School of Physics and Astronomy
ASP4200/ASP4020

ASP4200/ASP4020 Computational Astrophysics

A crash course in astrophysical fluid dynamics

Daniel J. Price

October 8, 2024

You can view the online lectures associated with this material at:
https://www.youtube.com/watch?v=MOjPtBFGtSI&list=PLMzuj51UjsPTZjHd6XKB4PYbqYDsEBKwH&index=5

Contents

1

Compressible hydrodynamics

Before we solve anything on the computer we need to understand what it is we're trying to do. I will therefore start with a quick introduction to compressible hydrodynamics. It will be a whirlwind tour because we want to focus on the computational aspects in this course. But first we need to go over the basic maths and physics of the fluid equations in order to understand our computational approach.

Let's start with compressible hydrodynamics. What do we even mean by "hydrodynamics"? We mean a particular set of partial differential equations — the equations of hydrodynamics. The inviscid version of these equations are called the Euler equations, after Leonard Euler (1707–1783). If we add viscosity these would be called the Navier-Stokes equations. But the Navier-Stokes equations are usually applied to incompressible flow, whereas in astrophysical fluid dynamics we deal almost entirely with compressible and nearly inviscid flow.

1.1 Equations of hydrodynamics

The fluid equations[1] are simply mathematical expressions of physical rules. These rules are simple: mass, momentum and energy must be conserved.

[1]When we refer to 'fluids', this could be solid, liquid or gas. In this course we are mostly concerned with gas because there are not too many liquids in space.

1.1.1 Conservation of mass

Conservation of mass is expressed by the *continuity equation*

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{v}) = 0, \tag{1.1}$$

where ρ is the density and $\boldsymbol{v}$ is the velocity (a vector) and $\nabla\cdot$ is the vector calculus divergence operator[2], here acting on the vector $\rho\boldsymbol{v}$. Equation (1.1) expresses the "rule" that mass must be conserved. Expanding the second term, we can also write this equation in the form

$$\frac{\partial \rho}{\partial t} + (\boldsymbol{v} \cdot \nabla)\rho = -\rho(\nabla \cdot \boldsymbol{v}). \tag{1.2}$$

1.1.2 Conservation of momentum

Our next rule is conservation of momentum. It is expressed by the 'momentum equation'

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v} = -\frac{\nabla P}{\rho} + \boldsymbol{a}_{\rm ext}, \tag{1.3}$$

where P is the pressure and $\boldsymbol{a}_{\rm ext}$ refers to any external acceleration (e.g. gravity) applied to the fluid. The main term of interest that accelerates the fluid is the $-\nabla P/\rho$ term, namely the acceleration caused by gradients (i.e. differences) in pressure.

1.1.3 Conservation of energy

Finally, we have conservation of energy. We are going to express this by writing down an equation for the internal energy per unit mass, u, given by

$$\frac{\partial u}{\partial t} + (\boldsymbol{v} \cdot \nabla)u = -\frac{P}{\rho}(\nabla \cdot \boldsymbol{v}) + \Lambda_{\rm heat} - \Lambda_{\rm cool}, \tag{1.4}$$

where $\Lambda_{\rm heat}$ and $\Lambda_{\rm cool}$ represent any external heating and cooling applied to the fluid. As we will see later, one can write down equivalent equations expressing conservation of energy in terms of other variables, e.g. the total specific energy $e \equiv \frac{1}{2}v^2 + u$ instead of u.

[2]See vector calculus revision notes if you are not confident working in vector notation

1.1.4 Summary of fluid equations

In summary, the fluid equations expressing conservation of mass, momentum and energy are given by

$$\begin{aligned}
\frac{\partial \rho}{\partial t} + (\boldsymbol{v} \cdot \nabla)\rho &= -\rho(\nabla \cdot \boldsymbol{v}), && (1.5)\\
\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v} &= -\frac{\nabla P}{\rho} + \boldsymbol{a}_{\rm ext}, && (1.6)\\
\frac{\partial u}{\partial t} + (\boldsymbol{v} \cdot \nabla)u &= -\frac{P}{\rho}(\nabla \cdot \boldsymbol{v}) + \Lambda_{\rm heat} - \Lambda_{\rm cool}. && (1.7)
\end{aligned}$$

Here all of our derivatives are written as partial, or *Eulerian* derivatives.

1.1.5 Advection and the Lagrangian time derivative

You may notice already that the left hand side of our equations look similar. This is no accident — the '$\boldsymbol{v} \cdot \nabla$' terms simply express the fact that the fluid is moving and that the fluid properties (ρ, $\boldsymbol{v}$ and u) are properties that are 'carried' or *advected* by the flow.

We can simplify these expressions by defining the 'co-moving' or 'Lagrangian' time derivative. We will use $\mathrm{d}/\mathrm{d}t$ to distinguish this operator from the *Eulerian* time derivative $\partial/\partial t$. We define the *Lagrangian* time derivative according to

$$\frac{\mathrm{d}}{\mathrm{d}t} \equiv \frac{\partial}{\partial t} + \boldsymbol{v} \cdot \nabla \tag{1.8}$$

So for example, applying this derivative to the density we would have

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} \equiv \frac{\partial \rho}{\partial t} + (\boldsymbol{v} \cdot \nabla)\rho. \tag{1.9}$$

The way to understand the Lagrangian derivative is as the time derivative of a co-moving observer. What does this mean? An easy way to think about the Lagrangian derivative is as follows: It is the operator such that the time derivative of the position $\boldsymbol{x} \equiv [x, y, z]$ equals the velocity, i.e.

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{v}. \tag{1.10}$$

In particular, it does not make sense to use the partial/Eulerian time derivative here, since $\partial \boldsymbol{x}/\partial t \neq \boldsymbol{v}$.

Lagrangian and Eulerian observers

The Lagrangian derivative corresponds to the time derivative seen by an observer co-moving with the flow. Imagine you have a river, and a flotilla of equally-spaced boats are floating along the river at some (constant) speed $\boldsymbol{v}_{\rm river}$ and one wishes to know how the density of boats, $\rho_{\rm b}$, changes with time. For an observer sitting on one of the boats (a *Lagrangian*, or *co-moving* observer), the local density of boats is constant, so $\mathrm{d}\rho_{\rm b}/\mathrm{d}t = 0$. However for an observer sitting on the bank watching the flotilla pass, the density is zero, rises to a constant and then fades to zero with time as the flotilla passes. This *Eulerian* observer sees $\partial\rho_{\rm b}/\partial t \neq 0$ — more precisely, $\partial\rho_{\rm b}/\partial t = -\boldsymbol{v}_{\rm river}\cdot\nabla\rho_{\rm b}$.

1.1.6 Equations of hydrodynamics in Lagrangian form

With the aid of the Lagrangian derivative, our fluid equations simplify to

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\rho(\nabla\cdot\boldsymbol{v}), \tag{1.11}$$

$$\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = -\frac{\nabla P}{\rho} + \boldsymbol{a}_{\rm ext}, \tag{1.12}$$

$$\frac{\mathrm{d}u}{\mathrm{d}t} = -\frac{P}{\rho}(\nabla\cdot\boldsymbol{v}) + \Lambda_{\rm heat} - \Lambda_{\rm cool}. \tag{1.13}$$

Much simpler! For the remainder of these notes we will assume no external forces ($\boldsymbol{a}_{\rm ext} = 0$) and no heating or cooling ($\Lambda_{\rm heat} = \Lambda_{\rm cool} = 0$).

1.2 Equation of state

We have 5 equations, evolving ρ, $\boldsymbol{v}$ and u, but 6 unknowns (ρ, $\boldsymbol{v}$, u and P). So the equations cannot be solved unless we provide an additional equation to act as a *closure relation*. This should be an equation that connects the remaining unknown P to known quantities, namely ρ and u. This is known as the *equation of state*.

1.2.1 Ideal gas

For an ideal gas we have $P = nk_{\rm B}T$, where n is the number density. Expressing this in terms of the mass density ρ gives

$$P = \frac{\rho k_{\rm B} T}{\mu m_u}, \tag{1.14}$$

where k_B is the Boltzmann constant, T is the temperature, μ is the mean molecular weight (e.g. $\mu = 2$ for a gas made of molecular Hydrogen) and m_u is the atomic mass unit (i.e. mass of a proton, or more precisely 1/12 the mass of a Carbon-12 atom)[3].

However, the fluid motion is not influenced by the temperature or the composition, only by the pressure. Thus it is more helpful to express the equation of state in terms of the internal energy, u, instead of T. We write

$$P = (\gamma - 1)\rho u, \tag{1.15}$$

where γ is the *ratio of specific heats*. Comparing (1.14) and (1.15), we find

$$u \equiv \frac{1}{(\gamma - 1)} \frac{k_\mathrm{B} T}{\mu m_u}. \tag{1.16}$$

From this equation we see that using the internal energy means we don't need to worry about the composition. That is, we only need to supply μ if we wanted to interpret our results in terms of temperature. The value of γ depends on the number of degrees of freedom in the gas. For a monatomic gas [a gas where the particles are single atoms], we have

$$u = \frac{3}{2} n k_\mathrm{B} T = \frac{3}{2} \frac{k_\mathrm{B} T}{\mu m_u}, \tag{1.17}$$

so we deduce that $\gamma = 5/3$ for this case.

1.2.2 Solving the energy equation with no external heating or cooling

A simpler equation of state can be employed in the case where there is no external heating or cooling. In this case our energy equation (1.13) becomes simply

$$\frac{\mathrm{d}u}{\mathrm{d}t} = -\frac{P}{\rho}(\nabla \cdot \boldsymbol{v}) = \frac{P}{\rho^2}\frac{\mathrm{d}\rho}{\mathrm{d}t}. \tag{1.18}$$

where in the last step we substituted the continuity equation in the form (1.11). If we then assume the equation of state in the form (1.15) we have

$$\frac{1}{u}\frac{\mathrm{d}u}{\mathrm{d}t} = \frac{(\gamma - 1)}{\rho}\frac{\mathrm{d}\rho}{\mathrm{d}t}, \tag{1.19}$$

[3] I personally avoid use of the gas constant $\mathcal{R}$ in astrophysics since i) there are two different definitions and ii) we rarely need to deal with moles.

giving

$$\frac{\mathrm{d}\ln u}{\mathrm{d}t} = (\gamma - 1)\frac{\mathrm{d}\ln\rho}{\mathrm{d}t} = \frac{\mathrm{d}\ln\rho^{(\gamma-1)}}{\mathrm{d}t}. \tag{1.20}$$

Dropping the dt and integrating both sides we have

$$\int \mathrm{d}\ln u = \int \mathrm{d}\ln\rho^{(\gamma-1)}, \tag{1.21}$$

$$\therefore \ln u = \ln\rho^{(\gamma-1)} + C, \tag{1.22}$$

$$\therefore u = \tilde{K}\rho^{(\gamma-1)}, \tag{1.23}$$

giving, using (1.15)

$$P = K\rho^{\gamma}, \tag{1.24}$$

where C and $\tilde{K}$ and K are arbitrary constants. Equation (1.24) is referred to as a *polytropic* equation of state and is the exact solution to the energy equation for an adiabatic gas with no irreversible heating or cooling. In this case we do not need to solve (1.13) on the computer, one can simply use the analytic solution (1.24) directly.

1.2.3 Isentropic gas

We can also consider (1.24) as referring to an *isentropic* (constant entropy) gas because from the first law of thermodynamics we have

$$T\mathrm{d}S = \mathrm{d}U + P\mathrm{d}V, \tag{1.25}$$

where S and U refer to the entropy and thermal energy *per unit volume*. Using $V = m/\rho$ and hence $\mathrm{d}V = -m/\rho^2\mathrm{d}\rho$ gives the time evolution for the entropy *per unit mass* as

$$T\frac{\mathrm{d}s}{\mathrm{d}t} = \frac{\mathrm{d}u}{\mathrm{d}t} - \frac{P}{\rho^2}\frac{\mathrm{d}\rho}{\mathrm{d}t}, \tag{1.26}$$

which equals zero if the gas is adiabatic, since $\mathrm{d}u/\mathrm{d}t = P/\rho^2\mathrm{d}\rho/\mathrm{d}t$. Now consider the Lagrangian time derivative of the quantity $K = P/\rho^{\gamma}$, which gives

$$\frac{\mathrm{d}K}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{P}{\rho^{\gamma}}\right) = \frac{\gamma-1}{\rho^{\gamma-1}}\left(\frac{\mathrm{d}u}{\mathrm{d}t} - \frac{P}{\rho^2}\frac{\mathrm{d}\rho}{\mathrm{d}t}\right). \tag{1.27}$$

So assuming $K = \text{const}$ is equivalent to stating that there is no change in entropy.

1.2.4 Isothermal gas

In the case where $\gamma = 1$ in (1.24) we obtain $P \propto \rho$. From our ideal gas law (1.14) we can see that this corresponds to $T =$ const.. So we refer to this as an *isothermal* — constant temperature — equation of state. It is convenient to write (1.24) in the form

$$P = c_s^2 \rho, \tag{1.28}$$

where c_s is the (constant) isothermal sound speed. Why we identify this constant with the sound speed will become clear in Section 1.4.4. For numerical work this is the simplest possible closure — one simply needs to specify the value of c_s to close the equation set.

Isothermal gas as a singular limit

The limit of $\gamma \to 1$ is known as a *singular limit* in that analytic solutions for $\gamma \neq 1$ do not simply reduce to the $\gamma = 1$ case. One can observe this in our equations (1.15) and (1.27) where taking $\gamma \to 1$ gives nonsense. Despite this our result (1.24) gives the correct limit when $\gamma \to 1$. The other famous singular limit in the fluid equations is that solutions to the fluid equations *with* viscosity — the Navier Stokes equations — do not simply reduce to solutions of the fluid equations *without* viscosity — the Euler equations — as the viscosity tends to zero.

1.3 Classification of partial differential equations

Partial differential equations are classified into three different kinds: elliptic, parabolic or hyperbolic. They can also be a mix of all three. While there is a mathematical definition, we are more interested in the physics of the equations. Briefly:

Elliptic. Example: Poisson's equation for the gravitational field

$$\nabla^2 \Phi = 4\pi G \rho. \tag{1.29}$$

The key ingredient is that there is no time in the equation. This implies *instant action*. That is, since there is no time involved, elliptic equations require information to be propagated instantaneously across the entire computational domain. For example, if the density changes, the gravitational potential Φ must also change. Everywhere in space. Instantly. This is expensive since it requires some form of global communication in the computational domain.

Parabolic. Example: The heat equation

$$\frac{\partial u}{\partial t} = \kappa \nabla^2 u. \tag{1.30}$$

Here there is a single time derivative and the physics corresponds to that of *diffusion*. Dimensional analysis shows that $[\kappa] = L^2/T$. So therefore the time for information to propagate is $T \propto L^2/\kappa$. This means that, for a given resolution length Δx, numerical solutions will only be stable with a timestep $\Delta t \lesssim \Delta x^2/\kappa$. Since this scales $\propto h^2$ it quickly becomes prohibitive at high spatial resolution, unless implicit time-stepping schemes are employed.

Hyperbolic. Example: The wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \nabla^2 u. \tag{1.31}$$

Here there are second derivatives in space and time. The corresponding physics is of propagating *waves*. Since $[c] = L/T$ the time for information to propagate is $T \propto L/c$. So for a given resolution length Δx, numerical solutions will only be stable with a timestep $\Delta t \lesssim \Delta x/c$. Since this scales $\propto h$ it is the most efficient to solve in terms of timestep size.

Do penguins get cold? If a tree falls in a forest and nobody hears it, does it make a sound? What nationality is Les Murray? These are life's mysteries. But a mystery we can solve is: *Which of the above kinds of partial differential equations are the fluid equations?* We just need to solve the equations...

1.4 Linear solutions

1.4.1 Linear perturbation analysis

To answer the previous question, we need to reduce our set of first order partial differential equations into a single second order equation. We first assume perturbations around some constant background state ρ_0, $\boldsymbol{v}_0$, P_0, according to

$$\rho = \rho_0 + \delta\rho, \tag{1.32}$$
$$\boldsymbol{v} = \boldsymbol{v}_0 + \delta\boldsymbol{v} = \delta\boldsymbol{v}, \tag{1.33}$$
$$P = P_0 + \delta P, \tag{1.34}$$

where in the second line we assumed $\boldsymbol{v}_0 = 0$, implying a background state at rest. We further define the ratio between the pressure and density perturbations according to

$$c_{\rm s}^2 \equiv \frac{\delta P}{\delta \rho}, \tag{1.35}$$

which has the dimensions of a speed squared, the meaning of which will become clear. For the moment this is just a definition. Perturbing the equations, we have

$$\frac{\partial}{\partial t}(\rho_0 + \delta\rho) + \delta\boldsymbol{v}\cdot\nabla(\rho_0 + \delta\rho) = -(\rho_0 + \delta\rho)\nabla\cdot(\delta\boldsymbol{v}), \tag{1.36}$$

$$(\rho_0 + \delta\rho)\left[\frac{\partial\delta\boldsymbol{v}}{\partial t} + (\delta\boldsymbol{v}\cdot\nabla)\delta\boldsymbol{v}\right] = -\nabla\delta P. \tag{1.37}$$

The derivatives $\partial\rho_0/\partial t$ and $\nabla\rho_0$ are zero because we assumed a constant background state. We then assume that perturbations to the background state ($\delta\rho$, $\delta\boldsymbol{v}$) are small, implying that terms involving two sets of perturbations, such as $\delta\boldsymbol{v}\cdot\nabla\delta\rho$ or $\delta\rho\nabla\cdot\delta\boldsymbol{v}$ are doubly small and hence negligible. This is obviously only true for a small perturbation to the density or velocity, but is *not* true in general. Linearising our equations this way allows us to solve them analytically, providing a useful insight into the underlying physics.

Neglecting second order terms, and using our definition (1.35) to relate δP to $\delta\rho$ we have

$$\frac{\partial\delta\rho}{\partial t} = -\rho_0\nabla\cdot(\delta\boldsymbol{v}), \tag{1.38}$$

$$\rho_0\frac{\partial\delta\boldsymbol{v}}{\partial t} = -c_{\rm s,0}^2\nabla\delta\rho. \tag{1.39}$$

Notice that we should also assume $c_{\rm s} = c_{\rm s,0} + \delta c_{\rm s}$, but that the $\delta c_{\rm s}$ term would be doubly small when multiplying the density perturbation. Hence why we write $c_{\rm s,0}$ in (1.39) and also brought it out the front of the gradient term.

We desire to obtain a single equation in just one of the variables $\delta\rho$ or $\delta\boldsymbol{v}$. We can eliminate the $\delta\boldsymbol{v}$ terms by taking $\partial/\partial t$ (1.38), and $\nabla\cdot$ (1.39). This gives

$$\frac{\partial^2\delta\rho}{\partial t^2} = -\rho_0\frac{\partial}{\partial t}\nabla\cdot(\delta\boldsymbol{v}), \tag{1.40}$$

$$\rho_0\nabla\cdot\frac{\partial\delta\boldsymbol{v}}{\partial t} = -c_{\rm s,0}^2\nabla^2\delta\rho. \tag{1.41}$$

Using the second equation on the right hand side of the first, we obtain

$$\frac{\partial^2\delta\rho}{\partial t^2} = c_{\rm s,0}^2\nabla^2\delta\rho. \tag{1.42}$$

You should immediately notice that this is just the wave equation (1.31) expressed in terms of the density perturbation. Hence small perturbations in density travel as *waves* with speed c_s, making the meaning of our definition (1.35) clear. Since physically these waves are sound waves, we refer to c_s as the *sound speed*.

We can also now answer the mystery: The Euler equations are *hyperbolic*. Also yes, penguins get cold, yes, and Les Murray was Hungarian.

1.4.2 Solving the wave equation

We can solve the wave equation by assuming solutions of the form

$$\delta\rho = De^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}, \tag{1.43}$$

where D is the (constant) perturbation amplitude, $\boldsymbol{k}$ is the wave vector and ω is the angular frequency. Taking time derivatives with the assumed form we find

$$\frac{\partial\delta\rho}{\partial t} = -i\omega De^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} = -i\omega\delta\rho, \tag{1.44}$$

$$\frac{\partial^2\delta\rho}{\partial t^2} = i^2\omega^2 De^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} = -\omega^2\delta\rho. \tag{1.45}$$

Similarly, for the spatial derivatives we have

$$\nabla\delta\rho = ikDe^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} = ik\delta\rho, \tag{1.46}$$

$$\nabla^2\delta\rho = i^2k^2De^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} = -k^2\delta\rho. \tag{1.47}$$

Using (1.45) and (1.47) in (1.42) we have

$$-\omega^2\delta\rho = -c_{\mathrm{s},0}^2k^2\delta\rho. \tag{1.48}$$

Dividing both sides by $-\delta\rho$ gives the so called *dispersion relation*

$$\omega^2 = c_\mathrm{s}^2k^2. \tag{1.49}$$

relating angular frequency to wavenumber. Since c_s is positive, the solutions assuming propagation in the x direction $\boldsymbol{k} = [k_x, 0, 0]$ are given by

$$\omega = \pm c_\mathrm{s}k_x, \tag{1.50}$$

giving solutions for the density perturbation in the form of *travelling waves*

$$\delta\rho = D\exp[ik_x(x \pm c_\mathrm{s}t)] = D\cos[k_x(x \pm c_\mathrm{s}t) + \phi_0], \tag{1.51}$$

where ϕ_0 is some arbitrary initial phase. That is, the solution is just a sinusoidal perturbation translated, or *travelling*, to the left or right at speed c_s.

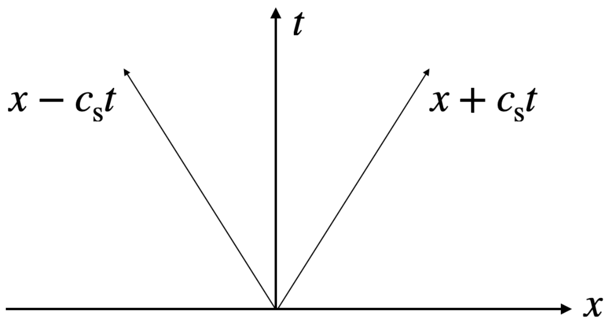

Figure 1.1: Characteristic curves for the wave equation, assuming a perturbation initially at $x = 0$. The density perturbation is constant along characteristics.

1.4.3 Characteristics

Another way to think about the solution (1.51) is that in a coordinate system with $x' = x + c_s t$ or $x' = x - c_s t$, or equivalently where the observer is moving with $\mathrm{d}x'/\mathrm{d}t = \pm c_s$, the density perturbation would remain unchanged. These curves (Figure 1.1) are known as the *characteristics* of the wave equation and are a defining feature of hyperbolic partial differential equations.

1.4.4 Sound speed in an adiabatic gas

If we assume an adiabatic equation of state $P = (\gamma - 1)\rho u$ with no change in entropy (i.e. $\delta u = P/\rho^2 \delta\rho$) we have

$$\delta P = (\gamma - 1) u \delta\rho + (\gamma - 1)\rho\delta u, \tag{1.52}$$

$$= \left[\frac{P}{\rho} + (\gamma - 1)\frac{P}{\rho}\right]\delta\rho, \tag{1.53}$$

$$= \frac{\gamma P}{\rho}\delta\rho, \tag{1.54}$$

Since we defined the sound speed via $c_s^2 \equiv \delta P/\delta\rho$, this implies that the sound speed in an adiabatic gas is given by

$$c_s = \sqrt{\frac{\gamma P}{\rho}}. \tag{1.55}$$

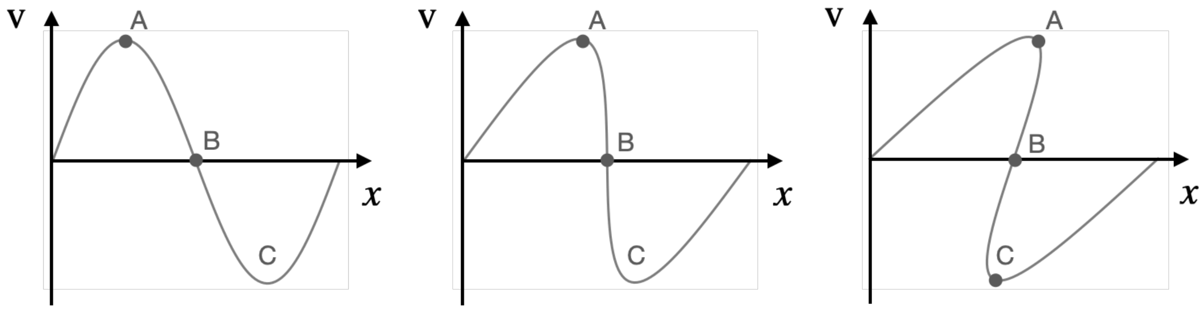

Figure 1.2: Non-linear behaviour of Burgers equation (1.57). Each Lagrangian observer A, B and C simply retains their initial velocity, which is positive for A, zero for B and negative for C. This results in steepening of the wave profile (centre panel) and ultimately a double-valued velocity field (right panel). Since the latter is unphysical in gas the end result is the formation of a discontinuity or shock wave.

1.5 Non-linear solutions

1.5.1 Steepening and shock formation

The best way to think about the non-linear behaviour of the fluid equations is to consider the terms in (1.5)–(1.7) that we neglected during our linear analysis. In particular, the most important term we dropped in the linear analysis is the $(\boldsymbol{v} \cdot \nabla)\boldsymbol{v}$ term. We can consider this term alone while neglecting the other terms by considering the simpler equation

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v} = 0. \tag{1.56}$$

This is known as *Burgers equation*. This equation becomes trivial when expressed in terms of the Lagrangian derivative, giving simply

$$\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = 0. \tag{1.57}$$

That is, the velocity is constant for an observer moving with $\mathrm{d}\boldsymbol{x}/\mathrm{d}t = \boldsymbol{v}$. This should already seem familiar, it is the same statement that we have seen for our wave solution, indicating that the velocity is constant along the characteristics of Burgers equation. We refer to quantities that are constant along characteristics as *Riemann invariants*.

We can visualise the non-linear behaviour by considering, as previously, a sinusoidal perturbation (Figure 1.2, left). Consider three Lagrangian observers A, B and C located at the maximum, zero and minimum of the sine wave, respectively. Since each observer simply maintains their initial velocity, one may observe that the wave starts to change shape, or *steepen* (Figure 1.2, centre). At some later time A and C will overtake each

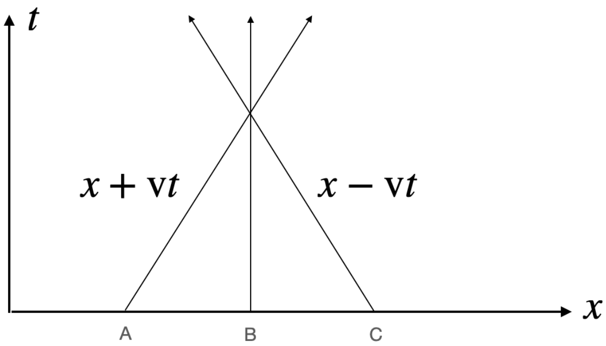

Figure 1.3: Characteristics for the example shown in Figure 1.2. The point at which characteristics cross corresponds to the formation of a shock if the fluid is collisional.

other resulting in a double-valued velocity field (Figure 1.2, right). Whether or not this occurs in practice depends on the microphysics of the fluid. In a gas or liquid at this point molecules would physically collide, resulting in dissipation and loss of energy. But there are plenty of situations where the double valued solution is the right one, namely when the fluid is *collisionless*. Examples of collisionless fluids are stars in a galaxy[4], cold dark matter, or large dust grains in protoplanetary discs. In this case one would almost always model the fluid as we have just done, namely with Lagrangian particles.

Figure 1.3 shows the same situation expressed in terms of characteristics. The point at which characteristics cross corresponds to formation of a shock in the gas, which is just a fancy name for a solution which is *discontinuous*. The problem with discontinuities is that it violates our assumption of having *differential equations*. That is, at the point of shock formation, the derivatives right hand sides of (1.5)–(1.7) become infinite, and our equations are no longer solvable in their present form! How can we proceed?

1.5.2 Integral vs differential form

The key is to remember that our set of differential equations is not the whole story. These equations just express higher principles, namely the conservation of mass, momentum and energy. We can apply these same principles to find solutions even when our differential equations would seem to fail us. Essentially we need to *integrate* our equations to remove the spatial derivatives. We can achieve this by writing our equations in the general 'conservative' form

$$\frac{\partial}{\partial t}(\text{thing}) + \nabla \cdot (\textbf{flux of thing}) = 0. \tag{1.58}$$

[4]Galaxies are mostly empty space. So when galaxies collide the stars themselves do not actually collide, the galaxies just pass right through each other while exerting mutual gravitational forces. Nevertheless there are so many stars in a galaxy it is valid to model them as a continuous density field, i.e. as a collisionless fluid.

We already saw that the continuity equation can be written in this form, namely

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{v}) = 0, \tag{1.59}$$

where ρ is the mass per unit volume and hence $\rho \boldsymbol{v}$ is the mass flux through a unit volume. Integrating this equation over a volume V gives

$$\frac{\partial}{\partial t} \int_V \rho \mathrm{d}V + \int_V \nabla \cdot (\rho \boldsymbol{v}) \mathrm{d}V = 0. \tag{1.60}$$

We can then use Gauss' theorem to write the second term as a surface integral, giving

$$\frac{\partial}{\partial t} \int_V \rho \mathrm{d}V = -\oint_{\partial V} \rho \boldsymbol{v} \cdot \mathrm{d}S. \tag{1.61}$$

Physically this simply expresses the fact that mass in a volume only changes because of the mass flux in or out of the bounding surface. The key consideration for us is that there are no spatial derivatives in (1.61), so the equation expressed in this 'integral form' should have no problems with discontinuous solutions, unlike (1.59).

1.5.3 Euler equations in conservation form

Applying the same logic to the momentum and energy equations, we find that (1.11)–(1.13) written in conservation form become

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{v}) = 0, \tag{1.62}$$

$$\frac{\partial (\rho \boldsymbol{v})}{\partial t} + \nabla \cdot (P\boldsymbol{I} + \rho \boldsymbol{v}\boldsymbol{v}) = 0, \tag{1.63}$$

$$\frac{\partial (\rho e)}{\partial t} + \nabla \cdot [(\rho e + P)\boldsymbol{v}] = 0, \tag{1.64}$$

where $e \equiv \frac{1}{2}v^2 + u$ is the specific total energy and $\boldsymbol{I}$ is the identity matrix. Notice that the quantity $\boldsymbol{v}\boldsymbol{v}$ in the momentum equation is actually a tensor rather than a vector[5].

[5] We have adopted dyadic notation '$\boldsymbol{v}\boldsymbol{v}$' but often it becomes convenient to switch to tensor notation. An ugly alternative is to use outer product notation, $\boldsymbol{v} \otimes \boldsymbol{v}$. The choice is a matter of style and convenience, but don't ask me to referee your papers if you choose the latter.

Using Gauss' theorem on the flux terms, our equations in integral form are given by

$$\frac{\partial}{\partial t}\int_V \rho \, \mathrm{d}V = -\oint_{\partial V} \rho \boldsymbol{v} \cdot \mathrm{d}S, \tag{1.65}$$

$$\frac{\partial}{\partial t}\int_V \rho \boldsymbol{v} \, \mathrm{d}V = -\oint (P\boldsymbol{I} + \rho \boldsymbol{v}\boldsymbol{v}) \cdot \mathrm{d}S, \tag{1.66}$$

$$\frac{\partial}{\partial t}\int_V \rho e \, \mathrm{d}V = -\oint (\rho e + P)\boldsymbol{v} \cdot \mathrm{d}S. \tag{1.67}$$

1.5.4 Shock jump conditions

The simplest case is a stationary shock, for which the left hand side of the equations are zero. In this case we just have constant flux across the discontinuity and simply need to match conditions on either side of the jump. If we consider a one dimensional discontinuous jump in ρ, v_x and u, with region "1" on one side of the jump, and region "2" on the other side, then the conditions are

$$\rho_1 v_1 = \rho_2 v_2, \tag{1.68}$$

$$P_1 + \rho_1 v_1^2 = P_2 + \rho_2 v_2^2, \tag{1.69}$$

$$\rho_1 \left(\frac{1}{2}v_1^2 + u_1 + \frac{P_1}{\rho_1}\right) v_1 = \rho_2 \left(\frac{1}{2}v_2^2 + u_2 + \frac{P_2}{\rho_2}\right) v_2. \tag{1.70}$$

1.5.5 Adiabatic shocks

Assuming an adiabatic equation of state $P = (\gamma - 1)\rho u$ and using (1.68) gives the energy jump condition (1.70) in the form

$$\frac{1}{2}v_1^2 + \frac{\gamma P_1}{(\gamma - 1)\rho_1} = \frac{1}{2}v_2^2 + \frac{\gamma P_2}{(\gamma - 1)\rho_2}. \tag{1.71}$$

We can then combine equations to try to solve for the density jump across the shock. From (1.68) we have

$$v_2^2 = \frac{\rho_1^2}{\rho_2^2} v_1^2. \tag{1.72}$$

Inserting this into (1.69) gives

$$P_1 + \rho_1 v_1^2 = P_2 + \frac{\rho_1^2}{\rho_2} v_1^2. \tag{1.73}$$

and hence

$$P_2 = P_1 + \rho_1 v_1^2 \left(1 - \frac{\rho_1}{\rho_2}\right). \tag{1.74}$$

Finally, inserting this into (1.70) gives

$$\frac{1}{2}v_1^2 + \frac{\gamma P_1}{(\gamma-1)\rho_1} = \frac{1}{2}\frac{\rho_1^2}{\rho_2^2}v_1^2 + \frac{\gamma}{(\gamma-1)\rho_2}\left[P_1 + \rho_1 v_1^2\left(1 - \frac{\rho_1}{\rho_2}\right)\right], \tag{1.75}$$

$$\therefore \frac{1}{2}v_1^2\left(1 - \frac{\rho_1^2}{\rho_2^2}\right) + \frac{c_{s,1}^2}{(\gamma-1)} = \frac{\rho_1}{\rho_2(\gamma-1)}c_{s,1}^2 + \frac{\gamma}{(\gamma-1)}\frac{\rho_1}{\rho_2}v_1^2\left(1 - \frac{\rho_1}{\rho_2}\right). \tag{1.76}$$

Multiplying by $2(\gamma - 1)$ and collecting terms gives

$$(\gamma-1)v_1^2\left(1 - \frac{\rho_1}{\rho_2}\right)\left(1 + \frac{\rho_1}{\rho_2}\right) + 2c_{s,1}^2\left(1 - \frac{\rho_1}{\rho_2}\right) = \frac{2\gamma\rho_1}{\rho_2}v_1^2\left(1 - \frac{\rho_1}{\rho_2}\right). \tag{1.77}$$

Cancelling the common factor, dividing by $c_{s,1}^2$ and defining the upstream Mach number $\mathcal{M}_1^2 = v_1^2/c_{s,1}^2$ we find

$$\mathcal{M}_1^2(\gamma-1)\left(1 + \frac{\rho_1}{\rho_2}\right) + 2 = \frac{2\gamma\rho_1}{\rho_2}\mathcal{M}_1^2,$$

$$\therefore \mathcal{M}_1^2(\gamma-1) + 2 = \frac{\rho_1}{\rho_2}\mathcal{M}_1^2\left[2\gamma - (\gamma-1)\right], \tag{1.78}$$

$$= \frac{\rho_1}{\rho_2}\mathcal{M}_1^2(\gamma+1), \tag{1.79}$$

giving our final expression for the density jump in the form

$$\frac{\rho_2}{\rho_1} = \frac{\mathcal{M}_1^2(\gamma+1)}{2 + (\gamma-1)\mathcal{M}_1^2}. \tag{1.80}$$

For a very strong shock $\mathcal{M}_1 \gg 1$ and we have

$$\frac{\rho_2}{\rho_1} = \frac{(\gamma+1)}{2/\mathcal{M}_1^2 + (\gamma-1)} \rightarrow \frac{(\gamma+1)}{(\gamma-1)} \text{ as } \mathcal{M}_1 \rightarrow \infty. \tag{1.81}$$

For a monatomic gas ($\gamma = 5/3$) this implies that the maximum density jump for an infinite strength shock is

$$\frac{\rho_2}{\rho_1} = \frac{(\frac{5}{3}+1)}{(\frac{5}{3}-1)} = 4. \tag{1.82}$$

For a diatomic gas such as air, $\gamma = 1.4$ and hence $\rho_2/\rho_1 \rightarrow 6$ as $\mathcal{M}_1 \rightarrow \infty$.

A little more algebra shows that one can also write down the pressure and temperature jumps according to

$$\frac{P_2}{P_1} = \frac{2\gamma\mathcal{M}_1^2 - (\gamma - 1)}{\gamma + 1}, \tag{1.83}$$

and

$$\frac{T_2}{T_1} = \frac{[2\gamma\mathcal{M}_1^2 - (\gamma - 1)]\,[2 + (\gamma - 1)\mathcal{M}_1^2]}{(\gamma + 1)^2\mathcal{M}_1^2}. \tag{1.84}$$

Taking the limit $\mathcal{M}_1 \to \infty$ shows that both $P_2/P_1 \to \infty$ and $T_2/T_1 \to \infty$ despite the limited jump in density. Physically the limited density jump occurs because we have assumed that all of the heat generated by the shock (more on this below) is trapped. It is a completely different situation for a shock that is allowed to cool (see Section 1.5.7).

1.5.6 Shocks and irreversibility

Notice in particular that the change in kinetic energy, from (1.68), is given by

$$\frac{\frac{1}{2}v_2^2}{\frac{1}{2}v_1^2} = \frac{1}{16}. \tag{1.85}$$

This shows that kinetic energy is lost across the jump. In an adiabatic shock this must be converted to heat (u) since total energy is conserved. The curious thing about this is that it implies an *irreversible*, dissipative process. Yet we started with a set of equations defined to have no dissipation in them. Another way to think about this is that we assumed *no* irreversible processes — our differential form of the energy equation (1.13) implies zero change in entropy! But it is clear from (1.85) that the entropy must change across a shock. So how can irreversibility arise from fundamentally reversible and dissipationless equations? You may also recall we did not consider any explicit viscosity or other dissipation when formulating our equations…

Why a reversible set of equations becomes irreversible

How irreversibility arises in the fluid equations is subtle but easily understandable. Figure 1.2 shows what would happen in a situation where there was no physical dissipation at all — the solution would just become double-valued. The irreversibility arises when we mandate that the velocity field must remain single-valued. Physically this occurs in a gas because molecules *actually collide*, which is an irreversible process. Such collisions produce viscosity, but on large scales we can safely ignore this macroscopic viscosity. The issue in a shock is that there is an infinitely short length scale involved. What matters is that viscosity occurs on *some* length scale, usually far below the resolution scale in simulations (this is also why the exact details of the shock capturing scheme are less important than having *some form of dissipation* applied at the shock front). The weird thing is that it does not matter *how* the viscosity occurs, there just needs to be *some* dissipation on *some* length scale. Conservation of energy across a shock jump requires it!

One may also think about where the entropy change arises from. When mandating a single-valued velocity field, we effectively have information loss at the shock front, and can no longer evolve the solution backwards in time to obtain our initial conditions. Information loss and irreversibility both imply an increase in entropy.

1.5.7 Isothermal shocks

For an isothermal shock the jump conditions are

$$\rho_2 v_2 = \rho_1 v_1, \tag{1.86}$$

$$\rho_2 v_2^2 + P_2 = \rho_1 v_1^2 + P_1, \tag{1.87}$$

$$T_2 = T_1 \quad \text{or} \quad c_\mathrm{s}^2 = \text{const}, \tag{1.88}$$

Proceeding as previously, the first equation gives $v_2^2 = \rho_1^2 v_1^2/\rho_2^2$. Using this in second expression and using $P_2 = c_\mathrm{s}^2 \rho_2$ and $P_1 = c_\mathrm{s}^2 \rho_1$ gives

$$\frac{\rho_1^2}{\rho_2} v_1^2 + c_\mathrm{s}^2 \rho_2 = \rho_1 v_1^2 + c_\mathrm{s}^2 \rho_1. \tag{1.89}$$

Dividing by $c_\mathrm{s}^2 \rho_2$ and defining $\mathcal{M}_1^2 \equiv v_1^2/c_\mathrm{s}^2$ gives

$$\frac{\rho_1}{\rho_2} \mathcal{M}_1^2 \left(\frac{\rho_1}{\rho_2} - 1 \right) = \left(\frac{\rho_1}{\rho_2} - 1 \right), \tag{1.90}$$

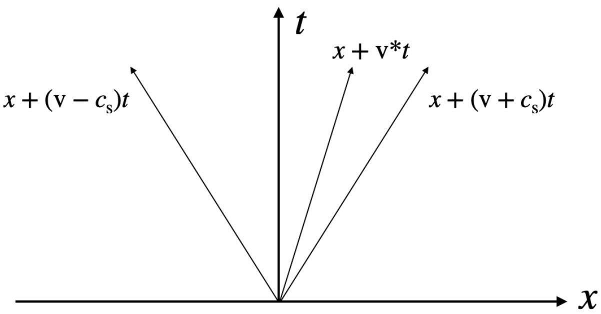

Figure 1.4: Characteristic curves for the fluid equations in 1D, assuming a perturbation initially at $x = 0$. A disturbance excites a left-going sound wave and a right-going sound wave. A contact discontinuity propagates at the post-shock speed.

where cancelling the factor in brackets gives

$$\frac{\rho_1}{\rho_2}\mathcal{M}_1^2 = 1, \tag{1.91}$$

and therefore

$$\frac{\rho_2}{\rho_1} = \mathcal{M}_1^2. \tag{1.92}$$

Hence in this case, as $\mathcal{M}_1 \to \infty$ then the density jump is also infinite. Physically this is because the fluid is allowed to radiate all of the heat that is generated at the shock front if we assume an isothermal equation of state.

It can be useful to consider adiabatic and isothermal equations of state to model the two extremes in any realistic astrophysical environment — adiabatic assumes all heat is trapped while isothermal assumes all heat is radiated. Just like in politics, the truth is in between.

1.5.8 Riemann invariants

A final question is to ask: What are the characteristics and Riemann invariants for the full set of fluid equations? That is, not just for Burgers equation or the linearised equations. We start with the fluid equations in the form

$$\frac{\partial \rho}{\partial t} + (\boldsymbol{v}\cdot\nabla)\rho + \rho(\nabla\cdot\boldsymbol{v}) = 0, \tag{1.93}$$

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{v} + \frac{\nabla P}{\rho} = 0. \tag{1.94}$$

Using $P = K\rho^\gamma$ and $c_\mathrm{s}^2 = \gamma P/\rho$ we have

$$\frac{1}{\rho}\frac{\partial \rho}{\partial t} = \frac{2}{\gamma - 1}\frac{1}{c_\mathrm{s}}\frac{\partial c_\mathrm{s}}{\partial t}, \tag{1.95}$$

and similarly

$$\frac{\nabla \rho}{\rho} = \frac{2}{\gamma - 1}\frac{\nabla c_\mathrm{s}}{c_\mathrm{s}}. \tag{1.96}$$

Using these expressions in (1.93) and (1.94) we obtain

$$\frac{\partial}{\partial t}\left(\frac{2}{\gamma - 1}c_\mathrm{s}\right) + \boldsymbol{v}\cdot\nabla\left(\frac{2}{\gamma - 1}c_\mathrm{s}\right) - c_\mathrm{s}(\nabla\cdot\boldsymbol{v}) = 0, \tag{1.97}$$

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{v} + c_\mathrm{s}\nabla\left(\frac{2}{\gamma - 1}c_\mathrm{s}\right) = 0. \tag{1.98}$$

Adding Equations (1.97) and (1.98), assuming propagation in the x direction gives

$$\frac{\partial}{\partial t}\left(v_x + \frac{2}{\gamma - 1}c_\mathrm{s}\right) + (v_x + c_\mathrm{s})\frac{\partial}{\partial x}\left(v_x + \frac{2}{\gamma - 1}c_\mathrm{s}\right) = 0, \tag{1.99}$$

while subtracting (1.97) from (1.98) gives

$$\frac{\partial}{\partial t}\left(v_x - \frac{2}{\gamma - 1}c_\mathrm{s}\right) + (v_x - c_\mathrm{s})\frac{\partial}{\partial x}\left(v_x - \frac{2}{\gamma - 1}c_\mathrm{s}\right) = 0. \tag{1.100}$$

If we then define the quantities

$$Q \equiv v_x + \frac{2}{\gamma - 1}c_\mathrm{s}, \tag{1.101}$$

$$R \equiv v_x - \frac{2}{\gamma - 1}c_\mathrm{s}, \tag{1.102}$$

then our equations become

$$\frac{\partial Q}{\partial t} + (v_x + c_\mathrm{s})\frac{\partial Q}{\partial x} = 0, \tag{1.103}$$

$$\frac{\partial R}{\partial t} + (v_x - c_\mathrm{s})\frac{\partial R}{\partial x} = 0, \tag{1.104}$$

which are equivalent to

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = 0; \qquad \frac{\mathrm{d}R}{\mathrm{d}t} = 0, \tag{1.105}$$

for observers moving with $\mathrm{d}x/\mathrm{d}t = (v_x + c_\mathrm{s})$ and $\mathrm{d}x/\mathrm{d}t = (v_x - c_\mathrm{s})$, respectively.

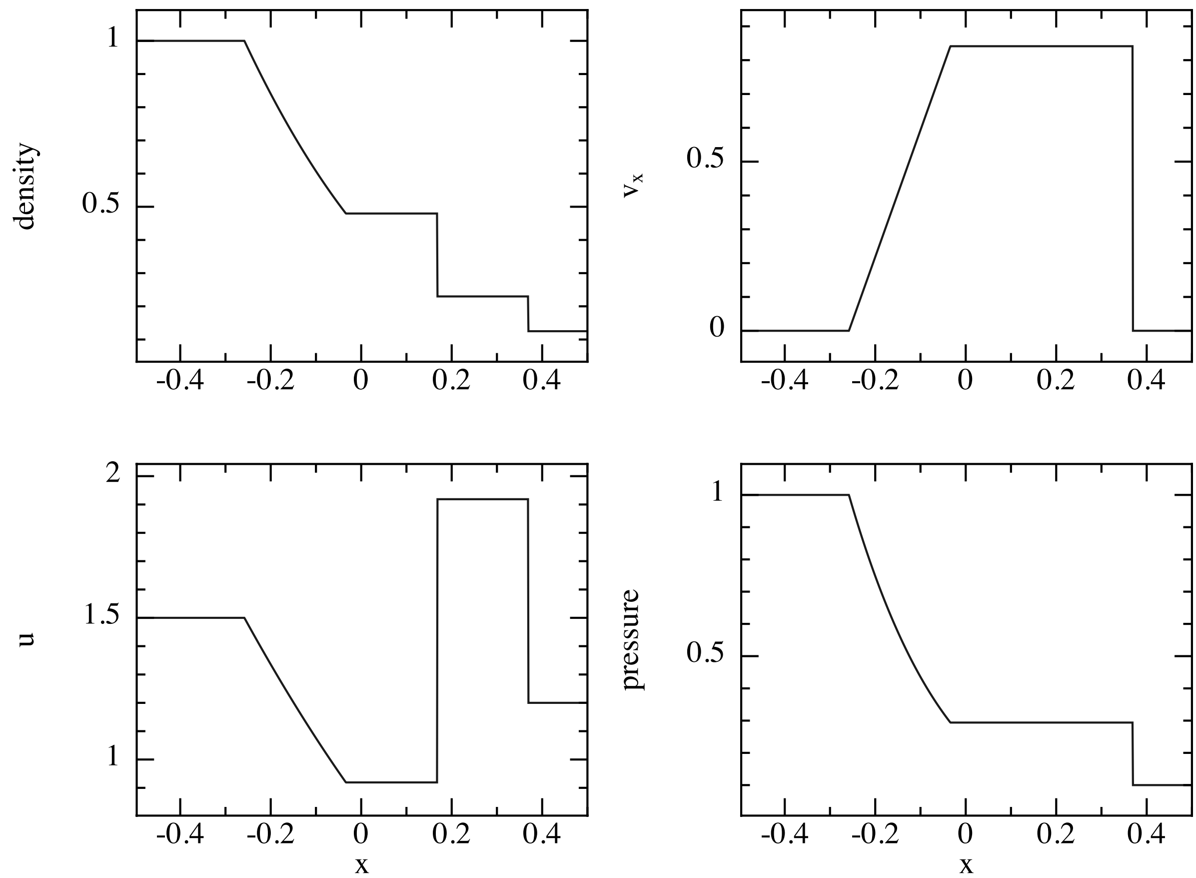

Figure 1.5: Exact solution for a hydrodynamic shock, where initially there was is a high pressure, high density region for $x < 0$ and a low pressure, low density region for $x \geq 0$, with a discontinuous jump at $x = 0$. A shock propagates into the undisturbed medium as the right-going sound wave, the left-going sound wave is seen as a rarefaction wave. A contact discontinuity — a jump in density and internal energy at constant pressure — propagates at the post-shock speed.

This shows that the quantities R and Q are constant, or *invariant* along characteristics. Hence these are the Riemann invariants for the fluid equations. Figure 1.4 shows the corresponding characteristics. Figure 1.5 shows an example of a 'shock tube' solution to the fluid equations, evolving from an initial discontinuous jump in density and pressure placed at $x = 0$. The three characteristics are evident in the structure of the solution — showing left and right-propagating sound waves and a contact discontinuity.

2

Magnetohydrodynamics

2.1 Equations of magnetohydrodynamics

Plenty of gas in space is ionised, hence we need to consider the effects of electric and magnetic fields generated by the movement of charged ions. Magnetic effects are important in many astrophysical phenomena, to the point where it's a long-standing joke in astronomy talks to ask "what about magnetic fields?".

We begin with Maxwell's equations of electromagnetism (Maxwell, 1865):

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \tag{2.1}$$

$$\nabla \times \boldsymbol{B} = \mu_0 \boldsymbol{J} + \frac{1}{c^2}\frac{\partial \boldsymbol{E}}{\partial t}, \tag{2.2}$$

$$\nabla \cdot \boldsymbol{E} = \frac{\rho_c}{\epsilon_0}, \tag{2.3}$$

$$\nabla \cdot \boldsymbol{B} = 0. \tag{2.4}$$

In all but the most extreme environments, we can consider the non-relativistic limit of Maxwell's equations ($v \ll c$) in order to neglect Maxwell's famous displacement current term in (2.2). But one should already apply some caution here: there are *two* non-relativistic limits, the *magnetic limit* where $v \ll c$ and $cE \ll B$ and the *electric limit* where $v \ll c$ and $cE \gg B$ (see e.g. Le Bellac and Lévy-Leblond 1973).

2.1.1 Deriving the MHD equations — a simple way

To derive the induction equation the traditional way[1] we consider the magnetic limit, where the Lorentz transformations transforming $\boldsymbol{E}$ and $\boldsymbol{B}$ to a co-moving frame become

$$\boldsymbol{E}' = \boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}, \tag{2.5}$$

$$\boldsymbol{B}' = \boldsymbol{B}. \tag{2.6}$$

In this limit we can neglect the displacement current term to give Ampére's law

$$\nabla \times \boldsymbol{B} = \mu_0 \boldsymbol{J}. \tag{2.7}$$

The usual procedure to derive the equations of magnetohydrodynamics (MHD) is then to consider a 'generalised Ohm's law' in the comoving frame given by

$$\boldsymbol{J}' = \sigma \boldsymbol{E}', \tag{2.8}$$

where σ is the electrical conductivity. Inserting these in (2.1) we have

$$\frac{\partial \boldsymbol{B}'}{\partial t} = -\nabla \times (\boldsymbol{E}' - \boldsymbol{v} \times \boldsymbol{B}). \tag{2.9}$$

Then from the generalised Ohm's law giving $\boldsymbol{E}' = \frac{\boldsymbol{J}'}{\sigma}$, we we find an evolution equation for the magnetic field in the form

$$\frac{\partial \boldsymbol{B}}{\partial t} = \nabla \times (\boldsymbol{v} \times \boldsymbol{B}) - \nabla \times \left(\frac{\boldsymbol{J}'}{\sigma} \right). \tag{2.10}$$

We can simplify the second term using Ampere's law (2.7) and defining the Ohmic resistivity $\eta \equiv 1/(\mu_0 \sigma)$, giving

$$\frac{\partial \boldsymbol{B}}{\partial t} = \nabla \times (\boldsymbol{v} \times \boldsymbol{B}) - \nabla \times (\eta \nabla \times \boldsymbol{B}). \tag{2.11}$$

This is known as the *induction equation*. In the limit of perfect conductivity $\sigma \to \infty$, known as *ideal magnetohydrodynamics*, we have simply

$$\frac{\partial \boldsymbol{B}}{\partial t} = \nabla \times (\boldsymbol{v} \times \boldsymbol{B}). \tag{2.12}$$

This gives the equation needed to update the magnetic field given the existing magnetic field and the fluid velocity. We still need an equation giving the force on the fluid — often this is derived by simply stating that the force on the fluid is equal to $\boldsymbol{J} \times \boldsymbol{B}$, but we can do better than this, as we will see below.

[1] we call something traditional in astrophysics when we are about to trash it...

2.1.2 Deriving the MHD equations — the right way

The above procedure is not terribly satisfactory. While physically plausible, we plucked the generalised Ohm's law out of thin air, and one has to be careful with transforming quantities to and from the co-moving frame. A better way (e.g. Pandey and Wardle, 2008) is to begin by describing a two-fluid mixture consisting of equally (but oppositely) charged ions (i) and electrons (e). In this case conservation of momentum is expressed by two separate equations for momentum conservation in each component of the fluid. In conservative form, we have

$$\frac{\partial}{\partial t}(\rho_e \boldsymbol{v}_e) + \nabla\cdot(\rho_e \boldsymbol{v}_e\boldsymbol{v}_e + P_e\boldsymbol{I}) = -en_e\left(\boldsymbol{E} + \boldsymbol{v}_e\times\boldsymbol{B}\right) + K(\boldsymbol{v}_i - \boldsymbol{v}_e), \quad (2.13)$$

$$\frac{\partial}{\partial t}(\rho_i \boldsymbol{v}_i) + \nabla\cdot(\rho_i \boldsymbol{v}_i\boldsymbol{v}_i + P_i\boldsymbol{I}) = +en_i\left(\boldsymbol{E} + \boldsymbol{v}_i\times\boldsymbol{B}\right) - K(\boldsymbol{v}_i - \boldsymbol{v}_e), \quad (2.14)$$

where $\rho_{e,i}$ and $\boldsymbol{v}_{e,i}$ are the electron and ion densities and velocities, respectively. Each fluid feels the electromagnetic force $\boldsymbol{F} = q_c(\boldsymbol{E} + \boldsymbol{v}\times\boldsymbol{B})$ where q_c is the charge density, proportional to the number density $n_{e,i}$ and the charge on the species, either $+e$ or $-e$ in this case, where e is the electron charge. We have assumed that the ions and electrons are coupled by a collisional term, with coefficient K. We will see a simpler example of this kind of coupling for dust-gas mixtures in Chapter 3.

Equations of motion

We can then add these equations assuming charge neutrality ($n_e = n_i$) to give

$$\frac{\partial}{\partial t}(\rho\boldsymbol{v}) + \nabla\cdot(\rho\boldsymbol{v}\boldsymbol{v} + P\boldsymbol{I}) = en_e(\boldsymbol{v}_i - \boldsymbol{v}_e)\times\boldsymbol{B} - \nabla\cdot(\rho\boldsymbol{w}_i\boldsymbol{w}_i + \rho\boldsymbol{w}_e\boldsymbol{w}_e) \quad (2.15)$$

where we have defined the velocity of the total fluid mixture, known as the *barycentric* velocity, according to $\boldsymbol{v} = (\rho_e\boldsymbol{v}_e + \rho_i\boldsymbol{v}_i)/\rho$, where $\rho \equiv \rho_e + \rho_i$ is the total mass density, and also defined drift velocities with respect to the barycentre according to $\boldsymbol{w}_i \equiv \boldsymbol{v}_i - \boldsymbol{v}$, $\boldsymbol{w}_e \equiv \boldsymbol{v}_e - \boldsymbol{v}$. Writing $\boldsymbol{J} = en_e(\boldsymbol{v}_i - \boldsymbol{v}_e)$ and assuming that the drift velocities are small, i.e. that $\boldsymbol{v} \approx \boldsymbol{v}_i$ — reasonable since in general $\rho_e \ll \rho_i$ — we find a momentum equation given by

$$\frac{\partial}{\partial t}(\rho\boldsymbol{v}) + \nabla\cdot(\rho\boldsymbol{v}\boldsymbol{v} + P\boldsymbol{I}) = \boldsymbol{J}\times\boldsymbol{B}. \quad (2.16)$$

This is identical to the usual momentum equation (1.3) with an additional $\boldsymbol{J}\times\boldsymbol{B}$ force perpendicular to the magnetic field direction, as stated above but here derived clearly by summing the individual forces on ions and electrons.

Induction equation and non-ideal MHD

We can then derive the induction equation using the remaining information by writing the electron momentum equation (2.13) in Lagrangian form to give

$$\frac{\mathrm{d}\boldsymbol{v}_e}{\mathrm{d}t} = -\frac{\nabla P_e}{\rho_e} - \frac{e}{m_e}\left(\boldsymbol{E} + \boldsymbol{v}_e \times \boldsymbol{B}\right) + \frac{K}{\rho_e}(\boldsymbol{v}_i - \boldsymbol{v}_e). \tag{2.17}$$

If we then neglect electron inertia by assuming that $\mathrm{d}\boldsymbol{v}_e/\mathrm{d}t \approx 0$ we can then obtain an expression for $\boldsymbol{E}$ in the form

$$\boldsymbol{E} = -\frac{\nabla P_e}{e n_e} - \boldsymbol{v}_e \times \boldsymbol{B} + \frac{K m_e}{e}(\boldsymbol{v}_i - \boldsymbol{v}_e), \tag{2.18}$$

giving, using $\boldsymbol{v}_e = \boldsymbol{v}_i - \boldsymbol{J}/(e n_e)$

$$\boldsymbol{E} = -\frac{\nabla P_e}{e n_e} - \boldsymbol{v}_i \times \boldsymbol{B} + \frac{\boldsymbol{J} \times \boldsymbol{B}}{e n_e} + \frac{\boldsymbol{J}}{\sigma}, \tag{2.19}$$

where we have defined the conductivity $\sigma \equiv e^2 n_e/(m_e K)$. This is a more general version of the 'Ohm's law' we plucked from the sky in the previous derivation. Our more general form of the induction equation is thus obtained by substituting (2.19) into Maxwell's equation (2.1) to give, assuming as previously that $\boldsymbol{v} \approx \boldsymbol{v}_i$,

$$\frac{\partial \boldsymbol{B}}{\partial t} = \nabla \times (\boldsymbol{v} \times \boldsymbol{B}) - \nabla \times \left[\frac{\boldsymbol{J}}{\sigma} + \frac{\boldsymbol{J} \times \boldsymbol{B}}{e n_e} + \frac{\nabla P_e}{e n_e}\right]. \tag{2.20}$$

The terms in square brackets are known as *non-ideal MHD* terms and are important when the fluid is only partially ionised. They correspond to Ohmic dissipation, the Hall effect, and the Biermann battery term, respectively. We will not consider these terms further in this crash course, but you should know that they exist. They are important in solar physics and also in star formation, since molecular clouds are weakly ionised. In general one should also consider a third, neutral component to the fluid mixture which gives rise to an additional term related to ion-neutral drift which in astrophysics is referred to as *ambipolar diffusion* (Beware: in plasma physics ambipolar diffusion refers to the effect of electron inertia, which is different).

The Biermann battery

The Biermann battery, after Biermann (1950) is so-named because it is the only term to generate a magnetic field when there is none already present, hence useful in studies of magnetogenesis in the Universe. However, this term is problematic to include in numerical codes because including the electron pressure gradient directly in this manner does not guarantee positive definite contribution to the entropy, resulting in 'the Biermann catastrophe in numerical MHD' (Graziani et al., 2015).

2.1.3 Induction equation in Lagrangian form

There are many possible ways to write the induction equation. To write this equation with a Lagrangian time derivative we can expand the term on the right hand side using vector identities to give

$$\frac{\partial \boldsymbol{B}}{\partial t} = \nabla \times (\boldsymbol{v} \times \boldsymbol{B}) = \boldsymbol{v}(\nabla \cdot \boldsymbol{B}) - \boldsymbol{B}(\nabla \cdot \boldsymbol{v}) + (\boldsymbol{B} \cdot \nabla)\boldsymbol{v} - (\boldsymbol{v} \cdot \nabla)\boldsymbol{B}. \tag{2.21}$$

Using the definition of the Lagrangian time derivative $\mathrm{d}\boldsymbol{B}/\mathrm{d}t \equiv \partial \boldsymbol{B}/\partial t + (\boldsymbol{v} \cdot \nabla)\boldsymbol{B}$ and assuming $\nabla \cdot \boldsymbol{B} = 0$ from (2.4) we find

$$\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t} = (\boldsymbol{B} \cdot \nabla)\,\boldsymbol{v} - \boldsymbol{B}\,(\nabla \cdot \boldsymbol{v})\,. \tag{2.22}$$

We can simplify further by writing the equation in terms of $\boldsymbol{B}/\rho$ using the continuity equation, giving

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\boldsymbol{B}}{\rho}\right) = \left(\frac{\boldsymbol{B}}{\rho} \cdot \nabla\right)\boldsymbol{v}. \tag{2.23}$$

A few other possible formulations are given in Section 2.4 in the context of trying to satisfy the $\nabla \cdot \boldsymbol{B}$ constraint in numerical codes.

2.1.4 MHD equations in Lagrangian form

In summary, we can write the equations of ideal MHD in Lagrangian form according to

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\rho(\nabla \cdot \boldsymbol{v}), \tag{2.24}$$

$$\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = -\frac{\nabla P}{\rho} + \frac{\boldsymbol{J} \times \boldsymbol{B}}{\rho}, \tag{2.25}$$

$$\frac{\mathrm{d}u}{\mathrm{d}t} = -\frac{P}{\rho}(\nabla \cdot \boldsymbol{v}), \tag{2.26}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\boldsymbol{B}}{\rho}\right) = \left(\frac{\boldsymbol{B}}{\rho} \cdot \nabla\right)\boldsymbol{v}. \tag{2.27}$$

Comparison to (1.5)–(1.7) shows they are the same as the equations of hydrodynamics but with an additional $\boldsymbol{J} \times \boldsymbol{B}$ force on the fluid due to the magnetic field, and an additional equation derived from Maxwell's equations governing the magnetic field evolution.

2.1.5 Magnetic force as the gradient of a stress tensor

We can also write the magnetic force as the gradient of a stress tensor. Using Ampére's law (2.7) we can expand the magnetic force to give

$$\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = -\frac{\nabla P}{\rho} + \frac{(\nabla \times \boldsymbol{B}) \times \boldsymbol{B}}{\mu_0 \rho}. \tag{2.28}$$

Then, from the vector identity

$$\nabla(\boldsymbol{A} \cdot \boldsymbol{B}) = \boldsymbol{A} \times (\nabla \times \boldsymbol{B}) + \boldsymbol{B} \times (\nabla \times \boldsymbol{A}) + (\boldsymbol{A} \cdot \nabla)\boldsymbol{B} + (\boldsymbol{B} \cdot \nabla)\boldsymbol{A}, \tag{2.29}$$

we have

$$\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = -\frac{\nabla P}{\rho} - \frac{\nabla(\frac{1}{2}B^2) - (\boldsymbol{B} \cdot \nabla)\boldsymbol{B}}{\mu_0 \rho}, \tag{2.30}$$

giving

$$\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = -\frac{1}{\rho}\nabla \cdot \left[(P + \frac{1}{2}B^2/\mu_0)\boldsymbol{I} - \frac{\boldsymbol{B}\boldsymbol{B}}{\mu_0}\right], \tag{2.31}$$

where we the last term arises from the product rule $\nabla \cdot (\boldsymbol{B}\boldsymbol{B}) = (\boldsymbol{B} \cdot \nabla)\boldsymbol{B} + \boldsymbol{B}(\nabla \cdot \boldsymbol{B}) = (\boldsymbol{B} \cdot \nabla)\boldsymbol{B}$ since $\nabla \cdot \boldsymbol{B} = 0$. In tensor notation we can write this as

$$\frac{\mathrm{d}v_i}{\mathrm{d}t} = -\frac{1}{\rho}\frac{\partial P}{\partial x^i} + \frac{1}{\rho}\frac{M_{ij}}{\partial x^j}, \tag{2.32}$$

where $M_{ij} \equiv 1/\mu_0 \left[B_i B_j - \frac{1}{2}B^2\delta_{ij}\right]$ is the Maxwell stress tensor.

2.1.6 MHD equations in Lagrangian conservative form

We can also include the hydrodynamic pressure in the stress tensor to write the MHD equations in the 'Lagrangian conservative form'. In tensor notation we find

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\rho\frac{\partial v^i}{\partial x^i}, \tag{2.33}$$

$$\frac{\mathrm{d}v_i}{\mathrm{d}t} = -\frac{1}{\rho}\frac{\partial S_{ij}}{\partial x^j}, \tag{2.34}$$

$$\frac{\mathrm{d}e}{\mathrm{d}t} = -\frac{1}{\rho}\frac{\partial(S_{ij}v^j)}{\partial x^i}, \tag{2.35}$$

where $S_{ij} \equiv (P + \frac{1}{2}B^2/\mu_0)\delta_{ij} - B_i B_j/\mu_0$ and $e \equiv \frac{1}{2}v^2 + u + \frac{1}{2}B^2/(\mu_0\rho)$ is the total specific energy.

2.1.7 MHD equations in Eulerian conservative form

As for the equations of hydrodynamics, one may also rewrite the MHD equations in Eulerian conservative form, giving

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{v}) = 0, \tag{2.36}$$

$$\frac{\partial}{\partial t}(\rho \boldsymbol{v}) + \nabla \cdot \left[\rho \boldsymbol{v}\boldsymbol{v} - \frac{\boldsymbol{B}\boldsymbol{B}}{\mu_0} + \left(P + \frac{1}{2}\frac{B^2}{\mu_0} \right) \boldsymbol{I} \right] = 0, \tag{2.37}$$

$$\frac{\partial}{\partial t}(\rho e) + \nabla \cdot \left[\left(\rho e + P + \frac{1}{2} B^2/\mu_0)\boldsymbol{I} - \frac{\boldsymbol{B}\boldsymbol{B}}{\mu_0} \right) \cdot \boldsymbol{v} \right] = 0, \tag{2.38}$$

$$\frac{\partial \boldsymbol{B}}{\partial t} + \nabla \cdot [\boldsymbol{v}\boldsymbol{B} - \boldsymbol{B}\boldsymbol{v}] = 0, \tag{2.39}$$

where e is the total specific energy as above. Using $\boldsymbol{E} = -\boldsymbol{v} \times \boldsymbol{B}$, the total energy equation can also be written in the form

$$\frac{\partial}{\partial t}(\rho e) + \nabla \cdot \left[\left(\frac{1}{2}\rho v^2 + \rho u + P \right) \boldsymbol{v} + \frac{\boldsymbol{E} \times \boldsymbol{B}}{\mu_0} \right] = 0. \tag{2.40}$$

The quantity $(\boldsymbol{E} \times \boldsymbol{B})/\mu_0$ is referred to as the *Poynting flux*, which from the above equation can be seen to represent the energy flow associated with the electromagnetic field.

2.2 Waves in ideal MHD

2.2.1 Linear perturbations to the MHD equations

As for the equations of hydrodynamics, a simple way to understand the new physics in a set of equations is to consider the propagation of small perturbations. We start with MHD equations written in the form

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{v}) = 0, \tag{2.41}$$

$$\rho \left[\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v} \right] = -\nabla P - \frac{1}{\mu_0} \left[\nabla \left(\frac{1}{2} B^2 \right) + (\boldsymbol{B} \cdot \nabla)\boldsymbol{B} \right], \tag{2.42}$$

$$\frac{\partial \boldsymbol{B}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{B} = (\boldsymbol{B} \cdot \nabla)\boldsymbol{v} - \boldsymbol{B}(\nabla \cdot \boldsymbol{v}), \tag{2.43}$$

we perturb the fluid variables using

$$\rho = \rho_0 + \delta\rho, \tag{2.44}$$

$$\boldsymbol{v} = \boldsymbol{v}_0 + \delta\boldsymbol{v} = \delta\boldsymbol{v}, \tag{2.45}$$

$$\boldsymbol{B} = \boldsymbol{B}_0 + \delta\boldsymbol{B}, \tag{2.46}$$

$$\delta P = c_{\mathrm{s}}^2 \delta\rho. \tag{2.47}$$

Inserting these into our equations and neglecting second order terms, we have

$$\frac{\partial \delta\rho}{\partial t} = -\rho_0 \nabla \cdot (\delta\boldsymbol{v}), \tag{2.48}$$

$$\rho_0 \frac{\partial \delta\boldsymbol{v}}{\partial t} = -c_{\mathrm{s},0}^2 \nabla\delta\rho - \frac{1}{\mu_0}\left[\nabla(\boldsymbol{B}_0 \cdot \delta\boldsymbol{B}) - (\boldsymbol{B}_0 \cdot \nabla)\delta\boldsymbol{B}\right], \tag{2.49}$$

$$\frac{\partial \delta\boldsymbol{B}}{\partial t} = (\boldsymbol{B}_0 \cdot \nabla)\,\delta\boldsymbol{v} - \boldsymbol{B}_0 \nabla \cdot (\delta\boldsymbol{v}). \tag{2.50}$$

Assuming perturbations of the form

$$\delta\rho = D e^{i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)}), \tag{2.51}$$

$$\delta\boldsymbol{v} = \boldsymbol{v} e^{i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)}) \tag{2.52}$$

$$\delta\boldsymbol{B} = \boldsymbol{b} e^{i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)}), \tag{2.53}$$

we find, for the time derivatives

$$\frac{\partial \delta\rho}{\partial t} = -i\omega D e^{i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)}, \tag{2.54}$$

$$\rho_0 \frac{\partial(\delta\boldsymbol{v})}{\partial t} = -i\omega\rho_0 \boldsymbol{v} e^{i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)}), \tag{2.55}$$

$$\frac{\partial(\delta\boldsymbol{B})}{\partial t} = -i\omega \boldsymbol{b} e^{i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)}), \tag{2.56}$$

while for the spatial derivatives we have

$$\nabla \cdot (\delta\boldsymbol{v}) = \frac{\partial(\delta v_x)}{\partial x} + \frac{\partial(\delta v_y)}{\partial y} + \frac{\partial(\delta v_z)}{\partial z} = i(k_x v_x + k_y v_y + k_z v_z) e^{i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)} \tag{2.57}$$

$$= i(\boldsymbol{v} \cdot \boldsymbol{k}) e^{i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)}, \tag{2.58}$$

$$\nabla(\delta\rho) = i\boldsymbol{k} D e^{i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)}, \tag{2.59}$$

$$\nabla(\boldsymbol{B}_0 \cdot \delta\boldsymbol{B}) = i(\boldsymbol{B}_0 \cdot \boldsymbol{b})\boldsymbol{k} e^{i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)}, \tag{2.60}$$

$$(\boldsymbol{B}_0 \cdot \nabla)\delta\boldsymbol{B} = i(\boldsymbol{B}_0 \cdot \boldsymbol{k})\boldsymbol{b} e^{i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)}, \tag{2.61}$$

$$(\boldsymbol{B}_0 \cdot \nabla)\delta\boldsymbol{v} = i(\boldsymbol{B}_0 \cdot \boldsymbol{k})\boldsymbol{v} e^{i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)}, \tag{2.62}$$

$$\boldsymbol{B}_0 \nabla \cdot (\delta\boldsymbol{v}) = i\boldsymbol{B}_0(\boldsymbol{v} \cdot \boldsymbol{k}) e^{i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)}. \tag{2.63}$$

Substituting into (2.48)–(2.50) and dividing by the common factor of $e^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}$ we have

$$-\omega D = -\rho_0(\boldsymbol{v}\cdot\boldsymbol{k}), \tag{2.64}$$

$$-\omega\boldsymbol{v} = -\frac{c_{\rm s}^2 D}{\rho_0}\boldsymbol{k} - \frac{1}{\mu_0\rho_0}\left[(\boldsymbol{B}_0\cdot\boldsymbol{b})\boldsymbol{k} - (\boldsymbol{B}_0\cdot\boldsymbol{k})\boldsymbol{b}\right], \tag{2.65}$$

$$-\omega\boldsymbol{b} = (\boldsymbol{B}_0\cdot\boldsymbol{k})\boldsymbol{v} - \boldsymbol{B}_0(\boldsymbol{v}\cdot\boldsymbol{k}). \tag{2.66}$$

Substituting $D = \rho_0(\boldsymbol{v}\cdot\boldsymbol{k})/\omega$ and $\boldsymbol{b} = -(\boldsymbol{B}_0\cdot\boldsymbol{k})\boldsymbol{v}/\omega + \boldsymbol{B}_0(\boldsymbol{v}\cdot\boldsymbol{k})/\omega$ into (2.65) gives

$$-\omega\boldsymbol{v} = -c_{\rm s}^2\frac{(\boldsymbol{v}\cdot\boldsymbol{k})\boldsymbol{k}}{\omega} - \frac{1}{\mu_0\rho_0}\left[-\frac{(\boldsymbol{B}_0\cdot\boldsymbol{k})(\boldsymbol{B}_0\cdot\boldsymbol{v})\boldsymbol{k}}{\omega} + \frac{B_0^2(\boldsymbol{v}\cdot\boldsymbol{k})\boldsymbol{k}}{\omega}\right. \tag{2.67}$$

$$\left. + \frac{(\boldsymbol{B}_0\cdot\boldsymbol{k})^2\boldsymbol{v}}{\omega} - \frac{(\boldsymbol{B}_0\cdot\boldsymbol{k})(\boldsymbol{v}\cdot\boldsymbol{k})\boldsymbol{B}_0}{\omega}\right] \tag{2.68}$$

Multiplying both sides by $-\omega$ and defining a new quantity $v_{\rm A}$ — the *Alfvén speed* — according to

$$v_{\rm A} \equiv \sqrt{\frac{B_0^2}{\mu_0\rho_0}}, \tag{2.69}$$

we have

$$\omega^2\boldsymbol{v} = \left(c_{\rm s}^2 + v_{\rm A}^2\right)(\boldsymbol{v}\cdot\boldsymbol{k})\boldsymbol{k} + \frac{(\boldsymbol{B}_0\cdot\boldsymbol{k})}{\mu_0\rho_0}\left[(\boldsymbol{B}_0\cdot\boldsymbol{k})\boldsymbol{v} - (\boldsymbol{B}_0\cdot\boldsymbol{v})\boldsymbol{k} - (\boldsymbol{v}\cdot\boldsymbol{k})\boldsymbol{B}_0\right]. \tag{2.70}$$

Our remaining problem is to eliminate the $(\boldsymbol{B}_0\cdot\boldsymbol{v})$ term. We can obtain an expression for this term by taking the dot product of (2.70) with $\boldsymbol{B}_0$ to give

$$\omega^2(\boldsymbol{B}_0\cdot\boldsymbol{v}) = \left(c_{\rm s}^2 + v_{\rm A}^2\right)(\boldsymbol{v}\cdot\boldsymbol{k})(\boldsymbol{B}_0\cdot\boldsymbol{k}) + \frac{(\boldsymbol{B}_0\cdot\boldsymbol{k})}{\mu_0\rho_0}\left[-(\boldsymbol{v}\cdot\boldsymbol{k})\boldsymbol{B}_0^2\right]. \tag{2.71}$$

Recognising $v_{\rm A}^2 \equiv B_0^2/(\mu_0\rho_0)$ and simplifying, we find

$$\omega^2(\boldsymbol{B}_0\cdot\boldsymbol{v}) = c_{\rm s}^2(\boldsymbol{v}\cdot\boldsymbol{k})(\boldsymbol{B}_0\cdot\boldsymbol{k}). \tag{2.72}$$

Finally, using this expression in (2.70) and taking an additional factor of $(\boldsymbol{B}_0\cdot\boldsymbol{k})$ out in front of the square brackets, we find

$$\omega^2\boldsymbol{v} = \left(c_{\rm s}^2 + v_{\rm A}^2\right)(\boldsymbol{v}\cdot\boldsymbol{k})\boldsymbol{k} + \frac{(\boldsymbol{B}_0\cdot\boldsymbol{k})^2}{\mu_0\rho_0}\left[\boldsymbol{v} - \frac{c_{\rm s}^2}{\omega^2}(\boldsymbol{v}\cdot\boldsymbol{k})\boldsymbol{k} - \frac{(\boldsymbol{v}\cdot\boldsymbol{k})}{(\boldsymbol{B}_0\cdot\boldsymbol{k})}\boldsymbol{B}_0\right]. \tag{2.73}$$

2.2.2 Transverse and longitudinal waves

We can find two possible families of solutions to (2.73). For the case where $\boldsymbol{v} \cdot \boldsymbol{k} = 0$ — corresponding to velocity perturbations purely transverse to the wave vector, hence *transverse waves* — we have

$$\omega^2 \boldsymbol{v} = \frac{(\boldsymbol{B}_0 \cdot \boldsymbol{k})^2}{\mu_0 \rho_0} \boldsymbol{v}. \tag{2.74}$$

Using $\boldsymbol{B}_0 \cdot \boldsymbol{k} \equiv |B_0||k| \cos\theta$ and dividing each equation by the relevant component of $\boldsymbol{v}$ on both sides gives

$$\omega^2 = \frac{B_0^2}{\mu_0 \rho_0} k^2 \cos^2\theta, \tag{2.75}$$

$$= k^2 v_{\mathrm{A}}^2 \cos^2\theta. \tag{2.76}$$

These waves are known as *Alfvén waves* after Alfvén (1942) and correspond to transverse oscillations travelling along magnetic field lines, like a wiggle on a rope. Alfvén waves are non-compressive since $\boldsymbol{v} \cdot \boldsymbol{k} = 0$ corresponds to $\nabla \cdot \boldsymbol{v} = 0$ in our original set of equations.

For the case where $\boldsymbol{v} \cdot \boldsymbol{k} \neq 0$, we can take the dot product of (2.73) with $\boldsymbol{k}$ and divide both sides by a factor of $\boldsymbol{v} \cdot \boldsymbol{k}$ to give

$$\omega^2 = \left(c_{\mathrm{s}}^2 + v_{\mathrm{A}}^2\right) k^2 + \frac{(\boldsymbol{B}_0 \cdot \boldsymbol{k})^2}{\mu_0 \rho_0} \left[1 - \frac{c_{\mathrm{s}}^2 k^2}{\omega^2} - 1\right]. \tag{2.77}$$

Multiplying by ω^2 and substituting $(\boldsymbol{B}_0 \cdot \boldsymbol{k})^2 = B_0^2 k^2 \cos^2\theta$ we find

$$\omega^4 - \left(c_{\mathrm{s}}^2 + v_{\mathrm{A}}^2\right) k^2 \omega^2 + c_{\mathrm{s}}^2 v_{\mathrm{A}}^2 k^4 \cos^2\theta = 0, \tag{2.78}$$

which is a quadratic for ω^2. Using the quadratic formula we find

$$\omega^2 = \frac{k^2}{2} \left[c_{\mathrm{s}}^2 + v_{\mathrm{A}}^2 \pm \sqrt{\left(c_{\mathrm{s}}^2 + v_{\mathrm{A}}^2\right)^2 - 4 c_{\mathrm{s}}^2 v_{\mathrm{A}}^2 \cos^2\theta}\right], \tag{2.79}$$

— the *fast* (+) and *slow* (-) magnetosonic waves. If $B_0 = 0$ these are just sound waves.

2.2.3 Summary of MHD wave types

In summary the wave speeds for all possible types of MHD waves are given by

$$v^2 \equiv \frac{\omega^2}{k^2} = \begin{cases} v_{\mathrm{A}}^2 \cos^2\theta & (\boldsymbol{v} \cdot \boldsymbol{k}) = 0; \\ \frac{1}{2}\left[c_{\mathrm{s}}^2 + v_{\mathrm{A}}^2 \pm \sqrt{\left(c_{\mathrm{s}}^2 + v_{\mathrm{A}}^2\right)^2 - 4 c_{\mathrm{s}}^2 v_{\mathrm{A}}^2 \cos^2\theta}\right] & (\boldsymbol{v} \cdot \boldsymbol{k}) \neq 0, \end{cases} \tag{2.80}$$

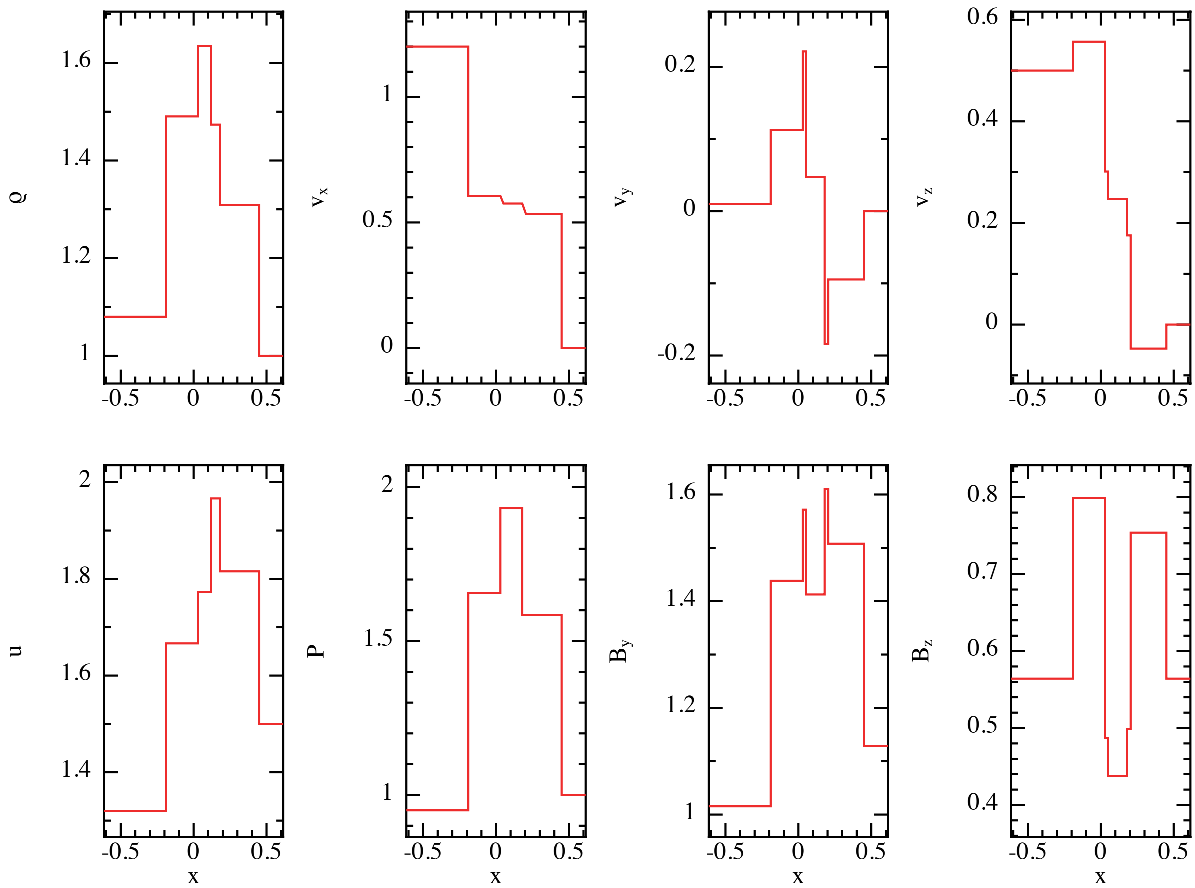

Figure 2.1: Structure of a hydro-magnetic shock, showing seven different possible discontinuities: the contact discontinuity at $x = 0.12$, with three characteristics corresponding to the slow, Alfvén and fast waves propagating left and right from the contact. Discontinuities corresponding to the Alfvén wave are called *rotational discontinuities* and can be seen to not change the pressure.

where $\boldsymbol{B}_0 \cdot \boldsymbol{k} = |B_0||k| \cos\theta$. Hence in MHD we have three possible types of waves: slow, Alfvén and fast. This implies that for a general disturbance there will be three characteristic waves travelling in each direction, meaning the corresponding shock structures are also more complex than for hydrodynamics (Figure 2.1).

2.2.4 Maximum wave speed in MHD

When the angle between the magnetic field and the wave vector is 90°, we have $\boldsymbol{B}_0 \cdot \boldsymbol{k} = 0$, i.e. $\cos\theta = 0$. To find maximum possible speed we take the positive root, giving

$$\omega^2 = \frac{k^2}{2}\left[c_{\rm s}^2 + v_{\rm A}^2 + \sqrt{(c_{\rm s}^2 + v_{\rm A}^2)^2}\right] = k^2(c_{\rm s}^2 + v_{\rm A}^2), \tag{2.81}$$

giving the maximum wave speed as

$$v_{\rm wave} = \sqrt{v_{\rm A}^2 + c_{\rm s}^2}, \tag{2.82}$$

where $v_{\rm A} \equiv \sqrt{B_0^2/(\mu_0 \rho_0)}$ is the Alfvén speed.

How to win a Nobel prize in 318 words

The paper by Alfvén (1942) is remarkable for its brevity — just over 300 words and seven equations to present the new kind of hydromagnetic waves and the associated Alfvén speed. A fine example of scientific writing, for which — amongst other things — Hannes Alfvén was awarded the 1970 Nobel Prize in Physics.

2.3 Flux freezing and Alfvén's theorem

A simple way to think about the non-linear behaviour of the MHD equations is by considering the magnetic flux through an arbitrary surface co-moving with the fluid,

$$\Phi = \int_c \boldsymbol{B}(\boldsymbol{x}, t) \cdot {\rm d}\boldsymbol{S}. \tag{2.83}$$

Considering the same surface at some later time $t + \delta t$, the change in flux is given by

$$\delta\Phi = \int_{c'} \boldsymbol{B}(\boldsymbol{x}, t + \delta t) \cdot {\rm d}\boldsymbol{S} - \int_c \boldsymbol{B}(\boldsymbol{x}, t) \cdot {\rm d}\boldsymbol{S}. \tag{2.84}$$

From the diagram (Figure 2.2) it is obvious that the flux at time $t + \delta t$ must be equal to the flux through the original surface minus any flux that has escaped through the sides of the cylinder swept out by the surface as it moves with the fluid. That is

$$\int_{c'} \boldsymbol{B}(\boldsymbol{x}, t + \delta t) \cdot {\rm d}\boldsymbol{S} = \int_c \boldsymbol{B}(\boldsymbol{x}, t) \cdot {\rm d}\boldsymbol{S} + \int_{\rm side} \boldsymbol{B}(\boldsymbol{x}, t) \cdot {\rm d}\boldsymbol{S} \tag{2.85}$$

$$= \int_c \boldsymbol{B}(\boldsymbol{x}, t) \cdot {\rm d}\boldsymbol{S} - \int \boldsymbol{B}(\boldsymbol{x}, t) \cdot ({\rm d}\boldsymbol{l} \times \boldsymbol{v}\delta t) \tag{2.86}$$

where the surface normal to the side of the cylinder corresponds to $({\rm d}\boldsymbol{l} \times \boldsymbol{v}\delta t)$, where ${\rm d}\boldsymbol{l}$ is the vector direction of a closed loop around the surface. From the vector identity $\boldsymbol{A} \cdot (\boldsymbol{B} \times \boldsymbol{C}) = \boldsymbol{B} \cdot (\boldsymbol{C} \times \boldsymbol{A})$ we have $\boldsymbol{B} \cdot ({\rm d}\boldsymbol{l} \times \boldsymbol{v}\delta t) = (\boldsymbol{v} \times \boldsymbol{B}) \cdot {\rm d}\boldsymbol{l}\delta t$, giving

$$\int_{c'} \boldsymbol{B}(\boldsymbol{x}, t + \delta t) \cdot {\rm d}\boldsymbol{S} = \int_c \boldsymbol{B}(\boldsymbol{x}, t) \cdot {\rm d}\boldsymbol{S} - \delta t \oint (\boldsymbol{v} \times \boldsymbol{B}) \cdot {\rm d}\boldsymbol{l}. \tag{2.87}$$

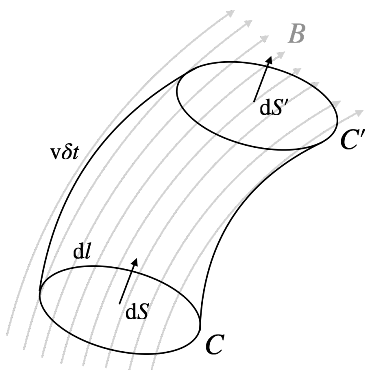

Figure 2.2: Magnetic flux through a surface moving with the fluid

Hence the change in flux, from (2.84) and using Stokes' theorem to write the line integral as a surface integral, is given by

$$\frac{\delta\Phi}{\delta t} = \int_c \frac{\boldsymbol{B}(\boldsymbol{x}, t+\delta t) - \boldsymbol{B}(\boldsymbol{x}, t)}{\delta t} \cdot \mathrm{d}\boldsymbol{S} - \int_c \nabla \times (\boldsymbol{v} \times \boldsymbol{B}) \cdot \mathrm{d}\boldsymbol{S}. \tag{2.88}$$

giving, taking the limit of $\delta t \to 0$

$$\frac{\mathrm{d}\Phi}{\mathrm{d}t} = \int \left[\frac{\partial \boldsymbol{B}}{\partial t} - \nabla \times (\boldsymbol{v} \times \boldsymbol{B})\right] \cdot \mathrm{d}\boldsymbol{S}. \tag{2.89}$$

Hence for ideal MHD, $\mathrm{d}\Phi/\mathrm{d}t = 0$ and we say that the flux is *frozen-in* to the fluid. This is known as Alfvén's theorem after Alfvén (1943). Matter is forced to flow along the magnetic field lines, if the field is strong, or the field lines are pushed around and tangled by the fluid, if the field is weak. Hence the main nonlinear behaviour in MHD is the *control of the direction of fluid flow by the magnetic field*. This is most obvious in nature in the solar atmosphere where the flow is directed by the magnetic field (Figure 2.3). This behaviour is also the basis for attempts to achieve controlled fusion on Earth by confining plasma with strong magnetic fields.

2.3.1 Magnetic flux in a closed volume

By the no-monopoles condition, the net magnetic flux through any surface bounding a closed volume must be zero, since from Green's theorem we can write

$$\Phi = \int_{\partial V} \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S} = \int_V (\nabla \cdot \boldsymbol{B}) \mathrm{d}V = 0, \tag{2.90}$$

where $\nabla \cdot \boldsymbol{B} = 0$ from (2.4).

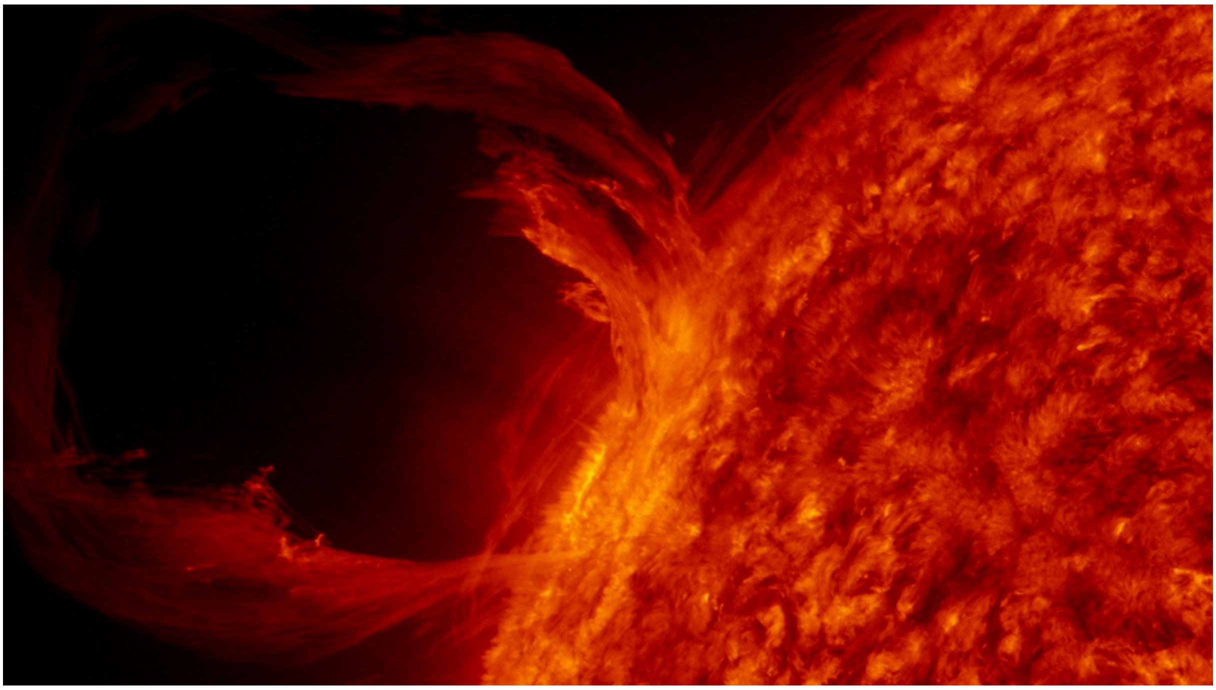

Figure 2.3: Solar prominence observed by the Solar Dynamics Observatory. Material lifted off the Sun's surface is channeled by the strong magnetic fields in the solar atmosphere. Credit: NASA/Goddard Space Flight Center Scientific Visualization Studio

2.4 Divergence constraint

2.4.1 The divergence constraint as an initial condition

How does the Maxwell equation $\nabla \cdot \boldsymbol{B} = 0$ (2.4) enter the MHD equations? Taking the divergence of the induction equation (2.20) we have

$$\frac{\partial}{\partial t}(\nabla \cdot \boldsymbol{B}) = \nabla \cdot [\nabla \times (\boldsymbol{v} \times \boldsymbol{B})], \tag{2.91}$$

giving, since the divergence of a curl equals zero,

$$\frac{\partial}{\partial t}(\nabla \cdot \boldsymbol{B}) = 0. \tag{2.92}$$

That is, if $\nabla \cdot \boldsymbol{B} = 0$ initially, it should remain so for all time. Hence the 'no-monopoles' condition enters as a constraint on the *initial conditions*. Obviously (2.92) is *not* guaranteed when the induction equation is solved on a computer. This requires thought.

2.4.2 Dirac's symmetric formulation of electromagnetism

Dirac (1931) first posited the existence of magnetic monopoles in the universe, famously showing this leads to a requirement that electric charge is quantised. While real magnetic monopoles remain undetected, it turns out to be useful in numerical codes to consider Maxwell's equations in 'symmetrised form', giving

$$\frac{1}{c^2}\frac{\partial \boldsymbol{E}}{\partial t} = \nabla \times \boldsymbol{B} - \mu_0 \boldsymbol{J}, \tag{2.93}$$

$$\frac{\partial \boldsymbol{B}}{\partial t} = -\nabla \times \boldsymbol{E} - \mu_0 \boldsymbol{J}_m, \tag{2.94}$$

$$\nabla \cdot \boldsymbol{E} = \frac{\rho_c}{\epsilon_0}, \tag{2.95}$$

$$\nabla \cdot \boldsymbol{B} = \mu_0 \rho_m, \tag{2.96}$$

where ρ_m is the magnetic charge density and $\boldsymbol{J}_m = \rho_m \boldsymbol{v}$ is the *magnetic current* arising from the movement of magnetic charges. Taking the divergence of (2.94), using the fact that the divergence of a curl is zero and substituting for $\boldsymbol{J}_m$ we find

$$\frac{\partial}{\partial t}(\nabla \cdot \boldsymbol{B}) + \nabla \cdot [\boldsymbol{v}(\nabla \cdot \boldsymbol{B})] = 0, \tag{2.97}$$

which has the same form as a continuity equation for the magnetic charge, i.e.

$$\frac{\partial \rho_m}{\partial t} + \nabla \cdot (\rho_m \boldsymbol{v}) = 0, \tag{2.98}$$

which, by analogy with the usual continuity equation implies conservation of $\int \rho_m \mathrm{d}V$. That is, using this form of the induction equation we would still conserve the total magnetic flux even if $\nabla \cdot \boldsymbol{B} \neq 0$, since the conserved quantity is

$$\Phi = \int \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S} = \int (\nabla \cdot \boldsymbol{B}) \mathrm{d}V. \tag{2.99}$$

Equation (2.94) may seem unusual, but it is actually just the induction equation we already wrote in Lagrangian form, since substituting for $\boldsymbol{J}_m$ and $\boldsymbol{E}$ in (2.94) gives

$$\frac{\partial \boldsymbol{B}}{\partial t} = \nabla \times (\boldsymbol{v} \times \boldsymbol{B}) - \boldsymbol{v}(\nabla \cdot \boldsymbol{B}). \tag{2.100}$$

which can be rearranged to give (2.22). This formulation of the induction equation using (2.100) or (2.22) is usually referred to as the '8-wave' formulation, after Powell (1994) and Powell et al. (1999). It does not actively remove $\nabla \cdot \boldsymbol{B}$ errors, but conserving monopoles does help to prevent them from becoming large in the first place.

2.4.3 Conservative vs orthogonal forces

A similar issue arises in the momentum equation. If we consider the magnetic force as the gradient of a stress tensor, we find

$$\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = -\frac{1}{\rho}\nabla \cdot \left[(P + \frac{1}{2}B^2/\mu_0)\boldsymbol{I} - \frac{\boldsymbol{B}\boldsymbol{B}}{\mu_0}\right], \tag{2.101}$$

$$= -\frac{\nabla P}{\rho} + \frac{\boldsymbol{J}\times\boldsymbol{B}}{\rho} + \frac{\boldsymbol{B}(\nabla\cdot\boldsymbol{B})}{\mu_0\rho}. \tag{2.102}$$

Hence the issue arises in numerical simulations as to whether the additional term proportional to $\nabla\cdot\boldsymbol{B}$ should be included, or not. The monopole force is directed along magnetic field lines and can be catastrophic in Lagrangian codes (e.g. Brackbill and Barnes, 1980; Phillips and Monaghan, 1985). In general one has a choice either to exactly conserve momentum, or to use a force which is exactly perpendicular to $\boldsymbol{B}$, but not both. Tóth (2002) showed that achieving both *is* possible in principle, but to my knowledge no actual implementation of such a scheme exists.

2.5 Enforcing the divergence-free condition

In numerical codes there are three general approaches to enforcing the $\nabla\cdot\boldsymbol{B} = 0$ condition:

i) ignore and hope for the best;

ii) reformulate the MHD equations to try to prevent $\nabla\cdot\boldsymbol{B} \neq 0$; or

iii) clean away $\nabla\cdot\boldsymbol{B}$ errors.

With approach i) one is hoping for simulation accuracy over sufficiently short timescales that numerical errors from the initial conditions do not have time to build. Like driving a car with no bolts in the wheel hubs and hoping you can drive to your destination before the wheels fall off. It's more common than you might imagine — I can supply references.

2.5.1 The vector potential

The simplest approach to prevention is to reformulate the MHD equations using the vector potential $\boldsymbol{B} = \nabla\times\boldsymbol{A}$. Integrating the induction equation (2.12) gives

$$\frac{\partial\boldsymbol{A}}{\partial t} = \boldsymbol{v}\times(\nabla\times\boldsymbol{A}) + \nabla\phi, \tag{2.103}$$

where ϕ is an arbitrary scalar. The most common application of the vector potential is within 'constrained transport' schemes for MHD on fixed meshes. The main issue with using the vector potential is that one must choose a gauge (value of ϕ). The most natural choice is the Lorenz gauge $\nabla \cdot \boldsymbol{A} = 0$, but this simply pushes the problem of obtaining a divergence-free field one level higher! The vector potential is also unnatural when written using the Lagrangian time derivative, giving

$$\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t} = \boldsymbol{v} \times (\nabla \times \boldsymbol{A}) + (\mathrm{v} \cdot \nabla)\boldsymbol{A} + \nabla\phi, \tag{2.104}$$

or, in tensor notation,

$$\frac{\mathrm{d}A^j}{\mathrm{d}t} = v^i \frac{\partial A^i}{\partial x^j} + \nabla\phi. \tag{2.105}$$

Another issue is that the right hand side depends on the absolute velocity which is not Galilean-invariant, although this can be fixed by choosing $\phi = -\boldsymbol{v} \cdot \boldsymbol{A}$ to give

$$\frac{\mathrm{d}A^j}{\mathrm{d}t} = -A^i \frac{\partial v^i}{\partial x^j}. \tag{2.106}$$

Nevertheless, the vector potential has not found widespread application in Lagrangian codes.

2.5.2 Euler potentials

An alternative is to consider the general decomposition of a vector field into three scalar fields according to

$$\boldsymbol{B} = \nabla\alpha \times \nabla\beta + \nabla\gamma. \tag{2.107}$$

Then, for a divergence-free field we have

$$\boldsymbol{B} = \nabla\alpha \times \nabla\beta. \tag{2.108}$$

This formulation is known as the *Euler potentials* (Stern, 1970). They are well suited to Lagrangian codes because the induction equation for ideal MHD (2.12) expressed in these variables reduces simply to

$$\frac{\mathrm{d}\alpha}{\mathrm{d}t} = 0; \qquad \frac{\mathrm{d}\beta}{\mathrm{d}t} = 0, \tag{2.109}$$

corresponding physically to the advection of magnetic field lines by Lagrangian particles. There are several (major) caveats to their practical use. First, not every possible field can be represented in this manner. Second, one cannot in general invert (2.108) for α and β from a given $\boldsymbol{B}$. Third, and related to the previous two is that the magnetic helicity —

$$\mathcal{H} \equiv \int (\boldsymbol{A} \cdot \boldsymbol{B})\, \mathrm{d}V, \tag{2.110}$$

a conserved quantity related to the topology of the magnetic field — is identically zero in the Euler potentials representation, since $\boldsymbol{A} = \alpha\nabla\beta$ and therefore $\boldsymbol{A} \cdot \boldsymbol{B} = 0$. This is a mathematical way of saying that only fields with simple geometries can be represented with the Euler potentials, which precludes important physical phenomena such as magnetic dynamos. Nevertheless, used within their limitations they offer a simple way to enforce the $\nabla \cdot \boldsymbol{B} = 0$ constraint (see e.g. Price and Bate 2007).

2.5.3 Projection methods

Brackbill and Barnes (1980) proposed to correct the magnetic field according to

$$\boldsymbol{B} = \boldsymbol{B}^* - \nabla\psi, \tag{2.111}$$

where $\boldsymbol{B}^*$ is a non-divergence-free field and $\boldsymbol{B}$ is the corrected, divergence-free field. Taking the divergence gives

$$\nabla^2\psi = \nabla \cdot \boldsymbol{B}^*, \tag{2.112}$$

which is an elliptic equation similar to Poisson's equation that can be solved for ψ (or $\nabla\psi$). The main problem is that solving elliptic equations is expensive, since it implies instant action — a small error in the magnetic field in one part of the domain results in a correction to the magnetic field everywhere. More subtle is that to achieve a truly divergence-free field, one must also satisfy (2.112) exactly, meaning that the numerical operators used to discretise the gradient and divergence in the ∇^2 term must match those used to compute the divergence and gradient in the right hand sides of (2.111) and (2.112).

2.5.4 Parabolic and hyperbolic cleaning

Dedner et al. (2002) proposed a more general formulation of divergence cleaning by adding a term to the induction equation according to

$$\frac{\partial \boldsymbol{B}}{\partial t} = \nabla \times (\boldsymbol{v} \times \boldsymbol{B}) - \nabla\psi. \tag{2.113}$$

For example, one can define $\psi = -\eta\nabla\cdot\boldsymbol{B}$ to give

$$\frac{\partial}{\partial t}(\nabla\cdot\boldsymbol{B}) = \nabla^2(\eta\nabla\cdot\boldsymbol{B}), \tag{2.114}$$

which is just the heat equation, hence *parabolic* cleaning. Solving the heat equation still brings numerical difficulties since the timestep constraint is in general proportional to the resolution length squared (i.e. stability requires $\Delta t < C_0\Delta x^2/\eta$). Even better is to evolve ψ with its own equation according to

$$\frac{\partial\psi}{\partial t} = -c_h^2(\nabla\cdot\boldsymbol{B}) - \frac{\psi}{\tau}, \tag{2.115}$$

where c_h is a quantity with the dimensions of speed and τ is a damping timescale. Taking $\partial/\partial t$ and $\nabla\cdot$ of (2.113) and ∇^2 of (2.115) and combining terms gives

$$\frac{\partial^2(\nabla\cdot\boldsymbol{B})}{\partial t^2} - c_h^2\nabla^2(\nabla\cdot\boldsymbol{B}) - \frac{c_h}{\tau}\frac{\partial(\nabla\cdot\boldsymbol{B})}{\partial t} = 0, \tag{2.116}$$

showing that in this case $\nabla\cdot\boldsymbol{B}$ propagates according to a damped wave equation. The same equation can be derived for ψ. Typically one sets the damping timescale τ equal to several times the resolution length divided by the cleaning speed c_h. Typically one chooses the cleaning speed to be equal to the maximum wave speed in the problem, i.e.

$$c_h = \sqrt{c_{\rm s}^2 + v_{\rm A}^2}. \tag{2.117}$$

The above means that divergence cleaning adds almost *no extra cost* to the numerical scheme, since there is no additional timestep constraint. Figure 2.4 shows an example of divergence cleaning in practice, showing the wave-like (hyperbolic) propagation of $\nabla\cdot\boldsymbol{B}$ and the reduction of these errors to zero on a short timescale by including the damping term.

2.5.5 Constrained hyperbolic/parabolic cleaning

Divergence cleaning as described above is not guaranteed to reduce the errors in the magnetic field and can sometimes make them worse (e.g. Balsara and Kim, 2004). The reason is that one has introduced an additional field (ψ) and should account for the energy exchanged between the $\boldsymbol{B}$-field and the ψ-field. Tricco and Price (2012) showed that the conserved total energy including the cleaning field is given by

$$E = \int\left[\frac{1}{2}\rho v^2 + \rho u + \frac{1}{2\mu_0}B^2 + \frac{1}{2\mu_0}\frac{\psi^2}{c_h^2}\right]{\rm d}V. \tag{2.118}$$

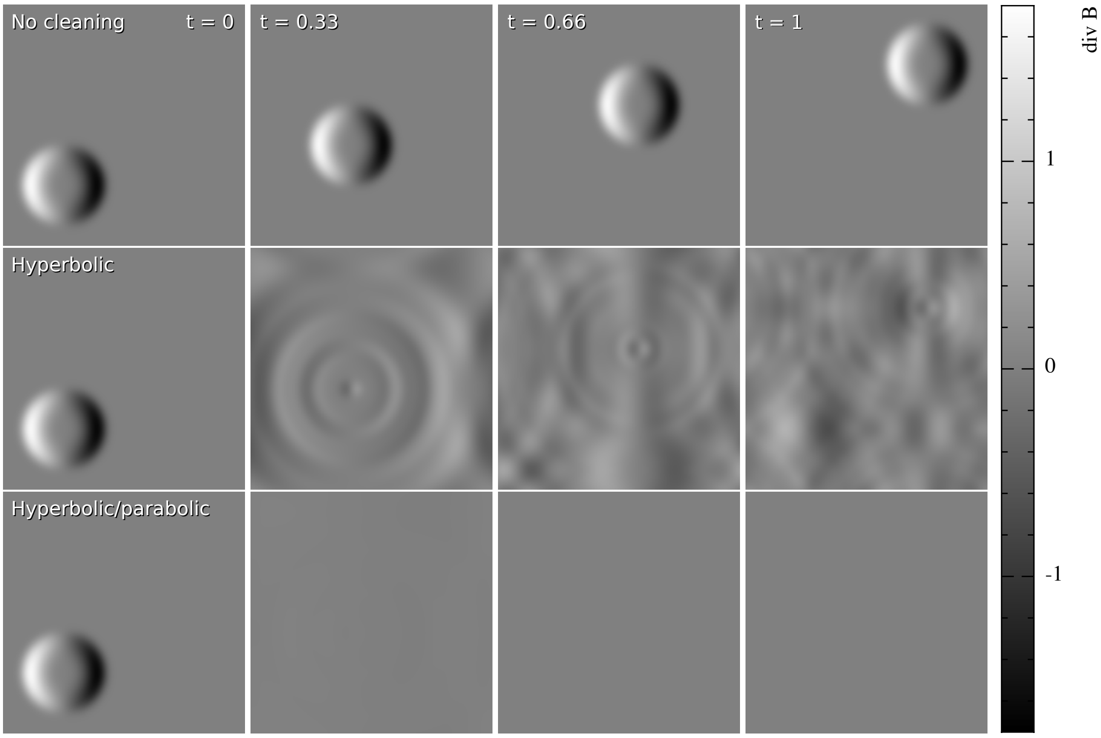

Figure 2.4: Divergence cleaning in action. The top row shows $\nabla \cdot \boldsymbol{B}$ when evolved with the 8-wave formulation (Eqs. 2.22 or 2.100), showing advection of $\nabla \cdot \boldsymbol{B}$ by the flow. Hyperbolic divergence cleaning with no damping (middle row) produces wave-like propagation of the divergence errors, whereupon adding the damping term in (2.115) produces rapid dissipation of the divergence errors to zero (bottom row). Credit: Tricco and Price (2012).

We must ensure that any numerical scheme conserves the above energy (in the absence of the damping term) — doing so guarantees a cleaning that can only ever conserve or dissipate the magnetic energy. Hence we require $\mathrm{d}E/\mathrm{d}t = 0$ using (2.118). The easiest way to take the Lagrangian time derivative into the integral is to notice that $\mathrm{d}/\mathrm{d}t\ (\rho \mathrm{d}V) = 0$ since $\rho\mathrm{d}V$ represents a mass element[2]. Then we require

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\int \left[\frac{1}{2}v^2 + u + \frac{B^2}{2\mu_0\rho} + \frac{\psi^2}{2\mu_0\rho c_{\mathrm{h}}^2}\right]\rho\mathrm{d}V \tag{2.119}$$

$$= \int_V \left[\boldsymbol{v}\cdot\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} + \frac{\mathrm{d}u}{\mathrm{d}t} + \frac{\boldsymbol{B}}{\mu_0\rho}\cdot\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t} - \frac{B^2}{2\mu_0\rho^2}\frac{\mathrm{d}\rho}{\mathrm{d}t} + \frac{\psi}{\sqrt{\rho}c_{\mathrm{h}}}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\psi}{\sqrt{\rho}c_{\mathrm{h}}}\right)\right]\rho\mathrm{d}V \tag{2.120}$$

$$= 0.$$

Since cleaning terms enter only in the $\mathrm{d}\boldsymbol{B}/\mathrm{d}t$ and $\mathrm{d}\psi/\mathrm{d}t$ equations, these terms must balance. Hence, we require, using $(\mathrm{d}\boldsymbol{B}/\mathrm{d}t)_{\mathrm{clean}} = -\nabla\psi$,

$$\int \left[\frac{\psi\sqrt{\rho}}{c_{\mathrm{h}}}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\psi}{\sqrt{\rho}c_{\mathrm{h}}}\right) - \frac{\boldsymbol{B}}{\mu_0}\cdot\nabla\psi\right]\mathrm{d}V = 0. \tag{2.121}$$

Integrating the second term by parts, we find

$$\int_V \left[\frac{\psi\sqrt{\rho}}{c_{\mathrm{h}}}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\psi}{\sqrt{\rho}c_{\mathrm{h}}}\right) + \frac{\psi}{\mu_0}\nabla\cdot\boldsymbol{B}\right]\mathrm{d}V + \frac{1}{\mu_0}\int_{\partial V}\psi\boldsymbol{B}\cdot\mathrm{d}\boldsymbol{S} = 0, \tag{2.122}$$

where we assume that $\psi \to 0$ on the boundary of the domain to cancel the surface integral. Hence, to conserve energy, we require (Tricco et al., 2016)

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\psi}{\sqrt{\rho}c_{\mathrm{h}}}\right) = -\frac{c_h}{\sqrt{\rho}}(\nabla\cdot\boldsymbol{B}). \tag{2.123}$$

which generalises (2.115) to the case where the density and cleaning speed are variable in space and time. We can then add a damping term as previously, giving

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\psi}{\sqrt{\rho}c_{\mathrm{h}}}\right) = -\frac{c_h}{\sqrt{\rho}}(\nabla\cdot\boldsymbol{B}) - \frac{1}{\tau}\frac{\psi}{\sqrt{\rho}c_{\mathrm{h}}}, \tag{2.124}$$

where one could also define an intermediate variable $\phi \equiv \psi/(\sqrt{\rho}c_h)$. Equivalently, one can expand (2.124) to give

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\psi}{c_h}\right) = -c_h(\nabla\cdot\boldsymbol{B}) - \frac{1}{2}\frac{\psi}{c_h}(\nabla\cdot\boldsymbol{v}) - \frac{\psi}{\tau c_h}. \tag{2.125}$$

As we have demonstrated, this version of the cleaning equations satisfy conservation of energy between the magnetic field and the cleaning field when no damping is applied

[2]Consider, for example, converting the integral to a sum over masses

($\tau \to \infty$). When damping is applied one can insert (2.124) into (2.120) to show that

$$\frac{\mathrm{d}E}{\mathrm{d}t} = -\frac{1}{\mu_0}\int_V \left(\frac{\psi^2}{\tau c_{\mathrm{h}}^2}\right)\mathrm{d}V, \tag{2.126}$$

which shows that energy is always removed and never added by the damping term, since the integral (2.126) is negative definite[3]. As long as the above properties are also satisfied by the discrete implementation, we can guarantee a stable divergence cleaning scheme. The only remaining issue is to ensure that the cleaning speed is sufficiently high to keep divergence errors acceptably small.

[3]One could also add the removed energy as heat which would ensure overall energy conservation. In practice, the energy loss is negligible anyway and so the choice of whether to discard or heat using the energy removed from the magnetic field is inconsequential.

3

Dust-gas mixtures

3.1 Equation set

In astrophysics and particularly in planet formation we model the fluid as a *mixture* of gas and dust. That is we model gas and dust as two separate fluids coupled by a drag term. This is the simplest example of a 'multiphase flow'. Conservation of mass and momentum in each species generalises the equations of hydrodynamics to the following set of equations

$$\frac{\partial \rho_{\rm g}}{\partial t} + \nabla \cdot (\rho_{\rm g} \boldsymbol{v}_{\rm g}) = 0, \tag{3.1}$$

$$\frac{\partial \rho_{\rm d}}{\partial t} + \nabla \cdot (\rho_{\rm d} \boldsymbol{v}_{\rm d}) = 0, \tag{3.2}$$

$$\rho_{\rm g} \left[\frac{\partial \boldsymbol{v}_{\rm g}}{\partial t} + (\boldsymbol{v}_{\rm g} \cdot \nabla)\, \boldsymbol{v}_{\rm g} \right] = -\nabla P_{\rm g} + K\,(\boldsymbol{v}_{\rm d} - \boldsymbol{v}_{\rm g}) + \rho_{\rm g} \boldsymbol{a}_{\rm ext}, \tag{3.3}$$

$$\rho_{\rm d} \left[\frac{\partial \boldsymbol{v}_{\rm d}}{\partial t} + (\boldsymbol{v}_{\rm d} \cdot \nabla)\, \boldsymbol{v}_{\rm d} \right] = -K\,(\boldsymbol{v}_{\rm d} - \boldsymbol{v}_{\rm g}) + \rho_{\rm d} \boldsymbol{a}_{\rm ext}, \tag{3.4}$$

$$\rho_{\rm g} \left[\frac{\partial u_{\rm g}}{\partial t} + (\boldsymbol{v}_{\rm g} \cdot \nabla)\, u_{\rm g} \right] = -P_{\rm g} (\nabla \cdot \boldsymbol{v})_{\rm g} + K\,(\boldsymbol{v}_{\rm d} - \boldsymbol{v}_{\rm g})^2, \tag{3.5}$$

where K is the drag coefficient, the subscripts g and d denote gas and dust, respectively, and $\boldsymbol{a}_{ext}$ is any external acceleration term (e.g. gravity). There are two important things: First, we implicitly assumed that dust-dust collisions are sufficiently rare so that there is no pressure in the dust. Hence the pressure gradient causes acceleration of the gas but not the dust. Second, we assumed that gas and dust exchange momentum via a 'drag' term $K(\boldsymbol{v}_\mathrm{d} - \boldsymbol{v}_\mathrm{g})$. At the moment K is just an arbitrary constant but in general would depend on fluid quantities — we will derive some possible expressions later. Key for the moment is that this term has opposite sign on the gas compared to the dust, so that the total momentum is conserved even though the gas and dust can exchange momentum with each other.

3.2 Physics of drag: the stopping time

To understand the new physics introduced by our more general set of equations, the best approach is to try to consider the effect of the drag terms in isolation. For example, consider a mixture with uniform density and velocities for each phase. In this case the spatial gradients are all zero, densities are constant in both space and time, and our equation set simplifies to

$$\frac{\partial \boldsymbol{v}_\mathrm{g}}{\partial t} = \frac{K}{\rho_\mathrm{g}}\left(\boldsymbol{v}_\mathrm{d} - \boldsymbol{v}_\mathrm{g}\right), \tag{3.6}$$

$$\frac{\partial \boldsymbol{v}_\mathrm{d}}{\partial t} = -\frac{K}{\rho_\mathrm{d}}\left(\boldsymbol{v}_\mathrm{d} - \boldsymbol{v}_\mathrm{g}\right). \tag{3.7}$$

Subtracting these gives

$$\frac{\partial(\boldsymbol{v}_\mathrm{d} - \boldsymbol{v}_\mathrm{g})}{\partial t} = -K\left(\frac{1}{\rho_\mathrm{d}} + \frac{1}{\rho_\mathrm{g}}\right)\left(\boldsymbol{v}_\mathrm{d} - \boldsymbol{v}_\mathrm{g}\right), \tag{3.8}$$

Defining

$$\Delta\boldsymbol{v} \equiv \boldsymbol{v}_\mathrm{d} - \boldsymbol{v}_\mathrm{g}, \tag{3.9}$$

and noticing that the prefactor on the right hand side has dimensions of inverse time, we can write this in the simple form

$$\frac{\partial \Delta\boldsymbol{v}}{\partial t} = -\frac{\Delta\boldsymbol{v}}{t_\mathrm{s}}, \tag{3.10}$$

where we see that the new physics is in the form of a new *timescale*, defined according to

$$t_\mathrm{s} \equiv \frac{\rho_\mathrm{g}\rho_\mathrm{d}}{K(\rho_\mathrm{g} + \rho_\mathrm{d})}. \tag{3.11}$$

Equation (3.10) is just a separable differential equation with solution (see box) given by

$$\Delta \boldsymbol{v} = \Delta \boldsymbol{v}_0 \exp\left[-\frac{(t-t_0)}{t_\mathrm{s}}\right]. \tag{3.12}$$

Solving for Δv

Assuming K independent of $\Delta \boldsymbol{v}$, this is a separable differential equation for each component, e.g. considering x direction

$$\frac{1}{\Delta v_x}\frac{\mathrm{d}\Delta v_x}{\mathrm{d}t} = -\frac{1}{t_\mathrm{s}}, \tag{3.13}$$

giving

$$\begin{aligned} \int_{\Delta v_x^0}^{\Delta v_x} \mathrm{d}(\ln \Delta v_x') &= -\frac{1}{t_\mathrm{s}}\int_{t_0}^{t} \mathrm{d}t' \\ \therefore \ln \Delta v_x - \ln \Delta v_x^0 &= -\frac{t-t_0}{t_\mathrm{s}} \\ \therefore \ln(\Delta v_x/\Delta v_x^0) &= -\frac{t-t_0}{t_\mathrm{s}} \\ \therefore \Delta v_x &= \Delta v_x^0 \exp\left[-\frac{(t-t_0)}{t_\mathrm{s}}\right]. \end{aligned}$$

Since the solution is the same for each component, we have the vector solution in the form (3.12)

From the above we see that t_s is the characteristic timescale on which the differential motion between the gas and dust is reduced to zero. The two species drag each other to the barycentric (centre-of-mass) velocity given by

$$\boldsymbol{v} = \frac{\rho_\mathrm{g}\boldsymbol{v}_\mathrm{g} + \rho_\mathrm{d}\boldsymbol{v}_\mathrm{d}}{\rho_\mathrm{g}+\rho_\mathrm{d}}. \tag{3.14}$$

One may already notice that adding (3.6) and (3.7) simply tells us that $\partial \boldsymbol{v}/\partial t = 0$.

The solution (3.12) can also be used to write down the velocity of each phase. Some straightforward algebra using (3.14) and (3.9) gives

$$\boldsymbol{v}_\mathrm{g}(t) = \boldsymbol{v} + \frac{\rho_\mathrm{d}}{\rho_\mathrm{g}+\rho_\mathrm{d}}\Delta \boldsymbol{v}(t), \tag{3.15}$$

$$\boldsymbol{v}_\mathrm{d}(t) = \boldsymbol{v} - \frac{\rho_\mathrm{g}}{\rho_\mathrm{g}+\rho_\mathrm{d}}\Delta \boldsymbol{v}(t). \tag{3.16}$$

Hence we see that the velocity of each phase exponentially decays to the barycentric velocity on the timescale t_s, which we hence refer to as the *stopping time*.

3.2.1 Stability condition for dust-gas evolution

Consider solving equation (3.10) numerically over a discrete time interval Δt. For example, using the Forward Euler method one would write

$$\frac{\Delta \boldsymbol{v}^{n+1} - \Delta \boldsymbol{v}^{n}}{\Delta t} = -\frac{\Delta \boldsymbol{v}^{n}}{t_{\rm s}}, \tag{3.17}$$

where n is the old timestep and $n+1$ is the next timestep. If we suppose that $\Delta \boldsymbol{v}^{n+1} = A\Delta \boldsymbol{v}^{n}$ then we require that $|A| < 1$ to prevent the solution from growing exponentially in time. From (3.17) we find

$$\frac{A-1}{\Delta t} = -\frac{1}{t_{\rm s}}, \tag{3.18}$$

giving

$$|A| = \left| 1 - \frac{\Delta t}{t_{\rm s}} \right|, \tag{3.19}$$

and hence a stability condition of the form

$$\Delta t \leq t_{\rm s}. \tag{3.20}$$

This stability requirement is general and will be true for any numerical method that solves (3.1)–(3.5) explicitly (that is, with the drag terms on the right hand side computed from the previous timestep). Hence in general one requires implicit timestepping methods to solve the dust-gas equations when K is large.

3.2.2 Stokes number

In summary, the basic physics of our new drag terms is the introduction of a new *timescale* — the stopping time. The gas and dust will drag each other towards their mutual barycentric velocity on this timescale. What is important is to consider this timescale with respect to other timescales in the problem. For dust in protoplanetary discs, the relevant timescale is the orbital period, so we define the dimensionless *Stokes number* according to

$$S_{\rm t} \equiv t_{\rm s}\Omega, \tag{3.21}$$

where $\Omega = \sqrt{GM/r^3}$ is the Keplerian angular speed.

3.2.3 Epstein and Stokes prescriptions for drag

To evaluate t_s we need to write down a physical prescription for the drag coefficient K. The simplest case is to consider spherical dust grains. There are two regimes:

Epstein drag Epstein (1924) considered the case when the grain size is much smaller than the mean free path of the gas, $s \lesssim \lambda_\mathrm{mfp}$. In this case gas molecules simply randomly bump into dust grains and bounce off in 'specular collisions'. In this case the net force on a single grain of radius s is given by

$$\boldsymbol{F}_\mathrm{drag} = -2\pi s^2 \rho_\mathrm{g} (\Delta v)^2 \times \left[\frac{1}{2\sqrt{\pi}}\left\{\left(\frac{1}{\psi}+\frac{1}{2\psi^3}\right)e^{-\psi^2} + \left(1+\frac{1}{\psi^2}-\frac{1}{4\psi^4}\right)\sqrt{\pi}\mathrm{erf}(\psi)\right\}\right]\Delta\boldsymbol{v}, \tag{3.22}$$

where $\psi \equiv \sqrt{\gamma/2}|\Delta v|/c_\mathrm{s}$. Fortunately this terrifying expression simplifies at low Mach number ($\psi \ll 1$; the *subsonic Epstein* regime) to simply

$$\boldsymbol{F}_\mathrm{drag} = -\frac{4\pi}{3}\rho_\mathrm{g} s^2 c_\mathrm{s} \Delta\boldsymbol{v}. \tag{3.23}$$

Notice in particular that the drag on a *single* grain is proportional to its surface area ($\boldsymbol{F}_\mathrm{drag} \propto s^2$). For our purposes, we need to consider the *collective* drag on the fluid, so we must multiply this force by the number density of grains, in order to find the drag force *per unit volume*. The number density of grains is given by

$$n = \frac{\rho_\mathrm{d}}{m_\mathrm{d}} = \frac{\rho_\mathrm{d}}{\frac{4}{3}\pi s^3 \rho_\mathrm{gr}}, \tag{3.24}$$

where ρ_gr is the *intrinsic grain density*. For example $\rho_\mathrm{gr} \approx 3$ g/cm^3 for typical silicate grains. Thus the drag force per unit volume is given by

$$\boldsymbol{F}_{\mathrm{drag},V} = \boldsymbol{F}_\mathrm{drag} \times n = \frac{\rho_\mathrm{g}\rho_\mathrm{d} c_\mathrm{s}}{\rho_\mathrm{gr} s}\Delta\boldsymbol{v}. \tag{3.25}$$

Comparing this to the term in our equations, we have

$$\boldsymbol{F}_{\mathrm{drag},V} = -K\Delta\boldsymbol{v} = \frac{\rho_\mathrm{g}\rho_\mathrm{d} c_\mathrm{s}}{\rho_\mathrm{gr} s}\Delta\boldsymbol{v}, \tag{3.26}$$

and hence the stopping time in the subsonic Epstein regime is given by

$$t_\mathrm{s} \approx \frac{\rho_\mathrm{gr} s}{\rho c_\mathrm{s}}, \tag{3.27}$$

where $\rho \equiv \rho_\mathrm{g} + \rho_\mathrm{d}$. The key point is that the multiplication by number density means that $t_\mathrm{s} \propto s$; the stopping time increases with grain size.

Stokes drag occurs when the grain size exceeds the mean free path $s \gtrsim \lambda_{\rm mfp}$. In this case the grain acts like a sphere obstructing the fluid flow. In this case the stopping time is given by

$$t_{\rm s} \approx \frac{\rho_{\rm gr} s}{\rho |\Delta \boldsymbol{v}| C_{\rm D}}, \tag{3.28}$$

where $C_{\rm D}$ is a coefficient that scales according to whether or not the flow around the sphere is turbulent according to (Fassio and Probstein, 1970; Whipple, 1972)

$$C_{\rm D} = \begin{cases} 24 R_{\rm e}^{-1}; & R_{\rm e} < 1, \\ 24 R_{\rm e}^{-0.6}; & 1 \leq R_{\rm e} \leq 800, \\ 0.44; & R_{\rm e} > 800, \end{cases} \tag{3.29}$$

where the Reynolds number is defined according to

$$R_{\rm e} \equiv \frac{2s|\Delta \boldsymbol{v}|}{\nu}. \tag{3.30}$$

Although the physics of the drag differs, the key point is that $t_{\rm s} \propto s$ also in the Stokes regime.

We find therefore that small stopping times correspond to small grains, and large stopping times correspond to large grains. This reflects our intuitive experience — smoke particles hang around in the air while rocks fall to the ground. What we define as small or large depends on the other timescales in the problem. In other words, we refer to *small grains* as those with $S_{\rm t} \ll 1$ and *large grains* as those with $S_{\rm t} \gg 1$.

3.3 Waves in a dust-gas mixture

3.3.1 Dispersion relation

Although our equation set is in general non-linear and complicated, simplified solutions provide a great deal of insight into the physics of the equations. As previously, we can obtain linear solutions by starting with perturbations of the form

$$\rho_{\rm g} = \rho_{\rm g}^0 + \delta \rho_{\rm g}, \tag{3.31}$$

$$\rho_{\rm d} = \rho_{\rm d}^0 + \delta \rho_{\rm d}, \tag{3.32}$$

$$\boldsymbol{v}_{\rm g} = \delta \boldsymbol{v}_{\rm g}, \tag{3.33}$$

$$\boldsymbol{v}_{\rm d} = \delta \boldsymbol{v}_{\rm d}, \tag{3.34}$$

$$\delta P_{\rm g} = c_{\rm s}^2 \delta \rho_{\rm g}. \tag{3.35}$$

Substituting these into our equations and keeping only first order terms, we find

$$\frac{\partial \delta\rho_{\rm g}}{\partial t} = -\rho_{\rm g}^0(\nabla\cdot\boldsymbol{v}_{\rm g}), \tag{3.36}$$

$$\frac{\partial \delta\rho_{\rm d}}{\partial t} = -\rho_{\rm d}^0(\nabla\cdot\boldsymbol{v}_{\rm d}), \tag{3.37}$$

$$\rho_{\rm g}^0\frac{\partial \boldsymbol{v}_{\rm g}}{\partial t} = -c_{\rm s}^2\nabla\delta\rho_{\rm g} + K\left(\boldsymbol{v}_{\rm d}-\boldsymbol{v}_{\rm g}\right), \tag{3.38}$$

$$\rho_{\rm d}^0\frac{\partial \boldsymbol{v}_{\rm d}}{\partial t} = -K\left(\boldsymbol{v}_{\rm d}-\boldsymbol{v}_{\rm g}\right), \tag{3.39}$$

where for convenience we have dropped the δ when referring to velocity perturbations, simply retaining the assumption that the velocity amplitudes are small. Taking the time derivative of (3.36) and (3.37) and adding these two equations, we find

$$\frac{\partial^2\delta\rho_{\rm g}}{\partial t^2} + \frac{\partial^2\delta\rho_{\rm d}}{\partial t^2} = -\rho_{\rm g}^0\frac{\partial}{\partial t}(\nabla\cdot\boldsymbol{v}_{\rm g}) - \rho_{\rm d}^0\frac{\partial}{\partial t}(\nabla\cdot\boldsymbol{v}_{\rm d}), \tag{3.40}$$

while taking the divergence of (3.38) and (3.39) and adding them gives

$$\rho_{\rm g}^0\frac{\partial}{\partial t}(\nabla\cdot\boldsymbol{v}_{\rm g}) + \rho_{\rm d}^0\frac{\partial}{\partial t}(\nabla\cdot\boldsymbol{v}_{\rm d}) = -c_{\rm s}^2\nabla^2\delta\rho_{\rm g}, \tag{3.41}$$

which on substitution into (3.40) gives

$$\frac{\partial^2\delta\rho_{\rm g}}{\partial t^2} + \frac{\partial^2\delta\rho_{\rm d}}{\partial t^2} = c_{\rm s}^2\nabla^2\delta\rho_{\rm g}. \tag{3.42}$$

The remaining information can be extracted by subtracting (3.38) from (3.39) and taking the divergence to give

$$\rho_{\rm d}^0\frac{\partial}{\partial t}(\nabla\cdot\boldsymbol{v}_{\rm d}) - \rho_{\rm d}^0\frac{\partial}{\partial t}(\nabla\cdot\boldsymbol{v}_{\rm g}) = c_{\rm s}^2\nabla^2\delta\rho_{\rm g} - 2K(\nabla\cdot\boldsymbol{v}_{\rm d}) + 2K(\nabla\cdot\boldsymbol{v}_{\rm g}). \tag{3.43}$$

Substituting using (3.36) and (3.37) and their time derivatives, we find

$$-\frac{\partial^2\delta\rho_{\rm d}}{\partial t^2} + \frac{\partial^2\delta\rho_{\rm g}}{\partial t^2} = c_{\rm s}^2\nabla^2\delta\rho_{\rm g} + 2\frac{K}{\rho_{\rm d}^0}\frac{\partial\delta\rho_{\rm d}}{\partial t} - 2\frac{K}{\rho_{\rm g}^0}\frac{\partial\delta\rho_{\rm g}}{\partial t}, \tag{3.44}$$

giving

$$\frac{\partial^2\delta\rho_{\rm d}}{\partial t^2} = \frac{\partial^2\delta\rho_{\rm g}}{\partial t^2} - c_{\rm s}^2\nabla^2\delta\rho_{\rm g} - 2\frac{K}{\rho_{\rm d}^0}\frac{\partial\delta\rho_{\rm d}}{\partial t} + 2\frac{K}{\rho_{\rm g}^0}\frac{\partial\delta\rho_{\rm g}}{\partial t}. \tag{3.45}$$

Substituting this expression in (3.42) and dividing by two, we find

$$\frac{\partial^2\delta\rho_{\rm g}}{\partial t^2} = c_{\rm s}^2\nabla^2\delta\rho_{\rm g} + \frac{K}{\rho_{\rm d}^0}\frac{\partial\delta\rho_{\rm d}}{\partial t} - \frac{K}{\rho_{\rm g}^0}\frac{\partial\delta\rho_{\rm g}}{\partial t}. \tag{3.46}$$

The key step to eliminating the remaining term involving $\delta\rho_{\rm d}$ is to take an additional time derivative, giving

$$\frac{\partial^3 \delta\rho_{\rm g}}{\partial t^3} = \frac{\partial}{\partial t}\left(c_{\rm s}^2 \nabla^2 \delta\rho_{\rm g}\right) + \frac{K}{\rho_{\rm d}^0}\frac{\partial^2 \delta\rho_{\rm d}}{\partial t^2} - \frac{K}{\rho_{\rm g}^0}\frac{\partial^2 \delta\rho_{\rm g}}{\partial t^2}, \tag{3.47}$$

upon which we can substitute using (3.42) to obtain

$$\frac{\partial^3 \delta\rho_{\rm g}}{\partial t^3} + K\left(\frac{1}{\rho_{\rm d}^0} + \frac{1}{\rho_{\rm g}^0}\right)\frac{\partial^2 \delta\rho_{\rm g}}{\partial t^2} - \frac{\partial}{\partial t}\left(c_{\rm s}^2 \nabla^2 \delta\rho_{\rm g}\right) - \frac{K}{\rho_{\rm d}^0} c_{\rm s}^2 \nabla^2 \delta\rho_{\rm g} = 0. \tag{3.48}$$

Finally, we can assume a perturbation of the form $\delta\rho_{\rm g} = De^{i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)}$ giving

$$\nabla^2 \delta\rho_{\rm g} = -k^2 \delta\rho_{\rm g}, \tag{3.49}$$

$$\frac{\partial}{\partial t}(\nabla^2 \delta\rho_{\rm g}) = i\omega k^2 \delta\rho_{\rm g}, \tag{3.50}$$

$$\frac{\partial^2 \delta\rho_{\rm g}}{\partial t^2} = -\omega^2 \delta\rho_{\rm g}, \tag{3.51}$$

$$\frac{\partial^3 \delta\rho_{\rm g}}{\partial t^3} = i\omega^3 \delta\rho_{\rm g}. \tag{3.52}$$

Inserting these expressions and dividing by $\delta\rho_{\rm g}$, and using our definition of $t_{\rm s}$ (noting also that $K/\rho_{\rm d}^0 \equiv \rho_{\rm g}^0/(\rho t_{\rm s})$) we find

$$i\omega^3 - \frac{\omega^2}{t_{\rm s}} - ic_{\rm s}^2 \omega k^2 + \frac{\rho_{\rm g}^0}{\rho t_{\rm s}} c_{\rm s}^2 k^2 = 0. \tag{3.53}$$

Multiplying by $-i/\omega$ we can write the dispersion relation in the comprehendible form

$$(\omega^2 - k^2 c_{\rm s}^2) + \frac{i}{\omega t_{\rm s}}(\omega^2 - \tilde{c}_{\rm s}^2 k^2) = 0, \tag{3.54}$$

where we have defined the *modified* sound speed as $\tilde{c}_{\rm s}^2 = (\rho_{\rm g}^0/\rho)c_{\rm s}^2$ (Miura and Glass, 1982).

3.3.2 Interpretation

The above dispersion relation is relatively straightforward to interpret. In the limit where $t_{\rm s} \to \infty$, corresponding to completely decoupled dust-and-gas ($K = 0$) the second term vanishes and we simply obtain

$$\omega^2 = c_{\rm s}^2 k^2, \tag{3.55}$$

which is identical to the regular hydrodynamical solution, namely propagation of undamped sound waves in the gas. In the opposite limit of $t_s \to 0$, corresponding to completely coupled dust-and-gas ($K = \infty$) the first term vanishes — you can see this by multiplying the whole expression by t_s — and we find

$$\omega^2 = \tilde{c}_s^2 k^2. \tag{3.56}$$

This corresponds to undamped sound waves propagating in a perfectly coupled *mixture*, where the only effect of dust is to slow the wave speed, since the sound speed is multiplied by the gas fraction. Essentially the fluid is 'weighed down' by the dust component which contributes to the inertia but not to the pressure. A key point however is that this limit of $t_s \to 0$ is perfectly well behaved and also corresponds to no damping. So both the $t_s \to 0$ and $t_s \to \infty$ limits are sensible and need to be handled correctly by numerical codes.

In between these limits we will obtain solutions for ω with imaginary components. Since $\delta\rho = De^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}$ having an imaginary component means that the *amplitude* of the wave will change, since the term in involving ω becomes $e^{|\omega|t}$. This could either be instability (exponential growth of the amplitude) or damping (exponential decay of the amplitude). In the absence of other forces it turns out that the solutions to (3.54) always *damp* the amplitude (a proof can be found in appendix A of Laibe and Price 2011). The strongest damping occurs when the term in the denominator is highest, namely when

$$\omega t_s \approx 1. \tag{3.57}$$

That is, when t_s is comparable to the wave period. This is when maximum damping of waves occurs.

3.4 Dust and gas with one fluid

We have seen from the dispersion relation that both the $t_s \to 0$ and $t_s \to \infty$ limits are sensible and realised in nature. The $t_s \to \infty$ limit is straightforward in our equation set (3.1)–(3.4): Simply uncouple the gas and dust by setting $K \to 0$ and you're done. However, the $t_s \to 0$ limit is problematic, because it corresponds to $K \to \infty$ which blows up the right hand side of the equations. Another consideration is that in general with explicit time integration one will require a stability condition of the form (3.20) which means it is impossible to represent the $t_s \to 0$ limit accurately with an explicit discretisation of (3.1)–(3.4).

3.4.1 Equation set

To show the limit of perfect coupling more clearly, let us change perspective. Instead of trying to describe our mixture as separate gas and dust fluids coupled by a drag term, one may consider a simple change of variables: From $\rho_{\rm g}$, $\rho_{\rm d}$, $\boldsymbol{v}_{\rm g}$ and $\boldsymbol{v}_{\rm d}$ to ρ, ϵ, $\boldsymbol{v}$ and $\Delta\boldsymbol{v}$ according to

$$\rho = \rho_{\rm g} + \rho_{\rm d}, \tag{3.58}$$

$$\epsilon = \rho_{\rm d}/\rho, \tag{3.59}$$

$$\boldsymbol{v} = \frac{\rho_{\rm g}\boldsymbol{v}_{\rm g} + \rho_{\rm d}\boldsymbol{v}_{\rm d}}{\rho_{\rm g} + \rho_{\rm d}}; \qquad \boldsymbol{v}_{\rm g} = \boldsymbol{v} - \frac{\rho_{\rm d}}{\rho}\Delta\boldsymbol{v}, \tag{3.60}$$

$$\Delta\boldsymbol{v} = \boldsymbol{v}_{\rm d} - \boldsymbol{v}_{\rm g}; \qquad \boldsymbol{v}_{\rm d} = \boldsymbol{v} + \frac{\rho_{\rm g}}{\rho}\Delta\boldsymbol{v}. \tag{3.61}$$

where ϵ is the dust fraction. Expressed in terms of these variables, equations (3.1)–(3.2) become (Laibe and Price, 2014)

$$\frac{{\rm d}\rho}{{\rm d}t} = -\rho\left(\nabla\cdot\boldsymbol{v}\right), \tag{3.62}$$

$$\frac{{\rm d}\epsilon}{{\rm d}t} = -\frac{1}{\rho}\nabla\cdot\left(\frac{\rho_{\rm g}\rho_{\rm d}}{\rho}\Delta\boldsymbol{v}\right), \tag{3.63}$$

$$\frac{{\rm d}\boldsymbol{v}}{{\rm d}t} = -\frac{\nabla P_{\rm g}}{\rho} - \frac{1}{\rho}\nabla\cdot\left(\frac{\rho_{\rm g}\rho_{\rm d}}{\rho}\Delta\boldsymbol{v}\Delta\boldsymbol{v}\right), \tag{3.64}$$

$$\frac{{\rm d}\Delta\boldsymbol{v}}{{\rm d}t} = -\frac{\Delta\boldsymbol{v}}{t_{\rm s}} + \frac{\nabla P_{\rm g}}{\rho_{\rm g}} - \left(\Delta\boldsymbol{v}\cdot\nabla\right)\boldsymbol{v} + \frac{1}{2}\nabla\cdot\left(\frac{\rho_{\rm d}-\rho_{\rm g}}{\rho}\Delta\boldsymbol{v}\Delta\boldsymbol{v}\right). \tag{3.65}$$

where

$$\frac{\rm d}{{\rm d}t} \equiv \frac{\partial}{\partial t} + \left(\boldsymbol{v}\cdot\nabla\right). \tag{3.66}$$

Although we have retained the same information — we have not yet made any approximations — already the $t_{\rm s} \to 0$ limit looks easier to handle. We can also see the physics of the coupling more clearly, since we have simply generalised our equation for the differential velocity evolution from (3.10). Notice also how (3.62) and (3.64) are just the regular fluid equations with an anisotropic pressure term and with a slight reinterpretation of quantities, since now ρ refers to the *total* mass density rather than just the gas density. Similarly, $\boldsymbol{v}$ is the barycentric velocity of the mixture rather than the gas velocity.

3.4.2 Terminal velocity approximation

We have not yet made any approximations, but we can make the $t_{\rm s}$ limit clearer by doing so. In the limit of small $t_{\rm s}$ and correspondingly small $\Delta\boldsymbol{v}$ we can neglect all but the first two terms on the right hand side of (3.65), giving the so-called *terminal velocity approximation* (Youdin and Goodman, 2005) in which

$$\Delta \boldsymbol{v} \approx t_{\rm s} \frac{\nabla P_{\rm g}}{\rho_{\rm g}}. \tag{3.67}$$

In this limit our equations simplify further to give (Laibe and Price, 2014)

$$\frac{{\rm d}\rho}{{\rm d}t} = -\rho \left(\nabla \cdot \boldsymbol{v}\right), \tag{3.68}$$

$$\frac{{\rm d}\boldsymbol{v}}{{\rm d}t} = -\frac{\nabla P}{\rho}, \tag{3.69}$$

$$\frac{{\rm d}\epsilon}{{\rm d}t} = -\frac{1}{\rho}\nabla \cdot \left(\epsilon t_{\rm s} \nabla P\right), \tag{3.70}$$

which are just the regular fluid equations with an additional equation evolving the dust fraction. The limit $t_{\rm s} \to 0$ is now trivial and corresponds to simply $\epsilon =$constant, or a constant dust-to-gas ratio. Indeed, one may question how the equations differ *at all* from the regular fluid equations when $t_{\rm s} \to 0$? The answer is in the equation of state, instead of $P = (\gamma - 1)\rho u$, here we have $P = (\gamma - 1)\rho_{\rm g} u = \rho_{\rm g}/\rho \times (\gamma - 1)\rho u$. So the sound speed is modified by the gas fraction, exactly as in our dispersion relation (3.54).

The downside to employing the one fluid formulation is that the $t_{\rm s} \to \infty$ limit has become correspondingly hard! However, solving (3.68)–(3.70) in numerical codes is relatively straightforward, and these equations are well behaved in the small grain / small Stokes number limit.

4

Relativistic hydrodynamics

4.1 Introduction to relativistic fluid dynamics

What is different about the fluid equations when we consider velocities close to the speed of light ($v \to c$) and strong gravitational fields such as those around black holes?

4.1.1 Newtonian hydrodynamics with self-gravity

Consider the Newtonian equations in the presence of self-gravity — gravity of the fluid on itself — written in Lagrangian form, are given by

$$\frac{d\rho}{dt} = -\rho(\nabla \cdot \boldsymbol{v}), \tag{4.1}$$

$$\frac{d\boldsymbol{v}}{dt} = -\frac{\nabla P}{\rho} - \nabla\Phi, \tag{4.2}$$

$$\frac{de}{dt} = -\frac{1}{\rho}\nabla \cdot (P\boldsymbol{v}), \tag{4.3}$$

alongside Poisson's equation for the gravitational potential Φ,

$$\nabla^2\Phi = 4\pi G\rho, \tag{4.4}$$

which in Newtonian mechanics is the *field equation* for the gravitational field, relating the density, ρ, to the gravitational potential, Φ. Recall that we can solve Poisson's equation analytically for simple cases. For example, if we neglect the contribution of the fluid but consider gravity from a central point mass, we find

$$\Phi = -\frac{GM}{r}, \tag{4.5}$$

where r is the distance from the point mass.

4.1.2 Relativistic equations of hydrodynamics

The relativistic hydrodynamics equations, in tensor notation, are given by

$$\frac{d\rho^*}{dt} = -\rho^* \frac{\partial v^i}{\partial x^i}, \tag{4.6}$$

$$\frac{dp_i}{dt} = -\frac{1}{\rho^*}\frac{\partial(\sqrt{-g}P)}{\partial x^i} + \frac{\sqrt{-g}}{2\rho^*}\left(T^{\mu\nu}\frac{\partial g_{\mu\nu}}{\partial x^i}\right), \tag{4.7}$$

$$\frac{de}{dt} = -\frac{1}{\rho^*}\frac{\partial(\sqrt{-g}Pv^i)}{\partial x^i} - \frac{\sqrt{-g}}{2\rho^*}\left(T^{\mu\nu}\frac{\partial g_{\mu\nu}}{\partial t}\right), \tag{4.8}$$

where ρ^*, p_i and e are the conserved or co-moving density, specific momentum and specific energy, respectively, $g_{\mu\nu}$ is the metric tensor, $T_{\mu\nu}$ is the energy-momentum tensor, and $\sqrt{-g}$ is the determinant of $g_{\mu\nu}$. Instead of Poisson's equation we must solve the Einstein Field Equations

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}, \tag{4.9}$$

where $R_{\mu\nu}$ is the Ricci tensor containing second derivatives of the metric and $R \equiv g^{\mu\nu}R_{\mu\nu}$ is the Ricci scalar. We adopt the Einstein summation convention where a repeated index implies a summation, giving for example our expression for $\nabla \cdot \boldsymbol{v}$ in the form

$$\frac{\partial v^i}{\partial x^i} \equiv \sum_{i=1}^{3} \frac{\partial v^i}{\partial x^i} = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}. \tag{4.10}$$

We also adopt the standard convention that Greek indices μ, ν run over four dimensions (0 to 3) while Latin indices i, j run from 1 to 3. For a perfect fluid (no viscosity or conductivity) the energy-momentum tensor is given by

$$T^{\mu\nu} = \left(\rho c^2 + \rho u + P\right) U^\mu U^\nu + P g^{\mu\nu}, \tag{4.11}$$

where $U^\mu = \frac{dx^\mu}{d\tau}$ is the *four velocity* and τ is the *proper time*. Here ρ is the *primitive* density, which differs from the conserved density ρ^*.

4.1.3 Key differences

Comparing our two sets of equations, we can pick out some key differences. The most important differences in moving to the relativistic case are

1. We must evolve a conserved mass, momentum and energy, each of which differ from their Newtonian counterpart.

   $$(\rho, \boldsymbol{v}, e) \rightarrow (\rho^*, p_i, e);$$

2. We require a procedure to solve for the *primitive* variables (ρ, v^i, u) from the *conserved* variables (ρ^*, p_i, e);

3. The Newtonian gravitational acceleration $(-\nabla\Phi)$ in (4.2) is replaced by spatial gradients of the metric $(\partial g_{\mu\nu}/\partial x^i)$ in (4.7) and there is a prefactor multiplying this term that involves the velocity and other fluid quantities, in the form of the stress-energy tensor $T^{\mu\nu}$.

4. In general, we must solve Einstein's equations to obtain the metric in place of Poisson's equation. However, similar to Poisson's equation we can use exact solutions for simple cases like a central point mass (the Schwarzschild metric) or a spinning central point mass (the Kerr metric). Importantly we also have an exact solution for the metric in flat space ('no gravity') in the form of the Minkowski metric.

4.1.4 The metric tensor

In Einstein's theory of both special and general relativity the key quantity is the metric tensor, $g_{\mu\nu}$. We shall refer to $g_{\mu\nu}$ as simply *the metric*, which is related to the invariant distance interval ds according to

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu. \tag{4.12}$$

For example, in flat space with one time and three space dimensions, we have

$$ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2,$$

implying a metric tensor (a 4×4 matrix) given by

$$g_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

This is the famous *Minkowski* metric[1]. In this case our 'Gravitational forces' are zero, since $\partial g_{\mu\nu}/\partial x^i = 0$. Hence if we want to perform fluid dynamics in *special relativity*, where the velocity of the fluid approaches the speed of light but there is no 'gravity' — space is flat — we can simply employ the Minkowski metric in (4.6)–(4.8). In this case we can also see that the determinant of the metric $\sqrt{-g} = 1$. It turns out this is true whenever a cartesian coordinate system is employed, so our equations can be further simplified in this case.

4.1.5 Special relativistic hydrodynamics

Hence, in the case of special relativity (Minkowski metric), our fluid equations simplify to a more familiar form

$$\frac{d\rho^*}{dt} = -\rho^*(\nabla\cdot\boldsymbol{v}), \tag{4.13}$$

$$\frac{d\boldsymbol{p}}{dt} = -\frac{\nabla P}{\rho^*}, \tag{4.14}$$

$$\frac{de}{dt} = -\frac{1}{\rho^*}\nabla\cdot(P\boldsymbol{v}), \tag{4.15}$$

where the specific momentum p_i can be written as a regular vector $\boldsymbol{p}$ in special relativity because there is no difference between raised and lowered indices (v^i vs. v_i) when space is flat. The only difference in (4.13)–(4.15) from non-relativistic hydrodynamics is the procedure required to solve for $\rho, \boldsymbol{v}$ and u from the quantities evolved on the left hand side ρ^*, $\boldsymbol{p}$ and e. This is more complicated than it seems, because (as we will show later) the conserved quantities are defined according to

$$\rho^* = \gamma\rho, \tag{4.16}$$

$$\boldsymbol{p} = w\gamma\boldsymbol{v}, \tag{4.17}$$

$$e = \boldsymbol{p}\cdot\boldsymbol{v} + \frac{(c^2+u)}{\gamma} = \frac{1}{\gamma}(c^2 + wv^2 + u), \tag{4.18}$$

where $\gamma = 1/\sqrt{1-v^2/c^2}$ is the Lorentz factor and $w \equiv c^2 + u + P/\rho$ is the specific enthalpy. Alternatively one can write the conserved energy in the form

$$e = w\gamma - \frac{P}{\gamma\rho}. \tag{4.19}$$

The key difficulty is that the velocity enters in all three equations via the Lorentz factor. So the inversion for ρ, $\boldsymbol{v}$ and u is tricky but nevertheless possible, as we will show in

[1] In the above we have also assumed a metric *signature* of the form $(-,+,+,+)$, which is the most common convention. However some famous textbooks do assume $(+,-,-,-)$ which is a crime against humanity.

Section 4.1.7. The main reason we need to perform the inversion is in order to evaluate the right hand side of our equations (4.13)–(4.15) which depend on the primitive quantities P and $\boldsymbol{v}$. The equation of state is as usual, e.g. $P = (\gamma_{\rm ad} - 1)\rho u$ where $\gamma_{\rm ad}$ is the adiabatic index (e.g. $\gamma_{\rm ad} = 5/3$ for a monatomic gas). We compute P from primitive quantities.

4.1.6 General relativistic hydrodynamics

General relativistic hydrodynamics corresponds to the case of a metric that is not simply the Minkowski metric. The equations in this case are given by (4.6)–(4.8). The main difference from special relativity is that we must be careful about the difference between raised and lowered quantities, and also with the definition of velocity. The rule is to raise and lower indices using the metric, for example the lowered (covariant) version of the four-velocity U_μ is related to the raised (contravariant) version U^ν according to

$$U_\mu = g_{\mu\nu} U^\nu, \tag{4.20}$$

where the four velocity is the derivative of x^μ with respect to proper time, i.e.

$$U^\mu \equiv \frac{{\rm d}x^\mu}{{\rm d}\tau}. \tag{4.21}$$

By contrast, our fluid velocity v^i relates to the coordinate time t and is defined as

$$v^\mu \equiv \frac{{\rm d}x^\mu}{{\rm d}t}, \tag{4.22}$$

where $v^0 \equiv {\rm d}t/{\rm d}t = 1$. We can also write $v^\mu = U^\mu / U^0$ since

$$U^0 = \frac{dt}{d\tau}. \tag{4.23}$$

The conserved quantities in this case are given by

$$\rho^* = \sqrt{-g}\, U^0 \rho, \tag{4.24}$$

$$p_i = w U_i = w U^0 g_{i\mu} v^\mu, \tag{4.25}$$

$$e = w p_i v^i + \frac{c^2 + u}{U^0} = \frac{1}{U^0}\left(c^2 + w v_i v^i + u\right). \tag{4.26}$$

From the metric we can write U^0 in terms of the velocity according to

$${\rm d}s^2 = -c^2 {\rm d}\tau^2 = g_{\mu\nu} {\rm d}x^\mu {\rm d}x^\nu, \tag{4.27}$$

giving

$$U^0 = \frac{1}{\sqrt{-g_{\mu\nu}v^\mu v^\nu/c^2}}. \tag{4.28}$$

When considering flows where the gravitational field is dominated by a central point mass (e.g. a supermassive black hole), where in Newtonian gravity we would use $\Phi = -\frac{GM}{r}$, the equivalent solution in general relativity is the Schwarzschild metric. In spherical polar coordinates we have (Schwarzschild, 1916)

$$ds^2 = -\left(1 - \frac{2GM}{rc^2}\right) c^2 dt^2 + \frac{dr^2}{\left(1 - \frac{2GM}{rc^2}\right)} + r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{4.29}$$

We can avoid coordinate issues at the poles by writing the metric in cartesian coordinates where $\sqrt{-g} = 1$. From (4.29) we have

$$g_{\mu\nu} = \begin{pmatrix} -f & 0 & 0 & 0 \\ 0 & f^{-1}\left[1 - \frac{R_s}{r}\frac{(y^2+z^2)}{r^2}\right] & f^{-1}\frac{R_s}{r}\frac{xy}{r^2} & f^{-1}\frac{R_s}{r}\frac{xz}{r^2} \\ 0 & f^{-1}\frac{R_s}{r}\frac{xy}{r^2} & f^{-1}\left[1 - \frac{R_s}{r}\frac{(x^2+z^2)}{r^2}\right] & f^{-1}\frac{R_s}{r}\frac{yz}{r^2} \\ 0 & f^{-1}\frac{R_s}{r}\frac{xz}{r^2} & f^{-1}\frac{R_s}{r}\frac{yz}{r^2} & f^{-1}\left[1 - \frac{R_s}{r}\frac{(x^2+y^2)}{r^2}\right] \end{pmatrix}, \tag{4.30}$$

where $f = (1 - R_s/r)$ and $R_s \equiv 2GM/c^2$ is the Schwarzschild radius. Hence we can easily compute the terms $\partial g_{\mu\nu}/\partial x^i$ required for the right hand side of the momentum equation (4.7). The time derivative of the metric required in (4.8) is zero in this case since the Schwarzschild metric is a *stationary* solution to the Einstein Field Equations (4.9).

4.1.7 Solving for primitive quantities

While there is more than one way to do this (see e.g. Noble et al., 2006), a simple and robust method involves solving for the enthalpy, w (Tejeda, 2012). For an ideal gas equation of state $P = (\gamma_{\rm ad} - 1)\rho u$ we have, in units where $c = 1$

$$w = 1 + \frac{\gamma_{\rm ad} P(w)}{(\gamma_{\rm ad} - 1)\rho(w)}, \tag{4.31}$$

from which we can rearrange our expressions (4.24)–(4.26) to give the required $\rho(w)$ and $P(w)$. For example we can easily rearrange (4.24) to give

$$\rho(w) = \frac{\rho^*}{\sqrt{-g}U^0(w)}, \tag{4.32}$$

and the procedure is similar but slightly more laborious for $P(w)$.

4.2 Newtonian self-gravity as a perturbation

One of the main uses of GR hydro is for simulating the close passage of stars to supermassive black holes. For this problem one can consider the self-gravity of the star as a small perturbation to the background (Kerr) spacetime. Then we can write the metric in the form

$$\tilde{g}_{\mu\nu} = g_{\mu\nu} + h_{\mu\nu}, \tag{4.33}$$

where $|h_{\mu\nu}| \ll |g_{\mu\nu}|$ is a small perturbation. We then assume a scalar perturbation to the metric, Φ, in the usual form (e.g. Bardeen, 1980; Mukhanov et al., 1992; Bertschinger, 2011)

$$h_{00} = -2\Phi; \qquad h_{ij} = 2\Phi/c^2\delta_{ij}. \tag{4.34}$$

Our assumption that $|h_{\mu\nu}| \ll |g_{\mu\nu}|$ means that we neglect $h_{\mu\nu}$ in comparison to $g_{\mu\nu}$ always, but do not assume that the gradient of $h_{\mu\nu}$ is small. We thus use $g_{\mu\nu}$ when computing dot products. The only change to the method is then an additional term in the momentum equation in the form

$$\left(\frac{\mathrm{d}p_i}{\mathrm{d}t}\right)_{\mathrm{sg}} = \frac{1}{2}U^0\frac{v^\mu v^\nu}{c^2}\frac{\partial h_{\mu\nu}}{\partial x^i}, \tag{4.35}$$

where we have neglected terms of $\mathcal{O}(v/c)^2$. The only non-zero terms are therefore

$$\left(\frac{\mathrm{d}p_i}{\mathrm{d}t}\right)_{\mathrm{sg}} = -U^0\frac{\partial\Phi}{\partial x^i} + U^0\delta_{jk}\frac{v^j v^k}{c^2}\frac{\partial\Phi}{\partial x^i} = -U^0\frac{\partial\Phi}{\partial x^i}\left(1 - \delta_{jk}\frac{v^j v^k}{c^2}\right). \tag{4.36}$$

The first term is the usual Newtonian acceleration term, while the second is a special relativistic correction of order $(v/c)^2$. Since self-gravity is only important at distances far from the black hole, we neglect the special relativistic correction term, and one may neglect the prefactor to simply use

$$\left(\frac{\mathrm{d}p_i}{\mathrm{d}t}\right)_{\mathrm{sg}} = -\frac{\partial\Phi}{\partial x^i}. \tag{4.37}$$

This approximation breaks down close to the black hole, but within the tidal radius $[r_t = r_*(M_{\mathrm{BH}}/m_*)^{1/3}]$ — by definition — the corrections due to self-gravity are negligible anyway. The Newtonian potential Φ can be computed in the usual manner, for example by solving Poisson's equation via a treecode (e.g. Barnes and Hut 1986; Hernquist and Katz 1989).

A possible improvement to the above would be to take account of the special relativistic corrections in the self-gravity term.

Bibliography

Alfvén, H.: 1942, 'Existence of Electromagnetic-Hydrodynamic Waves'. Nature **150**, 405–406.

Alfvén, H.: 1943, 'On the Existence of Electromagnetic-Hydrodynamic Waves'. *Arkiv for Matematik, Astronomi och Fysik* **29B**, 1–7.

Balsara, D. S. and J. Kim: 2004, 'A Comparison between Divergence-Cleaning and Staggered-Mesh Formulations for Numerical Magnetohydrodynamics'. ApJ **602**, 1079–1090.

Bardeen, J. M.: 1980, 'Gauge-invariant cosmological perturbations'. Phys. Rev. D **22**, 1882–1905.

Barnes, J. and P. Hut: 1986, 'A hierarchical O(N log N) force-calculation algorithm'. Nature **324**, 446–449.

Bertschinger, E.: 2011, 'One gravitational potential or two? Forecasts and tests'. *Philosophical Transactions of the Royal Society of London Series A* **369**, 4947–4961.

Biermann, L.: 1950, 'Über den Ursprung der Magnetfelder auf Sternen und im interstellaren Raum (miteinem Anhang von A. Schlüter)'. *Zeitschrift Naturforschung Teil A* **5**, 65.

Brackbill, J. U. and D. C. Barnes: 1980, 'The effect of nonzero product of magnetic gradient and B on the numerical solution of the magnetohydrodynamic equations'. *J. Comp. Phys.* **35**, 426–430.

Dedner, A., F. Kemm, D. Kröner, C.-D. Munz, T. Schnitzer, and M. Wesenberg: 2002, 'Hyperbolic Divergence Cleaning for the MHD Equations'. *J. Comp. Phys.* **175**, 645–673.

Dirac, P. A. M.: 1931, 'Quantised singularities in the electromagnetic field,'. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **133**(821), 60–72.

Epstein, P. S.: 1924, 'On the Resistance Experienced by Spheres in their Motion through Gases'. *Physical Review* **23**, 710–733.

Fassio, F. and R. F. Probstein: 1970, 'Dusty hypersonic flows'. *AIAA Journal* **8**, 772–779.

Graziani, C., P. Tzeferacos, D. Lee, D. Q. Lamb, K. Weide, M. Fatenejad, and J. Miller: 2015, 'The Biermann Catastrophe in Numerical Magnetohydrodynamics'. ApJ **802**, 43.

Hernquist, L. and N. Katz: 1989, 'TREESPH - A unification of SPH with the hierarchical tree method'. ApJS **70**, 419–446.

Laibe, G. and D. J. Price: 2011, 'DUSTYBOX and DUSTYWAVE: two test problems for numerical simulations of two-fluid astrophysical dust-gas mixtures'. MNRAS **418**, 1491–1497.

Laibe, G. and D. J. Price: 2014, 'Dusty gas with one fluid'. MNRAS **440**, 2136–2146.

Le Bellac, M. and J. Lévy-Leblond: 1973, 'Galilean electromagnetism'. *Il Nuovo Cimento B (1971-1996)* **14**(2), 217–234.

Maxwell, J. C.: 1865, 'A Dynamical Theory of the Electromagnetic Field'. *Philosophical Transactions of the Royal Society of London* **155**, 459–512.

Miura, H. and I. I. Glass: 1982, 'On a dusty-gas shock tube'. *Royal Society of London Proceedings Series A* **382**, 373–388.

Mukhanov, V. F., H. A. Feldman, and R. H. Brandenberger: 1992, 'Theory of cosmological perturbations'. Phys. Rep. **215**(5-6), 203–333.

Noble, S. C., C. F. Gammie, J. C. McKinney, and L. Del Zanna: 2006, 'Primitive Variable Solvers for Conservative General Relativistic Magnetohydrodynamics'. ApJ **641**, 626–637.

Pandey, B. P. and M. Wardle: 2008, 'Hall magnetohydrodynamics of partially ionized plasmas'. MNRAS **385**, 2269–2278.

Phillips, G. J. and J. J. Monaghan: 1985, 'A numerical method for three-dimensional simulations of collapsing, isothermal, magnetic gas clouds'. MNRAS **216**, 883–895.

Powell, K. G.: 1994, 'An approximate Riemann solver for magnetohydrodynamics (that works in more than one dimension)'. Technical Report ICASE 94-24, NASA Langley Research center (available from http://techreports.larc.nasa.gov/ltrs/dublincore/1994/icase-1994-24.html).

Powell, K. G., P. L. Roe, T. J. Linde, T. I. Gombosi, and D. L. de Zeeuw: 1999, ‘A Solution-Adaptive Upwind Scheme for Ideal Magnetohydrodynamics’. *J. Comp. Phys.* **154**, 284–309.

Price, D. J. and M. R. Bate: 2007, ‘The impact of magnetic fields on single and binary star formation’. MNRAS **377**, 77–90.

Schwarzschild, K.: 1916, ‘Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie’. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), 1916, Seite 189-196.*

Stern, D. P.: 1970, ‘Euler Potentials’. *Am. J. Phys.* **38**, 494–501.

Tejeda, E.: 2012, ‘An analytic Kerr-accretion model as a test solution for a new GR SPH code’. Ph.D. thesis, International School for Advanced Studies, Trieste, Italy, Via Bonomea 265, 34136 Trieste, Italy.

Tóth, G.: 2002, ‘Conservative and Orthogonal Discretization for the Lorentz Force’. *J. Comp. Phys.* **182**, 346–354.

Tricco, T. S. and D. J. Price: 2012, ‘Constrained hyperbolic divergence cleaning for smoothed particle magnetohydrodynamics’. *J. Comp. Phys.* **231**, 7214–7236.

Tricco, T. S., D. J. Price, and M. R. Bate: 2016, ‘Constrained hyperbolic divergence cleaning in smoothed particle magnetohydrodynamics with variable cleaning speeds’. *J. Comp. Phys.* **322**, 326–344.

Whipple, F. L.: 1972, ‘On certain aerodynamic processes for asteroids and comets’. In: A. Elvius (ed.): *From Plasma to Planet*. p. 211.

Youdin, A. N. and J. Goodman: 2005, ‘Streaming Instabilities in Protoplanetary Disks’. ApJ **620**(1), 459–469.