13. Temperature balance and gravitational collapse

The gas between stars, commonly referred to as the interstellar medium, and the gas between galaxies, called the intergalactic medium, is subject to a number of heating processes, which must be balanced by cooling, otherwise the temperature of the medium would perpetually increase. Some heating processes operate continuously, like heating by absorption of radiation. If the absorption is done by free particles, like free electrons, all the energy of the absorbed photon is available as heat. If the absorption is done by exciting or ionizing an atom, much of the photon energy is stored as potential energy in the atom. In case of ionization only the kinetic energy of the free electron can be thermalized, so the photon energy minus the binding energy of the atomic electron is available as heat. Gas can also be collisionally heated by energetic particles called cosmic rays.

Other heating processes like shock heating operate intermittendly. Let us estimate the heating contributed by supernova remnant shocks. Measurements indicate that typically every $t_{\rm SN} =$ 50 years a supernova explodes in the Galaxy and the typical explosion energy is $E = 10^{51}$ ergs. The Galaxy can for our purposes be approximated as a disk with radius 10.000 pc (parallax second=parsec, 1 pc = 3 Light-years) and thickness 500 pc, so its volume is $V_{\rm gal} \simeq 1.5 \cdot 10^{11}$ pc^3. From the jump conditions at strong ($M \to \infty$) hydrodynamical shocks we derived the post-shock temperature as a function of the shock velocity as

$$kT_2 = \frac{2\,(\gamma-1)}{(1+\gamma)^2}\, m\, V_s^2 \qquad \Rightarrow \quad T_2 \simeq (3\cdot 10^5\ \mathrm{K}) \left(\frac{V_s}{100\ \mathrm{km/s}}\right)^2 \tag{13.1}$$

From the Taylor-Sedov blast-wave solution we know that the shock velocity and radius scale as

$$V_s(t) = \frac{dr_s}{dt} = \frac{2}{5}\, x_0 \left(\frac{E}{\rho_u\, t^3}\right)^{1/5} \qquad r_s(t) = x_0 \left(\frac{E\, t^2}{\rho_u}\right)^{1/5} \tag{13.2}$$

where x_0 marks the shock location in the self-similar coordinates. Let us for simplicity just use $x_0 = 1$. Let us also take the average value for the gas density in the interstellar medium.

$$\text{On average :} \qquad n \simeq 1\ \mathrm{atom/cm^3} \qquad \rho \simeq 2\cdot 10^{-24}\ \mathrm{g/cm^3} \tag{13.3}$$

We can now ask ourselves: how often is the interstellar medium heated to at least a million degrees temperature? According to Eq.13.1 this would require the region be overrun by a shock with at least the critical velocity $V_c \simeq 180$ km/s. The blastwave velocity is larger than this critical value as long as the blastwave radius is smaller than the critical value

$$r_c \simeq \left(\frac{2^2\, E}{\rho_u\, 5^2\, V_c^2}\right)^{1/3} \simeq 6\cdot 10^{19}\ \mathrm{cm} \simeq 20\ \mathrm{pc} \tag{13.4}$$

for our parameters. The volume occupied by the remnant then is

$$V(r_c) \simeq \frac{4\pi}{3}\,r_c^3 \simeq (3.4\cdot 10^4\ \mathrm{pc}^3)\left(\frac{E}{10^{51}\ \mathrm{erg}}\right)\left(\frac{n}{\mathrm{atom/cm}^3}\right)^{-1} \qquad (13.5)$$

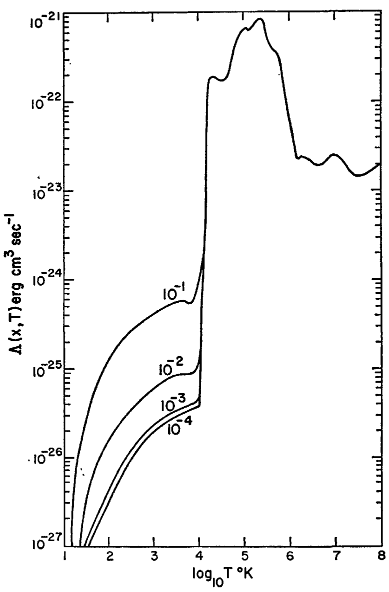
For each location the probability P to be overrun by the blastwave is identical to that being inside the volume $V(r_c)$ and given by the ratio of $V(r_c)$ and the total volume of the Galaxy, $V_{\rm gal}$. The average time between two encounters with shock waves that would heat to at least a million degrees then is

$$t_{rep} = t_{SN}\,\frac{V_{\rm gal}}{V(r_c)} \simeq (2\cdot 10^8\ \mathrm{years})\left(\frac{n}{\mathrm{atoms/cm}^3}\right)\left(\frac{T}{10^6\ \mathrm{K}}\right) \qquad (13.6)$$

where I have added the dependence on the gas density and the temperature, to which we gas is to be heated. Two hundred million years may seem a long time, but one has to compare this to the time it takes to cool the gas again.

13.2 The cooling of interstellar gas

Interstellar gas cools by radiating. Free electrons radiate on account of the acceleration in the Coulomb fields of ions, which we call bremsstrahlung. More efficient in many circumstances is line radiation from heavier atoms and ions, which includes so-called forbidden lines, which cannot be observed in the laboratory. These arise from states that have a very long lifetime with respect to radiative transitions. In the lab the atoms will all be collisionally de-excitated, but that doesn't work in space on account of the extremely low density, thus allowing the radiative transition to operate.



GURE 2. The interstellar cooling function $\Lambda(x,T)$ for various values of the fractional ionization x. The labels refer to the values of x.

The cooling function Λ is based on the probability of all transitions. The real energy loss rate per atom depends in addition on the number of interaction partners, i.e. the density of gas. The rate of change in the thermal energy density of gas also depends on how many particles per volume element loose energy, i.e. in total it scale with the square of the density.

$$\frac{\partial \epsilon}{\partial t} + \ldots = -\frac{\rho^2}{m_p^2}\,\Lambda(T) \qquad \text{or} \qquad \frac{\partial T}{\partial t} = -\frac{\rho}{m}\,\frac{\gamma - 1}{k}\,\Lambda(T) \tag{13.7}$$

At temperatures between 10^6 K and 10^8 K the cooling function is approximately constant and the rate of change of the temperature is

$$\frac{\partial T}{\partial t} \simeq -(2 \cdot 10^{-7}\ \mathrm{K/s})\,\left(\frac{n}{\mathrm{atoms/cm^3}}\right) \tag{13.8}$$

The timescale on which gas cools significantly is then estimated as

$$\tau_{\mathrm{cool}} = \frac{T}{|\dot{T}|} \simeq (1.5 \cdot 10^5\ \mathrm{years})\,\left(\frac{n}{\mathrm{atoms/cm^3}}\right)^{-1}\,\left(\frac{T}{10^6\ \mathrm{K}}\right) \tag{13.9}$$

Comparing Eq.13.6 and 13.9 we see that between two encounter with a SNR blastwave the gas will cool and likely reach an equilibrium state, unless the gas density is very low. That leaves us with the question: what are temperature equilibrium states for interstellar gas?

13.3 Temperature equilibria for gas

An equilibrium may exist between the continuous heating by energetic particles and radiation on one side and radiative cooling on the other side. Once out of equilibrium, gas has two ways of returning to equilibrium, (i) by an imbalance of heating versus cooling with the right sign and (ii) by expansion or compression. In the absence of large-scale expanding or compressing flows and under a hydrostatic equilibrium, what would be stable temperature equilibria? Are they possible at all temperatures?

An equilibrium is characterized by an exact balance of the heating and cooling terms.

$$\frac{k}{\gamma - 1}\,\dot{T} = \frac{1}{n}\,(\mathcal{H} - \mathcal{C}) = \frac{\mathcal{H}}{n} - n\,\Lambda = 0 \tag{13.10}$$

The heating arises from interactions with external agents, energetic particles and radiation, and will depend on density (in the equation for ϵ) but not on temperature. The heating per atom ($\frac{\mathcal{H}}{n}$) is usually constant.

We now wish to see, whether or not the equilibrium is stable. If the system is subjected to a temperature perturbation, does it return to the equilibrium temperature or would it move away? Let us denote the equilibrium temperature as T_0. Then

$$\frac{\mathcal{H}}{n} = n\,\Lambda(T_0) \tag{13.11}$$

Suppose an infinitesimal perturbation of the temperature $T = T_0 + \delta T$. Then

$$\frac{k}{\gamma - 1}\,\dot{T} = \frac{\mathcal{H}}{n} - n\,\Lambda(T_0 + \delta T) = n\,\Lambda(T_0) - n\,\Lambda(T_0 + \delta T) = -n\,\delta T\,\frac{\partial \Lambda}{\partial T}\Big|_{T_0} \tag{13.12}$$

So the heating/cooling imbalance imposed by a temperature perturbation is directed opposite to it, that is correcting the perturbation, if the cooling function is increasing with temperature. If the cooling function were falling off with temperature, the equilibrium would thus be unstable. The cooling function has a positive gradient for all temperature below 10^5 K, and steep gradients below 1000 K and around 10^4 K. This is why in the interstellar medium we find most of the gas at low temperatures, molecular hydrogen at about 40 K, atomic hydrogen between 50 K and a few hundred Kelvin, and ionized gas at 10^4 K, when it is ionized and heated by the strong radiation field of young massive stars (HII regions), and at 10^6 K. The phases in which the interstellar gas is observed can thus be understood in terms of temperature equilibria. Dense gas tends to have a lower temperature in equilibrium, where the cooling function is lower, because $\rho\,\Lambda$ must balance the heating. There is also a hot dilute gas, that doesn't have a characteristic temperature, because there is not stable temperature equilibrium. Typically the interstellar gas is in pressure equilibrium (Dense gas tends to have lower temperature!), so the different phases of the gas satisfy

$$n_{\mathrm{H}_2}\,T_{\mathrm{H}_2} \simeq n_{\mathrm{H}I}\,T_{\mathrm{H}I} \simeq n_{\mathrm{hot}}\,T_{\mathrm{hot}} \tag{13.13}$$

Gas at a few million degrees therefore tends to have a low density $n \approx 10^{-3} - 10^{-2}$ atoms/cm^3 and the cooling time (cf. Eq.13.9) is of the order of 10^8 years, so the gas is not in equilibrium.

13.4 Gravitational collapse

In homework problem 3 we have already treated the Jeans limit for the gravitational instability of a gas cloud. In real circumstances the gravitational instability very often arises as a consequence of a temperature instability. A density perturbation, that we can think of as a sound wave, creates a region of slightly higher density, so according to (13.10) the temperature balance is violated, cooling sets in, and compression is no longer adiabatic. Normally a compression would go along with an increase in pressure that provides a restoring force. In our case the cooling will somewhat reduce the pressure and thus the restoring force. This is in addition to the increase in self-gravity that we consider in the Jeans limit.

The energy loss or gain by emission or absorption of radiation is very important for gravitational collapse. Very early in the evolution of the universe, roughly before recombination at a redshift of $z \simeq 1000$, the ambient radiation was so strong that it was the dominant source of pressure. The Jeans mass for radiation is much larger than the total energy in the universe. Because matter and radiation were in thermodynamic equilibrium, the radiation could effectively prevent the growth of density perturbations. This is why the fluctuations in the microwave background, which represents density perturbations at $z \simeq 1000$, are of the order of a few parts in a million. After recombination matter and radiation were no longer in thermodynamic equilibrium, but evolved separately. However, the temperature had dropped to below a few thousand Kelvin, so the hydrogen atoms were in their ground states. Heavier nuclei had not yet been produced in stars, and so dust and smaller molecules did not exist, which today account for the bulk of cooling at low temperatures. Gravitational collapse was therefore largely adiabatic and therefore slow. Compare here our hydrostatic equilibrium calculation for a white dwarf with equation of state $P = P_0\, \rho^{5/3}$ which is the same as that of hydrogen gas given the appropriate choice of P_0. In (3.30) we found

$$\frac{5}{2}\, P_0\, \rho(0)^{2/3} = \frac{G\, M(R)}{R} \qquad \text{and} \quad M = x\, \rho(0)\, R^3 \approx \rho(0)\, R^3$$
$$\Rightarrow \qquad \frac{5}{2}\, \frac{P_0}{R} \simeq G\, M^{1/3} \tag{13.14}$$

Cooling corresponds to a reduction of P_0 because for an ideal gas $T = (m/k)\, P_0\, \rho^{2/3}$. Equation (13.14) tells us that following a decrease in P_0 the system will assume a new equilibrium configuration with smaller size R, but not continue collapsing. Cooling is therefore an important condition for gravitational collapse. The efficacy of cooling depends on density, which explains why the first phases of structure formation, the creation of galaxy clusters and superclusters, are very slow and still ongoing, whereas the formation of stars out of a dense molecular cloud takes only about a million years.

If cooling was very efficient, how quickly could a structure form? The structure could be anything, a star, a galaxy, or a cluster of galaxies, the difference being their mass and initial density. Let us consider a spherically symmetric gas cloud of mass M that is gravitationally unstable. The most unstable scenario is given if the pressure is exactly zero, so I can neglect radiative cooling. The situation is know a the free-fall limit.

The important principle is energy conservation between the potential energy and the kinetic energy associated with the radial infall. It suffices to follow the fall of one atom at the surface of the collapsing gas cloud. The gravitationally effective mass is then always the total mass of the cloud, and the acceleration is purely radial, for the angular momentum is zero. Then we have for the gravitational acceleration

$$\ddot{R} = -G\, \frac{M}{R^2} \qquad \Rightarrow\ \dot{R}\, \ddot{R} = -G\, \frac{M}{R^2}\, \dot{R}$$
$$\Rightarrow\ \frac{d}{dt}\left(\frac{\dot{R}^2}{2} - \frac{G\, M}{R}\right) = 0 \qquad \Rightarrow\ \frac{\dot{R}^2}{2} = \frac{G\, M}{R} + C \tag{13.15}$$

where C is a constant that we can fix using an initial condition, for example that the system was at rest at some earlier time when it had the radius R_0. Then the remaining differential equation can be solved using standard methods.

$$\dot{R}^2 = 2\,G\,M\left(\frac{1}{R} - \frac{1}{R_0}\right) \qquad \Rightarrow \frac{dR}{dt} = \pm\sqrt{\frac{2\,G\,M}{R_0}}\sqrt{\frac{R_0}{R} - 1} \tag{13.16}$$

Given that the acceleration is always negative, the minus-sign is the only relevant case.

$$\Rightarrow \sqrt{\frac{2\,G\,M}{R_0}}\,(t - t_0) = \int_R^{R_0} \frac{dr}{\sqrt{\frac{R_0}{r} - 1}} \tag{13.17}$$

To solve the integral on the RHS a variable transformation $y = \sqrt{r/R_0}$ yields the form

$$\int_R^{R_0} \frac{dr}{\sqrt{\frac{R_0}{r} - 1}} = R_0 \int_{R/R_0}^1 dy\,\frac{y^2}{\sqrt{1 - y^2}} \simeq R_0 \int_0^1 dy\,\frac{y^2}{\sqrt{1 - y^2}} \tag{13.18}$$

where I have used $R \ll R_0$ to set the lower limit to zero. A second transformation $y = \sin\theta$ with $dy = \sqrt{1 - y^2}\,d\theta$ yields

$$\sqrt{\frac{2\,G\,M}{R_0}}\,(t - t_0) = 2\,R_0 \int_0^{\pi/2} d\theta\;\sin^2\theta = \frac{\pi}{2}\,R_0 \tag{13.19}$$

Then

$$t - t_0 = \frac{\pi}{2}\sqrt{\frac{R_0^3}{2\,G\,M}} = \sqrt{\frac{3\pi}{32}\,\frac{1}{G\,\bar{\rho}}} \simeq \frac{1}{G\,\bar{\rho}} \tag{13.20}$$

where $\bar{\rho}$ is the average mass density of the collapsing cloud.

The collapse timescale of gas with density $n = 100\ \mathrm{H-atoms/cm^3}$ is about 5 million years, which is relevant for the collapse of a dense cloud of interstellar gas to be turned into stars. For a cluster of galaxies, the free-fall timescale of gas with density $n = 10^{-4}\ \mathrm{H-atoms/cm^3}$ is about 5 billion years or 40% of the age of the universe. This is another reason why stars form faster than galaxies and clusters of galaxies.

There is another issue: the free-fall time scales with the average gas density. Suppose the initial gas cloud has a higher density in the center than in its outer layers. The free-fall time for the central part is then shorter than that of the envelope, so the center will collapse away from the envelope, thus effectively providing the zero-pressure environment that we had assumed in our calculation. The free-fall collapse is therefore not homologous or self-similar, it doesn't preserve the radial density profile. This is an interesting aspect of supernovae explosions of massive stars, in which the core collapses to a neutron star or black hole, and incoming fresh material is undergoing a violent fusion reaction, all while the outer gas envelope hasn't moved.