

HANDBOOK OF NUMERICAL ANALYSIS

VOLUME

XVIII

Series Editors Q. Du, R. Glowinski, M. Hintermüller, E. Süli

Handbook of Numerical Methods for Hyperbolic Problems Applied and Modern Issues

Volume Editors Rémi Abgrall and Chi-Wang Shu

NORTH-HOLLAND

Handbook of Numerical Analysis

Volume 18

Handbook of Numerical Methods for Hyperbolic Problems

Applied and Modern Issues

Handbook of Numerical Analysis

Series Editors

Qiang Du

Columbia University, New York, United States of America

Roland Glowinski

University of Houston, Texas, United States of America

Michael Hintermüller

Humboldt University of Berlin, Germany

Endre Süli

University of Oxford, United Kingdom

Handbook of Numerical Analysis

Volume 18

Handbook of Numerical Methods for Hyperbolic Problems

Applied and Modern Issues

Edited by

Rémi Abgrall

Institut für Mathematik & Computational
Science, Universität Zürich,
Zürich, Switzerland

Chi-Wang Shu

Division of Applied Mathematics
Brown University
Providence, RI, United States



North-Holland

An imprint of Elsevier
elsevier.com

North-Holland is an imprint of Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom

Copyright © 2017 Elsevier B.V. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

ISBN: 978-0-444-63910-3

ISSN: 1570-8659

For information on all North-Holland publications
visit our website at <https://www.elsevier.com/>



Working together
to grow libraries in
developing countries

www.elsevier.com • www.bookaid.org

Publisher: Zoe Kruze

Acquisition Editor: Kirsten Shankland

Editorial Project Manager: Hannah Colford

Production Project Manager: Stalin Viswanathan

Cover Designer: Victoria Pearson

Typeset by SPi Global, India

Contents

Contributors	xv
Editors' Introduction	xvii

1. Cut Cells: Meshes and Solvers	1
<i>M. Berger</i>	
1 Introduction	1
2 Brief Early History	3
3 Mesh Generation	5
4 Data Structures and Implementation Issues	8
5 Finite Volume Methods for Cut Cells	10
5.1 Steady-State Solution Techniques	12
5.2 Explicit Time-dependent Solution Techniques	13
5.3 Viscous Flows	17
6 Conclusions	18
Acknowledgements	18
References	18
 2. Inverse Lax–Wendroff Procedure for Numerical Boundary Treatment of Hyperbolic Equations	 23
<i>C.-W. Shu and S. Tan</i>	
1 Introduction	24
2 Problem Description and Interior Schemes	27
3 Numerical Boundary Conditions for Static Geometry	28
3.1 One-Dimensional Scalar Conservation Laws: Smooth Solutions	28
3.2 One-Dimensional Scalar Conservation Laws: Solutions Containing Discontinuities	31
3.3 Two-Dimensional Euler Equations in Static Geometry	34
4 Moving Boundary Treatment for Compressible Inviscid Flows	38
5 Numerical Results	42
6 Conclusions and Future Work	48
References	49

3. Multidimensional Upwinding	53
<i>P. Roe</i>	
1 Introduction	53
2 Why Multidimensional Methods?	57
2.1 Dimensional Splitting and One-Dimensional Upwinding	58
3 Oblique Wave Methods	61
4 “Corner Transport” Methods	62
5 Edges or Corners?	64
6 When “Upwinding” Is Not Needed	65
7 Bicharacteristic Methods	67
8 Residual Distribution	69
8.1 The N Scheme	70
8.2 The NN Scheme	71
8.3 Systems of Equations	72
8.4 Unsteady Problems	72
8.5 Wave Models	73
8.6 Elliptic–Hyperbolic Splitting	73
9 The Poisson Formulas	75
9.1 Application to the Euler Equations	76
10 Concluding Remarks	78
References	78
4. Bound-Preserving High-Order Schemes	81
<i>Z. Xu and X. Zhang</i>	
1 Introduction	82
2 A Bound-Preserving Limiter for Approximation Polynomials	83
2.1 First-Order Monotone Schemes	83
2.2 The Weak Monotonicity in High-Order Finite Volume Schemes	84
2.3 A Simple and Efficient Scaling Limiter	86
2.4 SSP High-Order Time Discretizations	91
2.5 Extensions and Applications	92
3 Bound-Preserving Flux Limiters	93
3.1 Basic Idea and Framework	93
3.2 Decoupling for the Flux Limiting Parameters	95
4 Concluding Remarks	97
Acknowledgements	98
References	98
5. Asymptotic-Preserving Schemes for Multiscale Hyperbolic and Kinetic Equations	103
<i>J. Hu, S. Jin, and Q. Li</i>	
1 Introduction	104
2 Basic Design Principles of AP Schemes—Two Illustrative Examples	105

2.1	The Jin–Xin Relaxation Model	105
2.2	The BGK Model	107
3	AP Schemes for General Hyperbolic and Kinetic Equations	110
3.1	AP Schemes Based on Penalization	111
3.2	AP Schemes Based on Exponential Reformulation	115
3.3	AP Schemes Based on Micro–Macro Decomposition	117
4	Other Asymptotic Limits and AP Schemes	118
4.1	Diffusion Limit of Linear Transport Equation	118
4.2	High-Field Limit	119
4.3	Quasi-Neutral Limit in Plasmas	120
4.4	Low Mach Number Limit of Compressible Flows	121
4.5	Stochastic AP Schemes	122
5	Conclusion	123
	Acknowledgements	123
	References	124
6.	Well-Balanced Schemes and Path-Conservative Numerical Methods	131
	<i>M.J. Castro, T. Morales de Luna, and C. Parés</i>	
1	Introduction	132
2	Path-Conservative Numerical Schemes	138
3	Some Families of Path-Conservative Numerical Schemes	142
3.1	Godunov Method	142
3.2	Simple Riemann Solvers	142
3.3	Roe Methods	144
3.4	Functional Viscosity Matrix Methods	145
3.5	Other Path-Conservative Methods	148
4	High-Order Schemes Based on Reconstruction of States	148
5	Well-Balanced Schemes	151
5.1	Well-Balanced Property for SRSs	154
5.2	Well-Balanced HLL Scheme	155
5.3	Well-Balanced SRSs	156
5.4	Roe Method	156
5.5	Well-Balanced Functional Viscosity Matrix Methods	157
5.6	Generalized Hydrostatic Reconstruction	158
5.7	Well-Balanced Methods for a Subset of Stationary Solutions	160
5.8	High-Order Well-Balanced Schemes	161
6	Convergence and Choice of Paths	166
	Acknowledgements	169
	References	169
7.	A Practical Guide to Deterministic Particle Methods	177
	<i>A. Chertock</i>	
1	Introduction	178
2	Description of the Particle Method	181

2.1	Particle Approximation of the Initial Data	182
2.2	Time Evolution of Particles	183
2.3	Particle Function Approximations	186
3	Remeshing for Particle Distortion	190
3.1	Particle Weights Redistribution	191
3.2	Particle Merger—A Local Redistribution Technique	193
4	Applications to Convection–Diffusion Equations	194
4.1	Particle Methods for Convection–Diffusion Equations	195
	Acknowledgements	198
	References	198
8.	On the Behaviour of Upwind Schemes in the Low Mach Number Limit: A Review	203
	<i>H. Guillard and B. Nkonga</i>	
1	Introduction	204
2	The Multiple Low Mach Number Limits of the Compressible Euler Equations	205
2.1	Incompressible Limit	205
2.2	Acoustic Limit	207
2.3	Acoustic–Incompressible Interactions	208
2.4	Finite Volume Schemes	213
2.5	The Diagnosis	215
2.6	The Remedies	217
3	Numerical Illustrations	221
3.1	Order 1: Quadrangular Cartesian Grids	221
3.2	Order 1: Vertex-Centred Triangular Meshes	223
3.3	Order 1: Cell-Centred Triangular Meshes	224
3.4	Order 2: Vertex-Centred Triangular Meshes	226
4	Conclusion	228
	References	228
9.	Adjoint Error Estimation and Adaptivity for Hyperbolic Problems	233
	<i>P. Houston</i>	
1	Introduction	234
2	Error Representation for Linear Problems	235
2.1	Abstract Framework	236
2.2	Stabilized FEMs for the Linear Transport Equation	240
3	A Posteriori Error Estimation	243
4	Nonlinear Hyperbolic Conservation Laws	247
5	Practical Implementation and Adaptive Mesh Refinement	251
5.1	Numerical Approximation of the Dual Problem	251
5.2	Adaptive Mesh Refinement	252
5.3	Bibliographical Comments	253
6	Applications	254
6.1	Inviscid Flow Around a BAC3-11 Airfoil	254

6.2	Criticality Problems	255
6.3	Bifurcation Problems	255
7	Concluding Remarks and Outlook	256
	Acknowledgements	258
	References	258
10.	Unstructured Mesh Generation and Adaptation	263
	<i>A. Loseille</i>	
1	Introduction	264
1.1	Outline	266
2	An Introduction to Unstructured Mesh Generation	266
2.1	Surface Mesh Generation	266
2.2	Volume Mesh Generation	267
3	Metric-Based Mesh Adaptation	269
3.1	Metric Tensors in Mesh Adaptation	271
3.2	Techniques for Enhancing Robustness and Performance	272
3.3	Metric-Based Error Estimates	274
3.4	Controlling the Interpolation Error	276
3.5	Geometric Estimate for Surfaces	277
3.6	Boundary Layers Metric	278
4	Algorithms for Generating Anisotropic Meshes	280
4.1	Insertion and Collapse	280
4.2	Optimizations and Enhancement for Unsteady Simulations	282
5	Adaptive Algorithm and Numerical Illustrations	283
5.1	Adaptive Loop	284
5.2	A Wing-Body Configuration	285
5.3	Transonic Flow Around a M6 Wing	286
5.4	Direct Sonic Boom Simulation	288
5.5	Boundary Layer Shock Interaction	290
5.6	Double Mach Reflection and Blast Prediction	294
6	Conclusion	296
	References	297
11.	The Design of Steady State Schemes for Computational Aerodynamics	303
	<i>F.D. Witherden, A. Jameson, and D.W. Zingg</i>	
1	Introduction	304
2	Equations of Gas Dynamics and Spatial Discretizations	305
3	Time-Marching Methods	308
3.1	Model Problem for Stability Analysis of Convection Dominated Problems	308
3.2	Multistage Schemes for Steady State Problems	309
3.3	Implicit Schemes for Steady State Problems	312
3.4	Acceleration Methods	318
3.5	Multigrid Methods	323
3.6	RANS Equations	327

4	Newton–Krylov Methods	332
4.1	Background	332
4.2	Methodology	335
5	Conclusions	344
	References	345
12.	Some Failures of Riemann Solvers	351
	<i>R. Abgrall</i>	
1	Introduction	351
2	Real Gas Effects	353
2.1	Mixture of Gases	353
2.2	Nonlinear Equation of State	355
3	Multidimensional Effects	356
4	Accuracy Effects	358
	References	360
13.	Numerical Methods for the Nonlinear Shallow Water Equations	361
	<i>Y. Xing</i>	
1	Overview	362
2	Mathematical Model	363
3	Numerical Methods	364
3.1	Numerical Methods for the Homogeneous Equations	364
3.2	Well-Balanced Methods	368
3.3	Positivity-Preserving Methods	374
4	Shallow Water-Related Models	377
4.1	Shallow Water Flows Through Channels With Irregular Geometry	377
4.2	Shallow Water Equations on the Sphere	378
4.3	Two-Layer Shallow Water Equations	379
5	Conclusion Remarks	380
	Acknowledgements	380
	References	380
14.	Maxwell and Magnetohydrodynamic Equations	385
	<i>C.-D. Munz and E. Sonnendrücker</i>	
1	Introduction	386
2	Maxwell's Equations	386
2.1	The Model	386
2.2	Generalised Maxwell's Equations: Correcting the Fields	387
2.3	Cartesian Grids: Finite Difference and Spectral Methods	389
2.4	FV Schemes	390
2.5	Discontinuous Galerkin Schemes	393
2.6	Finite Element Methods	394

3	Magnetohydrodynamics	396
3.1	The Model	396
3.2	Discretization	397
4	Conclusion	398
	References	398
15.	Deterministic Solvers for Nonlinear Collisional Kinetic Flows: A Conservative Spectral Scheme for Boltzmann Type Flows	403
	<i>I.M. Gamba</i>	
1	Introduction	404
1.1	Kinetic Evolution Models	404
1.2	Binary Collisional Models and Double Mixing Convolution Forms	405
1.3	Classical Elastic Collisional Transport Theory: The Boltzmann Equation	408
1.4	Deterministic Solvers for Integral Equations of Boltzmann Type	409
2	The Landau and Boltzmann Operators Relation Through Their Double Mixing Convolutional Forms	410
2.1	The Grazing Collision Limit	413
3	A Conservative Spectral Method for the Collisional Form	415
3.1	Choosing a Computational Cut-Off Domain Ω_L	416
3.2	Fourier Series, Projections and Extensions	418
3.3	A Conservative Spectral Method for the Homogeneous Boltzmann Equation	419
3.4	Conservation Method—An Extended Isoperimetric Problem	421
3.5	Discrete in Time Conservation Method: Lagrange Multiplier Method	425
4	Local Existence, Convergence and Regularity for the Semidiscrete Scheme	426
4.1	Local Existence	427
4.2	Uniform Propagation of Numerical Unconserved Moments	428
4.3	Uniform L_k^2 Integrability Propagation	429
4.4	Uniform Semidiscrete H_k Sobolev Regularity Propagation	430
5	Final Comments and Conclusions	431
	Acknowledgements	432
	References	432
16.	Numerical Methods for Hyperbolic Nets and Networks	435
	<i>S. Čanić, M.L. Delle Monache, B. Piccoli, J.-M. Qiu, and J. Tambača</i>	
1	Introduction	436
2	Examples of Nets and Networks	437

2.1	Examples of Hyperbolic Nets	438
2.2	Examples of Hyperbolic Networks	446
3	Numerics for Nets and Networks	450
3.1	Finite Volume Methods	451
3.2	Discontinuous Galerkin Methods	451
	References	460
17.	Numerical Methods for Astrophysics	465
	<i>C. Klingenberg</i>	
1	Introduction	466
2	Astrophysical Scales for Astrophysical Phenomena	466
2.1	Spatial Scales	466
2.2	Density and Temporal Scales	466
3	Equations Used in Astrophysical Modelling	467
3.1	Source Terms	468
3.2	Additional Force Terms	468
3.3	Equation of State	469
4	Numerical Methods	469
4.1	Finite Difference Methods	469
4.2	Finite Volume Methods	470
4.3	Discontinuous Galerkin Method	471
4.4	N-Body Method	472
4.5	Grid-Free Method: Smoothed Particle Hydrodynamics	472
5	High-Performance Computing	473
6	Astrophysical Codes	473
7	Conclusion	475
	Acknowledgement	475
	References	475
18.	Numerical Methods for Conservation Laws With Discontinuous Coefficients	479
	<i>S. Mishra</i>	
1	Introduction	480
1.1	Conservation Laws With Coefficients	481
2	Motivating Examples	481
2.1	Multiphase Flows in Porous Media	481
2.2	Traffic Flow	482
2.3	Other Examples of Scalar Conservation Laws With Discontinuous Flux	483
2.4	Wave Propagation in Heterogeneous Media	484
2.5	Systems of Conservation Laws With Singular Source Terms	484
2.6	Flows as Perturbations of Discontinuous Steady States	485
3	A Brief Review of Available Theoretical Results	485

4 Numerical Schemes	491
4.1 Aligned Schemes	491
4.2 Staggered Schemes	493
4.3 Higher-Order Schemes	494
4.4 Extensions and Other Approaches	495
5 Numerical Experiments	496
5.1 Numerical Experiment 1	496
5.2 Numerical Experiment 2	497
5.3 Numerical Experiment 3	498
6 Summary and Open Problems	500
Acknowledgement	502
References	502
 19. Uncertainty Quantification for Hyperbolic Systems of Conservation Laws	 507
<i>R. Abgrall and S. Mishra</i>	
1 Introduction	508
1.1 Numerical Methods	509
1.2 Uncertainty Quantification	510
2 Random Fields and Random Entropy Solutions	511
2.1 Modelling of Random Inputs	512
2.2 Random Entropy Solutions	514
3 sG Method for UQ	515
3.1 Generalized Polynomial Chaos	516
3.2 Standard sG Method	517
3.3 gPC Expansions in the Entropy Variables	518
4 Stochastic Collocation Methods	520
4.1 Standard Stochastic Collocation Method	520
4.2 Stochastic Finite Volume Methods	521
5 Monte Carlo and Multilevel Monte Carlo Methods	524
5.1 Monte Carlo Method	524
5.2 Multilevel Monte Carlo Finite Volume Method	526
6 Numerical Experiments	529
6.1 Compressible Euler Equations	529
6.2 Uncertain Orszag–Tang Vortex	530
6.3 UQ for the Lituya Bay Mega-Tsunami	533
6.4 A Random Kelvin–Helmholtz Problem	534
7 Measure-Valued and Statistical Solutions	536
8 Conclusion and Perspectives	538
Acknowledgements	540
References	540
 20. Multiscale Methods for Wave Problems in Heterogeneous Media	 545
<i>A. Abdulle and P. Henning</i>	
1 Introduction	546

2 Numerical Methods for the Wave Equation in Heterogeneous Media Without Scale Separation	549
2.1 Approach 1—Harmonic Coordinate Transformations	551
2.2 Approach 2—MsFEM Using Limited Global Information	553
2.3 Approach 3—Flux-Transfer Transformations	555
2.4 Approach 4—Localized Orthogonal Decomposition	558
2.5 The Case of General Initial Values: G-Convergence and Perturbation Arguments	561
3 Numerical Methods for the Wave Equation in Heterogeneous Media With Scale Separation	563
3.1 Effective Model and Numerical Homogenization Method for Short-Time Wave Propagation	564
3.2 Effective Model and Numerical Homogenization Method for Long-Time Wave Propagation	569
Acknowledgement	574
References	574

Index	577
--------------	------------

Contributors

Numbers in Parentheses indicate the pages on which the author's contributions begin.

- A. Abdulle** (545), ANMC, Section de Mathématiques, École polytechnique fédérale de Lausanne, Lausanne, Switzerland
- R. Abgrall** (351, 507), Institute of Mathematics, University of Zürich, Zürich, Switzerland
- S. Čanić** (435), University of Houston, Houston, TX, United States
- M. Berger** (1), Courant Institute, New York University, New York, NY, United States
- M.J. Castro** (131), Universidad de Málaga, Málaga, Spain
- A. Chertock** (177), North Carolina State University, Raleigh, NC, United States
- M.L. Delle Monache** (435), Rutgers University—Camden, Camden, NJ, United States
- I.M. Gamba** (403), The University of Texas at Austin, Austin, TX, United States
- H. Guillard** (203), Université Côte d'Azur, Inria, CNRS, LJAD, France
- P. Henning** (545), KTH Royal Institute of Technology, Stockholm, Sweden
- P. Houston** (233), School of Mathematical Sciences, University of Nottingham, Nottingham, United Kingdom
- J. Hu** (103), Purdue University, West Lafayette, IN, United States
- A. Jameson** (303), Stanford University, Stanford, CA, United States
- S. Jin** (103), University of Wisconsin-Madison, Madison, WI, United States
- C. Klingenberg** (465), Institut für Mathematik, Universität Würzburg, Würzburg, Germany
- Q. Li** (103), University of Wisconsin-Madison, Madison, WI, United States
- A. Loseille** (263), INRIA Saclay-Ile de France, France
- S. Mishra** (479, 507), Seminar for Applied Mathematics, ETH Zürich, Zürich, Switzerland
- T. Morales de Luna** (131), Universidad de Córdoba, Córdoba, Spain
- C.-D. Munz** (385), Institute for Aerodynamics and Gas Dynamics, University of Stuttgart, Stuttgart, Germany
- B. Nkonga** (203), Université Côte d'Azur, Inria, CNRS, LJAD, France
- C. Parés** (131), Universidad de Málaga, Málaga, Spain
- B. Piccoli** (435), Rutgers University—Camden, Camden, NJ, United States

- J.-M. Qiu** (435), University of Houston, Houston, TX, United States
- P. Roe** (53), University of Michigan, Ann Arbor, MI, United States
- C.-W. Shu** (23), Brown University, Providence, RI, United States
- E. Sonnendrücker** (385), Max-Planck Institute for Plasma Physics; Mathematics Center, TU Munich, Garching, Germany
- J. Tambača** (435), University of Zagreb, Zagreb, Croatia
- S. Tan** (23), Brown University, Providence, RI, United States
- F.D. Witherden** (303), Stanford University, Stanford, CA, United States
- Y. Xing** (361), University of California Riverside, Riverside, CA, United States
- Z. Xu** (81), Michigan Technological University, Houghton, MI, United States
- X. Zhang** (81), Purdue University, West Lafayette, IN, United States
- D.W. Zingg** (303), University of Toronto Institute for Aerospace Studies, Toronto, ON, Canada

Editors' Introduction

R. Abgrall* and C.-W. Shu[†]

**Institut für Mathematik, Universität Zürich, Zürich, Switzerland*

[†]Brown University, Providence, RI, United States

These two volumes represent the volumes 17 and 18 of Handbook of Numerical Analysis. It is entirely devoted to the numerical methods designed for approximating the solution of hyperbolic equations, or of equations that write as a sum of operators where the most important, in terms of the behaviour of the solution, is the hyperbolic one. An example is the Navier–Stokes equations with high Reynolds number where the solution behaviour is essentially dictated by the hyperbolic operator (here the Euler system), except in boundary layers because of the boundary conditions.

Hyperbolic partial differential equations appear often in applications. The most important application, already mentioned, is fluid dynamics, including specific flows such as multiphase flows, magnetohydrodynamics, water waves, etc. Other application areas include Maxwell equations, kinetic equations, traffic flow models and networks, etc. The solutions of hyperbolic partial differential equations often involve discontinuities, making mathematical analysis and numerical simulations difficult. In the past few decades there has been a large amount of literature in the design, analysis and application of various numerical algorithms for solving hyperbolic equations. The current volumes attempt to have experts in different types of algorithms write concise summaries so that the readers can find a variety of algorithms under different situations and become familiar with their relative advantages and limitations.

This is a formidable task. We had to make choices because the field has grown tremendously since the early ages dating back to von Neumann in the United States and researchers from the former Soviet Union such as Rusanov and Godunov. This field has grown up for various reasons. The demand on diverse high tech areas ranging from airplanes and rockets, to the nuclear and car industries as well as more recently the green industry, to name just a few, necessitates to master better and better tools to improve performance. If it was possible in the early ages to rely on analytical solutions and experimental facilities only, this is no longer the case because of various constraints: economical, technological (weight, etc.), energy consumption, etc. This evolution has needed improved algorithms, i.e., more and more

accurate as well as more and more robust ones. Hence the research on algorithm has grown up and then exploded since the early 1970s.

In parallel, and also triggered by the same needs, computers have been more and more powerful from scalar, to vectors, then parallel and now massively parallel and hybrid architectures. This evolution of technology has also had a strong impact on the algorithms development.

Because of its success, it is now possible to compute more and more complicated problems, both in terms of geometry and physics.

There is still a lot to do to improve and understand the numerical methods designed for hyperbolic problems. The aim of these two volumes is to give a picture of the current state of the art.

In order to introduce the subject, we have asked Professor Dafermos from Brown University to provide a short summary on the theory of hyperbolic equations. Then, if one looks at the table of content, one would realize that we have tried to cover not only the classical topics, such as the finite volume method and the Riemann solvers that are the building blocks of many of the algorithms, but also less standard methods. Examples include algorithms for computing sharp transition propagated by linearly degenerate waves. Other examples are given by the ENO/WENO family. In that case we have tried to go over the classical description, by giving some analysis of the methods. Other high-order methods are also considered such as the discontinuous Galerkin (DG) ones, the more recent hybrid DG schemes, high-order finite element methods, front-tracking methods, methods for Lagrangian hydrodynamics, entropy stable schemes, etc. Time discretisation is also considered, as well as more specialized problems like the simulation of flows with low Mach numbers, level set techniques, numerical methods for Hamilton–Jacobi equations, etc.

Unfortunately, it is not possible, even in two quite thick volumes, to provide an exhaustive coverage of the state of the art. Even though the table of content seems to be exhaustive, many topics are still missing. For example, we have chosen to be quite restrictive on the subject of time stepping: there is no coverage on ADER and IMEX methods. The handling of problems with source terms is touched by [chapters 5 and 6](#) (well-balanced schemes and asymptotic-preserving schemes), but there is no direct coverage on stiff source terms. If we have a chapter on methods for Cartesian meshes, there is no direct coverage on the application of immersed boundary methods. Similarly we have chosen to consider the problem of meshing in a specific way; there is no direct coverage on adaptive mesh refinement (AMR). The problem on boundary conditions is considered in [chapter 2](#) this volume (volume 18) and chapter 19 previous volume (volume 17) of the handbook (SAT-SPB schemes and inverse Lax–Wendroff procedure), but much more could have been said. It was simply impossible to cover the whole field, and we apologize for this.

To end this introduction, we would like to thank all the contributors to these volumes, as well as the referees. Both have been extremely efficient.

ACKNOWLEDGEMENTS

R.A. has been supported in part by SNF grant # 200021_153604. C.-W.S. has been supported in part by NSF grant DMS-1418750.

Chapter 15

Deterministic Solvers for Nonlinear Collisional Kinetic Flows: A Conservative Spectral Scheme for Boltzmann Type Flows

I.M. Gamba

The University of Texas at Austin, Austin, TX, United States

Chapter Outline

1 Introduction	404	3.3 A Conservative Spectral Method for the Homogeneous Boltzmann Equation	419
1.1 Kinetic Evolution Models	404	3.4 Conservation Method—An Extended Isoperimetric Problem	421
1.2 Binary Collisional Models and Double Mixing Convolution Forms	405	3.5 Discrete in Time Conservation Method: Lagrange Multiplier Method	425
1.3 Classical Elastic Collisional Transport Theory: The Boltzmann Equation	408	4 Local Existence, Convergence and Regularity for the Semidiscrete Scheme	426
1.4 Deterministic Solvers for Integral Equations of Boltzmann Type	409	4.1 Local Existence	427
2 The Landau and Boltzmann Operators Relation Through Their Double Mixing Convolutional Forms	410	4.2 Uniform Propagation of Numerical Unconserved Moments	428
2.1 The Grazing Collision Limit	413	4.3 Uniform L_k^2 Integrability Propagation	429
3 A Conservative Spectral Method for the Collisional Form	415	4.4 Uniform Semidiscrete H_k Sobolev Regularity Propagation	430
3.1 Choosing a Computational Cut-Off Domain Ω_L	416		
3.2 Fourier Series, Projections and Extensions	418		

5 Final Comments and	Acknowledgements	432
Conclusions	References	432

ABSTRACT

We present an overview of deterministic solvers for the Boltzmann and Landau equations inspired by their Fourier space representation as weighted convolutional forms, where the later can be obtained as a grazing collision limit of the former. This presentation offers an introduction to the area and elaborates on recent results for conservative spectral Lagrangian schemes applied to several applications ranging from homogeneous flows for Coulomb potentials given by the Landau equation by an approximating of a corresponding Boltzmann model with grazing transition scattering rates, to a full conservative approach for Vlasov–Poisson–Landau system for electron–ion dynamics. This conservative method is enforced by a Lagrangian constrained optimization problem that and conservation correction estimates that give place to semidiscrete error estimates and long-time convergence to statistical equilibrium states given by *Maxwellians* distributions.

Keywords: Nonlinear integral equations, Rarefied gas flows, Boltzmann and Landau Fokker Plank equations, Deterministic numerical approximations to kinetic equations, Conservative spectral methods

2010 MSC: 45E99, 35A22, 76X05, 76P05, 82C05, 65C20, 65C30

1 INTRODUCTION

1.1 Kinetic Evolution Models

The numerical solutions of kinetic evolution transport given by integral equations of Boltzmann type needs the underlying understanding of the problem to be approximated: the evolution of a probability density function usually described by a Hamiltonian particle transport encountering interactions. Thus, before we discuss different aspects of deterministic solvers for such models, we introduced the basic notions associated to kinetic transport evolution models.

Our starting point is to recall that complex particle model systems with exchangeability properties yield the propagation of chaos property, that is, the particle system can be models by the evolution of independent and identically distributed (iid) continuous random variables or probability density measures. These models appear in many contexts of classical and quantum statistical physics, and more recently in novel applications to social dynamics of multiagent interactions defined by some multiplicatively interacting stochastic processes. The models we shall be considered share is a unified general framework of material transport dynamics can be derived from the so-called *Master Equations* derived for the time evolution of a particle system modelled as being in exactly one of the countable number of admissible states at any given time, where switching between states are treated probabilistically when interactions occur. Such evolution is described by associating a

probability density of states through discrete or continuous random variables. Interactions may be of “mean field” type when macroscopic forces depending on, either particle distribution averages, or particles of the same kind usually refer as *collisions*. Typical examples of “mean field” type interactions result in the so-called *collisionless* systems of transport models such as Vlasov–Poisson or Vlasov–Maxwell systems for plasma dynamics of charged particles.

However, when interactions due to *collisions* occur and the “switching” of states are described by a time-independent operator, the model represents a *kinetic evolution* and the process is *Markovian*. Such process may also include *birth and death* rates, meaning that probability density is injected in (birth) or taken away from (death) the system, and so the process is not in equilibrium. Examples of such models are the transport dynamics of classical kinetic collisional transport given by Boltzmann or Landau–Fokker–Plank type equations that may include mean field interactions to obtain a Vlasov–Poisson or Vlasov–Maxwell collisional plasma transport system (Chapman and Cowling, 1970; Graham and Méléard, 1999).

1.2 Binary Collisional Models and Double Mixing Convolution Forms

Rigorous justification of the propagation of chaos or *Stosszahlansatz* relies on contemporary ergodic theory and related areas in probability theory (Chapman and Cowling, 1970; Pulvirenti et al., 2014). We assume here its validity, implying that the system of N -particle interactions can be reduced to a closed form involving products of a single point probability density function (*pdf*) denoted by $f(x, v, t)$ solving a nonlocal, linear or multilinear structure in state space defined by $v \in \mathbb{R}^n$. In the case of binary interactions such *pdf* satisfies the following nonlocal weak form

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^d} f(x, v, t) \varphi(v) dv &:= \int_{\mathbb{R}^d} Q(f, f)(x, v, t) \varphi(v) dv \\ &= \kappa \int_{\mathbb{R}^d \times \mathbb{R}^d} f(x, v, t) f(x, v - u, t) \left(\int_{\Omega} (\varphi(v') - \varphi(v)) B(u, \sigma) d\sigma \right) du dv \end{aligned} \quad (1)$$

where the $u = v - v_*$ is the relative position for any interacting states pairs (v, v_*) changing into (v', v'_*) , for v fixed and v' determined by an *interaction law* with $v' = v'(v, v_*, \sigma)$ and $v'_* = v'_*(v, v_*, \sigma)$, with arbitrary $\sigma \in \Omega$ a manifold that determines the postcollisional relative position $u' = v' - v'_*$, for an arbitrary state $v_* \in \mathbb{R}^n$. Such manifold is the sphere S^{n-1} when the interaction correspond to particles characterized by indistinguishable spheres interacting by conserving centre of mass and local energies. The parameter κ quantifies the scaled mean free path in between interactions, and it is assumed to be or order of unity in rarefied regimes. The material derivative d/dt refers to the Lagrangian formulation of Hamiltonian dynamics of mixing (x, v) states, as in classical to particle plasma physics, and they are viewed as the divergence-free dynamics of

space-momentum mixing. The corresponding interacting pairs transition probability rates from (v, v_*) to (v', v'_*) are quantified by the collision kernel $B(u, \sigma)$. This equation is nonlocal and linear in the case when $f(x, v - u, t)$ is replaced by a known probability density. Thus, we define the *Kac Master equation* formulation as a *double mixing convolution* structure

$$\frac{d}{dt} \int_{\mathbb{R}^N} f(x, v, t) \varphi(v) dv = \int_{\mathbb{R}^d} Q(f, f)(x, v, t) \varphi(v) dv \quad (2)$$

$$= \kappa \int_{\mathbb{R}^d \times \mathbb{R}^d} f(x, v, t) f(x, v - u, t) G_{\varphi, B}(u, v) du dv. \quad (3)$$

The weight function $G_{\varphi, B}(u, v)$ is a mixing form of premixing and postmixing positions in v -space that depends only on the state variable v and its relative position state u . This weight function models the σ -average of the interaction on the Ω manifold, i.e.

$$G_{\varphi, B}(u, v) = \int_{\Omega} (\varphi(v') - \varphi(v)) B(v, u, \sigma) d\sigma, \quad (4)$$

that depends on the σ -average of the test function φ on the exchange law of states multiplied to the transition probability interaction rates $B(u, \sigma)$. These weight functions $G_{\varphi, B}(u, v)$ are often nonlinear and encode most of the information about, not only, the dynamics of interactions but also the regularity of the solution to such equation as much the decay rate to equilibrium states. More precisely,

- (1) The interaction law $v' = v'(v, v_*, \sigma)$; $v'_* = v'_*(v, v_*, \sigma)$, microreversible or not, determines the space of *collision invariants*: all those $\varphi(v)$ that nullify the weight function $G_{\varphi, B}(u, v)$. These collision invariants select the properties of the stationary states.
- (2) Propagation of chaos assumption and time irreversibility: *decorrelation before* the next interaction is encoded in the difference $\varphi(v') - \varphi(v)$ of the weight function and presets stability for the flow. In particular, when the interaction is microreversible and the transition probability rate function $B(u, \sigma)$ has symmetric properties consistent with such microreversibility, then setting $\varphi(v) = \log f(v)$, the monotonicity of the logarithmic function yields the *H-theorem* (Cercignani et al., 1994)

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^N} f(x, v, t) \log(v) dv &= \int_{\mathbb{R}^d} Q(f, f)(x, v, t) \log(v) dv \\ &= \int_{\mathbb{R}^d \times \mathbb{R}^d} f(x, v, t) f(x, v - u, t) \left[\int_{\Omega} (\log(v') - \log(v)) B(v, u, \sigma) d\sigma \right] du dv \leq 0. \end{aligned} \quad (5)$$

In the particular case of elastic interactions conserving centre of mass and local energy, both items (1) and (2) imply that the only stationary state is a Gaussian distribution in v -space, called the equilibrium Maxwellian

distribution, defined by the moments corresponding to the collision invariants associated with the initial data $f_0(v) \geq 0$ for a.e. $v \in \mathbb{R}^d$ and $\int_{\mathbb{R}^d} f_0(v)(1 + |v|^2) dv < \infty$. In the limit as $t \rightarrow +\infty$, we expect that $f(v, t)$ converges to the *equilibrium Maxwellian* distribution, i.e.

$$f(t, v) \rightarrow M[m_0, u_0, T_0](v) := m_0(2\pi T_0)^{-d/2} \exp\left(-\frac{|v - u_0|^2}{2T_0}\right), \quad (6)$$

where, if $m_0 > 0$ is the density mass, and the *moments or observables* are defined by

$$m_0 := \int_{\mathbb{R}^d} f_0(v) dv, \quad u_0 := \frac{1}{m_0} \int_{\mathbb{R}^d} f_0(v) dv, \quad T_0 := (dm_0)^{-1} \int_{\mathbb{R}^d} |v - u_0|^2 f_0(v) dv$$

while $f(v, t) = 0$ for a.e. $(v, t) \in \mathbb{R}^d \times \mathbb{R}^+$ if $m_0 = 0$. The quantities m_0, u_0 and T_0 are the density mass, mean and variance, associated to probability density $f(v, t)$.

- (3) The transition probability interaction rate operator $B(u, \sigma)$, or *collision kernels*, encodes not only regularity but also quantitative properties of solutions, as well as decay rates to equilibrium. For example, in the classical particle physics dynamics case, the dependence of $|u| = |v - v_*|$ relates to intermolecular potentials rates between pairs of interacting particles. In addition, the dependence on σ encodes the rate of collisions depending on the direction of the phase variables before and after the interaction.
- (4) *The double mixing convolution structure is changed into a weighted convolution by the Fourier transform:* We observed in [Gamba and Tharkabhushanam \(2009\)](#) that if the interaction law satisfies that the post-pre difference of states $v' - v$ depends only on the relative variable $u = v - v_*$ and σ , i.e., $v' - v = \omega(u, \sigma)$ (like in most particle systems of elastic or inelastic interactions) then, when testing the collisional integral with $\varphi(v) = \exp(-iv \cdot \zeta)$ in (1) and (2), it yields an identity for the Fourier transformed equation just in the v -variable, classically defined by $\hat{\cdot}(\zeta) = \mathcal{F}_{v \rightarrow \zeta}(f(v))(\zeta) = (2\pi)^{d/2} \int_{\mathbb{R}^d} e^{-i\zeta \cdot v} f(v) dv$, to obtain

$$\hat{Q}(f, f)(\zeta) = \frac{1}{(\sqrt{2\pi})^d} \int_{\mathbb{R}^d} e^{-i\zeta \cdot v} Q(f, f)(v) dv = \kappa \int_{\mathbb{R}^d} \mathcal{F}(f(v)f(v - \mathbf{u}))(\zeta) \mathcal{G}_B(\zeta, u) du$$

with the weight function $\mathcal{G}_B(\zeta, u) = \int_{\sigma \in \Omega} [e^{-i\frac{1}{2}\zeta \cdot \omega(u, \sigma)} - 1] B(u, \sigma) d\sigma.$

(7)

Consequently, $\hat{Q}(f, f)(\zeta)$ is also a *weighted convolution* of $\hat{f}(\zeta)$. In particular, the Boltzmann evolution in Fourier space is

$$\frac{d}{dt} \hat{f}(\zeta) = \hat{Q}(f, f)(\zeta) = \kappa \int_{\xi \in \mathbb{R}^d} \hat{G}_B(\zeta, \xi) \hat{f}(\xi) \hat{f}(\zeta - \xi) d\xi \quad \text{with} \quad \hat{G}_B(\zeta, \xi) = \mathcal{F}_{u \rightarrow \xi} \mathcal{G}_B(\zeta, u).$$
(8)

Remark. Both $\mathcal{G}_B(\zeta, u)$ and $\hat{G}_B(\zeta, \xi)$ can be viewed as *symbol* of the multi-linear integral operator (as the analogue to symbols of PDE's).

1.3 Classical Elastic Collisional Transport Theory: The Boltzmann Equation

Eqs. (1) and (2) are exactly the weak (or Maxwell) formulation of the collisional Boltzmann equation for elastic, inelastic interaction or collisional dynamics given by

$$\begin{aligned} v' &= v + \frac{\beta}{2}(|u|\sigma - u) \text{ and } v'_* = v - \frac{\beta}{2}(|u|\sigma - u) \text{ with relative velocity } u = v - v_*, \\ B(u, \sigma) &= |u|^\lambda b(\hat{u} \cdot \sigma), \text{ with } d < \lambda \leq 1, \quad \hat{u} = \frac{u}{|u|} \text{ and} \\ \cos \theta &= \frac{(\hat{u}, \sigma)}{|u|}, \text{ with } \theta \text{ the scattering angle, } 1/2 < \beta \leq 1, \end{aligned} \quad (9)$$

with the scattering direction $\sigma = u'/|u|$ the scattering direction given by the postcollisional relative velocity u' . The classical elastic case if for $\beta = 1$ (i.e. local energy conservation). In particular $v' - v = \frac{\beta}{2}(|u|\sigma - u) = \omega(u, \sigma)$.

In addition, a standard assumption is that the space dynamics in between the interactions evolve according to Hamiltonian dynamics for the evolving pair in x -space/ v -phase space given by $(x(t), v(t))$, position and velocity, respectively, when

$$\dot{x} = \partial_v H(x, v), \text{ and } \dot{v} = -\partial_x H(x, v) \quad (10)$$

we have the following classical dynamics of *rarefied* transport associated to the Liouville equation in between interactions and the collisional or interacting non-local form given by the Masters equation framework, written in strong form.

Two important cases in the space inhomogeneous setting are *binary* and *linear* interactions, discussed next.

- (i) *The nonlinear Boltzmann transport equation for binary interactions:* modelling monoatomic gases corresponds to binary collisional forms with Hamiltonian dynamics (10) between interactions defined for $f = f(x, v, t)$, written in strong form, is

$$\begin{aligned} \frac{\partial f}{\partial t} + v \cdot \nabla_x f &= Q(f, f)(x, v, t) \\ &= \kappa \int \int_{(u, \sigma) \in \mathbb{R}^d \times S^{d-1}} B_\lambda(|u|, \hat{u} \cdot \sigma) [J_\beta f(x, v, t) f(x, v' - u, t) \\ &\quad - f(x, v, t) f(x, v - u, t)] d\sigma du \end{aligned} \quad (11)$$

with $B_\lambda(|u|, \hat{u} \cdot \sigma) = |u|^\lambda b(\hat{u} \cdot \sigma)$ with $-d < \lambda \leq 1$, where $'v$ denotes a pre-collisional state with respect to v . The term $J_\beta = \frac{\partial(v, v_*)}{(v, v_*)}$ is the Jacobian of the “post” to “pre”-variable transformation and $|J_1| = 1$. The angular function $b(\hat{u} \cdot \sigma) = b(\cos \theta)$ may or may not be an integrable function on the sphere S^{d-1} . If integrable, i.e., $\int_{S^{d-1}} b(\hat{u} \cdot \sigma) d\sigma < K$, we will say it satisfies the *Grad cut-off assumption*.

- (ii) *The linear Boltzmann equation* (Forward Kolmogorov equation): The strong formulation of linear evolution of $f = f(x, k, t)$, a pdf, given by

$$D_t f = Q(f)(x, k, t) = \int S_{k, k'}(k \leftarrow k') f' dk' - f(x, k, t) \int S_{k', k}(k' \leftarrow k) dk'. \quad (12)$$

When the Hamiltonian dynamics are included, as for the case of charged transport with a repulsive potential $\Phi(x, t)$, this linear (12) models the dynamics of electron transport along an electronic band energy surface $\varepsilon(k)$ given by the Hamiltonian $H = \varepsilon(k) - \Phi(x)$ according to (10). It yields the material derivative

$$D_t f = \partial_t f + \frac{1}{\hbar} \partial_k \varepsilon(k) \cdot \partial_x f + \frac{1}{\hbar} \partial_x \Phi(x, t) \cdot \partial_k f \quad (13)$$

corresponding to $\dot{x} = \frac{1}{\hbar} \partial_k \varepsilon(k)$ and $\dot{v} = \frac{1}{\hbar} \partial_x \Phi(\vec{x}) = -\frac{q}{\hbar} \vec{E}(k, t)$. In addition, the potential function Φ , with its corresponding electric field $qE(x, t) = -\partial_x \Phi(x, t)$, takes into account mean field effects of the total system, and is determined by the Poisson equation for charges. The band energy function $\varepsilon(k)$ is an eigenvalue of the Bloch decomposition associated to the quantum crystallographic calculation. These equations model *hot electron collisional transport along divergence-free surfaces* ($\varepsilon(k)$, $\Phi(x, t)$) in nano-scale semiconductor devices where magnetic forces are negligible.

Both types of collisional models, nonlinear (i) and linear (ii), appear in problems that range from electron/hole transport in a crystal lattice with a linear collisional transport (Cheng et al., 2009, 2012; Morales Escalante and Gamba, 2016; Morales-Escalante et al., 2015), to classical gas dynamics problems (Aoki et al., 1993; Aristov, 2001; Bobylev et al., 2000; Brilliantov and Pöschel, 2004; Chapman and Cowling, 1970; Gamba and Tharkabhushanam, 2009, 2010; Gamba et al., 2004; Munafo et al., 2014; Sone, 2007); to flow of self-interacting particle systems, network dynamics in social, economic and information systems (Bobylev et al., 2009; Duffie et al., 2009; Ringhofer, 2010).

1.4 Deterministic Solvers for Integral Equations of Boltzmann Type

In recent years, there has been a development of deterministic solvers for kinetic transport equations, whole first high-dimensional simulations were performed

by Monte Carlo sampling methods for particle systems (Bird, 1994). There are essentially three alternatives to Monte Carlo approach for the computation of the Boltzmann equation: conservative finite element methods, conservative Spectral–Lagrangian methods and discrete velocities methods (DVM). Recent references for the conservative spectral methods can be found in Alonso et al. (2016) and Cheng et al. (2009). While in this review presentation we shall focus mostly on the conservative spectral method for nonlinear binary interactions, the linear collisional transport as in the case of collisional plasma simulations for semiconductor transport or Vlasov–Poisson–Maxwell dynamics can be performed by discontinuous Galerkin schemes, where conservation and positivity propagation is achieved by enhancing basis functions for conservation and reconstruction fluxes for positivity. This is possible for linear collisional forms that have only one conserve quantity: density mass. See this type of work in Cheng et al. (2009, 2012), Morales-Escalante et al. (2015), Morales Escalante and Gamba (2016) and references therein.

The propagation of numerical positivity for a probability distribution function $f(v, t)$ defined in all v -space while having its numerical mass, mean and variance preserved for each time step remains a very difficult task. The available reconstruction methods for conservation while preserving positivity propagation that worked so well in the Vlasov–Boltzmann equation for linear interactions (Cheng et al., 2012) fail to work in the nonlinear collisional setting as they yield overdetermined systems of discrete equations with no available solutions to guaranty such properties.

2 THE LANDAU AND BOLTZMANN OPERATORS RELATION THROUGH THEIR DOUBLE MIXING CONVOLUTIONAL FORMS

The *binary interaction* problem is the focus of the rest of this chapter. Its particular conservation and positivity propagation involves the preservation of several averaged quantities that are difficult to approximate while keeping the approximate solution positive throughout the flow computational time. Our goal is to present a method for numerically solve the Boltzmann equation with a constrained minimization problem that secures the $d + 2$ collision invariants conservation property and converge to the unique equilibrium Maxwellian characterized by the moments of initial state $f_0(v)$.

The Landau–Fokker–Plank equation (Landau, 1937; Landau and Lifschitz, 1980) is a limiting model for the Boltzmann equation used to describe binary elastic collisions (9) *that only result in very small deflections of particle trajectories*. Such limit is necessary in the case for *Coulomb potentials of the form $|u|^{-3}$* , where the classical formulation of the Boltzmann operator is not well posed. However, without loss of generality, one can consider the general form of a potential $|u|^\lambda$, with $-3 \leq \lambda \leq 1$ in the grazing collision regime. In particular, the strong form of the Landau–Fokker–Planck equation, written in $3 - d$ (omitting the variables x and t for simplicity), is

$$\begin{aligned}
 \partial_t f(v) &= Q_L(f, f)(v) \\
 &= \kappa \operatorname{div}_v \left(\int_{\mathbb{R}^3} |u|^{\lambda+2} \left(I - \frac{u \otimes u}{|u|^2} \right) (f(v-u) \nabla_v f(v) - f(v) (\nabla_v f)(v-u)) du \right) \\
 &= \kappa \operatorname{div}_v (\mathcal{D}_{ij}(v) \nabla_v f(v) - \mathcal{E}_i f(v)),
 \end{aligned} \tag{14}$$

with $\mathcal{D}_{ij}(v) := \int_{\mathbb{R}^3} |u|^{\lambda+2} \left(I - \frac{u \otimes u}{|u|^2} \right) f(v-u) du$ and $\mathcal{E}_i(v) = \int_{\mathbb{R}^3} |u|^{\lambda+2} \left(I - \frac{u \otimes u}{|u|^2} \right) (\nabla_v f)(v-u) du$.

Written in weak form, this operator is a *double mixing convolution*

$$\begin{aligned}
 &\int_{\mathbb{R}^3} Q_L(f, f) \phi(v) dv \\
 &= \kappa \int_{\mathbb{R}^6} f(v) f(v-u) \left(-4|u|^\lambda u \cdot \nabla \phi + |u|^{\lambda+2} \left(I - \frac{u \otimes u}{|u|^2} : D^2 \phi \right) \right) dv du,
 \end{aligned} \tag{15}$$

with a local weight function $G_L(u, v) = \left(-4|u|^\lambda u \cdot \nabla \phi + |u|^{\lambda+2} \left(I - \frac{u \otimes u}{|u|^2} : D^2 \phi \right) \right)$.

Its Fourier transform takes the form of the weighted convolution ([Gamba and Haack, 2014](#))

$$\hat{Q}_L(f, f)(\zeta) = \kappa \int_{\mathbb{R}^3} \mathcal{F}\{f(v)f(v-u)\}(\zeta) \left(4i|u|^\lambda (u \cdot \zeta) - |u|^{\lambda+2} |\zeta^\perp|^2 \right) du, \tag{16}$$

where $\zeta^\perp = \zeta - (\zeta \cdot u)/|u|^2 u$, the orthogonal component of ζ to u . Thus, the corresponding representation with weight function $\mathcal{G}_L(u, \zeta)$, as in (7), is now given by the local functions in (u, ζ) -space, the weight function is $\mathcal{G}_L(u, \zeta) = |u|^\lambda (4i(u \cdot \zeta) - |u|^2 |\zeta^\perp|^2)$.

Applying the Fourier transform to the difference of the Boltzmann (1), (9), (11) and Landau operators (1) and (16), as done in [Gamba and Haack \(2014\)](#), yields the difference of *weighted convolutions*

$$(\hat{Q}_B - \hat{Q}_L)(f, f)(\zeta) = \int_{u \in \mathbb{R}^3} (\mathcal{G}_B(\zeta, u) - \mathcal{G}_L(\zeta, u)) \mathcal{F}(f(v)f(v-u))(\zeta) du, \tag{17}$$

where $\mathcal{G}_B(\zeta, u)$ is represented in (7) for the case of classical elastic interactions, with an angular dependence of $\sigma \in S^2$. Considering a collision cross section that separates in a potential part depending on powers of $|u|$ and an angular part $b(\hat{u} \cdot \hat{\sigma})$, yields

$$\mathcal{G}_B(\zeta, u) = \int_{\sigma \in S^2} [e^{-i\frac{1}{2}\zeta \cdot \omega(u, \sigma)} - 1] |u|^\lambda b(\hat{u} \cdot \sigma) d\sigma, \quad \text{with } v' - v = \frac{1}{2}(|u|\sigma - u). \quad (18)$$

In order to be able to handle the approximation analysis and actual computations it is necessary to perform the following fundamental decomposition for spherical integrations of the weight function $\mathcal{G}_\sigma(\zeta, u)$, associated to the Fourier transform of the Boltzmann collision operator. This decomposition is done splitting the “polar” direction to the relative velocity u parametrized by the θ -angular parameter, and the corresponding azimuthal direction integration parametrized by a ϕ -angular parameter. The weight function $\mathcal{G}_\sigma(\zeta, u)$ form in (18) can be written as

$$\mathcal{G}_B(\zeta, u) = 2\pi |u|^\lambda \int_0^\pi b(\cos \theta) \sin \theta \left(e^{i\frac{1}{2}(1-\cos \theta)\zeta \cdot u} J_0\left(\frac{|u| \sin \theta |\zeta^\perp|}{2}\right) - 1 \right) d\theta, \quad (19)$$

with J_0 is the 0th Bessel function of first kind (see [Abramowitz and Stegun, 1964, 9.2.21](#)).

Remark. This formulation of the collisional integral does not separate the gain and loss terms. Cancellation potential singularity is possible in the grazing collision limit where the states v and v' are infinitesimally closed.

Defining σ with a pole in the direction of u , parametrized by $\sigma = \cos \theta \frac{u}{|u|} + \sin \theta \omega$, $\omega \in S^{d-2}$,

$$\mathcal{G}_B(\zeta, u) = |u|^\lambda \int_0^\pi \int_{S^{d-2}} b(\cos \theta) \sin \theta \left(e^{i\frac{1}{2}(1-\cos \theta)\zeta \cdot u} e^{-i\frac{1}{2}|u| \sin \theta (\zeta \cdot \omega)} - 1 \right) d\omega d\theta. \quad (20)$$

In the relevant case $d = 3$, the right-hand side of (20) can be written (see [Gamba and Haack, 2014](#) for the calculation)

$$2\pi |u|^\lambda \int_0^\pi b(\cos \theta) \sin \theta \left(e^{i\frac{1}{2}(1-\cos \theta)\zeta \cdot u} J_0\left(\frac{|u| \sin \theta |\zeta^\perp|}{2}\right) - 1 \right) d\theta. \quad (21)$$

The isotropic case when $b(\cos \theta)$ is constant, ζ can be used instead of u as the polar direction for σ , and so the weight (20) is a sinc-function ([Gamba and Tharkabhushanam, 2009](#)).

Finally, let \widehat{G}_b be the Fourier transform of G_b . By symmetry, \widehat{G}_b is real valued. Then, the convolution weights $\hat{G}_b(\zeta, \zeta)$ from (8), written in $3 - d$, where the integration with respect to u is performed in spherical coordinates $(r, \eta) \in \mathbb{R}^+ \times S^2$,

$$\begin{aligned} \hat{G}_b(\zeta, \zeta) &= 2\pi \int_{\mathbb{R}^3} |u|^\lambda e^{-i\zeta \cdot u} \int_0^\pi b(\cos \theta) \sin \theta \left[e^{i\frac{\zeta}{2} \cdot u(1-\cos \theta)} J_0\left(\frac{1}{2}|u| |\zeta^\perp| \sin \theta\right) - 1 \right] d\theta du \\ &= 2\pi \int_0^\infty \int_{S^2} r^{\lambda+2} \int_0^\pi b(\cos \theta) \sin \theta \left[e^{-ir(\zeta - \frac{\zeta}{2}(1-\cos \theta)) \cdot \eta} J_0\left(\frac{1}{2}r |\zeta^\perp| \sin \theta\right) - e^{-ir\zeta \cdot \eta} \right] \\ &\quad \times d\theta d\eta dr. \end{aligned}$$

Taking γ the polar angle with respect to ζ direction for the integration in $\eta \in S^2$

$$\begin{aligned} \hat{G}_b(\zeta, \zeta) = & 4\pi^2 \int_0^\infty r^{\lambda+2} \int_0^\pi \int_0^\pi b(\cos \theta) \sin \theta \sin \gamma J_0 \left(r \left| \zeta - \frac{\zeta \cdot \zeta}{|\zeta|^2} \zeta \right| \sin \gamma \right) \\ & \times \left[\cos \left(r \left(\zeta - \frac{\zeta}{2} (1 - \cos \theta) \right) \cdot \frac{\zeta}{|\zeta|} \cos \gamma \right) J_0 \left(\frac{1}{2} r |\zeta| \sin \gamma \sin \theta \right) \right. \\ & \left. - \cos \left(r \zeta \cdot \frac{\zeta}{|\zeta|} \cos \gamma \right) \right] d\theta d\gamma dr, \end{aligned} \quad (22)$$

This function can be precomputed. This is the actual weight in the *weighted convolutional* form in Fourier space (8).

2.1 The Grazing Collision Limit

This identity (21) developed in Gamba and Haack (2014) is used to get a detailed expansion of $e^{-i\frac{1}{2}\zeta \cdot (|u|\sigma - u)} - 1$ in powers of $\zeta \cdot (|u|\sigma - u)$, combined with the grazing collisions ansatz of short-range cut-off potentials such as of Rutherford-type potential satisfying the following properties (Villani, 1998): the family of angular kernels $b_\varepsilon(\theta) = c_\varepsilon b(\theta) \mathbf{1}_{\theta \geq \varepsilon}$, for $\varepsilon > 0$, satisfy

- G1.** $\int_{S^2} b(\hat{u} \cdot \sigma) d\sigma$ unbounded but $\lim_{\varepsilon \rightarrow 0} 4\pi \int_0^\pi b_\varepsilon(\cos \theta) \sin^2 \frac{\theta}{2} \sin \theta d\theta = \Lambda_0 < \infty$.
- G2.** $\int_0^\pi b_\varepsilon(\cos \theta) (1 - \cos \theta)^{2+k} \sin \theta d\theta \rightarrow 0$ for $k \geq 0$,
- G3.** $b_\varepsilon(\theta) \rightarrow 0$ uniformly on $\theta > \theta_0$; $\forall \theta_0 > 0$;

We note that the elastic interaction relation (9) implies $|v' - v|^2 = |u|^2 \sin^2 \frac{\theta}{2}$, yielding an angular singularity cancellation by means of the grazing collision ansatz *G1*. The constant c_ε depends on the singularity of the angular function $b(\theta)$ and $c_\varepsilon \rightarrow 0$ and $\varepsilon \rightarrow 0$. Hence conditions *G2* and *G3* indicate the interaction is “grazing”, meaning that, as $\varepsilon \rightarrow 0$, then because of the cut-off $\theta \rightarrow 0$ and so $v \approx v_*$.

A δ , ε -family of admissible angular singularities: In the sequel, we consider the $d = 3$ dimensional space. Introducing ε and δ reference parameters, in the notation of the ε -grazing and δ -singular angular function $b_\varepsilon^\delta(\cos \theta) \sin^{d-2}\theta$, define the functions $H_\delta(x)$ as the antiderivative from the area differential of the angular part of the differential cross section as follows

$$\begin{aligned} b_\varepsilon^\delta(\hat{u} \cdot \sigma) d\sigma &= -\frac{1}{2\pi H_\delta(\sin(\varepsilon/2))} \frac{1}{\sin^{4+\delta}(\theta/2)} \sin(\theta) \mathbf{1}_{\theta \geq \varepsilon} d\theta d\omega \\ &= -\frac{4}{2\pi H_\delta(\sin(\varepsilon/2))} \frac{1}{x^{1+\delta}} \frac{1}{x^2} \mathbf{1}_{x \geq \sin(\varepsilon/2)} dx d\omega, \end{aligned} \quad (23)$$

for $x = \sin(\theta/2)$. Thus, equating the last term above to the right-hand side of relation (21), one can explicitly calculate $H_\delta(x)$ as the antiderivative of $x^{-(1+\delta)}$, to obtain

$$H_\delta(x) = -\frac{x^{-\delta}}{\delta}, \text{ for } \delta > 0 \text{ and } H_0(x) = \log x, \text{ for } \delta = 0; \quad (24)$$

where the choice of the exponent δ must satisfy condition G3, for $H_\delta(x)$.

Finally, an expansion of the weight $\mathcal{G}_{b_\varepsilon^\delta}(u, \zeta)$ as calculated in (21), which recovers an ε and $\sin(\theta/2)$ free term, labelled $\mathcal{G}_L(u, \zeta)$ in Gamba and Haack (2014), so that the following theorem holds

Theorem 1. Assume that f_ε^δ satisfies

$$|\mathcal{F}\{f_\varepsilon^\delta(\mathbf{v}, t)f_\varepsilon^\delta(\mathbf{v} - \mathbf{u})\}(\zeta)| \leq \frac{A(\zeta, t)}{1 + |\mathbf{u}|^{3+a}}, \quad (25)$$

with $A(\zeta, t)$ uniformly bounded by $k(1 + |\zeta|)^{-3}$, k constant, and $a > 0$. Moreover, assume the angular scattering cross section $b_\varepsilon^\delta(\cos\theta)\sin\theta^{d-2}$ satisfies conditions G1, G2 and G3 with H_δ from (24), with $0 \leq \delta < 2$, and $\lambda = -3$. Then, the rate of convergence from the Boltzmann with grazing collisions to the Landau collision operator is given by

$$\|\widehat{\mathcal{Q}}_L(f_\varepsilon^\delta, f_\varepsilon^\delta) - \widehat{\mathcal{Q}}_{b_\varepsilon^\delta}(f_\varepsilon^\delta, f_\varepsilon^\delta)\|_{L^\infty} \leq O\left(\frac{|1 + (|\log(\sin(\varepsilon/2))| - 1) 1_{\{\delta=1\}}|}{|H_\delta(\sin(\varepsilon/2))|}\right) \rightarrow_{\varepsilon \rightarrow 0} 0. \quad (26)$$

The proof of this statement can be found in Gamba and Haack (2014), but basically consists in showing

$$\mathcal{G}_\sigma(u, \zeta) = \mathcal{G}_L(u, \zeta) + O(b_\varepsilon) \text{ and } \int_{\mathbb{R}^d} O(b_\varepsilon) \mathcal{F}(f(v)f(v-u))(\zeta) du \rightarrow_{\varepsilon \rightarrow 0} 0, \quad (27)$$

if $|\mathcal{F}\{f_\varepsilon^\delta(\mathbf{v}, t)f_\varepsilon^\delta(\mathbf{v} - \mathbf{u})\}(\zeta)|$ satisfies conditions (25), uniformly in time. This estimate will ensure the asymptotics of solutions to the Boltzmann equation for Coulombic interactions in the grazing collision limit to solutions of the Landau equation. It is important to observe that the rate of decay in ε depends on the choice of the angular singularity in b_ε^δ . In addition, it is numerically observed different entropy decay rates to equilibrium depending on the singularity strength measure by δ in the angular function $b_\varepsilon^\delta(\hat{\mathbf{u}} \cdot \boldsymbol{\sigma}) \sin^{d-2}\theta$.

The numerical implementations of a comparison of the Boltzmann equation for grazing collision limits and the corresponding reduced Landau equation were extensively performed in Gamba and Haack (2014). The conservative spectral method (see Sections 3 and 4) was implemented, where the weight function (22) was precomputed. Numerical simulations of such comparisons for different cross sections are in Fig. 1.

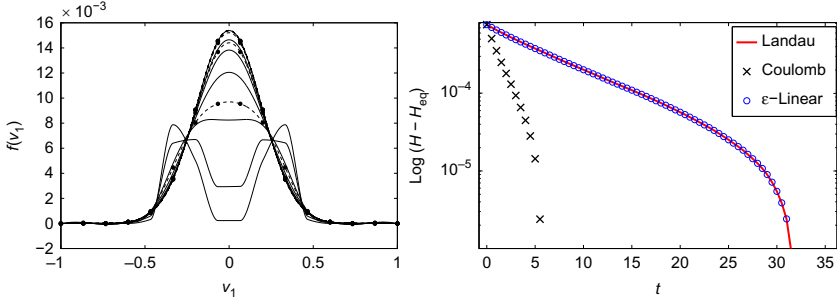


FIG. 1 Comparisons of solutions to Boltzmann using Rutherford cross section (Villani, 1998) and to Landau equations. *Left*: Slice of the distribution marginal function at times $t = 0, 9, 36, 81, 144, 225, 900$. *Solid lines*: Spline reconstruction of Landau equation solution. *Dashed lines with solid circles*: Spline reconstruction of Boltzmann equation. $\varepsilon = 10^{-4}$, $N = 16$. *Right*: Convergence of entropy to equilibrium: Log of entropy decay for Boltzmann solution with the Rutherford cross section (24) with $\delta = 0$ with crosses, and with ε -linear cross section with $\delta = 1$ with circles, and Landau solution with solid curve. $N = 16$, $\varepsilon = 10^{-4}$. When calculating the entropy H , we exclude grid points where the distribution is negative. *From Gamba, I.M., Haack, J.R., 2014. A conservative spectral method for the Boltzmann equation with anisotropic scattering and the grazing collisions limit. J. Comput. Phys. 270, 40–57.*

3 A CONSERVATIVE SPECTRAL METHOD FOR THE COLLISIONAL FORM

The following approach and its approximation analysis and error estimates for the solutions to the Cauchy problem Boltzmann equation were introduced in Alonso et al. (2016), Gamba and Tharkabhushanam (2009) and Gamba and Tharkabhushanam (2010) inspired by the work in Bobylev and Rjasanow (1999) and Pareschi and Russo (2000). We present here a short version of the strategy to be used. We recall the definition of the Lebesgue's spaces $L_k^p(\Omega)$ and the Hilbert spaces $H_k^\alpha(\Omega)$ for a measurable set Ω (for the purpose of this discussion without loss of generality Ω is either $(-L, L)^d$ or \mathbb{R}^d most of the time.)

$$L_k^p(\Omega) := \left\{ f : \|f\|_{L_k^p(\Omega)} := \left(\int_{\Omega} |f(v) \langle v \rangle^k|^p dv \right)^{\frac{1}{p}} < \infty \right\}, \text{ with } p \in [1, \infty), k \in \mathbb{R},$$

$$H_k^\alpha(\Omega) := \left\{ f : \|f\|_{H_k^\alpha(\Omega)} := \left(\sum_{\beta \leq \alpha} \|D^\beta f\|_{L_k^2(\Omega)}^2 \right)^{\frac{1}{2}} < \infty \right\}, \text{ with } \alpha \in \mathbb{N}^d, k \in \mathbb{R},$$

where $\langle v \rangle := \sqrt{1 + |v|^2}$. The standard definition is used for the case $p = \infty$,

$$L_k^\infty(\Omega) := \left\{ f : \|f\|_{L_k^\infty(\Omega)} := \text{esssup} \left| f(v) \langle v \rangle^k \right| < \infty \right\}, \text{ with } k \in \mathbb{R}.$$

It will be commonly used the following shorthand to ease notation when the domain Ω is clear from the context

$$\|\cdot\|_{L_k^p(\Omega)} = \|\cdot\|_{L_k^p} = \|\cdot\|_{p,k},$$

and the subindex k will be omitted in the norms for the classical spaces L^p and H^z . What follows is short presentation of the extended work in [Alonso et al. \(2016\)](#).

3.1 Choosing a Computational Cut-Off Domain Ω_L

Because the computation of this problem entices to numerically solve the evolution of a probability distribution function defined in the whole \mathbb{R}^d -space, it is relevant to discuss the choice of a computational cut-off domain in such a way that the numerical error for the flow evolution is negligible regarding this choice of computational window. This discussion is actually independent of the choice computational scheme and applies to existing as well new approaches such as the recently developed in [Zhang and Gamba \(2016\)](#). The following result is rigorous and applied to the Boltzmann equation for hard potentials, that is (11) with $1 \geq \lambda > 0$. Hence, for any $f(v, t)$ solution of the elastic homogeneous Boltzmann equation lying in $\mathcal{C}(0, T; H^z(\mathbb{R}^d))$, with a given initial state $f(v, 0) = f_0(v) \in H^z(\mathbb{R}^d)$. A natural question to ask is: can one secure the propagation of regularity and tail decay for the solution of the Boltzmann problem, uniformly in time? What are good functional spaces for probability distribution functions that are solutions of the Boltzmann flow problem? These questions have been recently addressed in [Bobylev et al. \(2004\)](#) and [Gamba et al. \(2009\)](#) and provide an answer in a suitable form for any computational approach of the space homogeneous elastic Boltzmann equation.

To address this problem, we introduce the following notation for exponentially weighted integrable functions. Define

$$L_{(r,2)}^1(\mathbb{R}^d) := \left\{ g : \|g\|_{L_{(r,2)}^1} := \int_{\mathbb{R}^d} |g(y)| e^{r|y|^2} dy < \infty \right\}, \text{ with } r > 0, \quad (28)$$

and analogous definition for the spaces $L_{(r,2)}^p(\mathbb{R}^d)$ with $p \in (1, \infty]$. These functional spaces are comprise of nonnegative elements in $L_{(r,2)}^1(\mathbb{R}^d)$ that are Gaussian (or Maxwellian) weighted regular probability densities, meaning that the probability density g not only has all its moments bounded but also grow as the moments of a Gaussian distribution. That also means the probability density $g(y)$ decays like $e^{-r|y|^2}$ with rate r for large $|y|$ in the sense of L^1 . In particular, one may view r^{-1} as the corresponding Gaussian or Maxwellian *tail temperature* of the density.

The following rigorous result was obtained for the elastic homogeneous Boltzmann initial value problem (Gamba et al., 2009) that shows if $f_0(v) \in L^1_{(r_0,2)}(\mathbb{R}^d)$, then solution $f(v,t) \in L^1_{(r,2)}(\mathbb{R}^d)$, for some $0 < r \leq r_0$, uniformly in t , where r only depends on a number k' -moments of the initial state f_0 , with $k' > 2$, as well as on the scattering kernel B (i.e. on the potential rate λ and the angular function $b(\hat{u} \cdot \sigma)$). In addition, if the angular cross section satisfies $b(\hat{u} \cdot \sigma) \in L^1+(\mathbb{S}^{d-1})$, then the propagation of $f(v,t) \in L^\infty_{(\bar{r},2)}(\mathbb{R}^d)$, with $0 < \bar{r} \leq r \leq r_0$, for all time $t > 0$ was also shown in Gamba et al. (2009), provided the initial data $f_0(v) \in L^\infty_{(r_0,2)}(\mathbb{R}^d)$. As a consequence, for any given initial state $f_0 \in L^1_{e^{2|v|^2}}(\mathbb{R}^d)$ or $f_0 \in L^\infty_{e^{2|v|^2}}(\mathbb{R}^d)$, there exists a rate $\beta = \beta(k'(f_0), \lambda, b)$, for which these exponentially weighted L^1 and L^∞ norms propagate uniformly in time.

This propagation property secures a stable numerical simulation of the Boltzmann equation, provided the numerical preserving the conservation laws or corresponding collision invariants hold. This property yields the convergence to the analytic solution of the initial value problem and its long-time behaviour converges to the equilibrium Maxwellian, as defined in (6). In fact we showed that it is sufficient to choose the domain $\Omega_L = (-L, L)^d$ large enough such that, at least, most of the mass and energy of the solution f will be contained in it throughout the simulation. One possible strategy for choosing the size of the simulation domain is as follows: assume without loss of generality a bounded initial datum f_0 with compact support and having zero momentum $\int f_0 v = 0$. Then,

$$f_0(v) \leq C_0 m_0 (2\pi T_0)^{-d/2} e^{-\frac{r_0|v|^2}{2T_0}}, \quad (29)$$

where $m_0 := \int f_0 dv$ is the initial mass, $T_0 := \int f_0 |v|^2 dv$ is the initial temperature, and $r_0 \in (0, 1]$ and $C_0 \geq 1$ are the stretching and dilating constants. The aforementioned analytical results secure that for some $r := r(f_0, \lambda, b) \in (0, r_0]$ and $C := C(f_0, \lambda, b) \geq 1$

$$f(t, v) \leq C_0 m_0 (2\pi T_0)^{-d/2} e^{-\frac{r|v|^2}{2T_0}} =: M(f_0, C, r), \quad t > 0.$$

A simple criteria to pick the segment length L of the simulation domain Ω_L are to ensure that most of the mass and kinetic energy (or variance) of f will remain in it throughout the numerical simulation. In other words, we want that, for some small number $\mu \ll 1$,

$$\int_{\Omega_L^c} f(v, t) \langle v \rangle^2 dv \leq \int_{\Omega_L^c} M(f_0, C, r) \langle v \rangle^2 dv \leq \mu \int_{\Omega_L} f_0(v) \langle v \rangle^2 dv = \mu(m_0 + T_0).$$

where μ is chosen to be understood as a domain cut-off error tolerance that *remains uniform in time* and solely depends on the initial state and Ω_L . Equivalently, one needs to choose the size of L , or equivalently the measure of Ω_L , such that

$$(m_0 + T_0)^{-1} \int_{\Omega_L^c} M(f_0, C, r) \langle v \rangle^2 dv \leq \mu \approx 0 \quad (30)$$

In order to minimize the computation effort, one should pick the smallest of such domains, that is, Ω_{L_*} , $L_* = \min \{L > 0 : \text{supp}(f_0) \subset \Omega_L, \Omega_L^c \text{ satisfies (30)}\}$.

The choice of μ in (30) depends on the knowledge of precise values for the constants C and r not so easy to determine for a generic initial data. Hence, in order to avoid overestimating the size of Ω_L , the simulation domain, it is best to set $r_0 = r = 1$ and choose $C = C_0 \geq 1$ as the smallest constant satisfying (29) (which always exists for any compactly supported and bounded f_0), and set

$$\max \left\{ f_0, m_0 (2\pi T_0)^{-d/2} e^{-\frac{|v|^2}{2T_0}} \right\} \leq M(f_0, C, 1),$$

with equality if and only if f_0 is the equilibrium Maxwellian as in (6) (in such a case $C = 1$). Then the use of classical *Normal Table for log-normal distributions* yields the error μ incurred in the simulation as a function of the chosen Ω_L , uniformly in time, for any simulation of the Boltzmann collisional model homogeneous in x -space.

Remark. In this deterministic approach, as much as with Monte Carlo methods like the Bird scheme (Bird, 1994), the x -space inhomogeneous Hamiltonian transport for nonlinear collisional forms is performed by time operator-splitting algorithms. That means, depending on the problem, the computational v -domain Ω_L can be updated with respect to the characteristic flow associated to underlying Hamiltonian dynamics.

3.2 Fourier Series, Projections and Extensions

In the implementation of any spectral method the single most important analytical tool is the Fourier transform defined by $\hat{f}(\zeta) := (2\pi)^{d/2} \int_{\mathbb{R}^d} f(v) e^{-i\zeta \cdot v} dv$, defined for any $f \in L^1(\mathbb{R}^d)$. Our goal is to approximate the collisional form in Fourier space, given by the weighted convolution in Fourier space (8), by making use on the approximant Fourier series in a rather simple and convenient way. Indeed, fixing a domain of work $\Omega_L := (-L, L)^d$ for $L > 0$, recall that for any $f \in L^2(\Omega_L)$ the *Fourier series* of f , denoted from now on by f_L is given by $f_L \sim 1/((2L)^d) \sum_{k \in \mathbb{Z}^d} \hat{f}_L(\zeta_k) e^{i\zeta_k \cdot v}$, where $\zeta_k := \frac{2\pi k}{L}$ are the spectral modes and $\hat{f}_L(\zeta_k)$ is the Fourier transform of f_L evaluated in such modes.

Next, define the operator $\Pi^N : L^2(\Omega_L) \rightarrow L^2(\Omega_L)$ as

$$(\Pi^N f_L)(v) := f_L^\Pi(v) = \left(\frac{1}{(2L)^d} \sum_{|k| \leq N} \hat{f}_L(\zeta_k) e^{i\zeta_k \cdot v} \right) \mathbf{1}_{\Omega_L}(v), \quad (31)$$

that is, the *orthogonal projection* on the “first N ” basis elements and see that for any integer α the derivative operator commutes with the projection operator. In Ω_L

$$\partial^\alpha (\Pi^N f_L)(v) = \left(\frac{1}{(2L)^d} \sum_{|k| \leq N} \widehat{\partial^\alpha f}(\zeta_k) e^{i\zeta_k \cdot v} \right) \mathbf{1}_{\Omega_L}(v) = (\Pi^N \partial^\alpha f)(v). \quad (32)$$

The Parseval’s theorem implies $\|\Pi^N f_L\|_{L^2(\Omega_L)} \leq \|f_L\|_{L^2(\Omega_L)}$ for any N ; and $\|\Pi^N f_L - f_L\|_{L^2(\Omega_L)} \searrow 0$ as $N \rightarrow \infty$. Because of the decay properties described in the procedure of choosing the computational domain Ω_L , we can avoid the expected aliasing effect for a classical Fourier approximation by series by using the classical extension theorem in Sobolev spaces as follows.

3.2.1 The Extension Operator

For fixed $\alpha_0 \geq 0$ we introduce the *extension operator* $E: L^2(\Omega_L) \rightarrow L^2(\mathbb{R}^d)$ such that $E: H^\alpha(\Omega_L) \rightarrow H^\alpha(\mathbb{R}^d)$ holds for any $\alpha \leq \alpha_0$. The construction of such operator (Stein, 1970) is well known having the following properties:

- E1.** Linear and bounded, with $\|Ef\|_{H^\alpha(\mathbb{R}^d)} \leq C_\alpha \|f\|_{H^\alpha(\Omega_L)}$ for $\alpha \leq \alpha_0$.
- E2.** $Ef = f$ a.e. in Ω_L . Furthermore, denoting f^\pm the positive and negative parts of f one has $(Ef)^\pm = Ef^\pm$, a.e. in \mathbb{R}^d .
- E3.** Outside Ω_L the extension is constructed using a reflexion of f near the boundary $\partial\Omega_L$. Thus, for any $\delta \geq 1$ we can choose an extension with support in $\delta\Omega_L$, the dilation of Ω_L by δ , and $\|Ef\|_{L^p(\delta\Omega_L \setminus \Omega_L)} \leq C_0 \|f\|_{L^p(\Omega_L \setminus \delta^{-1}\Omega_L)}$ for $1 \leq p \leq 2$, where the constant C_0 is independent of the support of the extension.
- E4.** In particular, properties E2 and E3 imply that for any $\delta \geq 1$, there is an extension such that $\|Ef\|_{L_k^p(\mathbb{R}^d)} \leq 2C_0 \delta^{2k} \|f\|_{L_k^p(\Omega_L)}$ for $1 \leq p \leq 2$, $k \geq 0$.

The case $\delta = 1$ is only possible using an extension by zero, that is, when $(Ef)(v) = f(v) \mathbf{1}_{\Omega_L}(v)$, and so α_0 is restricted to zero.

3.3 A Conservative Spectral Method for the Homogeneous Boltzmann Equation

After the cut-off domain Ω_L has been fixed, the projection operator is applied to both sides of Eq. (8) to obtain

$$\frac{\partial \Pi^N f}{\partial t}(v, t) = \Pi^N Q(f, f)(v, t), \text{ in } (0, T] \times \Omega_L.$$

Hence, for such a domain Ω_L and sufficiently large number of modes N , it is expected that the approximation $\Pi^N Q(f, f) \sim \Pi^N Q(\Pi^N f, \Pi^N f)$, in $(0, T] \times \Omega_L$ will be valid, leading to pose and solve the problem

$$\frac{\partial g}{\partial t}(v, t) = \Pi^N Q(g, g)(v, t), \text{ in } (0, T] \times \Omega_L,$$

with initial condition $g_0 = \Pi^N f_0$, and expect that it should be a good approximation to $\Pi^N f$. In other words we define the numerical solution to be $g_N := g$ and expect to show that this discrete solution will be a good approximation to the solution of the Boltzmann problem in the cut-off domain, that is, $g \approx f$ in Ω_L , provided the number of modes N used is sufficiently large. This formalism has been shown in [Alonso et al. \(2016\)](#), under some assumptions for the space homogeneous Boltzmann equation. To this end, we much study a modified problem, namely, the convergence towards f of the solution g of the semidiscrete problem

$$\frac{\partial g}{\partial t}(v, t) = Q_c(g, g)(v, t) \quad \text{in } (0, T] \times \Omega_L, \quad (33)$$

with initial condition $g_0 := g_0^N = \Pi^N f_0$. The operator $Q_c(g)$ is defined as the $L^2(\Omega_L)$ -closest function to $\Pi^N Q(Eg, Eg)$ having null mass, momentum and energy. Since the gain collision operator is global in velocity, it turns out that a good approximation to f will be obtained as long as Ω_L and N are sufficiently large. The extension operator E has a subtle job to do in the approximation scheme which is related precisely to the global behaviour of the gain collision operator. Since solutions of the approximation problem (33) lie in Ω_L , they are truncated versions of f . The gain operator does not possess higher derivatives in Ω_L when acting on truncated functions due to the singularity created in the boundary $\partial\Omega_L$. The extension smooths out the gain collision operator at the price of extending the domain. In the case of discontinuous solutions where only L^2 -error estimate is expected, the correct extension to use in the scheme is the extension by zero. We discuss this more carefully in the following sections.

Before continue with the discussion, we are now in position to summarize the main results on convergence, error estimates and asymptotic behaviour. These statements are stated in the following theorem. Rigorous proofs can be found at [Alonso et al. \(2016\)](#).

Theorem 2 (Error estimates and convergence to the equilibrium Maxwellian). *Fix an initial nonnegative initial data $f_0 \in (L^1_2 \cap L^2)(\mathbb{R}^d)$. Then, for any time $T > 0$ there exist a lateral size $L := L(T, f_0)$ and a number of modes $N_0 := N(T, L, f_0)$ such that*

1. **Semidiscrete existence and uniqueness:** *The semidiscrete problem (33) has a unique solution $g \in C(0, T; L^2(\Omega_L))$ for any $N \geq N_0$.*
2. **L^2_k -error estimates:** *if $f_0 \in (L^1 \cap L^2)_{k' + k + \frac{1}{2}}(\mathbb{R}^d)$ for some $k', k \geq 0$, then*

$$\sup_{t \in [0, T]} \|f - g\|_{L^2_k(\Omega_L)} \leq CL^{-\lambda k'} e^{cT}, \text{ for any } N \geq N_0,$$

where $N_0 := N(T, L, f_0, k)$, $C := C(k, f_0)$, $c := c(k, f_0)$ and f is the solution of the Boltzmann equation (11).

3. **H^α -error estimates:** For the smooth case $f_0 \in \left(L_2^1 \cap H_q^{\alpha_0}\right)(\mathbb{R}^d)$, with $\alpha_0 > 0$ and $q = \max\{k' + k, 1 + \frac{d}{2k}\}$, with $k' \geq 2$, it follows for any $\alpha \leq \alpha_0$

$$\sup_{t \in [0, T]} \|f - g\|_{H_k^\alpha(\Omega_L)} \leq C_{k'} e^{c_k T} \left(O\left(\frac{L^{\lambda k + |\alpha_0|}}{N^{|\alpha_0| - |\alpha|}}\right) + O\left(L^{-\lambda k'}\right) \right), \text{ for any } N \geq N_0,$$

where $N_0 := N(T, L, f_0, k, \alpha)$. And finally,

4. **Convergence to the equilibrium Maxwellian:** for every $\delta > 0$ there exist a simulation time $T := T(\delta) > 0$, corresponding lateral size $L := L(T, f_0)$ and baseline number of modes $N_0 := N_0(T, L, f_0, \alpha)$ such that for any $\alpha \leq \alpha_0$

$$\sup_{t \in [\frac{T}{2}, T]} \|\mathcal{M}_0 - g\|_{H^2(\Omega_L)} \leq \delta, \quad N \geq N_0,$$

where \mathcal{M}_0 is the equilibrium Maxwellian (6) having the same mass, momentum and kinetic energy of the initial configuration $f_0(v)$.

The sketch of the proof of [Theorem 2](#), presented next, relies on the control problem that enforces conservation at the numerical level.

3.4 Conservation Method—An Extended Isoperimetric Problem

This is the procedure that secures the conservation of the necessary collision invariants. This procedure is posed as a standard $L^2(\Omega_L)$ -optimization problem. Therefore, due to the truncation of the velocity domain the projection of $Q(f, f)$, namely $\Pi^N Q(f, f)$, does not preserve mass, momentum and energy. In order to accomplish these conservation properties, the problem is posed as constraints in a optimization problem to a conserved state. We denote, for the sake of brevity,

$$Q_u(f)(v) := \Pi^N(Q(Ef, Ef) \mathbf{1}_{\Omega_L})(v). \quad (34)$$

The indicator function $\mathbf{1}_{\Omega_L}(v)$ is due to the fact that the domain of $Q(Ef, Ef)$ most likely be larger than Ω_L , and thus the extension operator helps to avoid introducing spurious nonsmoothness within the domain Ω_L due to the domain cut-off, as described in [Section 3.2](#).

The conservation optimization problem consists into minimize, in the Banach space

$$\mathcal{B}^e = \left\{ X \in L^2(\Omega_L) : \int_{\Omega_L} X = \int_{\Omega_L} Xv = \int_{\Omega_L} X|v|^2 = 0 \right\},$$

the functional, defined for a computed and unconserved collision operator $Q_u(f)$, by

$$\mathcal{A}^e(X) := \int_{\Omega_L} (Q_u(f)(v) - X)^2 dv. \quad (35)$$

In other words, minimize the L^2 -distance to the projected collision operator subject to mass, momentum and energy conservation. The following lemma plays a fundamental role for error estimates as well as the convergence to the equilibrium Maxwellian (6).

Lemma 1 (Elastic Lagrange estimate). *The problem (35) has a unique minimizer given by*

$$X^* = Q_u(f)(v) - \frac{1}{2} \left(\gamma_1 + \sum_{j=1}^d \gamma_{j+1} v_j + \gamma_{d+2} |v|^2 \right),$$

where γ_j , for $1 \leq j \leq d+2$, are Lagrange multipliers associated with the elastic optimization problem. Furthermore, these Lagrange multipliers are given by

$$\gamma_1 = O_d \rho_u + O_{d+2} e_u; \quad \gamma_{j+1} = O_{d+2} \mu_u^j, \quad j = 1, 2, \dots, d; \quad \gamma_{d+2} = O_{d+2} \rho_u + O_{d+4} e_u.$$

The parameters ρ_u, e_u, μ_u^j are defined below in (38) and $O_r := O(L^{-r})$ only depends inversely on diameter $|\Omega_L|$. The minimized objective function is

$$\begin{aligned} \mathcal{A}(X^*) &= \|Q_u(f) - X^*\|_{L^2(\Omega_L)}^2 \leq C(d) \left(2\gamma_1^2 L^d + \left(\sum_{j=1}^d \gamma_{j+1}^2 \right) L^{d+2} + \gamma_{d+2}^2 L^{d+4} \right) \\ &\leq \frac{C(d)}{L^d} \left(\rho_u^2 + \frac{e_u^2}{L^{d+1}} + \sum_{j=2}^{d+1} \mu_j^2 \right) \end{aligned} \quad (36)$$

The proof of this lemma is constructive and very fundamental. When the objective function is an integral equation and the constraints are integrals, the optimization problem can be solved by forming the Lagrangian functional and finding its critical points. Indeed, set for all $j = 1, 2, \dots, d$,

$$\psi_1(X) := \int_{\Omega_L} X(v) dv; \quad \psi_{j+1}(X) := \int_{\Omega_L} v_j X(v) dv; \quad \psi_{d+2}(X) := \int_{\Omega_L} |v|^2 X(v) dv,$$

and define

$$\mathcal{H}(X, X', \boldsymbol{\gamma}) := \mathcal{A}(X) + \sum_{i=1}^{d+2} \gamma_i \psi_i(X) = \int_{\Omega_L} h(v, X, X', \boldsymbol{\gamma}) dv.$$

then, introduced $h(v, X, X', \boldsymbol{\gamma}) := (Q_u(f)(v) - X(v))^2 + \left(\gamma_1 + \sum_{j=1}^d \gamma_{j+1} v_j + \gamma_{d+2} |v|^2 \right) X(v)$. In order to find the critical points compute $D_X \mathcal{H}$ and $D_{\gamma_j} \mathcal{H}$ and note that the derivatives $D_{\gamma_j} \mathcal{H}$ just retrieve the constraint integrals. Hence, for multiple independent variables v_j and a single dependent function $X(v)$ the Euler–Lagrange equations are

$$D_2 h(v, X, X', \boldsymbol{\gamma}) = \sum_{j=1}^d \frac{\partial D_3 h}{\partial v_j}(v, X, X', \boldsymbol{\gamma}) = 0.$$

We used the fact that h is independent of X' . This gives the following equation for the conservation correction in terms of the Lagrange multipliers

$$2(X(v) - Q_u(f)(v)) + \gamma_1 + \sum_{j=1}^d \gamma_{j+1} v_j + \gamma_{d+2} |v|^2 = 0, \quad (37)$$

and therefore, $X^*(v) = Q_u(f)(v) - \frac{1}{2} \left(\gamma_1 + \sum_{j=1}^d \gamma_{j+1} v_j + \gamma_{d+2} |v|^2 \right)$.

Letting $g(v, \gamma) = \gamma_1 + \sum_{j=1}^d \gamma_{j+1} v_j + \gamma_{d+2} |v|^2$ and substituting (37) into the constraints $\psi_f(X^*) = 0$ yields

$$\begin{aligned} \rho_u &:= \int_{\Omega_L} Q_u(f)(v) dv = \frac{1}{2} \int_{\Omega_L} g(v, \gamma) dv \\ \mu_u^j &:= \int_{\Omega_L} v_j Q_u(f)(v) dv = \frac{1}{2} \int_{\Omega_L} v_j g(v, \gamma) dv, \quad j = 1, 2, \dots, d, \\ e_u &:= \int_{\Omega_L} |v|^2 Q_u(f)(v) dv = \frac{1}{2} \int_{\Omega_L} |v|^2 g(v, \gamma) dv. \end{aligned} \quad (38)$$

Identities (38) form a $d + 2$ by $d + 2$ system of linear equations that can be uniquely solved. Indeed, solving for the critical $\gamma_0, \gamma_{j+1}, j = 1, 2, \dots, d$ and γ_{j+2} yields

$$\gamma_1 = O_d \rho_u + O_{d+2} e_u; \quad \gamma_{j+1} = O_d \mu_u^j; \quad \gamma_{d+2} = O_{d+2} \rho_u + O_{d+4} e_u, \quad (39)$$

where $O_r := O(L^{-r})$. In particular, O_r depends inversely on $|\Omega_L|$. Substituting these values of critical Lagrange multipliers (39) into (37) gives explicitly the critical $X^*(v)$. Moreover, the objective function $\mathcal{A}^e(X)$ can be computed at its minimum as

$$\begin{aligned} \mathcal{A}^e(X^*) &= \|Q_u(f) - X^*\|_{L^2(\Omega_L)}^2 = \int_{\Omega_L} (Q_u(f)(v) - X^*(v))^2 dv \\ &= \frac{1}{4} \int_{\Omega_L} \left(\gamma_1 + \sum_{j=1}^d \gamma_{j+1} v_j + \gamma_{d+2} |v|^2 \right)^2 dv. \end{aligned}$$

Upon simplification, taking $\Omega_L = (-L, L)^d$ and expressing the Lagrange multipliers $\gamma_j, j = 1 \dots d + 2$ in terms of the unconserved moments $\rho_u, \mu_{u,j+1}, j = 1 \dots d$ and e_u from relation (39), one obtains

$$\begin{aligned} \|Q_u(f) - X^*\|_{L^2(\Omega_L)}^2 &\leq C(d) \left(2\gamma_1^2 L^d + \frac{2}{d} \left(\sum_{j=1}^d \gamma_{j+1}^2 \right) L^{2d-1} + \gamma_{d+2}^2 L^{d+4} \right) \\ &\leq \frac{1}{L^d} O \left(\left(\rho_u + \frac{e_u}{L^{(d+1)/2}} \right)^2 + \sum_{j=1}^{d+1} \mu_{j+1}^2 \right), \end{aligned}$$

which is just an $O(L^{-d})$ proportional to these unconserved moments that are shown to be uniformly bounded in time in [Lemma 3](#) in [Section 4.2](#). In a sense this result is better described as an *isomoment* estimate, which yields estimate (36). The strict convexity of \mathcal{A}^e implies that this critical point is the unique minimizer. This results clearly shows that the last estimate secures, after a few iterations of the conservation algorithm (to be fully described in the next section) that, for a fixed L and a number of modes N the solution of the minimization problem will converge to an approximate of the collision operator whose first $d + 2$ moments are null. This is the tool that will allows us to construct error estimates for the approximation to the true solution to the Boltzmann equation and the numerically calculated one for the spectral-conserved algorithm. In addition these *conservation correction estimates* from [Lemma 2](#) in (42), provide the necessary tool to prove that the numerical solution converges to the equilibrium Maxwellian (6) (see [Alonso et al., 2016](#) for details.) \square

Summarizing, we are now in conditions to define the *conserved projection operator* $Q_c(f)$ as follows.

Definition. For any fixed $f \in L^2(\Omega_L)$ the *conserved projection operator* $Q_c(f)$ is defined as the minimizer of problem (E) That is,

$$Q_c(f) := X^*. \quad (40)$$

From [Lemma 1](#), the minimized objective function (36) in the elastic optimization problem depends only on the nonconserved moments ρ_u , μ_u^i , and e_u of $Q_u(f)$. These quantities are to be approximating the $d + 2$ dimensional zero vector, therefore, the conserved projection operator is a perturbation of $Q_u(f)$ by a second order polynomial. Denoting the moments of a function f by

$$m_k(f) := \int_{\mathbb{R}^d} |f(v)| |v|^{\lambda k} dv. \quad (41)$$

Lemma 2 (Conservation correction estimate). *Fix $f \in L^2(\Omega_L)$, then the accuracy of the conservation minimization problem is proportional to the spectral accuracy. That is, for any $k, k' \geq 0$ and $\delta > 1$ there exists an extension E such that*

$$\begin{aligned} \|(Q_c(f) - Q_u(f))|v|^{\lambda k}\|_{L^2(\Omega_L)} &\leq \frac{C}{\sqrt{k+d}} L^{\lambda k} \|Q(Ef, Ef) - Q_u(f)\|_{L^2(\Omega_L)} \\ &\quad + \frac{\delta^{2\lambda k'}}{\sqrt{k+d}} O_{(d/2+\lambda(k'-k))}(m_{k'+1}(f)m_0(f) + Z_{k'}(f)), \end{aligned} \quad (42)$$

where C is a universal constant and $Z_{k'}(f)$ depending on the moments up to order k' .

3.5 Discrete in Time Conservation Method: Lagrange Multiplier Method

In this section we consider the discrete version of the conservation scheme. For such a discrete formulation, the conservation routine is implemented as a Lagrange multiplier method where the conservation properties of the discrete distribution are set as constraints. Let $M = N^d$, the total number of Fourier modes. For elastic collisions, $\rho = 0$, $\mathbf{m} = (m_1, \dots, m_d) = (0, \dots, 0)$ and $e = 0$ are conserved. Let $\omega_j > 0$ be the integration weights for $1 \leq j \leq M$ and define

$$\mathbf{Q}_u = (Q_{u,1} \ Q_{u,2} \ \dots \ Q_{u,M})^T$$

as the distribution vector at the computed time step, and

$$\mathbf{Q}_c = (Q_{c,1} \ Q_{c,2} \ \dots \ Q_{c,M})^T$$

as the corrected distribution vector with the required moments conserved. For the elastic case, let

$$\mathbf{C}_{(d+2) \times M}^e = \begin{pmatrix} \omega_j \\ v_1 \omega_j \\ \dots \\ v_d \omega_j \\ |v_j|^2 \omega_j \end{pmatrix} \quad 1 \leq j \leq M, \quad (43)$$

be the integration matrix, and $\mathbf{a}_{(d+2) \times 1}^e = \left(\frac{d}{dt} \rho \quad \frac{d}{dt} m_1 \quad \dots \quad \frac{d}{dt} m_d \quad \frac{d}{dt} e \right)^T$ be the vector of conserved quantities (note that \mathbf{a}^e is null $d + 2$ -dimensional null vector in the case of elastic theory, but may not be in general). With this notation in mind, the discrete conservation method can be written as a constrained optimization problem: find \mathbf{Q}_c such that is the unique solutions of

$$\mathcal{A}(\mathbf{Q}_c) = \{ \min \|\mathbf{Q}_u - \mathbf{Q}_c\|_2^2 : \mathbf{C}^e \mathbf{Q}_c = \mathbf{a}^e \text{ with } \mathbf{C}^e \in \mathbb{R}^{d+2 \times M}, \mathbf{Q}_u \in \mathbb{R}^M, \mathbf{a}^e \in \mathbb{R}^{d+2} \}.$$

The Lagrange multiplier is used to solve the minimization problem $\mathcal{A}(\mathbf{Q}_c)$. Let $\boldsymbol{\gamma} \in \mathbb{R}^{d+2}$ be the Lagrange multiplier vector. Then the scalar objective function to be optimized is given by

$$L(\mathbf{Q}_c, \boldsymbol{\gamma}) = \sum_{j=1}^M |Q_{u,j} - Q_{c,j}|^2 + \boldsymbol{\gamma}^T (\mathbf{C}^e \mathbf{Q}_c - \mathbf{a}^e). \quad (44)$$

Eq. (44) can be solved explicitly for the corrected distribution value and the resulting equation of correction be implemented numerically in the code. Indeed, taking the derivative of $L(\mathbf{Q}_c, \boldsymbol{\gamma})$ with respect to $Q_{c,j}$, for $1 \leq j \leq M$ and γ_i , for $1 \leq i \leq d + 2$

$$\frac{\partial \mathbf{L}}{\partial \mathbf{Q}_{c,j}} = 0, \quad j = 1, \dots, M \Rightarrow \mathbf{Q}_c = \mathbf{Q}_u + \frac{1}{2}(\mathbf{C}^e)^T \boldsymbol{\gamma}. \quad (45)$$

Moreover,

$$\frac{\partial \mathbf{L}}{\partial \gamma_i} = 0, \quad i = 1, \dots, d+2 \Rightarrow \mathbf{C}^e \mathbf{Q}_c = \mathbf{a}^e,$$

retrieves the constraints. Hence, one needs to solve for $\boldsymbol{\gamma}$ the following equation

$$\mathbf{C}^e (\mathbf{C}^e)^T \boldsymbol{\gamma} = 2(\mathbf{a}^e - \mathbf{C}^e \mathbf{Q}_u). \quad (46)$$

Now, since $\mathbf{C}^e (\mathbf{C}^e)^T$ is symmetric and \mathbf{C}^e is an integration matrix, then \mathbf{C}^e is also positive definite. As a consequence, the inverse of $\mathbf{C}^e (\mathbf{C}^e)^T$ exists and one can compute the value of $\boldsymbol{\gamma}$ simply by

$$\boldsymbol{\gamma} = 2(\mathbf{C}^e (\mathbf{C}^e)^T)^{-1} (\mathbf{a}^e - \mathbf{C}^e \mathbf{Q}_u).$$

Substituting $\boldsymbol{\gamma}$ into (45) and recalling that $\mathbf{a}^e = \mathbf{0}$ for the elastic case,

$$\begin{aligned} \mathbf{Q}_c &= \mathbf{Q}_u + (\mathbf{C}^e)^T \left(\mathbf{C}^e (\mathbf{C}^e)^T \right)^{-1} (-\mathbf{C}^e \mathbf{Q}_u) = \left[\mathbb{I} - (\mathbf{C}^e)^T \left(\mathbf{C}^e (\mathbf{C}^e)^T \right)^{-1} \mathbf{C}^e \right] \mathbf{Q}_u : \\ &= \Lambda_N(\mathbf{C}^e) \mathbf{Q}_u, \end{aligned} \quad (47)$$

where $\mathbb{I} = N \times N$ identity matrix. In the sequel, we regard this conservation routine as *Conserve*. Thus,

$$\text{Conserve}(\mathbf{Q}_u) = \mathbf{Q}_c = \Lambda_N(\mathbf{C}^e) \mathbf{Q}_u. \quad (48)$$

Define D_t to be any time discretization operator of arbitrary order. Then, the discrete problem that we solve reads

$$D_t \mathbf{f} = \Lambda_N(\mathbf{C}^e) \mathbf{Q}_u. \quad (49)$$

Thus, multiplying (49) by \mathbf{C}^e it follows the conservation of observables

$$D_t(\mathbf{C}^e \mathbf{f}) = \mathbf{C}^e D_t \mathbf{f} = \mathbf{C}^e \Lambda_N(\mathbf{C}^e) \mathbf{Q}_u = 0, \quad (50)$$

where we used the commutation $\mathbf{C}^e D_t = D_t \mathbf{C}^e$ valid since \mathbf{C}^e is independent of time, see [Gamba and Tharkabhushanam \(2009\)](#) for additional comments.

4 LOCAL EXISTENCE, CONVERGENCE AND REGULARITY FOR THE SEMIDISCRETE SCHEME

In this section we enunciate L_k^1 and L_k^2 estimates for the approximation solutions $\{g_N\}$ of the problem (33) in the elastic case. The reader must refer to reference [Alonso et al. \(2016\)](#) for rigorous details. For the purpose of this presentation we use several well-known results that require different integrability properties for the angular kernel b , by assuming $b(\hat{u} \cdot \sigma)$ bounded from $\sigma \in S^{d-1}$ like it is the case for hard spheres in three dimensions (generalization for $b \in L^1(S^{d-1})$ can be

made at the cost of technical work). What is important that we work with hard potentials and so $1 \geq \lambda > 0$ in (9). The theory for Maxwell molecules $\lambda = 0$ needs a slightly different approach.

Recall that we have imposed conservation of mass, momentum and energy by building the operator $Q_c(g)$ with a constrained minimization procedure. Thus,

$$\int_{\Omega_L} g(v, t) \psi(v) dv = \int_{\Omega_L} g_0(v) \psi(v) dv$$

for any collision invariant $\psi(v) = \{1, v, |v|^2\}$. However, due to velocity truncation, the approximating solution g in general may be negative in some small portions of the domain. This is precisely the technical difficulty that we have to overcome. In Section 4.1, we present the statement of the proof of convergence of the proposed approximation in the number of modes N in a time interval $(0, T(L)]$ where $T(L)$ is a time depending on the lateral size L of the velocity domain Ω_L . We find a control on the negative mass that can be formed in such interval characterized in terms of L . In Sections 4.2 and 4.3, we improve the estimates assuming that the approximating solutions behaves well, that is, its negative mass does not increases too fast in the time interval in question.

4.1 Local Existence

Since the natural space to study the spectral scheme is $L^2(\Omega_L)$, thus we start proving that the problem is well posed in this space. Due to velocity truncation, we do not have the standard a priori estimates in L^1 that help in the theory, however, the constrain method permits to extend the time where the scheme gives an accurate solution of the original Boltzmann problem.

Proposition 1. *Let $g_0 \in L^2(\Omega_L)$ and fix the domain $(0, T(L)] \times \Omega_L$ with $T(L) \sim (L^{d+2(\lambda+1)} \|g_0\|_{L^2(\Omega_L)})^{-1}$. Then the approximating problem (33) has a unique solution $g \in \mathcal{C}(0, T(L); L^2(\Omega_L))$ with initial condition g_0 .^a In addition, the approximating sequence $\{g_N\}$ to solutions of (33), with initial condition $g_{0N} = \Pi^N g_0$, converges strongly in $\mathcal{C}(0, T(L); L^2(\Omega_L))$ as $N \rightarrow \infty$. In particular,*

$$\sup_{t \in [0, T(L)]} \|Q(Eg_N, Eg_N) - Q_u(g_N)\|_{L^2(\Omega_L)} \rightarrow 0 \text{ as } N \rightarrow \infty, \quad (51)$$

and the strong limit \bar{g} is the unique solution of the equation

$$\frac{\partial \bar{g}}{\partial t} = Q(E\bar{g}, E\bar{g})I_{\Omega_L} - \frac{1}{2} \left(\bar{\gamma}_1 + \sum_{j=1}^d \bar{\gamma}_{j+1} v_j + \bar{\gamma}_{d+2} |v|^2 \right), \quad \bar{g}(0) = g_0. \quad (52)$$

The coefficients of the quadratic polynomial are given in Lemma 1 with parameters (38) evaluated at $Q(E\bar{g}, E\bar{g})$. Furthermore, the negative mass of g is quantified as

$$\sup_{t \in [0, T(L)]} \|g^-\|_{L^2(\Omega_L)} = \|g_0^-\|_{L^2(\Omega_L)} + O_{d/2+\lambda+2} \|g_0\|_{L^2(\Omega_L)}. \quad (53)$$

^aNote that g actually depends on N since Q_c depends on N . We omit this dependence to ease notation.

4.2 Uniform Propagation of Numerical Unconserved Moments

We assume now that a solution $g \in \mathcal{C}(0, T; L^2(\Omega_L))$ for problem (33) with initial condition $g_0 \in L^2(\Omega_L)$ exists. The *conservation scheme* and the following stability condition implies that moments up to order 2 are controlled by the initial datum.

4.2.1 Stability Condition

We denote $T_\epsilon \in [0, T]$ the time where the smallness relation for the negative mass and energy of g and the boundedness of sequence $\{g_N\} := \{g\}$ in L^2 holds, that is for some fixed $\epsilon > 0$,

$$\sup_{t \in [0, T_\epsilon]} \frac{\int_{\{g < 0\}} |g(v, t)| \langle v \rangle^2 dv}{\int_{\{g \geq 0\}} g(v, t) \langle v \rangle^2 dv} \leq \epsilon, \quad \sup_{N \in \mathbb{Z}^+} \sup_{t \in [0, T_\epsilon]} \|g(t)\|_{L^2(\Omega_L)} < \infty. \quad (54)$$

Remark. This stability condition even holds for the scheme to compute the Boltzmann equation with an anisotropic grazing-Coulomb collision cross section (23) as shown in Fig. 2 in the grazing collision limit approximating the Landau equation. The conservation Spectral–Lagrangian scheme secures that after 800 mean free times, the value of $\epsilon < 0.05$ in the relation (54), meaning that the relative proportion of negative energy is very small, and so the solutions keeps essentially positive, and so stabilizes the scheme.

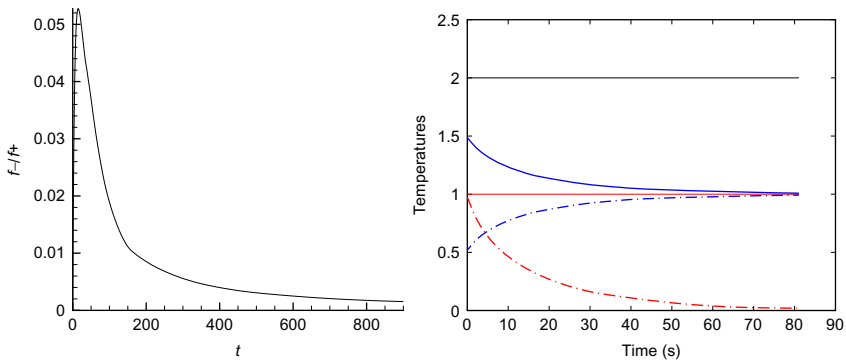


FIG. 2 Left: Ratio of energy in negative grid points to energy in positive grid points from the stability condition (54). The grazing parameter is $\varepsilon = 10^{-4}$, $N = 16$. Right: Temperatures evolution for the a benchmark component plasma system: solid blue, ion temperature T_i ; dash-dot blue, electron temperature T_e ; solid black, total conserved temperature $\bar{T} = T_i + T_e$; dash-dot red, temperature difference $|T_i - T_e|$ evolution. Left panel from Gamba, I.M., Haack, J.R., 2014. A conservative spectral method for the Boltzmann equation with anisotropic scattering and the grazing collisions limit. *J. Comput. Phys.* 270, 40–57. Right panel from Zhang, C., Gamba, I.M., 2016. A Conservative Scheme for Vlasov Poisson Landau Modeling Collisional Plasmas. *arXiv:1605.05787* (e-prints, section 7.1.2).

Indeed, for $k = \{0, 2\}$

$$\begin{aligned} \int_{\Omega_L} |g||v|^k &= \int_{\Omega_L} g_0|v|^k - 2 \int_{\Omega_L} g^-|v|^k \leq \int_{\Omega_L} g_0|v|^k \\ &+ 2\epsilon \int_{\Omega_L} g^+|v|^k \leq \int_{\Omega_L} g_0|v|^k + 2\epsilon \int_{\Omega_L} |g||v|^k. \end{aligned}$$

Hence, choosing $\epsilon \leq 1/4$ it follows

$$\int_{\Omega_L} |g(v,t)||v|^k dv \leq 2 \int_{\Omega_L} g_0|v|^k dv, \quad \text{for } t \in [0, T_\epsilon], \quad k = 1, 2; \quad (55)$$

and the following lemma holds.

Lemma 3 (Numerical moments bounds). *For any lateral size $L > 0$ and moment $k > 0$ there exist an extension E and a number of modes $N_0(T_\epsilon, L, k)$ such that*

$$\sup_{t \in [0, T_\epsilon]} \|g\|_{L_k^1(\Omega_L)} \leq C_k \left(\|g_0\|_{L_2^1}, m_{k'}(g_0) \right), \quad \forall N \geq N_0, \quad (56)$$

with $C_k(\cdot)$ a constant depending only on k , $\|g_0\|_{L_2^1}$, and $m_{k'}(g_0)$ with $k' = \max\{k, k_0\}$. The number $k_0 > 0$ it is uniquely determined by $\|g_0\|_{L_2^1}$.

The proof of this fundamental results relies on the conservative scheme estimate

$$\int_{\Omega_L} g(w,t)|v-w|^2 dw = \int_{\Omega_L} g_0(w)|v-w|^2 dw. \quad (57)$$

and condition (54) to obtain an uniform lower bound for the collision operator negative part. The following result implies stability of the conservative scheme as well as convergence to the equilibrium Maxwellian. This lower bound is shown in [Alonso et al. \(2016\)](#) with the assumption that the entropy $\int g(v) \log g(v) dv$ is bounded, since the numerical approximant $g(v)$ may not be positive at all its point of definition.

Lemma 4 (Lower bound for the discrete collision frequency). *Assume the uniform propagation of some moment $\frac{2+\mu}{\lambda}$, and that $\sup_{t \in [0, T_\epsilon]} \int_{\Omega_L} |g(w,t)||w|^{2+\mu} dw \leq C(g_0) < \infty$ for some $\mu > 0$. Then,*

$$\left(g_* |u|^\lambda \right) (v) \geq C(g_0) \langle v \rangle^\lambda, \quad (58)$$

with $C(g_0) > 0$ depending only on the mass, energy and the $\frac{2+\mu}{\lambda}$ -moment of g_0 .

4.3 Uniform L_k^2 Integrability Propagation

The result from [Lemma 4](#) is fundamental to obtain the following weighted Sobolev estimates for the approximate solution to the collisional equation.

Lemma 5 (L_k^2 -propagation estimates). *For any lateral size $L > 0$ and moment $k > 0$ there exist an extension E and a number of modes $N_0(T_e, L, k)$ such that*

$$\sup_{t \in [0, T_e]} \|g\|_{L_k^2(\Omega_L)} \leq \max \left\{ \|g_0\|_{L_k^2(\Omega_L)}, C_k(m_k(g_0)) \right\}, \quad N \geq N_0.$$

Moreover, the negative mass of g can be estimated as

$$\sup_{t \in [0, T_e]} \|g^-\|_{L^2(\Omega_L)} \leq e^{C(\|g_0\|_{L_2^1(\Omega_L)})T_e} \left(\|g_0^-\|_{L^2(\Omega_L)} + O_{d/2+\lambda k} \tilde{C}_k(m_{k+1}(g_0)) \max\{1, T_e\} \right), \quad N \geq N_0.$$

The constants C_k and \tilde{C}_k are independent of the asymptotic parameters T_e , L and N .

The next results gather the necessary information to estimate the propagation of higher order Sobolev regularity for the approximate solution, if initially so.

4.4 Uniform Semidiscrete H_k Sobolev Regularity Propagation

At last, we obtain the extend the discrete L_k^2 integrability estimates from [Lemma 5](#) to the derivatives of g . Indeed, the follow result has been shown as well.

Lemma 6 Assume $g_0 \in H_{k+2}^\alpha(\Omega_L)$ with $\alpha \in [0, \alpha_0]$ and $k \geq 0$. For any lateral size $L > 0$ there exist an extension E_{α_0} and a number of modes $N_0(T_e, L, k, \alpha)$ such that

$$\sup_{t \in [0, T_e]} \|g\|_{H_k^\alpha(\Omega_L)} \leq \max \left\{ \|g_0\|_{H_{k+2}^\alpha(\Omega_L)}, C_k(m_k(g_0)) \right\}, \quad N \geq N_0,$$

where $C_k(\cdot)$ depends on k and the k -moment of g_0 .

Remark. The initial restriction $\alpha \in [0, \alpha_0]$ is due to the fact that in general $Q(\text{Eg}, \text{Eg})$ possesses at most α_0 derivatives.

Finally, gathering the results of [Sections 3.4](#) and [3.5](#), with the results of global existence, L_k^1 and L_k^2 moment estimates, as well the higher order Sobolev regularity estimates from (6) of last section as well as spectral accuracy for the collisional integral shown in [Gamba and Tharkabhushanam \(2009\)](#), one can obtain both error estimates for the spectral scheme in the case of smooth and nonsmooth initial data, and convergence to the equilibrium Maxwellian (6). The first result removes the small negative mass and energy assumption (54) needed for the a priori estimates throughout the previous section. The results hold for any initial state $f_0(v)$ associated to the Cauchy problem for the Boltzmann equation, assumed to be $L^2(\mathbb{R}^d)$ and nonnegative (see [Alonso et al., 2016](#) for rigorous details).

5 FINAL COMMENTS AND CONCLUSIONS

The conservative spectral Lagrangian method for the Boltzmann equation was applied for a system of such equations in the modelling of a multienergy level gas (Munafo et al., 2014). In this case, the formulation of the numerical method accounts for both elastic and inelastic collisions. It was also used for the particular case of a mixture of monatomic gases without internal energy. The conservation of mass, momentum and energy during collisions is enforced through the solution of constrained optimization problem to keep the collision invariances associated to the mixtures (see Munafo et al., 2014, section 4.3). The effectiveness have been compared with the results obtained by means of the DSMC method and excellent agreement has been observed. More recently this conservative spectral Lagrangian approach has been implemented for a system of electron–ions in plasma modelled by a 2×2 system of Poisson–Vlasov–Landau equations (Zhang and Gamba, 2016), implemented by time-splitting methods staggering the time steps for advection of the Vlasov–Poisson system and the collisional system including recombinations. The constrained optimization problem is applied to the collisional step in a revised version from Gamba and Tharkabhushanam (2009) where the matrix \mathbf{C}^e defined in (43) was calculated in Fourier space given by the Fourier of the collision invariant polynomials to obtain a more accurate formulation. The benchmarking for the constrained optimization implementation for the mixing problem was done for an example of a space homogeneous system where the explicit decay the difference for electron and ion temperatures is known (Zhang and Gamba, 2016, section 7.1.2). Yet the used scheme captures the total conserved temperature being the sum of the Ions and electron temperatures, respectively (see Fig. 2, right side).

To end, we point out that the conservative spectral Lagrangian scheme for approximating solutions for the Boltzmann equation for elastic interactions converges to the equilibrium Maxwellian (6) if the equation is scalar, as is it shown in Theorem 2, part 4. One should note that it is the conservation subscheme the one that enforces the convergence to the equilibrium Maxwellian state by enforcing the collision invariants. This is exactly how the Boltzmann and H -theorems (Cercignani et al., 1994) work: the equilibrium Maxwellian (6) is proven to be the stationary state due to the conservation properties combined with the elastic collision law.

In other words for the case of inelastic collision (when the collision invariants are just $d + 1$) or for space inhomogeneous multicomponent Boltzmann systems flow models, it is not correct to assume that the stationary state is a Maxwellian distribution density (i.e. a Gaussian in v -space) as, for instance, asymptotic preserving schemes assume. Just the enforcing the conserved quantities for the system by the constrain minimization problem, the *Conservation Correction Estimate* of Lemma 2 will select the correct equilibrium states for each of the system components.

ACKNOWLEDGEMENTS

The author thanks Ricardo J. Alonso, Jeffrey Haack and S. Harsha Tharkabhushanam for their very valuable contributions. She has been partially supported by NSF under grants DMS-1413064 and NSF-RNMS 1107465. Support from the Institute of Computational Engineering and Sciences (ICES) at the University of Texas Austin is gratefully acknowledged.

REFERENCES

- Abramowitz, M., Stegun, I.A., 1964. Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables. National Bureau of Standards Applied Mathematics Series, No. 55 U.S. Government Printing Office, Washington, DC.
- Alonso, R.J., Gamba, I.M., Tharkabhushanam, S.H., 2016. Convergence and error estimates for the Lagrangian based conservative spectral method for Boltzmann equations. arXiv:1611.04171 (e-prints, submitted for publication).
- Aoki, K., Nishino, K., Sone, Y., Sugimoto, H., 1993. Numerical analysis of steady flows of a gas condensing on or evaporating from its plane condensed phase on the basis of kinetic theory: effect of gas motion along the condensed phase. In: Nonlinear PDE-JAPAN Symposium 2, 1991 (Kyoto, 1991) Lecture Notes Numer. Appl. Anal., No. 12. Kinokuniya, Tokyo, pp. 35–85.
- Aristov, V.V., 2001. Direct Methods for Solving the Boltzmann Equation and Study of Nonequilibrium Flows. Fluid Mechanics and its Applications, No. 60, Kluwer Academic Publishers, Dordrecht. ISBN 1-4020-0388-9, xviii+294. <http://dx.doi.org/10.1007/978-94-010-0866-2>.
- Bird, G.A., 1994. Molecular Gas Dynamics. Clarendon Press, Oxford.
- Bobylev, A.V., Rjasanow, S., 1999. Fast deterministic method of solving the Boltzmann equation for hard spheres. Eur. J. Mech. B. Fluids 18 (5), 869–887.
- Bobylev, A.V., Carrillo, J.A., Gamba, I.M., 2000. On some properties of kinetic and hydrodynamic equations for inelastic interactions. J. Stat. Phys. 98 (3–4), 743–773. ISSN 0022-4715.
- Bobylev, A.V., Gamba, I.M., Panferov, V.A., 2004. Moment inequalities and high-energy tails for Boltzmann equations with inelastic interactions. J. Stat. Phys. 116 (5–6), 1651–1682. ISSN 0022-4715.
- Bobylev, A., Cercignani, C., Gamba, I.M., 2009. On the self-similar asymptotic for generalized non-linear kinetic Maxwell models. Commun. Math. Phys. 291, 599–644. arXiv:math-ph/0608035.
- Brilliantov, N.V., Pöschel, T., 2004. Kinetic Theory of Granular Gases. Oxford Graduate Texts. Oxford University Press, Oxford. ISBN 0-19-853038-2, xii+329. <http://dx.doi.org/10.1093/acprof:oso/9780198530381.001.0001>.
- Cercignani, C., Reinhard, I., Pulvirenti, M., 1994. The Mathematical Theory of Dilute Gases. Applied Mathematical Sciences, No. 106. Springer-Verlag, New York.
- Chapman, S., Cowling, T.G., 1970. The Mathematical Theory of Non-Uniform Gases. An Account of the Kinetic Theory of Viscosity, Thermal Conduction and Diffusion in Gases, third ed. Cambridge University Press, London, xxiv+423.
- Cheng, Y., Gamba, I.M., Majorana, A., Shu, C.-W., 2009. A discontinuous Galerkin solver for Boltzmann-Poisson systems in nano devices. Comput. Methods Appl. Mech. Eng. 198 (37–40), 3130–3150. ISSN 0045-7825. <http://dx.doi.org/10.1016/j.cma.2009.05.015>.
- Cheng, Y., Gamba, I.M., Prof, J., 2012. Positivity-preserving discontinuous Galerkin schemes for linear Vlasov-Boltzmann transport equations. Math. Comput. 81 (277), 153–190. ISSN 0025-5718. <http://dx.doi.org/10.1090/S0025-5718-2011-02504-4>.

- Duffie, D., Malamud, S., Manso, G., 2009. Information percolation with equilibrium search dynamics. *Econometrica* 77 (5), 1513–1574. ISSN 0012-9682. <http://dx.doi.org/10.3982/ECTA8160>.
- Gamba, I.M., Haack, J.R., 2014. A conservative spectral method for the Boltzmann equation with anisotropic scattering and the grazing collisions limit. *J. Comput. Phys.* 270, 40–57.
- Gamba, I.M., Tharkabhushanam, S.H., 2009. Spectral-Lagrangian methods for collisional models of non-equilibrium statistical states. *J. Comput. Phys.* 228 (6), 2012–2036.
- Gamba, I.M., Tharkabhushanam, S.H., 2010. Shock and boundary structure formation by spectral-Lagrangian methods for the inhomogeneous Boltzmann transport equation. *J. Comput. Math.* 28, 430–460.
- Gamba, I.M., Panferov, V., Villani, C., 2004. On the Boltzmann equation for diffusively excited granular media. *Commun. Math. Phys.* 246 (3), 503–541.
- Gamba, I.M., Panferov, V., Villani, C., 2009. Upper Maxwellian bounds for the spatially homogeneous Boltzmann equation. *Arch. Ration. Mech. Anal* 194, 253–282.
- Graham, C., Méléard, S., 1999. Probabilistic tools and Monte-Carlo approximations for some Boltzmann equations. In: *ESAIM Proc., CEMRACS 1999 (Orsay)*, No. 10. Soc. Math. Appl. Indust., Paris, pp. 77–126. <http://dx.doi.org/10.1051/proc:2001010> (electronic).
- Landau, L.D., 1937. Kinetic equation for the case of Coulomb interaction. *Zh. Eks. Teor. Fiz.* 7, 203.
- Landau, L.D., Lifschitz, E.M., 1980. *Statistical Physics*, third ed. Butterworth-Heinemann.
- Morales Escalante, J.A., Gamba, I.M., 2016. Galerkin methods for Boltzmann-Poisson transport with reflection conditions on rough boundaries. [arXiv:1512.09210](https://arxiv.org/abs/1512.09210) (e-prints, submitted for publication).
- Morales-Escalante, J., Gamba, I.M., Cheng, Y., Majorana, A., Shu, C.-W., Chelikowsky, J., 2015. Discontinuous Galerkin deterministic solvers for a Boltzmann-Poisson model of hot electron transport by averaged empirical pseudopotential band structures. [arXiv:1512.05403](https://arxiv.org/abs/1512.05403) (e-prints, submitted for publication).
- Munafò, A., Haack, J.R., Gamba, I.M., Magin, T.E., 2014. A spectral-Lagrangian Boltzmann solver for a multi-energy level gas. *J. Comput. Phys.* 264, 152–176.
- Pareschi, L., Russo, G., 2000. Numerical solution of the Boltzmann equation. I. Spectrally accurate approximation of the collision operator. *SIAM J. Numer. Anal.* 37 (4), 1217–1245. ISSN 0036-1429. <http://dx.doi.org/10.1137/S0036142998343300>.
- Pulvirenti, M., Saffirio, C., Simonella, S., 2014. On the validity of the Boltzmann equation for short range potentials. *Rev. Math. Phys.* 26 (2), 1450001, 64. ISSN 0129-055X. <http://dx.doi.org/10.1142/S0129055X14500019>.
- Ringhofer, C., 2010. A level set approach to modeling general service rules in supply chains. *Commun. Math. Sci.* 8 (4), 909–930. ISSN 1539-6746. <http://projecteuclid.org.ezproxy.lib.utexas.edu/euclid.cms/1288725265>.
- Sone, Y., 2007. *Molecular Gas Dynamics. Modeling and Simulation in Science, Engineering and Technology*. Birkhäuser Boston Inc., Boston, MA. ISBN 978-0-8176-4345-4; 0-8176-4345-1, xiv+658.
- Stein, E.M., 1970. *Singular Integrals and Differentiability Properties of Functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, NJ, xiv+290.
- Villani, C., 1998. On a new class of weak solutions to the spatially homogeneous Boltzmann and Landau equations. *Arch. Ration. Mech. Anal.* 143 (3), 273–307. ISSN 0003-9527. <http://dx.doi.org/10.1007/s002050050106>.
- Zhang, C., Gamba, I.M., 2016. A conservative scheme for Vlasov Poisson Landau modeling collisional plasmas. [arXiv:1605.05787](https://arxiv.org/abs/1605.05787) (e-prints, submitted for publication).
- Zhang, C., Gamba, I.M., 2016. A conservative discontinuous Galerkin solver for space homogeneous Boltzmann equation. <http://www.ma.utexas.edu/users/gamba/papers/DGBE%5Fsubmit%5F2016.pdf> (submitted for publication).

Index

Note: Page numbers followed by “*f*” indicate figures, and “*t*” indicate tables.

A

- Acceleration methods, 318–323
 - enthalpy damping, 319–320
 - preconditioning, 320–321
 - residual averaging, 321–323
 - variable local time stepping, 318–319
- Accuracy effects, 358–360
- Acoustic–incompressible interactions
 - aerodynamical parameters, 208
 - barotropic flows, 209
 - classical iteration techniques, 209–210
 - Euler system, 208–209
 - filtered variable, 211–212
 - nonbarotropic flows, 209
 - pressure fluctuations, 209–212
 - quadratic operator, 209–210
 - velocity waves, 208
- Acoustic limit, 207–208
- Acoustic waves, 351
 - dimensional splitting, 58
- Active flux method, 76
- Adaptive algorithm
 - adaptive loop, 284–285
 - blast prediction, 294–296
 - boundary layer shock interaction, 290–294
 - direct sonic boom simulation, 288–290
 - double mach reflection, 294–296
 - transonic flow around M6 wing, 286–288
 - wing-body configuration, 285
- Adaptive loop
 - anisotropic quotients, 285
 - anisotropic ratio, 285
 - cavity-based operators, 284
 - Delaunay approach, 285
 - FEFLO, 284
 - frontal approach, 285
 - smoothing procedure, 284
 - for steady simulations, 284
 - for unsteady simulations, 284
 - WOLF, 284
- Adaptive mesh generator, 271–272
- Adaptive mesh refinement, 252–253
- Additional force terms, 468–469
- Additive-Schwarz approach, 338
- Aligned schemes, 491–493
- All-speed scheme, 218–220
- Alternating direction method, 313
- Ampère’s equation, 387, 395–397
- Anisotropic mesh
 - insertion and collapse, 280–282
 - unsteady simulations, 282–283
- Anisotropic mesh adaptation, 270
- Anisotropic mesh gradation, 272–273
- Anisotropic quotients
 - adaptive loop, 285
 - direct sonic boom simulation, 289–290, 289*f*, 291*t*
- Anisotropic ratio
 - adaptive loop, 285
 - direct sonic boom simulation, 289–290, 289*f*, 291*t*
- Antman–Cosserat curved rod model, 439*f*, 440–443, 442*f*
- Approximate-Schur approach, 338
- AREPO, 474
- Astrophysics
 - codes, 473–475
 - density and temporal scales, 466–467
 - equations
 - additional force terms, 468–469
 - source terms, 468
 - of state, 469
 - high-performance computing, 473
 - numerical methods
 - discontinuous Galerkin method, 471–472
 - finite difference methods, 469–470
 - finite volume methods, 470–471
 - grid-free method, 472–473
 - N-body method, 472
 - spatial scales, 466
- Asymptotic-preserving (AP) schemes
 - asymptotic limits
 - compressible flows, low mach number limit of, 121–122

Asymptotic-preserving (AP) schemes

(Continued)

high-field limit, 119–120

linear transport equation, diffusion limit of, 118–119

plasmas, quasi-neutral limit in, 120–121
stochastic AP schemes, 122–123Boltzmann equation, coupling of, 104–105
design principles

Bhatnagar–Gross–Krook (BGK) model, 107–110

Jin–Xin relaxation model, 105–107

hyperbolic and kinetic equations

exponential reformulation, 115–116

micro-macro decomposition, 117–118

penalization, 111–114

scales, 104–105, 104*f*

ATHENA, 474

AUSM, 307

B

Balance law, 152, 437–438, 468, 480–481

Baldwin–Lomax turbulence model, 333*f*

Barotropic flows, 209

Barrow's rule, 163–166

Batch sedimentation model, 483

Bhatnagar–Gross–Krook (BGK) model, 111–112, 114

implicit–explicit (IMEX) scheme, 109

kinetic flux vector splitting (KFVS) scheme, 108–109

kinetic theory, 107–108

Knudsen number, 107–108

nonstiff convection, 108

stiff collision, 108

Strang splitting, 109–110

Bicharacteristic methods, 67–68, 68*f*

Bilinear interpolation, 62

Binary collisional models

collision invariants, 406

collision kernel, 407

decorrelation, 406

equilibrium Maxwellian distribution, 406

Fourier transform, 407

Blast prediction, 294–296

Bloch waves techniques, 569–570

Body-fitted grids, 24

Boltzmann equation, 408–409

deterministic solvers, 409–410

doublemixing convolution, 411

hyperbolic and kinetic equations, 110–111

Landau–Fokker–Plank equation, 410–411

penalization method, 113–114

stiff term, 117

weighted convolution, 411–412

Boris–Book–Zalesak method, 93–94

Boris correction, 387–388

Boundary layer shock interaction, 290–294, 293–294*f*

Boundary layers mesh generation, 278–280

Boundary representation (BREP), 266–267, 267*f*

Bound-preserving flux limiting approach, 93–94

Bound-preserving property

flux limiters

decoupling for, 95–97

idea and framework, 93–95

limiter for approximation polynomials

extensions and applications, 92–93

first-order monotone schemes, 83–84

simple and efficient scaling limiter, 86–91

SSP high-order time discretizations, 91–92

weak monotonicity, in high-order finite volume schemes, 84–86

Boussinesq-type equation, 569–570

Burgers equation, 32*f*, 33, 33*t*

Butcher tableaux, 116

C

Cart3D, 5–7

Cartesian grids, 24, 389–390

Cartesian mesh methods, 2, 27, 38–39, 469–470

finite difference methods, 469–470

Cauchy–Kowalevski procedure, 150–151

Cavity-based operators, 284

Cell-centred triangular meshes, 224–226, 225*f*

Cell linking, 13–14

Cell residual, 69

Chapman–Enskog expansion, 111

Chebyshev polynomials, 147

Clarifier–thickener unit, 483

Collision operator, 111–112

Compressible Euler system, 205–221

Computational aerodynamics

gas dynamics and spatial discretizations, 305–308

Newton–Krylov methods

additional considerations and algorithm parameters, 340–342

globalization, 338–339

- GMRES, 334, 343*f*
- implicit methods, 335
- inexact, 335–336
- Jacobian-free, 336–337
- NASA Common Research Model, 343, 344*f*
- nonsymmetric linear systems, 334
- parallelization, 337–338
- parallel scaling algorithm, 343, 344*f*
- preconditioning, 337–338
- pseudo-transient continuation, 334
- Spalart–Allmaras one-equation turbulence model, 342
- steady and unsteady flows, 332
- time-marching methods
 - acceleration methods, 318–323
 - alternative approach, 304–305
 - implicit schemes, 312–318
 - model problem, 308–309
 - multigrid methods, 323–327
 - multistage schemes, 309–312
 - RANS equations, 327–332
- Computational domain, 491
- Computational fluid dynamics (CFD), 204, 264–265
- Computational pipeline, 264–265, 264–265*f*
- Computational uncertainty propagation, 510–511
- Conservation laws, 364–365, 447–448
- Conservation laws with discontinuous coefficients
 - clarifier–thickener unit, 483
 - flows, 485
 - interface connection, 489
 - interface entropy condition, 489
 - interior entropy condition, 487–488
 - ion etching, 483
 - Kruzkov entropy condition, 490
 - multiphase flows, in porous media, 481–482
 - nonuniqueness, 486
 - numerical experiments, 496–500, 496*f*, 500*f*
 - numerical methods
 - aligned schemes, 491–493
 - finite volume schemes, 495–496
 - front tracking scheme, 496
 - higher-order schemes, 494–495
 - multidimensional problems, 495
 - staggered schemes, 493–494
 - time-dependent coefficients, 495
 - oscillations, 486
 - scalar conservation law, 485
 - singular source terms, 484
 - traffic flow, 482–483
 - wave propagation, in heterogeneous media, 484
 - weak solution, 487
- Conservative flux difference approximation, 366–367
- Conservative spectral method
 - for collision method, 415–426
 - computational cut-off domain, 416–418
 - extension operator, 419
 - Fourier series, 418–419
 - projections and extensions, 418–419
 - extended isoperimetric problem, 421–424
 - for homogeneous Boltzmann equation, 418–419
 - equilibrium Maxwellian, 421
 - error estimation, 420–421
 - extension operator, 419–420
 - gain operator, 419–420
 - semidiscrete problem, 420
- Continuity equation, 218, 387
- Convection–diffusion equation
 - applications to, 194–197
 - diffusion-velocity particle method, 197
 - weighted particle method, 195–196
- Cordes-type condition, 552
- Coriolis force, 362–363
- Corner transport methods
 - bilinear interpolation, 62
 - gas-dynamic equations, 63–64
 - linear advection equation, 62, 63*f*
 - multidimensional Riemann problems, 63, 64*f*
- Courant–Friedrichs–Lewy (CFL) condition, 31, 321
- Courant, Isaacson and Rees (CIR) scheme, 54, 65
- C-property scheme, 133–134, 509
- upwind discretization, 134
- Curvilinear transformation, 24
- Cut-cells
 - Cartesian mesh methods, 2
 - data structure, 8–10
 - early history, 3–4
 - embedded boundary mesh, 2–3
 - finite volume methods
 - dual time stepping, 11
 - explicit time-dependent solution techniques, 13–17
 - stability problem, 10–11
 - steady-state problem, 10
 - steady-state solution techniques, 12–13
 - viscous flows, 17–18

Cut-cells (*Continued*)
 implementation issues, 8–10
 mesh generation, 1–2
 algorithm, 5–6
 Cart3D, 5–7
 Cartesian mesh generator, 5, 7, 8f
 geometry, 5
 triangulation, 8
 problem, 25
 Cylinder lift-off problem, 45–46, 47–48f, 47t

D

Dalton law, 353–354
 Darcy's law, 481–482
 Delaunay method, 268, 269f, 285
 Density scales, 466–467
 Diagonal matrix, 144
 Dichotomy approach, 280
 Diffusion limit, 118–119
 Diffusion-velocity particle method, 179, 197
 Dimensional splitting
 acoustic wave, 58
 Glimm's Random Choice method, 59
 Godunov-type methods, 58
 operator splitting, 60
 superfluous dissipation, 59–60, 60f
 Dimensional upwinding, 60–61
 Dirac's deltas, 138–139
 Direct sonic boom simulation
 anisotropic quotient, 289–290, 289f, 291t
 anisotropic ratio, 289–290, 289f, 291t
 final adapted mesh, 289–290, 289f, 291t
 initial surface mesh, 288, 288f, 293f
 interpolation error, 289–290
 pressure extractions, 289–290, 292f
 SSBJ design, Dassault Aviation, 288, 288f
 Discontinuous Galerkin (DG) method,
 367–368, 393–394, 451–460, 471–472
 applications
 to traffic network, 454–455
 to water channel, 455–456
 for nonlinear wave equation net problem,
 452–454
 numerical coupling conditions at junctions,
 456
 PDE-ODE coupling conditions, 456–460
 RK time discretization in time, 455
 Discrete continuity equation, 387
 Discrete velocities methods (DVM), 409–410
 Discretization, MHD equations, 397–398
 Dissipation, 55
 Dissipation matrix, 220–221

Divergence theorem, 470
 Double Mach reflection problem, 42–43, 43f,
 294–296, 295f
 density contours of, 43–44, 44f
 Dual control volume, 223
 Dual problem
 numerical approximation, 251–252
 Dual-weighted-residual (DWR) error
 estimation technique
 BAC3-11 airfoil, 254
 bifurcation problems, 255–256
 criticality problems, 255

E

Edge-based formula, 65–66
 Edge collapse, 280–282, 281f
 Eigenvalues, 152
 Eigenvectors, 152
 Einstein summation rule, 570–571
 Electrostatic plasma, 119–120
 Elliptic–hyperbolic splitting, 73–75, 75f
 Enthalpy damping, 319–320
 Entropy conservative methods, 148
 Entropy-stable methods, 148
 Entropy variables, 140
 ENZO, 474
 Equation of state, 469
 Error representation, 235
 abstract framework
 advantages and disadvantages, 237–238
 bilinear form and functional, 236–237
 consistent reformulation, 237
 discretization, 237–238
 dual problems, 239
 foregoing assumptions, 239
 linear operator, 236
 posteriori error estimation, 243–244
 stabilized FEMs
 bilinear form, 241
 definition, 241–242
 dual problems, 242
 linear functional, 241
 weak formulation, 240–241
 Essentially non-oscillatory (ENO) schemes,
 29, 83, 86, 391
 Euler code, 74–75
 Euler equations, 92–93, 95, 179, 304–306, 320
 micro–macro decomposition, 117
 multiple low MACH number limits,
 205–221
 potential limits, 216
 Euler–Poisson equations, 120–121

Exact C-property, 369. *See also* Well-balanced methods

Explicit time-dependent solution techniques
accuracy, 16

cancellation property, 14–15

cell merging, 13–14

dimensionally split, 16–17

flux redistribution, 17

h -box method, 14–15

wave propagation, 14

Exponential reformulation method, 115–116

F

Faraday equation, 395

Fast vortex transport, 76–78, 77*t*

Favre-averaged (FANS) equations, 308

FEFLO, 284

FEM. *See* Finite element method (FEM)

Filtered variable, 211–212

Final adapted mesh, 289–290, 289*f*, 291*t*

Finite difference method (FDM), 366–367, 564
astrophysics, 469–470

Finite-difference numerical homogenization
method

short-time wave propagation, 567–569

Finite element exterior calculus (FEEC),
394–395, 397–398

Finite element heterogeneous multiscale
method (FE-HMM)

FE-HMM-L, 569–574

short-time wave propagation, 565–566

Finite element method (FEM), 394–396, 442,
564

Finite-element numerical homogenization
method, 565–566

Finite volume (FV) discretization scheme,
390–393

advantages, 393

essentially non-oscillatory (ENO) schemes,
391

Gauss theorem, 390–391

Maxwell equations, 390

numerical flux, 391

predictor–corrector formulation, 392–393

Runge–Kutta schemes, 392–393

vector-valued functions, 392

Finite volume method (FVM), 213–214,
365–366, 451

astrophysics, 470–471

dual time stepping, 11

explicit time-dependent solution techniques,
13–17

stability problem, 10–11

steady-state problem, 10

steady-state solution techniques, 12–13
viscous flows, 17–18

First-order finite volume method, 509–510

First-order IMEX discretization, 112

First-order monotone schemes, 83–84

FLASH, 474

Flux function, 480

aligned schemes, 492

in numerical experiments, 496–500, 496*f*,
500*f*

Flux splitting scheme, 367, 391, 495–496

Flux tensor, 64–65

Flux-transfer transformations

Green's function, 557–558

transfer property, 555–556

Fokker–Planck–Landau equation, 113–114

collision operator, 113–114

Forward UQ. *See* Computational uncertainty
propagation

Free surface flows, 362–363

mathematical model, 363–364

Frontal approach, 268, 270*f*, 285

Front tracking scheme, 496

Functional viscosity matrix methods, 145–148

Chebyshev polynomials, 147

intermediate matrix, 145

local Lax–Friedrichs method, 145–146

path-conservative numerical schemes,
145–148

polynomial viscosity matrix (PVM), 147, 147*t*

Roe method, 145

well-balanced schemes, 157–158

G

GADGET, 473

Galaxies, 466

Gas-dynamic equations, 63–64

Gas dynamics and spatial discretizations,
305–308

Gauss law, 387, 395

Gauss–Legendre quadrature formula, 148

Gauss–Lobatto quadrature oints, 90

Gauss–quadrature rules, 371

Gauss–Seidel methods, 314–315

Gauss theorem, 390–391

Generalized hydrostatic reconstruction (GHR),
158–160

Generalized polynomial chaos (gPC)

expansion, 123, 516

entropy variables, 518–520

Geometric estimate, for surfaces, 277–278
 Ghost points
 high-order interior schemes, 24
 inflow boundary conditions, 28–29
 moving boundary treatment, 39, 39f
 outflow boundary conditions, 29–30
 Glimm's Random Choice method, 59
 GMRES, 334, 343f
 Godunov fluxes, 492
 Godunov method, 142
 dimensional splitting, 58
 relative errors, 134, 135t
 Godunov solver, 354–356
 Godunov's theorem, 84
 Goldstein–Taylor model, 122–123
 Gravity, 468
 Grazing collision limit, 413–414
 Grid-free method, 472–473

H

Harmonic coordinate transformations, 551–556
h-box method, 25
 HCUSP, 307
 Helm–Helmholtz decomposition, 556
 Heterogeneous multiscale methods (HMM), 564
 Higher-order schemes, 148–151, 494–495
 reconstruction of states, 148–151
 Higher order spatial accuracy, 510
 Higher order temporal accuracy, 510
 High-field limit, 119–120
 High-order well-balanced schemes
 MUSCL reconstruction, 163, 164–165f
 1d stationary problem, 163–166
 quadrature formula, 163
 reconstruction operator, 161
 shallow water system, 161
 stationary solutions, 161–162
 High-performance computing, 473
 Homogeneous equations, numerical methods
 discontinuous Galerkin (DG) methods,
 367–368
 finite difference methods, 366–367
 finite volume methods, 365–366
 residual distribution (RD) schemes, 368
 Hydrostatic pressure, 378
 Hydrostatic reconstruction, 369–370
 Hyperbolic and kinetic equations
 exponential reformulation, 115–116
 micro–macro decomposition, 117–118
 penalization, 111–114
 Hyperbolic conservation, 437
 one dimensional, 437–438

Hyperbolic correction, 388–389
 Hyperbolic nets, 436–437, 437f
 discontinuous Galerkin method, 451–460
 from 3D to 1D net problem, 439
 3D vs. 1D net problem, 442–443, 442f
 dynamic coupling conditions, 441
 finite volume methods, 451
 kinematic coupling conditions, 441
 nonlinear wave equation net problem,
 443–446, 444f
 weak formulation, 444–446, 445f
 one dimensional, 437–439
 vascular stent net problem, 439–443
 Hyperbolic networks, 436–437
 blood flow, 448–450
 data flow on telecommunication
 networks, 447–448
 discontinuous Galerkin method,
 451–460
 finite volume methods, 451
 of irrigation channels, 448
 models for vehicular traffic, 446–447
 one dimensional, 436–437

I

Immersed boundary method (IBM), 24
 Immersed interface method, 24
 Implicit–explicit (IMEX) scheme, 109
 Incomplete lower-upper (ILU), 337
 Cuthill–McKee ordering, 340
 Schur parallel preconditioner, 340
 Incompressible limit, 205–207
 Inexact Newton method, 335–336
 Inflow side, 70–71
 Inf-sup criterion, 219–220
 Initial surface mesh, 288, 288f, 293f
 Interface connection, 489–490
 Interface entropy condition, 489
 Interface pressures, 66
 Intergalactic medium, 466
 Interior entropy condition, 487–488
 Intermediate matrix, 145
 Interpolation error, 274–275
 control, 276–277
 direct sonic boom simulation, 289–290
 Interstellar medium (ISM), 467
 Inverse Lax–Wendroff procedure (ILW)
 Cartesian embedded boundary method,
 26–27
 fifth-order ILW, 42–48
 for inflow boundary conditions, 28–29
 interior schemes, 27

- moving boundary treatment, for
 - compressible inviscid flows, 38–42
- numerical boundary conditions, for static geometry
 - one-dimensional scalar conservation laws, 28–33
 - two-dimensional euler equations, 34–38
- Ion etching process, 483
- Irrigation channels, hyperbolic networks of, 448

J

- Jacobian-Free Newton–Krylov methods, 336–337
- Jacobian matrix, 35, 363–364
 - systems of equations, 72
- Jameson–Schmidt–Turkel (JST) scheme, 308
- Jin–Xin relaxation model, 105–107
- Jones–Wilkins–Lee (JWL), 355–356
- Journal of Computational Physics, 54
- Jump condition, 167

K

- Kapila multiphase model, 168
- Kelvin–Helmholtz problem, 534–536, 535*f*
- Kinetic evolution models, 404–405
- Kinetic flux vector splitting (KFVS) scheme, 108–109
- Knudsen number, 107–108
- k-point correlation marginal, 537

L

- Lagrange extrapolation, 25–26
 - for outflow boundary conditions, 29–30
- Lagrange interpolation polynomials, 89–90
- Lagrange multiplier method, 388–389, 425–426
- Lagrangian-type method, 178
- Landau–Fokker–Plank equation, 410–411
- Latitude–longitude mesh, 379
- Lax–Friedrich flux, 27, 83–84, 95, 366
- Lax–Friedrichs scheme, 495
- Lax–Oleinik entropy, 486
- Lax–Wendroff theorem, 352, 372–373
- Lebesgue constant, 89
- Legendre Gauss–Lobatto quadrature rule, 376
- Level set methods, 279, 355
- Lighthill–Whitham–Richards (LWR) model, 446, 482–483
- Linear advection equation, 62, 63*f*
- Linearity-preserving (LP), 71

- Linear stability, 31
- Linear transport equation, 118–119, 181
- Lituya Bay mega-tsunami, 533–534, 534*f*
- Local coordinate system, 34, 34*f*
- Localized orthogonal decomposition (LOD)
 - ε -independent error estimate, 559
 - fine scale discretization errors, 560
 - optimal error estimates, 560
 - regularity-independent estimate, 559
 - time-discretizations, 560–561
- Local Lax–Friedrichs method, 145–146
- Log-Euclidean framework, 273–274
- Long-time wave propagation
 - Boussinesq-type equation, 569–570
 - computational complexity, 574
 - dispersive effects, 573–574
 - fast Fourier transform algorithm, 573–574
 - finite difference approximation, 573–574
 - generalized FD-HMM scheme, 569–570
 - numerical homogenization methods, wave equations, 572–574
- Low-order coupling methods, 48–49
- LU decomposition methods, 313–314
- LWR model. *See* Lighthill–Whitham–Richards (LWR) model

M

- Mach number, 204, 228
 - acoustic limit, 207–208
 - in subcritical flow, 55–56, 56*f*
- Magnetohydrodynamics (MHD), 474
 - discretization, 397–398
 - model, 396–397
- Mantel integral, 67–68
- Mass conservation, 215
- Material waves, 351
- Mathematical model, 363–364
- Maxwellian function, 115–116
- Maxwell’s equations, 469
 - Boris correction, 387–388
 - Cartesian grids, 389–390
 - discontinuous Galerkin (DG) schemes, 393–394
 - finite element methods, 394–396
 - FV schemes, 390–393
 - hyperbolic correction, 388–389
 - model, 386–387
- Measure-valued solution, 536–537
- Menter SST models, 308
- Mesh-free particle method, 178

- Mesh generation, 1–2
 - algorithm, 5–6
 - Cart3D, 5–7
 - Cartesian mesh generator, 5, 7, 8*f*
 - geometry, 5
 - triangulation, 8
 - Metric-based mesh adaptation
 - boundary layers metric, 278–280
 - error estimates, 274–276
 - geometric estimate, for surfaces, 277–278
 - interpolation error, 276–277
 - metric tensors in, 271–272
 - robustness and performance, 272–274
 - Metric tensors, 271–272
 - Micro–macro decomposition, 117–118
 - Mimetic differencing, 64–65
 - Minmod function, 71–72
 - MLMC-FVM algorithm, 526–527
 - Momentum conservation, 215–216
 - Momentum equation, 218
 - Momentum interpolation method, 218–219
 - Monte Carlo method
 - error and complexity analysis, 525–526
 - probability space, 524
 - Moving boundary treatment, for compressible
 - inviscid flows
 - Cartesian mesh, 38–39
 - fifth-order boundary treatment, 39, 39*f*, 42
 - ghost points, 39, 39*f*
 - newly emerging points, 39, 39*f*
 - no-penetration boundary condition, 38–39
 - Moving water equilibrium, 372–374
 - Multidimensional effects, 356–358
 - Multidimensional physics, 69
 - Multidimensional problems, 495
 - Multidimensional Riemann problems, 63, 64*f*
 - Multidimensional system
 - balance laws with a discontinuous flux, 499
 - conservation laws with source term, 498
 - Multidimensional upwinding
 - bicharacteristic methods, 67–68
 - CIR scheme, 65
 - corners, 64–65
 - corner transport methods, 62–64
 - dimensional splitting, 58–60
 - dimensional upwinding, 60–61
 - edges, 64–65
 - oblique wave methods, 61–62
 - Poisson formulas, 75–78
 - residual distribution
 - elliptic–hyperbolic splitting, 73–75
 - NN scheme, 71–72
 - N scheme, 70–71
 - systems of equations, 72
 - unsteady problems, 72–73
 - wave models, 73
 - Multilevel Monte Carlo (MLMC) method
 - efficient implementation, 528
 - error and complexity analysis, 527–528
 - MLMC-FVM algorithm, 526–527
 - parallelization, 528
 - Multiphase flows, in porous media, 481–482
 - Multiple low MACH number
 - acoustic–incompressible interactions, 208–213
 - acoustic limit, 207–208
 - diagnosis, 215–217
 - finite volume schemes, 213–214
 - incompressible limit, 205–207
 - remedies, 217–221
 - Multiscale finite element method (MsFEM), 553–555
 - Multiscale methods, heterogeneous media
 - Aubin–Nitsche duality argument, 546–548
 - computational complexity, 548
 - finite element space, 546–548
 - finite element spaces, 546–548
 - higher order spatial approximations, 546–548
 - leap-frog method, 548
 - multiscale coefficient, 546–548
 - oscillatory hyperbolic problems, 546–548
 - MUSCL method, 163, 164–165*f*, 307
- ## N
- Naive explicit scheme, 105–106
 - NASA Common Research Model, 343, 344*f*
 - Navier–Stokes equations, 121–122, 196, 304–308, 363, 448–449
 - dimensional upwinding, 60–61
 - N-body method, 472
 - Newton–Krylov methods
 - additional considerations and algorithm parameters, 340–342
 - globalization, 338–339
 - GMRES, 334, 343*f*
 - implicit methods, 335
 - inexact, 335–336
 - Jacobian-free, 336–337
 - NASA common research model, 343, 344*f*
 - nonsymmetric linear systems, 334
 - parallelization, 337–338
 - parallel scaling algorithm, 343, 344*f*
 - preconditioning, 337–338
 - pseudo-transient continuation, 334

- Spalart–Allmaras one-equation turbulence model, 342
 - steady and unsteady flows, 332
 - NIRVANA, 474
 - NN scheme, 71–72
 - Nonbarotropic flows, 209
 - Nonflat bottom topography, 368–369
 - Nonlinear equation, of state, 355–356
 - Nonlinear hyperbolic conservation laws
 - dual problem, 249–250
 - formal adjoint problem, 249
 - Fréchet derivative, 248
 - Galerkin orthogonality property, 247–248, 250
 - mean-value linearization, 248
 - one-dimensional scalar hyperbolic equation, 249
 - PDE problem, 247–248
 - Nonlinear wave equation net problem, 443–446, 444*f*
 - discontinuous Galerkin method for, 452–454
 - weak formulation, 444–446, 445*f*
 - Nonstandard discretization, 370
 - Nonstiff convection, 108
 - Nonsymmetric linear systems, 334
 - Nonuniqueness, 486
 - No-penetration boundary condition, 37
 - for inviscid flows, 38–39
 - N scheme, 70–71
 - Numerical extrema, 82–83
 - Numerical fluxes, 93–94
 - aligned schemes, 491–492
 - higher-order schemes, 494
 - staggered schemes, 493
 - Numerical homogenization method
 - asymptotic expansion, 564
 - long-time wave propagation, 569–574
 - short-time wave propagation
 - FD-HMM, 567–569
 - FE-HMM, 565–566
 - Numerical illustrations
 - cell-centred triangular meshes, 224–226
 - quadrangular Cartesian grids, 221–222
 - vertex-centred triangular meshes, 223–224, 226–227
 - Numerical methods
 - astrophysics
 - discontinuous Galerkin method, 471–472
 - finite difference methods, 469–470
 - finite volume methods, 470–471
 - grid-free method, 472–473
 - N-body method, 472
 - homogeneous equations, 364–368
 - positivity-preserving methods, 374–377
 - well-balanced methods, 368–374
- ## O
- Oblique wave methods, 61–62
 - ODE system, 178
 - Ohm's law, 396–397
 - One-dimensional scalar conservation laws, 496–497
 - with discontinuous flux, 497–498, 497*f*
 - with flux function, 496–497, 496*f*
 - smooth solutions
 - inverse Lax–Wendroff procedure, for inflow boundary conditions, 28–29
 - Lagrange Extrapolation, for outflow boundary conditions, 29–30
 - linear stability, 31
 - simplified inverse Lax–Wendroff procedure, 30–31
 - solutions containing discontinuities, 31–33
 - well-posedness theory for, 501
 - One dimensional sweeps, 356
 - Onera M6 model, 266–267
 - Operator splitting, 60, 76–78
 - Optimal entropy connection, 496–498
 - Oscillations, 224, 486
 - Osher–Solomon scheme, 148
- ## P
- Parallelization, 473
 - Parallel scaling algorithm, 343, 344*f*
 - Parametric representation, 267, 268*f*
 - Parametrized flux limiter, 95
 - Partial differential equations (PDEs), 26, 234, 264–265
 - metric-based error estimates, 275
 - shallow water fluids, 132–133
 - Particle distortion, remeshing for, 190–194
 - Particle function approximation, 186–190
 - accuracy, 188–189
 - efficiency, 189–190
 - smooth vs. discontinuous solutions, 189
 - Particle merger, local redistribution technique, 193–194
 - Particle method, 178–179, 181–190
 - particle approximation of initial data, 182
 - particle function approximation, 186–190
 - accuracy, 188–189
 - efficiency, 189–190
 - smooth vs. discontinuous solutions, 189
 - time evolution of particles, 183–186

- Particle weights redistribution, 191–193
 - convolution, 191–192
 - interpolation, 192–193
 - Path-conservative methods, 369–370
 - Path-conservative numerical schemes, 138–142
 - convergence and choice, 166–169
 - entropy conservative methods, 148
 - entropy-stable methods, 148
 - functional viscosity matrix methods, 145–148
 - Godunov method, 142
 - Osher–Solomon scheme, 148
 - Roe methods, 144–145
 - simple Riemann solvers, 142–144
 - PDE-ODE coupling conditions
 - moving bottleneck simulation approaches, 457–460, 460*f*
 - toll-gates and flux constraints, 456–457
 - Penalization method
 - BGK operator, 111–112, 114
 - first-order IMEX discretization, 112
 - Fokker–Planck–Landau equation, 113–114
 - gas dynamics, 113–114
 - linear/simpler operator, 113–114
 - for nonlinear Boltzmann equation, 111–112
 - nonlinear hyperbolic system, 113
 - physical viscosity, 113–114
 - stiff relaxation, 113–114
 - Petrov–Galerkin problem, 558
 - Piecewise linear reconstruction, 494
 - PLUTO, 474
 - Point insertion, 280–282, 281*f*
 - Point smoothing, 282–283
 - Point values, 86
 - Poisson equation, 387–388
 - Poisson formulas, 75–78
 - Euler equations, application to, 76–78
 - Polymer flooding model, 495–496
 - Polynomial approximation, 393–394
 - Polynomial reconstruction, 358–359
 - Polynomial viscosity matrix (PVM), 147, 147*t*
 - Positivity-preserving methods, 470–471
 - mesh adaption technique, 374–375
 - SWEs, 374–375
 - Posteriori error estimation
 - classification, 243
 - error representation formula, 243–244
 - indicator, 246–247
 - linear advection, 246*f*
 - numerical scheme, 243–244
 - Type I, 244
 - Type II, 245–246
 - Postprocessing reconstruction, 374–375
 - Preconditioned dissipation, 217–218
 - Predictor–corrector formulation, 392–393
 - Pressure contour, 45, 46–47*f*
 - Pressure extractions, 289–290, 292*f*
 - Pressure fluctuations, 209–212, 222, 222–223*f*
 - cell-centred triangular meshes, 224–225, 225*f*
 - isolines, 226, 227*f*
 - vertex-centred triangular meshes, 226–227, 227*f*
 - Primitive variables, 37
 - Pseudo-transient continuation, 334
- Q**
- Quadrangular cartesian grids, 221–222
 - Quadratic surface model, 277–278
 - Quasi-neutral limit, in plasmas, 120–121
 - Quasi-neutral regime, 120
- R**
- RAMSES, 474
 - Random entropy solutions, 514–515
 - Rankine–Hugoniot conditions, 45, 139, 486
 - Rational viscosity matrix (RVM) methods, 148
 - Raviart–Thomas elements, 394–395
 - Real gas effects
 - gases, mixture of, 353–355
 - nonlinear equation, of state, 355–356
 - Reconstruction operator, 149
 - Reconstruction polynomial, 82–83
 - Residual averaging, 321–323
 - Residual distribution
 - elliptic–hyperbolic splitting, 73–75
 - NN scheme, 71–72
 - N scheme, 70–71
 - systems of equations, 72
 - to unsteady problems, 72–73
 - wave models, 73
 - Residual distribution (RD) scheme, 55–56, 368
 - Reynolds-averaged (RANS) equations, 308
 - Riccati equation, 320
 - Riemann invariants, 105–106
 - Riemann problem (RP), 142, 153, 364, 446–447
 - aligned schemes, 492
 - staggered schemes, 493–494
 - Riemann solvers, 470–471
 - accuracy effects, 358–360
 - multidimensional effects, 356–358
 - real gas effects
 - gases, mixture of, 353–355
 - nonlinear equation, of state, 355–356

Rieper fix, 220–221
 Rieper scheme, 226
 Rock permeability, 481–482
 Roe linearization, 144
 Roe matrix, 156–157, 328–329
 Roe methods, 213–214, 217–218, 226
 path-conservative numerical schemes, 144–145
 well-balanced schemes, 156–157
 Roe–Tukel scheme, 217–220
 Runge–Kutta discontinuous Galerkin (RKDG) method, 83
 Runge–Kutta (RK) method, 115–116, 311, 377, 392–393, 451–452
 Runge–Kutta (RK) time discretization, 455
 Rusanov scheme, 357

S

Saint-Venant equations. *See* Shallow water equations (SWEs)
 Saint-Venant system of equation, 448
 Scalar conservation law, 82, 93–94
 Second-order finite volume method, 471
 Second-order hyperbolic problems.
 See Multiscale methods, heterogeneous media
 Shallow water equations (SWEs)
 difficulties, 362–363
 mathematical model, 363–364
 numerical methods
 homogeneous equations, 364–368
 positivity-preserving methods, 374–377
 well-balanced methods, 368–374
 roles, 362–363
 shallow water
 flows, channels with irregular geometry, 377–378
 on sphere, 378–379
 two-layer equations, 379–380
 Shallow water-related models
 channels with irregular geometry, 377–378
 on sphere, 378–379
 two-layer shallow water equations, 379–380
 Shallow water system
 C-property, 134
 functional viscosity matrix methods, 145
 hydrostatic reconstruction technique, 160–161
 Roe methods, 157
 stationary solution, 160
 Shockwave, 61
 Short-time wave propagation
 FD-HMM, 567–569
 FE-HMM, 565–566
 macroscopic computational domain, 564
 periodic oscillatory tensors, 564
 Simple Riemann solvers (SRS), 142–144
 well-balanced property for, 154–156
 Simplified inverse Lax–Wendroff procedure, 30–31
 Simplified limiter, 90–91
 Singularmapping technique, 493
 SLH, 475
 Slope limiter procedure, 368, 370–371
 Smoothed particle hydrodynamics (SPH) method, 473
 Smoothing procedure, 284
 Smoothness indicators, 32–33
 Source terms, 67, 468
 Spalart–Allmaras model, 308, 341
 Spalart–Allmaras one-equation turbulence model, 342
 Spatial scales, 466
 Spectral element methods, 394, 564
 Sputtering yield, 483
 SSBJ design, Dassault Aviation, 288, 288f
 SSP high-order time discretizations, 91–92
 Staggered schemes, 493–494
 Standard explicit numerical method, 122
 Static geometry
 numerical boundary conditions for, 28–38
 two-dimensional Euler equations in, 34–38
 Stationary solutions, 133
 high-order well-balanced schemes, 161–162
 perturbation of, 135, 136–137f
 point values, 151
 well-balanced methods for, 160–161
 Steady-state solution techniques, 12–13
 Stiff collision, 108
 Stiff relaxation, 113–114, 119
 Stochastic AP schemes, 122–123
 Stochastic collocation methods, 511
 standard collocation method, 520–521
 stochastic finite volume method, 521–524
 Stochastic finite volume method, 521–524
 Stochastic Galerkin (sG) methods, 511, 517–518
 finite-dimensional noise assumption, 515
 generalized polynomial chaos (gPC), 516
 Strain-rate tensor, 73
 Strang splitting, 109–110
 Strong stability preserving (SSP) Runge–Kutta method, 91–94
 Superfluous dissipation, 59–60, 60f

Supersonic flow, business jet, 269–270
 Surface integrals, 394
 Surface mesh generation, 266–267, 267f
 Surface metric, 277–278
 Surface remeshing, 277–278
 Swanson–Tukel–Rossow implementation, 327

T

Temporal scales, 466–467
 Thermodynamic pressure, 206–207
 Time advancing scheme, 221
 Time-dependent coefficients, 495
 Time-marching methods
 acceleration methods, 318–323
 alternative approach, 304–305
 implicit schemes, 312–318
 model problem, 308–309
 multigrid methods, 323–327
 multistage schemes, 309–312
 RANS equations, 327–332
 Total variation bounded (TVB) limiter, 83, 368
 Total variation diminishing (TVD), 82
 Traffic flow, 482–483
 Transonic flow, Onera-M6 wing, 286–288, 287f
 Trapezium rule, 68
 Travelling wave, 167
 Two-dimensional Euler equations, in static
 geometry, 34–38
 fifth-order boundary treatment, 35
 ghost point, 38
 ILW procedure, 36–37
 Jacobian matrix, 35
 local coordinate system, 34, 34f
 no-penetration boundary condition, 37
 primitive variables, 37
 Two-layer shallow water equations, 379–380

U

Uncertainty quantification (UQ), 501
 challenges, 511
 compressible Euler Equations, 529–530
 entropy measure-valued solution, 536–537
 Lituya Bay mega-tsunami, 533–534, 534f
 Monte Carlo method
 error and complexity analysis, 525–526
 probability space, 524
 Multilevel Monte Carlo (MLMC) method
 efficient implementation, 528
 error and complexity analysis, 527–528
 MLMC-FVM algorithm, 526–527
 parallelization, 528

numerical methods, 509–510
 random entropy solutions, 514–515
 random fields
 concrete representations, 512–514
 covariance function, 513
 E-valued random variable, 511
 Karhunen–Loeve expansion, 513–514
 random Kelvin–Helmholtz problem,
 534–536, 535f
 sG methods, 517–518
 finite-dimensional noise assumption, 515
 generalized polynomial chaos (gPC), 516
 statistical solutions, 537–538
 stochastic collocation methods
 standard collocation method, 520–521
 stochastic finite volume method, 521–524
 uncertain Orszag–Tang Vortex, 530–533,
 531–532f
 water flooding, 510
 Uniform propagation
 L^2_k Integrability, 430
 lower bounds, 429
 numerical moments bounds, 429
 Semidiscrete H_k Sobolev Regularity, 430
 stability condition, 428–429
 Unit mesh, 270, 280
 Unsplit methods, 60–61
 Unsteady problems, 72–73
 Unstructured mesh
 adaptation
 boundary layers metric, 278–280
 error estimates, 274–276
 geometric estimate, for surfaces, 277–278
 interpolation error, 276–277
 metric tensors in, 271–272
 robustness and performance, 272–274
 anisotropic meshes
 insertion and collapse, 280–282
 unsteady simulations, 282–283
 generation
 surface mesh, 266–267
 volume mesh, 267–269
 numerical illustrations
 adaptive loop, 284–285
 blast prediction, 294–296
 boundary layer shock interaction,
 290–294
 direct sonic boom simulation, 288–290
 double mach reflection, 294–296
 transonic flow around M6 wing, 286–288
 wing-body configuration, 285
 Upstream mobility flux scheme, 496–498,
 496–497f

Upwind schemes

- multiple low MACH number
 - acoustic–incompressible interactions, 208–213
 - acoustic limit, 207–208
 - diagnosis, 215–217
 - finite volume schemes, 213–214
 - incompressible limit, 205–207
 - remedies, 217–221
- numerical illustrations
 - cell-centred triangular meshes, 224–226
 - quadrangular Cartesian grids, 221–222
 - vertex-centred triangular meshes, 223–224, 226–227
- relative errors, 134, 134*t*

V

- Vanishing diffusion limit, 167
- Variable local time stepping, 318–319
- Vascular stent net problem, 439–443
- Vector processing, 473
- Vector-valued functions, 392
- Velocity difference, 61
- Velocity divergence, 76
- Velocity waves, 208
- Vertex-based formula, 65–66
- Vertex-centred triangular meshes, 223–224, 226–227, 227*f*
- Vertex fluxes, 65
- Viscosity-free methods, 168
- Viscous profile, 167
- Vlasov equation, 179
- Vlasov–Maxwell solvers, 387
- Vlasov–Poisson–Fokker–Planck system, 119–120
- Volume mesh generation, 267–269
- Vorticity, business jet, 269–270

W

Wave equation

- in heterogeneous media without scale separation
 - engineering applications, 549–550
 - flux-transfer transformations, 555–558
- G-convergence and perturbation arguments, 561–563
- global fine scale computations, 553
- global fine scale quadrature rules, 553

- harmonic coordinate transformations, 551–553
- LOD, 558–561
- MsFEM, limited global information, 553–555
- operator upscaling, 551
- semidiscrete approximation, 550
- in heterogeneous media with scale separation
 - long-time wave propagation, 569–574
 - short-time wave propagation, 564–569
- homogenization of, 549

Wave models, 73

Wave propagation

- in heterogeneous media, 484
- in outer solar atmosphere, 485
- two-dimensional solar upper atmosphere, 499, 500*f*

Weak formulation, 393

Weak solution, 487

Weighted average flux (WAF) method, 148

Weighted essentially nonoscillatory (WENO) method, 25–26, 32–33, 83, 86, 365–367

Weighted particle method, 179, 195–196

Well-balanced HLL scheme, 155–156

Well-balanced methods, 151–166

- functional viscosity matrix methods, 157–158

generalized hydrostatic reconstruction (GHR), 158–160

high-order, 161–166

HLL scheme, 155–156

for moving water, 372–374

Roe method, 156–157

for SRSs, 154–156

for still water, 369–371

subset of stationary solutions, 160–161

Well-posedness theory, 501

Wing-body configuration, 285, 286*f*

WOLF, 284

Y

Yee scheme, 389

Yin–Yang mesh, 379

Z

Zero relaxation limit, 105–106

This page intentionally left blank

HANDBOOK OF NUMERICAL ANALYSIS

VOLUME

Series Editors

Q. Du, R. Glowinski, M. Hintermüller, E. Süli

XVIII

Handbook of Numerical Methods for Hyperbolic Problems

Applied and Modern Issues

Volume Editors Rémi Abgrall and Chi-Wang Shu

Hyperbolic partial differential equations arise in numerous applications, the most important of these being fluid dynamics, including specific flows, such as multiphase flows, magneto-hydrodynamics and water waves. Other application areas include electromagnetism, kinetic theory, astrophysics, and traffic flow models and networks.

Solutions to hyperbolic partial differential equations often exhibit discontinuities, which makes their mathematical analysis and numerical approximation difficult. Over the last few decades, a large body of literature has emerged on the design, analysis and application of various numerical algorithms for the approximate solution of hyperbolic equations. This is the second of two volumes in which experts in different types of algorithms provide concise summaries in order to acquaint the reader with a range of numerical techniques, in a variety of different situations, and survey their relative advantages and limitations. While the first volume addresses basic and fundamental questions concerning numerical methods for hyperbolic problems, this second volume focuses on more applied topics.



North-Holland

An imprint of Elsevier
elsevier.com

ISBN 978-0-444-63910-3



9 780444 639103