ASTR 4017/8007: Diffuse Matter in the Universe

Class 15 Notes: Photoionised regions: thermodynamics and ionisation fronts

We are now done with dust, and we we will move on to our next phase of the ISM: ionised gas. In today's class we will focus on the chemical and thermal properties of ionised gas. Our goal will be to understand why such gas tends to have temperatures near 10^4 K, and how its ionisation state is regulated. We will then begin to study the properties of the ionisation fronts that bound such regions, before moving on to a full treatment of the dynamics of photoionised regions in the next class.

I. Heating and cooling

We begin the class with a discussion of the processes responsible for heating and cooling ionised gas.

A. Photoionisation heating

The main source of energy in a photoionised region is, not surprisingly, the photons that keep the gas ionised. Consider the photoionisation process for an atom of species X of charge $+r$ initially that is ionised by a photon with energy $h\nu$:

$$X^{+r} + h\nu \to X^{+r+1} + e^{-}. \tag{1}$$

If I_0 is the ionisation potential for species X^{+r}, then the electron on the right hand side carries kinetic energy $h(\nu - \nu_0)$, where $\nu_0 = I_0/h$. The free electron will generally collide with other electrons and atoms and thermalise, adding this energy to the thermal reservoir of the gas.

The total rate of photoionisations per atom is, as we showed earlier in the class,

$$\zeta(X^{+r}) = \int_{\nu_0}^{\infty} \sigma_{\mathrm{pi}}(\nu) \frac{4\pi J_\nu}{h\nu}\, d\nu, \tag{2}$$

where $\sigma_{\mathrm{pi}}(\nu)$ is the photoionisation cross section at frequency ν. Thus the rate per unit volume at which photoionisation injects thermal energy, assuming that the ejected electron always thermalises, is

$$\Gamma_{\mathrm{pi}}(X^{+r}) = n(X^{+r}) \int_{\nu_0}^{\infty} \sigma_{\mathrm{pi}}(\nu) \frac{4\pi J_\nu}{h\nu} h(\nu - \nu_0)\, d\nu, \tag{3}$$

and the mean energy injected per photoionisation is

$$E_{\mathrm{pi}}(X^{+r}) = \frac{\Gamma_{\mathrm{pi}}(X^{+r})}{n(X^{+r})\zeta(X^{+r})}. \tag{4}$$

To estimate $E_{\mathrm{pi}}(X^{+r})$, let us approximate the stellar spectrum as that of a blackbody with colour temperature T_c. This isn't completely accurate, since of course the real stellar spectrum has all sort of complicated line features, but since we're interested in a broadband average these will largely be smoothed out, and the approximation is reasonable. We use this value to define a dimensionless version of E_{pi}:

$$\psi \equiv \frac{E_{\mathrm{pi}}(X^{+r})}{kT_c} = \frac{1}{kT_c} \frac{\int_{\nu_0}^{\infty} \sigma_{\mathrm{pi}}(\nu) \frac{J_\nu}{h\nu} h(\nu - \nu_0)\, d\nu}{\int_{\nu_0}^{\infty} \sigma_{\mathrm{pi}}(\nu) \frac{J_\nu}{h\nu}\, d\nu}. \tag{5}$$

Near the star, where there have not yet been any significant number of absorptions, we simply have $J_\nu \propto B_\nu(T_c)$, so

$$\psi = \psi_0 \equiv \frac{1}{kT_c} \frac{\int_{\nu_0}^{\infty} \sigma_{\mathrm{pi}}(\nu) \frac{B_\nu(T_c)}{h\nu} h(\nu - \nu_0)\, d\nu}{\int_{\nu_0}^{\infty} \sigma_{\mathrm{pi}}(\nu) \frac{B_\nu(T_c)}{h\nu}\, d\nu}. \tag{6}$$

As we move away from the star and photons are absorbed both by atoms and dust grains, the ionising spectrum changes and ψ does too. This occurs because $\sigma_{\mathrm{pi}}(\nu)$ is not constant, and photons at frequencies where $\sigma_{\mathrm{pi}}(\nu)$ is larger tend to be absorbed closer to the star than those for which it is smaller. If we are thinking about hydrogen, where the cross section is maximum at threshold and declines at higher frequencies, then photons near the threshold will tend to be absorbed closer to the star, while higher energy photons are absorbed further away. Thus ψ rises with distance from the star.

However, suppose that we neglect dust absorption and assume that every photon above frequency ν_0 is absorbed within the photoionised region. In this case we know that, if we average over the entire volume of the photoionised region, the total spectrum of all photons absorbed must match that emitted by the star. Thus the spatially-averaged value of ψ is

$$\langle \psi \rangle = \frac{1}{kT_c} \frac{\int_{\nu_0}^{\infty} \frac{B_\nu(T_c)}{h\nu} h(\nu - \nu_0)\, d\nu}{\int_{\nu_0}^{\infty} \frac{B_\nu(T_c)}{h\nu}\, d\nu}. \tag{7}$$

Note that $\sigma_{\mathrm{pi}}(\nu)$ does not appear here, because the cross section is irrelevant: when averaging over all absorptions, we must get back to the initial stellar spectrum.

Values of ψ_0 and $\langle \psi \rangle$ are straightforward to calculate numerically, although our estimates for ψ_0 are only as good as our knowledge of $\sigma_{\mathrm{pi}}(\nu)$, which is very good for some elements and ionisation states (e.g., H, He) and mediocre for others (e.g., high ionisation states of iron). Not surprisingly they are typically within a factor of 2 of unity, since most photons emitted by the star are near kT_c in energy, and for hot stars this is not far from the hydrogen ionisation threshold.

The local heating rate of course depends on the abundance of the species being ionised:

$$\Gamma_{\mathrm{pi}}(X^{+r}) = n(X^{+r}) \zeta(X^{+r}) \psi k T_c \tag{8}$$

If the region in question is in photoionisation equilibrium, however, this density also depends on the ionising flux. If α is the recombination rate coefficient for recombination from ionisation state X^{+r+1} to state X^{+r}, then we have

$$\zeta(X^{+r})n(X^{+r}) = \alpha n_e n(X^{+r+1}), \tag{9}$$

where n_e is the electron density. Thus the heating rate becomes

$$\Gamma_{\mathrm{pi}}(X^{+r}) = \alpha n_e n(X^{+r+1})\psi k T_c. \tag{10}$$

This might not seem like much of a simplification, since all we have done is replace $n(X^{+r})$ with $n(X^{+r+1})$. However, it is a big simplification for hydrogen, since we can typically assume that all but a tiny fraction of the hydrogen is ionised. Thus $n(\mathrm{H}^+) \approx n(\mathrm{H})$, where for clarity we use H to refer to all ionisation states, and H^0 to refer to neutral hydrogen. Thus we have

$$\Gamma_{\mathrm{pi}}(\mathrm{H}) \approx \alpha_B n_{\mathrm{H}} n_e \psi k T_c, \tag{11}$$

where we make the usual case B assumption.

Thus we have computed the rate of photoionisation heating due to hydrogen ionisation in terms of the stellar surface temperature, the local density, and atomic constants. In most ionised regions driven by stars, as opposed to black holes or other sources with harder spectra, this is the dominant heating source, with a small contribution from He.

B. Cooling processes

Now let us consider processes that cool the gas.

1. Recombination radiation

Every time an electron radiatively recombines with an ion, the photon that is emitted in the process carries away an energy equal to the sum of the ionisation potential and the kinetic energy of the electron. This represents a net loss from the thermal reservoir of the gas, cooling it. For hydrogen (which dominates by far) the rate per unit volume at which this process removes energy from the gas is

$$\Lambda_{\mathrm{rr}} = \alpha_{A,B} n_e n_{\mathrm{H}^+} \langle E_{\mathrm{rr}} \rangle, \tag{12}$$

where the A, B subscript indicates that we should use the appropriate α for case A or B, and where $\langle E_{\mathrm{rr}} \rangle$ is the mean energy lost per recombination.

To compute the mean energy loss, we must compute the mean electron kinetic energy, weighted by the rate at which electrons of that kinetic energy combine. This is yet another integration over the Boltzmann distribution. The recombination rate due to electrons with velocities between E and $E + dE$ is

$$\sigma_{\mathrm{rr}}(E)(2E/m_e)^{1/2} f_E, \tag{13}$$

where the factor $(2E/m_e)^{1/2}$ is the electron velocity, and f_E is our usual Boltzmann distribution. Thus the recombination rate-weighted electron energy is

$$\langle E\rangle_{\mathrm{rr}} = \frac{\int_0^\infty \sigma_{\mathrm{rr}}(E)(2E/m_e)^{1/2} f_E E\,dE}{\int_0^\infty \sigma_{\mathrm{rr}}(E)(2E/m_e)^{1/2} f_E\,dE} = \frac{\int_0^\infty \sigma_{\mathrm{rr}}(v) v^4 e^{-m_e v^2/2kT}\left(\frac{1}{2}m_e v^2\right)dv}{\int_0^\infty \sigma_{\mathrm{rr}}(v) v^4 e^{-m_e v^2/2kT}\,dv}. \tag{14}$$

To evaluate this integral, we must know $\sigma_{\mathrm{rr}}(E)$. Even for hydrogen this has a fairly nasty, complicated functional form obtained by applying the Milne relation to the photoionisation cross section as a function of frequency for each $n\ell$ level of the recombined atom, then summing over all levels. Thus a truly accurate solution would require that we evaluate this average numerically. However, we can get a reasonably good idea of the answer by approximating the cross section by a powerlaw, at least over some temperature range. We let

$$\sigma_{\mathrm{rr}}(E) \approx \sigma_0 (E/E_0)^\gamma, \tag{15}$$

and with this form for σ the integrals can be done analytically:

$$\langle E_{\mathrm{rr}}\rangle \approx \frac{\Gamma(3+\gamma)}{\Gamma(2+\gamma)} kT = (2+\gamma)kT, \tag{16}$$

where here $\Gamma(x)$ is the Gamma function.

So far so good, but what value of γ should we adopt for this approximation? We can obtain one from the temperature dependence of the recombination rate coefficients. The rate coefficient is given by the usual expression

$$\alpha_{A,B} = 4\pi \left(\frac{m_e}{2\pi kT}\right)^{3/2} \int_0^\infty v^3 \sigma_{\mathrm{rr}}(v) e^{-m_e v^2/2kT}\,dv. \tag{17}$$

If we plug in our powerlaw approximation for σ_{rr} here, we obtain

$$\alpha_{A,B} = \left(\frac{8kT}{\pi m_e}\right)^{1/2} \sigma_0 \left(\frac{kT}{E_0}\right)^\gamma \Gamma(2+\gamma) \propto T^{\gamma+1/2}. \tag{18}$$

Thus our approximate value for γ can be obtained simply by fitting the logarithmic slope of α_A or α_B as a function of T, i.e.

$$\gamma_{A,B} = \frac{d\ln\alpha_{A,B}}{dT} - \frac{1}{2}. \tag{19}$$

Detailed numerical fits are given in *Draine*, but a rough estimate is $\gamma_A \approx -1.21$, and $\gamma_B \approx -1.32$, with a very weak dependence on T. Plugging this in, we have

$$\langle E_{\mathrm{rr}}\rangle_A \approx 0.79kT \qquad \langle E_{\mathrm{rr}}\rangle_B \approx 0.68kT. \tag{20}$$

Note that this is considerably less than the mean energy per electron of $(3/2)kT$. This is because lower velocity electrons have larger recombination cross sections, so the average is skewed to the slower-moving part of the electron population.

Plugging this into our original formula for $\Lambda_{\rm rr}$ gives the rate of recombination cooling per unit volume. Note that the result has a similar functional form as the photoionisation heating rate, just a different coefficient. For case B,

$$\Lambda_{\rm rr} = \alpha_B n_e n_{\rm H}(0.68kT), \tag{21}$$

where we have again taken $n_{\rm H^+} \approx n_{\rm H}$.

2. Free-free cooling

   Free-free emission also removes energy from the gas. We computed the cooling rate earlier in the class, so we can just repeat that result here: the cooling rate per unit volume due to free-free emission in hydrogen is

   $$\Lambda_{\rm ff} = \frac{32\pi}{3}\left(\frac{2\pi}{3}\right)^{1/2}\langle g_{\rm ff}\rangle \frac{e^6}{m_e^2 c^3}\left(\frac{m_e}{kT}\right)^{1/2} n_e n_{\rm H}, \tag{22}$$

   where again $n_{\rm H^+} \approx n_{\rm H}$. Note that the density dependence is the same as for recombination cooling, while the temperature dependence is slightly different.

   It is useful to compute the the free-free cooling rate per recombination. For case B, the recombination rate per unit volume is $\alpha_{\rm B} n_e n_{\rm H}$, and

   $$\frac{\Lambda_{\rm ff}}{\alpha_{\rm B} n_e n_{\rm H}} \approx 0.54 T_4^{0.37} kT. \tag{23}$$

   Thus the net cooling rate per recombination including both free-free and recombination radiation is

   $$\Lambda_{\rm ff+rr} \approx n_e n_{\rm H}(0.68 + 0.54 T_4^{0.37})kT \approx n_e n_{\rm H}(1.22kT), \tag{24}$$

   where in the second step we have taken $T_4 = 1$. Recombination plus free-fee therefore removes an energy $1.22kT$ per recombination.

   It is worth pointing out that we should not be surprised that we can combine the free-free and recombination cooling rates in this way, since both are fundamentally the same process – free electrons encountering ions, the accelerating and radiating as a result. The difference is just that free-free describes the distant encounters where the energy radiated is not enough to leave the ion and electron bound, while recombination radiation describes the close encounters where the amount of energy radiated is enough to leave behind a bound atom at the end. Thus the two processes are both describing electron-ion collisions, just over different ranges of impact parameter.

3. Collisionally-excited line cooling

   Of course there is more in the world than just hydrogen, and heavy elements complicate this picture somewhat by introducing another source of cooling: collisionally-excited line emission. (In case A they can also provide an extra heat source by absorbing more ionising photons than the hydrogen would alone, but for case B this doesn't make a substantial difference.)

Collisionally-excited line emission can be important because many species likely to be present in photoionised regions have L-S levels of their ground electronic state that lie $\sim 10^4$ K above ground, and thus can be collisionally excited in photoionised gas at this temperature. Once excited these atoms decay via forbidden transitions (they must be forbidden, since the electronic state is unchanged from ground), and the resulting photons carry off energy. Almost all of them escape the ionised region, which is generally optically thin to these lines.

Computing the net cooling rate via this process requires a two-step procedure that in practice must be carried out numerically, which is what programs like Cloudy do. First, one must balance the photoionisation and recombination rates to obtain the fractions of each element in various ionisation states. Second, one must solve for the level populations in the various L-S levels using the procedure we outlined a few weeks ago. Once that is done, the rate of collisionally-excited line cooling per unit volume is

$$\Lambda_{\mathrm{ce}} = \sum_X \sum_i n(X, i) \sum_{j<i} A_{ij} E_{ij}, \tag{25}$$

where $n(X, i)$ is the number density of atoms of species and ionisation state X in level i, and A_{ij} and E_{ij} are the Einstein A and difference in energies for between levels i and j.

Of course the results of these calculations depend on both the gas density and temperature, as well as the abundance. Thus the numerical procedure to determine the equilibrium temperature in photoionised gas is generally to, for a given density and abundance of heavy elements, numerically generate the function $\Lambda_{\mathrm{ce}}(T)$ for different T values, and search for a value of T such that $\Gamma_{\mathrm{pi}} = \Lambda_{\mathrm{rr}} + \Lambda_{\mathrm{ff}} + \Lambda_{\mathrm{ce}}$.

At Solar metallicity and for ionising sources similar to those found in Orion, the results of numerical calculations are that the most important cooling species are N II, O II, O III, S II, and S III. Of course this depends both on the spectrum of the ionising star, which alters the abundances of the different ionisation states, and on the element abundances. These calculations also show that Λ_{ce} is larger than $\Lambda_{\mathrm{rr}} + \Lambda_{\mathrm{ff}}$ by a factor of a few.

C. Thermal equilibrium

We are now in a position to calculate equilibrium temperatures in photoionised regions. First consider the case of pure hydrogen. We have shown that photoionisations add an energy $\psi k T_c$ for each photoionisation, while recombinations and free-free remove an energy $1.22kT$. Thus the equilibrium temperature is simply $T = (\psi/1.22)T_c \approx T_c$, i.e., the gas has about the same temperature as the colour temperature of the star that is ionising it. Since any star hot enough to put out a reasonable number of ionising photons tends to have T_c of a few times 10^4 K, this is a prima facie explanation why photoionised regions tend to have this temperature.

As we just mentioned, collisionally excited lines increase the cooling rate by a factor of a few at roughly Solar metallicity, and so the equilibrium temperature winds up being smaller than T_c by a factor of a few. In the Orion nebula, for example, the ionising flux is dominated by the O7 star θ^1 Ori C, which has $T_c \approx 3.5\times10^4$ K, whereas the estimated gas temperature is about 8050 K. The difference between T and T_c is smaller for lower metallicities, so that typical H II region temperatures in low metallicity galaxies are closer to $15,000$ K, while those in super-solar metallicity galaxies may be as small as ~ 6000 K.

We can also ask how long it takes for gas to reach equilibrium. If the gas is initially cold relative to its equilibrium temperature (generally a good assumption, since H II regions tend to go off in molecular clouds with temperatures of $\sim$ 10 K), then initially the heating rate will greatly exceed the cooling rate. In this case we can estimate the time required to reach the equilibrium temperature simply by dividing the equilibrium energy by the rate at which photoionisations supply energy.

In equilibrium, the mean energy per particle in an H II region of temperature T is $(3/2)kT$. The heating rate per unit volume is

$$\Gamma_{\mathrm{pi}} = n_{\mathrm{H}^0}\zeta\psi kT_c, \tag{26}$$

where ζ is the photoionisation rate. The heating rate per particle is

$$\frac{\Gamma_{\mathrm{pi}}}{n_{\mathrm{H}^0}} = \zeta\psi kT_c. \tag{27}$$

Thus the time required to reach thermal equilibrium is

$$t_{\mathrm{eq}} = \frac{(3/2)kT}{\Gamma_{\mathrm{pi}}/n_{\mathrm{H}^0}} = \frac{3}{2\psi}\frac{T}{T_c}\zeta^{-1}. \tag{28}$$

The prefactor in front of ζ is of order unity (or slightly smaller if we include metals so that $T < T_c$), so we conclude that the thermal equilibration time is comparable to $1/\zeta$, which is just the ionisation time. Thus the gas reaches thermal equilibrium on the same timescale over which it ionises, and we can safely assume that all the ionised gas is at its equilibrium temperature.

II. Ionisation fronts

Thus far we have assumed that the medium is in thermal and photoionisation equilibrium. Now let us relax those assumptions, and study the time-dependent behaviour of the gas. Recall that the mean free path of the ionised photons through neutral gas is tiny. Optical depth unity corresponds to a column density of $\sim 10^{17}$ H atoms cm^{-2}, whereas the typical column density of the galactic disk is $\sim 10^{21}$ H atoms cm^{-2}, and regions of star formation where massive stars are prone to start ionising the gas are even denser. Thus we can approximate that the ionised region around the star will be a sphere with a very sharp boundary, and that, inside that region, the gas will be close to fully ionised. The boundary between these two regions is called an ionisation front.

A. Ionisation fronts without gas motion

We will begin by considering the simplest possible problem. We start with a pure hydrogen medium of uniform density $n_{\rm H}$ that is initially all neutral hydrogen, and at time zero we turn on an ionising source of ionising luminosity Q_0, and for now we hold the gas fixed in position. Thus there are no motions. We would like to compute the state of the gas as a function of the time t and its distance r from the ionising source.

Suppose that at some time t the radius is r_i. The instantaneous rate at which ionising photons are being produced is Q_0, and the instantaneous rate at which they are being destroyed by recombinations is (assuming case B) $(4/3)\pi r_i^3 \alpha_B n_{\rm H}^2$, i.e., just the volume of the ionised region multiplied by the recombination rate per unit volume. Thus net number of ionising photons added to the region (i.e., emitted minus removed by recombination) per unit time is

$$Q_0 - \frac{4}{3}\pi r_i^3 \alpha_B n_{\rm H}^2 = Q_0 \left[1 - \left(\frac{r_i}{r_S}\right)^3\right], \tag{29}$$

where $R_S = (3Q_0/4\pi\alpha_B n_{\rm H}^2)^{1/3}$ is the usual Strömgren radius.

The excess photons must escape from the ionised region and cause new ionisations at its boundary, allowing the region to expand. If the region of ionised gas expands a distance dr_i in a time dt, the number of new ionisations during this time is

$$4\pi r_i^2 n_{\rm H}\, dr_i. \tag{30}$$

This is just the volume of the shell of newly-ionised material multiplied by the density of neutral atoms inside it. This must balance the number of excess photons available for expanding the ionised region, which is just the excess we have computed times dt. Thus we require that

$$4\pi r_i^2 n_{\rm H} \frac{dr_i}{dt} = Q_0 \left[1 - \left(\frac{r_i}{r_S}\right)^3\right]. \tag{31}$$

To solve this equation it is helpful first to recast it in dimensionless form. We let $x = r_i/R_S$, and we let $\tau = t/t_{\rm rec}$, where $t_{\rm rec} = 1/\alpha_B n_{\rm H}$ is the recombination time, i.e., the time it takes an atom to recombine. With this change of variables, the equation nicely simplifies to

$$\frac{dx}{d\tau} = \frac{1 - x^3}{3x^2}, \tag{32}$$

which has the exact solution

$$x = \left(1 - e^{-\tau}\right)^{1/3} \qquad \Longrightarrow \qquad r_i = R_S \left(1 - e^{-\alpha_B n_{\rm H} t}\right)^{1/3}, \tag{33}$$

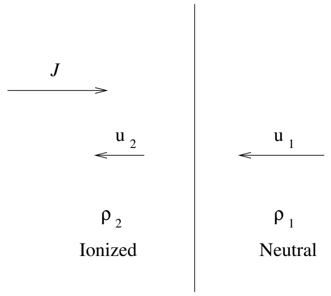
where we have set the constant of integration by requiring that $x = 0$ at $\tau = 0$. Note that we have implicitly taken α_B to be constant, but this is a reasonable assumption given that it depends only on temperature, and we have shown that the gas will reach its equilibrium temperature on a timescale comparable to the ionisation timescale.

At times $t \ll t_{\rm rec}$, the term in the exponential is small, and we have $r_i \propto t^{1/3}$ – this is exactly what we would expect if there were no recombinations, because it is equivalent the statement that the number of ionised atoms ($\propto r_i^3$) increases linearly in time. As the time approaches and then passes $t_{\rm rec}$, the expansion of the ionised region slows, and at late times the difference between the ionisation radius and R_S declines as $e^{-t/t_{\rm rec}}$.

B. General ionisation fronts

The calculation we have just done is artificial in the sense that we have magically assumed that the gas is at rest. However, if the ionised gas reaches temperatures of $\sim 10^4$ K, while the neutral gas is at ~ 10 K, clearly, such an equilibrium is artificial. The warm gas is overpressured relative to the cool gas, and this will induce motions. Our final goal for today is to understand the structure of ionisation fronts where the gas is not at rest.

Let us focus on a small patch of the interface between ionised and neutral gas, and let us adopt a coordinate system such that this interface is at rest. Let J be the flux per unit area of ionising photons arriving at the front. We denote gas in the neutral region ahead of the front with a subscript 1, and gas in the ionised region behind the front with subscript 2. The density in the neutral region is ρ_1, and the density in the ionised region is ρ_2. The velocity of gas approaching the ionisation front is u_1, and the velocity of gas flowing out of it on the ionised side is u_2. In the frame where the upstream, neutral gas is at rest, u_1 is simply the propagation speed of the front. We assume that gas is unmagnetised.



1. Jump conditions

As gas crosses the front, we must conserve mass, so we immediately require that

$$\rho_1 u_1 = \rho_2 u_2. \tag{34}$$

Moreover, we must conserve photons, so the rate at which neutral atoms cross the ionisation front and become ionised must match the rate at which ionising photons arrive at the front to ionise them. Thus

$$\rho_1 u_1 = \rho_2 u_2 = J\mu_i, \tag{35}$$

where μ_i is the mean mass per ion in the ionised gas. For pure hydrogen $\mu_i = m_{\rm H}$. For the standard helium abundance $\mu_i = 1.4 m_{\rm H}$ if the He remains neutral, and $\mu_i = (1.4/1.1)m_{\rm H}$ if it becomes once-ionised. The latter is usually the case.

If we imagine that the flux of ionising photons J is fixed by the source star, and that the upstream density ρ_1 is fixed by the material into which the I-front is propagating, then this means that we have solved for the front propagation speed: $u_1 = J\mu_i/\rho_1$.

We also require that momentum be conserved across the front. To see what this implies, we must write down the equation of momentum conservation. The x component of this reads:

$$\rho u \frac{\partial u}{\partial x} = -\frac{\partial p}{\partial x}, \tag{36}$$

where p is the gas pressure, u is the x component of the velocity, we have assumed that there is no magnetic or gravitational field present, and we have assumed that viscosity is negligible. This last assumption fails in a thin region near the front, but we can avoid that region. Rewriting this slightly, we have

$$\frac{\partial}{\partial x}\left(\rho u^2 + p\right) = 0. \tag{37}$$

In principle we should also include radiation pressure forces in this equation, but we will omit them because, except in special cases that we will handle later, they are small. If we integrate this equation across the front, then we find that it requires

$$\rho_1(u_1^2 + c_1^2) = \rho_2(u_2^2 + c_2^2), \tag{38}$$

where c_1 and c_2 are the sound speeds in regions 1 and 2, related to the pressures by $p_1 = \rho_1 c_1^2$ and similarly for region 2. We regard c_1 and c_2 as known, since they are fixed by the gas temperature, which is set by the radiative processes we discussed earlier.

Let us now combine mass and momentum conservation to see what they imply. We let $x = u_2/u_1 = \rho_1/\rho_2$, where we have used mass conservation to write $u_2/u_1 = \rho_1/\rho_2$. With this definition the equation of momentum conservation becomes

$$u_1 x^2 - (u_1^2 + c_1^2)x + c_2^2 = 0. \tag{39}$$

Thus the solution is

$$x = \frac{1}{2u_1^2}\left\{\left(u_1^2 + c_1^2\right) \pm \left[\left(u_1^2 + c_1^2\right)^2 - 4u_1^2 c_2^2\right]^{1/2}\right\}. \tag{40}$$

Since the ratio of velocities and densities had better be real, this means that we require

$$u_1^2 + c_1^2 > 2u_1 c_2. \tag{41}$$

This condition can be satisfied in two ways:

$$u_1 > u_R \equiv c_2 + \left(c_2^2 - c_1^2\right)^{1/2} \qquad \text{or} \qquad u_1 < u_D \equiv c_2 - \left(c_2^2 - c_1^2\right)^{1/2}. \tag{42}$$

Since it is usually the case that $c_2 \gg c_1$, since the ionised gas is much hotter than the neutral gas, we have

$$u_R \approx 2c_2 \qquad u_D \approx \frac{c_1^2}{2c_2}. \tag{43}$$

Note that u_1 is *not* a free variable: since $u_1 = J\mu_i/\rho_1$, it is fixed by the imposed flux of ionising photons J and the upstream density ρ_1. Thus it is entirely possible to have a value of u_1 that satisfies neither of these conditions. We'll get back to what happens in this case in a moment. For now, though, let's examine the solutions where u_1 does satisfy one of these conditions.

2. R-type fronts

First consider the case where $u_1 > u_R$, which occurs when $J > \rho_1 u_R/\mu_i$. This this occurs when ρ_1 is small, this is called the R type case, where R stands for rarefied. In this case the I-front travels at twice the ionised gas sound speed or faster. Since in general $u_1 > 2c_2 \gg c_1$, the ratio of densities x is approximately

$$x \approx \frac{1}{2}\left[1 \pm \sqrt{1 - \left(\frac{u_R}{u_1}\right)^2}\right]. \tag{44}$$

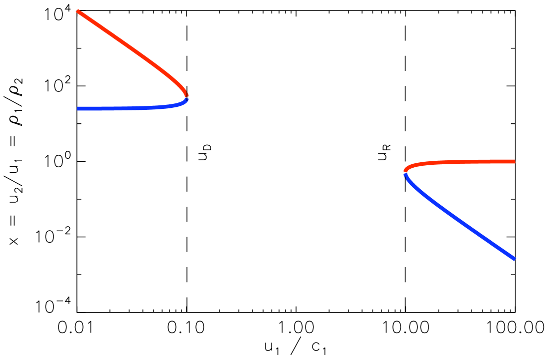
If we choose the positive root, then $x = \rho_1/\rho_2 > 1/2$. Thus the density is slightly higher in the downstream, ionised gas, but only slightly, and in the limit of a large ionising flux, so that $u_1 \gg c_2$, the density change goes to zero. We refer to this case as a strong R-type front. If on the other hand we choose the negative root, then $x < 1/2$, and for large ionising fluxes $x \ll 1/2$. Thus there is a large increase in density from the neutral to the ionised gas, and the ionisation front must also be a strong shock front. We refer to this as a weak R-type front. Note that if $u_1 = u_R \approx 2c_2$, the two cases are identical, and we refer to this as the R-critical case.

Physically the strong R-type front is the realistic one. Since $c_2 \gg c_1$, having $x \ll 1$ so that $\rho_1 \gg \rho_2$ requires that the pressure $p_2 = \rho_2 c_2^2$ be much, much greater in the ionised region than in the neutral region. This is possible only if there is some sort of piston pushing on the ionised gas to keep it accelerating into the neutral region, which doesn't tend to happen. For this reason the strong R-type front is the physically reasonable solution. It is characterized by a small to non-existent density jump at the front, coupled with a front that expands very quickly. We can see now that the case we solved of fixed gas that does not move is the limit of a very strong R-type front

3. D-type fronts

   Now let us consider the other possible solution: $u_1 < u_D$, which occurs when $J < \rho_1 u_D/\mu_1$. This occurs when ρ_1 is large, and the D stands for dense. Again, we have two roots to consider. For the positive root, called the weak D case, the ionised gas is much lower in density than the neutral gas, and it flows away from the front at a speed u_2 that is close to c_2. For the negative root, called the strong D case, the neutral and ionised regions are nearly in pressure balance. The density jump is not as large as in the weak D case. Both strong and weak D cases are astrophysically possible. As with R type fronts, if $u_1 = u_D$ then the strong and weak cases are the same, the the front is called D critical.

   We can summarize all possible states of the ionisation front in a single diagram:



   The curves on the left are the D type solutions, and those on the right are the R type, plotted for $c_2/c_1 = 5$. The red lines represent the positive roots, corresponding to weak D on the left and strong R on the right, while the blue are the negative roots, corresponding to strong D and weak R. The points where the two curves meet are the R- and D-crtical points. Recall that $x = u_2/u_1 = \rho_1/\rho_2$, so the y axis gives the size of the density jump, with values greater than unity indicating a density increase from neutral to ionised gas, and values less than unity indicating a density drop.

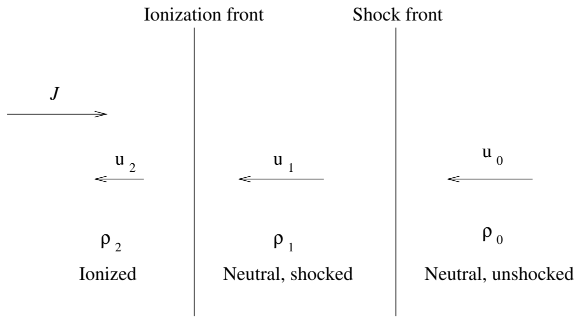
4. Intermediate velocities

   Recall that u_1 is fixed by the upstream density and the photoionising flux, $u_1 = J\mu_i/\rho_1$. This naturally raises the question: what happens if $u_1 > u_D$ and $u_1 < u_R$?

   The answer is that in this case there is only one way to satisfy the jump conditions, which is that ρ_1 must rise so that u_1 decreases to the point where it is equal to u_D. The way that this can happen is that a shock wave can run in front of the ionisation front, accelerating and compressing the upstream gas. The system then develops a two-front structure, in which a shock front precedes the I-front, and the postshock gas density is raised to the point where

$$\rho_1 = \frac{J\mu_i}{u_D}. \tag{45}$$

This is followed by a D type ionisation front, which ionises the gas.The configuration at this point is as shown below:



We let u_0 and ρ_0 be the velocity and density of the pre-shock material, and u_1 and ρ_1 be the density of the post-shock, pre-ionisation gas.

Once the density in the post-shock, pre-ionisation gas reaches the value of ρ_1 required to allow a D critical front to exist, we have $u_1 = u_D$ and

$$x = x_D \equiv \frac{1}{1 - \sqrt{1 - (c_1/c_2)^2}} \approx \frac{2c_2^2}{c_1^2}. \tag{46}$$

Thus the velocity of gas moving away from the ionisation front is

$$u_2 = x_D u_D \approx c_2. \tag{47}$$

At this point the post-shock gas moves away from the front at the ionised gas sound speed. To follow the behavior further, we must embed this model in a fuller picture of an expanding H II region.