in the ENO approach one constructs polynomials of order less than  $N$  that have three distinguishing features:

- they are *local*, that is, they are defined on sub-intervals  $I_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$
- they are conservative, that is

$$\frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} p_i(x) dx = q_i^n , \quad (20.13)$$

- they are *adaptive*, their coefficients depend on the behaviour of the data.

The local and adaptive nature of the interpolation method results in a representation polynomial that is essentially non-oscillatory (ENO).

A trivial example of a set of reconstructed polynomials is given by the set of constant polynomials

$$p_i(x) = \begin{cases} q_i^n & \text{if } x \in I_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] , \\ 0 & \text{otherwise ,} \end{cases} \quad (20.14)$$

resulting in a global, piece-wise constant, polynomial  $p_i(x)$ , for all  $x$  in a given interval  $I$ . This reconstruction satisfies trivially all three conditions above.

A first-degree polynomial in cell  $I_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$  is

$$p_i^{(1)}(x) = q_i^n + (x - x_i)\Delta_i^{(1)}, \quad (20.15)$$

where the slope, given as

$$\Delta_i^{(1)} = \Delta_i^{(1)}(\Delta_{i-\frac{1}{2}}, \Delta_{i+\frac{1}{2}}), \quad (20.16)$$

is a function of the two neighbouring gradients

$$\Delta_{i-\frac{1}{2}} = \frac{q_i^n - q_{i-1}^n}{\Delta x}, \quad \Delta_{i+\frac{1}{2}} = \frac{q_{i+1}^n - q_i^n}{\Delta x}. \quad (20.17)$$

A very simple choice for  $\Delta_i^{(1)}$  is the mean value

$$\Delta_i^{(1)} = \frac{1}{2}(\Delta_{i-\frac{1}{2}} + \Delta_{i+\frac{1}{2}}). \quad (20.18)$$

In the ENO approach one chooses the slope  $\Delta_i^{(1)}$  adaptively. For this first-degree polynomial one takes

$$\Delta_i^{(1)} = \begin{cases} \Delta_{i-\frac{1}{2}} & \text{if } |\Delta_{i-\frac{1}{2}}| \leq |\Delta_{i+\frac{1}{2}}|, \\ \Delta_{i+\frac{1}{2}} & \text{if } |\Delta_{i-\frac{1}{2}}| > |\Delta_{i+\frac{1}{2}}|. \end{cases} \quad (20.19)$$

Then, a global, piece-wise linear polynomial is defined as

$$h(x) = \begin{cases} p_i(x) & \text{if } x \in I_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}], \\ 0 & \text{otherwise.} \end{cases} \quad (20.20)$$

This reconstruction is equivalent to have considered two candidate polynomials  $q_L(x)$ ,  $q_R(x)$ , with corresponding two candidate stencils

$$S_L = \{q_{i-1}^n, q_i^n\} \text{ and } S_R = \{q_i^n, q_{i+1}^n\} \quad (20.21)$$

and then choosing the one with the smallest first divided difference. Higher-order ENO polynomials for each cell  $I_i$  can be constructed in an analogous manner.

Another choice for stencil-adaptive reconstructions is the WENO approach [331], [270] (Weighted Essentially Non-Oscillatory), whereby the sought polynomial  $p_i(x)$  is a weighted average of all polynomials that correspond to the candidate stencils in the ENO method. Further details of WENO are found in [493]. A modified WENO reconstruction method applicable to unstructured meshes in two and three space dimensions is proposed in [175], [176].

## 20.3 Second-Order Scheme for a Model Equation

Concerning accuracy, ADER is a unified approach, in that the same framework allows the choice of any desired order of accuracy. Here we illustrate the application of the ADER approach to construct a scheme of the form (20.2) of second order of accuracy. We follow the full general framework, as would be done for constructing an ADER scheme of any order of accuracy. The specific second-order scheme is presented in terms of the model advection equation with a source term

$$\partial_t q + \lambda \partial_x q = \beta q . \quad (20.22)$$

The flux function is  $f(q) = \lambda q$  and the source function is  $s(q) = \beta q$ , where  $\lambda$  and  $\beta \leq 0$  are known constants. We need to compute the numerical flux  $f_{i+\frac{1}{2}}$  and the numerical source  $s_i$  for the scheme (20.2).

### 20.3.1 Numerical Flux and Numerical Source

First we assume a reconstruction of the data in terms of first-degree polynomials  $p_i(x)$  of the form

$$p_i(x) = q_i^n + \Delta_i^{(1)}(x - x_i) , \quad p_{i+1}(x) = q_{i+1}^n + \Delta_{i+1}^{(1)}(x - x_{i+1}) , \quad (20.23)$$

with the slopes  $\Delta_i^{(1)}$  and  $\Delta_{i+1}^{(1)}$  chosen as in (20.19), for example. Then, to compute the numerical flux  $f_{i+\frac{1}{2}}$  according to (20.8) we solve the generalized Riemann problem  $GRP_1$

$$\left. \begin{aligned} & \text{PDE: } \partial_t q + \lambda \partial_x q = \beta q , \\ & \text{IC: } q(x, 0) = \begin{cases} p_i(x) & \text{if } x < 0 , \\ p_{i+1}(x) & \text{if } x > 0 . \end{cases} \end{aligned} \right\} \quad (20.24)$$

The solution of (20.24) takes the form

$$q_{i+\frac{1}{2}}(\tau) = q(0, 0_+) + \tau \partial_t q(0, 0_+) \quad (20.25)$$

and is completely determined by performing the following steps:

**Step (I): The leading term.** To compute the leading term  $q(0, 0_+)$  one solves the classical Riemann problem

$$\left. \begin{aligned} & \text{PDE: } \partial_t q + \lambda \partial_x q = 0 , \\ & \text{IC: } q(x, 0) = \begin{cases} p_i(0) \equiv q_i^n + \frac{1}{2} \Delta x \Delta_i^{(1)} & \text{if } x < 0 , \\ p_{i+1}(0) \equiv q_{i+1}^n - \frac{1}{2} \Delta x \Delta_{i+1}^{(1)} & \text{if } x > 0 , \end{cases} \end{aligned} \right\} \quad (20.26)$$

in which the initial condition consists of the boundary extrapolated values  $p_i(0)$  from the left and  $p_{i+1}(0)$  from the right. The similarity solution  $d_{i+\frac{1}{2}}^{(0)}(x/t)$  of this problem is

$$d_{i+\frac{1}{2}}^{(0)}(x/t) = \begin{cases} q_i^n + \frac{1}{2}\Delta x \Delta_i^{(1)} & \text{if } x/t < \lambda , \\ q_{i+1}^n - \frac{1}{2}\Delta x \Delta_{i+1}^{(1)} & \text{if } x/t > \lambda , \end{cases} \quad (20.27)$$

and the leading term is the Godunov state ( $x/t = 0$ )

$$q(0, 0_+) = d_{i+\frac{1}{2}}^{(0)}(0) = \begin{cases} q_i^n + \frac{1}{2}\Delta x \Delta_i^{(1)} & \text{if } \lambda > 0 , \\ q_{i+1}^n - \frac{1}{2}\Delta x \Delta_{i+1}^{(1)} & \text{if } \lambda < 0 . \end{cases} \quad (20.28)$$

**Step (II): The higher order term.** To compute the second term in (20.25) we do the following:

1. *Time derivative in terms of the spatial derivative.* The Cauchy–Kowalewski procedure uses the PDE (20.22) to express the time derivative in terms of the space derivative and the source term,

$$\partial_t q(x, t) = -\lambda \partial_x q + \beta q , \quad (20.29)$$

with the function  $q(x, t)$  evaluated at the leading term, so that the expansion (20.25) reads

$$q_{i+\frac{1}{2}}(\tau) = q(0, 0_+) + [-\lambda \partial_x q(0, 0_+) + \beta q(0, 0_+)]\tau . \quad (20.30)$$

2. *Evolution equation for the space derivative.* To determine the space derivative on the right-hand side of (20.30) we first construct an evolution equation for  $\partial_x q(x, t)$ ; it is easily seen that in fact  $\partial_x q(x, t)$  obeys the original evolution advection–reaction equation (20.22), namely

$$\partial_t(\partial_x q(x, t)) + \lambda \partial_x(\partial_x q(x, t)) = \beta \partial_x q(x, t) . \quad (20.31)$$

3. *Riemann problem for the spatial derivative.* We simplify (20.31) by neglecting the source term for the gradient. Then we pose the classical homogeneous Riemann problem for the spatial derivative:

$$\left. \begin{aligned} \text{PDE: } & \partial_t(\partial_x q(x, t)) + \lambda \partial_x(\partial_x q(x, t)) = 0 , \\ \text{IC: } & \partial_x q(x, 0) = \begin{cases} \Delta_i^{(1)} & \text{if } x < 0 , \\ \Delta_{i+1}^{(1)} & \text{if } x > 0 . \end{cases} \end{aligned} \right\} \quad (20.32)$$

We solve this Riemann problem to obtain the similarity solution

$$d_{i+\frac{1}{2}}^{(1)}(x/t) = \begin{cases} \Delta_i^{(1)} & \text{if } x/t < \lambda, \\ \Delta_{i+1}^{(1)} & \text{if } x/t > \lambda \end{cases} \quad (20.33)$$

and set

$$\partial_x q(0, 0_+) = d_{i+\frac{1}{2}}^{(1)}(0) = \begin{cases} \Delta_i^{(1)} & \text{if } \lambda > 0, \\ \Delta_{i+1}^{(1)} & \text{if } \lambda < 0. \end{cases} \quad (20.34)$$

**Step (III): Solution of  $GRP_K$ .** The final expression for the sought solution (20.25), (20.30) is

$$q_{i+\frac{1}{2}}(\tau) = \begin{cases} q_i^n + \frac{1}{2}\Delta x\Delta_i^{(1)} + \tau \left[ -\lambda\Delta_i^{(1)} + \beta(q_i^n + \frac{1}{2}\Delta x\Delta_i^{(1)}) \right], & \lambda > 0, \\ q_{i+1}^n - \frac{1}{2}\Delta x\Delta_{i+1}^{(1)} + \tau \left[ -\lambda\Delta_{i+1}^{(1)} + \beta(q_{i+1}^n - \frac{1}{2}\Delta x\Delta_{i+1}^{(1)}) \right], & \lambda < 0. \end{cases} \quad (20.35)$$

**Step (IV): Numerical flux.** Finally, the numerical flux in (20.2) is obtained according to (20.8), as

$$f_{i+\frac{1}{2}} = \begin{cases} \lambda \left( q_i^n + \frac{1}{2}(1-c)\Delta x\Delta_i^{(1)} + \frac{1}{2}r(q_i^n + \frac{1}{2}\Delta x\Delta_i^{(1)}) \right) & \text{if } \lambda > 0, \\ \lambda \left( q_{i+1}^n - \frac{1}{2}(1+c)\Delta x\Delta_{i+1}^{(1)} + \frac{1}{2}r(q_{i+1}^n - \frac{1}{2}\Delta x\Delta_{i+1}^{(1)}) \right) & \text{if } \lambda < 0, \end{cases} \quad (20.36)$$

where  $c = \lambda\Delta t/\Delta x$  is the Courant number and  $r = \Delta t\beta$  is a dimensionless reaction number.

**Remark:** the numerical flux (20.36) includes a contribution due to the source term. In other words, the flux knows of the source term, as one would expect. But note that there is still another contribution to the scheme resulting from the numerical source, still to be specified.

**Step (V): Numerical source.** To compute the numerical source  $s_i$  in the finite volume scheme (20.2) we need to select the function  $q_i(x, t)$  in the integrand of the volume integral (20.11) and an integration scheme. For a second order method we may apply the mid-point integration rule in space and time, so that one only requires  $q_i(x_i, \frac{1}{2}\Delta t)$ . At time  $t = 0$ , within the cell  $i$  we have the reconstructed polynomial  $p_i(x)$  given by (20.15)–(20.16). Taylor expanding the solution in time at the point  $x_i$  and applying the Cauchy–Kowalewski procedure we obtain

$$\left. \begin{aligned} q_i(x_i, \tau) &= q_i(x_i, 0_+) + \tau\partial_t q(x_i, 0_+) \\ &= q_i(x_i, 0_+) + \tau(-\lambda\partial_x q_i(x_i, 0_+) + \beta q_i(x_i, 0_+)) \end{aligned} \right\} \quad (20.37)$$

In the case considered we have

$$q_i(x_i, 0_+) = q_i^n = p_i(x_i), \quad \partial_x q_i(x_i, 0_+) = \Delta_i^{(1)}$$

so that the numerical source from (20.11) is

$$s_i = \beta \left( q_i^n + \frac{1}{2} \Delta t (-\lambda \Delta_i^{(1)} + \alpha q_i^n) \right). \quad (20.38)$$

### 20.3.2 The Scheme

The finite volume scheme (20.2) to solve the model advection–reaction equation (20.22) to second–order of accuracy is now completely determined, the numerical flux is given by (20.36) and the numerical source is given by (20.38). In full, for the case of  $\lambda > 0$  the scheme reads

$$\left. \begin{aligned} q_i^{n+1} &= q_i^n - c \left[ (q_i^n - q_{i-1}^n) + \frac{1}{2}(1-c)(\Delta_i^{(1)} - \Delta_{i-1}^{(1)}) \right] \\ &\quad - cr \left[ \frac{1}{2}(q_i^n - q_{i-1}^n) + \frac{1}{2}\Delta x(\Delta_i^{(1)} - \Delta_{i-1}^{(1)}) \right] \\ &\quad + r \left[ (1 + \frac{1}{2}r)q_i^n - \frac{1}{2}c\Delta x\Delta_i^{(1)} \right]. \end{aligned} \right\} \quad (20.39)$$

**Remark:** Note that if the slopes are all set to zero and we also set to zero the contribution of the source to the numerical flux, then we produce a first–order numerical scheme for (20.22) that consists of the first–order upwind method of Godunov for the advection part, along with an upwind discretization of the source term, namely

$$q_i^{n+1} = q_i^n - c(q_i^n - q_{i-1}^n) - r \left[ \frac{1}{2}cq_{i-1}^n + (1 - \frac{1}{2}c)q_i^n \right]. \quad (20.40)$$

This corresponds to one of the upwind schemes analyzed in [412] and [51] for solving advection equations with source terms.

**Remark:** By choosing the slopes  $\Delta_i^{(1)}$  appropriately, we can reproduce some well–known schemes, when  $\beta = 0$ . For example, for the case  $\lambda > 0$  in (20.22), the three classical second–order schemes of Warming–Beam, Lax–Wendroff and Fromm are identically reproduced by (20.39) by choosing the slope as follows

$$\Delta_i^{(1)} = \begin{cases} \Delta_{i-\frac{1}{2}} & \text{Warming–Beam ,} \\ \Delta_{i+\frac{1}{2}} & \text{Lax–Wendroff ,} \\ \frac{1}{2} (\Delta_{i-\frac{1}{2}} + \Delta_{i+\frac{1}{2}}) & \text{Fromm .} \end{cases} \quad (20.41)$$

**Remark:** the second–order ADER scheme with ENO reconstruction (20.15)–(20.19) is effectively a blend of the Warming–Beam and the Lax–Wendroff schemes.

## 20.4 Schemes of Arbitrary Accuracy

The procedure to construct ADER schemes of arbitrary accuracy is exactly the same as for the scheme of second-order of accuracy discussed in full detail in the preceding section. We thus describe the general case very succinctly.

### 20.4.1 The Numerical Flux

To compute the numerical flux  $f_{i+\frac{1}{2}}$  we solve the generalized Riemann problem  $GRP_K$

$$\left. \begin{aligned} \text{PDE: } & \partial_t q + \lambda \partial_x q = \beta q , \\ \text{IC: } & q(x, 0) = \begin{cases} p_i(x) & \text{if } x < 0 , \\ p_{i+1}(x) & \text{if } x > 0 , \end{cases} \end{aligned} \right\} \quad (20.42)$$

in which the reconstruction polynomials  $p_i(x)$  and  $p_{i+1}(x)$  are of arbitrary degree  $K$ . As discussed in chapter 19, the solution of (20.42) takes the form

$$q_{i+\frac{1}{2}}(\tau) = q(0, 0_+) + \sum_{k=1}^K \left[ \partial_t^{(k)} q(0, 0_+) \right] \frac{\tau^k}{k!} . \quad (20.43)$$

The leading term  $q(0, 0_+)$  is determined in exactly the same manner as for the second order scheme, the only difference being that the reconstructed polynomials  $p_i(x)$  and  $p_{i+1}(x)$  have obviously a difference form.

To compute the higher order terms one requires the determination of each time derivative  $\partial_t^{(k)} q(0, 0_+)$  as a function of space derivatives. As seen in chapter 19, application of the Cauchy–Kowalewski procedure to (20.22) yields

$$\partial_t^{(k)} q(x, t) = \sum_{n=0}^k \left[ \frac{k!}{(k-n)!n!} \beta^n (-\lambda)^{k-n} \right] \partial_x^{(k-n)} q(x, t) . \quad (20.44)$$

When the source term vanishes one has the simpler expression

$$\partial_t^{(k)} q = (-1)^k \lambda^k \partial_x^{(k)} q . \quad (20.45)$$

Solutions of classical Riemann problems for the spatial derivatives, as in (20.32)–(20.34), will give the spatial derivatives  $\partial_x^{(l)} q(0, 0_+)$  in (20.44), of any order  $l$ . Then, all temporal derivatives  $\partial_t^{(k)} q(0, 0_+)$ , as from (20.44), will be determined and thus the final solution (20.43) will be known.

Finally, according to (20.8) the numerical flux is obtained as

$$f_{i+\frac{1}{2}} = \frac{1}{\Delta t} \int_0^{\Delta t} \lambda \left[ q(0, 0_+) + \sum_{k=1}^K \left( \partial_t^{(k)} q(0, 0_+) \right) \frac{\tau^k}{k!} \right] d\tau . \quad (20.46)$$

In this simple case one can perform exact integration to obtain

$$f_{i+\frac{1}{2}} = \lambda \left[ q(0, 0_+) + \sum_{k=1}^K \left( \partial_t^{(k)} q(0, 0_+) \right) \frac{(\Delta t)^k}{(k+1)!} \right]. \quad (20.47)$$

For the homogeneous case,  $s(q) = 0$ , the explicit form of the numerical flux reads

$$f_{i+1/2} = \begin{cases} \lambda \left[ p_i(0) + \sum_{k=1}^K \left( (-1)^k \lambda^k \partial_x^{(k)} p_i(0) \right) \frac{(\Delta t)^k}{(k+1)!} \right] & \text{if } \lambda > 0, \\ \lambda \left[ p_{i+1}(0) + \sum_{k=1}^K \left( (-1)^k \lambda^k \partial_x^{(k)} p_{i+1}(0) \right) \frac{(\Delta t)^k}{(k+1)!} \right] & \text{if } \lambda < 0. \end{cases} \quad (20.48)$$

#### 20.4.2 The Numerical Source

To compute the numerical source  $s_i$  in the finite volume scheme (20.2) we need to select the function  $q_i(x, t)$  in the integrand of the volume integral (20.11) and an integration scheme. At time  $t = 0$ , within the cell  $i$  we have the reconstructed polynomial  $p_i(x)$ . To perform the integration we select a number of interior points in cell  $i$ , denoted by  $x_i^{(l)}$ ,  $l = 1, \dots, L$ . At each point  $x_i^{(l)}$  we write a time power series expansion

$$q(x_i^{(l)}, \tau) = q(x_i^{(l)}, 0_+) + \sum_{k=1}^K \left[ \partial_t^{(k)} q(x_i^{(l)}, 0_+) \right] \frac{\tau^k}{k!}. \quad (20.49)$$

The leading term is obtained from point values of the reconstructed polynomial, i.e.  $q(x_i^{(l)}, 0_+) = p_i(x_i^{(l)})$ . To compute the higher order terms we apply the Cauchy–Kowalewski procedure to express time derivatives as functions of space derivatives

$$\partial_t^{(k)} q(x_i^{(l)}, 0) = \sum_{n=0}^k \left[ \frac{k!}{(k-n)! n!} \beta^n (-\lambda)^{k-n} \right] \partial_x^{(k-n)} q(x_i^{(l)}, 0_+), \quad (20.50)$$

in which all spatial derivatives  $\partial_x^{(k-n)} q(x_i^{(l)}, 0_+)$  in the interior of cell  $i$  are evaluated using the reconstruction polynomial  $p_i(x)$ . The complete space–time volume integral in (20.11) is then evaluated, to the appropriate accuracy, to obtain the numerical source  $s_i$  for scheme (20.2).

#### 20.4.3 Summary

We have constructed a numerical method for solving the model scalar hyperbolic equation with a source term (20.22). One first applies a non–linear

reconstruction operator, TVD, ENO or WENO, to produce reconstruction polynomials  $p_i(x)$  of arbitrary degree  $K$ . The scheme to solve (20.22) is the finite volume one-step scheme (20.2), which requires the determination of the numerical flux  $f_{i+\frac{1}{2}}$  and the numerical source  $s_i$ . To construct the numerical flux one solves generalized Riemann problems (20.42), whose solutions are given by (20.43). Time integration as in (20.8) will give the numerical flux for (20.2). Then the computation of the numerical source  $s_i$  involves evaluation of the volume integral (20.11), which in turn requires the determination of the integrand; for this, one uses the Cauchy–Kowalewski method at a prescribed number of interior points  $x_i^{(l)}$ , as in (20.49)–(20.50), with the spatial derivatives computed from the available reconstructed polynomial  $p_i(x)$ .

The resulting numerical method is fully discrete, one-step, of arbitrary order of accuracy  $K + 1$  in space and time, requiring non-linear reconstruction only once per time step.

*Remark: stiff source terms.* The ADER methods studied so far work well for hyperbolic systems with non-stiff source terms. A new variant of the ADER approach has recently been put forward in [173] to solve hyperbolic equations with stiff source terms. The new scheme reconciles stiffness with high-order of accuracy in space and time. See also [171].

In the following section we show some selected numerical results to illustrate the performance of the high-order methods described in this chapter.

## 20.5 Sample Numerical Results

Here we present some sample ADER numerical results for the linear advection equation in one space dimension and for the linearized Euler equations in two space dimensions on unstructured meshes.

### 20.5.1 Long-Time Advection of Smooth Profiles

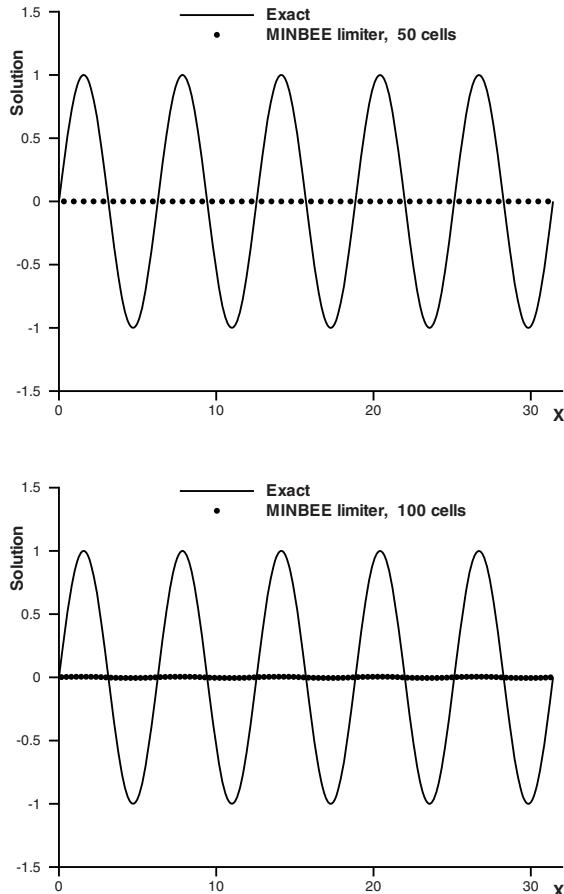
We consider the advection of smooth profiles by solving equation (20.22), with  $\beta = 0$  and  $\lambda = 1$ , in a domain  $[0, 10\pi]$  and a smooth initial profile given by

$$q(x, 0) \equiv q_0(x) = \sin(\pi x), \quad (20.51)$$

with periodic boundary conditions. The exact solution is  $q(x, t) = q_0(x - \lambda t)$ .

A very challenging problem is the advection of the solution for very long times. Conventional TVD methods are known to have great difficulties in resolving this type of advection problems, particularly for very long evolution times.

Fig. 20.2 shows the results from the TVD scheme WAF with the MINBEE limiter function, for two different meshes, at the output time  $t = 1000\pi$ .

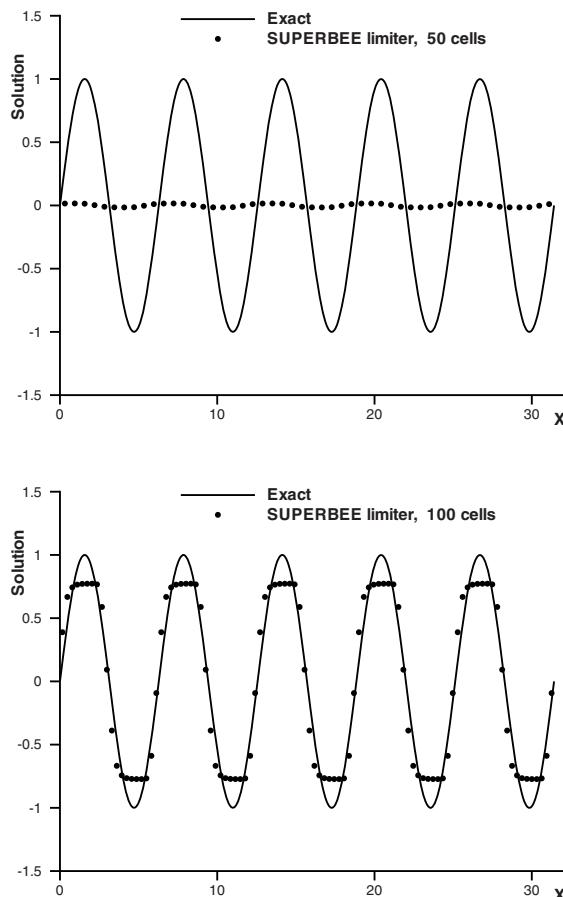


**Fig. 20.2. Linear advection.** Results from TVD scheme with MINBEE limiter (symbols) at time  $t = 1000\pi$  using meshes of 50 and 100 cells, with  $C_{cfl} = 0.95$ . Exact solution shown by full line (Courtesy of Dr. V. A. Titarev).

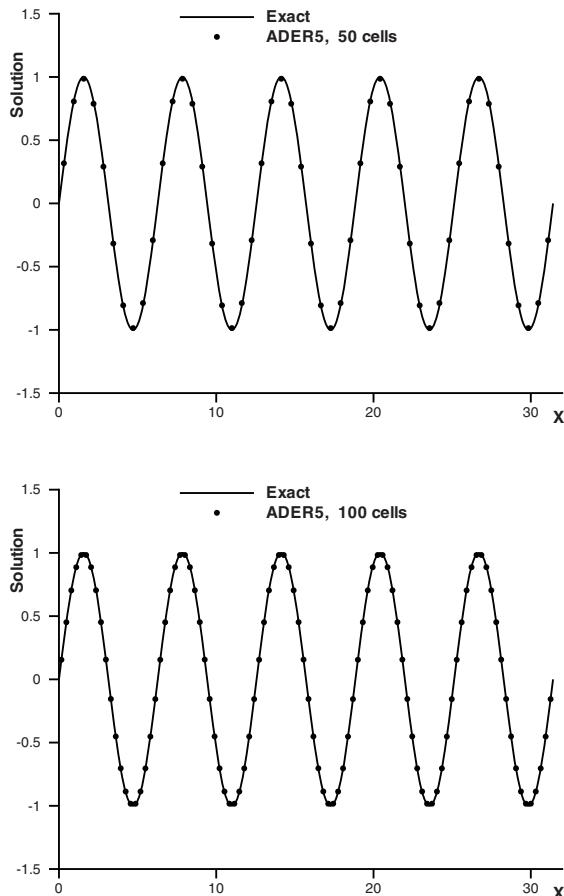
These results clearly illustrate the limitations of TVD methods for solving long-time evolution problems. The use of diffusive limiters, such as MINBEE (or MINMOD) tend to clip extrema very severely, reducing peak values very rapidly as time evolves.

Using compressive limiter functions in TVD methods tends to reduce the overall error but at the cost of *squaring* smooth features. Fig. 20.3 shows the results of the WAF TVD method with the SUPERBEE limiter, for two meshes, at the output time  $t = 1000\pi$ . Compare results with those of Fig. 20.2.

Methods of uniform high order of accuracy, such as the ADER methods introduced in this chapter, give better results for long-time evolution problems, particularly when having to resolve rapid but smooth variations of the solution. The true solution to a physical problem may be smooth but highly oscillatory. The numerical methods must be capable of resolving these physically meaningful oscillations, which should not be artificially suppressed. Fig. 20.4 shows results from the fifth order ADER method, for two meshes, at the output time  $t = 1000\pi$ . Compare with results for Fig. 20.2 and Fig. 20.3. The ADER results are clearly superior to those of TVD methods.



**Fig. 20.3. Linear advection.** Results from TVD scheme with SUPERBEE limiter (symbols) at time  $t = 1000\pi$  using meshes of 50 and 100 cells, with  $C_{cfl} = 0.95$ . Exact solution shown by full line (Courtesy of Dr. V. A. Titarev).



**Fig. 20.4. Advection of smooth profile.** Results from 5-th order ADER scheme (symbols) at time  $t = 1000\pi$  using meshes of 50 and 100 cells, with  $C_{cfl} = 0.95$ . Exact solution shown by full line (Courtesy of Dr. V. A. Titarev).

### 20.5.2 Convergence Rates

A most important task is the verification of the expected accuracy of a numerical method. All too often, claims are made regarding the accuracy of a particular scheme, which may well be ill-founded. We stress that even the construction of a *non-linear* second-order method for inhomogeneous equations is not a trivial task. This is even more challenging when using unstructured meshes.

Here we reproduce some results taken from [178], obtained from ADER schemes implemented in the frame of discontinuous Galerkin finite element methods. The results are for the two-dimensional linearized Euler equations, solved in a two-dimensional domain discretized with unstructured triangular meshes.

Table 20.1 shows empirical convergence rates for schemes of second to tenth order of accuracy in space and time. We remark that for the convergence rates study, *purposely distorted* meshes, such as those of Fig. 20.6, have been used. The first column of Table 20.1 shows the size  $h$  for the sequence of meshes considered in the study. The second column shows the corresponding number of elements; the third, fourth and fifth columns show errors in three different norms and the remaining three columns show the corresponding convergence rates. It is seen that the empirically obtained convergence rates are consistent with the theoretically expected accuracy.

Fig. 20.5 shows error against reciprocal of mesh size (top) and against CPU time (bottom) for schemes of second to tenth order of accuracy. In general it is seen that for the range of smaller errors it is distinctly better to use a high-order method than a lower-order one.

## 20.6 Concluding Remarks

We have presented an introduction to the family of ADER schemes for solving hyperbolic equations with source terms. These methods are a natural extension of the classic Godunov first-order upwind method. The basic ingredients of the method are (a) a high-order spatial representation of the data, instead of the piece-wise constant data of the Godunov method; (b) solution of the generalized Riemann problem at each cell interface, instead of the classical piece-wise constant data Riemann problem of the Godunov method; (c) the numerical flux is computed from a time-integral of the solution of the generalized Riemann problem at the interface; for the Godunov method this last step is trivial, as the similarity solution of the classical Riemann problem along the interface is constant. Source terms are treated in a very natural way; in addition to the numerical source computed to high order there is also a contribution of the source to the numerical flux, which comes in through the solution of the generalized Riemann problem.

The ADER methods are one-step, fully discrete schemes for which the reconstruction procedure is carried out only once per time step, and can be implemented in the framework of finite volumes and discontinuous Galerkin finite elements. The applicability of the ADER methods is not restricted to first-order systems, such as hyperbolic equations with source terms. For example, the ADER methodology has recently been applied to reaction-diffusion equations (parabolic), extending the concept of Riemann problem [535]; schemes of up to 10th order of accuracy in space and time are reported. See also the related work [224]. ADER, being of arbitrary order, allows the user to select the

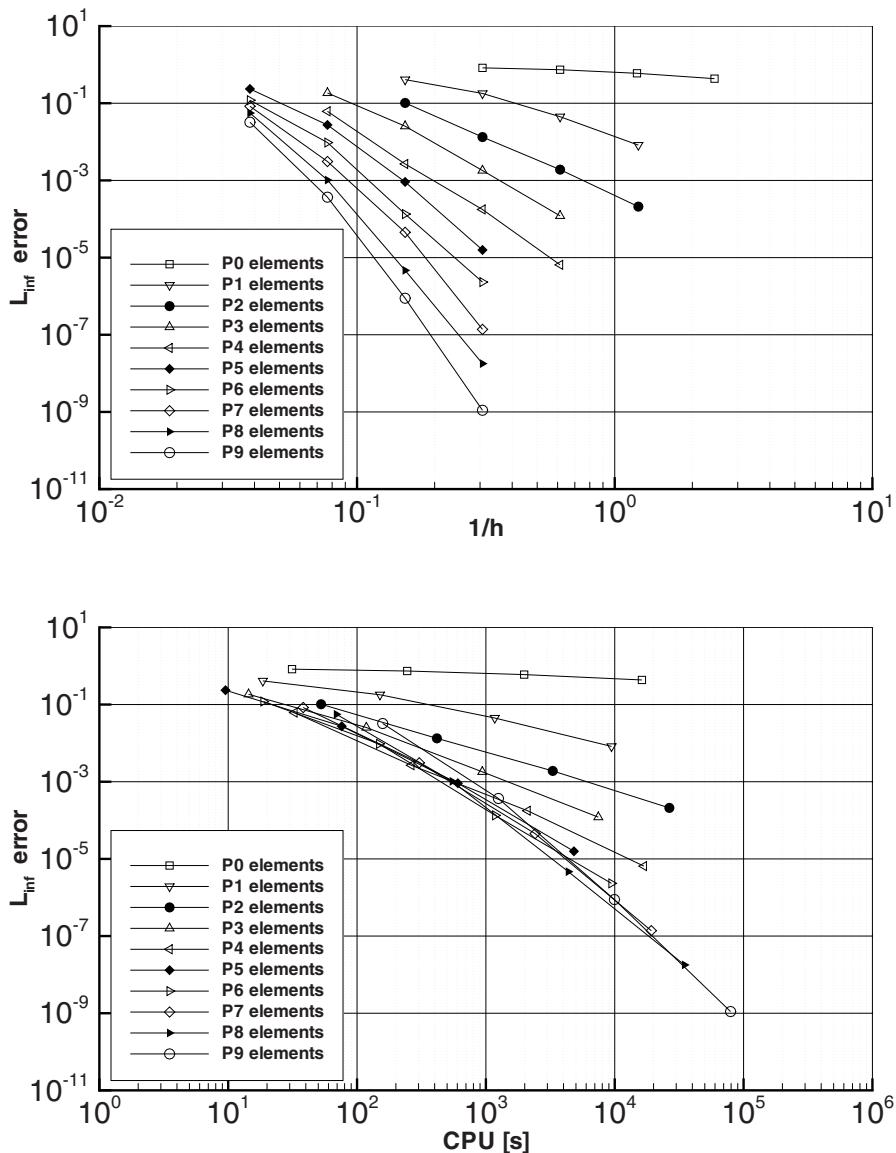
wanted level of complexity by choosing the desired order of accuracy. The full framework is the same, the first-order Godunov method being at the bottom of the hierarchy, for the finite volume framework.

In this chapter we have given a very basic introduction to the ADER methods in the framework of finite volumes, for solving one-dimensional hyperbolic systems with non-stiff source terms. We have also shown some numerical results that illustrate the limitation of second-order TVD methods, particularly for long-time evolution problems. We have also displayed some results of the ADER approach in the framework of discontinuous Galerkin finite element methods. Results of convergence rate studies verifies that the schemes, even on very distorted unstructured meshes do achieve the theoretically expected orders of accuracy; results for ADER schemes of up to 10th order of accuracy have been shown.

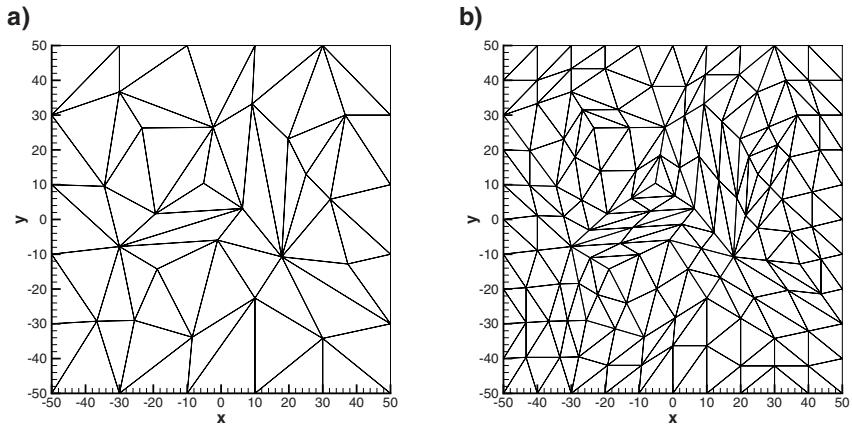
Some recent developments of the ADER approach for hyperbolic conservation laws are found in the following references: [89], [175], [176], [177], [419], [173], [171]. Current research concerns the extension of the ADER approach to solve non-linear hyperbolic systems with source terms, for which the differential, or principal part, cannot be expressed in conservation-law form. These systems are sometimes referred as non-conservative systems. Recent, still unpublished, communications on the subject are [85] and [174].

ADER-DG O2							
$h$	$N_d$	$L^\infty$	$L^1$	$L^2$	$O_{L^\infty}$	$O_{L^1}$	$O_{L^2}$
6.51	2,978	4.0950E-01	7.3710E+01	3.1185E+00			
3.26	11,913	1.7949E-01	2.1142E+01	1.0785E+00	1.2	1.8	1.5
1.63	47,653	4.4793E-02	3.9943E+00	2.2639E-01	2.0	2.4	2.3
0.81	190,610	8.2223E-03	6.3435E-01	3.7319E-02	2.4	2.7	2.6
ADER-DG O4							
$h$	$N_d$	$L^\infty$	$L^1$	$L^2$	$O_{L^\infty}$	$O_{L^1}$	$O_{L^2}$
13.03	2,478	1.8482E-01	3.1944E+01	1.3455E+00			
6.51	9,928	2.5445E-02	2.1268E+00	1.0901E-01	2.9	3.9	3.6
3.26	39,710	1.8028E-03	6.0709E-02	3.4818E-03	3.8	5.1	5.0
1.63	158,842	1.2070E-04	2.8285E-03	1.8062E-04	3.9	4.4	4.3
ADER-DG O6							
$h$	$N_d$	$L^\infty$	$L^1$	$L^2$	$O_{L^\infty}$	$O_{L^1}$	$O_{L^2}$
26.05	1,302	2.3490E-01	5.7109E+01	1.9630E+00			
13.03	5,208	2.7394E-02	4.1956E+00	1.8993E-01	3.1	3.8	3.4
6.51	20,832	9.1437E-04	4.6738E-02	2.3541E-03	4.9	6.5	6.3
3.26	83,328	1.5737E-05	1.0614E-03	5.6039E-05	5.9	5.5	5.4
ADER-DG O8							
$h$	$N_d$	$L^\infty$	$L^1$	$L^2$	$O_{L^\infty}$	$O_{L^1}$	$O_{L^2}$
26.05	2,232	8.3906E-02	2.3316E+01	8.4765E-01			
13.03	8,928	3.0905E-03	4.7899E-01	2.2030E-02	4.8	5.6	5.3
6.51	35,712	4.5237E-05	2.6535E-03	1.2786E-04	6.1	7.5	7.4
3.26	142,848	1.3771E-07	1.4681E-05	7.7184E-07	8.4	7.5	7.4
ADER-DG O9							
$h$	$N_d$	$L^\infty$	$L^1$	$L^2$	$O_{L^\infty}$	$O_{L^1}$	$O_{L^2}$
26.05	2,790	5.5839E-02	1.4153E+01	5.2238E-01			
13.03	11,160	1.0213E-03	1.6167E-01	7.4949E-03	5.8	6.5	6.1
6.51	44,640	4.6598E-06	4.9232E-04	2.5595E-05	7.8	8.4	8.2
3.26	178,560	1.7872E-08	1.3131E-06	7.0996E-08	8.0	8.6	8.5
ADER-DG O10							
$h$	$N_d$	$L^\infty$	$L^1$	$L^2$	$O_{L^\infty}$	$O_{L^1}$	$O_{L^2}$
26.05	3,410	3.2304E-02	8.6122E+00	3.2389E-01			
13.03	13,640	3.6709E-04	5.4707E-02	2.6230E-03	6.5	7.3	6.9
6.51	54,560	8.8532E-07	1.0717E-04	5.2679E-06	8.7	9.0	9.0
3.26	218,240	1.1050E-09	9.6184E-08	5.2430E-09	9.6	10.1	10.0

**Table 20.1.** ADER discontinuous Galerkin schemes. Convergence rates for schemes of second to tenth order of accuracy in space and time, as applied to the linearized Euler equations on the very irregular unstructured meshes depicted in Fig. 20.6. (Courtesy of Dr. M. Dumbser).



**Fig. 20.5.** ADER discontinuous Galerkin schemes. Error against reciprocal of mesh size (top) and against CPU time (bottom) for schemes of second to tenth order of accuracy in space and time. (Courtesy of Dr. M. Dumbser).



**Fig. 20.6.** Grid topologies for convergence study of ADER discontinuous Galerkin schemes as applied to the linearized Euler equations: (a) very coarse mesh (b) finer mesh (Courtesy of Dr. M. Dumbser).

## Concluding Remarks

Here we first summarize the main themes of this book, point out some related topics that have been excluded and give updated supplementary references. We also discuss potential practical applications of the methods studied and point out some current areas of research on numerical methods. Finally, we briefly describe the library NUMERICA.

### 21.1 Summary of Numerical Aspects

We have studied a family of finite volume methods for solving numerically systems hyperbolic conservation laws. The emphasis has been on the computation of the numerical flux, for which a large variety of methods have been studied, all of which consider the classical Riemann problem. Then, there are essentially two ways of making use of the Riemann problem, the *upwind* approach and the *centred* approach. In the *upwind* approach one solves the Riemann problem exactly, or approximately, in order to extract wave propagation information contained in the differential equations to be used explicitly in the computation of the numerical flux. In the *centred approach* one avoids the direct solution of the Riemann problem and thus no explicit information on wave propagation is utilized in the construction of the numerical flux. Most of the book has been devoted to the study of ways of solving the Riemann problem, see chapters 4, 9, 10, 11 and 12. An alternative upwind methodology to determine a numerical flux is offered by the Flux Vector Splitting (FVS) approach; FVS methods are studied in chapter 8, where three different splittings are considered. The Random Choice Method (RCM) of chapter 7 is somewhat special, in that it uses the exact solution of the classical Riemann problem, see chapter 4, and gives infinite resolution of discontinuities but does not extend to general systems with more than two independent variables. We have also studied the *centred approach*, starting with the classics, namely Lax–Wendroff, Lax–Friedrichs and Godunov centred. A more recent centred flux is the FORCE flux, introduced in chapter 7 for the one-dimensional case.

FORCE is related to the centred methods advocated by Tadmor and collaborators. The multidimensional version of FORCE is studied in chapter 17 for structured and unstructured meshes.

Numerical fluxes based on the explicit solution of the Riemann problem can be divided into two classes: *complete Riemann solvers* and *incomplete Riemann solvers*. In the first class, the underlying solution of the Riemann problem accounts for all the characteristic fields present in the structure of the exact solution of the Riemann problem. Otherwise we speak of incomplete Riemann solvers. It is well known that for computing isolated shock waves, the performance of complete Riemann solvers, incomplete Riemann solvers, FVS fluxes and centred fluxes is very similar. However, for intermediate waves, such as contact discontinuities and shear waves, there can be a big difference in resolution. Complete Riemann solvers are the most accurate. All other fluxes, with the exception of the FVS splitting of Liou and Steffen, contain excessive artificial viscosity. This also has a bearing on the accuracy of computations for more complex systems, such as the Navier–Stokes equations, in resolving shear layers for instance.

At the first-order level, the emphasis has been on fluxes that are *monotone*. Having established several ways of defining monotone numerical fluxes, we have then introduced ways of constructing higher order methods, with most of the material devoted to the second-order case. However, in view of Godunov’s theorem, proved in chapter 13, useful schemes of accuracy greater than one must be non-linear. We have introduced the TVD criterion for constructing non-linear schemes, see chapter 13. Then in chapter 14 we have extended, empirically, the TVD concept to non-linear systems.

The generalized Riemann problem is studied in chapter 19, which forms the basis for the construction of ADER type schemes introduced in chapter 20, for the case of model equations in one space dimension. The ADER schemes are explicit, fully discrete one-step non-linear and of arbitrary order of accuracy in space and time. The ADER schemes can be implemented in the frame of finite volumes or discontinuous Galerkin finite elements. The non-linear version of the ADER schemes is based on ENO and WENO reconstruction procedures for higher-order schemes. For second-order non-linear ADER schemes one can also enforce the TVD criterion. The ADER schemes have many features in common with the already classical ENO/WENO schemes, not studied in this book, and for which relevant references are given in chapter 20.

There are important topics at the level of numerical methods that have not been included. One example is the Piecewise Parabolic Method (PPM) of Colella and Woodward [137], [584]. This is an excellent method of the Godunov type and can be implemented following the basic material presented here. We have not included implicit methods. Useful references here are Harten [239]; Yee [595] and specially the VKI Lectures [593]. We have not included material related to multidimensional upwinding, for which the interested reader may consult the classical papers of Roe and collaborators [411]; Deconinck, Powell, Roe and Struijs [152]; van Leer [563]; Deconinck, Roe and Struijs [153];

García–Navarro, Hubbard and Priestley [197] and Baines and Hubbard [22]. More recent advances are reported, for example, in [113], [159] and [2], and in references therein. The upwind and centred numerical fluxes studied in this book can be used in the framework of the finite volume methods, as illustrated here. These fluxes can also be used in the framework of discontinuous Galerkin finite element methods, not studied here. Useful background on discontinuous Galerkin finite element methods is found, for example, in papers by Cockburn and collaborators, see for example [125], [123], [121], [126] and [128].

I note that there are by now many books that are related to the main themes of this textbook, some of them very recent. Examples include the books by Godlewski and Raviart [215]; Laney [298]; Kulikovskii, Pogorelov and Semenov [292]; Oran and Boris [366]; Toro [520]; LeVeque [311]; Guinot [229]; Ben–Artzi and Falcovitz [39]; Drikakis and Rider [165]. We also recommend the edited review book [545], which contains 97 refereed articles by leading authors on Godunov methods and applications. Two books on discontinuous Galerkin finite element methods are Cockburn, Karniadakis and Shu [122]; Hesthaven and Warburton [248]; Wesseling [581]; and Feistauer, Felcman and Straskraba [189].

## 21.2 Potential Applications

The methods studied in this book can be applied to a large variety of practical problems. The paper by Roe [415] contains some examples of very ambitious applications of upwind methods, in particular.

The **shallow water equations** are a popular hyperbolic system for modelling environmental problems. The methods studied here are applicable to these equations. Useful references are Marshall and Méndez [339]; Glaister [207]; Toro [504]; Watson, Peregrine and Toro [576]; Alcrudo and García–Navarro [4]; Bermúdez and Vázquez [51]; García–Navarro, Hubbard and Priestley [197]; Bermúdez et al. [50]; Fraccarollo and Toro [194]; Borthwick, Fujihara and Rogers [73], to name but a few. The recent review paper of Toro and García–Navarro [534] contains a large number of updated references on this topic. See also the recent textbook by Toro [520].

The **Navier–Stokes equations** are another system of practical interest, for which some of the methodology studied here could be useful. The steady incompressible Navier–Stokes equations may be formulated via the artificial compressibility approach of Chorin [109] to produce the so-called *Artificial Compressibility Equations*. Details of the derivation of these equations are given in chapter 1. The inviscid part of the equations constitute a hyperbolic system and one can therefore apply all the methods presented in this book. The approach can also be extended to solve the time-dependent incompressible Navier–Stokes equations. Useful, early references are the following: Marx

[345], [346]; Drikakis [162], [163]; Weinan and Shu [579]; Turkel and Arnone [550]; Hänel and Sharma [235]; Tamamidis, Zhang, Assanis [481]; Toro [514] and Drikakis et al. [164]. More recent, advanced methods are described in the paper by Bassi and collaborators [30] and in references therein. See also the recent textbook by Drikakis and Rider [165], and references therein.

The compressible Navier–Stokes equation, see chapter 1, are the obvious next extension of the methods presented in this book. A very useful reference here are the VKI Lectures by Hänel [234]. Classical papers on this subject are, for example, those of van Leer, Thomas and Roe [565]; Brown [82]; Toro and Brown [529]; Drikakis and Tsangaris [166] and Allmaras [5]. A word of caution is in order, when applying TVD methods to solve the Navier–Stokes equations. TVD methods are effectively a smart way of introducing numerical viscosity to control spurious oscillations near large gradients. As pointed out in chapter 13, section 13.9.2, in the presence of physical viscosity, the standard TVD condition is, strictly speaking, incorrect, in that it does not acknowledge the existence of physical viscosity. The net result is excessive artificial viscosity. For most problems, satisfactory results are reported in the literature, but for special problems concerned with fine details of the flow physics, one would expect the limitations of standard TVD methods for viscous problems to become apparent. Standard TVD methods are techniques that have been constructed for inviscid problems. For viscous problems, new TVD methods are needed. See [503] for possible extensions of TVD methods for viscous problems. Advanced methods are described in [31], [127] and [129], for example.

The Euler equations can be used to model wave propagation in **general compressible media**, gases being the most obvious example. For problems under extreme thermodynamic conditions, real gas effects become important and use of real gas equations of state must be considered. For Godunov-type schemes this requires the solution of the Riemann problem for a general equation of state. The media may be some material other than a gas, a liquid for instance, which at high pressure is compressible. See Iving, Causon and Toro [264] for details on Riemann solvers for compressible liquids. The review paper by Menikoff and Plohr [349] is highly recommended. See also the papers by Colella and Glaz [135]; Glaister [208], [209]; Saurel, Larini and Loraud [429]; Quartapelle et al. [398]; and Müller and Voss [356]. The techniques presented in this textbook may also be applied to model **wave propagation in solids**. Useful references are, amongst many others, Kim and Ballmann [284]; Lin and Ballmann [320], [323], [321], [322]. See the recent paper of Gavrilyuk, Favrie and Saurel [199]; and that of Titarev, Romenski and Toro [491].

**Detonation waves** in gases and solids is an extensive area of application of the numerical methods of this book. Classical references on detonation waves are, amongst many others, Oppenheim and Soloukhin [365]; Fickett and Davis [191]; Williams [583]; Clarke [117], [116]; Edwards, Thomas and

Williams [180]; Gelfand, Frolov and Nettleton [201] and Gilbert [204]. Useful references on numerical computation of detonation waves are Taki and Fujiwara [479]; Oran and Boris [367], [368]; Kailasanath, Oran, Boris and Young [276]; Colella, Majda and Roytburd [136]; Bourlioux, Majda and Roytburd [78]; Bourlioux and Majda [77]; Singh [448]; Singh and Clarke [449]; Clarke et al. [118] and Nikiforakis and Clarke [363]. See also the book by Oran and Boris [366] and references therein.

**Astrophysics** is a very active area of research in which numerical methods of the kind studied here are used. Common systems of equations used are the **relativistic Euler equations** and the equations of **ideal and relativistic magnetohydrodynamics**, or **MHD equations**. The exact solution of the Riemann problem for relativistic gas dynamics was first reported by Martí and Müller [342]. They extended the shock–tube problem solution of Thompson [488] and highlighted the importance of their work in the context of numerical relativistic hydrodynamics via Godunov–type methods. See also the papers by Schneider, Katscher, Rischke, Waldhauser, Maruhn and Munz [431]; Plohr and Sharp [386]; Marquina, Martí, Ibañez, Miralles and Donat [338]; Martí et al. [341], [342], [343] and Falle and Komissarov [187]. The reader interested in the simulation of astrophysical flows should consult the book by LeVeque, Mihalas, Dorfi and Müller [312].

Early papers on the MHD equations are, for example, Brio and Wu [80]; van Puten [566]; Dai and Woodward [146], [147]; Tanaka [482]; Barmin, Kulikovskiy and Pogorelov [29] and Falle and Komissarov [188]; and Balsara [26], [27]. More recent works are Dedner, Kemm, Kröner, Munz, Schnitzer and Wessenberg [154]; Gurski [230]; Honkkila and Janhunen [255]; Li [318]; Mignone and Bodo [351]; Klingenberg, Schmidt and Waagan [285]; Balsara, Rumpf, Dumbser and Munz [28]; Sangan [425]; Bodo, Masaglia, Mignone and Rossi [68]; Dumbser, Balsara, Toro and Munz [171]. An important recent work to mention here is the paper by Giacomazzo and Rezzolla [202]; they found the exact solution of the Riemann problem for relativistic magnetohydrodynamics, which is very valuable for assessing numerical schemes.

We finally mention the fact that **numerical relativity** is today a growing area of application of numerical methods. For a review on the subject see Lehner [306].

**Multiphase flow** has become a large area of application of numerical methods of the kind studied in this book. Such flows usually involve a *fluid* mixture of different materials, typically solid particles (solid phase) immersed in a gas (gaseous phase) or liquid (liquid phase). The physical concept of *phase* is generalized to *field* or *fluid*, to allow for mixtures of physical phases with different material properties, for example air and steam. In this manner the mathematical models for multifield flow may consider a large number of *fields* or *phases*. A distinguishing feature of these systems is that each field has its own velocity vector, unlike the case of multicomponent flow in which a single

velocity vector is needed, namely that of the *carrier fluid*. Useful background on multiphase flow is found, for example, in the textbooks Drew and Gulliver [160]; Drew, Marsden and Sirovich [161]; Jackson [265]; Gidaspow [203]; Ishii and Hibiki [261]; Stadtke [460]; Brennen [79]; Prosperetti and Tryggvason [388].

Multiphase flow models have, in general, the following distinguishing features: (i) the equations involve source terms, (ii) there may be non-conservative products and (iii) the equations may, in the absence of dissipative effects, be of mixed elliptic–hyperbolic type. The classical paper of Stewart and Wendroff [464] addresses some of these issues. Interest on non-conservative products is relatively new and will be addressed in more detail below. However it should be pointed out that, at present, there are some mathematical models for multiphase flow that are conservative, see Romenski et al. [417], for instance.

Regarding the application of Godunov-type numerical methods and related Riemann problems to compressible multiphase flow, perhaps the earliest reported attempt is found in [501]; see also [525]. Other works are reported by Sainsaulieu [421]; Saurel, Daniel and Loraud [427]; Saurel, Forestier, Veyret and Loraud [428]; Toumi and Kumbaro [549]; Toumi [547]; Toumi and Caruge [548]; Cortes, Debussche and Toumi [142]; Saurel and Abgral [426]; Stadtke et al. [461]; Castro and Toro [90]; Chang and Liou [102].

A number of papers have centred on the Baer–Nunziato equations [21]. Embid and Baer [183] carried out a comprehensive mathematical analysis of these equations. Some publications have addressed the challenging task of solving the Riemann problem for the Baer–Nunziato system. Andrianov and Warnecke [14] studied the *indirect* Riemann problem; that is they assumed a solution and found the corresponding initial conditions. Such solutions have proved useful in assessing numerical methods. Schwendeman, Wahle and Kapila [438] appear to be the first to have solved the *direct* Riemann problem for the Baer–Nunziato equations. This is more useful, as the solution can also be used to compute Godunov-type fluxes. Deledicque and Miltiadis [156] have also solved the direct Riemann problem for a Baer–Nunziato type model. Building upon the exact solution of [438] of the classical Riemann problem, Toro and Castro [531] solved the generalized Riemann problem for the Baer–Nunziato system. New path-conservative schemes for these equations in three space dimensions on unstructured meshes have recently been reported by Dumbser, Hidalgo, Castro, Par s and Toro [174].

**Free-surface two phase flow** models have also appeared recently, starting with the work of Pitman and Le [385]. Corresponding numerical methods for these equations have also been reported, see for example the works of Pelanti, Bouchut and Mangeney [376]; Rhebergen, Bokhove and van der Vegt [401] and Dumbser, Castro, Par s and Toro [172].

## 21.3 Current Research Topics

**Non-conservative products** is a topic that has attracted the attention of numerical analysts in recent years. The presence of non-conservative products means that the principal part of the equations cannot be written in *conservative form* or *divergence form*. At the mathematical (and physical) level this leads to the problem of defining weak solutions. Dal Maso, LeFloch and Murat [148] have addressed this issue from the mathematical point of view, see also the early work of Volpert [571]. Numerically, it is also difficult to treat non-conservative terms properly. The pioneering work of Toumi [546] and of Toumi and Kumbaro [549], based on the work [148], has recently lead to important developments of suitable numerical methods to deal with non-conservative systems. See the recent works of Parés and collaborators [375], [93], [94], [85] [172] and [174]. See also the work of Rhebergen, Bokhove and van der Vegt [401].

**Source terms** are invariably present in most systems of practical interest. In particular, multiphase flow systems are inherently inhomogeneous, the various phases are coupled through the presence of source terms. Dealing with source terms properly still represents a serious numerical challenge, although significant progress has been made in recent years. A significant step forward was Roe's proposition [412] of using wave propagation information (upwinding) to discretize the source terms. A fundamental requirement is that the numerical method be *well-balanced*. Glimm, Marshall and Plohr [214] were probably the first to address this issue. Other early works are those of Bermúdez and Vázquez [51]; Vázquez [569]; Greenberg and LeRoux [225]; Greenberg, LeRoux, Baraille and Noussair [226]; LeVeque [310] and Gosse [220]. Some of the more recent works on the subject include Gallardo, Parés and Castro [196]; Castro, Gallardo, López and Parés [92] and references therein. Some very recent advances on the subject are reported in Pardo et al. [94], Canestrelli et al. [85]; Dumbser et al. [172], [174].

In many applications the source terms are *stiff*. A new approach to deal with hyperbolic systems with stiff source terms is presented in Dumbser, Enaux and Toro [173] in the frame of high order ADER schemes. The method introduces a locally implicit numerical solver for the generalized Riemann problem, while retaining a globally explicit one-step method, reconciling in this manner high accuracy and stiffness.

**New Riemann solvers** remain the subject of current research. The challenge is to devise a method of solution that results in a *complete Riemann solver*, that is, a numerical flux in which all characteristic fields are accounted for. This is vital for resolving intermediate characteristic fields. Robustness and generality are two fundamental requirements on such Riemann solver. The MUSTA approach proposed by Toro [496] contains some potential that may lead to such an ideal Riemann solver. MUSTA is a predictor-corrector, or

*multi-stage*, approach, whereby the initial conditions are evolved in time for a number of stages, prior to a corrector stage that leads to the sought interface numerical flux. Further developments of the MUSTA approach are reported in [523], [524], [521], [495] and [357]. A promising variant of MUSTA, called EVILIN, is reported in [522], whereby a simple centred-type flux is used as predictor. In the corrector step one uses a simple linearization of the exact solution of the Riemann problem, resulting in a complete Riemann solver. EVILIN enjoys a degree of generality but the question of robustness remains the subject of further investigations.

## 21.4 The NUMERICA Library

*NUMERICA* [518] is a library of source codes for teaching, research and applications. The library is based on the contents of this book and [520]. The source codes of the library can be useful as *learning and teaching* tools. First, they can be used in a *self-study situation* as another way of reading the details of a numerical method and second, they can be used as a method of demonstrating, in a *teaching situation*, the main steps of a numerical technique and a possible way of coding the methods. The programs can also be utilized as *development programs*. First, they can most easily be converted into practical tools for *real* applications. They can also form the bases for much more ambitious applications, such as the ones discussed above. For details on *NUMERICA* the reader should contact the author directly: Professor E. F. Toro, Laboratory of Applied Mathematics, Department of Civil and Environmental Engineering, University of Trento, Italy; email: *toro@ing.unitn.it*; website: <http://www.ing.unitn.it/toro>. The library is freely available to academics and can be downloaded from the website. Non academics should contact the following email: *inquiries@numeritek.com*.

---

## References

1. R. Abgrall. How to Prevent Pressure Oscillations in Multicomponent Flow Calculations: A Quasiconservative Approach. *J. Comput. Phys.*, 125:150–160, 1996.
2. R. Abgrall and M. Mezine. Construction of Second-Order Accurate Monotone and Stable Residual Distribution Schemes for Steady Problems. *J. Comput. Phys.*, 195:474–507, 2004.
3. C. M. Albone. Report on the AGCFM Working Party on Software Quality Assurance. Technical Report AERO 2105, Royal Aircraft Establishment, Farnborough, UK, 1987.
4. A. Alcrudo and P. García-Navarro. A High Resolution Godunov-Type Scheme in Finite Volumes for the 2D Shallow Water Equations. *Int. J. Num. Meth. Fluids*, 16:489–505, 1993.
5. S. R. Allmaras. Contamination of Laminar Boundary Layers by Artificial Dissipation in Navier–Stokes Solutions. In *Numerical Methods for Fluid Dynamics*, 4, pages 443–449. Oxford University Press, 1993.
6. V. R. Ambati and O. Bokhove. Space–Time Discontinuous Galerkin Discretization of Rotating Shallow Water Equations. *J. Comput. Phys.*, 225(2):1233–1261, 2007.
7. D. A. Anderson, J. C. Tannehill, and R. H. Pletcher. *Computational Fluid Mechanics and Heat Transfer*. Hemisphere Publishing Corporation, 1984.
8. D. S. Anderson and J. J. Gottlieb. On Random Numbers for the Random Choice Method. Technical Report 258, UTIAS, University of Toronto, Canada, 1987.
9. J. D. Anderson. *Hypersonic and High Temperature Gas Dynamics*. Mc Graw–Hill, 1989.
10. J. D. Anderson. *Modern Compressible Flow*. Mc Graw–Hill, 1990.
11. J. L. Anderson, S. Preiser, and E. L. Rubin. Conservation Form of the Equations of Hydrodynamics in Curvilinear Coordinate Systems. *J. Comput. Phys.*, 2:279–287, 1968.
12. W. Anderson, J. L. Thomas, and B. van Leer. Comparison of Finite Volume Flux Vector Splittings for the Euler Equations. *AIAA Journal*, 24:1453–1460, 1986.

13. W. K. Anderson, J. L. Thomas, and L. Rumsey. Extension and Application of Flux–Vector Splitting to Calculations on Dynamic Meshes. *AIAA Journal*, 27:673–674, 1989.
14. N. Andrianov and G. Warnecke. The Riemann Problem for the Baer–Nunziato Two–Phase Flow Model. *J. Comput. Phys.*, 212:434–464, 2004.
15. T. M. Apostol. *Mathematical Analysis*. Addison–Wesley Publishing Company, 1974.
16. P. Arminjon and A. St-Cyr. Nessyahu–Tadmor Type Central Finite Volume Methods Without Predictor for 3D Cartesian and Unstructured Tetrahedral Grids. *Applied Numerical Mathematics*, 46:135–155, 2003.
17. P. Arminjon, A. St-Cyr, and A. Madrane. New Two– and Three Dimensional Non-Oscillatory Central Finite Volume Methods on Staggered Cartesian Grids. *Applied Numerical Mathematics*, 40:367–390, 2002.
18. P. Arminjon and R. Touma. Central Finite Volume Methods with Constrained Transport Divergence Treatment for Ideal MHD. *J. Comput. Phys.*, 204:737–759, 2005.
19. M. Arora and P. L. Roe. On Postshock Oscillations Due to Shock Capturing Schemes in Unsteady Flows. *J. Comput. Phys.*, 130:25–40, 1997.
20. E. H. Atta and J. Vadyak. A Grid Overlapping Scheme for Flowfield Computations About Multicomponent Configurations. *AIAA Journal*, 21(9):1271–1277, 1983.
21. M. R. Baer and J. W. Nunziato. A Two–Phase Mixture Theory for the Deflagration–to–Detonation Transition (DDT) in Reactive Granular Materials. *J. Multiphase Flow*, 12:861–889, 1986.
22. M. J. Baines and M. E. Hubbard. Multidimensional Upwinding with Grid Adaptation. In *Numerical Methods for Wave Propagation*. Toro, E. F. and Clarke, J. F. (Editors), pages 33–54. Kluwer Academic Publishers, 1998.
23. M. J. Baines and P. K. Sweby. A Flux Limiter Based on Fromm’s Scheme. Technical Report 9/84, Department of Mathematics, University of Reading, UK, 1984.
24. G. J. Ball and R. A. East. Shock and Blast Attenuation by Aqueous Foam Barriers: Influences of Barrier Geometry. *J. Shock Waves*, 9(1):37–47, 1999.
25. D. Balsara and C. W. Shu. Monotonicity Preserving Weighted Essentially Nonoscillatory Schemes with Increasingly High Order of Accuracy. *J. Comput. Phys.*, 160:405–452, 2000.
26. D. Balsara and D. Spicer. A Staggered Mesh Algorithm Using High Order Godunov Fluxes to Ensure Solenoidal Magnetic Fields in Magnetohydrodynamic Simulations. *J. Comput. Phys.*, 149:270–292, 1999.
27. D. Balsara and D. Spicer. Maintaining Pressure Positivity in Magnetohydrodynamic Simulations. *J. Comput. Phys.*, 149:133–148, 1999.
28. D. S. Balsara, T. Rumpf, M. Dubser, and C. D. Munz. Efficient, High Accuracy ADER–WENO Schemes for Hydrodynamics and Divergence–Free Magnetohydrodynamics. *J. Comput. Phys.*, 2008 (Submitted).
29. A. A. Barmin, A. G. Kulikovskii, and N. V. Pogorelov. Shock–Capturing Approach and Nonevolutionary Solutions in Magnetohydrodynamics. *J. Comput. Phys.*, 126:77–90, 1996.
30. F. Bassi, A. Crivellini, D. A. Di Pietro, and S. S. Rebay. An Artificial Compressibility Flux for the Discontinuous Galerkin Solution of the Incompressible Navier–Stokes Equations. *J. Comput. Phys.*, 218:794–815, 2006.

31. F. Bassi and S. S. Rebay. High-Order Accurate Discontinuous Finite Element Method for the Numerical Solution of the Compressible Navier-Stokes Equations. *J. Comput. Phys.*, 131:267–279, 1997.
32. P. Batten, N. Clarke, C. Lambert, and D. Causon. On the Choice of Wave Speeds for the HLLC Riemann Solver. *SIAM J. Sci. and Stat. Comp.*, 18:1553–1570, 1997.
33. P. Batten, M. A. Leschziner, and U. C. Goldberg. Average-State Jacobians and Implicit Methods for Compressible Viscous and Turbulent Flows Flows. *J. Comput. Phys.*, 137:38–78, 1997.
34. A. Bayliss and E. Turkel. Radiation-Boundary Conditions for Wave-Like Equations. *Comm. Pure and Appl. Maths.*, 33:708–725, 1980.
35. E. Becker. *Gas Dynamics*. Academic Press, 1968.
36. M. Ben-Artzi. Application of the Generalised Riemann Problem Method to 1-D Compressible Flows with Interfaces. *J. Comput. Phys.*, 65:170–178, 1986.
37. M. Ben-Artzi and J. Falcovitz. A Second Order Godunov-Type Scheme for Compressible Fluid Dynamics. *J. Comput. Phys.*, 55:1–32, 1984.
38. M. Ben-Artzi and J. Falcovitz. A High Resolution Upwind Scheme for Quasi 1-D Flows. In *Numerical Methods for the Euler Equations of Fluid Dynamics*, pages 66–83. INRIA, SIAM, 1985.
39. M. Ben-Artzi and J. Falcovitz. *Generalized Riemann Problems in Computational Fluid Dynamics*. Cambridge University Press, 2003.
40. M. Ben-Artzi and J. Li. Hyperbolic Balance Laws: Riemann Invariants and Hyperbolic Balance Laws. *Numerische Mathematik*, 106:69–425, 2007.
41. M. Ben-Artzi, J. Li, and G. Warnecke. A Direct Eulerian GRP Scheme for Compressible Fluid Flows. *J. Comput. Phys.*, 218(19–43), 2006.
42. G. Ben-Dor. *Shock Wave Reflection Phenomena*. Springer-Verlag, 1992.
43. J. A. Benek, J. L. Steger, and F. C. Dougherty. A Flexible Grid Embedding Technique with Application to the Euler Equations. AIAA Paper 83–1944, 1983.
44. F. Benkhaldoun and A. Chalabi. Characteristic Based Scheme for Conservation Laws with Source Terms. Technical Report 95.15, Centre National de la Recherche Scientifique, Université Paul Sabatier, UFR-MIG, 118 route de Narbone, 31062, Toulouse Cedex, France, 1995.
45. M. J. Berger. *Adaptive Mesh Refinement for Hyperbolic Partial Differential Equations*. PhD thesis, Dept. of Computer Science, Stanford University, California, 1982.
46. M. J. Berger and P. Colella. Local Adaptive Mesh Refinement for Shock Hydrodynamics. *J. Comput. Phys.*, 82:64–84, 1989.
47. M. J. Berger and R. J. LeVeque. Stable Boundary Conditions for Cartesian Grid Calculations. Technical Report ICASE 90–37, NASA Langley Research Center, USA, 1990.
48. M. J. Berger and J. Oliger. Adaptive Mesh Refinement for Hyperbolic Partial Differential Equations. *J. Comput. Phys.*, 53:484–512, 1984.
49. A. C. Berkenbosch, E. F. Kaasschieter, and J. H. M. ten Thije Boonkkamp. The Numerical Computation of One-dimensional Detonation Waves. Technical Report RANA 94–21, Eindhoven University of Technology, Department of Mathematics and Computer Science, The Netherlands, 1994.
50. L. Bermúdez, A. Dervieux, J. A. Desideri, and M. E. Vázquez. Upwind Schemes for the Two-Dimensional Shallow Water Equations with Variable Depth Using

- Unstructured Meshes. Technical Report 2738, INRIA, Rocquencourt, 78153 Le Chesnay, France, 1995.
51. L. Bermúdez and M. E. Vázquez. Upwind Methods for Hyperbolic Conservation Laws with Source Terms. *Computers and Fluids*, 23:1049–1071, 1994.
  52. H. Bershader and R. Hanson (Editors). *Shock Waves and Shock Tubes (Proc. 15th Int. Symp. on Shock Waves and Shock Tubes, Berkeley, California, USA 1985)*. Stanford University Press, 1986.
  53. C. Berthon, P. Charrier, and B. Dubroca. An HLLC Scheme to Solve The M1 Model of Radiative Transfer in Two Space Dimensions. *Journal of Scientific Computing*, 31(3):347–389, 2007.
  54. C. Berthon, F. Coquel, J. M. Hérard, and M. Uhlmann. An Approximate Solution of the Riemann Problem for a Realisable Second-Moment Turbulent Closure. *J. Shock Waves*, 11(2):245–269, 2002.
  55. M. Berzins, P. M. Dew, and R. M. Fuzerland. Developing Software for Time-Dependent Problems Using the Method of Lines and Differential Algebraic Integrators. *Applied Numerical Mathematics*, 5:375–397, 1989.
  56. M. Berzins and J. M. Ware. Solving Convection and Convection Reaction Problems using M.O.L. *Applied Numerical Mathematics*, 20:83–89, 1996.
  57. F. Bianco, G. Puppo, and G. Russo. High Order Central Schemes for Hyperbolic Systems of Conservation Laws. *SIAM J. Scientific Computing*, 21:294–322, 1999.
  58. S. J. Billett. Numerical Aspects of Convection Diffusion Problems. MSc. Thesis, College of Aeronautics, Cranfield Institute of Technology, UK, 1991.
  59. S. J. Billett. A Class of Upwind Methods for Conservation Laws. Ph.D thesis, College of Aeronautics, Cranfield University, UK, 1994.
  60. S. J. Billett and E. F. Toro. Restoring Monotonicity of Slowly-Moving Shocks Computed with Godunov-Type Schemes. *CoA Report 9207, Cranfield Institute of Technology, UK*, 1992.
  61. S. J. Billett and E. F. Toro. Implementing a Three-Dimensional Finite Volume WAF-Type Scheme for the Euler Equations. In *Proc. Third ECCOMAS Computational Fluid Dynamics Conference, 9–13 September 1996, Paris, France*, pages 732–738. Wiley, 1996.
  62. S. J. Billett and E. F. Toro. On the Accuracy and Stability of Explicit Schemes for Multidimensional Linear Homogeneous Advection Equations. *J. Comp. Phys.*, 131:247–250, 1997.
  63. S. J. Billett and E. F. Toro. WAF-Type Schemes for Multidimensional Hyperbolic Conservation Laws. *J. Comp. Phys.*, 130:1–24, 1997.
  64. S. J. Billett and E. F. Toro. Unsplit WAF-Type Schemes for Three Dimensional Hyperbolic Conservation Laws. In *Numerical Methods for Wave Propagation. Toro, E. F. and Clarke, J. F. (Editors)*, pages 75–124. Kluwer Academic Publishers, 1998.
  65. S. J. Billett and E. F. Toro. TVB Limiter Functions for Multidimensional Schemes. In *Proc. 6th Intern. Symposium on Computational Fluid Dynamics, Lake Tahoe, Nevada, USA*, pages 117–122, September 4–8, 1995.
  66. E. P. Boden. *An Adaptive Gridding Technique for Conservation Laws on Complex Domains*. PhD thesis, Department of Aerospace Science, Cranfield University, UK, 1997.
  67. E. P. Boden and E. F. Toro. A Combined Chimera–AMR Technique for Hyperbolic PDEs. In *Proc. Fifth Annual Conference of the CFD Society of Canada*, pages 5.13–5.18. University of Victoria, Canada, 1997.

68. G. Bodo, S. Masaglia, A. Mignone, and P. Rossi (Editors). *Jets from Young Stars III. Numerical MHD and Instabilities. Lecture Notes in Physics 754*. Springer, 2008.
69. D. L. Book, J. P. Boris, and K. Hain. Flux Corrected Transport II: Generalizations of the Method. *J. Comput. Phys.*, 18:248–283, 1975.
70. J. P. Boris and D. L. Book. Flux-Corrected Transport. I. SHASTA, A Fluid Transport Algorithm that Works. *J. Comput. Phys.*, 11:38–69, 1973.
71. J. P. Boris and D. L. Book. Flux-Corrected Transport III: Minimal-Error FCT Algorithms. *J. Comput. Phys.*, 20:397–431, 1976.
72. J. P. Boris and E. S. Oran. *Numerical Simulation of Reactive Flow*. Elsevier, Amsterdam–New York, 1987.
73. A. G. L. Borthwick, M. Fujihara, and B. D. Rogers. Godunov Solutions of Shallow Water Equations on Curvilinear and Quadtree Grids. In *Godunov Methods: Theory and Applications. Edited Review, E. F. Toro (Editor)*, pages 141–148. Kluwer Academic/Plenum Publishers, 2001.
74. F. Bouchut. *Nonlinear Stability of Finite Volume Methods for Hyperbolic Conservation laws and Well-Balanced Schemes*. Birkhäuser, 2004.
75. F. Bouchut, C. Klingenberg, and K. Waagan. A Multiwave Approximate Riemann Solver for Ideal MHD Based on Relaxation I - Theoretical Framework. *Numer. Math.*, 108(1):7–42, 2007.
76. A. Bourgeade, P. Le Floch, and Raviart P. A. An Asymptotic Expansion for the Solution of the Generalized Riemann Problem. Part 2: Application to the Euler Equations of Gas Dynamics. *Ann. Inst. Henri Poincaré. Analyse non Linéaire*, 6(6):437–480, 1989.
77. A. Bourlioux and A. J. Majda. Theoretical and Numerical Structure for Unstable Two-Dimensional Detonations. *Combustion and Flame*, 90:211–229, 1992.
78. A. Bourlioux, A. J. Majda, and V. Roytburd. Theoretical and Numerical Structure for Unstable One-Dimensional Detonations. *SIAM J. Appl. Math.*, 51:303–343, 1991.
79. C. E. Brennen. *Fundamentals of Multiphase Flow*. Cambridge University Press, 2005.
80. M. Brio and C. C. Wu. An Upwind Differencing Scheme for the Equations of Ideal Magnetohydrodynamics. *J. Comput. Phys.*, 75:400–422, 1988.
81. M. Brown. *Differential Equations and their Applications*. Springer–Verlag, 1975.
82. R. E. Brown. Numerical Solution of the 2D Unsteady Navier–Stokes Equations Using Viscous–Convective Operator Splitting. MSc. Thesis, Department of Aerospace Science, Cranfield University, UK, 1990.
83. R. Brun and L. Z. Dumitrescu (Editors). *Shock Waves @ Marseille (Proc. 19th Int. Symp. on Shock Waves, Marseille, France, 1993), Four Volumes*. Springer–Verlag, 1995.
84. M. Cáceres. Development of a Third-Order Accurate Scheme of the MUSCL Type for the Time-Dependent One-Dimensional Euler Equations. MSc. Thesis, Department of Aerospace Science, Cranfield University, UK, 1993.
85. A. Canestrelli, A. Saviglia, M. Dumbser, and E. F. Toro. Well-Balanced High-Order Centred Schemes for Non-Conservative Hyperbolic Systems. Application to Shallow Water with Fixed and Mobile Bed. (*Submitted*), 2008.
86. G. Capdeville. Towards a Compact High-Order Method for Non-linear Hyperbolic Systems, II. The Hermite–HLLC Scheme. *J. Comput. Phys.*, 227(22):9428–9462, 2008.

87. Carasso, Raviart, and Serre. *Proc. First International Conference on Hyperbolic Problems*. Springer, 1986.
88. J. Casper and H. L. Atkins. A Finite Volume High-Order ENO Scheme for Two-Dimensional Hyperbolic Systems. *J. Comput. Phys.*, 106:62–76, 1993.
89. C. E. Castro. *High-Order ADER FV/DG Numerical Methods for Hyperbolic Equations*. PhD thesis, Department of Civil and Environmental Engineering, University of Trento, Italy, 2007.
90. C. E. Castro and E. F. Toro. A Riemann Solver and Upwind Methods for a Two-Phase Flow Model in Non-Conservative Form. *Int. J. for Numer. Meth. in Fluids*, 60:275–307, 2006.
91. C. E. Castro and E. F. Toro. Solvers for the High-Order Riemann Problem for Hyperbolic Balance Laws. *J. Comput. Phys.*, 227:2481–2513, 2008.
92. M. J. Castro, J. M. Gallardo, J. A. López, and C. Parés. Well-Balanced High Order Extensions of Godunov’s Method for Semilinear Balance Laws. *SIAM Journal of Numerical Analysis*, 46:1012–1039, 2008.
93. M. J. Castro, P. G. LeFloch, M. L. Muñoz-Ruiz, and C. Parés. Why Many Theories of Shock Waves are Necessary: Convergence Error in Formally Path-Consistent Schemes. *J. Comput. Phys.*, 227:8107–8129, 2008.
94. M. J. Castro, A. Pardo, C. Parés, and E. F. Toro. On Some Fast Well-Balanced First Order Solvers for Nonconservative Systems. *Mathematics of Computation*, 2008 (submitted).
95. V. Casulli and E. F. Toro. Preliminary Results on a Semi-Lagrangian Version of the WAF Method. Technical Report 9018, College of Aeronautics, Cranfield Institute of Technology, UK, 1990.
96. A. Chakraborty and E. F. Toro. Development of an Approximate Riemann Solver for the Steady Supersonic Euler Equations. *The Aeronautical Journal*, 98:325–339, 1994.
97. S. R. Chakravarthy, D. A. Anderson, and M. D. Salas. The Split-Coefficient Matrix Method for Hyperbolic Systems of Gas Dynamics Equations. *AIAA Paper 80-0268, AIAA 19th Aerospace Science Meeting*, 1980.
98. A. Chalabi. Stable Upwind Schemes for Conservation Laws with Source Term. *IMA J. Numer. Anal.*, 12:217–241, 1992.
99. A. Chalabi. On Convergence of Numerical Schemes for Hyperbolic Conservation Laws with Stiff Source Terms. Technical Report 95.22, Centre National de la Recherche Scientifique, Université Paul Sabatier, UFR-MIG, 118 route de Narbone, 31062, Toulouse Cedex, France, 1995.
100. A. Chalabi. An Error Bound for the Polygonal Approximation of Hyperbolic Conservation Laws with Source Terms. *Computers and Mathematics with Applications*, 32:59–63, 1996.
101. A. Chalabi. On Convergence of Numerical Schemes for Hyperbolic Conservation Laws with Stiff Source Terms. *Math. of Comput.*, 66:527–545, 1997.
102. C. H. Chang and M. S. Liou. A Robust and Accurate Approach to Computing Compressible Multiphase Flow: Stratified Flow Model and AUSM+ up Scheme. *J. Comput. Phys.*, 225:840–873, 2007.
103. G. Q. Chen and E. F. Toro. Centred Schemes for Non-Linear Hyperbolic Equations. *Journal of Hyperbolic Differential Equations*, 1(1):531–566, 2004.
104. D. N. Cheney. Upwinding Convective and Viscous Terms via a Modified GRP Approach. MSc. Thesis, Department of Aerospace Science, Cranfield University, UK, 1994.

105. G. Chesshire and W. D. Henshaw. Composite Overlapping Meshes for the Solution of Partial Differential Equations. *J. Comput. Phys.*, 90:1–64, 1990.
106. G. Chesshire and W. D. Henshaw. A Scheme for Conservative Interpolation on Overlapping Grids. *SIAM J. Sci. Stat.*, 15(4):819–845, 1994.
107. B. L. Chessor, Jr. Evaluation of Computed Steady and Unsteady Quasi–One–Dimensional Viscous/Inviscid Interacting Internal Flow Fields Through Comparison with Two–Dimensional Navier–Stokes Solutions. MSc. Thesis, 1992, Mississippi State University, USA, 1992.
108. T. Chiang and Hoffmann K. Determination of computational time step for chemically reacting flows. AIAA 20th Fluid Dynamics, Plasma Dynamics and Lasers Conference, Buffalo, New York, June 12–14, USA, 1989, Paper AIAA 89–1855, 1989.
109. A. J. Chorin. A Numerical Method for Solving Viscous Flow Problems. *J. Comput. Phys.*, 2:12–26, 1967.
110. A. J. Chorin. Random Choice Solutions of Hyperbolic Systems. *J. Comput. Phys.*, 22:517–533, 1976.
111. A. J. Chorin. Random Choice Methods with Applications to Reacting Gas Flow. *J. Comput. Phys.*, 25:253–272, 1977.
112. A. J. Chorin and J. E. Marsden. *A Mathematical Introduction to Fluid Mechanics*. Springer–Verlag, 1993.
113. C. S. Chu and C. W. Shu. High Order Residual Distribution Conservative Finite Difference WENO Schemes for Convection–Diffusion Steady State Problems on Non–Smooth Meshes. *J. Comput. Phys.*, 224:992–1020, 2007.
114. D. K. Clarke, M. D. Salas, and H. A. Hassan. Euler Calculations for Multielement Airfoils Using Cartesian Grids. *AIAA J.*, 24:353–, 1987.
115. J. F. Clarke. Compressible Flow Produced by Distributed Sources of Mass: An Exact Solution. Technical Report CoA Report 8710, College of Aeronautics, Cranfield Institute of Technology, Cranfield, UK, 1987.
116. J. F. Clarke. Chemical Reactions in High Speed Flows. *Phil. Trans. Roy. Soc. London, A* 335:161–199, 1989.
117. J. F. Clarke. Fast Flames, Waves and Detonations. *Prog. Energy Comb. Sci.*, 15:241–271, 1989.
118. J. F. Clarke, S. Karni, J. J. Quirk, L. G. Simmons, P. L. Roe, and E. F. Toro. Numerical Computation of Two–Dimensional, Unsteady Detonation Waves in High Energy Solids. *J. Comput. Phys.*, 106:215–233, 1993.
119. J. F. Clarke and M. McChesney. *Dynamics of Relaxing Gases*. Butterworths, 1976.
120. J. F. Clarke and E. F. Toro. Gas Flows Generated by Solid–Propellant Burning. In R. Glowinski, B. Larouturou, and R. Teman, editors, *Proc. Intern. Confer. on Numerical Simulation of Combustion Phenomena*, number 241 in Lecture Notes in Physics, pages 192–205, INRIA, Sophia Antipolis, France, May 21–24 1985.
121. B. Cockburn, S. Hou, and C. W. Shu. The Runge–Kutta Local Projection Discontinuous Galerkin Finite Element Method for Conservation Laws IV: the Multidimensional Case. *Mathematics of Computation*, 54:545–581, 1990.
122. B. Cockburn, G. E. Karniadakis, and C. W. Shu. *Discontinuous Galerkin Methods*. Lecture Notes in Computational Science and Engineering. Springer, 2000.

123. B. Cockburn, S. Y. Li, and C. W. Shu. TVB Runge–Kutta Local Projection Discontinuous Galerkin Finite Element Method for Conservation Laws III: One Dimensional Systems. *J. Comput. Phys.*, 84:90–113, 1989.
124. B. Cockburn and C. Shu. Nonlinearly Stable Compact Schemes for Shock Calculations. *SIAM J. Numer. Anal.*, 31:607–627, 1994.
125. B. Cockburn and C. W. Shu. TVB Runge–Kutta Local Projection Discontinuous Galerkin Finite Element Method for Conservation Laws II: General Framework. *Mathematics of Computation*, 52:411–435, 1989.
126. B. Cockburn and C. W. Shu. The Runge–Kutta Local Projection P1–Discontinuous Galerkin Finite Element Method for Scalar Conservation Laws. *Mathematical Modelling and Numerical Analysis*, 25:337–361, 1991.
127. B. Cockburn and C. W. Shu. The Local Discontinuous Galerkin Method for Time–Dependent Convection Diffusion Systems. *SIAM Journal on Numerical Analysis*, 35:2440–2463, 1998.
128. B. Cockburn and C. W. Shu. The Runge–Kutta Discontinuous Galerkin Method for Conservation Laws V: Multidimensional Systems. *J. Comput. Phys.*, 141:199–224, 1998.
129. B. Cockburn and C. W. Shu. Runge–Kutta Discontinuous Galerkin Methods for Convection–Dominated Problems. *Journal of Scientific Computing*, 16:173–261, 2001.
130. E. Coddington and N. Levinson. *Theory of Ordinary Differential Equations*. McGraw–Hill, New York, 1955.
131. P. Colella. *An Analysis of the Effect of Operator Splitting and of the Sampling Procedure on the Accuracy of Glimm's Method*. PhD thesis, Department of Mathematics, University of California, USA, 1978.
132. P. Colella. Glimm's Method for Gas Dynamics. *SIAM J. Sci. Stat. Comput.*, 3(1):76–110, 1982.
133. P. Colella. A Direct Eulerian MUSCL Scheme for Gas Dynamics. *SIAM J. Sci. Stat. Comput.*, 6:104–117, 1985.
134. P. Colella. Multidimensional Upwind Methods for Hyperbolic Conservation Laws. *J. Comput. Phys.*, 87:171–200, 1990.
135. P. Colella and H. H. Glaz. Efficient Solution Algorithms for the Riemann Problem for Real Gases. *J. Comput. Phys.*, 59:264–289, 1985.
136. P. Colella, A. Majda, and V. Roytburd. Theoretical and Numerical Structure for Reacting Shock Waves. *SIAM J. Sci. Stat. Comput.*, 7:1059–1080, 1986.
137. P. Colella and P. R. Woodward. The Piecewise Parabolic Method (PPM) for Gas Dynamical Simulation. *J. Comput. Phys.*, 54:174–201, 1984.
138. P. Concus and W. Proskurowski. Numerical Solution of a Nonlinear Hyperbolic Equation by the Random Choice Method. *J. Comput. Phys.*, 30:153–166, 1979.
139. S. D. Conte and C. de Boor. *Elementary Numerical Analyis*. McGraw–Hill Kogakusha Ltd., 1980.
140. J. M. Corberán and M. Ll. Gascón. TVD Schemes for the Calculation of Flow in Pipes of Variable Cross–Section. *Mathl. Comput. Modelling*, 21:85–92, 1995.
141. J. Corner. *Theory of the Interior Ballistics of Guns*. Wiley, 1950.
142. J. Cortes, A. Debussche, and I. Toumi. A Density Perturbation Method to Study the Eigenstructure of Two–Phase Flow Equation Systems. *J. Comput. Phys.*, 147:1–22, 1998.
143. R. Courant and K. O. Friedrichs. *Supersonic Flow and Shock Waves*. Springer–Verlag, 1985.

144. R. Courant, E. Isaacson, and M. Rees. On the Solution of Nonlinear Hyperbolic Differential Equations by Finite Differences. *Comm. Pure. Appl. Math.*, 5:243–255, 1952.
145. M. G. Crandall and A. Majda. Monotone Difference Approximations for Scalar Conservation Laws. *Mathematics of Computations*, 34:1–34, 1980.
146. W. Dai and P. R. Woodward. An Approximate Riemann Solver for Ideal Magnetohydrodynamics. *J. Comput. Phys.*, 111:354–372, 1994.
147. W. Dai and P. R. Woodward. Extension of the Piecewise Parabolic Method to Multidimensional Ideal Magnetohydrodynamics. *J. Comput. Phys.*, 115:485–, 1994.
148. G. Dal Maso, P. G. LeFloch, and F. Murat. Definition and Weak Stability of Nonconservative Products. *J. Math. Pure Appl.*, 74:483–548, 1995.
149. S. F. Davis. TVD Finite Difference Schemes and Artificial Viscosity. Technical Report ICASE 84–20, NASA Langley Research Center, Hampton, USA, 1984.
150. S. F. Davis. Simplified Second-Order Godunov-Type Methods. *SIAM J. Sci. Stat. Comput.*, 9:445–473, 1988.
151. A. S. Dawes. *Natural Coordinates and High Speed Flows. A Numerical Method for Reactive Gases*. PhD thesis, College of Aeronautics, Cranfield Institute of Technology, UK, 1992.
152. H. Deconinck, K. G. Powell, P. L. Roe, and R. Struijs. Multi-dimensional Schemes for Scalar Advection. *AIAA*, 91(1532), 1991.
153. H. Deconinck, P. L. Roe, and R. Struijs. Multi-dimensional Generalisation of Roe’s Flux Difference Splitter for the Euler Equations. *Computers and Fluids*, 22:215–222, 1993.
154. A. Dedner, F. Kemm, D. Kröner, C. D. Munz, T. Schnitzer, and M. Wesenberg. Hyperbolic Divergence Cleaning for the MHD Equations. *J. Comput. Phys.*, 175:645–673, 2002.
155. L. Del Zanna, N. Bucciantini, and P. Londrillo. An Efficient Shock-Capturing Central Scheme for Multi-dimensional Relativistic Flows II. Magnetohydrodynamics. *Astronomy and Astrophysics*, 400:397–413, 2003.
156. V. Deledicque and M. V. Papalexandris. An Exact Riemann Solver for Compressible Two-Phase Flow Models Containing Nonconservative Products. *J. Comput. Phys.*, 222:217–245, 2007.
157. T. DeNeff and G. Moretti. Shock Fitting for Everyone. *Computers and Fluids*, 8:327–334, 1980.
158. J. M. Dewey. Mach Reflection Research—Paradox and Progress. In *Shock Waves (Proc. 18th Int. Symp. on Shock Waves, Sendai, Japan, 1991. Takayama, Editor)*, Vol. I, pages 113–120. Springer–Verlag, 1992.
159. J. Dobes and H. Deconinck. Second Order Blended Multidimensional Upwind Residual Distribution Scheme for Steady and Unsteady Computations. *Journal of Computational and Applied Mathematics*, 215:378–389, 2008.
160. D. A. Drew and R. Gulliver. *Particulate Flows: Processing and Rheology*. Springer, 1997.
161. D. A. Drew, J. E. Marsden, and L. Sirovich. *Theory of Multicomponent Fluids*. Springer, 1998.
162. D. Drikakis. A Characteristic-Based Method for Incompressible Flow. *Int. J. Numer. Meth. Fluids*, 19:667–685, 1994.
163. D. Drikakis. A Parallel Multiblock Characteristic-Based Method for Three-Dimensional Incompressible Flows. *Advances in Engineering Software*, 26:111–119, 1996.

164. D. Drikakis, O. P. Iliev, and D. P. Vassileva. A Nonlinear Multigrid Method for the Three-Dimensional Incompressible Navier–Stokes Equations. *J. Comput. Phys.*, 146:301–321, 1998.
165. D. Drikakis and W. Rider. *High-Resolution Methods for Incompressible and Low-Speed Flows*. Springer, 2005.
166. D. Drikakis and S. Tsangaris. On the Solution of the Compressible Navier–Stokes Equations Using Improved Flux Vector Splitting Methods. *Appl. Math. Modelling*, 17:282–297, 1993.
167. F. Dubois and G. Mehlman. A Non-Parameterized Entropy Fix for Roe’s Method. *AIAA Journal*, 31 (1):199–200, 1993.
168. J. K. Dukowicz. A General, Non-Iterative Riemann Solver for Godunov’s Method. *J. Comput. Phys.*, 61:119–137, 1985.
169. M. Dumbser. *Arbitrary High Order Schemes for the Solution of Hyperbolic Conservation Laws in Complex Domains*. PhD thesis, PhD thesis, Institut für Aero- un Gasdynamik, Universität Stuttgart, Germany, 2005.
170. M. Dumbser. Building Blocks for Arbitrary High Order Discontinuous Galerkin Schemes. *Journal of Scientific Computing*, 27:215–230, 2006.
171. M. Dumbser, D. Balsara, E. F. Toro, and C. D. Munz. A Unified Framework for the Construction of One-Step Finite-Volume and Discontinuous Galerkin Schemes. *J. Comput. Phys.*, 227:8209–8253, 2008.
172. M. Dumbser, M. J. Castro, C. Parés, and E. F. Toro. ADER Schemes on Unstructured Meshes for Nonconservative Hyperbolic Systems: Applications to Geophysical Flows. *Computers and Fluids*, 2008 (submitted).
173. M. Dumbser, C. Enaux, and E. F. Toro. Finite Volume Schemes of Very High Order of Accuracy for Stiff Hyperbolic Balance Laws. *J. Comput. Phys.*, 227(8):3971–4001, 2008.
174. M. Dumbser, A. Hidalgo, M. J. Castro, C. Parés, and E. F. Toro. FORCE Schemes on Unstructured Meshes II: Nonconservative Hyperbolic Systems. *Isaac Newton Institute for Mathematical Sciences, University of Cambridge, UK. Preprint NI09053-NPA, December 2008. Submitted to J. Comput. Phys.*, 2008.
175. M. Dumbser and M. Käser. Arbitrary High Order Non-Oscillatory Finite Volume Schemes on Unstructured Meshes for Linear Hyperbolic Systems. *J. Comput. Phys.*, 221(2):693–723, 2007.
176. M. Dumbser, M. Käser, V. A. Titarev, and E. F. Toro. Quadrature-Free Non-Oscillatory Finite Volume Schemes on Unstructured Meshes for Nonlinear Hyperbolic Systems. *J. Comput. Phys.*, 226(8):204–243, 2007.
177. M. Dumbser, M. Käser, and E. F. Toro. An Arbitrary High Order Discontinuous Galerkin Method for Elastic Waves on Unstructured Meshes V: Local Time Stepping and  $p$ -Adaptivity. *Geophysical Journal International*, 171:695–717, 2007.
178. M. Dumbser and C. D. Munz. ADER Discontinuous Galerkin Schemes for Aeroacoustics. *Comptes Rendus Mécanique*, 333:683–687, 2005.
179. P. Dutt. A Riemann Solver Based on a Global Existence Proof for the Riemann Problem. Technical Report ICASE 86-3, NASA Langley Research Center, USA, 1986.
180. D. H. Edwards, G. O. Thomas, and T. L. Williams. Initiation of Detonation by Steady Planar Incident Shock Waves. *Combustion and Flame*, 43:187–198, 1981.

181. B. Einfeldt. On Godunov–Type Methods for Gas Dynamics. *SIAM J. Numer. Anal.*, 25(2):294–318, 1988.
182. B. Einfeldt, C. D. Munz, P. L. Roe, and B. Sjögreen. On Godunov–Type Methods near Low Densities. *J. Comput. Phys.*, 92:273–295, 1991.
183. P. Embid and M. Baer. Mathematical Analysis of a Two–Phase Continuum Mixture Theory. *Continuum Mech. Thermodyn.*, 4:279–312, 1992.
184. B. Engquist and B. Gustafsson. *Proc. Third International Conference on Hyperbolic Problems, Vol. I and II*. Chartwell–Bratt, 1991.
185. B. Engquist and S. Osher. One Sided Difference Approximations for Nonlinear Conservation Laws. *Math. Comp.*, 36(154):321–351, 1981.
186. L. C. Evans. *Partial Differential Equations*. American Mathematical Society, 2002.
187. S. A. E. G. Falle and S. S. Komissarov. An Upwind Numerical Scheme for Relativistic Hydrodynamics with a General Equation of State. *Monthly Notices of the Royal Astronomical Society*, 278:586–602, 1996.
188. S. A. E. G. Falle and S. S. Komissarov. A Multidimensional Upwind Scheme for Magnetohydrodynamics. *Monthly Notices of the Royal Astronomical Society*, 297:265–277, 1997.
189. M. Feistauer, J. Felcman, and I. Straskraba. *Mathematical and Computational Methods for Compressible Flow*. Oxford University Press, 2003.
190. M. Fey, R. Jeltsch, and S Müller. The Influence of a Source Term, an Example: Chemically Reacting Hypersonic Flow. Technical Report 92–06, SAM, Eidgenössische Technische Hochschule (ETH), Zürich, Switzerland, 1992.
191. W. Fickett and W. Davis. *Detonation*. University of California Press, 1979.
192. C. A. J. Fletcher. *Computational Techniques for Fluid Dynamics, Vols. I and II*. Springer–Verlag, 1988.
193. L. Fraccarollo and E. F. Toro. A Shock-Capturing Method for Two Dimensional Dam–Break Problems. *Proceedings of the Fifth International Symposium in Computational Fluid Dynamics, Sendai, Japan*, 1993.
194. L. Fraccarollo and E. F. Toro. Experimental and Numerical Assessment of the Shallow Water Model for Two–Dimensional Dam–Break Type problems. *Journal of Hydraulic Research*, 33:843–864, 1995.
195. J. E. Fromm. A Method for Reducing Dispersion in Convective Difference Schemes. *J. Comput. Phys.*, 3:176–189, 1968.
196. M. Gallardo, C. Parés, and M. J. Castro. On a Well–Balanced High–Order Finite Volume Scheme for Shallow Water Equations with Topography and Dry Areas. *J. Comput. Phys.*, 227:574–601, 2007.
197. P. García-Navarro, M. E. Hubbard, and A. Priestley. Genuinely Multidimensional Upwinding for the 2D Shallow Water Equations. *J. Comput. Phys.*, 121:79–93, 1995.
198. P. H. Gaskell and A. K. C. Lau. Curvature Compensated Convective Transport: SMART, a New Boundedness Preserving Transport Algorithm. *Int. J. Numer. Meth. Fluids*, 8:617–641, 1988.
199. S. L. Gavrilyuk, N. Favrie, and R. Saurel. Modelling Wave Dynamics of Compressible Elastic Materials. *J. Comput. Phys.*, 227(5):2941–2969, 2007.
200. C. W. Gear. *Numerical Initial–Value Problems in Ordinary Differential Equations*. Prentice–Hall, Englewood Cliffs, NJ, 1971.
201. B. E. Gelfand, S. M. Frolov, and M. A. Nettleton. Gaseous Detonations–A Selective Review. *Prog. Energy Comb. Sci.*, 17:327–371, 1991.

202. B. Giacomazzo and L. Rezzolla. The Exact Solution of the Riemann Problem in Relativistic Magnetohydrodynamics. *Journal of Fluid Mechanics*, 562:223–259, 2006.
203. D. Gidaspow. *Multiphase Flow and Fluidization*. Academic Press, 1994.
204. R. B. Gilbert and R. A. Strehlow. Theory of Detonation Initiation Behind Reflected Shock Waves. *AIAA Journal*, 4:1777–1783, 1966.
205. M. B. Giles. Non-Reflecting Boundary Conditions for Euler Equation Calculations. *AIAA J.*, 28:2050–2058, 1990.
206. L. Giraud and G. Manzini. Parallel Implementations of 2D Explicit Euler Solvers. *J. Comput. Phys.*, 123:111–118, 1996.
207. P. Glaister. Difference Schemes for the Shallow Water Equations. Technical Report 9/87, Department of Mathematics, University of Reading, England, 1987.
208. P. Glaister. An Approximate Linearised Riemann Solver for the Euler Equations of Gas Dynamics. *J. Comput. Phys.*, 74:382–408, 1988.
209. P. Glaister. An Approximate Linearised Riemann Solver for the Three-Dimensional Euler Equations for Real Gases Using Operator Splitting. *J. Comput. Phys.*, 77:361–383, 1988.
210. H. M. Glaz, P. Colella, I. I. Glass, and R. L. Deschambault. A Detailed Numerical, Graphical and Experimental Study of Oblique Shock Wave Reflections. Technical Report 285, Institute for Aerospace Science, University of Toronto (UTIAS), 1985.
211. H. M. Glaz, P. Colella, I. I. Glass, and R. L. Deschambault. A Numerical Study of Oblique Shock-Wave Reflections with Experimental Comparisons. *Proc. Roy. Soc. London, A* 398:117–140, 1985.
212. J. Glimm. Solution in the Large for Nonlinear Hyperbolic Systems of Equations. *Comm. Pure. Appl. Math.*, 18:697–715, 1965.
213. J. Glimm, M. J. Graham, J. W. Grove, and B. J. Plohr. *Proc. Fifth International Conference on Hyperbolic Problems*. World Scientific, 1996.
214. J. Glimm, G. Marshall, and B. Plohr. A Generalized Riemann Problem for Quasi-One-Dimensional Gas Flows. *Advances in Applied Mathematics*, 5:1–30, 1984.
215. E. Godlewski and P. A. Raviart. *Numerical Approximation of Hyperbolic Systems of Conservation Laws*. Applied Mathematical Sciences, Vol. 118, Springer-Verlag, New York, 1996.
216. S. K. Godunov. A Finite Difference Method for the Computation of Discontinuous Solutions of the Equations of Fluid Dynamics. *Mat. Sb.*, 47:357–393, 1959.
217. S. K. Godunov, A. V. Zabrodin, and G. P. Prokopov. —. *J. Comp. Math. Phys. USSR 1 (1962)*, pages 1187–, 1962.
218. S. K. Godunov (Ed.). *Numerical Solution of Multi-Dimensional Problems in Gas Dynamics*. Nauka Press, Moscow, 1976.
219. J. B. Goodman and R. J. LeVeque. On the Accuracy of Stable Schemes for 2D Scalar Conservation Laws. *Math. Comp.*, 45(181):15–21, 1985.
220. L. Gosse. A Well-Balanced Flux-Vector Splitting Scheme Designed for Hyperbolic Systems of Conservation Laws with Source Terms. *Computers and Mathematics with Applications*, 39:135–159, 2000.
221. J. J. Gottlieb. Staggered and Non-Staggered Grids with Variable Node Spacing and Local Timestepping for Random Choice Method. *J. Comput. Phys.*, 78:160–177, 1988.

222. J. J. Gottlieb and C. P. T. Groth. Assessment of Riemann Solvers for Unsteady One-Dimensional Inviscid Flows of Perfect Gases. *J. Comput. Phys.*, 78:437–458, 1988.
223. S. Gottlieb and C. W. Shu. Total Variation Diminishing Runge–Kutta Schemes. Technical Report ICASE 96–50, NASA Langley Research Center, Hampton, USA, 1996.
224. G. Grassner, F. Lörcher, and C. D. Munz. A Contribution to the Construction of Diffusion Fluxes for Finite Volume and Discontinuous Galerkin Schemes. *J. Comput. Phys.*, 224:1049–1063, 2007.
225. J. M. Greenberg and A. Y. LeRoux. A Well–Balanced Scheme for the Numerical Processing of Source Terms in Hyperbolic Equations. *SIAM J. Numerical Analysis*, 33:1–16, 1996.
226. J. M. Greenberg, A. Y. LeRoux, R. Baraille, and A. Noussair. Analysis and Approximation of Conservation Laws with Source Terms. *SIAM J. Numerical Analysis*, 34:1980–2007, 1997.
227. D. F. Griffiths, A. M. Stuart, and H. C. Yee. Numerical Wave Propagation in an Advection Equation with a Nonlinear Source Term. *SIAM J. Numer. Anal.*, 29:1244–1260, 1992.
228. H. Grönig (Editor). *Shock Tubes and Waves (Proc. 16th Int. Symp. on Shock Tubes and Waves, Aachen Germany 1987)*. VCH Verlagsgessellschaft mbH, 1988.
229. V. Guinot. *Godunov–Type Schemes. An Introduction for Engineers*. Elsevier, 2003.
230. K. F. Gurski. An HLLC–Type Approximate Riemann Solver for Ideal Magnetohydrodynamics. *SIAM J. Sci. Comput.*, 25(6):2165–2187, 2004.
231. B. Haasdonk, D. Kröner, and C. Rhode. Convergence of a Staggered Lax–Friedrichs Scheme on Unstructured Grids in 2D. *Numerische Mathematik*, 88:459–484, 2001.
232. R. Haberman. *Mathematical Models*. SIAM, 1998.
233. J. M. Hammersley and D. C. Handscombe. *Monte Carlo Methods*. Chapman and Hall, 1964.
234. D. Hänel. On the Numerical Solution of the Navier–Stokes Equations for Compressible Flow. VKI Lecture Series 1989–04 on Computational Fluid Dynamics, 1989.
235. D. Hänel and S. D. Sharma. Simulation of Hydrodynamical Free–Surface Flow. In *Proc. Third ECCOMAS Computational Fluid Dynamics Conference, 9–13 September 1996, Paris, France*, pages 1019–1024. Wiley, 1996.
236. E. Harabetian. A Convergent Series Expansion for Hyperbolic Systems of Conservation Laws. *Trans Amer. Math. Soc.*, 294:383–424,, 1986.
237. R. Harris, Z. J. Wang, and Y. Liu. Efficient Quadrature–Free High–Order Spectral Volume Method on Unstructured Grids: Theory and 2D. *J. Comput. Phys.*, 227(3):1620–1642, 2008.
238. A. Harten. High Resolution Schemes for Hyperbolic Conservation Laws. *J. Comput. Phys.*, 49:357–393, 1983.
239. A. Harten. On a Class of High Resolution Total Variation Stable Finite Difference Schemes. *SIAM J. Numer. Anal.*, 21(1):1–23, 1984.
240. A. Harten. Preliminary Results on the Extension of ENO Schemes to Two–Dimensional Problems. In *International Conference on Non–Linear Hyperbolic Problems*, pages 23–51. Springer–Verlag, Berlin, 1987.

241. A. Harten. ENO Schemes with Subcell Resolution. *J. Comput. Phys.*, 83:148–184, 1989.
242. A. Harten, B. Engquist, S. Osher, and S. R. Chakravarthy. Uniformly High Order Accuracy Essentially Non-oscillatory Schemes III. *J. Comput. Phys.*, 71:231–303, 1987.
243. A. Harten and J. M. Hyman. Self Adjusting Grid Methods for One-Dimensional Hyperbolic Conservation Laws. *J. Comput. Phys.*, 50:235–269, 1983.
244. A. Harten, P. D. Lax, and B. van Leer. On Upstream Differencing and Godunov-Type Schemes for Hyperbolic Conservation Laws. *SIAM Review*, 25(1):35–61, 1983.
245. A. Harten and S. Osher. Uniformly High-Order Accurate Nonoscillatory Schemes I. *SIAM J. Numer. Anal.*, 24(2):279–309, 1987.
246. W. Heilig and H. Reichenbach. High Speed Interferometric Study of Unstationary Shock Wave Processes. Technical report, Fraunhofer Institut für Kurzzeitdynamik, Ernst-Mach-Institut, Freiburg im Breisgau, Germany, 1982.
247. P. W. Hemker and S. P. Spekreijse. Multiple Grid and Osher's Scheme for the Efficient Solution of the Steady Euler Equations. *Applied Numerical Mathematics*, 2:475–493, 1986.
248. J. S. Hesthaven and T. Warburton. *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis and Applications*. Springer, 2008.
249. F. B. Hildebrand. *Introduction to Numerical Analysis*. Tata McGraw-Hill Publishing Co. Limited, New Delhi, 1974.
250. R. Hillier. Numerical Modelling of Shock Wave Diffraction. In *Shock Waves @ Marseille, Vol. 4 (Brun and Dumitrescu, Editors)*, pages 17–26. Springer-Verlag, 1995.
251. C. Hirsch. *Numerical Computation of Internal and External Flows, Vol. I: Fundamentals of Numerical Discretization*. Wiley, 1988.
252. C. Hirsch. *Numerical Computation of Internal and External Flows, Vol. II: Computational Methods for Inviscid and Viscous Flows*. Wiley, 1990.
253. K. A. Hoffmann. *Computational Fluid Dynamics for Engineers*. Engineering Education Systems, Austin, Texas, USA, 1989.
254. M. Holt. *Numerical Methods in Fluid Dynamics*. Springer-Verlag, 1984.
255. V. Honkkila and P. Janhunen. HLLC Solver for Ideal Relativistic MHD. *J. Comput. Phys.*, 203(1):344–357, 2007.
256. H. Honma and I. I. Glass. Random Choice Solutions for Weak Spherical Shock-Wave Transitions of N-Waves in Air with Vibrational Excitation. Technical Report 253, UTIAS, University of Toronto, Canada, 1983.
257. H. Honma, M. Wada, and K. Inomata. Random Choice Solutions for Two-Dimensional and Axi-Symmetric Supersonic Flows. Technical Report S. P. Number 2, The Institute of Space and Astronautical Science, 1984.
258. L. Hörmander. *Lectures on Nonlinear Hyperbolic Differential Equations, Mathématiques et Applications 26*. Springer-Verlag, 1997.
259. T. Y. Hou and LeFloch P. Why Non-Conservative Schemes Converge to the Wrong Solutions: Error Analysis. *Math. of Comput.*, 62:497–530, 1994.
260. E. L. Ince and I. N. Sneddon. *The Solution of Ordinary Differential Equations*. Longman Scientific and Technical, 1987.
261. M. Ishii and T. Hibiki. *Thermo-fluid Dynamics of Two-Phase Flow*. Springer, 2006.

262. M. J. Ivings, D. M. Causon, and E. F. Toro. On Hybrid High-Resolution Upwind Methods for Multicomponent Flows. *ZAMM*, 77, Issue 9:645–668, 1997.
263. M. J. Ivings, D. M. Causon, and E. F. Toro. On Riemann Solvers for Compressible Liquids. Technical Report 97-4, Department of Mathematics and Physics, Manchester Metropolitan University, UK, 1997.
264. M. J. Ivings, D. M. Causon, and E. F. Toro. On Riemann Solvers for Compressible Liquids. *Int. J. Numer. Meth. Fluids*, 28:395–418, 1998.
265. R. Jackson. *The Dynamics of Fluidized Particles*. Cambridge University Press, 2000.
266. G. Jagadeesh, E. Arunan, and K. P. J Reddy (Editors). *Shock Waves: Proceedings of the 25th International Symposium on Shock Waves-ISSW25, Bangalore, India*. Universities Press, 2006.
267. A. Jameson and P. D. Lax. Conditions for the Construction of Multi-Point Total Variation Diminishing Difference Schemes. Technical Report MAE 1650, Department of Mechanical and Aerospace Engineering, University of Princeton, USA, 1984.
268. A. Jameson, W. Schmidt, and E. Turkel. Numerical Solution of the Euler Equations by Finite Volume Methods using Runge-Kutta Stepping Schemes. Technical Report 81-1259, AIAA, 1981.
269. A. Jeffrey. *Quasilinear Hyperbolic Systems and Waves*. Pitman, 1976.
270. G. S. Jiang and C. W. Shu. Efficient Implementation of Weighted ENO Schemes. *J. Comp. Phys.*, 126:202–228, 1996.
271. G. S. Jiang and E. Tadmor. Non-oscillatory Central Schemes for Multi-dimensional Hyperbolic Conservation Laws. *SIAM Journal of Scientific Computing*, 19 (6):1892–1917, 1998.
272. F. John. *Partial Differential Equations*. Springer-Verlag, 1982.
273. L. W. Johnson and R. D. Riess. *Numerical Analysis*. Addison-Wesley Publishing Company, 1982.
274. P. Jorgenson and E. Turkel. Central-Difference TVD and TVB Schemes for Time-Dependent and Steady State Problems. Technical Report ICOMP-91-27; AIAA-92-0053, NASA Lewis Research Center, Cleveland, Ohio, USA, 1992.
275. D. Kahaner, C. Moler, and S. Nash. *Numerical Methods and Software*. Prentice Hall, Englewood Cliffs, New Jersey, 1989.
276. K. Kailasanath, E. S. Oran, J. P. Boris, and T. R. Young. Determination of the Detonation Cell Size and the Role of Transverse Waves in Two-Dimensional Detonations. *Combustion and Flame*, 61:199–209, 1985.
277. S. Karni. Accelerated Convergence to Steady State by Gradual Far-Field Damping. *AIAA J.*, 30:1220–1228, 1992.
278. S. Karni. Multicomponent Flow Calculations Using a Consistent Primitive Algorithm. *J. Comput. Phys.*, 112(1):31–43, 1994.
279. S. Karni. Hybrid Multifluid Algorithms. Technical Report 95-001, Courant Mathematics and Computing Laboratory, 1995.
280. S. Karni and S. Čanić. Computations of Slowly Moving Shocks. *J. Comput. Phys.*, 136:132–139, 1997.
281. M. Käser. *Adaptive Methods for the Numerical Simulation of Transport Processes*. PhD thesis, Institute of Numerical Mathematics and Scientific Computing, University of Munich, Germany, 2003.

282. M. Käser. ADER Schemes for the Solution of Conservation Laws on Adaptive Triangulations. Mathematical Methods and Modelling in Hydrocarbon Exploration and Production. Springer-Verlag (to appear), 2004.
283. M. Käser and A. Iske. Adaptive ADER Schemes for the Solution of Scalar Non-Linear Hyperbolic Problems. *J. Comput. Phys.*, :-:, 2005.
284. K. S. Kim and J. Ballmann. Numerische Simulation Mechanischer Wellen in Geschichteten Elastischen Körpern. *Z. Angew. Math. Mech.*, 70:204–206, 1990.
285. C. Klingenberg, W. Schmidt, and K. Waagan. Numerical Comparison of Riemann Solvers for Astrophysical Hydrodynamics. *J. Comput. Phys.*, 227(1):12–35, 2007.
286. L. Knut-Andreas and S. Noelle. An Improved Quadrature Rule for the Flux Computation in Staggered Central Difference Schemes in Multidimensions. *J. Scientific Computing*, 18(1):69–81, 2003.
287. V. P. Kolgan. Application of the Principle of Minimum Derivatives to the Construction of Difference Schemes for Computing Discontinuous Solutions of Gas Dynamics (in Russian). *Uch. Zap. TsaGI, Russia*, 3(6):68–77, 1972.
288. V. P. Kolgan. Finite Difference Schemes for the Computation of Three-dimensional Solutions of Gas Dynamics and Calculation of the Flow over a Body under an Angle of Attack (in Russian). Technical Report Vol. 6, No 2, pp 1–6, *Uchenye Zapiski TsaGI* (Scientific Notes of the Central Institute of Aerodynamics, Russia), 1975.
289. V. P. Kolgan. Finite Difference Schemes for the Computation of Two-dimensional Discontinuous Solutions of Non-stationary Gas Dynamics (in Russian). Technical Report Vol. 6, No 1, pp 9–14, *Uchenye Zapiski TsaGI* (Scientific Notes of the Central Institute of Aerodynamics, Russia), 1975.
290. B. Koren and S. P. Spekreijse. Multigrid and Defect Correction for the Efficient Solution of the Steady Euler Equations. In *Notes on Numerical Fluid Mechanics*, Vol. 17, pages 87–100. Vieweg, 1987.
291. D. Kröner. *Numerical Schemes for Conservation Laws*. Wiley Teubner, 1997.
292. A. G. Kulikovskii, N. V. Pogorelov, and A. Y. Semenov. *Mathematical Aspects of Numerical Solution of Hyperbolic Systems*. Chapman and Hall, 2001.
293. A. Kurganov and G. Petrova. A Third-Order Semi-Discrete Genuinely Multidimensional Central Scheme for Hyperbolic Conservation Laws and Related Problems. *Numerische Mathematik*, 88(3):683–729, 2001.
294. A. Lafon and H. C. Yee. Dynamical Approach Study of Spurious Steady-State Numerical Solutions of Nonlinear Differential Equations, Part III. The Effects of Nonlinear Source Terms in Reaction-Convection Equations. *Intern. J. Computational Fluid Dynamics*, 6:1–36, 1996.
295. A. Lafon and H. C. Yee. Dynamical Approach Study of Spurious Steady-State Numerical Solutions of Nonlinear Differential Equations, Part IV. Stability vs. Methods of Discretizing Nonlinear Source Terms in Reaction-Convection Equations. *Intern. J. Computational Fluid Dynamics*, 6:89–123, 1996.
296. J. D. Lambert. *Computational Methods in Ordinary Differential Equations*. John Wiley and Sons, 1973.
297. L. D. Landau and E. M. Lifshitz. *Fluid Mechanics*. Pergamon Press, 1982.
298. C. B. Laney. *Computational Gasdynamics*. Cambridge University Press, 1998.
299. P. D. Lax. Weak Solutions of Nonlinear Hyperbolic Equations and Their Numerical Computation. *Comm. Pure. Appl. Math.*, VII:159–193, 1954.
300. P. D. Lax. The Formation and Decay of Shock Waves. *American Mathematical Monthly*, 79:227–241, 1972.

301. P. D. Lax. *Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves*. Society for Industrial and Applied Mathematics, Philadelphia, 1990.
302. P. D. Lax and B. Wendroff. Systems of Conservation Laws. *Comm. Pure Appl. Math.*, 13:217–237, 1960.
303. P. D. Lax and B. Wendroff. Difference Schemes for Hyperbolic Equations with High Order Accuracy. *Comm. Pure Appl. Math.*, XVII:381–393, 1964.
304. P. Le Floch and P. A. Raviart. An Asymptotic Expansion for the Solution of the Generalized Riemann Problem. Part 1: General Theory. *Ann. Inst. Henri Poincaré. Analyse non Linéaire*, 5(2):179–207, 1988.
305. P. Le Floch and L. Tatsien. A Global Asymptotic Expansion for the Solution of the Generalized Riemann Problem. *Ann. Inst. Henri Poincaré. Analyse non Linéaire*, 3:321–340, 1991.
306. L. Lehner. Numerical Relativity: a Review. *Class. Quantum Grav.*, 18:R25–R86, 2001.
307. B. P. Leonard. Simple High-Accuracy Resolution Program for Convective Modelling of Discontinuities. *Int. J. Numer. Methods in Fluids*, 8:1291–1318, 1988.
308. R. J. LeVeque. *Numerical Methods for Conservation Laws*. Birkhäuser Verlag, 1992.
309. R. J. LeVeque. Simplified Multidimensional Flux Limiter Methods. In M. J. Baines and K. W. Morton, editors, *Numerical Methods in Fluid Dynamics 4: Proceedings of the 1992 International Conference on Numerical Methods in Fluids*, pages 173–189, Reading, 1993.
310. R. J. LeVeque. Balancing Source Terms and Flux Gradients in High-Resolution Godunov Methods. *J. Comput. Phys.*, 146:346–365, 1998.
311. R. J. LeVeque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge University Press, 2002.
312. R. J. LeVeque, D. Mihalas, E. A. Dorfi, and E. Müller. *Computational Methods for Astrophysical Flow*. Springer-Verlag, 1998.
313. R. J. LeVeque and K. M. Shyue. Two-Dimensional Front Tracking Based on High Resolution Wave Propagation Methods. *J. Comput. Phys.*, 123:354–368, 1996.
314. R. J. LeVeque and H. C. Yee. A Study of Numerical Methods for Hyperbolic Conservation Laws with Stiff Source Terms. *J. Comput. Phys.*, 86:187–210, 1990.
315. D. W. Levy, G. Puppo, and G. Russo. Central WENO Schemes for Hyperbolic Systems of Conservation Laws. *Mathematical Models and Numerical Analysis*, 33:547–571, 1999.
316. D. W. Levy, G. Puppo, and G. Russo. A Fourth Order Central WENO Scheme for Multidimensional Systems of Hyperbolic Conservation Laws. *SIAM J. Scientific Computing*, 24:480–506, 2002.
317. K. M. Li and M. Holt. Numerical Solutions to Water Waves Generated by Shallow Underwater Explosions. *Phys. Fluids*, 24:816–824, 1981.
318. S. Li. An HLLC Riemann Solver for Magneto-Hydrodynamics. *J. Comput. Phys.*, 203(1):344–357, 2005.
319. NAG Library. NAG Library Routines, Mark 18, Routines D03PWF and D03PXF, 1996.

320. X. Lin and J. Ballmann. A Riemann Solver and a Second-Order Godunov Method for Elastic-Plastic Wave Propagation in Solids. *Int. J. Impact Engng.*, 13(3):463–478, 1993.
321. X. Lin and Ballmann J. Numerical Method for Elastic-Plastic Waves in Cracked Solids, Part 1: Anti-Plane Shear Wave. *Archive of Applied Mechanics*, 63:261–282, 1993.
322. X. Lin and Ballmann J. Numerical Method for Elastic-Plastic Waves in Cracked Solids, Part 2: Plain Strain Problem. *Archive of Applied Mechanics*, 63:283–295, 1993.
323. X. Lin and Ballmann J. Reconsideration of Chen’s Problem by Finite Difference Method. *Engineering Fracture and Mechanics*, 44(5):735–739, 1993.
324. T. J. Linde. *A Three-Dimensional Adaptive Multifluid MHD Model for the Heliosphere*. PhD thesis, Aerospace Engineering, University of Michigan, USA, 1998.
325. T. J. Linde, T. I. Gombosi, P. L. Roe, K. G. Powell, and D. L. DeZeeuw. The Heliosphere in Magnetized Local Interstellar Medium: Results of a 3D MHD Simulation. *J. Geophys. Res.*, 103(A2):1989–1904, 1998.
326. W. B. Lindquist. *Current Progress in Hyperbolic Systems: Riemann Problems and Computations*. American Mathematical Society, 1989.
327. M. S. Liou. Recent Progress and Applications of AUSM+. In *Sixteenth International Conference on Numerical Methods in Fluid Dynamics, Lecture Notes in Physics 515*, pages 302–307. Springer-Verlag, 1998.
328. M. S. Liou and C. J. Steffen. A New Flux Splitting Scheme. *J. Comput. Phys.*, 107:23–39, 1993.
329. R. Liska and B. Wendroff. Composite Schemes for Conservation Laws. *SIAM J. Numerical Analysis*, 35(6):2250–2271, 1998.
330. X. D. Liu and P. D. Lax. Positive Schemes for Solving Multi-Dimensional Hyperbolic Conservation Laws. Technical Report 95-003, Courant Mathematics and Computing Laboratory, New York University, USA, 1995.
331. X. D. Liu and T. Osher, S. Chan. Weighted Essentially Non-oscillatory Schemes. *J. Comput. Phys.*, 115:200–212, 1994.
332. C. Y. Loh, M. S. Liou, and W. H. Hui. An Investigation of Random Choice Method for Three Dimensional Steady Supersonic Flow. In *First Asian Computational Fluid Dynamics Conference*, pages 963–971. The Hong Kong University of Science and Technology, 1995.
333. J. Lorenz and H. J. Schroll. Stiff Well-Posedness for Hyperbolic Systems with Large Relaxation Terms. Technical Report 125, Institut für Geometrie un Praktische Mathematik, RWTH Aachen, Germany, 1996.
334. H. Luo, J. D. Baum, and R. Löhner. A Hermite WENO-Based Limiter for Discontinuous Galerkin Method on Unstructured Grids. *J. Comput. Phys.*, 225(1):686–713, 2007.
335. A. Majda and S. Osher. Numerical Viscosity and Entropy Condition. *Comm. Pure and Applied Mathematics*, 32:797–838, 1979.
336. J. C. Mandal and S. M. Desphande. Kinetic Flux Vector Splitting for the Euler Equations. *Computers and Fluids*, 23:447–, 1994.
337. M. J. Maron and R. J. Lopez. *Numerical Analysis*. Wadsworth, 1991.
338. A. Marquina, J. M. Martí, J. M. Ibañez, Miralles J. A., and R. Donat. Ultrarelativistic Hydrodynamics: High-Resolution Shock-Capturing Methods. *Astron. Astrophys.*, 258:566–, 1992.

339. E. Marshall and R. Méndez. Computational Aspects of the Random Choice Method for Shallow Water Equations. *J. Comput. Phys.*, 39:1–21, 1981.
340. G. Marshall and B. Plohr. A Random Choice Method for Two-Dimensional Steady Supersonic Shock Wave Diffraction Problems. *J. Comput. Phys.*, 56:410–427, 1984.
341. J. M. Martí, J. M. Ibañez, and J. A. Miralles. Numerical Relativistic Hydrodynamics: Local Characteristic Approach. *Phys. Rev.*, 43:3794–, 1991.
342. J. M. Martí and E. Müller. The Analytical Solution of the Riemann Problem for Relativistic Hydrodynamics. *Journal of Fluid Mechanics*, 258:317–333, 1994.
343. J. M. Martí, E. Müller, J. A. Font, J. M. Ibañez, and A. Marquina. Morphology and Dynamics of Relativistic Jets. Technical Report MPA 895, Max-Planck-Institut für Astrophysik, Garching, Germany, 1995.
344. Y. P. Marx. A Practical Implementation of Turbulence Models for the Computation of Three-Dimensional Separated Flows. *Int. J. Numer. Meth. Fluids*, 13:775–796, 1991.
345. Y. P. Marx. Evaluation of the Artificial Compressibility Method for the Solution of the Incompressible Navier–Stokes Equations. In *Notes on Numerical Fluid Mechanics*, Vol. 35, pages 201–210. Vieweg, 1991.
346. Y. P. Marx. Time Integration Schemes for the Unsteady Incompressible Navier–Stokes Equations. *J. Comput. Phys.*, 112:182–209, 1994.
347. J. H. Mathews. *Numerical Methods*. Prentice-Hall International, Inc., 1987.
348. C. Y. McNeil. *Efficient Upwind Algorithms for Solution of the Euler and Navier–Stokes Equations*. PhD thesis, Department of Aerospace Science, Cranfield University, UK, 1995.
349. R. Menikoff and B. J. Plohr. The Riemann Problem for Fluid Flow of Real Materials. *Reviews of Modern Physics*, 61:75–130, 1989.
350. I. S. Men'shov. Increasing the Order of Approximation of Godunov's Scheme Using the Generalized Riemann Problem. *USSR Comput. Math. Phys.*, 30(5):54–65, 1990.
351. A. Mignone and G. Bodo. An HLLC Riemann Solver for Relativistic Flows. *Monthly Notices of the Royal Astronomical Society*, 368(3):1040–1054, 2006.
352. A. R. Mitchell and D. F. Griffiths. *The Finite Difference Method in Partial Differential Equations*. John Wiley and Sons, 1980.
353. G. Moretti. The  $\lambda$ -scheme. *Computers and Fluids*, 7:191–205, 1979.
354. G. Moretti. Computation of Flows with Strong Shocks. *Ann. Rev. Fluid Mech.*, 19:313–337, 1987.
355. G. Moretti and G. Bleich. A Time-Dependent Computational Method for Blunt-Body Flows. *AIAA J.*, 4:2136–2141, 1966.
356. S. Müller and A. Voss. The Riemann Problem for the Euler Equations with Nonconvex and Nonsmooth Equation of State: Construction of Wave Curves. *SIAM Journal on Scientific Computing*, 28(2):651–681, 2006.
357. S. T. Munkejord, S. Evje, and T. Flatten. A Multi-Stage Centred Scheme Approach Applied to a Drift Flux Two-Phase Flow Model. *Int. J. Numer. Meth. Fluids*, 52:679–705, 2006.
358. C. D. Munz. A Tracking Method for Gas Flow into Vacuum Based on the Vacuum Riemann Problem. *Mathematical Methods in Applied Sciences*, 17:597–612, 1994.
359. C. D. Munz, R. Schneider, and O. Gerlinger. The Numerical Approximation of a Free Gas–Vacuum Interface. In *Numerical Methods for the Navier–Stokes Equations*, pages 181–190. Vieweg–Verlag, 1994.

360. S. Navarro-Martinez and O. R. Tutty. Numerical Simulation of Görtler Vortices in Hypersonic Compression Ramps. *Computers and Fluids*, 34(2):225–247, 2005.
361. A. Nemeć and M. J. Aftosmis. Adjoint Sensitivity Computations for an Embedded-Boundary Cartesian Mesh Method. *J. Comput. Phys.*, 227(4):2724–2742, 2008.
362. H. Nessyahu and E. Tadmor. Non-Oscillatory Central Differencing for Hyperbolic Conservation Laws. *J. Comput. Phys.*, 87:408–463, 1990.
363. N. Nikiforakis and J. F. Clarke. Quasi-Steady Structures in the Two Dimensional Initiation of Detonations. *Proc. Roy. Soc. Lond. A*, 452:2023–2042, 1996.
364. H. Olivier and H. Grönig. The Random Choice Method Applied to Two-Dimensional Shock Focusing and Diffraction. *J. Comput. Phys.*, 63:85–106, 1986.
365. A. K. Oppenheim and R. I. Soloukhin. Experiments in Gas Dynamics of Explosions. *Annual Review of Fluid Mechanics*, 5:31–58, 1973.
366. E. Oran and J. P. Boris. *Numerical Simulation of Reactive Flow*. Cambridge University Press, 2001.
367. E. S. Oran and J. P. Boris. Detailed Modelling of Combustion Systems. *Prog. Energy. Comb. Sci.*, 7:1–72, 1981.
368. E. S. Oran and J. P. Boris. *Numerical Simulation of Reactive Flow*. Elsevier, New York, 1987.
369. S. Osher. Riemann Solvers, the Entropy Condition and Difference Approximations. *SIAM J. Numer. Anal.*, 21(2):217–235, 1984.
370. S. Osher and S. Chakravarthy. Upwind Schemes and Boundary Conditions with Applications to Euler Equations in General Geometries. *J. Comput. Phys.*, 50:447–481, 1983.
371. S. Osher and S. Chakravarthy. High Resolution Schemes and the Entropy Condition. *SIAM J. Numer. Anal.*, 21(5):955–984, 1984.
372. S. Osher and F. Solomon. Upwind Difference Schemes for Hyperbolic Conservation Laws. *Math. Comp.*, 38,158:339–374, 1982.
373. Liu T. P. and J. A. Smoller. On the Vacuum State for the Isentropic Gas Dynamics Equations. *Advances in Applied Mathematics*, 1:345–359, 1980.
374. M. Pandolfi. Computation of Steady Supersonic Flows by a Flux-Difference Splitting Method. *Computers and Fluids*, 10:37–46, 1985.
375. C. Parés. Numerical Methods for Nonconservative Hyperbolic Systems: a Theoretical Framework. *SIAM Journal on Numerical Analysis*, 44:300–321, 2006.
376. M. Pelanti, F. Bouchut, and A. Mangeney. A Roe-Type Scheme for Two-Phase Shallow Granular Flows over Variable Topography. *Mathematical Modelling and Numerical Analysis*, 42:851–885, 2008.
377. R. B. Pember. Numerical Methods for Hyperbolic Conservation Laws with Stiff Relaxation I: Spurious Solutions. *SIAM J. Appl. Math.*, 53:1293–1330, 1993.
378. R. B. Pember. Numerical Methods for Hyperbolic Conservation Laws with Stiff Relaxation II: Higher-Order Godunov Methods. *SIAM J. Sci. Comput.*, 14:824–859, 1993.
379. L. L. Pennisi. *Elements of Complex Analysis*. Holt, Rinehart and Winston, 1963.
380. B. Perthame. Boltzmann Type Schemes for Gas Dynamics and the Entropy Property. *SIAM J. Numer. Anal.*, 27:1405–1421, 1990.

381. B. Perthame. Second-Order Boltzman Schemes for Compressible Euler Equations in one and two Space Dimensions. *SIAM J. Numer. Anal.*, 29:1992, 1992.
382. L. Pesch and J. J. W. van der Vegt. A Discontinuous Galerkin Finite Element Discretization of the Euler Equations for Compressible and Incompressible Fluids. *J. Comput. Phys.*, 227(11):5426–5446, 2008.
383. I. G. Petrovskii. *Partial Differential Equations*. London Cliffe Books Ltd., 1967.
384. J. Pike. Riemann Solvers for Perfect and Near-Perfect Gases. *AIAA Journal*, 31(10):1801–1808, 1993.
385. B. Pitman and L. Le. A Two-Fluid Model for Avalanche and Debris Flows. *Phil. Trans. R. Soc. A*, 363:1573–1601, 2005.
386. B. J. Plohr and D. H. Sharp. Riemann Problems and their Application to Ultra-Relativistic Heavy Ion Collisions. In *The VIII International Conference on Mathematical Physics* (Mebkhout and Séneor, editors), pages 708–713. World Scientific Pub. Co., Singapore, 1987.
387. K. H. Prendergast and K. Xu. Numerical Hydrodynamics from Gas-Kinetic Theory. *J. Comput. Phys.*, 109:53–66, 1993.
388. A. Prosperetti and G. Tryggvason. *Fundamentals of Multiphase Flow*. Cambridge University Press, 2007.
389. D. I. Pullin. Direct Simulation Methods for Compressible Inviscid Ideal Gas Flow. *J. Comput. Phys.*, 34:231–244, 1980.
390. N. Qin. A Comparative Study of Two Upwind Schemes as Applied to Navier-Stokes Solutions for Resolving Boundary Layers in Hypersonic Viscous Flow. Technical Report GU Aero Report 9120, Department of Aerospace Engineering, University of Glasgow, UK, 1992.
391. N. Qin and G. W. Foster. Study of Flow Interactions due to a Supersonic Lateral Jet Using High Resolution Navier-Stokes Solutions. In *AIAA-95-2151 Paper, 27th AIAA Fluid Dynamics Conference, June 1995, San Diego, USA*, 1995.
392. N. Qin and A. Redlich. Massively Separated Flows due to Transverse Sonic Jet in Hypersonic Laminar Stream. In *AIAA-96-1934 Paper, 27th AIAA Fluid Dynamics Conference, June 1996, New Orleans, USA*, 1996.
393. N. Qin and B. E. Richards. Sparse Quasi-Newton Method for High Resolution Schemes. In *Notes on Numerical Fluid Mechanics*, Vol. 20, pages 310–317. Vieweg, 1988.
394. N. Qin and B. E. Richards. Sparse Quasi-Newton Method for Navier-Stokes Solution. In *Notes on Numerical Fluid Mechanics*, Vol. 29, pages 474–483. Vieweg, 1990.
395. N. Qin, K. W. Scriba, and B. E. Richards. Shock-Shock, Shock-Vortex Interaction and Aerodynamic Heating In Hypersonic Corner Flow. *The Aeronautical Journal*, pages 152–160, 1991.
396. N. Qin, X. Xu, and B. E. Richards. Newton-Like Methods for Fast High Resolution Simulation of Hypersonic Viscous Flows. *Computing Systems in Engineering*, 3:429–435, 1992.
397. J. Qiu, B. C. Khoo, and C. W. Shu. A Numerical Study for the Performance of the Runge-Kutta Discontinuous Galerkin Method Based on Different Numerical Fluxes. *J. Comput. Phys.*, 212(2):540–565, 2006.
398. L. Quartapelle, L. Castelletti, A. Guardone, and G. Quaranta. Solution of the Riemann Problem of Classical Gasdynamics. *J. Comput. Phys.*, 190:118–140, 2003.

399. J. J. Quirk. An Adaptive Grid Algorithm for Computational Shock Hydrodynamics. PhD Thesis, College of Aeronautics, Cranfield Institute of Technology, UK, 1991.
400. J. J. Quirk. An Alternative to Unstructured Grids for Computing Gas Dynamic Flows Around Arbitrarily Complex Two Dimensional Bodies. *Computers and Fluids*, 23(1):125–142, 1994.
401. S. Rhebergen, O. Bokhove, and J. J. W. van der Vegt. Discontinuous Galerkin Finite Element Methods for Hyperbolic Nonconservative Partial Differential Equations. *J. Comput. Phys.*, 227:1887–1922, 2008.
402. R. D. Richtmyer and K. W. Morton. *Difference Methods for Initial Value Problems*. Interscience-Wiley, New York, 1967.
403. N. H. Risebro and A. Tveito. Front Tracking Applied to Nonstrictly Hyperbolic Systems of Conservation Laws. *SIAM J. Sci. Stat. Comput.*, 12:1401–1419, 1991.
404. A. Rizzi. Numerical Implementation of Solid-Body Boundary Conditions for the Euler Equations. *ZAMM*, 58:301–304, 1978.
405. P. J. Roache. *Computational Fluid Dynamics*. Hermosa Publishers, 1982.
406. T. W. Roberts. The Behavior of Flux Difference Splitting Schemes near Slowly Moving Shock Waves. *J. Comput. Phys.*, 90:141–160, 1990.
407. P. L. Roe. Approximate Riemann Solvers, Parameter Vectors, and Difference Schemes. *J. Comput. Phys.*, 43:357–372, 1981.
408. P. L. Roe. Numerical Algorithms for the Linear Wave Equation. Technical Report 81047, Royal Aircraft Establishment, Bedford, UK, 1981.
409. P. L. Roe. Some Contributions to the Modelling of Discontinuous Flows. In *Proceedings of the SIAM/AMS Seminar*, San Diego, 1983.
410. P. L. Roe. Generalized Formulation of TVD Lax–Wendroff Schemes. Technical Report ICASE 84–53, NASA Langley Research Center, Hampton, USA, 1984.
411. P. L. Roe. Discrete Models for the Numerical Analysis of Time–Dependent Multidimensional Gas Dynamics. *J. Comput. Phys.*, 63:458–476, 1986.
412. P. L. Roe. Upwind Differencing Schemes for Hyperbolic Conservation Laws with Source Terms. In *Proc. First International Conference on Hyperbolic Problems*, Carasso, Raviart and Serre (Editors), pages 41–51. Springer, 1986.
413. P. L. Roe. Remote Boundary Conditions for Unsteady Multidimensional Aerodynamic Computations. *Computers and Fluids*, 17:221–231, 1989.
414. P. L. Roe. Sonic Flux Formulae. *SIAM J. Sci. Stat. Comput.*, 13(2):611–630, 1992.
415. P. L. Roe. The Harten Memorial Lecture—New Applications of Upwinding. In *Numerical Methods for Wave Propagation*. Toro, E. F. and Clarke, J. F. (Editors), pages 1–31. Kluwer Academic Publishers, 1998.
416. P. L. Roe and J. Pike. Efficient Construction and Utilisation of Approximate Riemann Solutions. In *Computing Methods in Applied Science and Engineering*. North–Holland, 1984.
417. E. Romenski, E. D. Resnyanski, and E. F. Toro. Conservative Hyperbolic Formulation for Compressible Two–Phase Flow with Different Phase Pressures and Temperatures. *Quarterly of Applied Mathematics*, 65:259–279, 2007.
418. V. V. Rusanov. Calculation of Interaction of Non–Steady Shock Waves with Obstacles. *J. Comput. Math. Phys. USSR*, 1:267–279, 1961.
419. G. Russo, E. F. Toro, and V. A. Titarev. ADER–Runge–Kutta Schemes for Conservation Laws in One Space Dimension. In *Hyperbolic Problems: Theory,*

- Numerics and Applications. Benzone–Gavage S. and Serre D. (Editors), Vol. 1*, pages 929–936. Springer, 2008.
420. A. V. Safranov. Difference Method for Gasdynamical Equations Based on the Jump Conditions (in Russian). *Mathematical Modelling*, 20:76–84, 2008.
  421. L. Sainsaulieu. Finite Volume Approximation of Two Phase–Fluid Flows Based on an Approximate Roe–Type Riemann Solver. *J. Comput. Phys.*, 121:1–28, 1995.
  422. T. Saito and I. I. Glass. Application of Random Choice to Problems in Shock and Detonation Wave Dynamics. Technical Report UTIAS 240, Institute for Aerospace Studies, University of Toronto, 1979.
  423. D. A. Sánchez. *Ordinary Differential Equations and Stability Theory*. Dover Publications, Inc., 1968.
  424. R. H. Sanders and K. H. Prendergast. On the Origin the Three Kiloparsec Arm. *Astrophys. J.*, 188:489–500, 1974.
  425. A. Sangan. An HLLC Scheme for Ten–Moments Approximation Coupled with Magnetic Field. *Internat. J. Comput. Science and Math.*, 2(1/2):73–109, 2008.
  426. R. Saurel and R. Abgrall. A Multiphase Godunov Method for Compressible Multifluid and Multiphase Flows. *J. Comput. Phys.*, 150:425–467, 1999.
  427. R. Saurel, E. Daniel, and J. C. Loraud. Two–Phase Flows: Second Order Schemes and Boundary Conditions. *AIAA Journal*, 32(6):1214–1221, 1994.
  428. R. Saurel, A. Forestier, D. Veyret, and J.-C. Loraud. A Finite Volume Scheme for Two–Phase Compressible Flows. *Int. J. Numer. Meth. Fluids*, 18:803–819, 1994.
  429. R. Saurel, M. Larini, and J. C. Loraud. Exact and Approximate Riemann Solvers for Real Gases. *J. Comput. Phys.*, 112:126–137, 1994.
  430. M. Schleicher. Ein Einfaches und Effizientes Verfahren zur Loesung des Riemann–Problems. *Z. Flugwiss. Weltraumforsch.*, 17:265–269, 1993.
  431. V. Schneider, U. Katscher, D. H. Rischke, B. Waldhauser, J. A. Maruhn, and C. D. Munz. New Algorithms for Ultra–Relativistic Numerical Hydrodynamics. *J. Comput. Phys.*, 105:92–, 1993.
  432. H. J. Schroll and A. Tveito. A System of Conservation Laws with a Relaxation Term. Technical Report 103, Institut für Geometrie un Praktische Mathematik, RWTH Aachen, Germany, 1994.
  433. H. J. Schroll, A. Tveito, and R. Winther. An Error Bound for a Finite Difference Scheme Applied to a Stiff System of Conservation Laws. Technical Report 1994–3, Department of Informatics, University of Oslo, Norway, 1994.
  434. H. J. Schroll and R. Winther. Finite Difference Schemes for Scalar Conservation Laws with Source Terms. Technical Report 107, Institut für Geometrie un Praktische Mathematik, RWTH Aachen, Germany, 1994.
  435. T. Schwartzkopff. *Finite–Volumen Verfahren Hoher Ordnung und Heterogene Gebietszerlegung für die Numerische Aeroakustik*. PhD thesis, Institut für Aero- un Gasdynamik, Universität Stuttgart, Germany, 2005.
  436. T. Schwartzkopff, M. Dumbser, and C. D. Munz. Fast High Order ADER Schemes for Linear Hyperbolic Equations. *J. Comput. Phys.*, 197:532–539, 2004.
  437. T. Schwartzkopff, C. D. Munz, and E. F. Toro. ADER: High–Order Approach for Linear Hyperbolic Systems in 2D. *J. Scientific Computing*, 17:231–240, 2002.

438. D. W. Schwendeman, C. W. Wahle, and A. K. Kapila. The Riemann Problem and a High-Resolution Godunov Method for a Model of Compressible Two-Phase Flow. *J. Comput. Phys.*, 212:490–526, 2006.
439. F. W. Sears and G. L. Salinger. *Thermodynamics, Kinetic Theory and Statistical Thermodynamics*. Addison-Wesley, 1986.
440. C. C. L. Sells. . Technical Report RAE TR 80065, Royal Aircraft Establishment, Bedford, UK, 1980.
441. L. F. Shampine. *Numerical Solution of Ordinary Differential Equations*. Chapman and Hall, London and New York, 1994.
442. J. S. Shang. A Fractional-Step Method for Solving 3D, Time-Domain Maxwell Equations. *J. Comput. Phys.*, 118:109–119, 1995.
443. Z. C. Shi and J. J. Gottlieb. Random Choice Method for Two Dimensional Planar and Axisymmetric Steady Supersonic Flows. Technical Report 297, UTIAS, University of Toronto, Canada, 1985.
444. C. Shu and S. Osher. Efficient Implementation of Essentially Non-oscillatory Shock-Capturing Schemes II. *J. Comput. Phys.*, 83:32–78, 1989.
445. C. W. Shu. TVB Boundary Treatment for Numerical Solutions of Conservation Laws. *Math. Comp.*, 49(179):123–134, 1987.
446. C. W. Shu. TVB Uniformly High-Order Schemes for Conservation Laws. *Math. Comput.*, 49(179):105–121, 1987.
447. C. W. Shu and S. Osher. Efficient Implementation of Essentially Non-oscillatory Shock-Capturing Schemes. *J. Comput. Phys.*, 77:439–471, 1988.
448. G. Singh. *A Numerical Study of the Evolution of Piston Driven Detonation*. PhD thesis, Department of Aerospace Science, Cranfield University, UK, 1990.
449. G. Singh and J. F. Clarke. Transient Phenomena in the Initiation of a Mechanically Driven Plane Detonation. *Proc. Roy. Soc. London A*, 438:23–46, 1992.
450. G. D. Smith. *Numerical Solution of Partial Differential Equations*. Clarendon Press, Oxford, 1978.
451. J. Smoller. *Shock Waves and Reaction-Diffusion Equations*. Springer-Verlag, 1994.
452. G. A. Sod. A Numerical Study of a Converging Cylindrical Shock. *J. Fluid Mechanics*, 83:785–794, 1977.
453. G. A. Sod. A Survey of Several Finite Difference Methods for Systems of Nonlinear Hyperbolic Conservation Laws. *J. Comput. Phys.*, 27:1–31, 1978.
454. G. A. Sod. *Numerical Methods in Fluid Dynamics*. Cambridge University Press, 1985.
455. G. A. Sod. A Random Choice Method for the Stefan Problem. *Computers Math. Applic.*, 19:1–9, 1990.
456. W. Speares. A Finite Volume Approach to the Weighted Average Flux Method. MSc thesis, College of Aeronautics, Cranfield Institute of Technology, UK, 1991.
457. W. Speares and E. F. Toro. A High-Resolution Algorithm for Time-Dependent Shock Dominated Problems with Adaptive Mesh Refinement. *Z. Flugwiss. Weltraumforsch. (Journal of Flight Sciences and Space Research)*, 19:267–281, 1995.
458. S. P. Spekreijse. Multigrid Solution of Monotone Second-Order Discretizations of Hyperbolic Conservation Laws. *Math. Comp.*, 49(179):135–155, 1987.

459. S. P. Spekreijse. *Multigrid Solution of the Steady Euler Equations*. PhD thesis, Centre for Mathematics and Computer Science, Amsterdam, The Netherlands, 1988.
460. H. Stadtke. *Gasdynamic Aspects of Two-Phase Flow*. Wiley-VCH, Berlin, 2006.
461. H. Stadtke, G. Franchello, B. Worth, U. Graf, P. Romstedt, A. Kumbaro, J. Garca-Cascales, H. Pailliere, H. Deconinck, M. Ricchiuto, B. Smith, F. De Cachard, E. F. Toro, E. Romenski, and S. Mimouni. Advanced Three-Dimensional Two-Phase Flow Simulation Tools for Application to Reactor Safety (ASTAR). *Nuclear Engineering and Design*, 235:379–400, 2005.
462. J. L. Steger, F. C. Dougherty, and J. A. Benek. A Chimera Grid Scheme. In *Advances in Grid Generation, ASME FED, Vol. 5*, 1983.
463. J. L. Steger and R. F. Warming. Flux Vector Splitting of the Inviscid Gasdynamic Equations with Applications to Finite Difference Methods. *J. Comput. Phys.*, 40:263–293, 1981.
464. H. B. Stewart and B. Wendroff. Two-Phase Flow: Models and Methods. *J. Comput. Phys.*, 56:363–409, 1984.
465. J. J. Stoker. *Water Waves. The Mathematical Theory with Applications*. John Wiley and Sons, 1992.
466. G. Strang. On the Construction and Comparison of Difference Schemes. *SIAM J. Numer. Anal.*, 5(3):506–517, 1968.
467. R. C. Swanson and E. Turkel. On Central-Difference and Upwind Methods. Technical Report ICASE 90-44, NASA Langley Research Center, Hampton, Virginia, USA, 1990.
468. B. K. Swartz and B. Wendroff. AZTEC: A Front Tracking Code Based on Godunov’s Method. *Applied Numerical Mathematics*, 2:385–397, 1986.
469. P. K. Sweby. *Shock Capturing Schemes*. PhD thesis, Department of Mathematics, University of Reading, UK, 1982.
470. P. K. Sweby. High Resolution Schemes Using Flux Limiters for Hyperbolic Conservation Laws. *SIAM J. Numer. Anal.*, 21:995–1011, 1984.
471. P. K. Sweby. TVD Schemes for Inhomogeneous Conservation Laws. In *Notes on Numerical Fluid Mechanics, Vol. 24, Non-Linear Hyperbolic Equations—Theory, Computation Methods and Applications*, pages 599–607. Vieweg, 1989.
472. P. K. Sweby and M. J. Baines. On Convergence of Roe’s Scheme for the General Non-linear Wave Equation. *J. Comput. Phys.*, 56(1):135–148, 1984.
473. E. Tadmor. Convenient Total Variation Diminishing Conditions for Non-Linear Difference Schemes. *SIAM J. Numer. Anal.*, 25:1002–1014, 1988.
474. M. Takahiro and K. Kanya. A Multi-State HLL Approximate Riemann Solver for Ideal Magnetohydrodynamics. *J. Comput. Phys.*, 208(1):315–344, 2005.
475. Y. Takakura and E. F. Toro. Arbitrarily Accurate Non-Oscillatory Schemes for a Non-Linear Conservation Law. *J. Computational Fluid Dynamics*, 11:7–18, 2002.
476. Y. Takano. An Application of the Random Choice Method to a Reactive Gas with Many Chemical Species. *J. Comput. Phys.*, 67:173–187, 1986.
477. K. Takayama. Optical Flow Visualization of Shock Wave Phenomena (Paul Vieille Memorial Lecture). In *Shock Waves @ Marseille, Vol. 4 (Brun and Dumitrescu, Editors)*, pages 7–16. Springer-Verlag, 1995.
478. K. Takayama (Editor). *Shock Waves (Proc. 18th Int. Symp. on Shock Waves, Sendai, Japan, 1991), Two Volumes*. Springer-Verlag, 1992.

479. S. Taki and T. Fujiwara. Numerical Analysis of Two-Dimensional Nonsteady Detonations. *AIAA Journal*, 16:73–77, 1978.
480. C. K. W. Tam and J. C. Webb. Dispersion-Relation-Preserving Finite Difference Schemes for Computational Acoustics. *J. Comput. Phys.*, 107:262–281, 1993.
481. P. Tamamidis, G. Zhang, and D. N. Assanis. Comparison of Pressure-Based and Artificial Compressibility Methods for Solving 3D Steady Incompressible Viscous Flows. *J. Comput. Phys.*, 124(1):1–13, 1996.
482. T. Tanaka. Finite Volume TVD Scheme on an Unstructured Grid System for Three-Dimensional MHD Simulation of Inhomogeneous Systems Including Strong Background Potential Fields. *J. Comput. Phys.*, 111:381–, 1994.
483. L. Tatsien and Y. Wenc. *Boundary-Value Problems for Quasi-Linear Hyperbolic Systems*. Duke University Mathematics Series, 1985.
484. J. W. Thomas. *Numerical Partial Differential Equations, Texts in Applied Mathematics 22*. Springer-Verlag, 1998.
485. J. F. Thompson, F. C. Thames, and C. W. Mastin. Automatic Numerical Grid Generation of Body-Fitted Curvilinear Coordinate System for Field Containing any Number of Arbitrary Two-Dimensional Bodies. *J. Comput. Phys.*, 15:299–319, 1974.
486. J. F. Thompson, J. F. Warsi, and C. W. Mastin. *Grid Generation: Foundations and Applications*. North-Holland, 1982.
487. J. F. (Editor) Thompson. *Numerical Grid Generation*. North-Holland, 1982.
488. K. W. Thompson. The Special Relativistic Shock Tube. *J. Fluid Mechanics*, 171:365–, 1986.
489. K. W. Thompson. Time Dependent Boundary Conditions for Hyperbolic Systems. *J. Comput. Phys.*, 68:1–24, 1987.
490. V. A. Titarev. *Derivative Riemann Problem and ADER Schemes*. PhD thesis, Department of Mathematics, University of Trento, Italy, 2005.
491. V. A. Titarev, E. Romenski, and E. F. Toro. MUSTA-Type Upwind Fluxes for Non-Linear Elasticity. *Int. J. Numer. Meth. Engineering*, 73:897–926, 2008.
492. V. A. Titarev and E. F. Toro. ADER: Arbitrary High Order Godunov Approach. *J. Scientific Computing*, 17:609–618, 2002.
493. V. A. Titarev and E. F. Toro. Finite Volume WENO Schemes for Three-Dimensional Conservation Laws. *J. Comput. Phys.*, 201(1):238–260, 2004.
494. V. A. Titarev and E. F. Toro. ADER Schemes for Three-Dimensional Hyperbolic Systems. *J. Comput. Phys.*, 204:715–736, 2005.
495. V. A. Titarev and E. F. Toro. MUSTA Schemes for Multi-Dimensional Hyperbolic Systems: Analysis and Improvements. *Int. J. Numer. Meth. Fluids*, 49:117–147, 2005.
496. E. F. Toro. Multi-Stage Predictor-Corrector Fluxes for Hyperbolic Equations. Technical Report NI03037-NPA, Isaac Newton Institute for Mathematical Sciences, University of Cambridge, UK, 17th June, 2003.
497. E. F. Toro. A New Numerical Technique for Quasi-Linear Hyperbolic Systems of Conservation Laws. Technical Report CoA-8608, College of Aeronautics, Cranfield Institute of Technology, UK, 1986.
498. E. F. Toro. A Fast Riemann Solver with Constant Covolume Applied to the Random Choice Method. *Int. J. Numer. Meth. Fluids*, 9:1145–1164, 1989.
499. E. F. Toro. A Weighted Average Flux Method for Hyperbolic Conservation Laws. *Proc. Roy. Soc. London*, A423:401–418, 1989.

500. E. F. Toro. Random Choice Based Hybrid Schemes for one and Two-Dimensional Gas Dynamics. In *Non-linear Hyperbolic Equations – Theory, Computation Methods and Applications. Notes on Numerical Fluid Mechanics*, pages 630–639. Vieweg, Braunschweig, 1989.
501. E. F. Toro. Riemann–Problem Based Techniques for Computing Reactive Two-Phase Flows. In Dervieux and Larroutrou, editors, *Proc. Third. Intern. Confer. on Numerical Combustion*, number 351 in Lecture Notes in Physics, pages 472–481, Antibes, France, May 1989.
502. E. F. Toro. A Linearised Riemann Solver for the Time–Dependent Euler Equations of Gas Dynamics. *Proc. Roy. Soc. London*, A434:683–693, 1991.
503. E. F. Toro. Some Aspects of Shock Capturing Methods for Gas Dynamics. Technical Report COA–9112, College of Aeronautics, Cranfield Institute of Technology, UK, 1991.
504. E. F. Toro. Riemann Problems and the WAF Method for Solving Two-Dimensional Shallow Water Equations. *Phil. Trans. Roy. Soc. London Series A: Physical Sciences and Engineering*, A338:43–68, 1992.
505. E. F. Toro. The Weighted Average Flux Method Applied to the Euler Equations. *Phil. Trans. Roy. Soc. London Series A: Physical Sciences and Engineering*, A341:499–530, 1992.
506. E. F. Toro. The Weighted Average Flux Method Applied to the Time–Dependent Euler Equations. *Phil. Trans. Roy. Soc. London*, A341:499–530, 1992.
507. E. F. Toro. Viscous Flux Limiters. In J. B. Vos, A. Rizzi, and I. L. Ryhming, editors, *Notes on Numerical Fluid Dynamics, Vol. 35*, pages 592–600. Vieweg, 1992.
508. E. F. Toro. Defects of Conservative Approaches and Adaptive Primitive–Conservative Schemes for Computing Solutions to Hyperbolic Conservation Laws. Technical Report MMU 9401, Department of Mathematics and Physics, Manchester Metropolitan University, UK, 1994.
509. E. F. Toro. Direct Riemann Solvers for the Time–Dependent Euler Equations. *Shock Waves*, 5:75–80, 1995.
510. E. F. Toro. MUSCL–Type Primitive Variable Schemes. Technical Report MMU–9501, Department of Mathematics and Physics, Manchester Metropolitan University, UK, 1995.
511. E. F. Toro. On Adaptive Primitive–Conservative Schemes for Conservation Laws. In M. M. Hafez, editor, *Sixth International Symposium on Computational Fluid Dynamics: A Collection of Technical Papers*, volume 3, pages 1288–1293, Lake Tahoe, Nevada, USA, September 4–8, 1995.
512. E. F. Toro. Some IVPs for Which Conservative Methods Fail Miserably. In M. M. Hafez, editor, *Sixth International Symposium on Computational Fluid Dynamics: A Collection of Technical Papers*, volume 3, pages 1294–1299, Lake Tahoe, Nevada, USA, September 4–8, 1995.
513. E. F. Toro. On Glimm–Related Schemes for Conservation Laws. Technical Report MMU–9602, Department of Mathematics and Physics, Manchester Metropolitan University, UK, 1996.
514. E. F. Toro. Exact and Approximate Riemann Solvers for the Artificial Compressibility Equations. Technical Report 97–02, Department of Mathematics and Physics, Manchester Metropolitan University, UK, 1997.

515. E. F. Toro. On Two Glimm–Related Schemes for Hyperbolic Conservation Laws. In *Proc. Fifth Annual Conference of the CFD Society of Canada*, pages 3.49–3.54. University of Victoria, Canada, 1997.
516. E. F. Toro. *Riemann Solvers and Numerical Methods for Fluid Dynamics*. Springer–Verlag, 1997.
517. E. F. Toro. Primitive, Conservative and Adaptive Schemes for Hyperbolic Conservation Laws. In *Numerical Methods for Wave Propagation*. Toro, E. F. and Clarke (Editors), J. F., pages 323–385. Kluwer Academic Publishers, 1998.
518. E. F. Toro. *NUMERICA*: A Library of Source Codes for Teaching, Research and Applications. Numeritek Ltd., www.numeritek.com, 1999.
519. E. F. Toro. *Riemann Solvers and Numerical Methods for Fluid Dynamics, Second Edition*. Springer–Verlag, 1999.
520. E. F. Toro. *Shock–Capturing Methods for Free–Surface Shallow Flows*. Wiley and Sons Ltd., 2001.
521. E. F. Toro. A Multi–Stage Numerical Flux. *Appl. Numer. Math.*, 56:1464–1479, 2006.
522. E. F. Toro. Riemann Solvers with Evolved Initial Conditions. *Int. J. Numer. Meth. in Fluids*, 52:433–453, 2006.
523. E. F. Toro and Titarev V. A. Derivative Riemann Solvers for Systems of Conservation Laws and ADER Methods. *J. Comput. Phys.*, 212(1):150–165, 2006.
524. E. F. Toro and Titarev V. A. MUSTA Schemes for Systems of Conservation Laws. *J. Comput. Phys.*, 216:403–429, 2006.
525. E. F. Toro, S. J. Billett, and E. P. Boden. Advanced Modelling Techniques for Propulsion Systems. In *Proceedings of Conference on Energetic Materials and Propulsion Technology, Salsbury, Australia, April 18–19, 1996*, pages –, 1996.
526. E. F. Toro and S. J. Billett. Centred TVD Schemes for Hyperbolic Conservation Laws. Technical Report MMU–9603, Department of Mathematics and Physics, Manchester Metropolitan University, UK, 1996.
527. E. F. Toro and S. J. Billett. A Unified Riemann–Problem Based Extension of the Warming–Beam and Lax–Wendroff Schemes. *IMA J. Numer. Anal.*, 17:61–102, 1997.
528. E. F. Toro and S. J. Billett. Centred TVD Schemes for Hyperbolic Conservation Laws. *IMA J. Numerical Analysis*, 20:47–79, 2000.
529. E. F. Toro and R. E. Brown. The WAF Method and Splitting Procedures for Viscous, Shocked Flows. In *Proceedings of the 18th International Symposium on Shock Waves, Tohoku University, Sendai, Japan (K. Takayama, Editor)*, pages 1119–1126. Springer–Verlag, 1992.
530. E. F. Toro and C. E. Castro. The Derivative Riemann Problem for the Baer–Nunziato Equations. In *Hyperbolic Problems: Theory, Numerics and Applications. Vol. 1. Benzoni–Gavage and Serre (editors)*, pages 1045–1052. Springer, 2008.
531. E. F. Toro and C. E. Castro. The Derivative Riemann Problem for the Baer–Nunziato Equations. In *Hyperbolic Problems: Theory, Numerics and Applications HYPE2006, Vol. 1. S. Benzoni–Gavage and D. Serre (Editors)*, pages 1045–1052. Springer, 2008.
532. E. F. Toro and A. Chakraborty. Development of an Approximate Riemann Solver for the Steady Supersonic Euler Equations. *The Aeronautical Journal*, 98:325–339, 1994.

533. E. F. Toro and C. C. Chou. A Linearised Riemann Solver for the Steady Supersonic Euler Equations. *Int. J. Numer. Meth. Fluids*, 16:173–186, 1993.
534. E. F. Toro and P. García-Navarro. Godunov-Type Methods for Free-Surface Shallow Flows: A Review. *J. Hydraulic Research*, 45(6):736–751, 2007.
535. E. F. Toro and A. Hidalgo. ADER Finite Volume Schemes for Diffusion-Reaction Equations. *Applied Numerical Mathematics*, 59:73–100, 2009.
536. E. F. Toro, A. Hidalgo, and M. Dumser. FORCE Schemes on Unstructured Meshes I: Conservative Hyperbolic Systems. *Isaac Newton Institute for Mathematical Sciences, University of Cambridge, UK. Preprint NI09005-NPA, September 2008. Submitted to J. Comput. Phys.*, 2008.
537. E. F. Toro, R. C. Millington, and L. A. M. Nejad. Towards Very High-Order Godunov Schemes. In *Godunov Methods: Theory and Applications. Edited Review, E. F. Toro (Editor)*, pages 905–937. Kluwer Academic/Plenum Publishers, 2001.
538. E. F. Toro, M. Olim, and K. Takayama. Unusual Increase in Tsunami Wave Amplitude at the Okushiri Island: Mach Reflection of Shallow Water Waves. In *Proceedings of the 22nd International Symposium on Shock Waves, Vol. II. Ball, Hillier and Roberts(Editors)*, pages 1207–1212. University of Southampton, UK, 1999.
539. E. F. Toro and P. L. Roe. A Hybridised High-Order Random Choice Method for Quasi-Linear Hyperbolic Systems. In Grönig, editor, *Proc. 16th Intern. Symp. on Shock Tubes and Waves*, pages 701–708, Aachen, Germany, July 1987.
540. E. F. Toro and P. L. Roe. A Hybrid Scheme for the Euler Equations Using the Random Choice and Roe's Methods. In *Numerical Methods for Fluid Dynamics III. The Institute of Mathematics and its Applications Conference Series, New Series No. 17, Morton and Baines (Editors)*, pages 391–402. Oxford University Press, New York, 1988.
541. E. F. Toro, M. Spruce, and W. Speares. Restoration of the Contact Surface in the HLL-Riemann Solver. Technical Report CoA-9204, Department of Aerospace Science, College of Aeronautics, Cranfield Institute of Technology, UK, 1992.
542. E. F. Toro, M. Spruce, and W. Speares. Restoration of the Contact Surface in the HLL-Riemann Solver. *Shock Waves*, 4:25–34, 1994.
543. E. F. Toro and V. A. Titarev. Solution of the Generalised Riemann Problem for Advection-Reaction Equations. *Proc. Roy. Soc. London A*, 458:271–281, 2002.
544. E. F. Toro and V. A. Titarev. ADER Schemes for Scalar Hyperbolic Conservation Laws with Source Terms in Three Space Dimensions. *J. Comput. Phys.*, 202(1):196–215, 2005.
545. E. F. Toro (Editor). *Godunov Methods: Theory and Applications. Edited Review*. Kluwer Academic/Plenum Publishers, 2001.
546. I. Toumi. A Weak Formulation of Roe's Approximate Riemann Solver. *J. Comput. Phys.*, 102:360–373, 1992.
547. I. Toumi. An Upwind Numerical Method for Two-Fluid Two-Phase Flow Models. *Nuclear Science and Engineering*, 123:147–168, 1996.
548. I. Toumi and D. Caruge. An Implicit Second-Order Numerical Method for Three-Dimensional Two-Phase Flow Calculations. *Nuclear Science and Engineering*, 130:1–13, 1998.

549. I. Toumi and A. Kumbaro. An Approximate Linearized Riemann Solver for a Two-Fluid Model. *J. Comput. Phys.*, 124:286–300, 1996.
550. E. Turkel and A. Arnone. Pseudo-Compressibility Methods for the Incompressible Flow Equations. Technical Report ICASE 93–66, NASA Langley Research Center, USA, 1993.
551. A. Tveito and R. Winther. *Introduction to Partial Differential Equations*. Springer-Verlag, 1998.
552. G. D. van Albada, B. van Leer, and W. W. Roberts. A Comparative Study of Computational Methods in Cosmic Gas Dynamics. *Astron. Astrophysics*, 108:76–84, 1982.
553. J. J. W. van der Vegt and H. van der Ven. Space-Time Discontinuous Galerkin Finite Element Method with Dynamic Grid Motion for Inviscid Compressible Flows: General Formulation. *J. Comput. Phys.*, 182(2):546–585, 2002.
554. B. van Leer. Towards the Ultimate Conservative Difference Scheme I. The Quest for Monotonicity. *Lecture Notes in Physics*, 18:163–168, 1973.
555. B. van Leer. Towards the Ultimate Conservative Difference Scheme II. Monotonicity and Conservation Combined in a Second Order Scheme. *J. Comput. Phys.*, 14:361–370, 1974.
556. B. van Leer. MUSCL, A New Approach to Numerical Gas Dynamics. In *Computing in Plasma Physics and Astrophysics*, Max-Planck-Institut für Plasma Physik, Garching, Germany, April 1976.
557. B. van Leer. Towards the Ultimate Conservative Difference Scheme III. Upstream-Centered Finite Difference Schemes for Ideal Compressible Flow. *J. Comput. Phys.*, 23:263–275, 1977.
558. B. van Leer. Towards the Ultimate Conservative Difference Scheme IV. A New Approach to Numerical Convection. *J. Comput. Phys.*, 23:276–299, 1977.
559. B. van Leer. Towards the Ultimate Conservative Difference Scheme V. A Second Order Sequel to Godunov’s Method. *J. Comput. Phys.*, 32:101–136, 1979.
560. B. van Leer. Flux-Vector Splitting for the Euler Equations. Technical Report ICASE 82–30, NASA Langley Research Center, USA, 1982.
561. B. van Leer. Flux Vector Splitting for the Euler Equations. In *Proceedings of the 8th International Conference on Numerical Methods in Fluid Dynamics*, pages 507–512. Springer-Verlag, 1982.
562. B. van Leer. On the Relation Between the Upwind-Differencing Schemes of Godunov, Engquist-Osher and Roe. *SIAM J. Sci. Stat. Comput.*, 5(1):1–20, 1985.
563. B. van Leer. Progress in Multi-Dimensional Upwind Differencing. Technical Report CR-189708/ICASE 92–43, NASA, Sept. 1992.
564. B. van Leer. Upwind and High-Resolution Methods for Compressible Flow: from Donor Cell to Residual Distribution Schemes. Review Article. *Commun. Comput. Phys.*, 1(2):192–206, 2006.
565. B. van Leer, J. L. Thomas, and P. L. Roe. A comparison of numerical flux formulas for the euler and navier-stokes equations. AIAA 8th Computational Fluid Dynamics Conference, June 9–11, 1987, Honolulu, Hawaii. Paper AIAA–87–1104–CP, 1987.
566. M. H. P. M van Putten. A Numerical Implementation of MHD in Divergence Form. *J. Comput. Phys.*, 105:339–353, 1993.

567. M. O. Varner, W. R. Martindale, W. J. Phares, and J. C. Adams. Large Perturbation Flow Field Analysis and Simulation for Supersonic Inlets. Technical Report NASA CR-174676, NASA, USA, 1984.
568. M. E. Vázquez-Cendón. *Estudio de Esquemas Descentrados para su Aplicación a las Leyes de Conservación Hiperbólicas con Términos de Fuente*. PhD thesis, Departamento de Matemáticas Aplicadas, Universidad de Santiago de Compostela, España, 1994.
569. M. E. Vázquez-Cendón. Improved Treatment of Source Terms in Upwind Schemes for the Shallow Water Equations in Channels with Irregular Geometry. *J. Comput. Phys.*, 148:497–526, 1999.
570. M. Vinokur. An Analysis of Finite-Difference and Finite-Volume Formulations of Conservation Laws. *J. Comput. Phys.*, 81:1–52, 1989.
571. A. I. Volpert. The Space BV and Quasilinear Equations. *Math. USSR Sbornik*, 73:225–267, 1967.
572. L. Wang and D. J. Mavriplis. Implicit Solution of the Unsteady Euler Equations for High-Order Accurate Discontinuous Galerkin Discretizations. *J. Comput. Phys.*, 225(2):1994–2015, 2007.
573. S. Y. Wang, P. B. Buttler, and H. Krier. Non-Ideal Equations of State for Combusting and Detonating Explosives. *Prog. Energy Combust. Sci.*, 11:311–331, 1985.
574. R. F. Warming and R. W. Beam. Upwind Second Order Difference Schemes with Applications in Aerodynamic Flows. *AIAA Journal*, 24:1241–1249, 1976.
575. Z. U. A. Warsi. *Fluid Dynamics: Theoretical and Computational Approaches*. CRC Press Inc., 1992.
576. G. Watson, D. H. Peregrine, and E. F. Toro. Numerical Solution of the Shallow Water Equations on a Beach Using the Weighted Average Flux Method. In *Computational Fluid Dynamics 1992, Vol. 1*, pages 495–502. Elsevier, 1992.
577. N. P. Weatherill. Numerical Grid Generation. Lecture series 1990–06, von Karman Institute for Fluid Dynamics, 1990.
578. N. P. Weatherill. An Upwind Kinetic Flux Vector Splitting Method on General Mesh Topologies. *Intern. J. Numer. Meth. Engineering*, 37:623–643, 1994.
579. E. Weinan and C. W. Shu. A Numerical Resolution Study of High-Order Essentially Non-Oscillatory Schemes Applied to Incompressible Flow. Technical Report ICASE 92–39, NASA Langley Research Center, USA, 1992.
580. B. V. Wells, M. J. Baines, and P. Glaister. Generation of Arbitrary Lagrangian–Eulerian (ALE) Velocities, Based on Monitor Functions, for the Solution of Compressible Fluid Equations. *Int. J. Numer. Methods in Fluids*, 47:1375–1381, 2005.
581. P. Wesseling. *Principles of Computational Fluid Dynamics*. Springer, 2001.
582. G. B. Whitham. *Linear and Non-linear Waves*. John Wiley and Sons, 1974.
583. F. A. Williams. *Combustion Theory*. Benjamin Cummings Publishing Company, Inc Merlo Park, California, 1985.
584. P. Woodward and P. Colella. The Numerical Simulation of Two-Dimensional Fluid Flow with Strong Shocks. *J. Comput. Phys.*, 54:115–173, 1984.
585. K. Xu. Gas Kinetic Schemes for Unsteady Compressible Flow Simulations. VKI Fluid Dynamics Lecture Series, 1998–3, 1998.
586. K. Xu, L. Martinelli, and A. Jameson. Gas–Kinetic Finite Volume Methods, Flux–Vector Splitting and Artificial Diffusion. *J. Comput. Phys.*, 120:48–65, 1995.

587. K. Xu and K. H. Prendergast. Numerical Navier–Stokes Solutions from Gas–Kinetic Theory. *J. Comput. Phys.*, 114:9–17, 1994.
588. N. N. Yanenko. *The Method of Fractional Steps*. Springer–Verlag, New York, 1971.
589. G. Yang, D. M. Causon, D. M. Ingram, R. Saunders, and P. Batten. A Cartesian Cut Cell Method for Compressible Flows Part A: Static Body Problems. *Aeronautical Journal*, 101:47–56, 1997.
590. G. Yang, D. M. Causon, D. M. Ingram, R. Saunders, and P. Batten. A Cartesian Cut Cell Method for Compressible Flows Part B: Moving Body Problems. *Aeronautical Journal*, 101:57–65, 1997.
591. J. Y. Yang, M. H. Chen, I. N. Tsai, and J. W. Chang. A Kinetic Beam Scheme for Relativistic Gas Dynamics. *J. Comput. Phys.*, 136:19–40, 1997.
592. H. C. Yee. Construction of Explicit and Implicit Symmetric TVD Schemes and their Applications. *J. Comput. Phys.*, 68:151–179, 1987.
593. H. C. Yee. A Class of High–Resolution Explicit and Implicit Shock–Capturing Methods. Von Karman Institute of Fluid Dynamics, Lecture Series 1989–04, 1989.
594. H. C. Yee and P. K. Sweby. Non–Linear Dynamics and Numerical Uncertainties in CFD. Technical Report TM 110398, NASA Ames Research Center, Moffett Field, USA, 1996.
595. H. C. Yee, R. F. Warming, and A. Harten. Implicit Total Variation Diminishing (TVD) Schemes for Steady State Problems. *J. Comput. Phys.*, 57:327–360, 1985.
596. E. C. Zachmanoglou and D. W. Thoe. *Introduction to Partial Differential Equations*. Dover Publications, Inc. New York, 1986.
597. S. T. Zalesak. Fully Multidimensional Flux–Corrected Transport Algorithms for Fluids. *J. Comput. Phys.*, 31:335–362, 1979.
598. D. D. Zeeuw and K. G. Powell. An Adaptively Refined Cartesian Mesh Solver for the Euler Equations. *J. Comput. Phys.*, 104:56–68, 1993.
599. Y. B. Zeldovich and Y. P. Raizer. *Physics of Shock Waves and High Temperature Hydrodynamic Phenomena*. Academic Press, 1966.
600. G-C. Zha and E. Bilgen. Numerical Solution of Euler Equations by a New Flux Vector Splitting Scheme. *Int. J. Numer. Meth. Fluids*, 17:115–144, 1993.
601. T. Zhang and Y. Zheng. Conjecture on the Structure of Solutions of the Riemann Problem for Two Dimensional Gas Dynamics. *SIAM J. Math. Anal.*, 21(3):593–630, 1990.
602. H. W. Zheng, C. Shu, and Chew Y. T. An Object–Oriented and Quadrilateral–Mesh Based Solution Adaptive Algorithm for Compressible Multi–Fluid Flows. *J. Comput. Phys.*, 227(14):6895–6921, 2008.

---

# Index

- 123 problem, 129, 281  
adaptive scheme, 522  
ADER method, 655–659, 661, 663, 666, 667, 669, 671, 673, 674  
adiabatic exponent, 12  
advection, 18, 66  
amplification factor, 540, 559  
amplitude, 167  
AMR, 593  
approximate Riemann solver, 112  
arithmetic mean, 251  
artificial compressibility equations, 39, 40, 681  
artificial compressibility factor, 40  
artificial viscosity, 171, 487  
asymptotically stable system, 538  
Avogadro Number, 11  
backward difference, 165  
backward Euler method, 539  
Boltzmann approach, 265  
Boltzmann Constant, 11, 12  
boundary conditions, 42, 182, 222, 223  
boundary Riemann problem, 223  
Buckley–Leverett equation, 64  
bulk viscosity, 16  
Burgers’s equation, 30, 71  
caloric equation of state, 7  
calorically ideal gas, 7  
canonical equations of state, 8  
canonical form, 52  
Cartesian mesh, 599, 605, 609, 610, 618  
Castro–Toro  $GRP_K$  solver, 652  
Cauchy problem, 626, 627, 629, 633, 635  
Cauchy–Kowalewski, 628, 630–635, 637, 640, 643, 646, 647, 651, 652, 659, 660, 665, 667–669  
Cauchy–Riemann equations, 46  
cell average, 176  
cell centred methods, 176, 555  
cell Reynolds number, 487  
cells, 176  
centred rarefaction, 76  
centred scheme, 416, 597, 598, 601, 603, 604, 611, 612, 679–681  
CFL coefficient, 183, 221, 495  
CFL condition, 217, 218  
CFL number, 167, 168  
chain rule, 43  
characteristic curves, 47  
characteristic equation, 92, 93, 299  
characteristic field, 77, 95, 111  
characteristic form, 52  
characteristic limiting, 293, 508  
characteristic polynomial, 44, 90, 538  
characteristic speed, 48, 52, 66, 77  
characteristic variables, 51, 54, 188, 189  
characteristics, 533  
chemical species, 326  
chi-square statistics, 251  
CHIMERA, 593  
CIR scheme, 168, 171, 172, 177  
classical Riemann problem, 625, 629, 636–639, 642–644, 646, 647, 651–653

- clipping of extrema, 171, 460, 670  
 compact support, 63  
 complete Riemann solver, 316, 596, 598, 680  
 composite scheme, 599  
 compressive region, 70  
 computational space, 576  
 concave flux, 66  
 conservation laws, 3, 43, 61, 66  
 conservation of energy, 4, 24, 25  
 conservation of mass, 2, 4, 20, 21  
 conservation of momentum, 4, 21, 22  
 conservative scheme, 2, 175  
 conserved variables, 2, 43, 61  
 Consistency Condition, 175, 350, 417  
 consistency condition, 319–324, 601  
 contact discontinuity, 85, 96, 107, 118, 140  
 contact wave, 85  
 convective flux component, 279  
 convergence rates, 672, 673  
 convex flux, 66  
 Courant number, 167  
 Courant number coefficient, 183  
 covolume equation of state, 13, 223  
 covolume gases, 117, 143  
 data compatibility, 450  
 decoupled system, 52  
 density, 2  
 dependent variables, 42  
 detonation, 682, 683  
 detonation analogue, 78  
 detonation waves, 78  
 diagonal matrix, 106  
 diagonalisable matrix, 106  
 diagonalisable system, 51  
 diffusion, 18  
 diffusion number, 487  
 dimensional splitting, 543  
 discontinuous Galerkin finite elements, 673  
 discontinuous Galerkin method, 655, 657, 673, 674, 680, 681  
 dispersive equation, 174  
 dissociation, 16  
 Distribution Theory, 452  
 divergence operator, 2  
 domain of dependence, 60  
 domain of determinacy, 60  
 dot product, 2  
 Dumbser–Enaux–Toro  $GRPK$  solver, 652  
 eigenvalues, 44, 65, 92, 104, 110, 298, 538  
 elementary waves, 94  
 elliptic system, 45, 46  
 ENO, 439, 661, 662, 666, 669, 680  
 enthalpy, 8, 89  
 entropy, 5, 7, 93  
 entropy condition, 72, 96  
 entropy formulation, 94  
 entropy glitch, 227  
 entropy-violating shock, 73  
 equation of state, 67, 140  
 escape velocity, 140  
 Euler method, 539  
 expansive region, 70  
 explicit scheme, 167, 539  
 explosion, 586, 587, 590  
 far-field boundary conditions, 224  
 fictitious cell, 182, 222, 495  
 finite volume, 19, 176  
 finite volume method, 543, 655, 657–660, 665, 666, 668, 669, 673, 674, 679–681  
 flow parameter, 501, 552  
 Flux Difference Splitting, 265  
 flux function, 66, 175  
 flux vector, 43, 61  
 Flux Vector Splitting, 90, 265  
 FORCE flux, 598–601, 603, 605, 606, 608–610, 612, 613, 617, 620  
 forcing term, 486  
 forward difference, 164, 167  
 fractional step method, 535, 543, 544  
 free energy, 9  
 free surface, 33, 35  
 free-stream Mach number, 46  
 Fromm scheme, 174  
 front tracking schemes, 294  
 Gas Constant, 11  
 general equation of state, 682  
 Generalised Riemann Invariant, 81, 97, 98, 111, 122–124, 135, 141, 377,

- 385–387, 389, 390, 393, 395, 398, 399, 404
- generalized Riemann problem, 625, 627–630, 635, 640–643, 648, 649, 651–653, 655, 656, 659, 660, 663, 665, 667, 669, 673
- generalized solution, 63
- genuinely non-linear field, 77, 84, 85, 95, 107, 111, 382, 395
- geopotential, 37
- GFORCE flux, 602–604
- Godunov centred flux, 609, 611
- Godunov centred scheme, 443–445
- Godunov flux, 180, 181, 191, 219
- Godunov scheme, 216, 218, 597, 600, 603, 604, 612
- gradient operator, 2
- grid generation, 574
- grid speed, 167, 171
  
- Harten–Osher–Engquist–Chakravarthy *GRP<sub>K</sub>* solver, 651
- head of rarefaction, 75
- heat capacity, 9
- heat conduction, 18, 66
- heats of reaction, 78
- high resolution, 445
- high-order Riemann problem, 625
- high-resolution methods, 449
- HLL Riemann solver, 315, 320, 321, 329, 333
- HLLC flux, 322–327, 331–333, 335, 336
- HLLC Riemann solver, 315, 593
- HLLE Riemann solver, 315, 328, 329
- HLLEM Riemann solver, 329
- homentropic flow, 94
- homogeneity property, 89, 269, 270
- homogeneous equation, 533
- homogeneous system, 18, 26, 42, 531
- hyperbolic conservation laws, 2
- hyperbolic PDE, 55
- hyperbolic system, 45, 61, 91
- hyperbolicity in time, 105, 106
  
- ideal gases, 117
- implicit method, 539
- implosion, 586, 587, 589
- incipient cavitation case, 403
- incomplete Riemann solver, 316, 596, 598, 680
- incremental form of a scheme, 450
- independent variables, 42
- infinite resolution, 335
- inhomogeneous equation, 535, 536
- inhomogeneous problem, 486
- inhomogeneous system, 26, 78
- initial condition, 42
- initial-value problem, 47
- instability, 74, 221
- integral curves, 377, 381
- integral form, 62
- integration path, 381, 384, 385, 389
- intercell boundary, 177
- intercell flux, 180, 184
- internal energy, 3, 8, 88
- intersection point, 377, 384, 385, 389, 397, 403
- inverse of Jacobian matrix, 271
- inviscid Burgers's equation, 42, 64, 181
- ionisation, 16
- isentropic equations, 30
- isentropic flow, 94
- isentropic gas dynamics, 65
- isentropic law, 94
- isothermal compressibility, 9, 11
- isothermal equations, 30, 64
- isotropic medium, 15
  
- Jacobian matrix, 43, 61, 64, 88, 104, 106, 108, 270, 271, 275, 354, 392, 538
- Jacobian of the transformation, 577
- jump discontinuity, 55
  
- kinematic viscosity, 38
- kinetic energy, 3
  
- large-time step scheme, 424
- Lax Equivalence Theorem, 418
- Lax–Friedrichs flux, 185, 329, 599, 601–605, 609, 610
- Lax–Friedrichs scheme, 172, 184, 226, 244, 249
- Lax–Wendroff flux, 186, 599, 601–605, 609–611, 620
- Lax–Wendroff scheme, 173
- left eigenvector, 44, 92

- linear advection equation, 31, 33, 42, 47, 244  
 linear schemes, 415  
 linear system, 31, 42  
 linearly degenerate field, 77, 83, 95, 107, 111, 382, 395  
 long time evolution, 669, 671  
 low density, 227
- Mach number, 32, 101, 266, 277, 591–593  
 Mach reflection, 592, 593  
 Mach stem, 593  
 magnetohydrodynamics, 683  
 mass flux, 121, 124, 145, 146  
 material derivative, 19  
 mean-free path, 70  
 mesh function, 452  
 mesh generation, 574  
 method of lines, 439  
 metrics, 576  
 MHD equations, 683  
 modified equation, 171, 172, 249  
 mole, 11  
 momentum, 2  
 monatomic gases, 13  
 monotone flux, 680  
 monotone scheme, 172, 173, 226, 250, 521, 568  
 moving boundary, 223  
 multicomponent flow, 326  
 multidimensional flow, 326  
 multidimensional upwinding, 680  
 multiphase flow, 683–685  
 MUSTA approach, 604
- Navier–Stokes equations, 17, 680–682  
 Nessyahu–Tadmor scheme, 598, 599  
 Newton–Raphson, 127  
 Newtonian fluid, 15  
 non–conservative method, 174  
 non–conservative products, 685  
 non–conservative system, 674  
 non–convex flux, 67  
 non–linear scheme, 447  
 non–linear system, 42  
 non–reflecting boundary conditions, 224  
 normal velocity, 107, 550  
 NUMERICA, 334, 679, 686
- numerical dissipation, 333  
 numerical domain of dependence, 169, 171  
 numerical flux, 175, 177, 180, 626, 630, 636, 640, 642, 645, 646, 648, 651–653, 655, 658, 663, 665–669, 679–681  
 numerical source, 626, 628, 630, 633, 640, 642, 653, 655, 658–660, 663, 665, 666, 668, 669  
 numerical viscosity, 171, 172, 460, 602–604
- one–sided derivative, 164  
 one–sided differencing, 168, 188, 190  
 open–end boundary conditions, 224  
 ordinary differential equation, 47  
 oscillation–free scheme, 563
- parameter vector, 352, 355  
 particle velocity, 101, 102  
 passive scalar, 296, 317, 327, 333  
 path–conservative scheme, 684  
 phase angle, 167, 559  
 phase plane, 77  
 phase space, 77, 381  
 physical flux, 175  
 physical space, 576  
 physical variables, 2  
 Poisson equation, 39  
 polyatomic gases, 13  
 polytropic gas, 12  
 positively conservative solver, 329  
 PPM method, 680  
 Prandtl number, 17  
 pressure, 2  
 pressure positivity condition, 127, 143, 403  
 primitive variables, 2, 91, 109  
 primitive–variable formulation, 114
- quasi–linear system, 42
- radiation boundary conditions, 224  
 random sampling, 240, 241  
 range of influence, 61  
 Rankine–Hugoniot condition, 71, 79, 80, 99, 112, 123, 139, 321, 323, 593  
 rarefaction shock, 73, 75, 366

- rarefaction wave, 74, 76, 85, 96, 98, 107, 111, 118  
ratio of specific heats, 12, 13, 88  
reaction progress variable, 78  
reaction–diffusion equation, 673  
real gases, 14  
reconstruction, 626, 652, 656, 659–663, 666–669  
regular reflection, 592  
relative velocity, 99, 102, 120  
relativistic gas dynamics, 683  
Richtmyer scheme, 186, 226, 249  
Riemann approach, 265  
Riemann problem, 49, 55, 79, 83, 94, 116, 140, 177, 315, 317–320, 322, 327–329  
right eigenvector, 65  
right eigenvectors, 44, 92, 94, 104, 110, 298  
robustness, 226, 227, 282, 307, 334, 372, 405  
Roe matrix, 351, 353  
rotation matrix, 105, 109  
rotational invariance, 105, 108, 109  
Runge–Kutta method, 539  
Rusanov flux, 329
- scaling factor, 45, 92  
shallow water equations, 35, 681  
shear viscosity, 16, 38  
shear wave, 107, 111  
shock fitting, 174  
shock Mach number, 101, 103  
shock thickness, 71  
shock wave, 66, 70, 76, 84, 96, 98, 101, 107, 111, 112, 118, 174  
shock–tube problem, 95, 117  
signal velocities, 318  
similarity solution, 83, 95, 111, 215, 629, 636, 639, 641, 643, 644, 646  
slowly moving shock, 228, 405  
small perturbation equations, 32, 46  
smearing, 171  
Sod’s test problem, 129  
sonic boom, 70  
sonic flow, 32, 296, 333, 366, 377  
sonic point, 182, 281, 377, 382, 384, 385, 389, 395, 397  
sonic rarefaction, 118, 182, 296, 335, 366, 368, 369, 500  
sound speed, 10, 32, 44, 64, 65, 82, 83, 88, 93, 101, 266, 270  
source term, 23, 25, 47, 78, 486, 531, 684, 685  
species equation, 336  
specific internal energy, 5, 13  
split Riemann problem, 111, 112  
Split–Coefficient Matrix Scheme, 267  
spurious oscillations, 173, 226, 440, 445, 447, 501, 573  
stability, 167, 171  
stability condition, 168  
stability of ODEs, 538  
standard deviation, 251  
Star Region, 56, 58, 59, 80, 118, 129, 139, 144, 390  
start up errors, 593  
stationary contact, 282  
steady state, 40  
steady supersonic flow, 151  
stencil of a scheme, 169  
stiff ODEs, 538  
stiff source term, 685  
stiffness ratio, 538  
stream function, 38  
strict hyperbolicity, 45  
structured mesh, 597, 599, 605, 609  
subsonic flow, 32, 46  
substantial derivative, 19  
supersonic flow, 32, 46  
support of a scheme, 416  
Sutherland formula, 16
- tail of rarefaction, 75  
tangential velocity, 107, 111, 550  
tensor product, 4  
tensors, 4  
test functions, 63  
thermal conductivity, 17  
thermal equation of state, 6  
time level, 166, 177  
time step, 182, 183, 221  
Toro–Titarev  $GRP_K$  solver, 651  
total derivative, 67  
total energy, 2, 3  
traffic flow equation, 64, 66  
transonic rarefaction, 182

- transonic shock, 384
- transparent boundary conditions, 224
- Trapezoidal Method, 539, 540
- trial solution, 167
- true domain of dependence, 171
- truncation error, 417
- TVB schemes, 456
- TVD, 413–415, 425, 431, 438, 440, 451–458, 460, 462–464, 469–484, 486–488, 655, 661, 669–671, 674, 680, 682
- Universal Gas Constant, 11
- UNO schemes, 439
- unstructured mesh, 597, 599, 605, 617, 620, 674
- upwind differencing, 168
- upwind flux, 679–681, 685
- upwind scheme, 165, 175, 597, 598, 600, 603, 604, 612
- upwinding the source terms, 532
- vacuum, 116, 118, 127, 139, 142, 304
- vacuum condition, 143
- vacuum problem, 345
- van der Corput numbers, 250
- vector potential, 38
- velocity, 2
- viscosity, 33
- viscosity coefficient, 66
- viscous dissipation, 66
- viscous flux limiters, 487
- volume expansivity, 9, 11
- vorticity transport equation, 39
- vorticity vector, 38
- WAF method, 587, 593, 609, 669
- Warming–Beam method, 173
- wave, 49
- wave breaking, 70
- wave equation, 31
- wave jump, 58, 191
- wave length, 167
- wave number, 167
- wave speed estimate, 326–331, 335
- wave steepening, 66
- wave strength, 58
- wave–by–wave limiting, 508
- waves in solids, 682
- weak shock, 70
- weak solution, 63
- well–balanced scheme, 685
- WENO, 649, 661, 662, 669, 680
- x sweep, 546, 548
- y sweep, 546, 548
- z sweep, 548