

Numerical overcooling in shocks

Peter Creasey,^{1*} Tom Theuns,^{1,2} Richard G. Bower¹ and Cedric G. Lacey¹

¹*Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE*

²*Department of Physics, University of Antwerp, Campus Groenenborger, Groenenborgerlaan 171, B-2020 Antwerp, Belgium*

Accepted 2011 April 29. Received 2011 April 28; in original form 2010 November 11

ABSTRACT

We present a study of cooling in radiative shocks simulated with smoothed particle hydrodynamics and adaptive mesh refinement codes. We obtain a similarity solution for a shock-tube problem in the presence of radiative cooling, and test how well the solution is reproduced in GADGET and FLASH. Shock broadening governed by the details of the numerical scheme (artificial viscosity or Riemann solvers) leads to potentially significant overcooling in both codes. We interpret our findings in terms of a resolution criterion, and apply it to realistic simulations of cosmological accretion shocks on to galaxy haloes, cold accretion and thermal feedback from supernovae or active galactic nuclei (AGN). To avoid numerical overcooling of accretion shocks on to haloes that should develop a hot corona a particle or cell mass resolution of $10^6 M_{\odot}$ is required, which is within reach of current state-of-the-art simulations. At this mass resolution, thermal feedback in the interstellar medium of a galaxy requires temperatures of supernova- or AGN-driven bubbles to be in excess of 10^7 K at densities of $n_{\text{H}} = 1.0 \text{ cm}^{-3}$, in order to avoid spurious suppression of the feedback by numerical overcooling.

Key words: hydrodynamics – shock waves – methods: numerical – galaxies: formation – galaxies: ISM.

1 INTRODUCTION

Radiative cooling and shocks are two important ingredients in galaxy formation theory (White & Rees 1978). Whilst most codes used in astrophysics have facilities for handling both of these, the scales at which these operate and interact are challenging. We will start with a short tour of the processes and of the numerical codes we will use to simulate them. We describe a one-dimensional model problem of a radiatively cooling shock with an analytic solution which we model with two popular codes in astrophysical simulations, FLASH (Fryxell et al. 2000), an adaptive mesh refinement (AMR) code, and GADGET (Springel 2005), a smoothed particle hydrodynamics (SPH) code (Gingold & Monaghan 1977; Lucy 1977). The results of our simulations give appropriate criteria with which we can analyse the efficacy of our numerical schemes in a wide variety of astrophysical environments. We also investigate the mitigating factors such as the ratio of the cooling to dynamical times, which may enable a simulation to give approximately correct results when otherwise we would consider there to be insufficient resolution.

There has been considerable discussion on the treatment of discontinuities in SPH (Price 2008; Read, Hayfield & Agertz 2010), motivated by problems illustrated by Agertz et al. (2007), but the issues highlighted in this paper are of a different nature. Those pa-

pers focus on spurious forces introduced by the SPH scheme, whilst we focus on evaluating the errors introduced as we approach the resolution limit, which are to some extent unavoidable.

1.1 Astrophysical shocks

The science of astrophysics is an ideal domain for the investigation of shock fronts on a variety of scales. Stellar winds form shocks as they push in to the interstellar medium (ISM). On larger scales supernovae (SNe) form very high Mach number shocks as they plough into the surrounding gas and form remnants. On larger scales still, galactic winds, driven by starbursts and active galactic nuclei (AGN), shock against the intergalactic medium (IGM). In the context of galaxy formation we can also consider accretion shocks, where gravitationally accelerated infalling gas shocks to form a hot corona in the dark matter potential well.

Of particular interest to us in this paper are radiatively cooling shocks. To an extent, all the aforementioned shocks have radiative cooling; however, the cosmological accretion shocks and SNe are particularly topical. In galaxy formation simulations the SNe at early times form remnants well below the resolution of current simulations and need to be modelled with subgrid physics (see Kay et al. 2002 for a review of feedback methods). The cooling of the hot gas causes a transition from a thermally driven to a momentum-driven phase, losing a significant fraction of the SNe energy. A similar transition is thought to occur in the thermal to momentum transition of winds powered by an AGN (Booth & Schaye 2009).

*E-mail: p.e.creasey@durham.ac.uk

Cooling in accretion shocks may affect the fuelling of star formation in the host galaxy. If the gas is shocked to too high a temperature it will not cool over a Hubble time, preventing star formation (though non-spherical geometries may allow the gas to compress first and thus cool faster; see e.g. Birnboim & Dekel 2003). In a cosmological simulation, however, the resolution around these cooling regions may be so coarse as to resolve these cooling regions with only a few particles. In this paper we intend to probe the effect of limited numerical resolution in these cases, and how these may affect the outcome of the simulation.

1.2 Physical shock fronts

Before we concentrate on the numerical aspect of cooling in shocks, we begin by briefly considering the processes that occur in a real physical shock front. A shock front is a region where one of the usually conserved fluid properties, entropy, is allowed to change. It is worth considering why such a property is otherwise treated as a constant, and why shocks are a special case.

In the kinetic theory of gases, a gas is described as a large number of particles (e.g. atoms, molecules, ions) in constant random motion. The frequency of collisions defines a time-scale and also a typical length between collisions, the mean free path. If all processes acting on the gas happen on time-scales much greater than the time between collisions, then the classical theory of adiabatics tells us that there will be another conserved property which, for an ideal gas, is p/ρ^γ . Here, p , ρ and γ are the pressure, density and adiabatic index, respectively. This property is a function of the entropy, and in astrophysics is often used as a proxy.

It is worth recalling that processes which change the fluid entropy (e.g. shocks, radiative absorption, thermal conduction) will occur on time-scales of the order of, or shorter than, the period between collisions (or over lengths of the order of the collision length). Mechanisms which heat the gas on slower time-scales will be *adiabatic* processes, and will alter the thermal energy with only very small increases in entropy.¹

Now we come to shock fronts. A canonical example of a shock front would be a one-dimensional system where the upstream fluid travels supersonically with respect to the downwind fluid (i.e. faster than the thermal velocities of the particles), until it reaches the shock, where the majority of its mechanical energy (the bulk motion of the particles) is converted into thermal energy. This happens because the pairs of particles that collide can have very different velocities: particles in the shock front change their energy on a time-scale of the order of the collision time between a pair of upwind and downwind particles, which is much shorter than that between two downwind, or two upwind, particles. From this description we immediately see that physical shocks must occur over length-scales of the order of the mean free path, which is usually much smaller than other physical length-scales in the problem.

The mean free path (Δx) depends upon the number density of particles (n) and their collisional cross-section (σ), as

$$\Delta x = \frac{1}{n\sigma}. \quad (1)$$

¹ One can of course construct systems in which the time-scale of interest is long enough such that viscosity, thermal diffusion, etc. dominate the large scales too. These problems, however, have low Reynolds and Péclet numbers, respectively, and are the exception rather than the norm in computational astrophysics.

In the case of a partially or fully ionized gas, particles may interact on a shorter length-scale (Zel'Dovich & Raizer 1967), that of the plasma skin depth/plasma oscillation length:

$$\Delta x = c \left(\frac{4\pi n_e e^2}{m_e} \right)^{-1/2} \approx 10^6 \left(\frac{n_e}{1 \text{ cm}^{-3}} \right)^{-1/2} \text{ cm}. \quad (2)$$

Since the particles are not interacting via Coulomb collisions this is known as a ‘collisionless shock’; the mechanism of interaction is the plasma oscillation (coupling together charged particles). Care must be taken, however, as the post-shock gas may be out of thermal and ionizational equilibrium for the problem in question (something that would not be a concern if the collision length is small), making these cases challenging to simulate.

The trapping of relativistic ions between magnetic fields in the upstream and downstream phases and subsequent acceleration is also believed to be the origin of the power-law spectrum of high-energy cosmic rays, a Fermi acceleration process.

Finally, we should complete this discussion by mentioning turbulence as a source of entropy. In general the bulk oscillations of fluids will occur on scales much larger than the mean free path and is thus unable to change the entropy. Transfer of spectral energy to shorter wavelengths, however, implies that eventually bulk oscillations reach the scale of the mean free path and will be dissipated into thermal energy (Kolmogorov 1941).

1.3 Shocks in simulations, artificial viscosity

Now let us consider shocks in simulations. Almost exclusively, the resolution of simulations will be much coarser than a physical shock width. This is not necessarily a problem, however, as the bulk properties of the post-shock gas may be deduced from the conservation of energy and momentum, and the assumption that the shock process does not produce oscillations on scales much larger than the mean free path.

In this paper we will contrast two schemes for numerical hydrodynamics that are popular in cosmology: SPH and AMR. SPH (Gingold & Monaghan 1977; Lucy 1977; see Monaghan 2005 and Springel 2005 for recent reviews) is a (pseudo-) Lagrangian scheme in which the fluid is represented by a set of particles that move along with the flow. In this paper we will illustrate the behaviour of SPH using the `GADGET` SPH implementation of Springel (2005). AMR follows how fluid flows across a (stationary) computational mesh, whose cell size may be locally ‘refined’ or ‘de-refined’ based on some criterion. In this paper we use the `FLASH` code, a block-structured AMR implementation by Fryxell et al. (2000).

The physical process of kinetic energy dissipation by particle collisions is represented in the continuum approximation by a viscous term in the Navier–Stokes equations,

$$\frac{\partial}{\partial t}(\rho v_i) + \frac{\partial}{\partial x_j}(\rho v_i v_j) = \frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_j} \sigma_{ij}, \quad (3)$$

where

$$\sigma_{ij} = \eta \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} - \frac{2}{3} \frac{\partial v_k}{\partial x_k} \delta_{ij} \right) + \zeta \frac{\partial v_k}{\partial x_k} \delta_{ij} \quad (4)$$

is the viscous stress tensor, and η and ζ are known as the shear and bulk viscosity coefficients, respectively. These coefficients can be measured for real fluids; however, in most astrophysical flows they are so small that the viscous term is insignificant outside of shocks (i.e. the flows have high Reynolds number).

The variant of SPH used in this paper handles shocks with a prescription known as artificial viscosity (although Godunov-type

methods for SPH also exist; see Inutsuka 2002). Artificial viscosity was originally developed for grid codes (von Neumann & Richtmyer 1950), and uses the bulk viscosity term in equation (3), however, with the coefficient raised by several orders of magnitude. These larger values prevent the shocks generating large unphysical oscillations due to the coarseness of the sampling (see the numerical stability criterion of Friedrichs & Lax 1971). In SPH they also fulfil a second role of preventing particle interpenetration (see Bate 1995 for a thorough discussion). A number of artificial viscosity prescriptions are in use, the most common being that of Monaghan and Balsara (Balsara 1995), a Lax–Friedrichs style viscosity that is turned on for compressing flows. The implementation in our version of GADGET is based on signal velocities (Monaghan 1997).

In mesh codes shocks can be treated with artificial viscosity but more commonly a conservative Riemann solver (based upon Godunov’s scheme, Godunov & Ryabenki 1964) is used. Riemann solvers (see e.g. the HLL solver, Harten, Lax & van Leer 1983) give exact solutions in one-dimensional or planar shock problems with homogeneous pre- and post-shock fluids, but are somewhat diffusive in other cases. They are still the preferred method for grid codes, however, and the default used in FLASH is a directionally split Riemann solver (Colella & Woodward 1984). Oscillations near the discontinuities are controlled with a monotonicity constraint.

1.4 Radiative cooling

Radiative cooling is an essential ingredient in galaxy formation as it is the process which allows the baryons in dark matter haloes to dissipate thermal energy and thus collapse to form galaxies. Multiple cooling mechanisms are important in the astrophysical domain; however, in this paper we will primarily be interested in collisional line cooling and, at higher temperatures, thermal bremsstrahlung. The evolution of the specific thermal energy, u , due to cooling can be written as

$$\rho \dot{u}|_{\Lambda} = -\Lambda(T; Z)n^2, \quad (5)$$

where ρ is the density, T the temperature, Z the metallicity and n the particle number density (for brevity we will subsequently refer to the radiative component of the specific cooling rate $\dot{u}|_{\Lambda}$ as \dot{u}_{Λ}). When baryon–photon interactions with the cosmic microwave background (CMB) and an ionizing background are important, we have followed the prescriptions of Wiersma, Schaye & Smith (2009a; see also Fig. 7 below).

The implementation of cooling in our versions of GADGET and FLASH is performed by an adaptive time-step integration over each cell/particle. The effects of cooling are included in the hydrodynamic solver by *operator splitting*, i.e. the separation of the two processes A (radiative cooling) and B (shock heating) into individual steps,

$$\dot{X} = (A + B)X \quad (6)$$

$$X_{t+\Delta t} - X_t \approx A(\Delta t)B(\Delta t)X, \quad (7)$$

where the errors on the latter term of the order of the time-step Δt depend upon the commutator $[A, B]$. Since the physical process of shock heating should occur over a much shorter time-scale than that of radiative cooling, we can justify this being zero. The numerical implementation of shock heating will of course take a longer time-scale and thus would interact with the cooling if operator splitting was not introduced; however, there is no physical motivation to prefer such a scheme in this case.

2 RADIATIVELY COOLING SHOCKS, A MODEL PROBLEM

A typical test problem in numerical hydrodynamics is that of the formation of a one-dimensional shock in a ‘test tube’. In one form of this problem a tube is initialized with gas of constant polytropic index γ , the left and right halves converging with opposing velocities v_0 and $-v_0$. For a sufficiently high velocity v_0 a shock will form, creating a downstream region with higher temperature, pressure and density. This problem has a similarity solution for constant γ . In our set-up the gas is allowed to cool radiatively, and the downstream region can then cool to form a dense post-shock region. The initial conditions are thus

$$\rho(x, t_0) = \rho_0 \quad (8)$$

$$p(x, t_0) = p_0 \quad (9)$$

$$T(x, t_0) = T_0 \quad (10)$$

$$v(x, t_0) = \begin{cases} v_0, & x < 0 \\ -v_0, & x > 0. \end{cases} \quad (11)$$

We note that the symmetry of this problem makes it equivalent to the wall shock (where there is an immovable boundary at $x = 0$; Monaghan 1997). In order to minimize the amount of modification in our SPH code, we chose to set up the symmetric problem rather than implement an immovable boundary.

2.1 Similarity solution for a radiative one-dimensional shock

If the temperature dependence of the cooling rate is sufficiently simple, then this one-dimensional shock problem has a similarity solution even in the presence of cooling. Such is the case for a cooling function which is a piecewise linear function of the temperature, such that the rate of radiative cooling, $\dot{u}|_{\Lambda}$, of the specific energy, u , is given by a ‘cooling spike’:

$$\rho \dot{u}|_{\Lambda} = \begin{cases} 0, & T < T_0 \\ -\Lambda n^2(T - T_0)/T_0, & T_0 \leq T \leq \frac{1}{2}(T_1 + T_0) \\ -\Lambda n^2(T_1 - T)/T_0, & \frac{1}{2}(T_1 + T_0) \leq T \leq T_1 \\ 0, & T_1 \leq T, \end{cases} \quad (12)$$

where Λ is a positive constant and cooling is maximum at $T = \frac{1}{2}(T_1 + T_0)$. For simplicity, in all simulations discussed below we initialize the temperature to T_0 (where the cooling vanishes), so the initial gas is not cooling.

The gas is chosen to be pure atomic hydrogen, i.e.

$$\gamma = \frac{5}{3} \quad (13)$$

$$\rho = m_p n \quad (14)$$

$$p = nk_B T. \quad (15)$$

For comparison the reader should see the simulations of Hutchings & Thomas (2000) who used a more realistic astrophysical cooling function, at the expense of not having an analytic solution. For the cooling post-shock region we find analytic solutions in a

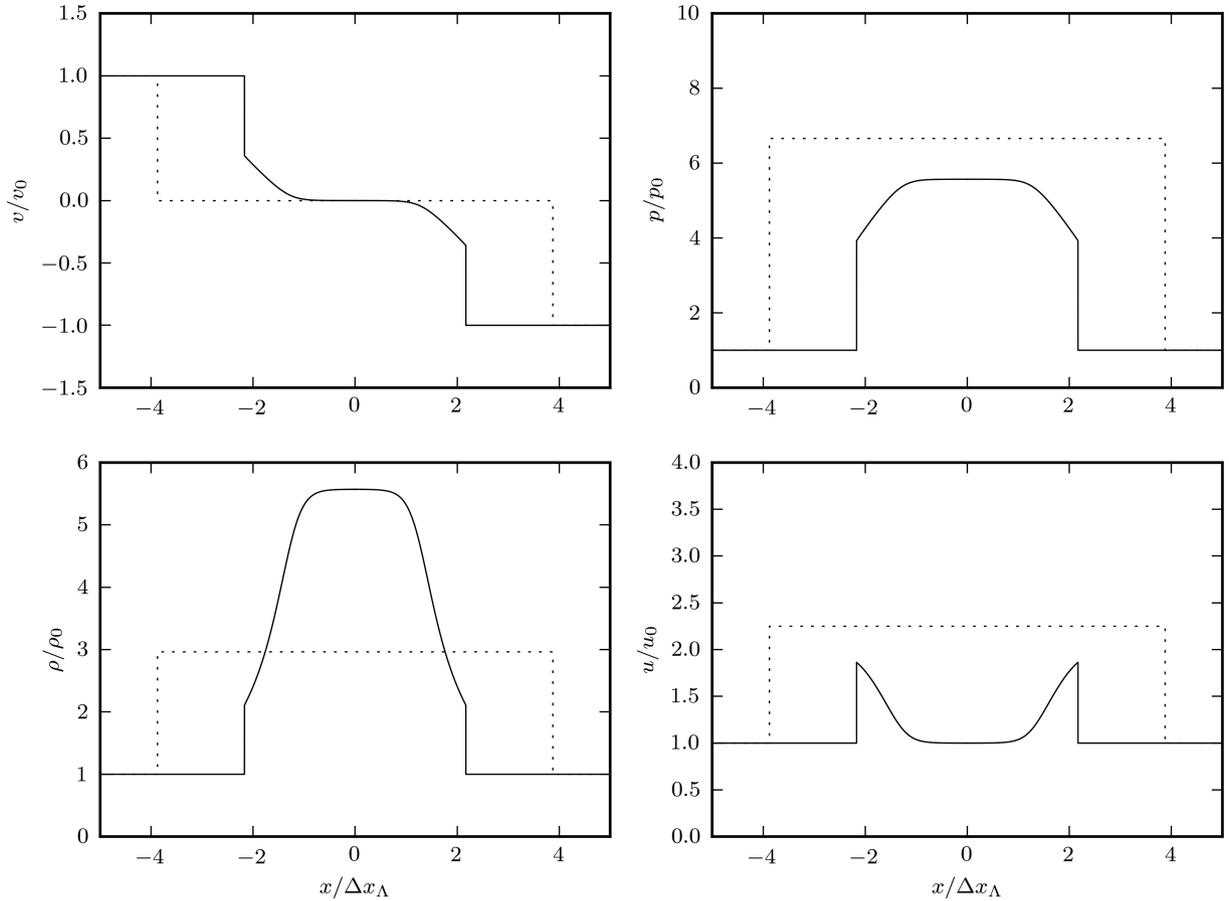


Figure 1. *Solid lines* represent the analytic solution for the colliding gas problem discussed in Section 2 when cooling is included. Incoming gas from the left and right shocks and then compresses and cools to form a cold dense region in the centre. For the example shown, the Mach number of the upstream gas is 1.5 with respect to the cold, central gas and the time is $5.1 \Delta x_\Lambda / c_0$ (see equation 17). For comparison, *dashed lines* show the solution without cooling at the corresponding time. At early times (i.e. $t \lesssim \Delta x_\Lambda / c_0$) the cooling profile is not of given form, as it has not had sufficient time to reach a stationary state (details of the similarity solution for cooling through a shock can be found in Appendix A1).

similarity variable $\lambda \equiv \rho/\rho_0$ of the form (see Appendix A1 for details)

$$T/T_0 = [(a+1)\lambda^{-1} - a\lambda^{-2}]$$

$$x - x_0 = \frac{-v_s k_B T_0}{(\gamma-1)\Delta n_0} \left[\frac{\gamma-a}{a-1} \log(1-\lambda^{-1}) + \frac{1-a\gamma}{(a-1)a^2} \log(1-a\lambda^{-1}) - \frac{a+1}{a} \lambda^{-1} - \frac{\gamma+1}{2} \lambda^{-2} \right]$$

$$a \equiv \frac{\rho_0 v_s^2}{p_0}.$$

The value of the shock velocity v_s and the final density in the cold, post-shock region ρ_0 can be found by imposing conservation of mass and momentum (see Appendix A2). The solution is shown in Fig. 1.

2.2 Shock stability

Chevalier & Imamura (1982) find that positive increasing linear cooling functions produce stable shocks. Applying this to the cooling function in equation (12) we see that we have stable shocks provided the post-shock temperature $T_s < \frac{1}{2}(T_1 + T_0)$ or $T_s > T_1$, which is the case for all the shocks we study later (we define

the shock temperature T_s as the temperature immediately after the shock, which is computed in Appendix A2). If the gas is shocked to $\frac{1}{2}(T_1 + T_0) < T_s < T_1$ then the shock may be unstable as the cooling function has a negative slope, $\partial_T(-\rho\dot{u}_\Lambda) < 0$. Intuitively this can be understood in terms of the length of the cooling region: if the length increases the shock velocity will be higher, causing the post-shock gas to be hotter, which increases the cooling time, which feeds back into a longer cooling region.

2.3 Numerical solution

2.3.1 Initial conditions

The similarity solution is described with two (dimensionless) parameters. The first is the ratio of the upper to the lower temperature in the cooling spike, which we will set to 20, i.e. $T_1 = 20T_0$. This is motivated by the temperature dependence of the radiative cooling function of an astrophysical plasma (see also Fig. 7), where individual elements contribute significantly to the cooling over approximately a 1-dex range in temperature.

The second parameter is the Mach number of the shock, which we will quote in the rest frame of the problem (rather than the rest frame of the post-shock gas, for example),

$$\mathcal{M} \equiv \frac{v_0}{c_0}, \quad (16)$$

where $c_0 \equiv (\gamma p_0/\rho_0)^{1/2}$ is the upstream sound speed. Our tests are performed at $\mathcal{M} = 4.70$ and 6.04 . The former has been chosen such that the shock reaches a temperature somewhat below the maximum of the cooling function, $(T_1 + T_0)/2$ (where the shock will be stable), and the latter such that the shock reaches a temperature somewhat above T_1 (where there is no cooling).

We plot positions in units of the cooling length,

$$\Delta x_\Lambda \equiv \frac{k_B T_0 c_0}{\Lambda n_0}. \quad (17)$$

Similarly, we express times in units of $\Delta x_\Lambda/c_0$. As observed in Monaghan (1997), numerical schemes (including both SPH and AMR) usually produce a transient unphysical thermal bump at $t = 0$ when there is no post-shock region. To avoid contamination by this transient, we run our simulation for a time $14.2\Delta x_\Lambda/c_0$ and $7.1\Delta x_\Lambda/c_0$ for the $\mathcal{M} = 4.7$ and 6.04 shocks, respectively (i.e. we simulate for several sound crossing times of the cooling region, to make sure it is in a stationary state).

For our SPH simulations we set up a long box, periodic along all boundaries. The particle mass is chosen to be

$$m_{\text{SPH}} = \rho_0(0.3\Delta x_\Lambda)^3, \quad (18)$$

(i.e. a mean interparticle spacing of $0.3\Delta x_\Lambda$). We note that this set-up creates a rarefaction wave that propagates inwards from the far edges of the computational volume (due to the discontinuity on this boundary), and thus we need a box long enough such that these rarefactions do not reach our domain of interest in the simulation time. The particles were set up with a ‘glass’ distribution (White 1994) to minimize relaxation effects in the pre-shock fluid (the SPH kernel allows a cubic lattice arrangement of particles to slightly reduce its density, and hence release some thermal energy, by relaxing to a glass-like state). We also raised the level of the bulk artificial viscosity constant, α , to 3 (from 1; see Springel 2005 for a complete description of the artificial viscosity prescription). We found this to be necessary to prevent ringing and the appearance of large scatter in the entropy of SPH particles in the post-shock region (see also Abel 2011).

For the AMR simulations we again set up a long box with cell spacing $0.3\Delta x_\Lambda$, with periodic boundaries in the y and z directions and inflowing gas along the (long) x -axis. No refinement was allowed, effectively making this a uniform Eulerian mesh.

We considered allowing an alternative refinement criterion; however, the standard FLASH refinement schemes will refine a shock to the maximum allowed level (since it contains a discontinuity), reducing it to the uniform mesh case. We refer the reader to the dashed lines in Figs 2 and 4 to compare resolutions.

We note that the use of inflowing boundary conditions in FLASH allowed us to avoid the rarefaction waves we created in SPH, and thus we could use a much shorter box (by a factor of 10). To set the scene we begin by looking at shocks in the absence of cooling.

2.3.2 Test without cooling

The test problems in the absence of cooling are compared in the upper panels of Figs 2 and 4 ($\mathcal{M} = 4.7$ and 6.04 , respectively). Provided we use the higher than usual value of the artificial viscosity ($\alpha = 3$) in GADGET, both the SPH and AMR codes handle this shock well (as expected), with the shock smeared out over a few times the resolution length h in SPH, and a few cells in FLASH. At GADGET’s default value for the artificial viscosity ($\alpha = 1$) we find that this is too high a Mach number shock to be handled (we do, however, return to the original value when we study the lower Mach number shocks in

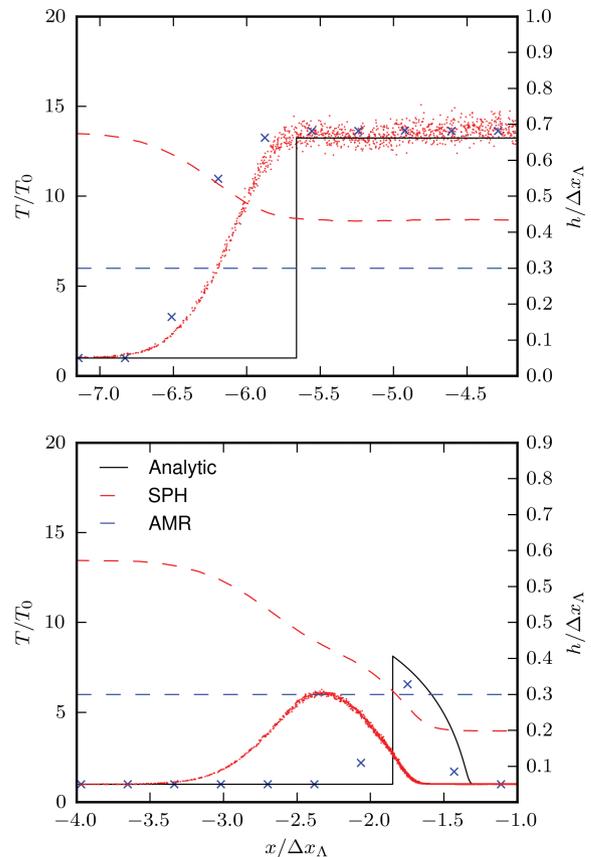


Figure 2. Upper panel plots the temperature in a Mach $\mathcal{M} = 4.7$ shock without cooling. Each blue cross represents a column of FLASH cells (the tube is orthogonal to the mesh), each red point represents a GADGET particle, the black line is the analytic solution. Red dashes denote the smoothing lengths of the GADGET particles, blue dashes the separation of FLASH cells (right axes). Incoming gas from the left (and right, not shown) collides to form a homogeneous hot, rarefied region in the centre. As expected, both codes reproduce the correct profile relatively well. The shock is seen to be spread over several cells (FLASH) or smoothing lengths (GADGET). Lower panel as for top panel but including cooling. The analytical solution shows that the gas shocks to a lower temperature (due to the smaller difference between the incoming gas velocity and the shock velocity), followed by a ‘cooling tail’ in the post-shock region. When simulated using GADGET, SPH particles shock in several steps before reaching their maximum temperature. As they do so, particles cool to some extent in the smoothed shock and hence reach a lower maximum temperature than the analytical solution (the SPH shock is also offset to the left of the analytic shock). In the FLASH run gas gets shocked to higher temperatures, closer to the analytical solution. Note that as the gas gets compressed the downstream SPH smoothing length is smaller than the FLASH cell size.

Section 3). In Fig. 3 we tested both altering the value of the artificial viscosity and adjusting the maximum time-step (between GADGET’s default adaptive scheme and a global minimum Courant step applied to all particles). The higher value of artificial viscosity was found to significantly reduce the scatter in the post-shock thermal energies, superior to a reduction in the global time-step. We would, however, expect that at very high Mach number shocks a more conservative time-step would be required.

2.3.3 Test with cooling

First let us consider the case of cooling for the $\mathcal{M} = 4.7$ shock. This should result in a gas temperature of less than $(T_1 + T_0)/2$,

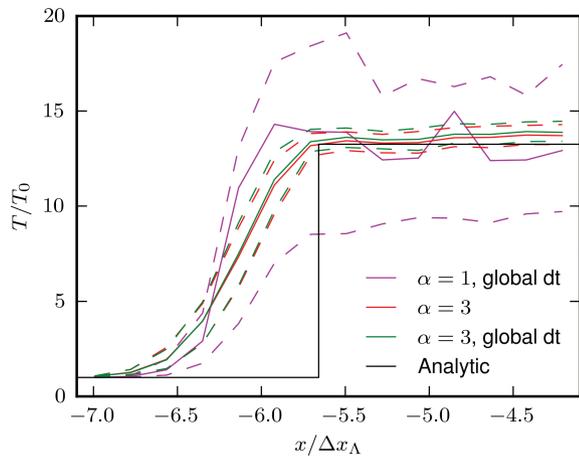


Figure 3. Comparison of the effects of altering the viscosity and time-step on the shock from the upper panel of Fig. 2. Red lines show the higher viscosity ($\alpha = 3$) scheme with adaptive time-steps, green lines the same viscosity but a global minimum time-step (set to the minimum Courant step of all particles) and purple lines the global minimum time-step but with the default viscosity ($\alpha = 1$), black line the analytic solution. Dashed lines show the 10th and 90th percentiles.

i.e. we are on the left side of the cooling spike. The initial collision of gas can result in higher temperatures and follows an evolution for which we have no analytic solution, before settling down to our stationary case.

In Fig. 2 (lower panel) we see the FLASH and GADGET representations of these shocks. Both codes reach a maximum temperature which is lower than that of the similarity solution. In SPH we attribute this to ‘pre-shocking’, i.e. particles will shock in several stages and cool as they are being shocked. In FLASH we attribute this to the cooling operation being applied after the hydrodynamics in a time-step, such that we do not record the post-shock temperature. Neither GADGET nor FLASH has the resolution to reproduce the cooling tail particularly well here, although the final cold state is achieved in both cases.

For our second cooling test we look at a more extreme case, $\mathcal{M} = 6.04$. This shocks the gas up to a temperature $T > T_1$ from which it cannot cool; hence, the analytic solution is the same as for a shock without cooling. In Fig. 4 we show the left-hand side of the shocked regions. Here the FLASH simulation reproduces the analytical result very well, but the GADGET simulation suffers from much more severe numerical overcooling through the pre-shock region, which prevents the gas from reaching the temperature from which it is unable to cool [due to our choice of $\Lambda(T)$]. As a result we see pile-up of high-density cold gas around $x = 0$, and the shocked region is left far behind that of the FLASH run.² As a result of this overcooling the SPH simulation fails to form any hot gas at all. We note, however, that this is a general problem and not specific to either GADGET or SPH.

2.4 Convergence study for GADGET results

As it stands, we can be confident that the results we have just given for SPH are not converged, as they have failed to reach our stable

²Note that if cooling is prevented, the shock speed will be much higher relative to the rest frame. This is easily understood in terms of conservation of mass, the gas is shocked to a lower density and a much larger region is required to contain it.

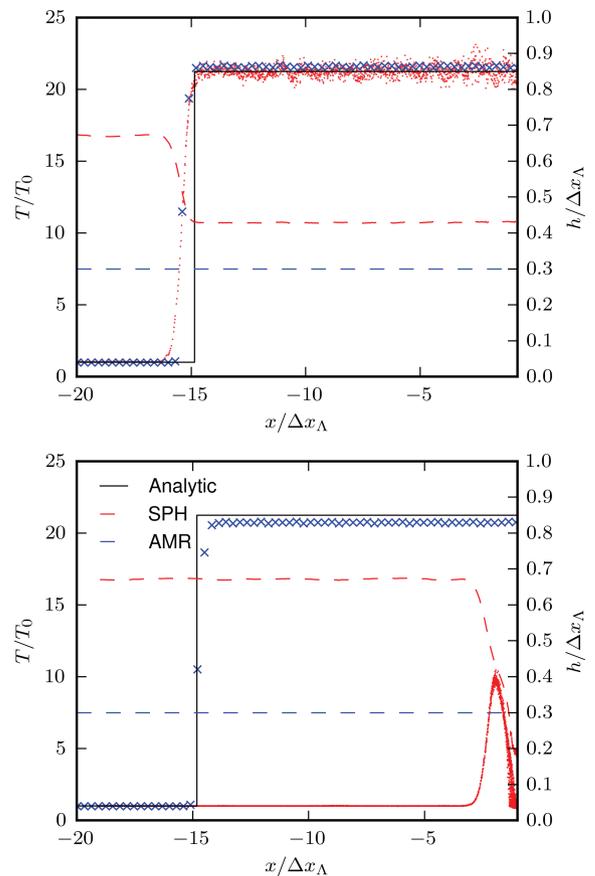


Figure 4. As for Fig. 2 but for an $\mathcal{M} = 6.04$ shock. Upper panel shows the case without cooling, with a higher post-shock temperature than Fig. 2. Lower panel, the case with cooling. Here the SPH particles shock over several smoothing lengths, allowing them time to cool. Unfortunately, this means they never reach the higher temperature where cooling vanishes and their temperatures decline to the pre-shock value, forming a cold dense region similar to that in Fig. 2. The shift of the position of the shock front is due to the conservation of mass; cooling allows the post-shock gas to be compressed down to a small region around $x = 0$. We note here that FLASH adequately captures the post-shock temperature even when cooling is included.

analytic solution. The problem we have attempted to solve involves no elements that an SPH code would not be expected to handle in the limit where the SPH resolution length $h \rightarrow 0$, and using a good prescription for artificial viscosity. The outstanding question here is thus only one of how much resolution is required; to this end we reran the $\mathcal{M} = 4.7$ simulation with a factor of 8 increase in the particle count³ (from $\approx 80\,000$ to $\approx 660\,000$). Let us first, however, make some general remarks about the problem.

Given the maximum temperature, T_s , of the radiating shock we can estimate the error ΔT of the SPH maximum temperature by estimating the radiative cooling over the shock (the physical shock is non-adiabatic and so occurs on time-scales many orders of magnitude shorter than that of the cooling, as discussed in Section 1). Assuming the width of the SPH shock is $\sim h$, the temperature

³One can, of course, successively reduce the width of the shock tube in a three-dimensional simulation to achieve the same scaling as a one-dimensional one, i.e. $h \propto N_{\text{SPH}}^{-1}$. In an astrophysical simulation, however, this option is usually unavailable as the shock will be embedded in an environment which needs to be simulated in full three-dimension.

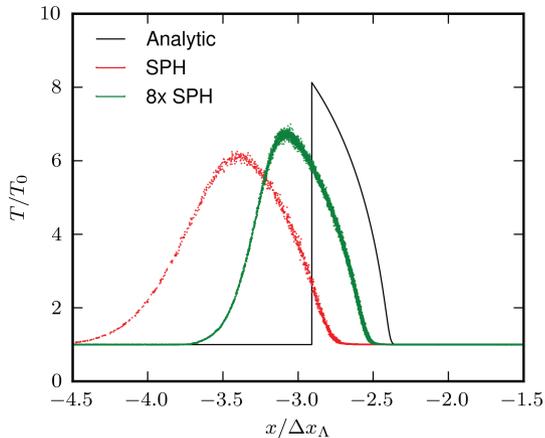


Figure 5. As for the lower panel of Fig. 2, but for two different SPH particle resolutions: *red points* are the SPH particles as for the lower panel of Fig. 2, whereas *green points* are for a two times better resolution run (i.e. a factor of 8 times more particles). The lower resolution run reproduces the temperature peak to within 25 per cent, for the higher it is around 15 per cent. When the simulation is close to convergence, we would expect $\Delta T \propto \Delta u \propto h$, the smoothing length, i.e. to get within 1 per cent of the temperature we would need a factor of $\sim 10^4$ more particles (compared to the lower resolution run). At higher resolutions the offset between the exact and simulated shock fronts is also reduced.

difference will be given, to first order, by

$$\frac{\Delta T}{T_0} \sim \frac{h}{v_0 k_B T_0} m_p |\dot{u}_\Lambda(T_s)| \sim \frac{h}{\Delta x_\Lambda}, \quad (19)$$

where we have assumed that all the mechanical energy has been converted into thermal energy, and that $T_s \gg T_0$. For larger smoothing lengths we expect ΔT to become sublinear in the smoothing length, since the cooling is weaker at lower temperatures (assuming we are on the left-hand side of the ‘cooling spike’).

If we apply this argument to Fig. 5 we see that increasing the number of particles by a factor of ~ 8 (i.e. reducing h by a factor of 2) reduces the temperature error by a factor of ~ 1.5 , suggesting that we are not quite in the linear regime. To reach a temperature within 1 per cent of the analytic temperature would seem to require increasing the particle count by a factor of $\sim 10^4$. Although this is (barely) possible for this particular case, such resolution is not feasible in cosmological calculations. Most shocks in cosmological simulations will occur at lower resolution than we have used in this test case. Therefore, we should seek an alternative solution involving a switch to prevent cooling during the shock process. We intend to explore such a switch in a further paper.

3 A RESOLUTION CRITERION FOR RADIATIVE SHOCKS

Having established the difficulty of modelling some shock problems with radiative cooling, we now wish to obtain a criterion against which we can judge simulations. Such a tool will allow us to identify those simulations where resolution is not a problem and those where more care is required. In the following section we will discuss the effects of resolution in quite a general way before deriving a metric from a simple model problem. We will frame our discussion in terms of SPH; however, there will be analogous arguments for mesh codes.

Let us take a general case of a numerical simulation of a radiative shock. We assume that we have pre-shock gas with velocity, specific internal energy and density v , u , ρ which passes through

a shock and comes to rest (v is the velocity of the incoming gas with respect to that of the post-shock gas). First we note that the SPH shock has a width which is some multiple of the smoothing length $h \propto (m_{\text{SPH}}/\rho)^{1/3}$ (for a mesh this would be the width of a cell), where the numerical factor will include some dependence on the artificial viscosity prescription. The change in thermal energy will be $\Delta u \propto v^2$ by energy conservation, and thus we can define a rate of ‘shock heating’,

$$\dot{u}|_{\text{shock}} = \Delta u / \Delta t \quad (20)$$

$$= k v^3 \left(\frac{m_{\text{SPH}}}{4\pi\rho/3} \right)^{-1/3}, \quad (21)$$

where k is some constant depending upon the details of the SPH scheme used (e.g. neighbour counts). This heating rate is entirely numerical, as can be seen by the presence of the SPH particle mass m_{SPH} : in the continuum approximation of the underlying fluid equations the shock heats the gas instantaneously, hence the heating rate is singular. As we reduce the particle mass, h decreases and the numerical rate at which particles are heated over the shock front increases.

By taking the ratio of the physical rate of gas cooling to the numerical rate at which the gas is shock heated (which, ideally, we would wish to be almost infinite), we can analyse the effects of shock resolution. Only if the absolute value of this dimensionless ratio is small do we expect the shock heating to overwhelm the cooling, i.e. we require

$$|\dot{u}_\Lambda| \frac{1}{c^3 \mathcal{M}^3} \left(\frac{m_{\text{SPH}}}{\rho} \right)^{1/3} < \eta_{\text{SPH}} \quad (22)$$

for the numerical solution to achieve close to the correct post-shock temperature, where η is a dimensionless parameter. Here we have written the velocity of the incoming gas as $v = \mathcal{M}c$ in terms of the Mach number and the upstream sound speed, $c \equiv c_0$. The equivalent for a mesh code can be written with the side length h of a cubic mesh cell written in terms of the mass enclosed, $m_{\text{AMR}} = \rho h^3$, to give

$$|\dot{u}_\Lambda| \frac{1}{c^3 \mathcal{M}^3} \left(\frac{m_{\text{AMR}}}{\rho} \right)^{1/3} < \eta_{\text{AMR}}. \quad (23)$$

In the subsequent section we attempt to determine a reasonable value of η which we can use to evaluate other simulations.

3.1 Heaviside cooling function

To achieve an extremely simple radiative shock we set up a wall shock (see Section 2) with a low Mach number $\mathcal{M} = 2$ and the piecewise cooling function

$$\dot{u}_\Lambda = \begin{cases} 0, & u \leq u_0 \\ -\tilde{\Lambda} \left(\frac{\rho}{\rho_0} \right), & u_0 < u, \end{cases} \quad (24)$$

where $\Delta x_0 \equiv (m_{\text{SPH}}/\rho_0)^{1/3}$ is the initial interparticle spacing, $\tilde{\Lambda} > 0$ a dimensionless cooling parameter, u_0 the initial specific internal energy and as in Section 2 we use $\gamma = 5/3$. In the SPH simulation the particles are initially arranged on a cubic lattice of dimensions $1024 \times 8 \times 8$ in units of Δx_0 (1024 referring to the long x direction). The simulations were all performed with periodic boundary conditions. Usually, a cooling function would be independent of the interparticle spacing; however, we chose to reuse our initial conditions whilst adjusting the dimensionless constant $\tilde{\Lambda}$, and in this way

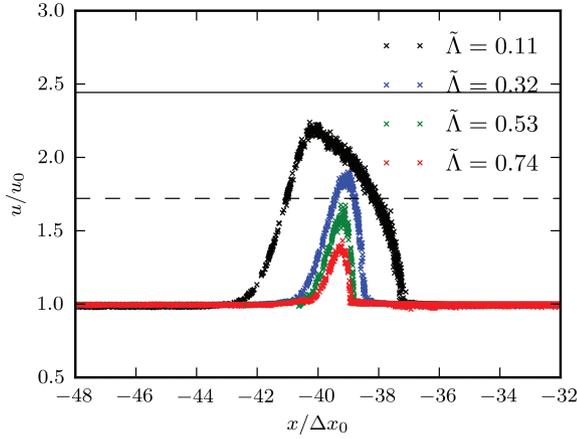


Figure 6. Specific thermal energy versus position for radiative shocks with a Heaviside cooling function, equation (24). Black, blue, green, red crosses are for SPH simulations with dimensionless cooling rates of $\tilde{\Lambda} = 0.11, 0.32, 0.53, 0.74$, respectively; positions are quoted in units of the initial (pre-shock) particle spacing Δx_0 , thermal energies in units of the initial thermal energy u_0 . Solid line indicates the analytic instantaneous post-shock thermal energy $u_s/u_0 = 2.44$. Dashed line indicates mid-point energy between the initial and final thermal energy $\frac{1}{2}(u_s + u_0)/u_0$. When the cooling rate is low ($\tilde{\Lambda} = 0.11$, black crosses), the numerical overcooling is small and the simulation gets close to the right post-shock temperature. Increasing the cooling rate degrades the accuracy of the numerical result. We use this to set a maximum cooling rate that the simulation can tolerate, for example by requiring that the simulated post-shock temperature be larger than half the correct result (horizontal dashed line).

scale the left-hand side of equation (22). This is equivalent to using a fixed cooling function but adjusting the interparticle spacing.

We now make a couple of observations. First, we note that we can calculate the instantaneous post-shock state using the equations derived in Appendix A1, to find $\rho_s/\rho_0 = 2.52$ and $u_s/u_0 = 2.44$. We note that this ratio is independent of the cooling parameter $\tilde{\Lambda}$. Increasing $\tilde{\Lambda}$ in the simulations, however, we expect the maximum post-shock temperature to fall as thermal energy is radiated away over the numerically broadened shock.⁴

In Fig. 6 we see the results of these simulations plotted at a time of $t = 141u_0^{-1/2}\Delta x_0$. We note that the particle distribution in the pre-shock region has also diverged from a lattice arrangement (if it were still a lattice the particles would appear at multiples of Δx_0) into something more glass-like. This is to be expected as the SPH equations of motion favour a large distance to the nearest neighbour for a given density, which can be achieved with a close-packed or glass-like arrangement. The position, velocity and Mach number of the shock at late times are independent of the cooling function, for fixed u_0 (provided the cooling function restores the thermal energy of the gas to u_0) as is shown in Appendix A1.

With a low cooling parameter ($\tilde{\Lambda} = 0.11$, black crosses) we see that the post-shock thermal energy reaches near the theoretical value, whilst with a high cooling parameter ($\tilde{\Lambda} = 0.74$, red crosses) we see that the simulation produces almost no hot gas. We take the mid-point of the thermal energies as a minimum value the code should reach to give at least approximately the correct answer. From Fig. 6 this corresponds to a cooling parameter of $\tilde{\Lambda} \approx 0.4$.

⁴ One might have expected the post-shock thermal energy ratio for a Mach 2 shock to be precisely $2\gamma(\gamma - 1) + 1 = 3.22$ as in the case without cooling; however, the immediate post-shock region is still in motion with respect to the final cold post-shock gas (hence $\tilde{\Lambda} = 0$ is a special case).

Substituting this maximum value into equation (22) allows us to evaluate the parameter η as

$$\eta = \tilde{\Lambda} u_0^{3/2} \frac{1}{c^3 \mathcal{M}^3} \left(\frac{\rho}{\rho_0} \right)^{2/3} \quad (25)$$

$$= \tilde{\Lambda} [\gamma(\gamma - 1)]^{-3/2} \mathcal{M}^{-3} \left(\frac{\rho}{\rho_0} \right)^{2/3} \quad (26)$$

$$\approx 0.08, \quad (27)$$

(where we have used the post-shock density $\rho = \rho_s = 2.52\rho_0$), or, using equation (22),

$$|\dot{u}_\Lambda| \frac{1}{c^3 \mathcal{M}^3} \left(\frac{m_{\text{SPH}}}{\rho} \right)^{1/3} < 0.08. \quad (28)$$

This is the value of η we will use throughout the remainder of this paper.

A similar analysis with FLASH yields an only slightly weaker criterion,

$$|\dot{u}_\Lambda| \frac{1}{c^3 \mathcal{M}^3} \left(\frac{m_{\text{AMR}}}{\rho} \right)^{1/3} < 0.09, \quad (29)$$

where m_{AMR} refers to the mass contained in a mesh cell (since the mass in cells varies we have taken m_{AMR} to be the value in the cell immediately to the right of the shock, for consistency with the evaluation of ρ).

It is worth discussing the differences between a grid and an SPH scheme when the adaptive capabilities are utilized. SPH has a resolution (smoothing) length which refines in areas of high density as $\rho^{-1/3}$. AMR on the other hand can have much more general refinement criteria, for example allowing higher resolution to be applied on features which need not be dense (e.g. shocks). As such AMR has something of an advantage when it comes to shocks, as almost all refinement schemes will refine over discontinuous variable to the maximum level, and hence there is no need to impose the refinement criterion equation (29). Of course it is possible that a region of space in the simulation is already maximally refined, yet even so fails to satisfy the criterion equation (29). We can then interpret this as a test of how well the finite resolution of an AMR simulation represents the physics in the problem.

Refining a simulation in a given volume V of a two-dimensional structure (such as a shock) to scale h in SPH requires $N_{\text{SPH}} \propto h^{-3}$ particles, whilst in AMR one would only need $N_{\text{cell}} \propto h^{-2}$ cells (we note that limitations on the refinement level between cells do not in general alter this scaling relation).

This can be contrasted with a sheet-like structure in a vacuum (e.g. a gravitationally collapsed disc of thickness $\ll h$), which will only require $N_{\text{SPH}} \propto h^{-2}$ particles, the same relation as AMR.⁵

Note that we have been concentrating on how well the simulations reproduce shocks in the presence of cooling. In practice we would also like the code to correctly reproduce the cooling tail, i.e. the cooling of the gas once it has passed through the shock (the right-hand side of Fig. 6). Clearly, here we would like to resolve the cooling length from equation (17), by requiring that $\Delta x_\Lambda \gtrsim h$ in SPH (or the cell size in mesh codes).

In the subsequent section we will apply the resolution criterion we derived to estimate in which areas of cosmological simulations numerical overcooling could be problematic.

⁵ As such SPH could be viewed as a refinement scheme optimized for gravitationally collapsed structures.

4 EFFECTS OF RESOLUTION ON GALAXY FORMATION

4.1 Galaxy formation simulations

In this section we apply the resolution criterion of equation (28) to different regions of temperature and density in the GIMIC SPH simulation of Crain et al. (2009). First we will plot the distribution of gas in temperature–density space and identify some environments of interest. We will then discuss the radiative shock resolution in this parameter space, but also explore mitigating factors which may allow us to have confidence in simulations even when they fail to accurately resolve the shocks.

The GIMIC simulations are zoomed resimulations of nearly spherical regions picked from the Millennium Simulation (Springel et al. 2005), including gas dynamics. Each sphere has a radius of $18 h^{-1}$ Mpc, and the SPH particle mass is $m_{\text{SPH}} \approx 10^6 h^{-1} M_{\odot}$. Radiative cooling in the simulation includes line cooling of 11 elements, Compton cooling with the CMB and thermal bremsstrahlung in the presence of a uniform but evolving ionizing background, as described in Wiersma et al. (2009a; see Fig. 7). The background cosmology, as for the Millennium Simulation, is $\Omega_m = 0.25$, $\Omega_{\Lambda} = 0.75$, $\Omega_b = 0.045$, $n_s = 1$, $\sigma_8 = 0.9$, $H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}$, $h = 0.73$. The enrichment of gas by nucleosynthesis in stars is described in Wiersma et al. (2009b). Photoheating, radiative and adiabatic cooling, and shocks induced by galactic winds and due to accretion result in gas occurring over a wide range of densities and temperatures, illustrated in Fig. 8. Five points A–E in this T – ρ space correspond to physical states where we want to investigate to what extent the simulation properly resolves radiative shocks (see also Table 1). For a general discussion on these diagrams see Wiersma et al. (2010). The simulation code described here was also used in the OWLS project (Schaye et al. 2010).

Point A is a typical IGM point outside virialized haloes, at low density and temperature. Here we see a very well-defined mild upward slope of temperature with ρ of the post-reionization gas. This gas is cooling due to adiabatic expansion of the universe and is being photoheated by the UV background. For a recombination coefficient $\propto T^{-0.7}$ this will at late times result in a temperature–

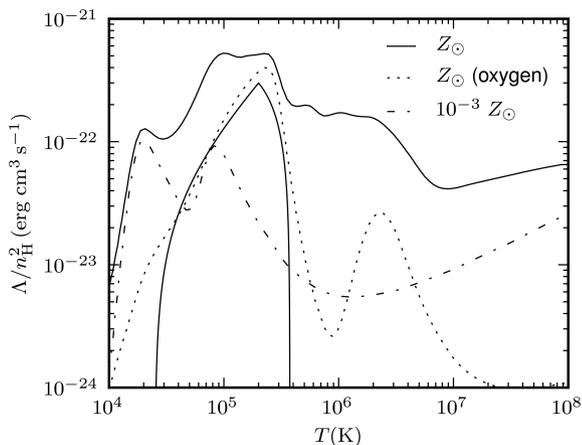


Figure 7. Cooling functions used in the GIMIC simulations at redshift 0. Upper solid and dot–dashed lines represent astrophysical cooling functions for a plasma with metallicity $[Z/Z_{\odot}] = 0$ and -3 (where square brackets denote the base-10 logarithm), respectively, in the presence of an ionizing background (Wiersma et al. 2009a). Lower solid line shows a cooling spike such as in Section 2, for comparison, on the same logarithmic scale. Dotted line shows cooling due to oxygen only, again assuming a solar abundance.

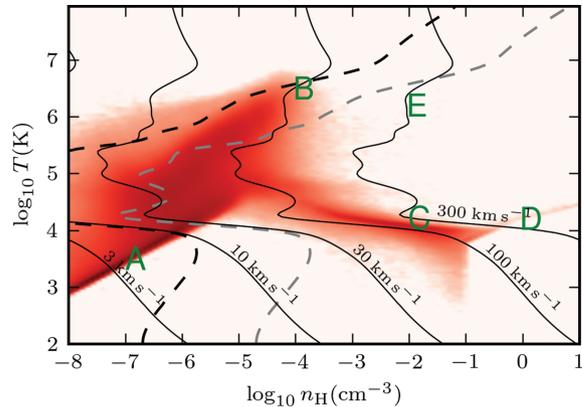


Figure 8. Resolution requirements for correctly representing shocks in different regions of a temperature–density diagram. Solid contours are labelled with the minimum values of the shock speed, $v = cM$, obtained from equation (28), that avoids excessive overcooling in the shock precursor, for simulations using an SPH particle mass of $m_{\text{SPH}} = 10^6 M_{\odot}$. The radiative cooling rate adopted is that of a plasma with solar abundances of elements, $[Z/Z_{\odot}] = 0$, as shown in Fig. 7. The overlaid red shaded region is the phase density in (T, n_{H}) space of SPH particles in a cosmological feedback simulation (see text). Bold letters refer to example environments described in Table 1. Heavy black and grey dashed lines refer to $t_{\text{cool}} = t_{\text{dyn}}$ and $t_{\text{cool}} = 0.1 t_{\text{dyn}}$, respectively. See text for discussion.

Table 1. Astrophysical shock environments identified in Fig. 8.

	n_{H} (cm^{-3})	T (K)	Λ ($\text{erg cm}^3 \text{ s}^{-1}$)	h (kpc)	v_{min} (km s^{-1})
A. IGM	10^{-7}	2000	10^{-24}	150	5
B. Hot halo	10^{-4}	2×10^6	10^{-22}	15	100
C. Cold halo	10^{-2}	10^4	5×10^{-24}	3	200
D. ISM (AGN)	10^0	10^4	5×10^{-24}	0.7	400
E. ISM (SNe)	10^{-2}	10^6	10^{-22}	3	400

density relation of $T \sim \rho^{1/1.7}$ (Hui & Gnedin 1997; Theuns et al. 1998).

Point B corresponds to gas heated in an accretion shock, falling into a galactic halo, or shocked by a galactic wind. Mechanical energy has been converted into thermal energy, and the density will jump by a factor of up to ~ 4 . When this gas cools, it will form the warm gas of point C which may condense to fuel star formation in a central galaxy (White & Rees 1978).

On the far right, $n_{\text{H}} > 10^{-1} \text{ cm}^{-3}$, is a sharp vertical feature in the distribution of SPH particles. This boundary delineates cold halo gas from gas which undergoes star formation in the model used in GIMIC. The denser gas represents a multiphase ISM, for which the imposed pressure–density relation in GIMIC is $p \propto n_{\text{H}}^{4/3}$, known as an effective equation of state for the ISM. Such a state is intended to mimic the physical pressure response in dense gas undergoing star formation (point D); compressing this gas results in significant star formation with associated feedback (see Schaye & Dalla Vecchia 2008 for motivation and details). The SPH density then represents a volume average density of star-forming gas, whereas the physical ISM lies in approximate pressure equilibrium, but with a hot and cold phase and corresponding variation in densities. In particular the simulation does not allow this gas to cool radiatively. Finally, point E represents the domain of Type II SNe that ignite in the hot (10^6 K) sparse phase of the ISM, generated by the activity of

previous generations of SNe. We note that there is little gas marked in this phase as the cooling time is short.

Now let us consider this simulation in terms of its ability to resolve the radiative shocks that occur in these five regions. Equation (28) suggests that a radiative shock of velocity v will be resolved if it satisfies

$$v > \left(\frac{|\dot{u}_\Lambda|}{0.08} \right)^{1/3} \left(\frac{m_{\text{SPH}}}{\rho} \right)^{1/9}, \quad (30)$$

i.e. there is a minimum shock velocity which can be resolved. Shocks below this velocity will tend to artificially radiate away their energy because there will be cooling through the (artificially extended) shock region. Shocks above this velocity will heat up the gas on such a short time-scale in the simulation that the cooling in the shock region will make little difference to the final result.

In Fig. 8 we plot contours of given v , the minimum shock velocity for which there is no significant overcooling in shocks. At each temperature and density a cooling rate is evaluated (using the cooling rate from Wiersma et al. 2009a, shown in Fig. 7, evaluated at solar metallicity, and in the absence of an ionizing background), which is combined with a particle mass of $m_{\text{SPH}} = 10^6 M_\odot$, to derive a minimum shock velocity which can be resolved. Note that equation (30) is very weakly dependent on particle mass, and thus changing mass resolution is a very ineffective way of shifting the contours. These contours represent the minimum velocity shock which can be resolved at each T , ρ . Any shocks at lower velocities will appear artificially colder due to resolution effects.

A key point, however, is that even if we fail to resolve the radiative shock, the cooling of the gas in many cases may be inevitable anyway. Indeed, there can be situations where other processes are occurring on much longer time-scales than the cooling, and for which having an incorrect thermal history of the gas is not a problem as far as the dynamics of the system is concerned.⁶ Establishing a general criterion for these is not trivial; here we will simply compare to the locally estimated dynamical time $t_{\text{dyn}} \equiv (G\rho)^{-1/2}$ as indicative of the time-scales for other processes. We assume that the simulation will cool adequately if $t_{\text{dyn}} \gg t_{\text{cool}}$ even in the case where radiative shocks are resolved poorly (we define $t_{\text{cool}} \equiv |\dot{u}_\Lambda|/u$). The heavy dashed black contour in Fig. 8 represents the line where $t_{\text{dyn}} = t_{\text{cool}}$. All points to the left of this represent $t_{\text{dyn}} < t_{\text{cool}}$, so certainly we would wish to completely resolve any shocks here. We have also included in dashed grey a line where $0.1t_{\text{dyn}} = t_{\text{cool}}$ to demonstrate a somewhat stronger limit. The necessity of resolving the thermal history to the right of this line is questionable, because the gas cooling time is so small in any case. Of course these simulations assume ionization equilibrium and optically thin gas, so the cooling rates may have been overestimated. In addition if one were to attempt to track the chemistry of the shocked gas, for example the formation and destruction of molecular hydrogen, then having a correct thermal history could still be important (Abel et al. 1997).

Now let us evaluate the resolution criterion of equation (30) for the five diverse environments of Table 1. First we take point A for radiative shocks in the IGM; here we can see that the low density and cooling rate combine to allow us to resolve all shocks above 5 km s^{-1} , almost certainly much exceeding our requirements.

For point B we have taken a value for gas heated by a virial shock to $2 \times 10^6 \text{ K}$. The higher density and cooling rate here push up

the minimum shock velocity we can resolve to around 100 km s^{-1} , comparable to the virial shock velocities v_{200} themselves for haloes of mass $\sim 10^{12} M_\odot$:

$$v_{200} = [10GH(z)M_{200}]^{1/3}, \quad (31)$$

(Mo, Mao & White 1998), where G is the gravitational constant, $H(z)$ the Hubble parameter and M_{200} is the virial mass of the halo. For $z \gtrsim 1$ we can approximate the Hubble parameter as $H(z) \approx H_0 \Omega_m^{1/2} (1+z)^{3/2}$ to see

$$v_{200} \approx 201 \text{ km s}^{-1} \left(\frac{h}{0.73} \right)^{1/3} \left(\frac{\Omega_m}{0.25} \right)^{1/6} \times \left(\frac{1+z}{3} \right)^{1/2} \left(\frac{M_{200}}{10^{12} M_\odot} \right)^{1/3}, \quad (32)$$

$$T_{200} = \frac{1}{3} \frac{\mu m_p}{k_B} v_{200}^2 \quad (33)$$

$$\approx 1.0 \times 10^6 \text{ K} \left(\frac{h}{0.73} \right)^{2/3} \left(\frac{\Omega_m}{0.25} \right)^{1/3} \times \left(\frac{1+z}{3} \right) \left(\frac{M_{200}}{10^{12} M_\odot} \right)^{2/3}, \quad (34)$$

where Ω_m is the matter density in units of the critical density and $h = H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1}$. For lower mass haloes the gas actually cools even faster and the shocks are more difficult to resolve; however, the cooling may be so short as to save the situation. We explore this situation further in Section 4.2.

For point C we consider the warm gas within galactic discs. The minimum shock velocity which can be resolved close to the star formation threshold is higher than for point B, because the cooling rate is higher. However, the cooling time is so much smaller than the dynamical time in the disc, so that we suspect that gas cooling will be inevitable in any case. This suggests that, although numerical overcooling is potentially severe here, it is unlikely to have any important effects on the evolution of the disc.

At a higher temperature than C we have point E, a fiducial point for a (Type II) SN blast wave shocking to hot ($\sim 10^6 \text{ K}$), rarefied ($n_H \sim 10^{-2} \text{ cm}^{-3}$) ISM (irradiated and inflated by the massive progenitor star). We have a high minimum resolved shock velocity due to the fast cooling of this gas, making its simulation problematic. We expect the gas to form a cold, dense shell (Cox 1972), and the lack of resolution to manifest itself primarily in an alteration of the onset of this phase. We discuss the implications for SN feedback further in Section 4.3.

Finally, for point D we have considered the case of an AGN outflow shocking into a dense ISM of $n_H \sim 1 \text{ cm}^{-3}$. The minimum resolved velocity is so high here that we can have little confidence in the simulated shock dynamics (excluding the most basic properties such as conservation of momentum). The gas is cooling fast compared to dynamical time-scales, yet we have similar concerns at point E about the artificial suppression of feedback.

4.2 Virial shocks

We now consider the effects of the resolution requirement equation (28) on the discussion on cold accretion and virial shocks around haloes. Here we are following the ideas of spherical collapse set out in White & Rees (1978). The basic question here is what is the fate of gas as it accretes on to a halo, and gets shocked as it converts its mechanical energy into thermal energy. If the cooling

⁶ Note that even if dynamics is not affected, there may be other consequences of numerically underestimating the amount of hot gas, for example when calculating the spectrum of cooling radiation.

time is short, then this hot phase will be a temporary one; however, if the cooling time is long, then a hydrostatic hot halo of gas will form within the halo, the scaling relations for which can be found in e.g. Mo et al. (1998).

The properties and stability of such spherical shocks have been studied by e.g. Birnboim & Dekel (2003), depending on mass and redshift. More massive haloes have hotter shocks with longer cooling times. At a given mass lower redshifts imply lower densities and hence slower cooling, and hence easier build-up of a hot halo. It must be recalled, however, that in this situation geometry will also play a role. If the gas accretion can achieve a configuration where it will penetrate farther into the halo (e.g. in filaments), it will shock at higher densities and generally have a shorter cooling time (the situation is complicated by the fact that the gas continues to accelerate, and so the shock will generally be hotter).

Here we first consider applying our resolution criteria to the spherical case. Assuming a spherical halo of mean density

$$\bar{\rho}_{200} = 200 \left(\frac{3H^2}{8\pi G} \right) \quad (35)$$

and virial mass M_{200} , we take the accreting gas to shock to the virial temperature⁷ T_{200} and virial velocity v_{200} given in equations (31) and (32), respectively. For the baryon density at the edge of the halo we use

$$\rho_b = \frac{1}{3} \frac{\Omega_b}{\Omega_m} \rho_{200}, \quad (36)$$

where the factor 1/3 is the ratio of edge to mean densities for an isothermal sphere of profile $\rho = \rho_0(r/r_0)^{-2}$. We can then apply our shock resolution criteria in terms of the maximum mass of SPH particles that do not suffer from numerical overcooling in the shock,

$$m_{\text{SPH,max}} = \eta^3 |\dot{u}_{\Lambda(T_{200})}^{-3}| v_{200}^9 \rho_b, \quad (37)$$

where our convergence tests suggest that $\eta \approx 0.08$.

Equation (37) defines curves in a plot of virial mass M_{200} versus redshift z , shown in Fig. 9 for a cooling rate appropriate for a plasma with solar abundance ratios but mean metallicity of $[Z/Z_{\odot}] = -3$ (we have chosen the lower metallicity as more representative of accreting gas that has yet to be enriched by several generations of star formation). In Fig. 10 we show the case where cooling is partly suppressed by the presence of a uniform ionizing background (see Wiersma et al. 2009a for details). Each thin coloured line represents the limiting particle mass required to prevent numerical overcooling in the corresponding spherical accretion shock. Clearly the resolution requirement is punitively strict (smallest m_{SPH}) for small haloes ($M_{200} \sim 10^8-10^{10} M_{\odot}$), especially at high redshifts ($z \sim 9$). Intuitively this can be understood because these masses correspond to virial temperatures near the peak of the cooling function, and at higher redshifts the mean baryon density (and thus collisional cooling rate) grows. In the presence of an ionizing background cooling is suppressed in lower mass haloes that cool mostly through hydrogen lines. At lower masses the ionizing radiation has a very large effect, and gas will be *photoheated* instead of cooling (Okamoto, Gao & Theuns 2008).

However, even though lack of resolution will lead to overcooling in some haloes, the cooling time in these haloes may be so short that the gas would cool quickly anyway. The grey area in the figure indicates where the dynamical time in the halo is shorter than the

⁷ The infalling gas has actually twice this energy, so if it shocks into the rest frame of the halo the temperature will be increased by a factor of 2; however, we will ignore this for now.

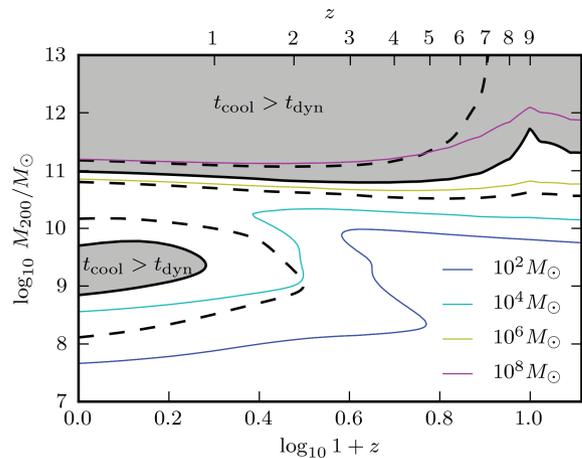


Figure 9. Contours of maximum SPH particle mass m_{SPH} required to prevent numerical overcooling at a virial shock, for gas $([Z/Z_{\odot}] = -3)$ accreting on to haloes of different virial masses M_{200} at a given redshift z . Coloured lines corresponding to $m_{\text{SPH}} = 10^8, 10^6, 10^4$ and $10^2 M_{\odot}$ limits are represented by the *thin maroon, yellow, cyan and blue lines*, respectively. *Black lines* compare cooling time to the dynamical time of the halo: $t_{\text{cool}} = t_{\text{dyn}}$ (*heavy solid line*), $t_{\text{cool}} = 2t_{\text{dyn}}$, $t_{\text{cool}} = \frac{1}{2}t_{\text{dyn}}$ (*heavy dashed lines*). The *shaded grey region* denotes $t_{\text{cool}} > t_{\text{dyn}}$. Numerical overcooling due to lack of resolution in regions where $t_{\text{cool}} \gtrsim t_{\text{dyn}}$ will likely affect the dynamics of the accreting gas; hence, SPH simulations would appear to need resolutions $\sim 10^6 M_{\odot}$.

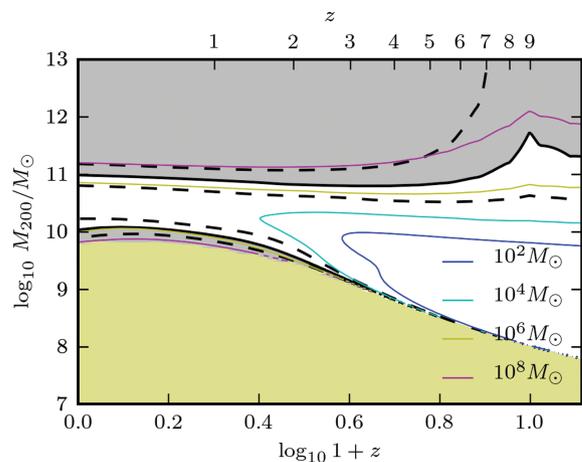


Figure 10. As for Fig. 9, but this time including a uniform ionizing background (see text for details). In the *gold region* (lower left of the figure), the gas is being heated rather than cooled, so resolution of the shock is of lesser importance.

cooling time: in this region we expect that numerical overcooling may prevent the formation of a hot halo. Conversely, in the white region, cooling is so fast that even if a hot halo were to form, it would quickly cool. The demarcation line between these scenarios follows closely the $\sim 10^6 M_{\odot}$ limiting SPH mass (yellow line). Simulation runs with that resolution or better will be able to form hot spherical haloes in situations where we would expect such a hot halo to form. At lower resolution, simulations may artificially suppress the formation of a hot halo due to numerical overcooling in the accretion shock.

Our considerations apply only at the virial radius. However, nearer the centre of haloes we expect this conclusion to remain valid, as the cooling time diminishes faster than the dynamical time.

The very high mass haloes have $2t_{\text{dyn}} < t_{\text{cool}}$ (heavy dashed black line) and we expect these to be in near hydrostatic equilibrium. As a result of these analyses we conclude that $10^6 M_{\odot}$ is a sensible upper limit for the gas particle mass in cosmological simulations intending to capture the evolution of protogalactic haloes, although lower masses enable more accurate resolution of the thermal history of gas in lower mass haloes.

4.3 Thermal feedback

Thermal feedback refers to the mechanism whereby injection of thermal energy into the ISM causes adiabatic expansion of the gas and subsequent suppression of star formation due to the diminished density. The simplest model to envisage is perhaps that of a single SN creating a hot, spherical, rarefied, blast wave. On larger scales, however, we expect to see analogous effects from star-forming regions and AGN. In this section we intend to consider our results in terms of thermal feedback in SPH. We will review the basic physics of thermal feedback and its role in galaxy evolution. We will then discuss its implementation in SPH and derive some quantitative criteria for accurately resolving it. Finally, we will relate our observations to the feedback experiments in other work.

We begin by considering the problem of simulating an SN blast wave. Here we are primarily concerned with the situation where we may artificially radiate away the thermal energy of the blast wave due to a lack of resolution. This would result in the premature transition from a thermally driven to a momentum-driven phase.

A concise overview of the evolution of an SN remnants can be found in Cox (1972). Essentially the blast wave will follow a Sedov–Taylor self-similar solution (Sedov 1959) until the shock temperature T_s falls to a value where the radiative cooling exceeds the cooling via adiabatic expansion. A full calculation is beyond the scope of the present paper; however, we can get close just by dimensional considerations:

$$k_B T_s = (E_0^2 m_p^3 n_H^4 \Lambda^6)^{1/11} \quad (38)$$

$$\approx k_B 4 \times 10^6 \text{ K}, \quad (39)$$

where $E_0 = 10^{51}$ erg is the SNe energy, $n_H = 1.0 \text{ cm}^{-3}$ and $\Lambda = 10^{-22} \text{ erg cm}^3 \text{ s}^{-1}$. The value of Cox (1972) is a factor of 2 smaller, at $2.0 \times 10^6 \text{ K}$.

In a simulation we would like to resolve the transition in the SN shock from being pressure driven to being momentum driven, which typically occurs for shock temperatures $T_s \sim 2.0 \times 10^6 \text{ K}$; numerical overcooling may cause the shock to transition too early, hence underestimating the feedback effect of the explosion on star formation in the surroundings. Using the Sedov similarity solution for a three-dimensional blast wave in a uniform cold medium of adiabatic index $\gamma = 5/3$ and density ρ_0 (which we will shortly take to be the density of the ISM), we can then write the pressure and temperature just inside the shock wave in terms of the shock velocity v_s , as

$$\rho_s = \frac{\gamma + 1}{\gamma - 1} \rho_0 \quad (40)$$

$$p_s = \frac{2}{\gamma + 1} \rho_0 v_s^2 \quad (41)$$

$$k_B T_s = \bar{\mu} \frac{p_s}{\rho_s} = 2 \frac{\gamma - 1}{(\gamma + 1)^2} \bar{\mu} v_s^2, \quad (42)$$

where $\bar{\mu}$ is the mean particle mass. Combining this with the resolution criterion of equation (28) we find that (excluding fairly pathological cooling functions) the cooling will be hardest to resolve at the lower temperatures, i.e. at $T_s = 2.0 \times 10^6 \text{ K}$.

Applying fiducial values for the ISM of $\bar{\mu} \approx 0.6 m_p$ at $T_s = 2.0 \times 10^6 \text{ K}$ yields a shock velocity, density and cooling rate of

$$v_s = 380 \text{ km s}^{-1} \quad (43)$$

$$\rho_s \approx 9 \times 10^{-24} \text{ g cm}^{-3} \quad (44)$$

$$\dot{u}|_{\Lambda} \approx -180 \text{ cm}^2 \text{ s}^{-3}, \quad (45)$$

and the corresponding radius of the blast wave

$$r_s = 1.15 \left(\frac{E}{\rho_0} \right)^{1/5} t^{2/5} \quad (46)$$

$$= 1.15^{5/3} \left(\frac{E}{\rho_0} \right)^{1/3} \left(\frac{2}{5} \right)^{2/3} v_s^{-2/3} \quad (47)$$

$$\approx 15 \text{ pc}, \quad (48)$$

where $E \sim 10^{51}$ erg is the thermal energy injected by the explosion. The corresponding limiting SPH particle mass that avoids numerical overcooling, evaluated from equation (28), is

$$m_{\text{SPH}} = 70 M_{\odot} \left(\frac{\rho_s}{9 \times 10^{-24} \text{ g cm}^{-3}} \right) \times \left(\frac{v_s}{380 \text{ km s}^{-1}} \right)^9 \left(\frac{|u_{\Lambda}|}{180 \text{ cm}^2 \text{ s}^{-3}} \right)^{-3}. \quad (49)$$

The small values of both the radius and the required minimal mass resolution imply that most cosmological simulations of galaxy formation are far from resolving individual SN explosions; however, detailed simulations of high- z dwarf galaxies do indeed already reach such extreme resolutions (e.g. Wise & Abel 2008). For a state-of-the-art mass resolution for a cosmological simulation of say $10^5 M_{\odot}$, a star particle really represents very many stars and hence also many SNe. Simply scaling up E_0 to represent the many SNe that go off does not really help much, as for example the blast radius only scales $\propto E^{1/3}$. In reality different SNe will go off in different places, and once the density of the ISM is decreased due to one explosion, another explosion in the lower density gas will have a much larger effect, eventually resulting in a percolating hot phase.

However, these small scales cannot yet be resolved in current cosmological simulations; hence, they fail to follow the transition from pressure to momentum-driven SN shells. One can try to model the expected effects by simply heating a small number of neighbouring SPH particles. In this case our resolution study indicates that the reheating temperature must be sufficiently high for numerical overcooling not to affect the dynamics. We can thus generalize the above calculation to find the minimum temperature for resolved thermal feedback for a given SPH particle mass. To perform this calculation we will need to associate a shock velocity with a single particle, which we will take to be the Sedov shock velocity for a blast wave of the same mass as the SPH particle:

$$v_s = 1.15^{5/2} \sqrt{\frac{\pi}{3\gamma(\gamma-1)}} \frac{4}{5} c_s \quad (50)$$

$$= 1.1 c_s. \quad (51)$$

Combined with equation (30) we find

$$u_{\text{fb}} \gtrsim \frac{\eta^{-2/3}}{1.1^2 \gamma(\gamma-1)} \left(\frac{m_{\text{SPH}}}{\rho} \right)^{2/9} (|\dot{u}_{\Lambda}|)^{2/3} \quad (52)$$

$$= 4 \left(\frac{m_{\text{SPH}}}{\rho} \right)^{2/9} (|\dot{u}_{\Lambda}|)^{2/3}, \quad (53)$$

the minimum thermal energy required to avoid numerical overcooling.

Dalla, Vecchia & Schaye (2008) argue that for thermal feedback to be effective it requires that the sound crossing time across an SPH particle, $t_s = h/c_s$, be smaller than the cooling time, $\tau_c = u/|\dot{u}_{\Lambda}|$. Using $\rho h^3 \sim m_{\text{SPH}}$, $t_s < \tau_c$ requires that

$$u_{\text{fb}} \gtrsim \left(\frac{m_{\text{SPH}}}{\rho} \right)^{2/9} (|\dot{u}_{\Lambda}|)^{2/3}, \quad (54)$$

which is identical apart from a numerical factor to equation (53). Our criterion is stronger as it takes into account that the code will in practice overestimate the cooling of the gas in shocks; the lower value simply requires there to be a shock.

In Fig. 11 we explore the parameter space for modelling thermal feedback in an SPH simulation with $10^6 M_{\odot}$ particles. At each density, there is a minimum temperature required to drive an adiabatic blast wave phase. The light grey region is defined by the sound crossing time argument of equation (54) and the dark grey is from equation (53). In the white region we expect effective thermal feedback; in the dark grey region it will be suppressed by overcooling in shocks, and finally in the light grey the code will be unable to produce a shock at all.

It is helpful to introduce some numbers. For gas with hydrogen number density $n_{\text{H}} = 1 \text{ cm}^{-3}$, a $m_{\text{SPH}} = 10^6 M_{\odot}$ particle would need to be heated to a temperature of $T_s \approx 5 \times 10^6 \text{ K}$ to be in the pressure-driven phase, according to equation (54). However, our resolution study suggests that we need a higher temperature of $\approx 10^7 \text{ K}$ to prevent excessive overcooling through the shock, implying that at

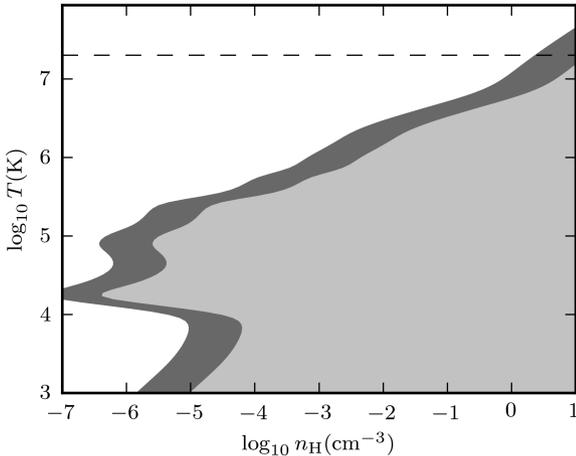


Figure 11. Minimum reheating temperature T required to avoid numerical overcooling as a function of hydrogen number density n_{H} , assuming an SPH resolution of $m_{\text{SPH}} = 10^6 M_{\odot}$ and solar metallicities. Cooling is so rapid in the *shaded* regions that the transition from thermally driven to momentum-driven expansion phases of SN bubbles is so fast that much of the injected energy will be lost to radiation. The *light grey shaded region* corresponds to equation (54), the *dark grey shaded* is the more demanding equation (53). The *white region* is where the reheating temperature is sufficient to force thermal feedback despite resolution concerns. *Dashed line* is an estimate of the specific energy of SNe from Kay et al. (2002).

this resolution the simulation cannot properly represent the effects of thermal feedback. Note also from equation (53) that this improves only very slowly with improved resolution, $\propto m_{\text{SPH}}^{2/9}$. Relating back to models of feedback, this is somewhat problematic as the specific energy of SNe⁸ is estimated to be only around $2 \times 10^7 \text{ K}$ (dashed line in Fig. 11) which can still be too low for the simulation code to properly follow the thermal evolution of the explosion.

Clearly, this has important consequences for prescriptions for SN-driven thermal feedback. In densities above $n_{\text{H}} = 1 \text{ cm}^{-3}$ the thermal feedback starts to reduce its effectiveness, even if we ignite all our SNe in a single time-step, yet this is one of the key environments where feedback is required.

One way to drive feedback at this resolution, whilst still maintaining a globally consistent initial mass function (IMF), is to stochastically inject the energy of SNe due to star formation. Since the mean thermal energy of a single particle can then be greater than $2 \times 10^7 \text{ K}$, we can remain in the effective region of Fig. 11. To balance the IMF other star-forming particles will need to receive less or no SNe energy. The alternative is to increase the resolution, but again the low exponent of m_{SPH} in equation (53) makes this quite prohibitive.

The issue is further complicated by the existence of a multiphase ISM not resolved by the simulation. This is the motivation for many of the prescriptions for feedback, such as applying a fraction of the SNe energy as kinetic energy (Navarro & White 1993), disabling the cooling of thermal bubbles (Gerritsen 1997) or releasing the energy of many accumulated SNe in one step. A more thorough discussion on all these methods can be found in Kay et al. (2002). Another approach is to model the net feedback effects by a subgrid model (for example an imposed equation of state without cooling as in Schaye et al. 2010), or to model the hot and cold phases separately by representing clouds in the cold phase as collisionless particles (Booth, Theuns & Okamoto 2007).

At lower densities it becomes easier to thermally drive a blast wave due to the reduced cooling rate, reinforcing the importance of simulating a multiphase ISM. For their star formation threshold of $n_{\text{H}} = 10^{-1} \text{ cm}^{-3}$, the high-resolution OWLS simulations of Schaye et al. (2010) can represent thermally driven SNe at the edges of discs, but not in more central regions. Indeed, at higher densities the required temperatures rapidly reach extreme values. Booth & Schaye (2009) note that in their simulations AGN feedback requires reheating temperatures $T_s > 10^8 \text{ K}$, as at lower temperature the energy is simply radiated away. We believe that this problem is not a physical one but one of resolution. At lower temperatures the density is higher, the cooling faster, and the cooling region behind the shock cannot be resolved, as is clear from Fig. 11.

4.4 Shocks at the sound speed

As an interesting aside it is worth considering that there will usually be an upper limit to the resolution required. If we assume that the weakest shock has a velocity of the order of the sound speed, $c_s = [\gamma(\gamma-1)u]^{1/2}$, then the minimum requirement for the particle mass for a given problem will be

$$m_{\text{SPH}} = (0.08)^3 [\gamma(\gamma-1)]^{9/2} \min_{x \in V} \{ |\dot{u}_{\Lambda}|^{-3} u^{9/2} \rho \}. \quad (55)$$

⁸ The specific energy of SNe is the energy released by SNe per unit mass of star-forming gas (Kay et al. 2002), i.e. if all SNe were to ignite simultaneously we would reach this mean temperature.

Unfortunately, such a limit will usually be very small indeed, at least for cosmological simulations, because of the low sound speed of cold, dense gas present in galactic discs. However, if one chooses to go down this path, then one can examine the following criteria. If we have a conventional collisional cooling function, then \dot{u}_Λ/ρ is independent of density, and we can make the additional assumptions that $\dot{u}_\Lambda \rightarrow 0$ as $u \rightarrow 0$, and for large u

$$\frac{|\dot{u}_\Lambda|}{\rho} \sim u^{1/2} \quad (56)$$

(thermal bremsstrahlung), giving

$$m_{\text{SPH}} \propto \min_{x \in V} \{\rho^{-2}\}, \quad (57)$$

i.e. the smallest particle mass is determined by the highest density in the problem. This analysis is of course not valid with Compton cooling via the CMB, or the presence of a UV background, as neither process is collisional.

5 CONCLUSIONS

In this paper we have examined the role of radiative cooling in shocks. We have found a general analytical solution for one-dimensional piecewise linear collisional cooling functions and compared it to numerical simulations of the same shock, performed with an SPH code (GADGET) and an AMR code (FLASH). These codes smear out the shock over several particles or cells, and such an artificial ‘pre-shock’ results in numerical overcooling which may prevent the formation of a hot post-shock region. We have estimated a general resolution criterion to avoid such overcooling, and applied it to the problems of virial shocks and the production of hot gaseous haloes. We have found that to avoid numerical overcooling of accretion shocks on to haloes that should develop a hot corona a particle or cell mass resolution of $10^6 M_\odot$ is required (Fig. 10), which is within reach of current state-of-the-art simulations.

Similarly, we have applied our estimates to thermal feedback from AGN or SNe blast waves, in the presence of radiative cooling. We have seen that the energy required to drive thermal feedback at a given mass scale, for current numerical results, is an order of magnitude higher than one would expect just from physical considerations. For cosmological simulations ($10^6 M_\odot$ gas particles) of an $n_{\text{H}} = 1.0 \text{ cm}^{-3}$ ISM we see (Fig. 11) that temperatures in excess of 10^7 K are required to effectively drive thermal feedback by avoiding spurious suppression of the feedback by numerical overcooling.

Although all of these issues can be rectified by increasing the resolution, the minimum thermal energy of injected feedback required to avoid artificial cooling scales weakly with decreasing particle mass, $\propto m_{\text{SPH}}^{2/9}$ (see equation 53). Consequentially, a potentially fertile region of study may be that of cooling switches, i.e. a criterion for disabling cooling through a shock. Such a switch would allow a simulation to resolve temperatures much closer to the physical temperatures of radiative shocks without requiring extreme resolutions. Unfortunately, it is not a straightforward problem to have a criterion that will consistently suppress cooling in the presence of shocks, yet does not affect cooling in regions where there are no shocks. Since we can never hope to completely remove resolution effects, it seems sensible to have a more limited aim, perhaps to capture the temperatures of shocks up to some maximum cooling rate. As such one might wish to suppress cooling, when the cooling time is greater than some fraction of the shock heating time. We intend to explore this avenue in a further paper.

Further work could include the effects of shock-induced non-collisional ionizational equilibrium (non-CIE) or non-thermalized

gas. Since the resolution can make such a significant modification to the thermal history of a gas, we expect that a criterion due to non-CIE may be quite strict.

ACKNOWLEDGMENTS

PC would like to acknowledge the support of an STFC studentship. The authors would like to thank Rob Crain for the reuse of simulation data, and Volker Springel for the provision of the SPH code GADGET. The authors thank Peter Thomas, Frazer Pearce, Justin Read, Tom Abel, Daniel Price, Romain Teyssier and the anonymous referee for encouragement and very useful suggestions. Some of the calculations for this paper were performed on the ICC Cosmology Machine, which is part of the DiRAC Facility jointly funded by STFC, the Large Facilities Capital Fund of BIS and Durham University. The FLASH software used in this work was in part developed by the DOE-supported ASC/Alliance Center for Astrophysical Thermonuclear Flashes at the University of Chicago.

REFERENCES

- Abel T., 2011, *MNRAS*, 413, 271
 Abel T., Anninos P., Zhang Y., Norman M. L., 1997, *New Astron.*, 2, 181
 Agertz O. et al., 2007, *MNRAS*, 380, 963
 Balsara D. S., 1995, *J. Comput. Phys.*, 121, 357
 Bate M., 1995, PhD thesis, Univ. Cambridge
 Birnboim Y., Dekel A., 2003, *MNRAS*, 345, 349
 Booth C. M., Schaye J., 2009, *MNRAS*, 398, 53
 Booth C. M., Theuns T., Okamoto T., 2007, *MNRAS*, 376, 1588
 Chevalier R. A., Imamura J. N., 1982, *ApJ*, 261, 543
 Colella P., Woodward P. R., 1984, *J. Comput. Phys.*, 54, 174
 Cox D. P., 1972, *ApJ*, 178, 159
 Crain R. A. et al., 2009, *MNRAS*, 399, 1773
 Dalla Vecchia C., Schaye J., 2008, *MNRAS*, 387, 1431
 Friedrichs K. O., Lax P. D., 1971, in *Proc. Natl Acad. Sci.*, 68, 1686
 Fryxell B. et al., 2000, *ApJS*, 131, 273
 Gerritsen J. P. E., 1997, PhD thesis, Groningen Univ., the Netherlands
 Gingold R. A., Monaghan J. J., 1977, *MNRAS*, 181, 375
 Godunov S. K., Ryabenki V. S., 1964, *Theory of Difference Schemes - An Introduction*. North Holland Pub. Co., Amsterdam
 Harten A., Lax P. D., van Leer B., 1983, *SIAM Rev.*, 25, 35
 Hui L., Gnedin N. Y., 1997, *MNRAS*, 292, 27
 Hutchings R. M., Thomas P. A., 2000, *MNRAS*, 319, 721
 Inutsuka S.-I., 2002, *J. Comput. Phys.*, 179, 238
 Kay S. T., Pearce F. R., Frenk C. S., Jenkins A., 2002, *MNRAS*, 330, 113
 Kolmogorov A., 1941, *Dokl. Akad. Nauk SSSR*, 30, 301
 Lucy L. B., 1977, *AJ*, 82, 1013
 Mo H. J., Mao S., White S. D. M., 1998, *MNRAS*, 295, 319
 Monaghan J. J., 1997, *J. Comput. Phys.*, 136, 298
 Monaghan J. J., 2005, *Rep. Progress Phys.*, 68, 1703
 Navarro J. F., White S. D. M., 1993, *MNRAS*, 265, 271
 Okamoto T., Gao L., Theuns T., 2008, *MNRAS*, 390, 920
 Price D. J., 2008, *J. Comput. Phys.*, 227, 10040
 Read J. I., Hayfield T., Agertz O., 2010, *MNRAS*, 405, 1513
 Schaye J., Dalla Vecchia C., 2008, *MNRAS*, 383, 1210
 Schaye J. et al., 2010, *MNRAS*, 402, 1536
 Sedov L. I., 1959, *Similarity and Dimensional Methods in Mechanics*. Academic Press, New York
 Springel V., 2005, *MNRAS*, 364, 1105
 Springel V. et al., 2005, *Nat*, 435, 629
 Theuns T., Leonard A., Efstathiou G., Pearce F. R., Thomas P. A., 1998, *MNRAS*, 301, 478
 von Neumann J., Richtmyer R. D., 1950, *J. Applied Phys.*, 21, 232
 White S. D. M., 1996, *Cosmology and Large Scale Structure*. Elsevier, Dordrecht
 White S. D. M., Rees M. J., 1978, *MNRAS*, 183, 341

- Wiersma R. P. C., Schaye J., Smith B. D., 2009a, MNRAS, 393, 99
 Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009b, MNRAS, 399, 574
 Wiersma R. P. C., Schaye J., Dalla Vecchia C., Booth C. M., Theuns T., Aguirre A., 2010, MNRAS, 409, 132
 Wise J. H., Abel T., 2008, ApJ, 685, 40
 Zel'Dovich Y. B., Raizer Y. P., 1967, Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena. Academic Press, New York

APPENDIX A: RADIATIVE SHOCKS WITH PIECEWISE LINEAR COOLING FUNCTIONS

A1 Similarity solution for a one-dimensional radiatively cooling shock

We start with an ideal gas with adiabatic index γ ,

$$p = (\gamma - 1)\rho u, \quad (\text{A1})$$

and a collisional radiative cooling function,

$$du|_{\Lambda} = -\rho f(u) dt. \quad (\text{A2})$$

These combine to give an evolution of

$$du = (\gamma - 1)\frac{u}{\rho}d\rho - \rho f(u) dt. \quad (\text{A3})$$

Stationary solutions of a post-shock cooling region satisfy integrals of the mass and momentum equations, i.e.

$$\rho(v - u_s) = \rho_0(v_0 - u_s) \quad (\text{A4})$$

$$p + \rho(v - u_s)^2 = p_0 + \rho_0(v_0 - u_s)^2, \quad (\text{A5})$$

where u_s is the shock velocity and ρ_0, p_0 and v_0 denote the density, pressure and velocity at some arbitrary downstream point. Thus, the density, velocity and thermal energy can be written in terms of a similarity variable λ :

$$\rho/\rho_0 = \lambda \quad (\text{A6})$$

$$\frac{v - u_s}{v_0 - u_s} = \lambda^{-1} \quad (\text{A7})$$

$$u/u_0 = (a + 1)\lambda^{-1} - a\lambda^{-2}, \quad (\text{A8})$$

with

$$a = \frac{\rho_0(v_0 - u_s)^2}{p_0}. \quad (\text{A9})$$

Now we assume we have a piecewise linear cooling function, i.e. we can solve each segment separately with the linear cooling function

$$f(u) = A(u - u_c), \quad (\text{A10})$$

where A is some constant and u_c denotes the ‘cold’ thermal energy where cooling vanishes. This gives an ordinary differential equation for x of the form

$$\frac{dx}{d\lambda} = \frac{v_0 - u_s}{A\rho_0} \left[\frac{\gamma(a + 1)\lambda^{-4} - (\gamma + 1)a\lambda^{-5}}{(a + 1)\lambda^{-1} - a\lambda^{-2} - u_c/u_0} \right], \quad (\text{A11})$$

which can be solved generally; however, in the case of $u_c = u_0$ we have the particularly simple case:

$$x - x_0 = \frac{v_c - u_s}{A\rho_c} \left[\frac{\gamma - a}{a - 1} \log(\lambda^{-1} - 1) + \frac{1 - a\gamma}{(a - 1)a^2} \log(1 - a\lambda^{-1}) - \frac{a + 1}{a} \lambda^{-1} - \frac{\gamma + 1}{2} \lambda^{-2} \right]$$

$$\lambda \in \left[\frac{a}{a + 1} \frac{\gamma + 1}{\gamma}, 1 \right],$$

the left-hand limit for λ coming from entropy considerations. An example cooling shock of this form can be seen in Fig. 1.

A2 Colliding gas

Assume two homogeneous flows collide from the left and right, with properties $\rho_0, p_0, \pm v_0$. With no cooling, a hot, static region is created in the centre, with properties p_c and ρ_c . The mass, momentum and energy equations are

$$(v_0 - u_s)\rho_0 = -\rho_c u_s \quad (\text{A12})$$

$$(v_0 - u_s)^2 \rho_0 + p_0 = \rho_c u_s^2 + p_c \quad (\text{A13})$$

$$\frac{p_0}{\rho_0} + \frac{1}{2}(\gamma - 1)v_0^2 = \frac{p_c}{\rho_c}, \quad (\text{A14})$$

where u_s is the velocity of the left moving shock in the rest frame. Eliminating p_c, ρ_c gives

$$u_s^2 + \frac{1}{2}(\gamma - 3)u_s v_0 = \frac{p_0}{\rho_0} + \frac{1}{2}(\gamma - 1)v_0^2, \quad (\text{A15})$$

so

$$u_s = -\frac{1}{4}(\gamma - 3)v_0 - \frac{1}{4}\sqrt{v_0^2(\gamma + 1)^2 + 16\frac{p_0}{\rho_0}}. \quad (\text{A16})$$

Assume now that there is cooling and that the gas in the centre cools to the temperature of the pre-shock gas (where cooling is assumed to vanish). In this case the mass, momentum and energy equations are

$$(v_0 - u_s)\rho_0 = -\rho_c u_s \quad (\text{A17})$$

$$(v_0 - u_s)^2 \rho_0 + p_0 = \rho_c u_s^2 + p_c \quad (\text{A18})$$

$$p_0/\rho_0 = p_c/\rho_c, \quad (\text{A19})$$

where these equations are only dependent on the cooling function via the thermal state at which cooling vanishes, p_0/ρ_0 . Eliminating p_c, ρ_c gives

$$u_s^2 \rho_0 - u_s v_0 \rho_0 - p_0 = 0. \quad (\text{A20})$$

The solution for the shock velocity $u_s = v_0/2 - \sqrt{(v_0/2)^2 + p_0/\rho_0}$ p_c, ρ_c can be found by substitution.

The conditions immediately to the right of the shock (v_s, ρ_s, p_s) can be found from the usual Rankine–Hugoniot relations,

$$(v_0 - u_s)\rho_0 = (v_s - u_s)\rho_s \quad (\text{A21})$$

$$(v_0 - u_s)^2 \rho_0 + p_0 = (v_s - u_s)^2 \rho_s + p_s \quad (\text{A22})$$

$$\frac{1}{2}(v_0 - u_s)^2 + \frac{\gamma}{\gamma - 1} \frac{p_0}{\rho_0} = \frac{1}{2}(v_s - u_s)^2 + \frac{\gamma}{\gamma - 1} \frac{p_s}{\rho_s}. \quad (\text{A23})$$

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.