

Dissertation
submitted to the
Combined Faculties for the Natural Sciences and for
Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

put forward by
Dipl.-Phys. Kevin Schaal
born in Tübingen

Date of oral examination: July 8th, 2016

Shocks in the Illustris Universe and Discontinuous Galerkin Hydrodynamics

Referees:

Prof. Dr. Volker Springel

Prof. Dr. Ralf Klessen

Zusammenfassung

Die Erforschung hochgradig nichtlinearer astrophysikalischer Systeme und Prozesse hängt in zunehmender Weise von numerischen Simulationen ab. Die Ziele dieser Arbeit liegen in der Entwicklung neuartiger Analysemethoden für kosmologische Simulationen und der Einführung neuer numerischer Methoden zur Verbesserung ihrer Genauigkeit.

Wir stellen die Implementierung eines Algorithmus zur Stoßwellendetektierung für den AREPO-Code mit bewegtem Gitter vor. Damit analysieren wir Stoßwellen in Illustris, einer hochmodernen kosmologischen Galaxienentstehungssimulation. Wir identifizieren Stoßwellen verschiedenster Art, zum Beispiel aufgrund von Akkretionsprozessen, Strukturkollisionen, schwarzen Löchern, und galaktischen Winden. Die stärksten Stoßwellen werden hierbei durch schwarze Löcher hervorgerufen. Zu späten Zeiten messen wir eine spezifische Stoßwellendissipationsrate, die relativ umgebungsunabhängig ist, von $10^{-1} \text{ erg g}^{-1} \text{ s}^{-1}$.

In einem weiteren Projekt beschreiben und implementieren wir eine diskontinuierliche Galerkin (DG) Methode höherer Ordnung zur Lösung der idealen Gasgleichungen auf einem strukturierten Gitter, welches lokal verfeinert werden kann. Durch Vergleich mit einem traditionellen Finiten-Volumen-Verfahren finden wir, dass DG geringere Diffusions- und Advektionsfehler aufweist, sowie für rotierende Systeme zu bevorzugen ist. Unsere Ergebnisse belegen ein großes Potenzial dieser Methode für astrophysikalische Anwendungen.

Abstract

The study of highly-nonlinear astrophysical systems and processes increasingly relies on simulation techniques. This thesis aims to develop new analysis methods for cosmological simulations and advance their accuracy by introducing novel numerical methods.

We present the implementation of a shock finding algorithm for the moving-mesh code AREPO, and use it to analyse shocks in Illustris, a cosmological state-of-the-art simulation of galaxy formation. We identify a large diversity of shock morphologies, including accretion shocks, merger shocks, as well as feedback shocks due to galactic winds and black holes; the latter produce the strongest shocks in the simulation. At late times, we measure a specific energy dissipation rate at shocks of $10^{-1} \text{ erg g}^{-1} \text{ s}^{-1}$, fairly independent of the cosmological environment.

In another project we formulate and implement a higher-order discontinuous Galerkin (DG) scheme on a structured mesh, which can be adaptively refined, for solving the equations of ideal hydrodynamics. We find that DG produces significantly less diffusion and smaller advection errors compared to a traditional finite volume scheme, and is superior in evolving rotating objects. These results demonstrate that DG offers a high potential for astrophysical applications.

List of publications

- 1.* **Schaal, K.**, Springel, V., Pakmor, R., Pfrommer, C., Nelson, D., Vogelsberger, M., Genel, S., Pillepich, A., Sijacki, D., & Hernquist, L.
Shock finding on a moving-mesh: II. Hydrodynamic shocks in the Illustris universe
submitted to MNRAS (arXiv:1604.07401)
2. Pfrommer, C., Pakmor, R., **Schaal, K.**, Simpson, C., & Springel, V.
Simulating cosmic ray physics on a moving mesh
submitted to MNRAS (arXiv:1604.07399)
3. Bauer, A., **Schaal, K.**, Chandrashekar, P., Pakmor, R., Klingenberg, C., & Springel, V.
Simulating Turbulence Using the Astrophysical Discontinuous Galerkin Code TENET
To appear in Proceedings of the SPPEXA symposium, Lecture Notes in Computational Science and Engineering (LNCSE), Springer (arXiv:1602.09079)
4. Pakmor, R., Springel, V., Bauer, A., Mocz, P., Munoz, D. J., Ohlmann, S. T., **Schaal, K.**, & Zhu, C.
Improving the convergence properties of the moving-mesh code AREPO
2015, MNRAS, 455, 1134
- 5.* **Schaal, K.**, Bauer, A., Chandrashekar, P., Pakmor, R., Klingenberg, C., & Springel, V.
Astrophysical hydrodynamics with a high-order discontinuous Galerkin scheme and adaptive mesh refinement
2015, MNRAS, 453, 4278
- 6.* **Schaal, K.**, & Springel, V.
Shock finding on a moving mesh - I. Shock statistics in non-radiative cosmological simulations
2015, MNRAS, 446, 3992
7. Geretshauer, R. J., Meru, F., **Schaal, K.**, Speith, R., & Kley, W.
Simulation of Pre-Planetesimal Collisions with Smoothed Particle Hydrodynamics II
2013, Springer, High Performance Computing in Science and Engineering 2012, p. 51-67

* part of this thesis.

Contents

I	Introduction	1
1	Introduction	3
1.1	Motivation	3
1.2	Cosmology	5
1.3	Hydrodynamic shocks in cosmological simulations	7
1.4	Discontinuous Galerkin hydrodynamics	9
1.5	Structure of the thesis	10
2	Hydrodynamics	13
2.1	Governing equations	13
2.1.1	Navier-Stokes equations	13
2.1.2	Euler equations	16
2.1.3	Ideal gas	19
2.1.4	Sound waves	23
2.2	Hydrodynamic shocks	25
2.2.1	Hydrodynamic discontinuities	26
2.2.2	The Rankine–Hugoniot jump conditions	27
2.2.3	Energy dissipation at shocks	30
2.2.4	Diffusive shock acceleration	32
2.3	Analytic test problems	34
2.3.1	Shock tube solution	34
2.3.2	Sedov-Taylor blast wave solution	40
3	Computational hydrodynamics	45
3.1	Finite volume method	45
3.1.1	The finite volume scheme	45
3.1.2	Limiting	47
3.1.3	Finite volume method on a moving Voronoi mesh	50
3.2	Discontinuous Galerkin method	54
3.2.1	The discontinuous Galerkin scheme	54
3.2.2	Legendre polynomials	58
3.2.3	Gaussian quadrature	59
3.2.4	Lobatto quadrature	60

3.3	Runge-Kutta time integration	61
3.4	The Courant-Friedrichs-Lewy condition	63
II	Hydrodynamic Shocks in non-radiative Simulations	67
4	Introduction	71
5	Methodology	75
5.1	The moving-mesh code AREPO	75
5.2	The Rankine–Hugoniot jump conditions	75
5.3	Shock-finding method for AREPO	76
5.3.1	Shock direction	76
5.3.2	Shock zone	78
5.3.3	Shock surface	78
5.3.4	Mach number calculation	79
5.4	Energy dissipation	79
6	Validation	81
6.1	Sedov–Taylor blast wave	81
6.2	Shock tubes	82
7	Shocks in non-radiative simulations	85
7.1	Simulation set-up	85
7.2	Shock finder assessment	85
7.3	Reionization modelling	87
7.4	General properties	89
7.5	Shock statistics	89
7.6	Galaxy cluster shocks	93
7.7	Cosmic ray acceleration	94
8	Methodology variations	97
9	Summary	103
III	Hydrodynamic Shocks in the Illustris Universe	107
10	Introduction	111
11	Analysing shocks in the Illustris universe	115
11.1	The Illustris project	115
11.2	Moving-mesh hydrodynamics	117
11.3	Shock finding	117

12 Shock statistics	121
12.1 Global shock statistics	121
12.2 Environmental dependence of the shock statistics	129
13 Shock morphologies across cosmic time	135
13.1 High-redshift accretion shocks	135
13.2 Galaxy and galaxy cluster shocks	142
14 Methodology variations	147
15 Resolution study	151
16 Discussion and summary	155
IV Discontinuous Galerkin Hydrodynamics	159
17 Introduction	163
18 Discontinuous Galerkin hydrodynamics	167
18.1 Euler equations	167
18.2 Solution representation	168
18.3 Initial conditions	169
18.4 Evolution equation for the weights	170
18.5 Timestep calculation	173
18.6 Angular momentum conservation	174
19 Slope limiting	177
19.1 Component-wise limiter	177
19.2 Characteristic limiter	178
19.3 Total variation bounded limiting	179
19.4 Positivity limiting	180
20 DG with AMR	183
20.1 Refinement criterion	184
20.2 Mesh refinement	185
20.3 Mesh derefinement	186
20.4 Limiting with AMR	187
20.5 Main simulation loop	188
21 Validation	191
21.1 Isentropic vortex	192
21.2 Shock tube	195
21.3 Sedov-Taylor blast wave	197
21.4 Square advection	199

Contents

21.5 Keplerian disc	202
21.6 KH instability	205
22 Summary	209
V Conclusions and Outlook	211
23 Discontinuous Galerkin hydrodynamics	213
24 Hydrodynamic shocks in cosmological simulations	217
Appendix	223
Supplementary material	223
List of Abbreviations	227
Table of Constants	229
Table of Symbols	231
Bibliography	250
Acknowledgments	251

Part I

Introduction

1

Introduction

1.1 Motivation

Hydrodynamics and hydrodynamical processes play an important role in many astrophysical systems. By far the most abundant atoms in the Universe are hydrogen and helium, gases that can be found in large amounts across all spatial scales. In order to obtain a deep understanding of the Universe, it is essential to take the physics of fluids into account.

In current research in the field of astrophysics, hydrodynamics plays an ever more important role. The basic formation mechanisms and processes of many objects, such as galaxies, black holes, stars, and planets are qualitatively understood at the level of coarse theoretical ideas. The aim of research within several branches has therefore shifted towards a more detailed modelling of the formation and evolution of astrophysical objects. This includes as a centerpiece the effects and dynamics of fluids, in particular in the non-linear regime, where fluid instabilities and hydrodynamic shocks occur.

Describing a system in the fluid limit is valid if the mean free path between particle collisions is much smaller than the characteristic size of the system. If this is not the case, kinetic gas theory has to be adopted, in which the solution is obtained by means of distribution functions. Fortunately, the fluid limit can be used for many astrophysical systems, and their hydrodynamic state at any location is given by macroscopic fluid quantities, such as density, pressure, and velocity. Although this is the case in many situations, there are technically no particle collisions, and particle interactions are mediated by magnetic fields, which are omnipresent in the Universe.

Astrophysical hydrodynamics is relevant for systems across all spatial scales. Galaxy clusters, which can contain thousands of galaxies, consist mainly of dark matter (around 85%), and hot gas ($\gtrsim 10\%$). Clusters grow in size by gas accretion and

in mergers with smaller objects. These processes can create hydrodynamic shocks and turbulence. Shocks are fluid discontinuities that separate two hydrodynamic states, e.g. the pristine gas outside of a cluster and the intracluster medium (ICM). They originate from supersonic collisions of fluid streams, and irreversibly transform kinetic energy into thermal energy. Perturbations in the cluster atmosphere can create vorticity in the gas and induce turbulence, which is a chaotic gas motion transporting energy from large scales to smaller scales, where it is dissipated by viscous processes. Both phenomena, shock waves and turbulence, heat the ICM and change irreversibly the thermal state of a cluster [see for example Vazza et al., 2010].

Another example where non-linear fluid phenomena are important is the stripping of gas in a galaxy-cluster interaction due to ram pressure. The relative movement of the galaxy within the halo atmosphere can result in a pushing and dragging on the gas of the galaxy [Gunn and Gott, 1972; Nulsen, 1982]. The pushing can result in a shock and trigger star formation in the interstellar medium (ISM). Moreover, a Kelvin-Helmholtz instability can occur at the galaxy outskirts and excite non-linear waves. As a consequence, gas billows are stripped off the infalling galaxy, depleting it of its gas. The detailed effect on the galaxy and its star formation history depend strongly on the infall parameters and galaxy properties [Steinhauser et al., 2016].

There are many other examples involving complex hydrodynamical processes, such as black hole outflows in the form of relativistic jets, star formation in turbulent molecular clouds, or the formation of planets in a protoplanetary disc. Their common feature is that the underlying governing equations are highly non-linear, and can not be solved straight-forwardly for complex systems. Hence, for studying these problems in detail, one has to adopt numerical simulations.

In a numerical hydrodynamic simulation the underlying equations are discretized, and the system is evolved step by step. This is commonly achieved by discretizing time and space, or alternatively, time and mass. Moreover, such simulations are ideally run on massively parallel systems, leveraging the power of thousands of computing cores simultaneously. Within the last two decades, simulation codes have greatly improved in both, performance and complexity, and their evolution will progress continuously in the future. In order to ensure constant advancement, new algorithms and methods have to be developed with the aim of improving accuracy, and obtaining a more complete physical model. Furthermore, with the ever increasing availability of parallel computing resources, the development of scalable algorithms has come into focus.

These aspects represent the motivation for this thesis, which consists of two projects. In the first project, we develop and implement an algorithm for accurately detecting and characterizing hydrodynamic shocks in cosmological simulations. Shocks are of particular interest in astrophysical contexts, since they can accelerate particles via the diffusive shock acceleration mechanism [DSA; Axford et al., 1977; Bell, 1978a,b; Blandford and Ostriker, 1978; Krymskii, 1977; Malkov and O’C Drury, 2001], giving rise to cosmic rays. These relativistic particles can provide a non-thermal pressure component, which is of great importance for galaxy formation [Girichidis et al., 2016; Pfrommer et al., 2016]. Moreover, studying shocks provides insights into the

dynamics and thermal states of astrophysical systems, representing a unique view on hydrodynamic simulations.

In the second project, we develop and implement a discontinuous Galerkin (DG) method for solving the hydrodynamic equations on a structured mesh which can be adaptively refined (adaptive mesh refinement, AMR). The DG method offers significant advantages compared to commonly used finite volume (FV) methods, and we investigate its potential for astrophysical applications.

1.2 Cosmology

A breakthrough in cosmology has been achieved in 1936, when Edwin Hubble published his famous velocity-distance relation of nearby galaxies. He found that the radial velocity of galaxies as seen from Earth is related to their distance by the simple law

$$v = H_0 d, \tag{1.1}$$

where the constant H_0 is known today as the Hubble constant and has a value of around $H_0 \approx 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$. Hence, the further away a galaxy, the faster it is moving away, and if the position of Earth in the Universe is not preferred relative to any other position, this can only be explained by an expansion of the Universe. Moreover, by extrapolating back in time, one can draw the conclusion that the Universe originated from an extremely condensed and hot state, a theory also known as the Big Bang. If we assume that the recession velocity of a galaxy is constant with time, a rough estimate for the age of the Universe is given by the inverse of the Hubble constant, which is around 13.8 billion years and known as the Hubble time.

Another milestone in the last century has been the discovery of the cosmic microwave background (CMB) in 1964. The CMB is the afterglow of the early Universe and its existence strongly favours the Big Bang model. Around 380.000 years after the creation of the Universe, its temperature had decreased to a point where protons and electrons of the primordial plasma could combine for the first time, a process known as recombination. The transformation of a charged plasma into neutral atoms results in the decoupling of photons, which are then able to travel freely. Almost 14 billion years later these photons can be detected by radio telescopes on Earth, although with a significantly decreased energy due to the expansion of space-time within this period of time. Today, the CMB is measured as a very homogenous and isotropic black body radiation, with a temperature of 2.73 K. However, two poles with slightly different temperatures can be observed, an effect known as the CMB dipole anisotropy. It originates from the movement of the Solar system relative to the observable Universe. Moreover, after subtracting the dipole contribution, high precision measurements of the CMB reveal tiny temperature fluctuations, at the level of one part per million, which encode primordial density fluctuations of the Universe. It is believed that these fluctuations have their origin in quantum

fluctuations, stretched to macroscopic scales by cosmic inflation. The latter refers to the exponential expansion within a tiny fraction of a second right after the Big Bang.

Pioneering measurements of the CMB have been taken by the space telescopes COBE, WMAP, and PLANCK. The resolution of the latter two is high enough in order to obtain a power spectrum of the fluctuations, from which cosmological parameters can be inferred. Those include the Hubble constant and the energy densities of the different components of the Λ CDM model, which is the very successful and widely favoured standard model of cosmology. According to CMB measurements, the Λ CDM Universe consists of around 75% dark energy, represented by the cosmological constant Λ , about 25% cold dark matter, which mostly interacts gravitationally with other particles, and around 5% ordinary baryonic matter. At the present epoch, the contribution of radiation to the total energy density is negligible.

As we have seen above, the relative expansion rate of today's Universe is given by the Hubble constant H_0 . However, this rate is time dependent, which can be expressed by writing

$$H(t) = \frac{\dot{a}}{a}, \quad (1.2)$$

where a is the scale factor of the Universe. This quantity is related to the redshift z measured in the spectra of a receding galaxy by $a = \frac{1}{1+z}$. Both quantities are monotonic functions of time, and they are commonly used in cosmological contexts for measuring the latter. At the present time their values are $z = 0$ and $a = 1$.

The time evolution of the Hubble parameter is given by Friedmann's equation, which reads

$$H^2(a) = H_0^2 [\Omega_{r0} a^{-4} + \Omega_{m0} a^{-3} + \Omega_{\lambda0} + \Omega_K a^{-2}]. \quad (1.3)$$

Here $\Omega_{\lambda0}$ is the dark energy density parameter, Ω_{r0} the radiation density parameter, and Ω_{m0} the total matter density parameter at the present epoch. The latter two are physical densities normalized by the critical density of the Universe,

$$\rho_{\text{cr}}(t) = \frac{3H(t)^2}{8\pi G}. \quad (1.4)$$

This density corresponds to the mean density in a flat, matter-dominated Universe. Flatness is, at least approximately, realized, since the curvature parameter Ω_K is measured to be close to zero.

As indicated by the CMB, the Universe is isotropic and homogeneous on large scales, properties which are known as the cosmological principle. However, the tiny primordial fluctuations grow with cosmic time due to the force of gravity. This gives rise to the formation of the cosmic 'sponge', and later on the cosmic web, which is the dark matter skeleton of the Universe. At the nodes of the cosmic web, dark matter haloes assemble mass by accretion processes and mergers with smaller haloes. During this process of hierarchical structure formation, the gas in the Universe follows the dark matter distribution on large scales. However, it can cool radiatively in the

densest regions and collapse. In this way, stars come into existence, and subsequently galaxies, and galaxy clusters.

1.3 Hydrodynamic shocks in cosmological simulations

Modern cosmological simulations evolve dark matter and baryonic components for a physical time of around 13.6 billion years, until the present epoch. Their initial conditions are created with the cosmological parameters inferred from the CMB. At the end of such a simulation, the simulated universe¹ can be compared with observations of the real Universe, which provides an important validation for the adopted physical models.

In this thesis, we design and implement an algorithm for detecting and characterizing hydrodynamic shocks in cosmological simulations. For this purpose, we use the state-of-the-art code AREPO [Springel, 2010], which evolves the gas component on an unstructured, moving Voronoi mesh. This approach offers many advantages compared to traditional methods, as we will outline in Section 3.1.3.

With the shock finder we analyse a non-radiative simulation, as well as one with a more complete physical model ('full physics simulation'). The former evolves the dark matter and gas in the universe by solving the laws of gravity and ideal hydrodynamics, without any other source terms. On the other hand, the analysed full physics simulation, which is the Illustris simulation [Vogelsberger et al., 2014c], also incorporates a reionizing background, cooling physics, stars, and black holes.

In both simulations, the hierarchical structure growth gives rise to a variety of hydrodynamic shocks. Additional shocks are present in the full physics run, reflecting the adopted physical models. It is instructive to analyse each of these simulations, and compare the results. This project aims at answering the following questions.

- **How can we construct a suitable algorithm for finding and characterizing hydrodynamic shocks in simulations on an unstructured-mesh?**

Shocks in cosmological simulations have been explored previously, as well as their detection with suitable algorithms. However, this has been mostly done for rather simple non-radiative simulations, as well as on structured meshes. How can a shock finding algorithm for an unstructured mesh be designed? Which ideas of established methods can be generalized and adopted for our purpose? Can the existing methodology for structured meshes be improved? In simulations, a hydrodynamic shock is smoothed, usually over a couple of cells. How can it be reconstructed to a surface? Which hydrodynamic quantity is suited for identifying the shock direction and calculating its Mach number? For simulations with a more complete physical model new challenges arise in the form of cooling physics and subgrid models, which intervene with the evolution of the ideal gas. How can we deal with this difficulty?

¹In this thesis we follow the common convention and use upper case if we refer to the real Universe, and lower case in the context of a simulated universe.

- **How well does the novel shock finding algorithm perform?**

Our shock finding algorithm is validated and confronted with various tests. How well are shocks with different geometries and morphologies detected by the algorithm? Can analytic calculations for simple test problems be reproduced? Which fraction of the kinetic energy that is irreversibly converted to thermal energy in a cosmological simulation is registered by our shock finder?

- **What can we learn by detecting and characterizing shocks in non-radiative simulations?**

Non-radiative simulations evolve dark matter and gas by solving the laws of gravity and ideal hydrodynamics. In these types of simulations, the temperature evolution of the gas is almost exclusively determined by the adiabatic expansion of the universe on the one hand, and by heating due to accretion shocks and merger shocks on the other hand. Hence, it is very instructive and interesting to study the arising shocks and their properties. Where can most of the shocks be found? Which kinds of shocks do occur, and what are their strengths? Which cosmological environments experience most of the shock heating? Is it possible to see the signature of accretion shocks around haloes in radial shock profiles? If this is the case, what is the typical distance of an accretion shock from the halo centre? Radio relics can be observed in several galaxy clusters in our Universe, and their occurrence is related to cluster mergers; can we discover the associated shock signatures within haloes in the simulation? The accelerated expansion of the Universe at late times strongly dilutes gas inside voids; can this effect be seen in the shock statistics? Moreover, shocks can accelerate cosmic ray particles by the mechanism of diffusive shock acceleration. Which energy fraction can potentially be channeled into the acceleration of cosmic rays? Is the resulting cosmic ray population expected to be energetically and hence dynamically relevant? In our thesis we introduce a novel shock finding algorithm. Do our results agree with previous studies? Are the methodology improvements we adopt significant? And finally, how sensitive are our results to implementation details? Do we obtain robust results with respect to the resolution of the analysed simulations?

- **What can we learn by detecting and characterizing shocks in state-of-the-art cosmological simulations?**

In contrast to non-radiative simulations, current cosmological state-of-the-art simulations also incorporate a model for reionization, gas cooling, star- and black hole formation, and take the associated feedback processes into account, in the form of stellar winds and outflows of active galactic nuclei (AGN). Depending on the environment and properties of the feedback processes, the outflows can be supersonic and give rise to shock waves. It is instructive to analyse these shocks in detail. Are they significant for the overall shock statistics and energetics, or do they play only a minor role? If the shock statistics differ significantly from those of non-radiative runs, what is the main origin of the difference?

Moreover, state-of-the-art simulations aim at producing realistic populations of stars and black holes. How does the amount of feedback energy compare to the energy which is processed by shocks waves? If we investigate the properties of feedback shocks, can we draw conclusions concerning their observability in the Universe? What are the physical values for specific dissipation rates of shocks in different environments? The detailed process of gas accretion is an active topic of current research. For massive haloes, theoretical models favour a hot accretion mode; in this case gas gets shock-heated to the virial temperature before it is accreted by a galaxy. However, there exists also a cold mode, which is predominantly present for low mass systems, in which cold gas can accrete without being shocked. Is it possible to observe the different accretion modes in the simulation by analysing the shock morphology around haloes? How does the gas velocity field around an accreting halo look in detail?

- **How much information about the dynamics of the gas is contained in hydrodynamic shocks?**

From a theoretical point of view, hydrodynamic shocks are a manifestation of the non-linearity of the fluid equations, and we hence expect that the most important flow features of a structure or system are visible in its shock signature. By revealing large-scale accretion shocks on to non-linear structures, can we reveal the whole cosmic web? Is it possible to infer the state of a halo by investigating the shock environment? E.g., can we infer from the Mach number field only, whether a halo is in a merging, post-merging, or relaxed state? The Mach number field in a simulation is compared to e.g. the density field a significantly reduced dataset, and can be processed and investigated much faster for massive cosmological simulations. Does this dataset have the potential to characterize the dynamical state of cosmological structures and systems? Moreover, is it possible to discover features in the gas dynamics which can be barely seen in the density and temperature field?

1.4 Discontinuous Galerkin hydrodynamics

In this second main project of the thesis, we implement a DG method for solving the equations of ideal hydrodynamics on an AMR grid, which is a structured mesh that can be adaptively refined. In the DG method, the solution is represented within every cell by a higher-order polynomial, which is given as a linear combination of basis functions. The global solution is generally discontinuous across cell interfaces, and evolved by solving equations for the weights of the basis functions. These properties result in a very efficient scheme, in particular for computations on massively parallel systems.

DG can be generalized straight-forwardly to arbitrary order², and has the potential to solve the equations of hydrodynamics faster, and to a higher precision than an FV scheme. Hence, it is very interesting to investigate this method in astrophysical contexts, and compare its performance to a traditional FV method in hydrodynamic test problems. This work tries to answer the following questions.

- **How can a Discontinuous Galerkin method for solving the Euler equations on a structured adaptive mesh be realized?**

DG is subject to current research in applied mathematics and different fields involved in high-performance computing. Nevertheless, for hydrodynamical applications on an adaptive mesh, DG is largely unexplored so far. What is a suitable mathematical formulation for DG with AMR? How can this formulation be realized in a numerical implementation? Specifically, how can we obtain a scheme which is accurate as well as robust, especially at the boundaries of cells with different sizes? How can the refinement and derefinement of cells be handled?

- **What potential does DG offer for astrophysical applications?**

As opposed to finite volume methods, which have a long tradition in computational astrophysics, DG methods are relatively unexplored in the field. What are the advantages and disadvantages of DG compared to FV in astrophysically relevant test problems? In particular, how well can the scheme handle rotating objects, and objects which are advected supersonically? How well does the scheme capture hydrodynamic shocks?

- **Which order of accuracy in space and time is most promising for DG hydrodynamics?**

Most FV codes in astrophysics have second order accuracy, meaning that the error decreases quadratically with increasing resolution. A second order FV method can be implemented conveniently, however, for most scenarios we suspect that second order is not the optimal order. DG can straight-forwardly be generalized to arbitrary order. Which order gives the best results for different test problems? Most interestingly, at which order is the computational cost minimal for achieving a given accuracy in smooth flows?

- **How can we achieve an accurate solution representation at discontinuities?**

While higher-order methods reveal their full potential in smooth flows, hydrodynamic discontinuities such as shocks can be represented more accurately with low order. At these locations it is therefore crucial to limit the solution and hence reduce the order of the DG scheme. How can the DG solution polynomial

²The order of a numerical scheme indicates the rate with which the error of the numerical solution decreases, when the resolution is increased. E.g., in a first order code, the error decreases linearly with increasing resolution.

be effectively limited? How can this be achieved without losing accuracy in smooth flows?

1.5 Structure of the thesis

This thesis is structured as follows.

Section 2 presents an introduction to hydrodynamics, which starts with the derivation of the Navier-Stokes equation in 2.1.1. The Euler equations of ideal hydrodynamics and their properties as hyperbolic conservation laws are discussed in 2.1.2, and the ideal gas law is derived in 2.1.3. Section 2.2 studies the Euler equations at hydrodynamic shocks. To this end, we derive the Rankine-Hugoniot jump conditions in 2.2.2, and calculate the dissipation of kinetic energy at shocks in 2.2.3. Last but not least, we present analytic solutions for the Riemann problem in 2.3.1, and for the problem of a strong hydrodynamic explosion in 2.3.2.

In Section 3 we give an introduction to computational hydrodynamics, and discuss the FV scheme in Section 3.1, as well as the DG scheme in Section 3.2. The limiting in an FV method is explained in 3.1.2, and the generalization to a moving mesh in 3.1.3. The DG equations are derived in 3.2.1, and we introduce Legendre polynomials and Gaussian quadrature in 3.2.2 and 3.2.3, respectively. At the end, we give an outline of explicit Runge-Kutta time integrators in Section 3.3, and discuss the CFL timestep criterion in Section 3.4.

The parts II, III, and IV are the main parts of this thesis, and they correspond to the publications [Schaal and Springel, 2015], [Schaal et al., 2016], and [Schaal et al., 2015], respectively.

Part II deals with hydrodynamic shocks in non-radiative simulations. It starts in Section 4 with an overview of the observability of cosmological shocks, and of existing shock finder implementations. Section 5 presents the methodology of our shock finding algorithm, which is validated in Section 6. In Section 7 we describe the analysed simulation and present our results. Those include the mean distance of accretion shocks from the centre of the biggest haloes, and an estimate for the acceleration of cosmic rays. In Section 8 we test the performance of our algorithm compared to previously used shock finding methods, and our results are summarized in Section 9.

In Part III of the thesis, we analyse shocks in a state-of-the-art cosmological simulation, the Illustris simulation. First of all, we give an introduction to this topic in Section 10, and introduce the Illustris simulation in Section 11. This includes the description of the adopted physical models, an outline of the challenges for our shock finding algorithm in these kinds of simulations, and the presentation of an improved methodology. Section 12 presents analysis results, first in a global sense, and hereafter for different cosmological environments. The former includes a comparison of the energy rate processed by shocks with the rates released by different feedback channels. In Section 13 we illustrate and interpret various shock morphologies that can be found in the Illustris universe. These include shocks around disk galaxies, and a

zoomed in view on a double accretion shock pattern, a feature that can be found relatively frequently in haloes. In Section 14 we demonstrate the robustness of our improved shock finding methodology, and in Section 15 we study how much our results depend on the resolution of the simulation. An overview of our findings is given in Section 16.

Part IV of the thesis is dedicated to discontinuous Galerkin hydrodynamics. In Section 17 we describe the most popular methods in astrophysics for solving the fluid equations, and motivate the exploration of novel methods. Section 18 presents a detailed derivation of the DG scheme, and outlines a stable choice for the timestep. Moreover, we demonstrate explicitly that DG can conserve angular momentum. Section 19 deals with limiters for obtaining an accurate solution at hydrodynamic discontinuities, as well as for achieving an overall stability. In Section 20 we generalize the DG scheme to AMR grids, and present the main simulation loop of our implementation. Section 21 validates our method and investigates its performance in various test problems, many of which are of great relevance in astrophysical contexts, e.g. the evolution of a cold Keplerian disc. Moreover, for most of the tests, we compare the results to those obtained with a standard second-order finite volume method. Our findings are summarized in Section 22.

Finally, in the last part of this thesis (Part V), we present the conclusions of our work, and give an outlook on possible developments for the future. This is done separately for each of our projects, for discontinuous Galerkin hydrodynamics in Section 23, and for hydrodynamic shocks in cosmological simulations in Section 24.

There are many great text books that deal with hydrodynamics and computational hydrodynamics, and lectures covering these fields are offered in many universities. The books and lecture notes we have used for writing the introductory Sections 2 and 3 are Landau and Lifshitz [1966]; LeVeque [2002]; Nolting [2013]; Pfrommer [2016]; Reinhardt [2009]; Springel [2016]; Toro [2009].

2

Hydrodynamics

2.1 Governing equations

2.1.1 Navier-Stokes equations

The fluid description is macroscopic, and the associated fluid limit is valid if the mean free path between particle collisions is much shorter than the characteristic size of the system. In the following, we derive the fluid equations from the distribution function of the particles in phase space.

The phase space distribution function $f(\mathbf{x}, \mathbf{v}, t)$ indicates the probability that a particle is located within the volume $d^3x d^3v$ centred at (\mathbf{x}, \mathbf{v}) , and it is normalized such that

$$\int \int f(\mathbf{x}, \mathbf{v}, t) d^3x d^3v = N, \quad (2.1)$$

where N is the total number of particles. We assume that particles are neither created nor destroyed, however, we take particle collisions into account which can change the trajectory of particles through phase space discontinuously. The continuity equation in this case reads

$$\frac{d}{dt} f(\mathbf{x}, \mathbf{v}, t) = \left. \frac{\partial f}{\partial t} \right|_c, \quad (2.2)$$

where the right-hand side expresses the collision between particles.

If the time derivative of the distribution function is expanded, the collisional Boltzmann equation is obtained, which reads

$$\frac{\partial f}{\partial t} + \dot{\mathbf{x}} \cdot \nabla f + \dot{\mathbf{v}} \cdot \nabla_v f = \left. \frac{\partial f}{\partial t} \right|_c. \quad (2.3)$$

In the fluid limit the velocity distribution of f approaches a Maxwell-Boltzmann distribution, enabling the transition to a macroscopic system. In this case it is appropriate to decompose the velocity into a mean velocity $\mathbf{v}_m = \langle \mathbf{v} \rangle$, and a random velocity \mathbf{v}_r ,

$$\mathbf{v} = \mathbf{v}_m + \mathbf{v}_r. \quad (2.4)$$

The local fluid density can be calculated from the distribution function by

$$\rho(\mathbf{x}, t) = \int m f(\mathbf{x}, \mathbf{v}, t) d^3v, \quad (2.5)$$

where m denotes the mass of a single particle. Using this expression, the mass-weighted mean of any other quantity q is given by

$$\langle q \rangle = \frac{1}{\rho} \int q m f(\mathbf{x}, \mathbf{v}, t) d^3v. \quad (2.6)$$

Macroscopic evolution equations for the density, momentum density, and internal energy density of the fluid can be obtained by taking moments of the Boltzmann equation and integrating over velocity space. The zeroth moment is calculated by multiplying equation (2.3) with m and integrating over d^3v , which yields

$$\frac{\partial}{\partial t} \int f m d^3v + \int (\mathbf{v} \cdot \nabla f) m d^3v + \int \mathbf{g} \cdot (\nabla_v f) m d^3v = \int \left. \frac{\partial f}{\partial t} \right|_c m d^3v. \quad (2.7)$$

Here we have also replaced $\dot{\mathbf{x}}$ by \mathbf{v} , and $\dot{\mathbf{v}}$ by the acceleration \mathbf{g} .

The first term simply corresponds to the local density change,

$$\frac{\partial}{\partial t} \int f m d^3v = \frac{\partial \rho}{\partial t}. \quad (2.8)$$

The second term can be written as a sum, and we obtain

$$\int (\mathbf{v} \cdot \nabla f) m d^3v = \sum_{i=1}^3 \frac{\partial}{\partial x_i} \int v_i m f d^3v \stackrel{(2.6)}{=} \sum_{i=1}^3 \frac{\partial}{\partial x_i} \rho \langle v_i \rangle = \nabla \cdot (\rho \mathbf{v}_m). \quad (2.9)$$

In the last step we have used that $\langle \mathbf{v} \rangle = \langle \mathbf{v}_m + \mathbf{v}_r \rangle = \langle \mathbf{v}_m \rangle$.

Applying Gauss' divergence theorem to the third term yields

$$\int \mathbf{g} \cdot (\nabla_v f) m d^3v = \int_{\partial V_v} m f(\mathbf{g} \cdot \mathbf{n}) d^2A_v, \quad (2.10)$$

where d^2A_v is the differential surface element of velocity space and \mathbf{n} its normal vector. This term vanishes if we integrate over the surface of the whole velocity space, since $f \rightarrow 0$ for $|\mathbf{v} \rightarrow \infty|$. Moreover, the integral of the right-hand side of the Boltzmann equation (2.9) is also zero, since collisions do not change the local density of the particles.

Therefore, the zeroth order of the Boltzmann equation reads

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}_m) = 0. \quad (2.11)$$

This equation is simply the mass continuity equation and indicates that mass changes locally only due to a mass flux. Moreover, the total mass of the system is conserved.

In a similar way, the first moment can be obtained by multiplying the Boltzmann equation with $m\mathbf{v}$ and integrating over d^3v , which results in

$$\frac{\partial \rho \mathbf{v}_m}{\partial t} + \nabla \cdot (\rho \mathbf{v}_m \mathbf{v}_m^T) = -\nabla p + \nabla \Pi + \rho \mathbf{g}. \quad (2.12)$$

Here we have introduced the pressure

$$p = \frac{1}{3} \rho \langle |\mathbf{v}_r|^2 \rangle, \quad (2.13)$$

and the viscous stress tensor

$$\Pi_{ij} = p \delta_{ij} - \rho \langle v_{r,i} v_{r,j} \rangle. \quad (2.14)$$

Equation (2.12) is known as *the* Navier-Stokes equation, and expresses that linear momentum changes locally due to a momentum flux, a pressure gradient, the presence of viscosity, and an external acceleration \mathbf{g} .

A connection of the viscous stress tensor to the mean velocities can only be derived heuristically. If one assumes that the tensor expressed in these velocities is symmetric, depends linearly on the velocity derivatives, and vanishes for a non-differential rotation of the fluid, it can be written in a general form as [Landau and Lifshitz, 1966]

$$\Pi = \eta \left[\nabla \mathbf{v}_m^\top + (\nabla \mathbf{v}_m^\top)^\top - \frac{2}{3} (\nabla \mathbf{v}_m) \mathbf{1} \right] + \xi (\nabla \mathbf{v}_m) \mathbf{1}. \quad (2.15)$$

Shearing motions of the fluid result in internal friction which dissipates momentum and kinetic energy of the flow. Macroscopically, these motions manifest themselves as a shear viscosity η . Similarly, if a fluid is expanded or compressed, dissipative processes occur, and they are described by the bulk viscosity ξ .

The second moment of the Boltzmann equation is obtained by multiplication with $m\mathbf{v}^2$, and integration over velocity space. It is given by

$$\frac{\partial \rho u}{\partial t} + \nabla(\rho u \mathbf{v}_m) = -p \nabla \mathbf{v}_m + (\Pi \nabla) \mathbf{v}_m - \nabla Q, \quad (2.16)$$

with the specific internal energy

$$u = \frac{1}{2} \langle |\mathbf{v}_r|^2 \rangle, \quad (2.17)$$

and the heat flux

$$\mathbf{Q} = \frac{1}{2} \rho \langle \mathbf{v}_r |\mathbf{v}_r|^2 \rangle. \quad (2.18)$$

Equation (2.16) describes the evolution of the internal energy of the fluid, which changes locally due to an internal energy flux, a velocity divergence, and the presence of a heat flux. If the small scale random velocities are isotropic, the heat flux is zero. On the other hand, the presence of a temperature gradient can skew the distribution and heat is transferred from hot to cold.

In this section we have derived the continuity equations of mass (2.11), momentum (2.12), and energy (2.16) of a fluid. They describe the evolution and relation between the hydrodynamic quantities velocity, density, pressure, and internal energy, and they are commonly referred to as Navier-Stokes equations.

2.1.2 Euler equations

If the terms of viscosity and thermal conduction are neglected in the Navier-Stokes equations, one obtains the Euler equations of ideal hydrodynamics. They build a system of hyperbolic partial differential equations, and simply represent the conservation of mass, momentum, and energy. The Euler equations read

$$\frac{\partial \rho}{\partial t} + \nabla(\rho \mathbf{v}) = 0, \quad (2.19)$$

$$\frac{\partial \rho \mathbf{v}}{\partial t} + \nabla(\rho \mathbf{v} \mathbf{v}^T) + \nabla p = 0, \quad (2.20)$$

$$\frac{\partial \rho e}{\partial t} + \nabla((\rho e + p) \mathbf{v}) = 0. \quad (2.21)$$

Compared to the Navier-Stokes equations, we have expressed the energy equation in terms of the total energy $e = u + \frac{1}{2} \mathbf{v}^2$, and moreover, we use for simplicity the notation $\mathbf{v}_m = \mathbf{v}$. The unknown quantities are density ρ , velocity \mathbf{v} , pressure p , and total energy per unit mass e . The set (ρ, \mathbf{v}, p) are the so-called primitive variables, and $(\rho, \rho \mathbf{v}, \rho e)$ the conserved variables. Since there are six unknowns but only five equations, the system has to be closed with an additional equation of state connecting

the pressure of the fluid with its density and internal energy. For this purpose, a common choice is the equation of state of an ideal gas,

$$p = \rho u(\gamma - 1), \quad (2.22)$$

where γ denotes the adiabatic index. We derive this equation explicitly in the framework of statistical mechanics in Section 2.1.3.

The Euler equations can also be written in compact form as

$$\frac{\partial \mathbf{u}}{\partial t} + \nabla \mathbf{f} = 0, \quad (2.23)$$

or

$$\frac{\partial \mathbf{u}}{\partial t} + \sum_{\alpha=1}^3 \frac{\partial \mathbf{f}_\alpha}{\partial x_\alpha} = 0, \quad (2.24)$$

highlighting the conservation properties of the system. Here we have introduced the state vector \mathbf{u} , which is

$$\mathbf{u} = \begin{pmatrix} \rho \\ \rho \mathbf{v} \\ \rho e \end{pmatrix} = \begin{pmatrix} \rho \\ \rho \mathbf{v} \\ \rho u + \frac{1}{2} \rho \mathbf{v}^2 \end{pmatrix}, \quad (2.25)$$

and the flux vectors

$$\mathbf{f}_1 = \begin{pmatrix} \rho v_1 \\ \rho v_1^2 + p \\ \rho v_1 v_2 \\ \rho v_1 v_3 \\ (\rho e + p)v_1 \end{pmatrix} \quad \mathbf{f}_2 = \begin{pmatrix} \rho v_2 \\ \rho v_1 v_2 \\ \rho v_2^2 + p \\ \rho v_2 v_3 \\ (\rho e + p)v_2 \end{pmatrix} \quad \mathbf{f}_3 = \begin{pmatrix} \rho v_3 \\ \rho v_1 v_3 \\ \rho v_2 v_3 \\ \rho v_3^2 + p \\ (\rho e + p)v_3 \end{pmatrix}. \quad (2.26)$$

The Euler equations are dimensionless, if they are rewritten in terms of a characteristic length, density, and velocity of the system, they take the exact same form. For this reason we omit units in the hydrodynamic test problems presented in this thesis.

Despite their simplicity, the Euler equations are non-linear and can give rise to rather complex flow dynamics, such as turbulent motions or hydrodynamic shocks. The latter represent discontinuities in the solution, which can develop even for smooth initial conditions.

On the other hand, in the absence of discontinuities it can be shown that the entropy is constant along particle trajectories,

$$\frac{ds}{dt} = 0. \quad (2.27)$$

Along flow lines, the gas changes therefore only adiabatically, $p \propto \rho^\gamma$. Moreover, if the solution is sufficiently smooth, we can apply the chain rule in equation (2.24) and the Euler equations can be written in quasilinear form

$$\frac{\partial \mathbf{u}}{\partial t} + \sum_{\alpha=1}^3 \mathbf{A}_\alpha \frac{\partial \mathbf{u}}{\partial x_\alpha} = 0, \quad (2.28)$$

with $\mathbf{A}_i(\mathbf{u}) = \frac{\partial \mathbf{f}_i(\mathbf{u})}{\partial \mathbf{u}}$. The matrices \mathbf{A}_i are called flux Jacobians and depend on the state vector \mathbf{u} and therefore on space and time. However, further insight into the mathematical system can be obtained if we *assume* that flux Jacobians are constant. Moreover, per definition of a hyperbolic system, the matrices \mathbf{A}_i are diagonalisable with real eigenvalues. Each matrix can be transformed to diagonal form \mathbf{D}_i via its eigendecomposition,

$$\mathbf{A}_i = \mathcal{R}_i \mathbf{D}_i \mathcal{L}_i \quad \Rightarrow \quad \mathbf{D}_i = \mathcal{L}_i \mathbf{A}_i \mathcal{R}_i, \quad (2.29)$$

with $\mathcal{L}_i = \mathcal{R}_i^{-1}$. The columns of the transformation matrix \mathcal{R}_i correspond to the right eigenvectors, and the rows of \mathcal{L}_i to the left eigenvectors of \mathbf{A}_i . We show all transformation matrices of the flux Jacobians in the Appendix.

It is not possible to diagonalize all $\mathbf{A}_i, i \in \{1, 2, 3\}$ simultaneously, and we hence study as an example the one-dimensional case

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{A}_1 \frac{\partial \mathbf{u}}{\partial x_1} = 0. \quad (2.30)$$

Multiplying this equation by $\mathcal{L}_1 =: \mathcal{L}_x$, inserting the identity $\mathcal{R}_x \mathcal{L}_x$, and using the constancy of the transformation matrices yields

$$\frac{\partial \mathcal{L}_x \mathbf{u}}{\partial t} + \mathcal{L}_x \mathbf{A}_1 \mathcal{R}_x \frac{\partial \mathcal{L}_x \mathbf{u}}{\partial x_1} = 0. \quad (2.31)$$

The variables obtained by the transformation of the state vector \mathbf{u} are called characteristic variables. With the definition $\mathbf{c} = \mathcal{L}_x \mathbf{u}$, equation (2.31) becomes

$$\frac{\partial \mathbf{c}}{\partial t} + \mathbf{D}_1 \frac{\partial \mathbf{c}}{\partial x_1} = 0 \quad \Leftrightarrow \quad \frac{\partial c_i}{\partial t} + \lambda_i \frac{\partial c_i}{\partial x_1} = 0, \quad i = 1, 2, 3, \quad (2.32)$$

representing a decoupled system of advection equations. The eigenvalues in the one-dimensional case are $\lambda_i = \{v - c, v, v + c\}$, with the velocity v and sound speed c of the gas. Sound waves are small adiabatic perturbations in the gas and we discuss them in Section 2.1.4. For each advection equation the solution simply corresponds to the advection of an initial shape,

$$c_i(x, t) = c_i(x - \lambda_i t, 0), \quad i = 1, 2, 3, \quad (2.33)$$

and the constant advection velocity is given by the eigenvalue. Hence, the eigenvalues correspond to the speeds of information propagation of the different waves. Moreover, the characteristic variables are conserved along space-time curves $X(t) = x_0 + \lambda_i t$, which are called characteristic curves or simply ‘characteristics’.

The solution for the state vector \mathbf{u} can be obtained by retransforming the characteristic variables, viz.

$$\mathbf{u} = \mathcal{R}_x \mathbf{c} = \sum c_i \mathbf{r}^i, \quad (2.34)$$

where \mathbf{r}^i are the rows of \mathcal{R}_x and therefore indicate the right eigenvectors. This linear combination highlights the nature of the solution \mathbf{u} , which is a superposition of the different waves. The wave strengths are given by the coefficients c_i , and the wave velocities by the eigenvalues λ_i .

The analysis above provides us intuition about the nature of the hydrodynamic equations. Moreover, the characteristic variables are the ‘natural’ variables of the Euler equations, and they can be used for designing powerful hydrodynamic solvers.

The characteristic variables are constant along the characteristic curves, and for systems of linear differential equations they can not collide in the $(x-t)$ -plane. However, this may happen in the non-linear Euler equations, resulting in the formation of shock waves. Shocks are therefore a manifestation of non-linearity, and we will study them in detail in Section 2.2. Note that at locations of discontinuities, such as shocks, the linearization of the Euler equations (2.28) can not be accomplished, and a different strategy has to be adopted for studying these phenomena.

2.1.3 Ideal gas

In many cases, an adequate approximation for the equation of state of a fluid consists of the ideal gas. This theoretical concept assumes elastically colliding point particles in random motion, without any additional interactions. In what follows, we give a derivation of the equation of state of a classical ideal gas by means of statistical mechanics.

Given all possible states of a statistical system, a general strategy for inferring thermodynamic properties is the calculation of a thermodynamic potential, e.g. the entropy. In a second step, the thermodynamic variables, such as pressure or temperature, can be obtained by calculating derivatives of the potential.

The possible states of an isolated system for which the particle number N , volume V , and total energy E is conserved can be represented with the density distribution of the microcanonical ensemble,

$$\rho_{\text{MCE}} = \begin{cases} \text{const} & \text{for } E < H(\mathbf{q}, \mathbf{p}) < E + \Delta \\ 0 & \text{else.} \end{cases} \quad (2.35)$$

Here $H(\mathbf{q}, \mathbf{p})$ is the Hamiltonian of the system depending on the generalized coordinates \mathbf{q} and momenta \mathbf{p} , and the measurement uncertainty of the macroscopic system

is taken into account in the form of an energy range with size Δ . Moreover, every allowed microscopic state occurs with the same probability, which represents the fundamental postulate of statistical mechanics. The numerical value of the constant probability is given by the normalization

$$\int \rho_{\text{MCE}} d\mathbf{q}d\mathbf{p} \stackrel{!}{=} 1. \quad (2.36)$$

The ensemble density can also be expressed in terms of the number of allowed microstates. In the energy range $[E, E + \Delta]$ the number of states is given by the integral

$$\Omega(E) = \int_{E < H(\mathbf{q}, \mathbf{p}) < E + \Delta} d\Omega, \quad d\Omega = \frac{d^n p d^n q}{(2\pi\hbar)^n}, \quad (2.37)$$

where $d\Omega$ is the number of states in the phase space volume $d^n p d^n q$, and n denotes the total number of degrees of freedom of the system. Hence, the density distribution of the microcanonical ensemble reads

$$\rho_{\text{MCE}} = \frac{1}{(2\pi\hbar)^n} \begin{cases} \frac{1}{\Omega(E)} & \text{for } E < H(\mathbf{q}, \mathbf{p}) < E + \Delta \\ 0 & \text{else.} \end{cases} \quad (2.38)$$

With the number of micro states, the entropy of a statistical system can be defined, viz.

$$S = k_{\text{B}} \ln \Omega(E). \quad (2.39)$$

The entropy indicates the lack of information we have about the microscopic realization of the macroscopic state. The number of microscopic states increases rapidly with increasing energy, motivating the following alternative definition of entropy:

$$S = k_{\text{B}} \ln \Sigma(E), \quad \Sigma(E) = \int_{H(\mathbf{q}, \mathbf{p}) < E} d\Omega. \quad (2.40)$$

The two definitions are indeed equivalent up to terms of the order $\mathcal{O}(N)$, however, we will use the second one since it is often more convenient for practical calculations.

For an ideal gas with N particles of mass m , the Hamiltonian takes the simple form

$$H = \sum_{i=1}^{3N} \frac{p_i^2}{2m}. \quad (2.41)$$

Moreover, the phase space volume including all states with energies smaller than the total energy E of the system is

$$\Sigma(E) = \int_{H < E} \prod_{i=1}^{3N} \frac{dp_i dq_i}{2\pi\hbar}. \quad (2.42)$$

The Hamiltonian is independent of the coordinates, and we obtain for the gas confined to the volume V

$$\int \prod_{i=1}^{3N} dq_i = V^N. \quad (2.43)$$

The integral over the momenta can be solved by using the Heaviside step function $\Theta(x)$, which is 1 for positive arguments and 0 for negative arguments. One obtains

$$\begin{aligned} & \int \prod_{i=1}^{3N} \Theta\left(E - \sum_{i=1}^{3N} \frac{p_i^2}{2m}\right) dp_i \\ &= \int \prod_{i=1}^{3N} \Theta\left(2mE - \sum_{i=1}^{3N} p_i^2\right) dp_i = \omega_{3N}(\sqrt{2mE}), \end{aligned} \quad (2.44)$$

with the volume of the n -dimensional sphere $\omega_n(R) = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} R^n$. This expression includes the gamma function, which is simply $\Gamma(n) = (n-1)!$ for a positive integer argument. For a large number of particles a single particle does not contribute significantly to the total system and we assume that $n = 3N$ is an even number, such that we can replace the gamma function with the factorial. The phase space volume integral is hence

$$\Sigma(E) = \left(\frac{V}{h^3}\right)^N \frac{\pi^{3N/2}}{\frac{3N!}{2}} (2mE)^{3N/2}. \quad (2.45)$$

The entropy is proportional to the logarithm of the above equation, and for large n we can use Stirling's approximation for the factorial which reads $\ln n! \approx n(\ln n - 1)$. Finally, we obtain for the entropy of the ideal gas

$$S(E, V, N) = k_B \ln \Sigma(E) = k_B N \left\{ \ln \left[V \left(\frac{4\pi m E}{3h^2 N} \right)^{3/2} \right] + \frac{3}{2} \right\}. \quad (2.46)$$

An expression connecting pressure, temperature, and volume can now easily be calculated from the relation

$$\left(\frac{\partial S}{\partial V}\right)_E = \frac{p}{T}, \quad (2.47)$$

resulting in

$$pV = Nk_{\text{B}}T. \quad (2.48)$$

In order to obtain an energy equation, (2.46) can be solved for E ,

$$E(S, V, N) = \frac{3h^2}{4\pi m} \frac{N}{V^{2/3}} \exp\left(\frac{2}{3} \frac{S}{Nk_{\text{B}}} - 1\right). \quad (2.49)$$

By using the relation

$$\left(\frac{\partial S}{\partial E}\right)_V = \frac{1}{T}, \quad (2.50)$$

we obtain the energy equation of the ideal gas,

$$E = \frac{3}{2}Nk_{\text{B}}T. \quad (2.51)$$

Equation 2.51 expresses the equipartition theorem of classical statistical systems; every degree of freedom contributes $1/2k_{\text{B}}T$ to the total energy. If there are f degrees of freedom per atom, we can combine equation (2.48) with (2.51) in order to obtain a more general equation of state, viz.

$$p = \rho u(\gamma - 1). \quad (2.52)$$

Here u denotes the thermal energy per unit mass, and we have introduced the adiabatic index $\gamma = \frac{f+2}{f}$. This index is $5/3$ for a monoatomic gas with 3 (translational) degrees of freedom per atom, and $7/5$ for a diatomic gas having 3 translational and 2 rotational degrees of freedom. Formally, the adiabatic index is defined as

$$\gamma = \frac{c_p}{c_v}, \quad (2.53)$$

where c_p and c_v are the specific heat at constant pressure and constant volume, respectively.

Next, we derive the equation $p \propto \rho^\gamma$, which holds for adiabatic processes. The first law of thermodynamics is

$$dE = TdS - pdV, \quad (2.54)$$

and for the case of constant entropy reads

$$dE = -pdV. \quad (2.55)$$

The energy of an ideal gas as a function of pressure and volume is obtained by combining (2.48) with (2.51), it is

$$E = \frac{1}{\gamma - 1} pV. \quad (2.56)$$

Inserting this expression into (2.55) yields

$$\frac{dp}{p} = -\gamma \frac{dV}{V}, \quad (2.57)$$

and hence

$$p \propto V^{-\gamma}. \quad (2.58)$$

We conclude that for adiabatic processes the function

$$s = \frac{p}{\rho^\gamma} \quad (2.59)$$

is constant. This quantity is also called entropic function and represents a convenient measure for the entropy per unit mass in hydrodynamic contexts. For the sake of convenience we will use the terms ‘entropic function’ and ‘entropy’ interchangeably in this thesis.

2.1.4 Sound waves

Sound waves are small adiabatic fluid oscillations carrying information at a characteristic speed c . They can be derived from the Euler equations by the ansatz of small perturbations in pressure and density,

$$p = p_0 + \tilde{p} \quad (2.60)$$

$$\rho = \rho_0 + \tilde{\rho}. \quad (2.61)$$

We transform to a frame in which the bulk velocity of the fluid vanishes and we can hence assume $v \ll c$. By inserting the ansatz into (2.19) and (2.20) and by neglecting quadratic terms of the perturbations we obtain

$$\frac{\partial \tilde{p}}{\partial t} + \rho_0 \nabla \mathbf{v} = 0, \quad (2.62)$$

$$\frac{\partial \mathbf{v}}{\partial t} + \frac{1}{\rho_0} \nabla \tilde{p} = 0. \quad (2.63)$$

Moreover, if we assume that the small oscillations are adiabatic, changes in pressure can be related to changes in density by

$$\tilde{p} = \left(\frac{\partial p}{\partial \rho} \right)_S \tilde{\rho}, \quad (2.64)$$

where the index S indicates that the derivative has to be calculated at constant entropy. Substituting the density in equation (2.62) yields

$$\frac{\partial \tilde{p}}{\partial t} + \rho_0 \left(\frac{\partial p}{\partial \rho} \right)_S \nabla \cdot \mathbf{v} = 0. \quad (2.65)$$

As a next step, the velocity field can be decomposed via Helmholtz's theorem into a curl-free and a divergence-free part. The latter vanishes, since taking the curl of both sides of equation (2.63) implies $\nabla \times \mathbf{v} = 0$ for oscillatory motions. Moreover, the curl-free field can be expressed in terms of a velocity potential ϕ , resulting in

$$\mathbf{v} = \nabla \phi. \quad (2.66)$$

Finally, by combining the equations (2.63), (2.65), and (2.66) we obtain the wave equation

$$\frac{\partial^2 \phi}{\partial t^2} - c^2 \Delta \phi = 0, \quad (2.67)$$

with the sound speed

$$c = \sqrt{\left(\frac{\partial p}{\partial \rho} \right)_S}. \quad (2.68)$$

Integration of equation (2.67) in the one-dimensional case ($\phi(x, y, z) = \phi(x)$) gives solutions for the velocity potential of the form

$$\phi = f_1(x - ct) + f_2(x + ct), \quad (2.69)$$

corresponding to translational motions of initial perturbations $f_1(x)$ and $f_2(x)$ in the positive and negative directions, respectively. By integrating equations (2.62), (2.63), and (2.66), solutions of the same form are obtained for the perturbations in pressure, density, and velocity.

For an ideal gas, the sound speed can be calculated by expressing the pressure in terms of the entropic functions s , viz.

$$p = s\rho^\gamma. \quad (2.70)$$

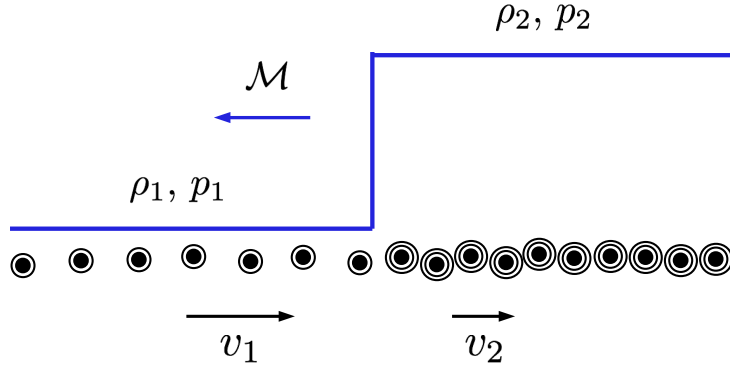


Figure 2.1: Basic picture of a collisional shock with gas streaming supersonically from left to right. At the position of the shock the gas is slowed down, compressed, and heated. Moreover, the gas particles cross the shock surface, which is moving to the left.

With the definition (2.68) of the sound speed one obtains

$$c = \sqrt{\gamma \frac{p}{\rho}} = \sqrt{\gamma \frac{k_B T}{\mu m_p}}. \quad (2.71)$$

2.2 Hydrodynamic shocks

The Euler equations of ideal hydrodynamics are non-linear, giving rise to non-linear waves and discontinuities in the hydrodynamic properties. Most importantly, waves can steepen into shock waves.

In Fig. 2.1 we sketch the basic picture of a shock. Gas particles stream in from the left and collide with slower gas particles. During the collision the gas gets abruptly compressed and heated, thereby creating a discontinuity surface moving to the left. This surface is called a shock surface, and it separates two distinct hydrodynamic states. In this thesis we follow the common convention and indicate pre-shock gas properties with the index 1, and post-shock properties with the index 2. Behind the shock, the gas has an increased density and pressure ($\rho_2 > \rho_1$ and $p_2 > p_1$). Moreover, the shock irreversibly converts kinetic energy of the pre-shock gas into thermal energy, and increases the entropy of the gas. We will refer to this mechanism as ‘energy dissipation’ throughout the thesis.

v_1 is the velocity of the incoming gas in the rest frame of the shock. It is usually expressed in units of the pre-shock sound speed, $\mathcal{M} = v_1/c_1$, which is called the Mach number of the shock. A shock can form for $\mathcal{M} > 1$, when the inflowing gas stream moves supersonically. Since information propagates in the gas at the speed of sound, this implies that, until the moment of collision, the pre-shock gas is not affected by the shock.

While the Euler equations in differential form give rise to shocks, a discontinuity in the solution can not be evolved further in time, resulting in a break down of the

mathematical system. In computational hydrodynamics this intrinsic problem can be circumvented by solving the equations in an integral form, e.g. by means of a finite volume method or discontinuous Galerkin method. An introduction to these methods will be given in Section 3.

In nature, shocks are not perfect discontinuities. They rather consist of a transition layer with a scale of the order of the mean free path of the gas particles. In this layer, the energy dissipation occurs by means of viscous effects and heat conduction. The latter two are absent in the Euler equations of ideal hydrodynamics, and the physics inside the transition layer is not included in our hydrodynamic simulations. Fortunately, the impact on the gas of the shock and the associated energy dissipation can be unambiguously characterized by the hydrodynamic quantities of the pre- and post-shock state. In this sense, a quantitatively correct treatment of shocks can be achieved by solving the Euler equations.

We want to note that most astrophysical shocks differ from classical hydrodynamic shocks with respect to the physics in the transition layer. In many astrophysical environments the mean free path of the particles is large compared to the transition layer of the shock. Hence, the abrupt deceleration of gas particles does not originate from collisions. This type of shock is called ‘collisionless shock’, and can be explained by the presence of magnetic fields giving raise to electromagnetic viscosity.

In what follows we outline the different types of discontinuities and derive the jump conditions for shock waves. Moreover, we calculate the energy dissipation at shocks and outline the basic mechanism for particle acceleration at shocks.

2.2.1 Hydrodynamic discontinuities

At a discontinuity surface one or several hydrodynamic quantities change abruptly. Nevertheless, due to the conservation of mass, momentum, and energy, we can assume that the mass, momentum, and energy fluxes are continuous across the discontinuity. We choose a coordinate system moving with the discontinuity, and with the x -axis pointing in the direction of the surface normal. In this case, the continuity of the mass flux is given by $\rho_1 v_{1x} = \rho_2 v_{2x}$, where the velocities are fluid velocities relative to the surface, and the indices 1 and 2 indicate pre- and post-shock quantities, respectively. In the following, we adopt the bracket notation from Landau and Lifshitz [1966], $\rho_1 v_{1x} - \rho_2 v_{2x} = [\rho v_x]$, and hence the flux continuity constraints read

$$[\rho v_x] = 0, \tag{2.72}$$

$$[p + \rho v_x^2] = 0, \quad [\rho v_x v_y] = 0, \quad [\rho v_x v_z] = 0, \tag{2.73}$$

$$[(\rho e + p)v_x] = 0. \tag{2.74}$$

Here we have used the momentum flux components of the flux vector in the x -direction, which is given by \mathbf{f}_1 in (2.26).

Two types of discontinuities can be distinguished. In the first type, the mass flux across the discontinuity is zero, implying $v_{1x} = v_{1y} = 0$, and $p_1 = p_2$. These

discontinuities are called ‘tangential discontinuities’, since the other hydrodynamic quantities, including the tangential velocities v_y and v_z , can be discontinuous.

For discontinuous tangential velocities a shearing flow is present which triggers a fluid instability, the so-called Kelvin-Helmholtz instability, resulting in turbulence and the disappearance of the discontinuity surface.

On the other hand, if the tangential velocities are continuous ($[v_y] = 0$ and $[v_z] = 0$), the discontinuity is called ‘contact discontinuity’. In the latter only the density and density-dependent quantities jump across the surface, giving a stable configuration.

The second type of discontinuity is obtained if the mass flux across the discontinuity surface is non-zero, $\rho_1 v_{1x} = \rho_2 v_{2x} \neq 0$. It follows from the momentum equations (2.73) that this non-zero mass flux results in continuous tangential velocities, $v_{1y} = v_{2y}$ and $v_{1z} = v_{2z}$. A jump occurs in all the other hydrodynamic variables, and the discontinuity is a hydrodynamic shock surface.

Across the shock, the flux continuity equations (2.72), (2.73), and (2.74) read

$$[\rho v_x] = 0, \tag{2.75}$$

$$[p + \rho v_x^2] = 0, \quad [v_y] = 0, \quad [v_z] = 0, \tag{2.76}$$

$$\left[u + \frac{p}{\rho} + \frac{v_x^2}{2} \right] = 0. \tag{2.77}$$

Here we have inserted $e = u + \mathbf{v}^2/2$ in the energy flux equation and canceled the non-zero and continuous mass flux ρv_x . Moreover, \mathbf{v}^2 has been replaced by v_x^2 , since v_y^2 and v_z^2 can be subtracted from both sides of the equation.

With these equations we can calculate the jump experienced by different hydrodynamic quantities at a shock. This calculation will be carried out in the next section.

2.2.2 The Rankine–Hugoniot jump conditions

In this section we derive the jumps experienced by different hydrodynamic quantities across a shock. The corresponding relations are known as the Rankine-Hugoniot jump conditions.

To begin with, we introduce the symbols $j := \rho_1 v_{1x} = \rho_2 v_{2x}$ for the mass flux, and $\bar{V} = 1/\rho$ for the specific volume. If we insert equation (2.75) in (2.76) and use these definitions, we straightforwardly obtain for the mass flux the relation

$$j^2 = \frac{p_2 - p_1}{\bar{V}_1 - \bar{V}_2}. \tag{2.78}$$

Moreover, the velocities in (2.77) can now be expressed as $v_{1x} = j\bar{V}_1$ and $v_{2x} = j\bar{V}_2$. Rewriting equation (2.77) with (2.78) then yields

$$u_1 - u_2 + \frac{1}{2}(\bar{V}_1 - \bar{V}_2)(p_1 + p_2) = 0. \tag{2.79}$$

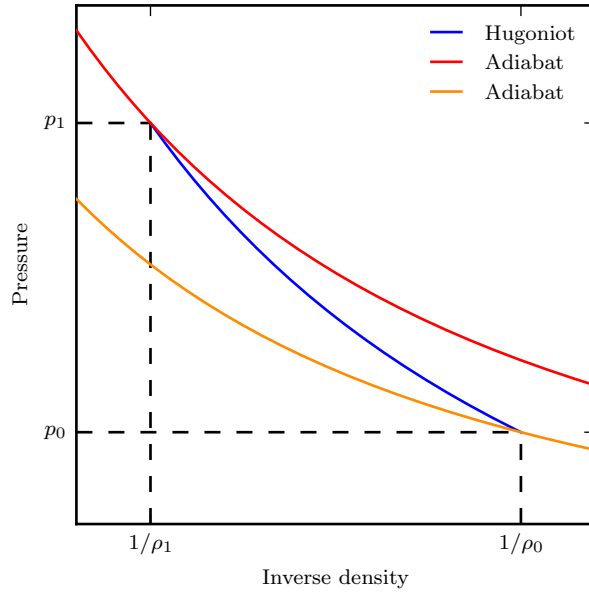


Figure 2.2: Comparison of a Hugoniot, which connects pre- and post-shock states, with adiabatic curves. While the latter are defined as curves of constant entropy, the entropy is increased for gas particles moving along the Hugoniot.

This equation connects pre- and post-shock quantities, and is known as the Hugoniot relation. The corresponding curve in the $\bar{V}p$ -plane is simply called Hugoniot, and we show a comparison with adiabatic curves in Fig. 2.2. An adiabat is defined as a curve of constant entropy, $p \propto \rho^\gamma$ or $p\bar{V}^\gamma = \text{const}$. If a sound wave passes a gas particle, it moves up and down on an adiabat. As can be seen from the figure, the start and end point of a Hugoniot lie on two different adiabats, and hence the entropy changes along a Hugoniot. The gas entropy of the post-shock state (p_2, ρ_2) is higher than the entropy of the pre-shock state (p_1, ρ_1) , indicating the irreversible process of energy dissipation.

If we combine the Hugoniot relation (2.79) with the equation of state $p = \rho u(\gamma - 1)$, we can obtain an equation for the density ratios, viz.

$$\frac{\rho_1}{\rho_2} = \frac{\bar{V}_2}{\bar{V}_1} = \frac{(\gamma + 1)p_1 + (\gamma - 1)p_2}{(\gamma + 1)p_2 + (\gamma - 1)p_1} \quad (2.80)$$

This equation can be used to eliminate \bar{V}_2 from the mass flux,

$$\begin{aligned} j^2 &= \frac{p_2 - p_1}{\bar{V}_1 - \bar{V}_2} = \frac{p_2 - p_1}{\bar{V}_1 \left(1 - \frac{(\gamma+1)p_1 + (\gamma-1)p_2}{(\gamma+1)p_2 + (\gamma-1)p_1} \right)} \\ &= \frac{(p_2 - p_1)[(\gamma + 1)p_2 + (\gamma - 1)p_1]}{\bar{V}_1(2p_2 - 2p_1)} = \frac{(\gamma + 1)p_2 + (\gamma - 1)p_1}{2\bar{V}_1} \end{aligned} \quad (2.81)$$

With (2.80) and (2.81) it is straightforward to derive the jump conditions for the density, velocity, and pressure. The pressure jump can be obtained from

$$\begin{aligned} v_1^2 &= M^2 c_1^2 = j^2 \bar{V}_1^2 = \frac{\bar{V}_1}{2} [(\gamma + 1)p_2 + (\gamma - 1)p_1] \frac{p_1}{p_1} \\ &= \frac{c_1^2}{2\gamma} [(\gamma - 1) + (\gamma + 1) \frac{p_2}{p_1}], \end{aligned} \quad (2.82)$$

resulting in

$$\frac{p_2}{p_1} = \frac{2\gamma \mathcal{M}^2}{\gamma + 1} - \frac{\gamma - 1}{\gamma + 1}. \quad (2.83)$$

For the density jump we get

$$\frac{\rho_2}{\rho_1} \stackrel{(2.80)}{=} \frac{(\gamma - 1) + (\gamma + 1) \frac{p_2}{p_1}}{(\gamma + 1) + (\gamma - 1) \frac{p_2}{p_1}} \stackrel{(2.83)}{=} \frac{(\gamma - 1) + 2\gamma \mathcal{M}^2 - (\gamma - 1)}{(\gamma + 1) + \frac{\gamma - 1}{\gamma + 1} [2\gamma \mathcal{M}^2 - (\gamma - 1)]}. \quad (2.84)$$

The mathematical expression on the right-hand side can be simplified, yielding

$$\frac{v_1}{v_2} = \frac{\rho_2}{\rho_1} = \frac{(\gamma + 1)\mathcal{M}^2}{(\gamma - 1)\mathcal{M}^2 + 2}. \quad (2.85)$$

We note that the density jump corresponds to the inverse jump of the velocities due to continuity of the mass flux.

The Rankine-Hugoniot jump conditions for the pressure (2.83) and density (2.85) can now be used to calculate additional hydrodynamic quantities, such as the temperature and the entropy. With $T \propto p/\rho$ and $s = p/\rho^\gamma$ we get

$$\frac{T_2}{T_1} = \frac{[2\gamma \mathcal{M}^2 - (\gamma - 1)][(\gamma - 1)\mathcal{M}^2 + 2]}{(\gamma + 1)^2 \mathcal{M}^2}, \quad (2.86)$$

$$\frac{s_2}{s_1} = \left(\frac{2\gamma \mathcal{M}^2}{\gamma + 1} - \frac{\gamma - 1}{\gamma + 1} \right) \left(\frac{(\gamma - 1)\mathcal{M}^2 + 2}{(\gamma + 1)\mathcal{M}^2} \right)^\gamma. \quad (2.87)$$

Fig. 2.3 shows the jump conditions as a function of Mach number for a gas with $\gamma = 5/3$. For $\mathcal{M} \rightarrow \infty$ the pressure, the temperature, and the entropy scale as \mathcal{M}^2 , whereas the density converges asymptotically towards $(\gamma + 1)/(\gamma - 1)$. This limit indicates the maximum gas compression that can occur at a strong shock. For a monoatomic gas ($\gamma = 5/3$), the maximum density ratio is 4, whereas for a diatomic gas ($\gamma = 7/5$) the equation of state is ‘softer’, resulting in a maximum compression ratio of 6.

For this thesis we are interested in detecting and characterizing shocks in cosmological simulations of the Universe. We hence have to develop an algorithm which can reliably detect shocks, measure their Mach numbers, and compute energy dissipation rates. Fig. 2.3 indicates which hydrodynamic quantities can be used for achieving

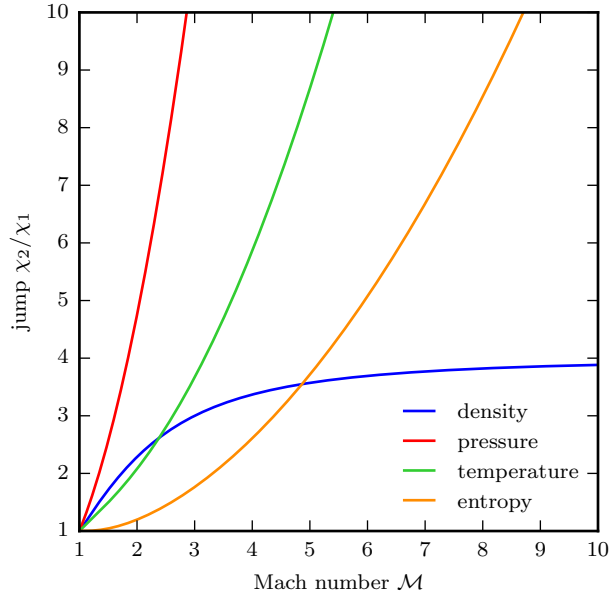


Figure 2.3: Visualization of the Rankine-Hugoniot jump conditions for a gas with $\gamma = 5/3$. For each hydrodynamic quantity we plot the ratio of the post-shock value to the pre-shock value (χ_2/χ_1) versus the Mach number of the shock.

this task. Clearly, the density jump flattens for high Mach numbers and is therefore not sensitive in the regime of strong shocks. Moreover, the entropy jump is also not suitable, since it is flat for low Mach numbers ($\mathcal{M} \lesssim 3$). On the other hand, the pressure and the temperature jump are sensitive over the whole Mach number range, and represent good candidates for measuring the Mach number in simulations. For the shock finding methodology adopted in this thesis we prefer the temperature jump; our algorithm is described and motivated in Part II of the thesis.

2.2.3 Energy dissipation at shocks

For smooth flows, the Euler equations conserve entropy along flow lines, $\frac{ds}{dt} = 0$. However, as we have seen in Section 2.2.2, the entropy of gas is increased at shocks. They *irreversibly* convert kinetic energy into thermal energy, and we term this process ‘energy dissipation’.

The energy dissipation at shocks can be inferred by comparing the thermal energy densities of shock-heated gas in the post-shock region and unshocked gas in the pre-shock region. Care has to be taken in this calculation, since the energy density increases at the shock not only due to irreversible processes, but also due to adiabatic compression.

Let us consider a box filled with an ideal gas of thermal energy density ϵ , density ρ , and pressure p , which obeys the ideal gas law $p = \epsilon(\gamma - 1)$. The hydrodynamic

quantities $\hat{\epsilon}$, $\hat{\rho}$, and \hat{p} after adiabatic compression can be calculated straight-forwardly, viz.

$$\tilde{\epsilon} = \frac{\tilde{p}}{\gamma - 1} = \epsilon \frac{\tilde{p}}{p} = \epsilon \left(\frac{\tilde{\rho}}{\rho} \right)^\gamma. \quad (2.88)$$

In the last step we have used the adiabatic equation $p \propto \rho^\gamma$.

The generated thermal energy flux at a shock therefore takes the form [Ryu et al., 2003]

$$f_{\text{th}} = \left[\epsilon_2 - \epsilon_1 \left(\frac{\rho_2}{\rho_1} \right)^\gamma \right] v_2 = (\epsilon_2 - \epsilon_1 R^\gamma) v_2, \quad \text{with } R := \frac{\rho_2}{\rho_1}. \quad (2.89)$$

The units of f_{th} are unit energy per unit shock surface area and unit time. In order to obtain shock dissipation rates in this thesis, we multiply this flux with the shock surface area, as explained in detail in Part II.

The energy for the thermalization process is provided by the kinetic energy of the supersonically incoming gas streams. In the frame of the shock, the kinetic energy flux is given by

$$f_\Phi = \frac{1}{2} \rho_1 v_1^3 = \frac{1}{2} \rho_1 (c_1 \mathcal{M})^3. \quad (2.90)$$

The generated thermal energy flux can be expressed as a fraction of this kinetic energy flux, viz.

$$f_{\text{th}} = \delta(\mathcal{M}) f_\Phi, \quad (2.91)$$

where $\delta(\mathcal{M})$ is the thermalization efficiency. The efficiency depends only on the Mach number and can be calculated as follows.

$$\begin{aligned} \delta(\mathcal{M}) &= \frac{f_{\text{th}}}{f_\Phi} = \frac{(\epsilon_2 - \epsilon_1 R^\gamma) v_2}{\frac{1}{2} \rho_1 v_1^3} = \frac{2(p_2 - p_1 R^\gamma) v_2}{(\gamma - 1) \rho_1 c_1^2 \mathcal{M}^2 v_1} = \frac{2(p_2 - p_1 R^\gamma)}{\gamma(\gamma - 1) \frac{\rho_2}{\rho_1} p_1 \mathcal{M}^2} \\ &= \frac{2}{\gamma(\gamma - 1) \mathcal{M}^2 R} \left(\frac{p_2}{p_1} - R^\gamma \right) \\ &= \frac{2}{\gamma(\gamma - 1) \mathcal{M}^2 R} \left(\frac{2\gamma \mathcal{M}^2 - (\gamma - 1)}{\gamma + 1} - R^\gamma \right). \end{aligned} \quad (2.92)$$

In this equation, the density jump R appears, which has been derived in Section 2.2.2, and reads

$$R \equiv \frac{\rho_2}{\rho_1} = \frac{(\gamma + 1) \mathcal{M}^2}{(\gamma - 1) \mathcal{M}^2 + 2}. \quad (2.93)$$

In Fig. 2.4 we plot the thermalization efficiency $\delta(\mathcal{M})$ as a function of Mach number for a gas with $\gamma = 5/3$. It increases steeply for Mach numbers between 2 and 5, and

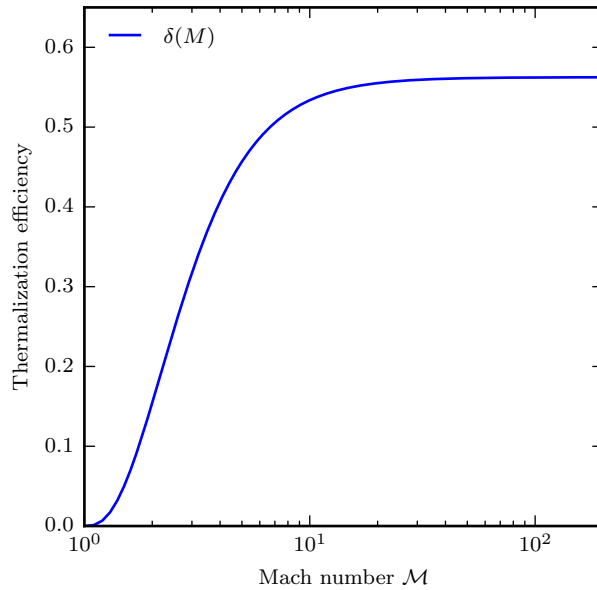


Figure 2.4: Thermalization efficiency as a function of Mach number.

an efficiency of around 50% is reached at $\mathcal{M} \approx 7$. For $\mathcal{M} \rightarrow \infty$ the thermalization efficiency asymptotes to the value 0.56.

Equations (2.92), (2.93) and Fig. 2.4 describe the thermalization efficiency at shocks without considering cosmic rays. The possibility of cosmic ray acceleration and the corresponding efficiencies are addressed in the next section.

2.2.4 Diffusive shock acceleration

Radio synchrotron emission in galaxy clusters provides direct evidence for the existence of relativistic electrons and magnetic fields [see for example Brunetti and Jones, 2014, for a review on non-thermal emission in galaxy clusters]. It is believed that hydrodynamic shocks can accelerate particles to relativistic energies via the process of diffusive shock acceleration [DSA; Axford et al., 1977; Bell, 1978a,b; Blandford and Ostriker, 1978; Krymskii, 1977; Malkov and O’C Drury, 2001]. Cosmological hydrodynamic shocks are collisionless, meaning that particle interactions are mediated by magnetic fields¹. In the DSA process, suprathermal particles leak from the post-shock region to the pre-shock region, against the hydrodynamic flow. Upstream, they can excite Alfvén waves, and scatter back downstream. In resonance with the Alfvén waves, they can cross the shock multiple times, thereby gaining more and

¹These interactions are not resolved in cosmological simulations, and the plasma is therefore approximated by simulating a collisional fluid.

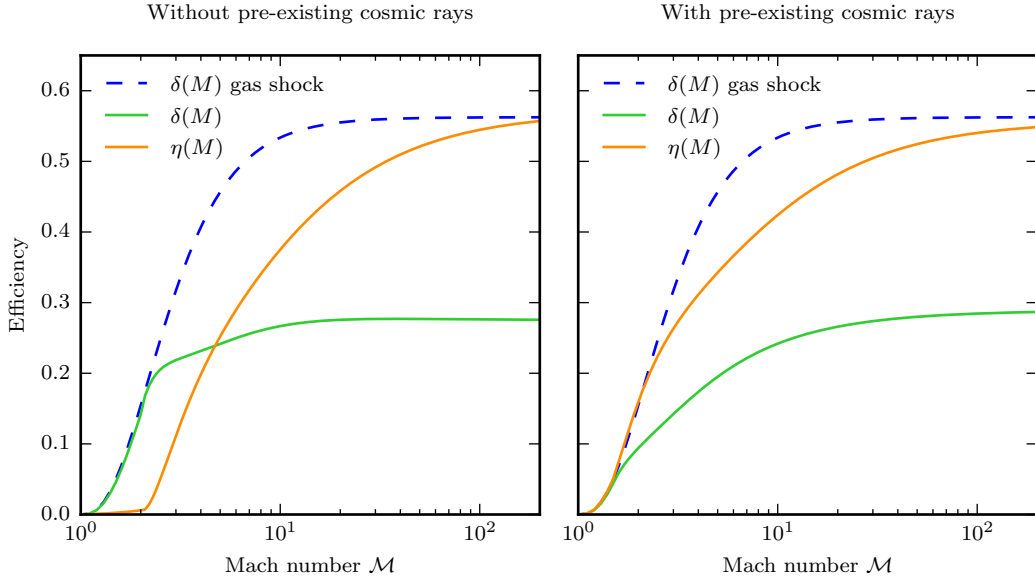


Figure 2.5: Efficiencies for a cosmic ray accelerating shock without pre-existing cosmic rays (left-hand side panel), and with pre-existing cosmic rays (right-hand side panel). The thermalization efficiency is shown in green, the cosmic ray acceleration efficiency in orange, and the blue line indicates the thermalization efficiency of a normal gas shock for reference.

more energy, before they are ejected as relativistic cosmic ray particles². Since many particles are accelerated at the same location, a significant fraction of the kinetic energy processed by a shock can be transferred to cosmic rays.

Understanding the DSA mechanism in detail is subject to current research, and involves numerical simulations on the microscopic level. Moreover, while DSA is efficient for protons, there is an injection problem for cosmic ray electrons. Electrons have a smaller mass and Larmor radius, and moving upstream is therefore more difficult for these particles. Simulations indicate that for the acceleration of electrons additional processes are important, such as shock surfing acceleration and shock drift acceleration [e.g. Guo et al., 2014; Matsumoto et al., 2012].

For estimating cosmic ray acceleration efficiencies in this thesis, we use results from DSA simulations of Kang and Jones [2007a,b]. By running simulations of shocks with different Mach numbers, which is the key parameter in the acceleration process, they could infer fitting functions for the acceleration efficiency. These are shown in Fig. 2.5. The left-hand panel indicates efficiencies for the case that no pre-existing cosmic rays are present, and the right-hand panel shows the efficiencies in the presence of pre-existing cosmic rays. Pre-existing cosmic rays have been accelerated in shocks at earlier times, and their reacceleration is more efficient compared to the acceleration from the thermal pool.

²The process of energy transfer to a cosmic ray particle is similar to a tennis ball gaining momentum and energy when the two players approach each other.

The energy flux channeled to cosmic rays is given by

$$f_{\text{cr}} = \eta(\mathcal{M})f_{\Phi}, \quad (2.94)$$

and as can be seen in the figure, this can be a significant fraction of the incoming kinetic energy flux, in particular at high Mach number shocks.

2.3 Analytic test problems

2.3.1 Shock tube solution

One simple way to study the evolution of non-linear waves in ideal hydrodynamics consists of evolving initial conditions in the form of two constant states. This one-dimensional initial value problem is known as the Riemann problem. It is given by a left state \mathbf{W}_L with the gas properties $(\rho_L, \mathbf{v}_L, p_L)$, a right state \mathbf{W}_R with $(\rho_R, \mathbf{v}_R, p_R)$, and a discontinuity at $x = 0$ separating these states.

Mathematical and numerical schemes designed for solving this problem are called Riemann solvers, and the first exact Riemann solver for non-linear hyperbolic systems has been introduced by Godunov [1959].

The Riemann problem is of great relevance in computational hydrodynamics. It can be used as a hydrodynamic test for assessing the correctness and quality of a hydrodynamic code, and moreover, a Riemann solver is an important building block of many state-of-the-art hydrodynamic solvers for the Euler equations.

In this section we follow Toro [2009] and outline the general strategy for obtaining a solution for the Riemann problem. Moreover, we discuss exemplarily the result obtained for the classical shock tube problem [Sod, 1978]. A Sod shock tube problem is a special Riemann problem with vanishing velocities in the initial conditions.

When the discontinuous initial conditions of the Riemann problem are evolved, three different waves emerge instantly, corresponding to the three eigenvalues $\{\lambda_1, \lambda_2, \lambda_3\} = \{v - c, v, v + c\}$ of the one-dimensional system. The eigenvalues indicate the wave speeds, and since the waves originate from $x = t = 0$, which is a point in space-time, the solution of the problem has to be self-similar, $\mathbf{u} = \mathbf{u}(x/t)$.

By analysing the eigenvectors and characteristics of the middle wave with speed $\lambda_2 = v$, one finds that the characteristics are parallel to $\frac{dx}{dt} = v$. This corresponds to the pure advection of a hydrodynamic profile, and the gas velocity is constant across the wave. Hence, the wave can be characterized as a contact discontinuity. For each of the outer waves, the characteristics can either diverge, corresponding to a rarefaction wave, or converge, corresponding to a shock wave.

In Fig. 2.6 we show a typical wave structure for a shock tube problem, where the left wave is a rarefaction, and the right wave a shock. The three waves separate four different and constant states. The states beyond the outer waves correspond to the states of the initial conditions and are denoted by \mathbf{W}_L and \mathbf{W}_R . The region between the outer waves is referred to as the star region, which is divided by the

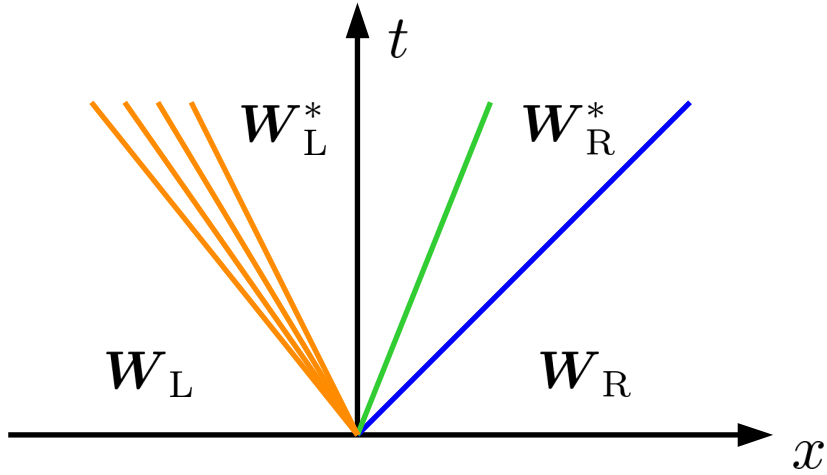


Figure 2.6: A typical wave structure for a shock tube problem. Four different hydrodynamic states are separated by, from left to right, a rarefaction wave (orange), a contact discontinuity (green), and a shock (blue).

contact into \mathbf{W}_L^* and \mathbf{W}_R^* . The pressure and the velocity of the gas is continuous across the contact discontinuity, $p_L^* = p_R^* =: p^*$ and $v_L^* = v_R^* =: v^*$.

In order to solve the Riemann problem, we have to consider the different occurring cases. To begin with, if the pressure in the star region is larger than the pressure in the left initial state, the left wave is compressive, that is to say a shock wave. In this case we can straight-forwardly connect the states \mathbf{W}_L and \mathbf{W}_L^* by the flux continuity equations (2.75), (2.76), and (2.77), and the Rankine-Hugoniot jump conditions as follows.

According to (2.78) the mass flux between the states is

$$j_L^2 = \frac{p^* - p_L}{1/\rho_L - 1/\rho_L^*}. \quad (2.95)$$

In this equation the unknown quantities are p^* and ρ_L^* . An expression for the latter is given by (2.80), viz.

$$\rho_L^* = \rho_L \frac{[(\gamma - 1)/(\gamma + 1) + p^*/p_L]}{(\gamma - 1)/(\gamma + 1)p^*/p_L + 1}. \quad (2.96)$$

Hence, ρ_L^* can be eliminated in the mass flux, yielding

$$j_l = \left[\frac{p^* + B_L}{A_L} \right]^{1/2}, \quad (2.97)$$

with the symbols

$$A_L = \frac{2}{(\gamma + 1)\rho_L}, \quad \text{and} \quad B_L = \frac{(\gamma - 1)}{(\gamma + 1)}p_L. \quad (2.98)$$

The pressure in the star region p^* depends also on the right state \mathbf{W}_R , and can therefore not simply be calculated from the jump across the left wave. For this reason, we first derive equations connecting the unknown velocity v^* to the left and right state, respectively. In a second step, these equations can be combined yielding an equation for the pressure in the star region.

The velocity in the star region can be calculated from the flux continuity equations across the shock. The continuity of the mass flux for a left shock wave reads

$$j_L = \rho_L \hat{v}_L = \rho_L^* \hat{v}^*. \quad (2.99)$$

Here the hats indicate that the velocities in this equation are the velocities relative to the moving shock surface, that is in the frame of the shock. They are connected to the velocities in the laboratory frame of reference by

$$\hat{v}_L = v_L - S_L, \quad \text{and} \quad \hat{v}^* = v^* - S_L, \quad (2.100)$$

where S_L denotes the speed of the shock. Inserting equation (2.99) into the momentum continuity equation

$$\rho_L \hat{v}_L^2 + p_L = \rho_L^* (\hat{v}^*)^2 + p^* \quad (2.101)$$

yields

$$\hat{v}^* = \hat{v}_L - \frac{p^* - p_L}{j_L} \quad \Leftrightarrow \quad v^* = v_L - \frac{p^* - p_L}{j_L}. \quad (2.102)$$

By inserting the expression (2.97) for the mass flux one obtains

$$v^* = v_L - f_L^s(p^*, \mathbf{W}_L), \quad (2.103)$$

with the function

$$f_L^s(p^*, \mathbf{W}_L) = (p^* - p_L) \left[\frac{A_L}{p^* + B_L} \right]^{1/2}. \quad (2.104)$$

If the pressure in the star region is smaller than the pressure in the left initial state, the left wave is a rarefaction wave. This wave does not represent a discontinuity like a contact or a shock wave, but it consists of a spatially extended continuous structure. In what follows, we will give a shortened derivation of this wave type.

While the characteristic variables are constant *along* the characteristics, there are also quantities which are constant *across* the wave structure. These quantities

are called Riemann invariants. For details the interested reader may be referred to Jeffrey [1976].

Across a rarefaction wave the Riemann invariants are the entropy s , and the quantity $v + 2c/(\gamma - 1)$. Constant entropy implies $p \propto \rho^\gamma$, and hence

$$\rho_L^* = \rho_L \left(\frac{p^*}{p_L} \right)^{(1/\gamma)}. \quad (2.105)$$

This equation can be rewritten in terms of the sound speed, yielding

$$c_L^* = c_L \left(\frac{p^*}{p_L} \right)^{\frac{\gamma-1}{2\gamma}}. \quad (2.106)$$

The equation for the second generalized Riemann invariant reads

$$v_L + \frac{2c_L}{\gamma - 1} = v^* + \frac{2c_L^*}{\gamma - 1}. \quad (2.107)$$

Solving for the velocity in the star region and substituting the sound speed via equation (2.106) leads to the result

$$v^* = v_L - f_L^r(p^*, \mathbf{W}_L), \quad (2.108)$$

where

$$f_L^r(p^*, \mathbf{W}_L) = \frac{2c_L}{(\gamma - 1)} \left[\left(\frac{p^*}{p_L} \right)^{\frac{\gamma-1}{2\gamma}} - 1 \right]. \quad (2.109)$$

Analogous to the calculations above, the right wave can be connected to the velocity and pressure of the star region. For a right shock wave one obtains

$$v^* = v_R + f_R^s(p^*, \mathbf{W}_R), \quad (2.110)$$

and a right rarefaction wave results in

$$v^* = v_R + f_R^r(p^*, \mathbf{W}_R). \quad (2.111)$$

Finally, by equating the velocities in the star region we can connect the left and right state of the initial conditions, viz.

$$f(p^*, \mathbf{W}_L, \mathbf{W}_R) \equiv f_L(p^*, \mathbf{W}_L) + f_R(p^*, \mathbf{W}_R) + v_R - v_L = 0, \quad (2.112)$$

with

$$f_L(p^*, \mathbf{W}_L) = \begin{cases} f_L^s(p^*, \mathbf{W}_L) & \text{if } p^* > p_L \text{ (left shock wave),} \\ f_L^r(p^*, \mathbf{W}_L) & \text{if } p^* \leq p_L \text{ (left rarefaction wave),} \end{cases} \quad (2.113)$$

and

$$f_{\text{R}}(p^*, \mathbf{W}_{\text{R}}) = \begin{cases} f_{\text{R}}^{\text{s}}(p^*, \mathbf{W}_{\text{R}}) & \text{if } p^* > p_{\text{R}} \text{ (right shock wave),} \\ f_{\text{R}}^{\text{r}}(p^*, \mathbf{W}_{\text{R}}) & \text{if } p^* \leq p_{\text{R}} \text{ (right rarefaction wave).} \end{cases} \quad (2.114)$$

With this algebraic equation we can calculate the pressure in the star region, and subsequently v^* with (2.110) or (2.111). However, no closed analytic solution exists for (2.112), and we have to adopt numerical tools. The derivative $f'(p)$ can be calculated analytically, such that the Newton-Raphson root finding scheme proves to be suitable. It reads

$$p_n = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})}, \quad (2.115)$$

and requires a starting value p_0 . This initial guess is however not very critical, since f in our problem is monotone and concave, and has only one root. The iteration is stopped when the relative pressure change is sufficiently small. In most of the cases this approach is successful. However, there are situations with very low density and pressure values, which require a different solution strategy. For inferring these so-called vacuum solutions, a detailed outline may be found in Toro [2009].

Once the pressure and the velocity of the star region is known, it is straight forward to calculate the full solution. For a shock wave, the density in the adjacent star region can be computed from equation (2.96). Lastly, the speed of the left shock wave is

$$S_{\text{L}} = v_{\text{L}} - \frac{j_{\text{L}}}{\rho_{\text{L}}}, \quad (2.116)$$

with the mass flux j_{L} given by (2.97).

The solution in the case of a left rarefaction consists of a continuous change of the hydrodynamic quantities between the characteristics with speeds $S_{\text{head}} = v_{\text{L}} - c_{\text{L}}$ and $S_{\text{tail}} = v^* - c_{\text{L}}^*$. The solution for a characteristic within the rarefaction fan can be calculated by combining its velocity $\frac{dx}{dt} = v - c$ with the equations (2.105) and (2.107) of the Riemann invariants. One obtains

$$\mathbf{W}_{\text{L}}^{\text{fan}} = \begin{cases} \rho = \rho_{\text{L}} \left[\frac{2}{(\gamma+1)} + \frac{\gamma-1}{(\gamma+1)c_{\text{L}}} \left(v_{\text{L}} - \frac{x}{t} \right) \right]^{\frac{2}{\gamma-1}}, \\ v = \frac{2}{\gamma+1} \left[c_{\text{L}} + \frac{\gamma-1}{2} v_{\text{L}} + \frac{x}{t} \right], \\ p = p_{\text{L}} \left[\frac{2}{\gamma+1} + \frac{\gamma-1}{(\gamma+1)c_{\text{L}}} \left(v_{\text{L}} - \frac{x}{t} \right) \right]^{\frac{2\gamma}{\gamma-1}}. \end{cases} \quad (2.117)$$

The results for a right shock wave and a right rarefaction wave are calculated analogously.

In order to obtain the solution of the Riemann problem at location x and time t , we can take advantage of its self-similar nature. The solution depends only on the speed $S = \frac{x}{t}$, and if we can infer the position of the corresponding line in space-time

with respect to the three different waves, the solution is known immediately. This task can be accomplished simply by comparing S to the speeds of the different waves, and is called sampling of the solution.

In Fig. 2.7 we show the solution for the classical Sod shock tube problem [Sod, 1978]. The initial conditions consist of the two constant states $\mathbf{W}_L = (\rho_L = 1, v_L = 0, p_L = 1)$ and $\mathbf{W}_R = (\rho_R = 0.125, v_R = 0, p_R = 0.1)$. We have computed the solution of this shock tube problem with an exact Riemann solver at different times, and show the result in space-time in the upper panels. This is a direct visualization of the wave structure, which consists of a left rarefaction wave, a contact discontinuity, and a right shock wave. Hence, the wave structure at hand is the same as shown in the sketch of Fig. 2.6. The three waves separate four constant states, which are clearly visible in the density plot. On the other hand, in the pressure plot only three different states can be seen, since the pressure across the contact discontinuity is constant. The bottom panels present the solution at $t = 0.228$. In this representation the shape of the rarefaction wave can be seen more easily. It is created by a fan of characteristics, resulting in a smooth change of density and pressure across the wave.

2.3.2 Sedov-Taylor blast wave solution

Another famous hydrodynamical test problem is the strong explosion, which creates a spherical shock known as Sedov-Taylor blast wave. This phenomenon occurs if a large amount of energy is released in a small region and on a short time scale. A famous astrophysical example are supernovae. These stellar explosions are highly energetic and can easily outshine a typical galaxy.

The Sedov-Taylor blast wave problem can be modelled hydrodynamically by injecting the energy E into a point-like region of homogeneous cold gas at rest. The gas has the pressure p_1 and density ρ_1 , corresponding to pre-shock quantities of the developing shock.

If we assume that the pressure of the cold pre-shock gas is negligible compared to the pressure behind the shock, $p_2 \gg p_1$, and that we can use the ideal gas law as a suitable equation of state, the problem has an analytic and self-similar solution. In what follows we outline the derivation of the solution following Landau and Lifshitz [1966].

With the assumptions above, the problem is fully specified by the pre-shock density ρ_1 and the injected energy E . Hence, the distance of the shock from the centre of the explosion has to be a combination of these quantities and of the time variable t . For dimensional reasons there is only one possible combination, and the radius of the spherical explosion as a function of time has to take the form

$$R(t) = \beta \left(\frac{Et^2}{\rho_1} \right)^{1/5}, \quad (2.118)$$

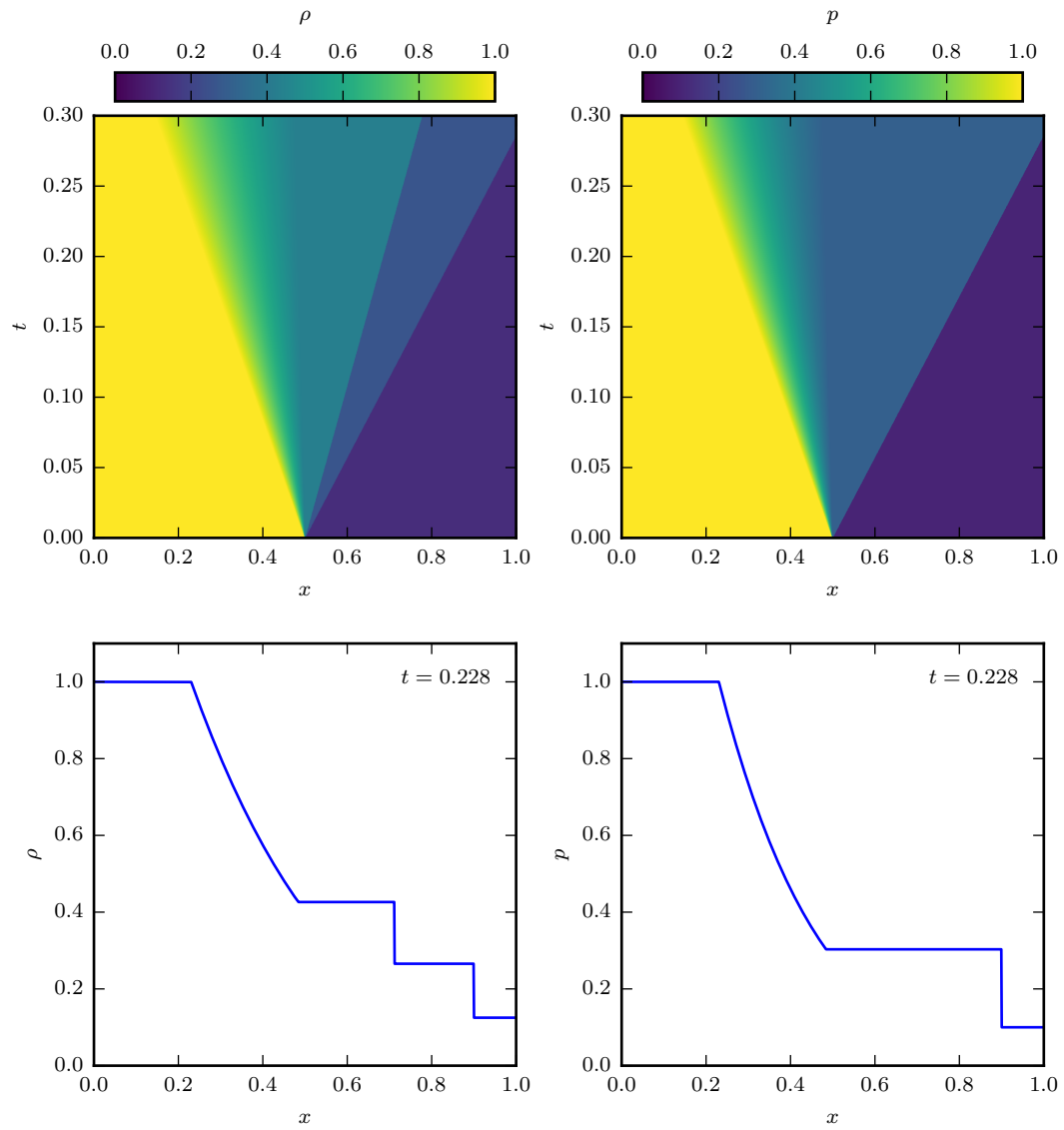


Figure 2.7: Solution for the density ρ and pressure p of the classical Sod shock tube problem. The upper panels show the solution in space-time, revealing the different waves discussed in this section. The initial conditions give rise to a left rarefaction wave, a central contact discontinuity, and a right shock wave. The lower panels show the solution at $t = 0.228$.

where β is a dimensionless constant. The speed of the shock surface is given by the derivative of (2.118), viz.

$$S = \frac{dR}{dt} = \beta \left(\frac{E}{\rho_1} \right)^{1/5} \frac{2}{5} \frac{1}{t^{3/5}} = \frac{2R}{5t}. \quad (2.119)$$

In order to solve the full problem, we can use the spherically symmetric version of the Euler equations. In spherical symmetry the flux becomes $\nabla \mathbf{f} = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 f_r)$ and the pressure gradient is $\nabla p = \frac{\partial p}{\partial r} \hat{\mathbf{e}}_r$. With these substitutions the mass conservation equation (2.19) takes the form

$$\frac{\partial v}{\partial t} + v \frac{\partial v}{\partial r} = -\frac{1}{\rho} \frac{\partial p}{\partial r}, \quad (2.120)$$

and the momentum conservation equation (2.20) becomes

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho v)}{\partial r} + \frac{2\rho v}{r} = 0. \quad (2.121)$$

We complete the set of equations by the entropy conservation law, which for spherical symmetry reads

$$\frac{dS}{dt} = 0 \quad \Leftrightarrow \quad \left(\frac{\partial}{\partial t} + v \frac{\partial}{\partial r} \right) \ln \frac{p}{\rho^\gamma} = 0. \quad (2.122)$$

The solution in the interior of the Sedov-Taylor shock wave can be obtained by integrating equations (2.120), (2.121), and (2.122) with appropriate boundary conditions. The latter are given by the post-shock quantities and can be obtained from the Rankine-Hugoniot jump conditions as follows.

In the limit of a strong shock ($p_2 \gg p_1$), it is

$$\frac{\rho_2}{\rho_1} = \frac{\gamma + 1}{\gamma - 1}. \quad (2.123)$$

In this case the velocities in the shock frame are

$$\frac{\hat{v}_2}{\hat{v}_1} = \frac{\rho_1}{\rho_2} = \frac{\gamma - 1}{\gamma + 1}. \quad (2.124)$$

The pre-shock gas is at rest, such that $\hat{v}_1 = -S$, where S is the speed of the shock. Moreover, \hat{v}_2 can be transformed to the laboratory frame of reference according to $\hat{v}_2 = v_2 - S$, and equation (2.124) becomes

$$v_2 = \frac{2}{\gamma + 1} S. \quad (2.125)$$

In Section 2.2.2 we have derived the equation

$$\dot{v}_1^2 = \frac{\bar{V}_1}{2}[(\gamma + 1)p_2 + (\gamma - 1)p_1]. \quad (2.126)$$

Compared to the post-shock pressure, the pressure of the unshocked gas is negligible and we can write

$$\dot{v}_1^2 = \frac{\bar{V}_1}{2}(\gamma + 1)p_2. \quad (2.127)$$

Solving for the post-shock pressure yields

$$p_2 = \frac{2}{\gamma + 1}\rho_1 S^2. \quad (2.128)$$

For the Sedov-Taylor blast wave problem, the equations (2.123), (2.125), and (2.128) represent the boundary conditions for the system of differential equations (2.120), (2.121), and (2.122). By inserting the shock speed (2.119), we can infer the scaling of the post-shock quantities with time. While the density jump is time independent, the post-shock velocity and pressure decrease with $t^{-3/5}$ and $t^{-6/5}$, respectively.

In order to obtain the full solution, it is advisable to utilize its self-similarity, and define the variable

$$\xi = \frac{r}{R(t)} = \frac{r}{\beta} \left(\frac{\rho_1}{Et^2} \right)^{1/5}. \quad (2.129)$$

The hydrodynamic quantities can then be rewritten as

$$v = \frac{2r}{5t}V, \quad \rho = \rho_1 G, \quad c^2 = \frac{4r^2}{25t^2}Z, \quad (2.130)$$

with the dimensionless variables $V(\xi)$, $G(\xi)$, and $Z(\xi)$.

One integral of the system of equations can be found by the following physical line of argumentation. The pressure in the unshocked gas is very small, and its total energy is hence negligible. In this case, the total energy within the Sedov-Taylor shock sphere is conserved. Since the solution is self-similar, the energy within any other smaller sphere centred on the point explosion and expanding with the velocity $v_n = 2r/(5t)$ is conserved as well. The latter conservation can be expressed by the equality between the energy leaving the sphere due to the energy flux and the increase of the enclosed energy,

$$4\pi r^2 v(\rho e + p) dt = 4\pi r^2 v_n \rho e dt.$$

Inserting the expressions (2.130) for the dimensionless quantities and simplifying the equation results in

$$Z(\xi) = \frac{\gamma(\gamma - 1)(1 - V)V^2}{2(\gamma V - 1)}. \quad (2.131)$$

The remaining equations of the full solution can be obtained by a straight-forward integration, and one obtains [Landau and Lifshitz, 1966]

$$\xi^5 = \left[\frac{2}{(\gamma + 1)V} \right]^2 N_1(V)^{n_1} N_2(V)^{n_2} \quad (2.132)$$

$$G(\xi) = \frac{\gamma + 1}{\gamma - 1} N_2(V)^{n_3} N_1(V)^{n_4} \left[\frac{\gamma + 1}{\gamma - 1} (1 - V) \right]^{n_5}. \quad (2.133)$$

with

$$N_1(V) = \frac{\gamma + 1}{7 - \gamma} [5 - (3\gamma - 1)V], \quad (2.134)$$

$$N_2(V) = \frac{\gamma + 1}{\gamma - 1} (\gamma V - 1), \quad (2.135)$$

$$n_1 = -\frac{13\gamma^2 - 7\gamma + 12}{(3\gamma - 1)(2\gamma + 1)}, \quad n_2 = \frac{5(\gamma - 1)}{2\gamma + 1}, \quad (2.136)$$

$$n_3 = \frac{3}{2\gamma + 1}, \quad n_4 = -\frac{n_1}{2 - \gamma}, \quad n_5 = -\frac{2}{2 - \gamma}. \quad (2.137)$$

Last but not least, the constant β can be calculated by demanding that the total energy within the Sedov-Taylor blast wave sphere amounts to E ,

$$E = \int_0^R \rho \left(\frac{v^2}{2} + \frac{c^2}{\gamma(\gamma - 1)} \right) 4\pi r^2 dr. \quad (2.138)$$

With the dimensionless quantities the expression reads

$$\beta^5 \frac{16\pi}{25} \int_0^1 G \left[\frac{V^2}{2} + \frac{Z}{\gamma(\gamma - 1)} \right] \xi^4 d\xi = 1. \quad (2.139)$$

This integral has to be evaluated numerically, which can be accomplished by the code provided in Kamm and Timmes [2007]. One obtains the values $\beta = 1.152$ for $\gamma = 5/3$, and $\beta = 1.033$ for $\gamma = 7/5$.

Fig. 2.8 shows the full self-similar solution of the Sedov-Taylor blast wave problem as a function of the dimensionless parameter $\xi = r/R$. The adiabatic index of the gas is $\gamma = 5/3$. For $\xi \rightarrow 0$, the pressure approaches a constant value, while the velocity scales as

$$\frac{v}{v_2} \propto \xi = \frac{r}{R}, \quad (2.140)$$

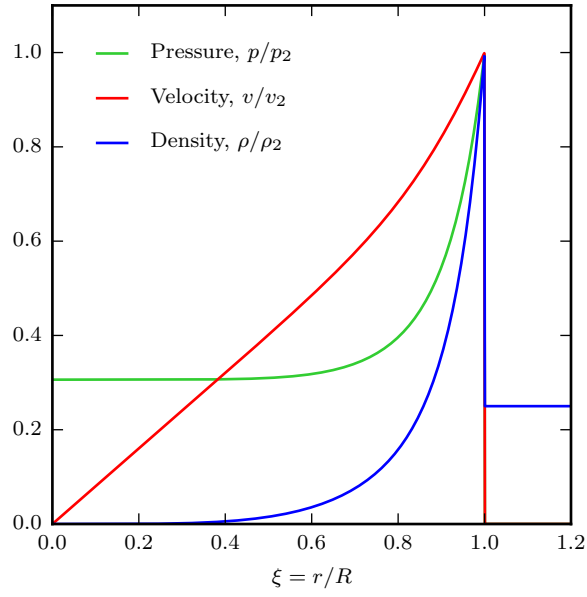


Figure 2.8: Self-similar solution of the Sedov-Taylor blast wave problem for an ideal gas with $\gamma = 5/3$. The hydrodynamic quantities are plotted with respect to their post-shock values.

and the density decreases with

$$\frac{\rho}{\rho_2} \propto \xi^{3/(\gamma-1)} = \left(\frac{r}{R}\right)^{3/(\gamma-1)}. \quad (2.141)$$

Each hydrodynamic quantity is plotted with respect to its post-shock value, such that all curves have the value 1 at the location of the discontinuity $\xi = 1$. Recall however that the density jump is time independent, whereas the post-shock velocity and post-shock pressure decrease with $t^{-3/5}$ and $t^{-6/5}$, respectively.

3

Computational hydrodynamics

3.1 Finite volume method

A finite volume method is a classical scheme for solving the hydrodynamic equations with the reputation of being robust and relatively easy to implement, at least for achieving second order accuracy. In these methods, the simulation volume is subdivided into cells with finite volumes, and their mean fluid values are evolved by solving the Euler equations in a weak, integral form. The strength of this method lies in the capability of capturing shocks and discontinuities, which is not possible by solving the Euler equations in differential form, as well as the strict conservation of the total mass, momentum, and energy inside the simulated volume.

3.1.1 The finite volume scheme

In this Section we derive the finite volume scheme for the one-dimensional Euler equations, and with an explicit first order time discretization. The generalization to a higher-order method for three-dimensions can be accomplished straight-forwardly.

The hyperbolic conservation laws we are interested in have the form

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{f}}{\partial x} = 0, \quad (3.1)$$

with the vector of conserved quantities

$$\mathbf{u} = \begin{pmatrix} \rho \\ \rho v \\ \rho e \end{pmatrix}, \quad (3.2)$$

and the flux vector

$$\mathbf{f} = \begin{pmatrix} \rho v \\ \rho v^2 + p \\ (\rho e + p)v \end{pmatrix}. \quad (3.3)$$

The general idea of the finite volume method consists of discretizing space into grid cells with size Δx , and evolving the mean cell values. The vector of the mean conserved quantities of cell i ranging from $x_{i-\frac{1}{2}}$ to $x_{i+\frac{1}{2}}$ is given by

$$\mathbf{U}_i = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathbf{u}(x, t) dx. \quad (3.4)$$

In order to obtain a first order time accurate evolution equation we integrate the conservation law (3.1) over cell i and time step $\Delta t = t^{n+1} - t^n$, which yields

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathbf{u}(x, t^{n+1}) - \mathbf{u}(x, t^n) dx + \int_{t^n}^{t^{n+1}} \mathbf{f}(x_{i+\frac{1}{2}}, t) - \mathbf{f}(x_{i-\frac{1}{2}}, t) dt = 0. \quad (3.5)$$

The terms of the first integral in this expression can be replaced by (3.4). The second integral describes the change of the conserved variables due to fluxes across the cell boundaries, and couples cell i to the neighbouring cells. Following the idea of Godunov, these fluxes can be calculated with a so-called Riemann solver. A Riemann solver computes a solution for the Riemann problem, consisting of a constant left and constant right state separated by a hydrodynamic discontinuity.

As we have discussed in detail in Section 2.3.1, different waves originate at the jump and propagate outwards, however, the solution at the initial location of the discontinuity is independent of time. We can hence replace the time integral over the physical fluxes with numerical fluxes $\bar{\mathbf{f}}_{i\pm\frac{1}{2}}$, which can either be calculated analytically with an approximate Riemann solver, or numerically with an exact Riemann solver.

In either case, equation (3.5) becomes

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \frac{\Delta t}{\Delta x} \left(\bar{\mathbf{f}}_{i-\frac{1}{2}} - \bar{\mathbf{f}}_{i+\frac{1}{2}} \right). \quad (3.6)$$

This equation represents the explicit update of the mean conserved quantities inside cell i , such that the total mass, momentum, and energy summed up for all cells are conserved. The latter feature allows the finite volume method to capture the nature of the underlying equations, and holds as long as the boundary conditions of the simulation volume are conserving, e.g. periodic.

The spatial order of the finite volume scheme depends on the hydrodynamic states entering the Riemann solver. If simply the mean cell quantities are used, one obtains

a first order scheme, which means that the global error of the solution decreases linearly with decreasing cell size Δx . A second order scheme can be obtained by calculating a slope for every fluid quantity and every cell, based on the hydrodynamic states of neighbouring cells. With this information one obtains a linear reconstruction of the solution inside the cell, and most importantly, at the cell boundaries. If these improved solutions at the cell boundaries enter the Riemann solver, a scheme with quadratic convergence in space is achieved.

Higher order methods can be constructed by reconstructing the solution with higher order polynomials, e.g. in the piece-wise parabolic method (PPM) the reconstructed solution consists of parabolas. Moreover, the reconstruction should always be accomplished conservatively, such that the integrated fluid quantities do not change.

After reconstructing the numerical solution is evolved by calculating the fluxes and updating the mean conserved quantities. The last step is also known as ‘averaging’, since the reconstructed solution is not further taken into account and we are only interested in the new means.

The steps discussed above can be summarized as a Reconstruct-Evolve-Average (REA) scheme, and the hydrodynamic system is continuously evolved in time by consecutive REA steps. The generalization to three spatial dimensions can be accomplished straight-forwardly, either by a dimensional splitting, or by applying the fluxes in different directions simultaneously (‘unsplit method’). Moreover, in practice the order of the time integration scheme should be equal to the order of the spatial discretization. A possible choice for higher order time integrators are Runge-Kutta methods, which we will discuss in Section 3.3.

The finite volume approach for solving the hydrodynamic equations is generally very robust and can capture hydrodynamic shocks well. However, the scheme also incorporates a couple of inherent disadvantages. In the averaging step numerical errors are introduced into the solution which accumulate with time and manifest themselves as a diffusive behaviour of the simulated fluid, also known as ‘numerical diffusion’. Moreover, with increasing spatial order of the scheme the amount of information needed inside every cell for computing the reconstruction increases as well. While the second order FV method in one dimension has a three-point stencil, meaning that every cell requires information from the two neighbouring cells, the PPM method has a widened stencil of five points. The latter method is therefore difficult to implement and requires a lot of communication in parallel computing applications.

3.1.2 Limiting

In what follows we discuss the linear reconstruction of the solution for a scalar quantity in the one-dimensional case. A linear reconstruction is most commonly used

and results in a second order FV scheme. The reconstructed solution $\tilde{u}(x, t)$ at time t_n within the cell having the index i can be written as

$$\tilde{u}(x, t) = U_i^n + \sigma_i^n(x - x_i), \quad (3.7)$$

where x_i denotes the centre of the cell, and σ_i^n the slope of the quantity u in the cell. Possible choices for the latter are the centred slope

$$\sigma_i^n = \frac{U_{i+1}^n - U_{i-1}^n}{2\Delta x}, \quad (3.8)$$

or one-sided slopes in the form of

$$\sigma_i^n = \frac{U_i^n - U_{i-1}^n}{\Delta x} \quad \text{or} \quad \sigma_i^n = \frac{U_{i+1}^n - U_i^n}{\Delta x}. \quad (3.9)$$

The one-sided slopes represent upwind and downwind slopes, depending on the direction of information propagation.

Unfortunately, for all three choices a problem arises at hydrodynamic discontinuities. The reconstructed solution can ‘overshoot’ or ‘undershoot’, meaning that it has a greater or smaller value at the cell boundary compared to the reconstructed solution in the neighbouring cell. In the subsequent evolve and average steps, a new extremum in the cell average values can be introduced. Moreover, after several REA cycles the numerical solution oscillates near the discontinuity. This can result in an unstable numerical scheme, and it is therefore desirable to design the reconstruction such that spurious oscillations are avoided.

Oscillations in the numerical solution can be measured by the total variation, which is given by

$$TV(U^n) = \sum_i |U_{i+1}^n - U_i^n|, \quad (3.10)$$

and a numerical scheme is said to be total variation diminishing (TVD) if the total variation does not increase with time,

$$TV(U^{n+1}) \leq TV(U^n). \quad (3.11)$$

It can be shown that a TVD scheme is monotonicity preserving, which means that if

$$U_i^n \geq U_{i+1}^n \quad \forall i, \quad (3.12)$$

it follows that

$$U_i^{n+1} \geq U_{i+1}^{n+1} \quad \forall i. \quad (3.13)$$

In particular, numerical oscillations do not appear at monotonic discontinuities. Hence, we are looking for a linear reconstruction, such that the resulting scheme is

TVD. In the case of the linear advection equation, this can simply be achieved by limiting the slopes of the linearly reconstructed solution such that it does not over- or undershoot.

One example of a commonly used TVD limiter is the minmod limiter, for which the gradient of the reconstruction is calculated as

$$\sigma_i^n = \text{minmod} \left(\frac{U_i^n - U_{i-1}^n}{\Delta x}, \frac{U_{i+1}^n - U_i^n}{\Delta x} \right), \quad (3.14)$$

where

$$\text{minmod}(a, b) = \begin{cases} s \min(|a|, |b|) & s = \text{sign}(a) = \text{sign}(b) \\ 0 & \text{otherwise.} \end{cases} \quad (3.15)$$

In this approach both one-sided slopes are computed and compared, and the gradient of the cell is set to the slope with smaller absolute value. However, if the two slopes point in different directions, the gradient is set to zero.

In cells where the slope limiter is active, the convergence order of the scheme is decreased. Hence, the limiter should only be applied to problematic locations, such that the solution in smooth regions remains unaffected and converges with the maximum order. This can be achieved by adopting a bounded version of the minmod limiter, which modifies the solution only where its gradient is large. For details of this approach we refer the reader to Section 19.3.

As outlined in the discussion above, by constructing an appropriate slope limiter a TVD scheme can be obtained, and spurious numerical oscillations can be avoided. However, this only holds for the linear advection equation. In this case the reconstruction step is the only possible source of new extrema, and hence the only step in the REA cycle where oscillations can be introduced.

This is however not the case for non-linear differential equations, such as the Euler equations of ideal hydrodynamics. Due to the non-linearity of the equation its true solution is not TVD, and the total variation can increase arbitrarily within one timestep. As a consequence the numerical scheme can introduce new extrema also in the evolve step, and the TVD property can not be guaranteed. On the other hand, a TVD scheme is also not desired in locations where new extrema are generated in the true solution.

For the Euler equations it is a common choice to adopt a TVD slope limiter, such that the reconstruction step does not introduce spurious new extrema at hydrodynamic discontinuities. Nevertheless, for some combinations of numerical schemes and applications, also non-TVD slope limiters are used, which are less restrictive in limiting the solution. They account for the fact that isolating discontinuities, and confining the limiter to these locations, proves to be difficult. The choice of the slope limiter is therefore in many cases a tradeoff between avoiding spurious oscillations and avoiding the clipping of smooth extrema.

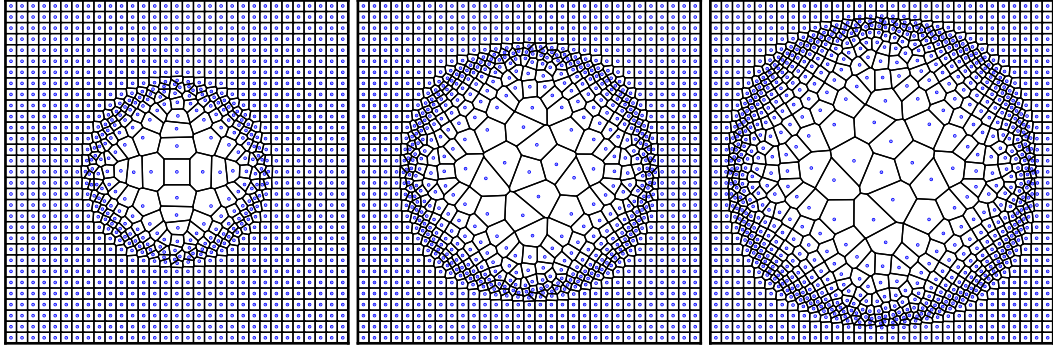


Figure 3.1: Illustration of the moving Voronoi mesh in a 2D Sedov blast wave simulation. A Voronoi cell contains all points in space which are closer to its own mesh-generating point than to all other mesh-generating points. The mesh-generating points (blue) are moved with the local fluid velocity, such that the resolution of the mesh adapts automatically.

3.1.3 Finite volume method on a moving Voronoi mesh

Finite volume methods on AMR grids, as well as SPH, are the most commonly used methods in computational astrophysics. Both methods have certain advantages, however, they are also very different, indicating that their strengths are quite different in nature. A finite volume method operating on a moving-mesh can combine advantages of each, and has been realized by Springel [2010] in the AREPO code.

For a mesh moving with the hydrodynamic flow, the mass fluxes across cell interfaces are very small, giving rise to an approximately Lagrangian scheme (‘quasi-Lagrangian’). These schemes have very small advection errors and are inherently Galilean invariant, meaning that the bulk velocity of the gas does not introduce any errors. Moreover, the mesh structure adapts automatically such that the mass per cell is roughly kept constant in a natural way. The automatic adaptivity offers another advantage, which is of great importance for this thesis. Gas at hydrodynamic shocks gets compressed and the resolution of the mesh increases in the immediate post-shock region. In this way, shocks are captured very well across only a couple of cells. All things considered, the moving-mesh approach combines the accuracy of a grid code with Lagrangian features.

A very suitable spatial discretization for a moving mesh is the Voronoi tessellation. It can be computed from a set of mesh-generating points, and each Voronoi cell contains all points in space which are closer to its own mesh-generating point than to any other mesh-generating point. The Voronoi tessellation consists of disjoint polygons of different shapes and sizes, and the mesh changes continuously when the mesh-generating points are moved. Figure 3.1 shows from left to right the time evolution of the moving Voronoi mesh in a 2D Sedov blast wave simulation. The mesh-generating points, which are advected with the local fluid velocity, are shown in blue. Even though we do not plot any hydrodynamic quantities, the mesh clearly reveals the spherical shock solution, highlighting its automatic adaptivity.

In the following, we give a brief outline for constructing a finite volume method on the moving Voronoi mesh, and discuss AREPO's second order FV scheme for solving the equations of ideal hydrodynamics. In compact form, the Euler equations in 3D can be written as

$$\frac{\partial \mathbf{u}}{\partial t} + \nabla \mathbf{f} = 0, \quad (3.16)$$

where \mathbf{u} is the vector of the primitive variables, and \mathbf{f} the hydrodynamic flux.

In the finite volume scheme we are interested in the mean conserved quantities of the cells, and for cell i those are

$$\mathbf{U}_i = \int_{V_i} \mathbf{u} \, dV. \quad (3.17)$$

If the cell moves and changes its shape the volume of the integration is time-dependent, $V = V(t)$, and the total derivative of \mathbf{U}_i follows from Reynold's transport theorem, viz.

$$\frac{d\mathbf{U}_i}{dt} = \int_{V_i(t)} \frac{\partial \mathbf{u}}{\partial t} + \nabla(\mathbf{u} \mathbf{w}^\top) \, dV. \quad (3.18)$$

Here \mathbf{w} denotes the velocity field of the transforming cell. By inserting the conservation law (3.16) and using Gauss' divergence theorem one obtains

$$\frac{d\mathbf{U}_i}{dt} = - \int_{\partial V_i(t)} [\mathbf{f}(\mathbf{u}) - \mathbf{u} \mathbf{w}^\top] \, d\mathbf{n}. \quad (3.19)$$

For a Voronoi cell, which in the three-dimensional case is a polyhedron, we can express the right-hand side of (3.19) as a sum over the interfaces, viz.

$$\frac{d\mathbf{U}_i}{dt} = - \sum_j A_{ij} \mathbf{f}_{ij}. \quad (3.20)$$

\mathbf{f}_{ij} denotes the flux between the cells i and j , and is given by

$$\mathbf{f}_{ij} = \frac{1}{A_{ij}} \int_{A_{ij}} [\mathbf{f}(\mathbf{u}) - \mathbf{u} \mathbf{w}^\top] \, d\mathbf{A}_{ij}. \quad (3.21)$$

In these expressions \mathbf{A}_{ij} denotes the oriented area of the interface between cell i and neighbour j , and moreover, the velocity \mathbf{w} has only to be known for the interfaces.

Equations (3.20) and (3.21) represent the general formulation of a finite volume method on a moving Voronoi mesh, and from this point it is straight-forward to formulate a specific FV realization. In what follows, we summarize the key features of the second-order FV method adopted in the AREPO code. A detailed description of the scheme can be found in Springel [2010], and recent improvements in Pakmor et al. [2016].

In order to achieve a method with a second order accuracy in space, the solution inside every Voronoi cell has to be reconstructed linearly. Let ϕ_i be a mean quantity inside Voronoi cell i , and ϕ_j cell average values of the neighbouring cells. For obtaining the gradient $\langle \nabla \phi \rangle_i =: \boldsymbol{\sigma}_i$ of the reconstruction, one can make the ansatz

$$\phi_j = \phi_i + \boldsymbol{\sigma}_i(\mathbf{s}_j - \mathbf{s}_i), \quad (3.22)$$

where \mathbf{s} is the centre of mass of a cell. This equation expresses that the value of the reconstructed solution for cell i should be equal to the mean value ϕ_j , at the centre of neighbour j . If the number of neighbours is larger than 3, which is generally the case for a 3D Voronoi cell, this approach results in an overdetermined problem. Therefore, the optimal gradient for cell i in the sense of equation (3.22) has to be computed by means of a least square optimization, viz.

$$\sum_{\text{neighb. } j} g_j [\phi_j - \phi_i - \boldsymbol{\sigma}_i(\mathbf{s}_j - \mathbf{s}_i)]^2 \stackrel{!}{=} \min, \quad (3.23)$$

where $g_j = A_{ij}/|\mathbf{s}_j - \mathbf{s}_i|^2$ represents an appropriate scalar weight for neighbour j . The result in form of the gradient we are looking for can be obtained straight-forwardly with a least square fit.

In order to stabilize the finite volume scheme at hydrodynamic discontinuities, the gradient is slope limited,

$$\boldsymbol{\sigma}'_i = \alpha_i \boldsymbol{\sigma}_i, \quad (3.24)$$

with $\alpha_i \in [0, 1]$. The amount of limiting depends on all neighbouring cells j and is computed via

$$\alpha_i = \min_{\text{neighb. } j} (1, \psi_{ij}), \quad (3.25)$$

where the scalar ψ_{ij} is given by

$$\psi_{ij} = \begin{cases} (\phi_i^{\max} - \phi_i)/\Delta\phi_{ij} & \text{for } \phi_{ij} > 0 \\ (\phi_i^{\min} - \phi_i)/\Delta\phi_{ij} & \text{for } \phi_{ij} < 0 \\ 1 & \text{for } \Delta\phi_{ij} = 0. \end{cases} \quad (3.26)$$

ϕ_i^{\max} and ϕ_i^{\min} are the maximum and minimum mean values among the neighbouring cells, and $\Delta\phi_{ij} = \boldsymbol{\sigma}_i \cdot (\mathbf{f}_{ij} - \mathbf{s}_i)$ is the change of the quantity ϕ between the centre \mathbf{s}_i of the cell and the centre \mathbf{f}_{ij} of the interface. Since overshoots and undershoots are only avoided with respect to the maximum and minimum mean values among all neighbours, this limiter is not strictly TVD and allows for some oscillations. On the other hand, the overall accuracy can be higher in many applications, since it is less dissipative in smooth regions.

After reconstructing the hydrodynamic quantities, the numerical solution is evolved by computing the fluxes (3.21) numerically. For the moving Voronoi mesh this can be achieved by solving a one-dimensional Riemann problem at the centre of every interface, with the additional complication of non-zero interface velocities \mathbf{w} . In order to retain Galilean invariance, the Riemann problem at an interface is transformed to the rest frame of the interface and solved with a standard Riemann solver. The resulting flux is transformed back to the laboratory frame of reference, and applied to the mean conserved variables of the adjacent cells, which represents the averaging step of the FV method. For the transformation the velocity of the interface at its centre, where the Riemann problem is solved, has to be computed. This can be accomplished given the velocities and positions of the mesh-generating points adjacent to the interface, and the calculation may be found in section 3.3 of Springel [2010].

Before the average step is executed, the mesh-generating points are advected with the local fluid velocity, representing the movement of the Voronoi mesh. Additionally, a small correction velocity pointing towards the centre of mass of the cell is taken into account. Generally, the position of a mesh-generating point is close to the centre of mass of the Voronoi cell, but their locations do not coincide. With the correction the mesh can be regularized, and cells are kept roundish.

After the mesh-generating points have been moved, the Voronoi mesh has to be recomputed. This is accomplished in AREPO by first constructing the Delaunay tessellation of the mesh-generating points. In a second step, the Voronoi tessellation can be straight-forwardly obtained as the dual graph¹. A 2D Delaunay tessellation is a triangulation such that for every triangle the circumcircle contains no other mesh-generating point. In the three-dimensional case, the Delaunay tessellation consists of tetrahedra, and every circumsphere does not contain any point other than the four points defining its tetrahedron.

The Delaunay tessellation is constructed in AREPO by an incremental insertion algorithm with $N \log(N)$ complexity, where N is the number of points. In this approach, one mesh-generating point after the other is inserted into a valid Delaunay tessellation. A new point is connected with the corners of the element containing it, and if the Delaunay property is locally not fulfilled anymore, interfaces are flipped until the property is restored. The computational time spent for constructing Voronoi tessellations during a hydrodynamical simulation is about 40%. This value decreases significantly in a cosmological simulation, for which the calculation of gravity is dominating. Moreover, other Lagrangian schemes such as SPH require a particle neighbour search in every timestep, which has the same complexity of $N \log(N)$. Hence, considering the advantages of the moving-mesh approach, the additional computational cost for constructing the Voronoi mesh can be well tolerated.

¹In the Voronoi diagram every interface corresponds to a connection of the Delaunay tessellation, and the Voronoi cell corners are located at the centres of the circumcircles/circumspheres of the Delaunay tessellation.

For the time integration of equation (3.20), Heun’s explicit Runge-Kutta method (3.55) is used, which provides second order accuracy. In contrast to other second order RK methods, Heun’s method requires evaluations of the right-hand side of the differential equation only at the beginning and at the end of a timestep. In this way, the Voronoi mesh has to be constructed only once per timestep. In order to render local timestepping possible, Heun’s method is used in a slightly modified version. The intermediate step of the modified method is obtained by a time extrapolation using the gradients and the Euler equations. Technically, this method is a hybrid between Heun’s method and the MUSCL-Hancock scheme.

After the publication of the AREPO code paper [Springel, 2010], a couple of other moving Voronoi mesh codes were implemented [Duffell and MacFadyen, 2011; Schaal, 2013; Yalinewich et al., 2015], although only for 2D hydrodynamics. Due to the mathematical properties and physical advantages of this approach, we suspect that the popularity of moving-mesh codes will increase significantly in the near future.

3.2 Discontinuous Galerkin method

The discontinuous Galerkin method belongs to the class of finite element methods, which are used for solving a vast range of partial differential equations. In these methods a mesh is generated by subdividing space into individual elements, and basis functions defined on a reference element are used for representing the solution. The error of the numerical solution is minimized by variational methods, resulting in equations for every element. The equations are coupled across element interfaces and form an overall system of equations which has to be solved numerically.

In DG, the equation of interest is solved in a weak formulation with respect to the basis functions. As opposed to the solution in a continuous Galerkin method, the solution in DG is generally discontinuous across element interfaces. DG was introduced by Reed and Hill [1973], and later generalized to non-linear problems in [Cockburn et al., 1989, 1990; Cockburn and Shu, 1989, 1991, 1998]. In this section we outline the basic mathematical concept of DG, and argue that this concept offers a high potential for applications in computational hydrodynamics.

3.2.1 The discontinuous Galerkin scheme

For the sake of simplicity we introduce the DG scheme for a scalar conservation law,

$$\frac{\partial u}{\partial t} + \nabla \mathbf{f}(u) = 0, \quad (3.27)$$

where \mathbf{f} is the flux and $u = u(\mathbf{x}, t)$ the scalar function we are looking for. The basic idea of DG is to represent the numerical solution within every cell as a linear combination of a finite set of orthogonal and normalized basis functions $\{\phi_j(\mathbf{x})^K \mid j = 1, \dots, N\}$. A common choice for the latter are Lagrange polynomials or Legendre polynomials. For the DG scheme implemented in this thesis we use

Legendre polynomials, which are introduced in Section 3.2.2. The numerical solution inside cell K is given by

$$u(\mathbf{x}, t)^K = w_1(t)^K \phi_1(\mathbf{x})^K + w_2(t)^K \phi_2(\mathbf{x})^K + \dots + w_N(t)^K \phi_N(\mathbf{x})^K. \quad (3.28)$$

In this linear combination the weights only depend on time, whereas the basis functions only depend on the spatial coordinate. In the following we drop the space and time dependencies as well as the cell index K in our notation. Due to the orthonormality of the basis functions, the weight w_j can be obtained by the projection

$$w_j = \frac{1}{|K|} \int_K u \phi_j \, dV, \quad (3.29)$$

where $|K|$ denotes the volume of the cell. Given initial conditions $u(\mathbf{x}, t = 0)$, this equation can be used for calculating initial weights of the DG scheme.

Next, we derive an evolution equation for the weights of a cell. To this end, we multiply the partial differential equation (3.27) with the basis function ϕ_j and integrate over the cell volume,

$$\int_K \left[\frac{\partial}{\partial t} u + \nabla \mathbf{f}(u) \right] \phi_j \, dV = 0. \quad (3.30)$$

By using the product rule of the nabla operator we can write

$$\int_K \frac{\partial u}{\partial t} \phi_j \, dV + \int_K \nabla [\mathbf{f}(u) \phi_j] \, dV - \int_K \mathbf{f}(u) \nabla \phi_j \, dV = 0. \quad (3.31)$$

In this equation the second term can be transformed with Gauss' theorem into a surface integral,

$$\int_K \nabla [\mathbf{f}(u) \phi_j] \, dV = \oint_{\partial K} \phi_j \mathbf{f}(u) \mathbf{n} \, dS, \quad (3.32)$$

where \mathbf{n} is the surface normal vector. Moreover, in the first term of (3.31) the basis function ϕ_j depends only on space and the partial time derivative can be pulled in front of the integral and written as a total derivative. With (3.29) we finally obtain the evolution equation for the weights, which reads

$$|K| \frac{d}{dt} w_j + \oint_{\partial K} \phi_j \mathbf{f}(u) \mathbf{n} \, dS - \int_K \mathbf{f}(u) \nabla \phi_j \, dV = 0. \quad (3.33)$$

For every cell K , this represents a system of ordinary differential equations,

$$\frac{dw_j^K}{dt} + R_K = 0, \quad j = 1, \dots, N. \quad (3.34)$$

The system is coupled, since the function R_K depends on the weights of cell K , as well as on weights of neighbouring cells. (3.34) can be readily integrated with an explicit Runge-Kutta (RK) scheme, which is introduced in Section 3.3.

The integrals of R_K can be computed numerically, e.g. by means of Gaussian quadrature. This integration technique is explained in Section 3.2.3. With this approach, the integrands have to be evaluated only at discrete points, which are called quadrature points. For the volume integral this can be accomplished straightforwardly, since the flux f as well as the derivatives of the basis functions can be calculated analytically. However, the situation proves to be more difficult for the surface integral. At the boundary between two cells the global solution is generally discontinuous, and the flux has to depend on both, the solution left and the solution right to the discontinuity. In this sense, one has to calculate an ‘average’ flux at the quadrature points of the interfaces. In the case of the Euler equations, the exact flux can be computed by solving the Riemann problem as outlined in Section 2.3.1.

The presented explicit RK DG scheme can be implemented for an arbitrary order of accuracy in space and time. The spatial order of the scheme simply depends on the choice of basis functions; if polynomials with degree k are adopted, the Gaussian quadrature rule with $k + 1$ quadrature points per dimension should be used, resulting in a numerical scheme of the order $p = k + 1$. For achieving a suitable time integration accuracy, one can simply use an RK method of the same order. We conclude that in the DG framework a higher order scheme can be obtained straight-forwardly. This is one of the main advantages of the method.

A higher order scheme is favourable in regions where the solution is smooth, since it can be represented more accurately with the same number of degrees of freedom (DOF). This effect is demonstrated in Fig. 3.2, which shows piecewise approximations of the sine function with different approximation orders. In all approximations the degrees of freedom for achieving an optimal representation of the sine function are equal. For every cell we compute the approximating polynomial by minimizing the volume integral over the squared difference to the target function (L^2 -projection).

As can be seen from the figure, a better result is obtained for higher order solution representations. Moreover, the result improves significantly when second order polynomials are used instead of linear functions, whereas the improvements are minimal for polynomials higher than second order. In general, the optimal order of a numerical scheme is strongly problem-dependent, and for an optimal choice one has to consider the computing time to solution, the targeted accuracy, as well as memory requirements.

The discontinuous nature of DG gives rise to a couple of advantages. Most importantly, for numerically evolving the weights of a cell only information from neighbouring cells is required. Unlike in a finite volume method, the stencil does therefore not widen with increasing order. This feature is of great value for high performance computing applications, since less time is spent on communication and hence more time can be devoted to calculating the solution.

Moreover, for hydrodynamic applications a discontinuous solution representation can potentially capture shocks very well. At locations of these hydrodynamic discon-

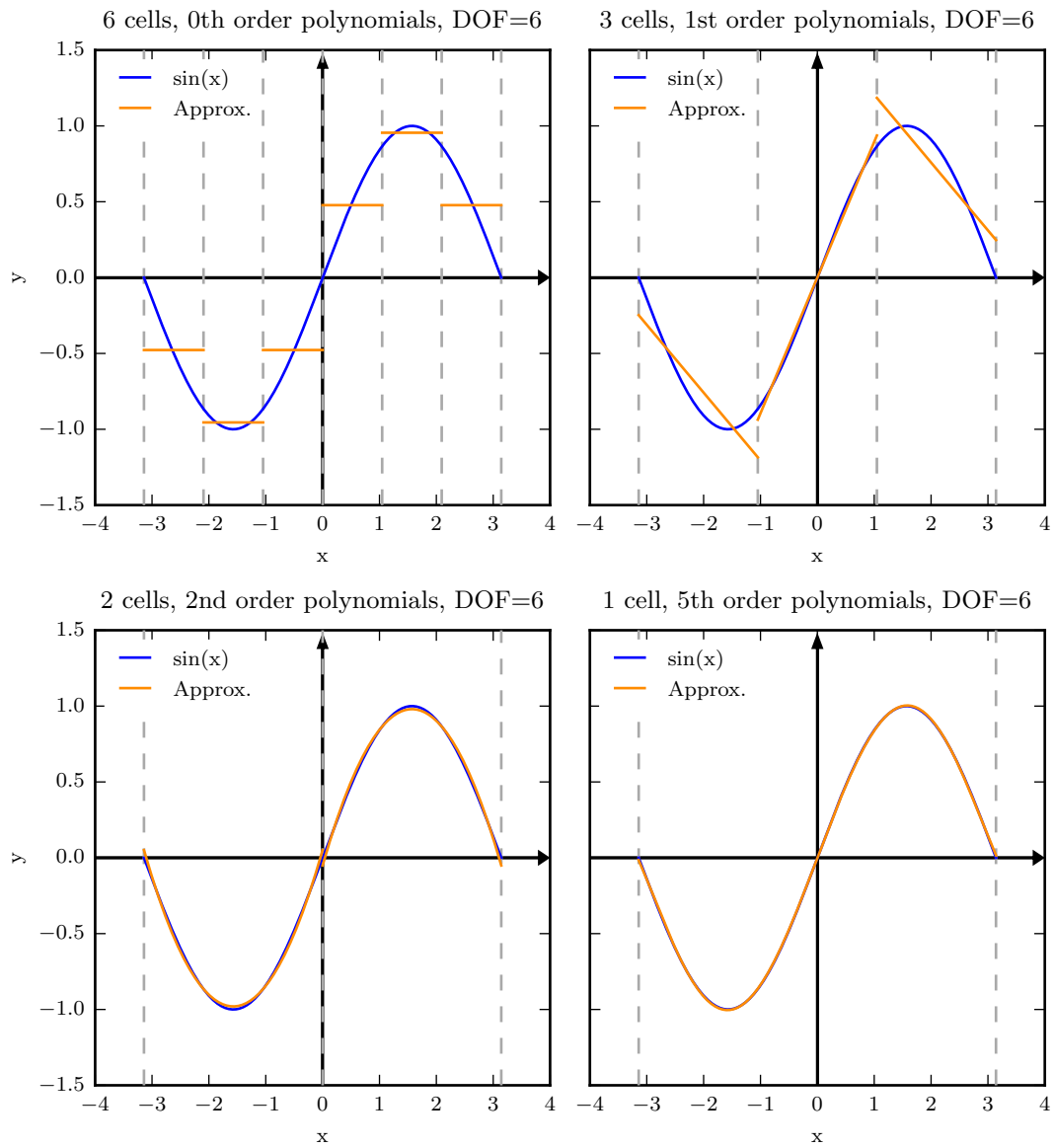


Figure 3.2: Piecewise approximations of the sine function with polynomials of different order. The number of used cells as well as the order of the adopted polynomials are indicated above each panel. While the number of degrees of freedom is the same for each approximation, it becomes more accurate if higher order functions are used. This illustrates the potential of higher order methods for problems with smooth solutions.

tinuities the higher order terms of the solution can be limited such that an accurate representation is achieved.

In Part IV we give an introduction to DG in an astrophysical context, and derive the scheme for solving the Euler equations on a structured adaptive mesh. By self-consistently solving for the higher order moments of the solution, DG offers several advantages compared to FV methods, which will be demonstrated in hydrodynamic test problems.

3.2.2 Legendre polynomials

For the DG scheme presented in this thesis we use Legendre polynomials as basis functions. They are solutions of Legendre's differential equation, which reads

$$\frac{d}{d\xi} \left[(1 - \xi^2) \frac{d}{d\xi} P_n(\xi) \right] + n(n+1)P_n(\xi) = 0, \quad n \in \mathbb{N}_0. \quad (3.35)$$

Given the first two polynomials $P_0(\xi) = 1$ and $P_1(\xi) = \xi$, the higher order polynomials can be computed with the recursion formula

$$P_n(\xi) = \frac{1}{n} [(2n-1)\xi P_{n-1}(\xi) - (n-1)P_{n-2}(\xi)]. \quad (3.36)$$

The DG scheme also requires knowledge of the derivative of the basis functions. In the case of Legendre polynomials these functions can be obtained straight-forwardly by differentiating (3.36), yielding

$$\frac{dP_n(\xi)}{d\xi} = \frac{1}{n} \left[(2n-1)(P_{n-1}(\xi) + \xi \frac{d}{d\xi} P_{n-1}(\xi)) - (n-1) \frac{d}{d\xi} P_{n-2}(\xi) \right]. \quad (3.37)$$

We display the first Legendre polynomials in Fig. 3.3, which are

$$\begin{aligned} P_0(\xi) &= 1 & P_1(\xi) &= \xi \\ P_2(\xi) &= \frac{1}{2}(3\xi^2 - 1) & P_3(\xi) &= \frac{1}{2}(5\xi^3 - 3\xi) \\ P_4(\xi) &= \frac{1}{8}(35\xi^4 - 30\xi^2 + 3) & P_5(\xi) &= \frac{1}{8}(63\xi^5 - 70\xi^3 + 15\xi). \end{aligned} \quad (3.38)$$

They are pairwise orthogonal on the interval $[-1,1]$,

$$\int_{-1}^1 P_i(\xi) P_j(\xi) d\xi = \begin{cases} 0 & \text{if } i \neq j \\ \frac{2}{2i+1} & \text{if } i = j. \end{cases} \quad (3.39)$$

and have the boundary values

$$P_n(1) = 1, \quad \text{and} \quad P_n(-1) = (-1)^n. \quad (3.40)$$

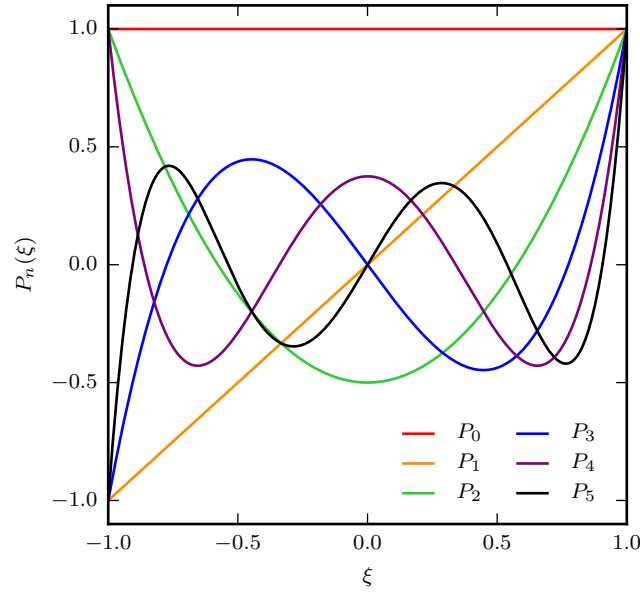


Figure 3.3: Legendre polynomials.

For practical reasons we use scaled Legendre polynomials as basis functions,

$$\tilde{P}_n(\xi) = \sqrt{2n+1}P_n(\xi), \quad (3.41)$$

such that

$$\int_{-1}^1 \tilde{P}_i(\xi)\tilde{P}_j(\xi) d\xi = \begin{cases} 0 & \text{if } i \neq j \\ 2 & \text{if } i = j. \end{cases} \quad (3.42)$$

If basis functions with a maximum polynomial degree of k are adopted, the resulting numerical scheme has a spatial order of $p = k + 1$. For example, for obtaining a third order scheme for a one-dimensional problem we can use the set $\{\tilde{P}_0(\xi), \tilde{P}_1(\xi), \tilde{P}_2(\xi)\}$. In our DG scheme the set of basis functions is the same for every cell, and the DG integrals for cell K can be solved conveniently by transforming them to the reference interval $[-1, 1]$.

For three-dimensional problems the basis functions are defined on $[-1, 1]^3$, and can be obtained as products of the one-dimensional scaled Legendre polynomials. In this case the basis functions for a third order scheme are $\{\tilde{P}_0, \tilde{P}_1(\xi_1), \tilde{P}_1(\xi_2), \tilde{P}_1(\xi_3), \tilde{P}_2(\xi_1), \tilde{P}_2(\xi_2), \tilde{P}_2(\xi_3), \tilde{P}_1(\xi_1)\tilde{P}_1(\xi_2), \tilde{P}_1(\xi_1)\tilde{P}_1(\xi_3), \tilde{P}_1(\xi_2)\tilde{P}_1(\xi_3)\}$.

3.2.3 Gaussian quadrature

The integrals of the DG equation (3.33) are solved by means of numerical integration, usually by a Gaussian quadrature rule. The numerical integration of a function

$f : [-1, +1] \rightarrow \mathbb{R}$ with the Gaussian quadrature rule of n points is given by the weighted sum

$$\int_{-1}^{+1} f(\xi) d\xi \approx \sum_{q=1}^n f(\xi_q^{1D}) \omega_q^{1D}. \quad (3.43)$$

Here, $\xi_q^{1D} \in (-1, +1)$ are the Gaussian quadrature nodes and ω_q^{1D} are the corresponding weights. The rule is exact for polynomials of degree up to $2n - 1$. The integration rule for multiple dimensions can be constructed by using tensor products of the n Gauss points, e.g. the integration of a 2D function $f : [-1, +1]^2 \rightarrow \mathbb{R}$ is given by

$$\begin{aligned} & \int_{-1}^{+1} \int_{-1}^{+1} f(\xi_1, \xi_2) d\xi_1 d\xi_2 \\ & \approx \sum_{q=1}^n \sum_{r=1}^n f(\xi_{1,q}^{1D}, \xi_{2,r}^{1D}) \omega_q^{1D} \omega_r^{1D} = \sum_{q=1}^{n^2} f(\xi_q^{2D}) \omega_q^{2D}. \end{aligned} \quad (3.44)$$

In our DG scheme we use the Gauss-Legendre quadrature rule, and the one-dimensional nodes are given as the roots of the Legendre polynomial $P_n(\xi)$. We calculate them numerically by means of the Newton-Raphson method. As starting values of the iterative root finding, approximate expressions for the roots can be used [see for example Lether and Wenston, 1995],

$$\xi_q \approx \left(1 - \frac{1}{8n^2} + \frac{1}{8n^3} \right) \cos \left(\pi \frac{4q-1}{4n+2} \right), \quad q = 1, \dots, n. \quad (3.45)$$

Furthermore, the corresponding weights can be calculated as [Abramowitz and Stegun, 2012]

$$\omega_q = \frac{2}{(1 - \xi_q^2) P_n'(\xi_q)^2}, \quad q = 1, \dots, n. \quad (3.46)$$

With this approach, we can compute and store the necessary quadrature data in the initialization routine of our DG code for arbitrary spatial order.

3.2.4 Lobatto quadrature

As will be outlined later on in the detailed description of our DG scheme, an additional quadrature rule is needed, which includes also the endpoints of the integration interval². These rules are known as Lobatto quadrature, and we denote their nodes and weights by $\hat{\xi}_q$ and $\hat{\omega}_q$, respectively. They are very similar to Gaussian quadrature rules, however, due to the additional restriction of including the integration boundaries, they are exact for polynomials of degree $2n - 3$, instead of $2n - 1$.

²This type of quadrature rule is needed for constructing the positivity limiter, see Section 19.4.

For our purpose, we use the Gauss-Lobatto-Legendre (GLL) quadrature rule, where the nodes are the roots $\hat{\xi}_q$ of the function $(1 - \xi^2)P'_{n-1}(\xi)$, and the corresponding weights are given by [Abramowitz and Stegun, 2012]

$$\hat{\omega}_q = \frac{2}{n(n-1)P_{n-1}(\hat{\xi}_q)^2}, \quad q = 2, \dots, n-1. \quad (3.47)$$

The weights of the endpoints are equal, $\hat{\omega}_1 = \hat{\omega}_n$, and the sum of the weights is $\sum_{q=1}^n \hat{\omega}_q = 2$.

3.3 Runge-Kutta time integration

Explicit Runge-Kutta methods are a popular class of methods for solving initial value problems consisting of an ordinary differential equation

$$\frac{dw}{dt} + R(t, w) = 0 \quad (3.48)$$

and the initial value

$$w(t=0) = w_0. \quad (3.49)$$

The right-hand side of the differential equation ($-R(t, w)$) corresponds to the slope of the solution $w(t)$ we are looking for. The most simple Runge-Kutta method is the explicit Euler scheme,

$$w^{n+1} = w^n - \Delta t^n R(t^n, w^n), \quad (3.50)$$

where w^{n+1} and w^n are the numerical solutions at times t^{n+1} and t^n , respectively, and $\Delta t^n = t^{n+1} - t^n$ is the timestep. The numerical solution is exact if the analytic solution $w(t)$ is a linear function, otherwise the local error within every integration step is of the order $(\Delta t)^2$,

$$w^{n+1} - w(t^{n+1}) \leq \mathcal{O}((\Delta t)^2). \quad (3.51)$$

The number of integration steps for the whole time range is proportional to $1/\Delta t$, such that the cumulative error is of first order. Hence, the explicit Euler scheme is a first order Runge-Kutta method.

A second order method can be constructed by improving the estimate of the slope across the timestep, e.g. by computing the slope in the middle between t^n and t^{n+1} , viz.

$$w^{n+1} = w^n - \Delta t^n R\left(t^n + \frac{1}{2}\Delta t^n, w^n - \frac{1}{2}\Delta t^n R(t^n, w^n)\right). \quad (3.52)$$

This is the so-called midpoint method, which represents a two-stage second order Runge-Kutta method. The two stages consist of calculating k_1 and k_2 , which are given by

$$k_1 = -R(t^n, w^n), \quad (3.53)$$

and

$$k_2 = -R\left(t^n + \frac{1}{2}\Delta t^n, w^n + \frac{1}{2}\Delta t^n k_1\right). \quad (3.54)$$

It can be shown that for constructing an explicit RK method of order s , at least s stages are required. If such a method has exactly s stages it is called ‘optimal’.

Another optimal second order Runge-Kutta scheme is Heun’s method, which reads

$$w^{n+1} = w^n - \Delta t^n \left(\frac{1}{2}R(t^n, w^n) + \frac{1}{2}R(t^n + \Delta t^n, w^n - \Delta t^n R(t^n, w^n)) \right). \quad (3.55)$$

In this case a slope is estimated by averaging the slopes at the beginning and at the end of the timestep; the latter is obtained by means of a forward Euler step.

The general form of an s -stage RK method is

$$w^{n+1} = w^n + \Delta t^n \sum_{i=1}^s b_i k_i, \quad (3.56)$$

where the factors b_i weight the sum over the solution derivatives k_i . These are evaluations of $R(t, w)$, viz.

$$k_i = -R(t^n + c_i \Delta t^n, w^n + \Delta t^n (a_{i1} k_1 + a_{i2} k_2 + \dots + a_{i,i-1} k_{i-1})), \quad (3.57)$$

with c_i being the nodes of the timestep interval, and a_{ij} the so-called RK matrix. The consistency of the scheme requires

$$\sum_{i=1}^s b_i = 1. \quad (3.58)$$

For constructing an RK method one can expand the unknown analytic solution $w(t)$ into a Taylor series up to a desired order, and compare it to equation (3.56). This approach results in condition equations for the coefficients $a_{ij}, b_i \in \mathbb{R}$, and $c_i \in [0, 1]$, which fully specify the RK scheme. The coefficients can then be calculated by hand or with the assistance of computer algebra, which is more convenient for high order methods.

The time integration in the AREPO code is accomplished by Heun’s method (3.55), although with some modifications for coping with local time steps as outlined in Pakmor et al. [2016]. For the moving-mesh code this method offers the advantage that the right-hand side of the ordinary differential equation is evaluated at the

beginning and end of a timestep, and not in between. In this case the mesh has to be constructed only once per timestep, which reduces the computational cost significantly.

Another benefit of Heun’s method is a property called ‘strong stability preserving’ (SSP). SSP RK methods are convex combinations of forward Euler steps, and they hence share several mathematical properties with the explicit Euler method. With an appropriate timestep the explicit Euler method is a total variation diminishing (TVD) time discretization. This means that in combination with a spatial TVD discretization, the discrete numerical solution is total variation diminishing. SSP RK methods are higher order time discretizations which maintain the TVD property, and are hence an important building block for a numerical scheme which can effectively suppress oscillations.

An SSP RK method is also a requirement for the positivity limiter we adopt in our DG scheme. This limiter ensures the positivity of density and pressure of the numerical solution, which is a crucial feature of a robust code. Hence, for our second order DG implementation we use Heun’s method, and for the third and fourth order version the SSP RK methods listed in the appendix (tables A3 and A4).

3.4 The Courant-Friedrichs-Lewy condition

A numerical scheme converges if it is stable and consistent. Consistency means that the local approximation of the solution is correct up to the targeted order, and stability indicates that the error made in an individual timestep does not grow too fast. A necessary condition for stability is the limitation of oscillations, as outlined in Section 3.1.2, another one is the Courant-Friedrichs-Lewy (CFL) timestep condition.

If we recall that in the evolve step of an FV method a time-independent flux is calculated for every interface, it is clear that the timestep size of the scheme has to be restricted in order to achieve stability. Every flux is computed by solving the Riemann problem across the corresponding interface, and the timestep has to be chosen small enough, such that the solutions of Riemann problems of opposing interfaces do not interact. The Riemann problem gives rise to different waves and the appropriate timestep can be expressed in terms of the CFL number, which is given by

$$\text{CFL} = \frac{S_{\max} \Delta t}{\Delta x}, \quad (3.59)$$

where S_{\max} denotes the maximum wave speed involved. Hence, the CFL number measures the maximum distance the physical information can travel within one timestep in units of Δx . For a linearized hyperbolic system the wave speeds correspond to the Eigenvalues $\lambda_i = \{v - c, v, v + c\}$ of the flux Jacobian, and we can write

$$\text{CFL} = \frac{(|v| + c) \Delta t}{\Delta x}. \quad (3.60)$$

Solving for the timestep results in

$$\Delta t = \text{CFL} \frac{\Delta x}{|v| + c}. \quad (3.61)$$

If the CFL number is set to $\text{CFL} = 0.5$, the maximum distance the fastest wave can travel is halfway through the cell, and an interaction between two approaching waves is effectively prevented. However, one can also choose a slightly larger value, since the fluxes in the finite volume method are independent of time over one timestep, and the fastest wave can possibly reach and influence the opposing interface only for $\text{CFL} = 1$. For explicit time integration schemes the value of the CFL number should in general be set to a value $\text{CFL} < 1$.

In the finite volume scheme on the moving Voronoi mesh, the timestep is larger. The bulk velocity v is negligibly small in the moving frame, and the timestep becomes

$$\Delta t = \text{CFL} \frac{R\Delta x}{c}, \quad (3.62)$$

where $R = \left(\frac{3V}{4\pi}\right)^{1/3}$ is the effective radius of the Voronoi cell and measures its size. For supersonic bulk velocities ($v > c$), the timestep on the moving-mesh is significantly larger compared to the one on the static grid, making up for a fraction of the additional computational cost of the mesh construction.

For the discontinuous Galerkin method, the timestep for achieving stability also depends on the order $p = k + 1$ of the scheme [see Cockburn and Shu, 1989, 1998, and Section 18.5], viz.

$$\Delta t = \frac{\text{CFL}}{2k + 1} \frac{\Delta x}{S_{\max}}. \quad (3.63)$$

Hence, with increasing order the timestep decreases, and more steps have to be computed for reaching the end of the simulation. Nevertheless, as we will show in Section 21.1, for smooth flows higher order DG schemes are significantly more computationally efficient.

The different timesteps discussed above are associated with a single cell, and a global simulation timestep can be determined as the minimum timestep of all cells. Moreover, in order to cope with multiscale problems, it is also possible to evolve cells with larger timesteps less frequently than cells with shorter timesteps. This approach of individual time steps can be achieved by sorting the cells according to their timesteps within a timestep hierarchy, and is adopted in the AREPO code. For details the interested reader may be referred to Section 7.2 of Springel [2010].

In this section we have motivated and qualitatively derived the CFL condition for the finite volume method. It is a necessary condition for the stability of a numerical scheme. Formally, it states that the numerical domain of dependence has to contain the physical domain of dependence of the underlying differential equation for $\Delta t \rightarrow 0$ and $\Delta x \rightarrow 0$. Hence, the timestep can be increased significantly if an implicit time

3.4 The Courant-Friedrichs-Lewy condition

integration scheme is adopted; in this case the numerical domain of dependence consists of the whole simulation volume.

Part II

Hydrodynamic Shocks in non-radiative Simulations

Abstract

Cosmological shock waves play an important role in hierarchical structure formation by dissipating and thermalizing kinetic energy of gas flows, thereby heating the Universe. Furthermore, identifying shocks in hydrodynamical simulations and measuring their Mach number accurately are critical for calculating the production of non-thermal particle components through diffusive shock acceleration. However, shocks are often significantly broadened in numerical simulations, making it challenging to implement an accurate shock finder. We here introduce a refined methodology for detecting shocks in the moving-mesh code AREPO, and show that results for shock statistics can be sensitive to implementation details. We put special emphasis on filtering against spurious shock detections due to tangential discontinuities and contacts. Both of them are omnipresent in cosmological simulations, for example in the form of shear-induced Kelvin-Helmholtz instabilities and cold fronts. As an initial application of our new implementation, we analyse shock statistics in non-radiative cosmological simulations of dark matter and baryons. We find that the bulk of energy dissipation at redshift zero occurs in shocks with Mach numbers around $\mathcal{M} \approx 2.7$. Furthermore, almost 40% of the thermalization is contributed by shocks in the warm hot intergalactic medium, whereas $\approx 60\%$ occurs in clusters, groups, and smaller haloes. Compared to previous studies, these findings revise the characterization of the most important shocks towards higher Mach numbers and lower density structures. Our results also suggest that regions with densities above and below $\delta_b = 100$ should be roughly equally important for the energetics of cosmic ray acceleration through large-scale structure shocks.

4

Introduction

The collapse of dark and baryonic matter during hierarchical large-scale structure formation releases gravitational energy and transforms it into kinetic energy. The bulk of the kinetic energy of the gas gets dissipated by cosmological shocks, heating the gas in virialized haloes (e.g the intracluster medium, ICM) as well as in the warm-hot intergalactic medium (WHIM). Cosmological hydrodynamic shocks are collisionless; they are established due to plasma interactions by means of magnetic fields. They can themselves amplify magnetic fields and accelerate particles via diffusive shock acceleration [DSA; Axford et al., 1977; Bell, 1978a,b; Blandford and Ostriker, 1978; Krymskii, 1977; Malkov and O’C Drury, 2001] up to relativistic energies, producing cosmic rays.

Directly observing cosmological shocks is challenging, especially outside cluster cores where the X-ray emission is weak. An obvious approach is to look for jumps in the thermal gas quantities. In this way, and using exquisite X-ray data from the *Chandra* telescope, the first merger shocks have been confirmed in the bullet cluster [$\mathcal{M} \approx 3$; Markevitch, 2006; Markevitch et al., 2002] and in the train-wreck cluster [$\mathcal{M} \approx 2.1$; Markevitch et al., 2005]. Maps of the gas density and the temperature can be inferred from the luminosity and the spectrum of the X-ray radiation, respectively. Both are necessary in order to calculate a pressure map and confirm a shock. Furthermore, it is possible to directly measure a pressure jump by means of the thermal Sunyaev–Zel’dovich signal [Sunyaev and Zeldovich, 1980]. For example, steep pressure gradients have been detected inside R_{500} in the nearby Coma cluster [Planck Collaboration X, 2013]. The location of the gradients coincides with temperature jumps, and two shocks with Mach numbers around $\mathcal{M} \approx 2$ were reported in this way.

Shocks can also be observed indirectly at radio wavelengths. Diffusively shock-accelerated cosmic ray electrons in merger and accretion shocks produce synchrotron

radiation, so-called radio gischt [Battaglia et al., 2009; Ensslin et al., 1998; Pinzke et al., 2013]. This phenomenon has been observed in several clusters [see e.g. Bonafede et al., 2009; Brüggen et al., 2012; Clarke and Ensslin, 2006; van Weeren et al., 2010, for a review of shocks in cluster outskirts and the associated features]. Another radio source triggered by shocks is the radio phoenix [Enßlin and Brüggen, 2002; Enßlin and Gopal-Krishna, 2001]. In this scenario, a shock overruns fossil radio plasma initially produced by an active galactic nucleus (AGN), compressing the plasma and reviving its radio emission. A radio phoenix can reveal large-scale accretion shocks and has been reported for the Perseus Cluster [Pfrommer and Jones, 2011].

The motivation for observing shocks is manifold. First of all, supersonic flows and their associated shocks allow the study of thermalization patterns and energetics of phenomena at a broad range of spatial scales. This includes accretion shocks onto clusters, mergers of galaxies and galaxy clusters, winds and jets of AGN, as well as stellar winds and supernovae. Secondly, observations of supernova remnants provide evidence for the creation of non-thermal cosmic ray particles at these locations. Cosmic ray protons can collide with thermal protons of the interstellar medium producing pions, which subsequently decay and release γ -radiation. The pion decay and hence the acceleration of cosmic ray protons has been confirmed for several supernova remnants [e.g. Fermi Collaboration, 2013; Giuliani et al., 2011].

DSA and associated processes such as the modification of the shock structure due to cosmic ray back-reaction or magnetic field amplification have been investigated analytically [e.g. Amato and Blasi, 2006; Blasi, 2002; Drury, 1983; Malkov, 1997], as well as with numerical simulations [e.g. Ellison et al., 1996; Ferrand et al., 2014; Kang and Jones, 2007a; Kang and Ryu, 2013; Vladimirov et al., 2006]. Furthermore, DSA can be simulated bottom-up by resolving the micro physics with particle-in-cell (PIC) methods [e.g. Amano and Hoshino, 2007, 2010; Riquelme and Spitkovsky, 2011]. Alternatively, less costly hybrid methods [e.g. Caprioli and Spitkovsky, 2014; Quest, 1988] can be used, where the ions are treated kinetically and the electrons are modelled as a fluid. While basic predictions of radio, X-ray and γ -ray emission of DSA models can be confirmed by observations of supernova remnants [e.g. Edmon et al., 2011; Reynolds, 2008], the detailed understanding of the non-linear acceleration mechanism requires further analytic and numerical work.

Additional insights will be provided by forthcoming observations with, for example, the Cherenkov Telescope Array [CTA; Actis et al., 2011] and the Square Kilometre Array (SKA). With the CTA it will be possible to study particle acceleration over larger energy ranges and with increased resolution compared to present observations. The SKA will presumably allow a detailed study of the magnetic field of galactic supernova remnants utilizing the effect of Faraday rotation. Furthermore, large-scale cosmological shocks are expected to be observed due to their synchrotron emission [Keshet et al., 2004].

Cosmological shocks in numerical simulations of large-scale structure formation were analysed comprehensively in previous studies, for example in Quilis et al. [1998], Miniati et al. [2000], Ryu et al. [2003], Pfrommer et al. [2006], Kang et al. [2007], Skillman et al. [2008], Vazza et al. [2009a], Planelles and Quilis [2013], and Hong

et al. [2014]. The detected shocks can be divided into two distinct classes, external and internal shocks [Ryu et al., 2003]. Strong external shocks form when previously cold and unshocked gas ($T \lesssim 10^4$) accretes from voids onto the cosmic web. They have typically high Mach numbers up to $\mathcal{M} \approx 100$, but dissipate comparatively little energy due to the low pre-shock density and temperature. Internal shocks on the other hand occur if previously shocked and thus hotter gas inside non-linear structures gets shock-heated further. Because of the smaller temperature ratios compared to external shocks, the Mach numbers of internal shocks are typically smaller ($\mathcal{M} \lesssim 10$). The pre-shock density and temperature of internal shocks are however high. This allows them to account for the bulk of the energy dissipation, especially shocks with Mach numbers in the range $2 \lesssim \mathcal{M} \lesssim 4$ contribute strongly.

A detailed characterization of the prevalence and strength of shocks in numerical simulations requires the implementation of an accurate shock finder. The first approaches in grid-based cosmological codes simply used the jump conditions on a cell-by-cell basis to identify shocked cells [Miniati et al., 2000; Quilis et al., 1998]. As a first improvement, Ryu et al. [2003] proposed a method in which the shock centres are identified in a two-step procedure. First, cells are considered to be in a shock zone if they simultaneously meet three different criteria meant to identify cells with some numerical shock dissipation. Within this zone, the shock centres were then determined by looking for the cells with the maximum compression. This more elaborate approach takes into account that the common numerical methods capture a shock discontinuity over a few cells, rather than exposing the full jump strength at a single cell interface.

In order to deal with three dimensional simulations, Ryu et al. [2003] calculated three different Mach numbers ($\mathcal{M}_x, \mathcal{M}_y, \mathcal{M}_z$) for each cell in the shock centre by evaluating the temperature jump across the shock zone in each coordinate direction. The maximum occurring Mach number was then assigned to the shock cell. A refinement to this method is to calculate the Mach number via $\mathcal{M} = (\mathcal{M}_x^2 + \mathcal{M}_y^2 + \mathcal{M}_z^2)^{1/2}$, thus minimizing projection effects [Vazza et al., 2009a]. Furthermore, Vazza et al. [2009a] showed that by using the velocity jump instead of the temperature jump slightly less scatter in the calculation was achieved with their code. The use of coordinate-splitting can be avoided by characterizing the direction of shock propagation with the local temperature gradient [Skillman et al., 2008]. In this way, a single Mach number can be calculated and the result of the shock finder becomes independent of the orientation between the shock and the underlying grid. The shock-finding implementation of Skillman et al. [2008] additionally filters tangential discontinuities and contacts by evaluating the pre- and post-shock temperature and density.

Quite different shock detection methodologies have been developed for Lagrangian smoothed-particle hydrodynamics (SPH) codes. To this end, Keshet et al. [2003] measured the entropy increase of each particle between different snapshots, and Pfrommer et al. [2006] measured the entropy injection rate on a per-particle basis during the simulation. As the entropy production is directly sourced by the artificial

viscosity used for shock capturing in SPH, this allows an estimate of the Mach number of a shock. In another SPH shock-finding method, Hoeft et al. [2008] proposed to use the local entropy gradient for determining associated pre- and post-shock regions, and then to calculate the Mach number across the associated jump.

In a recent code comparison project, Vazza et al. [2011] reported reasonable agreement of different codes with respect to energy dissipation and shock abundance as a function of Mach number. However, significant differences have also been detected. Especially the detailed comparison of grid-based shock finders with the SPH-based techniques revealed some apparent inconsistencies in the shock morphologies and in various features in the gas phase-space diagrams. These discrepancies in the results of the different shock finder implementations highlight the computational challenges involved in accurate numerical shock detection. As we will demonstrate in this work, a shock finder can be very sensitive to implementation details, and it is hence crucial to improve these methods further, for example by more carefully removing false positive shock detections associated with tangential and contact discontinuities.

This is the goal of this part of the thesis, which has the following structure. We describe and validate our new methodology for finding shocks in the moving-mesh code AREPO in Sections 5 and 6, respectively. The shock finder is then applied to non-radiative simulations in Section 7, and differences to previous studies are discussed in Section 8. Finally, we summarize our results in Section 9.

5

Methodology

5.1 The moving-mesh code AREPO

The non-radiative cosmological simulations analysed in this work and the development of the shock detection method were carried out using the AREPO code [Springel, 2010]. In this cosmological hydrodynamical code, the gas physics is calculated on a moving Voronoi mesh. The mesh generating points are advected with the local velocity of the fluid in order to achieve quasi-Lagrangian behaviour. For solving the Euler equations on the unstructured Voronoi grid, a finite volume method is used in the form of a second-order unsplit Godunov scheme with an exact Riemann solver. With this approach the accuracy of a grid code can be combined with features of Lagrangian codes such as Galilean invariance and approximately constant mass per resolution element. Gravity exerted by the gas and the dark matter is computed with a Tree-PM method [Springel, 2005; Xu, 1995] in which long-range gravitational forces are calculated with a particle-mesh scheme, whereas short-range interactions are calculated in real space using a hierarchical multipole expansion organized with an octree [Barnes and Hut, 1986].

5.2 The Rankine–Hugoniot jump conditions

It is well known that the mass, momentum, and energy flux are continuous across a discontinuity in an ideal gas. If the mass flux happens to be zero, it follows that the normal component of the velocity and the pressure do not jump across the discontinuity (‘tangential discontinuities’). If additionally the tangential velocity is also continuous, a special discontinuity is present which is called a contact. A non-zero mass flux on the other hand implies that the tangential velocities are

continuous. In this case, a shock is present and the normal velocities as well as the other thermodynamic variables jump according to [Landau and Lifshitz, 1966]

$$\frac{\rho_2}{\rho_1} = \frac{v_1}{v_2} = \frac{(\gamma + 1)\mathcal{M}^2}{(\gamma - 1)\mathcal{M}^2 + 2}, \quad (5.1)$$

$$\frac{p_2}{p_1} = \frac{2\gamma\mathcal{M}^2}{\gamma + 1} - \frac{\gamma - 1}{\gamma + 1}, \quad (5.2)$$

$$\frac{T_2}{T_1} = \frac{[2\gamma\mathcal{M}^2 - (\gamma - 1)][(\gamma - 1)\mathcal{M}^2 + 2]}{(\gamma + 1)^2\mathcal{M}^2}, \quad (5.3)$$

$$\frac{s_2}{s_1} = \left(\frac{2\gamma\mathcal{M}^2}{\gamma + 1} - \frac{\gamma - 1}{\gamma + 1} \right) \left(\frac{(\gamma - 1)\mathcal{M}^2 + 2}{(\gamma + 1)\mathcal{M}^2} \right)^\gamma. \quad (5.4)$$

The quantities ρ , v , p , T , and $s = p/\rho^\gamma$ denote density, velocity in the shock frame, pressure, temperature, and the entropic function, respectively. The Mach number \mathcal{M} is the shock speed in the frame of the pre-shock gas, in units of the pre-shock sound speed c_1 . The indices 1 and 2 label the pre-shock and post-shock regions, respectively, and γ is the adiabatic index of the gas.

5.3 Shock-finding method for AREPO

We base the implementation of our shock finder on a number of previous ideas [Hong et al., 2014; Ryu et al., 2003; Skillman et al., 2008], augmented with some improvements. First of all, a shock zone is identified by appropriate criteria which put special emphasis on filtering spurious shocks such as tangential discontinuities and contacts. We then tag cells with maximum compression along the shock direction and inside the shock zone as shock surface cells. The Mach number for these cells is calculated with the temperature jump across the shock zone. Finally, we take care of overlapping shock zones which can be present in the case of colliding shocks.

5.3.1 Shock direction

For our method, the direction of shock propagation in each Voronoi cell has to be specified. In order to be consistent with the Mach number calculation (see Section 5.3.4), we use the unlimited temperature gradient for calculating the shock direction:

$$\mathbf{d}_s = -\frac{\nabla T}{|\nabla T|}, \quad (5.5)$$

where ∇T is computed with the second-order accurate gradient operator available in AREPO for Voronoi meshes.

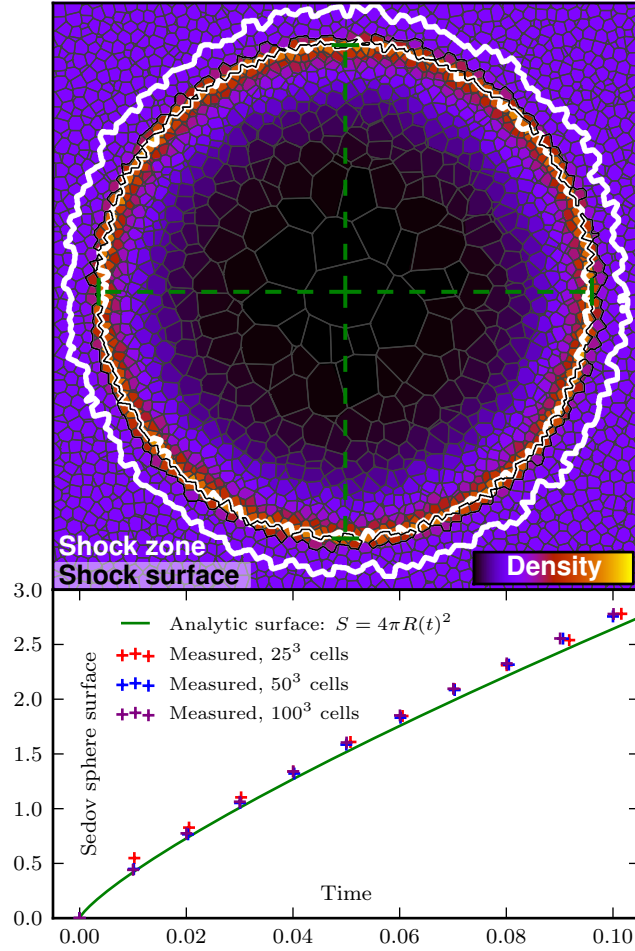


Figure 5.1: Top panel: cross-section of a three-dimensional Sedov blast wave simulation with 50^3 cells at $t = 0.08$ and energy $E = 1$. The colour map encodes the density field of the fluid, and the green dashed cross marks the analytically calculated extent of the Sedov blast wave at this time. The cells inside the white contours belong to the identified ‘shock zone’; they fulfil the three criteria described in Section 5.3.2. The black contours surround the cells that contain the reconstructed shock surface. These cells exhibit the minimum velocity divergence across the shock zone. Bottom panel: time evolution of the surface area of the spherical Sedov blast wave. We compare our approach to measure the shock surface in the test simulation (crosses) with the analytic evolution (solid line). For each cell in the shock surface, we assume an area contribution proportional to $V^{2/3}$ with a prefactor calibrated with shock tube simulations.

5.3.2 Shock zone

The first part of our algorithm consists of a loop over all cells. A cell is flagged as being in the *shock zone*, if the following criteria are met:

- (i) $\nabla \cdot \mathbf{v} < 0$,
- (ii) $\nabla T \cdot \nabla \rho > 0$,
- (iii) $\Delta \log T \geq \log \frac{T_2}{T_1} \Big|_{\mathcal{M}=\mathcal{M}_{\min}} \wedge \Delta \log p \geq \log \frac{p_2}{p_1} \Big|_{\mathcal{M}=\mathcal{M}_{\min}}$.

The first criterion is the standard compression criterion for shocks; whenever a shock is present, this condition is true. It also in principle filters tangential and contact discontinuities, however, this is not effective in real-world numerical simulations.

The second criterion is constructed such that spurious shock detections, potentially in a shear-flow or a cold front, are filtered out. Constant pressure implies that the density is inversely proportional to temperature and therefore these variables increase in opposite directions. At the same time, criterion (ii) holds in shocked cells.

The third criterion is a numerical guard against detecting spurious weak shocks. The first part of this protection mechanism introduces a lower boundary for the temperature jump, as in Ryu et al. [2003]. The second part demands a minimum pressure jump, which again discriminates against tangential discontinuities and contacts. Note that this part of criterion (iii) on its own may not be sufficient since gravitationally compressed cells are also able to fulfil it. $\Delta \log T$ and $\Delta \log p$ are calculated with the temperature and pressure of neighbouring cells along the shock direction. The logarithm is taken such that the calculation can be accomplished with a difference in order to avoid inaccurate divisions in low temperature and pressure regimes. In our analysis, the minimum Mach number is set to $\mathcal{M}_{\min} = 1.3$ as in Ryu et al. [2003]. We want to remark that the third criterion also rules out shocks with a slightly higher Mach number, since it is a local lower limit and the shock is broadened over a few cells. Note that we show in Section 6.2 that already $\mathcal{M} = 1.5$ shocks are fully captured.

In the following, we refer to the cells directly outside the shock zone in the direction of the positive temperature gradient as *post-shock region*, while the corresponding cells in the direction of the negative temperature gradient are referred to as *pre-shock region*.

5.3.3 Shock surface

After the determination of the shock zone, which has a typical thickness of 3–4 cells, we proceed with the construction of a shock surface consisting of a single layer of cells. For this purpose, rays are sent from each cell of the shock zone in the direction of the post-shock region (along the temperature gradient). When the first cell outside of the shock zone is reached, the post-shock temperature is recorded and the ray direction is reversed in order to find the pre-shock region. Furthermore, each ray

stores the velocity divergence of the cell from which it started. If a ray traverses a cell with a smaller divergence, the ray is discarded. For the rays reaching the pre-shock region, the Mach number is calculated via the temperature jump of equation (5.3) and assigned to the original cell of the ray. We call these cells with minimum velocity divergence (i.e. maximum compression) across the shock zone the *shock surface* cells. In this way, a Mach number is only calculated for cells in the shock surface. In the rare case that the direction of the temperature jump inferred from the pre- and post-shock temperatures is not consistent with the shock direction (given by the temperature gradient in the shocked cell), the detected feature is discarded.

In order to correctly treat overlapping shock zones of shocks propagating in opposite directions, we calculate in each step along a ray the scalar product of the shock direction of the original cell with the shock direction of the current cell. If the product is negative, the current temperature is recorded and the ray turns around or stops, depending on whether it was heading for the post- or pre-shock region, respectively. With this approach we ensure that even when the shock zones of two different shocks overlap we are usually able to distinguish them and calculate their correct Mach numbers.

For the sake of bookkeeping simplicity in the distributed memory parallelization of the algorithm, we send only one ray per shock zone cell combined with reverting its direction once, instead of simultaneously sending two separate rays in opposite directions. Since the shock surface is very close to the post-shock region (see Fig. 5.1), the maximum path a ray travels is only slightly larger than the thickness of the shock zone. Each ray starts at the centre of mass of a cell and thereafter propagates from cell interface to cell interface. The intersection between a ray and a Voronoi interface is calculated analytically. After all rays on the local MPI task are propagated for one cell, the rays leaving the local domain are communicated to the correct neighbouring task via a hypercube communication scheme.

5.3.4 Mach number calculation

Given the pre- and post-shock values, the Mach number can in principle be calculated with any of the equations (5.1)–(5.4). Note however that the Mach number calculation with the entropy jump has to be accomplished with a numerical root finder, for example a Newton–Raphson method. In Section 6.2, we investigate the quality of the practical results achieved with each of these Mach number determination methods and conclude that the temperature jump is best suited for the computation of the Mach number in AREPO, see also Fig. 6.1.

5.4 Energy dissipation

The thermal energy created at a shock can be expressed in terms of a generated thermal energy flux [Ryu et al., 2003]:

$$f_{\text{th}} = \left[\epsilon_2 - \epsilon_1 \left(\frac{\rho_2}{\rho_1} \right)^\gamma \right] v_2. \quad (5.6)$$

The indices 1 and 2 indicate the pre- and post-shock quantities, respectively, and ϵ denotes the thermal energy per unit volume. This flux can be expressed as a fraction of the incoming kinetic energy flux $f_\Phi = \frac{1}{2}\rho_1(c_1\mathcal{M})^3$:

$$f_{\text{th}} = \delta(\mathcal{M})f_\Phi. \quad (5.7)$$

The thermalization efficiency $\delta(\mathcal{M})$ can be calculated from the Rankine–Hugoniot jump conditions [Kang et al., 2007], yielding

$$\delta(\mathcal{M}) = \frac{2}{\gamma(\gamma-1)\mathcal{M}^2R} \left[\frac{2\gamma\mathcal{M}^2 - (\gamma-1)}{(\gamma+1)} - R^\gamma \right], \quad (5.8)$$

where R represents the density jump:

$$R \equiv \frac{\rho_2}{\rho_1} = \frac{(\gamma+1)\mathcal{M}^2}{(\gamma-1)\mathcal{M}^2 + 2}. \quad (5.9)$$

In our analysis, we use equation (5.7) for calculating the generated thermal energy flux of a cell in the shock surface. Given the area of the shock surface within a shocked cell (see Section 6.1), we are then also able to calculate the total dissipated energy per unit time.

Note that equations (5.8) and (5.9) describe the thermalization efficiency at shocks without considering cosmic rays. The possibility of cosmic ray acceleration and the corresponding efficiencies are addressed in Section 7.7.

6

Validation

6.1 Sedov–Taylor blast wave

We test the determination of the shock surface with simulations of three-dimensional point explosions. We performed runs with 25^3 , 50^3 , and 100^3 cells. In order to obtain an unstructured Voronoi mesh free of any preferred directions for the initial conditions, we distribute mesh-generating particles randomly in the unit box $(x, y, z) \in [0, 1]^3$. The mesh is then relaxed via Lloyd’s algorithm [Lloyd, 1982] such that a glass-like configuration is obtained. We then set up the initial conditions as follows: the whole box is filled with uniform gas of density $\rho_1 = 1$ and pressure $p = 10^{-4}$, the initial velocities are zero, and the adiabatic index is set to $\gamma = 5/3$. The energy $E = 1$ is injected into a single central cell of the grid.

We show a cross-section of the 50^3 simulation at $t = 0.08$ in the top panel of Fig. 5.1. At the corners of the box the initial glass-like grid is still visible. Note however that the cross-section of a three-dimensional Voronoi grid is in general no longer a Voronoi tessellation itself. The colour of the cells represents the density field of the fluid. The cells inside the white contours constitute the identified shock zone. The shock surface consists of the cells inside the black contour lines and features a position that agrees well with the expected position (extent of the green cross).

Determining the correct shock surface area is important for calculating the energy dissipation accurately. We describe in Section 6.2 how we measure this area from the shock surface cells. In the bottom panel of Fig. 5.1 we compare the time evolution of the measured surface area of the Sedov shock sphere with the analytic solution, which is given by $S(t) = 4\pi R^2(t)$ (green line), where $R(t) = \beta(Et^2/\rho_1)^{1/5}$ [Landau and Lifshitz, 1966]. The coefficient β can be calculated numerically. We obtain the value $\beta = 1.152$ for $\gamma = 5/3$ from the code provided in Kamm and Timmes [2007]. Our measurement tracks the expected scaling well but shows a small systematic

6 Validation

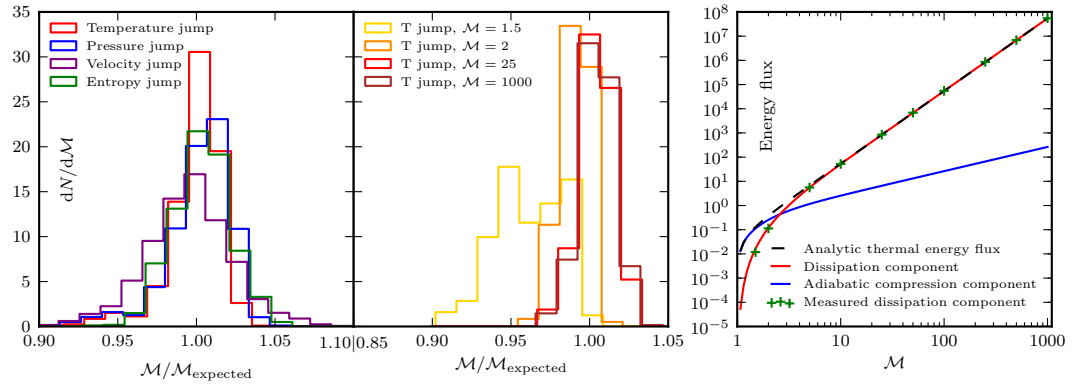


Figure 6.1: Left-hand panel: number of shock surface cells N per Mach number bin for ten different shock tube tests (see Table 6.1) for different calculation methods of the Mach number. We obtain the best results for the temperature jump of equation (5.3). Middle panel: Mach number distributions of single shock tubes obtained with the temperature jump method. Low Mach numbers such as $M \simeq 1.5$ are slightly underestimated due to mild post-shock oscillations. This effect vanishes for Mach numbers $M \geq 2$, where the correct value is found with an accuracy of 1 per cent. In both the left-hand and middle panels, each histogram is normalized such that the area under the curve is unity. Right-hand panel: thermal energy fluxes in shock tubes, separately for the adiabatic and dissipative components. We compare the measurement using the temperature jump with the analytic solution, finding excellent agreement.

overestimation of $\sim 5\%$. We suspect the primary cause of the offset does not lie in the shock surface estimation itself but rather appears because the simulated blast wave is slightly ahead of the analytic solution due to low resolution present at early times [Springel, 2010] in this self-similar problem.

6.2 Shock tubes

We also checked the accuracy of the Mach number estimate for the identified shock surface by performing numerous shock tube tests [Sod, 1978]. In view of our target applications, we chose to adopt a three-dimensional box $(x, y, z) \in [0, 100] \times [0, 20] \times [0, 20]$ in all the tests. Again, a hydrodynamic glass-like initial grid is used with 4×10^4 cells. The gas has an adiabatic index of $\gamma = 5/3$, and the initial position of the discontinuity is prepared at $x = 50$. The variables of the right state ($x > 50$) are set to $p_r = 0.1$, $\rho_r = 0.125$, and $v_r = 0$. The density and the velocity of the left state ($x < 50$) are $\rho_l = 1$ and $v_l = 0$, respectively. Furthermore, we assign a pressure p_l to the left state such that the shock has a specific Mach number, see Table 6.1. The third column of the table shows the simulation time at which the shock reaches $x = 75$. We apply our shock finder to the corresponding output file. Note that the shock finder in this test, in contrast to the Sedov-Taylor blast wave, is also confronted with rarefaction waves and contact discontinuities, which obviously should not be mistaken as shock features by the shock finder.

p_l	\mathcal{M}	t_{end}
0.81445190	1.5	14.43
1.9083018	2.0	10.83
15.357679	5.0	4.330
63.498622	10.0	2.165
400.51500	25.0	0.8660
1604.1492	50.0	0.4330
6418.6865	100.0	0.2165
40120.448	250.0	0.08660
160483.88	500.0	0.04330
641937.62	1000.0	0.02165

Table 6.1: Shock tube initial conditions. The pressure of the left state (p_l) is varied such that the shock has a specific Mach number. The right-hand column indicates the time t_{end} when the shock has traversed three quarters of the tube, at which point we measure its strength with our shock finder implementation.

The left-hand panel of Fig. 6.1 shows the quality of the Mach number determination for all considered Mach number calculation methods according to equations (5.1)–(5.4), except for the density jump method which is omitted because it is not sensitive for high Mach numbers because $\rho_2/\rho_1 \rightarrow (\gamma + 1)/(\gamma - 1)$ for $\mathcal{M} \rightarrow \infty$. We note that in order to apply the velocity jump method, the velocities have to be transformed into the lab frame [Vazza et al., 2009a].

The overall best results with AREPO for the shock tube tests are obtained with the temperature jump method according to equation (5.3). It performs very well for Mach numbers $\mathcal{M} \geq 2$, as can be seen in the middle panel of Fig. 6.1. For small Mach numbers ($\mathcal{M} < 2$), there are mild post-shock oscillations which cause the temperature jump method to underestimate the Mach number by a few percent. This systematic offset is present for all jump methods, unless the entropy jump is used, which is not perturbed by these adiabatic oscillations.

For calculating the energy dissipation, the correct shock surface area has to be determined. The area contribution S_i of a single cell to the whole shock surface is expected to scale with its volume according to $V_i^{2/3}$. Furthermore, S_i also depends on the shape of the cell. Cells in a shock are compressed normal to the shock direction and the degree of the compression depends on the strength of the shock. We therefore make the ansatz $S_i = \alpha F_i^\beta V_i^{2/3}$, where F_i is the maximum face angle of the Voronoi cell, which characterizes the shape of the cell. The definition of this quantity has been introduced in Vogelsberger et al. [2012] in the context of a mesh regularization switch. We calibrate the constants α and β with a least-square fit using the 10 shock tube problems described above, where the total area of the shock surface is expected to be equal to the cross-section of the tube ($S = 400$). Our calibration yields $\alpha = 1.074$ and $\beta = 0.4378$. By using these values, we obtain for the mean shock surface area of

the tubes $\langle S \rangle = 396.35 \pm 2.45$, which is accurate to within 1%. Note that in Fig. 5.1 we have demonstrated that also curved shock surfaces are measured to high accuracy.

With accurate Mach numbers combined with accurate shock surface areas in each shocked cell, we are able to calculate the dissipated energy on a cell-by-cell basis. We show this explicitly in the right-hand panel of Fig. 6.1. The shock finder measures the dissipative component of the thermal energy flux across a shock surface. The second component contributing to the total thermal energy flux is given by the adiabatic compression, which is present behind the shock. As can be seen from the figure, the adiabatic thermal energy flux is only relevant for small Mach numbers ($\mathcal{M} \leq 5$).

7

Shocks in non-radiative simulations

7.1 Simulation set-up

Besides full physics runs, the Illustris simulation suite [Genel et al., 2014; Vogelsberger et al., 2014b,c] contains also dark matter only as well as non-radiative runs. In this work we investigate shocks in the non-radiative runs, which include dark matter as well as gas, but no radiative cooling, star formation, and feedback. The cosmological parameters are consistent with the 9-year *Wilkinson Microwave Anisotropy Probe* (*WMAP9*) measurements [Hinshaw et al., 2013], and are given by $\Omega_m = 0.2726$, $\Omega_\Lambda = 0.7274$, $\Omega_b = 0.0456$, $\sigma_8 = 0.809$, $n_s = 0.963$, and $H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}$ with $h = 0.704$. Two simulations with a resolution of 2×455^3 and 2×910^3 were carried out in a periodic box having $75 h^{-1} \text{ Mpc}$ on a side, where the factor of 2 indicates that the same number of gas and dark matter elements were used. The adiabatic index of the gas is set to $\gamma = 5/3$. In the following, we refer to these runs as Illustris-NR-3 and Illustris-NR-2, where the latter is the one with the higher resolution. The corresponding dark matter mass resolutions are 4.008×10^8 and $5.010 \times 10^7 M_\odot$; the gas mass resolutions are kept fixed within a factor of 2 at 8.052×10^7 and $1.007 \times 10^7 M_\odot$ ('target gas mass') by the quasi-Lagrangian nature as well as the refinement and derefinement scheme of AREPO. For redshifts $z > 1$, the gravitational softening lengths of all mass components are equal and constant in comoving units, growing in physical units to 2840 and 1420 pc at $z = 1$ for the NR-3 and NR-2 run, respectively. For $z \leq 1$, the softening lengths of the baryonic mass components are kept fixed at these values.

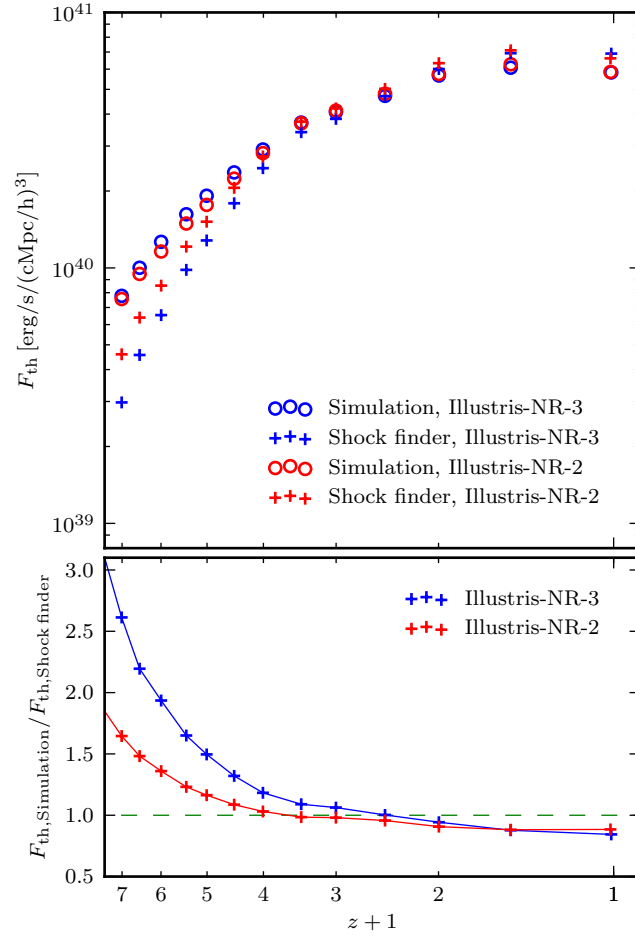


Figure 7.1: Comparison between the total energy dissipation rate found either with the shock finder, or inferred from two consecutive time steps of a non-radiative simulation. The absolute values are shown in the top panel, the ratio is given in the bottom panel. We find fairly good agreement for low redshifts, indicating that our shock finder properly accounts for all significant shock dissipation in the simulation. However, at high redshift not all the shock dissipation is recovered by the shock finder, an effect that diminishes greatly with better numerical resolution.

7.2 Shock finder assessment

First of all, we use the non-radiative runs for assessing the overall quality of our shock finder. To this end, we compare the total dissipated energy per unit time obtained with the shock finder with the dissipated energy measured between two consecutive time steps of the simulation. Under the assumption that the thermal energy during one time step changes only due to dissipation and adiabatic compression and expansion, we can write for every cell:

$$u_2 m_2 - u_1 m_1 = \Delta E_{\text{diss}} + \left[u_1 \left(\frac{\rho_2}{\rho_1} \right)^{\gamma-1} - u_1 \right] m_1. \quad (7.1)$$

Here u denotes the thermal energy per unit mass, m the mass, ρ the physical density, and the indices 1 and 2 correspond to quantities before and after the time step, respectively. We divide equation (7.1) by the time step size times the comoving box volume and solve for $F_{\text{th, Simulation}} = \Delta E_{\text{diss}} / (\Delta t V_c)$, the dissipated energy per time and volume.

The comparison with the shock finder measurement is shown in the top panel of Fig. 7.1. We find that our shock finder recovers the full amount of dissipated energy for low redshifts within 15% accuracy. For very high redshifts, a progressively larger deviation occurs and our shock detection results do not account for all the dissipated energy any more. The origin of this difference lies in the topology of early shocks, which are not yet pronounced and resolved well at high redshift. Instead, they are rather scattered and occupy a large fraction of the simulation volume. As can be seen in the bottom panel of Fig. 7.1, the deviation kicks in at progressively higher redshift for higher resolution simulations. We hence conclude that our shock finder statistics has an effective redshift completeness limit which depends on the resolution. We can trust the shock detection results from $z = 0$ up to $z \approx 4.0$ or $z \approx 5.0$ for the Illustris-NR-3 or Illustris-NR-2 runs, respectively.

7.3 Reionization modelling

The simulations Illustris-NR-3 and Illustris-NR-2 have no significant temperature floor and do not model cosmic reionization during their runtime. However, it is important to account for the nearly uniform heating of the ambient gas at the reionization redshift ($z \simeq 6 - 7$) in order to avoid overestimating the Mach numbers of shocks from voids onto filaments at late times. For this purpose, we use a temperature floor of 10^4 K for the shock finding carried out in post-processing, the same procedure as used by Ryu et al. [2003] and Skillman et al. [2008]. We can justify the simplicity of this approach by the marginal contribution made by shocks with low pre-shock temperature and density (voids onto filaments) to the dissipated energy, which is the main quantity of interest in our analysis. Nevertheless, reionization in post-processing could of course be modelled in a more sophisticated way, for example by a fitting function in the density–temperature plane [Vazza et al., 2009a].

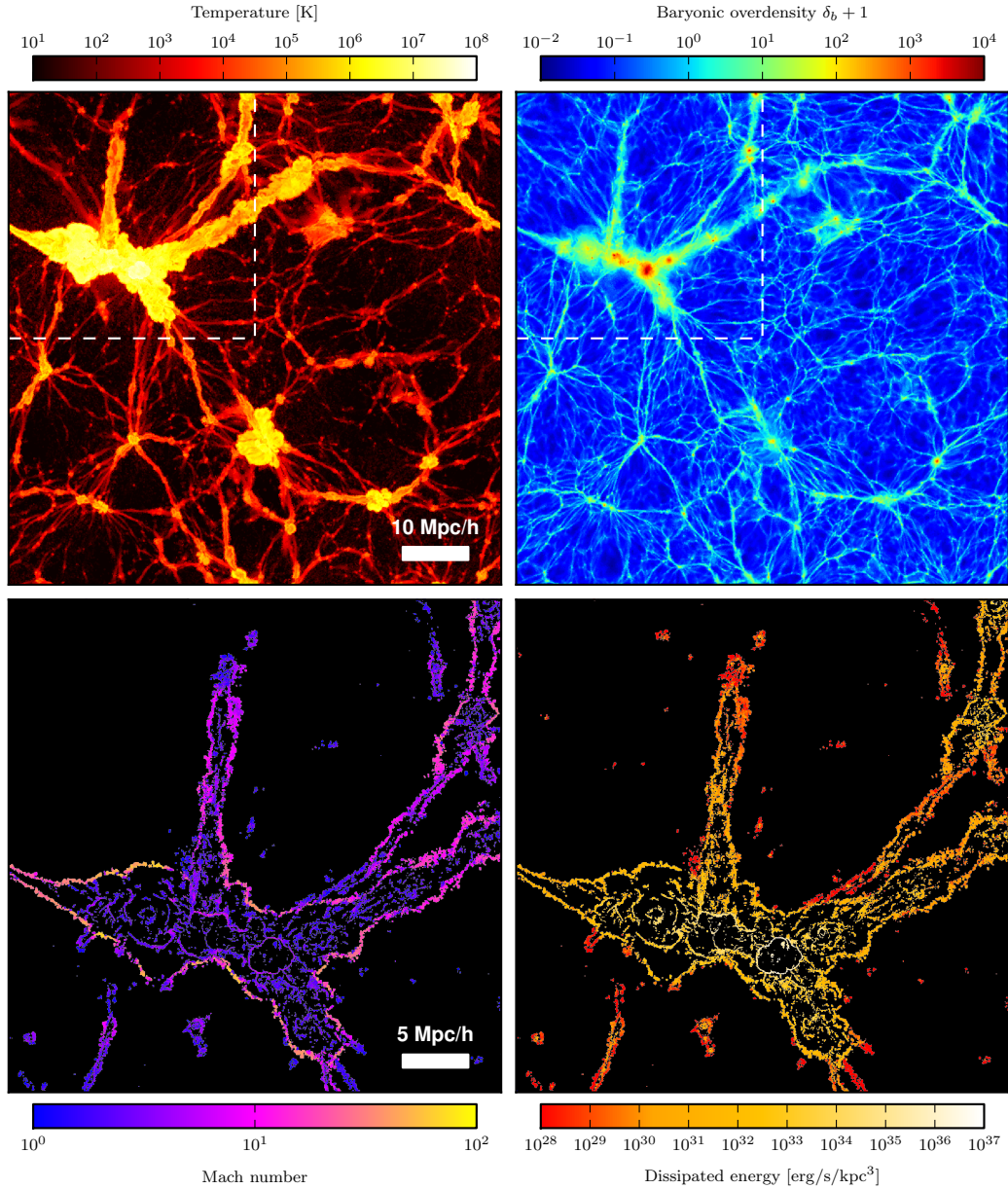


Figure 7.2: Top panels: projections of the mass-weighted temperature and mean baryonic overdensity of the Illustris-NR-2 run at redshift $z = 0$. The width and the height of the plots correspond to the full box size ($75 h^{-1} \text{Mpc}$). The projection in the z -direction has a depth of 150kpc and is centred onto the biggest halo in the simulation. Bottom panels: Mach number field weighted with the energy dissipation and mean energy dissipation rate density for the top left quarter of the box. Strong external shocks with Mach numbers up to $\mathcal{M} \sim 100$ onto the super-cluster are visible, as well as mostly weak shocks in the interior. However, most of the energy gets dissipated internally due to the higher pre-shock density and temperature. Note that we do not find many shocks inside the accretion shock onto the biggest halo, because here the gas motion is governed by subsonic turbulence, see also Fig. 7.5.

7.4 General properties

In Fig. 7.2 we present the state of the Illustris-NR-2 simulation at redshift $z = 0$. The projections were created by means of point sampling, and the shown quantities are the mass-weighted temperature, the mean baryonic overdensity, the Mach number weighted with the dissipated energy, and the mean dissipated energy density. The latter two are displayed only for the top left-hand quarter of the former projections, which show a supercluster including the biggest halo of the simulation. This halo has a mass of $M_{200, \text{cr}} = 3.2 \times 10^{14} M_{\odot}$, corresponding to a virial radius of $r_{200, \text{cr}} = 1.4 \text{ Mpc}$.

Note that due to the temperature floor applied in post-processing only the hottest filaments are detected by the shock finder and are hence present in the bottom panels. We can clearly observe the well-known fact [Ryu et al., 2003] that high Mach numbers ($\mathcal{M} \sim 10 - 100$) are generated at external shocks involving pristine pre-shock gas ($T_{\text{pre}} \lesssim 10^4 \text{ K}$), whereas gas previously processed by internal shocks ($T_{\text{pre}} > 10^4 \text{ K}$) experiences typically lower Mach numbers ($\mathcal{M} \lesssim 10$). The density inside the supercluster exceeds the density of the voids by several orders of magnitude and thus the energy dissipation is most effective internally. The highest dissipation rate in this projection is present in the accretion shock onto the biggest cluster, whereas we do not detect many shocks inside the accretion shock.

7.5 Shock statistics

Fig. 7.3 quantifies the shock distribution and the associated energy dissipation in the Illustris-NR-2 run. In the left-hand panel, the differential shock surface area normalized by the simulation volume is plotted as a function of the Mach number. We find redshift independently that 50% of the shocks have Mach numbers below $\mathcal{M} = 3$, and 75% below $\mathcal{M} = 6$. Towards lower redshift the cumulative area of shocks increases, especially for strong shocks. At redshift $z = 6$, shocks with a Mach number smaller than $\mathcal{M} = 12$ account for 99% of all the shocks, whereas at $z = 0$ all the shocks up to $\mathcal{M} = 35$ make up 99%. At low redshift, the accretion from previously unshocked gas onto hot filaments and cluster outskirts provides Mach numbers up to $\mathcal{M} \approx 100$. At redshift $z = 0$, the total shock surface area reaches a value of $S = 2.5 \times 10^{-1} \text{ Mpc}^2 / \text{Mpc}^3$ (integral of the blue curve).

The right-hand panel of Fig. 7.3 shows the differential thermal energy flux as a function of the Mach number. The total dissipated energy increases with time up to $z = 0.5$ and drops thereafter slightly to a value of $2.3 \times 10^{40} \text{ erg s}^{-1} \text{ Mpc}^{-3}$ (see also Fig. 7.1, but be aware of the factor h^3). The increase in time is expected due to an increasing number of shocks and the ever higher pre-shock densities and temperatures found inside structures. At low redshifts, this effect saturates and, furthermore, dark energy slows structure growth and dilutes the pre-shock gas inside voids, which leads to a drop of the thermal energy flux for high Mach numbers. The latter observation has also been pointed out by Skillman et al. [2008].

7 Shocks in non-radiative simulations

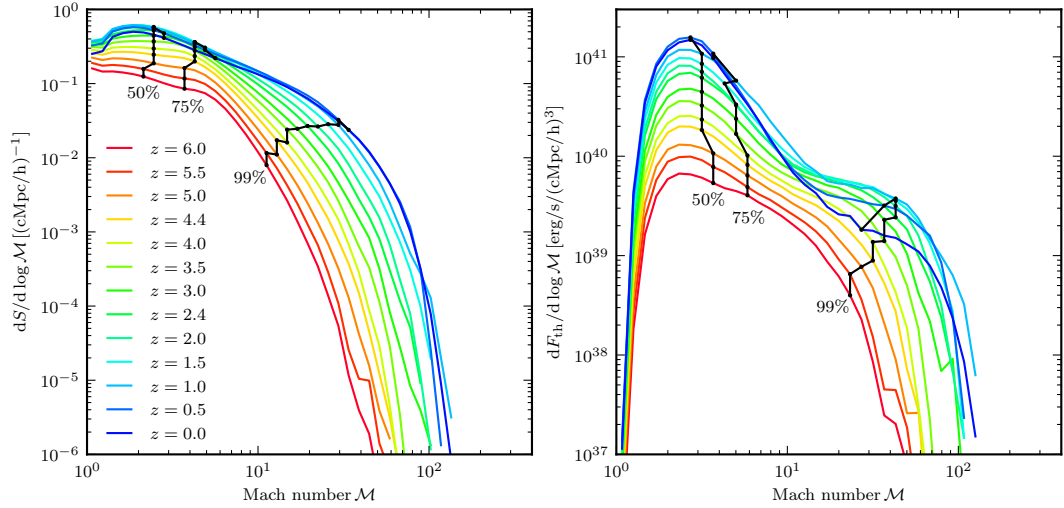


Figure 7.3: Left-hand panel: differential shock surface area as a function of Mach number for different redshifts. The black lines indicate the Mach numbers up to which a specific fraction of the total surface is included. Right-hand panel: distribution of the dissipated energy (generated thermal energy per time). We find that shocks with $\mathcal{M} < 6$ account for more than 75% of all the shocks as well as 75% of the total dissipated energy at all redshifts. Compared to former studies, the peak of our energy dissipation distribution at $z = 0$ is located at a considerable higher Mach number ($\mathcal{M} \approx 2.7$ instead of $\mathcal{M} \approx 2$).

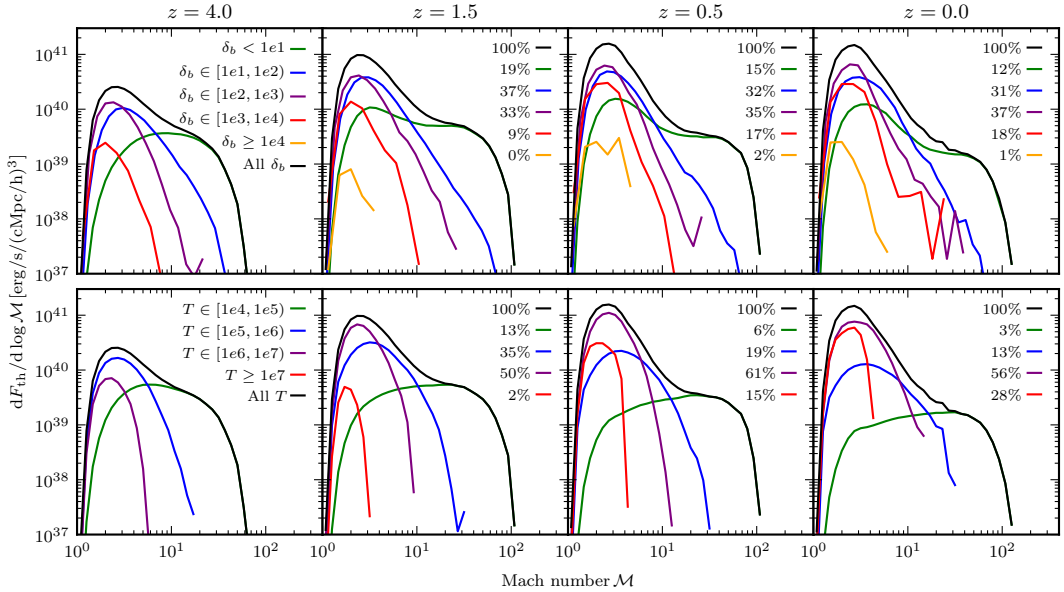


Figure 7.4: Contribution of different baryonic pre-shock overdensities (top panels) and pre-shock temperatures (bottom panels) to the overall energy dissipation in the Illustris-NR-2 run. We find that most of the energy dissipation at $z = 0$ is due to shocks with pre-shock overdensities in the range $\delta_b \in [10^1, 10^3]$. Furthermore, around 70% of the total dissipation is contributed by pre-shock gas with temperatures $T \in [10^5, 10^7]$, of which most is located in the WHIM.

	Temperature range [in K]			
	$[10^4, 10^5)$	$[10^5, 10^6)$	$[10^6, 10^7)$	$[10^7, \infty)$
$\delta_b < 10^1$	2%	4%	6%	0%
$\delta_b \in [10^1, 10^2)$	1%	6%	22%	3%
$\delta_b \in [10^2, 10^3)$	0%	2%	20%	15%
$\delta_b \in [10^3, 10^4)$	0%	1%	7%	10%
$\delta_b \geq 10^4$	0%	0%	1%	1%

Table 7.1: Contributions to the total energy dissipation for different pre-shock temperature and baryonic overdensity ranges at redshift $z = 0$. We find that $\approx 38\%$ of the energy gets dissipated by pre-shock gas of the WHIM ($T \in [10^5, 10^7)$, $\delta_b < 10^2$), and $\approx 57\%$ of the thermalization happens in clusters and groups ($\delta_b \geq 10^2$).

We find that 50% of the total energy dissipation occurs in shocks with $\mathcal{M} < 4$, and 75% in shocks with $\mathcal{M} < 6$. Mach numbers above $\mathcal{M} > 40$ do not contribute significantly to the dissipation. We find that the energy dissipation peaks for $z = 6$ at $\mathcal{M} \approx 2.3$ and shifts towards $\mathcal{M} \approx 2.7$ for $z = 0$. This peak position at redshift zero differs considerably from the value $\mathcal{M} \approx 2$ found by different previous studies [Pfrommer et al., 2006; Ryu et al., 2003; Skillman et al., 2008; Vazza et al., 2009a, 2011]. We believe that the origin of this discrepancy lies in the improved methodology we adopt, and we will elaborate more on this in Section 8.

Additional differences become apparent if we investigate the shock locations. In Fig. 7.4 we show the distribution of energy dissipation with respect to pre-shock densities (top panel) and pre-shock temperatures (bottom panel). Note however that this plot is not directly comparable to a similar investigation in Skillman et al. [2008] since their analysis is based on the inflowing kinetic energy. As expected, the relative contribution of shocks in the densest regions increases with time while shocks in low density gas become less important. At zero redshift, 68% of the energy dissipation is due to shocks with baryonic pre-shock overdensities $10^1 \leq \delta_b < 10^3$. On the other hand, only 19% of the dissipated energy heats cluster cores ($\delta_b \geq 10^3$).

The temperature contributions show the typical bimodal distribution consisting of external and internal shocks [Ryu et al., 2003]. Previously unshocked cold gas accretes in strong external shocks but accounts only for a small amount of the dissipated energy. On the other hand, gas which gets shocked multiple times and is thus located inside structures produces low Mach number shocks with high energy dissipation. At zero redshift, gas with pre-shock temperatures $10^5 \leq T < 10^7$ accounts for 69% of the energy dissipation. Moreover, in order to determine the contribution of the WHIM, we examine the density of the gas in this pre-shock temperature range and find that $\approx 54\%$ has a baryonic overdensity below $\delta_b = 100$. A detailed break-up of the dissipation rates in different bins of pre-shock density and temperature is given in Table 7.1.

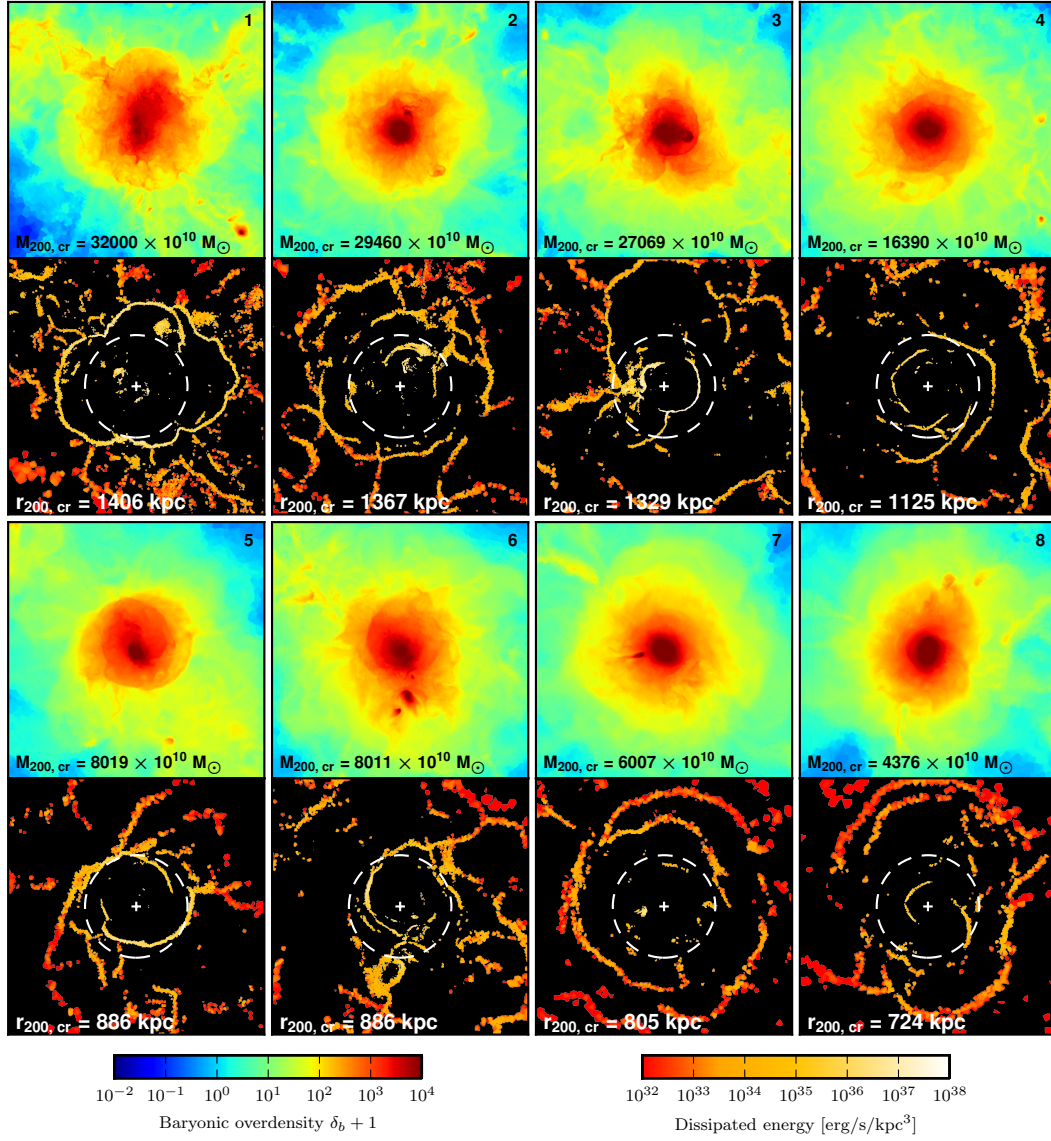


Figure 7.5: Zoom projections with a width of 100 kpc centred on some of the biggest haloes in the Illustris-NR-2 simulation. The top panels show the baryonic overdensity while the bottom panels indicate the energy dissipation. Accretion shocks onto the haloes can be found close to but outside of $r_{200, cr}$ (white circles). Inside the accretion shocks prominent merger shocks are present. We do not find many shocks due to complex flow patterns within clusters, unlike reported by previous studies.

We interpret our findings as follows. Almost 40% of the thermalization happens when gas from the WHIM gets shock heated, whereas shocks inside the accretion shocks of clusters and groups ($\delta_b > 100$) account for $\approx 60\%$ of the dissipation. Thus, the relative importance of the WHIM is significantly higher than what has been found in previous studies. Furthermore, the shocks we identify in clusters and groups are prominent merger shocks rather than stemming from halo-filling, complex flow patterns, as shown in the next section.

7.6 Galaxy cluster shocks

Fig. 7.5 shows zoom-projections of width 100 kpc around different massive haloes of the Illustris-NR-2 simulation. The eight chosen haloes are sorted by mass, and we show the baryonic overdensity as well as the energy dissipation by shocks. The white circle indicates the virial radius $r_{200, \text{cr}}$. The first halo is the biggest halo in the simulation and is also present in Fig. 7.2. As can be seen in the gas projection, there are turbulent motions inside the virial radius. However, we do not find many shocks inside this region, which points towards predominantly subsonic turbulence.

The third halo is in a state similar to the bullet cluster. The small subcluster ploughs through the gas of the big halo and produces a bow shock where a lot of energy is dissipated. The Mach number along the bow shock is $2 \lesssim \mathcal{M} \lesssim 3.5$, comparable to the value $\mathcal{M} = 3 \pm 0.4$ measured for the bullet cluster system [Markevitch, 2006; Markevitch et al., 2002]. The shocks in haloes four and five form interesting spiral structures. These shocks might point towards the following scenario: A minor-merger event triggers gas sloshing and the formation of a spiral cold front [Ascasibar and Markevitch, 2006; Markevitch and Vikhlinin, 2007; Roediger and Zuhone, 2012] which steepens to a shock while propagating outwards. A hint could also be the temperature map of halo four, which indicates that the spiral shock structures fade into contact discontinuities in the interior. In halo six, there are fine shock structures close to each other. In order to resolve these, it is crucial to handle overlapping shock zones as described in Section 5.3.3. The last two haloes are surrounded by prominent accretion shocks. Furthermore, halo eight underwent a merger event recently, as can be seen from the shock remnants inside the virial radius lying opposite to each other. Several of such remnants have been observed in form of double radio relics, as reported for example in Rottgering et al. [1997], Bonafede et al. [2009, 2012], and van Weeren et al. [2009]. The accretion shocks are located outside of the virial radius $r_{200, \text{cr}}$. This is expected since the kinetic and thermal pressure become roughly equal around $r_{200, \text{mean}}$ [Battaglia et al., 2012] which is $\approx 1.5 r_{200, \text{cr}}$ at redshift $z = 0$.

In Fig. 7.6, we present the radial shock distribution for several of the haloes of Fig. 7.5. For the biggest halo in the simulation (halo 1, red curve) the accretion shock is located at around $1.6 r_{200, \text{cr}}$. The green curve (halo 3) shows the bow shock of the bullet at approximately $0.6 r_{200, \text{cr}}$, however, no single radius for the accretion shock can be determined for this highly dynamical system. In the blue curve (halo 8) three bumps can be seen. The first one corresponds to the merger shocks inside

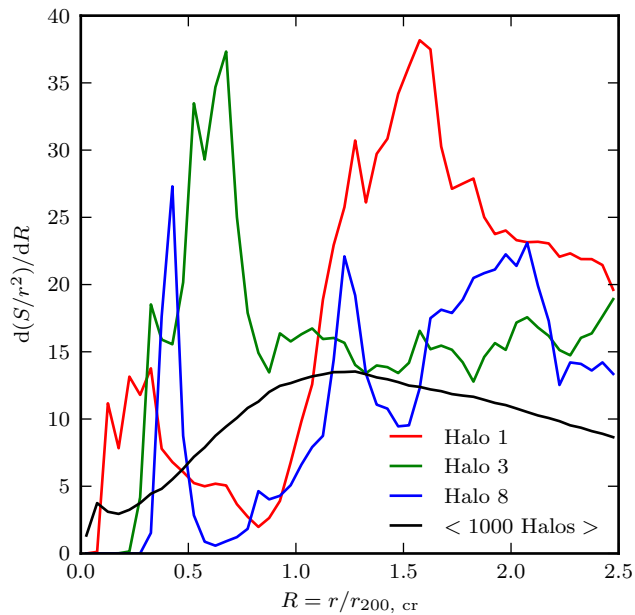


Figure 7.6: The radial shock surface distribution in several of the haloes of Fig. 7.5 (coloured lines, as labelled) as well as the average distribution of the 1000 largest haloes in the Illustris-NR-2 simulation, stacked at redshift $z = 0$. The average dissipation profile has a peak at around $R = 1.3$ in units of the virial radius, which can be interpreted as the typical radius of the accretion shocks.

the virial radius ($R = 0.4$). The two bumps outside the virial radius belong both to the accretion shock and are located at the minor ($R = 1.2$) and major axis ($R \approx 1.9$) of the projected accretion shock ellipsoid. In order to estimate an average radius for accretion shocks, we stack the distributions of the 1000 largest haloes in the simulation ($1.7 \times 10^{12} M_{\odot} \leq M_{200, \text{cr}} \leq 3.2 \times 10^{14} M_{\odot}$) and average them (black curve). We find a signal with a broad peak at $R = 1.3$, implying that the accretion shocks at redshift $z = 0$ can typically be found at around $1.3 r_{200, \text{cr}}$.

7.7 Cosmic ray acceleration

In this section, we discuss the role of the detected shocks in the Illustris-NR-2 run as cosmic ray sources. While our analysis is carried out by post-processing simulation outputs, cosmological simulations including cosmic rays and associated processes have also been presented, for example in Miniati et al. [2001], Pfrommer et al. [2007], Jubelgas et al. [2008], and Vazza et al. [2012].

Cosmic rays get injected in collisionless cosmological shocks by means of the DSA mechanism [e.g. Blandford and Ostriker, 1978; Malkov and O’C Drury, 2001], also known as first order Fermi acceleration. In this process, ions with high thermal energies can diffuse upstream after crossing a shock and gain in a repetitive way in multiple shock crossings more and more energy. The cosmic ray injection efficiency

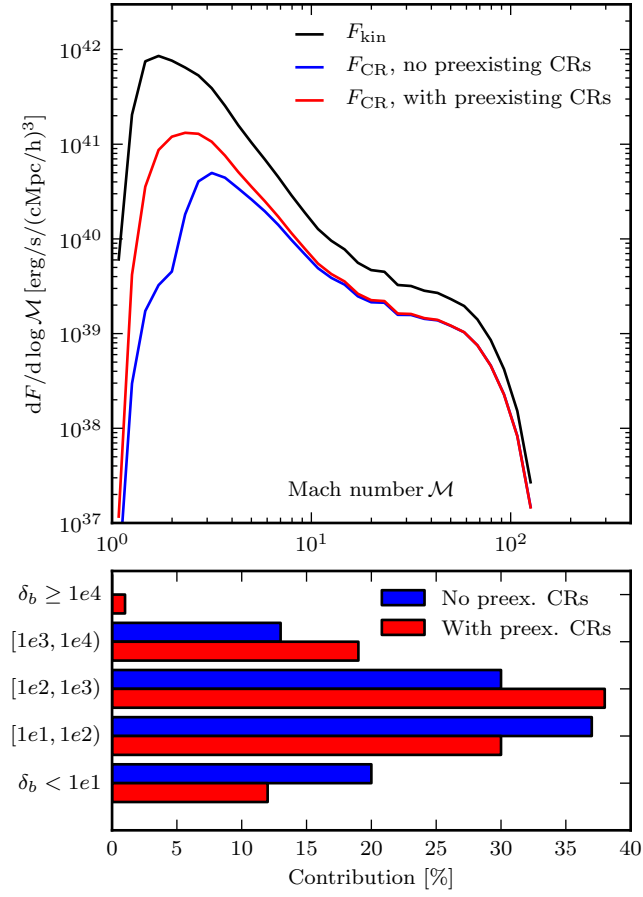


Figure 7.7: Top panel: Mach number dependent energy distribution for accelerating cosmic rays according to the DSA simulations of Kang et al. [2007]. The black line shows the total kinetic energy processed by shocks per time and volume, while the red and blue curves give the fractions expected for particle acceleration with and without pre-existing cosmic rays, respectively. Bottom panel: contribution of different baryonic pre-shock overdensities to the total energy used for cosmic ray acceleration.

depends strongly on the Mach number as well as on the level of Alfvén turbulence, and is most efficient if the magnetic field is parallel to the shock normal. Simulations of this DSA scenario were carried out by Kang et al. [2007], inferring upper limits for the cosmic ray acceleration efficiency at specific Mach numbers, in the case of a pre-existing as well as a non-pre-existing cosmic ray population. Furthermore, fitting functions $\eta(M)$ are provided for both cases and we use these for estimating the energy flux $f_{\text{CR}} = \eta(M)f_{\Phi}$ transferred into cosmic rays, where f_{Φ} is the processed kinetic energy flux.

The kinetic energy processed by shocks in the simulation as well as the estimated available energies for cosmic ray acceleration at redshift $z = 0$ are presented in the top panel of Fig. 7.7. The total inflowing kinetic energy amounts to $1.14 \times 10^{41} \text{ erg s}^{-1} \text{ Mpc}^{-3}$. Without a pre-existing cosmic ray population most of the

acceleration energy is available in shocks with Mach numbers $M \approx 3.2$ (blue curve). Furthermore, if we compare the integrals of the distributions, we estimate that $\approx 6\%$ of the total kinetic energy processed by shocks is transferred into cosmic rays. In the case of pre-existing cosmic rays, which are generated in shocks at earlier times, the DSA mechanism is more efficient, especially for low Mach numbers (red curve). For this scenario, we find a peak position at $M \approx 2.5$ and a kinetic energy transfer to cosmic rays amounting to $\approx 18\%$.

In order to determine the spatial origin of the cosmic rays, we show the contributions of different baryonic overdensities to the total cosmic ray energy in the bottom panel of Fig. 7.7. For both scenarios (with and without pre-existing cosmic rays) most of the energy transfer into cosmic rays is located in regions with densities $10 \leq \delta_b < 10^3$ ($\approx 68\%$). Without a pre-existing population, only 43% of the cosmic ray energy is generated inside clusters and groups ($\delta_b > 100$). On the other hand, pre-existing cosmic rays increase the efficiency for low Mach number shocks, which are mainly present inside dense structures. In this case we find that 58% of the energy for cosmic ray acceleration is provided by shocks inside clusters. We conclude that cosmological shocks could produce a significant amount of cosmic rays. Furthermore, the relevant shocks are shared in comparable proportions among both, regions with overdensities above and below $\delta_b = 100$.

8

Methodology variations

In this section, we turn to an investigation of the origin of the discrepancies between our results and previous studies. For this purpose, we in particular investigate the shock zone determination criteria as well as the Mach number calculation approach for shock detection in the Illustris-NR-2 run.

For calculating the Mach number of shock surface cells, we normally evaluate pre- and post-shock temperatures of cells just outside the shock zone. An exception to this are overlapping shock zones, in which case the pre-shock temperature is taken inside the combined shock zone and between the shock surfaces. Note that with this procedure, Mach numbers are calculated across a variable number of cells (variable jump range), depending on the extent of the imprint of the shock on the primitive variables. In Fig. 8.1 we show how the inflowing kinetic energy distribution as well as the dissipated energy distribution changes when we adopt a fixed jump range instead by calculating the Mach number always from the directly adjacent cells of a shock centre (short jump). Both distributions are shifted towards slightly lower Mach numbers. This finding is expected since a short jump does not fully enclose the broadened shock, and should hence lead to a smaller jump in the measured temperatures. For the dissipated energy distribution we find a peak shift from $\mathcal{M} \approx 2.7$ to ≈ 2.2 as well as a reduction of the total dissipated energy by $\approx 15\%$ when the short jump is adopted.

In the next test we compare our default implementation of shock zone finding to a method in which the requirement for a minimum pressure jump ($\Delta \log p \geq \log p_2/p_1|_{\mathcal{M}=\mathcal{M}_{\min}}$) is abandoned, and the criterion $\nabla T \cdot \nabla \rho > 0$ is replaced with $\nabla T \cdot \nabla S > 0$. These relaxed criteria have been frequently used in previous studies and only recently Hong et al. [2014] suggested a replacement of the temperature–entropy criterion.

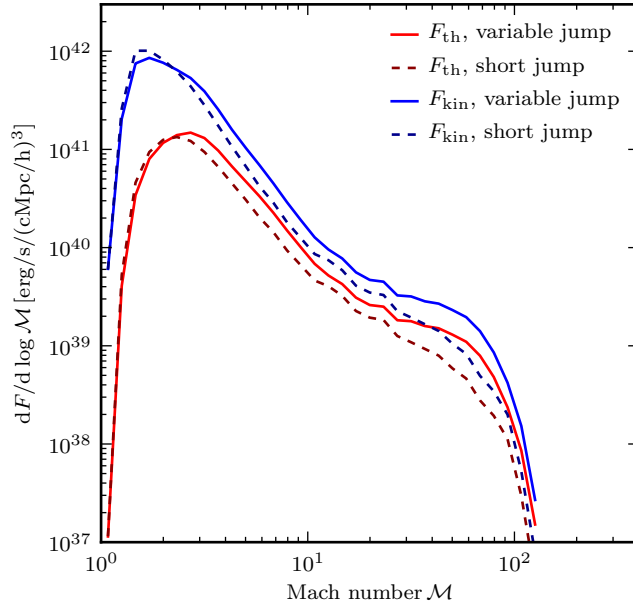


Figure 8.1: Effect on the inflowing kinetic energy and dissipated energy when we calculate the Mach number based just on neighbouring cells (short jump) instead of using our standard implementation with a variable jump range. The restriction to short jumps shifts the distributions to slightly lower Mach numbers.

In Fig. 8.2, we visualize the differences by applying the two methods to the Illustris-NR-2 simulation, omitting a temperature floor here for the purposes of this comparison test. With the old standard shock finding method (right-hand panel), many more low Mach number shocks are found inside the WHIM and inside galaxy clusters. The additional shocks increase the total energy dissipation and the relative contribution of clusters due to high pre-shock densities and temperatures.

We demonstrate in Fig. 8.3 that the additional shocks are spurious by applying the old method to a three-dimensional Kelvin–Helmholtz simulation in which no shocks are present. However, the old implementation is not able to filter tangential and contact discontinuities reliably in the shock zone determination, and thus false positive shocks are found along the density jump.

Depending on the applied scheme for assessing the shock strength (temperature jump, density jump, or velocity jump), some of the spurious shock detections are suppressed by the consistency check that requires a correct jump direction in the shock surface determination (Section 5.3.3). The energy dissipation inferred for the different methods and different jump evaluations is presented in Fig. 8.4. Note that the reionization model through a temperature floor is disabled here, and we therefore obtain a tail towards very high Mach numbers. It can be seen that the old standard shock finding method is very sensitive to the adopted jump quantity for inferring the Mach number, while our new method is rather stable. There is good

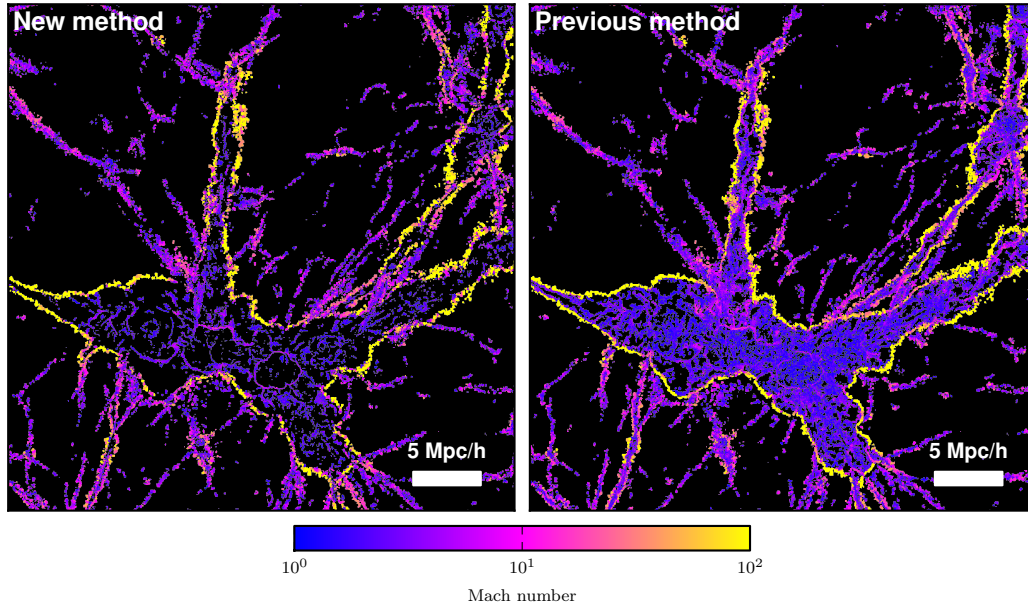


Figure 8.2: Demonstration of the difference between our default shock finder implementation (left-hand panel) and the old standard method (right-hand panel). With the latter implementation many more shocks are found inside the WHIM and inside clusters. These shocks are spurious as demonstrated in Fig. 8.3. Depending on the applied evaluation method for the shock strength, some of the spurious shocks can be suppressed as shown in Fig. 8.4. Nevertheless, if spurious shocks enter the analysis the relative contribution of clusters to the total dissipated energy is overestimated. Note that for this comparison we do not adopt a temperature floor and thus find more shocks and higher Mach numbers compared to Fig. 7.2.

agreement between the temperature jump and velocity jump methods, whereas the pressure jump evaluation shifts the Mach numbers towards higher values for very strong shocks.

The old standard method in combination with the temperature jump gives clearly wrong results that are associated with an overestimate of the energy dissipation, which can be clearly seen from the difference of the dashed and the solid red line. By using the velocity jump instead, spurious shocks can be removed if the jump direction of the normal velocities is not consistent with the shock direction (temperature gradient). However, there remain then still a significant number of cases with wrong shock detections, both with low and high Mach numbers (green dashed line). Note that the pressure jump across such spurious shocks is very small such that only a small Mach number is assigned if the pressure discontinuity is used to measure the shock strength. However, the distribution of energy dissipation at high Mach numbers is also altered compared to our new method, and perhaps surprisingly it is lower in this case (blue dashed line). The origin of this effect lies in the modification of the Mach number of real shocks by spurious shocks. This happens whenever the shock zone of a real shock gets modified (extended) by the shock zone of a spuriously detected shock nearby.

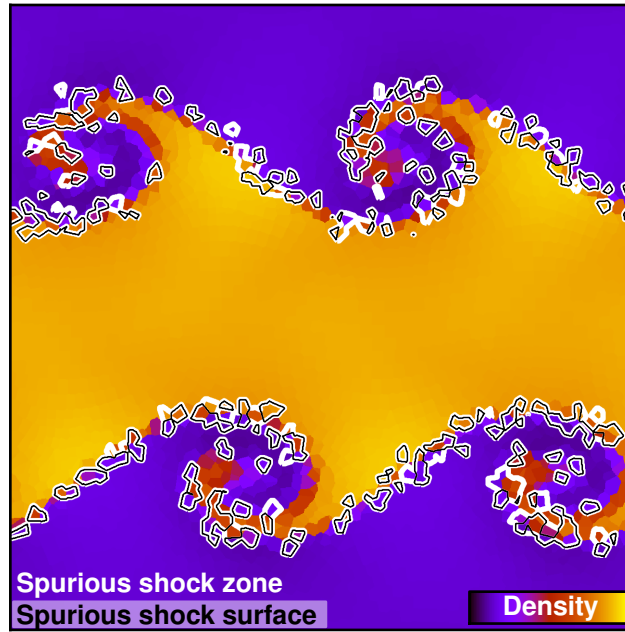


Figure 8.3: Application of the previous standard shock finding method to a three-dimensional Kelvin–Helmholtz instability simulation. This method does not filter tangential and contact discontinuities and thus finds spurious shocks even though this subsonic problem is formally free of shocks. Our new implementation on the other hand does not find shocked cells at all in the simulation volume, as desired.

We conclude that the use of a variable jump range gives a small improvement compared to a fixed small jump range. It is yet more important to properly filter against tangential and contact discontinuities already in the shock zone determination to robustly avoid spurious shock detections and distortions in the resulting shock statistics.

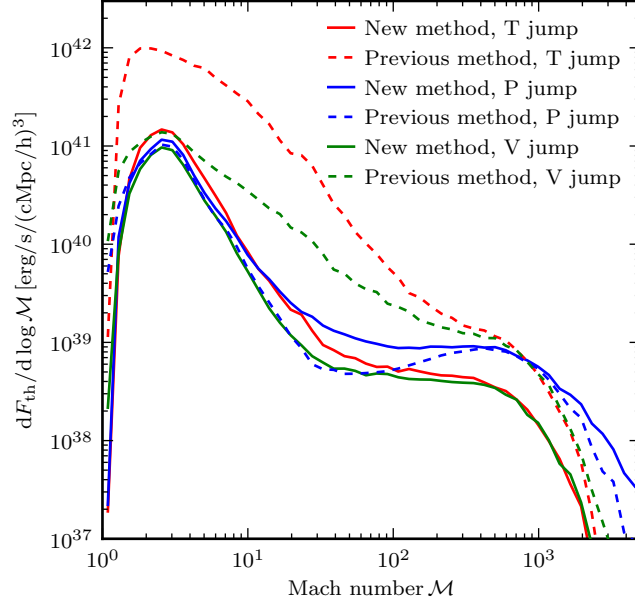


Figure 8.4: Comparison of our new method to a previous standard method which has no minimum pressure jump requirement and uses $\nabla T \cdot \nabla S > 0$ instead of $\nabla T \cdot \nabla \rho > 0$ as a shock zone criterion. The previous method does not filter tangential discontinuities in this case, see also Fig. 8.3. We furthermore investigate the dependence of these methods on the Mach number estimation by using instead of the temperature jump (our standard approach) also the pressure and velocity jumps, as labelled in the panel.

9

Summary

We have implemented a parallel shock finder for the unstructured moving-mesh code AREPO, based on ideas of previous work [Hong et al., 2014; Ryu et al., 2003; Skillman et al., 2008] combined with new refinements. Shocks are detected in a two-step procedure. First, a broad shock zone is determined by analysing local quantities of the Voronoi cells to identify regions of compression. In a second step, a shock surface is identified by finding cells with the maximum compression across the shock zone, followed by an estimate of the Mach number through measuring the temperature jump across the shock zone. In this way, the Mach number is calculated over a variable number of cells which adjusts to the numerical broadening of the particular shock.

Improvements to previous methods have been realized by handling overlapping shock zones and by carefully filtering out tangential discontinuities and contacts in the shock zone determination. Such discontinuities are abundantly present in cosmological simulations, for example in the form of cold fronts and in shear flows that mix gas of different specific entropy. We robustly suppress spurious shock detections by replacing the commonly used criterion $\nabla T \cdot \nabla S > 0$ with $\nabla T \cdot \nabla \rho > 0$ [as also adopted in Hong et al., 2014], and additionally by requiring a suitable minimum pressure jump threshold.

We have shown that shock finder results can be quite sensitive to implementation details. For example, allowing for a variable number of cells when evaluating the Mach number across shocks leads in general to slightly higher (and more correct) values compared to adopting a fixed number of cells. A still more important role is played by the choice of fluid quantities selected for the Mach number calculation. Especially when tangential and contact discontinuities are not cleanly filtered out, an unfortunate choice can here produce substantial distortions in the inferred shock statistics.

We have introduced a new test for assessing the overall performance of a shock finder in which the energy dissipation inferred from the shock detection is directly compared with the actual energy dissipation in a non-radiative cosmological simulation between two consecutive time steps. We find a good performance of our techniques for low redshifts $z \lesssim 4.0\text{--}5.0$ in non-radiative cosmological simulations, where we can identify all strong shocks reliably and account accurately and in a numerically converged way for the bulk of significant dissipation in shocks. At zero redshift, the energy dissipation measured with our shock finder for the adopted cosmological parameters is measured to be $2.3 \times 10^{40} \text{ erg s}^{-1} \text{ Mpc}^{-3}$.

Interestingly, we detected rich shock morphologies in the high-resolution non-radiative simulation Illustris-NR-2. In particular, high Mach number accretion shocks onto filaments and cluster outskirts nicely trace the cosmic web. Accretion shocks onto galaxy clusters dissipate a lot of energy, while merger shocks inside the clusters give hints about their recent formation history. We note that the merger shocks appear as rather prominent and distinct features, whereas we do not find complex flow shock patterns inside the cluster accretion shocks, suggesting that there the gas dynamics is mostly characterized by subsonic turbulence [which we expect to be well captured by AREPO; Bauer and Springel, 2012].

With our improved methodology, we find quantitatively revised results for the shock dissipation statistics in non-radiative cosmological simulations. Most of the thermalization happens in shocks with Mach numbers around $\mathcal{M} \approx 2.7$. Moreover, almost 40% is contributed by shocks in the WHIM and $\approx 60\%$ by shocks in clusters and groups. Compared to previous studies, these findings correspond to a shift in the energy dissipation spectrum towards higher Mach numbers and towards structures with lower densities. Also, we have found $R = 1.3 r_{200, \text{cr}}$ as a typical radius for accretion shocks onto galaxy clusters at redshift $z = 0$, based on identifying a peak when stacking the radial shock dissipation profiles of 1000 haloes of the Illustris-NR-2 simulation. Consequently, the accretion shock is expected to typically lie slightly outside the virial radius, a finding which is consistent with Battaglia et al. [2012]. We note however that the accretion shock position shows a high degree of temporal variability in any given halo.

Finally, we also investigated the expected energy transfer to cosmic rays in the identified large-scale structure shocks if acceleration efficiencies derived from DSA plasma simulations are adopted [Kang et al., 2007]. These simulations are set up with a magnetic field parallel to the shock normal direction and provide therefore upper limits for the acceleration efficiency. We obtain at redshift $z = 0$ an average cosmic ray energy injection rate of $7.0 \times 10^{39} \text{ erg s}^{-1} \text{ Mpc}^{-3}$ in the case of non-pre-existing cosmic ray populations, and a considerably larger value in the case of pre-existing cosmic rays. Considering these numbers, it is quite plausible that even for random magnetic field orientations a dynamically important cosmic ray population is produced in these shocks. Furthermore, we found that gas with pre-shock overdensities above and below $\delta_b = 100$ contribute roughly equally to the energy transfer into cosmic rays.

In future work, it will be interesting to couple the shock finder to hydrodynamical simulations that take cosmic rays self-consistently into account. Also, it should be

interesting to contrast the results obtained here for non-radiative simulations with an analysis of full physics simulations of galaxy formation that include radiative cooling and heating mechanisms, as well as prescriptions for star formation, stellar evolution, black hole growth, and associated feedback processes. These simulations feature interesting additional shocks, for example due to strong feedback-driven outflows. In the next part of the thesis, we will present an analysis of the corresponding shocks in the recent Illustris simulation [Vogelsberger et al., 2014b].

Part III

Hydrodynamic Shocks in the Illustris Universe

Abstract

Hydrodynamical shocks are a manifestation of the non-linearity of the Euler equations and play a fundamental role in cosmological gas dynamics. In this work, we identify and analyse shocks in the Illustris simulation, and contrast the results with those of non-radiative runs. We show that simulations with more comprehensive physical models of galaxy formation pose new challenges for shock finding algorithms due to radiative cooling and star forming processes, prompting us to develop a number of methodology improvements that robustly suppress spurious shock detections. We find in Illustris a total shock surface area which is about 1.4 times larger at the present epoch compared to non-radiative runs, and an energy dissipation rate at shocks which is higher by a factor of around 7. Remarkably, shocks with Mach numbers above and below $\mathcal{M} \approx 10$ contribute about equally to the total dissipation across cosmic time, and for $z < 3$ we measure a specific dissipation rate which is roughly constant and fairly independent of environment. This is in strong contrast to non-radiative simulations, for which most of the kinetic energy is thermalised in low Mach number shocks. By comparing the energetics of different feedback channels with the associated shocks we demonstrate that a large part of the difference arises due to black hole radio-mode feedback. We argue that the lack of observational signatures expected for such strong shocks could be used to constrain the modelling and parameter space of feedback implementations. We also provide an overview of the large diversity of shock morphologies across cosmic time, which includes complex networks of halo-internal shocks, shocks on to cosmic sheets, feedback shocks due to black holes and galactic winds, as well as ubiquitous accretion shocks on to non-linear structures. In high redshift systems more massive than $10^{12} M_{\odot}$ we discover the existence of a double accretion shock pattern in haloes. They are created when gas streams along filaments without being shocked at the outer accretion shock, but then forms a second, roughly spherical accretion shock further inside.

10

Introduction

Hydrodynamical shock waves play an important role in the evolution of the baryonic component of our Universe. During the process of hierarchical structure formation, shocks dissipate and thermalise kinetic energy in supersonic gas flows, thereby heating the cosmic web and structures therein. Gas inside dark matter haloes can be compressed and virialized by means of shocks, before radiative cooling in the densest regions allows it to collapse further, eventually leading to the formation of stars and accretion on to black holes. Supersonic gas motions are also created by feedback processes such as stellar winds, supernovae explosions, or relativistic jets from active galactic nuclei (AGN). The conversion of the released mechanical energy to thermal energy in the ambient gas is again mediated by thermalisation processes that ultimately involve hydrodynamical shocks.

Unlike in classical hydrodynamical shocks in which the kinetic energy of a fluid is thermalised via particle collisions, shocks on cosmological and galactic scales are *collisionless*. This is because the mean free path of particles in the corresponding environments is large compared to the transition layer between the pre-shock and post-shock regions. In this case, the energy transfer is mediated by electromagnetic viscosity [Kennel et al., 1985; Wentzel, 1974]. Interestingly, collisionless shocks are also sites of diffusive shock acceleration [DSA; Axford et al., 1977; Bell, 1978a,b; Blandford and Ostriker, 1978; Krymskii, 1977; Malkov and O’C Drury, 2001], creating relativistic cosmic rays. In this process, ions inside the converging flow can cross the shock multiple times by scattering off the magnetic field fluctuations in the pre- and post-shock regions, thereby gaining more and more energy with every crossing. The net result is a power-law momentum spectrum of energetic particles, thus creating a non-thermal cosmic ray particle component that may modify the gas dynamics in important ways. This in particular happens in supernova remnants, but may also be important for accretion shocks around dark matter haloes and active galactic nuclei.

Directly observing cosmological shocks is in general very difficult. While the central high density regions of galaxy clusters can be observed in X-ray, the gas temperature in these environments is high as well. This allows only for low Mach number shocks, and these are difficult to detect since the density contrast between the shocked and unshocked gas is small. Still, a number of bow shocks in merging clusters have been discovered by analysing *Chandra* observations [e.g. Dasadia et al., 2016; Macario et al., 2011; Markevitch et al., 2005; Russell et al., 2014], the most famous being the Mach number $\mathcal{M} \approx 3$ shock in the bullet cluster [Markevitch et al., 2002].

Fortunately, the synchrotron emission of accelerated or reaccelerated electrons can reveal the location of a cosmological shock more easily. Relativistic cosmic ray electrons have short life times, and their non-thermal emission is therefore confined to a region close to the accelerating shock. There is strong evidence that radio relics [e.g. Brüggen et al., 2011; Feretti, 2005; Ferrari et al., 2008], which can be found in peripheral cluster regions, are associated with merger shocks [Ensslin et al., 1998; Miniati et al., 2001; Pfrommer, 2008; Pfrommer et al., 2008; Pinzke et al., 2013; Skillman et al., 2013]. These sources are elongated and often curved, with length scales up to a few Mpc. Moreover, the emitted photons are highly polarised and indicate in many cases a magnetic field which is aligned with the relic [Brüggen et al., 2012]. Strikingly, for a few clusters the associated shock is directly observed in the form of an X-ray edge [e.g. Akamatsu and Kawahara, 2013; Bourdin et al., 2013; Markevitch, 2010; Ogrean and Brüggen, 2013; Owers et al., 2014].

Radio relics are very interesting objects to study with the current generation of radio observatories, but they are in principle only the brightest components of a much larger entity, the magnetised cosmic web. The cosmic web consists of non-linear structures in the form of a network of filaments and nodes, which is expected to be surrounded by more or less stationary accretion shocks. Based on current estimates for the magnetisation and electron acceleration efficiencies at these shocks, the expected synchrotron radiation will be in reach of the Square Kilometre Array (SKA) [Battaglia et al., 2009; Keshet et al., 2004; Vazza et al., 2015]. Consequently, in the near future it should become possible to observe filaments of the cosmic web and their surrounding shocks, providing a novel probe of cosmic large-scale structure formation and the simultaneous opportunity to test simulation models.

In this part of the thesis, we analyse shock waves in an advanced cosmological simulation of galaxy formation, the Illustris Simulation. It is the second part of a short series on shocks in numerical hydrodynamic simulations. In the first part, which corresponds to Part II of this thesis, we have introduced our shock finding algorithm which operates on the Voronoi mesh of the AREPO code [Springel, 2010], and analysed non-radiative cosmological simulations of gas and dark matter. Those simulations did not incorporate cooling or feedback of any kind, and hence the shocks which are present can mostly be classified as merger shocks and accretion shocks. In this case, the shock statistics features a bimodality where so-called external shocks enter with high Mach numbers and internal shocks inside structures with relatively low Mach numbers [Ryu et al., 2003]. The external shocks are created when pristine gas accretes on to the cosmic web and gets shock heated for the first time. While the

corresponding flows are highly supersonic, only a little kinetic energy is thermalized due to the low densities involved. On the other hand, internal flows are characterised by high temperatures and densities, which lead to very energetic low Mach number shocks.

Shocks in such non-radiative simulations have been analysed extensively in previous studies (see for example Hong et al. [2014]; Miniati et al. [2000]; Pfrommer et al. [2006]; Quilis et al. [1998]; Ryu et al. [2003]; Skillman et al. [2008]; Vazza et al. [2009a, 2010], and Vazza et al. [2011] for a code comparison project), but only few authors have explored runs including radiative physics and star formation [e.g Kang et al., 2007; Pfrommer et al., 2007; Planelles and Quilis, 2013]. It is therefore the aim of this part of the thesis to investigate shocks in a state-of-the-art cosmological simulation that includes gas and dark matter, as well as stars and black holes and their associated feedback processes. These feedback sources are expected to create powerful additional shocks, making it particularly interesting to compare the resulting shock statistics with those obtained for the non-radiative runs. We shall use the Illustris simulation suite [Genel et al., 2014; Vogelsberger et al., 2014b,c] for our study, focusing on the highest-resolution, full physics run (Illustris-1), also known as *the* Illustris simulation. This is among the presently most successful simulations of galaxy formation and uses an accurate mesh-based hydrodynamical method, making it particularly well suited for our purposes.

This part of the thesis is structured as follows. Section 11 recaps the analysed simulations, the simulation technique, and the shock finding method. Section 12 focuses on the interpretation of shock statistics, both in a global sense and within different environments. In Section 13, we present shock morphologies across cosmic time and discuss what they can reveal about the gas dynamics. Variations in our shock finding techniques that test our methodology are investigated in Section 14, and we study resolution effects in Section 15. Lastly, our findings are summarised and discussed in Section 16.

11

Analysing shocks in the Illustris universe

11.1 The Illustris project

At the center of the Illustris project is the cosmological hydrodynamical simulation Illustris-1 [Genel et al., 2014; Vogelsberger et al., 2014b,c]. It evolves around 1.2×10^{10} resolution elements of dark matter and baryons in a comoving periodic volume $75 h^{-1} \text{Mpc}$ on a side. With this setup a dark matter mass resolution of $6.26 \times 10^6 M_{\odot}$ and an initial gas mass resolution of $1.26 \times 10^6 M_{\odot}$ are achieved. The simulations of the Illustris project adopt the standard ΛCDM model and cosmological parameters according to the 9-year *Wilkinson Microwave Anisotropy Probe* (WMAP9) observations [Hinshaw et al., 2013], with the following values: $\Omega_{\text{m}} = 0.2726$, $\Omega_{\Lambda} = 0.7274$, $\Omega_{\text{b}} = 0.0456$, $\sigma_8 = 0.809$, $n_s = 0.963$, and $H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}$ where $h = 0.704$.

The simulations of Illustris were carried out with the moving-mesh code AREPO [Springel, 2010] and a galaxy formation physics implementation described and validated in detail in Vogelsberger et al. [2013] and Torrey et al. [2014]. In the following, we briefly summarise the key features of the physics model adopted for Illustris-1. The interstellar medium (ISM) is represented through a sub-resolution approach with an effective equation of state following Springel and Hernquist [2003]. In this model, a self-regulated ISM arises through the interplay of gas cooling, star formation, supernova (SN) feedback and galactic winds. Star particles are created stochastically in overdense regions according to the estimated local star formation rate, and represent single-age stellar populations characterized by a Chabrier [2003] initial mass function (IMF). Additionally, supernovae energy is injected as purely kinetic energy of wind particles. The latter are briefly decoupled from the hydrodynamic scheme until they leave the dense star-forming region and can deposit their momentum into the surrounding lower density gas. Black holes (BHs) and the associated energy feedback

processes from gas accretion are implemented following Springel et al. [2005] and Sijacki et al. [2007]. To this end, black hole sink particles are inserted at the potential minima of newly identified dark matter haloes that do not yet contain black holes. They grow in mass by Eddington-limited Bondi accretion and BH merger events.

For high accretion rates relative to the Eddington rate, the BH is assumed to be in a quasar-mode, and the model couples the feedback energy thermally to the surrounding gas. For low accretion rates, a different feedback channel, a radio-mode, is assumed in which mechanical feedback occurs through jets. Since the jets cannot be directly resolved, the hot radio bubbles inflated by them are instead modelled through an injection of heat energy in spherical volumes corresponding to the bubbles. Additionally, a scheme for radiative AGN feedback is adopted in which the elevated ionizing flux and its impact on the cooling rates is accounted for in the proximity of actively accreting black holes. The total radiative heating rate of the gas is calculated by the superposition of the AGN radiation field and a spatially uniform but time-dependent photoionizing UV background (UVB). The cooling rate in the simulations has three contributions: atomic cooling based on the equations describing the ionisation balance of hydrogen and helium [Katz et al., 1996], metal line cooling calculated from pre-computed cooling tables, and Compton cooling off the cosmic microwave background. Nine different elements are explicitly tracked in the simulation (H, He, C, N, O, Ne, Mg, Si, Fe), whose abundance is increased by stellar winds from AGB stars, core collapse supernovae (SN), and type-Ia SNe.

With this setup of the Illustris simulations, a significant number of basic properties of the galaxy population beyond those used in tuning the model (which are mostly the present-day stellar mass function and the cosmic star formation rate history) have been found to be in good agreement with observations. Remarkably, these include realistic morphologies and velocity structures for around 40000 well-resolved galaxies [Genel et al., 2014; Snyder et al., 2015; Torrey et al., 2015; Vogelsberger et al., 2014b,c]. The simulation also contains 32552 black holes at $z = 0$ with properties that match observed quasar luminosity functions and galaxy-BH correlations quite well [Sijacki et al., 2015]. The star formation histories and stellar masses of galaxies show a plausible diversity [Sparre et al., 2015], and the hydrogen reionization history of the universe can be successfully explained by the high-redshift galaxies forming in Illustris [Bauer et al., 2015]. All simulation data has recently been made publicly available [Nelson et al., 2015c], and a comprehensive list of further studies analysing the simulation data can be found online¹.

While the focus of this work is on the analysis of shocks in the highest-resolution full physics Illustris-1 simulation, we also show results for the lower resolution counterparts Illustris-2 and Illustris-3, and a few selected results for the non-radiative simulations Illustris-NR-2 and Illustris-NR-3. Here the numbers 2 and 3 indicate an 8 and 64 times lower mass resolution compared to Illustris-1, respectively. We note that a more extensive analysis of the non-radiative simulations can be found in Part II of the thesis.

¹ See <http://www.illustris-project.org/results/>

11.2 Moving-mesh hydrodynamics

The cosmological simulations analysed in this work were run with the AREPO code [Springel, 2010], which solves the Euler equations of ideal hydrodynamics on a moving Voronoi mesh by means of a second order finite volume method. This approach improves on traditional hydrodynamic solvers such as smoothed-particle hydrodynamics (SPH) and grid-based schemes by combining advantages of each. In particular, the Euler equations are solved accurately by using a finite volume method on the unstructured Voronoi mesh. The mesh-generating points are moved with the local velocity field, resulting in very small residual mass fluxes across cell interfaces and hence a quasi-Lagrangian and manifestly Galilean-invariant behaviour. This is desirable in the cosmological structure formation problem, because it allows the accurate tracking of highly supersonic flows without large advection errors and additional timestep constraints. Also, the natural adaptivity of the mesh resolution means that the mass per resolution element is automatically kept approximately constant.

Recently, Pakmor et al. [2016] introduced refinements for the moving-mesh scheme in the form of an improved gradient estimator and a suitable Runge-Kutta time integrator. They have shown that with these changes AREPO conserves angular momentum to a high degree and the convergence rate of the code is improved. We note, however, that the Illustris-1 simulation still used an older version of the code.

Moving-mesh simulations are well suited for the analysis of hydrodynamic shocks in cosmology. In a finite volume scheme, the solution is discontinuous across cell interfaces and no artificial viscosity is required, allowing shocks to be captured in principle over one or very few cells. Furthermore, with the moving-mesh approach the resolution at shocks tends to be increased because the compression of the gas also implies a compression of the grid cells transverse to the shock front, creating an effective enhancement of the resolution particularly in the post-shock region. This feature is very beneficial for constructing an accurate shock finding algorithm, as outlined in the subsequent section.

11.3 Shock finding

Constructing an accurate method for identifying hydrodynamic shocks in cosmological simulations proves to be challenging. They are numerically broadened across a few cells and local criteria are therefore not sufficient. Moreover, tangential discontinuities arising in shear flows or contact waves are omnipresent and must not be confused with shocks by the algorithm.

Our method for revealing shocks in moving-mesh simulations is described comprehensively in Part II of the thesis. It is based on a generalisation of previous algorithms [Hong et al., 2014; Ryu et al., 2003; Skillman et al., 2008] to an unstructured mesh and incorporates a number of additional improvements. Here we briefly outline the main steps of the algorithm.

First, we flag cells in regions around shocks, which we call *shock zones* [as in Skillman et al., 2008]. Specifically, a cell is considered part of a shock zone if and only if all of the following three criteria are met. (i) The gas inside the cell is compressed (negative velocity divergence), (ii) the temperature and density gradient must point in the same direction, and (iii) a minimum Mach number jump ($M \approx 1.3$) is exceeded.

In a second step, the shock is reconstructed as a single layer of cells containing the *shock surface*. The corresponding cells are given as those shock zone cells that exhibit the maximum compression along the shock normal, which is defined by means of the temperature gradient. Finally, the Mach number is estimated from the Rankine-Hugoniot temperature jump condition based on the pre- and post-shock temperature values directly outside of the shock zone.

Once the Mach number of the shock is known, its energy dissipation can be calculated, which is the fraction of kinetic energy *irreversibly* converted into thermal energy. The corresponding thermal energy flux f_{th} is given by the product of the thermalisation efficiency $\delta(\mathcal{M})$ [Kang et al., 2007; Ryu et al., 2003] and the inflowing kinetic energy flux,

$$f_{\text{th}} = \delta(\mathcal{M}) \frac{1}{2} \rho_1 (c_1 \mathcal{M})^3, \quad (11.1)$$

where ρ_1 and c_1 are the pre-shock density and pre-shock sound speed, respectively.

In Part II of the thesis, we have demonstrated that our algorithm can reliably characterise shocks in non-radiative simulations. However, the full physics run Illustris-1 incorporates cooling, star formation and feedback, which poses additional challenges for the shock finding algorithm. It turns out that this necessitates additional algorithmic measures in the detection of shocks in full physics runs, and we hence extend our methodology as follows.

First of all, we need to take care of the star-forming regions. Because the pressure of cells with active star formation is set artificially such that the gas follows an effective equation of state based on a sub-grid multiphase ISM model [Springel and Hernquist, 2003], the sharp transition in the thermodynamic properties between star-forming and non star-forming cells can potentially be interpreted as a shock by our algorithm, distorting our shock statistics. We therefore exclude all cells which lie in density-temperature phase space close to the effective equation of state. Specifically, we ignore detections for which either the pre-shock cell, the post-shock cell, or the cell with the shock fulfils $\rho > \rho_{\text{sfr}}$ and $T < 10^5 (\rho / \rho_{\text{sfr}})^{0.2}$ K, where ρ_{sfr} is the star formation threshold. The excluded region is indicated by a grey dashed line in the mass-weighted density-temperature histogram discussed in Section 12.1.

Furthermore, gas cooling and fragmentation due to self-gravity can sometimes result in large gradients of the primitive variables which can appear as correspondingly strong jumps when these gradients are not well-resolved locally. In such a case, the inferred jumps are not consistent with the Rankine-Hugoniot conditions across shocks and should hence not be taken into account. Addressing this issue is primarily important in high density regions. In these locations strong cooling can be present,

and moreover, measured energy dissipation rates of spurious detections contribute significantly to the overall dissipation.

We therefore adopt two further precautions in our shock finding algorithm when applied to full physics runs. Firstly, for all tentative shock detections we compare the pre-shock values of pressure, density, and velocity to the corresponding post-shock values, and ignore a detection if one of these variables change in the wrong direction. Recall that the shock direction in our algorithm is determined by means of the temperature gradient.

Secondly, for overdense regions with pre-shock densities $\delta_b > 1000$, we calculate in addition to the Mach number based on the temperature jump (\mathcal{M}_T), the Mach numbers based on the pressure jump (\mathcal{M}_p) and density jump (\mathcal{M}_ρ). We then check the mutual consistency of the Mach numbers, and only keep detections with $f^{-1}\mathcal{M}_p < \mathcal{M}_T < f\mathcal{M}_p$, and $f^{-1}\mathcal{M}_\rho < \mathcal{M}_T < f\mathcal{M}_\rho$. The latter filter is only applied to detections with $M_\rho < 3$ or $M_T < 3$, since the density jump is not sensitive for high Mach numbers.

With the parameter f introduced here, the tolerance for accepting detections that deviate from the expected jump conditions can be adjusted. From a theoretical point of view, if a jump in the hydrodynamic variables represents a shock, the Mach numbers inferred by the different jump conditions should be perfectly equal. However, this equality can of course not be expected when measuring shock parameters in numerical simulations due to the unavoidable discretization errors, and consequently, f should be set to a value well above unity. In this work we adopt $f = 2$.

The presence of the unwanted effects described above is demonstrated in Section 14, where we also show that our approach for coping with these limitations is effective. Most importantly, the adopted numerical parameter of $f = 2$ is a robust choice; varying it within the interval $[1.3, 4.0]$ gives very similar results. This finding indicates that spurious detections of shocks due to poorly resolved local cooling and fragmentation effects indeed strongly violate the Rankine-Hugoniot jump conditions, as otherwise they could not be filtered out so easily.

We want to note that for strongly radiating gas the jump conditions across shocks deviate from the classic Rankine-Hugoniot conditions derived for non-radiating gas. This is especially the case for regions where the cooling time scale is very small. In our algorithm the densest regions around star-forming gas are not taken into account, and we hence assume that we can achieve sufficient accuracy by using the unmodified Rankine-Hugoniot jump conditions.

When we analyse the non-radiative run Illustris-NR-2, an important issue arises due to its lack of cosmic reionization. This leads to void regions that are unrealistically cold at low redshift, and hence the strength of shocks in gas that streams out of voids and accretes on to the cosmic web is overestimated. In order to obtain more realistic Mach numbers for these shocks, we impose a temperature floor of 10^4 K on the gas in the post-processing of non-radiative simulations, similar as done in Part II and in previous studies [Ryu et al., 2003; Skillman et al., 2008].

12

Shock statistics

12.1 Global shock statistics

We start by investigating global shock patterns and statistics in the Illustris-1 simulation. Fig. 12.1 presents projections of the mean baryonic overdensity, the mass-weighted temperature, the dissipated energy weighted Mach number, and the mean dissipated energy density at different redshifts. For the density and temperature, the projection depths are 100 kpc, whereas for the quantities inferred from the shock finder we adopt a smaller value of 50 kpc in order to better show the thin shock surfaces. The extent of each panel corresponds to the width of the full simulation volume of $75 h^{-1} \text{Mpc}$ and is centered on one of the most massive objects at $z = 0$. The mass inside $r_{200, \text{cr}}$, the radius for which the mean overdensity of the halo equals 200 times the critical density of the universe, is $M_{200, \text{cr}} = 2.2 \times 10^{14} M_{\odot}$ for this central object, similar to the most massive cluster in the volume with $M_{200, \text{cr}} = 2.3 \times 10^{14} M_{\odot}$.

At redshift $z = 2$, corresponding to a lookback time of around 10 Gyr, the large scale picture is dominated by accretion onto collapsing structures and shocks due to moderate black hole feedback. At this time, a typical Mach number for accretion shocks in the cosmic web is $\mathcal{M} \approx 2$. The pre-shock gas consists of streams from voids with temperatures around $4 \times 10^3 \text{ K}$ (see also Fig. 12.2) and a sound speed of $c \approx 7 \text{ km s}^{-1}$, assuming no ionisation. A typical velocity in the shock frame for the primordial gas streams is therefore $v = c\mathcal{M} \approx 15 \text{ km s}^{-1}$. With time the Mach numbers of shocks around non-linear structures increase. There appear to be two main reasons for this trend. Firstly, the sound speed of voids is lowered by the adiabatic expansion of the universe (although much of this effect is compensated by heating from the UV background) and ever larger infall velocities are produced by the growing masses of objects, and secondly, the black hole feedback processes increase in strength with cosmic time.

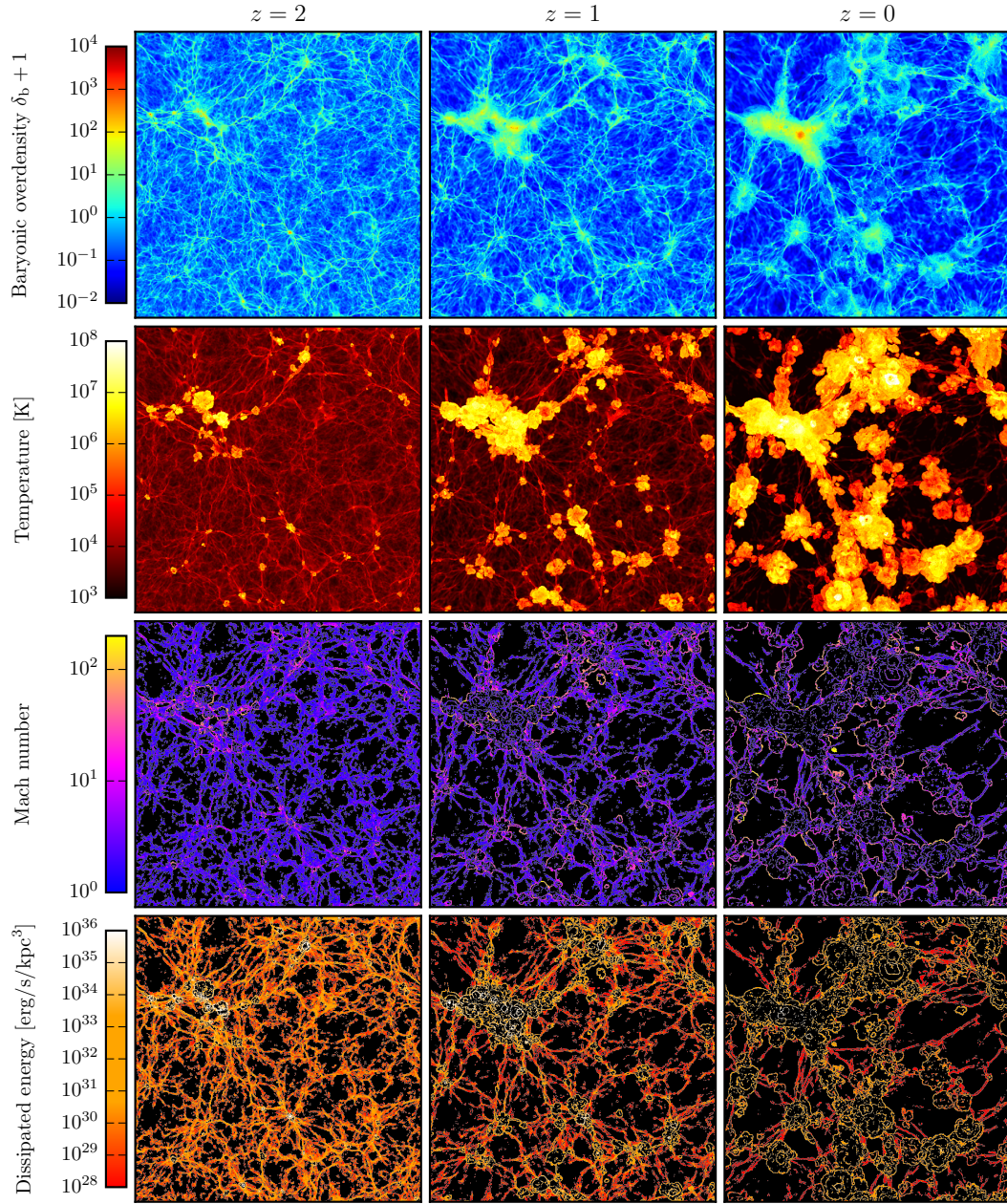


Figure 12.1: Projections of the baryonic overdensity, the temperature, the Mach number field, and the energy dissipation rate density for the full Illustris-1 box ($75 h^{-1} \text{Mpc}$ per side) at different redshifts. The z -coordinate of the projections is centred on one of the most massive clusters in the simulation which can be seen in the upper left of each panel ($M_{200, \text{cr}} = 2.2 \times 10^{14} M_{\odot}$ at $z = 0$). On large scales, accretion shocks on to the cosmic web dominate in the early universe, whereas at late times the radio mode feedback of black holes is omnipresent and creates hot bubbles that shock heat the ambient medium. When cold gas from voids gets shocked, Mach numbers of several hundreds can be reached, but the overall energy dissipated is dominated by the cosmic web. A movie of the shock dynamics in a sub box of Illustris-1 may be accessed online: <https://youtu.be/1CXCRv3i3uw>.

In particular, between $z = 2$ and $z = 1$, the black hole radio mode feedback starts driving hot bubbles that resemble Sedov-Taylor blast waves and feature Mach numbers of several tens in a significant fraction of the simulation volume. These shocks become even stronger at redshift zero when they run into cool and low density voids, with peak Mach numbers of several hundred. We can estimate the speed of a $\mathcal{M} = 200$ blast wave by neglecting the speed of the pre-shock gas in the lab frame and assuming a void temperature of 10^3 K. This temperature corresponds to a sound speed of $c \approx 3.4 \text{ km s}^{-1}$ (assuming a mean molecular weight of neutral gas, for simplicity), and we obtain for the velocity of the shock $v \approx 700 \text{ km s}^{-1}$. This value is consistent with the velocities inferred by Haider et al. [2016], who compared the diameter of the hot bubbles at different redshifts and reported expansion speeds of $500 - 1000 \text{ km s}^{-1}$.

The bottom panels of Fig. 12.1 show the energy dissipation at shocks, which depends on the Mach number, the pre-shock density, and the pre-shock sound speed. The latter two quantities are high in the interior of non-linear structures and the total energy dissipation at shocks is typically dominated by these locations [Ryu et al., 2003]. It can be seen that in the Illustris simulation a high energy dissipation rate is present inside the densest structures between $z = 2$ and $z = 1$, however it seems that at $z = 0$ the bulk of the thermalisation happens in more extended regions. Quantifying these qualitative observations is one of the main goals of this work and will be pursued in subsequent sections.

Large-scale feedback shocks which are omnipresent at $z = 0$ have the effect of converting black hole radio mode feedback energy into thermal energy of the gas. In this way, an extended warm hot intergalactic medium (WHIM) gas phase is created, with temperatures between 10^5 K and 10^7 K. In Fig. 12.2 we plot the gas volume fraction per logarithmic temperature bin as a function of temperature for different redshifts. At all times, void regions occupy the largest volume in the simulation, and at $z = 2$ the distribution peaks around a void temperature of 5×10^3 K. With increasing time, this peak shifts towards lower temperatures due to the expansion of the universe and the associated adiabatic gas cooling, reaching 10^3 K at $z = 0$. Interestingly, the volume distribution develops a bimodality at late times due to a monotonically increasing fraction of gas with temperatures above 10^5 K, revealing the extended WHIM created by the black hole radio mode feedback¹. At $z = 0$ this gas phase occupies around 40% of the volume in the simulated universe. Furthermore, it contains more than 60% of the total baryonic mass [Haider et al., 2016].

An important effect of AGN feedback from supermassive black holes is the quenching of star formation in massive systems at late times. In fact, the parameters of the adopted radio mode feedback model have been set in an attempt to match observations of the galaxy stellar mass function [e.g. Bernardi et al., 2013] and the cosmic star formation rate density [e.g. Behroozi et al., 2013]. Our results suggest that it would be potentially also very constraining to consider the effects of AGN

¹ The bimodality can clearly be attributed to feedback processes, since in non-radiative simulations the volume fraction of gas with $T > 10^5$ K is negligible at all redshifts (not shown in Fig. 12.2).

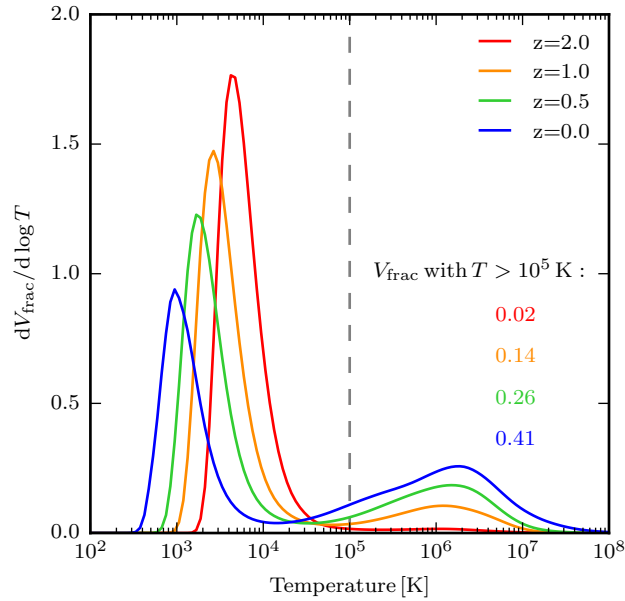


Figure 12.2: Differential volume fraction of the gas in the Illustris-1 simulation as a function of temperature with respect to the whole simulation volume. Due to black hole feedback processes the distribution develops a bimodality at late times; at redshift zero a significant volume fraction ($\approx 40\%$) is filled with gas hotter than 10^5 K. The positions of the left peaks indicate typical void temperatures in Illustris-1 for the different redshifts.

feedback on the predicted WHIM properties. Unfortunately, this gas phase is difficult to observe due to its low density and comparatively high temperature, resulting in low emission and absorption efficiencies, limiting at present the power of this approach.

Nevertheless, the strong Sedov-Taylor blast wave like shocks due to black hole feedback originate in clusters and groups, where the gas density is high. In this environment, as we will showcase in Section 13, these shocks can have higher Mach numbers and larger dissipation rates than cluster merger shocks. The latter can be observed in the local Universe in a few cases directly, and in many others as radio relics. Moreover, shocks associated with AGN have been observed and identified [e.g. Blanton et al., 2009; Cavagnolo et al., 2011; Gitti et al., 2011; Jetha et al., 2008; Nulsen et al., 2005a,b; Randall et al., 2011], however, they are typically of low strength with Mach numbers $\mathcal{M} \lesssim 2$ and located relatively close to the cluster centre, at distances of at most 200 – 300 kpc. To our knowledge no strong shocks with properties similar to the AGN feedback shocks in the Illustris simulation have been seen in observations thus far. Their absence hence indicates that the radio mode feedback channel is too strong in the simulation model. This shortcoming of the simulation has also been pointed out by other authors based on other lines of evidence. In particular, as a result of the redistribution of baryons due to the overly

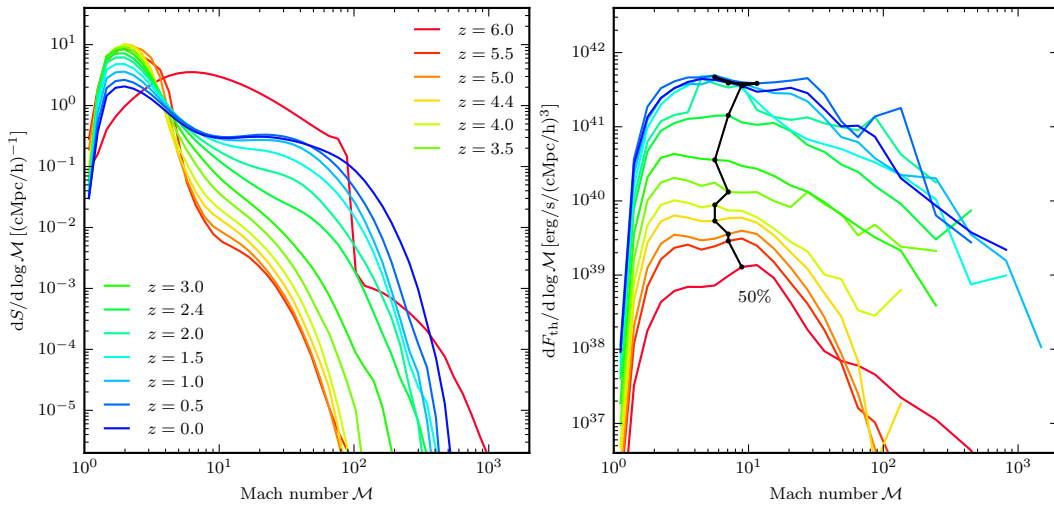


Figure 12.3: *Left panel:* Shock surface area per volume as a function of Mach number; this statistic is dominated by external shocks on to the cosmic web. The red line indicates the shock distribution on the eve of reionization, during which voids are heated to 10^4 K and the distribution is shifted towards lower Mach numbers. Subsequently, voids merge and cool adiabatically, resulting in a decrease of low Mach number shocks and a larger number of shocks with high Mach numbers. *Right panel:* Energy dissipation at shocks as a function of Mach number. At all redshifts, weak and strong shocks contribute similarly to the total thermalisation. This is in contrast to the dissipation in non-radiative runs, for which high Mach number shocks are less important. Consequently, the increased dissipation at the high Mach number end can be attributed to shocks created by feedback processes.

strong AGN feedback, the baryon fraction within clusters and groups at $z = 0$ is underpredicted compared to observations [Genel et al., 2014; Haider et al., 2016].

In Fig. 12.3 we show the cumulative shock surface (left-hand side panel) and the energy dissipation at shocks (right-hand side panel) as a function of Mach number for different redshifts. The shock surface is normalized by the simulation volume, and it has therefore units of an inverse length scale.

For $z \geq 2$ by far most of the shocks occur between voids and filaments of the cosmic web. The shock surface within haloes is insignificant compared to the total area of shocks, and shocks in the WHIM only contribute non-negligibly for $z < 2$ (see also Section 12.2). Therefore, the shock surface area is dominated by external shocks. The red curve ($z = 6$) indicates the distribution of shocks on the eve of reionization, which in Illustris happens almost instantaneously and is modelled by a uniform time-dependent ionising background following Faucher-Giguère et al. [2009].

The shock distribution is relatively broad and most of the shocks have Mach numbers around $\mathcal{M} = 10$. However, there are also very strong external shocks present with Mach numbers of up to $\mathcal{M} = 1000$. These highly supersonic flows are found between the coldest voids and strong filaments or halo outskirts. During the process of reionization the temperature in voids reaches $\sim 10^4$ K and the distribution is therefore shifted towards lower Mach numbers. Moreover, the total shock surface

area increases between $z = 6$ and $z = 5.5$ (given by the integrals of the curves) by around 7%. Thereafter, it decreases monotonically. We interpret this finding as the creation of a small population of newly created shock waves as the gas responds to the sudden heating during reionization. Note however that this mechanism may be strongly alleviated for a more realistic, gradual reionization history that is stretched out in time, similar to the models arising from radiative transfer methods applied to the Illustris simulation [Bauer et al., 2015].

After reionization ($z = 5.5$ to $z = 0$), the shock surface of low Mach number shocks progressively decreases, whereas the surface area of high Mach number shocks grows monotonically. This behaviour can be understood if one recalls that the shock surface statistic is dominated by accretion shocks from voids on to the cosmic web. As voids evolve they expand and merge into larger voids (see Fig. 12.1), which tends to decrease the abundance of low Mach number shocks. At the same time, their gas becomes colder with time [disregarding scenarios such as heating by TeV blazars, see Broderick et al., 2012; Pfrommer et al., 2012], giving rise to a stronger contrast between voids and the cosmic web, and therefore stronger shocks.

Interestingly, $\mathcal{M} \approx 4.2$ is a peculiar Mach number of the universe; the surface area of shocks with this strength stays roughly constant after reionization. At $z = 0$, the total shock surface density reaches a value of $8.5 \times 10^{-1} \text{ Mpc}^{-1}$ (integral of the blue curve). This number is around 3.4 times higher than what we have found for Illustris-NR-2 in Part II of the thesis. However, in the non-radiative run a large fraction of weak shocks on to the cosmic web is lost by the post-processing reionization modelling in form of a global temperature floor. We therefore argue that a comparison to Illustris-NR-2 without a temperature floor is more meaningful. In this case we find that in Illustris-1 the total shock surface area is about 1.4 times higher.

The energy dissipation at shocks as a function of Mach number is shown in the right-hand side panel of Fig. 12.3. This shock property is strongly dependent on the environment and high for regions with high densities ($f_{\text{th}} \propto \rho_1$). Although the shock Mach numbers decrease during the process of reionization, the energy dissipation increases significantly. This can be partly attributed to the triggering of new shocks, but also to the modification of existing ones. We expect that non-linear structures, such as filaments, dynamically respond to the heating by slightly expanding. In this way, the inflowing kinetic energy can be increased.

Throughout cosmic evolution, the shock distribution remains broad and a whole range of Mach numbers contribute comparably. For example, 50% of the thermalisation happens at shocks with strengths below $M \leq 6 - 10$, the other half in higher Mach number shocks. As opposed to the low Mach number shocks in the surface statistic, the low Mach number shocks in the energy statistic are shocks well inside non-linear structures, which we refer to as internal shocks. This distribution differs substantially from what is found in non-radiative simulations, where 50% of the dissipation occurs in $\mathcal{M} < 4$ shocks, and high Mach number shocks are present but do not contribute significantly to the total dissipation. Non-radiative simulations show a bimodality in the energy statistics, consisting of low Mach number internal

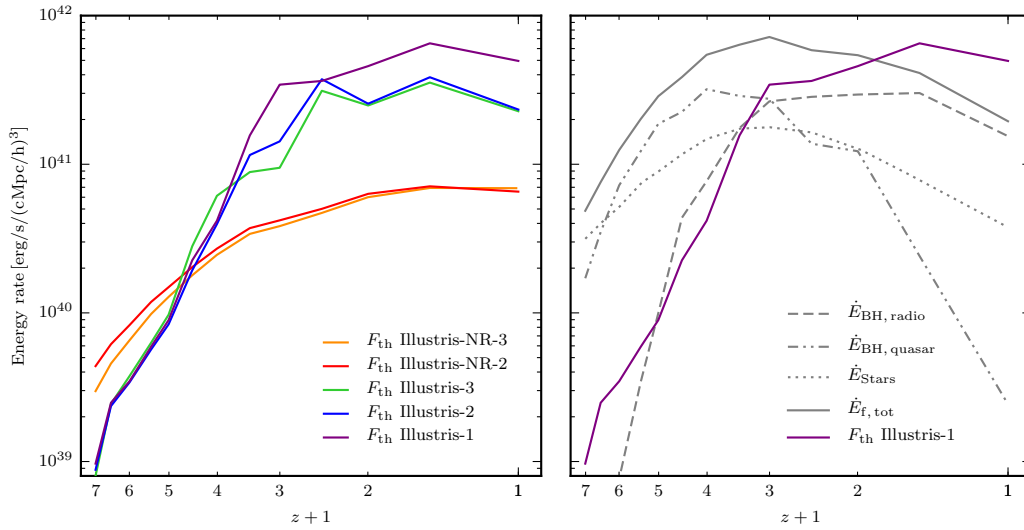


Figure 12.4: *Left panel:* Total energy dissipation at shocks as a function of redshift for the different Illustris runs. At late times, the thermalisation in the full physics runs is higher by a factor of around 4-8 compared to the non-radiative runs. Moreover, there is good agreement for the runs with resolution levels 2 and 3, whereas the full physics Illustris-1 run shows a stronger dissipation at low redshifts. This effect presumably originates from an increased black hole feedback activity in the highest resolution run. *Right panel:* Energy dissipation at shocks in the Illustris-1 simulation compared to the energy output of different feedback channels. For redshifts $z > 1$, the dissipation stays well below the total energy dumped into feedback processes of stars and black holes, indicating that not all feedback channels produce energetic shocks. On the other hand, there exists a strong correlation at redshifts $z < 4$ between the energy release by black holes in the radio-mode state and the energy rate dissipated at shocks.

shocks which dissipate most of the energy, and high Mach number external shocks processing cold primordial gas [Ryu et al., 2003; Vazza et al., 2011]. Since it can be a source of confusion we want to point out that the bimodality in the energy statistics of non-radiative simulations can not be directly mapped to the surface distribution of shocks. While high Mach number shocks are external shocks, most of the external shocks actually have low Mach numbers [see also Fig. 12 in Vazza et al., 2009a]. Moreover, the adopted modelling of reionization can erase weak shocks on to the cosmic web, hence a self-consistent interpretation can ultimately only be achieved if reionization is followed self-consistently during the simulation, as is the case for Illustris-1, albeit with a simplified reionization history.

In the dissipation statistics we measure for the full physics Illustris-1 run, the appearance of energetic high Mach number shocks prevents the development of a bimodality. This difference to the non-radiative simulations is striking, and as we will show later on, the additional shocks can be attributed to feedback processes.

In Fig. 12.4, we compare the total energy dissipation at shocks across cosmic time for the full physics and non-radiative runs at different numerical resolution. For these simulations the shock finder gives consistent results with respect to resolution. In

general, the global dissipation rate shows a shallow but strictly monotonic increase at early times before a saturation sets in. In the full physics runs, a strong increase in the total dissipation rate at shocks can be seen during reionization $z \approx 6$, which, as discussed above, presumably originates from the triggering of new shocks and the dynamical response of non-linear structures to the heating. The next four data points between $z = 5.5$ and $z = 4$ indicate a power-law behaviour, before the energy rate steepens and slightly diverges for the runs with different resolutions. We suggest that the latter feature marks the time when feedback processes from stars and black holes become important for the total energy dissipation at shocks.

From $z = 4$ to $z = 2$ the energy rate rises by more than an order of magnitude and bends sharply thereafter. This sharp bend coincides with the creation of the extended WHIM at $z = 2$ (recall Fig. 12.2), indicating that internal shocks are erased in this process. At $z = 0$, the simulations Illustris-1, Illustris-2, and Illustris-NR-2 reach final values of $1.7 \times 10^{41} \text{ erg s}^{-1} \text{ Mpc}^{-3}$, $8.1 \times 10^{40} \text{ erg s}^{-1} \text{ Mpc}^{-3}$, and $2.3 \times 10^{40} \text{ erg s}^{-1} \text{ Mpc}^{-3}$, respectively. In either case, due to the creation of feedback shocks the dissipation in the full physics runs at late times is higher by a factor of several compared to the non-radiative runs.

While the Illustris-3 and Illustris-2 runs give fairly similar results, the dissipation at shocks in the highest resolution run is considerably larger at low redshifts. As we show in Section 15, the difference originates from high Mach number shocks and correlates with a larger amount of feedback energy released in Illustris-1. Most importantly, between $z = 2$ and $z = 0.5$ the black hole radio mode feedback is stronger by a factor of around 2. This indicates that black hole accretion histories in Illustris are resolution dependent, and more gas can be funnelled to central regions if the gas dynamics are resolved with a higher resolution.

In order to put the values we measured for the energy dissipation at shocks in Illustris-1 into context, we compare the dissipation rates to other characteristic energy rates in the simulation. Specifically, we want to compare the kinetic energy dissipated at shocks with the energy output of black holes and stars. The energy release by AGN in Illustris-1 scales with the accretion rate \dot{M}_{BH} on to the black hole and is given by $\dot{E}_{\text{BH}} = \epsilon \epsilon_r \dot{M}_{\text{BH}} c^2$. Here, $\epsilon_r = 0.2$ is the radiative efficiency and $\epsilon \in \{\epsilon_m, \epsilon_f\}$ the efficiency with which the feedback couples to the surrounding gas. Depending on the accretion rate the BH is either in the radio feedback mode ($\epsilon = \epsilon_m = 0.35$) or quasar feedback mode ($\epsilon = \epsilon_f = 0.05$). The corresponding criteria are $\dot{M}_{\text{BH}} < \chi_{\text{radio}} \dot{M}_{\text{Edd}}$ and $\dot{M}_{\text{BH}} \geq \chi_{\text{radio}} \dot{M}_{\text{Edd}}$, respectively, where $\chi_{\text{radio}} = 0.05$ denotes the radio threshold and \dot{M}_{Edd} the Eddington accretion rate. Moreover, the injected supernovae feedback energy can be inferred by multiplying the cosmic star formation rate density with the SNII energy per stellar mass, which in Illustris-1 has the value of $\text{egy}_w = 1.09 \times 1.73 \times 10^{-2} \times 10^{51} \text{ erg}/M_{\odot}$ [Vogelsberger et al., 2014a].

In the right-hand panel of Fig. 12.4, we present the kinetic energy dissipated at shocks in Illustris-1 as well as the energy output of the different feedback channels. The grey lines indicate from top to bottom the black hole radio mode energy rate, the black hole quasar mode energy rate, the feedback energy rate dumped into stellar winds, and the sum of the feedback energy rates. Several interesting aspects

of the adopted galaxy formation model become apparent. First of all, due to the co-evolution of black holes and galaxies, the BH feedback increases at early times in parallel with an increasing supernovae energy rate. Later on, the black hole accretion rate declines, resulting in a decrease of the black hole quasar mode activity and an increase in radio mode activity. The latter trend coincides with an onsetting decline in the star formation rate density, highlighting the importance of the radio mode feedback for quenching star formation in the full physics runs.

For $z > 0.5$, the shock energy dissipation stays well below the total feedback energy, indicating that not all feedback channels drive energetic shocks. Especially the AGN quasar-mode feedback and supernova feedback offer more net energy than what is detected by shocks for $z > 3$. The former heats central regions of galaxies, and presumably much of its energy is radiated away in cooling processes before it can drive energetic shock waves. Galactic winds on the other hand are often launched inside hot haloes where the sound speed is high, making them unlikely to appear as supersonic flows, and in addition, they also suffer from strong radiative cooling losses. Nevertheless, as we will show in Section 13, there are some galaxies which show quasar mode feedback shocks and launch supersonic winds.

For redshifts $z < 4$, there is a remarkable correlation between the radio mode AGN feedback rate and the dissipation measured at shocks. We interpret this finding as an indication that this mode is not only able to drive high Mach number and energetic shocks, as seen in Fig. 12.1, but that these strong feedback blast waves also contribute significantly to the integrated energy statistics at late times.

Hence, it is the black hole feedback in the full physics runs which can explain a significant fraction of the overall higher energy dissipation compared to the non-radiative runs. Another source of additional energy dissipation at shocks in the full physics runs could be the presence of gas cooling via radiative processes. In this case, dense streams are formed as well as cold, concentrated gas blobs, which are difficult to disrupt and create ram-pressure shocks absent in the non-radiative simulations.

12.2 Environmental dependence of the shock statistics

In the previous section we have analysed the occurrence of shocks in the Illustris universe and the associated energy dissipation in a global sense as a function of Mach number and redshift. We are now turning to an investigation of the importance of different environments in the thermalisation processes during cosmic evolution. For this purpose, we define four disjoint regions based on the gas temperature T and overdensity $\delta_b = (\rho_b - \bar{\rho}_b)/\bar{\rho}_b$, where $\bar{\rho}_b$ is the mean baryon density of the Universe. Those environments are cold diffuse gas ($T < 10^5$ K and $\delta_b < 100$), hot diffuse gas ($T \geq 10^7$ K and $\delta_b < 100$), the WHIM ($10^5 \leq T/\text{K} < 10^7$ and $\delta_b < 100$)²,

² Some authors define the WHIM environment as $10^5 \leq T/\text{K} < 10^7$, independent of the gas overdensity. However, in the case of the full physics run our definition is very similar since by far most of the gas in this temperature range is located outside of haloes and therefore has overdensities below 100, as can be seen in Fig. 12.5.

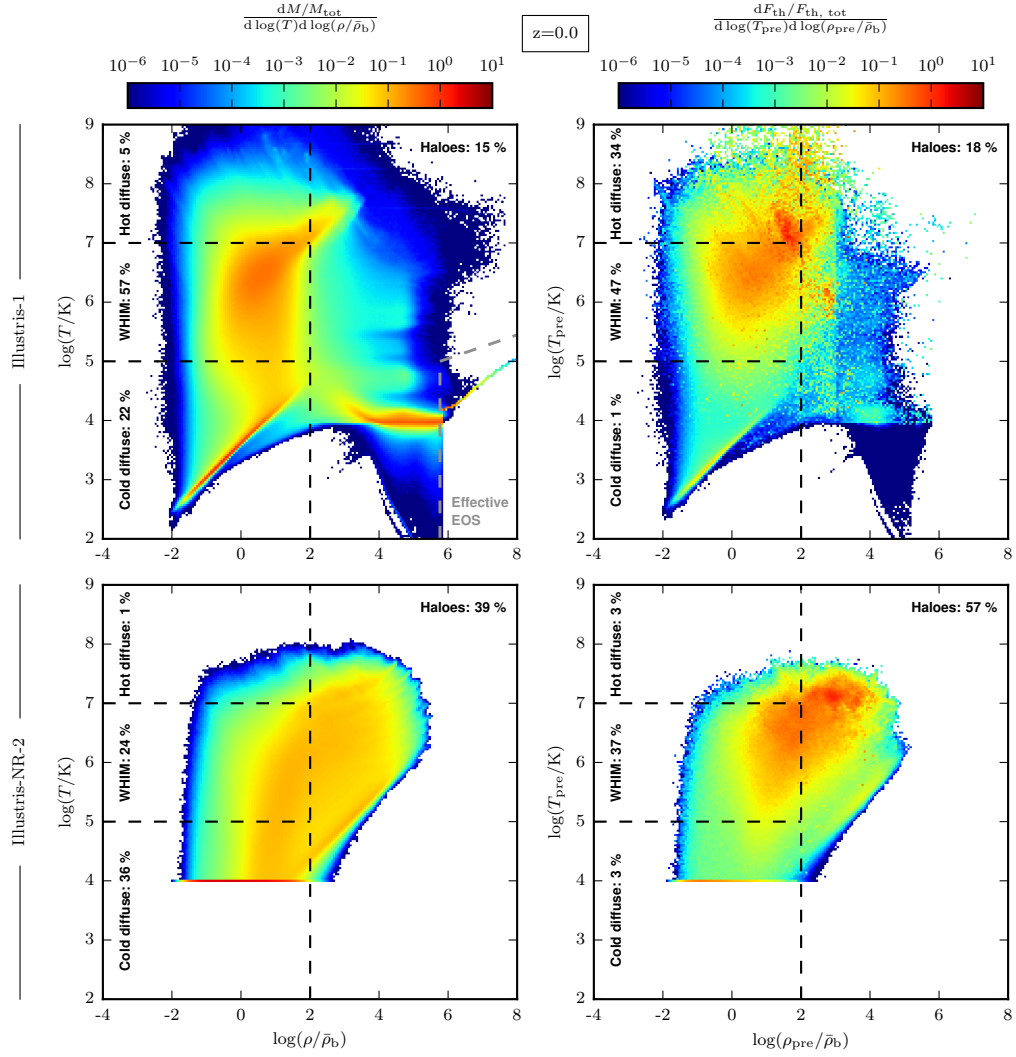


Figure 12.5: Gas density-temperature phase-space diagrams at redshift $z = 0$ weighted with gas mass (left-hand side panels) and energy dissipation at shocks (right-hand side panels). The latter are constructed with respect to pre-shock gas quantities. The top panels show results for the full physics Illustris-1 simulation, and the bottom panels for the non-radiative Illustris-NR-2 run. In the full physics run, black hole feedback processes transfer gas mass from haloes to the ambient medium, creating an extended WHIM region which contains around 60% of the gas mass at $z = 0$. Similarly, a large fraction of the energy dissipation is contributed by shocks in the WHIM, which dominate the dynamics of this gas phase. Remarkably, a considerable fraction of dissipation is also present in the hot diffuse phase, where relatively little mass resides.

and haloes ($\delta_b \geq 100$). Fig. 12.5 presents the distributions of gas mass (left-handed panels) and energy dissipation at shocks (right-handed panels) within the different regions at $z = 0$. The histograms that are weighted with energy dissipation are constructed with respect to pre-shock quantities, and all histograms are normalised to unity. The panels on the top and bottom show results obtained for the Illustris-1 and Illustris-NR-2 simulations, respectively. Note that for the non-radiative run, reionization is modelled in post-processing by simply adopting a temperature floor of 10^4 K. A large fraction of the mass and dissipated energy of the cold diffuse phase resides in the corresponding individual temperature bin.

Several features in the mass-weighted density-temperature phase diagram of Illustris-1 are apparent (top panel on the left-hand side). In the cold diffuse environment, most of the gas follows a mean equation of state, which is a power-law governed by the net effect of photoionisation-heating and adiabatic cooling due to the Hubble expansion [Hui and Gnedin, 1997]. Gas inside the WHIM is characterised by an increased temperature due to shock heating. A correlation can also be seen inside this phase [Davé et al., 2001], which is however less tight and exhibits more scatter. The latter observation is expected due to the variety in number, origin, and properties of shocks governing the dynamics of the WHIM.

Most of the gas above $\delta_b = 100$ is in a cool phase around 10^4 K within galaxies, where photoionization heating and radiative cooling are close to equilibrium [Vogelsberger et al., 2012]. Condensed gas above the star formation threshold (vertical grey dashed line) follows an effective equation of state, which represents a subgrid model for a pressurised ISM consisting of a hot and cold phase [Springel and Hernquist, 2003]. The bulk of gas mass in the Illustris simulation at redshift zero is contained in the WHIM ($\approx 60\%$), followed by the cold diffuse phase ($\approx 20\%$), and gas within haloes ($\approx 15\%$).

This gas mass distribution differs significantly from the distribution of the non-radiative Illustris-NR-2 run, in which $\approx 40\%$ of the gas is contained within haloes, and only $\approx 25\%$ inside the WHIM at $z = 0$. The difference arises mainly due to black hole feedback processes in Illustris-1, which transfer gas from haloes to the ambient medium, creating an extended WHIM phase. However, compared to observations the gas mass fractions inside massive haloes ($M \gtrsim 10^{13} M_\odot$) in Illustris-1 at the present epoch appears too low by a factor of several [Genel et al., 2014], indicating that the redistribution is too efficient. For a comprehensive analysis of the mass distribution in the Illustris simulation and comparison to previous work as well as observations, we refer the reader to Haider et al. [2016].

In the non-radiative simulations, for which radiative cooling and feedback is absent, the temperature evolution of the gas is dominated by shock heating on the one hand, and adiabatic cooling due to the Hubble flow on the other hand. Most of the energy dissipation at shocks at $z = 0$ takes place internally within haloes and filamentary structures, summing up to around 60% of the total dissipation. Those internal shocks typically have high pre-shock sound speeds and therefore low Mach numbers ($2 \lesssim \mathcal{M} \lesssim 4$). However, they dissipate strongly since the inflowing kinetic energy flux is large due to the involved high densities. The second relevant environment

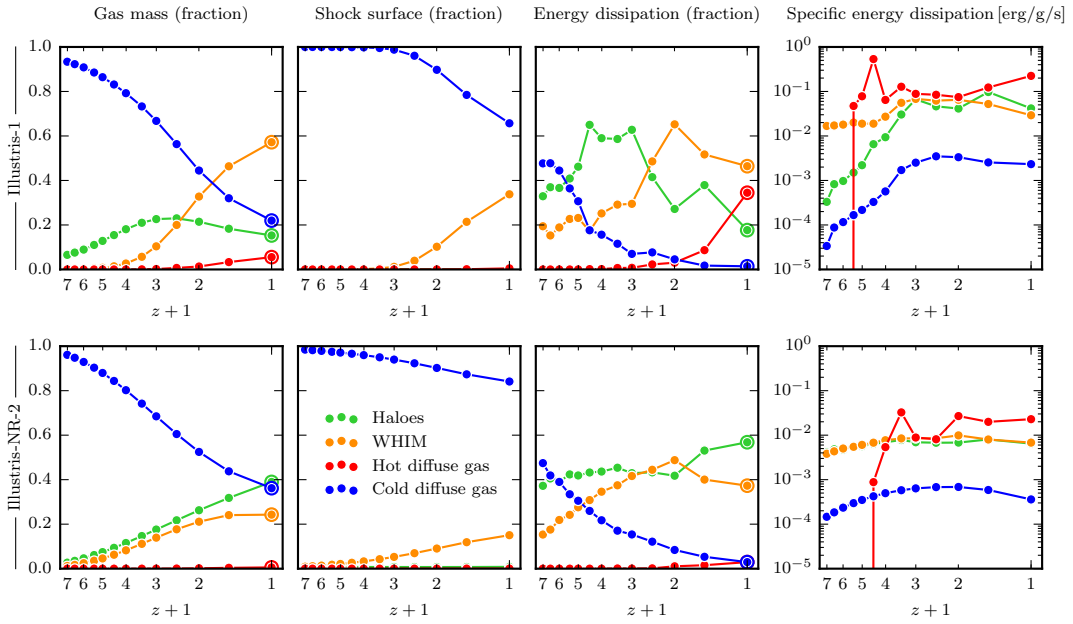


Figure 12.6: Time evolution of the contribution of different environments to, from left to right, the gas mass fraction, the shock surface fraction, the dissipation at shocks, and the dissipation per unit mass. In the full physics as well as in the non-radiative run most of the shock surface area resides in the cold diffuse phase, especially at high redshifts. Between $z = 3.5$ and $z = 2$, shocks inside haloes play the most important role in thermalising energy in Illustris-1, and their relative contribution is significantly higher compared to the non-radiative run. This clearly demonstrates the impact of non-radiative physics, which is most important in this environment. For $z < 1.5$, most of the energy is dissipated in the extended WHIM phase and even the hot diffuse phase contributes significantly. The fact that for the latter the shock surface area is very small indicates that this environment is created and heated by individual highly energetic shocks. Interestingly, except for the cold diffuse phase, the different environments have a similar and roughly constant specific energy dissipation for $z < 3$.

is the WHIM, where merger and accretion shocks contribute almost 40% to the total thermalisation by shocks. A more extended discussion of shock statistics and morphologies for the Illustris-NR-2 simulation can be found in Part II.

We note that for the full physics runs we exclude cells from our shock analysis which are in a region around the effective equation of state, as indicated by the grey dashed lines in the mass-weighted histogram. Moreover, we filter inconsistent jumps above $\delta_b = 1000$, and this density threshold can be seen slightly as an edge in the energy dissipation weighted histogram. Its visibility depends on the choice for the parameter f of our shock finding algorithm, however, as will be shown in Section 14, the choice does not influence the statistics significantly.

In the Illustris-1 simulation, about 80% of the dissipated shock energy is found inside the WHIM and the hot diffuse phase. Compared to the non-radiative run, differences in the WHIM potentially arise from photoionisation heating, radiative cooling, and feedback processes from stars and black holes. However, heating due

to photoionisation increases the gas temperature only to around 10^4 K, which is well below the temperature of the WHIM. Moreover, radiative cooling plays only a subdominant role in the evolution of the WHIM [Davé et al., 2001], since the bulk of the gas has low density. Supernova feedback can potentially dump thermal energy into these gas phases, depending on the implementation of this feedback channel and the numerical setup. In Illustris-1, winds are launched in star-forming regions within galaxies and couple to the gas in the near vicinity. These winds can drive shocks and increase the thermal energy of gas within the halo, as shown in Section 13. However, they typically do not reach the virial radius. Consequently, the influence of supernova feedback on the WHIM in Illustris-1 can be neglected. We therefore argue that the increased importance of the WHIM and the hot diffuse phase in thermalising kinetic energy originates from black hole feedback.

Compared to these phases, the cold diffuse phase is not important in this respect due to its low density. The dissipation within haloes amounts to 15%, which is low compared to the non-radiative run. Note however that in Illustris-1 the overall energy dissipation at late times is higher by a factor of around 8, so compared to the non-radiative simulation the total energy dissipation within haloes is higher.

In Figure 12.6 we compare several properties of the different environments across cosmic time. From left to right we show the gas mass fraction, the shock surface fraction, the energy dissipation rate, as well as the specific dissipation rate. For the shock related quantities the environments are defined with respect to the pre-shock region, and the circled data points at $z = 0$ correspond to the integrated phase diagrams of Fig. 12.5. Results for the full physics run are shown in the upper panels, while the lower panels indicate results obtained for the non-radiative run.

As expected, the gas mass fractions in the non-radiative run show a monotonic evolution, in which mass is transferred from low to high density regions. In contrast, black hole feedback redistributes baryons in Illustris-1 from haloes to the WHIM and the hot diffuse phase, decreasing the gas mass fraction of haloes at late times to a value below 20%. Moreover, in the full physics run a steeper decrease of the mass in the cold phase is present for $z < 2$. This can be explained by the fact that the WHIM does not only grow in mass, but also becomes more spatially extended, as we have seen in Fig. 12.2.

The shock surface fraction clearly demonstrates that in both kinds of simulations the shock heating of gas in the cold diffuse phase produces the largest contribution to the total shock surface area. Most of these shocks are external shocks, created when pristine gas from voids accretes on to non-linear structures. Only for $z < 2$ does the WHIM contribute a significant fraction of shock surface area, and furthermore, the cumulative surface area of shocks in haloes and in the hot phase are negligible in comparison.

While measuring the shock surface only reveals the locations of shocks, the energy dissipation quantifies the impact on the thermal history of the gas. Comparing Illustris-1 to Illustris-NR-2 in this respect reveals that between redshifts $z = 3.5$ and $z = 2$ a higher fraction of kinetic energy is thermalised within haloes in the full physics run. As can be seen from the left panels, this correlates with a higher halo

gas mass fraction in Illustris-1 at early times due to radiative cooling. With more available mass and denser streams more energy can potentially be dissipated. On the other hand, the increase within haloes is abrupt and also coincides with the onset of the black hole radio mode feedback at $z \approx 4 - 3$. Due to the subsequent mass transfer and expansion of the WHIM, a higher fraction of energy gets dissipated in this environment. Interestingly, shocks in the hot diffuse phase also contribute significantly, although there is only a small increase in the mass fraction and a negligible increase in the shock surface area in this environment. This suggests heating through few but very energetic shocks, pointing once more towards the radio-mode blast waves launched by black holes.

In the panels on the right-hand side we show the specific energy dissipation, which is the dissipation rate per unit mass. Remarkably, apart from the cold diffuse phase, all environments have within a factor of several similar and roughly constant rates for $z < 3$. This finding is non-trivial, since although the energy dissipation is proportional to the pre-shock density of the inflowing gas, there is also a strong dependence on the sound speed and the Mach number. The specific dissipation rate in the non-radiative run is around $10^{-2} \text{ erg g}^{-1} \text{ s}^{-1}$, and for Illustris-1, we measure a ten times higher value of approximately $10^{-1} \text{ erg g}^{-1} \text{ s}^{-1}$. In order to put these numbers in context we compare them to the power released in supernovae explosions with respect to the mean baryon density of the universe. This rate can be calculated for Illustris-1 as the product of the star formation rate density and the SNII energy per stellar mass, $P = \text{SFRD} \times \text{egy}_w \approx 0.1 \text{ M}_\odot/\text{yr}/\text{cMpc}^3 \times 1.09 \times 1.73 \times 10^{-2} \times 10^{51} \text{ erg}/\text{M}_\odot$, so that we obtain for the specific power a value of $P_s = P/\bar{\rho}_b = 3.2 \times 10^{38} \text{ erg yr}^{-1} \text{ M}_\odot^{-1} = 5.0 \times 10^{-3} \text{ erg g}^{-1} \text{ s}^{-1}$. Hence, in Illustris-1, the energy rate per unit mass of shocks is higher by a factor of 20 compared to the specific energy rate of stellar feedback.

13

Shock morphologies across cosmic time

The aim of this section is to discuss the different shock morphologies found in the Illustris-1 simulation. We mostly present a qualitative overview, with the intention to stimulate a more comprehensive analysis of some of the rich phenomenology exemplified here in future work.

13.1 High-redshift accretion shocks

We begin with an investigation of the morphology of accretion shocks. Figure 13.1 shows the shock environment around some of the biggest haloes in the simulation at $z = 4$. As we have seen in Section 12.1, the violent black hole radio mode feedback, which operates during low accretion rates, is still largely absent at this time. Moreover, we do not expect many shocks from the quasar mode feedback, and the stellar feedback operates on smaller scales compared to the one under consideration. We have checked the environment of the same haloes in the non-radiative Illustris-NR-2 run (not shown), and find shocks with very similar Mach numbers and locations. Consequently, the shocks seen in Fig. 13.1 are indeed accretion shocks.

The importance and character of shocks in the formation of gaseous haloes as well as their role in galaxy formation have been a lasting topic in the literature. In the classical two-stage galaxy formation theory [White and Rees, 1978] infalling gas gets shock-heated to the virial temperature before it can cool and settle into a disk, forming stars inside out. The pressurised atmosphere of the gaseous halo can either be created due to thermalisation in a strong accretion shock, or through a succession of smaller shocks inside the halo. The scenario in which the gas gets shock heated to the virial temperature before it cools and fragments into stars is termed the ‘hot/slow’ mode of gas accretion. If the cooling time is much shorter than the dynamical time, which is the case for low mass galaxies, gas can also be accreted in

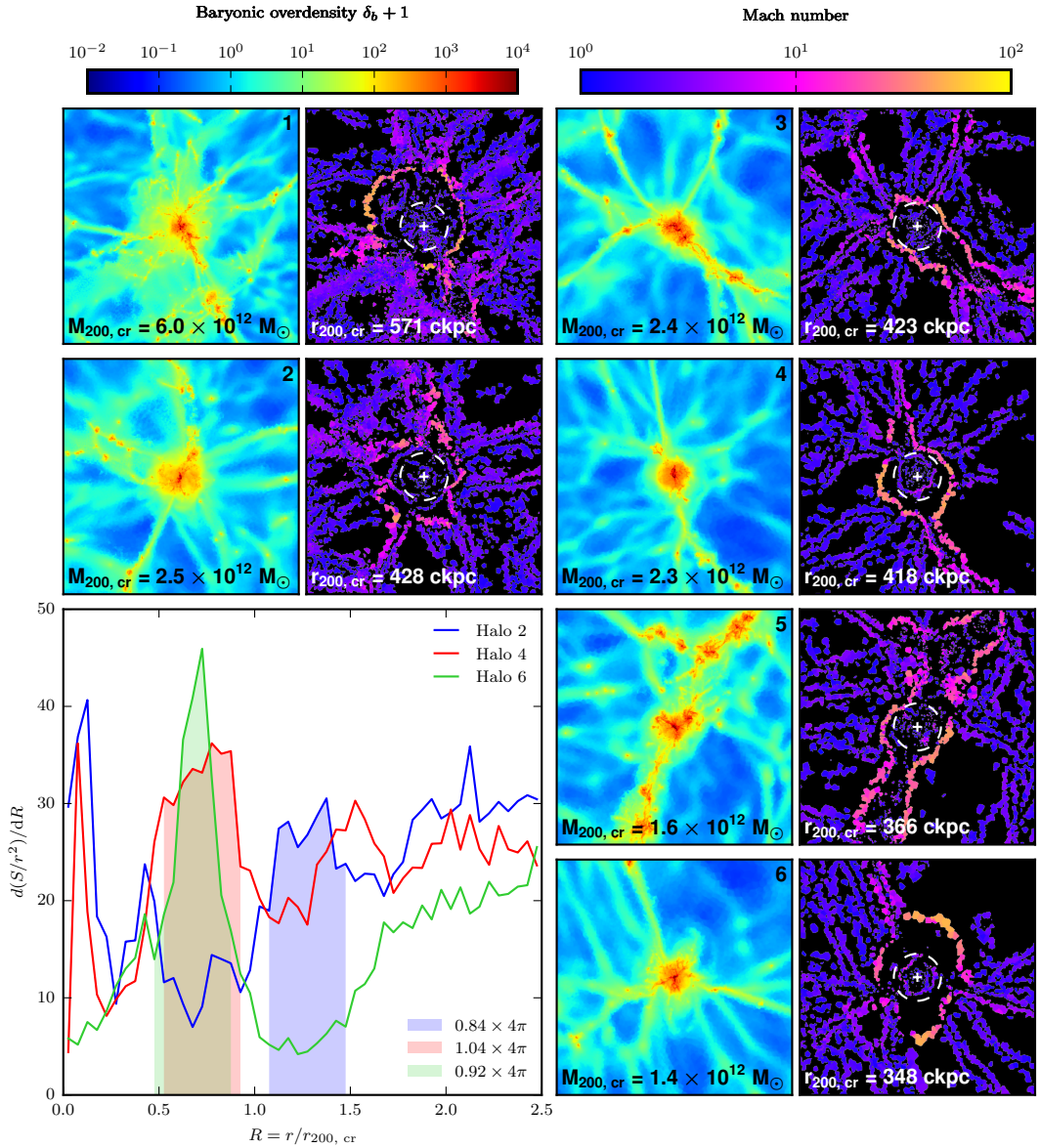


Figure 13.1: Shock environments around some of the largest haloes in the Illustris-1 simulation at $z = 4$. Shown are thin projections of the baryonic overdensity and the Mach number, with projection depths of $0.5 \times r_{200, cr}$ and $0.1 \times r_{200, cr}$, respectively. The white circles correspond to the virial radii. In all haloes strong accretion shocks can be seen, which are typically located outside of the virial radius. Moreover, these outer accretion shocks can get penetrated by filaments, depending on their size and hydrodynamic properties. Interestingly, in several cases a second inner accretion shock is formed, especially when the flow inside the penetrating filament is smooth (haloes 2, 4, and 6). The lower left panel shows the shock surface distribution as a function of radius for the corresponding haloes. In all three cases, the inner accretion shock covers a solid angle of approximately 4π , indicating that they are spherical.

a ‘cold/rapid’ mode without being shocked [Birnboim and Dekel, 2003; Fardal et al., 2001], something that is also present in the original theory of hierarchical galaxy formation [White and Frenk, 1991]. Furthermore, there may exist a third mode, in which an outer accretion shock gets penetrated by cold filaments, feeding the central galaxy with cold and unshocked gas [e.g. Dekel and Birnboim, 2006; van de Voort and Schaye, 2012].

The haloes shown in Fig. 13.1 have masses well above $M_{200, \text{cr}} > 10^{12} M_{\odot}$ and are among the most massive haloes at $z = 4$. The specific haloes have been selected for illustrating the wide variety of occurring shock features. For haloes in this mass range, theoretical models predict the existence of stable accretion shocks [Dekel and Birnboim, 2006]. The distance of the outer accretion shock from the halo centre as inferred from non-radiative simulations is on average around $1.3 \times r_{200, \text{cr}}$ (see Part II). However, the accretion shock can sometimes also be found significantly further outside ($\lesssim 2 \times r_{200, \text{cr}}$), especially if the halo is in a non-equilibrium state after a major merger event [Nelson et al., 2015a].

For all haloes in Fig. 13.1, the presence of an outer accretion shock is evident, with varying Mach number depending on the environment. The accretion shock is stronger if gas is infalling directly from voids, as it is the case for halo 4 in the bottom-right, and halo 6 in the top-right direction, with Mach numbers up to $\mathcal{M} = 100$. The corresponding accretion flows are highly supersonic due to their low sound speed. On the other hand, if the halo is surrounded by weak filaments whose gas is shock heated at the outer accretion shock, the typical Mach numbers are significantly lower (e.g. halo 2, $\mathcal{M} \approx 10$). In several locations, we see that strong filaments are able to penetrate the outer accretion shock. This preferentially happens for filaments with high gas densities and velocities, resulting in a large ram pressure $p_{\text{r}} \propto \rho v^2$.

Interestingly, although the masses of the haloes are similar, we find a considerable variety among their interior shock morphologies. Halo 1 dissipates thermal energy internally by a network of complex weak shocks. This network is likely induced by recent merger events, since it also coincides with an increased outer accretion shock pointing towards a non-equilibrium system. Strikingly, in haloes 2, 4, and 6, a second accretion shock located inside of the outer accretion shock can be seen. These shocks can potentially be formed if there is cold gas from filaments crossing the outer accretion shock unshocked.

The lower left panel of Fig. 13.1 shows the radial shock surface distribution of these haloes, and for each halo the inner accretion shock can clearly be seen as a peak in the distribution. Moreover, by integrating the peaks one obtains the solid angles covered by shocks at the corresponding radii. We integrate each curve at the location of the inner accretion shock over an interval of size $0.45 R$, which accounts for the deviation from perfect sphericity of the shocks. The integrals of the inner accretion shocks of haloes 2, 4, and 6 amount to $0.84 \times 4\pi$, $1.04 \times 4\pi$, and $0.92 \times 4\pi$, respectively, indicating that they are largely spherical.

Haloes 3 and 5 are also penetrated by strong filaments, but in these cases no inner spherical accretion shocks can be found. On the other hand, these haloes experience

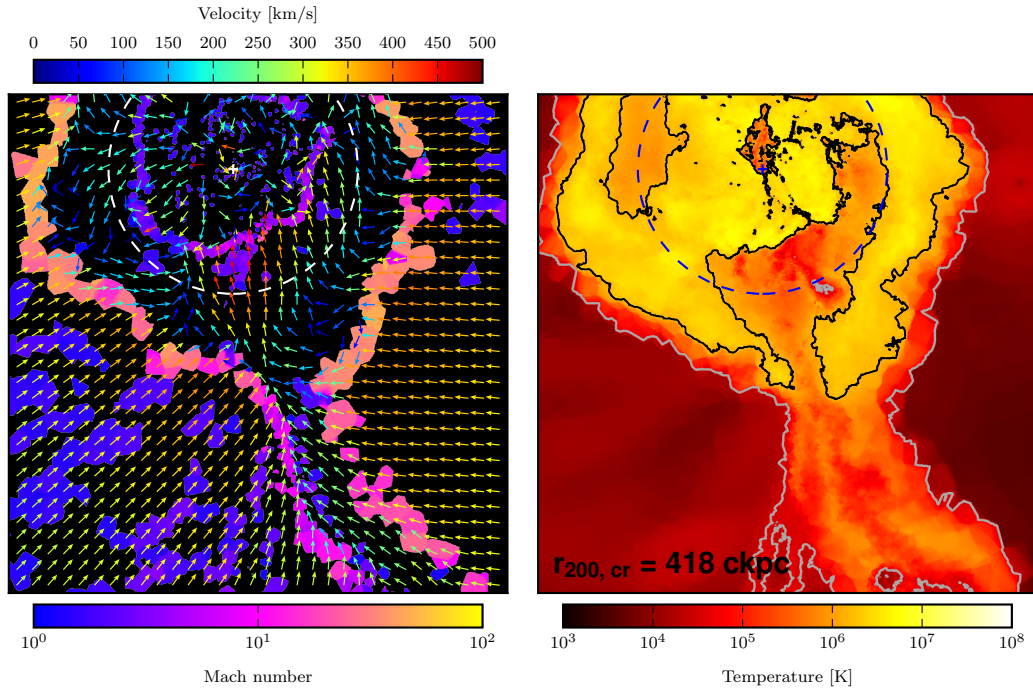


Figure 13.2: Zoom on to a halo with a double accretion shock (halo 4 from Fig. 13.1). The left-hand panel shows a thin projection of the Mach number field, as well as the direction and magnitude of the peculiar gas velocity field (coloured arrows). On the right-hand side panel, the mass-weighted mean temperature is projected, together with 2D contours at $T = 3 \times 10^4$ K (grey) and $T = 1.5 \times 10^6$ K (black) for visual guidance. The dashed circles correspond to the virial radius, $r_{200, \text{cr}} = 418$ ckpc. A vertical filament penetrates the outer accretion shock and channels cold gas into the interior. The cold stream coming from the bottom is accelerated and heated at a smaller radius by a strong shock ($\mathcal{M} \approx 10$). Moreover, since the cold phase mixes with gas shock heated at the outer accretion shock, a roughly spherical inner accretion shock is formed.

a more clumpy accretion flow compared to halo 4 and 6, which might prevent the formation of the inner spherical shock.

The existence of inner and outer accretion shocks is consistent with recent work by Nelson et al. [2015a], who analysed gas accretion on to $10^{12} M_{\odot}$ haloes at $z = 2$ by means of cosmological hydrodynamic zoom simulations¹. In addition to a strong virial shock at around $1.3 \times r_{200, \text{cr}}$, they report for their analysed haloes that inflows which are not shocked at the virialization boundary are significantly heated inside the virial radius. This happens at a distinct distance from the centre, typically at $\lesssim 0.5 \times r_{200, \text{cr}}$, and only thereafter the gas cools and accretes on to the central galaxy in the examined systems.

Fig. 13.2 shows a zoom on to halo 4 that focuses on projections of the Mach number and temperature field. Additionally, in the left-hand panel the direction and

¹Although this study also uses the AREPO code, very different simulation physics have been used compared to the Illustris runs.

magnitude of the peculiar gas velocity is visualised by the orientations and colors of arrows, respectively. The halo has a virial mass of $M_{200, \text{cr}} = 2.3 \times 10^{12} M_{\odot}$ and is surrounded by a strong outer accretion shock with Mach numbers of $\mathcal{M} = 40 - 50$. Such shock strengths are expected for relatively cool accretion flows from voids and weak filaments. Moreover, gas inside the highlighted filament is dense and fast enough to penetrate this outer shock and thermalise further in, where the halo pressure is higher. The accretion flow from the filament gets accelerated towards the potential minimum and reaches a velocity of around 500 km s^{-1} relative to the halo center. With a pre-shock temperature of 10^5 K and assuming full ionisation, this configuration corresponds roughly to a Mach number 10 shock, consistent with the value obtained by the shock finder. Moreover, given the location inside the virial radius, this shock can be considered very strong.

Interestingly, while the filament crosses the outer shock in the vertical direction, there are also inner accretion shocks in perpendicular directions, forming as a whole an approximately spherical inner accretion shock. This phenomenon can be understood by investigating the temperature map of the halo, which is shown in the right-hand panel of Fig. 13.2. The grey contour line at $T = 3 \times 10^4 \text{ K}$ corresponds roughly to the pre-shock temperature of the outer accretion shock, and the black contour line at $T = 1.5 \times 10^6 \text{ K}$ highlights the inner temperature structure of the halo. It can be seen that the cold stream from the filament does not only create a bow shock, it also flows around it and mixes with gas shock-heated at the outer accretion shock. In this way, the temperature and sound speed of the latter is decreased such that it gets shocked a second time in order to reach the virial temperature. The black contour line also indicates that the left part of the inner accretion shock forms mainly due to a cooling stream from the top; as can be seen in Fig. 13.1 there is a counterpart of the highlighted filament.

This discussion raises the question whether the interpreted shock surfaces change with simulation resolution, and to which degree they are modified by radiative physics. We show in Section 15 the halo of Fig. 13.2 for different resolutions and find that a very similar inner accretion shock is obtained in the Illustris-2 run. Moreover, it is also present in the non-radiative simulation Illustris-NR-2, although in this case the inner shock is less spherical and the shape is dominated by the two bow shocks in the directions of the filaments.

It would be instructive to infer the abundance of the double accretion shocks, ideally in an automated way. However, this proves to be difficult, since a very similar signature in the differential shock surface distribution is created by ellipsoidal accretion shocks, which peak at the radial positions of the semi-principal axes (see Fig. 6, Halo 8 in Part II). Moreover, the gas resolution in smaller mass haloes is too low at $z = 4$ in Illustris-1 for this kind of analysis. Nevertheless, it would be interesting in future simulations to shed light on the questions if, when, and to which extent gas inside virialized structures of different masses got shock-heated. This could be accomplished in zoom simulations by running the shock finder on the fly and recording a shock history for every gas parcel, a natural extension of a temperature

history based analysis by means of Monte Carlo tracer particles [Genel et al., 2013; Nelson et al., 2013].

13.2 Galaxy and galaxy cluster shocks

In Fig. 13.3 we show the shock environment around some of the biggest haloes in the simulation at $z = 2$. We have chosen these specific haloes in order to showcase the different occurring shock morphologies, and they are therefore not necessarily representative for a typical system. The cluster at the top accretes streams of substructures which induce a complex network of low Mach number shocks in the cluster interior. Due to the high central densities they are very effective in thermalising kinetic energy, as can be seen in the top panel on the right-hand side. Moreover, these merger shocks in combination with the accretion shocks at the outskirts are thought to be one of the main drivers of intracluster turbulence [e.g. Dolag et al., 2005; Miniati, 2014, 2015; Vazza et al., 2009b].

Merging clusters often emit unpolarised diffuse synchrotron emission over large scales, observed as giant radio haloes [Ferretti et al., 2012; Ferrari et al., 2008]. These observations indicate the presence of magnetic fields and relativistic electrons. The latter can acquire their energy via reacceleration due to MHD turbulence, or have a hadronic origin through creation in the decay chain of accelerated protons colliding with thermal protons (Enßlin et al. 2011; for a recent review see Brunetti and Jones 2014, and references therein). On the other hand, directly accelerating electrons from the thermal pool via second-order Fermi acceleration proves to be difficult from a theoretical point of view, since the required acceleration time is long. In any case, if only a small fraction of the energy dissipation at cosmological shocks can be used for particle acceleration, an energetically significant cosmic ray population is expected in the ICM [Pfrommer et al., 2007].

The galaxy cluster in the middle panel recently launched an AGN feedback shock, in the form of a thermal pressure dominated Sedov-Taylor like blast wave. The Mach numbers across the shock surface are fairly high, considering also the fact that it has not yet escaped the hot halo atmosphere. Moreover, this shock is highly energetic at its current location. We have measured the total energy dissipation inside the box $[4 \times r_{200, \text{cr}}]^3$ and find that it is around 2 times higher compared to the dissipation rate in the cluster at the top, and around 9 times higher than the dissipation of the cluster shown at the bottom of the figure. For black holes remaining predominantly in the radio mode, these kinds of feedback shocks are launched relatively frequently. We hence infer that these blast waves make an important contribution to the high value of the energy dissipation rate in haloes between $z = 3.5$ and $z = 2$.

The bottom panels show a galaxy cluster with corrugated shock surfaces. The surfaces are typically located where the ram pressure of the inflow matches the thermal pressure of the cluster. In this way, for homogeneous inflows, the shock locations trace equipotential lines of the cluster. In the displayed halo, three major filamentary streams are present, and the accretion shock shows a concavity in their

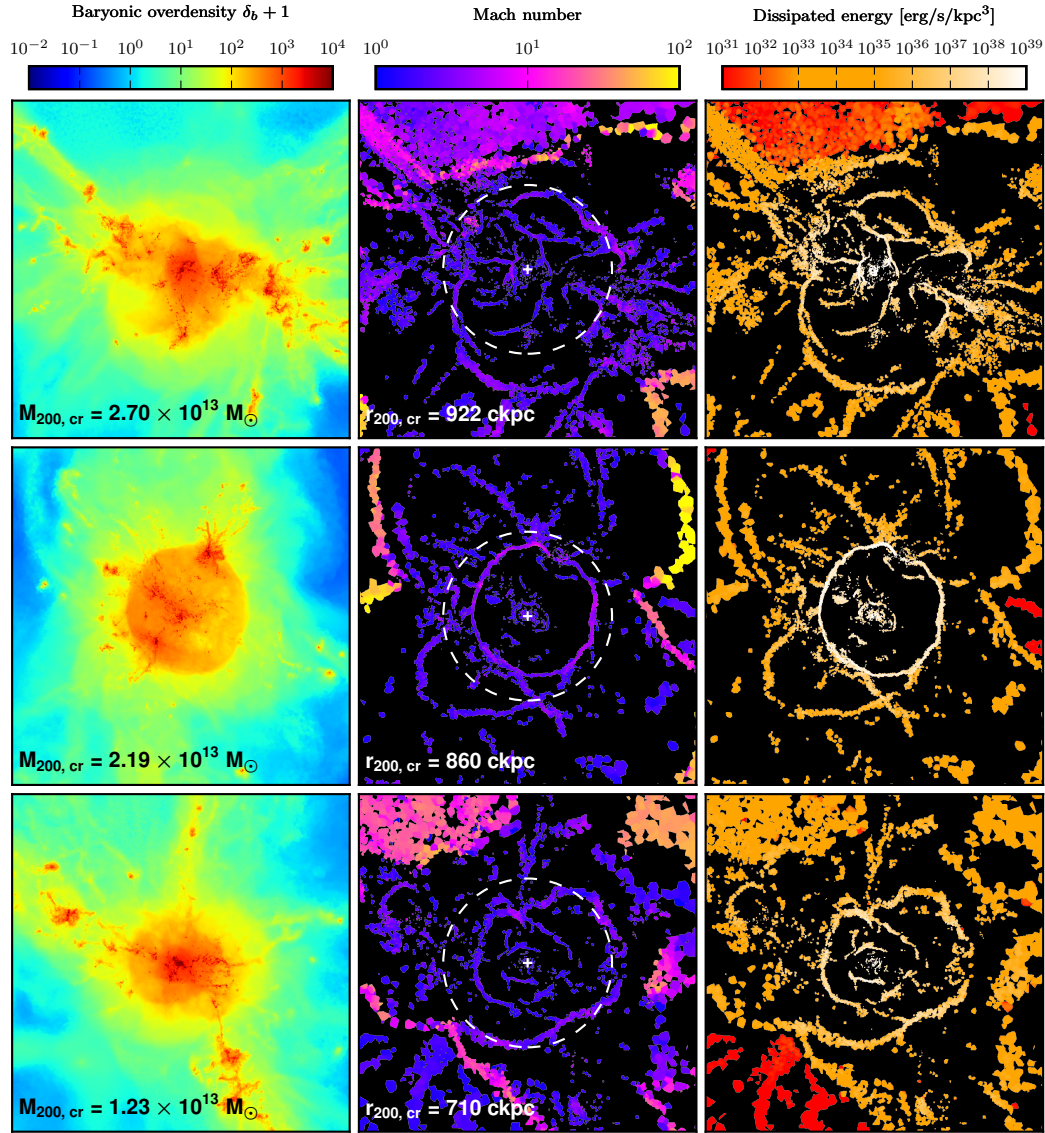


Figure 13.3: Zoom on to some of the most massive galaxy clusters at $z = 2$ in Illustris-1. Shown are projections of the baryonic overdensity, the Mach number, and the energy dissipation, with depths of $0.5 \times r_{200, \text{cr}}$ for the density, and $0.1 \times r_{200, \text{cr}}$ for the shock related quantities. In the halo of the first row, merger activities can be seen which induce a complex network of internal shocks. These shocks as well as the stimulated supersonic turbulence inside the cluster are potential cosmic ray particle accelerators. The black hole of the halo displayed in the middle row recently released a large amount of thermal energy, resulting in an energetic Sedov-Taylor like blast wave feedback shock. The halo at the bottom presumably underwent a major merger recently, as revealed by the arc-shaped merger remnants opposite to each other. Interestingly, we detect strong shocks on to cosmic sheets around the haloes at the top and at the bottom.

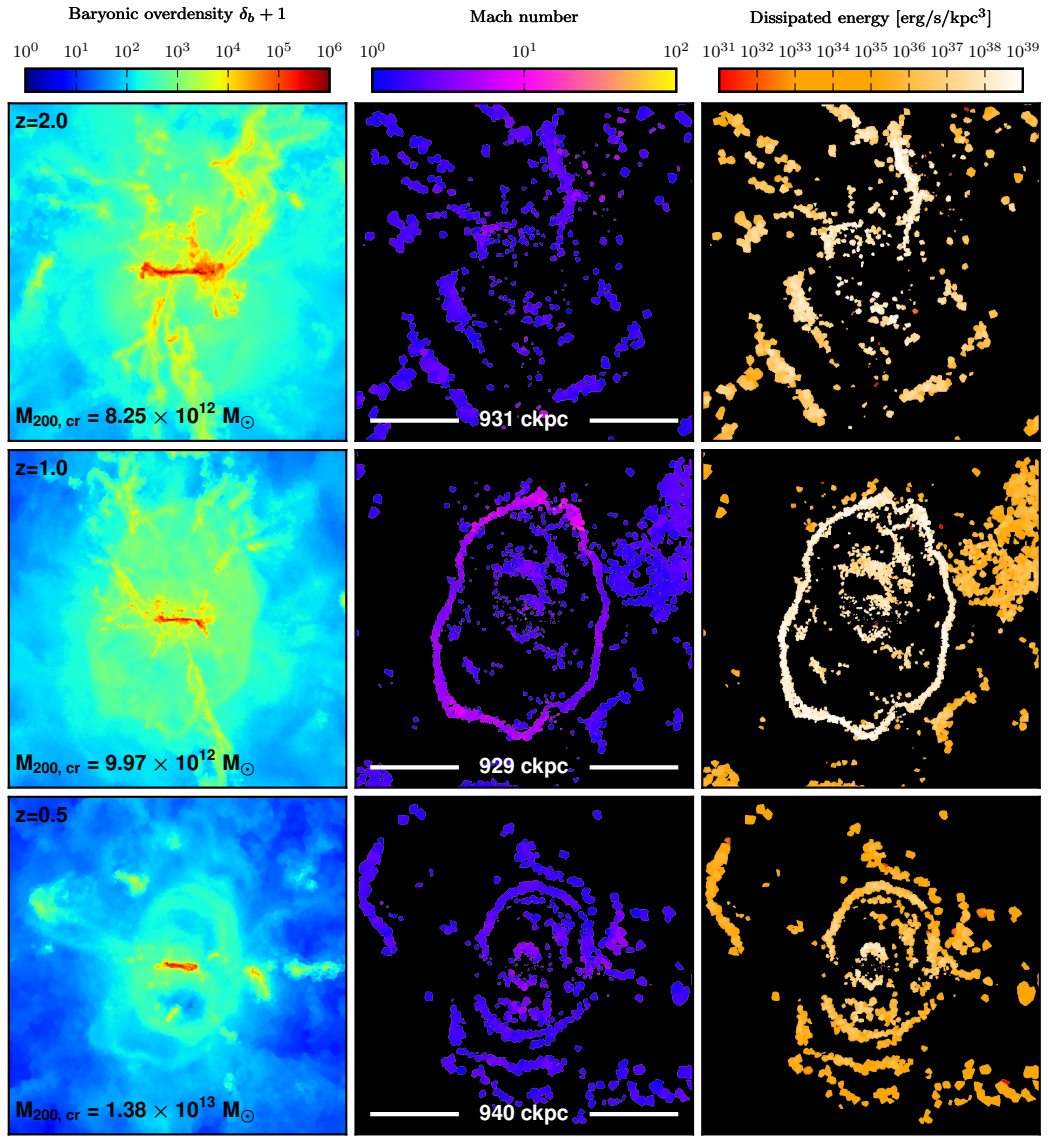


Figure 13.4: Zoom on to disc galaxies at three different redshifts. The projections show a square with a size of $1.5 \times r_{200, cr}$ per side and have a depth of $0.08 \times r_{200, cr}$. Different morphologies can be seen with, from top to bottom, complex inflow and outflow patterns, recent black hole activity, and supersonic stellar winds. Also on these scales the most energetic shocks are provided by AGN.

direction. Interestingly, the filaments at the upper left span a cosmic sheet, and we detect high Mach number accretion shocks at this location. A similar morphology can be seen for the halo in the top panels. Consequently, the large area of Zeldovich pancakes can provide a very interesting site for particle acceleration near clusters. In the central region of the galaxy cluster, merger remnants with positions opposite to each other are visible; we assume that this halo underwent a major merger event very recently.

In the local Universe, synchrotron radiation originating from merger remnants can be observed in the form of radio relics. Considering the strength and dissipation rate of the AGN feedback shock, we argue that this shock should also be observable. Up to now, most of the radio haloes and relics are observed at redshifts $z \lesssim 0.5$ [Feretti et al., 2012], however, this may largely be due to present sensitivity limits. But as we show in more detail below, in Illustris AGN feedback shocks are also created for lower redshifts, up to the present epoch. Since the shocks appear strong enough to create strong synchrotron emission, the lack of clear observational detections of such signals indicates that the radio mode feedback is likely too strong in Illustris-1. This argument shows that shock properties could be used in the future to constrain or validate the modelling and parameter space of feedback models used in state-of-the-art cosmological simulations.

The Illustris simulation contains around 40000 well-resolved galaxies with an encouragingly realistic mix of morphologies and kinematic properties [Genel et al., 2014; Snyder et al., 2015; Vogelsberger et al., 2014b,c]. In Fig. 13.4, we show zoomed in view of three late-type systems with masses of around $10^{13} M_{\odot}$ at, from top to bottom, redshifts $z = 2$, $z = 1$, and $z = 0.5$, in order to investigate shock signatures around such systems. The displayed images have a length of $1.5 \times r_{200, \text{cr}}$ on a side, i.e. the virial radius corresponds roughly to the distance between the corners and the centers of the images.

The spiral galaxy at $z = 2$ interacts with the circumgalactic medium (CGM) via a complex pattern of inflows and outflows, resulting in several shocks inside the hot halo. However, this example demonstrates the limitation of our shock finder in combination with Illustris. Poorly resolved gradients in the outskirts of the galaxy result in patches of shocks, rather than in well-defined shock surfaces. Moreover, since we filtered out detections of shocks near star-forming regions that are governed by the subgrid model, it is not possible to see shocks within the galaxy, for example around spiral arms.

Recent black hole activity is apparent for the galaxy at $z = 1$. Given the shock morphology we suspect that in this case the feedback originates from the quasar mode. The shape of the shock is less spherical compared to the blast waves created by the radio-mode feedback, and moreover, it has apparently been launched very centrally. This is unlikely to happen in the case of the radio mode feedback, for which a random injection within a sphere around the black hole is adopted. Nevertheless, the energy dissipation at this shock is exceedingly high, indicating that also on galactic scales black hole feedback drives the most energetic shocks. This is especially expected when the black hole reaches the end stages of a phase of exponential growth (during

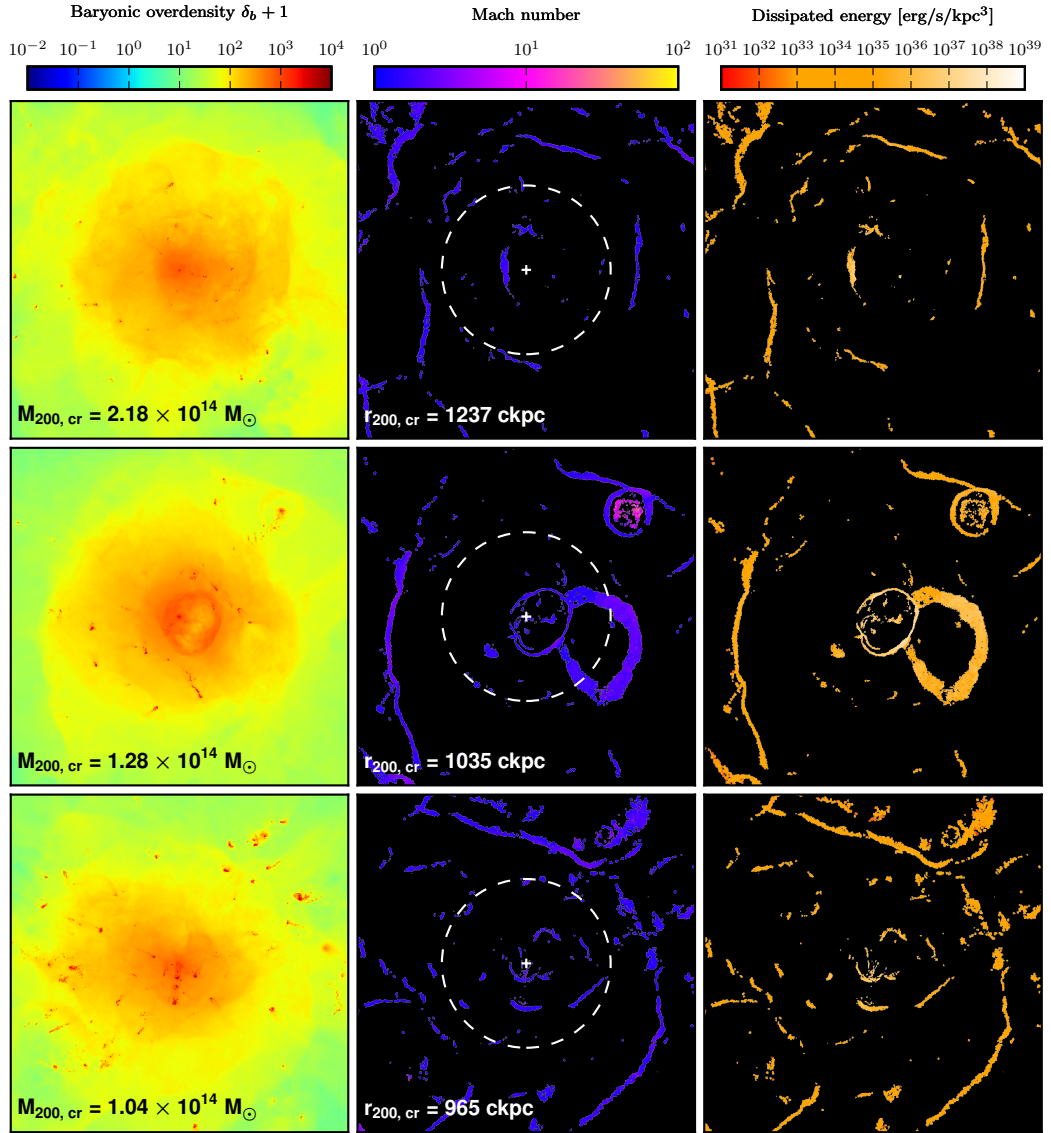


Figure 13.5: Projections of some of the most massive haloes at $z = 0$, with depths of $0.5 \times r_{200, cr}$ and $0.1 \times r_{200, cr}$ for the overdensity and the shock-related quantities, respectively. Due to black hole feedback the gas of these haloes has been substantially diluted, and most of the present shocks are remnants of this process. However, as can be seen in the middle panels, new black hole feedback shocks are still created at the present epoch. In the halo at the bottom the accretion of numerous substructures can be seen, giving rise to bow-shaped shocks.

which it shines as a quasar) and shuts down its own growth through energy feedback [Di Matteo et al., 2005].

Interestingly, the galaxy at $z = 0.5$ launches supersonic galactic winds, and we are able to detect the associated shocks. These winds are dominated by kinetic energy and become particularly effective when they are strongly collimated in disc-dominated systems. The role of supernova driven winds in the Illustris galaxy formation model [Vogelsberger et al., 2013] is to slow down star formation, suppress cosmological gas inflow [Nelson et al., 2015b], and enrich the CGM with metals [Bird et al., 2014; Marinacci et al., 2014; Suresh et al., 2015a,b]. As the Mach number and energy dissipation fields reveal, the heating through the winds can be effectively mediated by shocks. We note that the morphology of galactic wind shocks can be clearly distinguished from other feedback shocks through the typical consecutive arcs they produce well inside the virial radius.

Finally, in Fig. 13.5, we show projections of three of the most massive systems in Illustris-1 at $z = 0$. These cluster-sized haloes underwent violent black hole feedback processes for much of their cosmic evolution, such that their gas is heavily diluted at $z = 0$. This also manifests itself in an ambient WHIM gas phase that is spatially very extended. The shock morphologies in these environments are dominated by a small number of very fine surfaces, indicating that these systems are less dynamic and more relaxed compared to the massive haloes at higher redshifts. As can be seen in Fig. 12.4, the energy release rate of black holes decreases substantially for $z < 0.5$, but the shocks found around the second halo demonstrate that new strong feedback shock waves are still created at these times.

14

Methodology variations

Our shock-finding algorithm can potentially report spurious detections of shocks due to large but poorly resolved gradients in regimes which are dominated by gravitational forces, or where the equations of ideal hydrodynamics are modified by an imposed equation of state. This can for example be the case if gas enters gravitational free fall after rapid cooling, or for gas around the effective equation of state used to model the star-forming medium. The associated jumps in the hydrodynamic variables should not be confused with shocks. We have therefore introduced in Section 11.3 the filtering parameter f , which controls the amount of tolerance we allow for inconsistencies between the jumps inferred from different hydrodynamic quantities. In this section, we briefly test the dependence of our results on this nuisance parameter of our shock detection scheme.

In Fig. 14.1, we compare the shock environment around one of the biggest haloes at $z = 2$ when no shock filtering is applied (right-most panel) to the case when active filtering with our standard settings is applied (middle panel). Without the filter, there is a spray of loosely connected shocked cells with high Mach numbers whose locations coincide with the highly overdense gas of star-forming regions. A large fraction of these shocks is likely spurious and arises because our shock detection algorithm is misled by the transition to a non-ideal gas equation of state at the star formation threshold. If the filter is active, the detections involving gas states in the density-temperature regime $\rho > \rho_{\text{sfr}}$ and $T < 10^5(\rho/\rho_{\text{sfr}})^{0.2}$ K are excluded by construction. These cells are potentially part of the pressurised ISM, or may have been recently, and therefore undergo modifications by means of the subgrid model. Moreover, all detections are discarded when the directions of the pressure, density, or velocity jumps are inconsistent with the temperature gradient. Additionally, for shocks having pre-shock densities $\delta_b > 1000$ only those detections are kept which have consistent Mach numbers within a tolerance f , inferred by different Rankine-

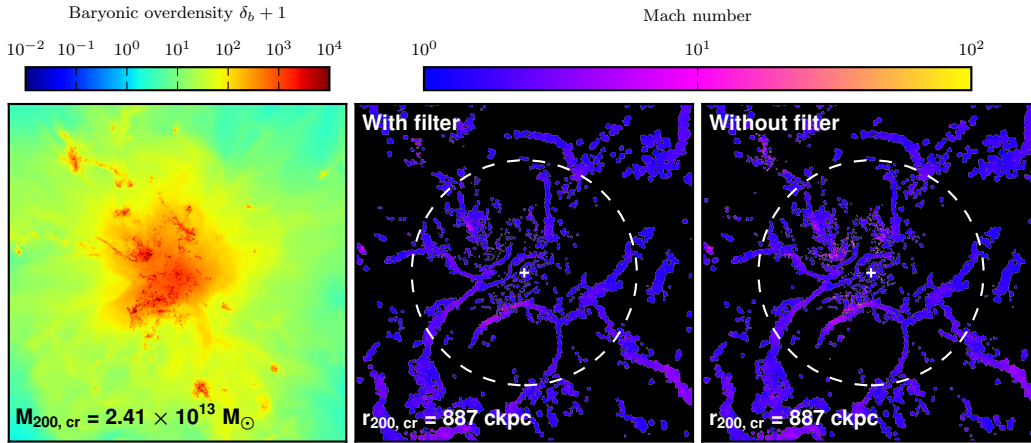


Figure 14.1: Demonstration of the impact of filtering against spurious shock detections due to non-hydrodynamic physics. The middle panel shows the shock environment around a halo at $z = 2$ obtained with our improved shock finding methodology. In the right-hand side panel the result without filtering star-forming gas and inconsistent jumps can be seen. The difference seems to be very subtle, however, the spurious detections dominate the total energy dissipation over a wide range of redshifts, as shown in Fig. 14.2. It can therefore be very important to further improve shock finder implementations developed for non-radiative simulations before applying them to full physics runs.

Hugoniot jump conditions. To be precise, the inequality $f^{-1}\mathcal{M}_p < \mathcal{M}_T < f\mathcal{M}_p$ and for $M_\rho < 3$ or $M_T < 3$ also $f^{-1}\mathcal{M}_\rho < \mathcal{M}_T < f\mathcal{M}_\rho$ has to be fulfilled, where \mathcal{M}_T , \mathcal{M}_p , and \mathcal{M}_ρ is the Mach number inferred from the temperature, pressure, and density jump condition, respectively. For this work we adopt the value $f = 2$, which, as can be seen in the middle panel, filters most of the unwanted detections.

While in the Mach number projections the spurious detections are only visible as a subtle feature, they can nevertheless significantly change the shock energy statistics due to their high pre-shock densities. In Fig. 14.2, we compare the energy dissipation at shocks as a function of Mach number with and without the filter. In the middle panel we give the results for our standard filter configuration as a reference; this plot is the same as in Fig. 12.3. The right-hand side panel shows results obtained if no filter is used. Especially for high redshifts and Mach numbers $\mathcal{M} > 10$, the measured energy dissipation is greatly enhanced in this case. The total dissipation (integral of the curve) without applied filter is higher by factors of 2.9, 8.5, and 1.4 at redshifts $z = 6$, $z = 3.5$, and $z = 0$, respectively.

In the left-hand panel of Fig. 14.2, we investigate the robustness of our filter with respect to the introduced parameter f . Shown is the contribution of haloes to the total dissipation rate for different filter configurations. The black curve is obtained if no filter is used at all. For the black dashed curve we filter gas around star forming regions, but we do not perform consistency checks. The grey curve shows the result obtained if the star forming region is filtered and in addition the consistency of the jump directions with the temperature gradient is demanded. Finally, the coloured curves indicate results with the full filter being active, but for different tolerance

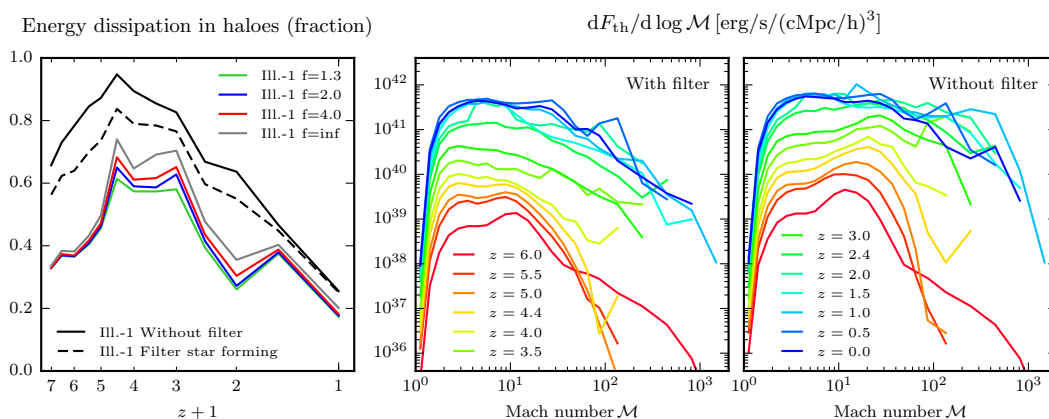


Figure 14.2: *Left panel:* Contribution of shocks in haloes ($\delta_b \geq 100$) to the total dissipation rate for different tolerance values f , as well as for the case when no filter, or only the filter for the star forming region is used. The result does not change significantly for $f \in [1.3, 4]$, and hence our choice of $f = 2$ is very robust. This furthermore indicates that unwanted spurious detections are highly inconsistent with the full set of Rankine-Hugoniot jump conditions. *Middle and right panels:* Energy dissipation at shocks as a function of Mach number with and without the improved shock finding methodology, respectively. Especially at high redshifts the spurious detections enhance the measured energy dissipation by a factor of several.

values f . Even though the parameter range $f \in [1.3, 4]$ is rather large, there are only changes at the 10% level in the results. This means that the spurious detections are strongly violating the consistency relations of the Rankine-Hugoniot jump conditions, allowing them to be reliably filtered with our approach and a parameter value of $f = 2$. Also note that it is not sufficient to only remove detections around star forming regions. The consistency condition is also required to obtain fully reliable results.

An important lesson from our investigation is that shock finders designed and tested with non-radiative simulation are unlikely to work reliably for full physics runs unless they are refined further. Applying them blindly risks that systematic biases in measures like the energy dissipation statistics are introduced. The severity of these biases will depend on the radiative and feedback physics realised in the simulation code, as well as on the particular shock finding technique that is used. For example, while Kang et al. [2007] find only minor effects in runs including cooling and star formation, the energy dissipation tail at high Mach numbers reported in Pfrommer et al. [2007] might partially originate from spurious shocks around the adopted subgrid star formation model.

15

Resolution study

In Fig. 15.1 we study resolution effects related to shock properties for the full physics runs Illustris-1, Illustris-2, and Illustris-3. Illustris-2 and Illustris-3 have an 8 and 64 times lower mass resolution compared to Illustris-1, respectively, but adopt the same physics model and evolve otherwise identical initial conditions. The first row shows the cumulative shock surface area as a function of Mach number for the different resolutions. The results across cosmic time are similar, with the trend that more shocks are found in higher resolution runs. This is especially true at high redshift. These findings are expected since finer and therefore more numerous structures are resolved with increasing resolution.

In the second row we investigate resolution effects on the energy dissipation at shocks as a function of Mach number. At $z = 6$ we find very good agreement, indicating that the difference in the cumulative shock surface area at this time is due to less energetic external shocks. At $z \leq 2$, shocks with $\mathcal{M} \gtrsim 5$ dissipate significantly more energy in the highest resolution run compared to Illustris-2 and Illustris-3. As can be seen in the left-hand side panel of Fig. 12.4, this results in a total energy dissipation which is higher by a factor of around 2 at late times.

We suspect that the origin of this discrepancy lies in the feedback energy released by black holes, for the following reasons. The energy output of quasar and radio AGN as a function of redshift is shown in the central panels of the third row, and for both channels significantly larger values are reached in the highest resolution run. Moreover, as can be seen in the row at the bottom, a notable hot diffuse phase is only created in Illustris-1. As we have argued in Section 12.2, this phase originates from few and highly energetic shocks. Nevertheless, we obtain overall reasonably robust results. Given that the simulations themselves have to produce resolution independent results, as well as the shock finding algorithm, this achievement is non-trivial.

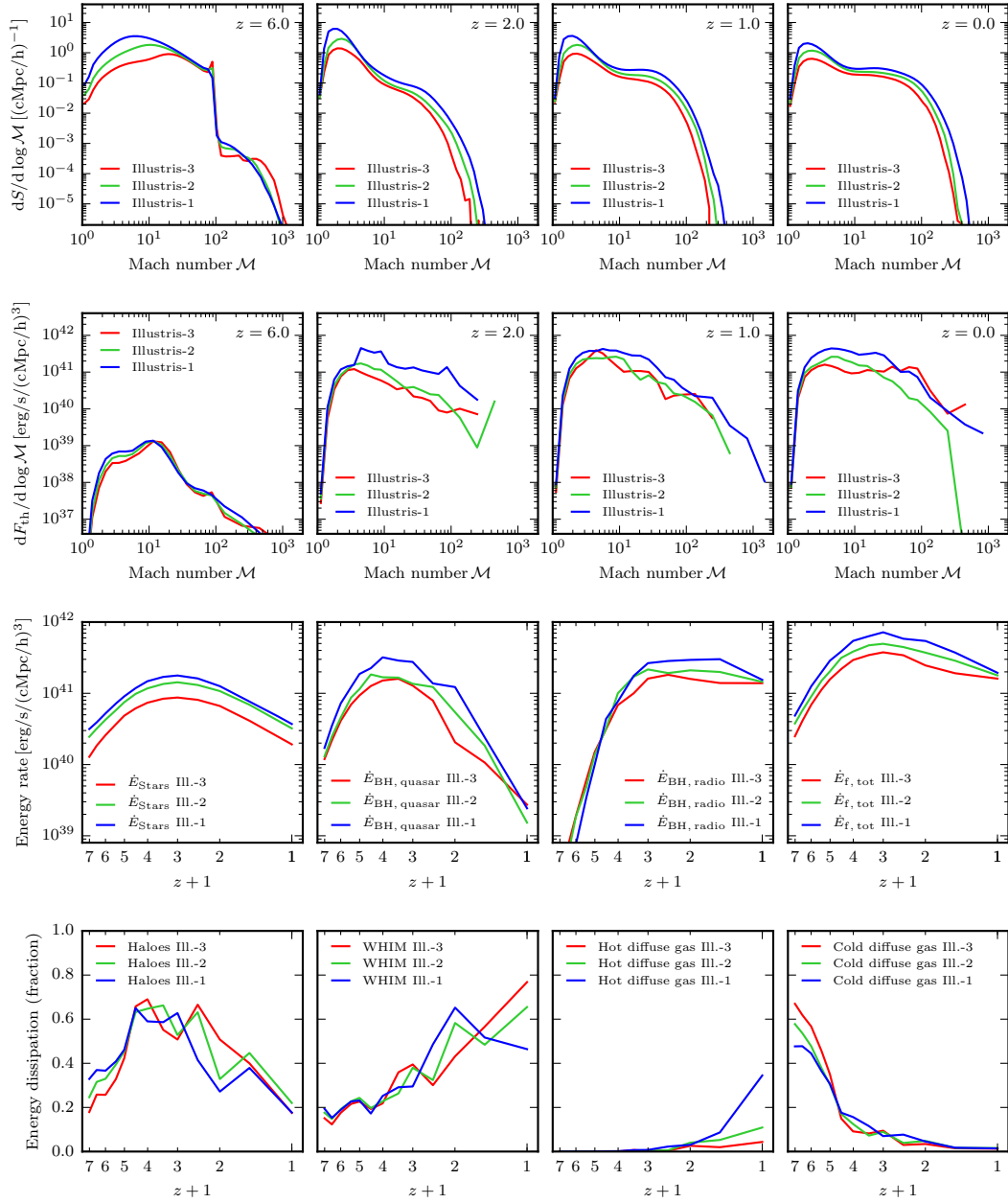


Figure 15.1: Convergence study of several quantities of the full physics runs. From top to bottom we show the cumulative shock surface as a function of Mach number, the energy dissipation at shocks as a function of Mach number, the energy released in different feedback channels as a function of redshift, and the corresponding energy dissipation within different environments as a function of redshift. A higher black hole feedback energy rate is present in Illustris-1 compared to the lower resolution runs, and we suspect this to be the main reason for the increased energy dissipation rate at shocks for $z \leq 2$.

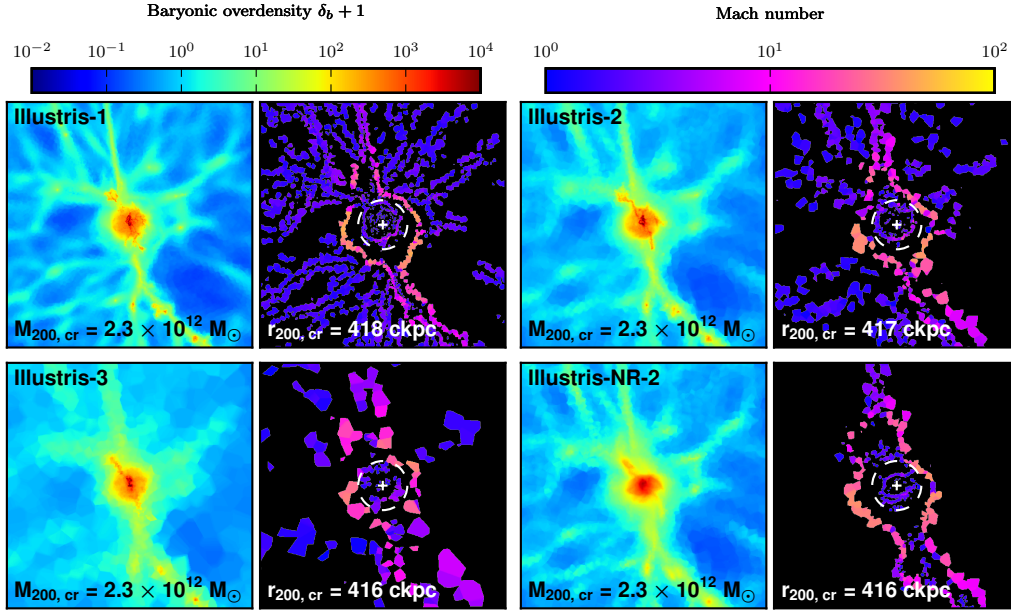


Figure 15.2: Resolution study for the halo with double accretion shock shown in Fig. 13.2. Although less details are resolved in the lower resolution runs, the inner accretion shock can still be seen. Moreover, it is also present in the non-radiative run Illustris-NR-2, providing evidence that this shock is truly an accretion phenomenon. In the full physics runs the shape of the inner shock is roughly spherical, whereas in the non-radiative run it consists of connected bow-shaped shocks, suggesting that cooling helps to stabilize the inner accretion shock and make them more spherical.

Fig. 15.2 shows halo 4 of Fig. 13.1, which can also be seen in the zoom of Fig. 13.2, for different simulations. It can be seen, as previously discussed, that fewer shocks are found for the lower resolution runs. Moreover, in the non-radiative simulation Illustris-NR-2, many shocks on to the cosmic web are erased by the global temperature floor imposed in post-processing to model reionization. In Illustris-3, the resolution is not high enough for detecting distinct shock surfaces at this scale and redshift ($z = 4$). Nevertheless, the detections at hand point towards the existence of a double accretion shock. On the other hand, in Illustris-1 and Illustris-2, the inner and the outer accretion shock are clearly visible, and moreover, the measured Mach numbers are similar. We hence expect that these surfaces are largely converged; running the simulation with an even higher resolution would result in shocks at very similar locations and with very similar properties.

The inner accretion shock is also present in the non-radiative run Illustris-NR-2, and is hence unrelated to feedback processes. Remarkably, the inner shock is less spherical in this simulation and consists of two connected bow-shaped shocks pointing in the direction of the filaments. This observation leads us to suggest that the sphericity in the full physics runs originates from cooling physics, as discussed in Section 13.1.

16

Discussion and summary

Cosmological hydrodynamical simulations are now able to successfully describe the co-evolution of dark matter and baryons in remarkable detail. As the modelling of the involved physical processes is refined, it is important to explore novel ways of accessing the information content provided by such simulations. Thus far, the analysis of properties of the diffuse gas has often been restricted to considerations of the most basic quantities, such as (spherically averaged) density and temperature values. However, gas dynamics has much more to offer. In this study, we have for the first time analysed the shock statistics in a state-of-the-art high-resolution simulation of galaxy formation, the Illustris project.

We have characterised the shocks in the Illustris universe by post-processing the simulations with AREPO's shock finding algorithm (see Part II). AREPO accurately solves the Euler equations of ideal hydrodynamics with a finite volume method on a moving-mesh and is able to capture shocks precisely. In combination with our shock finder we can reliably detect and characterise the hydrodynamic shocks and associated quantities.

From a theoretical point of view, shocks are of fundamental interest. They arise from the non-linearity of the hydrodynamic equations and are therefore a manifestation of one of the key characteristics of the governing equations. In order to understand the gas dynamics in a complex system, much can be learned by merely investigating the locations and strengths of shocks. Moreover, shock quantities provide a reduced dataset of the gas in hydrodynamic cosmological simulations, which can be used to simplify the analysis of large output files¹.

In this work we have focused on analysing shocks in the full physics Illustris-1 simulation. In order to put the results for the most important statistics into context,

¹In the case of Illustris-1, a snapshot consists of around 1.2×10^{10} resolution elements and has a size of about 2.5 terabytes.

we also compared with results obtained for the non-radiative run Illustris-NR-2. The differences in shock statistics and shock morphologies between the two kinds of simulations arise due to physics beyond ideal gas dynamics in Illustris-1. Most importantly, this simulation accounts for cosmic reionization, radiative cooling and heating of gas, star formation and black hole growth, as well as their associated feedback processes. In what follows, we summarise our most important findings.

- **Improvement of the shock finding methodology.** In the full physics runs, we have discovered spurious shock detections in star-forming regions and at the edges of cold, poorly resolved self-gravitating clumps. They originate from modifications of the gas state by the subgrid ISM model and cooling physics, respectively. We have shown that these spurious detections severely change the energy dissipation statistics if the problem is just ignored. We have thus introduced improvements to our shock finding routines that were not part of the method we developed originally for non-radiative runs. First of all, we excise star-forming regions in our shock finding algorithm. Secondly, we ignore detections for which the hydrodynamic jumps are inconsistent with the shock direction given by the temperature gradient, and last but not least, we demand that shocks in overdense regions consistently fulfill the Rankine-Hugoniot jump conditions for several different hydrodynamic quantities. We have demonstrated that these conservative filters are necessary to reliably find and characterise hydrodynamic shocks in the Illustris simulation. We anticipate similar problems for other shock finders developed for non-radiative runs, and hence caution against applying them in unmodified form to full physics runs.
- **Impact of reionization.** We find in Illustris-1 a significant increase in the measured shock surface area as well as in the total energy dissipation between redshifts $z = 6$ and $z = 5.5$. This coincides with reionization, which is modelled by a uniform UV background following Faucher-Giguère et al. [2009] and happens in the simulation almost instantaneously at $z \lesssim 6$. We interpret this finding as the signature of a temporary population of weak shocks associated with the dynamical response of the gas to the sudden heating during reionization, which for example lets gas stream off the shallow potential wells of very low mass dark matter haloes. Quantitatively, this result certainly depends on the details of the reionization transition.
- **Shock surface area.** Until $z \approx 1$ by far the largest shock surface area can be found between voids and the cosmic web. The cumulative shock surface statistic is therefore dominated by the number and temperature of voids. During structure formation, small voids merge to form larger ones, and the number of low Mach number shocks decreases. At the same time, voids are coupled to the Hubble flow and cool adiabatically, resulting in an increase in the occurrence of high Mach number shocks. Moreover, we find that the number of shocks with $\mathcal{M} \approx 4$ stays roughly constant with time. For $z < 1$ a significant number of shocks can also be found in the WHIM, with a maximum contribution

of around 34% at $z = 0$. After reionization, the total shock surface area decreases monotonically, nevertheless, at the present epoch we find a total shock surface area in Illustris-1 which is around 1.4 times larger compared to the corresponding non-radiative run.

- **Energy dissipation.** By measuring the energy dissipation rate at shocks the impact on the thermal history of the gas can be inferred. In non-radiative runs the overall energy dissipation grows with cosmic time, and when its Mach number dependence is considered a bimodality in the energy statistics can be seen; internal low Mach number shocks dissipate most of the energy, whereas external shocks have high Mach numbers but are much less energetic [Ryu et al., 2003]. Remarkably, we find for Illustris-1 that shocks with Mach numbers smaller and larger than $\mathcal{M} \approx 10$ contribute about equally to the overall dissipation, and no bimodality is observed. For this full physics simulation, we attribute the large population of high Mach number shocks to feedback processes.

Between $z = 3.5$ and $z = 2$, most of the energy dissipation occurs within haloes, where cooling processes create structures with large density contrasts. It is also around this time that the radio mode feedback of black holes first kicks in, which is associated with the low accretion rate state of the black holes. We show that this type of feedback drives energetic high Mach number shocks and creates an extended WHIM gas phase. We measure a total energy dissipation in Illustris-1 which is higher by a factor of around 8 for $z < 2$ compared to the corresponding non-radiative run. It is plausible that a large fraction of this additional dissipation originates from the black hole radio mode feedback.

As a result of the redistribution of baryons at low redshifts to the WHIM, this phase gains importance as a site for energetic shocks. At $z = 0$ around 80% of the dissipation occurs at shocks in the WHIM and in the hot diffuse phase ($T > 10^7$ and $\delta_b < 100$). This is in strong contrast to the corresponding non-radiative simulation, in which about 60% of the thermalisation happens inside haloes at the present epoch. At $z = 0$ the total dissipation rate in shocks in Illustris-1 amounts to $1.7 \times 10^{41} \text{ erg s}^{-1} \text{ Mpc}^{-3}$. Moreover, except for the cold diffuse gas phase, all environments have similar and approximately constant specific dissipation rates for $z < 3$. This value is approximately $10^{-1} \text{ erg g}^{-1} \text{ s}^{-1}$.

- **Shock morphologies.** We have investigated accretion shocks of haloes with mass above $10^{12} M_{\odot}$ at $z = 4$ when feedback processes influencing large scales are still largely absent. As expected for haloes in this mass range, we find accretion shocks which are typically located outside of the virial radius. However, we see filaments which can penetrate the accretion shock, and their gas streams enter the halo unshocked. Interestingly, we discover that the formation of a spherical accretion shock further inside the halo is still possible in this case. Our finding of an accretion mode involving both an inner and an outer accretion

shock is consistent with recent work by Nelson et al. [2015a], who analysed the thermal gas structure of $10^{12} M_{\odot}$ haloes at $z = 2$ in zoom simulations.

We have showcased different shock morphologies of galaxies and galaxy clusters at $z \leq 2$. Those include complex networks of weak shocks in cluster interiors, merger remnants, strong black hole feedback shocks, shocks on to cosmic sheets, as well as supersonic stellar winds and complex flows around galaxies. By visually comparing the properties of an AGN feedback shock with a merger remnant shock, we argue that the radio mode feedback is too strong in Illustris-1. This conclusion is consistent with the findings of Genel et al. [2014] and Haider et al. [2016], who report a baryon fraction inside clusters which is too small at the present epoch and also attribute this finding to the radio mode feedback channel. This also shows that the properties of shocks can be used in future state-of-the-art cosmological simulations for constraining the modelling and parameter space of feedback implementations.

These results underline the importance of feedback processes for shock statistics, and the large impact they have on the shocks that are present. As we expect the strongest shocks to be efficient accelerators that produce non-thermal particle distributions, shock statistics combined with upper limits on cosmic ray densities could be developed into a powerful constraint for viable feedback models. Future simulation models will also try to take the cosmic ray particle populations self-consistently into account. This means that further refinements of the shock detection methodology will then be needed in order to cope with mixtures of ordinary gas and cosmic rays.

Part IV

Discontinuous Galerkin Hydrodynamics

Abstract

Solving the Euler equations of ideal hydrodynamics as accurately and efficiently as possible is a key requirement in many astrophysical simulations. It is therefore important to continuously advance the numerical methods implemented in current astrophysical codes, especially also in light of evolving computer technology, which favours certain computational approaches over others. Here we introduce the new adaptive mesh refinement (AMR) code TENET, which employs a high order discontinuous Galerkin (DG) scheme for hydrodynamics. The Euler equations in this method are solved in a weak formulation with a polynomial basis by means of explicit Runge-Kutta time integration and Gauss-Legendre quadrature. This approach offers significant advantages over commonly employed second order finite volume (FV) solvers. In particular, the higher order capability renders it computationally more efficient, in the sense that the same precision can be obtained at significantly less computational cost. Also, the DG scheme inherently conserves angular momentum in regions where no limiting takes place, and it typically produces much smaller numerical diffusion and advection errors than a FV approach. A further advantage lies in a more natural handling of AMR refinement boundaries, where a fall-back to first order can be avoided. Finally, DG requires no wide stencils at high order, and offers an improved data locality and a focus on local computations, which is favourable for current and upcoming highly parallel supercomputers. We describe the formulation and implementation details of our new code, and demonstrate its performance and accuracy with a set of two- and three-dimensional test problems. The results confirm that DG schemes have a high potential for astrophysical applications.

17

Introduction

Through the availability of ever more powerful computing resources, computational fluid dynamics has become an important part of astrophysical research. It is of crucial help in shedding light on the physics of the baryonic part of our Universe, and contributes substantially to progress in our theoretical understanding of galaxy formation and evolution, of star and planet formation, and of the dynamics of the intergalactic medium. In addition, it plays an important role in planetary science and in solving engineering problems related to experimental space exploration.

In the astrophysics community, most numerical work thus far on solving the Euler equations of ideal hydrodynamics has been carried out with two basic types of codes. On the one hand, there is the broad class of Eulerian methods which utilize classical hydrodynamic solvers operating on fixed Cartesian grids [e.g. Stone et al., 2008], or on meshes which can adjust their resolution in space with the adaptive mesh refinement (AMR) technique [e.g. Bryan et al., 2014; Fryxell et al., 2000; Mignone et al., 2007; Teyssier, 2002]. On the other hand, there are pseudo-Lagrangian discretizations in the form of smoothed particle hydrodynamics (SPH), which are other flexible and popular tools to study many astrophysical problems [e.g. Springel, 2005; Wadsley et al., 2004].

Some of the main advantages and drawbacks of these methods become apparent if we recall the fundamental difference in their numerical approach, which is that grid codes discretize space whereas SPH decomposes a fluid in terms of mass elements. The traditional discretization of space used by Eulerian methods yields good convergence properties, and a high accuracy and efficiency for many problems. Furthermore, calculations can be straightforwardly distributed onto parallel computing systems, often allowing a high scalability provided there is only little communication between the cells. The discretization of mass used by SPH on the other hand results in a natural resolution adjustment in converging flows such that most of the computing

time and available resolution are dedicated to dense, astrophysically interesting regions. Moreover, Lagrangian methods can handle gas with high advection velocities without suffering from large errors. However, both methods have also substantial weaknesses, ranging from problems with the Galilean invariance of solutions in the case of grid codes, to a suppression of fluid instabilities and noise in the case of SPH.

This has motivated attempts to combine the advantages of traditional grid-based schemes and of SPH in new methods, such as moving mesh-codes [Duffell and MacFadyen, 2011; Springel, 2010] or in mesh-free methods that retain a higher degree of accuracy [Hopkins, 2015; Lanson and Vila, 2008] than SPH. Both of these new developments conserve angular momentum better than plain Eulerian schemes, while still capturing shocks accurately, a feature that is crucial for many applications in astrophysics. However, these new methods need to give up the simple discretization of space into regular grids, meaning also that their computational efficiency takes a significant hit, because the regular memory access patterns possible for simple structured discretizations are ideal for leveraging a high fraction of the peak performance of current and future hardware.

In fact, the performance of supercomputers has increased exponentially over the last two decades, roughly following the empirical trend of Moore’s law, which states that the transistor count and hence performance of computing chips doubles roughly every two years. Soon, parallel computing systems featuring more than 10 million cores and “exascale” performance in the range of 10^{18} floating point operations per second are expected. Making full use of this enormous compute power for astrophysical research will require novel generations of simulation codes with superior parallel scaling and high resilience when executed on more and more cores. Also, since the raw floating point speed of computer hardware grows much faster than memory access speed, it is imperative to search for new numerical schemes that focus on local computations within the resolution elements. Those allow for data locality in the memory leading to a fast data transfer to the CPU cache, and furthermore, the relative cost for communication with neighbouring resolution elements is reduced.

A very interesting class of numerical methods in this context are so-called discontinuous Galerkin (DG) schemes, which can be used for a broad range of partial differential equations. Since the introduction of DG [Reed and Hill, 1973] and its generalization to nonlinear problems [Cockburn et al., 1989, 1990; Cockburn and Shu, 1989, 1991, 1998], it has been successfully applied in diverse fields of physics such as aeroacoustics, electro magnetism, fluid dynamics, porous media, etc. [Cockburn et al., 2011; Gallego-Valencia et al., 2014]. On the other hand, in the astrophysics community the adoption of modern DG methods has been fairly limited so far. However, two recent works suggest that this is about to change. Mocz et al. [2014] presented a DG method for solving the magnetohydrodynamic (MHD) equations on arbitrary grids as well as on a moving Voronoi mesh, and Zanotti et al. [2015] developed an AMR code for relativistic MHD calculations. The present work is a further contribution in this direction and aims to introduce a novel DG-based hydrodynamical code as an alternative to commonly employed schemes in the field.

DG is a finite element method which incorporates several aspects from finite volume (FV) methods. The partial differential equation is solved in a weak formulation by means of local basis functions, yielding a global solution that is in general discontinuous across cell interfaces. The approach requires communication only between directly neighbouring cells and allows for the exact conservation of physical quantities including angular momentum. Importantly, the method can be straightforwardly implemented with arbitrary spatial order, since it directly solves also for higher order moments of the solution. Unlike in standard FV schemes, this higher order accuracy is achieved without requiring large spatial stencils, making DG particularly suitable for utilizing massive parallel systems with distributed memory because of its favourable compute-to-communicate ratio and enhanced opportunities to hide communication behind local computations.

In order to thoroughly explore the utility of DG in real astrophysical applications, we have developed TENET, an MPI-parallel DG code which solves the Euler equations on an AMR grid to arbitrary spatial order. In our method, the solution within every cell is given by a linear combination of Legendre polynomials, and the propagation in time is accomplished with an explicit Runge-Kutta (RK) time integrator. A volume integral and a surface integral have to be computed numerically for every cell in every timestep. The surface integral involves a numerical flux computation which we carry out with a Riemann solver, similar to how this is done in standard Godunov methods. In order to cope with physical discontinuities and spurious oscillations we use a simple minmod limiting scheme.

The goal of this part of the thesis is to introduce the concepts of our DG implementation and compare its performance to a standard FV method based on the reconstruct-evolve-average (REA) approach. In this traditional scheme, higher order information is discarded in the averaging process and recomputed in the reconstruction step, leading to averaging errors and numerical diffusion. DG on the other hand does not only update the cell-averaged solution every cell, but also higher order moments of the solution, such that the reconstruction becomes obsolete. This aspect of DG leads to important advantages over FV methods, in particular an inherent improvement of angular momentum conservation and much reduced advection errors, especially for but not restricted to smooth parts of the solutions. These accuracy gains and the prospect to translate them to a lower computational cost at given computational error form a strong motivation to investigate the use of DG in astrophysical applications.

This part of the thesis is structured as follows. In Section 18, we present the general methodology of our DG implementation for Cartesian grids. The techniques we adopt for limiting the numerical solution are described in Section 19, and the generalization to a mesh with adaptive refinement is outlined in Section 20. These sections give a detailed account of the required equations and discretization formulae for the sake of clarity and definiteness, something that we hope does not discourage interested readers. We then validate our DG implementation and compare it to a standard second order Godunov FV solver with two- and three-dimensional test

problems in Section 21. Finally, Section 22 summarizes our findings and gives a concluding discussion.

18

Discontinuous Galerkin hydrodynamics

18.1 Euler equations

The Euler equations are conservation laws for mass, momentum, and total energy of a fluid. They are a system of hyperbolic partial differential equations and can be written in compact form as

$$\frac{\partial \mathbf{u}}{\partial t} + \sum_{\alpha=1}^3 \frac{\partial \mathbf{f}_{\alpha}}{\partial x_{\alpha}} = 0, \quad (18.1)$$

with the state vector

$$\mathbf{u} = \begin{pmatrix} \rho \\ \rho \mathbf{v} \\ \rho e \end{pmatrix} = \begin{pmatrix} \rho \\ \rho \mathbf{v} \\ \rho u + \frac{1}{2} \rho \mathbf{v}^2 \end{pmatrix}, \quad (18.2)$$

and the flux vectors

$$\mathbf{f}_1 = \begin{pmatrix} \rho v_1 \\ \rho v_1^2 + p \\ \rho v_1 v_2 \\ \rho v_1 v_3 \\ (\rho e + p)v_1 \end{pmatrix} \quad \mathbf{f}_2 = \begin{pmatrix} \rho v_2 \\ \rho v_1 v_2 \\ \rho v_2^2 + p \\ \rho v_2 v_3 \\ (\rho e + p)v_2 \end{pmatrix} \quad \mathbf{f}_3 = \begin{pmatrix} \rho v_3 \\ \rho v_1 v_3 \\ \rho v_2 v_3 \\ \rho v_3^2 + p \\ (\rho e + p)v_3 \end{pmatrix}. \quad (18.3)$$

The unknown quantities are density ρ , velocity \mathbf{v} , pressure p , and total energy per unit mass e . The latter can be expressed in terms of the internal energy per unit

mass u and the kinetic energy of the fluid, $e = u + \frac{1}{2}\mathbf{v}^2$. For an ideal gas, the system is closed with the equation of state

$$p = \rho u(\gamma - 1), \quad (18.4)$$

where γ denotes the adiabatic index.

18.2 Solution representation

We partition the domain by non-overlapping cubical cells, which may be refined using AMR techniques as explained in later sections. Moreover, we follow the approach of a classical modal DG scheme, where the solution in cell K is given by a linear combination of $N(k)$ orthogonal and normalized basis functions ϕ_l^K :

$$\mathbf{u}^K(\mathbf{x}, t) = \sum_{l=1}^{N(k)} \mathbf{w}_l^K(t) \phi_l^K(\mathbf{x}). \quad (18.5)$$

In this way, the dependence on time and space of the solution is split into time-dependent weights, and basis functions which are constant in time. Consequently, the state of a cell is completely characterized by $N(k)$ weight vectors $\mathbf{w}_j^K(t)$.

The above equation can be solved for the weights by multiplying with the corresponding basis function ϕ_j^K and integrating over the cell volume. Using the orthogonality and normalization of the basis functions yields

$$\mathbf{w}_j^K = \frac{1}{|K|} \int_K \mathbf{u}^K \phi_j^K \, dV, \quad j = 1, \dots, N(k), \quad (18.6)$$

where $|K|$ is the volume of the cell. The first basis function is chosen to be $\phi_1 = 1$ and hence the weight \mathbf{w}_1^K is the cell average of the state vector \mathbf{u}^K . The higher order moments of the state vector are described by weights \mathbf{w}_j^K with $j \geq 2$.

The basis functions can be defined on a cube in terms of scaled variables $\boldsymbol{\xi}$,

$$\phi_l(\boldsymbol{\xi}) : [-1, 1]^3 \rightarrow \mathbb{R}. \quad (18.7)$$

The transformation between coordinates $\boldsymbol{\xi}$ in the cell frame of reference and coordinates \mathbf{x} in the laboratory frame of reference is

$$\boldsymbol{\xi} = \frac{2}{\Delta x^K}(\mathbf{x} - \mathbf{x}^K), \quad (18.8)$$

where Δx^K and \mathbf{x}^K are edge length and cell centre of cell K , respectively. For our DG implementation, we construct a set of three-dimensional polynomial basis

functions with a maximum degree of k as products of one-dimensional scaled Legendre polynomials \tilde{P} :

$$\{\phi_l(\boldsymbol{\xi})\}_{l=1}^{N(k)} = \left\{ \tilde{P}_u(\xi_1)\tilde{P}_v(\xi_2)\tilde{P}_w(\xi_3) \mid u, v, w \in \mathbb{N}_0 \wedge u + v + w \leq k \right\}. \quad (18.9)$$

The first few Legendre polynomials are shown in Section 3.2.2. The number of basis functions for polynomials with a maximum degree of k is

$$N(k) = \sum_{u=0}^k \sum_{v=0}^{k-u} \sum_{w=0}^{k-u-v} 1 = \frac{1}{6}(k+1)(k+2)(k+3). \quad (18.10)$$

Furthermore, when polynomials with a maximum degree of k are used, a scheme with spatial order $p = k + 1$ is obtained. For example, linear basis functions lead to a scheme which is of second order in space.

18.3 Initial conditions

Given initial conditions $\mathbf{u}(\mathbf{x}, t = 0) = \mathbf{u}(\mathbf{x}, 0)$, we have to provide an initial state for the DG scheme which is consistent with the solution representation. To this end, the initial conditions are expressed by means of the polynomial basis on cell K , which will then be

$$\mathbf{u}^K(\mathbf{x}, 0) = \sum_{l=1}^{N(k)} \mathbf{w}_l^K(0) \phi_l^K(\mathbf{x}). \quad (18.11)$$

If the initial conditions at hand are polynomials with degree $\leq k$, this representation preserves the exact initial conditions, otherwise equation (18.11) is an approximation to the given initial conditions. The initial weights can be obtained by performing an L^2 -projection,

$$\min_{\{w_{i,i}^K(0)\}_i} \int_K (u_i^K(\mathbf{x}, 0) - u_i(\mathbf{x}, 0))^2 dV, \quad i = 1, \dots, 5, \quad (18.12)$$

where $i = 1, \dots, 5$ enumerates the conserved variables. The projection above leads to the integral

$$\mathbf{w}_j^K(0) = \frac{1}{|K|} \int_K \mathbf{u}(\mathbf{x}, 0) \phi_j^K(\mathbf{x}) dV, \quad j = 1, \dots, N(k), \quad (18.13)$$

which can be transformed to the reference frame of the cell, viz.

$$\mathbf{w}_j^K(0) = \frac{1}{8} \int_{[-1,1]^3} \mathbf{u}(\boldsymbol{\xi}, 0) \phi_j(\boldsymbol{\xi}) d\boldsymbol{\xi}, \quad j = 1, \dots, N(k). \quad (18.14)$$

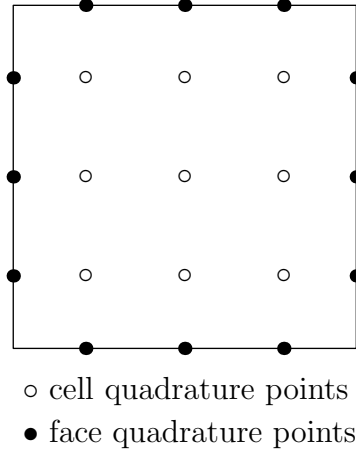


Figure 18.1: In the DG scheme, a surface integral and a volume integral have to be computed numerically for every cell, see equation (18.17). In our approach, we solve these integrals by means of Gaussian quadrature (Section 3.2.3). This example shows the nodes for our third order DG method (with second order polynomials) when used in a two-dimensional configuration. The black nodes indicate the positions where the surface integral is evaluated, which involves a numerical flux calculation with a Riemann solver. The white nodes are used for numerically estimating the volume integral.

We solve the integral numerically by means of tensor product Gauss-Legendre quadrature (hereafter called Gaussian quadrature) with $(k + 1)^3$ nodes:

$$\mathbf{w}_j^K(0) \approx \frac{1}{8} \sum_{q=1}^{(k+1)^3} \mathbf{u}(\boldsymbol{\xi}_q^{3D}, 0) \phi_j(\boldsymbol{\xi}_q^{3D}) \omega_q^{3D}, \quad j = 1, \dots, N(k). \quad (18.15)$$

Here, $\boldsymbol{\xi}_q^{3D}$ is the position of the quadrature node q in the cell frame of reference and ω_q^{3D} denotes the corresponding quadrature weight. The technique of Gaussian quadrature is explained in more detail in Section 3.2.3.

18.4 Evolution equation for the weights

In order to derive the DG scheme on a cell K , the Euler equations for a polynomial state vector \mathbf{u}^K are multiplied by the basis function ϕ_j^K and integrated over the cell volume,

$$\int_K \left[\frac{\partial \mathbf{u}^K}{\partial t} + \sum_{\alpha=1}^3 \frac{\partial \mathbf{f}_\alpha}{\partial x_\alpha} \right] \phi_j^K dV = 0. \quad (18.16)$$

Integration by parts of the flux divergence term and a subsequent application of Gauss' theorem leads to

$$\frac{d}{dt} \int_K \mathbf{u}^K \phi_j^K dV - \sum_{\alpha=1}^3 \int_K \mathbf{f}_\alpha \frac{\partial \phi_j^K}{\partial x_\alpha} dV + \sum_{\alpha=1}^3 \int_{\partial K} \mathbf{f}_\alpha n_\alpha \phi_j^K dS = 0, \quad (18.17)$$

where $\hat{n} = (n_1, n_2, n_3)^T$ denotes the outward-pointing unit normal vector of the surface ∂K . In the following, we discuss each of the terms separately.

According to equation (18.6), the first term is simply the time variation of the weights,

$$\frac{d}{dt} \int_K \mathbf{u}^K \phi_j^K dV = |K| \frac{d\mathbf{w}_j^K}{dt}. \quad (18.18)$$

The second and third terms are discretized by transforming the integrals to the cell frame and applying Gaussian quadrature (Fig. 18.1, Section 3.2.3). With this procedure the second term becomes

$$\begin{aligned} & \sum_{\alpha=1}^3 \int_K \mathbf{f}_\alpha(\mathbf{u}^K(\mathbf{x}, t)) \frac{\partial \phi_j^K(\mathbf{x})}{\partial x_\alpha} dV \\ &= \frac{(\Delta x^K)^2}{4} \sum_{\alpha=1}^3 \int_{[-1,1]^3} \mathbf{f}_\alpha(\mathbf{u}^K(\boldsymbol{\xi}, t)) \frac{\partial \phi_j(\boldsymbol{\xi})}{\partial \xi_\alpha} d\boldsymbol{\xi} \\ &\approx \frac{(\Delta x^K)^2}{4} \sum_{\alpha=1}^3 \sum_{q=1}^{(k+1)^3} \mathbf{f}_\alpha(\mathbf{u}^K(\boldsymbol{\xi}_q^{3D}, t)) \frac{\partial \phi_j(\boldsymbol{\xi})}{\partial \xi_\alpha} \Big|_{\boldsymbol{\xi}_q^{3D}} \omega_q^{3D}. \end{aligned} \quad (18.19)$$

Note that the transformation of the derivative $\partial/\partial x_\alpha$ gives a factor of 2, see equation (18.8). The volume integral is computed by Gaussian quadrature with $k+1$ nodes per dimension. These nodes allow the exact integration of polynomials up to degree $2k+1$.

The flux functions \mathbf{f}_α in the above expression can be evaluated analytically, which is not the case for the fluxes in the last term of the evolution equation (18.17). This is because the solution is discontinuous across cell interfaces. We hence have to introduce a numerical flux function $\bar{\mathbf{f}}(\mathbf{u}^{K-}, \mathbf{u}^{K+}, \hat{n})$, which in general depends

on the states left and right of the interface and on the normal vector. With this numerical flux, the third term in equation (18.17) takes the form

$$\begin{aligned}
& \sum_{\alpha=1}^3 \int_{\partial K} \mathbf{f}_\alpha n_\alpha(\mathbf{x}) \phi_j^K(\mathbf{x}) \, dS \\
&= \frac{(\Delta x^K)^2}{4} \int_{\partial[-1,1]^3} \bar{\mathbf{f}}(\mathbf{u}^{K-}(\boldsymbol{\xi}, t), \mathbf{u}^{K+}(\boldsymbol{\xi}, t), \hat{\mathbf{n}}(\boldsymbol{\xi})) \phi_j(\boldsymbol{\xi}) \, dS_{\boldsymbol{\xi}} \\
&\approx \frac{(\Delta x^K)^2}{4} \sum_{A \in \partial[-1,1]^3} \sum_{q=1}^{(k+1)^2} \bar{\mathbf{f}}(\mathbf{u}^{K-}(\boldsymbol{\xi}_{q,A}^{2D}, t), \mathbf{u}^{K+}(\boldsymbol{\xi}_{q,A}^{2D}, t), \hat{\mathbf{n}}) \phi_j(\boldsymbol{\xi}_{q,A}^{2D}) \omega_q^{2D}.
\end{aligned} \tag{18.20}$$

Here for each interface of the normalized cell, a two-dimensional Gaussian quadrature rule with $(k+1)^2$ nodes is applied. The numerical flux across each node can be calculated with a one-dimensional Riemann solver, as in ordinary Godunov schemes. For our DG scheme, we use the positivity preserving HLLC Riemann solver [Toro, 2009].

In order to model physics extending beyond ideal hydrodynamics, source terms can be added on the right-hand side of equation (18.16). Most importantly, the treatment of gravity is accomplished by the source term

$$\mathbf{s} = \begin{pmatrix} 0 \\ -\rho \nabla \Phi \\ -\rho \mathbf{v} \cdot \nabla \Phi \end{pmatrix}, \tag{18.21}$$

which by projecting onto the basis function ϕ_j inside cell K and discretizing becomes

$$\begin{aligned}
& \int_K \mathbf{s}(\mathbf{x}, t) \phi_j^K(\mathbf{x}) \, dV \\
&= \frac{|K|}{8} \int_{[-1,1]^3} \mathbf{s}(\boldsymbol{\xi}, t) \phi_j(\boldsymbol{\xi}) \, d\boldsymbol{\xi} \\
&\approx \frac{|K|}{8} \sum_{q=1}^{(k+1)^3} \mathbf{s}(\boldsymbol{\xi}_q^{3D}, t) \phi_j(\boldsymbol{\xi}_q^{3D}) \omega_q^{3D}.
\end{aligned} \tag{18.22}$$

We have now discussed each term of the basic equation (18.17) and arrived at a spatial discretization of the Euler equations of the form

$$\frac{d\mathbf{w}_j^K}{dt} + \mathbf{R}_K = 0, \quad j = 1, \dots, N(k), \tag{18.23}$$

which represents a system of coupled ordinary differential equations. For the discretization in time we apply an explicit strong stability preserving Runge-Kutta (SSP RK) scheme [Gottlieb et al., 2001] of the same order as the spatial discretization.

With a combination of an SSP RK method, a positivity preserving Riemann solver, and a positivity limiter (see Section 19.4), negative pressure and density values in the hydro scheme can be avoided. The Butcher tables of the SSP RK methods we use in our code are listed in the Appendix of the thesis.

18.5 Timestep calculation

If the Euler equations are solved with an explicit time integrator, the timestep has to fulfil a Courant-Friedrichs-Lewy (CFL) condition for achieving numerical timestep stability. For the DG scheme with explicit RK time integration, the timestep depends also on the order $p = k + 1$ of the scheme. We calculate the timestep Δt^K of cell K following Cockburn and Shu [1989] as

$$\Delta t^K = \frac{\text{CFL}}{2k+1} \left(\frac{|v_1^K| + c^K}{\Delta x_1^K} + \frac{|v_2^K| + c^K}{\Delta x_2^K} + \frac{|v_3^K| + c^K}{\Delta x_3^K} \right)^{-1}, \quad (18.24)$$

where $c = \sqrt{\gamma p / \rho}$ is the sound speed and the denominators in the parantheses are the edge lengths of the cell. Since we used only cubic cells, we have $\Delta x_1^K = \Delta x_2^K = \Delta x_3^K = \Delta x^K$. In this work, we apply a global timestep given by the minimum of the local timesteps among all cells. The CFL number depends on the choice of flux calculation but in practice should be set to a somewhat smaller value as formally required. This is because for the calculation of the timestep in an RK method of second order or higher, a mathematical inconsistency arises. At the beginning of the timestep, only the current velocity and sound speed of the gas are known, whereas in principle the maximum speed occurring during all RK stages within one step should be used for calculating the CFL timestep to be strictly safe. This information is not readily available, and the only straightforward way to cope with this problem is to reduce the CFL number to a value smaller than mathematically required. For the tests presented in this work, we decided to adopt a conservative choice of $\text{CFL} = 0.2$. Furthermore, if the positivity limiter is used, the hydro timestep Δt^K has to be modified and can be slightly more restrictive, as described in Section 19.4.

Source terms $\mathbf{s}(\mathbf{u}, t)$ on the right hand side of the Euler equations can induce additional timestep criteria. In this case, positivity of the solution can be enforced by determining the timestep such that $\rho(\mathbf{u}') > 0$ and $p(\mathbf{u}') > 0$, with $\mathbf{u}' = \mathbf{u} + 2\mathbf{s}(\mathbf{u}, t) \Delta t$ [Zhang, 2006]. For the gravity source term (18.21), this leads to the timestep limit

$$\Delta t_{\text{grav}}^K \leq \frac{1}{\sqrt{2\gamma(\gamma-1)}} \frac{c}{|\nabla\Phi|}. \quad (18.25)$$

The actual timestep has then to be chosen as the minimum of the hydrodynamical and gravity timesteps.

18.6 Angular momentum conservation

A welcome side effect of the computation of higher order moments of the solution in DG schemes is that angular momentum is inherently conserved. Without loss of generality, we show this conservation property explicitly for the two-dimensional case ($z = 0$). In this case, the angular momentum density is defined as

$$L = x\rho v_y - y\rho v_x, \quad (18.26)$$

and the flux momentum tensor in 2D is given by

$$\begin{pmatrix} f_{1,2} & f_{2,2} \\ f_{1,3} & f_{2,3} \end{pmatrix} = \begin{pmatrix} p + \rho v_1^2 & \rho v_1 v_2 \\ \rho v_1 v_2 & p + \rho v_2^2 \end{pmatrix}. \quad (18.27)$$

The conservation law for angular momentum can be conveniently derived from the Euler equations (18.1). Multiplying the x -momentum equation by y and the y -momentum equation by x , applying the product rule and subsequently subtracting the two equations gives

$$\frac{\partial L}{\partial t} + \frac{\partial}{\partial x}(x f_{1,3} - y f_{1,2}) + \frac{\partial}{\partial y}(x f_{2,3} - y f_{2,2}) = 0. \quad (18.28)$$

In order to obtain the angular momentum conservation law on a cell basis, this can be integrated over element K , resulting in

$$\frac{d}{dt} \int_K L \, dV + \int_{\partial K} x(f_{1,3} n_1 + f_{2,3} n_2) - y(f_{1,2} n_1 + f_{2,2} n_2) \, dS = 0. \quad (18.29)$$

On the other hand, for a DG scheme with order $p > 1$, we can choose the test function to be $\phi^K = y$ in the x -momentum equation and $\phi^K = x$ in the y -momentum equation in the weak formulation (18.17) of the Euler equations:

$$\frac{d}{dt} \int_K \rho v_1 y \, dV - \int_K f_{2,2} \, dV + \int_{\partial K} \bar{f}_2 y \, dS = 0, \quad (18.30)$$

$$\frac{d}{dt} \int_K \rho v_2 x \, dV - \int_K f_{1,3} \, dV + \int_{\partial K} \bar{f}_3 x \, dS = 0, \quad (18.31)$$

where \bar{f}_2 and \bar{f}_3 are the momentum components of the numerical flux function. By subtracting the above equations and using $f_{1,3} = f_{2,2} = \rho v_1 v_2$ we get the angular momentum equation of DG, viz.

$$\frac{d}{dt} \int_K L y \, dV + \int_{\partial K} (\bar{f}_3 x - \bar{f}_2 y) \, dS. \quad (18.32)$$

This equation is consistent with the exact equation (18.29), and hence DG schemes of at least second order accuracy are angular momentum conserving. However, there is one caveat to this inherent feature of DG. In non-smooth regions of the

solution, a limiting scheme has to be applied, which can slightly modify the angular momentum within a cell and hence lead to a violation of manifest angular momentum conservation. This is also the case for the simple limiters we shall adopt and describe in the subsequent section.

19

Slope limiting

The choice of the slope-limiting procedure can have a large effect on the quality of a hydro scheme, as we will demonstrate in Section 21.2. Often different limiters and configurations represent a trade-off between dissipation and oscillations, and furthermore, the optimal slope limiter is highly problem dependent. Consequently, the challenge consists of finding a limiting procedure which delivers good results for a vast range of test problems. For the DG scheme, this proves to be even more difficult. The higher order terms of the solution should be discarded at shocks and contact discontinuities if needed, while at the same time no clipping of extrema should take place in smooth regions, such that the full benefit of the higher order terms is ensured. In what follows, we discuss two different approaches for limiting the linear terms of the solution as well as a positivity limiter which asserts non-negativity of density and pressure.

19.1 Component-wise limiter

In order to reduce or completely avoid spurious oscillations, we have to confine possible over- and undershootings of the high order solution of a cell at cell boundaries compared to the cell average of neighbouring cells. For that purpose, the weights $w_{2,i}^K, w_{3,i}^K, w_{4,i}^K$ which are proportional to the slopes in the x -, y -, and z -directions,

respectively, are limited by comparing them to the difference of cell-averaged values, viz.

$$\begin{aligned}\tilde{w}_{2,i}^K &= \frac{1}{\sqrt{3}} \min\text{mod} \left(\sqrt{3}w_{2,i}^K, \beta(w_{1,i}^K - w_{1,i}^{W_K}), \beta(w_{1,i}^{E_K} - w_{1,i}^K) \right), \\ \tilde{w}_{3,i}^K &= \frac{1}{\sqrt{3}} \min\text{mod} \left(\sqrt{3}w_{3,i}^K, \beta(w_{1,i}^K - w_{1,i}^{S_K}), \beta(w_{1,i}^{N_K} - w_{1,i}^K) \right), \\ \tilde{w}_{4,i}^K &= \frac{1}{\sqrt{3}} \min\text{mod} \left(\sqrt{3}w_{4,i}^K, \beta(w_{1,i}^K - w_{1,i}^{B_K}), \beta(w_{1,i}^{T_K} - w_{1,i}^K) \right).\end{aligned}\tag{19.1}$$

Here, $\tilde{w}_{2,i}^K, \tilde{w}_{3,i}^K, \tilde{w}_{4,i}^K$ are the new weights, $W_K, E_K, S_K, N_K, B_K, T_K$ denote cell neighbours in the directions west, east, south, north, bottom, and top, respectively, and the minmod-function is defined as

$$\min\text{mod}(a, b, c) = \begin{cases} s \min(|a|, |b|, |c|) & s = \text{sign}(a) = \text{sign}(b) = \text{sign}(c) \\ 0 & \text{otherwise.} \end{cases}\tag{19.2}$$

Each component of the conserved variables, $i = 1, \dots, 5$, is limited separately and hence treated independently. The $\sqrt{3}$ -factors in equation (19.1) account for the scaling of the Legendre polynomial $\tilde{P}_1(\xi)$, and the parameter $\beta \in [0.5, 1]$ controls the amount of limiting. The choice of $\beta = 0.5$ corresponds to a total variation diminishing (TVD) scheme for a scalar problem but introduces more diffusion compared to $\beta = 1$. The latter value is less restrictive and may yield a more accurate solution.

If the limited weights are the same as the old weights, i.e. $\tilde{w}_{2,i}^K = w_{2,i}^K, \tilde{w}_{3,i}^K = w_{3,i}^K$, and $w_{4,i}^K = w_{4,i}^K$, we keep the original component of the solution $u_i^K(\mathbf{x}, t)$ of the cell. Otherwise, the component is set to

$$u_i^K(\mathbf{x}, t) = w_{1,i}^K + \tilde{w}_{2,i}^K \phi_2^K + \tilde{w}_{3,i}^K \phi_3^K + \tilde{w}_{4,i}^K \phi_4^K,\tag{19.3}$$

where terms with order higher than linear have been discarded. Alternative approaches to our simple limiter exist in the form of hierarchical limiters [e.g. Kuzmin, 2014, and references therein]. In these schemes the highest order terms are limited first before gradually lower order terms are modified one after another. In this way also higher order terms can be kept in the limiting procedure. Note that in equation (19.3) the cell-averaged values do not change, and thus the conservation of mass, momentum, and energy is unaffected. However, the limiter may modify the angular momentum content of a cell, implying that in our DG scheme it is manifestly conserved only in places where the slope limiter does not trigger. Clearly, it is a desirable goal for future work to design limiting schemes which can preserve angular momentum.

19.2 Characteristic limiter

An improvement over the component-wise limiting of the conserved variables can be achieved by limiting the characteristic variables instead. They represent advected quantities, and for the Euler equations we can define them locally by linearizing about the cell-averaged value.

The transformation matrices \mathcal{L}_x^K , \mathcal{L}_y^K , and \mathcal{L}_z^K are formed from the left eigenvectors of the flux Jacobian matrix based on the mean values of the conserved variables of cell K , $\bar{\mathbf{u}}^K = \mathbf{w}_1^K$. We list all matrices in the Appendix of the thesis. The slopes of the characteristic variables can then be obtained by the matrix-vector multiplications $\mathbf{c}_2^K = \mathcal{L}_x^K \mathbf{w}_2^K$, $\mathbf{c}_3^K = \mathcal{L}_y^K \mathbf{w}_3^K$, $\mathbf{c}_4^K = \mathcal{L}_z^K \mathbf{w}_4^K$, where \mathbf{w}_2^K , \mathbf{w}_3^K and \mathbf{w}_4^K denote the slopes of the conserved variables in the x -, y -, and z -directions, respectively. The transformed slopes are limited as in Section 19.1 with the minmod-limiting procedure, viz.

$$\begin{aligned}\tilde{\mathbf{c}}_2^K &= \frac{1}{\sqrt{3}} \text{minmod} \left(\sqrt{3} \mathbf{c}_2^K, \beta \mathcal{L}_x^K (\mathbf{w}_1^K - \mathbf{w}_1^{W_K}), \beta \mathcal{L}_x^K (\mathbf{w}_1^{E_K} - \mathbf{w}_1^K) \right), \\ \tilde{\mathbf{c}}_3^K &= \frac{1}{\sqrt{3}} \text{minmod} \left(\sqrt{3} \mathbf{c}_3^K, \beta \mathcal{L}_y^K (\mathbf{w}_1^K - \mathbf{w}_1^{S_K}), \beta \mathcal{L}_y^K (\mathbf{w}_1^{N_K} - \mathbf{w}_1^K) \right), \\ \tilde{\mathbf{c}}_4^K &= \frac{1}{\sqrt{3}} \text{minmod} \left(\sqrt{3} \mathbf{c}_4^K, \beta \mathcal{L}_z^K (\mathbf{w}_1^K - \mathbf{w}_1^{B_K}), \beta \mathcal{L}_z^K (\mathbf{w}_1^{T_K} - \mathbf{w}_1^K) \right).\end{aligned}\quad (19.4)$$

If the limited slopes of the characteristic variables are identical to the unlimited ones (i.e. $\tilde{\mathbf{c}}_2^K = \mathbf{c}_2^K$, $\tilde{\mathbf{c}}_3^K = \mathbf{c}_3^K$, and $\tilde{\mathbf{c}}_4^K = \mathbf{c}_4^K$), the original solution is kept. Otherwise, the new slopes of the conserved variables are calculated with the inverse transformation matrices $\mathcal{R}_x^K = (\mathcal{L}_x^K)^{-1}$, $\mathcal{R}_y^K = (\mathcal{L}_y^K)^{-1}$, and $\mathcal{R}_z^K = (\mathcal{L}_z^K)^{-1}$ via $\tilde{\mathbf{w}}_2^K = \mathcal{R}_x^K \tilde{\mathbf{c}}_2^K$, $\tilde{\mathbf{w}}_3^K = \mathcal{R}_y^K \tilde{\mathbf{c}}_3^K$, and $\tilde{\mathbf{w}}_4^K = \mathcal{R}_z^K \tilde{\mathbf{c}}_4^K$. In this case, the higher order terms of the solution are set to zero and the limited solution in the cell becomes

$$\mathbf{u}^K(\mathbf{x}, t) = \mathbf{w}_1^K + \tilde{\mathbf{w}}_2^K \phi_2^K + \tilde{\mathbf{w}}_3^K \phi_3^K + \tilde{\mathbf{w}}_4^K \phi_4^K. \quad (19.5)$$

The difference between the component-wise limiting of the conserved variables (Section 19.1) and the limiting of the more natural characteristic variables is demonstrated with a shock tube simulation in Section 21.2.

19.3 Total variation bounded limiting

The limiters discussed so far can effectively reduce overshootings and oscillations; however, they can potentially also trigger at smooth extrema and then lead to a loss of higher order information. Considering the goals of a higher order DG scheme, this is a severe drawback that can negatively influence the convergence rate of our DG code. In order to avoid a clipping of the solution at smooth extrema, the minmod

limiter in Sections 19.1 and 19.2 can be replaced by a bounded version [Cockburn and Shu, 1998], viz.

$$\text{minmodB}(a, b, c) = \begin{cases} a & \text{if } |a| \leq M(\Delta x^K)^2 \\ \text{minmod}(a, b, c) & \text{otherwise.} \end{cases} \quad (19.6)$$

Here, M is a free parameter which is related to the second derivative of the solution. The ideal choice for it can vary for different test problems. Furthermore, in the above ansatz, the amount of limiting depends on the resolution since $|a| \propto \Delta x^K$, which is not the case for the right hand side of the inequality. For reasons of simplicity and generality, we would however like to use fixed limiter parameters without explicit resolution dependence for the tests presented in this work. Hence we define $\tilde{M} = M\Delta x^K$, and use a constant value for \tilde{M} to control the strength of the bounding applied to the minmod limiter. With the minmod-bounded limiting approach, the high accuracy of the solution in smooth regions is retained while oscillations, especially in post-shock regions, can be eliminated.

19.4 Positivity limiting

When solving the equations of hydrodynamics for extreme flows, care has to be taken in order to avoid negative pressure or density values within the cells and at cell interfaces. A classical example is high Mach number flows, for which in the pre-shock region the total energy is dominated by the kinetic energy. Because the pressure is calculated via the difference between total and kinetic energies, it can easily become negative in numerical treatments without special precautions.

The use of an ordinary slope limiter tends to be only of limited help in this situation, and only delays the blow-up of the solution. In fact, it turns out that even with TVD limiting and arbitrarily small timesteps, it is not guaranteed in general that unphysical negative values are avoided. Nevertheless, it is possible to construct positivity preserving FV and DG schemes, which are accurate at the same time. Remarkably, the latter means that high order accuracy can be retained in smooth solutions and furthermore the total mass, momentum, and energy in each cell are conserved. For our DG code, we adopt the positivity limiting implementation following Zhang and Shu [2010].

For constructing this limiter, a quadrature rule including the boundary points of the integration interval is needed. One possible choice consists of m -point Gauss-Lobatto-Legendre (GLL) quadrature rules (Section 3.2.4), which are exact for polynomials of degree $k \leq 2m - 3$. Let $\xi_1^{1D} < \dots < \xi_{k+1}^{1D} \in [-1, 1]$ be the one-dimensional Gauss quadrature points and $\hat{\xi}_1^{1D} < \dots < \hat{\xi}_m^{1D} \in [-1, 1]$ the GLL quadrature points. Quadrature rules for the domain $[-1, -1]^3$ which include the interface

Gauss quadrature points can be constructed by tensor products of Gauss and GLL quadrature points:

$$\begin{aligned} S_x &= \{(\hat{\xi}_r^{1D}, \xi_s^{1D}, \xi_t^{1D}) : 1 \leq r \leq m, 1 \leq s \leq k+1, 1 \leq t \leq k+1\}, \\ S_y &= \{(\xi_r^{1D}, \hat{\xi}_s^{1D}, \xi_t^{1D}) : 1 \leq r \leq k+1, 1 \leq s \leq m, 1 \leq t \leq k+1\}, \\ S_z &= \{(\xi_r^{1D}, \xi_s^{1D}, \hat{\xi}_t^{1D}) : 1 \leq r \leq k+1, 1 \leq s \leq k+1, 1 \leq t \leq m\}. \end{aligned} \quad (19.7)$$

The union $S = S_x \cup S_y \cup S_z$ of these sets constitutes quadrature points of a rule which is exact for polynomials of degree k and furthermore contains all the points for which we calculate fluxes in equation (18.20). It can be shown that if the solution is positive at these union quadrature points, the solution averaged after one explicit Euler step will stay positive if a positivity preserving flux calculation and an adequate timestep are used.

The positivity limiter is operated as follows. For a DG scheme with polynomials of maximum order k , choose the smallest integer m such that $m \geq (k+3)/2$. Carry out the following computations for every cell K . First, determine the minimum density at the union quadrature points,

$$\rho_{\min}^K = \min_{\xi \in S} \rho^K(\xi). \quad (19.8)$$

Use this minimum density for calculating the factor

$$\theta_1^K = \min \left\{ \left| \frac{\bar{\rho}^K - \epsilon}{\bar{\rho}^K - \rho_{\min}^K} \right|, 1 \right\}, \quad (19.9)$$

where $\epsilon \approx 10^{-10}$ is a small number representing the target floor for the positivity limiter. Then, modify the higher order terms of the density by multiplying the corresponding weights with the calculated factor,

$$w_{j,1}^K \leftarrow \theta_1^K w_{j,1}^K, \quad j = 2, \dots, N(k). \quad (19.10)$$

At this point, the density at the union quadrature points is positive ($\geq \epsilon$), and as desired, the mean density $\bar{\rho}^K = w_{1,1}^K$ and therefore the total mass has not been changed.

In order to enforce pressure positivity at the union quadrature points $\xi \in S$, we compute the factor

$$\theta_2^K = \min_{\xi \in S} \tau^K(\xi), \quad (19.11)$$

with

$$\tau^K(\xi) = \begin{cases} 1 & \text{if } p^K(\xi) \geq \epsilon \\ \tau_* & \text{such that } p^K(\mathbf{u}^K(\xi) + \tau_*(\mathbf{u}^K(\xi) - \bar{\mathbf{u}}^K)) = \epsilon, \end{cases} \quad (19.12)$$

and limit all higher order terms ($j \geq 2$) for all components of the state vector:

$$w_{j,i}^K \leftarrow \theta_2^K w_{j,i}^K, \quad j = 2, \dots, N(k), \quad i = 1, \dots, 5. \quad (19.13)$$

The idea of the second case of equation (19.12) is to calculate a modification factor τ_* for the higher order terms $\mathbf{u}^K(\boldsymbol{\xi}) - \bar{\mathbf{u}}^K$, such that the new pressure at the quadrature point equals ϵ . Furthermore, this second case represents a quadratic equation in τ_* , which has to be solved carefully by minimizing round-off errors. In our implementation, we improve the solution for τ_* iteratively with a small number of Newton-Raphson iterations.

The outlined method ensures positivity of pressure and density of the mean cell state after the subsequent timestep under several conditions. First, a positivity preserving flux calculation has to be used, e.g. the Godunov flux or the Harten-Lax-van Leer flux is suitable. Secondly, when adopting an RK method, it should be strong stability preserving (SSP); these methods are convex combinations of explicit Euler methods, and hence a number of relevant properties and proofs valid for Euler's method also hold for these higher order time integrators. Lastly, the local timestep for the DG scheme has to be set to

$$\Delta t^K = \text{CFL} \cdot \min \left\{ \frac{1}{2k+1}, \frac{\hat{\omega}_1^{1D}}{2} \right\} \left(\frac{|v_1^K| + c^K}{\Delta x_1^K} + \frac{|v_2^K| + c^K}{\Delta x_2^K} + \frac{|v_3^K| + c^K}{\Delta x_3^K} \right)^{-1}, \quad (19.14)$$

where $\hat{\omega}_1^{1D}$ is the first GLL weight. Compared to Zhang and Shu [2010] we obtain an additional factor of 1/2 since our reference domain is $[-1, 1]$ and therefore the sum of the GLL weights is $\sum_{q=1}^m \hat{\omega}_q^{1D} = 2$. The first GLL weight is $\hat{\omega}_1^{1D} = 1$ for our second order DG (DG-2) method ($k = 1$) and $\hat{\omega}_1^{1D} = 1/3$ for our third order and fourth order method ($k = 2, 3$). Depending on the order, the timestep with positivity limiting can be slightly more restrictive. We conclude this section by remarking that the positivity limiting procedure is a local operation and does not introduce additional inter-cell communication in the DG scheme.

20

DG with AMR

Many astrophysical applications involve a high dynamic range in spatial scales. In grid-based codes, this multi-scale challenge can be tackled with the AMR technique. In this approach, individual cells are split into a set of smaller subcells if appropriate (see Fig. 20.2 for a sketch of the two-dimensional case), thereby increasing the resolution locally. We adopt a tree-based AMR implementation where cubical cells are split into eight subcells when a refinement takes place. This allows a particularly flexible adaption to the required spatial resolution.

Our implementation largely follows that in the RAMSES code [Teyssier, 2002]. The tree is organized into cells and nodes. The root node encompasses the whole simulation domain and is designated as level $l = 0$. A node at level l always contains eight children, which can be either another node on a smaller level $l' = l + 1$, or a leave cell. In this way, the mesh of leave cells is guaranteed to be volume filling. An example of such an AMR mesh in 2D is shown in Fig. 20.1.

To distribute the work load onto many MPI processes, the tree is split into an upper part, referred to as ‘top tree’ in the following, and branches that hang below individual top tree nodes. Every MPI process stores a full copy of the top tree, whereas each of the lower branches is in principle stored only on one MPI rank. However, some of the branch data are replicated in the form of a ghost layer around each local tree to facilitate the exchange of boundary information. The mesh can be refined and structured arbitrarily, with only one restriction. The level jump between a cell and its 26 neighbours has to be at most ± 1 . This implies that a cell can either have one or four split cells as direct neighbours in a given direction. In order to fulfil this level jump condition, additional cells may have to be refined beyond those where the physical criterion demands a refinement.

To simplify bookkeeping, we store for each cell and node the indices of the father node and the six neighbouring cells or nodes. In case of a split neighbour, the index

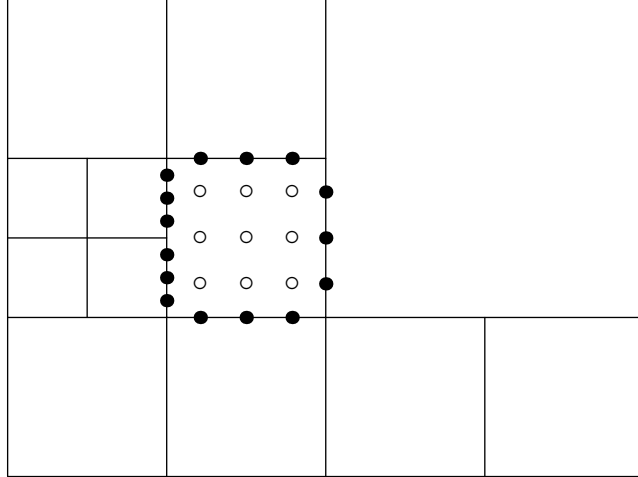


Figure 20.1: Quadrature points of an AMR boundary cell in our 2D, third order ($k = 2$) version of the code. Interfaces with neighbouring cells on a finer AMR level are integrated with the quadrature points of the smaller cells. In this way no accuracy is lost at AMR boundaries and the order of our DG code is unaffected.

of the node at the same level is stored instead. To make the mesh smoother and guarantee a buffer region around cells which get refined due to the physical refinement criterion, additional refinement layers are added as needed. In the AMR simulations presented in this work, we use one extra layer of refined cells around each cell flagged by the physical criterion for refinement.

20.1 Refinement criterion

Many different refinement strategies can be applied in an AMR code, for example, the refinement and derefinement criterion may aim to keep the mass per cell approximately constant. Another common strategy for refinement is to focus on interesting regions such as shocks, contact discontinuities, or fluid instabilities. In this work, we apply the mesh refinement simply at locations where the density gradient is steep. To be precise, cell K is refined if the following criterion is met:

$$\max(w_{2,0}^K, w_{3,0}^K, w_{4,0}^K) > \alpha \cdot w_t. \quad (20.1)$$

Here, $w_{2,0}^K$, $w_{3,0}^K$, and $w_{4,0}^K$ are the density changes divided by $\sqrt{3}$ along the x -, y -, and z - directions, respectively. The refinement is controlled by the target slope parameter w_t and a range factor α . We have introduced the latter with the purpose of avoiding an oscillating refinement-derefinement behaviour. In this work, we adopt the values $w_t = 0.01$ and $\alpha = 1.1$. The reason for using the density changes instead of the physical slopes is that in this way a runaway refinement can be avoided.

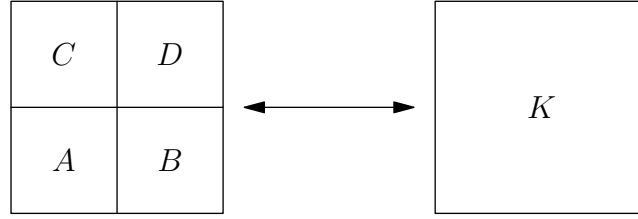


Figure 20.2: Refinement (arrow to the left) and coarsening (arrow to the right) of a cell in the 2D version of TENET. In both operations the solution on the new cell structure is inferred by an L^2 -projection of the current polynomial solution. In doing so, no information is lost in a refinement, and as much information as possible is retained in a derefinement.

Leave nodes of the AMR tree are kept refined if the following equation is true:

$$\max(w_{2,0}^L, w_{3,0}^L, w_{4,0}^L) > \frac{1}{\alpha} \cdot w_t. \quad (20.2)$$

The weights on the left hand side are the leave node weights calculated with a projection of the eight subcell solutions. If inequality (20.2) does not hold, the node gets derefined and the eight subcells are merged into one.

20.2 Mesh refinement

Let $K = \{\boldsymbol{\xi} | \boldsymbol{\xi} \in [-1, 1]^3\}$ be the cell which we want to refine into eight subcells, and A, B, C, \dots, H denote the daughter cells, as illustrated in Fig. 20.2 for the two-dimensional case. The refinement can be achieved without higher order information loss if the solution polynomial of the coarse cell is correctly projected onto the subcells. In the following, we outline the procedure for obtaining the solution on subcell $A = \{\boldsymbol{\xi} | \boldsymbol{\xi} \in [-1, 0] \times [-1, 0] \times [-1, 0]\}$, the calculations for the other subcells are done in an analogous way.

The weights of the solution on subcell A are obtained by solving the minimization problem

$$\min_{\{w_{i,i}^A\}_i} \int_A (u_i^K - u_i^A)^2 \, dV, \quad i = 1, \dots, 5, \quad (20.3)$$

where $\mathbf{u}^K = \sum_{l=1}^{N(k)} \mathbf{w}_l^K \phi_l^K$ is the given solution on the coarse cell and $\mathbf{u}^A = \sum_{l=1}^{N(k)} \mathbf{w}_l^A \phi_l^A$ are the polynomials of the conserved variables on cell A we are looking for. The minimization ansatz (20.3) is solved by projecting the solution of the coarse cell onto the basis functions of subcell A ,

$$\mathbf{w}_j^A = \frac{1}{|A|} \int_A \mathbf{u}^K \phi_j^A \, dV, \quad j = 1, \dots, N(k). \quad (20.4)$$

We can plug in the solution \mathbf{u}^K and move the time- but not space-dependent weights in front of the integral,

$$w_{j,i}^A = \sum_{l=1}^{N(k)} w_{l,i}^K \frac{1}{|A|} \int_A \phi_l^K \phi_j^A dV, \quad i = 1, \dots, 5, j = 1, \dots, N(k). \quad (20.5)$$

If we define the projection matrix

$$(\mathbf{P}_A)_{j,l} = \frac{1}{|A|} \int_A \phi_l^K \phi_j^A dV, \quad (20.6)$$

and introduce for each conserved variable $i = 1, \dots, 5$ the weight vector $\hat{\mathbf{w}}_i = (w_{1,i}, w_{2,i}, \dots, w_{N,i})^\top$, the weights of the subcell solution \mathbf{u}^A are simply given by the matrix-vector multiplications

$$\hat{\mathbf{w}}_i^A = \mathbf{P}_A \hat{\mathbf{w}}_i^K \quad i = 1, \dots, 5. \quad (20.7)$$

The matrix (20.6) can be computed exactly by transforming the integral to the reference domain of cell A ,

$$(\mathbf{P}_A)_{j,l} = \frac{1}{8} \iiint_{[-1,1]^3} \phi_j \left(\frac{\xi_1 - 1}{2}, \frac{\xi_2 - 1}{2}, \frac{\xi_3 - 1}{2} \right) \phi_l(\xi_1, \xi_2, \xi_3) d\xi, \quad (20.8)$$

and applying a Gaussian quadrature rule with $(k + 1)^3$ points. The projection integrals for the other subcells (B, C, \dots, H) are given in the Appendix of the thesis. The refinement matrices are the same for all the cells; we calculate them once in the initialization of our DG code.

20.3 Mesh derefinement

When the eight cells A, B, \dots, H are merged into one coarser cell K , some information is unavoidably lost. Nevertheless, in order to retain as much accuracy as possible, the derefinement should again be carried out by means of an L^2 -projection,

$$\min_{\{w_{l,i}^K\}_l} \int_A (u_i^K - u_i^A)^2 dV + \int_B (u_i^K - u_i^B)^2 dV + \dots + \int_H (u_i^K - u_i^H)^2 dV, \quad i = 1, \dots, 5. \quad (20.9)$$

Here, $\mathbf{u}^A, \mathbf{u}^B, \dots, \mathbf{u}^H$ denote the given solutions on the cells to be derefined, and \mathbf{u}^K is the solution on the coarser cell we are looking for. The minimization problem is solved by the weights

$$\mathbf{w}_j^K = \frac{1}{|K|} \left(\int_A \mathbf{u}^A \phi_j^K dV + \int_B \mathbf{u}^B \phi_j^K dV + \dots + \int_H \mathbf{u}^H \phi_j^K dV \right),$$

$$j = 1, \dots, N(k). \quad (20.10)$$

We insert the solutions on the cells A, B, \dots, H and use the definition of the refinement matrices,

$$\begin{aligned} w_{j,i}^K &= \frac{1}{|K|} \left(\int_A \sum_{l=1}^{N(k)} w_{l,i}^A \phi_l^A \phi_j^K dV + \dots + \int_H \sum_{l=1}^{N(k)} w_{l,i}^H \phi_l^H \phi_j^K dV \right) \\ &= \frac{1}{|K|} \left(\sum_{l=1}^{N(k)} w_{l,i}^A |A| (\mathbf{P}_A^\top)_{j,l} + \dots + \sum_{l=1}^{N(k)} w_{l,i}^H |H| (\mathbf{P}_H^\top)_{j,l} \right), \\ i &= 1, \dots, 5 \quad , \quad j = 1, \dots, N(k). \end{aligned} \quad (20.11)$$

If we again use the vector $\hat{\mathbf{w}}_i^K$ consisting of the weights of the solution for the conserved variable i , the weights of the solution on cell K can be computed with the transposed refinement matrices,

$$\hat{\mathbf{w}}_i^K = \frac{1}{8} \left(\mathbf{P}_A^\top \hat{\mathbf{w}}_i^A + \mathbf{P}_B^\top \hat{\mathbf{w}}_i^B + \dots + \mathbf{P}_H^\top \hat{\mathbf{w}}_i^H \right), \quad i = 1, \dots, 5. \quad (20.12)$$

Here we have used the fact that the cells to be merged have the same volume, $|A| = |B| = \dots = |H| = \frac{1}{8}|K|$.

The refinement and derefinition matrices are orthogonal, i.e. $\mathbf{P}_A \mathbf{P}_A^\top = \mathbf{I}$. Thus, if a cell is refined and immediately thereafter derefined, the original solution on the coarse cell is restored. This property is desirable for achieving a stable AMR algorithm; in our approach it can be shown explicitly by inserting the subcell weights from equation (20.7) of every subcell in equation (20.12). While the refinement of eight cells into a subcell preserves the exact shape of the solution, this is in general not the case for the derefinition. For example, if the solution is discontinuous across the eight subcells, it cannot be represented in a polynomial basis on the coarse cell. After derefining we limit the solution before further calculations are carried out, especially for asserting positivity.

20.4 Limiting with AMR

In the case of AMR boundary cells, the limiting procedure has to be well defined. For the limiting of cell K , the slope limiters described in Sections 19.1 and 19.2 need the average values of the neighbouring cells in each direction in order to adjust the

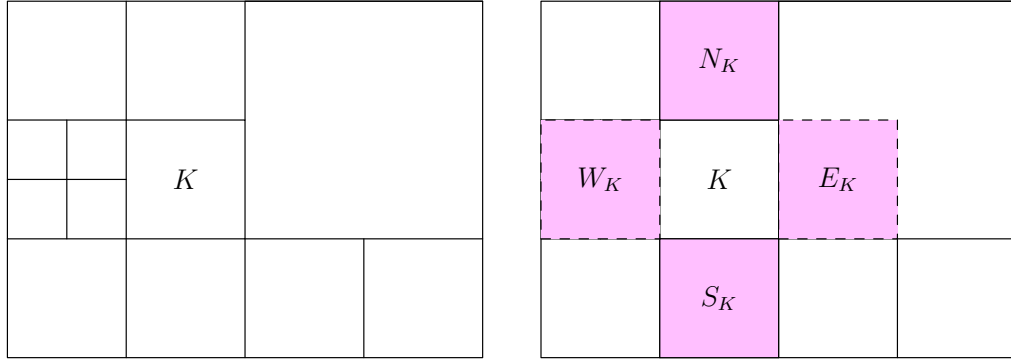


Figure 20.3: Definition of neighbours in the slope-limiting procedure at AMR boundaries, shown for the 2D version of TENET. The neighbours of cell K on the left are on a finer level. For the slope limiting of cell K , the node weights are used, which are calculated by projecting the solutions of the subcells onto the encompassing node volume W_K . The right neighbour of cell K on the other hand is here a coarser cell; in this case, the solution of this neighbour has to be projected onto the smaller volume E_K . In the slope limiting routine of TENET, this is only done in a temporary fashion without actually modifying the mesh.

slope of the conserved variables. However, if the cell neighbours are on a different AMR level, they are smaller or larger compared to the cell to limit, and a single neighbouring cell in a specific direction is not well defined.

Fortunately, there is a straightforward way in DG to cope with this problem. If a neighbouring cell is on a different level, the polynomials of the neighbour are projected onto a volume which is equal to the volume of the cell to refine, see Fig. 20.3. This can be done with the usual refinement and derefinement operations as described in Sections 20.2 and 20.3. By doing so, the limiting at AMR boundaries can be reduced to the limiting procedure on a regular grid. In AMR runs, we also slightly adjust the positivity limiting scheme. For cells which have neighbours on a higher level, the positivity limiter is not only applied at the locations of the union quadrature points (equation 19.7), but additionally at the quadrature points of the cell interfaces to smaller neighbours. By doing so, negative values in the initial conditions of the Riemann problems are avoided, which have to be solved in the integration of these interfaces.

20.5 Main simulation loop

Our new TENET code has been developed as an extension of the AREPO code [Springel, 2010], allowing us to reuse AREPO’s input-output routines, domain decomposition, basic tree infrastructure, neighbour search routines, and gravity solver. This also helps to make our DG scheme quickly applicable in many science applications. We briefly discuss the high-level organization of our code and the structure of the main loop in a schematic way, focusing on the DG part.

1. Compute and store quadrature points and weights, basis function values, and projection matrices.
2. Set the initial conditions by means of an L^2 -projection.
3. Apply slope limiter.
4. Apply positivity limiter.
5. While $t < t_{\max}$:
 - a) Compute timestep.
 - b) For every RK stage:
 - i. Calculate \mathbf{R}_K of the differential equation (18.23) (inner and outer integral).
 - ii. Update solution to the next RK stage.
 - iii. Update node data.
 - iv. Apply slope limiter.
 - v. Apply positivity limiter.
 - c) Do refinement and derefinement.
 - d) $t = t + \Delta t$.

The basis function values and the quadrature data are computed in a general way for arbitrary spatial order as outlined in Appendices 3.2.2, 3.2.3, and 3.2.4. For the time integration, our code can be provided with a general Butcher tableau (Table A1), specifying the desired RK method. By keeping these implementations general, the spatial order and time integrator can be changed conveniently.

The first step of an RK stage consists of calculating \mathbf{R}_K of the differential equation (18.23), including possible source terms. This involves computing the inner integral by looping over the cells and the outer integral by looping over the cell interfaces. After updating the solution weights for every cell, the hydrodynamical quantities on the AMR nodes have to be updated, such that they can be subsequently accessed during the slope-limiting procedure (Section 20.4). After the RK step, all cells are checked for refinement and derefinement, and the mesh is adjusted accordingly.

21

Validation

In this section, we discuss various test problems which are either standard tests or chosen for highlighting a specific feature of the DG method. For most of the test simulations, we compare the results to a traditional second order FV method (FV-2). For definiteness, we use the AREPO code to this end, with a fixed Cartesian grid and its standard solver as described in Springel [2010]. The latter consists of a second order unsplit Godunov scheme with an exact Riemann solver and a non-TVD slope limiter. Recently, some modifications to the FV solver of AREPO have been introduced for improving its convergence properties [Pakmor et al., 2016] when the mesh is dynamic. However, for a fixed Cartesian grid, this does not make a difference and the old solver used here performs equally well.

There are several important differences between FV and DG methods. In an FV scheme, the solution is represented by piecewise constant states, whereas in DG the solution within every cell is a polynomial approximation. Moreover, in FV a reconstruction step has to be carried out in order to recreate higher order information. Once the states at the interfaces are calculated, numerical fluxes are computed and the mean cell values updated. In the DG method, no higher order information is discarded after completion of a step and therefore no subsequent reconstruction is needed. DG directly solves also for the higher order moments of the solution and updates the weights of the basis functions in every cell accordingly.

For all the DG tests presented in this section, we use an SSP RK time integrator of an order consistent with the spatial discretization, and the fluxes are calculated with the HLLC approximate Riemann solver. Furthermore, if not specified otherwise, we use for all tests the positivity limiter in combination with the characteristic limiter in the bounded version with the parameters $\beta = 1$ and $\tilde{M} = 0.5$. Specifically, we only deviate from this configuration when we compare the limiting of the characteristic variables to the limiting of the conserved variables in the shock tube test in Sec-

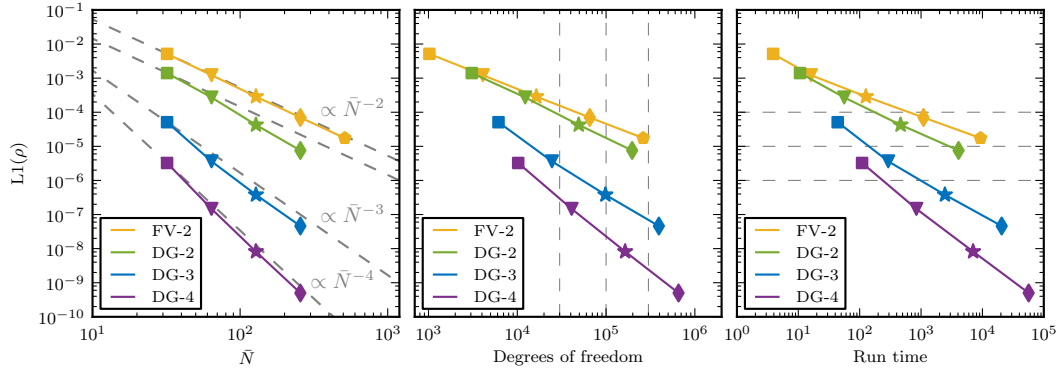


Figure 21.1: Left-hand panel: L1 error norm as a function of linear resolution for the two-dimensional isentropic vortex test. Each data point corresponds to a simulation, and different colours indicate the different methods. We find a convergence rate as expected (dashed line) or slightly better for all schemes, indicating the correctness of our DG implementation. Middle panel: the same simulation errors as a function of degrees of freedom (DOF), which is an indicator for the memory requirements. For our test runs the higher order DG methods are more accurate at the same number of DOF. Right-hand panel: L1 error norm versus the measured run time of the simulations. The second order FV implementation (FV-2) and the second order DG (DG-2) realization are approximately equally efficient in this test, i.e. a given precision can be obtained with a similar computational cost. In comparison, the higher order methods can easily be faster by more than an order of magnitude for this smooth problem. This illustrates the fact that an increase of order (p -refinement) of the numerical scheme can be remarkably more efficient than a simple increase of grid resolution (h -refinement).

tion 21.2, and when the angular momentum conservation of our code is demonstrated in Section 21.5 with the cold Keplerian disc problem. The higher order efficiency of our DG code is quantified in the test problem of Section 21.1, and the 3D version of our code is tested in a Sedov-Taylor blast wave simulation in Section 21.3. In Section 21.4, we show that advection errors are much smaller for DG compared to FV-2. Finally, the AMR capabilities of TENET are illustrated with a high-resolution Kelvin-Helmholtz (KH) instability test in Section 21.6.

21.1 Isentropic vortex

In this first hydrodynamical test problem, we verify the correctness of our DG implementation by measuring the convergence rate towards the analytic solution for different orders of the scheme. Additionally, we investigate the precision of the FV-2 and DG schemes as a function of computational cost.

An elementary test for measuring the convergence of a hydrodynamical scheme is the simulation of one-dimensional travelling waves at different resolutions [e.g. Stone et al., 2008]. However, the DG scheme performs so well in this test, especially at higher order, that the accuracy is very quickly limited by machine precision, making the travelling sound wave test impractical for convergence studies of our DG implementation. We hence use a more demanding setup, the stationary and

isentropic vortex in two dimensions [Yee et al., 1999]. The primitive variables density, pressure, and velocities in the initial conditions are

$$\begin{aligned} p(r) &= \rho^\gamma, \\ \rho(r) &= \left[1 - \frac{(\gamma - 1)\beta^2}{8\gamma\pi^2} \exp(1 - r^2) \right]^{\frac{1}{\gamma-1}}, \\ v_x(r) &= -(y - y_0) \frac{\beta}{2\pi} \exp\left(\frac{1 - r^2}{2}\right), \\ v_y(r) &= (x - x_0) \frac{\beta}{2\pi} \exp\left(\frac{1 - r^2}{2}\right), \end{aligned}$$

with $\beta = 5$, and the adiabatic index $\gamma = 7/5$. With this choice for the primitive variables, the centrifugal force at each point is exactly balanced by the pressure gradient pointing towards the centre of the vortex, yielding a time-invariant situation.

The vortex is smooth and stationary, and every change during the time integration with our numerical schemes can be attributed mainly on numerical truncation errors. In order to break the spherical symmetry of the initial conditions, we additionally add a mild velocity boost of $v_b = (1, 1)$ everywhere. The simulation is carried out in the periodic domain $(x, y) \in [0, 10]^2$ with the centre of the vortex at $(x_0, y_0) = (5, 5)$ and run until $t = 10$, corresponding to one box crossing of the vortex. We compare the obtained numerical solution with the analytic solution, which is given by the initial conditions. The error is measured by means of the L1 density norm, which we define for the FV method as

$$\text{L1} = \frac{1}{N} \sum_i |\rho'_i - \rho_i^0|, \quad (21.1)$$

where N is the total number of cells, and ρ'_i and ρ_i^0 denote the density in cell i in the final state and in the initial conditions, respectively. This norm should be preferred over the L2 norm, since it is more restrictive.

For DG the error is inferred by calculating the integral

$$\text{L1} = \frac{1}{V} \int_V |\rho'(x, y) - \rho^0(x, y)| dV, \quad (21.2)$$

with the density solution polynomial $\rho'(x, y)$ and the analytic expression $\rho^0(x, y)$ of the initial conditions. We compute the integral numerically with accuracy of order $p + 2$, where p is the order of the DG scheme. In this way, we assert that the error measurement cannot be dominated by errors in the numerical integration of the error norm.

The result of our convergence study is shown in the left-hand panel of Figure 21.1. For every method we run the simulation with several resolutions, indicated by different symbols. The L1 errors decrease with resolution and meet the expected convergence rates given by the dashed lines, pointing towards the correct implementation of

our DG scheme. At a given resolution, the DG-2 code achieves a higher precision compared to the FV-2 code, reflecting the larger number of flux calculations involved in DG. The convergence rates are equal, however, as expected. The higher order DG formulations show much smaller errors, which also drop more rapidly with increasing resolution.

In the middle panel of Fig. 21.1, we show the same plot but substitute the linear resolution on the horizontal axis with the number of degrees of freedom (DOF). This number is closely related to the used memory of the simulation. According to equation (18.10) the number of DOF per cell for 3D DG is $1/6p(p+1)(p+2)$. For two-dimensional simulations one obtains $1/2p(p+1)$ DOF for every cell. The total number of DOF is thus given by N , $3N$, $6N$, and $10N$ for FV-2, DG-2, DG-3, and DG-4, respectively. At the same number of DOF, as marked by the grey dashed lines, the FV-2 and DG-2 code achieve a similar accuracy. Moreover, when using higher order DG methods the precision obtained per DOF is significantly increased in this test problem.

In the right-hand panel of Fig. 21.1, we compare the efficiencies of the different schemes by plotting the obtained precision as a function of the measured run time, which directly indicates the computational cost.¹ By doing so, we try to shed light on the question of the relative speed of the methods (or computational cost required) for a given precision. Or alternatively, this also shows the precision attained at a fixed computational cost.

For the following discussion, we want to remark that the isentropic vortex is a smooth 2D test problem; slightly different conclusions may well hold for problems involving shocks or contact discontinuities, as well as for 3D tests. The FV-2 method consumes less time than the DG-2 method when runs with equal resolution are compared. On the other hand, the error of the DG-2 scheme is smaller. As Fig. 21.1 shows, when both methods are quantitatively compared at equal precision, the computational cost is essentially the same; hence, the overall efficiency of the two second order schemes is very similar for this test problem. Interestingly, the 64^2 -cell FV-2 run (yellow triangle) and the 32^2 -cell DG-2 simulation (green square) give an equally good result at almost identical runtime.

However, the higher order schemes (DG-3, DG-4) show a significant improvement over the second order methods in terms of efficiency, i.e. prescribed target accuracies as indicated by the grey dashed lines can be reached much faster by them. In particular, the third order DG (DG-3) scheme performs clearly favourably over the second order scheme. The fourth order DG (DG-4) method uses the SSP RK-4 time integrator, which has already five stages, in contrast to the time integrators of the lower order methods, which have the optimal number of p stages for order p . Moreover, the timestep for the higher order methods is smaller; according to equation (18.24), it is proportional to $1/(2p-1)$. Nevertheless, DG-4 is the most

¹ The simulations were performed with four MPI-tasks on a conventional desktop machine, shown is the wall-clock time in seconds. A similar trend could also be observed for larger parallel environments; while DG is a higher order scheme, due to its discontinuous nature the amount of required communication is comparable to that of a traditional second order FV method.

efficient method for this test. Consequently, for improving the calculation efficiency in smooth regions, the increase of the order of the scheme (p -refinement) should be preferred over the refinement of the underlying grid (h -refinement).

For FV methods, this principle is cumbersome to achieve from a programming point of view, because for every order a different reconstruction scheme has to be implemented, and moreover, there are no well-established standard approaches for implementing arbitrarily higher order FV methods. Higher order DG methods on the other hand can be implemented straightforwardly and in a unified way. If the implementation is kept general, changing the order of the scheme merely consists of changing a simple parameter. In principle, it is also possible to do the p -refinement on the fly. In this way, identified smooth regions can be integrated with large cells and high order, whereas regions close to shocks and other discontinuities can be resolved with many cells and a lower order integration scheme.

21.2 Shock tube

Due to the non-linearity of the Euler equations, the characteristic wave speeds of the solution depend on the solution itself. This dependence and the fact that the Euler equations do not diffuse momentum can lead to an inevitable wave steepening, ultimately producing wave breaking and the formation of mathematical discontinuities from initially smooth states [Toro, 2009, and references herein]. Such hydrodynamic shocks are omnipresent in many hydrodynamical simulations, particularly in astrophysics. The proper treatment of these shocks is hence a crucial component of any numerical scheme for obtaining accurate hydrodynamical solutions.

In a standard FV method based on the RSA approach, the solution is averaged within every cell at the end of each timestep. The solution is then represented through piecewise constant states that can have arbitrarily large jumps at interfaces, consequently allowing shocks to be captured by construction. Similarly, the numerical solution in a DG scheme is discontinuous across cell interfaces, allowing for an accurate representation of hydrodynamic shocks. We demonstrate the shock-capturing abilities of our code with a classical Sod shock tube problem [Sod, 1978] in the two-dimensional domain $(x, y) \in [0, 1]^2$ with 64^2 cells. The initial conditions consist of a constant state on the left, $\rho_l = 1$, $p_l = 1$, and a constant state on the right, $\rho_r = 0.125$, $p_r = 0.1$, separated by a discontinuity at $x = 0.5$. The velocity is initially zero everywhere, and the adiabatic index is chosen to be $\gamma = 7/5$.

In Figure 21.2 we show the numerical and analytic solution at $t_{\text{end}} = 0.228$ for DG-3, and for different limiting strategies. Several problems are apparent when a true discontinuity is approximated with higher order functions, i.e. the rapid convergence of the approximation at the jump is lost, the accuracy around the discontinuity is reduced, and spurious oscillations are introduced [Gibbs phenomenon, see for example Arfken and Weber, 2013]. In the bottom-left panel of Fig. 21.2, we can clearly observe oscillations at both discontinuities, the contact and the shock. Here the result is obtained without any slope limiter; merely the positivity limiter slightly

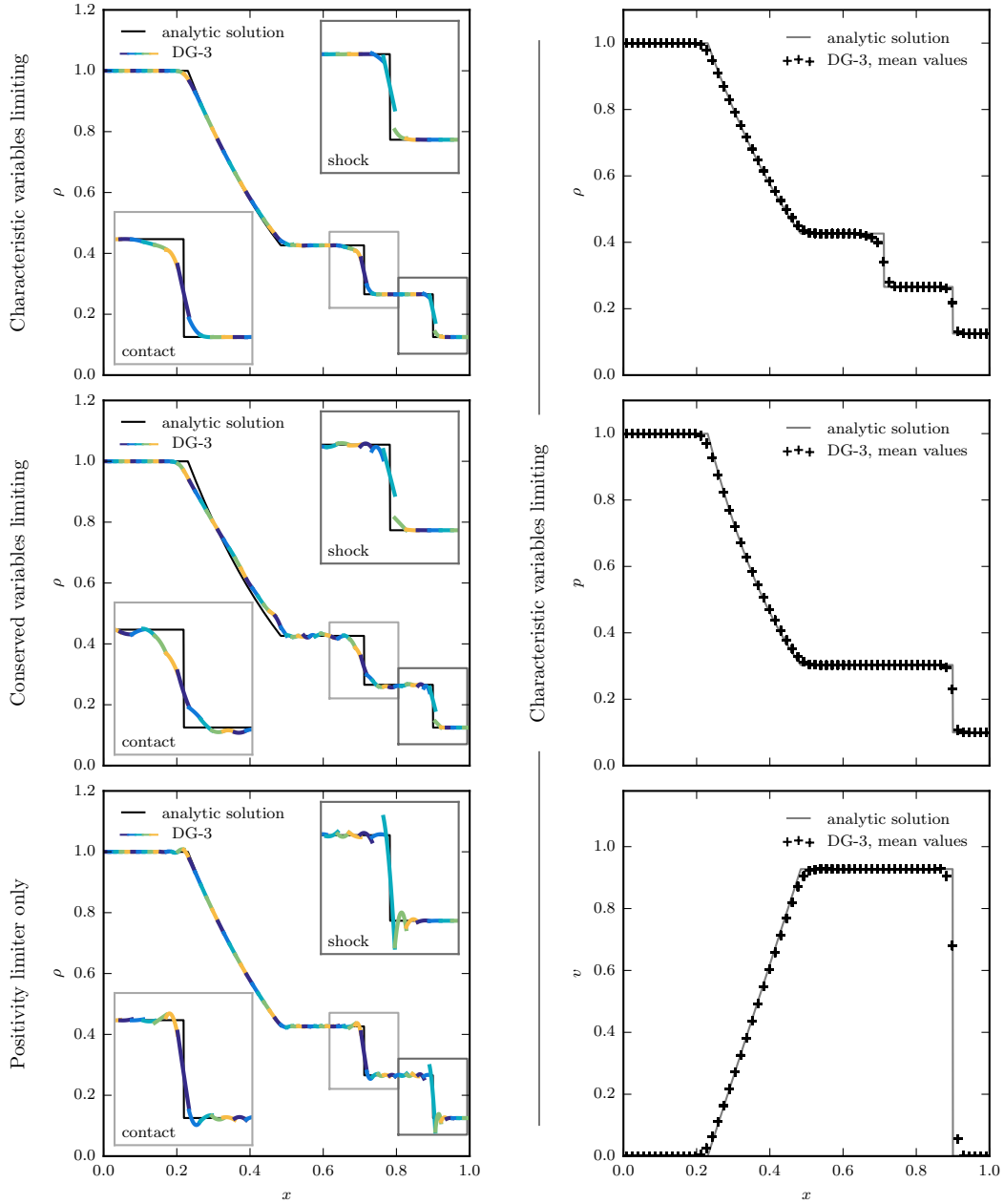


Figure 21.2: Left-hand panels: Sod shock tube problem calculated with DG-3 and different limiting approaches. We show the full polynomial DG solutions of the density, where the different colours correspond to different cells. The best result is achieved when the limiting is carried out based on the characteristic variables. As desired, the numerical solution is discontinuous and of low order at the shock. If the conserved variables are limited instead, the obtained solution is much less accurate, including numerical oscillations. The bottom-left panel shows the result without a slope limiter. The solution is of third order in every cell (parabolas) leading to an under- and overshooting at the shock as well as spurious oscillations. Right-hand panels: comparison of the mean cell values with the analytic solution for the characteristic variables limiter. Advection errors wash out the solution at the contact rapidly, until it can be represented smoothly by polynomials. Overall, we find very good general agreement with the analytic solution, especially at the position of the shock.

adjusts higher order terms such that negative density and pressure values in the calculation are avoided. Nevertheless, the obtained solution is accurate at large and has even the smallest L1 error of the three approaches tested.

The oscillations can however be reduced by limiting the second order terms of the solution, either expressed in the characteristic variables (upper-left panel) or in the conserved variables (middle-left panel). Strikingly, the numerical solution has a considerably higher quality when the characteristic variables are limited instead of the conserved ones, even though in both tests the same slope-limiting parameters are used ($\beta = 1$, $M = 0.5$). The former gives an overall satisfying result with a discontinuous solution across the shock, as desired. The contact discontinuity is less sharp and smeared out over ~ 5 cells. This effect arises from advection errors inherent to grid codes and can hardly be avoided. Nevertheless, the DG scheme produces less smearing compared to an FV method once the solution is smooth, see also Section 21.4. Hence, the higher order polynomials improve the sharpness of the contact. In the right-hand panels of Fig. 21.2, we compare the mean density, pressure, and velocity values of the numerical solution calculated with the characteristic limiter with the analytic solutions, and find good agreement. Especially, the shock is fitted very well by the mean values of the computed polynomials. To summarize, limiting of the characteristic variables is favourable over the limiting of the conserved variables. With the former, our DG code produces good results in the shock tube test thanks to its higher order nature, which clearly is an advantage also in this discontinuous problem.

21.3 Sedov-Taylor blast wave

The previous test involved a relatively weak shock. We now confront our DG implementation with a strong spherical blast wave by injecting a large amount of thermal energy E into a point-like region of uniform and ultra-cold gas of density ρ . The well-known analytic solution of this problem is self-similar and the radial position of the shock front as a function of time is given by $R(t) = R_0(Et^2/\rho)^{1/5}$ [see for example Padmanabhan, 2000]. The coefficient R_0 depends on the geometry (1D, 2D, or 3D) and the adiabatic index γ ; it can be obtained by numerically integrating the total energy inside the shock sphere. Under the assumption of a negligible background pressure, the shock has a formally infinite Mach number with the maximum density compression $\rho_{\max}/\rho = (\gamma + 1)/(\gamma - 1)$.

Numerically, we set up this test with 64^3 cells in a three-dimensional box $(x, y, z) \in [0, 1]^3$ containing uniform gas with $\rho = 1$, $p = 10^{-6}$, and $\gamma = 5/3$. The gas is initially at rest and the thermal energy $E = 1$ is injected into the eight central cells. Fig. 21.3 shows numerical solutions obtained with FV-2 as well as with DG-2 and DG-3 at $t = 0.05$. The density slices in the top panels indicate very similar results; only the logarithmic colour-coding reveals small differences between the methods in this test. Unlike Lagrangian particle codes, Cartesian grid codes have preferred coordinate directions leading to asymmetries in the Sedov-Taylor blast wave problem,

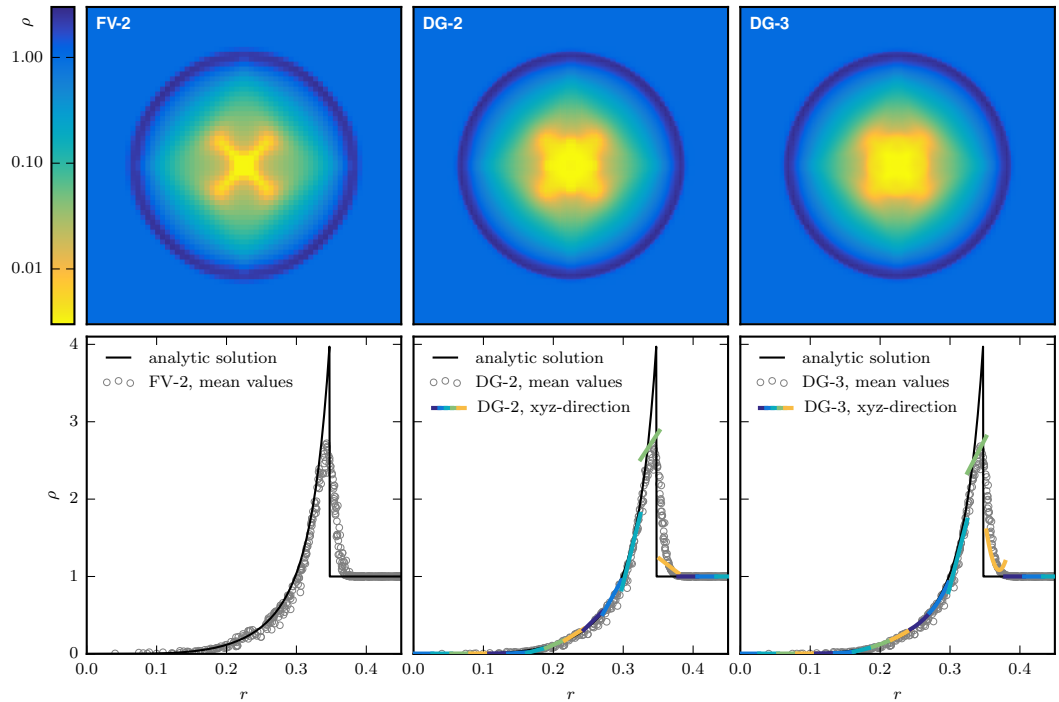


Figure 21.3: Three-dimensional Sedov-Taylor shock wave simulations at $t = 0.05$ calculated on a 64^3 -cell grid with FV-2 as well as with DG-2 and DG-3. The panels on top display central density slices ($z = 0$) on a logarithmic scale. In this test, Cartesian grid codes deviate noticeably from spherical symmetry in the central low-density region, and the shape of this asymmetry depends on details of the different methods. In the bottom panels, we compare the numerical results with the analytic solution. For each method, we plot the mean density of about every 200th cell, and for DG also the solution polynomials in the diagonal direction along the coordinates from $(0.5, 0.5, 0.5)$ to $(1, 1, 1)$. The obtained results are very similar in all three methods considered here, and DG does not give a significant improvement over FV for this test.

especially in the inner low-density region. FV-2 produces a characteristic cross in the centre. In DG the asymmetries are significantly weaker; only a mild cross and squared-shape trace of the initial geometry of energy injection are visible. These effects can be further minimized by distributing the injected energy across more cells, for a point-like injection; however, they cannot be completely avoided in grid codes.

In the bottom panels of Fig. 21.3, we compare the numerical results to the analytic solution obtained with a code provided by Kamm and Timmes [2007]; in particular we have $R_0 \approx 1.152$ for $\gamma = 5/3$ in 3D. Due to the finite and fixed resolution of our grid codes, the numerical solutions do not fully reach the analytic peak compression of $\rho_{\max} = 4$. More importantly, the DG method does not provide a visible improvement in this particular test, and furthermore, the results obtained with DG-2 and DG-3 are very similar. The reason behind these observations is the aggressive slope limiting due to the strong shock, suppressing the higher order polynomials in favour of avoiding over- and undershootings of the solution. We note that a better result could of

course be achieved by refining the grid at the density jump and thereby increasing the effective resolution. However, arguably the most important outcome of this test is that the DG scheme copes with an arbitrarily strong shock at least as well as a standard FV scheme, which is reassuring given that DG’s primary strength lies in the representation of smooth parts of the solution.

21.4 Square advection

SPH, moving mesh, and mesh-free approaches can be implemented such that the resulting numerical scheme is manifestly Galilean invariant, implying that the accuracy of the numerical solution does not degrade if a boost velocity is added to the initial conditions. On the other hand, numerical methods on stationary grids are in general not manifestly Galilean invariant. Instead, they produce additional advection errors when a bulk velocity is added, which have the potential to significantly alter, e.g., the development of fluid instabilities [Springel, 2010]. While the numerical solution is still expected to converge towards the reference solution with increasing resolution in this case [Robertson et al., 2010], this comes at the price of a higher computational cost in the form of a (substantial) increase of the number of cells and an accompanying reduction of the timestep size.

We test the behaviour of the DG method when confronted with high advection velocities by simulating the supersonic motion of a square-shaped overdensity in hydrostatic equilibrium [following Hopkins, 2015]. For the initial conditions, we choose a $\gamma = 7/5$ fluid with $\rho = 1$, $p = 2.5$, $v_x = 100$, and $v_y = 50$ everywhere, except for a squared region in the centre of the two-dimensional periodic box, $(x, y) \in [0, 1]^2$ with side lengths of 0.5, where the density is $\rho_s = 4$. The test is run with a resolution of 64^2 cells until $t = 10$, corresponding to 1000 transitions of the square in the x -direction and 500 in the y -direction.

In Fig. 21.4, we visually compare the results obtained with DG-2, DG-3, and the FV-2 scheme. Already at $t = 1$, the FV method has distorted the square to a round and asymmetric shape, and at $t = 10$, the numerical solution is completely smeared out². In comparison, DG shows fewer advection errors, and a better approximation of the initial shape can be sustained for a longer time. Especially, the run with third order accuracy produces a satisfying result. Note that due to the high advection speed the CFL timestep is very small, leading to around 1.7 and 3.3 million timesteps at $t = 10$ with DG-2 and DG-3, respectively.

In the bottom panels of Fig. 21.4 we show one-dimensional slices of the density solution polynomials from $(x, y) = (0, 0.5)$ to $(x, y) = (1, 0.5)$, and in grey the initial conditions as reference. With DG-2, an immediate over- and undershooting at the discontinuity can be observed owing to the adopted non-TVD slope limiter ($\beta = 1$,

²In our FV-2 method, the overdensity moves slightly faster in the direction of advection and is clearly not centred any more at $t = 10$. We have investigated this additional error and found that its occurrence depends on the choice of the slope limiter. Either way, the solution is completely washed out and the FV-2 method does not provide a satisfying result in this test.

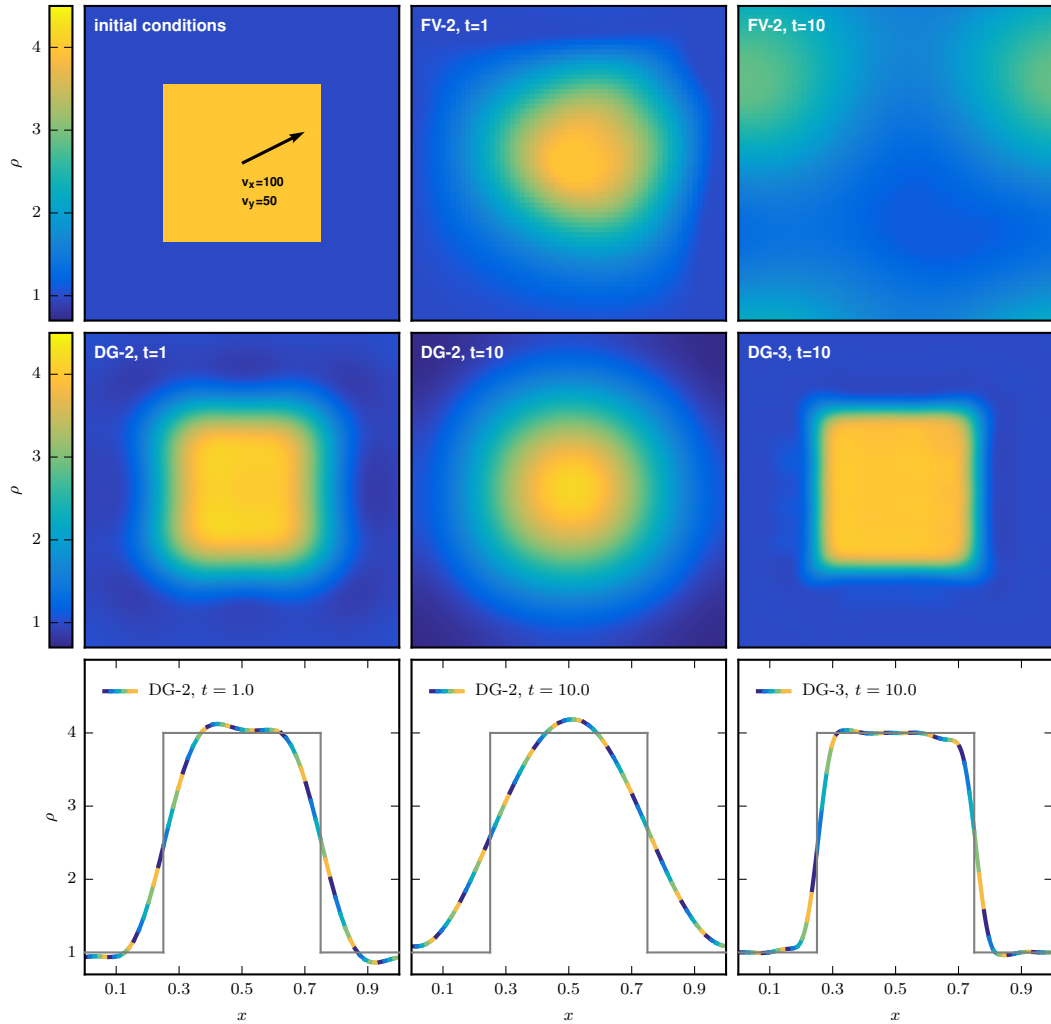


Figure 21.4: Density maps and centred slices for the advection test with 64^2 cells: a fluid with a square-shaped overdensity in hydrostatic equilibrium is advected supersonically, crossing the periodic box several hundred times. The FV-2 method shows large advection errors in this test and the square is smeared out completely by $t = 10$. On the other hand, the advection errors in the DG method become small once the solution is smooth. DG-3 shows less diffusion than DG-2 due to the higher order representation of the advected shape. The time evolution of the density errors for the three simulations is shown in Fig. 21.5.

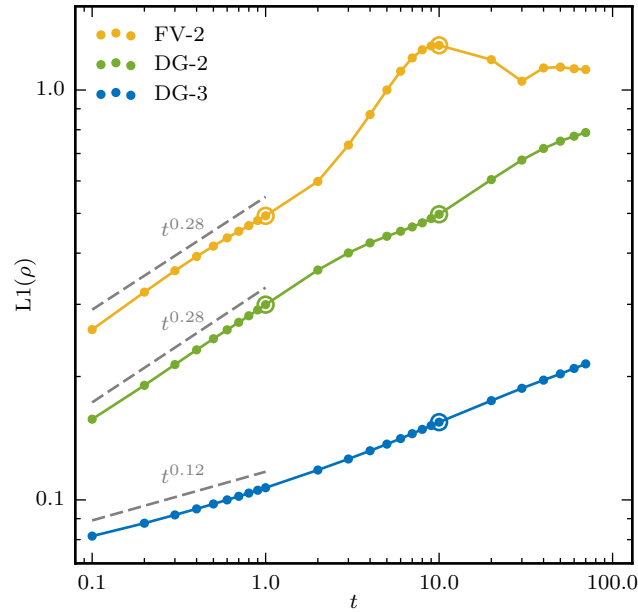


Figure 21.5: Time evolution of the L1 density error norm for the square advection test. The grey lines indicate the initial slopes of error growth, and the encircled data points correspond to the plots shown in Fig. 21.4. Interestingly, FV-2 and DG-2 have the same polynomial growth, though the absolute error is smaller by a factor of ≈ 1.5 for DG. On the other hand, the DG-3 method does not only decrease the absolute error in this test, but the error growth rate is also significantly smaller.

$\tilde{M} = 0.5$). More interestingly, while in the FV-2 method the solution is only washed out by diffusion, for DG-2 it is also steepened, which leads to a higher maximum density at $t = 10$ than present in the initial conditions. This difference arises from the updating of the slopes in DG instead of reconstructing them, and moreover, it vanishes if we use TVD limiter parameters. With the latter the slopes are reduced aggressively, resulting in pure diffusion with a result similar to the FV-2 method.

We quantify the quality of each scheme in this test by measuring the time evolution of the L1 density error norm, the result of which is presented in Fig. 21.5. The FV-2 and DG-2 codes show the same initial polynomial error growth, $\propto t^{0.28}$; however, the absolute error is much smaller for DG, i.e. the FV-2 error at $t = 1$ is reached with DG-2 only at a much later time, at $t = 10$. With the DG-3 code (quadratic basis functions), the error grows only as $\propto t^{0.12}$, leading to an acceptably small error even until $t = 70$, which corresponds to 23 million timesteps. Also with this higher order approximation the solution is transformed to a smooth numerical solution, but the mean cell values are less modified compared to the second order code with linear basis functions.

As demonstrated above, DG produces far smaller advection errors compared to an FV method. But how can this be understood, especially since the DG scheme is also not Galilean invariant? A powerful tool for studying the behaviour of discretizations in computational fluid dynamics is the modified equation analysis. Here, the discrete

equations are expanded by means of a Taylor series, leading to a so-called modified equation which includes the target equation to solve and additional terms. For example, the modified equation of the first order upwind method for the scalar advection equation is an advection-diffusion equation [LeVeque, 2002]. The scheme solves the advection equation to first order by construction, but at the same time it effectively solves this advection-diffusion equation to second order accuracy, explaining the large diffusion errors when adopting this simple scheme, and pointing towards an explanation for the observed diffusion in our FV-2 method. Such an analysis proves to be more difficult for DG; nevertheless Moura et al. [2014] recently accomplished a modified equation analysis for the second order DG scheme applied to the linear advection equation. In this case, the modified equation consists of a physical mode and an unphysical mode moving at the wrong speed, which is however damped very quickly. For upwind fluxes the leading error term of the physical mode is diffusive and of third order, which is better than naively expected for a second order method, and this is likely one of the keys for the improved behaviour of DG we find. A heuristic argument for the superiority of DG in this test is that after an initial smoothing of the global numerical solution, it is not only continuous inside every cell but also at the cell interfaces. If the solution is perfectly continuous across cell interfaces, the left and right state entering the Riemann solver are identical, and the calculated flux is always related by a simple Galilei boost to the flux calculated in any other frame of reference. In this case, no manifest differences in the conserved quantities in a cell due to the flux calculation of the surface integral (18.20) can arise under a Galilean transformation, implying that advection errors must be minimal. Nevertheless, some small averaging errors will arise in practice if the current profile cannot be represented exactly at an arbitrary grid position with the given set of cell basis functions.

21.5 Keplerian disc

Rotating gas discs are omnipresent in our Universe, for example in galaxies, accretion discs around black holes, or protostellar and protoplanetary discs. Numerically, such objects are ideally modelled either with a Lagrangian method or with a grid code which operates with a suitably tailored mesh geometry and furthermore accounts for part of the angular rotation in the solver [Masset, 2000]. If a simulation contains however several rotating objects at different locations and with different orientations, as for example is the case in cosmological galaxy formation simulations, no preferable grid geometry exists. In this case, Cartesian grids are often adopted, optionally with AMR, making it much more challenging to avoid spurious errors in differentially rotating gas flows.

These problems can be exposed by the demanding problem of a cold Keplerian disc, where real pressure forces are negligibly small and any spurious hydrodynamic forces from numerical errors result in readily visible perturbations of the system. We subject our DG code to this test following a similar setup as recently used in other works [Hopkins, 2015; Pakmor et al., 2016]. The ambient gas of the disc has

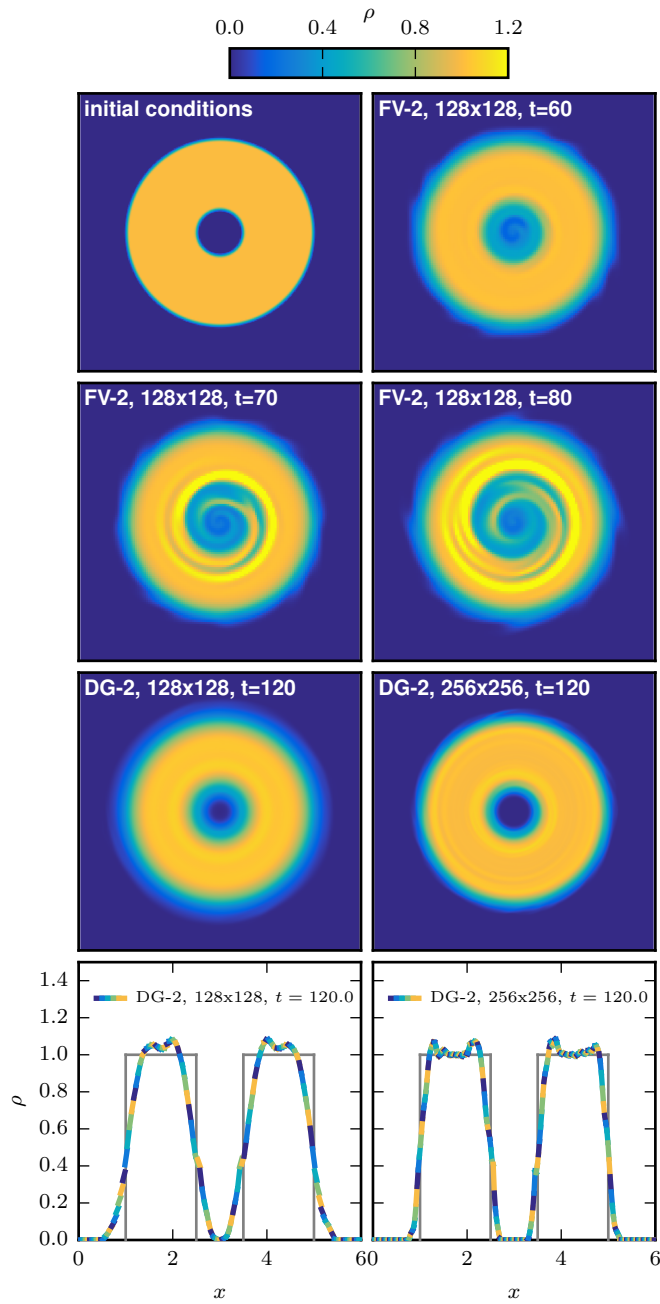


Figure 21.6: Density evolution in the cold Keplerian disc problem. The centrifugal force acting on the rotating disc is balanced by an external static central potential such that every fluid element is on a Keplerian orbit. This test is very challenging for Cartesian grid codes, as shear flows without pressure support are prone to numerical instabilities. Our FV-2 method is stable for around 10 orbits, at which point the disc gets disrupted eventually. In contrast, without a slope limiter DG conserves angular momentum accurately and can handle this problem very well.

negligible density and pressure, and is initially confined to lie between two radii. Every fluid element of the disc is rotating on a Keplerian orbit where the centrifugal forces are balanced by an external and static central gravitational potential. Gas self-gravity is not included. Analytically, the system is in perfect equilibrium and the initial state should not change in time. The difficulty for hydro codes applied to this test lies in the lack of pressure support, as well as the differential rotation which leads to shearing flows. Both can trigger numerical instabilities and the eventual disruption of the disc. In particular, it is clear that for codes which do not conserve angular momentum exactly, this point will inevitably be reached eventually. In SPH codes, angular momentum is conserved but angular momentum transport is caused by the use of artificial viscosity, which typically is active at a small level in strong shear flows. Using an improved switch for the viscosity, however, the Keplerian disc problem can be integrated accurately [Cullen and Dehnen, 2010].

For definiteness, the initial conditions we use for our DG test are as follows. We use the computational domain $(x, y) \in [0, 6]^2$ with periodic boundaries, and gas with an adiabatic index of $\gamma = 5/3$. The primitive variables are initially set to

$$p = p_0,$$

$$\rho(r') = \begin{cases} \rho_0 & \text{if } r' < 0.5 - \frac{\Delta r}{2} \\ \frac{\rho_D - \rho_0}{\Delta r} (r' - (0.5 - \frac{\Delta r}{2})) + \rho_0 & \text{if } 0.5 - \frac{\Delta r}{2} \leq r' < 0.5 + \frac{\Delta r}{2} \\ \rho_D & \text{if } 0.5 + \frac{\Delta r}{2} \leq r' < 2 - \frac{\Delta r}{2} \\ \frac{\rho_0 - \rho_D}{\Delta r} (r' - (2 - \frac{\Delta r}{2})) + \rho_D & \text{if } 2 - \frac{\Delta r}{2} \leq r' < 2 + \frac{\Delta r}{2} \\ \rho_0 & \text{if } r' \geq 2 + \frac{\Delta r}{2}, \end{cases}$$

$$v_x(x', y') = \begin{cases} -y'/r'^{3/2} & \text{if } 0.5 - 2\Delta r < r' < 2 + 2\Delta r \\ 0 & \text{else,} \end{cases}$$

$$v_y(x', y') = \begin{cases} x'/r'^{3/2} & \text{if } 0.5 - 2\Delta r < r' < 2 + 2\Delta r \\ 0 & \text{else,} \end{cases}$$

where the coordinates x' , y' , and r' are measured in a coordinate system with the origin at the centre $(x_0, y_0) = (3, 3)$ of the box. The values used for the background gas are $p_0 = 10^{-5}$ and $\rho_0 = 10^{-5}$. The disc has a density of $\rho_D = 1$ and a radial extent of $[r_{\min} = 0.5, r_{\max} = 2]$, with a small transition region of width $\Delta r = 0.1$. We adopt a time-independent external acceleration $\mathbf{a} = -\nabla\Phi$ with the components

$$a_x(x', y') = \begin{cases} -x'/r'^3 & \text{if } 0.5 - 0.5\Delta r < r' \\ -x'/[r'(r'^2 + \epsilon^2)] & \text{else,} \end{cases}$$

$$a_y(x', y') = \begin{cases} -y'/r'^3 & \text{if } 0.5 - 0.5\Delta r < r' \\ -y'/[r'(r'^2 + \epsilon^2)] & \text{else,} \end{cases}$$

where $\epsilon = 0.25$ smoothes the potential in the very inner regions in order to avoid a singularity there. The orbital period of the Keplerian disc depends on the radius and is given by $T = 2\pi r^{3/2}$.

We evolve the system until $t = 120$, corresponding to around 20 orbits at $r = 1$, and present the results in Fig. 21.6. When the FV-2 method is used, the edges of the disc are washed out immediately but the disc is stable for around 10 orbits. Like the majority of FV methods in use, our scheme does not manifestly preserve angular momentum, resulting in secular integration errors and an eventual breakdown of the quasi-static solution. The number of orbits the disc survives depends mainly on how carefully the problem has been set up, as well as on the choice of slope limiter and resolution used. However, the disruption of the disc is inevitable and can only be delayed with adjustments of these parameters.

On the other hand, DG schemes of second order and higher are inherently angular momentum conserving and can hence potentially handle this test problem much more accurately. To test for this, we have disabled the slope limiter of the DG scheme and use only the positivity limiter. This is because with our simple minmod limiter, angular momentum conservation is also violated and would result in a similar disruption as observed with FV-2. The construction of an improved angular momentum preserving limiting scheme is hence desirable and worthwhile subject for future work. With the positivity limiter alone, the DG-2 scheme can integrate the disc stably and gives good results at $t = 120$, corresponding to about 20 orbits at $r = 1$. Merely two fine rings with a slightly higher density can be observed in the inner and outer regions of the disc, which can be attributed to the gentle overshooting of the solution at the rims of the disc. We have also carried out this simulation with the DG-3 code. In this case, the disc also does not break down, but the solution shows some mild oscillations with amplitude up to 20 percent of the initial density; without applying a limiter these cannot be suppressed.

21.6 KH instability

The KH instability is one of the most important fluid instabilities in astrophysics, for example, it plays an important role in the stripping of gas from galaxies falling into a galaxy cluster. The instability is triggered by shear flows, often also involving fluids with different densities, and grows exponentially until the primary billows break, subsequently leading to a turbulent mixing of the two phases. The KH instability can be investigated with initial conditions consisting either of a sharp surface between two phases or with a transition region separating the layers smoothly. Analytic growth rates for the linear regime can be derived for both cases [Chandrasekhar, 1961; Hendrix and Keppens, 2014]; however, the thickness of a smooth transition layer sets a limit on the minimum wavelength which can become unstable in the linear phase. This suppression is desired if one wants to compare growth rates inferred from simulations with the analytic growth rate [McNally et al., 2012], since the underlying mesh can trigger the instability of waves at the resolution scale due to noise, a

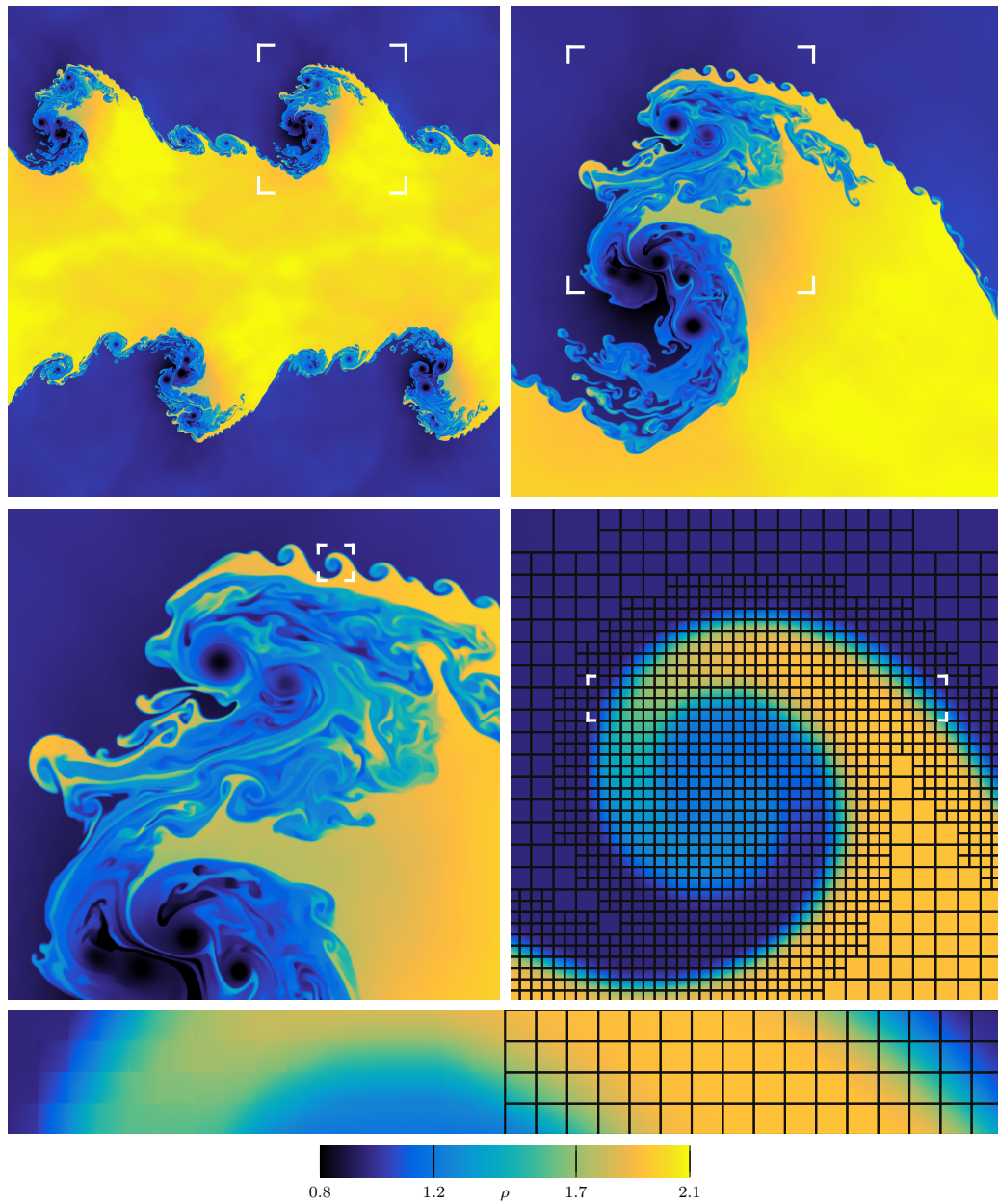


Figure 21.7: High-resolution KH simulation with DG-4 and AMR at time $t = 0.8$. The simulation starts with 64^2 cells (level 6) and refines down to level 12, corresponding to an effective resolution of 4096^2 . We illustrate the AMR levels in Fig. 21.8. The mesh refinement approach renders it possible to resolve fractal structures created by secondary billows on top of the large-scale waves. Furthermore, as can be seen in the bottom panel, the solution within every cell contains rich information, consisting of a third order polynomial. A movie of the simulation until $t = 2$ may be accessed online: <http://youtu.be/cTRQP6DSaqA>

numerical effect which does not vanish with increasing resolution. Nevertheless, we set up a sharp discontinuity and use the KH instability as a high-resolution test for the robustness of our AMR implementation and DG's capabilities of capturing small-scale turbulent structures.

The initial conditions are chosen as in Springel [2010]; in the periodic box $(x, y) \in [0, 1]^2$ the primitive variables at $t = 0$ are set to

$$\begin{aligned}
 p &= 2.5, \\
 \rho(x, y) &= \begin{cases} 2 & \text{if } 0.25 < y < 0.75 \\ 1 & \text{else,} \end{cases} \\
 v_x(x, y) &= \begin{cases} 0.5 & \text{if } 0.25 < y < 0.75 \\ -0.5 & \text{else,} \end{cases} \\
 v_y(x, y) &= w_0 \sin(4\pi x) \left\{ \exp \left[-\frac{(y - 0.25)^2}{2\sigma^2} \right] + \exp \left[-\frac{(y - 0.75)^2}{2\sigma^2} \right] \right\},
 \end{aligned}$$

with $w_0 = 0.1$, $\sigma = 0.05/\sqrt{2}$, and the adiabatic index $\gamma = 7/5$. The velocity perturbation in the y -direction supports the development of a specific single mode on large scales. We start with an initial resolution of 64^2 cells (level 6) and refine where the density slope is steep, as described in Section 20.1. The maximum refinement level is set to 12, corresponding to an effective resolution of 4096^2 cells. A sharp discontinuity between the two layers in combination with AMR leads to a fast transition into the non-linear regime and generates secondary billows early on.

We illustrate the state of the simulation at $t = 0.8$ in Fig. 21.7. The panel on the top left shows the density for the whole two-dimensional simulation box; in the following panels we zoom in repeatedly. The DG scheme shows only little diffusion and mixing thanks to the higher order and the avoidance of reconstruction steps, allowing the formation of very fine structures. Smaller KH instabilities arise on top of the large-scale waves demonstrating the fractal nature of this test problem. Self-similar instabilities are present across different scales, and an ending of this pattern is only set by the limited resolution.

The adaptive mesh used by the calculation is overlaid in the bottom-right panel, revealing three different AMR levels in this subbox. The density gradients are well resolved with the finest AMR level (level 12), whereas smooth regions are on smaller levels. Furthermore, technical features of the AMR mesh structure can be seen in the plot: the level difference between neighbouring cells is at most one, and the mesh smoothing algorithm introduces an additional cell layer around physically refined cells. Fig. 21.8 shows a map of the AMR levels of the whole simulation box at $t = 0.8$, which can be directly compared to the top-left panel of Fig. 21.7. The number of cells at the displayed instance is around 1.8 million cells, which corresponds to about 10 percent of the effective level 12 resolution 4096^2 , highlighting the efficiency gain possible with the AMR approach. Ideally, we would also like to use a corresponding

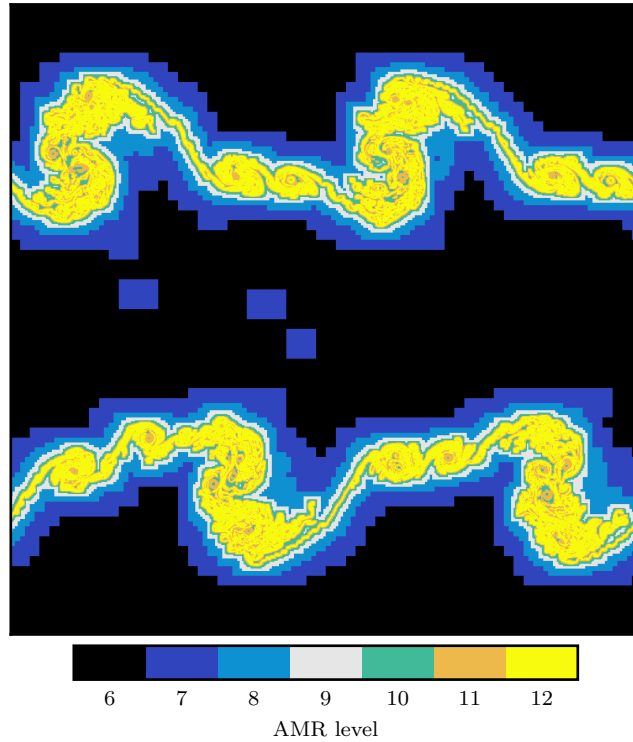


Figure 21.8: Map of the AMR levels of the KH simulation at $t = 0.8$. We here refine in space where the density gradient is steep, allowing us to capture interesting regions with high resolution and save computational time in places where the solution is smooth.

adaptiveness in time by utilizing local timesteps, something that is planned as a code extension in future work.

22

Summary

In this work, we have developed a 3D MPI-parallel higher order DG code with AMR for solving the Euler equations, called TENET, and investigated its performance in several astrophysically relevant hydrodynamic test problems. DG methods are comparatively new in astrophysics, despite a vast body of literature and long history of these approaches in the applied mathematics community. A number of highly attractive features of DG however suggest that it is timely to introduce these methods as standard tools in computational astrophysics. In particular, DG allows higher order to be reached without introducing communication beyond the immediate neighbouring cells, making it particularly well suited for parallel computing. Also, the method is characterized by an enhanced continuous memory access due to the increased number of DOF per cell, making it a better match for modern computer architectures, where floating point operations are “almost free” in comparison to slow and costly repeated memory accesses of small chunks. In addition, DG allows an easy and flexible way to reach arbitrarily higher order, quite unlike FV schemes. This makes it possible to also vary the order of a scheme in space, providing for additional flexibility in AMR approaches.

Our tests furthermore clearly show that it is computationally worthwhile to employ higher order methods. While in general it depends on the problem which order is optimal, we have found in our tests that third and fourth order DG implementations are typically computationally much more efficient compared with a second order code, at least in regions where the solution is reasonably smooth. Moreover, the numerical result is of higher quality also in non-smooth regions, especially shocks are represented very well, thanks to the discontinuous nature of the DG representation. They are typically captured within at most two to three cells.

Nevertheless, close to shocks and discontinuities, possible limitations of the DG scheme become apparent. In favour of a higher overall accuracy, our DG scheme is

not TVD, which can lead to over- and undershootings of the solution. Moreover, higher order methods tend to produce spurious oscillations at these locations. The detection of troubled cells and the prevention of this unwanted behaviour is an active topic of current research. The spurious oscillations in our idealized test problems could be prevented effectively by limiting the slopes and discarding higher order terms of the solution when appropriate. For future real-world problems, a more sophisticated method is desirable. A very promising concept in this respect is the so-called *hp*-adaption. In this approach, the grid is refined around shocks and discontinuities, while at the same time the degree of the polynomials is reduced, resulting in a locally more robust scheme with less oscillations. Recently, Sonntag and Munz [2014] presented a practical realization of this concept by incorporating a subgrid TVD FV scheme into troubled DG cells. The subgrid resolution is chosen such that it has the same number of DOF as a DG cell. In this way, the timesteps of the FV cells are similar to the timesteps of the larger surrounding DG cells, and the scheme can be applied without significant computational overhead.

The fundamental difference between DG and FV is that DG solves directly also for higher order moments of the solution, whereas in FV all higher order information is discarded in the implicit averaging at the end of each timestep, necessitating a subsequent reconstruction. This aspect of DG leads directly to two major advantages over traditional FV methods. First, DG produces significantly less advection errors, and furthermore, if the solution is smooth across cell interfaces, the numerical solution does not depend on the chosen Riemann solver. Secondly, DG inherently conserves not only mass, momentum, and energy, but also angular momentum. The conservation of angular momentum is very desirable for many astrophysical applications, e.g. simulations involving discs, or objects like stars or molecular clouds in rotation. There is however one caveat which has to be kept in mind. The conservation can be spoiled by the limiting procedure, which is reminiscent of the situation in SPH, where angular momentum is spuriously transported by artificial viscosity. Improving the limiter with the goal of global angular momentum conservation is hence desirable and a promising direction for future improvements of the DG implementation. Finally, DG can also be comparatively easily generalized to the AMR technique, and importantly, this can be done without loss of accuracy, unlike in standard FV approaches. The higher order is formally preserved due to the usage of ‘hanging nodes’, which are the quadrature points of interfaces between cells of different sizes.

This part of the thesis has focused on introducing our new code and highlighting its differences relative to FV schemes. Future developments will focus on more sophisticated shock capturing schemes, scaling improvements through Open-MP hybrid parallelization, as well as on incorporating magnetic fields and astrophysical source terms relevant in galaxy formation. The latter is greatly facilitated by the modular structure of TENET and its parent code AREPO. In a first science application of our DG code, we quantitatively analyse its capabilities of capturing driven turbulence structures [Bauer et al., 2016].

Part V

Conclusions and Outlook

23

Discontinuous Galerkin hydrodynamics

Discontinuous Galerkin (DG) schemes are powerful higher order methods for solving differential equations such as the hyperbolic conservation laws of ideal hydrodynamics. Nevertheless, discontinuous Galerkin hydrodynamics is relatively unexplored in astrophysical contexts, as well as for structured meshes that can be adaptively refined (adaptive mesh refinement, AMR). We have outlined the mathematical framework for a Runge-Kutta DG method operating on an AMR grid, including crucial details such as the refinement and derefinement procedure, the limiting at refinement boundaries, the handling of source terms, and the calculation of the timestep size. Moreover, we have implemented this DG formulation as an MPI-parallel module into the AREPO code¹ for solving the Euler equations, and tested its performance with various hydrodynamic test problems relevant for astrophysical applications.

Traditional grid-based hydrodynamic solvers in computational astrophysics employ the finite volume (FV) method, which evolves the mean mass, momentum, and energy for every grid cell with reconstruct-evolve-average (REA) cycles. In this scheme, the solution within every cell is reconstructed based on the mean fluid quantities of neighbouring cells, and evolved by means of a Riemann solver. At the end of every timestep the solution is averaged and the reconstructed higher order information is not further taken into account. Even though only the mean cell values are stored in FV, the adoption of a reconstruction step renders it possible to obtain a method which is higher than first order.

On the other hand, DG is a finite element method which stores and self-consistently evolves also the higher order moments of the solution. This can be accomplished by solving the Euler equations with higher order polynomial basis functions for every grid cell, and by computing the solutions of multiple Riemann problems at

¹We want to remark that in its standard mode AREPO operates on a moving Voronoi mesh, however, for this work we have utilized the fact that it can also construct AMR grids.

every cell interface. The global numerical solution consists of piece-wise continuous polynomials, and is in general discontinuous across cell interfaces. Similarly to the FV method, DG conserves the total mass, momentum, and energy of the fluid.

We have shown that DG offers several advantages compared to traditional FV methods. First of all, the generalization to a higher spatial order can be accomplished straight-forwardly by increasing the order of the basis functions and numerical integration rules. As a matter of fact, we have implemented the DG scheme for an arbitrary order in space and time. Given the desired order of the spatial scheme, as well as the coefficients of a Runge-Kutta method, our code computes all basis functions and numerical integration rules in an automated way, and adopts the correct time integration steps. A higher order method can have a higher computational efficiency, especially in locations where the solution is smooth. Strikingly, independent of the order of the DG scheme, only information from direct neighbouring cells are required for evolving the solution within a grid cell. This feature is greatly beneficial for parallel computing applications, since the amount of communication is small compared to the time spent for local computations.

The situation proves to be quite different for FV methods. In this case no simple framework for generalizing the scheme to higher order accuracy exists, and the implementation becomes cumbersome already for a third order code. Moreover, the stencil of the method gets wider with increasing order, e.g. every grid cell in a third order FV code requires information from the neighbours of neighbours.

The optimal order for solving a hydrodynamic problem is always problem dependent. For representing discontinuities, a low order scheme is beneficial, whereas smooth flows can be represented more efficiently with higher order methods. We have demonstrated that in the latter regime a great improvement in efficiency can be obtained by increasing the accuracy from second to third order. Moreover, with fourth order DG an even better result can be achieved compared to third order DG, however, the efficiency gain is smaller in this case. Hence, a third order scheme might be a good overall choice in many cases.

Near hydrodynamic discontinuities, higher order methods suffer from a severe drawback. Hydrodynamic schemes do not converge properly at these locations, and higher order terms tend to induce oscillations, an effect also known as Gibbs phenomenon. This limitation is present for higher order DG methods and higher order FV methods alike, and appears in particular at hydrodynamic shocks. An effective way of suppressing these oscillations consists of decreasing the convergence order of the scheme. This can effectively be accomplished by limiting the solution. However, this should only be done at problematic locations, and the difficulty lies in the identification of the latter. If the solution is limited, shocks can be captured accurately due to the discontinuous nature of DG and FV. We have investigated different minmod limiters for DG in shock tube tests, and find that limiting the solution expressed in characteristic variables gives a significantly better result compared to the limiting in conserved variables. We hence argue that computing and applying the transformation to the characteristic variables pays off if a high accuracy at hydrodynamic shocks is desired.

In regions where discontinuities are absent, DG exhibits a couple of advantages over FV, which are very relevant for astrophysical applications. By solving self-consistently for the second order moment of the solution, DG is able to conserve angular momentum inherently. We have demonstrated that due to this feature a second order DG code can evolve an ultra cold Keplerian disc on stable orbits, whereas this test problem can not be handled by a FV method on the same Cartesian grid. Moreover, for a DG method with third or higher order, the basis functions connect across cell interfaces at locations where the solution is smooth, such that it is continuous across more than one cell. In this case the Riemann solver does not introduce any inaccuracies when calculating the fluxes at cell interfaces, and advection errors become very small. Both features represent a significant improvement compared to current FV codes used for astrophysical applications.

In another project, which is not part of this thesis, the performance of our DG code has been tested by Bauer et al. [2016] in simulations of subsonic turbulence. No shocks are present in these runs, however, tangential discontinuities such as Kelvin-Helmholtz instabilities can form and the solution is in general not smooth. Interestingly, they report that similar results are achieved by a third order DG simulation with 128^3 cells, and by a second order FV simulation with 512^3 cells. The former run requires a factor of 6.4 less degrees of freedom², and moreover, it finished around 1.5 times faster. This clearly demonstrates that for subsonic turbulence simulations the third order DG method is favourable compared with the second order FV method.

Discontinuous Galerkin methods have been explored only recently in astrophysical contexts [Mocz et al., 2014; Schaal and Springel, 2015; Teukolsky, 2016; Zanotti et al., 2015], however, we believe that they can replace finite volume methods in the long term due to their mathematical properties and flexibility. Future developments in our code could be the improvement of the parallel efficiency, e.g. by a hybrid parallelization, the implementation of more sophisticated limiting strategies [see for example Sonntag and Munz, 2014], or an increase of the accuracy by solving generalized Riemann problems [Toro, 2009], which consist of smoothly varying states left and right to the discontinuity. Ultimately, it is desirable to implement a higher order DG scheme on top of AREPO’s moving-mesh discretization. Up to now this proves to be difficult, since a mathematical formulation which can retain higher order accuracy for changing cell shapes and neighbour topologies has yet to be found.

Nevertheless, based on the DG AMR code we have implemented, many astrophysical applications could be targeted in the near future, e.g. the simulation of galaxy clusters and the associated turbulence, or the modelling of accretion discs around protostars or black holes.

²The third order DG run uses 64 times fewer cells, and every cell has 10 degrees of freedom.

24

Hydrodynamic shocks in cosmological simulations

We have implemented an accurate shock finder operating on AREPO's moving Voronoi mesh and analysed non-radiative simulations as well as simulations with a more complete physical model. We have validated our implementation in shock tube and Sedov-Taylor blast wave tests, and find very good agreement with analytical solutions for both, the computed shock surface area, as well as the measured energy dissipation. Moreover, we have introduced a new test for assessing the performance of a shock finder in cosmological simulations. The total energy dissipation obtained from the shock finder can be directly compared with the irreversible increase of thermal energy between two consecutive timesteps of the simulation, and we can report a satisfying agreement within around 15%. This test proves to be very demanding, since shocks are smoothed across several cells, and the flow patterns in cosmological simulations are rather complex compared to idealized test problems. The obtained result indicates that our algorithm is able to reliably reconstruct shock surfaces and compute their Mach numbers in complex hydrodynamical systems.

The methodology of our algorithm is based on previous shock finders implemented in codes with structured meshes that can be adaptively refined, although augmented with the following improvements. Firstly, we also take into account that shocks from different directions can come close to each other and eventually collide; by comparing gradients along rays perpendicular to the shock surfaces we are able to distinguish the shocks and calculate the correct Mach number for each of them. Secondly, shocks are not the only existing hydrodynamic discontinuities, and we take special care that contact discontinuities and tangential discontinuities are not detected by our shock finder.

We have shown that the latter improvement is crucial for our shock finder, as well as for existing realizations in other codes. The number of shock detections within haloes is significantly reduced, and instead of ubiquitous complex flow patterns inside structures, as has been found by previous studies, we find rather distinct shock surfaces. This is even the case inside haloes where complex subsonic gas motions are present, e.g. subsonic turbulence, indicating that our shock finder is well able to distinguish between the different hydrodynamic discontinuities.

For the non-radiative simulations we find that most of the high Mach number shocks ($\mathcal{M} \gg 20$) are located between voids and the cosmic web, and most of the energy dissipation at shocks happens inside non-linear structures where the Mach number of shocks is typically low ($\mathcal{M} \approx 3$). On the whole, this picture is in agreement with previous studies, however, as a result of the improved shock finding methodology, halo-internal shocks are less abundant and we measure a thermalization peak at higher Mach numbers ($\mathcal{M} \approx 3$ instead of $\mathcal{M} \approx 2$). Similar to previous studies, we infer an energy amount transferred to cosmic rays which is expected to result in a dynamically relevant cosmic ray population; we find that about 5% – 20% of the total kinetic energy processed by shocks at the present epoch is channeled into cosmic ray acceleration. This corresponds to a physical value of around $10^{40} \text{ erg s}^{-1} \text{ Mpc}^{-3}$.

In a second project, we have investigated shocks in a state-of-the-art cosmological simulation, the Illustris simulation. It aims at reproducing the real Universe on large scales as accurately as possible, and incorporates additional physics such as radiative cooling, star- and blackhole formation, and the associated feedback processes. For the first time, we have detected and characterized shocks in a simulation of this size and physical completeness.

After preliminary analysis, we have discovered that our shock finder designed for non-radiative simulations can not be readily used for a full physics simulation such as Illustris. We have identified two reasons why this is the case. Firstly, these simulations adopt subgrid models, which aim at taking physics into account on spatial scales that are not resolved within the current resolution limits. In particular, the coevolution of the hot and cold phase of the interstellar medium is modelled on a cell basis by an effective equation of state. In this way, the hydrodynamic state of a cell can be modified in addition to the evolution due to the Euler equations. Secondly, gravity in combination with cooling physics can result in dense clumps with sharp edges, which should not be confused with shocks. We could straight-forwardly cope with the former problem by excluding star-forming regions from our analysis. On the other hand, the effect of cooling physics is not confined to a specific environment, and we had to introduce a crucial improvement to our methodology. Instead of measuring the shock jump with the Rankine-Hugoniot jump condition for a specific hydrodynamic quantity, as it is commonly done in shock finding algorithms, we calculate Mach numbers for multiple jump conditions. If those are not consistent with each other, the discontinuity does not represent a hydrodynamic shock, and is filtered by our algorithm. We have demonstrated that our improvements result in a robust algorithm for detecting shocks in the Illustris simulation. On the other hand, we have also shown that without these improvements spurious detections can

enter and highly bias the shock statistics. Hence, we argue that current shock finder implementations designed for non-radiative simulations should be augmented in a similar way, before they are used for the analysis of simulations with a more complete physical model.

Interestingly, by detecting and characterizing shocks in the Illustris simulation with our improved algorithm, we have measured a specific dissipation rate for $z < 3$ which is roughly constant and fairly independent of environment. This rate is around $10^{-1} \text{ erg g}^{-1} \text{ s}^{-1}$. Moreover, we have detected a total shock surface area which is larger by a factor of 1.4, and measured a total energy dissipation at shocks which is higher by a factor of around 7 at $z = 0$, compared to the non-radiative simulations. This finding clearly demonstrates that the influence of feedback shocks on the overall shock statistics can be rather substantial. We have investigated the origin of the difference and found that a large fraction can be attributed to the presence of high Mach number AGN feedback shocks. Furthermore, by comparing the properties of these shocks with merger shocks, which can be observed in several cases in the form of radio relics, we have argued that the black hole feedback in Illustris is modelled too violently.

For both kinds of simulations, the non-radiative run as well as the full physics Illustris run, we have illustrated many different systems and the corresponding shock morphologies. We have demonstrated that the Mach number and energy dissipation fields contain rich information about the gas dynamics of astrophysical systems, and reveal their dynamical state in many cases. Moreover, we have discovered a double accretion shock pattern around $10^{12} M_{\odot}$ haloes at $z = 4$, which expands the current picture of the gas accretion process of haloes. In the classical hierarchical galaxy formation theory [White and Frenk, 1991; White and Rees, 1978], two modes of gas accretion are discussed. In the hot mode of accretion, gas gets shock heated to the virial temperature close to the virial radius, forming a roughly spherical accretion shock, before it can cool and fragment into a disc. In the cold mode of gas accretion, the central galaxy is directly fed with cold gas, and no accretion shock can be found in this case. Later, a third mode of gas accretion has been reported. In this mode, the halo atmosphere supports an accretion shock, however, filaments can penetrate into the halo and channel cold gas into the central galaxy [Dekel and Birnboim, 2006]. Our finding corresponds to a fourth accretion mode, in which gas can cross the outer accretion shock without being shocked, but then forms a second accretion shock further inside. This inner accretion shock is roughly spherical, and we have proposed a scenario for the generation of the latter.

Our shock finder offers many potential future applications. First of all, more simulation analysis could be carried out, which aims at specific scientific questions. For example, one could investigate the occurrence of the different accretion modes as a function of halo mass and redshift, and compare the result with analytic expectations. Moreover, zoom simulations on to individual late-type galaxies potentially allow for the study of shocks in the disc, and the investigation of their role in the formation of spiral arms and stars.

The use of our shock finder could also be extended beyond simulation analysis. Shock properties are very sensitive to feedback processes, as we have demonstrated in this work, and might hence be used to constrain and validate feedback models adopted in cosmological simulations. This can be accomplished by modelling the radio synchrotron emission of feedback shocks, and comparing this emission to observations of radio telescopes such as LOFAR, or the forthcoming SKA. The latter will be of particular interest for shock analysis in simulations, since it will most likely be able to detect the synchrotron emission of the cosmic web, which is caused by relativistic electrons accelerated at the accretion shocks [Vazza et al., 2015].

Future state-of-the-art simulations will also take the self-consistent evolution of cosmic rays into account. Cosmic rays provide a non-thermal pressure component which is dynamically important for the formation of galaxies [Girichidis et al., 2016; Pfrommer et al., 2016], and their incorporation into simulations represents a milestone in the progress of cosmological simulation techniques. Cosmic rays are accelerated at supernova remnants and cosmological shocks, and the shock finder will be used to identify the locations of the latter. For detecting shocks in a two-fluid medium consisting of cosmic rays and thermal gas, we have to refine our shock finder methodology, something we already accomplished and reported recently in Pfrommer et al. [2016].

Appendix

Supplementary material

Eigenvectors of the Euler equations

The Eigenvalues of the flux Jacobian Matrix $\mathbf{A}_1 = \frac{\partial \mathbf{f}_1(\mathbf{u})}{\partial \mathbf{u}}$ are $\lambda_i = \{v_1 - c, v_1, v_1 + c, v_1, v_1\}$. The corresponding Eigenvectors are the columns of the matrix

$$\mathcal{R}_x = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ v_1 - c & v_1 & v_1 + c & 0 & 0 \\ v_2 & v_2 & v_2 & -1 & 0 \\ v_3 & v_3 & v_3 & 0 & 1 \\ h - cv_1 & k & h + cv_1 & -v_2 & v_3 \end{pmatrix}, \quad (\text{A.1})$$

with the specific kinetic energy $k = \frac{1}{2}(v_1^2 + v_2^2 + v_3^2)$ and the specific stagnation enthalpy $h = c^2/(\gamma - 1) + k$. The left eigenvectors of \mathbf{A}_1 are the rows of the Matrix $\mathcal{L}_x = \mathcal{R}_x^{-1}$. \mathcal{L}_x is the linear transformation operator from the conserved to the characteristic variables, $\mathbf{c} = \mathcal{L}_x \mathbf{u}$, and can be calculated as

$$\mathcal{L}_x = \begin{pmatrix} \beta(\phi + cv_1) & -\beta(\gamma_1 v_1 + c) & -\beta\gamma_1 v_2 & -\beta\gamma_1 v_3 & \beta\gamma_1 \\ 1 - 2\beta\phi & 2\beta\gamma_1 v_1 & 2\beta\gamma_1 v_2 & 2\beta\gamma_1 v_3 & -2\beta\gamma_1 \\ \beta(\phi - cv_1) & -\beta(\gamma_1 v_1 - c) & -\beta\gamma_1 v_2 & -\beta\gamma_1 v_3 & \beta\gamma_1 \\ v_2 & 0 & -1 & 0 & 0 \\ -v_3 & 0 & 0 & 1 & 0 \end{pmatrix}, \quad (\text{A.2})$$

where the definitions $\gamma_1 = \gamma - 1$, $\phi = \gamma_1 k$, and $\beta = 1/(2c^2)$ have been used. The eigenvector matrices for the flux Jacobians $\frac{\partial \mathbf{f}_2(\mathbf{u})}{\partial \mathbf{u}}$ and $\frac{\partial \mathbf{f}_3(\mathbf{u})}{\partial \mathbf{u}}$ are

$$\mathcal{R}_y = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ v_1 & v_1 & v_1 & 1 & 0 \\ v_2 - c & v_2 & v_2 + c & 0 & 0 \\ v_3 & v_3 & v_3 & 0 & -1 \\ h - cv_2 & k & h + cv_2 & v_1 & -v_3 \end{pmatrix}, \quad (\text{A.3})$$

$$\mathcal{L}_y = \begin{pmatrix} \beta(\phi + cv_2) & -\beta\gamma_1 v_1 & -\beta(\gamma_1 v_2 + c) & -\beta\gamma_1 v_3 & \beta\gamma_1 \\ 1 - 2\beta\phi & 2\beta\gamma_1 v_1 & 2\beta\gamma_1 v_2 & 2\beta\gamma_1 v_3 & -2\beta\gamma_1 \\ \beta(\phi - cv_2) & -\beta\gamma_1 v_1 & -\beta(\gamma_1 v_2 - c) & -\beta\gamma_1 v_3 & \beta\gamma_1 \\ -v_1 & 1 & 0 & 0 & 0 \\ v_3 & 0 & 0 & -1 & 0 \end{pmatrix}, \quad (\text{A.4})$$

and

$$\mathcal{R}_z = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ v_1 & v_1 & v_1 & -1 & 0 \\ v_2 & v_2 & v_2 & 0 & 1 \\ v_3 - c & v_3 & v_3 + c & 0 & 0 \\ h - cv_3 & k & h + cv_3 & -v_1 & v_2 \end{pmatrix}, \quad (\text{A.5})$$

$$\mathcal{L}_z = \begin{pmatrix} \beta(\phi + cv_3) & -\beta\gamma_1 v_1 & -\beta\gamma_1 v_2 & -\beta(\gamma_1 v_3 + c) & \beta\gamma_1 \\ 1 - 2\beta\phi & 2\beta\gamma_1 v_1 & 2\beta\gamma_1 v_2 & 2\beta\gamma_1 v_3 & -2\beta\gamma_1 \\ \beta(\phi - cv_3) & -\beta\gamma_1 v_1 & -\beta\gamma_1 v_2 & -\beta(\gamma_1 v_3 - c) & \beta\gamma_1 \\ v_1 & -1 & 0 & 0 & 0 \\ -v_2 & 0 & 1 & 0 & 0 \end{pmatrix}, \quad (\text{A.6})$$

respectively.

Runge-Kutta Butcher tables

Let $w(t)$ be the unknown scalar solution of the ordinary differential equation

$$\frac{dw}{dt} + R(t, w) = 0. \quad (\text{A.7})$$

The propagation of the numerical solution from timestep n to $n + 1$ with an s -stage explicit RK method can be written as

$$w^{n+1} = w^n + \Delta t^n \sum_{i=1}^s b_i k_i, \quad (\text{A.8})$$

with

$$k_i = -R(t^n + c_i \Delta t^n, w^n + \Delta t^n (a_{i1} k_1 + a_{i2} k_2 + \dots + a_{i,i-1} k_{i-1})). \quad (\text{A.9})$$

The RK scheme is fully specified by the weights b_i , the nodes c_i , and the RK matrix a_{ij} ; it can be represented in compact form by means of a Butcher tableau (Table A1).

For our DG code, we use SSP RK methods, which are convex combinations of forward Euler steps. In combination with a positivity preserving Riemann solver and the positivity limiter outlined in Section 19.4 negative pressure and density values can effectively be avoided in the hydro scheme. For the DG-2 code, we use SSP RK2

Heun’s method (Table A2), and for our third order code the SSP RK3 from Table A3. These methods are optimal in the sense that the number of stages is equal to the order of the scheme. It can be proven that a fourth order method with this feature does not exist [Gottlieb and Shu, 1998]; we therefore adopt the five-stage SSP RK 4 method tabulated in Table A4 [Spiteri and Ruuth, 2002].

c_1					
c_2	a_{21}				
c_3	a_{31}	a_{32}			
\vdots	\vdots	\vdots	\ddots		
c_s	a_{s1}	a_{s2}	\cdots	$a_{s,s-1}$	
	b_1	b_2	\cdots	b_{s-1}	b_s

Table A1: RK butcher tableau.

0	
1	1
	$\frac{1}{2}$ $\frac{1}{2}$

Table A2: SSP-RK2.

0		
1	1	
$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$
	$\frac{1}{6}$	$\frac{1}{6}$ $\frac{2}{3}$

Table A3: SSP-RK3.

0					
0.39175222700392	0.39175222700392				
0.58607968896779	0.21766909633821	0.36841059262959			
0.47454236302687	0.08269208670950	0.13995850206999	0.25189177424738		
0.93501063100924	0.06796628370320	0.11503469844438	0.20703489864929	0.54497475021237	
	0.14681187618661	0.24848290924556	0.10425883036650	0.27443890091960	0.22600748319395

Table A4: SSP-RK4.

DG refinement matrices

Below, we list the refinement matrices for merging the cells A, B, \dots, H into a coarser cell $K = \{\xi | \xi \in [-1, 1]^3\}$. We calculate the integrals by means of an exact numerical integration with $(k + 1)^3$ quadrature points before the main loop of our DG code.

Subcell $A = \{\xi | \xi \in [-1, 0] \times [-1, 0] \times [-1, 0]\}$:

$$(P_A)_{j,l} = \frac{1}{8} \iiint_{[-1,1]^3} \phi_j \left(\frac{\xi_1 - 1}{2}, \frac{\xi_2 - 1}{2}, \frac{\xi_3 - 1}{2} \right) \phi_l(\xi_1, \xi_2, \xi_3) d\xi. \quad (\text{A.10})$$

Subcell $B = \{\xi | \xi \in [0, 1] \times [-1, 0] \times [-1, 0]\}$:

$$(P_B)_{j,l} = \frac{1}{8} \iiint_{[-1,1]^3} \phi_j \left(\frac{\xi_1 + 1}{2}, \frac{\xi_2 - 1}{2}, \frac{\xi_3 - 1}{2} \right) \phi_l(\xi_1, \xi_2, \xi_3) d\xi. \quad (\text{A.11})$$

Subcell $C = \{\boldsymbol{\xi} | \boldsymbol{\xi} \in [-1, 0] \times [0, 1] \times [-1, 0]\}$:

$$(\mathbf{P}_C)_{j,l} = \frac{1}{8} \iiint_{[-1,1]^3} \phi_j \left(\frac{\xi_1 - 1}{2}, \frac{\xi_2 + 1}{2}, \frac{\xi_3 - 1}{2} \right) \phi_l(\xi_1, \xi_2, \xi_3) d\boldsymbol{\xi}. \quad (\text{A.12})$$

Subcell $D = \{\boldsymbol{\xi} | \boldsymbol{\xi} \in [0, 1] \times [0, 1] \times [-1, 0]\}$:

$$(\mathbf{P}_D)_{j,l} = \frac{1}{8} \iiint_{[-1,1]^3} \phi_j \left(\frac{\xi_1 + 1}{2}, \frac{\xi_2 + 1}{2}, \frac{\xi_3 - 1}{2} \right) \phi_l(\xi_1, \xi_2, \xi_3) d\boldsymbol{\xi}. \quad (\text{A.13})$$

Subcell $E = \{\boldsymbol{\xi} | \boldsymbol{\xi} \in [-1, 0] \times [-1, 0] \times [0, 1]\}$:

$$(\mathbf{P}_E)_{j,l} = \frac{1}{8} \iiint_{[-1,1]^3} \phi_j \left(\frac{\xi_1 - 1}{2}, \frac{\xi_2 - 1}{2}, \frac{\xi_3 + 1}{2} \right) \phi_l(\xi_1, \xi_2, \xi_3) d\boldsymbol{\xi}. \quad (\text{A.14})$$

Subcell $F = \{\boldsymbol{\xi} | \boldsymbol{\xi} \in [0, 1] \times [-1, 0] \times [0, 1]\}$:

$$(\mathbf{P}_F)_{j,l} = \frac{1}{8} \iiint_{[-1,1]^3} \phi_j \left(\frac{\xi_1 + 1}{2}, \frac{\xi_2 - 1}{2}, \frac{\xi_3 + 1}{2} \right) \phi_l(\xi_1, \xi_2, \xi_3) d\boldsymbol{\xi}. \quad (\text{A.15})$$

Subcell $G = \{\boldsymbol{\xi} | \boldsymbol{\xi} \in [-1, 0] \times [0, 1] \times [0, 1]\}$:

$$(\mathbf{P}_G)_{j,l} = \frac{1}{8} \iiint_{[-1,1]^3} \phi_j \left(\frac{\xi_1 - 1}{2}, \frac{\xi_2 + 1}{2}, \frac{\xi_3 + 1}{2} \right) \phi_l(\xi_1, \xi_2, \xi_3) d\boldsymbol{\xi}. \quad (\text{A.16})$$

Subcell $H = \{\boldsymbol{\xi} | \boldsymbol{\xi} \in [0, 1] \times [0, 1] \times [0, 1]\}$:

$$(\mathbf{P}_H)_{j,l} = \frac{1}{8} \iiint_{[-1,1]^3} \phi_j \left(\frac{\xi_1 + 1}{2}, \frac{\xi_2 + 1}{2}, \frac{\xi_3 + 1}{2} \right) \phi_l(\xi_1, \xi_2, \xi_3) d\boldsymbol{\xi}. \quad (\text{A.17})$$

List of Abbreviations

AMR	adaptive mesh refinement
AGN	active galactic nuclei
BH	black hole
CGM	circumgalactic medium
CMB	cosmic microwave background
CFL	Courant-Friedrichs-Lewy
DOF	degrees of freedom
DG	discontinuous Galerkin
DG-1	first order DG method
DG-2	second order DG method
DG-3	third order DG method
DG-4	fourth order DG method
DSA	diffusive shock acceleration
FV	finite volume
GLL	Gauss-Lobatto-Legendre
ICM	intracluster medium
IMF	initial mass function
ISM	interstellar medium
KH	Kelvin-Helmholtz
MHD	magnetohydrodynamics
MPI	Message Passing Interface
PIC	particle-in-cell
REA	reconstruct-evolve-average

List of Abbreviations

RK	Runge-Kutta
SKA	Square Kilometre Array
SN	supernova
SPH	smoothed-particle hydrodynamics
SSP	strong stability preserving
TVD	total variation diminishing
UVB	UV background
WHIM	warm-hot intergalactic medium

Table of Constants

Symbol	Value	Name
<hr/> <i>Physical constants</i> <hr/>		
c	$2.998 \times 10^{10} \text{ cm s}^{-1}$	speed of light in vacuum
e	$4.803 \times 10^{10} \text{ cm}^{3/2} \text{ g}^{1/2} \text{ s}^{-1}$	electron charge
G	$6.674 \times 10^{-8} \text{ cm}^3 \text{ g}^{-1} \text{ s}^{-2}$	gravitational constant
h	$6.626 \times 10^{-27} \text{ cm}^2 \text{ g s}^{-1}$	Planck constant
\hbar	$1.055 \times 10^{-27} \text{ cm}^2 \text{ g s}^{-1}$	reduced Planck constant
k_B	$1.381 \times 10^{-16} \text{ cm}^2 \text{ g s}^{-2} \text{ K}^{-1}$	Boltzmann constant
m_e	$9.109 \times 10^{-28} \text{ g}$	electron mass
m_p	$1.673 \times 10^{-24} \text{ g}$	proton mass
<hr/> <i>Astrophysical constants</i> <hr/>		
AU	$1.496 \times 10^{13} \text{ cm}$	astronomical unit
ly	$9.461 \times 10^{17} \text{ cm}$	light year
L_{\odot}	$3.839 \times 10^{33} \text{ erg s}^{-1}$	solar luminosity
M_{\odot}	$1.989 \times 10^{33} \text{ g}$	solar mass
pc	$3.086 \times 10^{18} \text{ cm}$	parsec

Table of Constants

R_{\odot}	6.955×10^{10} cm	solar radius
yr	3.156×10^7 s	year

Table of Symbols

Symbol	Formula	Name
<i>Ideal hydrodynamics</i>		
A	$\frac{\partial \mathbf{f}}{\partial \mathbf{u}}$	flux Jacobian
c	$\sqrt{\gamma p / \rho}$	sound speed
CFL		Courant-Friedrichs-Lewy number
\mathbf{c}	$\mathcal{L}_x \mathbf{u}$	vector of characteristic variables
E		total energy
e	$u + \frac{1}{2} \mathbf{v}^2$	total energy per unit mass
ϵ	ρu	thermal energy per unit volume
\mathbf{f}		flux vector
$\bar{\mathbf{f}}$		numerical flux function
γ		adiabatic index
j	ρv	mass flux
\mathcal{L}	\mathcal{R}^{-1}	matrix of left eigenvectors
λ		eigenvalue
m	$\int \rho dV$	mass
μ		mean molecular weight
N		number of particles

Table of Symbols

p	$\rho u(\gamma - 1)$	pressure
\mathcal{R}		matrix of right eigenvectors
ρ		density
s	p/ρ^γ	entropic function
T	$u(\gamma - 1)\mu m_p/k_B$	temperature
u		thermal energy per unit mass
\mathbf{u}	$(\rho, \rho \mathbf{v}, \rho e)^\top$	vector of conserved variables
\mathbf{U}	$\int \mathbf{u} dV$	vector of mean conserved quantities
V		volume
\bar{V}	$1/\rho$	specific volume
\mathbf{v}		velocity
\mathbf{W}	$(\rho, \mathbf{v}, p)^\top$	vector of primitive variables

Shock quantities

χ_1		pre-shock quantity
χ_2		post-shock quantity
$\delta(\mathcal{M})$	f_{th}/f_ϕ	thermalisation efficiency
\mathbf{d}_s	$-\nabla T/ \nabla T $	shock direction
f_{th}	$[e_{\text{th},2} - e_{\text{th},1}(\rho_2/\rho_1)^\gamma] v_2$	dissipated thermal energy flux
f_ϕ	$\frac{1}{2}\rho_1 v_1^3$	incoming kinetic energy flux
\mathcal{M}	v_1/c_1	Mach number
R	ρ_2/ρ_1	density jump

Discontinuous Galerkin

A, B, C, \dots, H daughter cells

α		refinement/deref. parameter
β		slope limiter parameter
Δx^K		edge length of K
Δt^K		timestep of K
N		number of basis functions
ϕ		basis function
k		maximum polynomial order
K		cell K
$ K $		volume of cell K
L		angular momentum density
\tilde{M}		slope limiter parameter
ω_q^{3D}		3D Gauss quadrature weight
$\hat{\omega}_q^{1D}$		1D GLL quadrature weight
p	$k + 1$	degree of the scheme
P_A, P_B, \dots, P_H		projection matrices
$P(\xi)$		Legendre polynomial
$\tilde{P}(\xi)$		scaled Legendre polynomial
Φ		external potential
\mathbf{w}		weight vector
w_t		refinement/deref. target slope
\mathbf{x}^K		cell centre of K
ξ	$\in [-1, 1]^3$	scaled variables
ξ_q^{1D}		1D Gauss quadrature node
$\hat{\xi}_q^{1D}$		1D GLL quadrature node
ξ_q^{3D}		3D Gauss quadrature node

Bibliography

- M. Abramowitz and I. Stegun. *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*. Dover Books on Mathematics. Dover Publications, 2012. ISBN 9780486158242.
- M. Actis, G. Agnetta, F. Aharonian, A. Akhperjanian, J. Aleksić, E. Aliu, D. Allan, I. Allekotte, F. Antico, L. A. Antonelli, and et al. Design concepts for the Cherenkov Telescope Array CTA: an advanced facility for ground-based high-energy gamma-ray astronomy. *Exp. Astron.*, 32:193–316, 2011.
- H. Akamatsu and H. Kawahara. Systematic X-Ray Analysis of Radio Relic Clusters with Suzaku. *PASJ*, 65, 2013.
- T. Amano and M. Hoshino. Electron Injection at High Mach Number Quasi-perpendicular Shocks: Surfing and Drift Acceleration. *ApJ*, 661:190–202, 2007.
- T. Amano and M. Hoshino. A Critical Mach Number for Electron Injection in Collisionless Shocks. *Physical Review Letters*, 104(18):181102, 2010.
- E. Amato and P. Blasi. Non-linear particle acceleration at non-relativistic shock waves in the presence of self-generated turbulence. *MNRAS*, 371:1251–1258, 2006.
- G. Arfken and H. Weber. *Mathematical Methods for Physicists*. Elsevier Science, 2013. ISBN 9781483288062.
- Y. Ascasibar and M. Markevitch. The Origin of Cold Fronts in the Cores of Relaxed Galaxy Clusters. *ApJ*, 650:102–127, 2006.
- W. I. Axford, E. Leer, and G. Skadron. The acceleration of cosmic rays by shock waves. *International Cosmic Ray Conference*, 11:132–137, 1977.
- J. Barnes and P. Hut. A hierarchical $O(N \log N)$ force-calculation algorithm. *Nature*, 324:446–449, 1986.
- N. Battaglia, C. Pfrommer, J. L. Sievers, J. R. Bond, and T. A. Enßlin. Exploring the magnetized cosmic web through low-frequency radio emission. *MNRAS*, 393:1073–1089, 2009.
- N. Battaglia, J. R. Bond, C. Pfrommer, and J. L. Sievers. On the Cluster Physics of Sunyaev-Zel’dovich and X-Ray Surveys. I. The Influence of Feedback, Non-thermal Pressure, and Cluster Shapes on Y-M Scaling Relations. *ApJ*, 758:74, 2012.

Bibliography

- A. Bauer and V. Springel. Subsonic turbulence in smoothed particle hydrodynamics and moving-mesh simulations. *MNRAS*, 423:2558–2578, 2012.
- A. Bauer, V. Springel, M. Vogelsberger, S. Genel, P. Torrey, D. Sijacki, D. Nelson, and L. Hernquist. Hydrogen reionization in the Illustris universe. *MNRAS*, 453:3593–3610, 2015.
- A. Bauer, K. Schaal, V. Springel, P. Chandrashekar, R. Pakmor, and C. Klingenberg. Simulating Turbulence Using the Astrophysical Discontinuous Galerkin Code TENET. *ArXiv e-prints*, 2016.
- P. S. Behroozi, R. H. Wechsler, and C. Conroy. The Average Star Formation Histories of Galaxies in Dark Matter Halos from $z = 0-8$. *ApJ*, 770:57, 2013.
- A. R. Bell. The acceleration of cosmic rays in shock fronts. I. *MNRAS*, 182:147–156, 1978a.
- A. R. Bell. The acceleration of cosmic rays in shock fronts. II. *MNRAS*, 182:443–455, 1978b.
- M. Bernardi, A. Meert, R. K. Sheth, V. Vikram, M. Huertas-Company, S. Mei, and F. Shankar. The massive end of the luminosity and stellar mass functions: dependence on the fit to the light profile. *MNRAS*, 436:697–704, 2013.
- S. Bird, M. Vogelsberger, M. Haehnelt, D. Sijacki, S. Genel, P. Torrey, V. Springel, and L. Hernquist. Damped Lyman α absorbers as a probe of stellar feedback. *MNRAS*, 445:2313–2324, 2014.
- Y. Birnboim and A. Dekel. Virial shocks in galactic haloes? *MNRAS*, 345:349–364, 2003.
- R. D. Blandford and J. P. Ostriker. Particle acceleration by astrophysical shocks. *ApJ*, 221:L29–L32, 1978.
- E. L. Blanton, S. W. Randall, E. M. Douglass, C. L. Sarazin, T. E. Clarke, and B. R. McNamara. Shocks and Bubbles in a Deep Chandra Observation of the Cooling Flow Cluster Abell 2052. *ApJ*, 697:L95–L98, 2009.
- P. Blasi. A semi-analytical approach to non-linear shock acceleration. *Astropart. Phys.*, 16:429–439, 2002.
- A. Bonafede, G. Giovannini, L. Feretti, F. Govoni, and M. Murgia. Double relics in Abell 2345 and Abell 1240. Spectral index and polarization analysis. *A&A*, 494:429–442, 2009.
- A. Bonafede, M. Brüggen, R. van Weeren, F. Vazza, G. Giovannini, H. Ebeling, A. C. Edge, M. Hoeft, and U. Klein. Discovery of radio haloes and double relics in distant MACS galaxy clusters: clues to the efficiency of particle acceleration. *MNRAS*, 426:40–56, 2012.

- H. Bourdin, P. Mazzotta, M. Markevitch, S. Giacintucci, and G. Brunetti. Shock Heating of the Merging Galaxy Cluster A521. *ApJ*, 764:82, 2013.
- A. E. Broderick, P. Chang, and C. Pfrommer. The Cosmological Impact of Luminous TeV Blazars. I. Implications of Plasma Instabilities for the Intergalactic Magnetic Field and Extragalactic Gamma-Ray Background. *ApJ*, 752:22, 2012.
- M. Brüggen, R. J. van Weeren, and H. J. A. Röttgering. Magnetic fields and shock waves in cluster outskirts. *Mem. Soc. Astron. Italiana*, 82:627, 2011.
- M. Brüggen, A. Bykov, D. Ryu, and H. Röttgering. Magnetic Fields, Relativistic Particles, and Shock Waves in Cluster Outskirts. *Space Sci. Rev.*, 166:187–213, 2012.
- G. Brunetti and T. W. Jones. Cosmic Rays in Galaxy Clusters and Their Nonthermal Emission. *International Journal of Modern Physics D*, 23:1430007-98, 2014.
- G. L. Bryan, M. L. Norman, B. W. O’Shea, T. Abel, J. H. Wise, M. J. Turk, D. R. Reynolds, and D. C. Collins. ENZO: An Adaptive Mesh Refinement Code for Astrophysics. *ApJS*, 211:19, 2014.
- D. Caprioli and A. Spitkovsky. Simulations of Ion Acceleration at Non-relativistic Shocks. I. Acceleration Efficiency. *ApJ*, 783:91, 2014.
- K. W. Cavagnolo, B. R. McNamara, M. W. Wise, P. E. J. Nulsen, M. Brüggen, M. Gitti, and D. A. Rafferty. A Powerful AGN Outburst in RBS 797. *ApJ*, 732:71, 2011.
- G. Chabrier. Galactic Stellar and Substellar Initial Mass Function. *PASP*, 115:763–795, 2003.
- S. Chandrasekhar. *Hydrodynamic and Hydromagnetic Stability*. Dover Books on Physics Series. Dover Publications, 1961. ISBN 9780486640716.
- T. E. Clarke and T. Ensslin. Mergers, relics and magnetic fields: The complex case of A2256. *Astron. Nachr.*, 327:553, 2006.
- B. Cockburn, S.-Y. Lin, and C.-W. Shu. TVB Runge Kutta Local Projection Discontinuous Galerkin Finite Element Method for Conservation Laws III: One-Dimensional Systems. *Journal of Computational Physics*, 84:90–113, 1989.
- B. Cockburn, S. Hou, and C.-W. Shu. The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. IV. The multidimensional case. *Mathematics of Computation*, 54:545–581, 1990.
- B. Cockburn, G. Karniadakis, and C. Shu. *Discontinuous Galerkin Methods: Theory, Computation and Applications*. Lecture Notes in Computational Science and Engineering. Springer Berlin Heidelberg, 2011. ISBN 9783642640988.

Bibliography

- B. Cockburn and C.-W. Shu. Tvb runge-kutta local projection discontinuous galerkin finite element method for conservation laws. ii. general framework. *Mathematics of Computation*, 52(186):411–435, 1989.
- B. Cockburn and C.-W. Shu. The runge-kutta local projection p^1 -discontinuous-galerkin finite element method for scalar conservation laws. *RAIRO-Modélisation mathématique et analyse numérique*, 25(3):337–361, 1991.
- B. Cockburn and C.-W. Shu. The Runge-Kutta Discontinuous Galerkin Method for Conservation Laws V. Multidimensional Systems. *Journal of Computational Physics*, 141:199–224, 1998.
- L. Cullen and W. Dehnen. Inviscid smoothed particle hydrodynamics. *MNRAS*, 408:669–683, 2010.
- S. Dasadia, M. Sun, C. Sarazin, A. Morandi, M. Markevitch, D. Wik, L. Feretti, G. Giovannini, F. Govoni, and V. Vacca. A Strong Merger Shock in Abell 665. *ApJ*, 820:L20, 2016.
- R. Davé, R. Cen, J. P. Ostriker, G. L. Bryan, L. Hernquist, N. Katz, D. H. Weinberg, M. L. Norman, and B. O’Shea. Baryons in the Warm-Hot Intergalactic Medium. *ApJ*, 552:473–483, 2001.
- A. Dekel and Y. Birnboim. Galaxy bimodality due to cold flows and shock heating. *MNRAS*, 368:2–20, 2006.
- T. Di Matteo, V. Springel, and L. Hernquist. Energy input from quasars regulates the growth and activity of black holes and their host galaxies. *Nature*, 433:604–607, 2005.
- K. Dolag, F. Vazza, G. Brunetti, and G. Tormen. Turbulent gas motions in galaxy cluster simulations: the role of smoothed particle hydrodynamics viscosity. *MNRAS*, 364:753–772, 2005.
- L. O. Drury. An introduction to the theory of diffusive shock acceleration of energetic particles in tenuous plasmas. *Reports on Progress in Physics*, 46:973–1027, 1983.
- P. C. Duffell and A. I. MacFadyen. TESS: A Relativistic Hydrodynamics Code on a Moving Voronoi Mesh. *ApJS*, 197:15, 2011.
- P. P. Edmon, H. Kang, T. W. Jones, and R. Ma. Non-thermal radiation from Type Ia supernova remnants. *MNRAS*, 414:3521–3536, 2011.
- D. C. Ellison, M. G. Baring, and F. C. Jones. Nonlinear Particle Acceleration in Oblique Shocks. *ApJ*, 473:1029, 1996.
- T. Enßlin, C. Pfrommer, F. Miniati, and K. Subramanian. Cosmic ray transport in galaxy clusters: implications for radio halos, gamma-ray signatures, and cool core heating. *A&A*, 527:A99, 2011.

- T. A. Enßlin and M. Brüggen. On the formation of cluster radio relics. *MNRAS*, 331:1011–1019, 2002.
- T. A. Enßlin and Gopal-Krishna. Reviving fossil radio plasma in clusters of galaxies by adiabatic compression in environmental shock waves. *A&A*, 366:26–34, 2001.
- T. A. Ensslin, P. L. Biermann, U. Klein, and S. Kohle. Cluster radio relics as a tracer of shock waves of the large-scale structure formation. *A&A*, 332:395–409, 1998.
- M. A. Fardal, N. Katz, J. P. Gardner, L. Hernquist, D. H. Weinberg, and R. Davé. Cooling Radiation and the Ly α Luminosity of Forming Galaxies. *ApJ*, 562:605–617, 2001.
- C.-A. Faucher-Giguère, A. Lidz, M. Zaldarriaga, and L. Hernquist. A New Calculation of the Ionizing Background Spectrum and the Effects of He II Reionization. *ApJ*, 703:1416–1443, 2009.
- L. Feretti. Radio Observations of Cluster Mergers. In L. O. Sjouwerman and K. K. Dyer, editors, *X-Ray and Radio Connections*, 2005.
- L. Feretti, G. Giovannini, F. Govoni, and M. Murgia. Clusters of galaxies: observational properties of the diffuse radio emission. *A&A Rev.*, 20:54, 2012.
- Fermi Collaboration. Detection of the Characteristic Pion-Decay Signature in Supernova Remnants. *Science*, 339:807–811, 2013.
- G. Ferrand, R. J. Danos, A. Shalchi, S. Safi-Harb, P. Edmon, and P. Mendygral. Cosmic Ray Acceleration at Perpendicular Shocks in Supernova Remnants. *ApJ*, 792:133, 2014.
- C. Ferrari, F. Govoni, S. Schindler, A. M. Bykov, and Y. Rephaeli. Observations of Extended Radio Emission in Clusters. *Space Sci. Rev.*, 134:93–118, 2008.
- B. Fryxell, K. Olson, P. Ricker, F. X. Timmes, M. Zingale, D. Q. Lamb, P. MacNeice, R. Rosner, J. W. Truran, and H. Tufo. FLASH: An Adaptive Mesh Hydrodynamics Code for Modeling Astrophysical Thermonuclear Flashes. *ApJS*, 131:273–334, 2000.
- J. P. Gallego-Valencia, J. Löbber, S. Müthing, P. Bastian, C. Klingenberg, and Y. Xia. Implementing a discontinuous galerkin method for the compressible, inviscid euler equations in the dune framework. *PAMM*, 14(1):953–954, 2014.
- S. Genel, M. Vogelsberger, D. Nelson, D. Sijacki, V. Springel, and L. Hernquist. Following the flow: tracer particles in astrophysical fluid simulations. *MNRAS*, 435:1426–1442, 2013.
- S. Genel, M. Vogelsberger, V. Springel, D. Sijacki, D. Nelson, G. Snyder, V. Rodriguez-Gomez, P. Torrey, and L. Hernquist. Introducing the Illustris project: the evolution of galaxy populations across cosmic time. *MNRAS*, 445:175–200, 2014.

Bibliography

- P. Girichidis, T. Naab, S. Walch, M. Hanasz, M.-M. Mac Low, J. P. Ostriker, A. Gatto, T. Peters, R. Wünsch, S. C. O. Glover, R. S. Klessen, P. C. Clark, and C. Baczynski. Launching Cosmic-Ray-driven Outflows from the Magnetized Interstellar Medium. *ApJ*, 816:L19, 2016.
- M. Gitti, P. E. J. Nulsen, L. P. David, B. R. McNamara, and M. W. Wise. A Chandra Study of the Large-scale Shock and Cool Filaments in Hydra A: Evidence for Substantial Gas Dredge-up by the Central Outburst. *ApJ*, 732:13, 2011.
- A. Giuliani, M. Cardillo, M. Tavani, Y. Fukui, S. Yoshiike, K. Torii, G. Dubner, G. Castelletti, G. Barbiellini, A. Bulgarelli, P. Caraveo, E. Costa, P. W. Cattaneo, A. Chen, T. Contessi, E. Del Monte, I. Donnarumma, Y. Evangelista, M. Feroci, F. Gianotti, F. Lazzarotto, F. Lucarelli, F. Longo, M. Marisaldi, S. Mereghetti, L. Pacciani, A. Pellizzoni, G. Piano, P. Picozza, C. Pittori, G. Pucella, M. Rapisarda, A. Rappoldi, S. Sabatini, P. Soffitta, E. Striani, M. Trifoglio, A. Trois, S. Vercellone, F. Verrecchia, V. Vittorini, S. Colafrancesco, P. Giommi, and G. Bignami. Neutral Pion Emission from Accelerated Protons in the Supernova Remnant W44. *ApJ*, 742:L30, 2011.
- S. K. Godunov. A finite difference method for the computation of discontinuous solutions of the equations of fluid dynamics. *Matematicheskii Sbornik*, 47:357–393, 1959.
- S. Gottlieb and C.-W. Shu. Total variation diminishing runge-kutta schemes. *Mathematics of computation of the American Mathematical Society*, 67(221):73–85, 1998.
- S. Gottlieb, C.-W. Shu, and E. Tadmor. Strong stability-preserving high-order time discretization methods. *SIAM review*, 43(1):89–112, 2001.
- J. E. Gunn and J. R. Gott, III. On the Infall of Matter Into Clusters of Galaxies and Some Effects on Their Evolution. *ApJ*, 176:1, 1972.
- X. Guo, L. Sironi, and R. Narayan. Non-thermal Electron Acceleration in Low Mach Number Collisionless Shocks. I. Particle Energy Spectra and Acceleration Mechanism. *ApJ*, 794:153, 2014.
- M. Haider, D. Steinhauser, M. Vogelsberger, S. Genel, V. Springel, P. Torrey, and L. Hernquist. Large-scale mass distribution in the Illustris simulation. *MNRAS*, 457:3024–3035, 2016.
- T. Hendrix and R. Keppens. Effect of dust on Kelvin-Helmholtz instabilities. *A&A*, 562:A114, 2014.
- G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley, M. R. Nolte, and Halpern. Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results. *ApJS*, 208:19, 2013.

- M. Hoeft, M. Brüggen, G. Yepes, S. Gottlöber, and A. Schwobe. Diffuse radio emission from clusters in the MareNostrum Universe simulation. *MNRAS*, 391:1511–1526, 2008.
- S. E. Hong, D. Ryu, H. Kang, and R. Cen. Shock Waves and Cosmic Ray Acceleration in the Outskirts of Galaxy Clusters. *ApJ*, 785:133, 2014.
- P. F. Hopkins. A new class of accurate, mesh-free hydrodynamic simulation methods. *MNRAS*, 450:53–110, 2015.
- L. Hui and N. Y. Gnedin. Equation of state of the photoionized intergalactic medium. *MNRAS*, 292:27, 1997.
- A. Jeffrey. *Quasilinear Hyperbolic Systems and Waves*. Research Notes in Mathematics. Pitman Publishing, 1976. ISBN 9780273001027.
- N. N. Jetha, M. J. Hardcastle, T. J. Ponman, and I. Sakelliou. Shock heating in the group atmosphere of the radio galaxy B2 0838+32A. *MNRAS*, 391:1052–1062, 2008.
- M. Jubelgas, V. Springel, T. Enßlin, and C. Pfrommer. Cosmic ray feedback in hydrodynamical simulations of galaxy formation. *A&A*, 481:33–63, 2008.
- J. Kamm and F. Timmes. On efficient generation of numerically robust sedov solutions. *Technical Report LA-UR-07-2849*, 2007.
- H. Kang and T. W. Jones. Self-similar evolution of cosmic-ray-modified quasi-parallel plane shocks. *Astropart. Phys.*, 28:232–246, 2007a.
- H. Kang and T. W. Jones. Self-similar evolution of cosmic-ray-modified quasi-parallel plane shocks. *Astroparticle Physics*, 28:232–246, 2007b.
- H. Kang and D. Ryu. Diffusive Shock Acceleration at Cosmological Shock Waves. *ApJ*, 764:95, 2013.
- H. Kang, D. Ryu, R. Cen, and J. P. Ostriker. Cosmological Shock Waves in the Large-Scale Structure of the Universe: Nongravitational Effects. *ApJ*, 669:729–740, 2007.
- N. Katz, D. H. Weinberg, and L. Hernquist. Cosmological Simulations with TreeSPH. *ApJS*, 105:19, 1996.
- C. F. Kennel, J. P. Edmiston, and T. Hada. A quarter century of collisionless shock research. *Washington DC American Geophysical Union Geophysical Monograph Series*, 34:1–36, 1985.
- U. Keshet, E. Waxman, A. Loeb, V. Springel, and L. Hernquist. Gamma Rays from Intergalactic Shocks. *ApJ*, 585:128–150, 2003.

Bibliography

- U. Keshet, E. Waxman, and A. Loeb. Searching for intergalactic shocks with the Square Kilometer Array. *New A Rev.*, 48:1119–1135, 2004.
- G. F. Krymskii. A regular mechanism for the acceleration of charged particles on the front of a shock wave. *Akademiia Nauk SSSR Doklady*, 234:1306–1308, 1977.
- D. Kuzmin. Hierarchical slope limiting in explicit and implicit discontinuous Galerkin methods. *Journal of Computational Physics*, 257:1140–1162, 2014.
- L. Landau and E. Lifshitz. *Course of Theoretical Physics, Vol. 6, Fluid Mechanics*. Pergamon Press, London, 1966.
- N. Lanson and J.-P. Vila. Renormalized meshfree schemes i: consistency, stability, and hybrid methods for conservation laws. *SIAM Journal on Numerical Analysis*, 46(4):1912–1934, 2008.
- F. Lether and P. Wenston. Minimax approximations to the zeros of $p_n(x)$ and gauss-legendre quadrature. *Journal of Computational and Applied Mathematics*, 59(2):245 – 252, 1995.
- R. LeVeque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2002. ISBN 9780521009249.
- S. Lloyd. Least squares quantization in pcm. *IEEE Trans. Inf. Theory*, 28(2):129–137, 1982.
- G. Macario, M. Markevitch, S. Giacintucci, G. Brunetti, T. Venturi, and S. S. Murray. A Shock Front in the Merging Galaxy Cluster A754: X-ray and Radio Observations. *ApJ*, 728:82, 2011.
- M. A. Malkov. Analytic Solution for Nonlinear Shock Acceleration in the Bohm Limit. *ApJ*, 485:638–654, 1997.
- M. A. Malkov and L. O’C Drury. Nonlinear theory of diffusive acceleration of particles by shock waves. *Rep. Progress Phys.*, 64:429–481, 2001.
- F. Marinacci, R. Pakmor, V. Springel, and C. M. Simpson. Diffuse gas properties and stellar metallicities in cosmological simulations of disc galaxy formation. *MNRAS*, 442:3745–3760, 2014.
- M. Markevitch. Chandra Observation of the Most Interesting Cluster in the Universe. In A. Wilson, editor, *Proceedings of the The X-ray Universe 2005, ESA SP-604*, volume 604, page 723, El Escorial, Madrid, Spain, 2006.
- M. Markevitch. Intergalactic shock fronts. *ArXiv e-prints*, 2010.
- M. Markevitch and A. Vikhlinin. Shocks and cold fronts in galaxy clusters. *Phys. Rep.*, 443:1–53, 2007.

- M. Markevitch, A. H. Gonzalez, L. David, A. Vikhlinin, S. Murray, W. Forman, C. Jones, and W. Tucker. A Textbook Example of a Bow Shock in the Merging Galaxy Cluster 1E 0657-56. *ApJ*, 567:L27–L31, 2002.
- M. Markevitch, F. Govoni, G. Brunetti, and D. Jerius. Bow Shock and Radio Halo in the Merging Cluster A520. *ApJ*, 627:733–738, 2005.
- F. Masset. FARGO: A fast eulerian transport algorithm for differentially rotating disks. *A&AS*, 141:165–173, 2000.
- Y. Matsumoto, T. Amano, and M. Hoshino. Electron Accelerations at High Mach Number Shocks: Two-dimensional Particle-in-cell Simulations in Various Parameter Regimes. *ApJ*, 755:109, 2012.
- C. P. McNally, W. Lyra, and J.-C. Passy. A Well-posed Kelvin-Helmholtz Instability Test and Comparison. *ApJS*, 201:18, 2012.
- A. Mignone, G. Bodo, S. Massaglia, T. Matsakos, O. Tesileanu, C. Zanni, and A. Ferrari. PLUTO: A Numerical Code for Computational Astrophysics. *ApJS*, 170:228–242, 2007.
- F. Miniati. The Matryoshka Run: A Eulerian Refinement Strategy to Study the Statistics of Turbulence in Virialized Cosmic Structures. *ApJ*, 782:21, 2014.
- F. Miniati. The Matryoshka Run. II. Time-dependent Turbulence Statistics, Stochastic Particle Acceleration, and Microphysics Impact in a Massive Galaxy Cluster. *ApJ*, 800:60, 2015.
- F. Miniati, D. Ryu, H. Kang, T. W. Jones, R. Cen, and J. P. Ostriker. Properties of Cosmic Shock Waves in Large-Scale Structure Formation. *ApJ*, 542:608–621, 2000.
- F. Miniati, T. W. Jones, H. Kang, and D. Ryu. Cosmic-Ray Electrons in Groups and Clusters of Galaxies: Primary and Secondary Populations from a Numerical Cosmological Simulation. *ApJ*, 562:233–253, 2001.
- P. Mocz, M. Vogelsberger, D. Sijacki, R. Pakmor, and L. Hernquist. A discontinuous Galerkin method for solving the fluid and magnetohydrodynamic equations in astrophysical simulations. *MNRAS*, 437:397–414, 2014.
- R. Moura, S. Sherwin, and J. Peiró. Modified equation analysis for linear advection with dg. Technical report, Research report, Imperial College London, 2014.
- D. Nelson, M. Vogelsberger, S. Genel, D. Sijacki, D. Kereš, V. Springel, and L. Hernquist. Moving mesh cosmology: tracing cosmological gas accretion. *MNRAS*, 429:3353–3370, 2013.
- D. Nelson, S. Genel, A. Pillepich, M. Vogelsberger, V. Springel, and L. Hernquist. Zooming in on accretion - I. The structure of halo gas. *ArXiv e-prints*, 2015a.

Bibliography

- D. Nelson, S. Genel, M. Vogelsberger, V. Springel, D. Sijacki, P. Torrey, and L. Hernquist. The impact of feedback on cosmological gas accretion. *MNRAS*, 448:59–74, 2015b.
- D. Nelson, A. Pillepich, S. Genel, M. Vogelsberger, V. Springel, P. Torrey, V. Rodriguez-Gomez, D. Sijacki, G. F. Snyder, B. Griffen, F. Marinacci, L. Blecha, L. Sales, D. Xu, and L. Hernquist. The illustris simulation: Public data release. *Astronomy and Computing*, 13:12–37, 2015c.
- W. Nolting. *Grundkurs Theoretische Physik 4: Spezielle Relativitätstheorie, Thermodynamik*. Springer-Lehrbuch. Springer Berlin Heidelberg, 2013. ISBN 9783662099735.
- P. E. J. Nulsen. Transport processes and the stripping of cluster galaxies. *MNRAS*, 198:1007–1016, 1982.
- P. E. J. Nulsen, D. C. Hambrick, B. R. McNamara, D. Rafferty, L. Birzan, M. W. Wise, and L. P. David. The Powerful Outburst in Hercules A. *ApJ*, 625:L9–L12, 2005a.
- P. E. J. Nulsen, B. R. McNamara, M. W. Wise, and L. P. David. The Cluster-Scale AGN Outburst in Hydra A. *ApJ*, 628:629–636, 2005b.
- G. A. Ogrean and M. Brüggen. First X-ray evidence for a shock at the Coma relic. *MNRAS*, 433:1701–1708, 2013.
- M. S. Owers, P. E. J. Nulsen, W. J. Couch, C.-J. Ma, L. P. David, W. R. Forman, A. M. Hopkins, C. Jones, and R. J. van Weeren. A Merger Shock in A2034. *ApJ*, 780:163, 2014.
- T. Padmanabhan. *Theoretical Astrophysics: Volume 1, Astrophysical Processes*. Theoretical Astrophysics. Cambridge University Press, 2000. ISBN 9780521566322.
- R. Pakmor, V. Springel, A. Bauer, P. Mocz, D. J. Munoz, S. T. Ohlmann, K. Schaal, and C. Zhu. Improving the convergence properties of the moving-mesh code AREPO. *MNRAS*, 455:1134–1143, 2016.
- C. Pfrommer. Simulating cosmic rays in clusters of galaxies - III. Non-thermal scaling relations and comparison to observations. *MNRAS*, 385:1242–1256, 2008.
- C. Pfrommer and T. W. Jones. Radio Galaxy NGC 1265 Unveils the Accretion Shock Onto the Perseus Galaxy Cluster. *ApJ*, 730:22, 2011.
- C. Pfrommer, V. Springel, T. A. Enßlin, and M. Jubelgas. Detecting shock waves in cosmological smoothed particle hydrodynamics simulations. *MNRAS*, 367:113–131, 2006.

- C. Pfrommer, T. A. Enßlin, V. Springel, M. Jubelgas, and K. Dolag. Simulating cosmic rays in clusters of galaxies - I. Effects on the Sunyaev-Zel'dovich effect and the X-ray emission. *MNRAS*, 378:385–408, 2007.
- C. Pfrommer, T. A. Enßlin, and V. Springel. Simulating cosmic rays in clusters of galaxies - II. A unified scheme for radio haloes and relics with predictions of the γ -ray emission. *MNRAS*, 385:1211–1241, 2008.
- C. Pfrommer, P. Chang, and A. E. Broderick. The Cosmological Impact of Luminous TeV Blazars. III. Implications for Galaxy Clusters and the Formation of Dwarf Galaxies. *ApJ*, 752:24, 2012.
- C. Pfrommer, R. Pakmor, K. Schaal, C. M. Simpson, and V. Springel. Simulating cosmic ray physics on a moving mesh. *ArXiv e-prints*, 2016.
- C. Pfrommer. The physics of galaxy clusters. *Lecture notes, University of Heidelberg*, 2016.
- A. Pinzke, S. P. Oh, and C. Pfrommer. Giant radio relics in galaxy clusters: reacceleration of fossil relativistic electrons? *MNRAS*, 435:1061–1082, 2013.
- Planck Collaboration X. Planck intermediate results. X. Physics of the hot gas in the Coma cluster. *A&A*, 554:A140, 2013.
- S. Planelles and V. Quilis. Cosmological shock waves: clues to the formation history of haloes. *MNRAS*, 428:1643–1655, 2013.
- K. B. Quest. Theory and simulation of collisionless parallel shocks. *J. Geophys. Res.*, 93:9649–9680, 1988.
- V. Quilis, J. M. A. Ibanez, and D. Saez. On the Role of Shock Waves in Galaxy Cluster Evolution. *ApJ*, 502:518, 1998.
- S. W. Randall, W. R. Forman, S. Giacintucci, P. E. J. Nulsen, M. Sun, C. Jones, E. Churazov, L. P. David, R. Kraft, M. Donahue, E. L. Blanton, A. Simionescu, and N. Werner. Shocks and Cavities from Multiple Outbursts in the Galaxy Group NGC 5813: A Window to Active Galactic Nucleus Feedback. *ApJ*, 726:86, 2011.
- W. H. Reed and T. Hill. Triangularmesh methodsfor the neutrontransportequation. *Los Alamos Report LA-UR-73-479*, 1973.
- H. Reinhardt. Thermodynamik und statistik. *Lecture notes, University of Tübingen*, 2009.
- S. P. Reynolds. Supernova Remnants at High Energy. *ARA&A*, 46:89–126, 2008.
- M. A. Riquelme and A. Spitkovsky. Electron Injection by Whistler Waves in Non-relativistic Shocks. *ApJ*, 733:63, 2011.

Bibliography

- B. E. Robertson, A. V. Kravtsov, N. Y. Gnedin, T. Abel, and D. H. Rudd. Computational Eulerian hydrodynamics and Galilean invariance. *MNRAS*, 401:2463–2476, 2010.
- E. Roediger and J. A. Zuhone. Fast simulations of gas sloshing and cold front formation. *MNRAS*, 419:1338–1349, 2012.
- H. J. A. Rottgering, M. H. Wieringa, R. W. Hunstead, and R. D. Ekers. The extended radio emission in the luminous X-ray cluster A3667. *MNRAS*, 290:577–584, 1997.
- H. R. Russell, A. C. Fabian, B. R. McNamara, A. C. Edge, J. S. Sanders, P. E. J. Nulsen, S. A. Baum, M. Donahue, and C. P. O’Dea. The bow shock, cold fronts and disintegrating cool core in the merging galaxy group RX J0751.3+5012. *MNRAS*, 444:629–641, 2014.
- D. Ryu, H. Kang, E. Hallman, and T. W. Jones. Cosmological Shock Waves and Their Role in the Large-Scale Structure of the Universe. *ApJ*, 593:599–610, 2003.
- K. Schaal and V. Springel. Shock finding on a moving mesh - I. Shock statistics in non-radiative cosmological simulations. *MNRAS*, 446:3992–4007, 2015.
- K. Schaal, A. Bauer, P. Chandrashekar, R. Pakmor, C. Klingenberg, and V. Springel. Astrophysical hydrodynamics with a high-order discontinuous Galerkin scheme and adaptive mesh refinement. *MNRAS*, 453:4278–4300, 2015.
- K. Schaal, V. Springel, R. Pakmor, C. Pfrommer, D. Nelson, M. Vogelsberger, S. Genel, A. Pillepich, D. Sijacki, and L. Hernquist. Shock finding on a moving-mesh: II. Hydrodynamic shocks in the Illustris universe. *ArXiv e-prints*, 2016.
- K. Schaal. Havoc-here’s another voronoi code. *Diploma thesis, University of Tübingen*, 2013.
- D. Sijacki, V. Springel, T. Di Matteo, and L. Hernquist. A unified model for AGN feedback in cosmological simulations of structure formation. *MNRAS*, 380:877–900, 2007.
- D. Sijacki, M. Vogelsberger, S. Genel, V. Springel, P. Torrey, G. F. Snyder, D. Nelson, and L. Hernquist. The Illustris simulation: the evolving population of black holes across cosmic time. *MNRAS*, 452:575–596, 2015.
- S. W. Skillman, B. W. O’Shea, E. J. Hallman, J. O. Burns, and M. L. Norman. Cosmological Shocks in Adaptive Mesh Refinement Simulations and the Acceleration of Cosmic Rays. *ApJ*, 689:1063–1077, 2008.
- S. W. Skillman, H. Xu, E. J. Hallman, B. W. O’Shea, J. O. Burns, H. Li, D. C. Collins, and M. L. Norman. Cosmological Magnetohydrodynamic Simulations of Galaxy Cluster Radio Relics: Insights and Warnings for Observations. *ApJ*, 765: 21, 2013.

- G. F. Snyder, P. Torrey, J. M. Lotz, S. Genel, C. K. McBride, M. Vogelsberger, A. Pillepich, D. Nelson, L. V. Sales, D. Sijacki, L. Hernquist, and V. Springel. Galaxy morphology and star formation in the Illustris Simulation at $z = 0$. *MNRAS*, 454:1886–1908, 2015.
- G. A. Sod. A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws. *J. Comput. Phys.*, 27(1):1 – 31, 1978.
- M. Sonntag and C.-D. Munz. Shock capturing for discontinuous galerkin methods using finite volume subcells. In J. Fuhrmann, M. Ohlberger, and C. Rohde, editors, *Finite Volumes for Complex Applications VII-Elliptic, Parabolic and Hyperbolic Problems*, volume 78 of *Springer Proceedings in Mathematics & Statistics*, pages 945–953. Springer International Publishing, 2014. ISBN 978-3-319-05590-9.
- M. Sparre, C. C. Hayward, V. Springel, M. Vogelsberger, S. Genel, P. Torrey, D. Nelson, D. Sijacki, and L. Hernquist. The star formation main sequence and stellar mass assembly of galaxies in the Illustris simulation. *MNRAS*, 447:3548–3563, 2015.
- R. J. Spiteri and S. J. Ruuth. A new class of optimal high-order strong-stability-preserving time discretization methods. *SIAM J. Numer. Anal.*, 40(2):469–491, 2002.
- V. Springel. The cosmological simulation code GADGET-2. *MNRAS*, 364:1105–1134, 2005.
- V. Springel. E pur si muove: Galilean-invariant cosmological hydrodynamical simulations on a moving mesh. *MNRAS*, 401:791–851, 2010.
- V. Springel and L. Hernquist. Cosmological smoothed particle hydrodynamics simulations: a hybrid multiphase model for star formation. *MNRAS*, 339:289–311, 2003.
- V. Springel, T. Di Matteo, and L. Hernquist. Modelling feedback from stars and black holes in galaxy mergers. *MNRAS*, 361:776–794, 2005.
- V. Springel. Fundamentals of simulation methods. *Lecture notes, University of Heidelberg*, 2016.
- D. Steinhauser, S. Schindler, and V. Springel. Simulations of ram-pressure stripping in galaxy-cluster interactions. *ArXiv e-prints*, 2016.
- J. M. Stone, T. A. Gardiner, P. Teuben, J. F. Hawley, and J. B. Simon. Athena: A New Code for Astrophysical MHD. *ApJS*, 178:137–177, 2008.
- R. A. Sunyaev and I. B. Zeldovich. Microwave background radiation as a probe of the contemporary structure and history of the universe. *ARA&A*, 18:537–560, 1980.

Bibliography

- J. Suresh, S. Bird, M. Vogelsberger, S. Genel, P. Torrey, D. Sijacki, V. Springel, and L. Hernquist. The impact of galactic feedback on the circumgalactic medium. *MNRAS*, 448:895–909, 2015a.
- J. Suresh, K. H. R. Rubin, R. Kannan, J. K. Werk, L. Hernquist, and M. Vogelsberger. On the OVI Abundance in the Circumgalactic Medium of Low-Redshift Galaxies. *ArXiv e-prints*, 2015b.
- S. A. Teukolsky. Formulation of discontinuous Galerkin methods for relativistic astrophysics. *Journal of Computational Physics*, 312:333–356, 2016.
- R. Teyssier. Cosmological hydrodynamics with adaptive mesh refinement. A new high resolution code called RAMSES. *A&A*, 385:337–364, 2002.
- E. Toro. *Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction*. Springer, 2009. ISBN 9783540498346.
- P. Torrey, M. Vogelsberger, S. Genel, D. Sijacki, V. Springel, and L. Hernquist. A model for cosmological simulations of galaxy formation physics: multi-epoch validation. *MNRAS*, 438:1985–2004, 2014.
- P. Torrey, S. Wellons, F. Machado, B. Griffen, D. Nelson, V. Rodriguez-Gomez, R. McKinnon, A. Pillepich, C.-P. Ma, M. Vogelsberger, V. Springel, and L. Hernquist. An analysis of the evolving comoving number density of galaxies in hydrodynamical simulations. *MNRAS*, 454:2770–2786, 2015.
- F. van de Voort and J. Schaye. Properties of gas in and around galaxy haloes. *MNRAS*, 423:2991–3010, 2012.
- R. J. van Weeren, H. J. A. Röttgering, J. Bagchi, S. Raychaudhury, H. T. Intema, F. Miniati, T. A. Enßlin, M. Markevitch, and T. Erben. Radio observations of ZwCl 2341.1+0000: a double radio relic cluster. *A&A*, 506:1083–1094, 2009.
- R. J. van Weeren, H. J. A. Röttgering, M. Brüggen, and M. Hoeft. Particle Acceleration on Megaparsec Scales in a Merging Galaxy Cluster. *Science*, 330:347–, 2010.
- F. Vazza, G. Brunetti, and C. Gheller. Shock waves in Eulerian cosmological simulations: main properties and acceleration of cosmic rays. *MNRAS*, 395:1333–1354, 2009a.
- F. Vazza, G. Brunetti, A. Kritsuk, R. Wagner, C. Gheller, and M. Norman. Turbulent motions and shocks waves in galaxy clusters simulated with adaptive mesh refinement. *A&A*, 504:33–43, 2009b.
- F. Vazza, G. Brunetti, C. Gheller, and R. Brunino. Massive and refined: A sample of large galaxy clusters simulated at high resolution. I: Thermal gas and properties of shock waves. *New A*, 15:695–711, 2010.

- F. Vazza, K. Dolag, D. Ryu, G. Brunetti, C. Gheller, H. Kang, and C. Pfrommer. A comparison of cosmological codes: properties of thermal gas and shock waves in large-scale structures. *MNRAS*, 418:960–985, 2011.
- F. Vazza, M. Brüggen, C. Gheller, and G. Brunetti. Modelling injection and feedback of cosmic rays in grid-based cosmological simulations: effects on cluster outskirts. *MNRAS*, 421:3375–3398, 2012.
- F. Vazza, C. Ferrari, A. Bonafede, M. Brüggen, C. Gheller, R. Braun, and S. Brown. Filaments of the radio cosmic web: opportunities and challenges for SKA. *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, art. 97, 2015.
- A. Vladimirov, D. C. Ellison, and A. Bykov. Nonlinear Diffusive Shock Acceleration with Magnetic Field Amplification. *ApJ*, 652:1246–1258, 2006.
- M. Vogelsberger, D. Sijacki, D. Kereš, V. Springel, and L. Hernquist. Moving mesh cosmology: numerical techniques and global statistics. *MNRAS*, 425:3024–3057, 2012.
- M. Vogelsberger, S. Genel, D. Sijacki, P. Torrey, V. Springel, and L. Hernquist. A model for cosmological simulations of galaxy formation physics. *MNRAS*, 436:3031–3067, 2013.
- M. Vogelsberger, S. Genel, D. Sijacki, P. Torrey, V. Springel, and L. Hernquist. Erratum: A model for cosmological simulations of galaxy formation physics. *MNRAS*, 438:3607–3607, 2014a.
- M. Vogelsberger, S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G. Snyder, S. Bird, D. Nelson, and L. Hernquist. Properties of galaxies reproduced by a hydrodynamic simulation. *Nature*, 509:177–182, 2014b.
- M. Vogelsberger, S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G. Snyder, D. Nelson, and L. Hernquist. Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe. *MNRAS*, 444:1518–1547, 2014c.
- J. W. Wadsley, J. Stadel, and T. Quinn. Gasoline: a flexible, parallel implementation of TreeSPH. *New A*, 9:137–158, 2004.
- D. G. Wentzel. Cosmic-ray propagation in the Galaxy - Collective effects. *ARA&A*, 12:71–96, 1974.
- S. D. M. White and C. S. Frenk. Galaxy formation through hierarchical clustering. *ApJ*, 379:52–79, 1991.
- S. D. M. White and M. J. Rees. Core condensation in heavy halos - A two-stage theory for galaxy formation and clustering. *MNRAS*, 183:341–358, 1978.
- G. Xu. A New Parallel N-Body Gravity Solver: TPM. *ApJS*, 98:355, 1995.

Bibliography

- A. Yalinewich, E. Steinberg, and R. Sari. RICH: Open-source Hydrodynamic Simulation on a Moving Voronoi Mesh. *ApJS*, 216:35, 2015.
- H. C. Yee, N. D. Sandham, and M. J. Djomehri. Low-Dissipative High-Order Shock-Capturing Methods Using Characteristic-Based Filters. *Journal of Computational Physics*, 150:199–238, 1999.
- O. Zanotti, F. Fambri, and M. Dumbser. Solving the relativistic magnetohydrodynamics equations with ADER discontinuous Galerkin methods, a posteriori subcell limiting and adaptive mesh refinement. *MNRAS*, 452:3010–3029, 2015.
- X. Zhang and C.-W. Shu. On positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes. *Journal of Computational Physics*, 229:8918–8934, 2010.
- X. Zhang. Maximum-principle-satisfying and positivity-preserving high order schemes for conservation laws. *Dissertation, U. of Science and Technology of China*, 2006.

Acknowledgments

I am very grateful to many people for making this work possible and supporting me during the last three years.

First of all, I would like to warmly thank my supervisor Prof. Volker Springel for accepting me as a PhD student and providing excellent support for my research. His door was always open and he took time whenever I needed his help. Moreover, I enjoyed many insightful discussions and had a great time working with him.

At this point I would also like to thank Prof. Ralf Klessen, who kindly agreed to referee this thesis.

Thanks are also due to Prof. Lars Hernquist, Prof. Mark Vogelsberger, Prof. Christian Klingenberg, and their research groups for a fruitful collaboration and lively scientific exchange. In particular, I want to thank the former two for inviting me to the US and hosting me during my stay.

I am grateful for the support by the Klaus Tschira foundation and the International Max Planck Research School for Astronomy and Cosmic Physics at the University of Heidelberg, and I would like to thank the Heidelberg Institute for Theoretical Studies and its employees for providing an outstanding work environment.

I had a great time in the theoretical astrophysics group at HITS and I wish to say thank you to all members, especially to my officemates Christoph Pfrommer, Rüdiger Pakmor, Andreas Bauer, Christian Arnold, Rainer Weinberger, Jolanta Zjupa, as well as Federico Marinacci and Martin Sparre.

Finally, I want to express my deepest gratitude to my parents Sigmar and Sonja. Without your love and support this thesis would not have been possible.

Eigenständigkeitserklärung

Hiermit versichere ich, die vorliegende Arbeit selbstständig verfasst zu haben und keine Quellen oder Hilfsmittel außer den angegebenen verwendet zu haben.

Heidelberg, den 9. Mai 2016,

Kevin Schaal