10 Basic gas dynamics

Gravity is the dominant driver behind cosmic structure formation (e.g. Mo et al., 2010), but at small scales hydrodynamics in the baryonic components becomes very important, too. In this section we very briefly review the basic equations and some prominent phenomena related to gas dynamics in order to make the discussion of the numerical fluid solvers used in galaxy evolution more accessible. For a detailed introduction to hydrodynamics, the reader is referred to the standard textbooks on this subject (e.g. Landau & Lifshitz, 1959; Shu, 1992).

10.1 Euler and Navier-Stokes equations

The gas flows in astrophysics are often of extremely low density, making internal friction in the gas extremely small. In the limit of assuming internal friction to be completely absent, we arrive at the so-called ideal gas dynamics as described by the Euler equations. Most calculations in cosmology and galaxy formation are carried out under this assumption. However, in certain regimes, viscosity may still become important (for example in the very hot plasma of rich galaxy clusters), hence we shall also briefly discuss the hydrodynamical equations in the presence of physical viscosity, the Navier-Stokes equations, which in a sense describe *real* fluids as opposed to ideal ones. Phenomena such as fluid instabilities or turbulence are also best understood if one does not neglect viscosity completely.

10.1.1 Euler equations

If internal friction in a gas flow can be neglected, the dynamics of the fluid is governed by the Euler equations:

$$\frac{\partial \rho}{\partial t} + \nabla(\rho \mathbf{v}) = 0, \tag{10.1}$$

$$\frac{\partial}{\partial t}(\rho \mathbf{v}) + \nabla(\rho \mathbf{v}\mathbf{v}^T + P) = 0, \tag{10.2}$$

$$\frac{\partial}{\partial t}(\rho e) + \nabla[(\rho e + P)\mathbf{v}] = 0, \tag{10.3}$$

where $e = u + \mathbf{v}^2/2$ is the total energy per unit mass, and u is the thermal energy per unit mass. Each of these equations is a continuity law, one for the mass, one for the momentum, and one for the total energy. The equations hence form a set of hyperbolic conservation laws. In the form given above, they are not yet complete,

however. One still needs a further expression that gives the pressure in terms of the other thermodynamic variables. For an ideal gas, the pressure law is

$$P = (\gamma - 1)\rho u, \tag{10.4}$$

where $\gamma = c_p/c_v$ is the ratio of specific heats. For a monoatomic gas, we have $\gamma = 5/3$.

10.1.2 Navier-Stokes equations

Real fluids have internal stresses, due to *viscosity*. The effect of viscosity is to dissipate relative motions of the fluid into heat. The Navier-Stokes equations are then given by

$$\frac{\partial \rho}{\partial t} + \nabla(\rho \mathbf{v}) = 0, \tag{10.5}$$

$$\frac{\partial}{\partial t}(\rho \mathbf{v}) + \nabla(\rho \mathbf{v}\mathbf{v}^T + P) = \nabla \Pi, \tag{10.6}$$

$$\frac{\partial}{\partial t}(\rho e) + \nabla[(\rho e + P)\mathbf{v}] = \nabla(\Pi \mathbf{v}). \tag{10.7}$$

Here Π is the so-called viscous stress tensor, which is a material property. For $\Pi = 0$, the Euler equations are recovered. To first order, the viscous stress tensor must be a linear function of the velocity derivatives (Landau & Lifshitz, 1959). The most general tensor of rank-2 of this type can be written as

$$\Pi = \eta \left[\nabla \mathbf{v} + (\nabla \mathbf{v})^T - \frac{2}{3}(\nabla \cdot \mathbf{v})\mathbf{1} \right] + \xi(\nabla \cdot \mathbf{v})\mathbf{1}, \tag{10.8}$$

where $\mathbf{1}$ is the unit matrix. Here η scales the traceless part of the tensor and describes the shear viscosity. ξ gives the strength of the diagonal part, and is the so-called bulk viscosity. Note that η and ξ can in principle be functions of local fluid properties, such as ρ, T, etc.

Incompressible fluids

In the following we shall assume constant viscosity coefficients. Also, we specialize to incompressible fluids with $\nabla \cdot \mathbf{v} = 0$, which is a particularly important case in practice. Let's see how the Navier-Stokes equations simplify in this case. Obviously, ξ is then unimportant and we only need to deal with shear viscosity. Now, let us consider one of the components of the viscous shear force described by equation (10.6):

$$\begin{aligned} \frac{1}{\eta}(\nabla \Pi)_x &= \frac{\partial}{\partial x}\left(2\frac{\partial v_x}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial v_x}{\partial y} + \frac{\partial v_y}{\partial x}\right) + \frac{\partial}{\partial z}\left(\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}\right) \\ &= \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) v_x = \nabla^2 v_x, \end{aligned} \tag{10.9}$$

where we made use of the $\nabla \cdot \mathbf{v} = 0$ constraint. If we furthermore introduce the *kinematic viscosity* ν as

$$\nu \equiv \frac{\eta}{\rho}, \tag{10.10}$$

we can write the equivalent of equation (10.6) in the compact form

$$\frac{\mathrm{D}\,\mathbf{v}}{\mathrm{D}\,t} = -\frac{\nabla P}{\rho} + \nu \nabla^2 \mathbf{v}, \tag{10.11}$$

where the derivative on the left-hand side is the Lagrangian derivative,

$$\frac{\mathrm{D}}{\mathrm{D}\,t} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla. \tag{10.12}$$

We hence see that the motion of individual fluid elements responds to pressure gradients and to viscous forces. The form (10.11) of the equation is also often simply referred to as the Navier-Stokes equation.

10.1.3 Scaling properties of viscous flows

Consider the Navier-Stokes equations for some flow problem that is characterized by some characteristic length L_0, velocity V_0, and density scale ρ_0. We can then define dimensionless fluid variables of the form

$$\hat{\mathbf{v}} = \frac{\mathbf{v}}{V_0}, \qquad \hat{\mathbf{x}} = \frac{\mathbf{x}}{L_0}, \qquad \hat{P} = \frac{P}{\rho_0 V_0^2}. \tag{10.13}$$

Similarly, we define a dimensionless time, a dimensionless density, and a dimensionless Nabla operator:

$$\hat{t} = \frac{t}{L_0/V_0}, \qquad \hat{\rho} = \frac{\rho}{\rho_0}, \qquad \hat{\nabla} = L_0 \nabla. \tag{10.14}$$

Inserting these definitions into the Navier-Stokes equation (10.11), we obtain the dimensionless equation

$$\frac{D\hat{\mathbf{v}}}{D\hat{t}} = -\frac{\hat{\nabla}\hat{P}}{\hat{\rho}} + \frac{\nu}{L_0 V_0}\,\hat{\nabla}^2 \hat{\mathbf{v}}. \tag{10.15}$$

Interestingly, this equation involves one number,

$$\mathrm{Re} \equiv \frac{L_0 V_0}{\nu}, \tag{10.16}$$

which characterizes the flow and determines the structure of the possible solutions of the equation. This is the so-called Reynolds number. Problems which have similar Reynolds number are expected to exhibit very similar fluid behavior. One then has *Reynolds-number similarity*. In contrast, the Euler equations ($\mathrm{Re} \to \infty$) exhibit always scale similarity because they are invariant under scale transformations.

One intuitive interpretation one can give the Reynolds number is that it measures the importance of inertia relative to viscous forces. Hence:

$$\mathrm{Re} \approx \frac{\text{inertial forces}}{\text{viscous forces}} \approx \frac{D\mathbf{v}/Dt}{\nu\nabla^2\mathbf{v}} \approx \frac{V_0/(L_0/V_0)}{\nu V_0/L_0^2} = \frac{L_0 V_0}{\nu}. \tag{10.17}$$

If we have $\mathrm{Re} \sim 1$, we are completely dominated by viscosity. On the other hand, for $\mathrm{Re} \to \infty$ viscosity becomes unimportant and we approach an ideal gas.

10.2 Shocks

An important feature of hydrodynamical flows is that they can develop shock waves in which the density, velocity, temperature and specific entropy jump by finite amounts (e.g. Toro, 1997). In the case of the Euler equations, such shocks are true mathematical discontinuities. Interestingly, shocks can occur even from perfectly smooth initial conditions, which is a typical feature of hyperbolic partial differential equations. In fact, acoustic waves with sufficiently large amplitude will suffer from wave-steeping (because the slightly hotter wave crests travel faster than the colder troughs), leading eventually to shocks. Of larger practical importance in astrophysics are however the shocks that occur when flows collide supersonically; here kinetic energy is irreversibly transferred into thermal energy, a process that also manifests itself with an increase in entropy.

In the limit of vanishing viscosity (i.e. for the Euler equations), the differential form of the fluid equations breaks down at the discontinuity of a shock, but the integral form (the *weak formulation*) remains valid. In other words this means that the flux of mass, momentum and energy must remain continuous at a shock front. Assuming that the shock connects two piecewise constant states, this leads to the Rankine-Hugoniot jump conditions (Rankine, 1870). If we select a frame of reference where the shock is stationary ($v_s = 0$) and denote the pre-shock state with (v_1, P_1, ρ_1), and the post-shock state as (v_2, P_2, ρ_2) (hence $v_1, v_2 > 0$), we have

$$\rho_1 v_1 = \rho_2 v_2, \tag{10.18}$$

$$\rho_1 v_1^2 + P_1 = \rho_2 v_2^2 + P_2, \tag{10.19}$$

$$(\rho_1 e_1 + P_1) v_1 = (\rho_2 e_2 + P_2) v_2. \tag{10.20}$$

For an ideal gas, the presence of a shock requires that the pre-shock gas streams supersonically into the discontinuity, i.e. $v_1 > c_1$, where $c_1^2 = \gamma P_1/\rho_1$ is the sound speed in the pre-shock phase. The Mach number

$$\mathcal{M} = \frac{v_1}{c_1} \tag{10.21}$$

measures the strength of the shock ($\mathcal{M} > 1$). The shock itself decelerates the fluid and compresses it, so that we have $v_2 < v_1$ and $\rho_2 > \rho_1$. It also heats it up, so

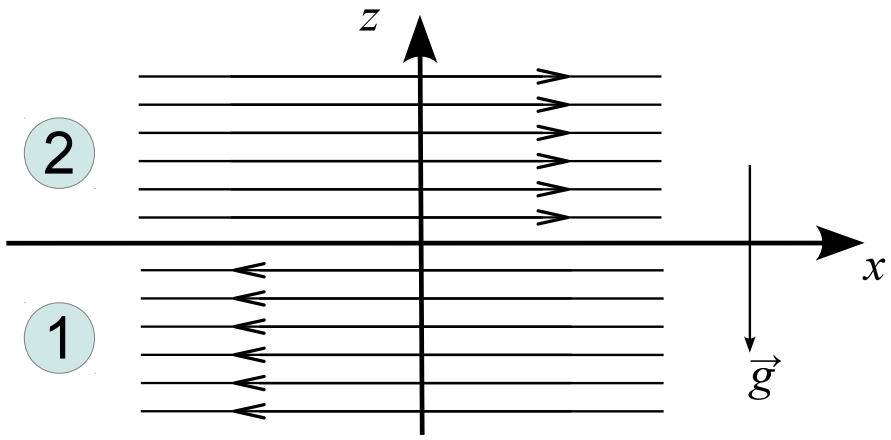

Figure 10.1: Geometry of a generic shear flow.

that $T_2 > T_1$, and makes the postshock flow subsonic, with $v_2/c_2 < 1$. Manipulating equations (10.18) to (10.20), we can express the relative jumps in the thermodynamic quantities (density, temperature, entropy, etc.) through the Mach number alone, for example:

$$\frac{\rho_2}{\rho_1} = \frac{(\gamma+1)\mathcal{M}^2}{(\gamma-1)\mathcal{M}^2+2}. \tag{10.22}$$

10.3 Fluid instabilities

In many situations, gaseous flows can be subject to fluid instabilities in which small perturbations can rapidly grow, thereby tapping a source of free energy. An important example of this are Kelvin-Helmholtz and Rayleigh-Taylor instabilities, which we briefly discuss in this section.

10.3.1 Stability of a shear flow

We consider a flow in the x-direction, which in the lower half-space $z < 0$ has velocity U_1 and density ρ_1, whereas in the upper half-space the gas streams with U_2 and has density ρ_2. In addition there can be a homogeneous gravitational field $\mathbf{g}$ pointing into the negative z-direction, as sketched in Figure 10.1.

The stability of the flow can be analysed through perturbation theory. To this end, one can for example treat the flow as an incompressible potential flow, and carry out an Eigenmode analysis in Fourier space. With the help of Bernoulli's theorem one can then derive an equation for a function $\xi(x,t) = z$ that describes the z-location of the interface between the two phases of the fluid. Details of this calculation can for example be found in Pringle & King (2007). For a single perturbative Fourier mode

$$\xi \;=\; \hat{\xi}\exp[i(kx-\omega t)], \tag{10.23}$$

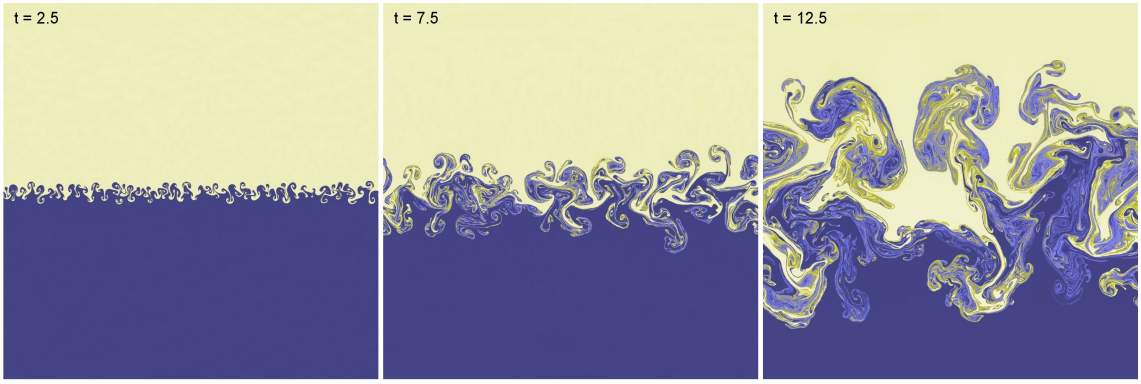

Figure 10.2: A growing Rayleigh-Taylor instability in which a lighter fluid (blue) is covered by a heaver fluid (yellow).

one then finds that non-trivial solutions with $\hat{\xi} \neq 0$ are possible for

$$\omega^2(\rho_1 + \rho 2) - 2\omega k(\rho_1 U_1 + \rho_2 U_2) + k^2(\rho_1 U_1^2 + \rho_2 U_2^2) + (\rho_2 - \rho 1)kg = 0, \quad (10.24)$$

which is the *dispersion relation*. Unstable, exponentially growing mode solutions appear if there are solutions for ω with positive imaginary part. Below, we examine the dispersion relation for a few special cases.

Rayleigh-Taylor instability

Let us consider the case of a fluid at rest, $U_1 = U_2 = 0$. The dispersion relation simplifies to

$$\omega^2 = \frac{(\rho_1 - \rho_2)kg}{\rho_1 + \rho_2}. \quad (10.25)$$

We see that for $\rho_2 > \rho_1$, i.e. the denser fluid lies on top, unstable solutions with $\omega^2 < 0$ exist. This is the so-called Rayleigh-Taylor instability. It is in essence buoyancy driven and leads to the rise of lighter material underneath heavier fluid in a stratified atmosphere, as illustrated in the simulation shown in Figure 10.2. The free energy that is tapped here is the potential energy in the gravitational field. Also notice that for an ideal gas, arbitrarily small wavelengths are unstable, and those modes will grow fastest. If on the other hand we have $\rho_1 > \rho_2$, then the interface is stable and will only oscillate when perturbed.

Kelvin-Helmholtz instability

If we set the gravitational field to zero, $g = 0$, we have the situation of a pure shear flow. In this case, the solutions of the dispersion relation are given by

$$\omega_{1/2} = \frac{k(\rho_1 U_1 + \rho_2 U_2)}{\rho_1 + \rho_2} \pm ik\frac{\sqrt{\rho_1 \rho_2}}{\rho_1 + \rho_2}|U_1 - U_2|. \quad (10.26)$$

Interestingly, in an ideal gas there is an imaginary growing mode component for every $|U_1 - U_2| > 0$! This means that a small wave-like perturbation at an interface will

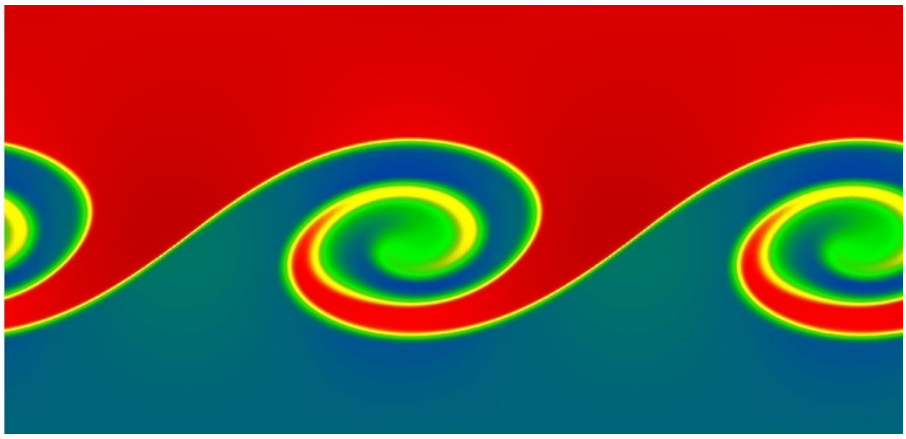

Figure 10.3: Characteristic Kelvin-Helmholtz billows arising in a shear flow.

grow rapidly into large waves that take the form of characteristic Kelvin-Helmholtz "billows". In the non-linear regime reached during the subsequent evolution of this instability the waves are rolled up, leading to the creation of vortex like structures, as seen in Figure 10.3. As the instability grows fastest for small scales (high k), the billows tend to get larger and larger with time.

Because the Kelvin-Helmholtz instability basically means that any sharp velocity gradient in a shear flow is unstable in a freely streaming fluid, this instability is particularly important for the creation of fluid turbulence. Under certain conditions, some modes can however be stabilized against the instability. This happens for example if we consider shearing with $U_1 \neq U_2$ in a gravitational field $g > 0$. Then the dispersion relation has the solutions

$$\omega = \frac{k(\rho_1 U_1 + \rho_2 U_2)}{\rho_1 + \rho_2} \pm \frac{\sqrt{-k^2 \rho_1 \rho_2 (U_1 - U_2)^2 - (\rho_1 + \rho_2)(\rho_2 - \rho_1) k g}}{\rho_1 + \rho_2}. \tag{10.27}$$

Stability is possible if two conditions are met. First, we need $\rho_1 > \rho_2$, i.e. the lighter fluid needs to be on top (otherwise we would have in any case a Rayleigh-Taylor instability). Second, the condition

$$(U_1 - U_2)^2 < \frac{(\rho_1 + \rho_2)(\rho_1 - \rho_2) g}{k \rho_1 \rho_2} \tag{10.28}$$

must be fulfilled. Compared to the ordinary Kelvin-Holmholtz instability without a gravitational field, we hence see that sufficiently small wavelengths are stabilized below a threshold wavelength. The larger the shear becomes, the further this threshold moves to small scales.

The Rayleigh-Taylor and Kelvin-Helmholtz instabilities are by no means the only fluid instabilities that can occur in an ideal gas (Pringle & King, 2007). For example, there is also the Richtmyer-Meshov instability, which can occur when an interface is suddenly accelerated, for example due to the passage of a shock wave. In self-gravitating gases, there is the Jeans instability, which occurs when the internal gas

pressure is not strong enough to prevent a positive density perturbation from growing and collapsing under its own gravitational attraction. This type of instability is particularly important in cosmic structure growth and star formation. If the gas dynamics is coupled to external sources of heat (e.g. through a radiation field), a number of further instabilities are possible. For example, a thermal instability (Field, 1965) can occur when a radiative cooling function has a negative dependence on temperature. If the temperature drops somewhere a bit more through cooling than elsewhere, the cooling rate of this cooler patch will increase such that it is cooling even faster. In this way, cool clouds can drop out of the background gas.

10.4 Turbulence

Fluid flow which is unsteady, irregular, seemingly random, and chaotic is called *turbulent* (Pope, 2000). Familiar examples of such situations include the smoke from a chimney, a waterfall, or the wind field behind a fast car or airplane. The characteristic feature of turbulence is that the fluid velocity varies significantly and irregularly both in position and time. As a result, turbulence is a statistical phenomenon and is best described with statistical techniques.

If the turbulent motions are subsonic, the flow can often be approximately treated as being incompressible, even for an equation of state that is not particularly stiff. Then only solenoidal motions that are divergence free can occur, or in other words, only shear flows are present. We have already seen that such flows are subject to fluid instabilities such as the Kelvin-Helmholtz instability, which can easily produce swirling motions on many different scales. Such vortex-like motions, also called *eddies*, are the conceptual building blocks of Kolmogorov's theory of incompressible turbulence (Kolmogorov, 1941), which yields a surprisingly accurate description of the basic phenomenology of turbulence, even though many aspects of turbulence are still not fully understood.

10.4.1 Kolmogorov's theory of incompressible turbulence

We consider a fully turbulent flow with characteristic velocity U_0 and length scale L_0. We assume that a quasi-stationary state for the turbulence is achieved by some kind of driving process on large scales, which in a time-averaged way injects an energy ϵ per unit mass. We shall also assume that the Reynolds number Re is large. We further imagine that the turbulent flow can be considered to be composed of eddies of different size l, with characteristic velocity $u(l)$, and associated timescale $\tau(l) = l/u(l)$.

For the largest eddies, $l \sim L_0$ and $u(l) \sim U_0$, hence viscosity is unimportant for them. But large eddies are unstable und break up, transferring their energy to somewhat smaller eddies. This continues to yet smaller scales, until

$$\mathrm{Re}(l) = \frac{l u(l)}{\nu} \tag{10.29}$$

becomes of the order of unity, where ν is the kinematic viscosity. For these eddies, viscosity will be very important so that their kinetic energy is dissipated away. We will see that this transfer of energy to smaller scales gives rise to the *energy cascade* of turbulence. But several important questions are still unanswered:

1. What is the actual size of the smallest eddies that dissipate the energy?

2. How do the velocities $u(l)$ of the eddies vary with l when the eddies become smaller?

Kolmogorov's hypotheses

Kolmogorov conjectured a number of hypotheses that can answer these questions. In particular, he proposed:

- For high Reynolds number, the small-scale turbulent motions ($l \ll L_0$) become statistically isotropic. Any memory of large-scale boundary conditions and the original creation of the turbulence on large scales is lost.

- For high Reynolds number, the statistics of small-scale turbulent motions has a universal form and is only determined by ν and the energy injection rate per unit mass, ϵ.

From ν and ϵ, one can construct characteristic Kolmogorov length, velocity and timescales. Of particular importance is the *Kolmogorov length*:

$$\eta \equiv \left(\frac{\nu^3}{\epsilon}\right)^{1/4}. \tag{10.30}$$

Velocity and timescales are given by

$$u_\eta = (\epsilon\nu)^{1/4}, \qquad \tau_\eta = \left(\frac{\nu}{\epsilon}\right)^{1/2}. \tag{10.31}$$

We then see that the Reynolds number at the Kolmogorov scale is

$$\mathrm{Re}(\eta) = \frac{\eta u_\eta}{\nu} = 1, \tag{10.32}$$

showing that they describe the dissipation range. Kolmogorov has furthermore made a second similarity hypothesis, as follows:

- For high Reynolds number, there is a range of scales $L_0 \gg l \gg \eta$ over which the statistics of the motions on scale l take a universal form, and this form is *only* determined by ϵ, *independent* of ν.

In other words, this also means that viscous effects are unimportant over this range of scales, which is called the *inertial range*. Given an eddy size l in the inertial range, one can construct its characteristic velocity and timescale just from l and ϵ:

$$u(l) = (\epsilon l)^{1/3}, \qquad \tau(l) = \left(\frac{l^2}{\epsilon}\right)^{1/3}. \tag{10.33}$$

One further consequence of the existence of the inertial range is that here the energy transfer rate

$$T(l) \sim \frac{u^2(l)}{\tau(l)} \tag{10.34}$$

of eddies to smaller scales is expected to be scale-invariant. Indeed, putting in the expected characteristic scale dependence we get $T(l) \sim \epsilon$, i.e. $T(l)$ is equal to the energy injection rate. This also implies that we have

$$\epsilon \sim \frac{U_0^3}{L_0}. \tag{10.35}$$

With this result we can also work out what we expect for the ratio between the characteristic quantities of the largest and smallest scales:

$$\frac{\eta}{L_0} \sim \left(\frac{\nu^3}{\epsilon L_0^4}\right)^{1/4} = \left(\frac{\nu^3}{U_0^3 L_0^3}\right)^{1/4} = \mathrm{Re}^{-\frac{3}{4}}, \tag{10.36}$$

$$\frac{u_\eta}{U_0} \sim \left(\frac{\epsilon\nu}{U_0^4}\right)^{1/4} = \left(\frac{U_0^3\nu}{L_0 U_0^4}\right)^{1/4} = \mathrm{Re}^{-\frac{1}{4}}, \tag{10.37}$$

$$\frac{\tau_\eta}{\tau} \sim \left(\frac{\nu U_0^2}{\epsilon L_0^2}\right)^{1/2} = \left(\frac{\nu U_0^2 L_0}{U_0^3 L_0^2}\right)^{1/2} = \mathrm{Re}^{-\frac{1}{2}}. \tag{10.38}$$

This shows that the Reynolds number directly sets the dynamic range of the inertial range.

10.4.2 Energy spectrum of Kolmogorov turbulence

Eddy motions on a length-scale l correspond to wavenumber $k = 2\pi/l$. The kinetic energy ΔE contained between two wave numbers k_1 and k_2 can be described by

$$\Delta E = \int_{k_1}^{k_2} E(k)\,\mathrm{d}k, \tag{10.39}$$

where $E(k)$ is the so-called energy spectrum. For the inertial range in Kolmogorov's theory, we know that $E(k)$ is a universal function that only depends on ϵ and k. Hence $E(k)$ must be of the form

$$E(k) = C\,\epsilon^a\,k^b, \tag{10.40}$$

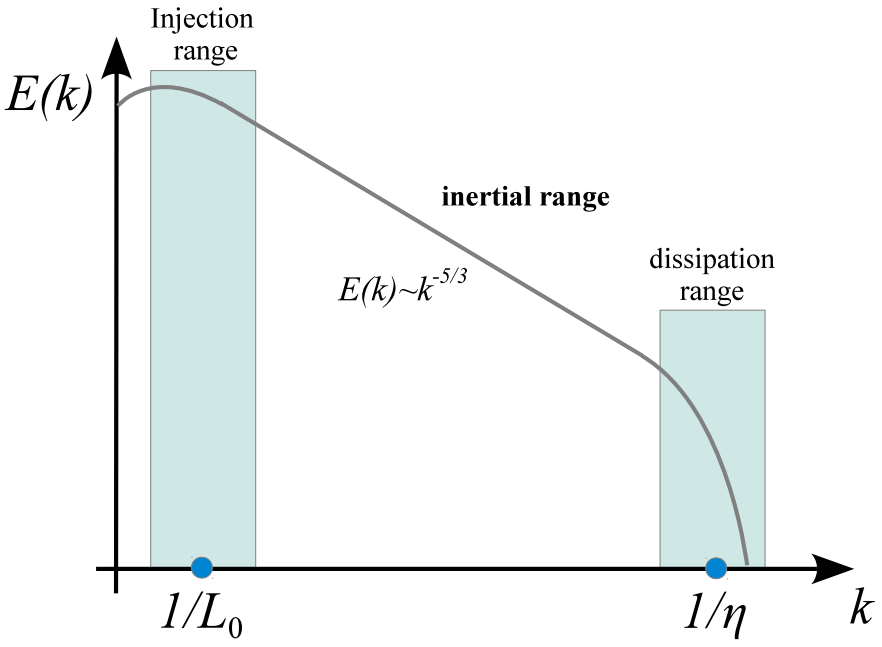

Figure 10.4: Schematic energy spectrum of Kolmogorov turbulence.

where C is a dimensionless constant. Through dimensional analysis it is easy to see that one must have $a = 2/3$ and $b = -5/3$. We hence obtain the famous $-5/3$ slope of the Kolmogorov energy power spectrum:

$$E(k) = C\,\epsilon^{2/3}\,k^{-5/3}. \tag{10.41}$$

The constant C is universal in Kolmogorov's theory, but cannot be computed from first principles. Experiment and numerical simulations give $C \simeq 1.5$ (Pope, 2000).

Actually, if we recall Kolmogorov's first similarity hypothesis, it makes the stronger claim that the statistics for all small scale motion is universal. This means that also the dissipation part of the turbulence must have a universal form. To include this in the description of the spectrum, we can for example write

$$E(k) = C\,\epsilon^{2/3}\,k^{-5/3} f_\eta(k\eta), \tag{10.42}$$

where $f_\eta(k\eta)$ us a universal function with $f_\eta(x) = 1$ for $x \ll 1$, and with $f_\eta(x) \to 0$ for $x \to \infty$. This function has to be determined experimentally or numerically. A good fit to different results is given by

$$f_\eta(x) = \exp\left(-\beta[(x^4 + c^4)^{1/4} - c]\right), \tag{10.43}$$

with $\beta_0 \sim 5.2$ and $c \sim 0.4$ (Pope, 2000).

11 Eulerian hydrodynamics

Many physical theories are expressed as partial differential equations (PDEs), including some of the most fundamental laws of nature, such as fluid dynamics (Euler and Navier Stokes equations), electromagnetism (Maxwell's equations) or general relativity/gravity (Einstein's field equations). Broadly speaking, partial differential equations (PDE) are equations describing relations between partial derivatives of a dependent variable with respect to several independent variables. For example,

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \tag{11.1}$$

is a PDE for the function $u = u(x, t)$. The independent variables are here x and t, the dependent variable is u (i.e. the function value). Unlike for ordinary differential equations (ODEs), there is no simple unified theory for PDEs. Rather, there are different types of PDEs which exhibit special features.

Given that the fundamental laws of classical physics are expressed in terms of partial differential equations, it is clear that such equations are of fundamental importance in all of physics. The solutions of PDEs can be extremely complex, encoding many physical phenomena. Simulation techniques are crucial to uncover and understand parts of the solution spaces of PDEs. But depending on the PDEs, numerically solving them can be very tricky, and is quite generally more difficult than solving systems of ordinary differential equations.

11.1 Types of PDEs

It is useful to distinguish different properties of PDEs in order to help classifying them. Some of the most important characteristics include:

- **Order of the PDE:** This is simply the highest derivative appearing in the PDE. A "second-order PDE" will have up to second partial derivatives.
- **Linearity:** A PDE (or system of PDEs) is linear if all terms are linear in the unknown function (the dependent variable) and its partial derivatives. Linear PDEs are (unsurprisingly) much simpler to solve than non-linear ones.
- **Homogeneity:** If all its terms contain the independent variable or its derivative, the PDE is said to be homogenous. Otherwise it contains "source-terms" and is inhomogenous.

In addition, PDEs can be classified according to some less obvious features which are however reflecting the nature of the physics that the PDEs describe. For example, the hydrodynamic fluid equations are local conservation laws, hence they are expressed as so-called *hyperbolic* conservation laws. What this means will be discussed later in more detail.

The most important linear and homogenous PDEs

In order to introduce the different characters of PDEs, let us look at the **three most impost important linear and homogenous PDEs**. These arguably are:

1. Laplace equation:
$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0 \tag{11.2}$$

2. Heat conduction equation
$$\frac{\partial u}{\partial t} = \lambda^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) \tag{11.3}$$

3. Wave equation
$$\frac{\partial^2 u}{\partial t^2} = c^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) \tag{11.4}$$

These equations look quite similar, but their type and character is fundamentally different. In fact, one needs different techniques for solving them. One thing they have in common however is that, thanks to their linearity, if u_1 and u_2 are two solutions, then all linear combinations $c_1 u_1 + c_2 u_2$ are also solutions, which is the superposition principle.

Often, PDEs are characterized in terms of properties that are called

- elliptic
- parabolic
- hyperbolic

We now discuss how this general characterization of the PDE-type is defined. For linear 2nd-order PDEs of the form (a, b, c not all zero)

$$a\frac{\partial^2 u}{\partial x^2} + b\frac{\partial^2 u}{\partial x \partial y} + c\frac{\partial^2 u}{\partial y^2} + d\frac{\partial u}{\partial x} + e\frac{\partial u}{\partial y} + f\,u = g, \tag{11.5}$$

the classification is based on the discriminant

$$D = b^2 - 4ac. \tag{11.6}$$

The following cases are distinguished:

$$D = \begin{cases} < 0 & \text{elliptic,} \\ 0 & \text{parabolic,} \\ > 0 & \text{hyperbolic.} \end{cases} \tag{11.7}$$

This is done in analogy to conic sections, a connection that can be made explicit by making the replacements $\partial^2 u/\partial x^2 \to X^2$, $\partial^2 u/\partial x \partial y \to XY$, etc.

PDEs that fall in the same class of these types often exhibit similar mathematical and physical properties, and also require similar solution strategies.

Examples for types of 2nd-order PDEs:

Let's go back to our 'three most important PDEs' from above and consider only two independent variables for now.

- For the wave equation we have:

$$\frac{\partial^2 u}{\partial t^2} - c_s^2 \frac{\partial^2 u}{\partial x^2} = 0. \tag{11.8}$$

  So here we can identify $a = 1$, $b = 0$ and $c = -c_s^2$. Hence $D = 4c_s^2 > 0$. This equation is therefore *hyperbolic*.

- The heat conduction equation takes the form:

$$\frac{\partial u}{\partial t} - \lambda^2 \frac{\partial^2 u}{\partial x^2} = 0. \tag{11.9}$$

  Here we have $a = 0$, $b = 0$ and $c = -\lambda^2$, hence $D = 0$. This is a *parabolic* equation.

- Finally, the Laplace equation is

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \tag{11.10}$$

  yielding $a = 1$, $b = 0$, and $c = 1$. Therefore $D = -4 < 0$, and the equation is *elliptic*.

General linear PDEs of 2nd-order with more unknowns than two can be written as

$$\sum_{i,j=1}^{n} a_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_i b_i \frac{\partial u}{\partial x_i} + cu + d = 0. \tag{11.11}$$

Here the eigenvalues of the coefficient matrix a_{ij} can be used to determine the type, according to the following scheme:

- elliptic: all eigenvalues positive, or all negative

- parabolic: one zero eigenvalue, the others all positive or all negative
- hyperbolic: one negative eigenvalue the rest all positive, or one positive the rest all negative

Other cases (multiple eigenvalues of different sign) are sometimes called ultra-hyperbolic. Note that in 2D, this definition should be the same as our previous one. We can readily check that. To make contact with the previous notation, the coefficient matrix is then

$$a_{ij} = \begin{pmatrix} a & b/2 \\ b/2 & a \end{pmatrix}. \tag{11.12}$$

The characteristic equation for the eigenvalues then has solutions

$$\lambda_{1/2} = \frac{(a+c) \pm \sqrt{(a+c)^2 - 4ac + b^2}}{2}. \tag{11.13}$$

This means, for example, that we have two real eigenvalues with opposite signs if $D = b^2 - 4ac > 0$, which corresponds to the hyperbolic case, consistent with our previous simpler definition.

Linear systems of first-order PDEs

Linear systems of first-order homogeneous PDEs can be written as

$$\frac{\partial u_i}{\partial t} + \sum_j A_{ij} \cdot \frac{\partial u_i}{\partial x_j} = 0, \tag{11.14}$$

where $\mathbf{A} = (A_{ij})$ is a coefficient matrix, and one independent variable has been singled out as 'time' t. We may also use a more compact notation of the form

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{A} \cdot \frac{\partial \mathbf{u}}{\partial \mathbf{x}} = 0 \tag{11.15}$$

for this. Here, if $\mathbf{A}$ has only *real eigenvalues* and is diagonalizable, then the PDE system is called *hyperbolic* as well.

This definition can also be extended to non-linear PDEs of the form

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial}{\partial \mathbf{x}} \left(\mathbf{F}(\mathbf{u}) \right) = 0, \tag{11.16}$$

where $\mathbf{F}(\mathbf{u})$ is some function of $\mathbf{u}$. Such PDEs are called conservation laws (compare with the continuity equation), and $\mathbf{F}$ is the flux. One can also write this equation in quasi-linear form by invoking the chain rule:

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{F}}{\partial \mathbf{u}} \cdot \frac{\partial \mathbf{u}}{\partial \mathbf{x}} = 0, \tag{11.17}$$

or more compactly as

$$\frac{\partial \mathbf{u}}{\partial t} + \overline{\mathbf{A}} \cdot \frac{\partial \mathbf{u}}{\partial \mathbf{x}} = 0, \tag{11.18}$$

where $\overline{\mathbf{A}} = \frac{\partial \mathbf{F}}{\partial \mathbf{u}}$ is the Jacobian of the flux. Now, if $\overline{\mathbf{A}}$ is diagonalizable and has real eigenvalues, we still call the (non-linear) set of equations a hyperbolic PDE system.

Character and boundary conditions of different PDE types

The different classes of PDEs show qualitatively different behavior in their solutions. Broadly speaking, the following observations can be made:

- **Hyperbolic PDEs in physics** typically describe dynamical processes of systems that start with some known initial conditions at time t_0. Solutions can develop steep regions or real discontinuities with time, even if they start with perfectly smooth initial states. To specify the initial conditions completely, one needs $u(x, t_0)$ and $\frac{\partial u}{\partial x}(x, t_0)$ and higher derivatives if present, as well as boundary conditions.

- **Parabolic PDEs** are often of second-order and describe slowly changing processes (for example diffusion). Solutions become here *smoother* with time. For describing the problem, one needs the initial state $u(x, t_0)$ as well as boundary conditions.

- **Elliptic PDEs** often describe static problems without time dependence, or equilibrium states of some kind. An important example is the Poisson equation for the gravitational or electrostatic field. The solutions of elliptic problems tend to be as smooth as allowed by the boundary conditions and source terms.

For all types of PDEs, boundary conditions are very important and need to be specified to determine the solution uniquely. If the value of the sought function is specified with a fixed value on the boundaries, one calls this *Dirichlet* boundary conditions. If instead the value of the derivative of u is prescribed on the boundary (usually in the normal direction with respect to the boundary), one has so-called *von Neumann* boundary conditions.

11.2 Solution schemes for PDEs

Unfortunately, for partial differential equations one cannot give a general solution method that works equally well for all types of problems. Rather, each type requires different approaches, and certain PDEs encountered in practice may even be best addressed with special custom techniques built by combining different elements from standard techniques. Important classes of solution schemes include the following:

- **Finite difference methods:** Here the differential operators are approximated through finite difference approximations, usually on a regular (cartesian) mesh, or some other kind of structured mesh (for example a polar grid). An example we already previously discussed is Poisson's equation treated with iterative (multigrid) methods.

- **Finite volume methods:** These may be seen as a subclass of finite difference methods. They are particularly useful for hyperbolic conservation laws. We

shall discuss examples for this approach in applications to fluid dynamics later in this section.

- **Spectral methods:** Here the solution is represented by a linear combination of functions, allowing the PDE to be transformed to algebraic equations or ordinary differential equations. Often this is done by applying Fourier techniques. For example, solving the Poisson equation with FFTs, as we discussed earlier, is a spectral method.

- **Method of lines:** This is a semi-discrete approach where all derivatives except for one are approximated with finite differences. The remaining derivative is then the only one left, so that the remaining problem forms a set of ordinary differential equations (ODEs). Very often, this approach is used in time-dependent problems. One here discretizes space in terms of a set of N points x_i, and for each of these points one obtains an ODE that describes the time evolution of the function at this point. The PDE is transformed in this way into a set of N coupled ODEs. For example, consider the heat diffusion equation in one dimension,

$$\frac{\partial u}{\partial t} + \lambda \frac{\partial^2 u}{\partial x^2} = 0. \tag{11.19}$$

  If we discretize this into a set of points that are spaced h apart, we obtain N equations

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} + \lambda \frac{u_{i+1} + u_{i-1} - 2u_i}{h^2} = 0. \tag{11.20}$$

  These differential equations can now be integrated in time as an ODE system. Note however that this is not necessarily stable. Some problems may require upwinding, i.e. asymmetric forms for the finite difference estimates to recover stability.

- **Finite element methods:** Here the domain is subdivided into "cells" (elements) of fairly arbitrary shape. The solution is then represented in terms of simple, usually polynomial functions on the element, and then the PDE is transformed to an algebraic problem for the coefficients in front of these simple functions. This is hence similar in spirit to spectral methods, except that the expansion is done in terms of highly localized functions on an element by element basis, and is truncated already at low order.

In practice, many different variants of these basic methods exist, and sometimes also combinations of them are used.

11.3 Simple advection

First-order equations of hyperbolic type are particularly useful for introducing the numerical difficulties that then also need to be addressed for more complicated nonlinear conservation laws (e.g. Toro, 1997; LeVeque, 2002; Stone et al., 2008). The

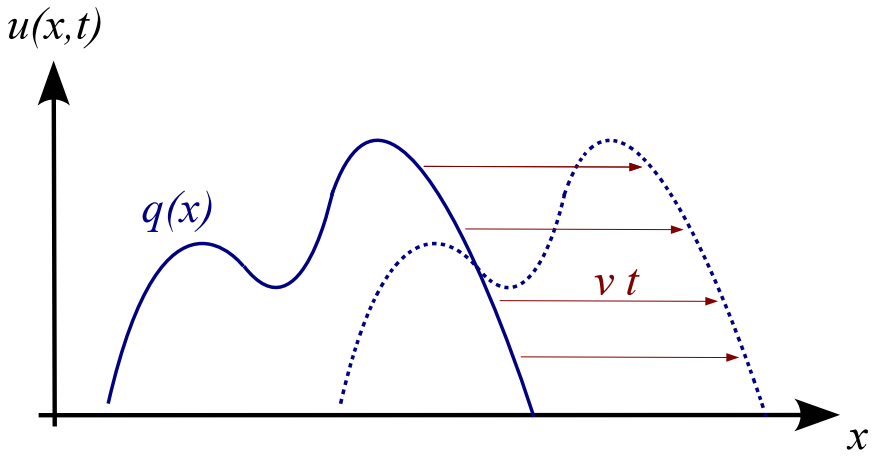

Figure 11.1: Simple advection with constant velocity to the right.

simplest equation of this type is the *advection equation* in one dimension. This is given by

$$\frac{\partial u}{\partial t} + v \cdot \frac{\partial u}{\partial x} = 0, \tag{11.21}$$

where $u = u(x,t)$ is a function of x and t, and v is a constant parameter. This equation is hyperbolic because the so-called coefficient matrix[1] is real and trivially diagonalizable.

If we are given any function $q(x)$, then

$$u(x,t) = q(x - vt) \tag{11.23}$$

is a solution of the PDE, as one can easily check. We can interpret $u(x, t = 0) = q(x)$ as initial condition, and the solution at a later time is then an exact copy of q, simply translated by $v\,t$ along the x-direction, as shown in Fig. 11.1.

Points that start at a certain coordinate x_0 are advected to a new location $x_{\mathrm{ch}}(t) = vt + x_0$. These so-called *characteristics* (see Fig. 11.2), which can be viewed as mediating the propagation of information in the system, are straight lines, all oriented in the downstream direction. Note that "downstream" refers to the direction in which the flow goes, whereas "upstream" is from where the flow comes.

Let's now assume we want to solve the advection problem numerically. (Strictly speaking this is of course superfluous as we have an analytic solution in this case, but we want to see how well a numerical technique would perform here.) We can approach this with a straightforward discretization of u on a special mesh, using for

[1] A linear system of first-order PDEs can be written in the generic form

$$\frac{\partial u_i}{\partial t} + \sum_j A_{ij} \frac{\partial u_i}{\partial x_j} = 0, \tag{11.22}$$

where A_{ij} is the coefficient matrix.

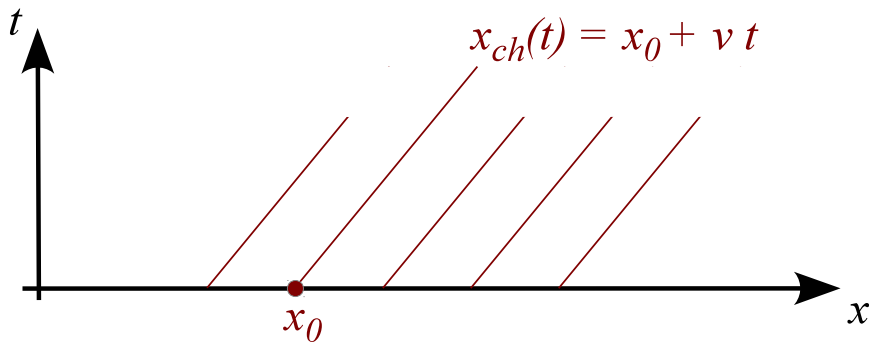

Figure 11.2: A set of flow characteristics for advection to the right with constant velocity v.

example the method of lines. This gives us:

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} + v\frac{u_{i+1} - u_{i-1}}{2h} = 0. \tag{11.24}$$

If we go one step further and also discretize the time derivative with a simple Euler scheme, we get

$$u_i^{(n+1)} = u_i^{(n)} - v\frac{u_{i+1}^{(n)} - u_{i-1}^{(n)}}{2h}\Delta t. \tag{11.25}$$

This is a complete update formula which can be readily applied to a given initial state on the grid. The big surprise is that this turns out to be quite violently unstable! For example, if one applies this to the advection of a step function, one invariably obtains strong oscillatory errors in the downstream region of the step, quickly rendering the numerical solution into complete garbage. What is the reason for this fundamental failure?

- First note that all characteristics (signals) propagate downstream in this problem, or in other words, information strictly travels in the flow direction in this problem.
- But, the information to update u_i is derived both from the upstream (u_{i-1}) and the downstream (u_{i+1}) side.
- According to how the information flows, u_i should not really depend on the downstream side at all, which in some sense is causally disconnected. So let's try to get rid off this dependence by going to a one-sided approximation for the spatial derivative, of the form:

  $$\frac{\mathrm{d}u_i}{\mathrm{d}t} + v\frac{u_i - u_{i-1}}{h} = 0. \tag{11.26}$$

  This is called *upwind differencing*. Interestingly, now the stability problems are completely gone!

- But there are still some caveats to observe: First of all, the discretization now depends on the sign of v. For negative v, one instead has to use

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} + v\frac{u_{i+1} - u_i}{h} = 0. \tag{11.27}$$

  The other is that the solution is not advected in a perfectly faithful way, instead it is quite significantly smoothed out, through a process one calls *numerical diffusion*.

We can actually understand where this strong diffusion in the 1st-order upwind scheme comes from. To this end, let's rewrite the upwind finite difference approximation of the spatial derivative as

$$\frac{u_i - u_{i-1}}{h} = \frac{u_{i+1} - u_{i-1}}{2h} - \frac{u_{i+1} - 2u_i + u_{i-1}}{2h}. \tag{11.28}$$

Hence our stable upwind scheme can also be written as

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} + v\frac{u_{i+1} - u_{i-1}}{2h} = \frac{vh}{2}\,\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2}. \tag{11.29}$$

But recall from equation (7.3) that

$$\left(\frac{\partial^2 u}{\partial x^2}\right)_i \simeq \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2}, \tag{11.30}$$

so if we define a diffusion constant $D = (vh)/2$, we are effectively solving the following problem,

$$\frac{\partial u}{\partial t} + v \cdot \frac{\partial u}{\partial x} = D\frac{\partial^2 u}{\partial x^2}, \tag{11.31}$$

and not the original advection problem. The diffusion term on the right hand side is here a byproduct of the numerical algorithm that we have used. We needed to add this numerical diffusion in order to obtain stability of the integration.

Note however that for better grid resolution, $h \to 0$, the diffusion becomes smaller, so in this limit one obtains an ever better solution. Also note that the diffusivity becomes larger for larger velocity v, so the faster one needs to advect, the stronger the numerical diffusion effects become.

Besides the upwinding requirement, integrating a hyperbolic conservation law with an explicit method in time also requires the use of a sufficiently small integration timestep, not only to get sufficiently good accuracy, but also for reasons of *stability*. In essence, there is a maximum timestep that may be used before the integration brakes down. How large can we make this timestep? Again, we can think about this in terms of information travel. If the timestep exceeds $\Delta t_{\max} = h/v$, then the updating of u_i would have to include information from u_{i-2}, but if we don't do this, the updating will likely become unstable.

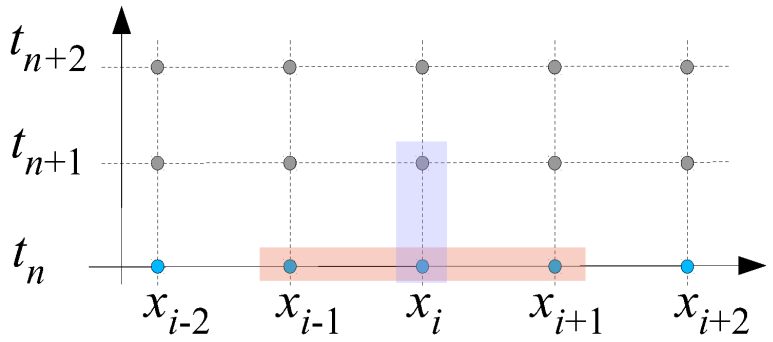

Figure 11.3: A discretization scheme for the continuity equation in one spatial dimension. The red and blue boxes mark the stencils that are applied for calculating the spatial and time derivatives.

This leads to the so-called *Courant-Friedrichs-Levy* (CFL) timestep condition (Courant et al., 1928), which for this problem takes the form

$$\Delta t \leq \frac{h}{v}. \tag{11.32}$$

This is a necessary but not sufficient condition for any explicit finite different approach of the hyperbolic advection equation. For other hyperbolic conservation laws, similar CFL-conditions apply.

Hyperbolic conservation laws

We now consider a hyperbolic conservation law, such as the continuity equation for the mass density of a fluid:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0. \tag{11.33}$$

We see that this is effectively the advection equation, but with a spatially variable velocity $\mathbf{v} = \mathbf{v}(\mathbf{x})$. Here $\mathbf{F} = \rho \mathbf{v}$ is the mass flux.

Let's study the problem in one spatial dimension, and consider a discretization both of the x- and t-axis. This corresponds to

$$\frac{\rho_i^{(n+1)} - \rho_i^{(n)}}{\Delta t} + \frac{F_{i+1}^{(n)} - F_{i-1}^{(n)}}{2\Delta x} = 0, \tag{11.34}$$

leading to the update rule

$$\rho_i^{(n+1)} = \rho_i^{(n)} + \frac{\Delta t}{2\Delta x}\left(F_{i-1}^{(n)} - F_{i+1}^{(n)}\right). \tag{11.35}$$

This is again found to be highly unstable, for the same reasons as in the plain advection problem: we have not observed in 'which direction the wind blows', or in other words, we have ignored in which direction the local characteristics point. For example, if the mass flux is to the right, we know that the characteristics point also

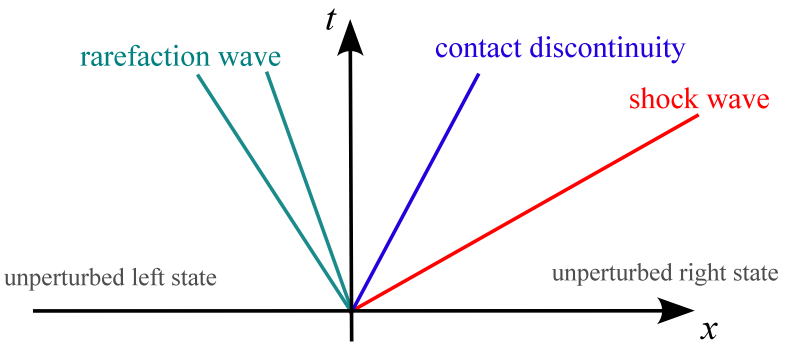

Figure 11.4: Wave structure of the solution of the Riemann problem. The central contact wave separates the original fluid phases. On the left and the right, there is either a shock or a rarefaction wave.

to the right. The upwind direction is therefore towards negative x, and by using only this information in making our spatial derivative one-sided, we should be able to resurrect stability.

Now, for the mass continuity equation identifying the local characteristics is quite easy, and in fact, their direction can simply be inferred from the sign of the mass flux. However, in more general situations for systems of non-linear PDEs, this is far less obvious. Here we need to use a so-called Riemann solvers to give us information about the local solution and the local characteristics (Toro, 1997). This then also implicitly identifies the proper upwinding that is needed for stability.

11.4 Riemann problem

The Riemann problem is an initial value problem for a hyperbolic system, consisting of two piece-wise constant states (two half-spaces) that meet at a plane at $t = 0$. The task is then to solve for the subsequent evolution at $t > 0$.

An important special case is the Riemann problem for the Euler equations (i.e. for ideal gas dynamics). Here the left and right states of the interface, can, for example, be uniquely specified by giving the three "primitive" variables density, pressure and velocity, viz.

$$U_L = \begin{pmatrix} \rho_L \\ P_L \\ \mathbf{v}_L \end{pmatrix}, \quad U_R = \begin{pmatrix} \rho_R \\ P_R \\ \mathbf{v}_R \end{pmatrix}. \tag{11.36}$$

Alternatively one can also specify density, momentum density, and energy density. For an ideal gas, this initial value problem can be solved analytically (Toro, 1997), modulo an implicit equation which requires numerical root-finding, i.e. the solution cannot be written down explicitly. The solution always contains characteristics for three self-similar waves, as shown schematically in Fig. 11.4. Some notes on this:

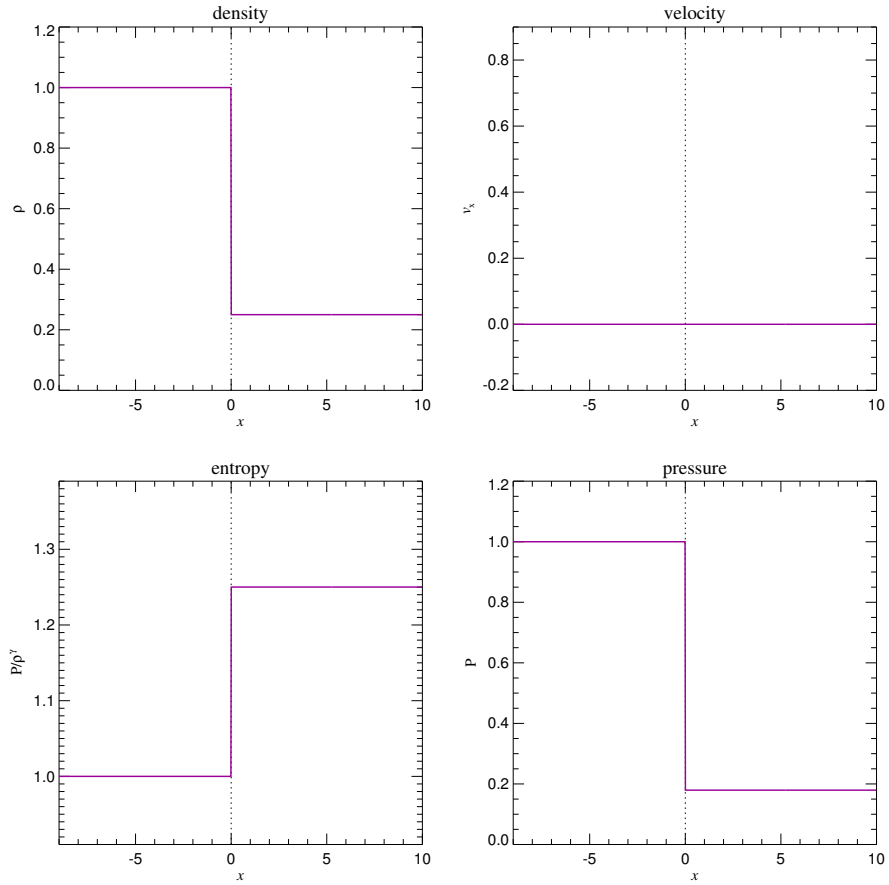

Figure 11.5: Initial state of an example Riemann problem, composed of two phases in different states that are brought into contact at $x = 0$ at time $t = 0$. (Since $v_x = 0$, the initial conditions are actually an example of the special case of a Sod shock-tube problem.)

- The middle wave is always present and is a contact wave that marks the boundary between the original fluid phases from the left and right sides.

- The contact wave is sandwiched between a shock or a rarefaction wave on either side (it is possible to have shocks on both sides, or rarefactions on both sides, or one of each). The rarefaction wave is not a single characteristic but rather a rarefaction fan with a beginning and an end.

- These waves propagate with constant speed. If the solution is known at some time $t > 0$, it can also be obtained at any other time through a suitable scaling transformation. An important corollary is that at $x = 0$, the fluid quantities $(\rho^\star, P^\star, \mathbf{v}^\star)$ are *constant in time* for $t > 0$.

- For $\mathbf{v}_L = \mathbf{v}_R = 0$, the Riemann problem simplifies and becomes the 'Sod shock tube' problem.

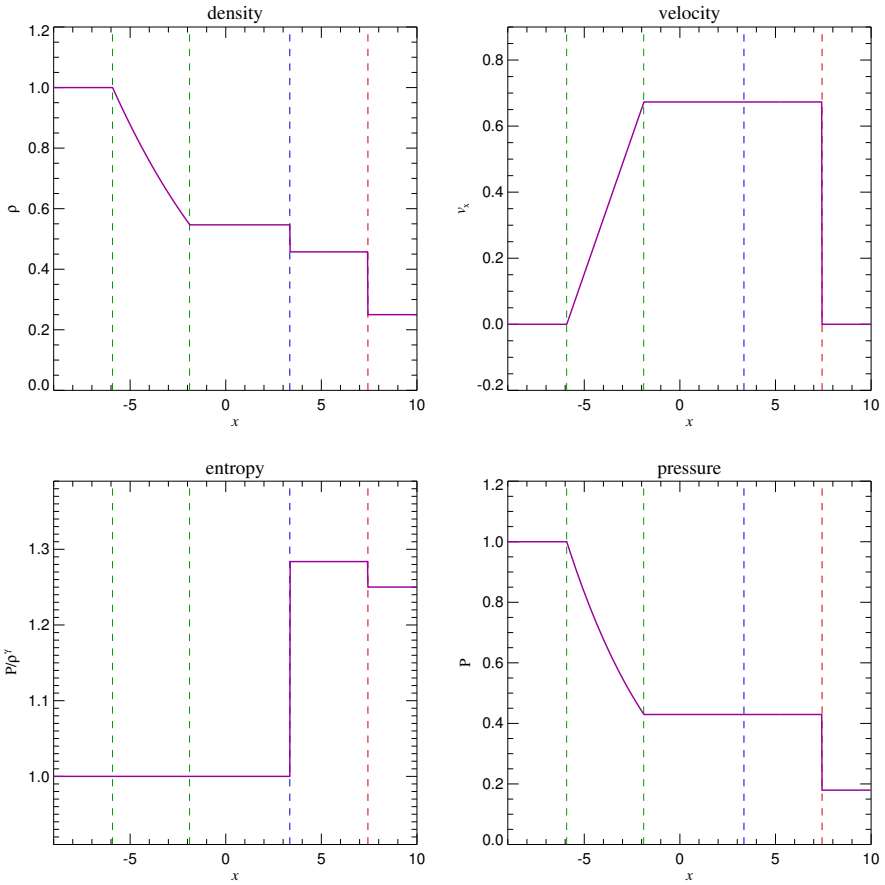

Figure 11.6: Evolved state at $t = 5.0$ of the initial fluid state displayed in Fig. 11.5. The blue dashed line marks the position of the contact wave, the green dashed lines give the location of the rarefaction fan, and the red dashed line marks the shock.

Let's consider an example how this wave structure looks in a real Riemann problem. We consider, for definiteness, a Riemann problem with $\rho_L = 1.0$, $P_L = 1.0$, $v_L = 0$, and $\rho_R = 0.25$, $P_R = 0.1795$, $v_R = 0$ (which is of Sod-shock type). The adiabatic exponent is taken to be $\gamma = 1.4$. We hence deal at $t = 0.0$ with the initial state displayed in Figure 11.5. After time $t = 5.0$, the wave structure formed by a rarefaction to the left (location marked in green), a contact in the middle (blue) and a shock to the right (red) can be nicely seen in Figure 11.6.

Some general properties of the waves appearing in the Riemann problem can be summarized as follows:

- *Shock:* This is a sudden compression of the fluid, associated with an irreversible conversion of kinetic energy to heat, i.e. here entropy is produced. The density, normal velocity component, pressure, and entropy all change discontinuously at a shock.

- *Contact discontinuity:* This traces the original separating plane between the two fluid phases that have been brought into contact. Pressure as well as the normal velocity are constant across a contact, but density, entropy and temperature can jump.
- *Rarefaction wave:* This occurs when the gas (suddenly) expands. The rarefaction wave smoothly connects two states over a finite spatial region; there are no discontinuities in any of the fluid variables.

11.5 Finite volume discretization

Let's now take a look how Riemann solvers can be used in the finite volume discretization approach to the PDEs of fluid dynamics. Recall that we can write our hyperbolic conservation laws as

$$\frac{\partial \mathbf{U}}{\partial t} + \nabla \cdot \mathbf{F} = 0. \tag{11.37}$$

Here $\mathbf{U}$ is a state vector and $\mathbf{F}$ is the flux vector. For example, the Euler equations of section 10.1.1 can be written in the form

$$\mathbf{U} = \begin{pmatrix} \rho \\ \rho\mathbf{v} \\ \rho e \end{pmatrix}, \qquad \mathbf{F} = \begin{pmatrix} \rho\mathbf{v} \\ \rho\mathbf{v}\mathbf{v}^T + P \\ (\rho e + P)\mathbf{v} \end{pmatrix}, \tag{11.38}$$

with the specific energy $e = u + \mathbf{v}^2/2$ and u being the thermal energy per unit mass. The ideal gas equation gives the pressure as $P = (\gamma - 1)\rho u$ and provides a closure for the system.

In a finite volume scheme, we describe the system through the averaged state over a set of finite cells. These cell averages are defined as

$$\mathbf{U}_i = \frac{1}{V_i} \int_{\text{cell } i} \mathbf{U}(\mathbf{x})\, \mathrm{d}V. \tag{11.39}$$

Let's now see how we could devise an update scheme for these cell-averaged quantities.

1. We start by integrating the conservation law over a cell, and over a finite interval in time:
$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathrm{d}x \int_{t_n}^{t_{n+1}} \mathrm{d}t \left(\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} \right) = 0. \tag{11.40}$$
2. This gives
$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathrm{d}x \left[\mathbf{U}(x, t_{n+1}) - \mathbf{U}(x, t_n) \right] + \int_{t_n}^{t_{n+1}} \mathrm{d}t \left[\mathbf{F}(x_{i+\frac{1}{2}}, t) - \mathbf{F}(x_{i-\frac{1}{2}}, t) \right] = 0. \tag{11.41}$$

In the first term, we recognize the definition of the cell average:

$$\mathbf{U}_i^{(n)} \equiv \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathbf{U}(x, t_n) \mathrm{d}x. \tag{11.42}$$

Hence we have

$$\Delta x \left[\mathbf{U}_i^{(n+1)} - \mathbf{U}_i^{(n)}\right] + \int_{t_n}^{t_{n+1}} \mathrm{d}t \left[\mathbf{F}(x_{i+\frac{1}{2}}, t) - \mathbf{F}(x_{i-\frac{1}{2}}, t)\right] = 0. \tag{11.43}$$

3. Now, $\mathbf{F}(x_{i+\frac{1}{2}}, t)$ for $t > t_n$ is given by the solution of the Riemann problem with left state $\mathbf{U}_i^{(n)}$ and right state $\mathbf{U}_{i+1}^{(n)}$. At the interface, this solution is *independent* of time. We can hence write

$$\mathbf{F}(x_{i+\frac{1}{2}}, t) = \mathbf{F}^\star_{i+\frac{1}{2}}, \tag{11.44}$$

where $\mathbf{F}^\star_{i+\frac{1}{2}} = \mathbf{F}_{\text{Riemann}}(\mathbf{U}_i^{(n)}, \mathbf{U}_{i+1}^{(n)})$ is a short-hand notation for the corresponding Riemann solution sampled at the interface. Hence we now get

$$\Delta x \left[\mathbf{U}_i^{(n+1)} - \mathbf{U}_i^{(n)}\right] + \Delta t \left[\mathbf{F}^\star_{i+\frac{1}{2}} - \mathbf{F}^\star_{i-\frac{1}{2}}\right] = 0. \tag{11.45}$$

Or alternative, as an explicit update formula:

$$\mathbf{U}_i^{(n+1)} = \mathbf{U}_i^{(n)} + \frac{\Delta t}{\Delta x} \left[\mathbf{F}^\star_{i-\frac{1}{2}} - \mathbf{F}^\star_{i+\frac{1}{2}}\right]. \tag{11.46}$$

The first term in the square bracket gives the flux that flows from left into the cell, the second term is the flux out of the cell on its right side. The idea to use the Riemann solution in the updating step is due to Godunov, that's why such schemes are often called *Godunov schemes*.

It is worthwhile to note that we haven't really made any approximation in the above (yet). In particular, if we calculate $\mathbf{F}_{\text{Riemann}}$ analytically (and hence exactly), then the above seems to account for the correct fluxes for arbitrarily long times. So does this mean that we get a perfectly accurate result even for very large timesteps? This certainly sounds too good to be true, so there must be a catch somewhere.

Indeed, there is. First of all, we have assumed that the Riemann problems are independent of each other and each describe infinite half-spaces. This is not true once we consider finite volume cells, but it is still ok for a while as long t_{n+1} is close enough to t_n such that the waves emanating in one interface have not yet arrived at the next interface left or right. This then leads to a CFL-timestep criterion, were $\Delta t \leq \Delta x / c_{\max}$ and $c_{\max}$ is the maximum wavespeed.

Another point is more subtle and comes into play when we consider more than one timestep. We assumed that the $\mathbf{U}_i^{(n)}$ describe piece-wise constant states which can then be fed to the Riemann solver to give us the flux. However, even when this is true initially, we have just seen that after one timestep it will not be true anymore. By ignoring this in the subsequent timestep (which is done by performing an averaging step that washes out the cell substructure that developed as part of the evolution during the previous timestep) we make some error.

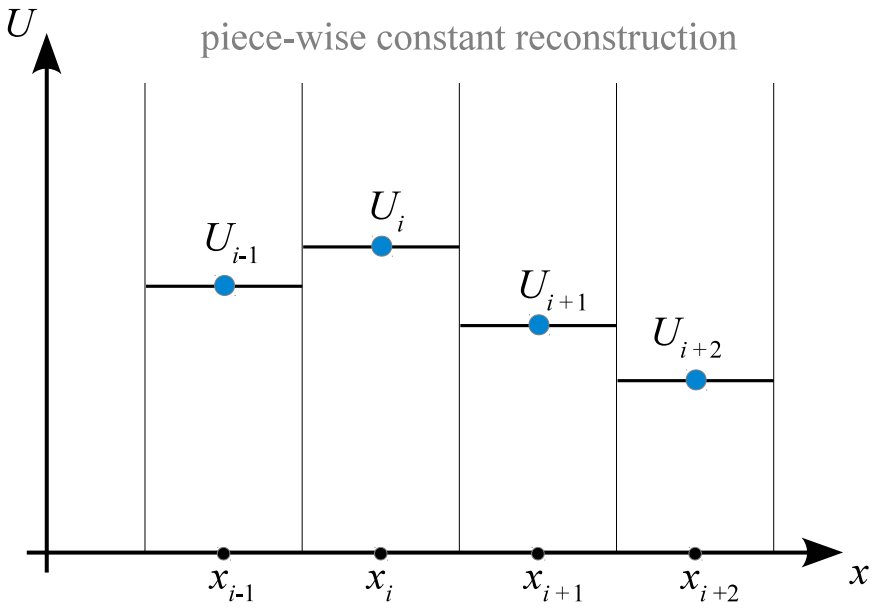

Figure 11.7: Piece-wise constant states of a fluid forming the simplest possible reconstruction of its state based on a set of discrete values U_i known at spatial positions x_i.

11.6 Godunov's method and Riemann solvers

It is useful to introduce another interpretation of common finite-volume discretizations of fluid dynamics, so-called Reconstruct-Evolve-Average (REA) schemes. We also use this here for a short summary of Godunov's important method, and the way Riemann solvers come into play in it.

An REA update scheme of a hydrodynamical system discretized on a mesh can be viewed as a sequence of three steps:

1. *Reconstruct:* Using the cell-averaged quantities (as shown in Fig. 11.7), this defines the run of these quantities everywhere in the cell. In the sketch, a piecewise constant reconstruction is assumed, which is the simplest procedure one can use and leads to 1st order accuracy.

2. *Evolve:* The reconstructed state is then evolved forward in time by Δt. In Godunov's approach, this is done by treating each cell interface as a piecewise constant initial value problem which is solved with the Riemann solver exactly or approximately. This solution is formally valid as long as the waves emanating from opposite sides of a cell do not yet start to interact. In practice, one therefore needs to limit the timestep Δt such that this does not happen.

3. *Average:* The wave structure resulting from the evolution over timestep Δt is spatially averaged in a conservative fashion to compute new states $\mathbf{U}^{n+1}$ for each cell. Fortunately, the averaging step does not need to be done explicitly;

instead it can simply be carried out by accounting for the fluxes that enter or leave the control volume of the cell. Then the whole cycle repeats again.

What is needed for the *evolve* step is a prescription to either exactly or approximately solve the Riemann problem for a piece-wise linear left and right state that are brought into contact at time $t = t_n$. Formally, this can be written as

$$\mathbf{F}^\star = \mathbf{F}_{\text{Riemann}}(\mathbf{U}_L, \mathbf{U}_R). \tag{11.47}$$

In practice, a variety of approximate Riemann solvers $\mathbf{F}_{\text{Riemann}}$ are commonly used in the literature (Rusanov, 1961; Harten et al., 1983; Toro, 1997; Miyoshi & Kusano, 2005). For the ideal gas and for isothermal gas, it is also possible to solve the Riemann problem exactly, but not in closed form (i.e. the solution involves an iterative root finding of a non-linear equation).

There are now two main issues left:

- How can this be extended to multiple spatial dimensions?
- How can it be extended such that a higher order integration accuracy both in space and time is reached?

We'll discuss these issues next.

11.7 Extensions to multiple dimensions

So far, we have considered *one-dimensional* hyperbolic conservation laws of the form

$$\partial_t \mathbf{U} + \partial_x \mathbf{F}(\mathbf{U}) = 0, \tag{11.48}$$

where ∂_t is a short-hand notation for $\partial_t = \frac{\partial}{\partial t}$, and similarly $\partial_x = \frac{\partial}{\partial x}$. For example, for isothermal gas with soundspeed c_s, the state vector $\mathbf{U}$ and flux vector $\mathbf{F}(\mathbf{U})$ are given as

$$\mathbf{U} = \begin{pmatrix} \rho \\ \rho u \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} \rho u \\ \rho u^2 + \rho c_s^2 \end{pmatrix}, \tag{11.49}$$

where u is the velocity in the x-direction.

In three dimensions, the PDEs describing a fluid become considerably more involved. For example, the Euler equations for an ideal gas are given in explicit form as

$$\partial_t \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ \rho e \end{pmatrix} + \partial_x \begin{pmatrix} \rho u \\ \rho u^2 + P \\ \rho uv \\ \rho uw \\ \rho u(\rho e + P) \end{pmatrix} + \partial_y \begin{pmatrix} \rho v \\ \rho uv \\ \rho v^2 + P \\ \rho vw \\ \rho v(\rho e + P) \end{pmatrix} + \partial_z \begin{pmatrix} \rho w \\ \rho uw \\ \rho vw \\ \rho w^2 + P \\ \rho w(\rho e + P) \end{pmatrix} = 0, \tag{11.50}$$

where $e = e_{\rm therm} + (u^2 + v^2 + w^2)/2$ is the total specific energy per unit mass, $e_{\rm therm}$ is the thermal energy per unit mass, and $P = (\gamma - 1)\rho\, e_{\rm therm}$ is the pressure. These equations are often written in the following notation:

$$\partial_t \mathbf{U} + \partial_x \mathbf{F} + \partial_y \mathbf{G} + \partial_z \mathbf{H} = 0. \tag{11.51}$$

Here the functions $\mathbf{F}(\mathbf{U})$, $\mathbf{G}(\mathbf{U})$ and $\mathbf{H}(\mathbf{U})$ give the flux vectors in the x-, y- and z-direction, respectively.

11.7.1 Dimensional splitting

Let us now consider the three dimensionally split problems derived from equation (11.51):

$$\partial_t \mathbf{U} + \partial_x \mathbf{F} = 0, \tag{11.52}$$

$$\partial_t \mathbf{U} + \partial_y \mathbf{G} = 0, \tag{11.53}$$

$$\partial_t \mathbf{U} + \partial_z \mathbf{H} = 0. \tag{11.54}$$

Note that the vectors appearing here have still the same dimensionality as in the full equations. They are *augmented* one-dimensional problems, i.e. the transverse variables still appear but spatial differentiation happens only in one direction. Because of this, these additional transverse variables do not make the 1D problem more difficult compared to the 'pure' 1D problem considered earlier, but the fluxes appearing in them still need to be included.

Now let us assume that he have a method to solve/advance each of these one-dimensional problems. We can for example express this formally through time-evolution operators $\mathcal{X}(\Delta t)$, $\mathcal{Y}(\Delta t)$, and $\mathcal{Z}(\Delta t)$, which advance the solution by a timestep Δt. Then the full time advance of the system can for example be approximated by

$$\mathbf{U}^{n+1} \simeq \mathcal{Z}(\Delta t)\mathcal{Y}(\Delta t)\mathcal{X}(\Delta t)\mathbf{U}^n. \tag{11.55}$$

This is one possible dimensionally split update scheme. In fact, this is the exact solution if equations (11.52)-(11.53) represent the linear advection problem, but for more general non-linear equations it only provides a first order approximation. However, higher-order dimensionally split update schemes can also be easily constructed. For example, in two-dimensions,

$$\mathbf{U}^{n+1} = \frac{1}{2}[\mathcal{X}(\Delta t)\mathcal{Y}(\Delta t) + \mathcal{Y}(\Delta t)\mathcal{X}(\Delta t)]\mathbf{U}^n \tag{11.56}$$

and

$$\mathbf{U}^{n+1} = \mathcal{X}(\Delta t/2)\mathcal{Y}(\Delta t)\mathcal{X}(\Delta t/2)\mathbf{U}^n \tag{11.57}$$

are second-order accurate. Similarly, for three dimensions the scheme

$$\mathbf{U}^{n+1} = \mathcal{X}(\Delta t/2)\mathcal{Y}(\Delta t/2)\mathcal{Z}(\Delta t)\mathcal{Y}(\Delta t/2)\mathcal{X}(\Delta t/2)\mathbf{U}^n \tag{11.58}$$

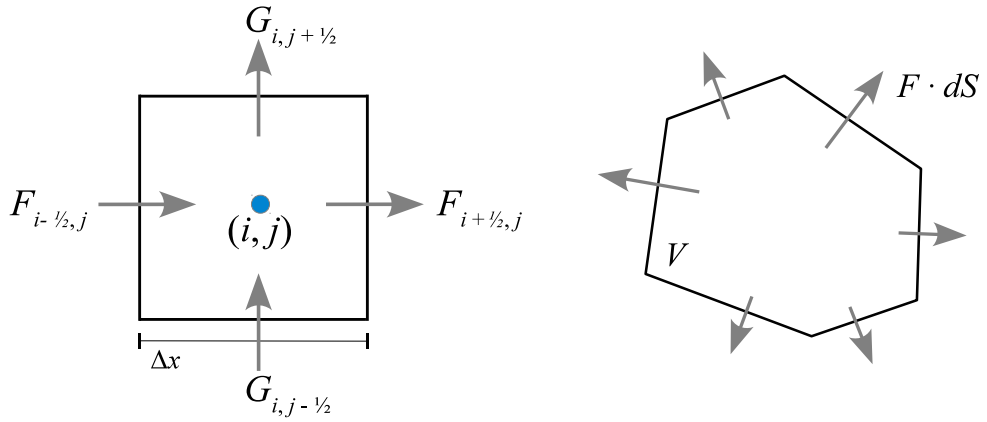

Figure 11.8: Sketch of unsplit finite-volume update schemes. On the left, the case of a structured Cartesian grid is shown, the case on the right is for an unstructured grid.

is second-order accurate. As a general rule of thumb, the time evolution operators have to be applied alternatingly in reverse order to reach second-order accuracy. We see that the dimensionless splitting reduces the problem effectively to a sequence of one-dimensional solution operations which are applied to multi-dimensional domains. Note that each one-dimensional operator leads to an update of $\mathbf{U}$, and is a complete step for the corresponding augmented one-dimensional problem. Gradients, etc., that are needed for the next step then have to be recomputed before the next time-evolution operator is applied. In practical applications of mesh codes, these one-dimensional solves are often called *sweeps*.

11.7.2 Unsplit schemes

In an unsplit approach, all flux updates of a cell are applied simultaneously to a cell, not sequentially. This is for example illustrated in 2D in the situations depicted in Figure 11.8. The unsplit update of cell i, j in the Cartesian case is then given by

$$U_{i,j}^{n+1} = U_{i,j}^{n} + \frac{\Delta t}{\Delta x}\left(\mathbf{F}_{i-\frac{1}{2},j} - \mathbf{F}_{i+\frac{1}{2},j}\right) + \frac{\Delta t}{\Delta y}\left(\mathbf{G}_{i,j-\frac{1}{2}} - \mathbf{G}_{i,j+\frac{1}{2}}\right). \tag{11.59}$$

Unsplit approaches can also be used for irregular shaped cells like those appearing in unstructured meshes (see Fig. 11.8). For example, integrating over a cell of volume V and denoting with $\mathbf{U}$ the cell average, we can write the cell update with the divergence theorem as

$$\mathbf{U}^{n+1} = \mathbf{U}^{n} - \frac{\Delta t}{V}\int \mathbf{F} \cdot \mathbf{dS}, \tag{11.60}$$

where the integration is over the whole cell surface, with outwards pointing face area vectors $\mathbf{dS}$.

11.8 Extensions for high-order accuracy

We should first clarify what we mean with higher order schemes. Loosely speaking, this refers to the convergence rate of a scheme in smooth regions of a flow. For example, if we know the analytic solution $\rho(x)$ for some problem, and then obtain a numerical result ρ_i at a set of N points at locations x_i, we can ask what the typical error of the solution is. One possibility to quantify this would be through a L1 error norm, for example in the form

$$L1 = \frac{1}{N} \sum_i |\rho_i - \rho(x_i)|, \tag{11.61}$$

which can be interpreted as the average error per cell. If we now measure this error quantitatively for different resolutions of the applied discretization, we would like to find that L1 decreases with increasing N. In such a case our numerical scheme is converging, and provided we use sufficient numerical resources, we have a chance to get below any desired absolute error level. But the *rate of convergence* can be very different between different numerical schemes when applied to the same problem. If a method shows a $L1 \propto N^{-1}$ scaling, it is said to be first-order accurate; a doubling of the number of cells will then cut the error in half. A second-order method has $L1 \propto N^{-2}$, meaning that a doubling of the number of cells can actually reduce the error by a factor of 4. This much better convergence rate is of course highly desirable. It is also possible to construct schemes with still higher convergence rates, but they tend to quickly become much more complex and computationally involved, so that one eventually reaches a point of diminishing return, depending on the specific type of problem. But the extra effort one needs to make to go from first to second-order is often very small, sometimes trivially small, so that one basically should always strive to try at least this.

A first step in constructing a 2nd order extension of Godunov's method is to replace the piece-wise constant with a piece-wise linear reconstruction. This requires that one first estimates gradients for each cell (usually by a simple finite difference formula). These are then slope-limited if needed such that the linear extrapolations of the cell states to the cell interfaces do not introduce new extrema. This slope-limiting procedure is quite important; it needs to be done to avoid that real fluid discontinuities introduce large spurious oscillations into the fluid.

Given slope limited gradients, for example $\nabla\rho$ for the density, one can then estimate the left and right states adjacent to an interface $x_{i+\frac{1}{2}}$ by spatial extrapolation from the centers of the cells left and right from the interface:

$$\rho^L_{i+\frac{1}{2}} = \rho_i + (\nabla\rho)_i \frac{\Delta x}{2}, \tag{11.62}$$

$$\rho^R_{i+\frac{1}{2}} = \rho_{i+1} - (\nabla\rho)_{i+1} \frac{\Delta x}{2}. \tag{11.63}$$

The next step would in principle be to use these states in the Riemann solver. In doing this we will ignore the fact that our reconstruction has now a gradient over

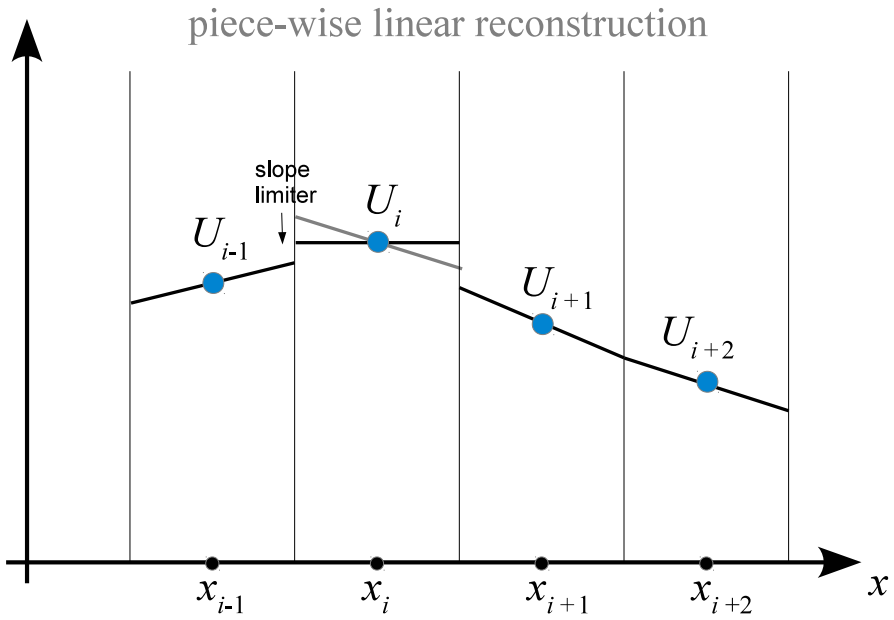

Figure 11.9: Piece-wise linear reconstruction scheme applied to a fluid state represented through a regular mesh.

the cell; instead we still pretend that the fluid state can be taken as piece-wise constant left and right of the interface as far as the Riemann solver is concerned. However, it turns out that the spatial extrapolation needs to be augmented with a temporal extrapolation one half timestep into the future, such that the flux estimate is now effectively done in the middle of the timestep. This is necessary both to reach second-order accuracy in time and also for stability reasons. Hence we really need to use

$$\rho^L_{i+\frac{1}{2}} = \rho_i + (\nabla\rho)_i \frac{\Delta x}{2} + \left(\frac{\partial\rho}{\partial t}\right)_i \frac{\Delta t}{2}, \tag{11.64}$$

$$\rho^R_{i+\frac{1}{2}} = \rho_{i+1} - (\nabla\rho)_{i+1} \frac{\Delta x}{2} + \left(\frac{\partial\rho}{\partial t}\right)_{i+1} \frac{\Delta t}{2}, \tag{11.65}$$

for extrapolating to the interfaces. More generally, this has to be done for the whole state vector of the system, i.e.

$$\mathbf{U}^L_{i+\frac{1}{2}} = \mathbf{U}_i + (\partial_x\mathbf{U})_i \frac{\Delta x}{2} + (\partial_t\mathbf{U})_i \frac{\Delta t}{2}, \tag{11.66}$$

$$\mathbf{U}^R_{i+\frac{1}{2}} = \mathbf{U}_{i+1} - (\partial_x\mathbf{U})_{i+1} \frac{\Delta x}{2} + (\partial_t\mathbf{U})_{i+1} \frac{\Delta t}{2}. \tag{11.67}$$

Note that here the quantity $(\partial_x\mathbf{U})_i$ is a (slope-limited) *estimate* of the gradient in cell i, based on finite-differences plus a slope limiting procedure. Similarly, we somehow need to estimate the time derivative encoded in $(\partial_t\mathbf{U})_i$. How can this be

done? One way to do this is to exploit the Jacobian matrix of the Euler equations. We can write the Euler equations as

$$\partial_t \mathbf{U} = -\partial_x \mathbf{F}(\mathbf{U}) = -\frac{\partial \mathbf{F}}{\partial \mathbf{U}} \partial_x \mathbf{U} = -\mathbf{A}(\mathbf{U}) \partial_x \mathbf{U}, \tag{11.68}$$

where $\mathbf{A}(\mathbf{U})$ is the Jacobian matrix. Using this, we can simply estimate the required time-derivative based on the spatial derivatives:

$$(\partial_t \mathbf{U})_i = -\mathbf{A}(\mathbf{U}_i) (\partial_x \mathbf{U})_i. \tag{11.69}$$

Hence the extrapolation can be done as

$$\mathbf{U}^L_{i+\frac{1}{2}} = \mathbf{U}_i + \left[\frac{\Delta x}{2} - \frac{\Delta t}{2}\mathbf{A}(\mathbf{U}_i)\right] (\partial_x \mathbf{U})_i, \tag{11.70}$$

$$\mathbf{U}^R_{i+\frac{1}{2}} = \mathbf{U}_{i+1} + \left[-\frac{\Delta x}{2} - \frac{\Delta t}{2}\mathbf{A}(\mathbf{U}_{i+1})\right] (\partial_x \mathbf{U})_{i+1}. \tag{11.71}$$

This procedure defines the so-called MUSCL-Hancock scheme (van Leer, 1984; Toro, 1997; van Leer, 2006), which is a 2nd-order accurate extension of Godunov's method.

Higher-order extensions such as the piece-wise parabolic method (PPM) start out with a higher order polynomial reconstruction. In the case of PPM, parabolic shapes are assumed in each cell instead of piece-wise linear states. The reconstruction is still guaranteed to be conservative, i.e. the integral underneath the reconstruction recovers the total values of the conserved variables individually in each cell. So-called ENO and WENO schemes (e.g. Balsara et al., 2009) use yet higher-order polynomials to reconstruct the state in a conservative fashion. Here many more cells in the environment need to be considered (i.e. the so-called *stencil* of these methods is much larger) to robustly determine the coefficients of the reconstruction. This can for example involve a least-square fitting procedure (Ollivier-Gooch, 1997).