Notes on astrophysical hydrodynamics

Filippo Fraternali

Kapteyn Astronomical Institute, University of Groningen
Email address: fraternali@astro.rug.nl

Contents

CHAPTER 1

Introduction

The aim of this Chapter is to introduce the reader to the astrophysical relevance of hydrodynamics. We start with a brief description of the gaseous component of the interstellar medium (ISM) of galaxies. This is meant to be a very general introduction as more details can be found in other tests [Tie05, Dra11, CFN19]. The ISM is an example of an astrophysical gas that can be studied with the use of hydrodynamics. However, the application of hydrodynamics extends to the study of other astrophysical systems, including stellar interiors and the medium between galaxies, which we call **intergalactic medium** (IGM). In fact, hydrodynamics and magneto-hydrodynamics are fundamental tools for our theoretical understanding of galaxy formation and evolution as well as star formation. In the second part of this Chapter we describe an important physical property of astrophysical gases and the ISM in particular, which is the fact that they can be considered *collisional*. At the end of the Chapter, we give a brief summary of the most important equilibria in astrophysical gases.

1.1. The interstellar medium

With the term **interstellar medium** we indicate the non-stellar baryonic component of galaxies, literally located "between the stars". In practice, for star-forming galaxies, this medium is mostly distributed in a disc roughly co-spatial with the stellar disc, however, as we see below, it may be more extended in the radial direction. For early-type galaxies the ISM is mostly made of hot gas, although small amounts of molecular and neutral gas are also detected in these systems. The ISM is also referred to as **diffuse matter**. Below we describe its major components dividing them in terms of the atomic or molecular nature of the gas and their physical state (neutral or ionised). Note the we focus here on the *gaseous* ISM, leaving the description of the interstellar dust to other sources [e.g. Dra03].

1.1.1. Atomic neutral gas. It is mostly composed of hydrogen, with a percentage in atoms of helium of about 10% which translates into a **atomic mass** $\mu \approx 1.3$. [1] Chemical elements heavier than helium, called metals in astronomy, are rare in the ISM of any galaxy and do not appreciably contribute to the mass budget. They are, however, very important for several aspects of the ISM physical state and evolution, in particular the cooling processes, as we see below.

In nearby galaxies, the atomic neutral hydrogen is mainly observed through the **21-cm line**, i.e. the hyperfine transition from parallel to anti-parallel spins of the electron and the proton in the hydrogen atom at its lowest energy state. The Einstein coefficient of hyperfine spontaneous emission is $A_{10} \simeq 2.88 \times 10^{-15}\,\mathrm{s}^{-1}$,

[1] The atomic (or molecular) mass is the average mass of gas particles in a medium given in units of the proton mass (nearly equivalent to the hydrogen mass), m_p.

which translates into a spontaneous emission timescale of $\tau_{21-\mathrm{cm}} \simeq 10^7\,\mathrm{yr}$. This is a very long time and it is only thanks to the high abundance of hydrogen atoms that this emission can be observed. The emission coming from this transition is often simply called H I emission and it is a very effective way to study the distribution of gas in local galaxies. In the ISM, collisions[2] are frequent and the excited state of the H I transition can be collisionally induced. Given that the collision time is much shorter than the time of spontaneous emission, the hyperfine excitation and most of the de-excitation of hydrogen are collisional. About 1/1000 de-excitations are, however, spontaneous and they produce the emission of photons at $\lambda_{\mathrm{H\,I}} \simeq 21.106\,\mathrm{cm}$ or $\nu_{\mathrm{H\,I}} \simeq 1.420406\,\mathrm{GHz}$. As a consequence of the frequent collisions, the H I emission can be considered *thermal*.

In the Milky Way's disc, around the Solar circle, there is about 1 particle of hydrogen per cubic centimetre, this is a number that is not much different in other galaxies and it is useful to remember[3]. However, the density of H I is usually given in term of column density, which is a quantity that is more directly observable, especially in external galaxies. The typical column densities of a galactic disc are in the range $N_{\mathrm{H\,I}} \approx 10^{20} - 10^{21}\,\mathrm{atoms\,cm^{-2}}$, which corresponds to about $1 - 10\,\mathrm{M_\odot\,pc^{-2}}$. The total H I mass of the Milky Way is $M_{\mathrm{H\,I}} \approx 4 \times 10^9\,\mathrm{M_\odot}$. The H I disc extends to $R \approx 25\,\mathrm{kpc}$ and has a scale-height at the Solar neighbourhood of $h_{\mathrm{H\,I}} \approx 100 - 200\,\mathrm{pc}$ (but see 3.1). The neutral medium is constituted by two **phases**, we see in §6.5 the physical reasons that likely cause such phases. Gas phases are characterized by distinct typical temperatures and densities, but similar pressure as the whole interstellar medium is roughly at pressure equilibrium (1.3). The cold phase ($\langle T \rangle \simeq 70\,\mathrm{K}$) is called **cold neutral medium** (CNM) and it is traditionally considered to be distributed in compact "clouds". The thermal speed of a gas particle can be written

$$v_{\mathrm{th}} \equiv \sqrt{\frac{k_{\mathrm{B}} T}{m}}, \tag{1.1}$$

where T is the temperature of the gas and m is the mass of the particle. For the CNM we obtain $v_{\mathrm{CNM}} \approx 1\,\mathrm{km\,s^{-1}}$ (see Table 1). A large fraction, nearly half, of the atomic hydrogen in the Milky Way is likely at higher temperatures approaching $T \sim 10^4\,\mathrm{K}$. This phase is called **warm neutral medium** (WNM) and its typical temperature is about $\langle T_{\mathrm{WNM}} \rangle \simeq 6000 - 8000\,\mathrm{K}$ [HT03]. If we use eq. 1.1 to calculate its typical speed we find $v_{\mathrm{WNM}} \approx 6 - 8\,\mathrm{km\,s^{-1}}$.

The observations of the H I (21-cm) emission lines in the ISM of a galaxy can be used to measure the typical random speeds of particles in this medium. The observed lines in normal spiral galaxies are roughly Gaussian with typical velocity dispersion (measure of the random motions along the line of sight) of $\sigma_{\mathrm{H\,I}} \sim 10\,\mathrm{km\,s^{-1}}$, usually decreasing from the inner disc ($\sigma_{\mathrm{H\,I}} \approx 15 - 20\,\mathrm{km\,s^{-1}}$) to the outer discs ($\sigma_{\mathrm{H\,I}} \approx 6 - 8\,\mathrm{km\,s^{-1}}$). These values cannot be produced by thermal broadening alone, in particular for the cold phase. In fact, this large value of the gas velocity dispersion with respect to the thermal expectations is the main evidence that in the ISM there a large amount of **turbulence** (Chap. 8). In the outer parts

[2]Hereafter with term **collision** we refer to any interaction between gas particles that changes their momentum vector and/or their energy state.

[3]Just for reference, the volume density of the Earth atmosphere is $\approx 3 \times 10^{19}\,\mathrm{cm^{-3}}$.

TABLE 1. Physical properties of the gaseous ISM in the Milky Way

Property	CNM	WNM	HII reg	WIM	HIM	Mol
Density (cm^{-3})	30	0.3	$10-10^4$	0.1	10^{-3}	$\gtrsim 10^2$
Temperature (K)	$30-120$	8000	$\approx 10^4$		$\gtrsim 10^6$	~ 10
Thermal speed[a] ($\mathrm{km\,s}^{-1}$)	~ 1	8	~ 10		$\gtrsim 100$	~ 0.1
Observed speed ($\mathrm{km\,s}^{-1}$)	few-10	$8-15$	$15-25$		-	~ 1
Total Mass ($10^9\,\mathrm{M}_\odot$)	2	2	0.1	1	$\lesssim 0.1$	1
Scale-height[b] (kpc)	0.1	$0.2-0.3$	-	1	3	0.06
Maximum Radius (kpc)	$\lesssim 25$	25	~ 15	-	-	~ 10
Filling factor	0.01	0.25	0.02	0.25	≈ 0.5	0.001
Main observables	21-cm		recomb/forbid		X-rays	CO

[a] Thermal speeds are for hydrogen atoms (neutral or ionised). [b] These are given in the inner disc, roughly withing the Solar radius.

of galaxy discs, the measured velocity dispersion often reaches values of $6-8\,\mathrm{km\,s}^{-1}$ (similar to that of the WNM) and turbulence may be less important.

1.1.2. Photoionised gas. Regions of the atomic ISM that are illuminated by photons with energy $h\nu > 13.6\,\mathrm{eV}$ end up being partially or totally ionised. Typical regions of this sort are called **H II regions** and are produced by massive O/B stars. The classical theory predicts that *all* neutral atoms in a sphere around an UV-emitting star are ionised (given the high cross section of the transition) out to the so-called **Strömgren radius**. Thus, in an homogeneous ISM, H II regions would be spheres (Strömgren spheres) with sharp edges and no ionizing photon should escape their boundaries. In a more realistic ISM, H II regions are, in fact, rather inhomogeneous and have a variety of shapes. This is due to the "porosity" and the anisotropy of the ISM on the small scales and in particular, in star-forming regions (see also §1.1.4).

Photoionised gas can be observed through **recombinations lines**, which are produced by electrons cascading down some subsequent energy levels after having been captured by an ion. The emission wavelength corresponding to a recombination line transition in an hydrogen atom is readily derived using the Rydberg formula:

$$\frac{1}{\lambda} = R_\mathrm{H}\left(\frac{1}{n_i^2} - \frac{1}{n_j^2}\right), \tag{1.2}$$

with $R_\mathrm{H} \simeq 1.1\times10^5\ \mathrm{cm}^{-1}$, n_i and n_j are respectively the lower and upper levels expressed in electron number. The most useful lines in local galaxies are Balmer lines: Hα (6563 Å), Hβ (4861 Å) and so forth, which are transitions from the third, forth etc. level to the second electron level ($n_i = 2$ in eq. 1.2). The Lyman series ($n_i = 1$) lines are widely used in high-redshift galaxies given that, from their restframe ultraviolet (UV), they end up in the optical band at $z > 2$.

A second class of emission lines extensively used to detect and study photoionised gas are the **forbidden lines**. They are called "forbidden" because they are normally not observed on Earth experiments as their emission requires extremely low densities. A prime example is given by the transitions of singly and doubly ionised oxygen. The first ionization potential of oxygen is similar to that

of hydrogen, thus in photoionised regions oxygen is vastly in the form of O^+. The ground levels of O^+ can be then *collisionally* excited and reach the higher levels. The transitions between those excited levels and the ground level are very improbable with typical probability of about $1\,s^{-1}$, against probabilities of $\sim 1\times10^8\,s^{-1}$ for normal dipole transitions (like Lyα and Hα). However, under the physical conditions of an ISM nebula (low density environment), these transitions can become much more probable than the collisional de-excitation. This occurs below (or close to) the so-called **critical density** of the transition, which for the O^+ lines is $n\approx10^3-10^4\,cm^{-3}$. Being the gas in H II regions typically below these densities (Table 1), oxygen forbidden lines are very commonly observed and they are in fact the dominant cooling process of photoionised regions. Forbidden line transitions are typically indicated with square brackets, e.g. [O II].

Finally, photoionised gas like any plasma with free electrons can emit for **bremsstrahlung**. This radiation is due to 'collisions' (encounters) between electrons and ions and, at typical temperatures of H II regions ($T\simeq10^4$), shines at radio wavelengths. H II regions are therefore very bright in radio continuum. This process refers to the thermal radio emission of H II regions, the presence of relativistic electrons and magnetic fields in the ISM are then at the origin of non-thermal synchrothron emission, which is also commonly observed in galaxies.

In a local star-forming galaxy, the luminosity given by recombination lines is typically dominated by H II regions. However, these nebulae do not dominate the mass as most of the photoionised gas is instead in a diffuse component permeating the whole ISM. In the Milky Way, the total mass of photoionised gas in HI regions is of the order of $10^8\,M_\odot$ with a filling factor[4] of about $f\approx0.02$. The diffuse **warm ionized medium** (WIM) has instead a total mass of the order $M_{\rm WIM}\simeq10^9\,M_\odot$ and much larger filling factor of $f\approx0.25$ (Table 1). In external galaxies the WIM gas is also called **diffuse ionized gas**. The ionizing radiation for the WIM can come from different sources. Some UV photons escape H II regions and propagate much further than the typical Strömgren radius. Additional photon flux can come from B stars that are not powerful enough to produce H II regions and from highly evolved stars. Moreover, in the outer parts of the disc or at a relatively large distance from the plane, neutral gas can be photoionised by the extragalactic UV background. The typical scale-height of the WIM disc of a spiral galaxy is $h_{\rm WIM}\approx1\,kpc$.

1.1.3. Collisionally ionised gas. At very high temperatures, hydrogen can be ionised simply by collisions with other hydrogen atoms. This certainly occurs when the temperature (T) of the gas is such that

$$k_{\rm B}T>13.6\,eV\simeq2.18\,erg, \tag{1.3}$$

which is for $T>T_{\rm ion}\simeq1.6\times10^5\,K$ (sometimes called **ionization temperature**). One would expect that only above these temperatures the ionization of the ISM should become effective. In practice, however, this is not the case as significant collisional ionisation can take place at much lower temperatures. This is due to the fact that the probability distribution of particle speeds is Maxwellian (see §1.2.2) with a characteristic tail of particles at very high speeds. Given that the competing process (radiative recombination) is relatively inefficient, this tail of (rare) fast electrons is sufficient to keep most of the gas ionised at temperatures as low as

[4]This is calculated as the volume occupied by the H II regions divided by the volume of the whole ISM.

$T \approx 2 - 3 \times 10^4\,\mathrm{K}$. Note that these temperatures are still higher than those of the neutral atomic (and also the photoionised) ISM (see Table 1) therefore the collisional ionisation of hydrogen requires the presence of powerful accelerations to *supersonic* speeds that bring to the development of *shock waves*, phenomena that we describe in detail in Chap. 4.

Collisionally ionised gas in disc galaxies emits most of its radiation at soft X-ray wavelengths. The base emission is thermal bremsstralung given that the medium (in particular H and He) is essentially fully ionised. Metals, however retain several of their electrons up to temperatures of $\sim 10^7\,\mathrm{K}$, therefore metal lines strongly contribute to the overall emission (if the metallicity of the gas is sufficiently high). In fact, metal lines fully dominate the soft X-ray emission (below 1 keV) if the gas is metal rich (like the Galactic ISM). Two types of line emissions are present: forbidden lines and the so-called **resonance lines**. The latter are collisionally excited transitions with a very high probability of spontaneous de-excitation, typical species are O VII, O VIII, Ne IX. The hot medium in the Milky Way is referred to as **hot inter-cloud medium** (HIM), its total mass is very uncertain, probably between 10^7 and $10^8\,\mathrm{M}_\odot$. The filling factor is close to unity and the scaleheight is about $h_{\mathrm{HIM}} \simeq 3\,\mathrm{kpc}$. These values vary substantially from galaxy to galaxy.

Apart from this hot medium, generically produced by stellar feedback, galaxies are surrounded by atmospheres of hot gas putatively formed by gravitationally shock-heated intergalactic medium but with probable contributions from stellar and AGN feedback. The general idea is that the intergalactic gas continually falls into dark matter potential wells, at the bottom of which galaxies reside. This infall of gas happens at velocities close to the free-fall velocities which are typically highly supersonic with respect to the IGM temperature. The gas is thus shock heated to a temperature close to the **virial temperature**[5] of the system and forms an extended **corona** around the central galaxy. Although clear observational evidence is still lacking, galactic coronae are thought to extend out to the virial radius and contain a mass of gas of the order of (or even larger than) the total stellar mass of the galaxy.

1.1.4. Molecular gas. The transition between atomic and molecular gas occurs a very low temperatures, typically below $30-40\,\mathrm{K}$. The molecular gas is mostly composed by H_2 that, unfortunately, does not have appreciable emission in any part of the electromagnetic spectrum. In particular, being a symmetric molecule, it does not have a permanent electric dipole and the rotational transitions are limited to quadropoles. The brightest line would be the $J : 2 \to 0$ (selection rule $\Delta J = \pm 2$) emission at 28.2 μm but the lifetime of the $J = 2$ level is typically of $3 \times 10^{10}\,\mathrm{s}$, which makes it almost undetectable. Other two possible transitions are the rovibrational transitions, in particular $v : 1 \to 0$ at 2.1 μm, and electronic transitions but they both require very high temperatures (6600 K and $\sim 10^5\,\mathrm{K}$ respectively). This is obsviously a serious limitation if we aim to trace a gas component at typical temperatures of $T \sim 10\,\mathrm{K}$ (Table 1). At such high temperatures the H_2 molecules can also be collisionally dissociated, although the main dissociation channel for

[5]The virial temperature of a galaxy can be defined as

$$T_{\mathrm{vir}} \equiv \frac{\mu m_{\mathrm{p}}}{2k_{\mathrm{B}}} v_{\mathrm{vir}}^2, \tag{1.4}$$

where v_{vir} is the circular speed at the virial radius.

molecular hydrogen in a typical disc galaxy is photodissociation, which occurs with photons at $h\nu > 11.2\,\mathrm{eV}$.

As a consequence, the main molecule used for detection of molecular gas in the ISM of galaxies is carbon mono-oxide (CO). CO is relatively abundant, but about 10^{-4} times less than H_2. Its main transition is that of $J : 1 \rightarrow 0$, observable at radio wavelengths at $\lambda \simeq 2.6\,\mathrm{mm}$. Other molecules are: CS, NH_3, HCN, OH, HCO^+ and H_2O. Their transitions, together with CO, carry out most of the *cooling* in dense molecular regions (although the fine-structure emission of singly ionised atomic carbon C^+ also contributes). Most of the molecular gas in the Milky Way is located in giant clouds. The temperatures of these clouds are very low, $T_{\mathrm{MC}} \simeq 10 - 15\,\mathrm{K}$. The typical mass of a giant molecular cloud (GMC) is about $M_{\mathrm{GMC}} \sim 10^5 - 10^6\,\mathrm{M}_\odot$ with densities $n_{\mathrm{H}_2} \gtrsim 10^2\,\mathrm{cm}^{-3}$. The interior of these clouds shows a hierarchy of structures with a filamentary and nearly fractal pattern down to small structures called **dense cores**, where star formation takes place. The total mass of molecular gas in the Milky Way is $M_{\mathrm{H}_2} \sim 1 \times 10^9\,\mathrm{M}_\odot$ (Table 1).

1.1.5. Other components. Magnetic fields are ubiquitously observed in star-forming galaxies in particular through the polarization of radio synchrotron emission; we discuss their origin and physical properties in Chap. 9. The strength of the field can be measured using the Faraday rotation (towards pulsars and/or quasars) which gives $\langle n_{\mathrm{e}} B_{||} \rangle$, where n_{e} is the electron density, $B_{||}$ is the component of the magnetic field parallel to the line of sight and the average is meant along the path-length of the object. This value is usually combined with that of the dispersion measure that allows us to estimate the mean electron density $\langle n_{\mathrm{e}} \rangle$ along the same path length. The average value of the magnetic field in the Milky Way, estimated in this fashion, is $\approx 3 \times 10^{-6}\,\mathrm{G}$.

Measures of magnetic field strengths have also been performed using the Zeeman effect of some emission lines, in particular the 21-cm line. These measures usually return higher values of $\langle B \rangle \approx$ few $\times\, 10^{-5}\,\mathrm{G}$. This is due to the fact that most of the emission comes from the densest clouds and in these the magnetic field is more intense for reasons that are explained in §9.2.

Interstellar dust is also a very important component of the ISM. Although highly subdominant in mass (about one hundredth of the atomic neutral gas), it plays a central role in several physical processes. These include the absorption and scatter of stellar light, cooling and heating (via photoelectric heating for instance) and the formation of molecules, in particular H_2. It is typically observed as (thermal) emission at mid and far infrared wavelengths (depending mostly on its temperature) and in absorption in the UV-optical bands.

Finally, the Galactic ISM is continuously crossed through by highly energetic (relativistic) particles. For historic reasons, we refer to these particles as **cosmic rays**. They impact on the Earth's atmosphere producing showers of secondary particles that can be detected from the ground. In the ecosystem of the ISM they are very important for the heating of dense regions like the cores of molecular clouds.

1.2. Astrophysical gases as collisional systems

The interstellar medium and, to some extent, the intergalactic medium are essentially **collisional systems** of non-ideal particles. 'Collisional' means that interactions between particles are frequent, with typical time scales much lower

than other relevant times, in particular the dynamical time or the crossing time of the system (eq. 2.65). In these conditions, the basic principles of statistical mechanics apply and the gas can be described using *macroscopic* thermodynamic variables. The variables often used to this end are: Temperature (T), Pressure (P), Density (ρ) and Entropy (S). We will use them extensively in these notes, see also §D.1.

A collisional system of *ideal* particles has a very important property: the distribution of the velocities of its particles is Maxwellian (see §1.2.2). Unfortunately, the ISM as a fluid is largely *non-ideal* in the sense that the particles have more degrees of freedom than the spatial ones. In particular, all atoms have electrons that can be brought to excited energy levels or even ejected (ionisation) as a consequence of collisions. Molecules have rotation and vibrational degrees of freedom that add up to the electronic ones. The outcome is a system that may have a number of *inelastic collisions* that can, in principle, have the effect of bringing the system away from a Maxwellian distribution. In the following, we calculate the typical collisional times for ISM particles, from which it becomes clear that in most astrophysical situations the Maxwellian approximation can be considered valid.

1.2.1. Collision time. The **mean freepath length** of a particle travelling in a fluid with density n is

$$\ell \equiv \frac{1}{nA}, \tag{1.5}$$

where A is the cross section of the particle in question. From eq. 1.5 one can derive the **collision time**, which is the mean time between two successive collisions:

$$\tau_{\rm coll} \approx \frac{1}{nAv}, \tag{1.6}$$

where v is the typical speed of the particle. We now show that $\tau_{\rm coll}$ in the ISM is typically much lower than other characteristic times.

Let us first consider an ionised gas of free electrons and nuclei (only ions for simplicity). In this case, the encounters are regulated by the Coulomb force. The typical interaction between an electron and an ion can be estimated by assuming that the ion is at rest. This assumption is justified by the fact that its mass is much larger than that of electrons and thus its velocity is much lower. The closest encounter (lowest impact parameter) corresponds to the ("classical") capture of the electron by the ion, which fulfills the equation

$$\frac{Ze^2}{b_{\rm min}^2} \sim \frac{m_{\rm e} v^2}{b_{\rm min}}, \tag{1.7}$$

where Z is the atomic number, e is the electric charge, $m_{\rm e}$ and v are the mass and velocity of the electron and $b_{\rm min}$ is the minimum value of the so called **impact parameter**, distance at which the encounter occurs. We find that

$$b_{\rm min} = \frac{Ze^2}{m_e v^2} \tag{1.8}$$

and the cross section of a general interaction between electrons and ions will then be larger that the one that produces the electron capture

$$A({\rm e,i}) > \pi b_{\rm min}^2 = \frac{\pi Z^2 e^4}{m_{\rm e}^2 v^4}. \tag{1.9}$$

Then, recalling the definition of collision time (eq. 1.6), we can write an upper limit to the time between two encounters as

$$\tau_{\rm coll}({\rm e,i}) < \frac{m_{\rm e}^2 v^3}{\pi Z^2 e^4 n_{\rm i}} \tag{1.10}$$

where $n_{\rm i}$ is the number density of ions, with which electrons are interacting.

We now make the assumption that the typical velocities of the electrons are close to their thermal speeds (eq. 1.1). This leads to

$$\tau_{\rm coll}({\rm e,i}) < \frac{m_{\rm e}^{1/2} (k_{\rm B}T)^{3/2}}{\pi Z^2 e^4 n_{\rm i}} \tag{1.11}$$

and, assuming a typical electron temperature of $10^4\,{\rm K}$, we finally obtain

$$\tau_{\rm coll}({\rm e,i}) < 3\times10^5 \left(\frac{T}{10^4\,{\rm K}}\right)^{3/2} \left(\frac{n_{\rm i}}{1\,{\rm cm^{-3}}}\right)^{-1} {\rm s} \tag{1.12}$$

i.e. a very short time, as we see in a moment. In realistic conditions, collisions with $b > b_{\rm min}$ are much more frequent than those at $b \approx b_{\rm min}$ and they contribute much more to the momentum redistribution (these are all elastic collisions) of the gas particles. Thus, the actual characteristic time must be lowered by a factor $1/(\ln\Lambda)$, with $\ln\Lambda$ the Coulomb logarithm, and it becomes

$$\tau_{\rm coll}({\rm e,i}) \approx 1\times10^4 \left(\frac{T}{10^4\,{\rm K}}\right)^{3/2} \left(\frac{n_{\rm i}}{1\,{\rm cm^{-3}}}\right)^{-1} {\rm s}. \tag{1.13}$$

The time in eq. 1.13 is not just a collision time but also as a *deviation* time (every $\tau_{\rm coll}$ *each* particle in the gas is deviated in its motion). Coulomb interactions are indeed rather strong and affect the motions of particles in important ways. The extreme case of an interaction at $b \approx b_{\rm min}$ corresponds to a 90 degree bent of the electron and thus changes completely its path and momentum making it effectively lose the memory of its past direction of motion. This has two main consequences. First, the collision time roughly corresponds to the time an electron can travel on a specific direction (the mean free path will be the distance); beyond that time its motion will effectively proceed at random. Second, the velocities of the electrons are affected significantly by these interactions and thus the typical time that the electrons take to *thermalise* (reach thermal equilibrium) will not be much larger that $\tau_{\rm coll}({\rm e,i})$ (see also below).

The Coulomb interactions also apply to particles of the same kind. So interactions between electrons can be treated analogously to the above (with a repulsive force) and we obtain a collision time of

$$\tau_{\rm coll}({\rm e,e}) \approx 1\times10^4 \left(\frac{T}{10^4\,{\rm K}}\right)^{3/2} \left(\frac{n_{\rm e}}{1\,{\rm cm^{-3}}}\right)^{-1} {\rm s}. \tag{1.14}$$

Likewise, for interactions between protons (ions), the same reasoning leads to

$$\tau_{\rm coll}({\rm p,p}) \sim \left(\frac{m_{\rm p}}{m_{\rm e}}\right)^{1/2} \tau_{\rm coll}({\rm e,e}), \tag{1.15}$$

given the dependence on the square root of the particle mass (eq. 1.11).

As mentioned, the above collision times are characteristic of *elastic* collisions that conserve the total momentum and kinetic energy, i.e. the energy is not transferred to another form. These are the interactions that make the gas *collisional*

and produce Maxwellian distributions of velocities. In order for this to take place in astrophysical gases, one must have that these elastic-collision times are much shorter than the typical time scales of *inelastic* collisions. One of such times is the excitation time of a species. Let us take for instance the oxygen atom, which is the most common atom in the ISM after hydrogen and helium. In the photoionised medium ($T \sim 10^4\,\mathrm{K}$), most oxygen is singly ionised and collisions will be able to excite it to an upper electronic level (see §1.1.2). Excitation of O^+ has the following characteristic time:

$$t_{\rm ex}(O^+) \approx 7 \times 10^{10} \left(\frac{n_{O^+}}{10^{-3} n_{\rm e}} \right)^{-1} \ \mathrm{s}. \tag{1.16}$$

Given that for a fully ionised gas $n_{O^+} \lesssim 10^{-3} n_{\rm e}$, this time is clearly several orders of magnitude larger than the elactic time in eqs. 1.13 and 1.14. Note that the typical hydrogen recombination time is even longer than this: $t_{\rm rec} \approx 4 \times 10^{12}\,\mathrm{s}$.

We end this discussion about Coulomb interactions by pointing out that the actual thermalisation of ions and electrons will, in fact, take a longer time than those shown in eqs. 1.13 and 1.15. If we imagine a gas having different temperatures between ions (protons in particular) and electrons, every collision will change the velocity of a proton by $\Delta v_{\rm p} \sim (m_{\rm e}/m_{\rm p})^{1/2}\, \Delta v_{\rm e}$ and the time that it takes to bring the two components at the same temperature (**equipartition time**) will be of the order

$$\tau_{\rm eq}({\rm p,e}) \sim \left(\frac{m_{\rm p}}{m_{\rm e}} \right) \tau_{\rm coll}({\rm e,p}) \tag{1.17}$$

thus about three order of magnitude larger than the collision time but still much smaller than typical inelastic times (eq. 1.16). We can conclude that for ionised gas, equipartition between the various components will be always valid in the ISM. Possible exceptions are peculiar situations like when a shock wave has just established itself (Chap. 4) or in very rarefied gases (e.g. in the IGM).

Let us now consider the neutral atomic gas where collisions occur essentially between two neutral hydrogen atoms (at low enough temperatures). We can calculate an upper limit to the collision time in this case by assuming a cross section equal to the classical Bohr cross section, $A > \pi a_0^2$ where $a_0 = 5.3 \times 10^{-9}\,\mathrm{cm}$ is the Bohr radius[6]. This leads to a cross section of about $10^{-16}\mathrm{cm}^2$. Obviously this is the minimum possible cross section for H-H collisions and it leads to a typical time scale

$$\tau_{\rm coll}(H,H) < \frac{1}{\pi a_0^2 \upsilon n_{\rm H}}, \tag{1.18}$$

[6]We remind that the Bohr radius, defined as $a_0 = \hbar/(m_{\rm e} c \alpha)$ (with α fine structure constant and $\hbar$ reduced Planck constant), can be seen in quantum mechanics as the most probable distance between the electron and the proton.

where $n_{\rm H}$ is the H I number density. Using again the thermal speed (eq. 1.1) as a typical speed of the particles in the neutral medium[7], we obtain

$$\tau_{\rm coll}({\rm H},{\rm H}) < \frac{m_{\rm H}^{1/2}}{\pi a_0^2 (k_{\rm B} T_{\rm H\,I})^{1/2} n_{\rm H}}, \tag{1.19}$$

where $T_{\rm H\,I}$ is the temperature of the H I gas and the hydrogen mass $m_{\rm H} \simeq m_{\rm p}$. Using a typical temperature and density of the CNM (see 1.1.1) we obtain

$$\tau_{\rm coll}({\rm H},{\rm H}) < 5 \times 10^9 \left(\frac{T_{\rm H\,I}}{70\,{\rm K}}\right)^{-1/2} \left(\frac{n_{\rm H\,I}}{30\,{\rm cm}^{-3}}\right)^{-1} {\rm s} \tag{1.20}$$

Under more realistic conditions, the value of the cross section $A({\rm H},{\rm H})$ is significantly larger than the one obtained using the Bohr radius. Quantum mechanical effects are quite important and by taking them into account one finds $A({\rm H},{\rm H}) \approx$ few $\times\, 10^{-15}{\rm cm}^2$, with a slight dependence on the particle speed. As a consequence, the collision time at $T = 70\,{\rm K}$ is

$$\tau_{\rm coll}({\rm H},{\rm H}) \approx 1 \times 10^8 \left(\frac{n_{\rm H\,I}}{30\,{\rm cm}^{-3}}\right)^{-1} {\rm s}. \tag{1.21}$$

In the WNM, the fraction of electrons starts to become important again and the typical collisional time is set by collisions between electrons and neutral atoms:

$$\tau_{\rm coll}({\rm e},{\rm H}) \sim 10^7 \left(\frac{T_{\rm e}}{10^4\,{\rm K}}\right) \left(\frac{n_{\rm H\,I}}{1\,{\rm cm}^{-3}}\right)^{-1} {\rm s}. \tag{1.22}$$

As for the ionised gas, the time scales in eqs. 1.21 and 1.22 can be consider short when compared to any other characteristic times. The most important inelastic process in H I regions is the excitation of C^+ ions through collision with free electrons, with typical excitation times orders of magnitude larger than the above. We conclude that, in the ISM, the (elastic) collision times are always short, compared to the times attributable to inelastic collisions or to the dynamical times (these latter usually of the order of Megayears or larger). This has the important consequences that we now describe.

1.2.2. Maxwellian distribution of the velocities. The results of the previous section show that collisions dominate the motion of particles within the ISM. Under these conditions, we can expect particles to have probability distributions of their peculiar velocities that approach Maxwellian functions. The Maxwellian

[7]This passage is not strictly justified because what we are set to prove is that the thermal speed *is* in fact the typical speed of the medium, as a result of the Maxwellian distribution. However, it is fair to assume that the particles are unlikely to have speeds much larger (or smaller) than this value. If it were much larger for instance, strong shocks would occur and the nature of the medium would change, it particular it would not remain neutral.

probability distribution $f(v)$ of speeds between v and $v + \mathrm{d}v$ is[8]

$$f(v)dv = 4\pi \left(\frac{m}{2\pi k_\mathrm{B} T}\right)^{3/2} v^2 \exp\left(-\frac{mv^2}{2k_\mathrm{B}T}\right) dv, \tag{1.24}$$

where m and T are the mass and temperature of the particles. The most important implication of this is that, being collisions so frequent as established above, the temperatures of the *different* particles in the same portion of gas turn out to be the *same*. One can, in fact, show that, in typical conditions of the ISM, the temperatures of the subcomponents (e.g. electrons versus ions) that make up the gas do not differ by more than 1 % from each other [Spi78].

The Maxwellian function (eq. 1.24) also allows us to derive characteristic speeds:

$$v_\mathrm{peak} = \sqrt{\frac{2k_\mathrm{B}T}{m}}, \quad v_\mathrm{rms} = \sqrt{\frac{3k_\mathrm{B}T}{m}}, \quad \langle v \rangle = \int_0^\infty v f(v) dv = \sqrt{\frac{8k_\mathrm{B}T}{\pi m}} \tag{1.25}$$

that we call **peak speed**, **r.m.s. speed** and **mean speed**. They are all of the same order of magnitude as the thermal speed (eq. 1.1) and, as we see in §2.10, of the sound speed. Note that, despite the fact that the temperatures of different particles composing a gas are the same, their peculiar velocities may differ greatly. So, for instance, having electrons and ions at the same temperature $T_\mathrm{e} = T_\mathrm{i}$ implies that $v_\mathrm{e} \gg v_\mathrm{i}$.

In the end, the fact that astrophysical gases can be assumed to have (locally) a single temperature is what allows us to describe this gas using relatively simple hydrodynamics. Indeed, hereafter, we mostly deal with a *one-component ideal gas*, we write equations using thermodynamic variables (T, P, ρ, etc.) without specifying which particular component of the fluid these variables refer to; they do, in fact, refer to an "average" particle. Before moving to the main assumption of hydrodynamics, we briefly describe the most relevant equilibria that we encounter in astrophysical gases and the ISM in particular.

1.3. Equilibria in the ISM

The following terms for the various equilibria are used in this course. Note, however, in other texts one may also find different nomenclature/definitions.

- Thermal equilibrium. We say that a portion of gas is at thermal equilibrium if it is (and remains for some relevant time scale) at the same temperature. A patch of ISM constituted by a single phase (e.g. CNM) can be approximated as in thermal equilibrium and one can associate a thermal speed to the gas given by its typical temperature. The ISM as a whole is *not* in thermal equilibrium as the temperatures of its different components vary greatly (Table 1). Large astronomical structures roughly in thermal equilibrium are coronae of hot gas

[8] In order to derive the Maxwellian distribution one starts from the velocity distribution per spatial direction (j), which, due to the randomness of particles motions, is a normalised Gaussian function (eq. C.6):

$$f(v_j)dv_j = \sqrt{\frac{m}{2\pi k_\mathrm{B}T}} \exp\left(-\frac{mv_j^2}{2k_\mathrm{B}T}\right) dv_j, \tag{1.23}$$

where v_j is the j^th component of the velocity. To obtain the distribution as a function of the magnitude of the velocity (speed), eq. 1.23 must be integrated in all directions by writing $f(v)dv^3 = f(v_x, v_y, v_z)dv_x dv_y dv_z$, with $f(v_x, v_y, v_z) = f(v_x)f(v_y)f(v_z)$ and $dv_x dv_y dv_z = v^2 \sin\theta \, d\theta \, d\phi \, dv$, where the latter becomes $4\pi v^2 dv$ for isotropic systems.

around galaxies (§1.1.3) and hot halos of galaxy clusters. The temperature in these cases is relatively constant with radius although mild gradients are occasionally observed. Smaller structures like portion of CNM or WNM/WIM are also roughly at constant temperature. Thermal equilibrium usually occurs when heating processes of the gas balance cooling processes (see §2.8), we discuss this further in §6.5.

- Pressure equilibrium. This is one of the most important equilibria in astrophysical gases as it allows components with temperatures orders of magnitude different from each other to co-exist in galaxies. Basically, all main components of the ISM are somewhat in pressure equilibrium, which is almost a necessary condition to survive. In the Milky Way, the mean ISM pressure is

$$\langle P \rangle \approx 2 \times 10^{-13} \,\mathrm{dyne\,cm^{-2}}, \tag{1.26}$$

also sometimes expressed in units of the Boltzmann constant: $\langle P/k_\mathrm{B} \rangle \approx 10^3 \,\mathrm{cm^{-3}\,K}$.

Figure 1 shows the location of the various components of the ISM in the density-temperature (n, T) plane. They are clearly not distributed at random but closely follow the diagonal lines, which represent lines at constant pressure. The few components out of equilibrium are H II regions and molecular clouds. The former are in surplus of pressure with respect to the surrounding ISM: their extra pressure comes from the conversion of radiation into internal energy due to photoionisation (see 1.1.2). As a consequence of this, H II regions tend to expand in the surrounding ISM (a description of this phenomenon, not treated in these notes, can be found in [DW97]). Molecular clouds (in Fig. 1 also called dark clouds) are instead in a special state as they feel the effect of their own gravity. In them, the self gravity of the gas tends to compress the clouds and this causes the pressure to increase, with respect to the ISM value.

At first sight, it would seem that the coronal gas in Fig. 1 is slightly out of equilibrium. However, more recent measurements of the density of hot gas in the Milky Way and other galaxies would shift these points towards the left part of the diagram, at $n \sim 10^{-3} \,\mathrm{cm^{-3}}$. Pressure equilibrium is sometimes referred to as *mechanical* equilibrium.

- Hydrostatic equilibrium. This is the equilibrium of a gas that balances a gravity force (its own gravity and/or an external field) with its own pressure force. Several gaseous systems, for instance compact molecular clouds, are in a state close to hydrostatic equilibrium with their own gravity. Galactic coronae and hot halos of galaxy clusters are in hydrostatic equilibrium with the gravitational potential of the dark matter halos. Finally, the whole ISM in a galaxy disc has random motions that balance the vertical gravity pull and settles in a state of equilibrium. Thus hydrostatic equilibrium determines the thickness of the gaseous discs and each component of the ISM (Table 1) tends to have its own typical thickness (being them characterized by different temperatures as well). We discuss this in details in §3.1. In a turbulent gas (where turbulent motions largely overcome the thermal motions) an "equivalent" hydrostatic equilibrium is established between gravity and the so-called "turbulent pressure" (see §8.3.4).
- Thermodynamical equilibrium. When the density of a portion of gas is larger than the critical density ($n > n_\mathrm{crit}$, see §1.1.2) of a specific transition (electronic, rotational etc.), collisions dominate the excitations and de-excitations with respect to spontaneous emission and absorption. In these conditions, **local thermodynamical equilibrium** (LTE) is nearly satisfied and the level populations are

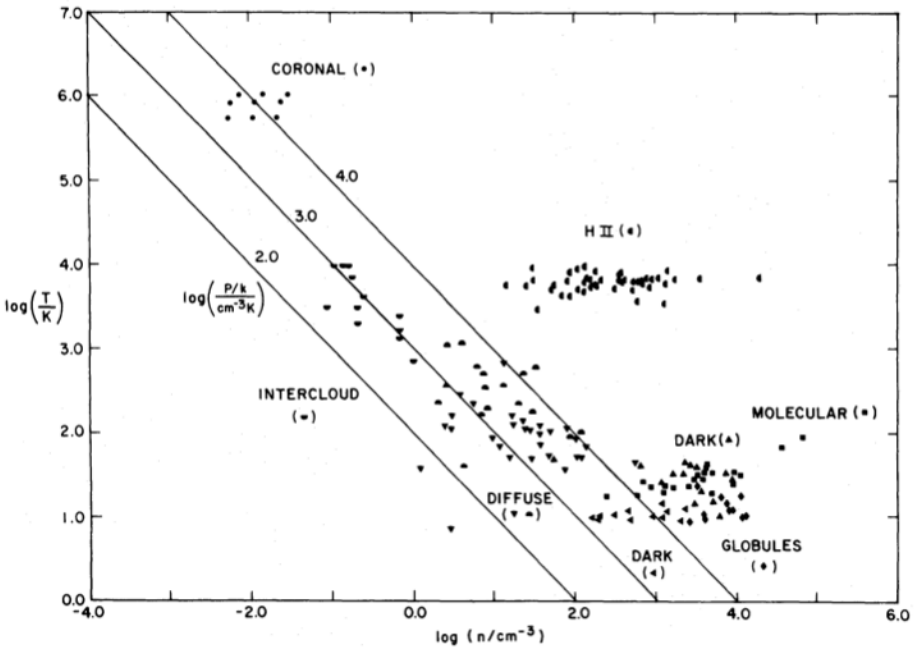

FIGURE 1. Temperatures and densities of the main components of the gaseous ISM, from [Mye78].

given by the Boltzmann law:

$$\frac{N_1}{N_0} = \frac{g_1}{g_0} \exp\left(-\frac{\Delta E}{k_{\rm B} T_{\rm ex}}\right), \tag{1.27}$$

where N_1 and N_0 are the populations of the excited and ground levels, g_1 and g_0 are their statistical weights, ΔE is the energy jump and $T_{\rm ex}$ is the excitation temperature, which, in these conditions, is close to the actual kinetic temperature of the medium.

Consider for instance the hyperfine transition of H I at 21 cm. In this case, the excitation temperature $T_{\rm ex}$ is called *spin* temperature ($T_{\rm sp}$) and is taken as representative for the kinetic temperature of the gas whenever the volume density exceeds the critical density. This assumption is valid in the CNM, whereas in the WNM (typical densities $n < n_{\rm crit}$) the excitation temperature is only a lower limit to the actual temperature.

CHAPTER 2

Fluids as continua

In the previous Chapter we have seen that the typical collision times between particles in astrophysical gases are short with respect to other relevant timescales. As a consequence, the mean free paths are also short, i.e. particles interact with each other on spatial scales much smaller that the relevant scales of the macroscopic system. This is the definition of a collisional system and it is crucial for the definition of a fundamental concept of fluid dynamics: the fluid element. More generally, it allows us to treat the fluid as *a continuum* and write the fundamental equations of hydrodynamics. In this Chapter, we first encounter the base concepts of hydrodynamics and we derive the three fundamental conservation equations of ideal fluids. Finally, we see how sound waves propagate in a fluid and we intruduce the concept of sound speed.

2.1. Fluid element

A **fluid element** is a portion of fluid of size l *much larger* than the mean free path and *much smaller* than the characteristic macroscopic scale of the fluid:

$$\ell \ll l \ll L, \tag{2.1}$$

where ℓ is the mean free path (eq. 1.5) and L is the typical scale over which any macroscopic physical quantity (q) vary appreciably:

$$L \sim \frac{q}{\nabla q}, \tag{2.2}$$

where ∇q is the spatial derivative (gradient) of q. If inequalities like those in eq. (2.1) hold everywhere in the fluid then it is possible to define a fluid element of size l.

We can have an idea of the scales involved in the above definition by estimating the mean free path in the ISM. In §1.2.1 we have seen that the longest collision time in the ISM is given by the typical interaction that takes place in the CNM (eq. 1.21). Thus, ℓ will likely always be smaller or equal to the mean free path of the interaction between two atoms of neutral hydrogen:

$$\ell < \ell(\mathrm{H},\mathrm{H}) < \frac{1}{A_{(\mathrm{H},\mathrm{H})} n} \sim 1 \times 10^{13} \left(\frac{n_{\mathrm{H}}}{30\,\mathrm{cm}^{-3}}\right)^{-1} \mathrm{cm}, \tag{2.3}$$

where we have used $A \sim 10^{-15}\,\mathrm{cm}^2$. As expected, the above is a very short length compared to the typical scales over which quantities vary in the ISM, of the order of several parsecs or more. This estimate is only meant to give an idea of the scales that we are dealing with in this context. In fact, the fluid element has no size whatsoever, it is only a *conceptual entity* and it is never explicitly used in the calculations.

The above considerations imply that a fluid element must contain a very large number of particles that continuously interact with each other ($\ell \ll l$) to the point that their motion within the element becomes completely random. We indicate the velocity of the i-th individual particle as $\boldsymbol{w}_i$, while the fluid element move at a velocity $\boldsymbol{u}$. As a consequence of the above, the fluid element has two fundamental properties. First, despite it following the global motion of the fluid with a velocity $\boldsymbol{u}(\boldsymbol{r}, t)$, the elementary particles in its interior continuously change directions and speeds, because of collisions (§1.2.1). However, seen at the scale of the fluid element, all these random motions of the internal particles average to zero and the only dominating motion is the motion of the fluid itself. This implies that at the end, also the elementary particles have, on average, the net velocity of the fluid element $\boldsymbol{u}$. Second, the collisions within the fluid element effectively bring the fluid to a state of *maximum entropy*. Statistical mechanics then tells us that this condition refers to the macroscopic state that corresponds to the *maximum number* of microscopic configurations. The fundamental consequence of this is that a fluid element can be described with thermodynamic variables (T, P, etc.) and the fluid has an equation of state.

In the end, under the above conditions, a fluid can be treated as a continuum, which means that we can write derivatives and differential equations of its macroscopic quantities, despite it being made of *discrete* particles. This is justified by the fact that every fluid element is defined by its thermodynamic variables and it is (with respect to the fluid at large) infinitely small in size. This continuum approach is at the foundations of **hydrodynamics**.

Systems for which the above conditions are not verified ($\ell \gg L$) are called **collision-less systems**. Typical astrophysical examples are stellar systems (a galaxy or a cluster of stars) but this condition can also occur for gases, usually not interstellar gases. These systems can be described by the Jeans equations that resemble very closely some conservation equations of the ideal gases that we derive in the next sections. However, collisionless systems lack an equation of state. Intermediate situations in which $\ell \sim L$ are the most difficult to threat and occur in astrophysical systems like the asteroid belt in the Solar system.

In these notes, we essentially only deal with astrophysical fluids and interstellar gases in particular that are collisional and for which the continuum approach is valid. These fluids, however, have properties substantially different from standard terrestrial fluids, in particular the following:

(1) there is a general lack of interactions with solid bodies as those between planets, stars (not including winds or stellar feedback) and the ISM are typically negligible;
(2) gravity is often important, this can either be an external potential or self gravity of the gaseous systems, as it happens, for instance, in star formation;
(3) the fluid (specifically gas) is compressible (§2.5) and this can make its physical description very complex;
(4) the gas is often non neutral, i.e. it is made of electrons and ions; moving charges produce electric currents and magnetic fields and in a number of circumstances the latter cannot be neglected.

We now proceed to define the mathematical framework that allows us to write the differential (conservation) equations of the most fundamental properties of an ideal fluid with the aim to describe its motion and physical state.

2.2. Advection

Consider a fluid element moving at velocity $\boldsymbol{u}$ and a generic physical (macroscopic) quantity, function of position and time $q = q(\boldsymbol{r}, t)$. We can write the global variation of this quantity with time as

$$\frac{\mathrm{D}q(\boldsymbol{r},t)}{\mathrm{D}t} = \lim_{\delta t \to 0} \frac{q(\boldsymbol{r}+\delta\boldsymbol{r}, t+\delta t) - q(\boldsymbol{r},t)}{\delta t}, \tag{2.4}$$

where the symbol D for the derivative is used to indicate that it is a *full* derivative. This kind of derivative is also called **Lagrangian** (or **material**) **derivative**.

The r.h.s. of eq. (2.4) can be modified by adding and substracting the quantity $q(\boldsymbol{r}, t+\delta t)$, which leads to

$$\frac{\mathrm{D}q(\boldsymbol{r},t)}{\mathrm{D}t} = \left(\frac{\partial q}{\partial t}\right)_{(\boldsymbol{r},t)} + \lim_{\delta t \to 0} \frac{q(\boldsymbol{r}+\delta\boldsymbol{r}, t+\delta t) - q(\boldsymbol{r}+\delta\boldsymbol{r},t)}{\delta t}, \tag{2.5}$$

where the first derivative is taken at the *fixed position* $\boldsymbol{r}$. The numerator of the second term of the r.h.s. can be expanded as follows:

$$\begin{aligned} q(\boldsymbol{r}+\delta\boldsymbol{r}, t+\delta t) &\simeq q(\boldsymbol{r}, t+\delta t) + \delta\boldsymbol{r}\cdot(\boldsymbol{\nabla}q)_{(\boldsymbol{r},t+\delta t)} + ... \\ &\simeq q(\boldsymbol{r}, t+\delta t) + \delta\boldsymbol{r}\cdot\left[\boldsymbol{\nabla}q_{(\boldsymbol{r},t)} + \delta t\left(\frac{\partial\boldsymbol{\nabla}q}{\partial t}\right)_{(\boldsymbol{r},t)} + ...\right] + ... \end{aligned} \tag{2.6}$$

and taking only the first order terms we obtain

$$\frac{\mathrm{D}q(\boldsymbol{r},t)}{\mathrm{D}t} = \left(\frac{\partial q}{\partial t}\right)_{(\boldsymbol{r},t)} + \lim_{\delta t \to 0} \frac{\left[\delta\boldsymbol{r}\cdot(\boldsymbol{\nabla}q)_{(\boldsymbol{r},t)}\right]}{\delta t}, \tag{2.7}$$

where it is easy to see that the second term on r.h.s. simply becomes $\boldsymbol{u}\cdot\boldsymbol{\nabla}q$.

Equation 2.7 can be rewritten more compactly and in generic terms as

$$\frac{\mathrm{D}}{\mathrm{D}t} = \frac{\partial}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla}, \tag{2.8}$$

which is the definition of the Lagrangian derivative and describes the variation of a quantity on a fluid element as the fluid moves. Note that $\boldsymbol{u}$ here is the velocity of the fluid. The operator $\boldsymbol{u}\cdot\boldsymbol{\nabla}$ is called **advection operator** (or advective operator)[1]. Advection thus makes physical quantities of fluid elements vary as a consequence of the fluid bulk motion.

2.3. Eulerian and Lagrangian description

The changes of physical properties of the fluid can be described in two different ways depending on whether the derivatives are taken "at fixed fluid elements" or at fixed positions. Equation 2.8 clearly shows the difference between the Lagrangian ($\mathrm{D}/\mathrm{D}t$) and the so-called **Eulerian derivative** ($\partial/\partial t$), the latter being executed at a fixed position in space. The Lagrangian derivative adds to the variation at a fixed position the variation due to the "transport" of the fluid element by the bulk motion of the fluid itself (advection). Thus, in the Lagrangian description of the fluid, the variations in the physical properties refer to the specific fluid elements, in the Eulerian description, they refer to fixed positions in space.

[1]Note that there is a variety of terms used in different texts, in particular the second term on the r.h.s. of eq. 2.8 is also called convective operator or convective derivative.

We are now in the position to describe some kinematic concepts of the fluid motion. We define **streamlines** as the flow lines of the fluid as a whole, they are tangential to the velocity $\boldsymbol{u}$ at each point in space. Streamlines give a depiction of the instantaneous (at a specific time) motions within the fluid. Another useful concept it that of **particle paths** or **pathlines**. These are the paths of the fluid elements[2] and allows to visualise their locations in time. A third concept is that of **streaklines**, which show the instantaneous positions of all fluid elements that have passed through a particular point in space in the past. Streamlines can be constructed very simply: if one takes two snapshots of the fluid and joins the locations of every fluid elements (supposing we can "recognise" the various fluid elements) at the two times, this would immediately build the fluid streamlines. Particle paths are instead more tricky because one has to follow the motion of the fluid elements from the beginning to the end of the flow. The same is needed for the streaklines. For **stationary fluids** (i.e. fluids whose motion pattern does not change with time) streamlines and particles paths coincide.

2.3.1. Eulerian and Lagrangian hydrodynamic codes. Eulerian and Lagrangian descriptions also correspond to two different numerical implementations of hydrodynamics. Eulerian hydrodynamic codes are called *grid* codes (for instance in astrophysics: Athena, ENZO, Ramses). They are characterised by a grid fixed in space and the fluid moves through this grid. In practice the hydrodynamic solver calculates, at each time step, the fluxes of the fluid through the boundaries of the grid cells. The main drawbacks of this approach are difficulties to reach high spatial resolutions and the fact that, in general, particle paths cannot be traced as there are no particles (fluid elements). The problem of resolution is solved by using a technique called **Adaptive Mesh Refinement** (AMR). This refines the grid by splitting the cells in subcells in regions of interest that are defined by the user or calculated by the code given certain prescriptions (*refinement criteria*). For instance, high density regions are often chosen to be more refined than low density regions. The second problem can be solved by calculating the paths of a (relatively) small number of *trace particles*.

On the contrary, Lagrangian codes are particle based. They are largely used in astrophysics (e.g. the Gadget code) because one can trace individual particles and they can be very easily combined with N-body codes, which allows to easily include live gravity (changing with time) in the calculation. Their main drawbacks are that they suffer from Poisson noise and have difficulties in properly capturing shocks and hydrodynamic instabilities. The first problem is solved with a technique called **Smoothed Particle Hydrodynamics** (SPH), while shocks are forced to occurred by adding a so-called *artificial viscosity*. Finally, the capturing of fluid instabilities requires other ad-hoc prescriptions.

2.4. Continuity equation

We now derive the most important conservation equations of an ideal fluid starting from the conservation of mass[3]. Consider a portion of fluid of density ρ contained in a volume V. Because of mass conservation, the time derivative of the

[2]Note that here the term "particles" is somewhat improperly used: it is not referred to atoms or molecules but to the "elements" of the fluid.

[3]The reader may also be interested in an alternative derivation that starts from the Boltzmann equation, this can be found in [Shu92].

fluid at fix position must be equal to the flux (flow) of fluid through the surface S that surrounds the volume V:

$$\frac{\partial}{\partial t}\int_V \rho \mathrm{d}V = -\int_S (\rho \boldsymbol{u})\cdot \mathrm{d}\boldsymbol{S}, \tag{2.9}$$

where $\mathrm{d}\boldsymbol{S}$ is the surface element vector perpendicular[4] to S and directed (by convention) outside the volume V. Given that the volume does not change with time, we can move the derivative inside the integral on the l.h.s. of eq. 2.9. While on the r.h.s., we can make use of the divergence theorem (§C.4.3) thus obtaining two volume integrals. The resulting identity is obviously valid for every volume V, thus eliminating the integrations we obtain

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla}\cdot(\rho \boldsymbol{u}) = 0, \tag{2.10}$$

which is a **mass conservation equation**, often referred to as the **continuity equation** of the fluid.

Equation 2.10 states that mass is not created nor destroyed within the fluid. This sounds like an obvious condition, however in astrophysical fluids is not always satisfied as there may be situations when mass is in fact removed or added to the fluid at some locations. Typical cases are stellar winds or supernovae (sources of gas) or star formation regions in the ISM where gas is converted into stars thus disappearing from the fluid state (sinks of gas). Therefore, when needed, **source** or **sink terms** can be added on r.h.s. of the continuity equation.

By developing the second term of eq. 2.10 one can rewrite the same equation using the Lagrangian derivative to obtain

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} + \rho\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0. \tag{2.11}$$

This equation gives a practical definition of an **incompressible fluid**, which is a fluid in which the density of fluid elements does not change in time during its motion: $D\rho/Dt = 0$. Incompressible fluids always satisfy the condition

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0, \tag{2.12}$$

i.e. the velocity field is solenoidal. Incompressibility is a typical property of some liquids and *subsonic* gases. As we will see, it dramatically simplifies calculations and, thus, it is sometimes assumed valid, with some caveats, also for astrophysical gases.

2.5. Euler equation

The second fondamental conservation equation that we derive is that of momentum conservation or force equation, often referred to as the **Euler equation**. Here, we introduce it in term of force equation, the momentum equation is derived in §7.2. The starting point is Newton second law, which can be written *for a fluid element* of mass $\rho\delta V$ as

$$\rho\delta V\,\frac{\mathrm{D}\boldsymbol{u}}{Dt} = \boldsymbol{F}_\mathrm{P} - \rho\delta V\,\boldsymbol{\nabla}\Phi, \tag{2.13}$$

[4]Note that other texts use the notation $\boldsymbol{n}\mathrm{d}S$, which is the unity vector perpendicular to the surface multiplied by the infinitesimal surface element. The two notations are equivalent.

where the first term on the r.h.s. is the pressure force and the second is the gravity force (for the moment we are considering that these are the only two forces acting on the element). We define the pressure force as

$$\boldsymbol{F}_{\rm P} \equiv -\int_S P{\rm d}\boldsymbol{S} = -\int_V \boldsymbol{\nabla} P{\rm d}V, \tag{2.14}$$

where P is the (thermal) pressure and ${\rm d}\boldsymbol{S}$ is, again, a surface element vector, normal to the surface where the force is acting. The second equality in eq. 2.14 has been obtained thanks to a corollary of the divergence theorem (see eq. C.20), while the minus sign is due to the convention that the force is positive if the pressure acts from the fluid element towards the external medium. This happens when the internal pressure is higher than the external, i.e. when its gradient (through the surface) is negative. For an infinitesimally small volume δV the force becomes $\boldsymbol{F}_{\rm P} = -P\delta\boldsymbol{S} = -\boldsymbol{\nabla} P\delta V$ and, eliminating the volume from eq. 2.13, one obtains

$$\rho\frac{{\rm D}\boldsymbol{u}}{{\rm D}t} = -\boldsymbol{\nabla} P - \rho\boldsymbol{\nabla}\Phi, \tag{2.15}$$

which is the Euler equation in Lagrangian form. If we now use the definition of the Lagrangian derivative (eq. 2.8) we obtain the same equation in Eulerian form

$$\frac{\partial\boldsymbol{u}}{\partial t} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} = -\frac{1}{\rho}\boldsymbol{\nabla} P - \boldsymbol{\nabla}\Phi, \tag{2.16}$$

where we have also divided all term by ρ.

Equations 2.16 or 2.15 are three equations (one for each coordinate) that add up to the continuity equation (eq. 2.10) to form a system of four equations. In general, this system has six unknowns: ρ, $\boldsymbol{u}$ (three components), P and Φ. However, the potential can, in some situations, be consider as a parameter if self gravity can be neglected, i.e. if an external potential dominates. For instance, if we are dealing with accretion discs or proto-planetary discs we can assume that the gravity due to the black hole or the star dominates. Similarly in galaxies, often the dark matter halo or the stellar components can dominate the potential over the gas. In cases in which the gas self gravity cannot be ignored one needs to add to the system the **Poisson equation**, which in general terms reads

$$\nabla^2\Phi = 4\pi G(\rho + \rho_{\rm ext}), \tag{2.17}$$

where ρ and $\rho_{\rm ext}$ are gas and (non-gas) matter densities, respectively. Unfortunately this equation is often of non-trivial solution. In any case, even for a fixed and parametrised external potential, we have an unknown too many to close the system, something the we discuss in the next section.

2.6. Equation of state

The fact that a gaseous system is collisional (§1.2.1) allows us to make use of an equation of state and, in particular, the **equation of state of ideal gases** that we can write in the form

$$\frac{P}{\rho} = \frac{kT}{\mu m_{\rm p}}, \tag{2.18}$$

where μ is the atomic (or molecular) weight[5]. This equation will be extensively used in several situations and it is a fundamental property of ideal collisional systems. On the contrary, collisionless systems do not have an equation of state and their treatment often requires assumptions on the initial conditions of the motion. Unfortunately in the present context, eq. 2.18 does not allow us to close the system of equations described in §2.5 because it adds another unknown variable that was not present before: the temperature (T) of the fluid.

In general, for ideal gases, the pressure depends on two thermodynamic variables, so for instance

$$P = P(\rho, T) \qquad \text{or} \qquad P = P(\rho, S), \tag{2.19}$$

where S is the entropy. A system of this type is called **baroclinic**. In several circumstances, however, one can assume that the fluid is such that the pressure can be considered a function of the density alone:

$$P = P(\rho). \tag{2.20}$$

In this case the fluid is called **barotropic**. Peculiar cases of barotropic systems are the *isothermal* flows ($P \propto \rho$) and the **adiabatic** flows with

$$P = K\rho^\gamma, \tag{2.21}$$

where γ is the **adiabatic index**, equal to 5/3 for a mono-atomic gas, and K is a constant. More generally, one can use a power-law relation (see §3.2.1).

2.7. Energy equation

The third equation of conservation in fluid dynamics is the energy equation. To derive it we start by writing the energy density as

$$\epsilon \equiv \rho\left(\frac{1}{2}u^2 + \mathcal{U} + \Phi\right), \tag{2.22}$$

where $\mathcal{U}$ is the specific (per unit mass) internal energy[6] and Φ is the gravitational potential. If we take a (Lagrangian) time derivative of this energy we obtain

$$\frac{\mathrm{D}\epsilon}{\mathrm{D}t} = \frac{\epsilon}{\rho}\frac{\mathrm{D}\rho}{\mathrm{D}t} + \rho\frac{\mathrm{D}}{\mathrm{D}t}\left(\frac{1}{2}u^2 + \mathcal{U} + \Phi\right). \tag{2.23}$$

The first term on the r.h.s. can be substituted using the continuity equation (eq. 2.11), whereas the second term can be handled as follows.

The derivative of the specific kinetic energy is

$$\frac{\mathrm{D}}{\mathrm{D}t}\left(\frac{1}{2}u^2\right) = \boldsymbol{u}\cdot\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} \tag{2.24}$$

and we can use the Euler equation to rewrite the term on r.h.s. For the specific internal energy we can make use of the first law of thermodynamics:

$$\mathrm{d}\mathcal{U} + P\mathrm{d}\mathcal{V} = \delta Q, \tag{2.25}$$

[5]The atomic (molecular) weight is the average mass of particles in the fluid in units of the proton mass (or the hydrogen mass for a neutral gas) and m_p is the mass of the proton. So for instance in a neutral gas of hydrogen and helium ($\sim 10\%$ of particles) the atomic weight is $\mu \simeq 1.3$, in a molecular cloud, mostly made of H_2, it is $\mu \simeq 2.3$.

[6]Note that whenever a *specific* quantity is multiplied by a density one obtains the volume density of that quantity. This can be easily seen by recalling the definition of specific quantities as q/M, with q a generic quantity and M the mass of the system, and that $\rho = M/V$.

where $\mathcal{V} = 1/\rho$ is the specific volume. We remind that eq. 2.25 states that, for *reversible transformations* only, the energy (heat) transferred by the fluid per unit mass and time to the external environment is equal to its change in internal energy per unit mass plus the work done per unit mass as its volume changes. For *irreversible transformations* (e.g. viscous fluids, Chap. 7) one has the inequality $\mathrm{d}\mathcal{U} + P\mathrm{d}\mathcal{V} > \delta Q$ as the energy is partially dissipated. We neglect these effects for the moment. Given the definition of specific volume, we can rewrite eq. 2.25 as

$$\frac{\mathrm{D}\mathcal{U}}{\mathrm{D}t} = \frac{P}{\rho^2}\frac{\mathrm{D}\rho}{\mathrm{D}t} + \dot{Q}, \tag{2.26}$$

where $\dot{Q}$ is, for the moment, just a generic rate of heat exchange with the external medium and we discuss it in detail in §2.8. Again, we can use the continuity equation for the first term on the r.h.s. of eq. 2.26.

In the end, writing down all the above terms together we obtain

$$\frac{\mathrm{D}\epsilon}{\mathrm{D}t} = -\epsilon\boldsymbol{\nabla}\cdot\boldsymbol{u} - \boldsymbol{u}\cdot\boldsymbol{\nabla}P - \rho\boldsymbol{u}\cdot\boldsymbol{\nabla}\Phi - P\boldsymbol{\nabla}\cdot\boldsymbol{u} + \rho\dot{Q} + \rho\frac{\partial\Phi}{\partial t}. \tag{2.27}$$

We can rearrange the terms and use the definition of lagrangian derivative (eq. 2.8) for the l.h.s. term to obtain the final expression for the equation of **conservation of energy density** in the fluid:

$$\frac{\partial\epsilon}{\partial t} + \boldsymbol{\nabla}\cdot\left[(\epsilon + P)\,\boldsymbol{u}\right] = \rho\dot{Q} \tag{2.28}$$

where, for simplicity, we have considered non-time-variable gravitational potentials and neglected the term $\partial\Phi/\partial t$.

The physical meaning of eq. 2.28 can be unveiled by integrating each member over a generic volume containing a portion of the fluid and using the divergence theorem (§C.4.3) for the second term of the l.h.s. This leads to

$$\frac{\partial}{\partial t}\int_V \epsilon\mathrm{d}V + \int_S \epsilon\boldsymbol{u}\cdot\mathrm{d}\boldsymbol{S} + \int_S P\boldsymbol{u}\cdot\mathrm{d}\boldsymbol{S} = \int_V \rho\dot{Q}\mathrm{d}V. \tag{2.29}$$

The first term on the l.h.s. of eq. 2.29 is clearly the variation of the total energy within the volume[7]. The second term is the flux of energy through the surface that encloses the volume V. This flux is carried by the motion of the fluid itself, it is therefore an *advective* energy transfer. The third term is the work done (per unit of time) by pressure forces on the fluid within that surface. To understand this, we recall that the work can be defined as $\mathrm{d}W \equiv \boldsymbol{F}\cdot\mathrm{d}\boldsymbol{r}$ and its time variation (considering now the pressure force, $\boldsymbol{F} = \boldsymbol{F}_\mathrm{p}$) can be written as

$$\frac{\mathrm{d}W}{\mathrm{d}t} = \boldsymbol{F}_\mathrm{p}\cdot\boldsymbol{u} = -\int_S P\boldsymbol{u}\cdot\mathrm{d}\boldsymbol{S}, \tag{2.30}$$

where we have used the definition of pressure force from eq. 2.14.

In conclusion, the energy equation states that the variation of energy within a generic volume is due either to energy leaving (or entering) that volume because of the fluid motion and/or to the fluid doing (or being subject to) some work on (by) the external medium. The leftover term on the r.h.s. describes potential heat exchanges with the external medium. In the next section we list the possible reasons for this heat exchange in astrophysical fluids.

[7]Note that we can move the derivative outside the integral because the volume that we are considering does not change with time.

2.8. Heat exchange

The exchange of heat between the fluid and the external environment, represented by the generic terms $\dot{Q}$ in eq. 2.28, can take several forms. In this section we look through the four main ways in which an astrophysical fluid can exchange heat; two of these: thermal conduction and radiation will eventually be included in the energy equation as new terms on the r.h.s.

2.8.1. Thermal conduction. Also called **heat conduction**, it occurs whenever there are temperature gradients within a fluid and is generated by the collisions between elementary particles (atoms, molecules and free electrons). Let us consider the idealised case of two similar fluids at different temperatures and *in contact* with each other. Within the single fluid, as we have seen (§1.2.1), collisions are frequent and they thermalise the systems assuring the onset of a Maxwellian distribution of the particle speeds (§1.2.2). However, at the interface between the two fluids, collisions will also occur between particles belonging to the two different fluids. Assuming that the vast majority of these collisions are elastic, this phenomenon has the consequence of transferring energy from the fluid at higher temperature to that at lower temperature. In practice, particles of the fluid at lower temperature close to the contact surface will acquire faster random motions raising (locally) the fluid temperature, while on the other side of the contact surface, the temperature will decrease. The macroscopic effect is a smoothing of the initial temperature gradient between the fluids and it is the essence of thermal conduction. Note that, given its nature, thermal conduction is a *local* phenomenon that occurs only in regions where there are strong gradients and these regions are expected to be relatively thin: of the order of the mean free path of the fluid particles.

In mathematical terms we can assume the *flux* of thermal conduction to be linearly dependent on the temperature gradient

$$\boldsymbol{q} = -\kappa \boldsymbol{\nabla} T, \tag{2.31}$$

where κ is a coefficient called **thermal conductivity** and the minus sign is justified by the fact that heat flows away from hot bodies and so in the opposite direction with respect to the temperature gradient. Equation 2.31 can be seen as a Taylor expansion in T where the first term is obviously zero because thermal conduction must vanish when there is no temperature gradient and we are neglecting the Laplacian term in T. In general, κ in eq. 2.31 is however not a constant as it varies with temperature and pressure. L. Spitzer [Spi78] derived what is considered the standard formula for heat conduction in the ISM with κ being a strong function of the temperature:

$$\kappa = -\kappa_{\mathrm{Sp}} T^{5/2}, \tag{2.32}$$

with the constant

$$\kappa_{\mathrm{Sp}} = \frac{1.84 \times 10^{-5}}{\ln \Lambda} \,\mathrm{erg\,s^{-1}\,cm^{-1}\,K^{-7/2}} \tag{2.33}$$

and the Coulomb logarithm

$$\ln \Lambda \simeq 29.7 + \ln \left[\left(\frac{n_{\mathrm{e}}}{1\,\mathrm{cm^{-3}}} \right)^{-1/2} \left(\frac{T_{\mathrm{e}}}{10^6\,\mathrm{K}} \right) \right], \tag{2.34}$$

where n_{e} and T_{e} are the electron density and temperature, respectively. The value of $\ln \Lambda$ is thus always around 30 in astrophysical conditions [CM77].

In several astrophysical situations one has to deal with a *plasma* and thus with non-negligible magnetic fields present in the fluid. The effect of the magnetic field in this context is that particles encounter a much stronger resistance to move perpendicularly to the field lines (Chap. 9). This makes thermal conduction both *anisotropic* and weaker than one would expect from eq. 2.32. Taking account of the magnetic field in the calculations is rather complex and the simplified solution of multiplying the r.h.s. of 2.32 by a **dimming factor** f is often adopted, leading to

$$\boldsymbol{q} = -f\,\kappa_{\rm Sp} T^{5/2} \boldsymbol{\nabla} T, \tag{2.35}$$

with $f = 1$ in the absence of magnetic fields. The range of possible values for f commonly used in astrophysics is $0.01 - 0.2$, with a possible upper limit at $f \simeq 0.2$ [NM01], thus the suppression introduced by the magnetic field is quite strong! This is obviously not a perfect solution as ideally one should include anisotropic thermal conduction in full magneto-hydrodynamic (MHD) simulations. Some recent codes do indeed offer this possibility.

The flux of heat conduction enters the energy equation through a divergence operation:

$$\frac{\partial \epsilon}{\partial t} + \boldsymbol{\nabla} \cdot [(\epsilon + P)\boldsymbol{u}] = -\boldsymbol{\nabla} \cdot \boldsymbol{q} + \rho \dot{Q}_{\rm (others)}, \tag{2.36}$$

where $Q_{\rm (others)}$ are the other sources of heat exchange that we have not yet considered. To understand eq. 2.36 recall that the two terms on the l.h.s. subject to the divergence operator are also *fluxes*. Indeed thermal conduction is sometime more explicitly incorporated with these fluxes and moved to the l.h.s.[8]

2.8.2. Convection. This is a *macroscopic* phenomenon that causes the motion of large portions of fluid in the presence of temperature gradients. The typical condition for the development of convection is a strong temperature gradient in a fluid stratified by the gravity field, this is what happens in the interiors of some stars for instance or in a cold room when a radiator (without fan) is turned on. We discuss the criteria for the development of convection when we analyse the fluid instabilities (Chap. 6). Here, we just mention that in the case when convection does not affect the whole fluid and one can identify a convective region, the heat transfer through this region can be treated as an *effective conduction*. In practice, as for conduction one has, in the convective region, a (macroscopic) transfer of heat from the hotter to the colder part of the fluid.

2.8.3. Radiation. This is the most important and effective way of transferring heat to the external medium in astrophysical fluids. It occurs because particles emit and absorb photons due to a variety of processes that we have partially outlined in Chap. 1. We distinguish between two opposite scenarios.

[8]Note that for a static fluid ($\boldsymbol{u} = 0$) and neglecting gravity we have $\epsilon = U = (3/2) n k_{\rm B} T$. If we then assume a constant κ for the conduction term and neglect other heat sources we obtain something reminiscent of the familiar heat diffusion equation $\frac{\partial T}{\partial t} - k\,\nabla^2 T = 0$, where k is the thermal conductivity. This equation is commonly used for heat transfer in solids.

2.8.3.1. *Radiative cooling.* When particles of the fluid emit radiation that manages to cross its volume boundaries, we say that the fluid is experiencing **radiative cooling**. Photons escaping the volume where the fluid is located are effectively removing kinetic energy from the particles that make up the fluid. This is very easy to see in the case of transitions caused by inelastic collisions (e.g. forbidden lines in H II regions, §1.1.2). In this case, part of the kinetic energy of the electron colliding with, say, the oxygen ion is "used" to shift a bound electron to an higher energy level. A spontaneous decay of this electron with the emission of a photon completes therefore the transformation of part of the fluid internal energy into radiation.

A common way to describe radiative cooling is to introduce a **cooling function** Λ that will in general depend on at least two thermodynamic variables (at fixed metallicity), for instance

$$\Lambda = \Lambda(T, \rho). \tag{2.37}$$

In the idealised conditions of **collisional ionisation equilibrium** (CIE) the cooling function can be expressed as an only function of the temperature of the gas (again at fixed metallicity). These conditions are valid when the density of the gas are low ($n \ll n_{\rm crit}$; §1.1.2) and there is no contribution from photoionisation.

More specifically, in CIE the atoms are assumed to be always in their fundamental state. As a consequence of a collision (in particular with free electrons) they can either be excited or ionised. Any collisional ionisation is followed by a recombination and the electron cascades down to the fundamental state through the various emissions. Collisional excitations are also promptly balanced by radiative de-excitations (collisional de-excitations are negligible). Clearly, the CIE condition is *not* akin to local thermodynamical equilibrium (§1.3) where the different states are populated according to the Boltzmann law. As mentioned, there is also no contribution from photoionisation. Finally, the gas is also assumed to be optically thin and so every photon emitted can escape from the system.

The CIE cooling function has been calculated for gases at different temperature and metallicities (Fig. 1). The conditions of CIE are typically fulfilled at temperature $T > 10^4\,\mathrm{K}$. The peak of the function at $\log(T/\,\mathrm{K}) \simeq 4.1$ is mostly due to the recombination of collisionally ionised hydrogen, while the peak at $T \approx 10^5\,\mathrm{K}$ (in the zero-metallicity function) is due to the recombination of helium. Between $\log(T/\,\mathrm{K}) \approx 4.1$ and $\log(T/\,\mathrm{K}) \approx 7$ and for metallicity $Z > 0$ the cooling is strongly dominated by metal forbidden and resonance lines. At $\log(T/\,\mathrm{K}) > 7$, all atoms including the heaviest ones are fully ionised and the remaining cooling is given only by the bremsstrahlung radiation (§1.1.2). When CIE conditions are not met (e.g. at interfaces between cold and hot gases) one should use non-equilibrium functions (NEQ) (see [SD93]). In this case, different cooling regimes emerge, e.g. isocoric and isobaric cooling and the cooling function is not only a function of T. NEQ functions tend to be very similar to the CIE function for $T > 10^{5.5}\,\mathrm{K}$, but they can depart from it quite dramatically at lower temperatures. For temperature $T < 10^4\,\mathrm{K}$ (mostly neutral hydrogen) the cooling function strongly depends on the ionisation fraction of the gas. Moreover, at much lower temperatures, molecules start to form and the treatment of cooling becomes much more complex. We refer the interested reader to specific books on this topic, e.g. [Tie05].

2.8.3.2. *Radiative Heating.* Radiation produced by external sources (e.g. massive stars) can enter the volume of fluid of interest and interact with the fluid particles transferring directly or indirectly its energy to the gas. This process is

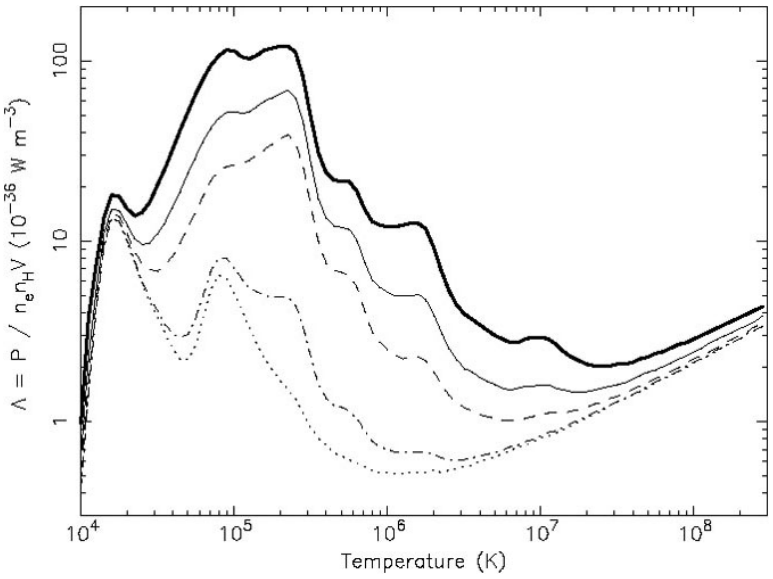

FIGURE 1. Cooling functions assuming collisional ionisation equilibrium for 5 different metallicities from zero (dotted line) to Solar (thick solid line); from [SD93].

called **radiative heating** and it can increase the internal energy of the fluid and its temperature. The typical radiation that produces this heating is constituted by ionising photons (e.g., in UV and X-rays) that hand over their surplus energy to the newly freed electrons. Another very effective form of heating is the **photoelectric heating** when an electron is "extracted" from a large molecule or a dust grain. Moreover, given their very large cross sections, dust particles can be heated by absorbing a photon, without changing their electronic configuration but just their velocity and thus the dust temperature. Hot dust is indeed typically observed around strong sources of radiation like star-forming regions or AGNs.

In general, radiative heating is more difficult to parameterise than cooling as it depends not only on two variables but also on the positions in the ISM, e.g. regions close to UV-emitting stars and on the ISM composition (e.g. metallicity and dust content). Different regions and phases have completely different heating functions and the details of these phenomena go beyond the scope of these notes. Again, we refer the reader to more specific books on the subject, e.g. [Tie05] or to specialised articles presenting astrophysical photoionisation codes, e.g. [FCG$^+$17].

2.8.4. High energy particles. The ISM of our Galaxy is constantly "pierced" by elementary particles streaming through it at relativistic velocities, so-called cosmic rays (§1.1.5). They are largely produced in supernova remnants and their paths is randomised by swirling around the Galactic magnetic field. These particles do constitute an important source of heating especially effective in regions where electromagnetic radiation cannot penetrate because it is shielded by a thick layer of gas. Typical regions of this kind are the very inner cores of molecular clouds. There, cosmic rays are in fact the main heating mechanism. The treatment of cosmic-ray heating requires a parameterisation similar from that of radiative heating.

2.8.5. Energy equation with heat terms. We now modify the energy equation derived in §2.7 to take into account thermal conduction and radiative heat exchanges. Both contributions are added to the r.h.s. of eq. 2.28 as follows:

$$\frac{\partial \epsilon}{\partial t} + \boldsymbol{\nabla} \cdot [(\epsilon + P)\boldsymbol{u}] = -\boldsymbol{\nabla} \cdot \boldsymbol{q} - \rho \mathcal{L}(\rho, T, Z, \boldsymbol{x}...), \tag{2.38}$$

where with $\mathcal{L}$ we indicate the *energy loss per unit time and mass* due to radiation (possibly also including cosmic rays). As mentioned, this rate has, in general, a complex dependency on density, temperature, metallicity and location. The $\mathcal{L}$ function can be split in the contribution of cooling and heating:

$$\mathcal{L} = \Lambda - \Gamma, \tag{2.39}$$

where with Λ and Γ we indicate generic cooling and heating functions. At temperatures above $T > 10^4\,\mathrm{K}$ we can often assume to be in conditions such that collisional ionisation equilibrium is valid and so the energy equation becomes

$$\frac{\partial \epsilon}{\partial t} + \boldsymbol{\nabla} \cdot [(\epsilon + P)\boldsymbol{u}] = -\boldsymbol{\nabla} \cdot \boldsymbol{q} - n_\mathrm{e} n_\mathrm{t} \Lambda_\mathrm{CIE}(T, Z) + \rho\Gamma, \tag{2.40}$$

where n_e and n_t are the number densities of electrons and atoms (neutrals or ions) respectively and Λ_CIE is the cooling function in Fig. 1. Note that the units of this function are $[\Lambda] = \mathrm{erg\,s^{-1}\,cm^3}$, which justifies the need to multiply it by the square of the numeric density to obtained an energy density per unit time.

2.8.6. Cooling time. We can use the equation of energy density conservation (eq. 2.40) to estimate one of the most important timescales in astrophysical hydrodynamics: the **cooling time**. If we consider a static fluid and we ignore thermal conduction, we have

$$\dot{\epsilon} \approx -n_\mathrm{e} n_\mathrm{t} \Lambda_\mathrm{CIE}(T), \tag{2.41}$$

where we are assuming that the gas is at a certain constant metallicity Z. Being the gas at rest, its total energy is essentially internal energy and we can use the relation for an ideal gas:

$$\epsilon \approx U = \frac{3}{2} n k_\mathrm{B} T = \frac{3}{2}(n_\mathrm{e} + n_\mathrm{t}) k_\mathrm{B} T, \tag{2.42}$$

where we have indicated with n the *total* number density of particles. Then a characteristic time for the radiative cooling process can be estimated as:

$$t_\mathrm{cool} \equiv \frac{\epsilon}{|\dot{\epsilon}|} = \frac{3}{2} \frac{n}{n_\mathrm{e} n_\mathrm{t}} \frac{k_\mathrm{B} T}{\Lambda_\mathrm{CIE}(T)} \simeq \frac{3 k_\mathrm{B} T}{n_\mathrm{e} \Lambda_\mathrm{CIE}(T)}, \tag{2.43}$$

where, in the last equality, we have assumed that the gas is fully ionised ($n_\mathrm{t} = n_\mathrm{i}$; negligible density of neutral atoms) and that $n_\mathrm{e} \simeq n_\mathrm{i}$[9]. This assumption is good for $T \gtrsim 3 \times 10^4\,\mathrm{K}$.

Equation 2.43 is widely used in astrophysical contexts to estimate the typical time beyond which radiative losses in a collisionally-ionised hot medium cannot be neglected. Note, however, that this time is only indicative as astrophysical fluids are dynamic systems and their cooling time can significantly change during their evolution (because temperature and density variations). We see an example of this when we follow the evolution of supernova remnants (Chap. 5). There the cooling time initially exceeds the lifetime of a SNR by orders of magnitude. However, as

[9]The exact value for fully ionised astrophysical plasmas is $n_\mathrm{e} \simeq 1.2 n_\mathrm{i}$ mostly because of the contribution of helium.

the shell expands and loses energy, its temperature drops and the cooling time also decreases quickly, soon becoming of the order of the age of the system. At that point, radiative cooling cannot be anymore neglected in the calculation and indeed it significantly affects the subsequent evolution.

2.9. Other forms of the energy equation

Let us consider the Euler equation in the form of eq. 2.16 and take the dot product of every member by $\rho\boldsymbol{u}$ to obtain

$$\rho\boldsymbol{u}\cdot\frac{\partial\boldsymbol{u}}{\partial t}+\rho\boldsymbol{u}\cdot(\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u})=-\boldsymbol{u}\cdot\boldsymbol{\nabla}P-\rho\boldsymbol{u}\cdot\boldsymbol{\nabla}\Phi. \tag{2.44}$$

The various terms of this equation can be manipulated given the following two identities:

$$\frac{\partial}{\partial t}\left(\frac{1}{2}\rho u^2\right)=\rho\boldsymbol{u}\cdot\frac{\partial\boldsymbol{u}}{\partial t}+\frac{1}{2}u^2\frac{\partial\rho}{\partial t} \tag{2.45}$$

$$\boldsymbol{\nabla}\cdot\left(\frac{1}{2}\rho u^2\boldsymbol{u}\right)=\rho\boldsymbol{u}\cdot(\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u})+\frac{1}{2}u^2\boldsymbol{\nabla}\cdot(\rho\boldsymbol{u}). \tag{2.46}$$

We substitute the two l.h.s. terms in eq. 2.44 with the expressions obtained from eqs. 2.45 and 2.46. At this point, the second r.h.s. terms of eqs. 2.45 and 2.46 vanish because of the continuity equation (eq. 2.10) and we are left with

$$\frac{\partial\epsilon_{\rm K}}{\partial t}+\boldsymbol{\nabla}\cdot(\epsilon_{\rm K}\boldsymbol{u})=-\boldsymbol{u}\cdot\boldsymbol{\nabla}P-\rho\boldsymbol{u}\cdot\boldsymbol{\nabla}\Phi, \tag{2.47}$$

where we have indicated with $\epsilon_{\rm K}\equiv(1/2)\rho u^2$ the kinetic energy density. Equation 2.47 is called the **work equation** and expresses the conservation of kinetic energy in the fluid.

We can get a grasp of its meaning by integrating it over a generic volume V. The first term on the l.h.s. is the variation of the kinetic energy within the volume. This is determined by the loss of kinetic energy via advection through the surface S enclosing the volume (second term of the l.h.s.) plus two other terms. The first of these two is the work done on the surface by the pressure, notice that, unlike in eq. 2.30, this is only the mechanical work (pushing or contrasting a push) at the surface without the contribution of compression ($P\boldsymbol{\nabla}\cdot\boldsymbol{u}$) because this latter has no effect on the variation of the kinetic energy (we will encounter this compression term just below). The last term is non-negligible only if the potential varies appreciably across the fluid volume.

We can also derive an analogous conservation equation for the internal energy. Consider the conservation of the total energy density (eq. 2.28) and subtract from it the work equation (eq. 2.47). Neglecting for simplicity the contribution of the gravitational potential and thus taking $\epsilon=\epsilon_{\rm K}+\rho\mathcal{U}$, we obtain

$$\frac{\partial(\rho\mathcal{U})}{\partial t}+\boldsymbol{\nabla}\cdot[(\rho\mathcal{U}+P)\boldsymbol{u}]-\boldsymbol{u}\cdot\boldsymbol{\nabla}P=-\boldsymbol{\nabla}\cdot\boldsymbol{q}-\rho\mathcal{L}. \tag{2.48}$$

Then, expanding the first and second term and making use of the continuity equation (eq. 2.10), one finds

$$\rho\frac{D\mathcal{U}}{Dt}+P\boldsymbol{\nabla}\cdot\boldsymbol{u}=-\boldsymbol{\nabla}\cdot\boldsymbol{q}-\rho\mathcal{L}, \tag{2.49}$$

which is the specific **internal energy equation**. It is easy to see that this equation is essentially the first law of thermodynamics for reversible transformations written

for the fluid element as the second term on the l.h.s. is the variation of the volume due to the compression (or expansion) of the fluid (recall eq. 2.26).

2.9.1. The heat equation. We remind that the first law of thermodynamics can be expressed in terms of exact differentials using entropy instead of heat exchange. Considering all specific quantities this takes the form

$$T\mathrm{d}\mathcal{S} = \mathrm{d}\mathcal{U} + P\mathrm{d}\mathcal{V}, \tag{2.50}$$

which is valid for *all* transformations, both reversible and irreversible. Substituting $1/\rho$ to the specific volume, multiplying by ρ the whole equation and differentiating with respect to time, the r.h.s. of 2.50 becomes the l.h.s. of eq. 2.49. In this way we obtain one of the most used forms of the energy equation in a fluid, often referred to as the **heat equation**:

$$\rho T\frac{\mathrm{D}\mathcal{S}}{\mathrm{D}t} = -\boldsymbol{\nabla}\cdot\boldsymbol{q} - \rho\mathcal{L}, \tag{2.51}$$

where, as mentioned, $\mathcal{S}$ is the specific entropy[10].

In the absence of thermal conduction, eq. 2.51 simplifies to

$$T\frac{D\mathcal{S}}{Dt} = -\Lambda + \Gamma. \tag{2.52}$$

A transformation is said to be **adiabatic** when there is no exchange of heat with the external environment, i.e. all the terms on the r.h.s. of eq. 2.52 are equal to zero[11]. As a consequence, in an adiabatic system the entropy of fluid elements does not change with time:

$$\frac{D\mathcal{S}}{Dt} = \frac{\partial\mathcal{S}}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\mathcal{S} = 0. \tag{2.53}$$

Flows of this kind are sometimes called **isentropic**, although this term is better referred to systems with spatially constant entropy, i.e. when all the elements have the same entropy ($\boldsymbol{\nabla}\mathcal{S} = 0$). This terminology can generate some confusion as in the latter case, the global (constant) entropy can in principle change with time.

2.10. Sound waves

Let us consider a homogeneous medium with constant pressure (P_0) and density (ρ_0). We further assume that the medium is initially stationary or, equivalently, we put ourselves in the rest-frame where this is true ($\boldsymbol{u}_0 = 0$). We consider infinitesimal adiabatic perturbations of all these quantities for a generic fluid element such that[12]

$$\begin{aligned} P &= P_0 + \Delta P \\ \rho &= \rho_0 + \Delta\rho \\ \boldsymbol{u} &= \Delta\boldsymbol{u}. \end{aligned} \tag{2.54}$$

[10]Note that $\mathrm{D}\mathcal{S}/\mathrm{D}t$ is the variation of specific entropy of a fluid element, $T(\mathrm{D}\mathcal{S}/\mathrm{D}t)$ is the variation of specific heat while $\rho T(\mathrm{D}\mathcal{S}/\mathrm{D}t)$ is the variation of the heat density (heat per unit volume).

[11]Note that the r.h.s. of eq. 2.52 can also be null when there is complete balance between heating and cooling. Such a system can exchange heat with the environment but it is in a state of *equilibrium* between cooling and heating, see also §1.3.

[12]Note that these perturbations are Lagrangian because they refer to the fluid element, however in the case of a homogeneous medium these are the same as the Eulerian perturbations, e.g. if $\boldsymbol{\nabla}P = 0$ then $\Delta P \equiv \delta P + \delta\boldsymbol{r}\cdot\boldsymbol{\nabla}P = \delta P$.

We substitute the above perturbations on the continuity (eq. 2.10) and Euler (eq. 2.16) equations to obtain

$$\begin{cases} \dfrac{\partial \Delta\rho}{\partial t} + \rho_0 \boldsymbol{\nabla} \cdot \Delta \boldsymbol{u} = 0 \\ \dfrac{\partial \Delta \boldsymbol{u}}{\partial t} = -\dfrac{1}{\rho_0} \boldsymbol{\nabla} \Delta P \end{cases} \tag{2.55}$$

where we have neglected all the second order terms (linear perturbation analysis).

Let us now write the gas pressure as a function of density and entropy $P = P(\rho, \mathcal{S})$. Given that our perturbations are infinitesimal and adiabatic we can write the perturbation in pressure as[13]

$$\Delta P = \left(\frac{\partial P}{\partial \rho} \right)_{\mathcal{S}} \Delta\rho, \tag{2.56}$$

where the derivative is taken at constant entropy. We now define the **sound speed** c_s of the fluid[14] such that

$$c_\mathrm{s}^2 \equiv \left(\frac{\partial P}{\partial \rho} \right)_{\mathcal{S}}. \tag{2.57}$$

Let us then take the time derivative of the first equation 2.55 and the divergence of the second equation, neglecting the second order terms and subtract the second equation from the first to obtain

$$\boldsymbol{\nabla}^2 \Delta\rho - \frac{1}{c_\mathrm{s}^2} \frac{\partial^2 \Delta\rho}{\partial t^2} = 0, \tag{2.58}$$

which is the *wave equation* for **sound waves** moving at speed c_s. For mono-dimensional perturbations, in which all quantities depend on one coordinate, say x, we have a so-called *plane wave* and eq. 2.58 can be simply rewritten as

$$\frac{\partial^2 \Delta\rho}{\partial x^2} - \frac{1}{c_\mathrm{s}^2} \frac{\partial^2 \Delta\rho}{\partial t^2} = 0. \tag{2.59}$$

The general solution of eq. 2.59 is given by

$$\Delta\rho = f(x - c_\mathrm{s} t) + g(x + c_\mathrm{s} t) \tag{2.60}$$

where f and g are generic functions. This solution is periodic and it can be written in terms of a Fourier series. In general, in linear perturbation analysis, the so-called superposition principle is valid, i.e. the sum of two solutions is a solution, thus the terms of our Fourier series will *all* be solutions. The problem can be, therefore, greatly simplified by studying only the simplest (first) mode of the series:

$$\Delta\rho = \widetilde{\Delta\rho} e^{\mathrm{i}(kx - \omega t)} \tag{2.61}$$

where $\widetilde{\Delta\rho}$ is the amplitude of the density perturbation, k is the wavenumber and ω is the frequency of the perturbation. The positive sign for the spatial part and negative sign for the temporal part are conventions.

If we now use $\Delta\rho$ from eq. 2.61 and substitute into eq. 2.59 we obtain the expression

$$c_\mathrm{s} = \frac{\omega}{k}, \tag{2.62}$$

---

[13] In general terms for $P = P(\rho, \mathcal{S})$ one has $\mathrm{d}P = \left(\frac{\partial P}{\partial \rho}\right)_{\mathcal{S}} \mathrm{d}\rho + \left(\frac{\partial P}{\partial \mathcal{S}}\right)_{\rho} \mathrm{d}\mathcal{S}$.

[14] Note that this is by construction the *adiabatic* sound speed (see §2.10.1).

which is the **dispersion relation** of a sound wave. Note that these waves are therefore non dispersive as their phase speed coincide with their group speed[15].

Analogously to what done for perturbations in density, one could derive an equation for perturbations in velocity. Such an equation will have a similar general solution as that in eq. 2.60, whose behaviour can again be studied by simply considering

$$\Delta u = \widetilde{\Delta u} e^{\mathrm{i}(kx-\omega t)}, \tag{2.63}$$

with $\widetilde{\Delta u}$ amplitude of the velocity perturbation. Note that the wavenumber and the frequency in eqs. 2.61 and 2.63 will naturally be the *same* because this is the same oscillation manifesting itself in the density and in the velocity. Equation 2.63 shows a fundamental property of sound waves, i.e. that they are **longitudinal** or **compression waves**: the oscillation takes place along the direction of the wave propagation (here we are considering only one direction but the same relation can be obtained in more dimensions). This means that a sound wave is nothing more that a (somewhat infinitesimal) compression of the medium ($\Delta\rho$ perturbation) that propagates at the sound speed.

We can now substitute eqs. 2.61 and 2.63 into the first eq. of 2.55 to obtain the following equality for the amplitude of the velocity oscillation:

$$\widetilde{\Delta u} = \frac{\widetilde{\Delta\rho}}{\rho_0} c_\mathrm{s}, \tag{2.64}$$

from which it is clear that $\widetilde{\Delta u} \ll c_\mathrm{s}$. This result shows that the amplitude of the oscillations in a sound wave are much smaller than the sound speed itself. So a fluid element can oscillate very little while transferring very quickly this oscillation to the adjacent fluid element and that is what determines the propagation speed of the wave. We recall that particles in the fluid are moving quite rapidly because of their thermal motions. Therefore, it should not come as a surprise that the sound speed is of the order of the thermal speed (eqs. 1.25); it is in fact this thermal motion that transfers momentum and energy forward to other particles. The amplitude of the velocity oscillation simply determines the strength of the waves and how far it will reach. In fact, in realistic circumstances, every wave will damp beyond a certain distance, stronger waves travel *at the same speed* as weak waves but damp further away.

Given the above, we can now define the **crossing time** of a gaseous system as

$$t_\mathrm{cross} \equiv \frac{\mathrm{L}}{c_\mathrm{s}}, \tag{2.65}$$

where L is the size of the system. The crossing time is thus the time a sound wave takes to go from one side of the fluid to the other. In practical terms, we can consider the crossing time as the time it takes for a *hydrodynamic information* to travel in a fluid. If a drop or an increase in pressure (or density) occurs somewhere in the fluid, sound waves will try to re-establish the equilibrium. However, this cannot be achieved more quickly than in a crossing time, because the carrier of the information that this change has occurred travels at c_s. Thus, the crossing time is a measure of the time the fluid takes to establish pressure equilibrium or to react

[15] We remind that the **phase speed** (ω/k) is the speed of a point in the wave, for instance a peak or a trough while the **group speed** ($d\omega/dk$) is the speed of the centroid of the wave, which is also the speed at which energy propagates.

to changes of its pressure. Instead, the fluid will not be able to react to phenomena taking place at faster pace (e.g. the passage of a supersonic wave, Chap. 4).

2.10.1. Adiabatic and isothermal sound speed. We can expect the actual sound speed of a fluid to be confined between the formulations for the two extreme regimes for a barotropic fluid (§2.6): adiabatic and isothermal. The adiabatic sound speed can be calculated from eq. 2.57 and, recalling that in this regime $P = K\rho^\gamma$, we obtain

$$c_\mathrm{s} \equiv \sqrt{\left(\frac{\partial P}{\partial \rho}\right)_\mathcal{S}} = \sqrt{\gamma \frac{P}{\rho}} = \sqrt{\gamma \frac{kT}{\mu m_\mathrm{p}}}, \tag{2.66}$$

where, as usual, γ is the adiabatic index and we recall that $\gamma = C_\mathrm{p}/C_\mathrm{V}$, with C_p and C_V specific heats at constant pressure and volume, respectively (§D.1.4). In the isothermal regime the same equation gives instead

$$c_\mathrm{s} = \sqrt{\left(\frac{\partial P}{\partial \rho}\right)_T} = \sqrt{\frac{kT}{\mu m_\mathrm{p}}}, \tag{2.67}$$

with a factor $\sqrt{\gamma}$ of difference.

The type of sound waves that develop in a medium do not necessarily reflect the unperturbed state of this medium. They instead depend on the way the medium *reacts to a perturbation*. For instance, if the medium tends to radiate away its energy then the waves will likely be closer to isothermal than adiabatic. An interesting example is the Earth's atmosphere, which at first approximation is nearly isothermal. However, the measured sound speed in the atmosphere (at sea level) is $344\,\mathrm{m\,s^{-1}}$, which is almost exactly what is obtained by using eq. 2.66 for the adiabatic sound speed (with $T = 293\,\mathrm{K}$, $\gamma = 7/5$ and $\mu = 29$)[16].

2.10.2. Sound waves in the ISM. We have no secure way to determine whether the sound speed of the ISM should be closer to adiabatic or isothermal as this will depend on its temperature and other conditions. Table 1 shows the sound speeds of the different phases of the ISM calculated using eqs. 2.67 and 2.66. Note that they are very similar the to thermal speeds calculated in §1.1.1.

Phase	T (K)	μ	Isothermal c_s ($\mathrm{km\,s^{-1}}$)	Adiabatic c_s ($\mathrm{km\,s^{-1}}$)
CNM	70	1.3	0.7	0.9
WNM	8000	1.3	7.1	9.2
WIM	1×10^4	0.6	11.7	15.1
HIM	1×10^6	0.6	117.3	151.4

TABLE 1. Approximate sound speeds of the various components of the ISM, also described in §1.1.

We remind that typical r.m.s. velocities measured in the atomic neutral medium with the H I emission line is of the order of $10\,\mathrm{km\,s^{-1}}$, reaching also higher values in the central regions of galaxies. This broadening of the line is largely ascribed to turbulence (Chap. 8), but it is interesting to note that it is not far from the sound

[16]Note that the main compound of air is N_2 with molecular mass $\mu_{N_2} = 28.02$.

speed of the WNM. In general we will consider as supersonic any motions that exceeds the sound speeds in Table 1. To this end, we can then define the **Mach number** as

$$\mathcal{M} \equiv \frac{v}{c_{\rm s}}, \tag{2.68}$$

where v is the speed of a portion of the fluid[17]. At supersonic speeds, $\mathcal{M}$ is larger than 1. The typical Mach numbers of phenomena that we describe in these notes (e.g. material put in motion by supernova explosions) are $\mathcal{M} \gg 1$.

[17]Note that, in Earth's applications, this could be the speed of a body, e.g. an airplane.

CHAPTER 3

Hydrostatic equilibrium

In this Chapter, we see a few examples of astrophysical systems in hydrostatic equilibrium. We use some of the concepts developed in the previous chapter, in particular the Euler equation (§2.5) written for static systems with different geometry.

3.1. Gas layer

We start with the derivation of the density profile of an isothermal layer in hydrostatic equilibrium. As mentioned in §1.3, this is the equilibrium in which the pressure force of a gaseous system equates the gravity force, either its own or external. We first take the Euler equation (eq. 2.16) and impose an equilibrium condition. For this, we must have that the time derivatives must be zero ($\partial/\partial t = 0$) and the velocities must be null ($\boldsymbol{u} = 0$) thus the whole l.h.s. of eq. 2.16 is null. Notice that this is analogous to say that there is no net acceleration on the generic fluid element, from eq. 2.15. Thus, in general, for a system in **hydrostatic equilibrium** we have

$$\boldsymbol{\nabla} P = -\rho \boldsymbol{\nabla} \Phi, \tag{3.1}$$

where P is the pressure and Φ is the gravitational potential.

3.1.1. Self-gravitating layer. Consider now a homogeneous layer of gas (slab), infinite in x and y (Cartesian coordinates) and stratified in z under the effect of *its own gravity*. We aim to derive its vertical density profile $\rho(z)$. To this end, we write a system of equations that comprises the Euler equation, the Poisson equation and a convenient equation of state. For simplicity we assume that the gas is *isothermal*, which we have seen to be a reasonable assumption for any of the ISM phases (Chap. 1). The system of equations reads therefore as follows:

$$\begin{cases} \dfrac{\mathrm{d}P}{\mathrm{d}z} = -\rho \dfrac{\mathrm{d}\Phi}{\mathrm{d}z} \\ \dfrac{\mathrm{d}^2\Phi}{\mathrm{d}z^2} = 4\pi G \rho \\ P = c_\mathrm{s}^2 \rho, \end{cases} \tag{3.2}$$

where the gravitational potential Φ is determined *only* by the gas.

We note that, if we aim to determine the thickness of a galaxy ISM, as we see below, a more realistic configuration would be that of a rotating disc (instead of an infinite slab). In that case we can rewrite the above equations in cylindrical coordinates with the difference that now the potential and other quantities would be function of R and z. Luckily, for any given R, the first of eqs. 3.2 remains unaltered

expect for the substitution of partial derivatives in z. The Poisson equation instead leads to[1]

$$\frac{\partial^2\Phi}{\partial z^2} = 4\pi G\rho - \frac{1}{R}\frac{\partial}{\partial R}\left(R\frac{\partial\Phi}{\partial R}\right), \tag{3.3}$$

thus with an extra term. However, if we restrict ourselves to regions of space close to the plane of the disc ($z \approx 0$), the second term on the r.h.s. of eq. 3.3 contains the radial derivative of the **circular speed**, which is defined by the relation

$$v_{\rm c}^2(R,0) \equiv \left(R\frac{\partial\Phi}{\partial R}\right)_{z=0}, \tag{3.4}$$

where the derivative is evaluated at $z = 0$. This is very handy as the circular speed is nearly constant with R in large areas of the discs of spiral galaxies, i.e. the rotation curves are *flat*. Thus the second term of the r.h.s. of eq. 3.3 vanishes or nearly so and therefore, the system of eqs. 3.2 provides a reasonably good approximation for the vertical distribution of the gas also in the case of a self-gravitating galaxy disc.

To proceed with our derivation of the vertical density of the isothermal slab, we substitute the third of eqs. 3.2 into the first and integrate between $\rho_0 \equiv \rho(z=0)$ on the l.h.s. and between $\Phi_0 \equiv \Phi(z=0)$ on the other to obtain

$$\rho(z) = \rho_0 \exp\left[-\frac{\Phi(z)-\Phi_0}{c_{\rm s}^2}\right], \tag{3.5}$$

which is a general solution of problems of hydrostatic equilibrium for isothermal gases, where one usually imposes $\Phi_0 = 0$. Then, we introduce the quantity

$$\psi \equiv \frac{(\Phi-\Phi_0)}{c_{\rm s}^2}, \tag{3.6}$$

which allows us to rewrite the Poisson equation as

$$\frac{{\rm d}^2\psi}{{\rm d}z^2} = Ae^{-\psi}, \tag{3.7}$$

where we have defined

$$A \equiv \frac{4\pi G\rho_0}{c_{\rm s}^2}. \tag{3.8}$$

Let us then use of the following identity:

$$\frac{\rm d}{{\rm d}z}\left(\frac{{\rm d}\psi}{{\rm d}z}\right)^2 = 2\frac{{\rm d}\psi}{{\rm d}z}\frac{{\rm d}^2\psi}{{\rm d}z^2} \tag{3.9}$$

to change the l.h.s. of 3.7 and obtain

$${\rm d}\left(\frac{{\rm d}\psi}{{\rm d}z}\right)^2 = 2Ae^{-\psi}{\rm d}\psi, \tag{3.10}$$

which can be easily integrated to give

$$\left(\frac{{\rm d}\psi}{{\rm d}z}\right)^2 = -2Ae^{-\psi} + C. \tag{3.11}$$

The constant C in eq. 3.11 can be readily determined given that $\psi(0) = 0$ and also $({\rm d}\psi/{\rm d}z)_{(z=0)}$ must be null. The second boundary condition comes from the

[1]Recall the Laplacian operator in cylindrical coordinates (eq. C.22).

fact that, at $z = 0$, there cannot be net accelerations as that is the locus where $g_z = -\mathrm{d}\Phi/\mathrm{d}z$ changes sign. Thus $C = 2A$ and eq. 3.11 becomes

$$\frac{\mathrm{d}\psi}{(1 - e^{-\psi})^{1/2}} = \sqrt{2A}\mathrm{d}z, \tag{3.12}$$

which is an integral that can be solved with the substitution $y^2 = e^{-\psi}$ leading to $d\psi = -(2/y)dy$ and

$$-\frac{\mathrm{d}y}{y(1 - y^2)^{1/2}} = \sqrt{\frac{A}{2}}\mathrm{d}z, \tag{3.13}$$

with solution

$$\mathrm{arcsech}\,(y) = \sqrt{\frac{A}{2}}z + C_1. \tag{3.14}$$

At $z = 0$, we have $\Phi = 0$ and thus $y^2 = 1$. Then, given that $\mathrm{arcsech}\,(y)$ is defined only for positive y and $\mathrm{arcsech}\,(1) = 0$, we obtain $C_1 = 0$. We can therefore substitute back $y = (\rho/\rho_0)^{1/2}$ to find the final expression for the density:

$$\rho(z) = \rho_0\,\mathrm{sech}^{\,2}\left(\frac{z}{h}\right), \tag{3.15}$$

where we have defined the **scaleheight** of the gas layer as

$$h \equiv \frac{c_\mathrm{s}}{\sqrt{2\pi G\rho_0}}. \tag{3.16}$$

Equation 3.15 is a generic solution of an infinite slab (or a rotating disc in a potentail with constant v_c) in hydrostatic equilibrium with its own gravity. Can we, however, use it to calculate the scaleheight of the gaseous disc of a spiral galaxy? In the inner regions of galaxy discs, this is not warranted because the potential is dominated by the stellar component (the gas is not self gravitating), at least in nearby galaxies. In high-z galaxies one can argue that there may be a phase when the potential is dominated by the gas while the stellar component is building up. In the outer parts of the nearby discs, instead, the gas could, in principle, become dominant as the stellar density fades out more quickly than the gas density. However, the outer regions are characterized by large amounts of dark matter and it can be shown that, despite this dark matter being distributed in a roughly spherical structure with respect to the gaseous disc, it dominates the vertical force in the outer parts of most galaxies even in the region close to the galactic plane [Oll95]. Thus, to study the scaleheight of galaxies we must investigate hydrostatic equilibrium in the presence of an external potential.

3.1.2. Gas layer in an external potential. The general expression of the Poisson equation is

$$\frac{\mathrm{d}^2\Phi}{\mathrm{d}z^2} = 4\pi G(\rho + \rho_*) \tag{3.17}$$

where ρ_* is the *non-gaseous* density that, in galaxy discs, will be the stellar density in the inner parts and the dark matter density in the outer parts. We assume that $\rho_* \gg \rho$ so that we can neglect the self gravity of the gas. Moreover, let us assume that the gas disc[2] has a scaleheight (h) smaller than that of the stellar disc (h_*).

[2]We consider here the neutral atomic gas as it is the dominant and the most radially extended component (§1.1.1). However, the same calculation is applicable to other gas components.

This is true for most galaxies, for instance in the Solar neighbourhood the H I disc has an observed thickness of $h_{\rm H\,I} \simeq 100 - 200\,{\rm pc}$ while $h_* \simeq 300 - 1000\,{\rm pc}$, where the latter range comprises the scaleheights of thin and thick discs. In this case, we can consider ρ_* almost constant *within the gas layer* and thus eq. 3.17 is readily integrated leading to

$$\Phi(z) = 2\pi G\rho_* z^2 + C_2 z + C_3, \tag{3.18}$$

with $C_2 = C_3 = 0$ for the aforementioned boundary conditions. Finally, substituting this in eq. 3.5 one obtains

$$\rho(z) = \rho_0\ \exp\left(-\frac{z^2}{2h^2}\right), \tag{3.19}$$

where we have defined

$$h \equiv \frac{c_{\rm s}}{\sqrt{4\pi G\rho_*}} \tag{3.20}$$

as the scaleheight of the gas layer due to the external potential produced by a constant density ρ_*. At the solar circle ($R = R_\odot$) the stellar density is $\rho_* \simeq 0.07\,{\rm M_\odot\,pc^{-3}}$ and, if we take $c_{\rm s} = 8\,{\rm km\,s^{-1}}$ for the WNM (also similar to the turbulent velocity dispersion; chapter 8), we obtain $h(R_\odot) \approx 130\,{\rm pc}$, in reasonably good agreement with observations.

In the outer parts of the discs, we must substitute the stellar with the dark matter density in eq. 3.16. The dark matter density can be estimated assuming that it is distributed in a sphere and it contributes to most of the rotation velocity at those radii. If the Galactic potential is not too dissimilar from that of an isothermal sphere[3], the mass within the spherical radius r will be

$$M(r) = 4\pi\rho_0 r_0^2 r, \tag{3.22}$$

with ρ_0 dark matter density at the radius r_0. Due to the second Newton theorem, the circular speed of a spherical system at a certain radius is, in general,

$$v_{\rm c}^2(r) = \frac{G\,M(r)}{r}, \tag{3.23}$$

which becomes a constant if $M(r)$ is given by eq. 3.22. This allows us to derive the density of an isothermal dark matter halo as

$$\rho_{\rm iso}(r) \simeq 0.00118\left(\frac{v_{\rm c}}{200\,{\rm km\,s^{-1}}}\right)^2\left(\frac{r}{25\,{\rm kpc}}\right)^{-2}\,{\rm M_\odot\,pc^{-3}}. \tag{3.24}$$

If we substitute this value to ρ_* in eq. 3.20, the scaleheight of the outer disc turns out to be $h(25\,{\rm kpc}) \simeq 1\,{\rm kpc}$, increasing linearly with radius for constant $c_{\rm s}$. The observed scaleheight of the outer H I layer in the Milky is indeed $\approx 1\,{\rm kpc}$.

---

[3]We recall [CFN19] that power-law density profiles have the form

$$\rho(r) = \rho_0\left(\frac{r}{r_0}\right)^{\alpha}, \tag{3.21}$$

where ρ_0 is the density at the radius r_0. For an isothermal sphere, $\alpha = 2$ and the profile leads to a circular speed constant with radius, as we also see in eq. 3.23.

3.2. Self-gravitating spheres

We consider the hydrostatic equilibrium of spheres of gas where the internal pressure balances the *self gravity* of the gas. This condition applies, as we see, to gas clouds and to the interior of stars as well.

3.2.1. Polytropes. Consider a non-rotating sphere of gas where gravity is balanced by internal pressure alone (absence of rotation). The hydrostatic equilibrium of the sphere is described by the equation

$$\frac{1}{\rho}\frac{\mathrm{d}P}{\mathrm{d}r} = -\frac{\mathrm{d}\Phi}{\mathrm{d}r}, \tag{3.25}$$

which is similar to the first of eqs. 3.2 written now with the spherical radius r.

Since $\rho > 0$, P always increases as Φ decreases. This is the definition of a monotonic function and we can write $P = P(\Phi)$. If this is the case, then the following is also true:

$$\frac{\mathrm{d}P}{\mathrm{d}r} = \frac{\mathrm{d}P}{\mathrm{d}\Phi}\frac{\mathrm{d}\Phi}{\mathrm{d}r}, \tag{3.26}$$

which implies that $\rho = -\mathrm{d}P/\mathrm{d}\Phi$ and therefore also the density can be written as $\rho = \rho(\Phi)$. If both density and pressure can be written as functions of the potential, then pressure can also be written as a function of density alone. We have thus found that non-rotating spheres of gas in hydrostatic equilibrium must have a barotropic equation of state (§2.6).

In stellar structure theory, the usual assumption that one makes is that the relation between pressure and density can be parameterised as a power law called *polytrope* and written as

$$P = K\rho^{1+1/n}, \tag{3.27}$$

where n is the polytropic index. We now substitute eq. 3.27 into eq. 3.25. Integrating from a generic ρ to $\rho = 0$ (i.e. the surface of the star) we obtain an expression for the density as

$$\rho = \left[\frac{\Phi_{\mathrm{out}} - \Phi}{(n+1)K}\right]^n, \tag{3.28}$$

where Φ_{out} is the gravitational potential at the surface. Eq. 3.28 is valid for any density including the central density of the star ρ_{c}, which corresponds to a central potential Φ_{c}. This allows us to write the density as

$$\rho = \rho_{\mathrm{c}}\left(\frac{\Phi_{\mathrm{out}} - \Phi}{\Phi_{\mathrm{out}} - \Phi_{\mathrm{c}}}\right)^n = \rho_{\mathrm{c}}\theta^n, \tag{3.29}$$

where we have defined the dimensionless variable

$$\theta \equiv \left(\frac{\Phi_{\mathrm{out}} - \Phi}{\Phi_{\mathrm{out}} - \Phi_{\mathrm{c}}}\right). \tag{3.30}$$

Next we use the Poisson equation (eq. 2.17 with $\rho_{\mathrm{ext}} = 0$) and substitute the expression for the density found above on the r.h.s. We set another dimensionless variable

$$\xi \equiv \left(\frac{4\pi G\rho_{\mathrm{c}}}{\Phi_{\mathrm{out}} - \Phi_{\mathrm{c}}}\right)^{1/2} r \tag{3.31}$$

and write the Laplacian in spherical coordinates (eq. C.23). Remembering that we only have dependencies on r, after some passages with obtain

$$\frac{1}{\xi^2}\frac{\mathrm{d}}{\mathrm{d}\xi}\left(\xi^2\frac{\mathrm{d}\theta}{\mathrm{d}\xi}\right) = -\theta^n, \tag{3.32}$$

which is the **Lane-Emden equation** of polytropic index n. This equation is integrated with boundary condition $\theta(\xi = 0) = 0$ (obvious from eq. 3.30) and $(\mathrm{d}\theta/\mathrm{d}\xi)_{(\xi=0)} = 0$ to avoid divergence of the density in the centre (see explanation in §3.2.2). Analytical solutions of the Lane-Emden equation exist for $n = 0$, $n = 1$ and $n = 5$, the other indexes require numerical integrations. In the following we consider the particular case of an isothermal sphere that we can see as the case for $n \to \infty$.

3.2.2. Isothermal sphere. Consider a sphere of gas whose temperature does not change with space and time. This assumption is a good representation of dense *molecular* clouds as we see below, where the temperature of the gas is always of the order of 10 K due to the high cooling rate of molecules. In such a situation, $P = c_\mathrm{s}^2\rho$, with c_s constant and thus eq. 3.25 can be integrated to obtain

$$\rho(r) = \rho_\mathrm{c}\exp\left[-\frac{\Phi(r) - \Phi_\mathrm{c}}{c_\mathrm{s}^2}\right], \tag{3.33}$$

where ρ_c is again the central density (compare with eq. 3.5).

We introduce the following two dimensionless quantities:

$$\psi = \frac{\Phi - \Phi_\mathrm{c}}{c_\mathrm{s}^2} \tag{3.34}$$

$$\xi = \left(\frac{4\pi G\rho_\mathrm{c}}{c_\mathrm{s}^2}\right)^{\frac{1}{2}} r \tag{3.35}$$

that allow us to rewrite the Poisson's equation with dimensional quantities as

$$\frac{1}{\xi^2}\frac{\mathrm{d}}{\mathrm{d}\xi}\left(\xi^2\frac{\mathrm{d}\psi}{\mathrm{d}\xi}\right) = e^{-\psi} \tag{3.36}$$

which is the isothermal Lane-Emden equation (compare with eq. 3.32).

As above, to solve eq. 3.32 we impose boundary conditions 1) $\psi(0) = 0$ and 2) $\psi'(0) = 0$. The second condition defines the presence of a core in the inner density distribution. We can see this if we consider the dependencies for $\xi \to 0$:

$$\frac{\mathrm{d}\psi}{\mathrm{d}\xi} \propto \frac{M(\xi)}{\xi^2} \sim \frac{\rho_\mathrm{c}\xi^3}{\xi^2} \to 0 \tag{3.37}$$

where $M(\xi)$ is the mass contained within the radius ξ and the second equality is obviously valid for cored density distributions: ρ_c constant at small ξ. Unfortunately, by imposing this condition we obtain an equation that cannot be solved analytically. We can, however, look for asymptotic solutions for large and small ξ. At large ξ, neglecting the inner boundary conditions, one finds the classical solution for a singular isothermal sphere that can be written as

$$\psi_\mathrm{out}(\xi) = \ln\left(\frac{\xi^2}{2}\right), \tag{3.38}$$

which corresponds to

$$\rho_\mathrm{out}(\xi) = \rho_\mathrm{c}e^{-\psi} = 2\rho_\mathrm{c}\xi^{-2} \tag{3.39}$$

and obviously does not fulfill the requirement of an inner core.

We search for a solution at small radii by rewriting the above equations for small ξ. Due to the first boundary condition, the r.h.s. of eq. 3.32 becomes

$$e^{-\psi} \xrightarrow[\xi\to 0]{} 1 - \psi + O(\psi^2), \tag{3.40}$$

where $O(\psi^2)$ is of second order in ψ and will be neglected. We can make a further change of variables and introduce

$$\chi \equiv \xi\psi. \tag{3.41}$$

We then substitute ψ in eq. 3.32 with this variable to find an analogous equation but in χ that reads

$$\frac{\mathrm{d}^2\chi}{\mathrm{d}\xi^2} + \chi = \xi, \tag{3.42}$$

which can be solved by imposing similar boundary conditions: $\chi(0) = 0$ and $\chi'(0) = 0$, this latter trivially verifiable by taking the derivative of χ by ξ.

Equation 3.42 can be solved by considering its complementary equation and then the particular integral leading to a general solution of this form

$$\chi = a\cos\xi + b\sin\xi + \xi, \tag{3.43}$$

which, once substituted in eq. 3.42 and using the above boundary conditions, returns the following values for the constants: $a = 0$ and $b = -1$. Thus, we have found a solution of eq. 3.32:

$$\psi = 1 - \frac{\sin\xi}{\xi}, \tag{3.44}$$

which, by construction, is only valid for small ξ. Finally, recalling the relation between ρ and ψ (eqs. 3.33 and 3.34) and eq. 3.40 we obtain

$$\frac{\rho}{\rho_\mathrm{c}} = \frac{\sin\xi}{\xi} \simeq 1 - \frac{\xi^2}{6}, \tag{3.45}$$

where we have used the Taylor expansion of $\sin\xi$ truncated at the second order. Equation 3.45 represents the behaviour at low radii of the solution of the isothermal Lane-Emden equation. When combined with the solution at large radii (singular isothermal sphere) we obtain the curve shown in Fig. 1 for an isothermal cloud in hydrostatic equilibrium with its own gravity.

3.2.3. Bonnor-Ebert mass. In general, the mass of a spherical cloud within a certain radius r can be estimated as

$$M(r) = 4\pi \int_0^r r'^2 \rho(r')\mathrm{d}r'. \tag{3.46}$$

If we consider the density distribution in Fig. 1, for large r, the density goes as r^{-2} and this implies that the mass goes to infinity for $r \to \infty$. In realistic situations, however, there is no physical reason to extend this integration to infinity. A molecular cloud, for instance, is not in isolation in the Universe, instead it is embedded in an external (high density) medium and, as essentially any structure in the ISM, has to be in nearly pressure equilibrium[4] with this medium. We can use this condition

[4]Note that this condition is only valid *at the surface* of the cloud, in the inner parts the pressure increases to much higher values to contrast the self gravity.

to define a *truncation* radius $r_{\rm t}$ as the radius at which the pressure of the cloud equals the external pressure, i.e. $P(r_{\rm t}) = c_{\rm s}^2 \rho_{\rm t} = P_{\rm ext}$.

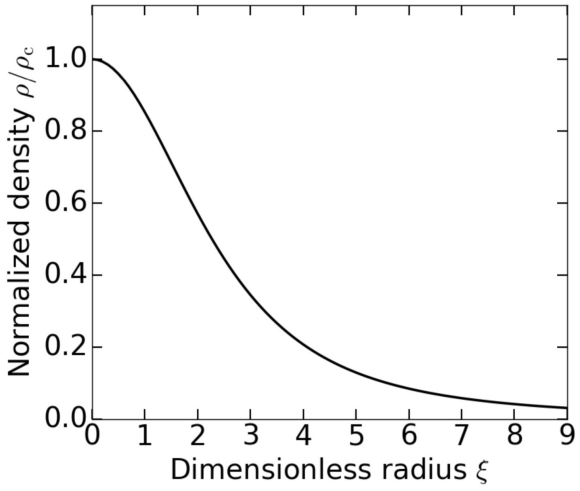

FIGURE 1. The solid line shows the solution of the Lane-Emden equation for the density distribution of an isothermal sphere in hydrostatic equilibrium with its own gravity. From [CFN19].

Using our dimensionless variable ξ (eq. 3.34) we can then write the mass of the cloud as

$$M = 4\pi \left(\frac{c_{\rm s}^2}{4\pi G \rho_{\rm c}} \right)^{3/2} \int_0^{\xi_{\rm t}} \xi^2 \rho(\xi) {\rm d}\xi, \tag{3.47}$$

where the function within the integral can be rewritten as a derivative using the isothermal Lane-Emden equation (eq. 3.36) and recalling that $e^{-\psi} = \rho(\xi)/\rho_{\rm c}$. Integrating in ξ we obtain

$$M = \frac{1}{\sqrt{4\pi\rho_{\rm c}}} \frac{c_{\rm s}^3}{G^{3/2}} \left(\xi^2 \frac{{\rm d}\psi}{{\rm d}\xi} \right)_{\xi=\xi_{\rm t}}, \tag{3.48}$$

where the derivative is evaluated at the edge of the cloud. We multiply eq. 3.48 above and below by $\rho_{\rm t}^{1/2}$ and define a dimensionless mass as

$$m \equiv \left(\frac{P_{\rm ext}^{1/2} G^{3/2}}{c_{\rm s}^4} \right) M, \tag{3.49}$$

where we have used $c_{\rm s}^2 \rho_{\rm t} = P_{\rm ext}$. The final expression for the dimensionless mass of an isothermal sphere[5] becomes

$$m = \frac{1}{2\sqrt{\pi}} \left(\frac{\rho_{\rm c}}{\rho_{\rm t}} \right)^{-1/2} \left(\xi^2 \frac{{\rm d}\psi}{{\rm d}\xi} \right)_{\xi=\xi_{\rm t}}. \tag{3.50}$$

Equation 3.50 shows that the mass of these isothermal clouds is a function of the ratio between central and truncation densities and of the truncation radius $\xi_{\rm t}$

[5]These are also called **Bonnor-Ebert spheres**, i.e. isothermal spheres in hydrostatic equilibrium.

(the term in the second parenthesis is a function of ξ_t). However, given that the density of the cloud is a monotonic function of ξ, ξ_t corresponds to the density contrast $\rho_t/\rho_c = \rho(\xi_t)/\rho_c$. Thus, the term in the second parentesis of eq. 3.50 can also be expressed in terms of ρ_c/ρ_t. Fig. 2 shows the dimensionless mass of isothermal spheres in hydrostatic equilibrium as a function of the ρ_c/ρ_t ratio. The mass has initially a linear dependence with the log of the density ratio but it then turns over. The peak occurs at the critical value $(\rho_c/\rho_t)_{crit} = 14$ corresponding to $m_{crit} = 1.18$. For larger ratios the mass decreases and then, after a slight increase, reaches the asymptotic value of $\sqrt{2/\pi}$.

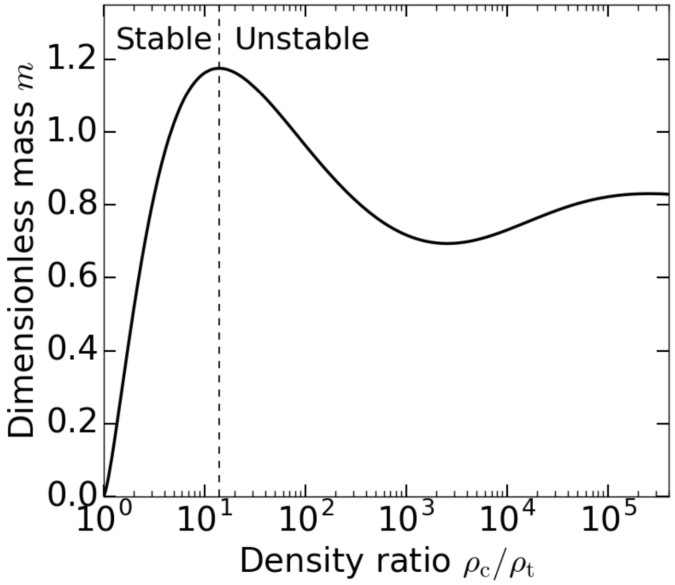

FIGURE 2. Dimensionless mass of isothermal clouds in hydrostatic equilibrium as a function of the ratio between their central and truncation density (ρ_c/ρ_t). From [CFN19].

The mass corresponding to the maximum in Fig. 2 is called **Bonnor-Ebert mass** and it can be written as

$$M_{BE} = \frac{1.18 c_s^3}{\rho_t^{1/2} G^{3/2}}. \tag{3.51}$$

This represents the maximum mass that an isothermal cloud can have to be in hydrostatic equilibrium and stable. Larger masses cannot be in equilibrium as they do not appear in Fig. 2, whereas it can be shown that masses with density contrasts larger than 14 (right side of the plot, beyond the peak) despite being in equilibrium are, however, *unstable* (see [SP05]). Thus, if subject to gravitational perturbations they tend to collapse under the pull of their own gravity.

We can calculate the typical Bonnor-Ebert mass in a molecular cloud in the Milky Way. We consider that the cloud has an outer density $n_t = \rho_t/(\mu_{MG} m_p) \approx 10^2\,\mathrm{cm}^{-3}$ (where $\mu_{MG} \simeq 2.3$ is the typical molecular mass in a Galactic molecular cloud), which corresponds to putting a $T = 10\,\mathrm{K}$ cloud in pressure equilibrium with

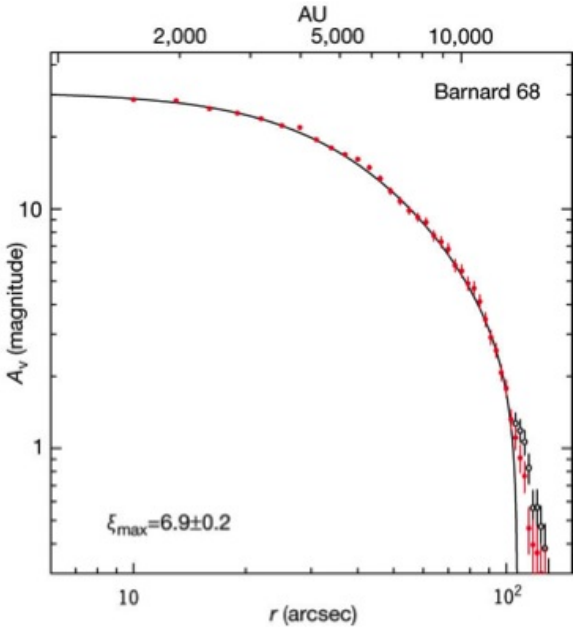

FIGURE 3. Extinction in V-band as a function of radius from the centre of the dense core Barnard 68. Given that extinction is produced by dust and it is function of the dust density, this profile essentially corresponds to a gas density profile (assuming a gas-to-dust ratio constant with radius). The solid line represents the fit with a Bonnor-Ebert sphere with $\xi_{\rm t} = 6.9$, which corresponds to a state at the verge of gravitational collapse [ALL01].

a medium at $P_{\rm ISM}/k_{\rm B} = 10^3\,{\rm K\,cm^{-3}}$ (average pressure of the Galactic ISM). The Bonnor-Ebert mass so obtained becomes

$$M_{\rm BE} \simeq 11.9 \left(\frac{c_{\rm s}}{0.19\,{\rm km\,s^{-1}}}\right)^{3/2} \left(\frac{n_{\rm t}}{10^2\,{\rm cm^{-3}}}\right)^{-1/2} {\rm M}_\odot, \tag{3.52}$$

where we used the isothermal sound speed relative to a temperature $T = 10\,{\rm K}$. This mass is of the same order of magnitude of the very compact **dense cores** that are seen in the densest regions of large molecular clouds. These are the regions where new stars are continuously forming. Figure 3 shows the fit of a Bonnor-Ebert sphere on the well-known dense core Barnard 68. The fit returns a dimensionless truncation radius of $\xi_{\rm t} = 6.9$, which corresponds to $\rho_{\rm c}/\rho_{\rm t} = 16.5$, a condition slightly beyond the critical ratio. The fact that the fit with the theory is so good means that distribution of density in this cloud is exactly as we should expect for an isothermal sphere in hydrostatic equilibrium and confined by an external pressure. Moreover, its mass is essentially at the critical mass that distinguishes between stability and instability. This is a beautiful confirmation of the Bonnor-Ebert theory and is a indication that this dense core is at the verge of gravitational collapse to form a star.

3.3. Spherical equilibrium in an external potential

Consider now a spherical gaseous system in hydrostatic equilibrium in a gravitational potential *not* dominated by the gas. A typical situation of this kind occurs for the hot gas in elliptical galaxies or in galaxy clusters where the potential is dominated by the dark matter. The hydrostatic equilibrium equation in spherical coordinates can be written (using also the second Newton theorem) as

$$\frac{\mathrm{d}P}{\mathrm{d}r} = -\rho \frac{GM_{\mathrm{tot}}}{r^2}, \tag{3.53}$$

where ρ is the gas density, while $M_{\mathrm{tot}}(r)$ is the total mass (mostly dark matter) contained within a sphere of radius r. We substitute the pressure P using the equation of state of ideal gases (§2.6) and obtain

$$M_{\mathrm{tot}}(r) = -\frac{r k_{\mathrm{B}} T}{G\mu m_{\mathrm{p}}} \left(\frac{\mathrm{d}\ln\rho}{\mathrm{d}\ln r} + \frac{\mathrm{d}\ln T}{\mathrm{d}\ln r} \right) \tag{3.54}$$

where $T = T(r)$ and $\rho = \rho(r)$ are gas temperature and pressure at the radius r.

Equation 3.54 allows to estimate the mass profiles of systems in hydrostatic equilibrium. Considering the examples of elliptical galaxies or galaxy clusters, one can determined the gas density using the X-ray surface brightness. Moreover, if the observations allow it, one can also estimate the temperature profile from the fitting of the X-ray spectra at different distances from the centre of the galaxy (or the cluster). This is all is needed to have the *total* mass profile. In cases in which the temperature profiles is not available one can assume that the hot gas is nearly isothermal (not a bad approximation) and eq. 3.54 simplifies. Clusters of galaxies are estimated (with this and other methods) to have total masses of typically $10^{14} - 10^{15}\,\mathrm{M}_{\odot}$ and dark matter profiles similar to the cosmological expectation [NFW96]. The hot gas has temperatures of few$\times 10^7$ K to 10^8 K and contributes about 10% of the total mass.

CHAPTER 4

Shock waves

In this Chapter, we study the phenomena occurring when a portion of the fluid is accelerated to supersonic speeds. A highly supersonic motion (Mach number $\mathcal{M} \gg 1$) always leads to the formation of a discontinuity in the thermodynamic variables. This discontinuity is called **shock wave** or simply **shock** and it is at the origin of a profound modification of the fluid properties. In thermodynamic terms, a shock wave is an *irreversible change* of the properties of the fluid. The portion of the fluid "perturbed" by the passage of a shock wave ends up having higher temperatures and densities than the original fluid. Its internal energy increases dramatically and, depending on its ability of releasing this energy to the external environment (through radiation), the shock is classified as non-radiative (often called adiabatic) or radiative. We start our investigation by exploring the mechanism at the basis of the shock formation.

4.1. The formation of a shock

Imagine a *non-infinitesimal* perturbation that propagates in a fluid. We can visualise this as a wave that has non-negligible physical size, unlike the infinitesimal waves that we have considered in §2.10. Such a perturbation will have a peak (high density) and a trough (low density) at two different locations in the fluid and these will not necessary move at the same speed. For instance, if the sound speed in the fluid is in the adiabatic regime, we have

$$c_{\rm s} = \sqrt{\gamma \frac{K\rho^\gamma}{\rho}} \propto \rho^{\frac{\gamma-1}{2}} \tag{4.1}$$

and so, in our exampl,e $c_{\rm s}(\text{peak}) > c_{\rm s}(\text{average}) > c_{\rm s}(\text{trough})$ for any adiabatic index γ. Thus, the peak (high density gas) goes faster than the rest and this deforms the morphology of the wave increasing the density gradient between peak and trough eventually leading to a density discontinuity. At that point, the peak cannot proceed further because it is blocked by viscous forces. However, it will keep moving at a speed faster than the trough and in the subsequent evolution the trough will be effectively engulfed by the discontinuity. The end result is the propagation of a shock wave.

4.2. Non-radiative shock

The above example shows how a shock wave can naturally develop when a portion of the fluid (the peak of the non-infinitesimal wave in that case) is moving at supersonic speeds (i.e. faster than the average sound speed). Consider now a configuration in which the shock has already developed and the system is in a

stationary state. Let us assume that the shock propagates in the direction[1] x and let us put ourselves in the frame where the shock is *at rest*. We identify the unperturbed region with subscript $_0$ and the perturbed region with $_1$ (Fig. 1). Each region is characterised by thermodynamic variables that remain constant within the region but may vary across the shock. For our present purposes, the shock is an infinitely thin layer, its thickness is in fact of the order of the mean particle free path ℓ (see §2.1).

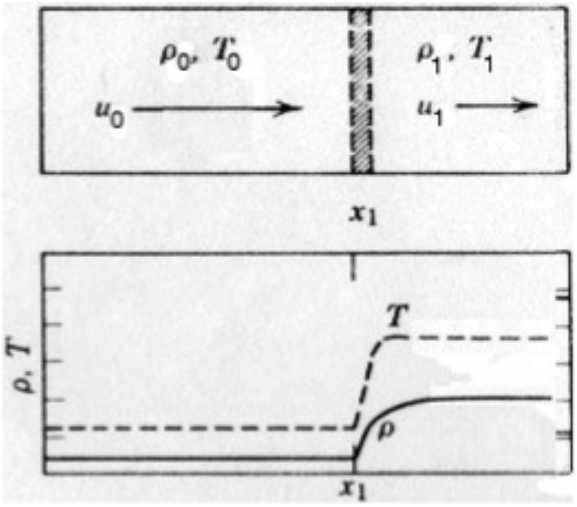

FIGURE 1. *Top:* Sketch of a non-radiative (adiabatic) shock, the unperturbed and shocked adiabatic regions are indicated with subscripts $_0$ and $_1$, respectively. *Bottom:* behaviour of gas temperature (T) and density (ρ) through the shock. Adapted from [Shu92].

4.2.1. Jump conditions. We use the conservation equations derived in §2.5 and §2.7 and write them across the shock. The assumption of stationarity allows us to get rid of all the time derivatives. The continuity equation (2.10) written along the x dimension becomes then

$$\frac{\mathrm{d}}{\mathrm{d}x}(\rho u) = 0, \tag{4.2}$$

where we have used the full derivative given that there is no dependence on time. Equation 4.2 simply states that the quantity ρu does not vary across the shock, it is in fact a spatial constant despite ρ and u may separately vary while crossing the shock. This leads us to the first of the so-called **jump conditions** stating that

$$\rho_0 u_0 = \rho_1 u_1, \tag{4.3}$$

i.e. the product of density and velocity is the same before and after the perturbation has taken place. We remind that we are considering the shock at rest, thus the velocity u is taken, on both sides, with respect to the shock.

[1] A treatment including oblique shocks can be found in [Shu92].

Another jump condition can be obtained in a similar manner from the Euler equation (2.16). We can first rewrite this equation in 1D neglecting the time derivative:

$$\rho u \frac{\mathrm{d}u}{\mathrm{d}x} + \frac{\mathrm{d}P}{\mathrm{d}x} + \rho \frac{\mathrm{d}\phi}{\mathrm{d}x} = 0, \tag{4.4}$$

where the last term on the l.h.s. can be clearly neglected as we do not expect any gravitational potential to change significantly across the shock, being this latter very thin. The first term of eq. 4.4 can be rewritten given the equality

$$\frac{\mathrm{d}(\rho u^2)}{\mathrm{d}x} = \rho u \frac{\mathrm{d}u}{\mathrm{d}x} + u \frac{\mathrm{d}(\rho u)}{\mathrm{d}x}, \tag{4.5}$$

where the second term on the r.h.s. is zero because of eq. 4.2. Hence, the second jump condition reads

$$\rho_0 u_0^2 + P_0 = \rho_1 u_1^2 + P_1. \tag{4.6}$$

Finally, a third condition is derived from the energy equation (2.28). This for an adiabatic fluid ($\dot{Q} = 0$) in 1-D reads

$$\frac{\mathrm{d}}{\mathrm{d}x} \left[(\epsilon + P)\, u \right] = 0, \tag{4.7}$$

where $\epsilon = (1/2)\rho u^2 + U$ (neglecting gravity). We recall that for an ideal gas the internal energy density is proportional to the pressure of the gas (see §D.1.2):

$$U = \frac{P}{\gamma - 1} \tag{4.8}$$

and therefore we can rewrite eq. 4.7 as

$$\frac{\mathrm{d}}{\mathrm{d}x} \left[\left(\frac{1}{2} u^2 + \frac{\gamma}{\gamma - 1} \frac{P}{\rho} \right) \rho u \right] = 0, \tag{4.9}$$

which leads, using again eq. 4.2, to the third jump condition that we write explicitly below together with the previous two:

$$\begin{cases} \rho_0 u_0 & = \rho_1 u_1 \\ \rho_0 u_0^2 + P_0 & = \rho_1 u_1^2 + P_1 \\ \dfrac{1}{2} u_0^2 + \dfrac{\gamma}{\gamma - 1} \dfrac{P_0}{\rho_0} & = \dfrac{1}{2} u_1^2 + \dfrac{\gamma}{\gamma - 1} \dfrac{P_1}{\rho_1} \end{cases} \tag{4.10}$$

The above equations are called **Rankine-Hugoniot jump conditions** and represent, respectively from top to bottom, the conservation of matter, momentum and energy-density fluxes across a shock. The latter equation in particular can be written in two slightly different ways. One reads

$$\frac{1}{2} u_0^2 + \frac{1}{\gamma - 1} c_{\mathrm{s},0}^2 = \frac{1}{2} u_1^2 + \frac{1}{\gamma - 1} c_{\mathrm{s},1}^2, \tag{4.11}$$

explicitly showing the conversion between ordered kinetic energy into random motions as a consequence of the passage of the shock. Indeed, as we see in §4.2.3, in eq. 4.11, $u_0 \gg c_{\mathrm{s},0}$ and $u_2 \ll c_{\mathrm{s},2}$. Moreover, the last of eqs. 4.10 can also be written as

$$\frac{1}{2} u_0^2 + \mathcal{H}_0 = \frac{1}{2} u_1^2 + \mathcal{H}_1, \tag{4.12}$$

where we have introduced the **specific enthalpy** (§D.1.5)

$$\mathcal{H} = \mathcal{U} + \frac{P}{\rho}. \tag{4.13}$$

Equation 4.12 describes the conversion from kinetic energy into enthalpy. In a similar way, the second condition of eqs. 4.10 can also be seen as a conversion of ram pressure into thermal pressure.

4.2.2. Solution of the jump condition equations. The system of equations 4.10 has three unknowns: ρ_1, P_1, T_1 and thus it is solvable. The first equation of the system 4.10 leads to a relation between the density and the velocity ratios between perturbed and unpertubed gas:

$$\frac{\rho_1}{\rho_0} = \frac{u_0}{u_1}. \tag{4.14}$$

We then make use of the definition of Mach number (eq. 2.68) and write it for the unperturbed medium:

$$\mathcal{M}_0^2 = \frac{u_0^2}{c_{\mathrm{s},0}^2} = \frac{\rho_0 u_0^2}{\gamma P_0}, \tag{4.15}$$

where we have used the adiabatic sound speed (§2.10.1). We divide all the terms of the second jump condition by P_0 and obtain

$$\frac{P_1}{P_0} = \gamma M_0^2 + 1 - \gamma M_0^2 \frac{\rho_0}{\rho_1}. \tag{4.16}$$

Next we manipulate the energy jump condition in a similar way (we divide it by P_0/ρ_0) to obtain

$$\frac{1}{2}\gamma M_0^2 + \frac{\gamma}{\gamma - 1} = \frac{1}{2}\gamma M_0^2 \frac{u_1^2}{u_0^2} + \frac{\gamma}{\gamma - 1}\frac{P_1}{P_0}\frac{\rho_0}{\rho_1}. \tag{4.17}$$

We substitute 4.16 into 4.17 and change the ratio of velocities squared using eq. 4.14. The result is a second order equation:

$$(\gamma + 1)M_0^2 x^2 - 2(\gamma M_0^2 + 1)x + (\gamma - 1)M_0^2 + 2 = 0, \tag{4.18}$$

where we have defined x as

$$x \equiv \frac{\rho_0}{\rho_1}. \tag{4.19}$$

The solutions of eq. 4.18 are

$$x_1 = \frac{(\gamma + 1)M_0^2}{(\gamma + 1)M_0^2} = 1 \qquad \text{and} \tag{4.20}$$

$$x_2 = \frac{(\gamma - 1)M_0^2 + 2}{(\gamma + 1)M_0^2}. \tag{4.21}$$

Equation 4.20 is not a shock condition, whereas eq. 4.21 represents the solution for the density jump across the shock for non-radiative shocks.

For large Mach numbers (condition often met in astrophysical shocks), eq. 4.21 leads to the following density jump for non-radiative shocks:

$$\frac{\rho_1}{\rho_0} \simeq \frac{\gamma + 1}{\gamma - 1} = 4, \tag{4.22}$$

where the last identity is valid for a mono-atomic ideal fluid ($\gamma = 5/3$). Equation 4.22 shows that the compression of the perturbed medium is always limited to a

factor four (achieved for $\mathcal{M}_0 \to \infty$). Note that this is independent of the shock speed as long as it is much larger than $c_{\rm s,0}$. The reason for this upper limit to the compression lies in the very nature of the non-radiative shock, which is that the perturbed medium reacts adiabatically to the passage of the shock. This condition forces the medium to store the extra heat (it cannot radiate it away) and to increase its internal energy. However, high internal energy density means effectively high pressure, therefore the medium can efficiently react to attempts of compression, opposing further increase of its density[2]. The value of the pressure for the perturbed medium can be calculated from eq. 4.16 and, assuming $M_0^2 \gg 1$, we get

$$P_1 = \frac{3}{4}\rho_0 u_0^2. \tag{4.23}$$

In the next section, we derive the velocity of the perturbed medium but, to better interpret it, it is useful to change the frame of reference.

4.2.3. Velocities in an external frame. To derive the jump conditions (eq. 4.10), we have made the useful assumption that the shock is at rest, this is handy mathematically but clearly impractical. In most astrophysical situations, it is convenient to consider an external frame of reference at rest with the unperturbed fluid with respect to which the shock moves at velocity $v_{\rm s}$. This requires a simple coordinate transformation to external velocities (indicated here with $'$) and defined as

$$\begin{aligned} u_0' &= u_0 + v_{\rm s} \equiv 0 \implies u_0 = -v_{\rm s} \\ u_1' &= u_1 + v_{\rm s} = \frac{1}{4}u_0 + v_{\rm s} = \frac{3}{4}v_{\rm s}, \end{aligned} \tag{4.24}$$

where in the second equality of the second equation we have used the jump conditions coming from eqs. 4.22 (strong shocks) and 4.3. Equations 4.24 show that the perturbed medium is moving behind the shock at a speed which is lower than the shock speed. In other words, it follows the shock but it cannot fully keep up with it, lagging behind as the shock proceeds into the unperturbed fluid. The consequence of this is that the size of the perturbed region will tend to increase with time. This is also an intuitive consequence of the restriction on the compression ($\rho_1 = 4\rho_0$) that we found above.

4.2.4. Properties of the shocked medium. Summarizing the above findings, the relevant quantities in the perturbed medium produced by a non-radiative shock with large Mach number and $\gamma = 5/3$ are the following:

$$\rho_1 = 4\rho_0 \tag{4.25}$$

$$u_1' = \frac{3}{4}v_{\rm s} \tag{4.26}$$

$$P_1 = \frac{3}{4}\rho_0 v_{\rm s}^2 \tag{4.27}$$

$$T_1 = \frac{3}{16}\frac{\mu m_{\rm p}}{k}v_{\rm s}^2, \tag{4.28}$$

where the last equation was obtained from eqs. 4.27 and 4.25 using the equation of ideal gases (eq. 2.18). Equation 4.28 is particularly important as it shows how the

[2]This is true because also the temperature is very high (see below), we discuss what happens if the temperature drops in §4.3.

gas temperature can jump to very high values being it proportional to the square of the shock speed. The large increase in temperature as a consequence of the passage of a shock is at the basis of the formation of the collisionally ionised gas (see 1.1.3).

Let us now calculate the sound speed in the perturbed medium as

$$c_{\mathrm{s},1} = \sqrt{\gamma \frac{P_1}{\rho_1}} = \sqrt{\frac{5}{16} v_{\mathrm{s}}^2} \simeq 0.56 v_{\mathrm{s}}, \tag{4.29}$$

where we have considered an atomic ideal gas ($\gamma = 5/3$). Equation 4.29 shows that, as a consequence of the shock, clearly a dramatic change has happened in the fluid. Its unperturbed sound speed ($c_{\mathrm{s},0}$) was, by definition, much lower than the shock speed v_{s} given that $M_0 \gg 1$. The passage of the shock has instead left behind a fluid in which the sound speed is of the order of the shock speed, much greater than before. In essence, the shock has converted the *ordered* motion of the shock front into a *chaotic* motion of the particles in the perturbed medium. We discuss this further in the next section where we analyze the jump in entropy.

A second important consequence of eq. 4.29 is that the shocked region is subsonic with respect to the shock itself. This can be seen by recalling that the velocity of the perturbed medium with respect to the shock is $|u_1| = 0.25 v_{\mathrm{s}}$, lower than its sound speed. Thus, the perturbed gas is in *causal connection* with the shock. Being the shocked medium perturbed by the shock itself, this latter conclusion should not come as a surprise.

4.2.5. Entropy jump. In general, the specific entropy of an ideal gas can be written (see §D.1.3) as

$$\mathcal{S} = c_{\mathrm{V}} \ln(P \rho^{-\gamma}) + \text{constant}, \tag{4.30}$$

where c_{V} is the specific heat at constant volume. We can use this equation to calculate the entropy in the shocked medium with respect to the original entropy

$$\mathcal{S}_1 - \mathcal{S}_0 = c_{\mathrm{V}} \ln \left[\frac{P_1}{P_0} \left(\frac{\rho_0}{\rho_1} \right)^{\gamma} \right], \tag{4.31}$$

where we have assumed that c_{V} and γ do not vary across the shock. These assumptions are true for ideal gases, also called *thermally perfect* gases[3]. Then, given that the pressure of the perturbed gas can be written $P_1/P_0 = 2\gamma M_0^2/(\gamma+1)$ (eqs. 4.16 with $M_0 \gg 1$ and 4.22) and using eq. 4.22 we obtain the expression for the entropy variation across a strong shock:

$$\mathcal{S}_1 - \mathcal{S}_0 = c_{\mathrm{V}} \ln \left[\frac{2\gamma}{\gamma+1} \left(\frac{\gamma-1}{\gamma+1} \right)^{\gamma} M_0^2 \right] = c_{\mathrm{V}} \ln \left(\frac{5}{4^{8/3}} M_0^2 \right), \tag{4.32}$$

where we have used $\gamma = 5/3$ in the last equality.

The argument of the logarithm in eq. 4.32 is obviously larger than one and this equation establishes that, as a consequence of the passage of the shock, there is an important increase of entropy. In general, the first law of thermodynamics tells us that

$$\mathrm{d}\mathcal{S} \geq \frac{\delta Q}{T}, \tag{4.33}$$

[3]In a *real* gas they may change, for instance γ could change if the medium is initially molecular but molecules are dissociated by the passage of the shock.

the equal sign being valid only for reversible transformations. It is clear that the passage of a shock is a prime example of an *irreversible* transformation. Although region $_0$ and region $_1$ are *individually* adiabatic (see §4.3), the passage from one to the other is *not*, as the entropy of the system increases. Equation 4.31 can be rewritten recalling the definition of an adiabatic fluid (eq. 2.21) as

$$\mathcal{S}_1 - \mathcal{S}_0 = c_{\rm V} \ln\left(\frac{K_1}{K_0}\right) \tag{4.34}$$

where K_0 and K_1 are the two adiabatic constants of the unperturbed and perturbed fluid. Clearly these two constants have to be different from each other because $\mathcal{S}_1 > \mathcal{S}_0$. In other words, the system has passed from one adiabatic to another, through an irreversible transformation that has produced no exchange of heat with the external medium but a net increase of its entropy.

4.3. Radiative shock

In this section, we describe a situation analogous to the previous strong shock, in which, however, the perturbed medium can lose internal energy through radiation. We first investigate under what conditions the shocked medium can be considered adiabatic.

4.3.1. Radiative cooling of the perturbed gas. As we have seen, the internal energy of the gas increases dramatically as a consequence of the passage of a shock. Depending on its state and temperature the shocked medium will tend to radiate part of this excess energy. We can calculate the timescale after which radiation losses can not be neglected by recalling the definition of cooling time (eq. 2.43). The temperature behind the shock is given by eq. 4.28:

$$T_1 = \frac{3}{16}\frac{\mu m_{\rm p}}{k} v_{\rm s}^2 \simeq 1.4\times10^5 \left(\frac{\mu}{0.6}\right)\left(\frac{v_{\rm s}}{100\,{\rm km\,s^{-1}}}\right)^2 {\rm K}, \tag{4.35}$$

using typical values of interstellar shocks.

A shock wave propagating at a speed $v_{\rm s} = 1000\,{\rm km\,s^{-1}}$ will bring the perturbed medium to a temperature exceeding 1×10^7 K. Under these conditions, the medium is largely ionised (nearly all atoms have lost their electrons) and an important contribution to the emission comes from bremsstrahlung. The collisional ionization equilibrium cooling function (see §2.8.3.1) has a value of $\Lambda_{\rm CIE}(10^7\,{\rm K}) \approx$ few $\times$ $10^{-23}\,{\rm erg\,s^{-1}\,cm^3}$ (at Solar metallicity) and the cooling time becomes

$$t_{\rm cool}(v_{\rm s} = 1000\,{\rm km\,s^{-1}}) \approx 10^7 \left(\frac{n_{\rm e}}{\rm cm^{-3}}\right)^{-1} {\rm yr}. \tag{4.36}$$

This is a very long time compared to typical processes of the galactic ISM, longer for instance than the lifetime of a supernova remnant (see Chapter 5). If instead the shock is moving at $v_{\rm s} \approx 100\,{\rm km\,s^{-1}}$ the temperature will be $T_1 \sim 10^5 K$. This is the temperature where the cooling function peaks to a value of about $\Lambda_{\rm CIE}(10^5\,{\rm K}) \approx 1\times10^{-21}\,{\rm erg\,s^{-1}\,cm^3}$ and the cooling time will plunge down to a mere

$$t_{\rm cool}(v_{\rm s} = 100\,{\rm km\,s^{-1}}) \simeq 2\times10^3 \left(\frac{n_{\rm e}}{\rm cm^{-3}}\right)^{-1} {\rm yr}, \tag{4.37}$$

four orders of magnitude lower despite only two orders or magnitude in temperature and one in speed! When the cooling time is so low, radiative cooling cannot be

neglected in most astrophysical phenomena. Thus, the gas perturbed by shocks with speeds of $\sim 100\,\mathrm{km\,s^{-1}}$ will likely have important energy losses.

4.3.2. Cooling length. Let us now estimate how far from the shock the perturbed medium will start to significantly lose energy through radiation. As we have seen, the perturbed medium flows away from the shock at a speed $u_1 = v_\mathrm{s}/4$. As it does, it will start radiating away part of its internal energy, decreasing its temperature and progressively decreasing its cooling time (see §2.8.3.1). In an idealised case, the shocked gas keeps cooling until it reaches the same temperature as the unperturbed medium. We calculate at what distance from the shock the radiative losses will be so significant that the medium can be expected to have temperatures comparable to the pre-shock conditions. This distance is

$$l_\mathrm{cool} = u_1 t_\mathrm{cool} = \frac{1}{4} v_\mathrm{s} t_\mathrm{cool}, \tag{4.38}$$

where t_cool is the cooling time and u_1 is the velocity of the medium with respect to the shock. If we take a shock speed of $v_\mathrm{s} = 100\,\mathrm{km\,s^{-1}}$ and thus the cooling time from eq. 4.37, we obtain

$$l_\mathrm{cool} \simeq 0.05 \left(\frac{n_\mathrm{e}}{\mathrm{cm^{-3}}}\right)^{-1} \mathrm{pc}, \tag{4.39}$$

i.e. a very small **cooling length** compare to most astrophysical scales (it is for instance much smaller than the radius of a SNR, which is typically tens of parsec). The cooling length in eq. 4.39 is, in fact, so small that it can be considered infinitesimal. The consequence of this is that the adiabatic region (subscript $_1$) shrinks to a tiny layer followed by another (also very small) transition region where the gas temperature decreases quickly down to the unperturbed value. This last region (beyond the transition) is therefore *isothermal* (subscript $_2$) with respect to the unperturbed gas (Fig. 2). The system can be therefore described with a new set of jump conditions that we derive in the next section.

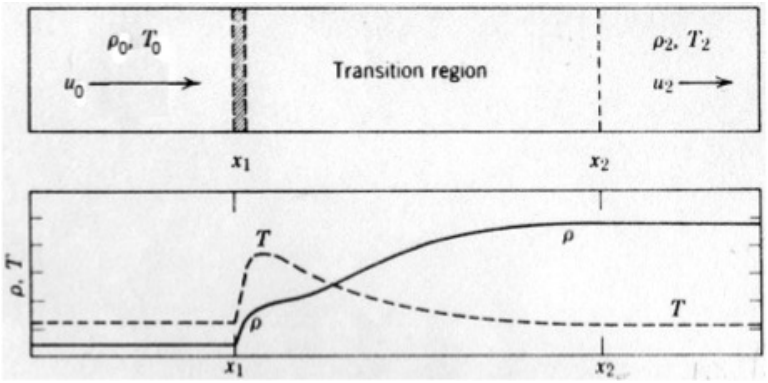

FIGURE 2. Sketch of a radiative shock, the unperturbed and isothermal regions are indicated with subscripts $_0$ and $_2$, respectively. Adapted from [Shu92].

4.3.3. Jump conditions for radiative shocks. The basic assumptions made for the adiabatic shocks are still valid here: the system is stationary, the shock is planar and perpendicular to x, and changes in gravity across the shock can be neglected. Given these assumptions, the first two jump conditions (eq. 4.10) remain unchanged. The last condition however is not valid anymore as the shocked medium is not adiabatic. One could modifying it by taking into account the cooling term on the r.h.s. (see §2.8.5) but this does not lead to easily solvable equations. As mentioned, a common approach is instead to consider that the radiative cooling is so efficient that it brings the shocked medium back to its initial temperature. This type of shock is called **isothermal shock** and it should be seen as a limit case of the radiative shock. The jump conditions are simply

$$\begin{cases} \rho_0 u_0 & = \rho_2 u_2 \\ \rho_0 u_0^2 + P_0 & = \rho_2 u_2^2 + P_2 \\ T_0 & = T_2, \end{cases} \tag{4.40}$$

where we have indicated with the subscript $_2$ the region beyond the cooling length (eq. 4.39).

In analogy to what done for the adiabatic system we can rewrite the above equations to include ratios between the thermodynamic variables and the Mach number, now being

$$M_0^2 = \frac{\rho_0 u_0^2}{P_0}, \tag{4.41}$$

because we are dealing with an isothermal transformation. Thus, diving by P_0 the second of eqs. 4.40 one obtains

$$M_0^2 + 1 = M_0^2 \frac{\rho_0}{\rho_2} + \frac{P_2}{P_0}. \tag{4.42}$$

Then from the last of eqs. 4.40 we have[4]

$$\frac{P_0}{\rho_0} = \frac{P_2}{\rho_2} \implies \frac{P_2}{P_0} = \frac{\rho_2}{\rho_0} \tag{4.43}$$

and setting $x \equiv \rho_2/\rho_0$ we obtain a second-order equation that reads

$$x^2 - (M_0^2 + 1)x + M_0^2 = 0, \tag{4.44}$$

with the only physically meaningful solution

$$\frac{\rho_2}{\rho_0} = M_0^2 \gg 1. \tag{4.45}$$

This solution is completely different from the one obtained for an adiabatic shock where the density in the perturbed medium could not increase by more than a factor four (for mono-atomic ideal gases). The compression in the radiative case (isothermal limit) can, instead, continue almost indefinitely to a very large value. This is essentially due to the fact that the medium is radiating away its internal energy and thus it does not efficiently counteract the density increase.

Fig. 2 shows the behaviour of the density as a function of the position x. Note that the shock always starts as adiabatic (density jump of a factor 4) then as the medium cools the density can increase to much higher values. As mentioned,

[4]Note that we are implicitly assuming that the atomic weight of the gas remains constant, $\mu_2 = \mu_0$.

the transition region can be very small, for instance for shock speeds of about $100\,\mathrm{km\,s^{-1}}$, and its presence is, therefore, neglected in the calculations. The temperature across shock and transition region shows opposite behaviour with respect to density. As a consequence, the pressure of the shocked medium remains at a rather similar value (isobaric transformation) throughout the whole region beyond the shock (cf. eq. 4.23):

$$P_2 = P_0 \frac{\rho_2}{\rho_0} = P_0 M_0^2 = \rho_0 u_0^2, \tag{4.46}$$

which is, as we said, much larger than the unperturbed pressure.

4.3.4. Other physical quantities. Moving to an external frame of reference, in which the unperturbed medium is at rest, we have again (see eq. 4.24) that $u_0 = -v_\mathrm{s}$ and then

$$u_2' = u_2 + v_\mathrm{s} = \frac{1}{M_0^2} u_0 + v_\mathrm{s} \approx v_\mathrm{s}, \tag{4.47}$$

where the last equality is valid because $M_0^2 \gg 1$. Thus the shocked medium follows the shock at nearly the same speed. This is an important result: as we saw in §4.2 the non-radiative shocked medium can not keep up with the shock and the region beyond the shock tends to expand. Now, instead, the shocked gas can pile up right behind the shock because its density can increase to very high values. Therefore, we can expect the perturbed gas to be mostly located in a *thin* post-shock region.

Summarizing, the relevant quantities in the perturbed medium produced by a radiative (isothermal) shock with large Mach number are the following:

$$\rho_2 = M_0^2 \rho_0 \tag{4.48}$$

$$u_2' = v_\mathrm{s} \tag{4.49}$$

$$P_2 = \rho_0 v_\mathrm{s}^2, \tag{4.50}$$

while the temperature is unchanged by construction.

The last two quantities that we discuss are the sound speed and the Mach number. The sound speed, again by construction, is unchanged and so $c_{\mathrm{s},2} = c_{\mathrm{s},0}$. Nevertheless, it is interesting to estimate the Mach number of the shocked medium, which is

$$M_2^2 = \frac{u_2}{c_{\mathrm{s},2}} = \frac{u_0}{c_{\mathrm{s},0}\, M_0^2} = \frac{1}{M_0} \ll 1. \tag{4.51}$$

Thus the shocked medium is highly subsonic, in particular it is causally connected to the shock as its speed with respect to this latter is much smaller than its sound speed.

CHAPTER 5

Evolution of supernova remnants

In this Chapter, we describe the phenomena that occur at the end of stellar evolution when massive stars explode as supernovae (SNe) releasing a large amount of energy in the surrounding ISM. The focus is on the hydrodynamic effects that these explosions have on the surrounding ISM. In this sense, we are focusing here on supernova remnants (SNR) produced by type II (core collapse) SNe, which typically explode in gas-rich regions. However, our treatment also applies to type Ia SNe that happen to go off within a galactic disk. These explosions typically release an energy of $E_{\rm SN} \approx 10^{51}\,{\rm erg}$ in the form of fast ejecta in a very short time. We follow the various phases of the evolution of their remnants making use of most of the theoretical machinery that we have developed in the previous chapters. One of our main goals is to estimate the amount of energy that gets converted into radiation during the expansion of the SNR as opposed to the energy that will be released in the ISM as kinetic. We start from the initial phase of free expansion.

5.1. Free expansion

The mass ejected in a typical explosion is of the order of a few Solar masses (up to tens of $\rm M_\odot$ for very massive stars). The typical ejection velocity is $\sim 10^4\,{\rm km\,s^{-1}}$. At the beginning, the velocity is so high that the surrounding ISM cannot react to the expansion of the supernova ejecta. This phase is characterised by free expansion and the velocity remains roughly constant. The reaction of the ISM becomes non-negligible when the engulfed ISM mass is comparable to the mass of the ejecta ($M_{\rm ej}$). We can write this condition as

$$\frac{4\pi}{3} r_{\rm F}^3 n_0 \mu m_{\rm p} \sim M_{\rm ej}, \tag{5.1}$$

where we have indicated with $r_{\rm F} = r(t_{\rm F})$ the radius of the ejecta at the time ($t_{\rm F}$) when the free phase ends and with n_0 and μ, respectively, the density and the atomic/molecular weight of the unperturbed ISM. Substituting for typical numbers of SNRs we find

$$r_{\rm F} \simeq 3.6 \left(\frac{M_{\rm ej}}{5\,{\rm M_\odot}}\right)^{1/3} \left(\frac{\mu n_0}{1\,{\rm cm^{-3}}}\right)^{-1/3} {\rm pc}. \tag{5.2}$$

The time at which this radius is reached is

$$t_{\rm F} \simeq 375 \left(\frac{M_{\rm ej}}{5\,{\rm M_\odot}}\right)^{1/3} \left(\frac{\mu n_0}{1\,{\rm cm^{-3}}}\right)^{-1/3} \left(\frac{v_{\rm ej}}{10^4\,{\rm km\,s^{-1}}}\right)^{-1} {\rm yr}, \tag{5.3}$$

showing that this is a very short phase. All SNRs associated with explosions directly observed in the Milky Way should have overcome this phase or being about to overcome it.[1]

5.2. Sedov phase

The interaction between fast ejecta and the ISM produces the formation of a shock wave. In practice, the highly supersonic speeds of the ejecta are slowed down by the surrounding medium, while the shocked gas is heated up to very high temperatures. This gas will lie behind the shock surface (we are implicitly assuming a spherical geometry for the explosion) forming a **shell** (see later). We can use the theory of shock waves developed in Chapter 4 to investigate the properties of this shell as a function of time. Here we show that the shell is initially nearly adiabatic and this has important consequences for the evolution of the supernova remnant. We start by calculating the temperature of the shocked gas using eq. 4.35:

$$T_1 \simeq 1.4\times10^5\left(\frac{\mu}{0.6}\right)\left(\frac{v_\mathrm{s}}{100\,\mathrm{km\,s^{-1}}}\right)^2 \sim 10^9\,\mathrm{K}, \tag{5.4}$$

where we have simply substituted the velocity of the ejecta. At these very high temperatures, the cooling function is totally dominated by the bremsstrahlung (all atoms are fully ionised) and we can take it as $\Lambda_\mathrm{CIE} \approx \text{few}\times10^{-22}\,\mathrm{erg\,s^{-1}\,cm^3}$, leading to a cooling time of $t_\mathrm{cool}(T=10^9\,\mathrm{K}) \approx 10^8(n_\mathrm{e}/1\,\mathrm{cm^{-3}})^{-1}$ yr. This very long time reassures us that the system can be considered, at least initially, *adiabatic* and thus energy losses via radiation can be safely neglected at this stage.

5.2.1. Energy equipartition. An adiabatic evolution implies that the energy released by the blast must go either into kinetic or internal energy of the shocked medium[2], i.e. it is not radiated away. We can estimate these two energies *per particle* in the shocked medium as follows. The internal energy per particle is

$$U_\mathrm{p} = \frac{3}{2}k_\mathrm{B}T_1 = \frac{9}{32}\mu m_\mathrm{p} v_\mathrm{s}^2, \tag{5.5}$$

where we have substituted eq. 4.35 for the temperature behind the shock. The kinetic energy per particle can also be calculated using the theory of adiabatic shocks as

$$K_\mathrm{p} = \frac{1}{2}\mu m_\mathrm{p} u_1'^2 = \frac{9}{32}\mu m_\mathrm{p} v_\mathrm{s}^2, \tag{5.6}$$

where the last equality comes from the fact that the shocked medium is moving (in an external frame of reference) at $(3/4)v_\mathrm{s}$ (see §4.2.3). We can conclude that there is an *equipartition* between internal and kinetic energies in the region behind the shock and we can use this result to write our equation of energy conservation.

Before doing this, we estimate the size of the shocked region and whether or not we can consider it as a relatively thin shell trailing behind the shock front. Let

[1]The last observed supernova explosion is indeed SN 1604, also called Kepler star. The Cassiopeia A SNR is thought to be generated by an explosion occurred around the end of the seventeen century but there is no historical record of the supernova. Other SNe may have exploded in the Galaxy more recently. For instance, supernova SNR G1.9+0.3 likely exploded in the second half of the nineteen century and has been discovered only now (as a SNR) with the Chandra X-ray space telescope.

[2]Note that in the (previous) free phase most of the energy was in the form of kinetic energy of the ejecta.

us call r_s the radius of the sphere out to the shock front and r_{in} the inner radius of the shocked material (Fig. 1; we assume that the shock itself is infinitely thin). We consider that all the ISM originally contained in a sphere $(4\pi r_s^3)/3$ has been incorporated by the SNR and thus we obtain the mass of the engulfed material multiplying this volume by the density of the unperturbed ISM ρ_0. In an adiabatic strong shock the density increases behind the shock by a factor four (for $\gamma = 5/3$). If this material lies in a shell we can then write its mass as

$$\frac{4\pi}{3}(r_s^3 - r_{in}^3)\rho_1 = \frac{16\pi}{3}(r_s^3 - r_{in}^3)\rho_0 = \frac{4\pi}{3}r_s^3\rho_0, \tag{5.7}$$

where ρ_1 is the density of the shocked medium and the last equality comes from the fact that all the engulfed mass must be in the shell. The above equation leads to a ratio between the two radii of

$$r_{in} = \left(\frac{3}{4}\right)^{1/3} r_s \simeq 0.91 r_s, \tag{5.8}$$

thus the shell can be considered thin enough to allow us to neglect its size with respect to the size of the SNR. In the following, we use only one radius to describe the evolution of the SNR: r_s describes both the location of the shock front and of the shell of shocked ISM.

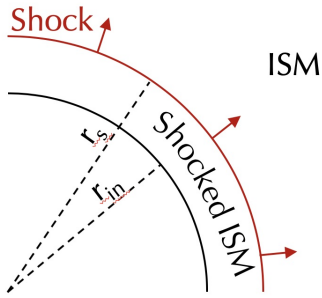

FIGURE 1. Simple sketch of the evolution of a SNR shock through the ISM, the sizes of the various regions are not to scale.

5.2.2. Evolution of the shell. In the light of the above findings, we can now write the equation of energy conservation in the shell behind the shock. Due to equipartition, the total energy will simply be twice the kinetic energy of the shell. Given that the shell evolves adiabatically, the total energy must be equal, in this phase, to the total energy release from the SN explosion (E_{SN}):

$$2 \cdot \frac{4\pi}{3} r_s^3 n_0 \cdot \frac{9}{32}\mu m_p v_s^2 = E_{SN}. \tag{5.9}$$

We can rewrite v_s as the temporal derivative of r_s obtaining the equation

$$r_s^{3/2}\frac{dr_s}{dt} = A, \tag{5.10}$$

where the constant is

$$A = \left(\frac{4E_{SN}}{3\pi\mu n_0 m_p}\right)^{1/2}. \tag{5.11}$$

Equation 5.10 can be easily integrated from the radius $r_{\rm F}$ and time $t_{\rm F}$ (end of the free phase) to generic radii and times. However, given that $t_{\rm F}$ is very small we can derive a simpler formula that applies to times much larger than $t_{\rm F}$ by setting the lower boundaries of these integrals to zero. With this prescription we obtain the following equation for the radius of the shell

$$r_{\rm s} \simeq 0.39 \left(\frac{E_{\rm SN}}{10^{51}\,{\rm erg}}\right)^{1/5} \left(\frac{\mu}{0.6}\right)^{-1/5} \left(\frac{n_0}{1\,{\rm cm}^{-3}}\right)^{-1/5} \left(\frac{t}{\rm yr}\right)^{2/5} {\rm pc}, \tag{5.12}$$

thus a power law in time with exponent 2/5. Finally, taking the time derivative of eq. 5.12 we obtain an equation for the velocity of the shell

$$v_{\rm s} \simeq 1.5 \times 10^5 \left(\frac{E_{\rm SN}}{10^{51}\,{\rm erg}}\right)^{1/5} \left(\frac{\mu}{0.6}\right)^{-1/5} \left(\frac{n_0}{1\,{\rm cm}^{-3}}\right)^{-1/5} \left(\frac{t}{\rm yr}\right)^{-3/5} {\rm km\,s}^{-1} \tag{5.13}$$

from which we see that the SNR is now decelerating ($\mathrm{d}v_{\rm s}/\mathrm{d}t < 0$) as expected due to its strong interaction with the external medium.

We note that the last two equations show very mild dependencies on important parameters of the model such as the supernova energy and the ISM density. This makes these equations quite robust in predicting the evolution of SNR shells even in the presence of significant differences in the density of the surrounding medium. For instance, a change in density of two orders of magnitude corresponds to a factor ~ 2.5 in radius and velocity.

5.2.3. Interior solution. As mentioned above, eqs. 5.12 and 5.13 refer to the evolution of the shell, but being this *relatively thin*[3], the same equations describe the evolution of the outer shock in the ISM. This is however an approximation as in reality, the interior of a blast wave does not have a well defined *shell* as its thermodynamic variables change substantially with radius. This is described by the so called *Sedov-Taylor* solution found by L. I. Sedov [Sed59] and J. L. Taylor [Tay50]. Sedov found an analytical solution, whereas Taylor perfomed one of the first numerical integration of differential equations in history. The calculation to find this solution are quite involved. Here we outline the general idea of the procedure.

First, one defines a self-similar variable as follows

$$\xi \equiv r \left(\frac{E_{\rm SN}}{\rho_0}\right)^{-1/5} t^{-2/5}. \tag{5.14}$$

With this prescription, the gas flow appears *the same* at every position and time. The main fluid equations can now be re-written in terms of dimensionless density, velocity and pressure

$$\rho'(\xi) = \frac{\rho(x,t)}{\rho_1}, \qquad u'(\xi) = \frac{r_{\rm s}}{r u_1} u(r,t), \qquad P'(\xi) = \frac{r_{\rm s}^2}{r^2 P_1} P(r,t), \tag{5.15}$$

where ρ_1, u_1 and P_1 are respectively the density, velocity and pressure of the shocked medium obtained from the jump conditions (eq. 4.10). At this point, one introduces these variable in the continuity, Euler and energy conservation (adiabatic system) equations, all written in spherical coordinates, and solves the system.

[3]We see later (§5.3) that the shell actually becomes much thinner when radiative losses are taken into account.

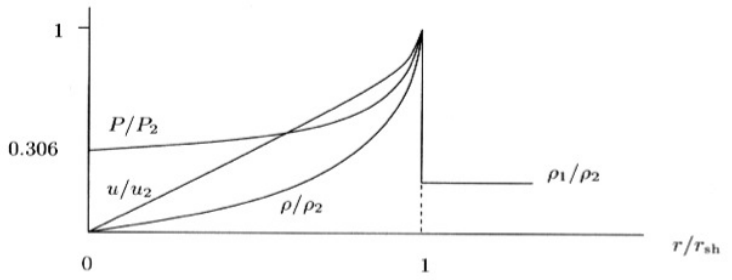

FIGURE 2. Interior of a spherical blast wave evolving adiabatically: Sedov solution. All variables are self similar and normalised, see text. From [Shu92], the subscripts $_1$ and $_2$ are here used for the unperturbed and perturbed media, respectively (they correspond to $_0$ and $_1$ used in these notes).

The final solution is shown in Fig. 2 where we see the trend of the three variables in the interior of the blast wave, obtained with $\gamma = 5/3$. The shock is at $r/r_s = 1$, immediately behind the shock the density shows the factor four jump characteristic of $\gamma = 5/3$ adiabatic shocks. It then sharply decreases reaching very low densities in the inner regions. The spatial region where ρ is relatively high (say more than twice the unperturbed medium) is of the order of 10% of r_s. This is the region that we refer to as *the shell*, although we now see that it is not clearly separated from the rest of the interior. The velocity of the shocked medium is higher close to the shock and then decreases almost linearly towards the centre. Finally, the pressure behaves rather differently: it decreases behind the shock and then remains quite flat in the interior. This trend is quite important because it shows that the internal energy of the interior of the blast wave is rather constant with r and quite high. The constancy of pressure and the decrease of density imply that the temperature must increase to very high values towards the centre. This, in turn, also implies that the sound speed is very high.

As a consequence of the above, the very interior a spherical blast wave is referred to as the *hot* **bubble**. In the case of a SNR the density of this bubble is very low but it still has a dynamical effect, this can be appreciated in the next phase of the evolution when radiative cooling from the shell cannot anymore be neglected.

5.2.4. End of the adiabatic phase. When radiative losses cannot be neglected anymore, the Sedov phase ends and the evolution of the SNR changes in nature. We can estimate the time when this transition takes place by writing the evolution of the temperature of the shell in the Sedov phase. This is the temperature of the perturbed medium for an adiabatic shock (eq. 4.35) where the shock speed is given by eq. 5.13:

$$T_s \simeq 5 \times 10^6 \left(\frac{E_{SN}}{10^{51}\,\mathrm{erg}}\right)^{2/5} \left(\frac{\mu}{0.6}\right)^{3/5} \left(\frac{n_0}{1\,\mathrm{cm}^{-3}}\right)^{-2/5} \left(\frac{t}{10^4\,\mathrm{yr}}\right)^{-6/5} \mathrm{K}. \tag{5.16}$$

The temperature is rapidly decreasing with time and we can take $T \approx 10^6\,\mathrm{K}$ as the characteristic threshold at which radiative losses can no longer be neglected. This is the temperature at which the cooling function starts to rise (§2.8.3.1) and below

which we can expect a runaway cooling (as the temperature decreases the cooling time also decreases causing the temperature to drop even faster). Thus by setting $T_{\rm A} = T(t_{\rm A}) = 10^6\,{\rm K}$ where $t_{\rm A}$ is the time at which the adiabatic phase ends, we obtain

$$t_{\rm A} \approx 4 \times 10^4\,{\rm yr}, \tag{5.17}$$

which is a relatively short time given that the total lifetime that a SNR can reach is $\sim 10^6\,{\rm yr}$, as we see in §5.3.2.

To have a confirmation of the timescale just calculated let us estimate the cooling time for a SNR shell with a temperature of $T = 10^6\,{\rm K}$. This is

$$t_{\rm cool}(10^6\,{\rm K}) = \frac{3k_{\rm B}T}{n_{\rm e,1}\Lambda_{\rm CIE}(T)} \simeq \frac{3k_{\rm B}T}{2n_0\Lambda_{\rm CIE}(T)} \approx 5 \times 10^4 \left(\frac{n_0}{1\,{\rm cm}^{-3}}\right)^{-1}\,{\rm yr}, \tag{5.18}$$

where $n_{\rm e,1}$ is the electron density of the perturbed medium and we have assumed that unperturbed and shocked media share the same atomic weight[4]. Therefore, the estimated age of the adiabatic phase is similar to the cooling time at the end this phase. This is a convincing confirmation that beyond $t_{\rm A}$ radiative losses become important and we cannot neglect them in the calculation.

Using eqs. 5.12 and 5.13 we find that at $t_{\rm A}$ the SNR will have a radius of about $r_{\rm s}(t_{\rm A}) \simeq 27\,{\rm pc}$ and a shock speed $v_{\rm s}(t_{\rm A}) \simeq 267\,{\rm km\,s}^{-1}$. With these timescales in mind we can also rewrite eqs. 5.12 and 5.13 in more convenient forms:

$$r_{\rm s} \simeq 15.7 \left(\frac{E_{\rm SN}}{10^{51}\,{\rm erg}}\right)^{1/5} \left(\frac{\mu}{0.6}\right)^{-1/5} \left(\frac{n_0}{1\,{\rm cm}^{-3}}\right)^{-1/5} \left(\frac{t}{10^4\,{\rm yr}}\right)^{2/5}\,{\rm pc}, \tag{5.19}$$

$$v_{\rm s} \simeq 612 \left(\frac{E_{\rm SN}}{10^{51}\,{\rm erg}}\right)^{1/5} \left(\frac{\mu}{0.6}\right)^{-1/5} \left(\frac{n_0}{1\,{\rm cm}^{-3}}\right)^{-1/5} \left(\frac{t}{10^4\,{\rm yr}}\right)^{-3/5}\,{\rm km\,s}^{-1}. \tag{5.20}$$

5.2.5. Energy losses during the Sedov phase. The definition of the adiabatic phase is that energy is conserved and so there should be no energy loss. However in real circumstances, *some* energy is always lost. Indeed, we do observe SNRs in this phase, which naturally shows that they are emitting radiation thus losing at least a fraction of their internal energy. We can estimate the amount of energy radiated as the integral in time of the energy loss rate

$$E_{\rm rad} = \int_0^{t_{\rm A}} \dot{E}{\rm d}t, \tag{5.21}$$

where the rate of radiation losses can be written as

$$\dot{E}(t) \approx 4\pi r_{\rm s}^2 \Delta r_{\rm s} n_{\rm i,1} n_{\rm e,1} \Lambda_{\rm CIE}[T(t)], \tag{5.22}$$

where $n_{\rm i,1}$ and $n_{\rm e,1}$ are the ion and electron densities in the shocked medium and we can assume them to be $n_{\rm i,1} \simeq n_{\rm e,1} \simeq n_1/2 = 2n_0$ for non-changing atomic weight[5]. $\Lambda_{\rm CIE}$ is a function of T and consequently a function of time but it is always of the order of a few $\times 10^{-23}$ (see §2.8.3.1) in the range of temperatures that we are

[4] In general one has that, for an adiabatic shock, $\rho_1 = 4\rho_0$, thus $n_1\mu_1 m_{\rm p} = 4n_0\mu_0 m_{\rm p}$ and $n_1 = 4(\mu_0/\mu_1)n_0$. In a fully ionised gas, such as the perturbed medium here, we also have $n_1 \simeq 2n_{\rm e,1}$ and so $n_{\rm e,1} \simeq 2(\mu_0/\mu_1)n_0$.

[5] See footnote 4.

considering here ($10^6 \lesssim T < 10^8\,\mathrm{K}$). Δr_s is about $0.1 r_\mathrm{s}$ (eq. 5.12) and thus we can estimate the energy losses to be

$$E_\mathrm{rad}(t_\mathrm{A}) \approx 3 \times 10^{49}\,\mathrm{erg}. \tag{5.23}$$

Equation 5.23 has been obtained using non-changing atomic weights between the unperturbed and shocked medium. For $\mu_1 \simeq 0.5\mu_0$ (ionization of the medium by the shock) $E_\mathrm{rad}(t_\mathrm{A})$ increases by a factor ≈ 3. We can conclude that, in general, no more than about 10% of the initial energy of the blast is radiated during the Sedov phase, i.e. the shell evolves quasi-adiabatically.

5.3. Radiative phase

When radiative losses become important, the evolution of a SNR dramatically changes. The shock quickly tends to the limit of an isothermal shock and this has fundamental consequences both for the structure of the shell and for its evolution with time. Let us first consider its structure. The reader will recall that in §4.3 we saw that the transition from an adiabatic to an isothermal shock is characterised by two dramatic changes in the properties of the shocked material. First, the density can increase up to very large values of the order of the shock Mach number squared and second, the velocity of the shocked material closely approaches the shock speed. In terms of a SNR, these two properties translate into the formation of the so-called **thin shell**, much thinner than before because the gas can pile up (larger density) behind the shock and because this gas now clings to the shock itself ($u_2' \simeq v_\mathrm{s}$, see eq. 4.48). As a consequence, the SNR clearly becomes separated in two zones: a *radiative* and thin shell (roughly at the same temperature as the unperturbed ISM) and a hot bubble that remains adiabatic.

The energy equation cannot be used anymore unless we include the radiative losses but this would make it rather complicated to treat analytically. We should therefore resort to use the momentum equation for the shell. The momentum of the shell varies essentially because of the pressure forces acting on it:

$$\frac{\mathrm{d}}{\mathrm{d}t}(M_\mathrm{s} v_\mathrm{s}) = 4\pi r_\mathrm{s}^2 (P_\mathrm{b} - P_0) \simeq 4\pi r_\mathrm{s}^2 P_\mathrm{b}, \tag{5.24}$$

where M_s is the total mass of the shell, P_b is the pressure of the bubble and P_0 is the pressure of the external ISM that we can safely neglect. Given the above considerations about the structure of an isothermal shell we can assume that all the material engulfed by the shock remains in the shell. This gives us a handy way to write the total mass of the shell as the whole ISM engulfed by the shock, leading to

$$\mu n_0 m_\mathrm{p} \frac{\mathrm{d}}{\mathrm{d}t}(r_\mathrm{s}^3 v_\mathrm{s}) = 3 r_\mathrm{s}^2 P_\mathrm{b}. \tag{5.25}$$

Next, we need to find an expression for P_b.

The bubble is adiabatic (it still has a very high temperature) and from the Sedov interior solution we also know that it is roughly isobaric (the pressure is constant with radius). This allows us to write its pressure as

$$P_\mathrm{b} = K \rho_\mathrm{b}^{5/3} \approx K \left(\frac{3 M_\mathrm{b}}{4\pi r_\mathrm{s}^3} \right)^{5/3}, \tag{5.26}$$

where ρ_b and M_b are the average density and the total mass of the bubble respectively and we have assumed that the radius of the bubble is the same of the radius

of the shell r_s (clearly a good approximation for this thin-shell phase). Given that the gas perturbed by the shock remains, in this phase, in the shell, we can assume that the mass of the bubble does not increase after the Sedov phase. This implies that we can write

$$r_\mathrm{s}^3\rho_\mathrm{b} = r_\mathrm{s,A}^3\rho_\mathrm{b,A} \implies \frac{\rho_\mathrm{b}}{\rho_\mathrm{b,A}} = \left(\frac{r_\mathrm{s,A}}{r_\mathrm{s}}\right)^3, \tag{5.27}$$

where $\rho_\mathrm{b,A}$ and $r_\mathrm{s,A}$ are, respectively, the density of the bubble and the radius of the shell at the end of the adiabatic phase. Consequently, we can write the pressure of the bubble as

$$P_\mathrm{b} = P_\mathrm{b,A}\left(\frac{r_\mathrm{s,A}}{r_\mathrm{s}}\right)^5 \tag{5.28}$$

We now recall that the pressure can be related to the internal energy density so that, at the end of the adiabatic phase, $P_\mathrm{b,A} = (2/3)U_\mathrm{b,A}$ for $\gamma = 5/3$. The total internal energy at that time will just be half the initial energy (because of equipartition, §5.2.1) and thus the pressure can be rewritten as[6]

$$P_\mathrm{b} = \frac{2}{3}\frac{E_\mathrm{SN}/2}{\frac{4}{3}\pi r_\mathrm{s,A}^3}\left(\frac{r_\mathrm{s,A}}{r_\mathrm{s}}\right)^5 = \frac{E_\mathrm{SN}}{4\pi}r_\mathrm{s,A}^2 r_\mathrm{s}^{-5} \tag{5.29}$$

and, finally, eq. 5.25 can be written as

$$r_\mathrm{s}^3\frac{\mathrm{d}}{\mathrm{d}t}(r_\mathrm{s}^3 v_\mathrm{s}) = C, \tag{5.30}$$

where

$$C \equiv \left(\frac{3}{4\pi}\frac{E_\mathrm{SN}}{\mu n_0 m_\mathrm{p}}r_\mathrm{s,A}^2\right) \tag{5.31}$$

is a constant and we can look for solutions of the kind $r_\mathrm{s} = At^\alpha$. This self-similar formulation of the solution leads to a velocity $v_\mathrm{s} = \mathrm{d}r_\mathrm{s}/\mathrm{d}t = A\alpha t^{\alpha-1}$ and equation 5.30 can be rewritten as

$$A^7\alpha(4\alpha-1)t^{7\alpha-2} = C \tag{5.32}$$

that returns as only possible solution $\alpha = 2/7$ in order for the l.h.s. to be a constant. With this solution we also find an expression for the normalization constant as

$$A = \left(\frac{49}{2}C\right)^{1/7} = \left(\frac{147}{8\pi}\frac{E_\mathrm{SN}r_\mathrm{s,A}^2}{m_\mathrm{p}\mu n_0}\right)^{1/7} \tag{5.33}$$

and thus the final evolution of the radius of the SNR in the radiative phase is

$$r_\mathrm{s} \simeq 37\left(\frac{E_\mathrm{SN}}{10^{51}\,\mathrm{erg}}\right)^{1/7}\left(\frac{\mu}{0.6}\right)^{-1/7}\left(\frac{n_0}{1\,\mathrm{cm}^{-3}}\right)^{-1/7}\left(\frac{t}{10^5\,\mathrm{yr}}\right)^{2/7}\,\mathrm{pc} \tag{5.34}$$

and the velocity of the thin shell is

$$v_\mathrm{s} \simeq 115\left(\frac{E_\mathrm{SN}}{10^{51}\,\mathrm{erg}}\right)^{1/7}\left(\frac{\mu}{0.6}\right)^{-1/7}\left(\frac{n_0}{1\,\mathrm{cm}^{-3}}\right)^{-1/7}\left(\frac{t}{10^5\,\mathrm{yr}}\right)^{-5/7}\,\mathrm{km\,s}^{-1}, \tag{5.35}$$

[6]It is interesting to note that the total internal energy of a system that expands adiabatically is $\mathsf{U} \propto Ur^3 \propto Pr^3$. Thus if the mass of the system does not change, the internal energy always decreases as $\mathsf{U} \propto r^{-2}$ as a consequence of an expansion (for $\gamma = 5/3$).

both expressions with very low dependences on energy input and density, again demonstrating that the size of the SNR and its expansion velocity is very mildly affected by these variables and much more by the thermodynamic state (adiabatic or radiative) or the shell. The self-similar solution is not a proper integration of eq. 5.30 and we should check whether or not the radius and the velocity here determined agree with those found above for the end of the Sedov phase. If we use eqs. 5.34 and 5.35 and $t_{\rm A} = 5 \times 10^4\,{\rm yr}$ we obtain $r_{\rm s}(t_{\rm A}) \simeq 29\,{\rm pc}$ and $v_{\rm s}(t_{\rm A}) \simeq 222\,{\rm km\,s^{-1}}$ in fairly good agreement with eqs. 5.19 and 5.20.

5.3.1. Simplified solution: snowplough. In cases where the pressure of the bubble can be neglected, the evolution of the SNR can take an easier form than the one determined in §5.3. This situation is likely to take place in every SNRs at the end of their evolution when holes in the shell, thermal conduction between interior and shell or even radiative cooling from the interior can de-pressurise the bubble. In this circumstance, eq. 5.24 becomes very simple

$$\frac{\mathrm{d}}{\mathrm{d}t}(M_{\rm s} v_{\rm s}) = 0, \tag{5.36}$$

and can be readily integrated

$$r_{\rm s}^4 = 4 r_{\rm s,A}^3 v_{\rm s,A} t - 3 r_{\rm s,A}^4 \tag{5.37}$$

and for $t \gg t_{\rm A}$ we have the asymptotic solution

$$r_{\rm s} \simeq 1.1\, r_{\rm s,A} \left(\frac{t}{t_{\rm s,A}}\right)^{1/4} \simeq 38 \left(\frac{t}{10^5\,{\rm yr}}\right)^{1/4}\ {\rm pc} \tag{5.38}$$

and

$$v_{\rm s} \simeq 93 \left(\frac{t}{10^5\,{\rm yr}}\right)^{-3/4}\ {\rm km\,s^{-1}}. \tag{5.39}$$

This phase is often referred to as **snowplough phase** given that the SNR keeps accumulating ISM mass and it is simply slowed down by this accumulation. From eq. 5.36 the velocity has to decrease proportionally to the mass engulfed by the shell in order for the momentum to be conserved.

5.3.2. End of the radiative phase. The SNR keeps expanding at an ever decreasing speed until it mixes with the surrounding ISM. This expansion will certainly cease when the velocity of the shell becomes of the same order of the typical random speeds in the ISM: $\sigma_{\rm HI} \sim 10\,{\rm km\,s^{-1}}$. In order to find this time ($t_{\rm R}$), we use eq. 5.35 and impose equality with $\sigma_{\rm HI}$:

$$v_{\rm s}(t_{\rm R}) = 115 \left(\frac{t}{10^5\,{\rm yr}}\right)^{-5/7} \simeq 10\,{\rm km\,s^{-1}} \implies t_{\rm R} \simeq 3.1 \times 10^6\,{\rm yr} \tag{5.40}$$

and consequently $r_{\rm s}(t_{\rm R}) \simeq 98\,{\rm pc}$. Thus, the maximum age of a SNR is of the order of a million years and its maximum radius is $\approx 100\,{\rm pc}$.

5.4. Energy budget

As we have seen, a SNR starts its expansion as an ejection of high speed material (ejecta) in the ISM and thus with a large amount of kinetic energy. During its evolution, most of this initial energy is lost through radiation. However, a fraction of it survives until the end of the radiative phase. This is due to the fact that the shell cannot reach zero expansion velocity as there is a substrate velocity in

the ISM (turbulence velocity mostly but also thermal velocity of the gas) and the SNR effectively mixes with the surrounding medium when it reaches this speed. At that point, the remaining kinetic energy is given to the ISM and may be the most important source of energy to feed its turbulence (Chap. 8). We can estimate the amount of energy left at the end of the radiative phase as follows

$$E_{\rm K}(t_{\rm R}) \simeq \frac{1}{2} M_{\rm s} v_{\rm s,R}^2 = \frac{2\pi}{3} m_{\rm p} \mu n_0 r_{\rm s,R}^3 v_{\rm s,R}^2, \tag{5.41}$$

where $v_{\rm s,R}$ is the velocity at $t_{\rm s,R}$ (from eq. 5.35). We can then compare this energy with the kinetic energy at the end of the adiabatic phase, which, as we have established (§5.2.1), must equal half of the initial energy of the blast:

$$E_{\rm K}(t_{\rm A}) = \frac{E_{\rm SN}}{2} \simeq \frac{2\pi}{3} m_{\rm p} \mu n_0 r_{\rm s,A}^3 v_{\rm s,A}^2. \tag{5.42}$$

As a consequence, the efficiency of the SNR in transferring kinetic energy to the ISM is

$$\eta = \frac{E_{\rm K}(t_{\rm R})}{E_{\rm SN}} \simeq \frac{1}{2} \left(\frac{r_{\rm s,R}}{r_{\rm s,A}} \right)^3 \left(\frac{v_{\rm s,R}}{v_{\rm s,A}} \right)^2 \simeq 0.03 \tag{5.43}$$

using the values given throughout this Chapter. Clearly the efficiency that we can expect is of the order of a few percent. From considerations based on incompressible turbulence (see Section 8.3.5) this energy may be just enough to keep the turbulence of the ISM at the observed value. Stellar winds (for which see [CFN19]) can also contribute to this energy input.

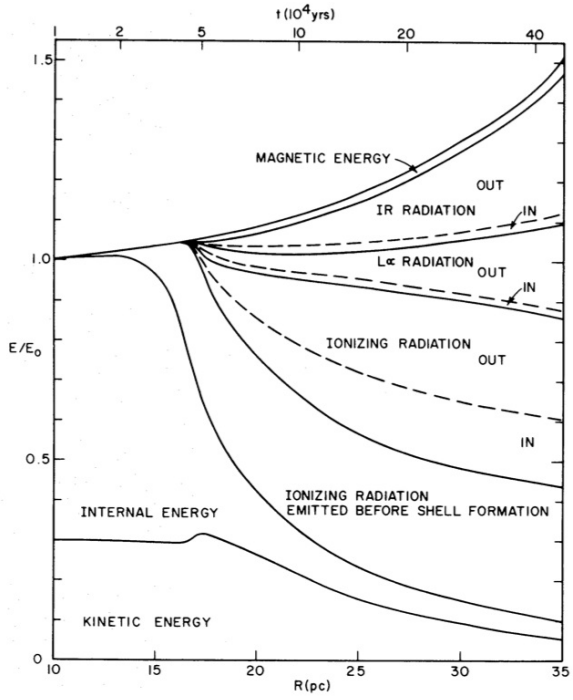

FIGURE 3. Energy budget throughout the adiabatic and radiative phases of the expansion of a SNR, from a simulation of [Che74].

Figure 3 shows the schematic time evolution of the various energies of the SNR as derived through hydrodynamic simulations that include radiative cooling. At the beginning, the vast majority of the energy is kinetic as the ejecta do not feel the reaction of the surrounding medium (this phase is not shown in Fig. 3). Half of this energy is lost by the progressive engulfing of external material and the formation of an adiabatic shock. Near (30% versus 70%) equipartition between kinetic (ordered) and internal (random) energies arises as the SNR enters the adiabatic phase. Radiative losses from the shell, already begun in the Sedov phase, cannot be neglected any longer after a time equal to $t_{\rm A} \simeq \text{a few} \times 10^4\,\text{yr}$ and the shell rapidly loses most of its internal energy. The kinetic energy, still present in the shell, is slowly lost because of the conservation of the momentum. At the end of the radiative phase a fraction of this kinetic energy will still be present in the shell and be given to the surrounding ISM. Note that the total energy of the system tends to increase with time in these simulations (upper envelop in Fig. 3). This is due to the fact that, as the SNR evolves, it sweeps up more and more ISM effectively incorporating its internal energy.

5.5. Evolution of a SNR in a realistic ISM

In this final section, we describe events that are expected to occur during the evolution of the SNR in a realistic ISM.

5.5.1. Rayleigh-Taylor instability. In §6.2 we describe the Rayleigh-Taylor instability and we mention that it can occur not only in the presence of a gravitational field but also when there is a prolonged acceleration. Except for the initial free expansion, the SNR is always decelerating throughout its evolution. This deceleration corresponds to an acceleration towards the interior and thus to a gravitational field from the bubble to the shell. The shell is much denser than the bubble and this is clearly a stable configuration according to the Rayleigh-Taylor criterion. However, there is a moment in which the acceleration is reversed. This happens at the time of the thin shell formation. The adiabatic shell is not a proper shell but more like a zone behind the shock where the gas struggles to keep up with the shock following it at $(3/4)v_{\rm s}$. When this gas starts radiating, the shocked material goes from $3/4v_{\rm s}$ to $\approx v_{\rm s}$ in a relative short time producing the formation of a thin shell right behind the shock[7]. The acceleration of the shell material from the interior to the shell, also helped by the bubble pressure, can be seen as a gravity field oriented inwards. This time, the denser medium is *above* the rarefied medium of the bubble and this is the ideal condition for Rayleigh-Taylor instabilities to take place.

Fingers of shell material will protrude inside the bubble and these are seen in various SNR, the most remarkable of all being the Crab Nebula. We can estimate the size (width) of these fingers as follows. The acceleration $a_{\rm thin}$ can be calculated taking the velocity of the shell at the end of the adiabatic phase ($v_{\rm s} \sim 200\,\text{km}\,\text{s}^{-1}$) and a typical time for shell formation of about $\Delta t_{\rm thin} \sim 10^4\,\text{yr}$. The dispersion relation for Rayleigh-Taylor instability gives

$$\frac{\omega}{k} = \pm\sqrt{\frac{g}{k}\frac{\rho_{\rm s}-\rho_{\rm b}}{\rho_{\rm s}+\rho_{\rm b}}} \simeq \pm i\sqrt{\frac{a_{\rm thin}}{k}}, \tag{5.44}$$

[7]Note that although the velocity of the shocked isothermal gas is not exactly $v_{\rm s}$, the small difference is compensated by the fact that the shock itself is decelerating. Thus, in practice, the velocity of thin shell in the radiative phase is essentially coincident with that of the shock.

where $\rho_{\rm s}$ and $\rho_{\rm b}$ are the densities of the shell and the bubble respectively and the last equality is true because the density of the shell is much large than $\rho_{\rm b}$. Given the definition of k, we can therefore estimate that the typical width of the perturbations (fingers) will be

$$\lambda \sim \frac{a_{\rm thin}}{\omega^2} \sim v_{\rm s}\Delta t_{\rm thin} \approx 2\,{\rm pc}. \tag{5.45}$$

5.5.2. Evolution of a SNR in a galactic disc. In all the above we have considered the external medium to be fully isotropic. Here, we ask ourselves whether the variation of the density of the ISM in the vertical direction is relevant for the evolution of a SNR. The fact that the scale height is of the same order of the size of the SNR may indicate that some effects should be expected. In practice, this is hardly the case for shells produced by single supernovae whereas it becomes very important when we are dealing with *supershells* produced by the combined action of winds and explosions of several stars. Let us assume that the vertical distribution of the disk is Gaussian (§3.1). If the supernova explodes exactly in the plane of the disc, at the end of its evolution, in the vertical direction, it will encounter a mean density of

$$\rho(z = r_{\rm s,R}) = \rho_0 e^{-\frac{z^2}{2h^2}} \simeq \rho_0 e^{-\left(\frac{90^2}{2\times 150^2}\right)} \simeq 0.8\rho_0, \tag{5.46}$$

where $h \simeq 150\,{\rm pc}$ is the thickness of the ISM at the Solar Neighbourhood (§3.1). Equation 5.46 shows that the vertical change of ISM density makes little difference in the SNR evolution, especially considering that both in the Sedov and in the radiative phases the dependence on the density is very small (1/5 and 1/7 power respectively). This situation may modify slightly if the SN explodes with some offset with respect to the plane of the disc.

5.5.3. Expansion in a non-homogeneous ISM. The real ISM is not a homogeneous medium in any respect. There are indeed regions at lower and higher densities, in particular relatively compact *clouds* at lower temperatures characterizing the cold neutral medium (CNM) (Chap. 1). Let us consider the simple case of one such cloud that is entering the supernova shell. Given that the cloud has a density different from the surrounding medium the shock will propagate through it at a different speed. From the shock theory we have that the pressure in the shocked medium inside the cloud is

$$P_{1,\rm c} = \rho_{\rm c} v_{\rm s,c}^2, \tag{5.47}$$

where we used the subscript c to indicated the cloud. If the cloud survives once entered the shock/shell it will do so by reaching a pressure equilibrium with the ambient. Thus we can expect the cloud pressure to be not too different from the pressure of the perturbed ISM ($P_1 = \rho_0 v_{\rm s}^2$) and so

$$\rho_{\rm c} v_{\rm s,c}^2 \sim \rho_0 v_{\rm s}^2 \implies v_{\rm s,c} \sim \left(\frac{\rho_0}{\rho_{\rm c}}\right)^{1/2} v_{\rm s} \sim 0.1 v_{\rm s} \tag{5.48}$$

where the last identity holds if the cloud is CNM and the average ISM is WNM/WIM.

An important consequence of the above simple calculation is that we expect different regions of SNRs to propagate at different velocities in a realistic non-homogeneous ISM. These differences in velocities are in fact observed in several SNRs (e.g. Cas A). Moreover, given that the temperature of the shocked gas is $T_1 \propto v_{\rm s}^2$ we can expect the temperature of a denser region (engulfed cloud) to drop

more quickly than the temperature of the surrounding medium. In other words, the SNR shock can become *locally* radiative and suffer important losses of energy even during the overall adiabatic phase. Finally, the destiny of the cloud is likely to be deformed and destroyed both by Kelvin-Helmholtz instability and by thermal conduction.

5.5.4. Non-thermal emission. Not all the energy emission from SNR is thermal and due to the increase of temperature and density of the shocked medium. An important component is constituted by non-thermal emission and in particular by the *synchrotron radiation* produced by relativistic electrons gyrating around the magnetic field lines of the medium. The electrons are most probably accelerated by the explosion and possibly re-accelerated by the passage through the shock (Fermi mechanism). The magnetic field already present in the medium is enhanced by the compression due to the passage of the shock (§9.2). In the end, the synchrotron radiation turns out to be one of the most prominent emissions of SNRs and a very efficient way to detected them both in our Galaxy and in external galaxies. The amount of energy released by this process is, however, limited due that most of the emission is at radio-wavelengths that are not carrying much energy compared to X-ray and optical bands.

CHAPTER 6

Instabilities

One of the most important aspects in the study of the dynamics of fluids, and gases in particular, is the analysis of the conditions under which a pertubation can generate instabilities. In the ISM, as in any astrophysical gaseous medium, perturbations are always present. Such perturbations can be of different kinds and involve different variables of the fluid such as density and temperature. In this chapter, we present the **linear perturbation analysis** of the most common instabilities that can develop as a consequence of *small perturbations*, i.e. that can be considered infinitesimal with respect to typical variations of the quantity in question. The goal of the linear analysis is to determine what are the conditions for an instability to grow. Once the instability starts to develop, the subsequent evolution pertains to the **non-linear theory** and it is typically studied using numerical simulations.

Instabilities are considered at the root of several fundamental astrophysical phenomena. Prime examples are the gravitational (Jeans) instability (§6.3) that is thought to play a role in the formation of stars and galaxies and the thermal instability (§6.5) that is at the origin of the phases of the ISM (§1.1). We also discuss convection (§6.1), very important in stellar interiors but with relevance also in galactic coronae; Kelvin-Helmholtz, Rayleigh-Taylor (§6.2) and rotational instabilities (§6.4).

6.1. Convection

Consider a fluid stratified in the vertical direction (z) under the action of a gravitational field pointing downwards[1]. We take the fluid to be in hydrostatic equilibrium i.e. the following equation holds between its pressure and an external gravitational potential (see §3.1):

$$\frac{\mathrm{d}P}{\mathrm{d}z} = -\rho\frac{\mathrm{d}\Phi}{\mathrm{d}z} = -\rho|g|, \tag{6.1}$$

where $g \equiv -\mathrm{d}\Phi/\mathrm{d}z$ is the acceleration that we assume constant with z at least in the small region of interest that we are exploring. The minus sign is due to the fact that the gradient of the pressure must be negative as it decreases with z.

Take a non-infinitesimal portion (a blob) of the fluid and consider a perturbation that moves it upward by an infinitesimal amount Δz. Let P and $\mathcal{S}$ be the pressure and specific entropy at a certain height z, both of the fluid blob and of the surrounding ambient medium. We assume that the displacement to $z+\Delta z$ happens slowly so that the process can be consider adiabatic. As a consequence, the blob at $z' = z + \Delta z$ will then have a slightly different pressure $(P')^2$ but the same entropy

[1]Note that the treatment and the solutions that we present here is analogous to what one gets for a spherical system stratified in the radius r.

[2]Note that P' is the pressure of the ambient medium at z'.

(adiabatic transformation) as it started with. The reason that justifies this assumption of an adiabatic transformation is rooted in the profound difference between the establishment of two very different kinds of equilibria: the pressure and thermal equilibrium (see §1.3). The former is easily established as a consegnence of the development of sound waves and it requires a time of the order of the crossing time (eq. 2.65). The thermal equilibrium is instead established by thermal conduction (in the absence of radiation). Thermal conduction acts in the region where thermal gradients are present (surface of the blob) and becomes less and less efficient as the gradient gets smoother. As a consequence, in general, the thermal equilibrium takes much longer to be established than the pressure equilibrium and the drifting of a fluid blob can indeed be considered almost adiabatic. In practice, the blob conserves its entropy during its motion but gets compressed or expands (slightly) in order to achieve pressure equilibrium with the ambient medium.

If we express the density of the blob in terms of pressure and specific entropy at the location z', this density will be $\rho(P', \mathcal{S})$, while at the same location the density of the ambient is, by construction, $\rho(P', \mathcal{S}')$. Given that gravity is pointing downwards the condition of *instability* then becomes[3]

$$\rho(P', \mathcal{S}) < \rho(P', \mathcal{S}'), \tag{6.3}$$

i.e. the blob has to be lighter than the ambient gas. If this is not the case, the blob will fall back to its original position and thus its displacement will simply end up in an oscillation that will quickly damp in time. If the blob is indeed lighter it will be *buoyant* at z' and proceed upwards. In this latter case thus, the upward displacement causes a further upward push, which is a typical *unstable* configuration.

In the search for an instability criterion, it is handy to make use of specific volumes instead of densities. This leads us to the following condition for instability:

$$\mathcal{V}(P', \mathcal{S}') - \mathcal{V}(P', \mathcal{S}) < 0. \tag{6.4}$$

Equation 6.4 can be divided and multiplied by $\mathcal{S}' - \mathcal{S}$, and by Δz, thus obtaining

$$\frac{\mathcal{V}(P', \mathcal{S}') - \mathcal{V}(P', \mathcal{S})}{\mathcal{S}' - \mathcal{S}} \frac{\mathcal{S}' - \mathcal{S}}{\Delta z} \Delta z < 0. \tag{6.5}$$

Then, recalling that primed quantities are at $z + \Delta z$ and non-primed are at z, and that Δz is positive because we are displacing the blob upwards, one can take the

[3]The reason for this is **buoyancy**, which is a modern way to describe the **Archimedes principle**. Consider a body immersed in a fluid, which is subject to a gravity field along z (in our situation the body is a portion of the fluid itself). Given that the pressure of the fluid is also stratified with z and increases toward lower z (for a gravity pointing downwards), the body will feel an upward pressure force. This **buoyant force** can be written

$$\boldsymbol{f}_{\rm buoy} = \rho V \frac{\mathrm{d}\Phi}{\mathrm{d}z} \boldsymbol{e}_z, \tag{6.2}$$

where V is the volume of the body, ρ is the density of the fluid and $\boldsymbol{e}_z$ is the unity vector along z. Because of hydrostatic equilibrium, this force is exactly equal (with opposite sign) to the gravity force acting on a portion of the fluid with the same volume V of the body in that position. The gravity force acting on the body is instead $-\rho_{\rm body} V (\mathrm{d}\Phi/\mathrm{d}z)\boldsymbol{e}_z$, where $\rho_{\rm body}$ is the body density. Then, clearly, if $\rho_{\rm body} < \rho$, the body becomes *buoyant* and it moves upward, otherwise it sinks down in the potential well. Archimedes expressed this by saying that the buoyant push on the body equals the *weight* of the fluid that it displaces. In the practical case of a body in water, if the weight of the body is smaller than the weight of the displaced water, then the push points upward and the body floats.

limit of equation 6.5 to obtain the following condition for instability:

$$\left(\frac{\partial \mathcal{V}}{\partial \mathcal{S}}\right)_P \frac{\mathrm{d}\mathcal{S}}{\mathrm{d}z} < 0, \tag{6.6}$$

where the partial derivative is taken at constant pressure.

The first derivative on the l.h.s. of eq. 6.6 can be written as

$$\left(\frac{\partial \mathcal{V}}{\partial \mathcal{S}}\right)_P = \left(\frac{\partial \mathcal{V}}{\partial T}\right)_P \left(\frac{\partial T}{\partial \mathcal{S}}\right)_P = \left(\frac{\partial \mathcal{V}}{\partial T}\right)_P \frac{T}{C_\mathrm{p}}, \tag{6.7}$$

where we have used the definition of specific heat at constant pressure (see D.1.4). For essentially every gas, the derivative of the gas volume with temperature at constant pressure is positive: the volume increases as the temperature increases (easy to see using the equation of state of ideal gases, eq. 2.18). Thus in conclusion, given that T and C_p are always positive, we obtain the final condition for **convection instability**:

$$\frac{\mathrm{d}\mathcal{S}}{\mathrm{d}z} < 0, \tag{6.8}$$

which establishes that convection will only occur in the presence of a negative entropy gradient. Note that this condition does not depend explicitly on the properties of the gravitational field (but see also later). When the condition for instability is fulfilled, a *convective zone* will form in the medium. In this zone, large portions of the fluid move from (inner) regions at high entropy to (outer) regions at low entropy and vice versa. These motions are aimed to flatten the negative entropy gradient. Relatively flat entropy gradients in systems in hydrostatic equilibrium can be seen as regions of potential (or past) convection instabilities.

The above condition (eq. 6.8) would be very useful to tag convectively unstable regions if we had detailed entropy profiles of astrophysical systems. Unfortunately, entropy is not an easy quantity to measure as it requires the simultaneous estimate of two other variables, typically temperature and density. Its behaviour as a function of position is mostly known for the gaseous halos of some galaxy clusters where it can be traced quite accurately (and it generally increases with radius). In most situations, e.g. for the hot halos of galaxies, it would be useful to have a condition involving the temperature of the gas; we see now how to obtain such a condition. We express the specific entropy as a function of pressure and temperature $\mathcal{S} = \mathcal{S}(P, T)$. Then the derivative in eq. 6.8 can be developed as

$$\frac{\mathrm{d}\mathcal{S}}{\mathrm{d}z} = \left(\frac{\partial \mathcal{S}}{\partial P}\right)_\mathrm{T} \frac{\mathrm{d}P}{\mathrm{d}z} + \left(\frac{\partial \mathcal{S}}{\partial T}\right)_\mathrm{P} \frac{\mathrm{d}T}{\mathrm{d}z}, \tag{6.9}$$

where the first derivative on the r.h.s. can be changed using the following Maxwell relation[4]

$$\left(\frac{\partial \mathcal{S}}{\partial P}\right)_{\mathrm{T}} = -\left(\frac{\partial \mathcal{V}}{\partial T}\right)_{\mathrm{P}}. \tag{6.12}$$

The second derivative on the r.h.s. of eq. 6.9 can be readily changed by using the definition of specific heat (eq. D.17) to obtain

$$\frac{\mathrm{d}\mathcal{S}}{\mathrm{d}z} = -\left(\frac{\partial \mathcal{V}}{\partial T}\right)_{\mathrm{P}} \frac{\mathrm{d}P}{\mathrm{d}z} + \frac{C_{\mathrm{p}}}{T}\frac{\mathrm{d}T}{\mathrm{d}z} < 0. \tag{6.13}$$

We can now use the condition for hydrostatic equilibrium (eq. 6.1) to modify the first term of the r.h.s. and write

$$\frac{dT}{dz} < -\rho T \left(\frac{\partial \mathcal{V}}{\partial T}\right)_{\mathrm{P}} \frac{|g|}{C_{\mathrm{p}}}, \tag{6.14}$$

where we used the absolute value of the acceleration to assure that $\mathrm{d}\Phi/\mathrm{d}z > 0$ (acceleration pointing downwards). Finally, using the law of ideal gases (§D.1.1) we have that $(\partial\mathcal{V}/\partial T)_{\mathrm{P}} = 1/(\rho T)$ and thus we obtain a condition for convection instability involving the temperature gradient:

$$\frac{dT}{dz} < -\frac{|g|}{C_{\mathrm{p}}}. \tag{6.15}$$

Equation 6.15 establishes that for positive and flat gradients of the temperature, in systems in hydrostatic equilibrium, there will always be convection *stability*. The condition for instability arises when and where the temperature decreases with z (or r for spherical systems). However, the temperature gradient has to be very steep and to fulfill this condition

$$\left|\frac{dT}{dz}\right| > \frac{|g|}{C_{\mathrm{p}}}. \tag{6.16}$$

The most typical astrophysical situations in which eq. 6.15 is satisfied are the interiors of massive stars. Here the CNO cycle brings the inner layers to very high temperatures and a sharp temperature gradient can form between the inner and the outer regions. A portion of the star becomes convectively unstable producing an efficient mixing of the various layer. The process is often treated making used of the so-called *mixing length*, defined as the path that a portion of the layer travels before achieving equilibrium, i.e., roughly, before it finds an ambient at the same entropy as its own.

---

[4]This is obtained by recalling the first law of thermodynamics written using the Gibbs free energy: $\mathrm{d}\mathcal{G} = -\mathcal{S}\mathrm{d}T + \mathcal{V}\mathrm{d}P$ and considering that, in general, one can write $\mathcal{G} = \mathcal{G}(T, P)$ and therefore

$$\mathrm{d}\mathcal{G} = \left(\frac{\partial \mathcal{G}}{\partial T}\right)_{\mathrm{P}} \mathrm{d}T + \left(\frac{\partial \mathcal{G}}{\partial P}\right)_{\mathrm{T}} \mathrm{d}P, \tag{6.10}$$

which gives a way to write $\mathcal{S}$ and $\mathcal{V}$ as partial derivatives of $\mathcal{G}$. The above relation is finally obtained using the identity

$$\left[\frac{\partial}{\partial P}\left(\frac{\partial \mathcal{G}}{\partial T}\right)_{\mathrm{P}}\right]_{\mathrm{T}} = \left[\frac{\partial}{\partial T}\left(\frac{\partial \mathcal{G}}{\partial P}\right)_{\mathrm{T}}\right]_{\mathrm{P}}. \tag{6.11}$$

6.2. Kelvin-Helmholtz and Rayleigh-Taylor

In this section, we describe two of the most important hydrodynamic instabilities that can occur in astrophysical fluids. We treat these two instabilities together given their similarity and the fact that a general solution can be found, albeit with some simplifying assumptions (see also [CC14]).

Consider two fluids stratified in z and in direct *contact*, with the plane at $z = 0$ being their contact surface[5]. The two fluids are infinite in x and y and there are no relevant motions in the y direction. We assume that the velocity fields of the fluids are both irrotational ($\boldsymbol{\nabla} \times \boldsymbol{u} = 0$) and solenoidal ($\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$). We recall that, given the continuity equation (eq. 2.11) the latter assumption corresponds to assume that the fluids are incompressible. This is an obvious limitation for the application of this treatment to astrophysical gases but given that we are only exploring the linear regime of the perturbation it is not too problematic. Indeed, this simplified analysis gives results (in the linear part) that are in very good agreement with both experiments and hydrodynamical simulations. Finally, we assume an external gravitational field with constant acceleration pointing downward. The two fluids have, in general, different initial densities, ρ and ρ' respectively for the fluid below and above $z = 0$. They are also moving with two velocities along x that are, in general, not equal ($u_0 \neq u_0'$), see Fig. 1.

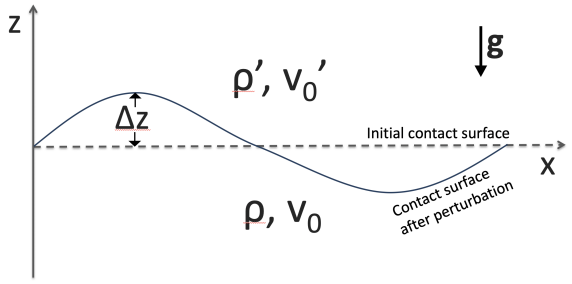

FIGURE 1. Sketch of the two fluids in contact and the onset of a perturbation that can lead to Kelvin-Helmholtz and/or Rayleigh-Taylor instabilities.

Consider now an infinitesimal perturbation that takes place exactly at the interface (contact surface) of the two fluids. Perturbing the interface has the consequence of changing slightly the positions and velocities of the fluid elements there located (Fig. 1). Given that the unperturbed elements are by definition at $z_0 = 0$, the perturbation can be written, for instance in the lower fluid, as

$$\begin{cases} z = & \Delta z \\ \boldsymbol{u} = & \boldsymbol{u}_0 + \Delta \boldsymbol{u}. \end{cases} \tag{6.17}$$

We note that being Δz an infinitesimal displacement in the z direction, it can depend only on the spatial coordinate x and on time, $\Delta z = \Delta z(x, t)$, and *not* on z. The perturbed velocity will have, instead, the following components:

$$\boldsymbol{u} = (u_0 + \Delta u_{\rm x},\ \Delta u_{\rm z}), \tag{6.18}$$

[5]We define as **contact surface** the locus of points where all thermodynamical variables are discontinuous, except the pressure. The typical situation is that of two adjacent fluids with different temperatures and densities that are in pressure equilibrium (at their contact surface).

given that the unperturbed velocity is different from zero only along x. The components of the velocity perturbation along x and z will, in general, depend on both spatial coordinates and time, so, for instance, $\Delta u_{\mathrm{x}} = \Delta u_{\mathrm{x}}(x, z, t)$.

One of our hypothesis is that the velocity field is irrotational, this implies that one can define a scalar quantity whose gradient generates the velocity field. This scalar can be seen as a *velocity potential*[6] ψ:

$$\boldsymbol{\nabla} \times \boldsymbol{u} = 0 \implies \exists\, \psi \text{ such that } \boldsymbol{u} = -\boldsymbol{\nabla}\psi. \tag{6.19}$$

The perturbation in velocity can be then conveniently rewritten as a perturbation in the velocity potential:

$$\psi = -u_0 x + \Delta\psi, \tag{6.20}$$

where the first term of the r.h.s. is the unperturbed velocity potential (we could have also indicated this term with ψ_0). A spatial derivation of eq. 6.20 returns

$$\boldsymbol{u} = -\boldsymbol{\nabla}\psi = u_0 \hat{e}_{\mathrm{x}} - \boldsymbol{\nabla}\Delta\psi \tag{6.21}$$

or, using a component notation (only x and z as usual),

$$\boldsymbol{u} = \left(u_0 - \frac{\partial \Delta\psi}{\partial x},\, -\frac{\partial \Delta\psi}{\partial z}\right), \tag{6.22}$$

which specifies the perturbations in eq. 6.18 in terms of derivatives of the velocity potential. By symmetry, there will be another equation like the above for the top fluid that reads

$$\boldsymbol{u}' = -\boldsymbol{\nabla}\psi' = \left(u_0' - \frac{\partial \Delta\psi'}{\partial x},\, -\frac{\partial \Delta\psi'}{\partial z}\right), \tag{6.23}$$

where ψ' is the velocity potential that generates $\boldsymbol{u}'$ and $\Delta\psi'$ is its perturbation.

Let us now write the equation that describes the velocity of a generic perturbed fluid element at the interface, again we first consider the bottom fluid. Any fluid element will experience a change in time of its vertical position that must correspond to the perturbation of its velocity in the same z direction. We can write this as

$$\frac{\mathrm{D}\Delta z}{\mathrm{D}t} = \Delta u_{\mathrm{z}}, \tag{6.24}$$

where we have used the Lagrangian derivative given that we are following the displacement of a fluid element. We substitute the Lagrangian derivative using its relation with the Eulerian derivative (eq. 2.8) and we use equations 6.22 and 6.23. Applying the same reasoning to the fluid above we obtain the following equations for a generic perturbed fluid element at the interface:

$$\begin{cases} \dfrac{\partial \Delta z}{\partial t} + u_0 \dfrac{\partial \Delta z}{\partial x} = & -\dfrac{\partial \Delta\psi}{\partial z} \\ \dfrac{\partial \Delta z}{\partial t} + u_0' \dfrac{\partial \Delta z}{\partial x} = & -\dfrac{\partial \Delta\psi'}{\partial z}, \end{cases} \tag{6.25}$$

respectively for the fluid below (top eq.) and above (bottom eq.). Note that eqs. 6.25 refer to two generic fluid elements in contact with each other at the contact surface (belonging to fluids above and below the surface, primed and non-primed respectively) and separated by an infinitesimal distance (as they have infinitely small sizes, cf. §2.1). For this reason, we used the same Δz and not a $\Delta z'$ for the

[6] We recall that $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} A) = 0, \forall A$, see §C.4.2.

top fluid. Note also that, instead, given that the velocities are, in general, different in the two fluids ($\boldsymbol{u} \neq \boldsymbol{u'}$), so are their potentials and perturbations.

Equations 6.25 can be see then as two linear equations describing the velocities of fluid elements at the interface as a consequence of a perturbation. The general solution of this system is not known but, being it periodic (as the perturbation), we can decompose it in a Fourier series. Therefore, the general behaviour of the perturbation (see also discussion in §2.10) is simply determined by studying the following solutions:

$$\Delta z = \widetilde{\Delta z}\, e^{i(kx-\omega t)} \tag{6.26}$$

$$\Delta \psi = \widetilde{\Delta \psi}\, e^{i(kx-\omega t)+ihz} \tag{6.27}$$

$$\Delta \psi' = \widetilde{\Delta \psi}\, e^{i(kx-\omega t)+ih'z}, \tag{6.28}$$

where k and ω are the usual wave number and orbital frequency of the oscillation (k being along x), while h and h' are generic wave numbers along z that we have to determine.

We make use of the assumption that the velocity field is solenoidal and thus the divergence of the velocity is zero everywhere. We recall that this, by use of eq. 2.11, is the same as to assert that the fluid is incompressible. As mentioned, such an assumption is nearly valid in the linear regime of the small perturbations. The compressible nature of the gas may become important in the further development of the instability and the onset of turbulence, however this analysis does not aim to describe that phase. The linear perturbation analysis mainly addresses the question of whether, given the initial conditions of the fluid and the type of perturbation considered, the instability will develop or not. Thus, without further hesitation, we write

$$\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0 \implies \nabla^2 \psi = 0, \tag{6.29}$$

given the definition of the Laplacian operator (eq. C.11). Using our definition of velocity potential and its perturbation, eq. 6.29 becomes

$$\frac{\partial^2 \Delta\psi}{\partial x^2} + \frac{\partial^2 \Delta\psi}{\partial z^2} = 0, \tag{6.30}$$

with obviously an analogous equation for the upper fluid. Substituting now eqs. 6.27 and 6.28 one obtains that $h = \pm ik$ and, analogously, $h' = \pm ik$ for the upper fluid. Finally, to choose between the plus or the minus sign, we impose that the perturbation goes to zero for z going to infinity. At negative z (lower fluid), the perturbation must be $\Delta\psi \propto e^{kz}$ for $z \to -\infty$, which implies $h = -ik$. At positive z (upper fluid), the perturbation must go to zero for $z \to +\infty$ (thus $\Delta\psi \propto e^{-kz}$), which implies $h = +ik$.

In the end, our three perturbations (eqs. 6.26, 6.27 and 6.28) luckily depend only on k and ω (constituents of the dispersion relation that we aim to find). If we substitute these solutions in the equations of the velocity of the fluid at the interface (eq. 6.25) we obtain

$$\begin{cases} i(ku_0 - \omega)\widetilde{\Delta z} = & -k\widetilde{\Delta\psi} \\ i(ku_0' - \omega)\widetilde{\Delta z} = & -k\widetilde{\Delta\psi}', \end{cases} \tag{6.31}$$

i.e. a system of two equations with three unknowns ($\widetilde{\Delta z}$, $\widetilde{\Delta\psi}$, $\widetilde{\Delta\psi}'$). We clearly need another equation.

The third equation can be obtained using the condition of pressure equilibrium at the interface. This imposition is implicit in the definition of contact surface and it is a necessary condition to assure the survival and the segregation of the two fluids before the development of the perturbation. We start by writing the Euler equation that, for instance, for the bottom fluid, reads

$$-\frac{\partial \boldsymbol{\nabla}\Delta\psi}{\partial t} + \boldsymbol{\nabla}\frac{1}{2}u^2 = -\frac{1}{\rho}\boldsymbol{\nabla}P - \boldsymbol{\nabla}\Phi, \tag{6.32}$$

where we have used the fact that the velocity field is irrotational (and the identity in eq. C.17). We note that eq. 6.32 has gradient terms for all quantities[7] expect for the first term on the r.h.s. This term can be changed by imposing the incompressibility of the fluid[8] to obtain an equation that can be straightforwardly integrated in the *spatial* domain leading to

$$-\frac{\partial \Delta\psi}{\partial t} + \frac{1}{2}u^2 + \frac{P}{\rho} + \Phi = f(t), \tag{6.33}$$

where $f(t)$ is a constant in space that could, however, be a function of time. The l.h.s. of eq. 6.33 is the so-called **Bernoulli function** plus a perturbation term (first term). In the absence of perturbations, one obtains the well-known **Bernoulli principle**, which states that for incompressible fluids the Bernoulli function $B = u^2/2 + P/\rho + \Phi$ is a constant of the flow. The Bernoulli principle has numerous practical applications in hydraulics and aerodynamics. For $z \to -\infty$ the perturbation vanishes and thus $\Delta\psi \to 0$. This means that all quantities in eq. 6.33 tend to the unperturbed values and so $f(t) = C$ (constant), i.e. it cannot depend on time because u_0, P_0 and Φ_∞ are all constants.

Let us now write eq. 6.33 at the interface between the two fluids. We expand the specific kinetic energy to obtain

$$\frac{1}{2}u^2 = \frac{1}{2}\boldsymbol{u}\cdot\boldsymbol{u} = \frac{1}{2}(u_0\hat{e}_\mathrm{x} - \boldsymbol{\nabla}\Delta\psi)\cdot(u_0\hat{e}_\mathrm{x} - \boldsymbol{\nabla}\Delta\psi) = \frac{1}{2}u_0^2 - u_0\frac{\partial\Delta\psi}{\partial x} + ..., \tag{6.34}$$

neglecting the second-order terms. At the end, reordering the terms to isolate the pressure of the gas, we have that at the interface the pressure of the bottom fluid reads

$$P = \rho\left[\frac{\partial\Delta\psi}{\partial t} + u_0\frac{\partial\Delta\psi}{\partial x} - g\Delta z\right] + \rho\left(C - \frac{1}{2}u_0^2\right) \tag{6.35}$$

where C is the integration constant introduced above. Note that we have substituted the potential Φ with $g\Delta z$, where g is the gravitational acceleration that we take as positive: $g > 0$. The sign of this term is then justified if we set null the potential at $z = 0$ ($\Phi_0 = 0$). Given that gravity points downward, the potential will be then be positive for $\Delta z > 0$ (above the interface) and negative for $\Delta z < 0$.

Given the way eq. 6.35 has been obtained it is clear that there exists a correspondent equation for the top fluid characterised by all primed quantities. Therefore

[7] The first term on the l.h.s. can be simply modified by exchanging derivatives.

[8] Note, however, that here the imposition is that $\boldsymbol{\nabla}\rho = 0$, which is not strictly the same as $\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0$.

the condition of pressure equilibrium at the interface ($P = P'$) can be written as

$$\rho\left[\frac{\partial\Delta\psi}{\partial t} + u_0\frac{\partial\Delta\psi}{\partial x} - g\Delta z\right] = \rho'\left[\frac{\partial\Delta\psi'}{\partial t} + u_0'\frac{\partial\Delta\psi'}{\partial x} - g\Delta z\right] + K, \tag{6.36}$$

where K is a constant that includes the two integration constants and the unperturbed kinetic energies. However eq. 6.36 must be valid also in the unperturbed medium (for vanishing perturbations) and therefore we must have that $K = 0$.

Equation 6.36 represents our third conservation equation to close the system in 6.31. If we substitute in this equation the Fourier modes from eqs. 6.26, 6.27 and 6.28 we obtain the third equation in k and ω. The final system of equations that we aim to solve is the following:

$$\begin{cases} i(ku_0 - \omega)\widetilde{\Delta z} = -k\widetilde{\Delta\psi} \\ i(ku_0' - \omega)\widetilde{\Delta z} = k\widetilde{\Delta\psi}' \\ i(ku_0 - \omega)\rho\widetilde{\Delta\psi} = i(ku_0' - \omega)\rho'\widetilde{\Delta\psi}' + g(\rho - \rho')\widetilde{\Delta z}. \end{cases} \tag{6.37}$$

We look for a dispersion relation that links the wave frequency and wave number by solving the system of eqs. 6.37. This leads to a second order equation in ω/k:

$$(\rho + \rho')\left(\frac{\omega}{k}\right)^2 - 2(\rho u_0 + \rho' u_0')\frac{\omega}{k} + \rho u_0^2 + \rho' u_0'^2 - \frac{g}{k}(\rho - \rho') = 0 \tag{6.38}$$

that can be easily solved to obtain the final dispersion relation for Kelvin-Helmholtz and Rayleigh-Taylor instabilities:

$$\frac{\omega}{k} = \frac{\rho u_0 + \rho' u_0'}{\rho + \rho'} \pm \sqrt{\frac{g}{k}\frac{\rho - \rho'}{\rho + \rho'} - \frac{\rho\rho'}{(\rho + \rho')^2}(u_0 - u_0')^2}. \tag{6.39}$$

We assume that the wave number k is real and ask ourselves under what conditions the wave frequency ω is imaginary. Indeed, given the definition of the solutions (e.g. eq. 6.26), $\mathrm{Im}(\omega) > 0$ is the condition for the development of an instability, because the perturbation does not oscillate but exponentially increases with time.

We can now consider a number of situations and discuss their predictions for the behaviour of the perturbations using eq. 6.39.

(1) Static medium, higher density below ($u_0 = u_0' = 0$, $\rho > \rho'$): The wave frequency is always real and the perturbation leads to oscillations (no instability) that follow the equation

$$\frac{\omega}{k} = \pm\sqrt{\frac{g}{k}\frac{\rho - \rho'}{\rho + \rho'}} \tag{6.40}$$

and are called **gravity waves**. In the case of $\rho \gg \rho'$ (for instance a lake: water in contact with air) the dispersion relation simply becomes

$$\omega = \pm\sqrt{gk}, \tag{6.41}$$

which shows that this wave is *dispersive* with a group speed equal to

$$\frac{\mathrm{d}\omega}{\mathrm{d}k} = \pm\frac{1}{2}\sqrt{\frac{g}{k}} = \pm\frac{1}{2}\frac{\omega}{k}, \tag{6.42}$$

i.e. half of the phase speed.

(2) Static medium, higher density above ($u_0 = u_0' = 0$, $\rho < \rho'$): in this case the wave frequency is never real and we have

$$\frac{\omega}{k} = \pm i \sqrt{\frac{g}{k} \frac{\rho' - \rho}{\rho + \rho'}}. \tag{6.43}$$

This is the typical situation for the development of the **Rayleigh-Taylor instability**. This instability develops as a simple exponential growth of the perturbation amplitude and the development of "fingers" from the upper fluid that penetrate the lower one and viceversa. Note that the gravitational acceleration can be substituted by a generic acceleration pointing in the opposite direction (equivalence principle). Thus we can also expect the development of the Rayleigh-Taylor instability in the presence of strong accelerations. This is what happens in supernova remnants during a particular phase of their evolution (the thin shell formation; see §5.5.1).

(3) Relative motion, higher density below ($u_0 \neq u_0'$, $\rho > \rho'$): This case is stable from the point of view of the Rayleigh-Taylor condition but, given the negative sign of the second term inside the square root in eq. 6.39, the wave frequency can become imaginary. This instance occurs under the condition

$$(u_0 - u_0')^2 > \frac{g}{k} \frac{\rho^2 - \rho'^2}{\rho \rho'} \tag{6.44}$$

and so, when the velocity difference between the two fluids is high enough to overcome the r.h.s. of eq. 6.44, we will have the development of the **Kelvin-Helmholtz instability**. Clearly, if gravity is not very important this instability will always develop whenever two fluids are in contact and in motion with respect to each other. Another situation that facilitates the onset of this instability is the presence of very high wave numbers k (i.e., perturbations with short wave lengths) although, in real fluids, these small perturbations can be damped by viscosity (Chap. 7). In Earth environments, the Kelvin-Helmholtz instability occurs quite often. Examples are water falling from a tap (and moving with respect to the surrounding air) or smoke flow out of a burning cigar. In astrophysics, typical situations are interfaces of cold clouds and filaments moving through an ambient gas and the rims of stellar/AGN jets.

(4) Relative motion, higher density above ($u_0 \neq u_0'$, $\rho < \rho'$): this is the most general case in which both Rayleigh-Taylor and Kelvin-Helmholtz instabilities can develop. Note indeed that both terms within the square root in eq. 6.39 become simultaneously negative. The dominance of one over the other will depend on the reciprocal importance of gravity (on the one side) and velocity (on the other). Note also that, in general, the development of the Rayleigh-Taylor instability causes the two fluids to be in relative motion with each other (fingers penetrating into the other fluid) and this is the natural circumstance in which the Kelvin-Helmholtz instability occurs. Thus, the two instabilities will often be present together in a fully developed Rayleigh-Taylor. An example is the "mushroom" generated by a nuclear bomb explosion.

For all the above cases, eq. 6.39 also allows us to estimate a timescale of the instability. Recalling the equation of the Fourier mode (eq. 6.26), the amplitude of

a perturbation grows with a typical exponential timescale

$$\tau \equiv \frac{1}{\mathrm{Im}[\omega]}, \tag{6.45}$$

where $\mathrm{Im}[\omega]$ is the imaginary part of the frequency. Thus, for instance, for the Rayleigh-Taylor instability we have

$$\tau = \sqrt{\frac{\rho' + \rho}{gk(\rho' - \rho)}}, \tag{6.46}$$

from which we see, for instance, that smaller perturbations (larger k) grow more quickly.

Until now, we have totally neglected the first term in eq. 6.39 given that it is always real. However, whenever $u_0 \neq u_0'$ this term produces an oscillation and the instability will have to develop *on top* of this oscillation. We can write down explicitly the values of the real and imaginary parts of ω in the case of no gravity ($g = 0$) and setting for simplicity $u_0' = u_0 + \Delta u$ (with Δu¿0) to obtain

$$\mathrm{Re}[\omega] = \left(\frac{\rho u_0 + \rho'(u_0 + \Delta u)}{\rho + \rho'}\right) k \quad \text{and} \quad \mathrm{Im}[\omega] = \left(\frac{\sqrt{\rho\rho'}\Delta u}{\rho + \rho'}\right) k, \tag{6.47}$$

which are quite similar to each other in particular if $\rho \sim \rho'$ and if $\Delta u \gg u_0$. Thus, in these conditions one has oscillations with exponentially growing amplitude, a condition that is sometime referred to as *overstability* and likely eventually leads to instability with a timescale not too different from the instability timescale.

6.3. Gravitational instability

In this Section, we derive the dispersion relation of one of the most important hydrodynamical instabilities that occurs in astrophysical fluids, which plays a relevant role in the formation of structures (stars, galaxies, clusters) in the Universe. We follow the derivation presented by Jeans at the beginning of the last century [Jea02]. We consider an homogeneous medium at constant pressure (P_0), density (ρ_0) and with a constant gravitational potential (Φ_0). We assume that the medium is initially stationary ($\boldsymbol{u}_0 = 0$) and consider infinitesimal perturbations of all these quantities for the generic fluid element such that

$$\begin{aligned} P &= P_0 + \Delta P \\ \rho &= \rho_0 + \Delta\rho \\ \boldsymbol{u} &= \Delta\boldsymbol{u}\Phi \qquad\qquad = \Phi_0 + \Delta\Phi. \end{aligned} \tag{6.48}$$

As in the case of sound waves (§2.10) the above perturbations are Lagrangian because they refer to the fluid element.

We substitute the perturbations in eqs. 6.48 into the continuity (eq. 2.10), the Euler (eq. 2.16) and the Poisson equations to obtain

$$\begin{cases} \dfrac{\partial\Delta\rho}{\partial t} + \rho_0 \boldsymbol{\nabla}\cdot\Delta\boldsymbol{u} = 0 \\ \dfrac{\partial\Delta\boldsymbol{u}}{\partial t} = -\dfrac{1}{\rho_0}\boldsymbol{\nabla}\Delta P - \boldsymbol{\nabla}\Delta\Phi \\ \nabla^2\Delta\Phi = 4\pi G\Delta\rho, \end{cases} \tag{6.49}$$

where we have neglected all the second order terms[9]. Let us write the pressure perturbation as $\Delta P = c_s^2 \Delta\rho$, where c_s is the sound speed, and take the time derivative of the first equation in 6.49 and the divergence of the second. Substituting the Laplacian of $\Delta\Phi$ in the second equation with the r.h.s. of the Poisson equation, neglecting again second order terms, one obtains

$$\begin{cases} \dfrac{\partial^2 \Delta\rho}{\partial t^2} + \rho_0 \boldsymbol{\nabla} \cdot \dfrac{\partial \Delta \boldsymbol{u}}{\partial t} = 0 \\ \rho_0 \boldsymbol{\nabla} \cdot \dfrac{\partial \Delta \boldsymbol{u}}{\partial t} = -c_s^2 \nabla^2 \Delta\rho - 4\pi G \rho_0 \Delta\rho. \end{cases} \tag{6.50}$$

Finally, by subtracting the second from the first equation in 6.50 we find an expression with only density perturbations:

$$\frac{\partial^2 \Delta\rho}{\partial t^2} = c_s^2 \nabla^2 \Delta\rho + 4\pi G \rho_0 \Delta\rho. \tag{6.51}$$

As usual, the solutions of eq. 6.51 can be expressed as a summation of solutions (Fourier modes) of the type

$$\Delta\rho = \widetilde{\Delta\rho} e^{i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t)}, \tag{6.52}$$

where $\boldsymbol{k}$ and ω are the wavenumber vector and the frequency of the oscillation. Thanks to the superposition principle, eq. 6.52 is a solution of eq. 6.51 and we can study the behaviour of the perturbation simply by using this solution. Substituting it then in eq. 6.51 one obtains the following dispersion relation:

$$\omega^2 = k^2 c_s^2 - 4\pi G \rho_0, \tag{6.53}$$

where $k = |\boldsymbol{k}|$. The condition for instability is that ω be imaginary, i.e. that the r.h.s. of eq. 6.53 is negative. We substitute the wavenumber with the wavelength ($\lambda = 2\pi/k$) of the perturbation to find that the condition for **gravitational instability** is

$$\lambda > \lambda_J \equiv \sqrt{\frac{\pi c_s^2}{G\rho}}, \tag{6.54}$$

where we have dropped the subscript $_0$ in the density to make the criterion generically applicable to a medium with density ρ. The r.h.s. of eq. 6.54 is called **Jeans length** and the criterion for gravitational instability states that in a medium of density ρ and sound speed c_s, if structures exist with sizes (much) larger than λ_J they are subject to their own self gravity. This latter will overcome the reaction of the internal pressure and these structures should undergo gravitational collapse. On the contrary, structures with sizes of the order of the Jeans length will settle in hydrostatic equilibrium with their own gravity, while in those much smaller than λ_J, the self gravity can be neglected. In this latter case, the survival of such structures is possible only if there is an external medium that confines them (they are called **pressure-confined** clouds), otherwise they would tend to disperse.

[9]Note that the unperturbed part of these equations leads to an unphysical conclusion: the density of the medium should be null. The neglect of this contradiction is the famous *swindle* that Jeans employed to carry out the calculations. Remarkably, despite this apparent incongruity, the perturbed equations lead to a correct solution. A way to overcome this issue would be to consider a medium with a slight density/pressure gradient to match the gravity field through the condition of hydrostatic equilibrium (Chap. 3).

The typical timescale for (unimpeded) gravitational collapse is the so-called **free-fall time**. This can be estimated using Newton second law for a spherical system:

$$\frac{\mathrm{d}^2 r}{\mathrm{d}t^2} = G\frac{M(r)}{r^2} \tag{6.55}$$

where $M(r)$ is the mass within the radius r. If we consider a homogeneous sphere of density ρ the free-fall time turns out to be

$$t_{\rm ff} = \sqrt{\frac{3\pi}{32G\rho}} \simeq 0.5\sqrt{\frac{1}{G\rho}}. \tag{6.56}$$

It is then interesting to note that the Jeans length can be expressed in terms of the free-fall time as

$$\lambda_{\rm J} \approx c_{\rm s}\, t_{\rm ff}. \tag{6.57}$$

To understand the relevance of this latter equation we remind that, for any fluid, the crossing time of a certain portion of the fluid is its size divided by the sound speed (eq. 2.65). Thus the Jeans criterion for gravitational collapse can be rewritten in terms of timescales as

$$t_{\rm cross} > t_{\rm ff}, \tag{6.58}$$

which effectively states that when the free-fall time is shorter that the time the fluid pressure takes to react to perturbations (crossing time) the collapse cannot be halted.

The Jeans criterion is often expressed in terms of mass instead of length. In order to do this, one only needs to write the mass of a structure of constant density ρ and size of the Jeans length:

$$\begin{aligned} M_{\rm J} \equiv \frac{4\pi}{3}\rho\left(\frac{\lambda_{\rm J}}{2}\right)^3 = \frac{\pi}{6}\left(\frac{\pi}{G}\right)^{3/2}\frac{c_{\rm s}^3}{\rho^{1/2}} = \frac{\pi}{6}\left(\frac{\pi k_{\rm B}}{\mu m_{\rm p} G}\right)^{3/2}\frac{T^{3/2}}{\rho^{1/2}} \\ \simeq 29.5\left(\frac{\mu}{2.3}\right)^{-2}\left(\frac{n}{10^2\,\mathrm{cm}^{-3}}\right)^{-1/2}\left(\frac{T}{10\,\mathrm{K}}\right)^{3/2}\mathrm{M}_\odot, \end{aligned} \tag{6.59}$$

where, for the second equality of eqs. 6.59, we have used the relation between sound speed and temperature for isothermal gases (eq. 2.67). Equation 6.59 is called **Jeans mass** and the criterion for gravitational instability in this case will be

$$M > M_{\rm J}. \tag{6.60}$$

In general terms, we can thus expect that, if a gaseous structure has a mass larger than its Jeans mass, self gravity will be important and it will collapse rapidly.

The value that we obtained in 6.59 of a few tens of Solar masses is typical for densities and temperatures of the molecular gas in the Milky Way (§1.1.4). The observed masses of Galactic giant molecular clouds (GMCs) are, instead, of order $10^5 - 10^6\,\mathrm{M}_\odot$, thus vastly exceeding the Jeans mass. This would imply that these structures cannot be in hydrostatic equilibrium and should collapse to form stars. The latter is, to some extent, true, as molecular clouds are the sites of star formation in galaxies. However, we do not observe them to collapse as fast as in a free-fall time. Instead, there seem to be other forces that counteract gravity in GMCs other than the (thermal) pressure force. The most obvious candidates are magnetic fields (magnetic pressure) and, even more likely, turbulence (turbulent pressure).

Indeed, GMCs are observed to be very turbulent, i.e. their observed typical random speeds (of order of $1\,\mathrm{km\,s^{-1}}$ as measured from the widths of line profiles) largely exceed the sound speed expected at their typical temperature $T \sim 10\,\mathrm{K}$, which is $\sim 0.1\,\mathrm{km\,s^{-1}}$. The Jeans mass is, however, of the order of the mass of the smallest structures inside GMCs called dense cores (§1.1.4) that are extremely dense clouds often hosting young stellar objects at their interior.

6.4. Rotational instability

Consider a rotating fluid in an axisymmetric potential. For simplicity, let us assume that the fluid has collapsed in a infinitely thin layer, roughly true for galaxy discs and that its density is constant with radius. Furthermore let us take the fluid as incompressible ($\boldsymbol{\nabla}\cdot\boldsymbol{u}=0$). A fluid element in this disc will obey the equation of motion given by the Euler equation in Lagrangian form (eq. 2.15), which can be written in cylindrical coordinates as

$$\begin{cases} \ddot{R} - R\Omega^2 = -\dfrac{1}{\rho}\dfrac{\mathrm{d}P}{\mathrm{d}R} - \dfrac{\mathrm{d}\Phi}{\mathrm{d}R} \\ 2\dot{R}\Omega + R\dot{\Omega} = 0, \end{cases} \tag{6.61}$$

where Ω is the angular speed ($\Omega \equiv \dot{\varphi}$) and the second equation simply states the conservation of the specific angular momentum:

$$\frac{1}{R}\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{J}) = \frac{1}{R}\frac{\mathrm{d}}{\mathrm{d}t}(R^2\Omega) = \frac{1}{R}(2R\dot{R}\Omega + R^2\dot{\Omega}) = 0. \tag{6.62}$$

Note that $\mathcal{J} \equiv R^2\Omega$ is the total specific angular momentum for the assumed two-dimensional problem or it would be the z-component of the angular momentum in a tree-dimensional setup.

We can rewrite the first of eqs. 6.61 by substituting the specific angular momentum $\mathcal{J}$ and isolating the radial acceleration to obtain

$$\ddot{R} = \frac{\mathcal{J}^2}{R^3} - \frac{1}{\rho}\frac{\mathrm{d}P}{\mathrm{d}R} - \frac{\mathrm{d}\Phi}{\mathrm{d}R} = 0, \tag{6.63}$$

which is null for any fluid element in the unperturbed disc because they are in pure rotation and their centrifugal and pressure forces balance the gravitational attraction, i.e. there is not net radial acceleration.

Consider now a radial perturbation that moves material from an unperturbed position R to $R+\Delta R$, with ΔR positive. For simplicity, we can consider this perturbation axi-symmetric and make the assumption that it takes place by conserving angular momentum. We ask ourselves under what conditions such a perturbation can develop an instability. The radial displacement of the gas brings it to a new radius and leads to the following new expression for the (perturbed) radial acceleration:

$$\ddot{R}_{\mathrm{pert}} = \frac{\mathcal{J}(R)^2}{(R+\Delta R)^3} - \frac{1}{\rho}\left(\frac{\mathrm{d}P}{\mathrm{d}R}\right)_{(R+\Delta R)} - \left(\frac{\mathrm{d}\Phi}{\mathrm{d}R}\right)_{(R+\Delta R)}, \tag{6.64}$$

where we have imposed the conservation of angular momentum: $\mathcal{J}(R)$ has not changed. The "normal" material (in rotational balance) at radius $R+\Delta R$ will follow the equation

$$\frac{\mathcal{J}(R+\Delta R)^2}{(R+\Delta R)^3} - \frac{1}{\rho}\left(\frac{\mathrm{d}P}{\mathrm{d}R}\right)_{(R+\Delta R)} - \left(\frac{\mathrm{d}\Phi}{\mathrm{d}R}\right)_{(R+\Delta R)} = 0, \tag{6.65}$$

which is analogous to eq. 6.63 but for $R+\Delta R$. By subtracting the last two equations we obtain the radial acceleration acting on the perturbed gas that moved from R to $R+\Delta R$:

$$\ddot{R}_{\rm pert} = \frac{\mathcal{J}(R)^2}{(R+\Delta R)^3} - \frac{\mathcal{J}(R+\Delta R)^2}{(R+\Delta R)^3}. \tag{6.66}$$

Given that in the above radial perturbation the gas has been moved outwards we can expect rotational stability if the acceleration of eq. 6.66 is negative. In this case, indeed, the displaced gas will go back to the original radius and oscillate around its initial position. Equation 6.66 can be rewritten as

$$\ddot{R}_{\rm pert} = \frac{\mathcal{J}(R)^2 - \mathcal{J}(R+\Delta R)^2}{(R+\Delta R)^3} = -\frac{\mathcal{J}(R+\Delta R)^2 - \mathcal{J}(R)^2}{\Delta R}\frac{\Delta R}{(R+\Delta R)^3}, \tag{6.67}$$

where we have simply multiplied and subtracted by ΔR (positive) and the second fraction of the last term is also positive. We can now take ΔR infinitely small and the first fraction of the last term becomes a derivative of the square of the specific angular momentum in R. In conclusion, the criterion for **rotational stability** becomes simply

$$\frac{{\rm d}\mathcal{J}^2}{{\rm d}R} = \frac{{\rm d}(R^2\Omega)^2}{{\rm d}R} > 0, \tag{6.68}$$

which is called **Rayleigh criterion** and it is verified essentially in all astrophysical discs. The reader can consider examples like a typical galaxy rotation curve where either $\Omega \sim$ constant (rising part) or $\Omega \propto 1/R$ (flat part). In both cases, the criterion in eq. 6.68 is obviously satisfied. Moreover, even when the velocity displays a Keplerian fall ($\Omega \propto R^{-3/2}$), typical, for instance, of accretion and protoplanetary discs, we will have $\mathcal{J}^2 \propto R$ and thus stability.

The rotation instability criterion can also be seen in terms of the so-called **epicycle frequency**. This is the frequency that a test particle has if perturbed from its circular motion and it reads

$$\kappa^2 \equiv \frac{1}{R^3}\frac{\rm d}{{\rm d}R}(R^2\Omega)^2 = R\frac{{\rm d}\Omega^2}{{\rm d}R} + 4\Omega^2. \tag{6.69}$$

For real κ, a perturbation of the circular motion of a particle will lead to oscillations around the path of the circular orbit and to a rosetta-type of orbit (but this is only true for collision-less systems). An imaginary κ would, instead, result in an instability. Note that the epicycle frequency is defined for a collision-less system but the criterion that we found is exactly the same. This is due to the fact that the pressure term cancels out in our derivation because we are not considering hydrodynamical effects like the compression of the medium at $R+\Delta R$ but only the conservation of angular momentum.

The Rayleigh criterion is perfectly justified for ideal fluids but in real fluids other effects may play important roles. One of these is viscosity that can stabilise discs also when they are potentially unstable (see Chapter 7). The other important ingredient of astrophysical fluids that can dramatically change this problem is the magnetic field. It can be shown that, if a *weak*[10] magnetic field along z is present, then the criterion for stability becomes

$$\frac{{\rm d}\Omega^2}{{\rm d}R} > 0, \tag{6.70}$$

[10]In fact, in the presence of a strong magnetic field the instability tends to disappear.

which makes both galaxy discs (in the region with constant rotation) and accretion discs unstable. This instability is called **magneto-rotational instability**. It was originally proposed by Chandrasekhar [Cha60] and subsequently "re-discovered" by Balbus & Hawley [BH91]. If gaseous discs are prone to develop this instability, the various rings of rotating gas can mix with each other. This, in real conditions, may mean the development of *turbulence* and indeed magneto-rotational instability is advocated as a possible cause of turbulence in accretion discs and, potentially, in the outer parts of spiral galaxies. This turbulence, especially in accretion discs, is then often considered to be the cause of a form of viscosity called **anomalous viscosity**. This is different from the molecular/atomic viscosity that we see in Chap. 7 as the "particles" that exchange momentum between different parts of the fluid are *macroscopic* (portions of fluid). Despite no clear proof of the existence of anomalous viscosity in accretion discs has been found, this idea is widely accepted in the community. The main reason is heuristic: indeed without this macroscopic form of viscosity, accretion disc would not fulfill their primary task of channeling low-angular momentum material towards the central regions and feed the inner black hole.

6.5. Thermal instability

One of the most relevant instabilities in the ISM of galaxies and in general in astrophysical fluids is the **thermal instability**. This is a condition that can occur when a portion of fluid is subject to a change in its entropy content as a consequence of radiative cooling and/or heating. To derive a criterion for this instability let us consider a fluid element initially at *thermal* equilibrium, i.e. in a condition such that cooling and heating are perfectly balanced[11]. We use the heat equation (eq. 2.51) written for the fluid element:

$$\frac{\mathrm{D}\mathcal{S}}{\mathrm{D}t} = -\frac{\mathcal{L}(\mathcal{S}, A)}{T}, \tag{6.71}$$

where $\mathcal{L} = \Lambda - \Gamma$ is the energy loss per unit time and mass due to radiation and it is null at equilibrium: $\mathcal{L}_0 = \mathcal{L}(\mathcal{S}_0, A_0) = 0$. Note that in eq. 6.71, we have taken $\mathcal{L}$ dependent on the specific entropy and on another thermodynamical variable A that we are not specifying for the moment.

Let us consider a perturbation that increases the specific entropy by a small amount:

$$\mathcal{S} = \mathcal{S}_0 + \Delta\mathcal{S} \tag{6.72}$$

and substitute it into eq. 6.71. On the l.h.s., only the derivative of the perturbed part remains, whereas we can expand the r.h.s. as a Taylor series around the region of equilibrium entropy $\mathcal{S}_0$. This leads to:

$$\frac{\mathrm{D}\Delta\mathcal{S}}{\mathrm{D}t} = -\frac{\mathcal{L}(\mathcal{S}_0, A_0)}{T_0} - \left[\frac{\partial}{\partial\mathcal{S}}\frac{\mathcal{L}(\mathcal{S}, A)}{T}\right]_{A,\,\mathcal{S}=\mathcal{S}_0} \Delta\mathcal{S} + ..., \tag{6.73}$$

neglecting higher order terms. The first term on the r.h.s. of eq. 6.73 is null by construction ($\mathcal{L}_0 = 0$) and the partial derivative is calculated at constant A and

[11] Note that in §1.3 we have defined as a system at thermal equilibrium one at constant temperature, the equilibrium between cooling and heating provides the physical explanation of how such a situation can take place.

evaluated at $\mathcal{S} = \mathcal{S}_0$. We bring $\Delta\mathcal{S}$ in eq. 6.73 to the l.h.s. and we should obviously have

$$\frac{1}{\Delta\mathcal{S}}\frac{\mathrm{D}\Delta\mathcal{S}}{\mathrm{D}t} = \frac{\mathrm{D}\ln\Delta\mathcal{S}}{\mathrm{D}t} < 0 \tag{6.74}$$

for stability, i.e. if the entropy variation of the fluid element, assumed positive, decreases with time then the entropy tends to go back to its unperturbed value.

Using equation 6.71, the above prescription for the entropy becomes the condition that the partial derivative on the r.h.s. should be larger than zero. We can develop this derivative as follows:

$$\left[\frac{\partial}{\partial\mathcal{S}}\frac{\mathcal{L}(\mathcal{S},A)}{T}\right]_{A,\,\mathcal{S}=\mathcal{S}_0} = \frac{1}{T_0}\left[\frac{\partial\mathcal{L}(\mathcal{S},A)}{\partial\mathcal{S}}\right]_{A,\,\mathcal{S}=\mathcal{S}_0} + \mathcal{L}_0\left[\frac{\partial}{\partial\mathcal{S}}\left(\frac{1}{T}\right)\right]_{A,\,\mathcal{S}=\mathcal{S}_0}, \tag{6.75}$$

where the second term on the r.h.s. is obviously null. This leads to a simple criterion for the **thermal stability** of the fluid:

$$\left(\frac{\partial\mathcal{L}}{\partial\mathcal{S}}\right)_A > 0, \tag{6.76}$$

where we have omitted the redundant equilibrium condition.

We are now in the position to render A explicit. Let us first take $A = P$ (the fluid pressure), eq. 6.76 can be rewritten as

$$\left(\frac{\partial\mathcal{L}}{\partial\mathcal{S}}\right)_P = \left(\frac{\partial\mathcal{L}}{\partial T}\right)_P\left(\frac{\partial T}{\partial\mathcal{S}}\right)_P = \frac{T}{C_\mathrm{p}}\left(\frac{\partial\mathcal{L}}{\partial T}\right)_P, \tag{6.77}$$

where we have used the definition of specific heat at constant pressure (D.1.4). Given that C_p and T are always positive, we obtain the final expression

$$\left(\frac{\partial\mathcal{L}}{\partial T}\right)_P > 0, \tag{6.78}$$

which is the *isobaric* **Field criterion** for thermal stability [Fie65]. In analogy to the above, taking $A = \rho$ and using the definition of heat capacity at constant volume we obtain

$$\left(\frac{\partial\mathcal{L}}{\partial T}\right)_\rho > 0, \tag{6.79}$$

which is the *isochoric* **Parker criterion** [Par53]. In general, we have seen that astrophysical gases tend to settle roughly in pressure equilibrium (§1.3) so the isobaric Field criterion is more often used. Moreover, the condition of the Parker criterion is rather difficult to violate in astrophysical fluids.

The Field criterion (eq. 6.78) quantitatively establishes whether a region in the ISM is thermally stable or not. The solid curve in Fig. 2 shows the locus of points at thermal (heating vs cooling) equilibrium ($\mathcal{L} = 0$) in the $\log T - \log\rho$ plane for the Galactic ISM. The dashed line shows the loci of constant pressure in the same plane, this is the mean pressure of the ISM and we remind that nearly all gaseous structures have this pressure (see §1.3). The two curves intersect in three points that are obvious points of equilibrium being at both pressure and thermal equilibrium. However, not all these points are of *stable* equilibrium. If a portion of the fluid finds itself in F or H and its temperature increases for some reason, it would move in this plot up along the dashed line (assuming that it remains in pressure equilibrium with the surroundings). This shift will bring it to a region where $\mathcal{L} > 0$ thus, for the Field criterion (eq. 6.78), it should be stable (it increases its $\mathcal{L}$ for increasing

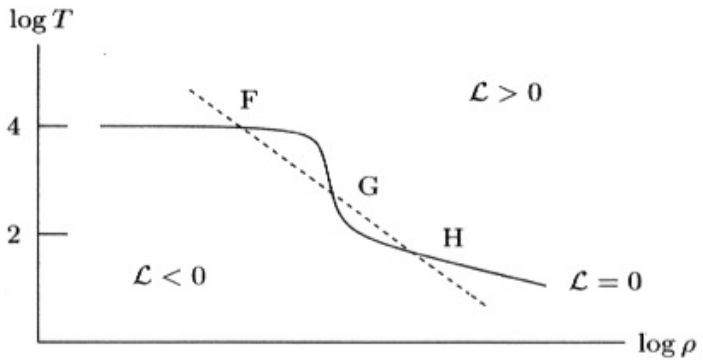

FIGURE 2. Equilibrium curves in the $\log T - \log \rho$ plane. The solid curve is the locus where there is equilibrium between cooling and heating ($\mathcal{L} = 0$), the dashed line shows points at constant pressure. The two curves intersect at three equilibrium points (see text). From [Shu92].

T). The fluid element would then get back to its previous temperature and will, at most, oscillate around that value. The same would happens if its temperature decreases.

Consider, instead, the case of a portion of the fluid starting from G, increasing temperature here (at constant pressure) means moving to a region where $\mathcal{L} < 0$. Thus G is an unstable point (for eq. 6.78) and all fluid elements that end up at these temperatures (and densities) should move away from it likely along the constant pressure curve. They will eventually reach F or H depending on whether, respectively, heating or cooling is dominating the instability. In other words, thermal instability causes the ISM to settle in well defined discrete phases, two in the example that we are describing here but, more generally, three [MO77] if we include also the hot medium at $T \sim 10^6$ K (see §1.1 for a description of these phases in the Galactic ISM). Gas at intermediate temperatures should not exist in normal conditions as it is *thermally unstable*. For cooling dominating over heating, the timescale of thermal instability is, essentially, the cooling time (§2.8.6).

In realistic situations, intermediate temperatures are actually present in regions of *interfaces* between different phases of the ISM, where hotter and colder phases are in contact. In these interface layers, thermal conduction (§2.8.1) plays an important role and gas at these "forbidden" temperatures can be present. In the halo of our Galaxy (and also around external galaxies) large amounts of gas at temperatures $\sim 10^5$ K has been detected via, in particular, the O VI absorption lines [SWS+03]. This gas is likely at the interface between cold-warm gas (the so-called **high-velocity clouds**) at temperatures $T \lesssim 10^4$ K and the hot gas of the Galactic corona at $T \sim 10^6$ K and it can survive at these intermediate temperatures as long as the other (stable) phases are present and thermal conduction and/or turbulent mixing operate efficiently.

We now look for a criterion for thermal instability that involves pressure and density instead of $\mathcal{L}$ and T, being $\mathcal{L}$ quite difficult to measure in practice. In order to do this, we take three thermodynamic variables and consider them to be functions of one another. We start from the following: $\mathcal{L} = \mathcal{L}(T, P)$, $T = T(P, \mathcal{L})$ and $P = P(\mathcal{L}, T)$ and employ a well-known property of three continuous function

with defined derivatives (see C.3.1) to obtain

$$\left(\frac{\partial \mathcal{L}}{\partial T}\right)_{\mathrm{P}} \left(\frac{\partial T}{\partial P}\right)_{\mathcal{L}} \left(\frac{\partial P}{\partial \mathcal{L}}\right)_{\mathrm{T}} = -1 \tag{6.80}$$

Using the Field criterion (eq. 6.78) for the first term in eq. 6.80 leads us to the following condition for stability:

$$\left(\frac{\partial T}{\partial P}\right)_{\mathcal{L}} \left(\frac{\partial P}{\partial \mathcal{L}}\right)_{\mathrm{T}} < 0 \tag{6.81}$$

This can be developed into

$$\left(\frac{\partial T}{\partial \rho}\right)_{\mathcal{L}} \left(\frac{\partial \rho}{\partial P}\right)_{\mathcal{L}} \left(\frac{\partial P}{\partial \rho}\right)_{\mathrm{T}} \left(\frac{\partial \rho}{\partial \mathcal{L}}\right)_{\mathrm{T}} < 0. \tag{6.82}$$

We can then write the analogous of eq. 6.80 for density $\rho = \rho(\mathcal{L}, T)$ instead of pressure and use the Parker criterion (eq. 6.78) to obtain

$$\left(\frac{\partial T}{\partial \rho}\right)_{\mathcal{L}} \left(\frac{\partial \rho}{\partial \mathcal{L}}\right)_{\mathrm{T}} < 0 \tag{6.83}$$

for stability. Finally, we use the condition in eq. 6.83 to simplify eq. 6.82 and obtain the following inequality[12]

$$\left(\frac{\partial \rho}{\partial P}\right)_{\mathcal{L}} \left(\frac{\partial P}{\partial \rho}\right)_{\mathrm{T}} > 0, \tag{6.84}$$

whose second term is simply P/ρ leading to the final condition for thermal stability:

$$\left(\frac{\partial \ln P}{\partial \ln \rho}\right)_{\mathcal{L}} > 0. \tag{6.85}$$

Thus, if we can define the locus of points where $\mathcal{L} = 0$ in the (ρ, P) plane we can easily separate between stable and unstable thermal equilibria.

Fig. 3 shows the above locus for neutral atomic gas with column densities from $N = 10^{20}\,\mathrm{cm}^{-2}$ (typical of the ISM of galaxy discs) to $N = 10^{18}\,\mathrm{cm}^{-2}$ (typical of regions outside the discs). If we consider a constant pressure of the ISM (a horizontal line at about $\log\left(\langle P/k_{\mathrm{B}}\rangle/\,\mathrm{K\,cm}^{-3}\right) \approx 3.3 - 3.4$) we find again three points of equilibrium. The left and right points are characterised by a derivative of $\ln P$ over $\ln \rho$ which is clearly positive and thus they are stable points. These two stable equilibrium configurations correspond to the two ISM phases that we called WNM and CNM in §1.1.1. In the middle point, the derivative is negative and the equilibrium is unstable. A fluid parcel that ends up there is prone to thermal instabilities. For instance, if cooling dominates, the fluid parcel would decrease its temperature (increase its density) and reach the rightmost equilibrium point (CNM). This is the basis of the **two-phase model** [WHM+95] for the galactic ISM and shows how thermal instability is responsible for the existence of these phases.

[12]Note that we could have reached the same conclusion by noting that i) the first derivative in eq. 6.82 is negative as one can see (at least for $\mathcal{L} = 0$) from Fig. 2 and ii) the last derivative in the same equation is positive for any astrophysical cooling function, i.e. the cooling(/heating) rate increases with the density.

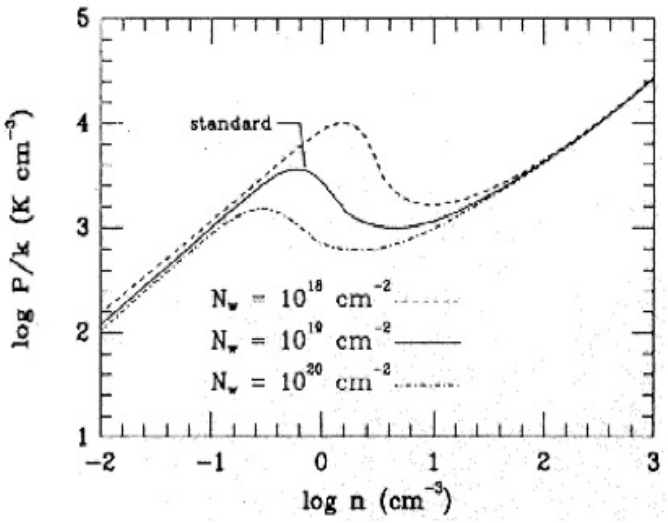

FIGURE 3. Locus of points at equilibrium between cooling and heating ($\mathcal{L} = 0$) in the $\log n - \log P$ plane for different column densities of the gas. From [WHM$^+$95].

CHAPTER 7

Viscosity

In this Chapter we make a step forward towards a more realistic description of fluid flows. We have seen that the interactions between particles that constitute the fluid have the fundamental importance of bringing the different constituents at the same temperature and make sure that their speed are well described by a Maxwellian distribution (§1.2.2). However, if two regions of the fluid with different properties are in contact, particles can also exchange energy and momentum with the adjacent part of the fluid. We have encountered a phenomenon of this kind when we have described the thermal conduction (§2.8.1). In that case, the two portions of the fluid have different temperatures, i.e. their elementary particles are moving at different typical (random) velocities at the microscopic level. Collisions at the interface produce a net exchange of heat from the hotter to the colder region and we have treated this phenomenon as a source term in the heat/energy equation (eq. 2.36).

We now consider another situation in which interactions between particles (atoms, electrons or molecules) produce important modifications in the fluid. This means releasing the assumption that the fluids that we are dealing with are ideal. In this new type of fluids, the interactions at a particle level can produce **stresses** (forces for unit area) that deform the flow and dissipate kinetic energy, a process akin to an internal *friction*. In the following sections, we see how these microscopic interactions can produce stresses of a kind very different from that of the thermal pressure considered until now, as they can act *orthogonally* to the fluid motion. All these phenomena go under the generic name of **viscosity** and the fluids (or the flows) are called **viscous** as opposed to the **inviscid** ones, analysed in the previous chapters. In order to treat the problem mathematically it is convenient to introduce a new notation for vectors in the fluid equations.

7.1. Notation by components

We use a notation by components to rewrite the relevant conservation equations found in Chap. 2. With this notation the gradient of a scalar, the divergence and the curl of a vector are written as

$$\begin{aligned} (\boldsymbol{\nabla} q)_i &\longrightarrow \frac{\partial q}{\partial x_i} \\ \boldsymbol{\nabla}\cdot\boldsymbol{a} &\longrightarrow \frac{\partial a_i}{\partial x_i} \\ (\boldsymbol{\nabla}\times\boldsymbol{a})_i &\longrightarrow \epsilon_{ijk}\frac{\partial a_k}{\partial x_j}, \end{aligned} \tag{7.1}$$

where we are writing, in the first and the third equation, only the i-th component and ϵ_{ijk} in the third equation is the Levi-Civita symbol (§C.4.5). We remind that,

in this notation, repeated indexes correspond to summations that are omitted for brevity, thus, for instance, $\partial a_i/\partial x_i$ means $\Sigma_{i=1}^{3}(\partial a_i/\partial x_i)$. With the same notation the advection operator (i-th component) acting on a vector is written as

$$(\boldsymbol{a}\cdot\boldsymbol{\nabla}\boldsymbol{a})_i \longrightarrow a_j\frac{\partial a_i}{\partial x_j}, \tag{7.2}$$

where it is now clear that the term $\partial a_i/\partial x_j$ (gradient of a vector) is a 2-dimensional tensor (see §2.2).

Using this component notation the continuity equation (eq. 2.10) becomes

$$\frac{\partial\rho}{\partial t}+\frac{\partial}{\partial x_i}(\rho u_i)=0, \tag{7.3}$$

while the i-th component of the Euler (force) equation (eq. 2.16) reads

$$\frac{\partial u_i}{\partial t}+u_j\frac{\partial u_i}{\partial x_j}=-\frac{1}{\rho}\frac{\partial P}{\partial x_i}-\frac{\partial\phi}{\partial x_i}. \tag{7.4}$$

Note the summation in j in the advection term (second on the l.h.s.) that returns the i-th component.

7.2. Momentum equation

We look for an equation of momentum conservation by developing the Eulerian time derivative of the momentum density as follows

$$\frac{\partial\rho u_i}{\partial t}=\rho\frac{\partial u_i}{\partial t}+u_i\frac{\partial\rho}{\partial t}. \tag{7.5}$$

We then substitute the continuity equation (7.3) and the Euler equation (7.4) to the first and second term of the r.h.s. Neglecting the potential term for simplicity, we obtain

$$\frac{\partial\rho u_i}{\partial t}=-\frac{\partial\Pi_{ij}}{\partial x_j}, \tag{7.6}$$

where

$$\Pi_{ij}\equiv\rho u_i u_j+P\delta_{ij} \tag{7.7}$$

is the **tensor of momentum flux density**. Equation 7.6 is a form of the Euler equation that describes the conservation of momentum. This can be easily seen if we perform the usual integration over a generic volume V (see e.g. §2.7) and apply the divergence theorem (eq. C.19), an operation that leads to

$$\frac{\partial}{\partial t}\int_V(\rho u_i)\mathrm{d}V=-\int_S(\rho u_i)u_j\mathrm{d}S_j-\int_S P\mathrm{d}S_i. \tag{7.8}$$

The l.h.s. of eq. 7.8 is the variation of $i-$th component of the momentum within our volume. This equals the amount of momentum ($i-$th component) that is advectively removed (or added) to the volume by the motion of the fluid (first term of the r.h.s.) plus the pressure force along i (second term). Note that, had we not excluded gravity, there would be here an additional term for the gravitational force.

So far, the fluid has been considered ideal and the variation of the momentum is only related to *standard* forces. The step that we take now is to extend our treatment to non-ideal fluids where forces other than pressure (and gravity) are at work. To understand the type of forces involved, consider a fluid where different portions in contact with each other move at different velocities (i-th component),

e.g. there is a gradient of u_i along a certain direction j, with $j \neq i$. Consider also that the microscopic interactions, due to the random motions of particles within the fluid, are not negligible. These motions will transfer momentum from the fast moving parts to the slow moving parts of the fluid. This produces a form of friction in the fluid that we call **shear viscosity**. As a consequence, the momentum of the fluid (say within the generic volume V) can now change also due to the effect of viscosity and we have to add a new term in eq. 7.8 that will represent these viscous forces.

We rewrite the momentum flux tensor as

$$\Pi_{ij} = \rho u_i u_j - \sigma_{ij}, \tag{7.9}$$

where we have defined the **stress tensor** as

$$\sigma_{ij} \equiv -P\delta_{ij} + \sigma'_{ij} \tag{7.10}$$

i.e. the sum of the pressure and the viscous stresses, these latter represented by σ'_{ij}: the **viscous stress tensor**. Note that the new stresses are not anymore acting only normally to the surfaces, i.e. they are non-zero *also* for $i \neq j$. The new generic force (i-th component) can be written as

$$\mathrm{d}F_i = \sigma_{ij}\mathrm{d}S_j, \tag{7.11}$$

from which we can see that σ_{ij} describes the i-th component of the force over an element of unit area with normal along the direction j. In the presence of pressure alone (ideal fluid) we must have $i = j$ (δ_{ij} is a diagonal tensor). In the case of viscous forces, the components depend on the form of the σ'_{ij} tensor, which we can expect to be non-diagonal. As a convention, the pressure force is positive if the fluid element is acting against the outside medium (see eq. 2.14) while the viscous stress tensor is positive if the element is subject to an external force. The reason for the pressure sign has been discussed in §2.5, the one of the stress tensor will become clear in the following.

Before searching for a form of the viscous stress tensor we rewrite the Euler equation with viscosity. Take eq. 7.6 with the inclusion of the viscous forces, expand the derivation on the l.h.s. and use the continuity equation to remove the time derivative of the density. We obtain

$$\rho\frac{\partial u_i}{\partial t} + \rho u_j\frac{\partial u_i}{\partial x_j} = -\frac{\partial P}{\partial x_i} + \frac{\partial \sigma'_{ij}}{\partial x_j}, \tag{7.12}$$

where again we have neglected gravity. We can also express this equation in a more compact Lagrangian form as

$$\rho\frac{\mathrm{D}u_i}{\mathrm{D}t} = \frac{\partial \sigma_{ij}}{\partial x_j}. \tag{7.13}$$

7.3. Viscous stress tensor

We now proceed to search for a mathematical form of the viscous stress tensor. The following is not a proper derivation of the tensor, which pertains to a more advanced text (for more details see [Pat84]). We instead content ourselves with a *justification* of its form based on a number of properties that this tensor should fulfill. The first of these properties is symmetry.

7.3.1. Symmetry of the stress tensor. We show that the stress tensor has to be symmetric by proving the physical incongruence of the opposite assumption. Let us postulate that $\sigma'_{ij} \neq \sigma'_{ji}$ and thus that a fluid element is subject to different forces along i acting across $\mathrm{d}S_j$ and along j acting across $\mathrm{d}S_i$. Being i orthogonal to j these different forces must produce a rotation of the fluid element. However, this is impossible because the fluid element is an abstract concept: we can take it as small as possible to the limit of infinitely small and thus cannot have a net rotation. This falsifies the hypothesis and justifies the assumption of symmetry of the stress tensor.

We can express the above mathematically as follows. If a portion of the fluid has a net rotation this would result in a torque expressed as

$$\boldsymbol{\tau} = I\frac{\mathrm{d}\boldsymbol{\omega}}{\mathrm{d}t}, \tag{7.14}$$

where I is the moment of inertia of the fluid portion and $\boldsymbol{\omega}$ is its angular velocity[1]. Now, I depends on the size (Δx) of the portion of fluid considered and if we reduce the size to make it infinitely small, I goes to zero. Maintaining $\tau = |\boldsymbol{\tau}| \neq 0$ would require the angular acceleration to go to infinity, this is obviously absurd so we must require that

$$\lim_{\Delta x \to 0} \tau = 0. \tag{7.15}$$

Let us now write an expression for the torque of a fluid element, we start from the definition of torque[2]:

$$\boldsymbol{\tau} \equiv \int_S \boldsymbol{r} \times \mathrm{d}\boldsymbol{F}, \tag{7.16}$$

the $i-$th component of which can be written

$$\tau_i = \int_S \epsilon_{ijk} x_j \mathrm{d}F_k = \int_S \epsilon_{ijk} x_j \sigma_{kl} \mathrm{d}S_l, \tag{7.17}$$

where we have taken $\boldsymbol{r} = (x_1, x_2, x_3)$ and the second equality is obtained by recalling the definition of the i-th component of the force (eq. 7.11). Equation 7.17 can be rewritten using the divergence theorem to obtain

$$\tau_i = \int_V \frac{\partial}{\partial x_l}(\epsilon_{ijk} x_j \sigma_{kl})\mathrm{d}V = \int_V \left(\epsilon_{ijk}\sigma_{kj} + \epsilon_{ijk} x_j \frac{\partial \sigma_{kl}}{\partial x_l}\right)\mathrm{d}V, \tag{7.18}$$

given that $\partial x_j / \partial x_l \equiv \delta_{jl}$. We now shrink the fluid element by letting its volume go to zero. The second term on the r.h.s. of eq. 7.18 will have dependencies on $x\sigma/L$, with L the scale over which σ varies appreciable, which in turn depends on the properties of σ and not on our limit for $x \to 0$. Thus this second term will tend to zero as $x \to 0$ and to have the requested $\tau_i = 0$ one simply needs that

$$\epsilon_{ijk}\sigma_{kj} = 0 \quad \Longrightarrow \quad \sigma_{kj} = \sigma_{jk}, \tag{7.19}$$

the implication being demonstrable using the definition of the Levi-Civita symbol (see C.4.5).

---

[1] We recall that the torque is $\boldsymbol{\tau} \equiv \mathrm{d}\boldsymbol{J}/\mathrm{d}t$ where the specific angular momentum $\boldsymbol{J}$ can be written as $\boldsymbol{J} = I\boldsymbol{\omega}$.

[2] This can be easily verified because $\boldsymbol{J} = \boldsymbol{r} \times \boldsymbol{p}$ and $\mathrm{d}\boldsymbol{J}/\mathrm{d}t = \boldsymbol{r} \times \mathrm{d}\boldsymbol{p}/\mathrm{d}t + \boldsymbol{p} \times \mathrm{d}\boldsymbol{r}/\mathrm{d}t$, with the second term being the cross product of $\boldsymbol{v}$ by itself and therefore null.

7.3.2. Isotropic and deviatoric part. We decompose the stress tensor into two parts, by adding and subtracting the same term:

$$\sigma_{ij} = \sigma^D_{ij} + \sigma^I_{ij} \equiv \left(\sigma_{ij} - \frac{1}{3}\sigma_{kk}\delta_{ij}\right) + \frac{1}{3}\sigma_{kk}\delta_{ij}, \tag{7.20}$$

where the first and second terms are called **deviatoric part** and **isotropic part** of the tensor; clearly, this operation can be done for any tensor. The deviatoric part in eq. 7.20 has null trace while, as expected, the isotropic part is a diagonal tensor. In physical terms, the isotropic tensor refers to normal forces (such as the pressure). These forces act normally to surfaces and produce *compression* or displacement (for incompressible fluids) in the direction of the variation of the momentum. On the contrary, the deviatoric part represents orthonormal forces. These act along orthogonal directions with respect to the variation of the momentum and produce *shear* in the fluid.

We know already a part of the tensor σ^I_{ij}, this is the (thermal) pressure of the gas. However this is not necessarily the only part as viscous forces can add more normal stresses. So in general we will have

$$\sigma^I_{ij} = -P\delta_{ij} + \sigma'^I_{ij}, \tag{7.21}$$

where σ'^I_{ij} will be made explicit below. We now find a form for the deviatoric part of the viscous stress tensor by focusing on the phenomenon of shear.

7.3.3. Form of shear stresses. As mentioned, shear occurs when two parts of the fluid in contact are moving at different velocities. For simplicity, let us assume that the motion of the fluid is only along the x-axis but the velocity changes along the z-axis, i.e. there is a velocity gradient along z ($\partial u_x/\partial z \neq 0$). Particle interactions along z will transfer momentum from the fast moving part of the fluid to the slow moving part. This transfer produces an acceleration (and deceleration) in x, therefore orthogonally to the direction of the momentum (or velocity) gradient, this is the basis of shear. The typical change in velocity along x will be

$$\Delta u_x \simeq \frac{\partial u_x}{\partial z}\ell, \tag{7.22}$$

where ℓ is the mean free path of the particles. The reason why ℓ appears in eq. 7.22 is that particles on average will move for a distance ℓ along z before interacting with another particle and exchange their momenta.

From the above, it is clear that the deviatoric part of the stress tensor (describing the shear) must depend on derivatives of the kind $\partial u_i/\partial u_j$, where $i \neq j$. This derivative constitutes a tensor and we can expect σ^D_{ij} to be some meaningful combination of these derivatives. To find the combination that is relevant for the shear we decompose the tensor $\partial u_i/\partial u_j$ in two parts, again by adding and subtracting the same quantity:

$$\frac{\partial u_i}{\partial u_j} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) + \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}\right) \equiv s_{ij} + r_{ij} \tag{7.23}$$

where the first tensor is symmetric and is called **strain tensor** s_{ij}, whereas the second is antisymmetric and is the **rotation tensor** r_{ij}.

In the previous Section we have established that σ_{ij} has to be symmetric. Thus it cannot be formed by compositions including the rotation tensor in eq. 7.23. This consideration is confirmed by analyzing the particular case of a rigid rotation and

recalling that in the presence of rigid rotations there is no shear. From the definition of rigid rotations we have that $\boldsymbol{v} = \boldsymbol{\omega} \times \boldsymbol{r}$ with $\boldsymbol{\omega}$ constant. The reader can verify that in this circumstance $r_{ij} = (1/2)(\boldsymbol{\nabla} \times \boldsymbol{u})_k = \omega_k$, which expresses rotations along the k direction. We are therefore left with the strain tensor that describes a *deformation* (also called *strain*) of the fluid as a consequence of stresses applied on it and we now have to find a way to link the stresses and the deformations.

7.3.4. Linear stress-strain relation. We recall the Hooke law for elastic deformations

$$\boldsymbol{F} = -K\boldsymbol{x}, \tag{7.24}$$

where K is the elasticity constant, which depends on the molecular properties of the material that we are considering. Equation 7.24 can be generalized to

$$F_i = -K_{ij}x_j, \tag{7.25}$$

where now K_{ij} is a tensor, not necessarily diagonal. In the following, we search for an analogue of eq. 7.25 that will link the 2-dimensional stress tensor (forces per unit area) to the strain tensor (deformations with respect to the initial positions). We require this relation to be linear and this implies, in general, the need for a four-order tensor in a relation that reads

$$\sigma^D_{ij} = A_{ijmn}s_{mn}, \tag{7.26}$$

where A_{ijmn} plays the role of a multidimensional constant.

We impose another crucial property for our four-order tensor: isotropy. This is same as to ask that the fluid is invariant for rotation, i.e. it is a so-called **Newtonian fluid**. In practice, astrophysical fluids can have properties that make them depart from isotropy, the most obvious being an ordered magnetic field. We neglect this eventuality here but we note that a dependence on the spatial direction of the viscous properties of a fluid in the presence of strong magnetic fields can be expected (as for thermal conduction, see §2.8.1). Under the assumption of isotropy, it can be shown that there is only a possible form for the linear A_{ijmn} tensor which reads[3]

$$A_{ijmn} = \alpha\delta_{im}\delta_{jn} + \beta\delta_{in}\delta_{jm} + \gamma\delta_{ij}\delta_{nm} \tag{7.27}$$

with α, β and γ parameters to be determined.

We substitute A_{ijmn} from eq. 7.27 in eq. 7.26 and impose the symmetry of s_{ij}:

$$\sigma^D_{ij} = \alpha s_{ij} + \beta s_{ij} + \gamma s_{kk}\delta_{ij} = 2\eta s_{ij} + \gamma s_{kk}\delta_{ij}, \tag{7.28}$$

where we have set $\alpha + \beta \equiv 2\eta$, with η another parameter that will be further described below. The trace of σ^D_{ij} is null by definition thus

$$\sigma^D_{ii} = 2\eta s_{ii} + 3\gamma s_{kk} = 0 \quad \Longrightarrow \quad \gamma = -\frac{2}{3}\eta. \tag{7.29}$$

We then recall the definition of the strain tensor (eq. 7.23) and finally derive the expression for the deviatoric part of the stress tensor:

$$\sigma^D_{ij} = \eta\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3}\frac{\partial u_k}{\partial x_k}\delta_{ij}\right). \tag{7.30}$$

[3]In two dimensions this tensor would be $\alpha\delta_{ij}$, in three dimensions it would be $\beta\epsilon_{ijk}$; α and β being constants.

The isotropic part of the same tensor must have the same shape but obviously it can depend only on derivatives of u_i along the same direction i. We can thus introduce another parameter ζ and write the final form for the stress tensor including pressure and viscous forces as:

$$\sigma_{ij} = -P\delta_{ij} + \eta\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3}\frac{\partial u_k}{\partial x_k}\delta_{ij}\right) + \zeta\frac{\partial u_k}{\partial x_k}\delta_{ij}. \tag{7.31}$$

Equation 7.31 describes all the forces that can act on a fluid element in the presence of pressure and viscosity. In this equation, η is the **coefficient of shear viscosity** and ζ is the **coefficient of bulk viscosity** (or second viscosity). Both have units of $\mathrm{g\,cm^{-1}\,s^{-1}}$. For incompressible fluid the stress tensor simply becomes

$$\sigma_{ij} = -P\delta_{ij} + \eta\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \tag{7.32}$$

being the divergence of the velocity null ($\partial u_k/\partial x_k = 0$).

7.4. Navier-Stokes equation

The stress tensor just found (eq. 7.31) can be put into the Euler equation (eq. 7.4) to derive a new expression that will describe the motion of the fluid in the presence of viscous forces. In practice, one has to take the spatial derivative (divergence) of the stress tensor with respect to x_j. Both η and ζ would, in general, be functions of two variables, for instance pressure and temperature and thus have complex spatial dependence. However, these coefficients tend to vary only slightly with thermodynamic variables and they are often considered constant. We therefore proceed to take them out of the spatial derivative to obtain the following expression for the divergence of the viscous stress tensor:

$$\frac{\partial \sigma'_{ij}}{\partial x_j} = \eta\left(\frac{\partial}{\partial x_j}\frac{\partial u_i}{\partial x_j} + \frac{\partial}{\partial x_j}\frac{\partial u_j}{\partial x_i} - \frac{2}{3}\frac{\partial}{\partial x_i}\frac{\partial u_k}{\partial x_k}\right) + \zeta\frac{\partial}{\partial x_i}\frac{\partial u_k}{\partial x_k}, \tag{7.33}$$

where the first term on the r.h.s. is a Laplacian operator. The second term can be changed by inverting the order of the derivatives and it becomes a gradient of a divergence, akin to the third and the forth terms. Recalling eq. 7.12 we then obtain our final equation that can be now written in vectorial form:

$$\rho\frac{\partial \boldsymbol{u}}{\partial t} + \rho(\boldsymbol{u}\cdot\boldsymbol{\nabla})\boldsymbol{u} = -\boldsymbol{\nabla}P + \eta\left[\nabla^2\boldsymbol{u} + \frac{1}{3}\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{u})\right] + \zeta\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{u}), \tag{7.34}$$

where gravity has been neglected for simplicity. Equation 7.34 is the so-called **Navier-Stokes equation** and describes the motion of viscous (non-ideal) fluids.

We can divide the eq. 7.34 by the density to find a different version of the viscosity coefficient. This version is particularly useful when the fluid is incompressible and reads

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u}\cdot\boldsymbol{\nabla})\boldsymbol{u} = -\frac{1}{\rho}\boldsymbol{\nabla}P + \nu\nabla^2\boldsymbol{u}, \tag{7.35}$$

where we have set $\nu = \eta/\rho$. The ν coefficient (units $\mathrm{cm^2\,s^{-1}}$) is called **kinematic shear viscosity** because it describes changes in velocity. By extension, η is analogously referred to as **dynamic viscosity** for its relation to changes in momentum. The kinematic viscosity ν is clearly more useful in fluids that can be considered incompressible because the density is implicitly considered constant here. Fluid

properties can change dramatically when compared to each other in terms of kinematic of dynamic viscosities. An illuminating example is given by air and water. At $T = 288\,\mathrm{K}$ the dynamic viscosity of water is about $\eta_{\mathrm{water}} = 1.1\times10^{-2}\mathrm{g\,cm^{-1}\,s^{-1}}$ while that of air is about two orders of magnitude lower, so water is more viscous than air. However, the kinematic viscosity of water is $\nu_{\mathrm{water}} = 1.1 \times 10^{-2}\mathrm{cm^2\,s^{-1}}$ (being its density $1\,\mathrm{g\,cm^{-3}}$) whereas that of the air is about an order of magnitude larger, so air is kinematically more viscous!

7.4.1. The viscosity coefficients. To estimate the above coefficients in a gas we go back to the example of a fluid moving in the x direction but with a gradient in velocity along z. Let us assume that at $z < 0$ the gas is moving at u_x larger than u'_x, the velocity of the gas at $z > 0$. The bottom gas will transfer momentum in the x direction to the top gas, the flux of this momentum being

$$f = \rho u_x v_z, \tag{7.36}$$

where v_z is the vertical component of the particle random motions[4]:

$$v_z = \sqrt{\frac{k_{\mathrm{B}}T}{\mu m_{\mathrm{p}}}}. \tag{7.37}$$

If we consider now the gas at $z > 0$, we can write an equation for the momentum flux analogous to eq. 7.36 but with u'_x. Thus, the net transfer between the two portions will be

$$\text{net flux} = \rho(u_x - u'_x)\sqrt{\frac{k_{\mathrm{B}}T}{\mu m_{\mathrm{p}}}} \tag{7.38}$$

assuming that the two gases have the same temperature. Then, recalling eq. 7.22 and the definition of mean free path (eq. 1.5) we have

$$\text{net flux} = \rho\ell\sqrt{\frac{k_{\mathrm{B}}T}{\mu m_{\mathrm{p}}}}\frac{\partial u_x}{\partial z} = \frac{\sqrt{\mu m_{\mathrm{p}} k_{\mathrm{B}}T}}{A}\frac{\partial u_x}{\partial z}. \tag{7.39}$$

We can now imagine to insert this expression in the Navier-Stokes equation (eq. 7.34). Being this a density of momentum flux in a certain direction (x), it will enter the equation inside a spatial derivative (see eq. 7.6). In other words, eq. 7.39 is analogous to eq. 7.30 in the very simplified case of two gas layers in contact at $z = 0$. We have then obtained an expression for the coefficient of shear viscosity[5]

$$\eta \approx \frac{\sqrt{\mu m_{\mathrm{p}} k_{\mathrm{B}}T}}{A}, \tag{7.40}$$

with an important dependence on the cross section of the particles A and on the square root of the gas temperature. Note that this coefficient does not depend on density, this may seem counterintuitive because higher density for instance would result in a larger number of particles carrying momentum. However, it would also

[4]Note that every component is alike.

[5]A more careful calculation leads to $\eta = (5\sqrt{\pi})\sqrt{\mu m_{\mathrm{p}} k_{\mathrm{B}}T}/(64A)$, about 7 times smaller than our approximate estimate. Note, however, that both for the η and the ν coefficients the dependencies remain unchanged.

results on a decrease in ℓ. These two facts compensate cancelling out the density dependence. As a consequence, the kinematic viscosity can be written as

$$\nu = \frac{\mu m_{\rm p}}{\rho A}\sqrt{\frac{k_{\rm B}T}{\mu m_{\rm p}}} = \ell v_{\rm th}, \tag{7.41}$$

where we have indicated with $v_{\rm th}$ the thermal speed of the fluid particles.

The last viscosity coefficient to consider is the bulk viscosity. This coefficient is related to compressions of the fluid and thus to the effect that non-ideal processes can play as a consequence of this compression. In practice, we are considering microscopic degrees of freedom of the particles different from the standard translational degrees (those that give rise to the thermal pressure). For instance, as a consequence of compression, particles can be excited (or ionized) producing a new state that, coupled with the translational degrees will give rise to a *frictional* force that opposes further compression. Bulk viscosity is important in phenomena like the damping of sound waves, which are compression waves and, in practical situations, decrease in amplitude (due to bulk viscosity) as they travel through a medium. We now investigate the effect of viscosity in some of the properties of fluid described in the previous chapters.

7.5. Dissipation of kinetic energy

One of the most important consequences of the presence of viscosity in a fluid is that part of its kinetic energy is not anymore conserved but it gets dissipated. This is analogous to what happens in any mechanical device in the presence of friction between its part, i.e. in any realistic (non-idealized) situation. We can estimate this dissipation by writing the rate of variation of kinetic energy with time by components:

$$\frac{\partial}{\partial t}\left(\frac{1}{2}\rho u^2\right) = \rho u_i \frac{\partial u_i}{\partial t} + \frac{1}{2}u^2\frac{\partial \rho}{\partial t} \tag{7.42}$$

and substituting the continuity equation (7.3) and the Navier-Stokes equation (7.12) on the r.h.s. We look for divergence terms as we have done in other occasions, in particular we make use of the following relations:

$$\begin{aligned}
&\frac{\partial}{\partial x_j}\left(\frac{1}{2}\rho u^2 u_j\right) = \frac{1}{2}u^2\frac{\partial}{\partial x_j}(\rho u_j) + \rho u_j u_i \frac{\partial u_i}{\partial x_j}\\
&\frac{\partial}{\partial x_j}(P u_j) = P\frac{\partial u_j}{\partial x_j} + u_j\frac{\partial P}{\partial x_j}\\
&\frac{\partial}{\partial x_j}(u_i \sigma'_{ij}) = u_i\frac{\partial \sigma'_{ij}}{\partial x_j} + \sigma'_{ij}\frac{\partial u_i}{\partial x_j}.
\end{aligned} \tag{7.43}$$

Substituting these equations in 7.42 and a using vectorial notation we obtain the equation of conservation of kinetic energy density including viscosity:

$$\frac{\partial \epsilon_{\rm K}}{\partial t} + \boldsymbol{\nabla}\cdot[(\epsilon_{\rm K}+P)\boldsymbol{u} - \boldsymbol{u}\cdot\overleftrightarrow{\sigma}'] = P\boldsymbol{\nabla}\cdot\boldsymbol{u} - \Psi, \tag{7.44}$$

where we have indicated the stress tensor with a double arrow. The second term on the r.h.s. of eq. 7.44 is the **rate of viscous dissipation** defined as

$$\Psi \equiv \sigma'_{ij}\frac{\partial u_i}{\partial x_j}. \tag{7.45}$$

As usual, eq. 7.44 can be integrated over a generic volume to better understand the meaning of every term. The first term on the l.h.s. is the rate of variation of the total kinetic energy in the volume. The divergence of the first term inside the square brackets gives the advective flux of $\epsilon_{\rm K}$ through the surface that encloses the volume plus the work done by the kinetic energy on that surface, see the discussion for the case without viscosity (§2.9). But we have now two more terms. The first one is the divergence of the scalar product between the velocity of the fluid and the viscous stress tensor, after integration in V this becomes a flux of energy transported by viscous processes. To visualize this consider the simple example of a fluid with a change in the x-component of the velocity at $z = 0$. If the volume enclosing the fluid has the plane $z = 0$ as a boundary then the shear at this surface will transfer some kinetic energy to the top portion of the fluid (outside the volume). This is a *viscous transfer* of kinetic energy.

The last term that we are left with is the integration in V of the rate of viscous dissipation (eq. 7.45). This is clearly something that reduces the kinetic energy of the volume (minus sign) without bringing anything outside the volume (there is no flux). It therefore represents a *loss of kinetic energy* due to viscous dissipation. Note also that Ψ is obviously positive given the definition of η (eq. 7.41) and the Navier-Stokes equation (eq. 7.34). This and the fact that kinetic energy must decrease due to dissipation justify the minus sign used for σ'_{ij} with respect to the pressure, see eq. 7.31.

7.5.1. Other energy equations. We have seen that viscosity produces a net loss of kinetic energy within the volume of the fluid. Given that this energy does not leave the volume and it cannot be destroyed, it must obviously cause an increase of the internal energy. The equation that expresses the conservation of (specific) internal energy (see 2.49), in the absence of radiative exchange of heat and thermal conduction, will thus become

$$\rho \frac{\mathrm{D}\mathcal{U}}{\mathrm{D}t} + P\boldsymbol{\nabla}\cdot\boldsymbol{u} = \Psi, \tag{7.46}$$

which, we recall, is an expression of the first law of thermodynamics. Note however that, in this case, the term on the r.h.s. does *not* represent an exchange of heat with the external medium but instead a increase of internal energy due to irreversible (viscous) processes. This is not in contradiction with the first law as, for irreversible processes and no heat exchange, this reads $\mathrm{d}\mathcal{U} + P\mathrm{d}\mathcal{V} > \delta Q = 0$.

We can rewrite equation 7.46 using the internal energy density ($U = \rho\mathcal{U}$) to obtain

$$\frac{\partial U}{\partial t} + \boldsymbol{\nabla}\cdot(U\boldsymbol{u}) = -P\boldsymbol{\nabla}\cdot\boldsymbol{u} + \Psi - \boldsymbol{\nabla}\cdot\boldsymbol{q} - \rho\mathcal{L}, \tag{7.47}$$

where we have now added also the exchange due to thermal conduction and radiation. Equation 7.47 can be now added to the conservation of kinetic energy density (eq. 7.44) to obtain the conservation of the *total* energy density in the presence of viscosity:

$$\frac{\partial \epsilon}{\partial t} + \boldsymbol{\nabla}\cdot[(\epsilon + P)\boldsymbol{u} - \boldsymbol{u}\cdot\overset{\leftrightarrow}{\sigma}{}'] = -\boldsymbol{\nabla}\cdot\boldsymbol{q} - \rho\mathcal{L}. \tag{7.48}$$

Note that energy can be transported by viscous stresses: $\boldsymbol{\nabla}\cdot(\boldsymbol{u}\cdot\overset{\leftrightarrow}{\sigma}{}')$, but the viscous dissipation does *not* produce a change in total energy but only a transfer of energy from kinetic to internal.

Finally, using again the first law of thermodynamics, this time in the form with entropy ($\mathrm{d}\mathcal{U} + P\mathrm{d}\mathcal{V} = T\mathrm{d}\mathcal{S}$), we can also write the *general* heat equation with viscosity and heat exchanges:

$$\rho T \frac{\mathrm{D}\mathcal{S}}{\mathrm{D}t} = -\boldsymbol{\nabla}\cdot\boldsymbol{q} - \rho\mathcal{L} + \Psi, \tag{7.49}$$

which states that the specific entropy of a volume of fluid changes as a consequence of thermal conduction at its surface, emission (or absorption) of radiation towards (or from) the outside world and viscous dissipation of kinetic energy. As one may have expected, the presence of viscosity produces an increase of the fluid entropy.

7.5.2. Rate of viscous dissipation. We seek a form of the viscous dissipation term in the above equations that can help to estimate its magnitude. Given the definition of the rate of viscous dissipation (eq. 7.45) we make the following transformations:

$$\Psi = \sigma'_{ij}\frac{\partial u_i}{\partial x_j} = \frac{1}{2}\sigma'_{ij}\frac{\partial u_i}{\partial x_j} + \frac{1}{2}\sigma'_{ij}\frac{\partial u_i}{\partial x_j} = \frac{1}{2}\sigma'_{ij}\frac{\partial u_i}{\partial x_j} + \frac{1}{2}\sigma'_{ji}\frac{\partial u_j}{\partial x_i}, \tag{7.50}$$

where the last equality is obtained simply by exchanging the indexes. We can then use the fundamental property of symmetry of the viscous stress tensor to obtain

$$\Psi = \frac{1}{2}\sigma'_{ij}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \tag{7.51}$$

and, recalling the expression that we have derived for the viscous stress tensor (eq. 7.31), we obtain

$$\Psi = \frac{1}{2}\eta\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)^2 + \left(\zeta - \frac{2}{3}\right)\left(\frac{\partial u_k}{\partial x_k}\right)^2, \tag{7.52}$$

which, for incompressible fluids, simply gives

$$\Psi = \frac{1}{2}\eta\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)^2. \tag{7.53}$$

In conclusion, viscosity modifies the energy transfer within the fluid with the consequence that part of its kinetic energy is now dissipated into heat. The rate at which dissipation occurs depends on the viscosity coefficients multiplied by the square of spatial derivatives of the fluid velocity (eq. 7.53).

7.6. Conservation of vorticity

We now derive and discuss a second fundamental result of the introduction of viscosity in the fluid motion. We introduce the concept of **vorticity of the flow** defined as

$$\boldsymbol{\omega} \equiv \boldsymbol{\nabla}\times\boldsymbol{u}. \tag{7.54}$$

Thus, $\boldsymbol{\omega}$ describes the property of the flow of being *rotational*. In practical terms, this property can be seen as follows. Imagine a fluid flow like that of a river; if we introduce in the flow a small floating arrow this will be transported downstream by the fluid motion. In an *irrotational* flow the arrow will point always in the same direction while in the presence of vorticity it will spin in a direction given by the orientation of $\boldsymbol{\omega}$ and with a speed given by its magnitude.

We now take the curl of all terms in the equation of Navier-Stokes (eq. 7.34). This operation leads to the following expression:

$$\frac{\partial \boldsymbol{\omega}}{\partial t} + \boldsymbol{\nabla} \times [(\boldsymbol{u} \cdot \boldsymbol{\nabla})\boldsymbol{u}] = -\boldsymbol{\nabla} \times \left(\frac{1}{\rho}\boldsymbol{\nabla} P\right) + \nu \boldsymbol{\nabla} \times (\nabla^2 \boldsymbol{u}), \tag{7.55}$$

where we have considered the fluid incompressible for simplicity[6]. The second term on the l.h.s. can be changed by recalling eq. C.17 to simply obtain $\boldsymbol{\nabla} \times (\boldsymbol{\omega} \times \boldsymbol{u})$. The first term on the r.h.s. is zero for any barotropic fluid given that the density is stratified with the pressure. Finally, the last term on the r.h.s. can be modified by applying twice eq. C.18 to obtain

$$\boldsymbol{\nabla} \times (\nabla^2 \boldsymbol{u}) = 0 - \boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{\omega}) = 0 + \nabla^2 \boldsymbol{\omega}, \tag{7.56}$$

given that the divergence of a curl is always null (see §C.4.2).

Finally, we obtain the equation for the vorticity in the presence of viscosity, which reads

$$\frac{\partial \boldsymbol{\omega}}{\partial t} + \boldsymbol{\nabla} \times (\boldsymbol{\omega} \times \boldsymbol{u}) = \nu \nabla^2 \boldsymbol{\omega}. \tag{7.57}$$

This equation is a generalization of the Kelvin theorem in the presence of viscosity. It states that the vorticity in these conditions is not conserved in the flow but it can change with time. The consequence of this will be discussed later, first let us digress a little on the case without vorticity, which provides a physical understanding of eq. 7.57.

7.6.1. Kelvin circulation theorem. Let us consider a closed curve C in a fluid and consider the line integral of the velocity along this curve. We define this integral as **circulation**:

$$\Gamma \equiv \oint_C \boldsymbol{u} \cdot \mathrm{d}\boldsymbol{l} = \int_S (\boldsymbol{\nabla} \times \boldsymbol{u}) \cdot \mathrm{d}\boldsymbol{S} = \int_S \boldsymbol{\omega} \cdot \mathrm{d}\boldsymbol{S}, \tag{7.58}$$

where we have applied the Stokes theorem (§C.4.3). From this we see that Γ simply represents the (flux of) vorticity that threads any open surface S that has C as a contour (boundary).

If we now take the full derivative of the circulation we can find the variation of this vorticity with time as the fluid moves:

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}t} = \int_S \frac{\partial \boldsymbol{\omega}}{\partial t} \cdot \mathrm{d}\boldsymbol{S} + \int \boldsymbol{\omega} \cdot \frac{\delta \mathrm{d}\boldsymbol{S}}{\delta t}. \tag{7.59}$$

The first term on the r.h.s. of eq. 7.59 is just the spatial integration of the time variation of the vorticity, the second term represents instead the change *in the surface* as the fluid moves (see Fig. 1). Indeed, the curve C moves with the fluid at a speed $\boldsymbol{u} = \delta\boldsymbol{r}/\delta t$ and thus $\delta\boldsymbol{r} = \boldsymbol{u}\delta t$. The variation in the surface (surface added to S by the motion) will then be the cross product between this $\delta\boldsymbol{r}$ and $\delta\boldsymbol{l}$, the unit vector along the curve C. As any surface element vector, $\delta \mathrm{d}\boldsymbol{S}$ will be orthogonal to this surface and therefore $\delta \mathrm{d}\boldsymbol{S} = \boldsymbol{u}\delta t \times \delta\boldsymbol{l}$. As a consequence, we have

$$\int \boldsymbol{\omega} \cdot \frac{\delta \mathrm{d}\boldsymbol{S}}{\delta t} = \oint \boldsymbol{\omega} \cdot \boldsymbol{u} \times \delta\boldsymbol{l} = \oint \boldsymbol{\omega} \times \boldsymbol{u} \cdot \delta\boldsymbol{l} = \int_S \boldsymbol{\nabla} \times (\boldsymbol{\omega} \times \boldsymbol{u}) \cdot \mathrm{d}\boldsymbol{S}, \tag{7.60}$$

[6]Note that the first term on the r.h.s. is essentially zero for incompressible fluids (see also §6.2), but we prefer to leave it explicit here because it shows that, in the absence of viscosity, the Kelvin theorem is also valid for compressible fluids as long as they are barotropic.

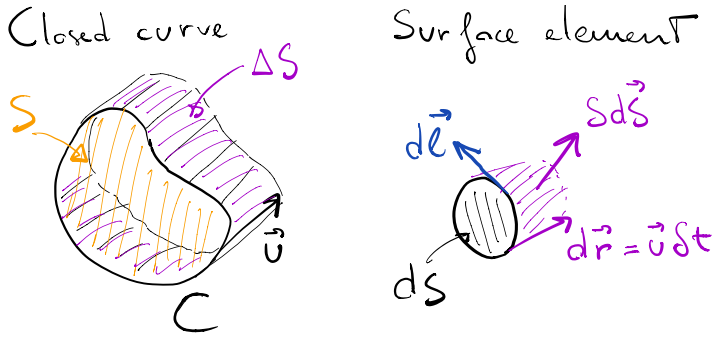

FIGURE 1. Sketch of the Kelvin circulation theorem, the left side shows the motion of the fluid elements located along the closed curve C, contour of a generic surface S. The right plot shows the evolution of the generic surface element.

where we have used a property of cross and dot products[7] (second identity) and the Stokes theorem for the third identity.

We now have two terms involving an integration over a surface that has the curve C as its boundary and thus we can express the variation of the circulation as follows:

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}t} = \int_S \left[\frac{\partial \boldsymbol{\omega}}{\partial t} + \boldsymbol{\nabla} \times (\boldsymbol{\omega} \times \boldsymbol{u})\right] \cdot \mathrm{d}\boldsymbol{S} = 0, \tag{7.61}$$

where the last equivalence is valid in the absence of viscosity (see eq. 7.57). Equation 7.61 is called **Kelvin circulation theorem** and states that, in an inviscid flow, the vorticity that threads a surface that moves with the fluid (Γ) is conserved with the fluid motion. If we represent the vorticity field with *vortex lines*, the number of these lines that thread any element of area remains unchanged with time, as the fluid moves. In practice, this means that if some vorticity is present from the beginning this will be transported downstream by the flow. If there is no vorticity, it will never be generated in the absence of an external cause.

This situation changes if viscosity cannot be neglected. In this case, as shown by the combination of eqs. 7.61 and eq. 7.57, the circulation is not conserved and the motion of the fluid can change the vorticity vector. This means that vorticity can be destroyed and/or generated within the flow thanks to forces orthogonal to the surfaces, which are prime properties of a viscous flow. We note that, although we have excluded it from this calculation, gravity would not play any role here given that it enters the Euler equation as a gradient that would vanish once we take its curl.

7.7. Flow around a body

We now consider the flow of a fluid past a solid body. This system has little astrophysical applications but it is useful to understand the effect of the introduction of viscosity in a flow. It will also allow us to introduce the concept of the generation of turbulent flows that are the subject of Chap. 8.

[7]This property is always valid for any three vectors: $\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{b} \cdot (\boldsymbol{c} \times \boldsymbol{a}) = \boldsymbol{c} \cdot (\boldsymbol{a} \times \boldsymbol{b})$.

Consider a solid body at rest and an ideal fluid that is moving around it from right to left (Fig. 2, top). Let us assume that the fluid is irrotational, i.e. there is no initial vorticity, and incompressible ($\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$). Given that $\boldsymbol{\nabla} \times \boldsymbol{u} = 0$ we can define a potential that generates the velocity field ($\boldsymbol{u} = -\boldsymbol{\nabla}\psi$) in analogy to what done in §6.2. The equation that describes the motion of the fluid in this case is simply a Laplace equation of the kind

$$\nabla^2 \psi = 0 \tag{7.62}$$

that requires two boundary conditions to be solved. The first condition can be given for $\boldsymbol{x} \to \infty$ and it may simply be that the velocity along x is a constant: u_{x}. The second condition that is usually set in these situations is the following:

$$\hat{\boldsymbol{n}} \cdot \boldsymbol{\nabla}\psi = 0, \tag{7.63}$$

where $\hat{\boldsymbol{n}}$ is the unit vector normal to the surface of the solid. This condition assures that the fluid does not enter the solid: **no-penetration condition**. Under this assumption, the flow past the body looks like the one shown by the streamlines in Fig. 2 (top panel). The fluid flows around the body and resumes its constant velocity as it has passed it. Away from the body and, in particular, immediately downstream it there is no evidence of its presence. There is no generation of vorticity in agreement with the Kelvin theorem and the flow is completely *laminar*.

Now, let us examine the situation in which viscosity is present (Fig. 2, middle and bottom panels). The Navier-Stokes equation introduces a term with the second derivative of the velocity $\nabla^2 \boldsymbol{u} \neq 0$, which translates, if we make the same assumption of no vorticity present upstream ($\boldsymbol{\nabla} \times \boldsymbol{u} = 0$), into equations with the third derivative of the velocity potential ($\nabla^3 \psi$). We therefore should impose a third boundary condition together the above two, this is taken to be

$$\hat{t} \cdot \boldsymbol{\nabla}\psi = 0, \tag{7.64}$$

where $\hat{t}$ is the unit vector tangent to the surface of the solid. The expression in eq. 7.64 is called **no-slip condition** and represents the idealized situation in which the elements of the fluid are at rest at the surface of the solid[8]. Starting from the solid surface, the fluid now develops a gradient in velocity that reaches the laminar value at large distances from the solid. The no-slip condition leads to solutions that are fully compatible with experiments and it is therefore widely used to characterize viscous flows.

We conclude that the main consequence of the introduction of viscosity (and the non-slip condition) is the creation of a velocity gradient close to the fluid's surface. This gradient produces a so-called **boundary layer** whose size depends on the viscous properties of the fluid. In this layer, the Kelvin theorem is no longer valid and vorticity is not conserved. In particular, the non laminarity of the flow can generate **vortexes** that leave the boundary layer and propagate downstream. If the fluid is not too viscous in the downstream flow the Kelvin theorem will then be again applicable and the vortexes can survive there for some time. The evolution of the fluid downstream can either remain roughly laminar (but with

[8]This condition may seems exaggerated because, if on the one hand the fluid is obviously decelerated close to the solid because of viscosity, on the other hand putting it at rest seems a contradiction with the fact that there cannot be accumulation of mass at the solid (the fluid is still assumed incompressible). However, this should be seen as a limit condition that applies to fluid elements that almost *coincide* spatially with the surface of the solid, thus effectively to an infinitely thin layer.

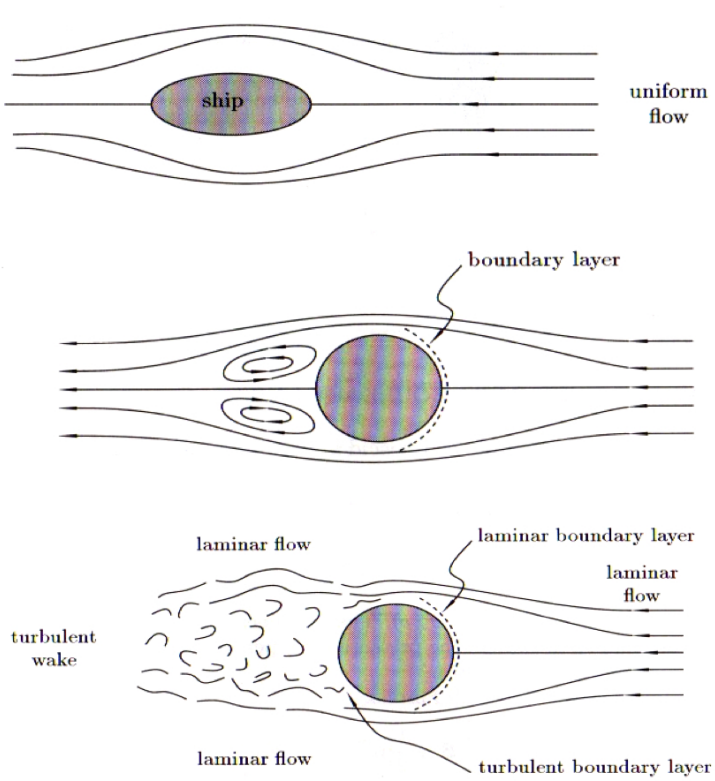

FIGURE 2. Flows past a solid body for ideal (top panel) and non-ideal fluids. From [Shu92].

possible vorticity) or become turbulent, depending on properties that we describe in the next Section.

7.7.1. Reynolds number. The onset of a turbulent flow from an initially laminar flow is determined by the so-called **Reynolds number** (Re) of the flow. This is a dimensionless quantity defined as the ratio between the inertial and the viscous forces in the Navier-Stokes equation in orders of magnitude:

$$\frac{(\boldsymbol{u}\cdot\boldsymbol{\nabla})\boldsymbol{u}}{\nu\nabla^2\boldsymbol{u}} \quad \Longrightarrow \quad \mathrm{Re}\equiv\frac{LV}{\nu}, \tag{7.65}$$

where L is the typical size of the fluid in question (or of the turbulent region) and V is the variation in the turbulent velocity within L. It has been experimentally established that whenever the Reynolds number is larger than a certain critical value ($\mathrm{Re} > \mathrm{Re}_{\mathrm{crit}}$) the fluid is prone to develop turbulence. The critical value does not depend much on the type of fluid and it is $\mathrm{Re}_{\mathrm{crit}} \sim 10^3 - 10^4$.

Going back to our solid object, for large Reynolds numbers the flow coming out of the boundary layer will be (or it will quickly turn into) a turbulent flow. A **turbulent wake** then forms downstream the body, while away enough from it the flow remains laminar. The onset of this turbulent flow has the global effect of

draining part of the initial kinetic energy of the body transforming it from ordered motion (of the body) to chaotic motion (of the fluid). An effect of this kind is unavoidable in realistic situations and to reduce its impact on the motion of bodies (boats, airplanes, cars) is one of the main concerns of aerodynamic engineering. In practice, most of the modern designs aim to limit the effect of the boundary layer by shifting it as much as possible downstream. Note that in the case of a solid body, the size of the fluid that is relevant in terms of development of turbulence is roughly the size of the body itself or, better, the cross section of the body with respect to the flow. This is the length L in the equation of the Reynolds number (eq. 7.65). Larger and faster bodies are more likely to develop copious turbulent flows.

CHAPTER 8

Turbulence

In this chapter, we describe in some detail one of the most important features of astrophysical fluid, i.e. their tendency to develop and maintain turbulent flows. Turbulence is a type of *chaotic* kinetic energy, in which *portions* of the fluid are moving with respect to each other at velocities that are often larger than the typical thermal motions of elementary particles. The main evidence of turbulence in the ISM is provided by observations of emission line broadening. Emission and absorption lines experience three types of broadening: natural, thermal and turbulent. The natural broadening is typically negligible in ISM emission lines but it can become very important for absorption lines at high column densities. Let us focus here on emission lines. Spectroscopic observations of various ISM components reveal that these lines are typically broader than what one would expect by knowing the temperature of the emitting medium[1]. The energy source of turbulence in the ISM is not fully understood in most cases but it is quite probable that supernovae and stellar winds are playing an important role. One of the main characteristics of turbulence is that it gets dissipated over relatively short timescales, thus, to be ubiquitously observed in the ISM, it must be that the sources of turbulence are continuously *active*. In this chapter, we see how turbulence can be treated with a simple theory and we consider the merits and limitation of this theory in the application to astrophysical gases.

8.1. The Reynolds number in the ISM

At the end of Chap. 7, we have seen that a convenient way to determine whether a fluid is prone to develop turbulence or not is to calculate its Reynolds number (eq. 7.65). Experiments show that if the Reynolds number is larger than a few $\times 10^3$ and we are in the presence of disturbances/instabilities, turbulence will develop in the flow. We can thus try to estimate this number in the Galactic ISM, for this we need to use our estimate of the viscous shear coefficient (eq. 7.41). This leads to a kinematic viscosity of the ISM of

$$\nu_{\rm ISM} = \frac{c_{\rm s}}{An} \simeq 10^{20} \left(\frac{c_{\rm s}}{\rm km\,s^{-1}}\right) \left(\frac{n}{\rm cm^{-3}}\right)^{-1} {\rm cm^2\,s^{-1}}, \tag{8.1}$$

where we have used $A = 10^{-15}\,{\rm cm^{-2}}$ (from §1.2.1) and substituted the thermal speed with the sound speed given that they are approximately equal. With this, we can write an expression for the Reynolds number in the ISM:

$${\rm Re_{ISM}} \simeq 3 \times 10^5 \left(\frac{L}{100\,\rm pc}\right) \mathcal{M}_{\rm ISM} \left(\frac{n}{\rm cm^{-3}}\right), \tag{8.2}$$

[1]Note that in order to determine this, one must have independent and reliable estimates of the gas temperature.

where we have indicated with $\mathcal{M}_{\rm ISM}$ the Mach number of the ISM, typically of order unity or larger. Note that in eq. 8.2 we have taken as typical size of the system the thickness of the H I layer in the inner Milky way. This is the typical length over which the system can be considered homogeneous and isotropic in 3D, for larger distances, the properties of the medium (e.g. its density) vary substantially at least along the vertical direction.

Equation 8.2 shows that we can expect, in general, the ISM to be a turbulent medium. In other words, if there are reasons for turbulence to develop (e.g. instabilities) and to be maintained (energy sources) then we should expect to observe the signature of turbulence in the ISM. Do we see this signature? As mentioned, if we observe a typical nearby disc galaxy in the 21-cm line at high angular[2] and spectral resolution the profile that we obtain is roughly Gaussian with an intrinsic velocity dispersion ($\sigma_{\rm HI}$) ranging from $7\,{\rm km\,s^{-1}}$ to $20\,{\rm km\,s^{-1}}$. The highest values of the velocity dispersion are found in the inner parts of the discs, the low values in the outer parts, typically. Line broadenings much higher that these values are found infrequently and they are hardly attributable to gas turbulence but instead to large scale motions such as supernova remnant expansions (Chap. 5). The thermal speed of the WNM is about $7\,{\rm km\,s^{-1}}$ (§1.1.1) and anytime we observe velocity dispersions larger than this value we are in the presence of (supersonic) turbulence. Thus, most of the neutral atomic gas in galaxy discs is likely turbulent.

As a final note, we discuss the importance of viscosity in the ISM. Going back to the definition of the Reynolds number (eq. 7.65), it is clear that a large value of Re occurs when inertial forces (the advective term in the Euler equation) largely exceed the viscous forces. Equation 8.2 then shows that viscous forces are typically largely negligible in the global motion of the ISM. This means that viscosity has little importance on large scales (sizes of the discs)[3] in the ISM. However, as we see below, it is very important for the dissipation of turbulence on small scales.

Turbulence is one of the most complex subjects in fluid-dynamics and its study requires either very advanced mathematical methods of investigation or sophisticated numerical simulations. We consider here the two standard extreme situations that allow simplified mathematical treatments: 1) transition to turbulence and 2) fully developed turbulence. The first is a branch of statistical mechanism that investigates under what conditions turbulence can develop in an initially laminar flow. The second assumes a static situation in which turbulence has already fully developed and investigates the presence of basic laws that describe the *average* behaviour of the fluid.

---

[2]The angular resolution is important here because disc galaxies usually rotate very fast and different regions in the disc will have different velocities along the line of sight just because of this rotation. In order to measure the *intrinsic* velocity dispersion, it is therefore important that the *beam* (radio-frequency analogue of the point spread function) of the observations is much smaller that the typical angular scale over which rotation velocities (or other streaming motions) vary appreciably. In the absence of this condition, one has an effect called **beam smearing** and the observed broadening of the line will be contributed (or even dominated) by large-scale motions in the galaxy disc.

[3]Note that this fact is at the basis of the introduction of the so-called **anomalous viscosity** in problems like that of accretion discs around black holes, see brief discussion at the end of §6.4.

8.2. Transition to turbulence

In general terms, two ingredients are required for turbulence to develop in a fluid: 1) the development of an instability, 2) a high Reynolds number of the flow. If these conditions are met, it is likely that turbulence will develop and it will remain in the fluid as long as the original source of instability is available. We can also have cases where the onset of turbulence is driven, not by instabilities, but by viscosity alone. This is what happens in the flow past a body that we saw in §7.7 where turbulence unfolds out of the viscous boundary layer. We now consider another simple classical case of this kind that admits an analytical solution.

8.2.1. Poiseuille flow. Let us consider the flow of a fluid between two infinite (solid) slabs. The fluid is moving along the $x-$direction with uniform velocity in x and potentially a velocity gradient in the $z-$direction, $\boldsymbol{u} = u_{\mathrm{x}}(z)\hat{\boldsymbol{e}}_{\mathbf{x}}$. The system is infinite in y. We assume the fluid to be incompressible and that viscosity cannot be neglected. We write the equation of Navier-Stokes under these assumptions and search for a solution for the velocity vector $\boldsymbol{u}$ in equilibrium conditions:

$$(\boldsymbol{u}\cdot\boldsymbol{\nabla})\boldsymbol{u} = -\frac{1}{\rho}\boldsymbol{\nabla}P + \nu\nabla^2\boldsymbol{u}, \tag{8.3}$$

where we have removed the time derivative because of equilibrium and we have assumed that the flow is fast enough that gravity can be neglected.

The $x-$component of eq. 8.3 reads

$$u_{\mathrm{x}}\frac{\partial u_{\mathrm{x}}}{\partial x} = -\frac{1}{\rho}\frac{\partial P}{\partial x} + \nu\left(\frac{\partial^2 u_x}{\partial x^2} + \frac{\partial^2 u_x}{\partial z^2}\right), \tag{8.4}$$

which, given that the velocity is simply $\boldsymbol{u} = u_{\mathrm{x}}(z)\hat{\boldsymbol{e}}_{\mathbf{x}}$, becomes

$$\frac{1}{\eta}\frac{\partial P}{\partial x} = \frac{\mathrm{d}^2 u_x}{\mathrm{d}z^2}. \tag{8.5}$$

Analogously, the $z-$component of the Navier-Stokes equation is[4]

$$\frac{1}{\rho}\frac{\partial P}{\partial z} = 0, \tag{8.6}$$

which states that, as we may have expected, the pressure does not depend on z but only on x and we can write eq. 8.5 with a total, rather than a partial, derivative of the pressure.

The only way to satisfy eq. 8.5 (being u_{x} only function of z and P of x) is if the derivatives on the left and right hand sides return a constant. Thus, P will be a linear function of x and we can substitute the derivative with the ratio of the difference in pressure between the beginning and the end of the channel in x divided by the length of the channel:

$$\frac{\mathrm{d}}{\mathrm{d}z}\frac{\mathrm{d}u_{\mathrm{x}}}{\mathrm{d}z} = \frac{1}{\eta}\frac{\Delta P}{l_0}, \tag{8.7}$$

which can be then easily integrated to obtain

$$u_{\mathrm{x}} = -\frac{1}{2\eta}\frac{\Delta P}{l_0}z^2 + c_1 z + c_2. \tag{8.8}$$

[4]Note that this would be modified by an equation stating hydrostatic equilibrium if gravity cannot be ignored.

At this point we make use of the main boundary condition for viscous flows: the no-slip condition. We rewrite eq. 8.8 for points at the upper and lower boundaries of the channel ($z = \pm h$). These are two equations that can be subtracted from each other to obtain $c_1 = 0$ and summed to obtain the second constant. The final solution of the velocity in a **Poiseuille flow** reads

$$u_{\mathrm{x}} = -\frac{1}{2\eta}\frac{\Delta P}{l_0}(z^2 - h^2). \tag{8.9}$$

Thus the velocity of the flow along z will take a parabolic form with the highest velocity at the centre of the channel and a symmetric gradient on both sides[5].

The velocity gradient that we just found is what viscosity (through the imposition of the no-slip condition) does to the fluid. In a problem of this kind, the impact of viscosity is quite remarkable and influences a large portion of fluid although the gradient is clearly much higher close to the boundaries. We recall that, in the region where viscosity cannot be neglected, the Kelvin theorem is no longer verified (§7.6). This translates into the possibility that vorticity (in the form of eddies) develops in the flow. We now ask ourselves under what conditions these eddies can lead to a turbulent flow, in order to know this we calculate the Reynolds number of the Poiseuille flow.

First we estimate the typical velocity of the flow by integrating eq. 8.9 along z:

$$\langle u\rangle = \frac{\int_{-h}^{h} u_{\mathrm{x}}(z)\mathrm{d}z}{2h} = \frac{1}{3}\frac{h^2}{\eta}\frac{\Delta P}{l_0} \tag{8.10}$$

and we substitute it in the definition of the Reynolds number:

$$\mathrm{Re} \approx \frac{2h\langle u\rangle}{\nu} = \frac{2\Delta P}{3l_0}\frac{h^3\rho}{\eta^2} = \frac{2\Delta P}{3l_0}\frac{h^3}{\rho\nu^2}. \tag{8.11}$$

Thus, in this specific flow, Re has a relatively strong dependence on the viscosity coefficients. Note also that the larger the viscosity, the lower Re. This latter property does not refer only to the Poiseuille flow but it is true in general: very high viscosity tends to hinder the development of turbulence.

8.2.2. Stability analysis of the Poiseuille flow. We now consider the introduction of perturbations in velocity and pressure in the Navier-Stokes equation and we look for solutions of the usual kind (Fourier modes):

$$\Delta u = \widetilde{\Delta u}(z)e^{i(kx-\omega t)}. \tag{8.12}$$

This problem is mathematically very involved and we give here only the final solution. In order to obtain a physically meaningful solution, it is, this time, convenient to assume that the frequency of the perturbation (ω) is a real number and therefore to look for conditions that lead to an imaginary wave-number k. This effectively corresponds to assume that at one end of the channel (say at $x = 0$) a perturbation is set with time frequency ω real. This makes the perturbation oscillate with time but not necessarily in space, indeed it could either oscillate, increase or decrease downstream in the flow depending on the value of the imaginary part of k. In particular, for $\mathrm{Im}(k) < 0$ we have the development of an instability downstream

[5]Note that a similar solution exists also for a pipe or hose with a round section. In that case $u(R) = -\Delta P(R^2 - R_{\mathrm{pipe}}^2)/(4\eta l_0)$, R_{pipe} is the hose radius, l_0 its length ΔP the difference in pressure at the extremities. Once integrated, this equation leads to a well-known equation for the flow rate in hydraulics: $Q = 2\pi\rho\int_0^{R_{\mathrm{pipe}}} u(R)R\mathrm{d}R = \pi\Delta P R_{\mathrm{pipe}}^4/(8\nu l_0)$.

(for increasing x), for $\mathrm{Im}(k) \geq 0$ the flow remains laminar (unperturbed) as the perturbation will oscillate or be damped.

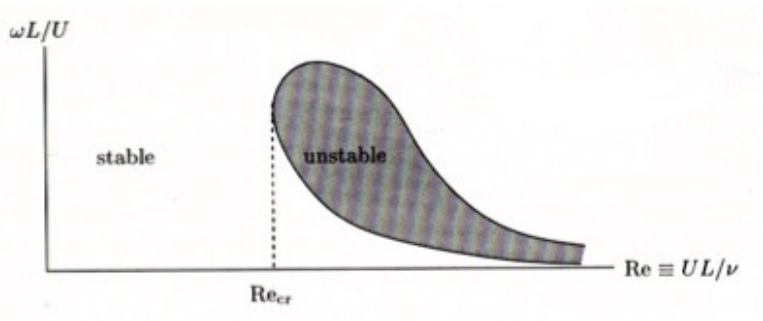

FIGURE 1. Locus of points where $\mathrm{Im}(k) \leq 0$ (shaded area) in the plane (Re, ω) for a Poiseuille flow. From [Shu92].

The graphical solution of the perturbation analysis is given in Fig. 1 where we show the locus of points where $\mathrm{Im}(k) < 0$ (shaded area) in the plane (Re, ω). The solid curve shows the demarcation between stable and unstable regions: $\mathrm{Im}(k) = 0$, and it is called **neutral stability curve**. Fig. 1 shows that for low Reynolds numbers the flow in a Poiseuille channel is always laminar (non turbulent) as expected. For $\mathrm{Re} \geq \mathrm{Re}_{\mathrm{crit}}$ turbulence can develop at least at some frequencies. At $\mathrm{Re} > \mathrm{Re}_{\mathrm{crit}}$ we enter a highly turbulent regime in which nearly every frequency up to a upper boundary are excited and the flow is very turbulent and chaotic. The maximum turbulent frequency ($\omega_{\max} \sim \langle u \rangle / h$) (top of the plot) is defined by the viscous damping of high frequencies that we discuss later. For $\mathrm{Re} \gg \mathrm{Re}_{\mathrm{crit}}$ the number of turbulent frequencies becomes smaller and smaller. This is an interesting result that can be understood as follows.

It is clear that, at least in the specific case of the Poiseuille flow, viscosity has a *dual role* in the onset of turbulence. On the one hand, it determines the gradient in velocity of the flow that is necessary for the instability to develop, on the other hand it enters the Reynolds number thus influencing the tendency of the fluid to develop turbulence. To have very high Reynolds numbers one needs to have low values of ν, thus for $\mathrm{Re} \gg \mathrm{Re}_{\mathrm{crit}}$ viscosity become progressively less important. This makes the instability tending to vanish, and so does, paradoxically, the turbulence. There is, in fact, an optimal range of values of ν where turbulence is fully developed at virtually all frequencies. Fig. 1 shows, however, that at low frequencies, turbulence never vanishes also for very high Reynolds numbers.

The behaviour of the Poiseuille flow can be extended to other configurations and fluids. The dual role of viscosity may not be present, in general, because the instabilities can be governed by other phenomena than viscosity (Chap. 6). For instance, the fluid can be subject to Kelvin-Helmholtz instabilities (§6.2) if there are important velocity gradients. In this case, the instability will develop (independently of viscosity) and, if Re is large, the non-linear regime of this instability will develop into a turbulent layer. Instead, for low Re (high viscosity), the flow will be unstable but not turbulent, meaning that viscosity will damp the development of turbulence. Note that in this case, it is quite likely that the viscosity will hamper the instability as well. Given that the Reynolds number of the ISM seems to be

largely above the critical value, we can conclude that once instabilities develop into non-linear regimes it is likely that turbulence will also develop.

8.2.3. Rotating fluid: Couette flow. We briefly consider the case of a rotating fluid with shear viscosity. In a laboratory, this is obtained by setting two concentric empty cylinders (with different radii) in motion at different angular speeds with the fluid being between them; this is called *Couette flow*. In astrophysical situations, this flow bears analogies with rotating galactic, proto-planetary or accretion discs.

We have seen that in the absence of viscosity, a rotational instability develops whenever the square of the angular momentum does not increase with radius (§6.4). We have also verified that this situation never occurs in astrophysical objects, it can however be produced in the laboratory. Figure 2 shows this condition schematically. Here the two axes are chosen conveniently to represent the Reynolds numbers at two radii (R and R') that could be the radii of the two cylinders or more generically two nearby radii as in §6.4 with $R' = R + \Delta R$ and $\Delta R > 0$. The curve $x = y$ represents the condition in which these two numbers are equal, which also corresponds to the specific angular momenta ($\mathcal{J} = \Omega R^2$) at the two radii to be equal. According to the Rayleigh criterion for rotational instability, if the angular momentum at the inner radius is lower than that at the outer radius ($y < x$) there will always be stability. For the opposite ($y > x$) we should expect rotational instability.

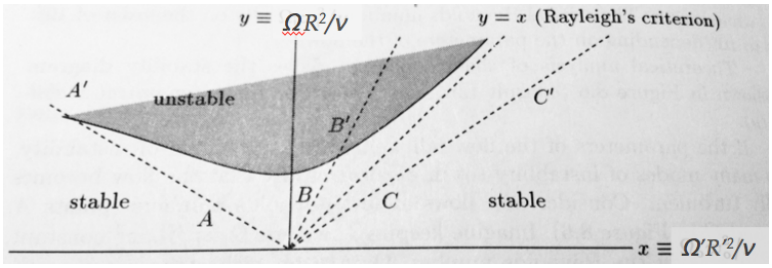

FIGURE 2. Development of instability and turbulence in a rotating fluid with viscosity: Couette flow. Adapted from [Shu92].

The introduction of viscosity causes the instability condition to substantially change, as shown by the shaded area in Fig. 2. For low Reynolds numbers (see point B) viscosity damps the instability and turbulence can not develop even if the flow is rotationally unstable according to the Rayleigh criterion. Instability (and turbulence) can only develop for relatively large Reynolds numbers if we enter the instability regions (point B'). For counter-rotation between the two radii (left side of the diagram), the region of instability is much larger. This is also the case of non-viscous flows as one can see by rewriting the instability criterion assuming that the sign of Ω' is different from Ω. Again, with the introduction of viscosity, instabilities are damped in a large fraction of flows with relatively low Reynolds numbers.

8.3. Fully developed turbulence

We now consider the case of turbulence **fully developed** in a fluid to a point that the (turbulent) flow can be consider stationary. We assume that the fluid is,

initially, incompressible, homogeneous and isotropic. This allows us to derive simple theoretical prescriptions, which are quite powerful to predict the general behaviour of relevant physical properties in the presence of turbulence. As usual, the assumption of incompressibility may be an oversimplification for the ISM. We discuss at the end of this chapter the potential consequences of releasing this assumption.

8.3.1. Kolmogorov theory. The assumption of incompressibility means that the velocity field of the fluid is solenoidal ($\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$). In general, we can decompose turbulence motions in two main modes: longitudinal and transverse. The first kind is obviously not allowed in incompressible fluids and so we are left with the transverse motions that can be described in terms of **eddies**[6]. The fact that the turbulence is fully developed means that we can expect the emergence of eddies of very different sizes. This is analogous to having $\mathrm{Re} > \mathrm{Re}_{\mathrm{crit}}$ in the Poiseuille flow (§8.2.1) without being $\mathrm{Re} \gg \mathrm{Re}_{\mathrm{crit}}$. In this regime, turbulence develops in a large range of frequencies or, equivalently, spatial scales. Analogously, the turbulent flow of fully-developed incompressible turbulence will be characterized by chaotic motions (eddies) at very different scales.

Let us start by describing the eddies at the largest possible scales. This is the physical scale over which turbulence has developed, or the size of the system if it is *all* turbulent, and we indicate this scale with L. At each point in the flow, the total speed of a fluid element can be written as

$$\boldsymbol{w} = \boldsymbol{u} + \boldsymbol{v}_{\mathrm{turb}}, \tag{8.13}$$

where $\boldsymbol{u}$ is our usual fluid speed and $\boldsymbol{v}_{\mathrm{turb}}$ is the random velocity due to turbulence. Let us define as V the typical speed of the largest eddies. This can be envisaged as the typical velocity difference between fluid elements located at a distance L. We are building a theory to describe the turbulent motion and thus that this difference must be representative of the turbulent part of the velocity, i.e. $V \sim |\Delta \boldsymbol{v}_{\mathrm{turb}}|$. However, given that turbulence is ultimately a consequence of the motion of the fluid (whether it develops because of shear or Kelvin-Helmholtz instability or other reasons) we can often expect that, at least on these large scales, $|\Delta \boldsymbol{v}_{\mathrm{turb}}| \sim |\Delta \boldsymbol{u}|$.

We use the definition of the Reynolds number (eq. 7.65) of the fluid to write

$$\mathrm{Re} = \frac{LV}{\nu} \sim \frac{L|\Delta \boldsymbol{u}|}{\nu} \gg 1, \tag{8.14}$$

where the last inequality is true by construction for a fluid with fully developed turbulence ($\mathrm{Re} > \mathrm{Re}_{\mathrm{crit}} \sim 10^3 - 10^4$). Given that this number is nothing else than the ratio between inertial (advective) forces and viscous forces (eq. 7.65), the above inequality means that, on large scales (L), the viscous forces are completely negligible for the motion of the fluid. In other terms, viscosity plays no role on the largest scales and in particular it *cannot* dissipate any appreciable kinetic energy of the fluid (see §7.5). The main assumption of the Kolmogorov theory is then that, given that kinetic energy cannot be dissipated by viscosity on large scales, it will be transferred to the smaller scales. This phenomenon produces a **cascade**, where kinetic energy is transferred to smaller and smaller eddies until one reaches a scale at which viscosity can no longer be neglected. At these small scales, dissipation will occur and the kinetic energy will turn into internal energy as we have seen in Sec. 7.5.1. Clearly, for a process like this to be stationary (one of the assumptions we

[6]Clearly, we are implicitly assuming that the fluid is not irrotational, $\boldsymbol{\nabla} \times \boldsymbol{u} \neq 0$.

made initially) one needs to continuously feed fresh kinetic energy into the largest eddies at the same rate at which it is transferred to the smaller eddies and eventually dissipated (at the smallest scales).

To write the above in formulae it is useful to consider a unit mass of fluid and to work with specific quantities. The specific kinetic energy ($\mathcal{E}$) for the large eddies can be simply written as

$$\mathcal{E} \sim V^2. \tag{8.15}$$

Given the typical timescale at these scales of $\tau = L/V$, the energy transferred to the smaller scales, in a unit of time, will be of the order

$$\dot{\mathcal{E}} \sim \frac{V^3}{L}. \tag{8.16}$$

As mentioned, this is also the rate at which one has to introduce (specific) kinetic energy on the largest eddies to keep the system stationary.

If we consider some intermediate scale λ with $\lambda < L$, the rate of specific energy transfer can be written, analogously to the above, as

$$\dot{\mathcal{E}} \sim \frac{v_\lambda^3}{\lambda}, \tag{8.17}$$

where we have indicated with v_λ the typical variation of $|\boldsymbol{v}_{\rm turb}|$ over a distance λ. An important assumption of the cascade process is that there is no significant loss of kinetic energy on scales larger than the dissipation scale (we discuss this below). Thus the rates in eqs. 8.17 and 8.16 are exactly the same and, if we equate them, we finally obtain the **Kolmogorov-Obukhov law** for the turbulent speed of an eddy of size λ:

$$v_\lambda = V \left(\frac{\lambda}{L}\right)^{1/3}. \tag{8.18}$$

This law states that the typical turbulent speed scales with the sizes of the eddies to the 1/3 power. Because of the shape of the relation, most of the kinetic energy in the Kolmogorov cascade is held on the largest eddies (as $\mathcal{E} \propto v_\lambda^2 \propto \lambda^{2/3}$), whilst most of the vorticity is on the smallest eddies (as $\omega \propto v_\lambda/\lambda \propto \lambda^{-2/3}$).

8.3.2. Dissipation scale. In §7.5.2 we have derived a form for the rate of viscous dissipation for an incompressible fluid (eq. 7.53). We can now use that equation to estimate the scale of turbulence dissipation. We call this scale λ_0 and approximate the dissipation rate as follows:

$$\Psi = \frac{1}{2}\eta \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)^2 \sim \eta \frac{v_{\lambda_0}^2}{\lambda_0^2}, \tag{8.19}$$

where v_{λ_0} is, again, the typical turbulent speed at these smallest eddies. We recall that, from the work equation (eq. 7.44), the variation of kinetic energy density in the fluid per unit time is equal to Ψ, in absence of other sources of energy transfer. Thus, in our present situation, we have

$$\rho\dot{\mathcal{E}} \sim \eta \frac{v_{\lambda_0}^2}{\lambda_0^2} \tag{8.20}$$

and, further substituting eq. 8.17 for $\dot{\mathcal{E}}$ at λ_0, we finally obtain that the **dissipation scale** is

$$\lambda_0 \sim \frac{\nu}{v_{\lambda_0}}. \tag{8.21}$$

How large is the above scale with respect to the other scales of the system? In §7.4.1 we have found that the coefficient of kinematic viscosity can be written as $\nu \sim \ell v_{\rm th}$, where ℓ is the mean particle free-path and $v_{\rm th}$ the thermal speed. Therefore we have that

$$\lambda_0 \sim \frac{c_{\rm s}}{v_{\lambda_0}}\ell, \tag{8.22}$$

where we have substituted the thermal speed with the sound speed. As mention, the typical (large-scale) turbulent speed (V) in the ISM is mildly supersonic, e.g. the H I line is broadened to $\approx 10\,{\rm km\,s^{-1}}$, which exceeds the velocity of the WNM by a factor ≈ 1.5 and the CNM sound speed by an order of magnitude. Given that the velocity of the eddies scales with $\lambda^{1/3}$ we can expect the velocity associated to the dissipation scale (v_{λ_0}) to be much lower than V. This means that the typical velocity of the dissipation-scale eddies will be likely subsonic. As a consequence, eq. 8.22 tells us the dissipation scale is larger than the mean free-path (ℓ is about $10^{13}\,{\rm cm}$ for HI at the average density of the CNM, eq. 2.3).

We can also use the Kolmogorov-Obukhov law written for the dissipation scale to derive an expression for the velocity of the λ_0-sized eddies. This equation can then be substituted in eq. 8.21 to obtain

$$\lambda_0 \sim \frac{\nu L^{1/3}}{V\lambda_0^{1/3}} \tag{8.23}$$

and thus, recalling the definition of the Reynolds number (eq. 7.65),

$$\lambda_0 \sim \frac{L}{{\rm Re}^{3/4}}, \tag{8.24}$$

which clearly shows that $\lambda_0 \ll L$ in any situation. In the particular case of the ISM, recalling the estimate of ${\rm Re} \sim 3\times10^5$ for $L = 100\,{\rm pc}$ we obtain $\lambda_0 \sim 10^{16}$, much smaller than L but larger than the mean free path. In general, we expect the dissipation scale in the ISM to be on sub-parsec scales. This poses serious problems if one aims to trace the turbulence cascade with hydrodynamical simulations as sub-parsec scales are out of reach of the most advanced simulations when modeling the entire galactic discs. Simulations of small portion of the disc can reach resolutions of the order of a parsec. In general therefore, the dissipation scale is not resolved in simulations and the dissipation that is always present (because of numerical effects) will occur on scales much larger than the physical scale (and thus at quite different speeds). A problematic consequence of this is that simulations reaching different resolutions can return quantitatively different results. If this is the case, we say that the **resolution convergence** has not been achieved. A test of any numerical investigation of this kind is therefore to perform simulations at progressively higher resolutions to test whether they tend to converge to the same results or not.

Another important consequence of eq. 8.24 is that between L and λ_0 there are typically several order of magnitudes. This interval of scales is called **inertial range** and confirms the applicability of the whole theory. The condition reads

$$L \gg \lambda \gg \lambda_0. \tag{8.25}$$

L is sometimes referred to as the energy scale because most of the energy is on these scales, λ is the range where the cascade takes place and at which the theory has its applicability (inertial range). Note, finally, that for increasing Reynolds number, λ_0 decreases (eq. 8.24) and therefore the inertial range expands.

8.3.3. Energy spectrum. The main feature of the Kolmogorov-Obukhov theory is the prediction that turbulence redistributes velocities (and energy) as *power laws* of the length scales. The identification of these power laws is a powerful tool to study turbulent fluids and provides clues about the type of turbulence that is present in the fluid. Equation 8.18 is not much used in practice, instead one usually employs energy as a function of the wave-number. For each eddy of size λ one can define, as usual, a wave-number $k = 2\pi/\lambda$. In a turbulent medium, all energy is essentially in the form of kinetic energy whose specific value will have the following dependencies:

$$\mathcal{E} \sim v_\lambda^2 \sim \dot{\mathcal{E}}^{2/3}\lambda^{2/3} \sim \dot{\mathcal{E}}^{2/3}k^{-2/3}, \tag{8.26}$$

where the second equality comes from the Kolmogorov-Obukhov law.

As mentioned, energy is stored in large eddies because of the dependence on $\lambda^{2/3}$ thus eq. 8.26 represents nearly *all* the energy[7] stored in eddies with sizes equal to or smaller than λ, or equivalently, with wave-numbers equal to or larger than k. With this definition in mind, we can now write an equation for the power spectrum of the energy as follows:

$$\mathcal{E} = \int_\infty^k E(k')\mathrm{d}k' \sim \dot{\mathcal{E}}^{2/3}k^{-2/3}, \tag{8.27}$$

where $E(k)$ is the **energy power spectrum** and the last equality comes from eq. 8.26. Clearly, for eq. 8.27 to be the verified one needs to have that the power spectrum is

$$E(k) \propto k^{-5/3}. \tag{8.28}$$

Then, $E(k)\mathrm{d}k$ will represent the specific kinetic energy stored in eddies with wave-numbers between k and $k+\mathrm{d}k$. Equation 8.28 shows that the energy power spectrum of Kolmogorov turbulence is a power law of k with exponent $-5/3$. This is a key result and it is often taken as reference for any other models of turbulent flows.

8.3.4. Turbulence in real gases. In a gas where the density varies significantly as a consequence of turbulent motions the assumptions made in the previous sections are no longer valid. In particular, one can not visualize chaotic motions only in terms of eddies but different structures must be present. The analytical treatment of so-called **compressible turbulence** is very involved and somewhat incomplete. A way to proceed is to assume that there are two types of **turbulent modes** (Helmholtz decomposition). The first is of the same kind as before and it is called solenoidal mode, it occurs when $\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0$ and it consists, as before, of eddies, so entirely transverse motions. The second mode is *compressible*, it occurs when $\boldsymbol{\nabla}\times\boldsymbol{u} = 0$ and it is entirely longitudinal. Unfortunately, one cannot decompose the medium fully in this way because, in general, there is a continuous transfer of energy between the two modes. An important consequence of this transfer is that

[7]The reader can verify that the energy stored in eddies between $\lambda/2$ and λ is larger than twice the energy stored in eddies between 0 and $\lambda/2$.

dissipation can take place at different scales. This is a rather inconvenient property of compressible turbulence.

A further complication is due to the fact that turbulence, in the ISM in particular, can be *supersonic*. As we saw in Chap. 4, supersonic motions produce shocks, which in turn compress the medium greatly and, under certain circumstances, can make it radiate efficiently. This is especially true for shocks at relative low speeds ($\sim 100\,\mathrm{km\,s^{-1}}$) that tend to be highly radiative (§4.3). Turbulence can cause a myriad of shocks in the ISM, although in most cases at speeds lower than $100\,\mathrm{km\,s^{-1}}$. This produces a consequent efficient dissipation of the kinetic energy. These modes are called **thermal modes** and the whole process is a shock cascade. In principle, these thermal modes could even dominate the turbulence dissipation in the ISM. However, it is not clear how much of this dissipation is effectively taking place. One way that the ISM has to prevent strong compressions (and thus efficient radiation and loss of kinetic energy) is the presence of magnetic fields. We see later (§9.6) that a magnetic field of the order of the one observed in the ISM can indeed greatly prevent gas compressions from shock waves.

Other components of the ISM can introduce further modes, these include gravity, the magnetic field itself and the change of excitation and ionisation of particles in the medium. To include all these effects in the theory, one necessarily needs to resort to hydrodynamical simulations. This is not, however, a *final* solution because simulations can have important limitations in this context, for instance, as mentioned, due to their inadequacy of resolving the dissipation scale. A general result of hydrodynamic simulations is that the energy power spectrum, in the presence of compressibility of the medium, tends to be steeper than the Kolmogorov value, $E(k) \propto k^{-\alpha}$, with $\alpha \approx 2$.

Finally, we mention that an assumption that it is often employed to simplify calculations is the introduction of a **turbulent pressure** defined as

$$P_{\mathrm{turb}} \equiv \rho \sigma_{\mathrm{turb}}^2, \tag{8.29}$$

where σ_{turb} is the typical turbulent speed of the medium. For instance, in the neutral atomic ISM this would be $\approx 10\,\mathrm{km\,s^{-1}}$, as observed with the 21-cm H I line, while in the molecular gas is $\approx 1\,\mathrm{km\,s^{-1}}$, as observed from the broadening of several emission lines. Despite all the complexity described above, this approach can capture global behaviour of turbulence and allow us to derive order-of-magnitude values of relevant physical properties. For instance, it is useful in the equilibrium of molecular clouds or in the study of the hydrostatic equilibrium of turbulent gaseous discs (§3.1).

8.3.5. The Kolmogorov theory in the ISM. In the light of the complications outlined in the last Section, it may appear unseemly to use the Kolmogorov theory to investigate the turbulent properties of the ISM. However, it is true that in some aspects this simplified approach works remarkably well, suggesting that the ISM as a whole does not behave too differently than an incompressible fluid, at least as far as the turbulent motion is concerned. For instance, structures in the ISM such as the Fourier-transform power spectra of the distribution of neutral hydrogen emission display power laws of the kind that one expects from a Kolmogorov-like theoretical description [ES04]. Moreover, the velocity dispersions of self-gravitating clouds are also shaped as power laws of their sizes with a similar slope [Lar81].

Here, we use the Kolmogorov theory to search for an estimate of the energy dissipated by turbulence. The typical velocity of the turbulent ISM is the observed broadening of a line, in particular the 21-cm line. Its value is, as mentioned, typically $\sigma_{\rm HI} \approx 10\,{\rm km\,s^{-1}}$ and, thus, the rate of energy dissipation per unit volume will be

$$\rho\frac{\sigma_{\rm ISM}^3}{L} \sim 5\times10^{-27}\mu\left(\frac{n}{1\,{\rm cm^{-3}}}\right)\left(\frac{\sigma_{\rm ISM}}{10\,{\rm km\,s^{-1}}}\right)^3\left(\frac{L}{100\,{\rm pc}}\right)^{-1}\,{\rm erg\,s^{-1}\,cm^{-3}}, \tag{8.30}$$

a value that should be considered a lower limit given that we are assuming incompressibility. Indeed, a compressible fluid like the ISM may have more modes of dissipation (e.g. thermal modes) and can require somewhat more energy injection.

We can now ask ourselves what energy source can provide turbulent energy in the ISM at least at the rate predicted by the Kolmogorov theory. The most likely source is stellar feedback, its energy input in the Milky Way can be calculated by assuming a supernova rate SNR $\approx 10^{-2}$ supernovae per year, an energy per supernova of $E_{\rm SN} \simeq 10^{51}\,{\rm erg}$ and the total volume of the (star-forming and turbulent) disc[8]. This leads to

$$\dot{K}_{\rm SN} \sim 10^{-25}\eta\left(\frac{E_{\rm SN}}{10^{51}\,{\rm erg}}\right)\left(\frac{\rm SNR}{10^{-2}{\rm yr^{-1}}}\right)\left(\frac{V}{\pi10^2\times0.2\,{\rm kpc}^3}\right)^{-1}\,{\rm erg\,s^{-1}\,cm^{-3}} \tag{8.31}$$

where η is the efficiency of transferring kinetic energy to the ISM. Thus, from this estimate, it appears that an efficiency of at least $\eta \sim 5\%$ would be required for supernovae to properly maintain the observed turbulence. Stellar winds from massive stars can also contribute some energy.

As we see in §5.4, the above efficiency is not straightforwardly achieved by supernova remnants. This, together with the fact that in the outer parts of discs the supernova rate decreases rapidly (while the turbulence seems still present) have brought to consider other possible energy sources. We list here a few of them, the contributions of which should be taken as very rough orders of magnitude: shocks in spiral arms ($\sim 10^{-26}-10^{-27}\,{\rm erg\,s^{-1}\,cm^{-3}}$), shear (gravitational energy) ($\sim 1\times10^{-27}\,{\rm erg\,s^{-1}\,cm^{-3}}$), magnetorotational instability ($\sim 10^{-28}-10^{-29}\,{\rm erg\,s^{-1}\,cm^{-3}}$). The latter in particular has been advocated to be important in the outer parts of galaxy discs (see §6.4).

[8]This latter is a very rough approximation and one may want to consider instead a radial dependence of the supernova energy injection.

CHAPTER 9

Astrophysical magneto-hydrodynamics

In this Chapter, we introduce and discuss a key property of astrophysical gases: the presence of a pervasive magnetic field. Magnetic fields in astrophysical gases arise due to the development of electrical currents due to slight differences in velocity between electrons and ions. Once magnetic fields have been established in a plasma they are very persistent. Magnetic fields have no sources (magnetic monopoles) and this is a the origin of their persistence. If such sources were present and could move freely, they would rush around to smooth inhomogeneities and make the field disappear. This is what happens for electric fields: if one such field establishes in the gas, *electrons* move from regions where there is a concentration of negative charges to regions with a concentration of positive charges and re-establish neutrality. In practice, because of this easiness of electrons to move around in gases, macroscopic electric fields never develop in large-scale astrophysical gases like the ISM and the IGM. The only astrophysical objects were this can take place are certain types of stars, like neutron stars. We start this Chapter by reviewing the fundamental assumptions at the basis of **magneto-hydrodynamics** (MHD): a fluid dynamic theory that incorporates magnetic fields. We then re-derive the Euler equation in MHD, describe the development in a plasma of MHD waves and see the effects of magnetic fields in astrophysical shocks.

9.1. Key assumptions

The introduction of magnetic fields in hydrodynamics can be done thanks to a number of assumptions and approximations that we outline here. These allow us to treat the **plasma** (gas with free electrical charges) as a continuum (as done before for the ideal gases, Chap. 2) subject to a new force generated by the magnetic field. The continuum assumption is applicable to astrophysical gases but its application can be questioned in some terrestrial fluids. The branch of physics that studies flows of media with charged particles that *do not* fulfill the MHD assumption is called **plasma physics** and is not treated in these notes.

We start from the Maxwell equations written in cgs units[1]:

$$\boldsymbol{\nabla} \cdot \boldsymbol{E} = 4\pi\rho_{\rm e} \tag{9.1}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0 \tag{9.2}$$

$$\boldsymbol{\nabla} \times \boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t} \tag{9.3}$$

$$\boldsymbol{\nabla} \times \boldsymbol{B} = \frac{4\pi}{c}\boldsymbol{J} + \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t}, \tag{9.4}$$

[1]We remind that in cgs, the units of the magnetic and the electric fields are the same. Also the permittivity and permeability of the vacuum are $\epsilon_0 = 1/(4\pi c)$ and $\mu_0 = 4\pi/c$, respectively.

where $\boldsymbol{E}$ and $\boldsymbol{B}$ are the electric and magnetic fields, $\rho_{\rm e}$ is the density of charges and $\boldsymbol{J}$ is the current. We modify these equations to adapt them to a fluid medium with properties akin to an astrophysical gas. In order to do this we take four steps (assumptions), described in the following sections.

9.1.1. Charge neutrality. The first assumption of MHD is **charge neutrality**. With this term we mean that a *large enough* portion of gas will have the same number of positive and negative charges and there will be no segregation between them. No segregation of charges implies no electric field, thus essentially this assumption leads to the unimportance of electric field in astrophysical gases mentioned above. In order to understand the conditions under which charge neutrality is fulfilled, let us assume that segregation has occurred in a region of the plasma and let us calculate the size of this region. If we have segregation of electrons and ions over a distance x, there will be an electric field that, from the first of the Maxwell equations, can be approximately written as

$$E \sim 4\pi\rho_{\rm e} x. \tag{9.5}$$

Electrons within this region will feel a Coulomb acceleration of the order

$$\ddot{x} = \frac{f_{\rm C}}{m_{\rm e}} = -\frac{eE}{m_{\rm e}} \sim -\frac{4\pi n_{\rm e} e^2}{m_{\rm e}} x, \tag{9.6}$$

where $f_{\rm C}$ is the Coulomb force, e and $m_{\rm e}$ are the elementary electric charge and the electron mass and $n_{\rm e}$ is the electron number density, such that $\rho_{\rm e} = n_{\rm e} e$.

The expression on the r.h.s. of eq. 9.6 is that of an harmonic oscillator with frequency

$$\omega_{\rm p} \equiv \left(\frac{4\pi n_{\rm e} e^2}{m_{\rm e}}\right)^{1/2}, \tag{9.7}$$

which is called **plasma frequency** and describes the reaction of the medium to the segregation of charges. The inverse of the plasma frequency is the time that the plasma needs to re-establish charge neutrality. So a potential segregation of charges could, in principle be maintained only for times shorter than $1/\omega_{\rm p}$. If we consider the thermal speed ($v_{\rm th}$) as the typical speed of the electrons in the plasma, the size of the segregation region is then

$$\lambda_{\rm D} \equiv \frac{v_{\rm th}}{\omega_{\rm p}} = \left(\frac{k_{\rm B} T_{\rm e}}{4\pi n_{\rm e} e^2}\right)^{1/2}, \tag{9.8}$$

which is the so-called **Debye length**. We can use typical values for the ISM of a galaxy in eq. 9.9 to realize that its value is infinitely small in typical astrophysical conditions:

$$\lambda_{\rm D} \simeq 10^3 \left(\frac{T_{\rm e}}{10^4\,{\rm K}}\right)^{1/2} \left(\frac{n_{\rm e}}{1\,{\rm cm}^{-3}}\right)^{-1/2} {\rm cm}. \tag{9.9}$$

We therefore conclude that the size of any realistic portions of an astrophysical fluid will greatly exceed the Debye length and therefore electric fields cannot be permanently established. In other words, we expect to have charge neutrality everywhere in the ISM of galaxies as well as in the IGM. We write this property of the plasma as follows:

$$\rho_{\rm e} = Z_{\rm i} e n_{\rm i} - e n_{\rm e} = 0, \tag{9.10}$$

where Z_i is the average number of positive charges in the ion particles and n_i is the ion number density. Note that the Debye length is so small that eq. 9.10 can be, in fact, considered valid within the single fluid element.

9.1.2. Negligible displacement current. Equation 9.4 states that magnetic fields can come both from electric currents or from variations of the electric field (displacement current). Here we show that the latter term can be neglected in most practical cases. Consider the ratio between the displacement current term and the curl of $\boldsymbol{B}$ in orders of magnitude:

$$\frac{\frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t}}{\boldsymbol{\nabla}\times\boldsymbol{B}} \sim \frac{EL}{cB\tau} \sim \frac{V}{c}\frac{E}{B}, \tag{9.11}$$

where L and τ are the typical distance and time over which, respectively, B and E vary appreciably. With these quantities we define a velocity (V) of the fluid relevant for variations of magnetic and electric fields (rightmost term in eq. 9.11). We can then approximate eq. 9.3 to derive an order-of-magnitude ratio between electric and magnetic fields:

$$\frac{E}{L} \sim \frac{B}{c\tau} \implies \frac{E}{B} \sim \frac{V}{c}. \tag{9.12}$$

Thus eq. 9.11 becomes

$$\frac{\frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t}}{\boldsymbol{\nabla}\times\boldsymbol{B}} \sim \left(\frac{V}{c}\right)^2 \ll 1, \tag{9.13}$$

i.e. negligible for any subluminal motion such as the motion of the fluids that we are investigating here. The fact that the second term on the r.h.s. of the forth of Maxwell equations (eq. 9.4) is much lower than the curl of $\boldsymbol{B}$ obviously implies that the first term must be of the order of this curl. Thus the last of the Maxwell equations simply reverts to the **Ampére law**:

$$\boldsymbol{\nabla}\times\boldsymbol{B} = \frac{4\pi}{c}\boldsymbol{J}. \tag{9.14}$$

Clearly, the result just found applies generically to a number of circumstances not involving relativistic motions. In fact, the typical situation in which the displacement current necessarily has to be introduced is in the description of electromagnetic waves. Equation 9.13 also implies a very important property of magnetic fields in fluids and astrophysical gases, in particular: it explicitly shows that the magnetic field is *generated by electric currents*. Thus, the observations of widespread magnetic fields in astrophysical objects imply the presence of widespread currents. A current in a plasma is produced by the relative motion between electrons and ions. In the following Section, we calculate the order of magnitude of this motion.

9.1.3. Single-component MHD. The density of current in a plasma can be written as follows:

$$\boldsymbol{J} = Z_\mathrm{i} e n_\mathrm{i}\boldsymbol{u}_\mathrm{i} - e n_\mathrm{e}\boldsymbol{u}_\mathrm{e} = -e n_\mathrm{e}(\boldsymbol{u}_\mathrm{e} - \boldsymbol{u}_\mathrm{i}) = -e n_\mathrm{e}\boldsymbol{u}'_\mathrm{e}, \tag{9.15}$$

where $\boldsymbol{u}_\mathrm{i}$ and $\boldsymbol{u}_\mathrm{e}$ are the velocities of the fluid ions and electrons, respectively, and we have defined with $\boldsymbol{u}'_\mathrm{e}$ the speed of electrons in the reference frame of the ions. The second equality of eq. 9.15 is a consequence of the charge neutrality (§9.1.1). Note that $\boldsymbol{u}_\mathrm{i}$ and $\boldsymbol{u}_\mathrm{e}$ are *not* the particle velocities but the velocities of fluid elements (§2.1). The actual *elementary particles* are moving at random and

at speeds that differ greatly between ions and electrons. However, in the fluid motion, random velocities cancel out and only contribute to the gas temperature (for ideal gases, see §1.2). The currents produced by these microscopic movements are also irrelevant for the global behaviour of the fluid because they cancel out as well. What matters globally is, instead, the motion of fluid elements. To see this we can make the assumption that we have ideally *two types* of fluid elements: positive (made up of ions) and negative (made up of electrons) elements. These two components have in principle different *fluid* velocities ($\boldsymbol{u}$) and this difference may produce a global large-scale current. An alternative way to look at $\boldsymbol{u}_\mathrm{i}$ and $\boldsymbol{u}_\mathrm{e}$ is to consider them as the velocities of, again ideally, *infinitely cold* particles, i.e. the velocities that elementary particles (ions and electrons) would have if we could take away their random (thermal) motions. We can call these "cold ions" and "cold electrons".

We now estimate the order of magnitude of the relative velocity between positive and negative fluid elements in the magnetised ISM of the Milky Way. This can be done by using eq. 9.15 and by writing $\boldsymbol{J}$ approximating (in orders of magnitudes) eq. 9.14. We obtain

$$\boldsymbol{u}'_\mathrm{e} \sim \frac{cB}{4\pi e n_\mathrm{e} L} \sim 5\times10^{-6}\left(\frac{B}{3\times10^{-6}\,\mathrm{G}}\right)\left(\frac{n_\mathrm{e}}{1\,\mathrm{cm}^{-3}}\right)^{-1}\left(\frac{L}{1\,\mathrm{pc}}\right)^{-1}\mathrm{cm\,s}^{-1}, \tag{9.16}$$

where we have indicated typical values for the various quantities. With the velocity in eq. 9.16, the displacement between fluid elements of different charges after 1 Myr will be of the order of one thousand kilometers, completely negligible in astrophysical scales. We conclude that, for the global motion of the fluid, there is no reason to distinguish between the fluid velocities of the ions ($\boldsymbol{u}_\mathrm{i}$) and that of the electrons ($\boldsymbol{u}_\mathrm{e}$). These velocities are effectively the same and thus we do not separate positive and negative fluid elements, i.e. the fluid can be described as a *single component* moving at velocity $\boldsymbol{u} \simeq \boldsymbol{u}_\mathrm{i} \simeq \boldsymbol{u}_\mathrm{e}$. On the other hand, it is the tiny difference between the two components shown in eq. 9.16 that is at the origin of the astrophysical magnetic fields and thus, from this point of view, it is of great importance. The possibility of describing the fluid as a single component with a global common motion for ions, electrons and also neutral particles is at the basis of the MHD treatment (**single-component MHD**).

9.2. The freezing of the magnetic field

The last assumption that is needed to finalise the MHD equations is that the so-called **electrical conductivity** of the fluid is very large. As we see below, this condition leads to a fundamental equation of MHD that describes the conservation of the magnetic field flux with the motion of the fluid. To see this we start by writing the density of currents as

$$\boldsymbol{J}' = -e n_\mathrm{e} \boldsymbol{u}'_\mathrm{e}, \tag{9.17}$$

where this time, we have explicitly indicated it with the prime ($'$) symbol, being this the current density in the reference system of the "cold ions" (fluid elements with positive electric charge).

Let us now consider the **Ohm law** written for a density of charged particles, which reads

$$\boldsymbol{J}' = \sigma \boldsymbol{E}', \tag{9.18}$$

where σ is the electrical conductivity[2]. We transform eq. 9.18 from the ion rest frame to an external frame of reference (the frame in which the fluid moves at velocity $\boldsymbol{u}$). Being $\boldsymbol{J}'$ dependent only on velocities and the electron density (eq. 9.15), in the regime of non-relativistic speeds ($u/c \ll 1$) using a Galilean transformation, we have $\boldsymbol{J} = \boldsymbol{J}'$. On the contrary, the electric field follows the **Lorentz transformation** that, for the non-relativistic case[3], leads to

$$\boldsymbol{J} = \sigma\left(\boldsymbol{E} + \frac{1}{c}\boldsymbol{u}_\mathrm{i} \times \boldsymbol{B}\right) = \sigma\left(\boldsymbol{E} + \frac{1}{c}\boldsymbol{u} \times \boldsymbol{B}\right), \tag{9.19}$$

where $\boldsymbol{u}_\mathrm{i}$ is the velocity of the "cold ions", which is the same as the fluid velocity for what established in §9.1.3, hence the second equality. Note that the reason why eq. 9.19 is different from eq. 9.18 is that a proportionality like the latter applies only to a medium at rest (so ions in their own rest frame). In the external (lab) frame where the ions are moving currents are created also by the presence of $\boldsymbol{B}$ (if we move a metal wire through a magnetic field a current is generated even if there is no voltage applied).

We use the above result to find the first fundamental equation of MHD. From the Ampére law (eq. 9.14) and the definition of current density (eq. 9.19) we have that

$$\boldsymbol{\nabla} \times \boldsymbol{B} = \frac{4\pi\sigma}{c}\left(\boldsymbol{E} + \frac{1}{c}\boldsymbol{u} \times \boldsymbol{B}\right). \tag{9.20}$$

We take the curl of this equation and make a substitution on the l.h.s. by recalling the identity C.18. We further assume that the conductivity σ is spatially constant (being a property of the whole fluid) and obtain the following fundamental equation:

$$\frac{\partial \boldsymbol{B}}{\partial t} + \boldsymbol{\nabla} \times (\boldsymbol{B} \times \boldsymbol{u}) = \frac{c^2}{4\pi\sigma}\nabla^2\boldsymbol{B} \equiv \eta\nabla^2\boldsymbol{B}, \tag{9.21}$$

where we have also introduced the **electric resistivity** $\eta \equiv c^2/(4\pi\sigma)$, inversely proportional to the conductivity.

Equation 9.21 is fully analogous to eq. 7.57 once we substitute the magnetic field with the vorticity and the electric resistivity with the viscosity coefficient. In Chap. 7 we have shown that, in the presence of viscosity, the vorticity is not conserved, i.e. the Kelvin theorem (eq. 7.61) is not fulfilled. Here, we encounter a very similar circumstance: if we define the **magnetic flux** as

$$\phi_\mathrm{B} \equiv \int_S \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S}, \tag{9.22}$$

we can now say that the magnetic flux (i.e. the number of magnetic field lines that threads a generic surface S *moving with the fluid*) can be conserved only if the resistivity term in eq. 9.21 is negligible. It is important therefore to have an handle

[2]This law comes from the continuum limit of the standard Ohm law of an electrical circuit $I = V/R$ (the intensity I equals the potential difference V divided by the resistance R of the material). The generalisation used here is sometimes referred to as the **Kirchoff law**. In general, the conductivity (inverse of the resistivity) can be a tensor and this can also be non-diagonal if the fluid is non-Newtonian (see §7.3.4).

[3]We remind that the Lorentz transformations of the orthogonal (with respect to $\boldsymbol{u}$) magnetic and electric fields are in general $\boldsymbol{B}'_\perp = \gamma(\boldsymbol{B}_\perp - \boldsymbol{u} \times \boldsymbol{E}_\perp/c)$ and $\boldsymbol{E}'_\perp = \gamma(\boldsymbol{E}_\perp + \boldsymbol{u} \times \boldsymbol{B}_\perp/c)$, where γ is the special-relativity coefficient $\gamma \equiv 1/\sqrt{1 + v^2/c^2}$. The parallel fields are unchanged: $\boldsymbol{B}'_\parallel = \boldsymbol{B}_\parallel$ and $\boldsymbol{E}'_\parallel = \boldsymbol{E}_\parallel$.

on the typical values that we can expect for the r.h.s. of eq. 9.21, a term that we can call **Ohmic dissipation** (in analogy to the viscous dissipation of Chap. 7).

9.2.1. Estimating the electrical conductivity. We can estimate the conductivity of a portion of the ISM as follows. Let us write again the equation of the current density as

$$\boldsymbol{J}' = \sigma \boldsymbol{E}' = -n_{\rm e} e \boldsymbol{u}'_{\rm e}, \tag{9.23}$$

where $\boldsymbol{u}'_{\rm e}$ is the fluid electron velocity in the rest frame of the ions. In the same reference frame, the force acting on the electrons is

$$m_{\rm e} \frac{{\rm d}\boldsymbol{u}'_{\rm e}}{{\rm d}t} = -e\boldsymbol{E}' - f_{\rm D}, \tag{9.24}$$

where $f_{\rm D}$ is the drag force between electrons and ions and we assume that other forces (e.g. pressure and gravity) are negligible. Here, we are implicitly considering that the medium is fully ionised and the neutral particles do not contribute the force budget. For collisions between electrons and ions we can write the **drag force** as

$$f_{\rm D} \sim m_{\rm e} \frac{u'_{\rm e}}{\tau_{\rm coll}(e,i)}, \tag{9.25}$$

where $\tau_{\rm coll}(e,i)$ is the collision time for ion-electron interactions: $\tau_{\rm coll}(e,i) = 1/(n_{\rm i} A_{\rm e,i} v_{\rm e})$, recalling §1.2.1, with $A_{\rm e,i}$ the cross section of ion-electron collisions (eq. 1.9).

If we assume that the electrons are in equilibrium (they are not accelerating), the l.h.s. of eq. 9.24 is zero and we obtain that

$$E' \sim \frac{m_{\rm e} u'_{\rm e}}{e\tau_{\rm coll}(e,i)} = \frac{m_{\rm e} n_{\rm i} A_{\rm e,i} u'_{\rm e} v_{\rm e}}{e}, \tag{9.26}$$

from which we can finally derive the conductivity using eq. 9.23:

$$\sigma \sim \frac{n_{\rm e} e^2 \tau_{\rm coll}(e,i)}{m_{\rm e}} = \frac{n_{\rm e} e^2}{m_{\rm e}^{1/2} n_{\rm i} A_{\rm e,i} (k_{\rm B} T_{\rm e})^{1/2}}, \tag{9.27}$$

where we have used the thermal speed for the motion of the electrons. Obviously, if we find ourselves in the opposite regime of a medium mostly neutral (typical, for instance, of a molecular cloud) the above equations have to be revised with the neutral particles in place of the ions and we would obtain

$$\sigma \sim \frac{n_{\rm e} e^2 \tau_{\rm coll}(e,n)}{m_{\rm e}} = \frac{n_{\rm e} e^2}{m_{\rm e} n_{\rm n} A_{\rm e,n} v_{\rm e}}, \tag{9.28}$$

where $A_{\rm e,n}$ is the cross section for neutral-electron collisions (§1.2.1), which is smaller than $A_{\rm e,i}$.

We can estimate the time that it would take to the ISM to lose its magnetic field via Ohmic dissipation ($\tau_{\rm OD}$) by rewriting eq. 9.21 in orders of magnitude:

$$\frac{B}{\tau_{\rm OD}} \sim \frac{c^2 B}{4\pi\sigma L^2} \sim \frac{m_{\rm e} c^2 B}{4\pi n_{\rm e} e^2 \tau_{\rm coll}(e,i) L^2}. \tag{9.29}$$

If we now take typical astronomical values for the quantities in this equation and use the collision time in eq. 1.13 we obtain

$$\tau_{\rm OD} \sim 1 \times 10^{22} \left(\frac{n_{\rm e}}{1\,{\rm cm}^{-3}}\right) \left(\frac{\tau_{\rm coll}(e,i)}{10^4\,{\rm s}}\right) \left(\frac{L}{\rm pc}\right)^2 {\rm yr}, \tag{9.30}$$

i.e. an enormously long time, from which we can conclude that the ISM is a perfect conductor.

What we just found allows us to safely modify eq. 9.21 to obtain

$$\frac{\partial \boldsymbol{B}}{\partial t} + \boldsymbol{\nabla} \times (\boldsymbol{B} \times \boldsymbol{u}) = 0, \tag{9.31}$$

which is the exact analogous of the Kelvin theorem once we substitute the vorticity ω with the magnetic field $\boldsymbol{B}$. In §7.6.1, we saw how the equation with vorticity implies the conservation of the vorticity through a surface moving with the fluid. Here we can extend this property to the magnetic flux (eq. 9.22) and write that

$$\frac{\mathrm{d}\phi_\mathrm{B}}{\mathrm{d}t} = 0, \tag{9.32}$$

which is the expression of the so-called magnetic flux *freezing* in the fluid. This implies that there is a tight relation between the strength of the magnetic field and the gas (and its motion).

Consider for instance a magnetised gas cloud with a roughly constant magnetic field of magnitude B. The flux through the equatorial surface of the cloud will be of order $\phi_\mathrm{B} \sim \pi B r^2$, with r radius of the cloud. Equation 9.31 then implies a proportionality of this kind

$$B \propto \frac{1}{r^2}. \tag{9.33}$$

Thus, if the cloud undergoes compression (for instance due to gravitational collapse) the field can dramatically increase with respect to the initial value. As a consequence, for instance, in the ISM of a spiral galaxy we can expect denser regions to have larger values of the magnitude of the magnetic field with respect to the mean value. The enhancement of the magnetic field strength in spiral arms, for instance, is likely caused by this phenomenon.

9.3. Euler equation with magnetic terms

The fundamental equation that governs the motion of charge particles is the equation of the **Lorentz force**:

$$\boldsymbol{f}_\mathrm{L} = q\left(\boldsymbol{E} + \frac{1}{c}\boldsymbol{u} \times \boldsymbol{B}\right), \tag{9.34}$$

where q is a generic charge moving at velocity $\boldsymbol{u}$. Let us now consider a fluid made of "cold ions" and "cold electrons" (positively charges and negatively charged fluid elements) and write the Lorentz equation describing the force that these particles are subject to:

$$\boldsymbol{f}_\mathrm{L} = Z e n_\mathrm{i}\left(\boldsymbol{E} + \frac{1}{c}\boldsymbol{u}_\mathrm{i} \times \boldsymbol{B}\right) - e n_\mathrm{e}\left(\boldsymbol{E} + \frac{1}{c}\boldsymbol{u}_\mathrm{e} \times \boldsymbol{B}\right). \tag{9.35}$$

We can use the assumption of charge neutrality (eq. 9.10) and the definition of current density (eq. 9.15) to obtain

$$\boldsymbol{f}_\mathrm{L} = -\frac{1}{c} e n_\mathrm{e}(\boldsymbol{u}'_\mathrm{e} \times \boldsymbol{B}) = \frac{1}{c}(\boldsymbol{J} \times \boldsymbol{B}) = \frac{1}{4\pi}(\boldsymbol{\nabla} \times \boldsymbol{B}) \times \boldsymbol{B}, \tag{9.36}$$

where the last equality comes from the last of Maxwell equations (eq. 9.4) and the assumption of negligible displacement current (§9.1.2).

The fact that we can describe the fluid as a single gas component (eq. 9.1.3) allows us also to write a *single* Euler equation valid for all the particles. We do this by simply adding the Lorentz force from eq. 9.36 to eq. 2.16 to obtain

$$\rho\frac{\partial \boldsymbol{u}}{\partial t} + \rho(\boldsymbol{u}\cdot\nabla)\boldsymbol{u} = -\boldsymbol{\nabla}P - \rho\boldsymbol{\nabla}\phi + \frac{1}{4\pi}(\boldsymbol{\nabla}\times\boldsymbol{B})\times\boldsymbol{B}, \tag{9.37}$$

which is a form of the Euler equation including the magnetic field and describes the global motion of ions, electrons and neutrals (if present).

A second version of this Euler equation with magnetic field forces can be obtained by using the eq. C.18 and it reads

$$\rho\frac{\partial \boldsymbol{u}}{\partial t} + \rho(\boldsymbol{u}\cdot\nabla)\boldsymbol{u} = -\boldsymbol{\nabla}P - \rho\boldsymbol{\nabla}\phi - \frac{1}{8\pi}\boldsymbol{\nabla}B^2 + \frac{1}{4\pi}\boldsymbol{B}\cdot\boldsymbol{\nabla}\boldsymbol{B}, \tag{9.38}$$

where the last two terms correspond to the **magnetic pressure** and the **magnetic tension** forces, respectively. The magnetic pressure can therefore be written as

$$P_{\rm mag} = \frac{B^2}{8\pi} \tag{9.39}$$

and describes the reaction of a magnetised medium to compression, in analogy to the thermal pressure. The magnetic tension, instead, describes the reaction to deformations. In practice, a magnetic field with some coherent orientation can be visualized in terms of force lines. These lines have a *tension* similar to the tension of hard strings. Any attempt to deform the field (as a consequence of the motion of the gas) will meet a reaction of the magnetic tension that opposes this deformation. In very simplified terms, magnetic pressure and tension tend to dominate the magnetic force in the presence of, respectively, chaotic and ordered magnetic fields.

9.3.1. Euler equation by components. It is useful to write the Euler equation with the magnetic field by components. We use the momentum form of the equation (§7.2), which reads

$$\frac{\partial(\rho u_i)}{\partial t} + \frac{\partial}{\partial x_j}(\rho u_i u_j) = -\frac{\partial P}{\partial x_i} - \rho\frac{\partial\phi}{\partial x_i} + \frac{\partial}{\partial x_j}T_{\rm ij}, \tag{9.40}$$

where we have introduced a tensor to describe the magnetic terms:

$$T_{ij} \equiv \frac{1}{4\pi}B_iB_j - \frac{1}{8\pi}B^2\delta_{ij}. \tag{9.41}$$

It is easy to prove that the spatial derivative of 9.41 in the j direction, leads to the expression for the Euler equation in eq. 9.38

9.4. Virial theorem with the magnetic field

We now derive an equation for the virial theorem for a portion of a magnetised fluid. We start by multiplying[4] all the terms in eq. 9.40 by x_i and integrating over a generic volume V. The first term on the l.h.s. leads to the definition of the **virial**:

$$\int_V x_i\frac{\partial\rho u_i}{\partial t}\mathrm{d}V = \frac{1}{2}\frac{\mathrm{d}^2 I}{\mathrm{d}t^2}, \tag{9.42}$$

where $I \equiv \int_V \rho x^2 \mathrm{d}V$ is the momentum of inertia.

[4]Note that in this way we derive the so-called *scalar* version of the virial theorem, whereas had we multiplied by x_j we would have obtained the tensorial virial theorem, of which the scalar version is the trace.

The second inertial term on the l.h.s. of eq. 9.40 multiplied by x_i and integrated can be developed as follows:

$$\int_V x_i \frac{\partial}{\partial x_j}\left(\rho u_i u_j\right) \mathrm{d}V = \int_V \frac{\partial}{\partial x_j}\left(\rho x_i u_i u_j\right) \mathrm{d}V - \int_V \rho u_i u_j \delta_{ij} \mathrm{d}V = -2\mathsf{K}, \tag{9.43}$$

with $\mathsf{K} = \int_V \rho u^2 \mathrm{d}V/2$ total kinetic energy of the volume of fluid. Note that for the last equality we have applied the divergence theorem to the first term and assumed the absence of a net flux of fluid in or out our volume, thus the second term vanishes.

Analogously to the above, the pressure term (first on the r.h.s. of eq. 9.40) can be manipulated to get

$$\int_V x_i \frac{\partial P}{\partial x_i} \mathrm{d}V = \int_V \frac{\partial P x_i}{\partial x_i} \mathrm{d}V - \int_V P \frac{\partial x_i}{\partial x_i} \mathrm{d}V = \int_S P\boldsymbol{x} \cdot \mathrm{d}\boldsymbol{S} - 2\mathsf{U}, \tag{9.44}$$

where $\mathsf{U} = (3/2) N k_\mathrm{B} T$ is the total internal energy within the volume and we have used the divergence theorem for the first term of the r.h.s.[5]

The potential term leads to the potential energy

$$\int_V \rho x_i \frac{\partial \phi}{\partial x_i} = -\mathsf{W}, \tag{9.45}$$

with the total potential energy of the volume of fluid

$$\mathsf{W} = -\frac{1}{2} G \int_V \int_V \frac{\rho(\boldsymbol{x})\rho(\boldsymbol{x}')}{|\boldsymbol{x} - \boldsymbol{x}'|} \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{x}'. \tag{9.46}$$

We omit the derivation of this last equation as it is reported in other texts (e.g. [BT08]).

Finally, the magnetic term reads

$$\int_V x_i \frac{\partial}{\partial x_j} T_{ij} \mathrm{d}V = \int_V \frac{\partial}{\partial x_j}\left(x_i T_{ij}\right) \mathrm{d}V - \int_V T_{ij} \delta_{ij} \mathrm{d}V = \int_S \left(\boldsymbol{x} \cdot \overleftrightarrow{T}\right) \cdot \mathrm{d}\boldsymbol{S} + \mathsf{M} \tag{9.47}$$

with $\mathsf{M} = \int_V (B^2/8\pi) \mathrm{d}V$ total magnetic energy within the volume. Note that we have indicated the tensor T_{ij} with a double arrow and used the tensorial version of the divergence theorem for the last equality.

We are now in the position to write the full virial equation for a magnetised gas, which reads

$$\frac{1}{2}\frac{\mathrm{d}^2 I}{\mathrm{d}t^2} = 2\mathsf{K} + 2\mathsf{U} + \mathsf{W} + \mathsf{M} - \int_S P\boldsymbol{x} \cdot \mathrm{d}\boldsymbol{S} + \int_S \left(\boldsymbol{x} \cdot \overleftrightarrow{\boldsymbol{T}}\right) \cdot \mathrm{d}\boldsymbol{S}, \tag{9.48}$$

where we see that the *virial* term equals the sum of twice the kinetic energy and the internal energy plus the gravitational and the magnetic energy plus two surface integrals. These latter two are often neglected being usually of negligible importance compared to the other quantities.

[5]Note that this equation is strictly valid only for an atomic perfect gas in fact to obtain it one has to use the equality $P = (\gamma - 1)U$ (with U internal energy density) with $\gamma = 5/3$. For different values of γ one would have a slightly different coefficient in front of U.

9.5. MHD waves

In this Section, we consider how perturbations in a magnetised fluid produce the formation of waves and we investigate what type of waves we can expect. Let us consider a uniform, homogeneous and stationary medium hosting a uniform magnetic field. We write the fundamental conservation equations (mass, momentum and magnetic flux) for this medium as follows[6]:

$$\begin{cases} \dfrac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \mathbf{u}) = 0 \\ \rho \dfrac{\partial \boldsymbol{u}}{\partial t} + \rho(\boldsymbol{u} \cdot \nabla)\boldsymbol{u} = -\boldsymbol{\nabla} P + \dfrac{1}{4\pi}(\boldsymbol{\nabla} \times \boldsymbol{B}) \times \boldsymbol{B} \\ \dfrac{\partial \boldsymbol{B}}{\partial t} + \boldsymbol{\nabla} \times (\boldsymbol{B} \times \boldsymbol{u}) = 0. \end{cases} \tag{9.49}$$

Consider then infinitesimal perturbations of all the physical quantities involved here. For a generic fluid element we have

$$P = P_0 + \Delta P \tag{9.50}$$

$$\rho = \rho_0 + \Delta \rho \tag{9.51}$$

$$\mathbf{u} = \Delta \mathbf{u} \tag{9.52}$$

$$\boldsymbol{B} = \boldsymbol{B}_0 + \Delta \boldsymbol{B}, \tag{9.53}$$

where we have set our rest-frame such that the medium is initially stationary ($\mathbf{u}_0 = 0$).

The introduction of these perturbations in eqs. 9.49 and the truncation to the first order lead to

$$\begin{cases} \dfrac{\partial \Delta \rho}{\partial t} + \rho_0 \boldsymbol{\nabla} \cdot (\Delta \boldsymbol{u}) = 0 \\ \rho_0 \dfrac{\partial \Delta \boldsymbol{u}}{\partial t} = -c_{\mathrm{s}}^2 \boldsymbol{\nabla} \Delta \rho + \dfrac{1}{4\pi}(\boldsymbol{\nabla} \times \Delta \boldsymbol{B}) \times \boldsymbol{B}_0 \\ \dfrac{\partial \Delta \boldsymbol{B}}{\partial t} + \boldsymbol{\nabla} \times (\boldsymbol{B}_0 \times \Delta \boldsymbol{u}) = 0, \end{cases} \tag{9.54}$$

where we have also made the usual substitution $\Delta P = c_{\mathrm{s}}^2 \Delta \rho$. We now multiply the first equation of 9.54 by c_{s}^2 and take its gradient, then we derive the second equation in time and subtract the two. We obtain an equation containing the time derivative of $\Delta \boldsymbol{B}$ that can be substituted using the third of the eqs. 9.54. This procedure leads to the following wave equation:

$$\frac{\partial^2 \Delta \boldsymbol{u}}{\partial t^2} - c_{\mathrm{s}}^2 \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \Delta \boldsymbol{u}) = \boldsymbol{v}_{\mathrm{A}} \times \{\boldsymbol{\nabla} \times [\boldsymbol{\nabla} \times (\boldsymbol{v}_{\mathrm{A}} \times \Delta \boldsymbol{u})]\}, \tag{9.55}$$

where we have set $\boldsymbol{v}_{\mathrm{A}} \equiv \boldsymbol{B}_0 / \sqrt{4\pi\rho_0}$. This quantity has the units of a velocity, it is called **Alfvén velocity** and defined, in general, as

$$\boldsymbol{v}_{\mathrm{A}} \equiv \frac{\boldsymbol{B}}{\sqrt{4\pi\rho}}. \tag{9.56}$$

Equation 9.55 is the equation of the **magneto-sonic (MHD) waves**.

As usual, we can write the solution of eq. 9.55 as a sum of Fourier modes of the type $\Delta \boldsymbol{u} = \widetilde{\Delta \boldsymbol{u}} e^{i(\boldsymbol{k} \cdot \boldsymbol{x} - \omega t)}$. The second derivative in time of this Fourier mode

[6]Note that this treatment is very similar to that used to derive the equation of sound waves in §2.10.

returns $-\omega^2$, whereas a space derivative (∇) returns $i\boldsymbol{k}$. Thus, for instance, the second term on the l.h.s. of eq. 9.55 modifies as follows:

$$\boldsymbol{\nabla}(i\boldsymbol{k}\cdot\Delta\boldsymbol{u}) = \boldsymbol{\nabla}(i\boldsymbol{k}\cdot\widetilde{\Delta\boldsymbol{u}}e^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}) = i\boldsymbol{k}\cdot\widetilde{\Delta\boldsymbol{u}}(i\boldsymbol{k}e^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}) = -(\boldsymbol{k}\cdot\Delta\boldsymbol{u})\boldsymbol{k}.$$

Using similar techniques and employing some fundamental identities of vector calculus (in particular relations C.12, C.13 and C.14) we finally obtain the following expression

$$\begin{aligned}(9.57)\qquad &-\omega^2\Delta\boldsymbol{u} + \left(c_\mathrm{s}^2 + v_\mathrm{A}^2\right)(\boldsymbol{k}\cdot\Delta\boldsymbol{u})\,\boldsymbol{k} + \\ &+ (\boldsymbol{v}_\mathrm{A}\cdot\boldsymbol{k})\left[(\boldsymbol{v}_\mathrm{A}\cdot\boldsymbol{k})\Delta\boldsymbol{u} - (\boldsymbol{v}_\mathrm{A}\cdot\Delta\boldsymbol{u})\boldsymbol{k} - (\boldsymbol{k}\cdot\Delta\boldsymbol{u})\boldsymbol{v}_\mathrm{A}\right] = 0,\end{aligned}$$

which is the **dispersion relation for MHD waves**.

We now analyze eq. 9.57 in terms of *modes* of propagation of the waves. To do this, we make simple assumptions about the orientations of the three main vectors ($\Delta\boldsymbol{u}$, $\boldsymbol{v}_\mathrm{A}$ and $\boldsymbol{k}$). The vector $\Delta\boldsymbol{u}$ describes the velocity oscillation of the wave and sets the direction of this oscillation. The vector $\boldsymbol{v}_\mathrm{A}$ is the Alfvén velocity (eq. 9.56) and is aligned with the magnetic field $\boldsymbol{B}_0$. Finally, the wavenumber $\boldsymbol{k}$ represents the direction in which the wave is propagating. We can identify three main modes for these MHD waves, bearing in mind that a generic wave will be a combination of these three modes.

(1) Mode 1: $\boldsymbol{k} \parallel \boldsymbol{B} \parallel \Delta\boldsymbol{u}$. If the three above vectors are parallel, all the scalar products in eq. 9.57 become standard products and this leads to a dramatic simplification of the dispersion equation that becomes

$$(9.58)\qquad \frac{\omega}{k} = c_\mathrm{s},$$

which is simply the dispersion relation of sound waves. These are normal longitudinal ($\boldsymbol{k} \parallel \Delta\boldsymbol{u}$) compression waves that propagates as if the magnetic field were not present.

(2) Mode 2: $\boldsymbol{k} \parallel \boldsymbol{B} \perp \Delta\boldsymbol{u}$. This is the definition of transverse waves, in which the propagation $\boldsymbol{k}$ is orthogonal to the oscillation ($\Delta\boldsymbol{u}$). This is a type of waves that cannot exist in the absence of a magnetic field. If we impose this condition in eq. 9.57 we find a physically meaningful solution with dispersion relation

$$(9.59)\qquad \frac{\omega}{k} = v_\mathrm{A},$$

which represents the so-called **Alfvén waves**. The important result here is that, in the presence of a magnetic field, there exists a completely new type of waves. These waves propagate *orthogonally* to the oscillation of the fluid elements and they do so at the Alfvén speed, which can be very different from the sound speed. This new type of waves are akin to the vibrations of guitar strings where the magnetic field lines behave like such strings.

(3) Mode 3: $\boldsymbol{k} \perp \boldsymbol{B}$, $\Delta\boldsymbol{u}$ unspecified. This condition leads to a slightly more complex dispersion relation that reads

$$(9.60)\qquad -\omega^2\Delta\boldsymbol{u} + \left(c_\mathrm{s}^2 + v_\mathrm{A}^2\right)(\boldsymbol{k}\cdot\Delta\boldsymbol{u})\,\boldsymbol{k} = 0$$

The phase speed of this mode is maximised for $\boldsymbol{k} \parallel \Delta\boldsymbol{u}$ (longitudinal waves), in which case we obtain

$$(9.61)\qquad \frac{\omega}{k} = \sqrt{c_\mathrm{s}^2 + v_\mathrm{A}^2},$$

which is the dispersion relation for the magneto-sonic waves. Clearly, these waves reach the highest velocities among all possible solutions for the other modes.

In the general case, we should expect to have a combination of the above modes. The general dispersion relation for $\Delta\boldsymbol{u}$ lying on the $\boldsymbol{B}$-$\boldsymbol{k}$ plane reads

$$\frac{\omega^2}{k^2} = \frac{1}{2}\left\{\left(v_{\rm A}^2 + c_{\rm s}^2\right) \pm \left[\left(v_{\rm A}^2 + c_{\rm s}^2\right)^2 - 4v_{\rm A}^2 c_{\rm s}^2 \cos\psi^2\right]^{1/2}\right\}, \tag{9.62}$$

where ψ is the angle between $\boldsymbol{B}$ and $\boldsymbol{k}$. These waves can be both longitudinal and transverse, the propagation speed will be a combination of $c_{\rm s}$ and $v_{\rm A}$ and it will depend on the direction of the magnetic field and on whether this latter is ordered or not. For generic orientation of $\boldsymbol{B}$, we expect the medium to show a spectrum of typical velocities ranging from *slow waves* (smaller speed between $c_{\rm s}$ and $v_{\rm A}$) and *fast waves* (propagation at $\sqrt{c_{\rm s}^2 + v_{\rm A}^2}$). Note that all the above speeds are both phase and group speeds.

9.6. Magnetic shocks

We end this chapter with a brief account of the consequence that we can expect having a magnetised medium when dealing with shock waves. Let us consider the simplified case of an ordered magnetic field aligned parallel to the shock itself (say along the z-axis if the shock propagates along x, see §4.2). This is an idealised configuration but it allows us to estimate the main differences that occur with respect to the non-magnetised case (Chap. 4). To proceed, we first write the jump conditions (§4.2.1) of this shock. The first equation (mass conservation through the shock) is obviously unchanged. The second condition comes, instead, from the Euler equation and it has to be modified. We start from the Euler equation including the magnetic force written in stationary conditions:

$$\rho(\boldsymbol{u}\cdot\nabla)\boldsymbol{u} = -\boldsymbol{\nabla}P - \boldsymbol{\nabla}\frac{B^2}{8\pi} + \frac{1}{4\pi}\boldsymbol{B}\cdot\boldsymbol{\nabla}\boldsymbol{B}, \tag{9.63}$$

where gravity is neglected for the same reasons as in §4.2.1. Given our assumption for the orientation of the magnetic field we do not expect any deformation of the field lines and thus any role of the magnetic tension. This can be explicitly shown by the fact that the unperturbed magnetic field is $\boldsymbol{B} = (0, 0, B_0)$ and thus the only derivative that can count in the dot product in the last term of the r.h.s. of eq. 9.63 is $\partial/\partial z$, which is null as in our system we only have variations along the x-axis.

Following the same procedure as in §4.2 we have that

$$\frac{\mathrm{d}}{\mathrm{d}x}(\rho u^2) - u\frac{\mathrm{d}\rho u}{\mathrm{d}x} = -\frac{\mathrm{d}}{\mathrm{d}x}\left(P + \frac{B^2}{8\pi}\right), \tag{9.64}$$

where the second term of the l.h.s. is zero from the first jump condition (§4.2). We are therefore in the position to write the jump conditions for this magnetised shock. We consider the case of a radiative shock (§4.3). These types of shocks are, indeed, the most interesting in the present context because they produce dissipation of kinetic energy. Our goal is to establish whether the introduction of the magnetic field has an impact on the amount of energy dissipation that we can expect. As in §4.3, we refer to the perturbed medium using the suffix $_2$ and the new jump

conditions read

$$\begin{cases} \rho_0 u_0 = \rho_2 u_2 \\ \rho_0 u_0^2 + P_0 + \dfrac{B_0^2}{8\pi} = \rho_2 u_2^2 + P_2 + \dfrac{B_2^2}{8\pi} \\ T_0 = T_2. \end{cases} \tag{9.65}$$

Equations 9.65 are analogous to the Hugoniot-Rankine conditions in eq. 4.40 with the added term of the magnetic pressure.

The presence of the magnetic pressure requires a forth equation. To derive this, we use the conservation of the magnetic flux (§9.2), which can be written as

$$B_0 \mathrm{d}S_0 = B_2 \mathrm{d}S_2, \tag{9.66}$$

where $\mathrm{d}S_0$ and $\mathrm{d}S_2$ are infinitesimal surfaces in the unperturbed and perturbed medium respectively. As a consequence of the passage through the shock, the fluid receives a compression only in the x-direction, thus both surfaces and volumes are effectively changing only if $\mathrm{d}x$ changes, for instance for the unperturbed medium $\mathrm{d}V_0 \propto \mathrm{d}S_0 \propto \mathrm{d}x_0$, and the same is valid for shocked medium, thus eq. 9.67 can be rewritten in this way:

$$B_0 \mathrm{d}x_0 = B_2 \mathrm{d}x_2 \tag{9.67}$$

We now employ the conservation of mass at the passage through the shock that we can write as

$$\rho_0 \mathrm{d}V_0 = \rho_2 \mathrm{d}V_2 \implies \rho_0 \mathrm{d}x_0 = \rho_2 \mathrm{d}x_2, \tag{9.68}$$

where the implication is due to the above proportionality: $\mathrm{d}V \propto \mathrm{d}x$. Combining eqs. 9.67 and 9.68 we obtain the last jump condition valid for a shock with a magnetic field aligned with the shock front and frozen in the fluid:

$$\frac{\rho_2}{\rho_0} = \frac{B_2}{B_0}. \tag{9.69}$$

Let us then inspect carefully the second of eqs. 9.65. On the l.h.s., for strong shocks, the dominant term must be the ram-pressure term (we remind that u_0 here is equal to the shock speed seen from a reference frame where the shock is at rest). The thermal pressure and, realistically, the magnetic pressure can be neglected in the unperturbed medium. On the r.h.s., the situation is instead very different, the first term is negligible because the shocked medium follows the shock at the same speed and $u_2 \approx 0$. Of the other two terms, the thermal pressure ($P_2 = c_{\mathrm{s}2}^2 \rho_2$) increases linearly with the density of the shocked medium (as the sound speeds are the same) whereas the magnetic pressure increases quadratically (from eq. 9.69). A fundamental property of a radiative shock (without $\boldsymbol{B}$) is that the density of the shocked medium can increase by a very large factor (§4.3.3). The dependencies found above then tell us that, as the density increases, we will likely reach a point at which the magnetic pressure dominates over the thermal pressure in the shocked medium, i.e. we will have a jump condition that reads

$$\rho_0 u_0^2 \approx \frac{B_2^2}{8\pi} = \left(\frac{\rho_2}{\rho_0}\right)^2 \frac{B_0^2}{8\pi}, \tag{9.70}$$

an equation from which we can derive a new ratio of densities using the definition of Alfvén velocity (eq. 9.56).

In the end, we conclude that, in the presence of a non-negligible magnetic field, the compression of a the gas produced by a radiative (isothermal) shock will be

$$\frac{\rho_2}{\rho_0} \approx \sqrt{\frac{2u_0^2}{v_{\rm A}^2}} = \sqrt{2}\mathcal{M}_{\rm A,0}, \tag{9.71}$$

where we have indicated with $\mathcal{M}_{\rm A,0} \equiv u_0/v_{\rm A}$ the **magnetic Mach number** (also called **Alfvén number**) in the unperturbed medium. The difference between this conclusion and that found for non-magnetised medium (§4.3.3) is apparent. In a non-magnetised medium, the density in the shocked gas can increase as the square of the Mach number (eq. 4.45). Instead, in the presence of a magnetic field this increase is limited to a linear proportionality with the magnetic Mach number. Moreover, in practical situations, $M_{\rm A,\,0}$ can often be much smaller than the Mach number (because $v_{\rm A} > c_{\rm s}$) and, as a consequence, the magnetic field can strongly limit the shock compression and the dissipation of kinetic energy.

APPENDIX A

Notation

Basic quantities

n	:	number density	particles cm^{-3}
v_{th}	:	thermal speed	$\mathrm{km\,s}^{-1}$
c_{s}	:	sound speed	$\mathrm{km\,s}^{-1}$

Thermodynamic variables

ρ	:	volume density	$\mathrm{g\,cm}^{-3}$
T	:	Temperature	K
P	:	Pressure	$\mathrm{dyne\,cm}^{-2}/\mathrm{erg\,cm}^{-3}$
$\mathcal{V}$	:	specific volume	$\mathrm{cm}^3\ \mathrm{g}^{-1}$
$\mathcal{S}$	:	specific entropy	$\mathrm{erg\,K}^{-1}\,\mathrm{g}^{-1}$

Energies

E	:	Total energy	erg
U	:	Internal energy	erg
K	:	Kinetic energy	erg
H	:	Enthalpy	erg
ϵ	:	Total energy density	$\mathrm{erg\,cm}^{-3}$
ϵ_{K}	:	Kinetic energy density	$\mathrm{erg\,cm}^{-3}$
U	:	Internal energy density	$\mathrm{erg\,cm}^{-3}$
$\mathcal{E}$	:	specific energy	$\mathrm{erg\,g}^{-1}$
$\mathcal{U}$	:	specific internal energy	$\mathrm{erg\,g}^{-1}$
$\mathcal{H}$	:	specific enthalpy	$\mathrm{erg\,g}^{-1}$
$\mathcal{G}$	:	specific Gibbs free energy	$\mathrm{erg\,g}^{-1}$

Energy rates

$\mathcal{L}$	:	Radiative energy loss per unit time and mass	$\mathrm{erg\,s}^{-1}\,\mathrm{g}^{-1}$
Λ	:	Radiative cooling	$\mathrm{erg\,s}^{-1}\,\mathrm{g}^{-1}$
Λ_{CIE}	:	Cooling function for collisional ionization equilibrium $\Lambda_{\mathrm{CIE}} = \Lambda_{\mathrm{CIE}}(T; Z)$	$\mathrm{erg\,s}^{-1}\,\mathrm{cm}^3$

Non-dimensional Numbers

$\mathcal{M}$	:	Mach number
Re	:	Reynolds number

Cartesian coordinates	x, y, z
Cylindrical coordinates	R, φ, z
Spherical coordinates	r, φ, θ

APPENDIX B

Constants and conversions

B.1. Physics constants

Constants	Symbol	Value	Units (cgs)
Speed of light	c	$2.99792458 \times 10^{10}$	$\mathrm{cm\,s^{-1}}$
Gravitational constant	G	6.6741×10^{-8}	$\mathrm{cm^{-3}\,g^{-1}\,s^{-2}}$
Boltzmann constant	k_B	$1.3806485 \times 10^{-16}$	$\mathrm{erg\,K^{-1}}$
Planck constant	h	$6.62607004 \times 10^{-27}$	$\mathrm{erg\,s}$
Mass of electron	m_e	$9.1093836 \times 10^{-28}$	g
Mass of proton	m_p	$1.67262190 \times 10^{-24}$	g
Bohr radius	a_0	5.3×10^{-9}	cm
Avogadro constant	N_A	$6.02214179 \times 10^{23}$	$\mathrm{mol^{-1}}$
Elementary charge	e	$4.80320425 \times 10^{-10}$	e.s.u. $\left(\mathrm{erg^{1/2}\,cm^{1/2}}\right)$

B.2. Astrophysical standard values

Constants	Symbol	Value	Units
Parsec	pc	3.086×10^{18}	cm
Astronomical unit	AU	1.496×10^{13}	cm
Solar mass	$\mathrm{M}_\odot$	1.989×10^{33}	g
Solar radius	$R_\odot$	6.955×10^{10}	cm
Sideral year	yr	3.1558×10^{7}	s

B.3. Conversions

	cgs	SI
1 eV	$1.60217657 \times 10^{-12}$ erg	$1.60217657 \times 10^{-19}$ J

APPENDIX C

Mathematical compendium

C.1. Coordinates

In cylindrical coordinates a positional vector is written as

$$\boldsymbol{x} = R\hat{e}_{\mathrm{R}} + z\hat{e}_{\mathrm{z}} \tag{C.1}$$

thus a velocity vector will read

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \dot{R}\hat{e}_{\mathrm{R}} + R\dot{\hat{e}}_{\mathrm{R}} + \dot{z}\hat{e}_{\mathrm{z}}. \tag{C.2}$$

The derivative of a the unity vector along R is

$$\dot{\hat{e}}_{\mathrm{R}} = \frac{\mathrm{d}\hat{e}_{\mathrm{R}}}{\mathrm{d}\varphi}\frac{\mathrm{d}\varphi}{\mathrm{d}t} = \dot{\varphi}\hat{e}_{\varphi} \tag{C.3}$$

and so

$$\dot{\boldsymbol{x}} = \dot{R}\hat{e}_{\mathrm{R}} + \dot{R}\varphi\hat{e}_{\varphi} + \dot{z}\hat{e}_{\mathrm{z}}. \tag{C.4}$$

If we now derive this expression in time and follow the same procedure as above we find

$$\ddot{\boldsymbol{x}} = (\ddot{R} - R\dot{\varphi}^2)\hat{e}_{\mathrm{R}} + (2\dot{R}\dot{\varphi} + R\ddot{\varphi})\hat{e}_{\varphi} + \ddot{z}\hat{e}_{\mathrm{z}}, \tag{C.5}$$

which is the acceleration in cylindrical coordinates.

C.2. Frequently used functions

The normalized Gaussian function is

$$G(x) = \frac{1}{\sqrt{2\pi}\,\sigma}\exp\left(-\frac{(x-x_0)^2}{2\sigma^2}\right), \tag{C.6}$$

where x_0 the central value and σ is the dispersion.

C.3. Useful relations

C.3.1. Cyclic relation. This property can be derived as follows. Take three continuous functions $x = x(y,z)$, $y = y(z,x)$ and $z = z(x,y)$ and write the differential of two of them:

$$dx = \left(\frac{\partial x}{\partial y}\right)_z \mathrm{d}y + \left(\frac{\partial x}{\partial z}\right)_y \mathrm{d}z \tag{C.7}$$

and

$$dy = \left(\frac{\partial y}{\partial z}\right)_x \mathrm{d}z + \left(\frac{\partial y}{\partial x}\right)_z \mathrm{d}x. \tag{C.8}$$

Now substitute the second in the first to obtain

$$dx = \left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial x}\right)_z \mathrm{d}x + \left[\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x + \left(\frac{\partial x}{\partial z}\right)_y\right] \mathrm{d}z. \tag{C.9}$$

The cancellation of the l.h.s. with the first term of the r.h.s. then returns the cyclic relation identity:

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1. \tag{C.10}$$

C.4. Vector calculus

C.4.1. Basic concepts. Laplacian.

$$\nabla^2 q \equiv \boldsymbol{\nabla} \cdot \boldsymbol{\nabla} q \tag{C.11}$$

C.4.2. Identities. Three useful properties of the vector calculus are

$$\boldsymbol{\nabla} \times (A\boldsymbol{a}) = \boldsymbol{\nabla} A \times \boldsymbol{a} + A\boldsymbol{\nabla} \times \boldsymbol{a} \tag{C.12}$$

$$\boldsymbol{\nabla} \times (\boldsymbol{a} \times \boldsymbol{b}) = \boldsymbol{a}(\boldsymbol{\nabla} \cdot \boldsymbol{b}) - \boldsymbol{b}(\boldsymbol{\nabla} \cdot \boldsymbol{a}) + \boldsymbol{b}\boldsymbol{\nabla}\boldsymbol{a} - \boldsymbol{a}\boldsymbol{\nabla}\boldsymbol{b} \tag{C.13}$$

$$\boldsymbol{\nabla}(\boldsymbol{a} \cdot \boldsymbol{b}) = \boldsymbol{a} \times (\boldsymbol{\nabla} \times \boldsymbol{b}) + \boldsymbol{b} \times (\boldsymbol{\nabla} \times \boldsymbol{a}) + \boldsymbol{a} \cdot \boldsymbol{\nabla}\boldsymbol{b} + \boldsymbol{b} \cdot \boldsymbol{\nabla}\boldsymbol{a} \tag{C.14}$$

The two following combinations of nabla operations return zero:

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} A) = 0 \qquad \forall A, \tag{C.15}$$

and

$$\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \boldsymbol{a}) = 0 \qquad \forall \boldsymbol{a}. \tag{C.16}$$

The first of these two indentities is fundamental in field theory because it allows us to define a *potential* that generates the field whenever the field is irrotational ($\boldsymbol{\nabla} \times \boldsymbol{a} = 0$). In that case one can impose that $\boldsymbol{a} = \pm\boldsymbol{\nabla} A$ and work with the scalar A instead of the vector $\boldsymbol{a}$. This largely simplifies the problem in most cases.

Other two very useful identities are

$$\boldsymbol{\nabla}\frac{1}{2}a^2 = \boldsymbol{a} \cdot \boldsymbol{\nabla}\boldsymbol{a} + \boldsymbol{a} \times (\boldsymbol{\nabla} \times \boldsymbol{a}), \tag{C.17}$$

and

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{a}) = \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{a}) - \nabla^2 \boldsymbol{a}. \tag{C.18}$$

C.4.3. Theorems. The divergence theorem states that, given a certain volume V contained within a surface S the following equivalence is fulfilled

$$\int_V (\boldsymbol{\nabla} \cdot \boldsymbol{a}) \mathrm{d}V = \int_S \boldsymbol{a} \cdot \mathrm{d}\boldsymbol{S} \tag{C.19}$$

for any vector quantity $\boldsymbol{a}$. As a convention, the surface element vector $\mathrm{d}\boldsymbol{S}$ is normal to the surface S and points outwards. The divergence theorem has the following corollary:

$$\int_V \boldsymbol{\nabla}(A) \mathrm{d}V = \int_S A \mathrm{d}\boldsymbol{S} \tag{C.20}$$

valid for every scalar A. To prove it just write the divergence theorem for the vector $A\boldsymbol{c}$ with $\boldsymbol{c}$ a constant vector.

The Stokes theorem states that, given a surface S bound by a closed curve C the following equivalance is fulfilled

$$\int_S (\boldsymbol{\nabla} \times \boldsymbol{a}) \cdot \mathrm{d}\boldsymbol{S} = \oint_C \boldsymbol{a} \cdot \mathbf{d}\boldsymbol{l} \tag{C.21}$$

for any vector field $\boldsymbol{a}$. Note that the surface S can be any open surface whose boundary is C.

C.4.4. Spatial derivatives in non-Cartesian coordinates. Laplacian in cylindrical coordinates:

$$\nabla^2 = \frac{1}{R}\frac{\partial}{\partial R}\left(R\frac{\partial}{\partial R}\right) + \frac{1}{R^2}\frac{\partial^2}{\partial \phi^2} + \frac{\partial^2}{\partial z^2} \tag{C.22}$$

Laplacian in spherical coordinates

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial}{\partial \theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial \phi^2} \tag{C.23}$$

C.4.5. Special tensors. The Kronecker symbol δ_{ij} identifies a second-order tensor defined as

$$\delta_{ij} \equiv \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases} \tag{C.24}$$

The Levi-Civita symbol ϵ_{ijk} identifies a third-order tensor with values 1, -1 and 0 depending on the following rules

$$\epsilon_{ijk} = \begin{cases} 1 & \text{if } ijk = 123,\, 231,\, 312 \\ -1 & \text{if } ijk = 132,\, 321,\, 213 \\ 0 & \text{if } ijk = 111,\, 222,\, 333. \end{cases} \tag{C.25}$$

The Levi-Civita symbol is useful to perform cross products by components. A cross product between two vectors can be written

$$\boldsymbol{v} = \boldsymbol{a} \times \boldsymbol{b} \implies v_i = \epsilon_{ijk} a_j b_k, \tag{C.26}$$

returning the $i-$component of the resulting vector.

APPENDIX D

Physics compendium

D.1. Thermodynamics

D.1.1. Equation of state of the ideal gases. There are several forms of this equation that are used interchangebly for different applications. The formula most used in these notes is

$$\frac{P}{\rho} = \frac{k_{\rm B}T}{\mu m_{\rm p}}, \tag{D.1}$$

where $k_{\rm B}$ is the Boltzman constant, $m_{\rm p}$ is the proton mass and μ is atomic/molecular *weight*, i.e. the avarage mass of a particle in the fluid in units of a proton mass. The use of density instead of volume is much handier in astrophysical contexts.

Other versions of the ideal gas equation often employ the volume and also the gas constant R in an equation of the form

$$PV = n_{\rm mol}RT, \tag{D.2}$$

where $n_{\rm mol}$ is the number of **moles** in the fluid, i.e. the number of chunks of fluid containing a number of particles equal to the Avogadro number $N_{\rm A}$. Equation D.2 can be also written using the Boltzman constant as

$$PV = Nk_{\rm B}T \tag{D.3}$$

where N is now the total number of particles. Comparing these two last equations we recover the identity

$$k_{\rm B} \equiv \frac{R}{N_{\rm A}}. \tag{D.4}$$

Equation D.2 can also be rewritten with specific quantities. Given that $V = M/\rho$ and $n_{\rm mol} = M/m_{\rm mol}$ with M total mass and $m_{\rm mol} = N_{\rm A}\mu m_{\rm p}$ mass of a mole, we obtain

$$\frac{P}{\rho} = \mathcal{R}T, \tag{D.5}$$

where $\mathcal{R}$ is the *specific* gas constant, $\mathcal{R} = R/m_{\rm mol}$.

D.1.2. Internal energy and pressure. We prove here the relation between internal energy and pressure. Take the first law of thermodynamics for adiabatic transformations ($\delta Q = 0$) written with specific variables

$$\mathrm{d}\mathcal{U} = -P\mathrm{d}\mathcal{V} = \frac{P}{\rho^2}\mathrm{d}\rho. \tag{D.6}$$

The adiabatic equation of state $P = K\rho^\gamma$ leads to

$$\mathrm{d}\rho = \frac{1}{\gamma K^{1/\gamma}}P^{1/\gamma}\mathrm{d}P. \tag{D.7}$$

Substituting this in the first law and integrating one finds

$$\mathcal{U} = \frac{1}{(\gamma - 1)} \frac{P}{\rho}. \tag{D.8}$$

D.1.3. Entropy. The equation of entropy as a function of other termodynamic variables can be derived by integrating the first law of thermodynamics written as

$$T \mathrm{d}\mathcal{S} = \mathrm{d}\mathcal{U} + P \mathrm{d}\mathcal{V} = \mathrm{d}\mathcal{U} - \frac{P}{\rho^2} \mathrm{d}\rho. \tag{D.9}$$

We use eq. D.5 to obtain

$$\mathrm{d}\mathcal{S} = \frac{\mathrm{d}\mathcal{U}}{\mathrm{d}T} \frac{\mathrm{d}T}{T} - \mathcal{R} \frac{\mathrm{d}\rho}{\rho} \tag{D.10}$$

and integrating[1] we have

$$\mathcal{S} = C_\mathrm{V} \ln T - \mathcal{R} \ln \rho + c = C_\mathrm{V} \ln \frac{P}{\rho} - \mathcal{R} \ln \rho + c' = C_\mathrm{V} \ln P - C_\mathrm{p} \ln \rho + c', \tag{D.11}$$

remembering that $C_\mathrm{p} = C_\mathrm{V} + \mathcal{R}$. At this point we multiply and divide the r.h.s. by C_V. Recalling that the adiabatic index can be writte as $\gamma = C_\mathrm{p}/C_\mathrm{V}$ one finally obtains the entropy equation

$$\mathcal{S} = C_\mathrm{V} \ \ln(P \rho^{-\gamma}) + \text{constant}. \tag{D.12}$$

D.1.4. Specific heats. The specific heat is the variation of the heat exchange per unit mass as a consequence of small temperature changes, in formulae:

$$c = \frac{\delta Q}{\mathrm{d}T} \tag{D.13}$$

and, being the heat not an exact differential, this equation cannot be integrated. However, one can use the first principle of thermodynamics to obtain

$$C_\mathrm{V} \equiv \left(\frac{\partial \mathcal{U}}{\partial T} \right)_V = \frac{\mathrm{d}\mathcal{U}}{\mathrm{d}T} \tag{D.14}$$

full differential because for ideal gases $\mathcal{U} = \mathcal{U}(T)$. And in the case of constant pressure

$$C_\mathrm{p} \equiv \left(\frac{\partial \mathcal{U}}{\partial T} + P \frac{\partial \mathcal{V}}{\partial T} \right)_P = C_\mathrm{V} + \mathcal{R}. \tag{D.15}$$

For reversible transformations when $\delta Q/T = \mathrm{d}\mathcal{S}$ the specific heats can also be defined as a function of entropy as

$$C_\mathrm{V} = T \left(\frac{\partial \mathcal{S}}{\partial T} \right)_\mathrm{V} \tag{D.16}$$

and

$$C_\mathrm{p} = T \left(\frac{\partial \mathcal{S}}{\partial T} \right)_\mathrm{P} \tag{D.17}$$

D.1.5. Enthalpy. The specific enthalpy reads

$$\mathcal{H} = \mathcal{U} + \frac{P}{\rho}. \tag{D.18}$$

[1]Note that we are simplifying the notation incorporating several terms in the constants, e.g. $c = -C_\mathrm{V} \ln T_0 + \mathcal{R} \ln \rho_0 + \mathcal{S}_\prime$.

D.1.6. Gibbs free energy. Gibbs free energy is defined as

$$\mathcal{G} \equiv \mathcal{H} - T\mathcal{S}. \tag{D.19}$$

The equivalent of the first law of thermodynamics with Gibbs free energy reads

$$\mathrm{d}\mathcal{G} = -\mathcal{S}\mathrm{d}T + \mathcal{V}\mathrm{d}P. \tag{D.20}$$

D.1.7. Maxwell's relations. These are relations that relate partial derivative of thermodynamical quantities. One of them is

$$\left(\frac{\partial \mathcal{S}}{\partial P}\right)_{\mathrm{T}} = -\left(\frac{\partial \mathcal{V}}{\partial T}\right)_{\mathrm{P}} \tag{D.21}$$

derived as follows. Take the Gibbs free energy as a function of temperature and pressure $\mathcal{G} = \mathcal{G}(T, P)$ and write the differential of G to obtain

$$\mathrm{d}\mathcal{G} = \left(\frac{\partial \mathcal{G}}{\partial T}\right)_{P} \mathrm{d}T + \left(\frac{\partial \mathcal{G}}{\partial P}\right)_{T} \mathrm{d}P. \tag{D.22}$$

We can recognise the two terms of the r.h.s. as the first law of thermodynamics written with the Gibbs free energy (eq. D.20). If we derive now the first term by T and the second term by P this leads to an identity

$$\left[\left(\frac{\partial \mathcal{G}}{\partial T}\right)_{\mathrm{P}}\right]_{\mathrm{T}} = \left[\left(\frac{\partial \mathcal{G}}{\partial P}\right)_{\mathrm{T}}\right]_{\mathrm{P}}. \tag{D.23}$$

Bibliography

[ALL01] J. F. Alves, C. J. Lada, and E. A. Lada, *Internal structure of a cold dark molecular cloud inferred from the extinction of background starlight*, Nature **409** (2001), 159–161.

[BH91] S. A. Balbus and J. F. Hawley, *A powerful local shear instability in weakly magnetized disks. I - Linear analysis. II - Nonlinear evolution*, ApJ **376** (1991), 214–233.

[BT08] J. Binney and S. Tremaine, *Galactic Dynamics: Second Edition*, Princeton University Press, 2008.

[CC14] C. Clarke and B. Carswell, *Principles of Astrophysical Fluid Dynamics*, January 2014.

[CFN19] A. Cimatti, F. Fraternali, and C. Nipoti, *Introduction to Galaxy Formation and Evolution: From Primordial Gas to Present-Day Galaxies*, 2019.

[Cha60] S. Chandrasekhar, *The Stability of Non-Dissipative Couette Flow in Hydromagnetics*, Proceedings of the National Academy of Science **46** (1960), 253–257.

[Che74] R. A. Chevalier, *The Evolution of Supernova Remnants. Spherically Symmetric Models*, ApJ **188** (1974), 501–516.

[CM77] L. L. Cowie and C. F. McKee, *The evaporation of spherical clouds in a hot gas. I - Classical and saturated mass loss rates*, ApJ **211** (1977), 135–146.

[Dra03] B. T. Draine, *Interstellar Dust Grains*, ARA&A **41** (2003), 241–289.

[Dra11] Bruce T. Draine, *Physics of the Interstellar and Intergalactic Medium*, 2011.

[DW97] J. E. Dyson and D. A. Williams, *The physics of the interstellar medium*, 1997.

[ES04] B. G. Elmegreen and J. Scalo, *Interstellar Turbulence I: Observations and Processes*, ARA&A **42** (2004), 211–273.

[FCG+17] G. J. Ferland, M. Chatzikos, F. Guzmán, M. L. Lykins, P. A. M. van Hoof, R. J. R. Williams, N. P. Abel, N. R. Badnell, F. P. Keenan, R. L. Porter, and P. C. Stancil, *The 2017 Release Cloudy*, Revista Mexicana de Astronomía y Astrofísica **53** (2017), 385–438.

[Fie65] G. B. Field, *Thermal Instability.*, ApJ **142** (1965), 531.

[HT03] C. Heiles and T. H. Troland, *The Millennium Arecibo 21 Centimeter Absorption-Line Survey. II. Properties of the Warm and Cold Neutral Media*, ApJ **586** (2003), 1067–1093.

[Jea02] J. H. Jeans, *The Stability of a Spherical Nebula*, Philosophical Transactions of the Royal Society of London Series A **199** (1902), 1–53.

[Lar81] R. B. Larson, *Turbulence and star formation in molecular clouds*, MNRAS **194** (1981), 809–826.

[MO77] C. F. McKee and J. P. Ostriker, *A theory of the interstellar medium - Three components regulated by supernova explosions in an inhomogeneous substrate*, ApJ **218** (1977), 148–169.

[Mye78] P. C. Myers, *A compilation of interstellar gas properties*, ApJ **225** (1978), 380–389.

[NFW96] Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White, *The Structure of Cold Dark Matter Halos*, ApJ **462** (1996), 563.

[NM01] R. Narayan and M. V. Medvedev, *Thermal Conduction in Clusters of Galaxies*, ApJL **562** (2001), L129–L132.

[Oll95] R. P. Olling, *On the Usage of Flaring Gas Layers to Determine the Shape of Dark Matter Halos*, AJ **110** (1995), 591.

[Par53] E. N. Parker, *Instability of Thermal Fields.*, ApJ **117** (1953), 431.

[Pat84] A. R. Paterson, *A First Course in Fluid Dynamics*, January 1984.

[SD93] R. S. Sutherland and M. A. Dopita, *Cooling functions for low-density astrophysical plasmas*, ApJS **88** (1993), 253–327.

[Sed59] L. I. Sedov, *Similarity and Dimensional Methods in Mechanics*, 1959.

[Shu92] F. H. Shu, *The physics of astrophysics. Volume II: Gas dynamics.*, 1992.

[SP05] S. W. Stahler and F. Palla, *The Formation of Stars*, 2005.

[Spi78] L. Spitzer, *Physical processes in the interstellar medium*, 1978.

[SWS+03] K. R. Sembach, B. P. Wakker, B. D. Savage, P. Richter, M. Meade, J. M. Shull, E. B. Jenkins, G. Sonneborn, and H. W. Moos, *Highly Ionized High-Velocity Gas in the Vicinity of the Galaxy*, ApJS **146** (2003), 165–208.

[Tay50] Geoffrey Taylor, *The Formation of a Blast Wave by a Very Intense Explosion. I. Theoretical Discussion*, Proceedings of the Royal Society of London Series A **201** (1950), no. 1065, 159–174.

[Tie05] A. G. G. M. Tielens, *The Physics and Chemistry of the Interstellar Medium*, August 2005.

[WHM+95] M. G. Wolfire, D. Hollenbach, C. F. McKee, A. G. G. M. Tielens, and E. L. O. Bakes, *The Neutral Atomic Phases of the Interstellar Medium*, ApJ **443** (1995), 152.