Thermal Balance of Photoionized Regions

Wednesday, January 19, 2011

CONTENTS:

1. Introduction
2. The Energy Equation
3. Photoionization and Recombination
  A. Photoionization heating
  B. Recombination cooling
4. Cooling Processes
  A. Free-free radiation
  B. Infrared fine structure lines
  C. Optical forbidden lines
  D. Ultraviolet lines
5. Results

---

1. Introduction

Up until now, we have treated the temperature of a photoionized region as a fixed variable with a fiducial value of 10^4 K. Of course, the true temperature of such a region is determined by the physical processes that add and remove thermal energy. Our objective here is to investigate these processes.

This material is covered in Chapter 3 of Osterbrock & Ferland (on which most of these notes are based).

2. The Energy Equation

We consider the thermal kinetic energy per unit volume U of the gas. This is

$$U = \frac{3}{2} nkT,$$

where n is the total number density of particles. For pure ionized hydrogen, this is $2n_H$ (one electron and one proton), whereas for pure ionized $H^+ + He^+$, it is $2(1+f_{He})n_H$. The kinetic energy changes according to all heating and cooling processes,

$$\dot{U} = G - L.$$

where G represents the energy input from all sources and L represents the energy lost from all sources. In our analysis here we will leave out hydrodynamic motions (including work done by gas pressure), but we will discuss this again later in the course. We will focus on the heating and cooling terms associated with atomic processes as these dominate in H II regions (since their timescale is short compared to the timescale of evolution of the H II region as a whole). The roster of processes relevant for most ionized nebulae include:

Heating: Usually this is controlled entirely by photoionization.

Cooling: This is a mix of processes:

- Recombination
- Free-free radiation
- IR fine structure lines
- Optical forbidden lines
- UV lines (occasionally, at high temperatures)

That is, we may write:

$$\dot{U} = G_{\rm pi} - L_{\rm rec} - L_{\rm ff} - L_{\rm line,IR} - L_{\rm line,opt} - L_{\rm line,UV}.$$

Often the nebula can be treated in steady state, i.e. the left-hand side is set to zero.

The photoionization and recombination are dominated by H and He and for a fixed source luminosity depend only weakly on the temperature. The free-free cooling is temperature-dependent but again is dominated by H and He. The optical and IR lines dominate cooling in most modern-day photoionized nebulae. They only exist for metals, hence the key role of metals in setting the temperature of ionized gas. Since these lines dominate the ultimate release of energy from the nebula, they also contain a significant fraction of the luminosity, even though their parent elements (N, O, S) make up only a sliver of the nebula's elemental composition.

3. Photoionization and Recombination

We now evaluate the heating and cooling rates associated with ionization and recombination processes. For simplicity, we will just treat hydrogen, with the understanding that helium is similar (and often comparably important in the inner regions of the nebula).

A. PHOTOIONIZATION HEATING

We have previously considered the rate of ionization of hydrogen atoms per unit volume,

$$\dot{n}_{\rm pi} = n({\rm H}^0)\int_{\nu_i({\rm H}^0)}^{\infty} \frac{4\pi J_\nu(r)}{h\nu} a_\nu \, d\nu.$$

The amount of kinetic energy injected by a photon when it ionized a hydrogen atom is simply its energy minus the ionization energy. Thus

$$G_{\mathrm{pi}} = n(\mathrm{H}^0)\int_{\nu_i(\mathrm{H}^0)}^{\infty} \frac{4\pi J_\nu(r)}{\nu}[\nu - \nu_i(\mathrm{H}^0)]a_\nu \, d\nu.$$

If the ionization rates are in steady state, $\dot{n}_{\mathrm{pi}}$ is in steady state with the recombination rate, i.e. in the interior of the nebula it is roughly $\alpha_B(\mathrm{H}^0,T)n_H^2$. We therefore have

$$G_{\mathrm{pi}} = \alpha_{\mathrm{B}}(\mathrm{H}^0,T)n_{\mathrm{H}}^2 \frac{\int_{\nu_i(\mathrm{H}^0)}^{\infty} \frac{4\pi J_\nu(r)}{\nu}[\nu - \nu_i(\mathrm{H}^0)]a_\nu \, d\nu}{\int_{\nu_i(\mathrm{H}^0)}^{\infty} \frac{4\pi J_\nu(r)}{h\nu} a_\nu \, d\nu}.$$

The ratio of integrals depends only on the spectrum of the local radiation field. If the spectrum is that of a central star emitting as a blackbody at $T_* << 13.6$ eV/k, then J_ν is a rapidly declining exponential, leading to a value for the integral of kT_*. In this case, we have

$$G_{\mathrm{pi}} = \alpha_{\mathrm{B}}(\mathrm{H}^0,T)n_{\mathrm{H}}^2 kT_*.$$

This heating rate depends on the density *squared*. That is because photoionization heating occurs whenever a proton and electron find each other and recombine.

Warning: This is only an approximation, both because of the approximation to the integral, and also because the actual radiation spectrum in the nebula differs from that of the star. In particular, photons close to ionization threshold are preferentially absorbed, so the spectrum of radiation at the outskirts of the nebula is harder than it is in the interior. Thus it is actually possible for T_* to increase outward.

Note: We have used the Case B recombination coefficient in our equations. This is appropriate in an optically thick nebula where recombinations directly to the ground state lead to immediate re-absorption of the photon: there is no net energy gain or loss. Of course, one should technically include this process in both G_{pi} and L_{rec}, and the accounting in Osterbrock & Ferland does this – but within our level of approximation there is no advantage to doing so.

B. RECOMBINATION COOLING

When a hydrogen atom recombines, kinetic energy disappears from the nebula and (along with the ionization energy of the atom) is radiated away. The amount of cooling is:

$$L_{\mathrm{rec}} = \alpha_{\mathrm{B}}(\mathrm{H}^0,T)n_{\mathrm{H}}^2\langle E_e\rangle,$$

where $\langle E_e\rangle$ is the mean energy of a recombining electron. The mean energy of the electrons is $3kT/2$, but since slower electrons are more Coulomb-focused and hence more easily captured by protons, $\langle E_e\rangle$ is somewhat lower, roughly $0.7kT$. The product $\alpha_B\langle E_e\rangle/kT$ is often denoted β_B. Then:

$$L_{\rm rec} = \beta_{\rm B}({\rm H}^0,T)n_{\rm H}^2 kT.$$

The net heating (or cooling) due to ionizations and recombinations is then:

$$G_{\rm pi} - L_{\rm rec} = \alpha_{\rm B}({\rm H}^0,T)n_{\rm H}^2 k\left[T_* - \frac{\beta_{\rm B}({\rm H}^0,T)}{\alpha_{\rm B}({\rm H}^0,T)}T\right] \approx \alpha_{\rm B}({\rm H}^0,T)n_{\rm H}^2 k(T_* - 0.7T).$$

Note that this is positive if T is small and negative if T is large, implying a stable equilibrium. If these were the only processes at work, the nebula would reach equilibrium at a temperature of $\sim 1.4T_*$, or $\sim 6\times 10^4$ K for a nebula surrounding an O star. However, this is almost never the case.

4. Cooling Processes

We are now ready to consider the additional cooling processes that affect H II regions. All of these will reduce the equilibrium temperature below the above analysis.

A. FREE-FREE RADIATION

Free-free radiation arises from the collision of electrons with ions (mainly protons). The computation of its emissivity will be a homework exercise; the result for fully ionized gas is:

$$L_{\rm ff} \sim (2\times 10^{-27}\ {\rm erg\ cm^3\ s^{-1}\ K^{-1/2}})T^{1/2}n_{\rm H}^2.$$

The luminosity is usually small compared to that associated with recombinations (except for very hot gas) and does not have a major effect on the thermal balance for H II regions. It is important for two reasons:

- Free-free radiation has a flat spectrum ($j_\nu \approx$constant) from the quantum cutoff ($h\nu \sim kT$) in the optical down to the very lowest frequencies (in the radio). While most of the radiation comes out in the optical (since that is where most of the bandwidth is), the continuum extends down into the radio where the central star (peaking in the UV) and dust (peaking in the IR) are negligible. Thus, *free-free radiation usually dominates the radio continuum of H II regions*.
- Free-free radiation has a larger temperature exponent than recombination cooling. Therefore in extremely hot gases (e.g. in galaxy clusters) it is a dominant contributor to energy loss.

B. INFRARED FINE STRUCTURE LINES

Another major cooling mechanism, which dominates at low temperatures, is excitation of the fine structure lines of the metals with partially filled p subshells (except np^3). As a simple and concrete (but not dominant) example, we will take [Ne II] λ12.8μm.

The ground configuration here has two levels: the true ground level ${}^2\mathrm{P}^o_{3/2}$, and the first excitation ${}^2\mathrm{P}^o_{1/2}$.

Such a two-level system is controlled by two processes: radiative decay,

$$\mathrm{Ne}^+({}^2\mathrm{P}^o_{1/2}) \rightarrow \mathrm{Ne}^+({}^2\mathrm{P}^o_{3/2}) + \gamma(12.8\mu\mathrm{m}),$$

and collisions, usually from the fast-moving electrons:

$$\mathrm{Ne}^+({}^2\mathrm{P}^o_{1/2}) + e^- \leftrightarrow \mathrm{Ne}^+({}^2\mathrm{P}^o_{3/2}) + e^-.$$

In such collisions, the electron spin usually does *not* change (remember, it only enters the Hamiltonian through relativistic corrections). Rather, the incoming electron either swaps places with one of the electrons in the atom, or a transfer of orbital angular momentum occurs.

To an order of magnitude, the downward collisional rate coefficient $q_\downarrow$ is given by a cross section times a velocity. A typical cross section might be $\sim 10^{-15}$ cm^2, and an electron velocity at 10^4 K is of order 10^8 cm s^{-1}, so we would expect $q_\downarrow \sim 10^{-7}$ cm^3 s^{-1}. This is indeed the correct order of magnitude. The total decay rate of the excited level is then the sum of the radiative rate (A) and the collisional rate ($q_\downarrow$). Defining the **critical density**

$$n_{\mathrm{cr}} = \frac{A}{q_\downarrow},$$

we see that the excited level ${}^2\mathrm{P}^o_{1/2}$ decays primarily via radiation for $n<n_{\mathrm{cr}}$, and primarily via collisional de-excitation for $n>n_{\mathrm{cr}}$. Note that $q_\downarrow$ and n_{cr} can depend on temperature. The critical density tends to be higher for the faster (i.e. shorter wavelength) transitions in heavier elements.

Of course, collisions can also excite the upper level. We can compute the excitation rate $q_\uparrow$ by using the principle of detailed balance. In thermal equilibrium, the ratio of the excited to the ground state would be given by Boltzmann statistics,

$$\frac{n_{\mathrm{exc}}}{n_{\mathrm{gnd}}} = \frac{g_{\mathrm{exc}}}{g_{\mathrm{gnd}}} e^{-\Delta E/kT},$$

where the g's are the statistical weights (sometimes denoted ω). They are simply given by $2J+1$. If we include only collisions, then equilibrium should be reached, which tells us the ratio of upward to downward collisional rates. This implies that:

$$q_\uparrow = q_\downarrow \frac{g_{\mathrm{exc}}}{g_{\mathrm{gnd}}} e^{-\Delta E/kT}.$$

The steady-state excited level population can be determined by setting the total upward and downward transitions to be equal:

$$n_{\text{gnd}} n_e q_\uparrow = n_{\text{exc}}(n_e q_\downarrow + A).$$

This simplifies to:

$$\frac{n_{\text{exc}}}{n_{\text{gnd}}} = \frac{n_e q_\uparrow}{n_e q_\downarrow + A},$$

or

$$\frac{n_{\text{exc}}}{n_{\text{ion}}} = \frac{n_e q_\uparrow}{n_e(q_\uparrow + q_\downarrow) + A}.$$

We may now compute the rate of photon emission from this ion, which is An_{exc}. The cooling rate is then determined by counting $h\nu$ of lost energy for each emission:

$$L_{\text{line}} = \frac{A n_e n_{\text{ion}} q_\uparrow}{n_e(q_\uparrow + q_\downarrow) + A} \Delta E.$$

This has two limiting cases of interest. If the electron density is low, $n \ll n_{\text{cr}}$, then the denominator is dominated by A, whereas if the density is high, the A in the denominator is negligible. This leads to:

$$L_{\text{line}} \to \begin{cases} q_\uparrow n_e n_{\text{ion}} \Delta E & n << n_{\text{cr}} \\ \dfrac{1}{1 + \frac{g_{\text{gnd}}}{g_{\text{exc}}} e^{-\Delta E/kT}} n_{\text{ion}} \Delta E & n >> n_{\text{cr}} \end{cases}.$$

In particular, at low density the line luminosity is proportional to density squared (and to the abundance of the ion in question). This is because each collisional excitation results in the emission of a photon. But at high density the line luminosity becomes simply proportional to density because the ions reach thermal equilibrium among their energy levels. Note that for the fine structure lines in ionized regions, $\Delta E \ll kT$. This limits their utility as coolants, since the energy emitted per photon is low, but the $e^{-\Delta E/kT}$ factor required to excite them provides no suppression, so they may dominate in regions that are too cool to excite the optical lines.

The np^2 and np^4 ions have three fine structure levels, so the analysis is more complex, but the physics is the same.

The major fine structure coolant in H II regions is [O III] λ52,88μm (due to the high abundance of oxygen, and – unlike carbon – the presence of an ionization stage with fine structure in the ground level), with critical densities of 510 and 3600 cm^{-3}. In neutral regions, [C II] λ158μm is significant. These latter species may be excited in collisions with H atoms as these are far more common than electrons in the neutral regions.

As a simple example, the [O III] λ52,88μm lines at low density in ionized gas have a cooling rate of

$$L_{\rm line,IR}(\mathrm{O}^{2+}) \sim (1\times10^{-21}\ \mathrm{erg\ cm^3\ s^{-1}})\frac{n(\mathrm{O}^{2+})}{n_{\rm H}}T_4^{-1/2}n_{\rm H}^2.$$

For solar oxygen abundances, this exceeds the free-free cooling rate by a factor of a few. Above the critical density, these lines rapidly become weak compared to the free-free.

C. OPTICAL FORBIDDEN LINES

The same principles that control the IR cooling lines also control the collisional excitation of the optical forbidden lines. The key difference is that the excitation energies are now several eV, so the Boltzmann factor $e^{-\Delta E/kT}$ implies that the cooling rates are steeply increasing functions of temperature.

Since the optical lines exist only for np^2, np^3, and np^4 ions, simple abundance considerations (and a preference for low-lying excited states) dictate that [N II], [O II], and [O III] will be the major coolants, with [Ne V] being important in He^{2+} zones:

Species	Level	Wavelengths	Excitation energy	Critical density
[N II]	$^1D_2^e$	6548, 6583 Å	1.90 eV	6.6×10^4 cm^{-3}
[O III]	$^1D_2^e$	4959, 5007 Å	2.51 eV	6.8×10^5 cm^{-3}
[O II]	$^2D_{5/2}^o$	3729 Å	3.33 eV	3.4×10^3 cm^{-3}
	$^2D_{3/2}^o$	3726 Å	3.33 eV	1.5×10^4 cm^{-3}
[Ne V]	$^1D_2^e$	3346, 3426 Å	3.62 eV	1.3×10^7 cm^{-3}

D. ULTRAVIOLET LINES

If the gas reaches very high temperatures, the allowed (and semiforbidden) UV lines can be excited. Examples would be gas at very low metallicity, which due to the absence of other cooling mechanisms can be heated to where Lyman-α emission is possible (activation energy 10.2 eV); or gas at very high density where the forbidden lines (either IR or optical) are suppressed. The latter case occurs in quasars and is responsible for the production of many of the prominent broad lines, e.g. C IV λ1548,1551Å, C III] λ1909Å, and Mg II λ2796,2803Å (excitation energies 4.4—8.0 eV).

5. Results

The thermal balance equation can be re-written as:

$$\frac{G_{\rm pi}-L_{\rm rec}}{n_{\rm H}^2} = \frac{L_{\rm ff}-L_{\rm line,IR}-L_{\rm line,opt}-L_{\rm line,UV}}{n_{\rm H}^2}.$$

The left-hand side is a decreasing function of temperature, while the right-hand side (when optical lines are significant) is increasing. Therefore a steady state temperature is reached. It is typically of the order of 10^4 K, depending logarithmically on properties such as T_* and the abundances.

As a dramatically oversimplified example, if the density is low and the cooling comes from the [O III] optical lines, then we find

$$3.6\times10^{-25}T_4^{0.3}(T_{*,4}-0.7T_4)=9\times10^{-20}\frac{n(\mathrm{O}^{2+})}{n_\mathrm{H}}T_4^{-1/2}e^{-2.9/T_4}.$$

The temperature T is determined by where the exponential factor balances out all others, hence there is a logarithmic dependence of T on other properties. In particular, increasing the oxygen abundance of the gas causes the temperature to go down until the right-hand side equals the left-hand side. Thus the luminosities in the [O III] lines depend only weakly on metallicity. (In fact, at low metallicity the [O III] lines can become stronger because the central star ionizes more of the O II to O III, and hence changes the ratios of metal coolants.)