III Ionized Hydrogen (HII) Regions

Ionized atomic Hydrogen regions, broadly termed "HII Regions", are composed of gas ionized by photons with energies above the Hydrogen ionization energy of 13.6eV. These objects include "Classical HII Regions" ionized by hot O or B stars (or clusters of such stars) and associated with regions of recent massive-star formation, and "Planetary Nebulae", the ejected outer envelopes of AGB stars photoionized by the hot remnant stellar core. While the physical origins of these types of gaseous nebulae are very different, the physics governing them is basically the same. We will refer to all of these as "HII Regions" generically in this section.

The UV, visible and IR spectra of HII regions are very rich in emission lines, primarily collisionally excited lines of metal ions and recombination lines of Hydrogen and Helium. HII regions are also observed at radio wavelengths, emitting radio free-free emission from thermalized electrons and radio recombination lines from highly excited states of H, He, and some metals (e.g., H109α and C lines).

Three processes govern the physics of HII regions:

1. **Photoionization Equilibrium,** the balance between photoionization and recombination. This determines the structure of the nebula and the rough spatial distribution of ionic states of the elements in the ionized zone.
2. **Thermal Balance** between heating and cooling. Heating is dominated by photoelectrons ejected from Hydrogen and Helium with thermal energies of a few eV. Cooling is dominated in most HII regions by electron-ion impact excitation of metal ions followed by emission of "forbidden" lines from low-lying fine structure levels. It is these cooling lines that give HII regions their characteristic spectra.
3. **Hydrodynamics**, including shocks, ionization and photodissociation fronts, and outflows and winds from the embedded stars.

The first two will be treated here in some detail, while the hydrodynamics of HII regions is properly the topic of the Radiative Hydrodynamics course (Astronomy 825).

III-1 Photoionization Equilibrium & Ionization Structure

Photoionization Equilibrium

Photoionization Equilibrium is the detailed balance between photoionization and recombination by electrons and ions. We will start by considering the case of a pure Hydrogen nebula. Such objects do not exist in nature, but they are useful for introducing the basic physics. A simple order-of-magnitude treatment will suffice at this step to motivate the details to follow.

At a given location within an HII region, the rate of photoionizations per unit volume is balanced by the rate of recombination per unit volume.

The **Volumetric Photoionization Rate** is:

$$n_{H^0}\int_{\nu_0}^{\infty}\frac{4\pi J_\nu}{h\nu}\,a_\nu d\nu = (\#\mathrm{H}^0 \text{ atoms/volume})\times(\text{flux of ionizing photons})\times(\text{photoionization cross-section})$$

The integral is taken over all photons with $h\nu \geq h\nu_0$, where $h\nu_0$ is the Ionization Potential of H^0, 13.59eV for ionization out of the 1s ground state.
The **Volumetric Recombination Rate** is:

$$n_e n_p \alpha(H^0, T) = (\#\ e^-/\text{volume}) \times (\#\ \text{protons/volume}) \times (\text{recombination coefficient})$$

The recombination coefficient depends weakly on the temperature of the electrons. Note that the photoionization and recombination rates have units of $cm^{-3}\ s^{-1}$ (i.e., volumetric rates).
To a first approximation, we will treat all ionizing photons as arising from a single central ionizing source (e.g., an O star), and we will ignore for now contributions from the diffuse radiation within the nebula. At a distance r from the central star, the flux of photons at frequency ν is:

$$4\pi J_\nu = \left(\frac{4\pi R_*^2}{4\pi r^2}\right) \times \pi F_\nu(0) = \frac{L_\nu}{4\pi r^2}$$

The first term in the middle above is the geometric dilution factor, and the second is the photosphere flux. The densities of the electrons, protons, and neutral Hydrogen atoms are all related via the **Hydrogen Neutral Fraction**, ξ, such that

$$n_e = n_p = (1-\xi) n_H$$
$$n_{H^0} = \xi\, n_H$$

The neutral fraction ξ is defined such that

$$\xi = 0 \Rightarrow \text{fully ionized}\ (n_e = n_p = n_H)$$
$$\xi = 1 \Rightarrow \text{fully neutral}\ (n_{H^0} = n_H;\ n_e = n_p = 0)$$

The ionization equilibrium condition at this location in the nebula is found by balancing the local photoionization and recombination rates:

$$\xi\ n_H \int_{\nu_0}^{\infty} \frac{L_\nu}{4\pi r^2} \frac{a_\nu}{h\nu} d\nu = (1-\xi)^2 n_H^2 \alpha(H^0, T)$$

For example, consider a region of gas with n_H=10 cm^{-3} located at r=5pc from a T_*=40,000 K star (an O6.5V star). To work this out numerically we need two numbers.

The first is the number of ionizing photons/second emitted by the central star:

$$\int_{\nu_0}^{\infty} \frac{L_\nu}{h\nu} d\nu = 6.6 \times 10^{48}\ \text{photons/sec}$$

The second is the **Photoionization Cross-Section** for the H^0 $1s^2S$ ground state:

$$a_\nu \approx a_{\nu_0} \left(\frac{\nu}{\nu_0}\right)^{-3} \quad \text{for}\ \nu \geq \nu_0$$

Where $a_{\nu 0}$=6×10^{-18} cm^2. Evaluating these quantities at r=5pc, the number of ionizations per second is

$$\int_{\nu_0}^{\infty} \frac{L_\nu}{4\pi r^2} \frac{a_\nu}{h\nu} d\nu \approx 10^{-8}\ \text{sec}$$

This implies a characteristic ionization timescale of ~10^8 sec or a little over 3 years.

For typical nebular conditions, the recombination coefficient is:

$$\alpha(H^0, T) \approx 4 \times 10^{-13}\ cm^3\ s^{-1}$$

This corresponds to a characteristic **Recombination Time** of

$$t_{rec} = \frac{1}{n_e \alpha(H^0, T)}$$

For $n_e \approx 10$ cm^{-3} (assuming nearly completely ionized H) the recombination time is $\sim 3\times10^{11}$ sec or $\sim 10^4$ years. In such a region, once a neutral Hydrogen atom is photoionized it stays ionized for a long time before recombining, and when it recombines it is relatively quickly photoionized again. This means that the instantaneous neutral fraction in this region should be very small. Putting all of the pieces together for this parcel of gas 5pc from the central star the neutral fraction ξ is:

$$\xi \times (10^{-8}) = (1-\xi)^2 n_H (4\times10^{-13})$$

$$\xi \approx 4\times10^{-4} \ll 1$$

Thus this gas parcel is nearly completely ionized.

The transition between ionized ($\xi << 1$) and neutral ($\xi \approx 1$) is very abrupt. The mean-free path of an ionizing photon with $\nu = \nu_0$ is

$$\ell_{\nu_0} = \frac{1}{a_{\nu_0} n_{H^0}}$$

At the location where $\xi = 0.5$ (50% neutral), $n_{H^0} = \xi\, n_H = 0.5\times10 = 5$ cm^{-3} and the ionization cross section at the ionization threshold is $a_{\nu_0} \approx 6\times10^{-18}$ cm^2. The resulting mean-free path is:

$$\ell_{\nu_0} \approx 3\times10^{16} \text{ cm} \approx 0.01 \text{ pc}.$$

The size of the transition region is therefore only ~0.2% of the size of the ionized nebula (at least 5pc in this case), hence ionized nebulae around stars should be very sharp-edged.

The radius of the fully ionized region is called the "Strömgren Radius", and an idealized spherical pure-Hydrogen HII region is often called a "Strömgren Sphere" (Strömgren1939, ApJ, 89, 529). While in practice very few HII regions resemble Strömgren spheres in detail, the Strömgren radius is an excellent estimate of the size scale of an HII region, and so it remains in use in a variety of contexts.

Ionization Structure

The idealized case treated above made a number of assumptions that we need to explore before proceeding further.

1. What photoionization cross-section should be used?

The photoionization cross-section for HI in the 1s ^{2}S ground state is used, ignoring ionizations out of any excited states. This is often called the **Nebular Approximation**.

When a proton recombines with an electron, the electron can end up in a highly excited state, followed quickly by a radiative cascade into the ground state.

For small principal quantum numbers, the transition probability is $A_{ij} \approx 10^8$ s^{-1}, corresponding to a lifetime in the excited state of $t_{ij} \approx 10^{-8}$ seconds! Even for n>30 quantum states, the lifetime is $\sim 10^{-4}$ seconds. Compared to the typical photoionization time of $\sim 10^8$ sec, excited H atoms have plenty of time to cascade into the 1s ^{2}S ground state following recombination.

To first order all photoionizations in a nebula are from the 1s^{2}S ground state.

The exception is when the electron cascades into the metastable 2s ^{2}S excited state. The only downward radiative transition out of this state is a highly forbidden 2-photon decay into the 1s ground

state, which has a transition probability $A_{2^2S,1^2S} = 8.23\ \mathrm{s}^{-1}$, implying a lifetime of ~0.12 seconds.

Even this timescale is short compared to the photoionization time under most conditions. As we'll see later, this 2-photon transition is an important contributor to the nebular continuum.

2. What happens to the electrons after being released photoionization?

Photoionization ejects an electron from H^0 with a kinetic energy of

$$\frac{1}{2} m_e \mathrm{v}^2 = h(\nu - \nu_0)$$

The "spectrum" of electron kinetic energies immediately after photoionization thus reflects the energy spectrum of the ionizing photons. However, electron–electron collisions are very efficient.

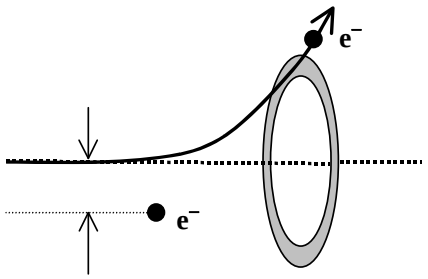


Figure III-1: Diagram of the geometry of an elastic electron–electron collision.

If the impact parameter, *b*, is very small there is a strong encounter resulting in scattering (Figure III-1). If *b* is large, there is no scattering. The borderline case occurs when

$$\tfrac{1}{2} m_e \mathrm{v}^2 \approx \frac{e^2}{b}$$
$$(K.E.) \approx (P.E.)$$

The cross-section is

$$\pi b^2 \approx \pi \left(\frac{e^2}{\frac{1}{2} m_e \mathrm{v}^2} \right)^2$$
$$\approx 4\pi \left(\frac{e^2}{m_e \mathrm{v}^2} \right)^2$$

For thermal electrons with $T_e \approx 10^4$K, $\frac{1}{2} m_e \mathrm{v}^2 = \frac{3}{2} kT_e$, and the cross-section is

$$\pi b^2 (10^4\ K) \approx 10^{-13}\ \mathrm{cm}^2$$

Thus the e-e collision cross-section is about 4 orders of magnitude *larger* than the photoionization cross-section at the ionization threshold ($a_0 \approx 6 \times 10^{-18}$ cm^2). The result is that the kinetic energies of the electrons will very quickly *thermalize* into a Maxwellian velocity distribution with $T_e \approx T_{kin}$ compared to other processes at work (recombination or ionization). This is the situation that Spitzer has referred to as *kinetic equilibrium*.

3. What is the Recombination Coefficient?

Because the photoelectrons quickly thermalize, the recombination coefficient into the n^2L state will be

$$\alpha_{n^2L}(T) = \int \sigma_{n^2L}(\mathrm{v})\mathrm{v}f(\mathrm{v})d\mathrm{v}$$

where f(v) is a Maxwellian distribution. The recombination cross section depends on velocity like

$$\sigma_{n^2L} \propto \mathrm{v}^{-2}$$

A thermal distribution of electrons results in a recombination coefficient with a temperature dependence of

$$\alpha_{n^2L} \propto T^{-1/2}$$

The recombination cross-section is $\sigma_{n^2L} \approx 10^{-20} - 10^{-21}$ cm^2 for T≈10^4 K. Compared to cross-sections of 6×10^{-18} cm^2 for photoionization and 10^{-13} cm^2 for e-e scattering, recombination is a very slow process and the electrons will have plenty of time to thermalize.

In the nebular approximation, a rapid cascade of the electron follows recombination into the 1s ^{2}S ground state. We can thus define a **Total Recombination Coefficient**:

$$\begin{aligned}\alpha_A &= \sum_{n,L} \alpha_{n^2L}(H^0,T) \\ &= \sum_n \sum_{L=0}^{n-1} \alpha_{n^2L}(H^0,T) \\ &= \sum_n \alpha_n(T)\end{aligned}$$

to describe the process. For typical nebular conditions (T≈10^4 K), $\alpha_A \approx 4\times10^{-13}$ cm^3 s^{-1}, and the recombination time, t_{rec} is:

$$t_{rec} \approx \frac{1}{n_e \alpha_A(T)} \approx 3\times10^{12}\, n_e^{-1} \text{ sec} \approx 10^5\, n_e^{-1} \text{ years}$$

The recombination time is generally long compared to either the ionization timescale or the electron thermalization timescale in most nebulae.

The Pure Hydrogen Nebula

We will now consider the case of a static, homogeneous, pure H nebula ionized by photons from a single star. In Nebular Approximation all photoionizations will occur from the ground state. The photoionization equilibrium condition for a pure H nebula will then a balance between photoionization of H^0 and recombination of electrons and protons back into H^0:

$$n_{H^0} \int_{\nu_0}^{\infty} \frac{4\pi J_\nu}{h\nu} a_\nu d\nu = n_e n_p \alpha_A(T)$$

For photons with $\nu \geq \nu_0$, the equation of radiative transfer is written

$$\frac{dI_\nu}{ds} = -n_{H^0} a_\nu I_\nu + j_\nu$$

where

I_ν = specific intensity of the radiation field
j_ν = local emission coefficient (erg cm^{-3} s^{-1} sr^{-1} Hz^{-1})

The radiation field, I_ν, consists of two parts

$$I_\nu = I_{\nu s} + I_{\nu d} = (stellar) + (diffuse)$$

The stellar radiation field at a given location in the nebula, r, is given by

$$4\pi J_{\nu s} = \pi F_{\nu s}(r)$$
$$= \pi F_{\nu s}(R_*)\left(\frac{R_*}{r}\right)^2 e^{-\tau_\nu(r)}$$

The optical depth at location r is the integral from the star to that location:

$$\tau_\nu(r) = \int_0^r n_{H^0} a_\nu ds$$

This optical depth can also be written in terms of the optical depth, τ_0, at the ionization threshold:

$$\tau_\nu(r) = \frac{a_\nu}{a_0}\tau_0(r)$$

The diffuse part of the radiation field is found by solving the transfer equation:

$$\frac{dI_{\nu d}}{ds} = -n_{H^0} a_\nu I_{\nu d} + j_\nu$$

In the limit that $kT \ll h\nu_0$, the only source of diffuse ionizing photons is from recombinations directly into the ground state:

$$j_\nu(T) = \frac{2h\nu^3}{c^2}\left(\frac{h^2}{2\pi mkT}\right)^{3/2} a_\nu e^{-h(\nu-\nu_0)/kT} n_e n_p$$

For $\nu \geq \nu_0$, $j_\nu(T)$ is strongly peaked at $\nu=\nu_0$, and hence:

$$4\pi\int_{\nu_0}^{\infty}\frac{j_\nu}{h\nu}d\nu \approx n_e n_p \alpha_{1s}(T)$$

Since $\alpha_{1s} < \alpha_A$, the diffuse ionizing radiation field is always weaker than the stellar radiation field. In practice it is solved for iteratively. There are two relevant approximations:

Optically Thin Nebulae

Here $J_{\nu d} \approx 0$, and we only consider the ionizing radiation from the photoionizing star.

Optically Thick Nebulae

In this approximation none of the ionizing photons can escape the nebula. This means that all ionizing photons in the diffuse radiation field eventually get absorbed elsewhere in the nebula, hence:

$$4\pi\int\frac{j_\nu}{h\nu}dV = 4\pi\int n_{H^0} a_\nu \frac{j_{\nu d}}{h\nu}dV$$

The integral above is evaluated over the entire ionized volume of the nebula. This leads to the **On-The-Spot (OTS) Approximation** in which each new ionizing photon emitted following recombination into the ground state is absorbed physically close to where it was created:

$$J_{\nu d}(r) = \frac{j_\nu(r)}{n_{H^0} a_\nu} = \frac{j_\nu(r)}{\kappa_\nu(r)}$$

Near the ionization threshold, ν_0, the cross-section a_ν is large and the mean-free path is very short, so this is a good approximation in general, not just in this idealized case.

Returning to the pure Hydrogen nebula, since most nebulae will be optically thick, making the OTS approximation allows us to simplify the equation of photoionization equilibrium to

$$n_{H^0}\left(\frac{R_*}{r}\right)^2\int_{\nu_0}^{\infty}\frac{\pi F_\nu(R_*)}{h\nu}a_\nu e^{-\tau_\nu}d\nu = n_e n_p \alpha_B(H^0,T)$$

Here we have introduced a new recombination coefficient, α_B:

$$\alpha_B(H^0,T) = \alpha_A(H^0,T) - \alpha_{1^2S}(H^0,T) = \sum_{n=2}^{\infty}\alpha_n(H^0,T)$$

α_B is the total recombination coefficient α_A less contributions from recombinations directly into the ground state. This form of the photoionization equilibrium condition means that the following conditions will prevail in an optically thick nebula:

1. Photoionization by the stellar radiation field is balanced by recombination into *excited* states of H.
2. Recombinations directly into the 1s ^{2}S ground state emit ionizing photons that are quickly reabsorbed by the nebula, and so have no net effect on the overall ionization balance.

In order to solve the photoionization equilibrium condition, two inputs are required:

1. The stellar spectrum, $\pi F_\nu(R_*)$, usually derived from model stellar atmospheres.
2. The density distribution in the nebula: $n_H(r) = n_e(r) + n_p(r)$

The equations of photoionization equilibrium are then integrated outwards from the surface of the exciting star until the nebula becomes mostly neutral. This transition point, which occurs very rapidly, defines the **Strömgren Radius, r_1**, of an idealized Strömgren sphere. If we are considering ionization of a plane-parallel slab of gas, the length scale is often called the Strömgren "depth" of the nebula. In either case, the transition of the nebula from ionized to neutral defines the fundamental size scale of the nebula. We will proceed to use the traditional spherically symmetric case for our discussion.

We can estimate the radius of a Strömgren sphere as follows. H is nearly completely ionized inside of $r=r_1$ ($n_e=n_p\approx n_H$), and essentially neutral for $r>r_1$ ($n_e=n_p\approx 0$). r_1 will be the radius of a spherical volume inside of which all ionizing photons are absorbed, that is $r_1 = r(\tau_{\nu 0}=\infty)$. Since the optical depth to ionizing photons τ_ν is defined as

$$\frac{d\tau_\nu}{dr} = n_{H^0}a_\nu$$

Imposing the boundary condition that $\tau_\nu\rightarrow\infty$ as $r\rightarrow r_1$ leads to an ionization equilibrium condition of:

$$4\pi R_*^2\int_{\nu_0}^{\infty}\frac{\pi F_\nu(R_*)}{h\nu}d\nu\int_0^{\infty}d(-e^{-\tau_\nu}) = \int_0^{r_1} n_e n_p \alpha_B(H^0,T)4\pi r^2 dr$$

The first integral on the left can be rewritten in terms of the luminosity spectrum of the star L_ν:

$$4\pi R_*^2\int_{\nu_0}^{\infty}\frac{\pi F_\nu(R_*)}{h\nu}d\nu = \int_{\nu_0}^{\infty}\frac{L_\nu}{h\nu}d\nu$$

This is just the total number of H-ionizing photons emitted per second by the star:

$$Q(H^0) \equiv \int_{\nu_0}^{\infty}\frac{L_\nu}{h\nu}d\nu$$

For a homogenous density distribution, the right-hand side of the ionization equilibrium condition is easily integrated over radius, reducing the equation to:

$$Q(H^0) = \frac{4\pi}{3}r_1^3 n_H^2 \alpha_B(H^0,T)$$

Or, in words,

$$\begin{pmatrix}\text{Total \# of ionizing}\\ \text{photons/sec emitted}\end{pmatrix}=\begin{pmatrix}\text{Total \# of recombinations}\\ \text{into excited levels of } H^0 \text{ per second}\end{pmatrix}$$

In the homogenous case, this can be solved for the Strömgren radius, r_1, given the density of the nebula, n_H, and the number of ionizing photons from the central star, $Q(H^0)$.
In practice, values of $Q(H^0)$ are tabulated for model stellar atmospheres for O and B stars (this is pointless for cool stars that emit little or no ionizing radiation), and one usually assumes some constant density and temperature to use for the left side of the equation. AGN[2] Table 2.3 lists values of representative r_1 computed for a variety of stellar effective temperatures. These are useful for simple order-of-magnitude estimates of the sizes of ionized regions around stars.

An implicit assumption of all of this is that we can treat photoionization equilibrium as a stationary (time-independent) problem, but is this true?

Consider the timescale for ionizing a Strömgren Sphere – if we were to turn on an ionizing source in a homogeneous pure Hydrogen region, the time to ionize the volume is

$$t_{ionize}=\frac{\text{\# of ions to create}}{\text{\# of ionizing photons / second}}=\frac{(4\pi/3)r_1^3 n_H}{Q(H^0)}$$

however, for photoionization equilibrium, this becomes

$$t_{ionize}=\frac{1}{n_H\alpha_B(H^0,T)}=t_{recomb}\,!$$

Not too surprisingly, the ionization time is the recombination time in the region. Putting in reasonable numbers for n_H and α_B, the ionization time is

$$t_{ionization}\approx 10^3 n_{100}^{-1}\ \text{yr}$$

where n_{100} is the Hydrogen density in units of 100 cm^{-3}. Compared to this, the lifetime of a typical ionizing O or B star is a few$\times 10^6$ years. The ionized gas region will be heated to ~104 K by photoionization, resulting in a higher pressure than the surrounding neutral medium. The relevant timescale for hydrodynamical expansion is the sound-crossing time for the region:

$$t_{sound}=\frac{r_1}{c_s}$$

where c_s is the sound speed:

$$c_s=\sqrt{\frac{2kT}{m_H}}\approx 13T_4^{1/2}\ \text{km sec}^{-1}$$

For typical nebular conditions (T_4=1, n_H=100 cm^{-3}),

$$t_{sound}=\frac{\left(3Q(H^0)/4\pi n_H^2\alpha_B(H^0,T)\right)^{1/3}}{\left(2kT/m_H\right)^{1/2}}\approx 2\times 10^5\frac{Q_{49}^{1/3}}{n_{100}^{2/3}}\ \text{yr}$$

where Q_{49} is Q(H0) in units of 10^{49} photons/sec. For typical nebular densities, both the OB star main-sequence lifetime and the sound crossing time are much longer than the ionization time, and our assumption of a stationary system is justified. However, at high densities (e.g., an embedded HII region in a very dense molecular region) the sound-crossing and recombination timescales become comparable and we must treat the system as a dynamic, time-dependent problem.

Nebulae with Hydrogen and Helium

The next level of complexity is to consider the effect of adding Helium to the nebula. Helium is the next most abundant element after Hydrogen, with typical abundances by number of He/H≈0.1 in the Galaxy. The atomic physics of He is made more complex by its two electrons and three possible ionic forms, He^0, He^+ and He^{++}.

The ionization potentials (I.P.'s) for H^0, He^0, and He^+, are as follows:

Element/Ion	I.P.
$H^0 \rightarrow H^+$	$h\nu_1$=13.6 eV
$He^0 \rightarrow He^+$	$h\nu_2$=24.6 eV
$He^+ \rightarrow He^{++}$	$h\nu_3$=54.4 eV

The photoionization cross-sections are plotted in Figure III-2. Despite the fact that $n_{He} \approx 0.1 n_H$, the cross-section for He^0 ionization at its threshold is ~10 times larger than the H^0 ionization cross-section at the same frequency. This means that ionizing photons with $h\nu \geq h\nu_2$ will primarily ionize He^0 instead of H^0. This has a significant impact upon the ionization equilibrium balance, and the nebular structure will be strongly dependent on the details of the stellar ionizing continuum spectrum.

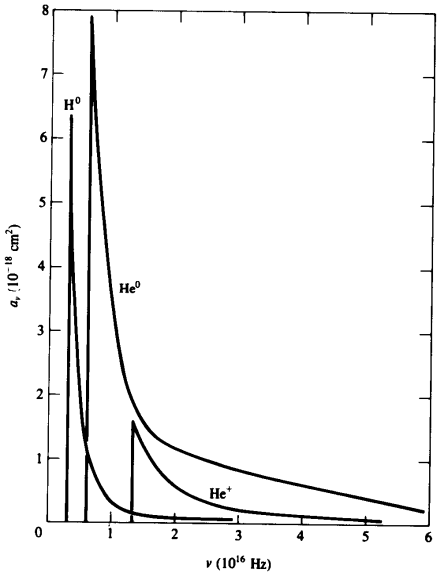


Figure III-2: Photoionization cross-sections for H^0, He^0, and He^+, reproduced from Osterbrock (AGN[2]).

For normal O and B stars, He^+ opacity in the stellar atmospheres produces a sharp, deep absorption edge in the emergent continuum for $h\nu \geq h\nu_3$, and so there is little or no He^{++} in classical HII regions. No such strong He^{++} edge exists in the atmospheres of the central stars of planetary nebulae, Wolf-Rayet stars (massive stars with tremendous wind-driven mass loss), or in Active Galactic Nuclei, and many of these objects show strong HeII recombination lines in their spectra as a result.

There are three regimes of stellar temperature that illustrate the interplay between H and He ionization in determining the detailed structure of HII regions.

"Cool" Stars ($T_* < 40000$ K):

These are stars later than about O6. These stars have many photons in the regime $h\nu_1 \leq h\nu \leq h\nu_2$ capable of ionizing H^0, but few with $h\nu \geq h\nu_2$ capable of ionizing He^0. He^0

will thus absorb most of the photons with $h\nu \geq h\nu_2$, resulting in a central He^+ zone surrounded by an H^+ zone with He^0 mixed in. The fraction of photons with $h\nu \geq h\nu_2$ absorbed by H is

$$y \approx \frac{n_{H^0} a_{\nu_2}(H^0)}{n_{H^0} a_{\nu_2}(H^0) + n_{He^0} a_{\nu_2}(He^0)}$$

and the fraction absorbed by He^0 is *(1–y)*.

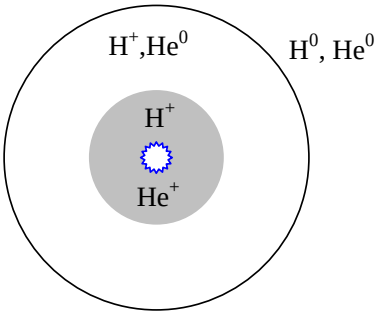


Figure III-3: Schematic of a nebula surrounding a cool ($T_* < 40000$ K) star.

Recombination of He^0 can result in the emission of photons with $h\nu \geq h\nu_1$ that can ionize H^0, so some H^+ will be mixed in with the He^+. Of these recombinations, ~3/4 are into triplet states of He^0 and the remaining ~1/4 are into singlet states. Since collisions with electrons can modify the mix of triplet and singlet states in He^0, the contribution to the H^0 ionizing radiation field is density-dependent with a critical density of $\sim 4\times10^4$ cm^{-3} (see AGN^2, section 2.4 for a detailed discussion).

Hot Stars (T_*=40,000 – 100,000 K):

These are stars earlier than O6 with a strong absorption edge in the stellar atmospheres above the He^+ ionization threshold (54.4eV). In these nebulae, both He^0 and H^0 are ionized throughout the volume of the nebula.

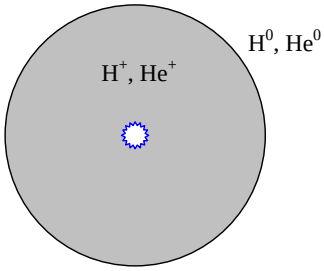


Figure III-4: Schematic of a nebula surrounding a hot star ($40{,}000\text{K} < T_* < 10^5$ K).

Near the edge of the ionized zone, the photons with the smallest cross-sections for photoionization also have the highest energies, thus higher energy photons have a greater mean-free path. Recall that for an H^+ Strömgren Sphere the width of the partially ionized transition zone at the edge was of order the mean-free path for ionizing photons. This means that the He^+ zone boundary is slightly fuzzier than the H^+ zone boundary because the radiation field at the edge is slightly hardened by radiative transfer through the nebula (see AGN^2, Fig 2.4).

Very Hot Stars ($T_*>10^5$ K):

These include the central stars of planetary nebulae and Wolf-Rayet stars. In this case an inner He^{++} zone forms immediately around the ionizing star. Some He^{++} recombinations emit photons capable of ionizing H^0, so that H^+ is mixed in with He^{++}.

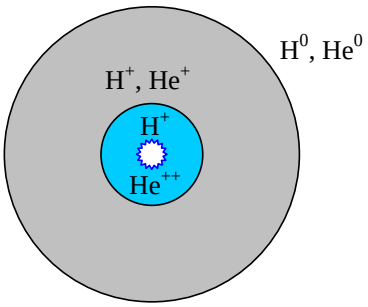


Figure III-5: Schematic of an $He^{++}/He^{+}/H^{+}$ nebula around a very hot star.

The effect of adding Helium to a pure Hydrogen nebula is a tight coupling between ionization and recombination from Helium and Hydrogen. That some Helium recombination photons can ionize Hydrogen means we must solve for coupled ionization equilibrium equations for both elements to derive the full structure of a nebula. This is best done numerically by photoionization equilibrium codes like CLOUDY.

The Effects of Metals

The important elements ("metals") to consider next are those species with abundances of $(X/H)\approx 10^{-3}$–10^{-4} by number, specifically C, N, O, S, Ne, Fe, and Si.

For any two successive stages of ionization (r to r+1) of element X, the ionization equilibrium equation is

$$n(X^{r})\int_{\nu_r}^{\infty}\frac{4\pi J_{\nu}}{h\nu}a_{\nu}(X^{r})d\nu = n(X^{r+1})n_e\alpha_G(X^{r},T)$$

where α_G is the recombination coefficient for the species. Each of these equations, together with the densities of all of the ionic species of X present:

$$n(X^{0})+n(X^{+})+n(X^{2+})+\cdots+n(X^{+n}) = n(X)$$

Assuming an overall abundance of element X, in principle you can completely determine the ionization equilibrium at each point in the nebula.

Metals have low abundances relative to Hydrogen and Helium ($X^r/H<<1$), so absorption by the different ionization stages of the metals does not significantly modify the radiation field of low-density nebulae. This is unlike what happens in dense stellar atmospheres or interiors where metals are important and sometimes dominant sources of opacity. Similarly, the low relative abundances of the metals mean that H is the primary source of electrons in the nebula, followed by electrons from He ionization. In typical HII regions, therefore, the electron density, n_e, is primarily coupled to the H+He density with negligible contributions from the metals.

There are, however, cases where radiative transfer in specific resonance lines could be important, especially at higher densities. One area of concern in detailed nebular modeling is the degree to which our neglect of the opacity of the metal species is leading to systematic errors in the model predictions. This may be an issue when one considered HII regions deeply embedded within dense molecular clouds, or at the boundaries of molecular clouds where high densities ($\sim 10^{4–6}$ cm^{-3}) prevail.

Ionization Cross-Sections

The relative gain in simplicity derived from being able to neglecting radiation field coupling is offset by the very complex electronic structure of the metal ions. Metal ions have many different thresholds

for photoionization, depending on which of the many electrons present is being kicked out of the atom. This results in the ionization cross-sections having "edges".

For example, consider the ionization of neutral oxygen, O^0. There are several terms of similar energy in the ground state, each of which contributes to the ionization cross-section:

$$\begin{aligned} O^0(2p^4\,{}^3P) + h\nu &\rightarrow O^+(2p^3\,{}^4S) + ks \text{ or } + kd \quad \text{for } h\nu \geq 13.6\,\text{eV} \\ &\rightarrow O^+(2p^3\,{}^2D) + ks \text{ or } + kd \quad \text{for } h\nu \geq 16.9\,\text{eV} \\ &\rightarrow O^+(2p^3\,{}^2P) + ks \text{ or } + kd \quad \text{for } h\nu \geq 18.6\,\text{eV} \end{aligned}$$

Schematically:

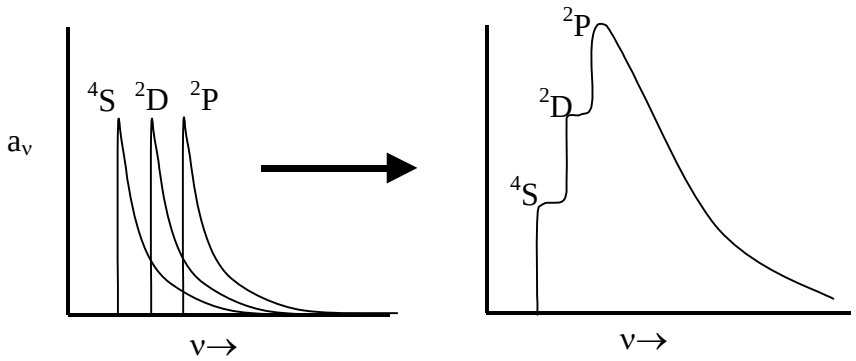


Figure III-6: Crude schematic of how individual ionization edges "add" to give the total photoionization cross-section for O^0, following Osterbrock (AGN^2).

An additional complicating factor is that many metal ions have complex photoionization resonance structures that can lead to dramatic changes in a_ν over small ranges of $\Delta\nu$ (e.g., autoionizing resonances). The treatment of metals requires detailed numerical calculations, embodied in photoionization equilibrium programs like CLOUDY.

Charge-Exchange Reactions

A final process that affects the coupling between the various ionic components of the gas in ionized nebulae is charge-exchange reactions. A particularly important reaction is due to collisions between O and H:

$$O^0(^3P) + H^+ \leftrightarrow O^+(^4S) + H^0(^2S)$$

This reaction is important near the outer ionization boundaries of optically thick nebulae (that is, most real ionization-bounded nebulae). The ionization potential of O^0 in the ^3P ground state is 13.62eV, while that of H^0 is 13.59eV, a difference of only 0.03eV, so this reaction occurs in near resonance, enhancing its efficiency. At temperatures large compared to the difference in the ionization potentials of the two species, an equilibrium is established in which the ratios of the ionic fractions of O and H are driven towards the ratios of their statistical weights:

$$\begin{aligned} \frac{n(O^0)}{n(O^+)} &= \frac{g(H^+)}{g(H^0)} \times \frac{g(O^0)}{g(O^+)} \times \frac{n(H^0)}{n_p} \\ &= \frac{9}{8}\frac{n(H^0)}{n_p} \end{aligned}$$

The result is that the H^+ and O^+ ionization boundaries of the nebula are coupled, and the O^+/O^0 transition region boundary steepens until it nearly matches that of the H^+/H^0 boundary. This is shown schematically in Figure III-7.

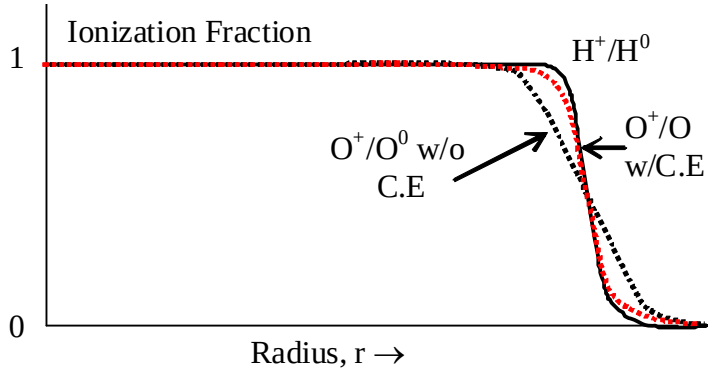


Figure III-7: Effect on O^+/O^0 ionization fraction (dotted curve) with and without charge exchange. With charge exchange, O^+/O^0 tracks H^+/H^0 closely (solid curve), but the edge is fuzzier without charge exchange.

Other charge exchange reactions are possible where there are close coincidences of ionization potentials for particular levels, and must be factored into the coupled equations describing the ionization equilibrium. This is why models of ionized nebulae (like CLOUDY) are large, fairly complex codes. A big part of getting the models "right" depends on making sure all of the relevant input physics is treated.

Dust

Dust grains introduce a number of important effects on the ionization equilibrium and structure of nebulae. In particular, dust strongly absorbs UV continuum and line photons, contributing to the optical depth that goes into the ionization equilibrium equations we have been considering:

$$\frac{d\tau_\nu}{dr} = n_{H^0} a_\nu (H^0) + n_{dust} \sigma_\nu (dust) + \cdots$$

When a dust grain absorbs a UV photon, the grain is heated to T≥50 K and then re-radiates the absorbed energy as Far-IR thermal continuum photons that escape from the nebula. In effect, ionizing photons are "destroyed" by dust grains, substantially modifying the ionizing radiation field (both stellar and diffuse), and changing the ionization balance and structure of the nebula. At the very least, the effect of mixing in dust *decreases* the size of the H^+ zone.

Other effects include absorption and subsequent re-radiation as Far-IR thermal continuum of the Lyman-series photons emitted by Hydrogen, and absorption of resonance-line photons capable of ionizing H that are emitted by other species (e.g., Helium). This modifies the recombination side of the photoionization-equilibrium equations, and has an impact upon the emergent recombination-line spectrum, as we'll see later in this chapter.

A counterbalancing effect is that the ionizing radiation field, especially in the presence of very hot stars, is sufficient to photoevaporate dust grains. The presence of dust is thus two-edged: dust can alter the ionization and recombination balance in a nebula, and the conditions can be hostile to the survival of the grains themselves. Sorting out these effects has been observationally very difficult, and as a consequence it is very difficult to simply "insert" dust into detailed photo-ionization models *a priori*, and predict clear effects that may be tested observationally.

A final, often overlooked effect of dust is that some elements are strongly depleted from the gas phase and locked into grains, changing the mix of metals in the gas-phase of the nebula. Similarly, photoevaporation of dust grains in a nebula might return elements to the gas-phase, enhancing the abundances. Since metals are the primary coolants in nebulae, how one treats dust depletion or destruction has potentially important implications for the results of model nebula calculations. The question of what abundances to use in nebular models, and the role of dust in determining those abundances, is a matter of considerable debate in the current research literature.

III-2 The Thermal Structure of Ionized Gas Regions

The thermal energy balance in a static ionized nebula is governed by the interplay between photoionization heating and cooling by recombination and other radiative losses, primarily in the form of line emission. The "nebular temperature" that we will speak of is the kinetic temperature of the free electrons, which electrons are quickly thermalized into a Maxwellian velocity distribution by electron-electron collisions.

Photoionization Heating

The principal source of heating in an ionized region is photoionization. The energy injected into the electron plasma by ionization is:

$$\tfrac{1}{2} m_e \mathrm{v}_e^2 = h(\nu - \nu_0)$$

where $h\nu_0$ is the **Ionization Potential** of the atom (usually H^0). The net photoionization heating rate (energy input/volume/sec) for hydrogen is $G(H^0)$:

$$G(H^0) = n_{H^0} \int_{\nu_0}^{\infty} \frac{4\pi J_\nu}{h\nu} h(\nu - \nu_0) a_\nu(H^0) d\nu$$

The three terms inside the integral are the number of ionizing photons from the star, the energy injected per photoionization, and the photoionization cross-section. The condition of photoionization equilibrium for a pure H nebula is:

$$n_{H^0} \int_{\nu_0}^{\infty} \frac{4\pi J_\nu}{h\nu} a_\nu(H^0) d\nu = n_e n_p \alpha_A(H^0, T)$$

The net heating rate therefore becomes:

$$G(H^0) = n_e n_p \alpha_A(H^0, T) \frac{\int_{\nu_0}^{\infty} \frac{J_\nu}{h\nu} h(\nu - \nu_0) a_\nu(H^0) d\nu}{\int_{\nu_0}^{\infty} \frac{J_\nu}{h\nu} a_\nu(H^0) d\nu}$$

The ratio of the integrals is the mean thermal energy of the photoelectrons, hence

$$G(H^0) = \tfrac{3}{2} k T_{init} \times n_e n_p \alpha_A(H^0, T)$$

T_{init} is the initial electron temperature. If the central source is a star with effective temperature T_*, then assuming that the stellar spectrum is approximately a blackbody, $T_{init} \approx T_*$. This gives us a reasonable limit on the initial temperature of the nebula.

Cooling

There are three principal sources of cooling in a normal HII region: Recombination Cooling, Free-Free continuum emission (bremsstrahlung), and line emission from collisionally excited ions.

Recombination Cooling:

The energy removed from the thermal electron plasma when electrons recombine with protons to form H^0 is:

$$L_R(H) = n_e n_p kT \beta_A(H^0, T)$$

Here, $\beta_A(H^0,T)$ is the energy-averaged recombination coefficient:

$$\beta_A(H^0,T) = \sum_{n=1}^{\infty}\sum_{L=0}^{n-1}\beta_{nL}(H^0,T)$$

and

$$\beta_{nL}(H^0,T) = \frac{1}{kT}\int_0^{\infty} \mathrm{v}\sigma_{nL}(H^0,T)\tfrac{1}{2}m_e\mathrm{v}^2 f(\mathrm{v})d\mathrm{v}$$

The recombination cross-section $\sigma_{nL}(H^0,T)$ depends on the relative velocities of the electrons and protons:

$$\sigma_{nL}(H^0,T) \propto \mathrm{v}^{-2}$$

Thus the recapture of slower electrons is favored over that of faster electrons. This has interesting implications for the nebular temperature. If recombination were the only cooling mechanism available to the gas, the resulting electron temperature, T_e, would actually be *slightly hotter* than T_{init} after recombination cooling. This is because the slower electrons are preferentially recombined out of the free electron plasma, skewing the velocity distribution of the remaining free electrons towards higher energies and hence higher temperatures. In effect, the temperature would be hotter than the effective temperature of the ionizing star!

In the OTS approximation, emission and absorption of photons resulting from recombinations directly into the $1s^2S$ ground state cancel out, so that

$$G_{OTS} = n_e n_p \alpha_B(H^0,T)\tfrac{3}{2}kT_{init}$$

$$L_{OTS} = n_e n_p kT\beta_B(H^0,T);$$

$$\beta_B(H^0,T) = \sum_{n=2}^{\infty}\beta_n(H^0,T)$$

Adding He^0 into the nebula is straightforward:

$$G(He^0) = n_e n_{He^+}\alpha_A(He^0,T)\frac{\int_{\nu_2}^{\infty}\frac{J_\nu}{h\nu}h(\nu-\nu_2)a_\nu(He^0)d\nu}{\int_{\nu_2}^{\infty}\frac{J_\nu}{h\nu}a_\nu(He^0)d\nu}$$

and

$$L_R = L_R(H) + L_R(He)$$

$$L_R(He) = n_e n_{He^+} kT\beta_A(He^0,T)$$

The contributions from other recombination cooling species (e.g., He^+ in a hot PNe) are included similarly.

Free-Free Cooling:

Thermal electrons can scatter off ions in the plasma and emit free-free (bremsstrahlung) radiation, contributing to the nebular cooling. The free-free cooling rate for an ion with nuclear charge Z is given by

$$L_{ff}(Z) = 4\pi j_{ff} = \frac{32\pi e^6 Z^2}{3^{3/2}hm_e c^3}\left(\frac{2\pi kT}{m_e}\right)^{1/2} n_e n_+ g_{ff}$$

Here n_+ is the number density of ions with nuclear charge Z in the gas. For an H+He nebula with no He^{++}, this is

$$n_+ = n_p + n_{He^+}$$

Evaluating all of the physical constants, the free-free cooling rate becomes:

$$L_{ff}(Z) = 1.42 \times 10^{-27} Z^2 T^{1/2} g_{ff} n_e n_+$$

g_{ff} is the free-free Gaunt Factor, which is a slowly varying function of density and temperature. For UV-to-NIR wavelengths and typical nebular conditions, g_{ff} has values of $1.0 < g_{ff} < 1.5$.

Overall, free-free cooling is fairly inefficient, but it does contribute enough to make the resulting nebular temperature cooler than T_{init} by a small amount when free-free emission and recombination are the only sources of cooling.

Collisionally-excited line emission

Metal ions like O^+, O^{++}, N^+, and others, while relatively underabundant compared to H or He, turn out to be the most important coolants in nebulae. Collisional excitation energies in the ground state fine structure levels of these ions have typical excitation potentials, χ, of a few eV. The thermal energies of the electrons, kT_e, are also of order a few eV for typical nebular temperatures of 10^4 K. This makes electron-ion impact excitation of the metal ions very efficient. By contrast, the first excited levels of H and He are at ~10eV above the ground state, so that collisional excitation of these elements is very inefficient at typical nebular densities and temperatures.

Proton-Ion and Ion-Ion impact excitation are inefficient because the Coulomb repulsion between the ions is so large. However, when ions become neutral, some important ion-ion collisional processes do occur (e.g., charge-exchange reactions between O^0 and H) that can contribute to the cooling.

Electron-Ion impact excitation of metal ions followed by the radiation of line emission is the dominant cooling mechanism in ionized nebulae with metallicities greater than a few percent of the solar value. This cooling mechanism is so efficient that it drives the electron temperature to well below the stellar photosphere temperature. For example, a nebula with roughly solar proportions of metals around a T_*=32,000 K star will a mean electron temperature of $T_e \approx 7000$K. A nebula with a lower metallicity will have a somewhat higher temperature for the same exciting star (fewer metals means fewer collisional coolants, and hence a higher equilibrium T_e). The abundance of metals relative to H in a nebula thus plays a crucial role in determining the thermal structure of an HII region. We will now consider the physics of this cooling mechanism in some detail.

Collisional Excitation

The cross-section for collisional excitation of an ion from a lower level to an upper level is $\sigma_{ul}(\upsilon)$ for an electron with velocity υ. Below the excitation threshold $h\nu_{ul}$ the excitation cross-section is identically zero. Just above the threshold, Coulomb focusing gives

$$\sigma_{lu}(\mathrm{v}) \propto \mathrm{v}^{-2}$$

The **collisional excitation cross-section** is usually expressed in terms of a dimensionless **Collision Strength, Ω**:

$$\sigma_{lu}(\mathrm{v}_l) = \frac{\pi\hbar^2}{m^2\mathrm{v}_l^2}\frac{\Omega(l,u)}{g_l};\ \text{for } \tfrac{1}{2}m\mathrm{v}_l^2 \geq h\nu_{ul}$$

[**Note:** Osterbrock in AGN[2] uses ω_1 instead of g_l for the statistical weight of the lower level, whereas I am endeavoring to adopt a consistent notation for statistical weights throughout these notes.]

Since $\sigma_{lu} \propto \mathrm{v}^{-2}$ near the threshold, Ω is approximately constant, or at least only a slowly varying function of v, away the threshold. Beware, however, as resonance effects can cause Ω to vary radically as one gets further away from the threshold energy.

The **collisional de-excitation cross section**, σ_{ul}, is related to σ_{lu} by a Milne Relation:

$$g_l \mathrm{v}_l^2 \sigma_{lu}(\mathrm{v}_l) = g_u \mathrm{v}_u^2 \sigma_{ul}(\mathrm{v}_u)$$

The velocities are related by energy conservation:

$$\tfrac{1}{2} m \mathrm{v}_l^2 = \tfrac{1}{2} m \mathrm{v}_u^2 + h\nu_{ul}$$

Collisional excitations are balanced by collisional de-excitations that produce electrons with the same velocity range, so that

$$n_e n_l \mathrm{v}_l \sigma_{lu}(\mathrm{v}_l) f(\mathrm{v}_l) d\mathrm{v}_l = n_e n_u \mathrm{v}_u \sigma_{ul}(\mathrm{v}_u) f(\mathrm{v}_u) d\mathrm{v}_u$$

In thermal equilibrium, the relative level populations of the upper and lower states would be those given by the Boltzmann equation:

$$\frac{n_u^*}{n_l^*} = \frac{g_u}{g_l} e^{-h\nu_{ul}/kT}$$

Where T is the kinetic temperature of the electrons. Combining all of these with the Milne Relation and the energy conservation constraint, we can write the collisional de-excitation cross-section in terms of the excitation collision strength, Ω_{lu}, as follows:

$$\sigma_{ul}(\mathrm{v}_u) = \frac{\pi \hbar^2}{m^2 \mathrm{v}_u^2} \frac{\Omega_{ul}}{g_u}$$

The Ω's are symmetric, $\Omega_{ul}=\Omega_{lu}$, so we only need to use Ω_{ul}.

The total volumetric collisional de-excitation rate is thus

$$\begin{aligned} n_e n_u q_{ul} &= n_e n_u \int_0^\infty \mathrm{v} \sigma_{ul}(\mathrm{v}) f(\mathrm{v}) d\mathrm{v} \\ &= n_e n_u \left(\frac{2\pi}{kT}\right)^{1/2} \frac{\hbar^2}{m^{3/2}} \frac{\gamma_{ul}(T)}{g_u} \end{aligned}$$

where γ_{ul}(T) is the **effective collision strength**[1], which is the collision strength Ω_{ul} averaged over a Maxwellian distribution for electrons with kinetic temperature T:

$$\gamma_{ul}(T) = \int_0^\infty \Omega_{ul}(E) e^{-E/kT} d(E/kT)$$

$$\text{where: } E = \tfrac{1}{2} m \mathrm{v}^2$$

Here the average has been conventionally rewritten as an integral in energy units. We can now solve for the q's:

$$q_{ul} = \frac{8.629 \times 10^{-6}}{T^{1/2}} \frac{\gamma_{ul}(T)}{g_u}$$

$$q_{lu} = \frac{g_u}{g_l} q_{ul} e^{-h\nu_{ul}/kT}$$

[1] An unfortunate but unavoidable notational issue: γ here is an effective collision strength, not the natural width we saw in the chapter on neutral hydrogen gas. Notational consistency across ISM subfields is virtually non-existent.

In general, the effective collision strength $\gamma_{ul}(T)$ is only a very weak function of temperature, and most of the temperature dependence of q_{ul} is in the $T^{-1/2}$ term. Tables of γ_{lu} computed for typical nebular temperatures of $T=10^4$ K are of particular use for HII regions and PNe[2]. However, it should be borne in mind that Ω_{ul} is not a smooth continuous function of energy: detailed quantum mechanical calculations show that there is considerable resonance structure, some of it very strong. Work by Anil Pradhan and his collaborators have found that these resonances can be numerous and potentially very strong, in some cases introducing order-of-magnitude differences with older mean cross-sections found in the literature.

Given these qualifications, it should be clear to the reader that while tabulated γ's are good starting points for most nebular work, they are subject to substantial revision, and may not be useful for other conditions beyond those found in nebulae (e.g., in high-temperature or high-density plasmas).

Collisional Cooling

If an electron collides with an ion and excites it, and then another electron comes along and de-excites the ion back into its original ground state before it can emit a photon, there is no net change in the energy content of the free-electron plasma. However, if the electron can radiatively de-excite before the next collision, it emits a photon that escapes from the nebula, leading to a net loss of energy from the free-electron plasma (i.e., you have converted electron kinetic energy into photons that escape).

Since the low-lying levels of most metal ions in nebulae arise from the same electronic configuration as the ground state, these lines are "forbidden" by the usual dipole selection rules. This is primarily true of visible and IR lines, but in the UV some collisional excitations populate levels that emit permitted lines. Since the nebula is optically thin in these lines, the photons escape from the nebula and carry off energy.

The radiant energy carried off by a given optically thin emission-line is

$$4\pi j_{ul} = n_u A_{ul} h\nu_{ul}$$

An estimate of the total cooling by collisionally excited line emission requires computation of the upper level populations for all low-lying levels for each ion present at a given location in the nebula. This is usually done numerically by nebular model codes, but we can gain some insight with some relatively simple arguments.

Consider a simple 2-level atom with ground state 1 and excited state 2. Statistical equilibrium within this atom is defined as the balance between the rate of collisional excitation out of the ground state on the one hand, and the rate of collisional and radiative de-excitations out of the excited state:

$$n_l n_e q_{lu} = n_u n_e q_{ul} + n_u A_{ul}$$

Because the emission lines are generally from forbidden transitions, we can ignore any contributions due to radiative excitation by absorbing photons from the radiation field (i.e., we don't need to consider radiative transfer. [Beware: this may not be strictly true of you are in high-density situations with permitted transitions instead of the usual low-density conditions and forbidden transitions]. In this case, solving for the relative level populations in statistical equilibrium gives:

[2] Osterbrock uses $\Omega(1,2)$ instead of $\gamma_{lu}(T)$ in his discussion in AGN[2]. For typical nebular conditions $\gamma_{lu}(T)\approx\Omega(1,2)$ to a reasonable approximation for some transitions, and one often sees the two used interchangeably literature. In general, however, they are not strictly identical, and some care should be taken when doing detailed calculations.

$$\frac{n_u}{n_l} = \frac{n_e q_{lu}}{A_{ul}}\left[1+\frac{n_e q_{ul}}{A_{ul}}\right]^{-1} = \frac{q_{lu}}{q_{ul}}\left[1+\left(\frac{A_{ul}}{n_e q_{ul}}\right)\right]^{-1}$$

There are two limiting cases of interest:

Low-Density Limit:

In this case, $n_e \to 0$, so that

$$\frac{n_u}{n_l} \to n_e \frac{q_{ul}}{A_{ul}} \propto n_e$$

This means that nearly all collisions are followed by radiative de-excitation and collisional de-excitation is relatively unimportant.

High-Density Limit:

In this case, $n_e \to \infty$, so that

$$\frac{n_u}{n_l} \to \frac{q_{lu}}{q_{ul}} = \frac{g_u}{g_l} e^{-h\nu_{ul}/kT}$$

This is just the thermal equilibrium value. At high densities collisional de-excitation completely dominates radiative de-excitation, and we say that the level populations have "thermalized".

In general, the cooling due to collisionally excited line emission from our 2-level atom is:

$$4\pi j_{ul} = n_u A_{ul} h\nu_{ul} = n_l \frac{g_u}{g_l} e^{-h\nu_{ul}/kT}\left[1+\left(\frac{A_{ul}}{n_e q_{ul}}\right)\right]^{-1} A_{ul} h\nu_{ul}$$

In the low-density limit, radiative de-excitation dominates, so that:

$$4\pi j_{ul} \to n_l n_e q_{lu} h\nu_{ul} \propto n_e^2$$

In the high-density limit, where collisions dominate and the levels thermalize:

$$4\pi j_{ul} \to n_u A_{ul} h\nu_{ul} \propto n_e$$

In both cases, we have recast n_1 in terms of $(n_1/n_e)\, n_e$, where (n_1/n_e) is proportional to the ionic abundance relative to H (since most of the electrons come from H^0). The dividing line between these limits is the **Critical Density** for the line, n_{crit}:

$$n_{crit} = \frac{A_{ul}}{q_{ul}}$$

As the density increases, the strength of a given collisionally excited line will increase as n^2 until it reaches the critical density, beyond which it will then only increase linearly with n. The overall effect is to "collisionally suppress" emission from the line at high densities ("high" now meaning greater than the critical density). The maximum cooling provided by a given collisionally-excited line therefore reaches a maximum when the density is at or near the critical density for that line, and dropping off in effectiveness rapidly above the critical density.

The simple 2-level atom description is a reasonable approximation for p^1 or p^5 ions like C^+, N^{++}, Si^+, and so forth, but ions with p^2, p^3, and p^4 ground-state configurations require at least 5 levels for a minimum description. The 5-level atom approximation works reasonably well for ions like O^{++}, N^+, and other common ionic species that are observed in gaseous nebulae.

The derivation of collisional cooling for an N-level atom is similar to the 2-level treatment above, except now we need to solve five coupled equations to get the five (unknown) level populations. For the j^{th} level, the condition of statistical equilibrium is:

$$\sum_{k \neq j} n_k n_e q_{kj} + \sum_{k>j} n_k A_{kj} = \sum_{k \neq j} n_j n_e q_{jk} + \sum_{k<j} n_j A_{jk}$$

Here all collisional excitations/de-excitations and radiative de-excitations that lead to population of the j^{th} state (left) are balanced against all collisional excitations/de-excitations and radiative de-excitations that de-populate the j^{th} state (right). In this case, the collisional line cooling from a N-level atom or ion is given by the sum over all the emission-lines arising from that atom

$$L_C = \sum_{j=1}^{N} n_j \sum_{k<j} A_{jk} h\nu_{jk}$$

Numerically, the procedure is to first solve the coupled statistical equilibrium equations for the level populations. These are then plugged into the cooling equation along with the transition probabilities (A) and line energies (hν) to derive the total cooling at a given density and temperature at a location within the nebula.

The critical density for a particular excited state, $n_{crit}(j)$ becomes, in general:

$$n_{crit}(j) = \frac{\sum_{k<j} A_{jk}}{\sum_{k \neq j} q_{jk}}$$

The sums are taken over all levels that interact with the j^{th} state. These are usually tabulated for different excited levels within an atom or ion of interest.

As with the 2-level atom, when $n_e < n_{crit}$ the cooling contribution from that level scales as the density squared. Similarly, above the critical density the cooling scales linearly with the density, leading to a net loss in cooling efficiency at high densities ("high" meaning substantially above n_{crit}). In general, a collisionally excited line will make its greatest contribution to cooling at or near its critical density.

Thermal Equilibrium

Bringing together all of the heating and cooling physics we have been discussing, the resulting thermal equilibrium in a nebula is given by

$$G = L_R + L_{ff} + L_C$$

Evaluating this equation entails solving the ionization equations at each point in the nebula to find out what ions are present, and then using these along with model atom calculations to evaluate the thermal balance at that location. The condition of thermal equilibrium is often written in the form:

$$(G - L_R) = L_{ff} + L_C$$

$(G–L_R)$ is the "Net Effective Heating Rate", with recombination losses already factored in. An example of a thermal balance calculation is shown in Figure III-8.

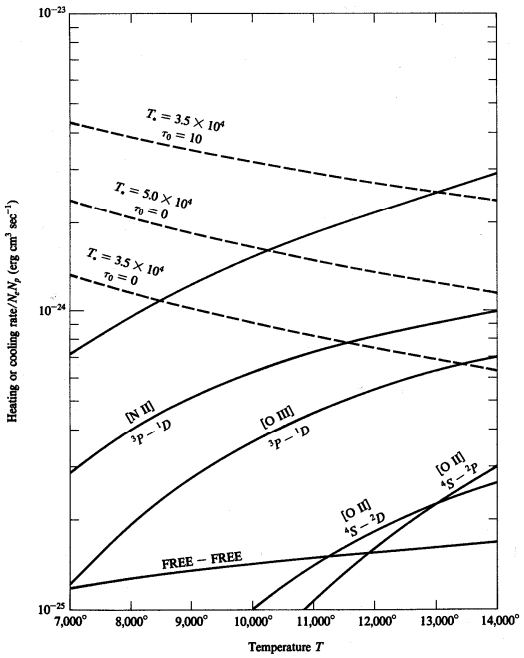


Figure III-8: Net effective heating rates ($G-L_R$) for various stellar spectra (dashed lines), and total radiative cooling rate ($L_{ff}+L_C$; upper solid line) for various nebular temperatures and assuming a density of $n_e=10^4$ cm^{-3}. The contributions from different metal-ion lines and free-free is shown below the cooling curve. The equilibrium temperature will be where the heating and cooling curves cross. [Reproduced from Osterbrock, AGN^2, Figure 3.3].

In the low-density limit, since each term is proportional to n_e and n_x, the resulting temperature tends to be independent of density, but strongly dependent on the relative abundances of the elements, n_X/n_H.

As the density increases, collisional de-excitation becomes important, and depending on the individual critical densities of the various cooling lines, various ionic coolants will decrease in effectiveness, increasing the equilibrium temperature.

Overall, the principal factors that influence the nebular temperature are:

1. The density through critical density effects suppressing different cooling lines
2. The ionizing radiation field determines the input photoelectron energies and the mix of ionic states present.
3. The elemental abundances determine the amounts of the various metal-line coolants present. In general, lower metallicity means hotter nebulae.

Dust can modify the radiation field and cause other second-order effects. For example, dust grains deplete the gas-phase abundances of important cooling ions, reducing the overall cooling efficiency for a given initial mix of metals. Also, since the absorption cross-section for dust increases with higher energies, it can act to lower the mean energy of the input photoelectrons. In short, dust makes things complicated, and modern photoionization models of nebulae are only just getting around to evaluating the effects of the dust content on the nebular conditions.

III-3 The Spectra of Ionized Hydrogen Regions

Because emission-line spectra can be detected relatively easily even at faint levels (the bright emission lines contrast with the fainter continuum level), we can measure the spectra of ionized gas nebulae in other galaxies. The emergent spectrum of HII regions gives us our principal probe of the physical conditions and gas-phase abundances in other galaxies.

The Recombination Spectrum

The recombination process in nebulae results in both line and continuum emission. In nebulae, we observe recombination lines of HI, HeI, and HeII at UV, visible, IR, and radio wavelengths. We also observe nebular continuum at all wavelengths, in the form of free-free, free-bound, and bound-bound (e.g., 2-photon) emission.

HI Recombination Lines

Recombination is the recapture of an electron into one of the excited states of the atom, followed by a "cascade" of radiative transitions into the ground state. The population of any given excited state has contributions from both direct recombination into that state and cascades from recombination into all higher levels. For typical nebular conditions, collisions are unimportant, and the level populations depend on the radiative transition probabilities for each cascade transition. As a consequence, the level populations that are setup by recombination are very far from the LTE populations. For a detailed treatment of the recombination process, particularly the calculation of cascade matrixes, see Osterbrock AGN[2].

There are two cases of interest in computing the recombination lines of H^0: optically thin and optically thick, called **Case A** and **Case B**, respectively.

Case A Recombination:

Case A assumes that the nebula is optically thin in all of the HI resonance lines arising from the 1s ground state (Lyα, Lyβ, etc.). Case A, however, is rarely true for real nebulae as the optical depth in the Lyman resonance lines is very high. For example,

$$\tau_{Ly\alpha} \approx 10^4 \tau_{912}$$

$$\tau_{Ly\beta} \approx 10^3 \tau_{912}$$

where τ_{912} is the optical depth at the HI ionization threshold (λ912Å or $h\nu$=13.6eV). Thus, photon-for-photon, the optical depth for UV resonance absorption is much larger than the ionization optical depth. This means that we cannot ignore contributions to the excited level populations by UV resonance absorption out of the 1s ground state, leading us to…

Case B Recombination:

In the case where the nebula is optically thick to UV Lyman resonance line absorption, each Lyman photon absorbed is quickly re-emitted ("resonantly scattered") at the same wavelength. In each subsequent scattering, however, there is a finite probability that the excited electron will choose a different path back to the ground state, with the result that the Lyman photon gets converted into a number of lower series line photons.

For example, if an HI atom absorbs a Lyβ photon, the electron is excited from the 1s state into the 3p state. The subsequent de-excitation has three possible paths back to the 1s state:

88% of the time the electron goes directly back to the 1s state and emits an Lyβ photon (resonant scattering)

12% of the time the electron makes a 3p–2s transition, resulting in the emission of a Balmer series Hα photon, followed (later) by 2-photon continuum emission from a forbidden 2s–1s transition back to the ground state.

The net effect is that after ~9 scatterings, each Lyβ photon is "degraded" into Hα plus 2-photon continuum photons. This enhances the total emissivity in the Hα line over the predictions of Case A recombination. A similar enhancement occurs for other higher-series HI lines (Balmer, Paschen, Bracket series, etc.) due to resonant absorption of higher-order Lyman series photons.

To compute the Case B emissivities, the downward transitions into the 1s ground state are omitted from the equations of statistical equilibrium. In practical terms, this amounts to replacing the recombination coefficients, $\alpha_{nn'}$, with *effective recombination coefficients*, such that for a given line:

$$\frac{4\pi j_{nn'}}{h\nu_{nn'}} = n_e n_p \alpha^{eff}_{nn'}(T) = \sum_{L=0}^{n-1} \sum_{L'=L\pm1} n_{nL} A_{nL,n'L'}$$

It is traditional to compute the emissivity of Hβ, the n=4-2 transition in the Balmer series, and then tabulate the strengths of the other HI recombination lines in terms of their emissivity relative to Hβ. As such, only the effective recombination coefficient of Hβ is often tabulated, and all others may be inferred from the relative line strengths (this is what you will find, for example, in the extensive recombination calculations of Hummer & Story 1987, MNRAS, 224, 801).

Nebular conditions most often favor Case B over Case A recombination.

Case A vs. Case B:

Predicted line emissivities for Case B are larger than the Case A predictions because the energy that would have come out as Lyman series emission-lines is now redistributed among the other series.

Line ratios within a given series (e.g., Balmer line ratios like Hα/Hβ) don't change very much between Case A and Case B. Between series, however, the effect is more pronounced. For example, for $T_e=10^4$K and low densities ($N_e<100$ cm^{-3}):

Case A	Case B
Hα/Hβ=2.86	Hα/Hβ=2.87
Paβ/Hβ=0.216	Paβ/Hβ=0.165

It is important to recognize that Case A and Case B represent extremes.

> **Case A:** Optically-thin nebula, all Lyman-series photons escape the nebulae.
>
> **Case B:** Optically-thick nebula, all Lyman-series photons are absorbed in the nebula and re-emitted as higher-order lines.

While observed conditions tend to favor Case B, if intermediate conditions prevail a detailed radiative transfer treatment is required. For a good overview, see AGN^2, section 4.5.

Sometimes the ratios of HI emission lines are observed to depart significantly from the Case B predictions (after correction for dust extinction). The usual sense of the departure is for the observed line ratios to be driven their Case B values towards their Case A values. These departures have been seen in everything from Active Galactic Nuclei to compact HII regions. Possible physical causes of these departures include:

1. High densities ($N_e=10^{8–12}$ cm^{-3}), leading to collisional excitation of HI (e.g., collisional excitation of Lyα) and altering the basic assumptions of the collisionless Case A/B recombination computations.
2. Optical depth effects, for example if the nebula starts becoming optically thick to Balmer series lines, or if the nebula is clumpy and Lyman-series photons escape in some regions but not in others. Balmer continuum and (perhaps) Balmer line opacity can be important, for example, in the broad-line regions of AGN.
3. Resonance fluorescence effects in which Lyman-series photons are resonantly absorbed by other species in the nebula (either due to wavelength coincidence with resonance lines in other species, or due to bulk motions of the gas Doppler shifting the resonance lines to the wavelengths of the Lyman photons). These Lyman-series photons become unavailable to be "degraded" into higher series photons, resulting in lower emissivities than predicted by Case B (thus moving them back towards Case A values).
4. Dust can similarly destroy Lyman-series photons before repeated resonant scattering with HI can degrade them into higher-series photons, also driving the line ratios away from the Case B values towards Case A.

Recombination Lines of He and Metals

Hydrogenic ions are those, like He^{++}, C^{6+}, etc. which recombine to have 1 electron and a nucleus with nuclear charge Z. The transition probabilities, line energies, line emissivities, and transition probabilities are all related to the corresponding values for Hydrogen via powers of Z as follows:

$$\alpha_{nL}(Z,T) = Z\alpha_{nL}(H^0, T/Z)$$
$$h\nu_{nn'}(Z) = Z^2 h\nu_{nn'}(H^0)$$
$$j_{nn'}(Z,T) = Z^3 j_{nn'}(H^0, T/Z^2)$$
$$A_{nn'}(Z) = Z^4 A_{nn'}(H^0)$$

This also means that the branching ratios are independent of Z.

The most common Hydrogenic ion is He^{++}, which is seen primarily in hot planetary nebulae, AGN, and Wolf-Rayet stars. Taking into account the Z-dependencies of the line emissivities ($j_{nn'}$), the He^{++} line spectrum for a PNe with T=20,000K looks like that for H^+ for T=5000K. The difference is that the emissivities are 2^3=8 times larger and the lines are at bluer wavelengths by a factor of 1/4. Thus, while He/H≈0.1 in nebulae, in PNe and the hot phases of nova outbursts, the HeII λ4686Å emission line can be as bright as the Hβ λ4861Å line!

Hydrogenic ions are relatively easy to work with, but the others are not. For example, He^+ recombination is made complicated by its complex 2-electron structure. As such, HeI/HI line strengths tend to be comparable to the He/H≈0.1 ratio in regions in which most of the He is He^+ and most of the H is H^+. Still, one can make analogous Case A and Case B calculations of the HeI recombination lines taking into account the effects of optical depth in the UV resonance lines of HeI. In general one finds that

HeI triplet lines always follow Case B assumptions fairly closely, primarily because the $2^3S–1^1S$ is dipole forbidden and therefore quite rare.

HeI singlet lines do make it into the 1^1S state, and Case B is usually a good approximation. The optical depths for the UV resonance lines are lower than those of HI by about a factor of 10 (the He/H ratio), and there are further complications because HeI $n^1P–1^1S$ emission lines are energetic enough to

ionize H^0. Further complications come from collisional excitation that can modify the Case B assumptions at relatively modest densities (under ~10^4 cm^{-3}).

In practice, if making back-of-the-envelope calculations, you can use the Case B recombination tables for HeI (like those in Osterbrock) for predicting the relative strengths of HI and HeI lines in low-density HII regions and PNe ("low" means $n_e \leq 10^4$ cm^{-3}). For more sophisticated calculations, like deriving He/H abundances, a more sophisticated approach that includes collisional and other effects is required. This is in part why nebular He abundances are so contentious; a detailed treatment is required to get a correct answer.

In dense nebulae (e.g., AGN and the early phases of novae and some SNRs), the Case B assumptions are no longer adequate and more sophisticated treatments are required, usually in the form of full numerical modeling of the regions of interest (e.g., using CLOUDY).

Nebular Continuum

In addition to discrete line emission bound-bound-bound transitions, several processes also lead to the production of continuum emission in nebulae at visible and radio wavelengths. We will treat each of these in turn, since different physics come into play in the different wavelength regimes.

UV, Visible, and IR Nebular Continuum

The principal sources of continuum emission in nebulae are:

1. Free-Bound Emission (recombination continuum)
2. Free-Free Emission (bremsstrahlung)
3. 2-Photon HI Continuum Emission

All three are primarily associated with Hydrogen and Helium, with negligible contributions from metals.

Free-Bound Emission:

Continuum radiation is emitted when a free electron with velocity v recombines into an excited level of HI with principal quantum numbers $n \geq n_1$. The energy of the photon is given by energy conservation:

$$h\nu = \tfrac{1}{2} m\mathrm{v}^2 + X_n = \tfrac{1}{2} m\mathrm{v}^2 + \frac{h\nu_0}{n^2}$$

$$\text{and} \quad h\nu \geq \frac{h\nu_0}{n_1^2}$$

The emission coefficient of this continuum radiation is given by:

$$4\pi j_\nu^{f-b} = n_e n_p \sum_{n=1}^{\infty} \sum_{L=0}^{n-1} \mathrm{v}\sigma_{nL}(\mathrm{v})\, f(\mathrm{v}) h\nu \frac{d\mathrm{v}}{d\nu}$$

The recombination cross-sections, σ_{nL} can be derived from the photoionization cross-sections using the Milne Relations. The complete spectrum is computed piece-wise for each quantum state of HI, with the result that the emergent continuum spectrum has large discontinuities (edges) at the ionization energies of the different excited states.

Free-Free Emission:

Free-free emission results from electrons scattering off the protons in an encounter that does *not* result in recombination. We have seen this before in the discussion of the diffuse radiation field in nebulae:

$$4\pi j_\nu^{ff} = n_e n_p \frac{32Z^2e^4h}{3m^2c^3}\left(\frac{\pi h\nu_0}{3kT}\right)^{1/2} e^{-h\nu/kT} g_{ff}(T,Z,\nu)$$

g_{ff} is the Gaunt Factor, and represents the quantum mechanical corrections to classical bremsstrahlung emission.

The bound-free and free-free terms are often combined into a single emission coefficient for HI recombination continuum:

$$4\pi j_\nu(HI) = n_e n_p \gamma_\nu(H^0,T)$$

Similar terms can be derived for HeI and HeII:

$$4\pi j_\nu(HeI) = n_e n_{He^+} \gamma_\nu(He^0,T)$$

$$4\pi j_\nu(HeII) = n_e n_{He^{++}} \gamma_\nu(He^+,T)$$

In typical nebulae He/H is 0.1 by number. If the He in the nebula is mostly He^{++} (e.g., in the cores of a PNe), the contribution to the continuum from HeII will be comparable to the HI continuum (modulo different edges due to different ionization thresholds for the different levels). In nebulae where most of the He is He^+, the contribution to the continuum from HeI is about 10% that of HI (i.e., about the relative abundances), again with edges at different wavelengths.

2-Photon HI Continuum:

The excited 2s ^{2}S state can be populated by direct recombination, cascades following recombination to excited states, and resonant scattering of Lyman series photons in Case B recombination (e.g., Lyβ). The 2s ^{2}S→1s ^{2}S radiative transition is forbidden by the dipole selection rules, but it can proceed via a 2-photon emission process with transition probability

$$A_{2^2S,1^2S} = 8.23\,\mathrm{s}^{-1}$$

The energies of the two photons emitted must add up to the Lyα photon energy:

$$h\nu' + h\nu'' = h\nu_{12} \approx 10.2\,\mathrm{eV}$$

The probability distribution of the photon frequencies is symmetric about the mean energy, $0.5h\nu_{12}$, which corresponds to a wavelength of λ=2431Å. If the distribution is expressed in wavelength units, the distribution will not be symmetric.

The emission coefficient is for 2-photon continuum is:

$$4\pi j_\nu(2q) = n_{2^2S} A_{2^2S,1^2S} 2hyP(y)$$

$$y = h\nu / h\nu_{12}$$

Here P(y) is the normalized probability per decay that one photon will be emitted with an energy in the range $h\nu = [y,(y+dy)]h\nu_{12}$.

The emission coefficient can be written in terms of the electron and proton density by working out all of the contributions to the population of the 2^2S level from direct recombination and cascades (using effective recombination coefficients to account for Case B). In addition, because the 2s ^{1}S state is so long-lived compared to all other excited HI states (~0.1 seconds compared to ~10^{-8} seconds), it is necessary to take into account collisional de-excitation due to electron and proton collisions.

The result is written in terms of γ_ν as was done did for the other continuum sources:

$$4\pi j_\nu(2q) = n_e n_p \gamma_\nu(2q)$$

where:

$$\gamma_\nu(2q) = g_\nu \frac{\alpha^{eff}_{2^2S}(H^0,T)}{1+\left(\frac{n_e q^e_{2^2S,2^2P} + n_p q^p_{2^2S,2^2P}}{A_{2^2S,1^2S}}\right)}$$

The *q*'s are the collision rates for collisions with electrons (q^e) and protons (q^p) respectively. The effective recombination coefficient (α^{eff}) and the frequency dependence, g_ν, are usually computed numerically for different conditions (see Tables 4.10–4.12 in AGN[2]).

The Nebular Continuum:

The emergent nebular continuum is the sum of the free-free, bound-free, and 2-photon contributions, resulting in a complex continuous spectrum crossed by sharp edges. An example is reproduced in Figure III-9 (from Figure 4.1 of AGN[2]) for T=10^4K, showing γ_ν for HI, HeI, HeII bound-free and 2-photon continuum. The continuum emission coefficient for He^{+2} is about an order of magnitude larger than that for H^+ on average, but this is compensated for by the nearly order of magnitude lower abundance of He relative to H. In PNe or highly ionized nebulae (like Nova outbursts), the H and He nebular continuum contributions can be comparable in strength.

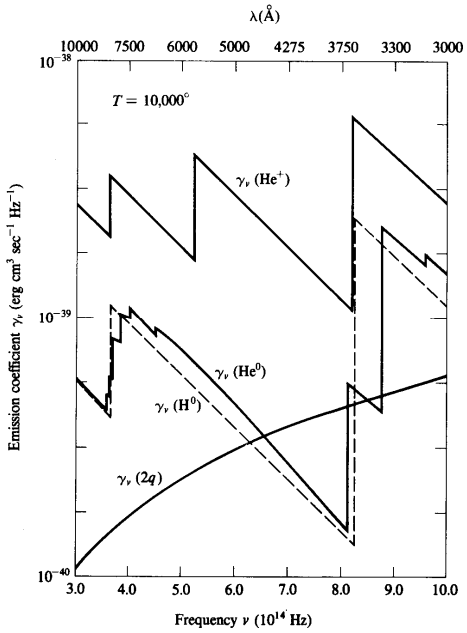


Figure III-9: Nebular continuum emission coefficients γ_ν for HI, HeI, HeII, and HI 2-photon continuum processes computed for low-densities at 10^4K (reproduced from Osterbrock AGN[2] Fig 4.1).

The edges that appear in the continuum spectra correspond to the $n\rightarrow\infty$ limits of the various line series (e.g., the H^+ continuum edge near λ=3640Å corresponds to the HI Balmer series limit). The reasons they appear as discontinuous edges are twofold:

At energies above an edge, captures into excited states have leftover energy that emerges in the continuum.

At energies below an edge, there are many possible excited states that can take up most of the electron's kinetic energy, and emission of energy into the continuum can be less effective by factors of ~10 or so.

Also notice that for typical nebular conditions (low density, T≈10^4 K), the 2-photon continuum can be the dominant contributor above the Balmer discontinuity at λ3640Å until ~5000Å.

Radio Continuum Emission

Nebulae emit a thermal radio continuum spectrum due to free-free emission. The emission coefficient is the same as at visible wavelengths, but hν<<kT and the free-free Gaunt Factor is no longer of order unity. Instead, at radio frequencies, the Gaunt Factor becomes:

$$g_{ff}(T,Z,\nu) = \frac{\sqrt{3}}{\pi}\left[\ln\left(\frac{8k^3T^3}{\pi^2Z^2e^4m\nu^2}\right)^{1/2} + \frac{5}{2}\gamma\right]$$

Here γ=Euler's constant (=0.577…). Numerically, this is written:

$$g_{ff}(T,Z,\nu) \approx \frac{\sqrt{3}}{\pi}\left[\ln\left(\frac{T^{3/2}}{Z\nu}\right) + 17.7\right]$$

For typical nebular conditions, $T=10^4$ K, $g_{ff}\approx 10$ at ν=1GHz for scattering off protons (Z=1). Compare this to $g_{ff}\approx 1-1.5$ for UV and visible wavelengths.

Free-free radiation is also *absorbed* at radio wavelengths. Using Kirchoff's Laws, the free-free absorption coefficient is:

$$\kappa_\nu^{ff} = n_e n_+ \frac{16\pi^2 Z^2 e^6}{(6\pi mkT)^{3/2}\nu^2 c} g_{ff}(T,Z,\nu)$$

Here κ_ν is an *effective* absorption coefficient that represents the difference between the true absorption coefficient and the stimulated emission coefficient.

The optical depth to free-free absorption is

$$\tau_\nu^{ff} = \int_{los} \kappa_\nu^{ff}\, ds$$

This can be estimated to ~5% precision by fitting power-laws in T and ν to the radio g_{ff}, resulting in the convenient formula:

$$\tau_\nu^{ff} = 8.24\times10^{-2} T^{-1.35} \nu_{GHz}^{-2.1} \int_{los} n_e n_+ ds$$

The integral on the left-hand side is the **Emission Measure**

$$EM = \int_{los} n_e n_+ ds$$

The Emission Measure has units of cm^{-6} pc, and is an important observable that can be derived from the surface brightness of the nebula.

At a sufficiently low frequency, the nebula is optically thick to free-free absorption. For example:

> For an HII region with $n_e \approx n_p \approx 100$ cm^{-3}, $T=10^4$ K, and a thickness of d=10 pc, $\tau^{ff}\approx 1$ at ν=200 MHz.
>
> For a PNe with $n_e \approx 3000$ cm^{-3} and d=0.1 pc, $\tau^{ff}\approx 1$ at ν=600 MHz.

A radiative transfer calculation helps to work out some of the properties of radio thermal continuum emission from ionized gas nebulae. Since we are at radio frequencies, we are in the Rayleigh-Jeans limit (hν<<kT), and the total spectral flux density of the nebula, S_ν of the nebula (intensity integrated over solid angle Ω) can be expressed in terms of the brightness temperature, $T_b(\nu)$:

$$S_\nu = \int_{source} I_\nu d\Omega = \frac{2k\nu^2}{c} \int_{source} T_b(\nu) d\Omega$$

For an isothermal nebula, the brightness temperatures at the extremes of free-free optical depth are:

$$T_b(\nu) \rightarrow T_e \tau_\nu^{ff} \quad \text{as } \tau_\nu^{ff} \rightarrow 0$$

$$T_b(\nu) \rightarrow T_e \quad \text{as } \tau_\nu^{ff} \rightarrow \infty$$

Since $\tau_\nu \propto \nu^{-2.1}$, each of these limits can occur in the same nebula at different frequencies. Thus,

$$\text{at low } \nu,\ \tau_\nu^{ff} \rightarrow \infty \text{ and } S_\nu \rightarrow \nu^2$$

$$\text{at high } \nu,\ \tau_\nu^{ff} \rightarrow 0 \text{ and } S_\nu \rightarrow \nu^{-0.1}$$

The transition between these two limits occurs at a critical frequency, ν_{crit}, that is found by setting $\tau_\nu=1$ at $\nu=\nu_{crit}$ in the equation for the free-free optical depth, giving

$$\nu_{crit}^{2.1} \approx 8.24 \times 10^{-2} n_e^2 L T_e^{-1.35} \text{ GHz}$$

where L is the thickness of the nebula along the line of sight. This is the "**Turn-Over Frequency**", which is a function of density and temperature.

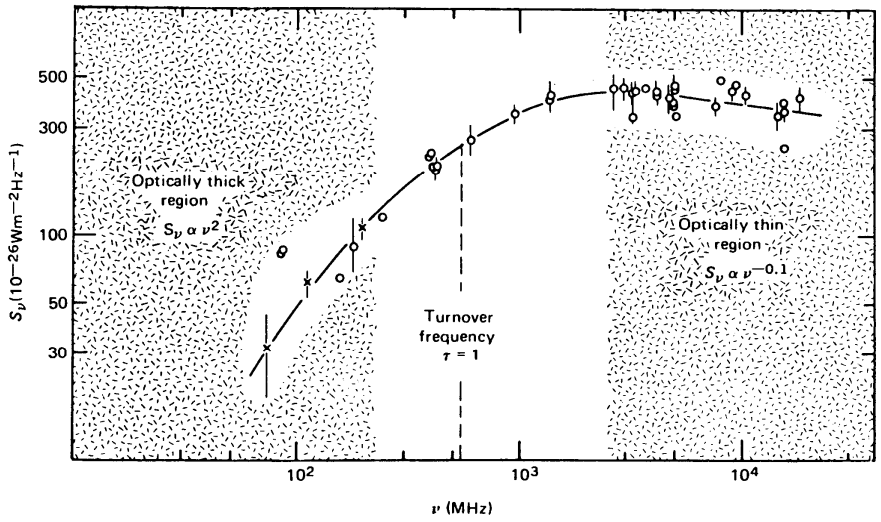


Figure III-10: Radio spectral flux density of the Orion Nebula HII region. Reproduced from Terzian & Parrish (1970, ApLett, 5, 261).

An example is shown in a radio flux-density spectrum of the Orion Nebula HII region reproduced in Figure III-10. The spectrum rises like ν^2 at low frequencies (10–100MHz) in the optically thick regime, and then flattens out and begins to fall like $\nu^{-0.1}$ at high frequencies (≥3 GHz) in the optically thin regime. Note that the "turn-over" frequency occurs where $\tau=1$, but not where the spectrum actually turns over (from increasing to decreasing flux density with frequency).

This behavior provides a way to estimate the nebular electron density, n_e. If you can observe a nebula down to a low enough frequency, the nebula becomes optically thick and the brightness temperature measures the nebular temperature, T_e directly. Then finding the turnover frequency, ν_{crit}, gives n_e.

Beware, however, that the calculations above all assume a homogenous nebula with a volume filling factor of 1. If the nebula is clumpy (and real nebulae usually are), the equations must be modified and a filling factor assumed. The typical filling factors for nebulae range between 0.01 to 0.5 for HII regions and PNe.

Radio Recombination Lines

In addition to the familiar recombination lines at UV, visible, and IR wavelengths, nebulae also emit lines arising in very highly excited states of H^0 that emit photons at radio frequencies.

At any instant, the level populations in highly excited states are small, usually 10^{-5} of the ground state populations. However, nebulae are so large that there is enough H^0 present that we can expect to observe lines from these levels.

Calculation of the frequencies is simple, as levels with large principle quantum number, n, differ very little in energy. While L–S coupling does act in these highly excited states, its effect is very small and the high-n levels are effectively degenerate in L. The frequencies of the levels are thus given by the Rydberg formula:

$$\nu = RcZ^2\left[\frac{1}{n^2} - \frac{1}{(n+\Delta n)^2}\right]$$

If the state is highly excited, n>>Δn, and we can expand the part in []'s to give

$$\nu_n \approx 2RcZ^2\,\frac{\Delta n}{n^3}$$

The Rydberg constant, R, depends on the species under consideration:

$$R_X = R_\infty\left(1 - \frac{m_e}{m_p}\right)$$

Thus, for HI, $R_H = 1.09678\times10^5$ cm^{-1}.

For radio recombination lines, the *lower* principle quantum number and its change, Δn, identify the transition responsible for the line. Examples are given in the table below.

Δn	Designation	Example
1	α transitions (Xnα)	H106α (n=107–106)
2	β transitions (Xnβ)	He109β (n=111–109)

The maximum quantum number is estimated by finding that excited state for which the typical interval for radiative de-excitation is shorter than the orbital timescale for the electron.

$$n_{\max} \approx \left[10^{20}\,\frac{T_e^{1.5}}{n_e}\,e^{26/T_e^{1/3}}\right]^{1/8.2}$$

For typical nebular conditions (N_e=1000 cm^{-3} and $T_e=10^4$ K) $n_{max}\approx740$. At this point, we enter a bizarre regime where the Hydrogen atom is on the borderline between the classical and quantum regimes (so-called mesoscale atomic physics).

Since hν<<kT, stimulated emission can be as important as spontaneous emission, quite unlike the case with the visible-wavelength recombination lines considered before. The line absorption coefficient, $\kappa_{\nu l}$, for a radio recombination line arising from an n′→n transition is:

$$\kappa_{\nu\ell} = \kappa^*_{\nu\ell}\left[1 - \frac{b_{n'}}{b_n}\,e^{-h\nu/kT_e}\right]$$

where κ^* is the absorption coefficient in the thermal equilibrium case. In the limit $h\nu<<kT_e$:

$$\kappa^*_{\nu\ell} = \frac{h^2\nu^2}{4\pi kT_e} f(\nu) n^*_n B_{nn'}$$

$$\approx 1.070\times10^7 \Delta n \frac{f_{nn'}}{n} n_e n_i T_e^{-2.5} e^{E_n/kT_e} f(\nu)$$

where f(ν) is the Doppler Broadening function for the nebula. For HI,

$$\frac{E_n}{kT_e} \approx \frac{1.579\times10^5}{n^2 T_e}$$

Assuming that the departure coefficients are slow functions of n for large n:

$$b_{n'} = b_n + \frac{db_n}{dn}\Delta n + \cdots$$

$$= b_n\left(1 + \frac{db_n}{dn}\frac{\Delta n}{b_n} + \cdots\right)$$

$$= b_n\left(1 + \Delta n \frac{d\ln b_n}{dn} + \cdots\right)$$

and, since $h\nu << kT_e$, the exponential can be expanded to lowest non-trivial order, and we can solve for the optical depth:

$$\tau_{\nu\ell} = n_{n'}\kappa_{\nu\ell} = n^*_{n'} b_{n'} \kappa^*_{\nu\ell}\left[\frac{h\nu}{kT_e} - \Delta n \frac{d\ln b_n}{dn}\right]$$

Since both terms in the []'s are small, this can be either >0 or <0, depending on the relative dominance of absorption and stimulated emission.

The emission coefficient, $j_{\nu l}$, for the line is given by

$$4\pi j_{\nu\ell} = n_{n'} A_{n'n} h\nu f(\nu) = \kappa^*_{\nu\ell} b_{n'} B_\nu(T_e)$$

Since it is possible for τ to be negative due to stimulated emission, the detailed behavior of the departure coefficients is what finally determines the emission-line strength.

In the LTE limit, the line and continuum intensities are:

$$I_{line} = B_\nu(T_e)\tau^*_{\nu\ell}$$
$$I_{cont} = B_\nu(T_e)\tau_{\nu,cont}$$

The line-to-continuum ratio is then

$$\frac{1}{I_C}\int_{line} I_{\nu\ell} d\nu = \int_{line}\frac{I_{\nu\ell}}{I_{\nu C}} d\nu = \int_{line}\frac{\tau^*_{\nu\ell}}{\tau_{\nu C}} d\nu = \int_{line}\frac{\kappa^*_{\nu\ell}}{\kappa_{\nu C}} d\nu$$

Since from the previous section

$$\kappa_{\nu C} \approx 0.08 n_e n_i \nu^{-2.1} T_e^{=1.35}$$

Using the κ^* derived above, we get

$$\int_{line}\frac{I_{\nu\ell}}{I_{\nu C}} d\nu \approx 1.3\times10^5 \Delta n \frac{f_{nn'}}{n} \nu^{2.1} T_e^{-1.15}\left[\frac{\exp\left(1.6\times10^5 \big/ n^2 T_e\right)}{\left(1 - N_{He}/N_H\right)}\right]$$

The fraction N_{He}/N_H takes into account contributions to the free-free continuum from interactions between electrons and He^+.

Observations of H and He radio recombination lines in Galactic HII regions has yielded $N_{He}/N_H \approx 0.08$, which implies that the He free-free correction factor is about 10%.

Why have we just gone through all of this?

At radio wavelengths, the line-to-continuum ratio is what is easily observed, and the relatively complicated relation above reduced to

$$\left(\frac{I_\ell}{I_C}\right) \propto \nu^{2.1} T_e^{-1.15}$$

Measurements of radio recombination lines can therefore provide accurate estimates of the electron temperature in nebulae. However, because the equation for the line absorption coefficient is sensitive to non-LTE effects, the relatively simple relations we have derived require subtle corrections to give accurate temperatures in practice. Thus far, radio recombination lines have only been observed in Galactic HII regions as they are too faint to be detected in nearby galaxies with current technologies.

III-4 Nebular Diagnostics

HII regions and PNe have spectra characterized by strong emission lines resulting from

1. Recombination cascades in H and He (and to a lesser degree in metal ions)
2. Electron-ion impact excitation of low-lying levels in metal ions.

Collisionally-excited metal lines dominate the cooling of the nebula, and can often have an integrated luminosity that far exceeds that of the recombination lines of H and He.

In this section, we will examine the emission-line spectra of ionized nebulae, and derive a number of important emission-line diagnostics that will permit measurement of the physical conditions within the nebula. The diagnostics take a number of forms:

The strengths of the H and He recombination lines are sensitive to both the total luminosity and the spectrum of the ionizing radiation field, and may be used as diagnostics of the global properties of the nebula.

The relative level populations of a given ion (e.g., O^{++} or S^{+}) depend critically on the local thermodynamic properties of the ionized gas, principally density and temperature. These relative level populations are manifest in the intensity ratios of lines emitted by the ion. We shall see how a number of readily observable line ratios provide powerful diagnostics of the nebular density and temperature in the line-forming regions.

The strengths of the metal ion lines relative to Hydrogen are sensitive to the relative ionic and total abundances of these metals in the nebula (e.g., O/H), as well as to the shape of the ionizing continuum spectrum. These provide a way to measure the gas-phase abundances of particular elements within the nebula. Observations of HII regions in our own galaxy and other galaxies have been crucial in working out the chemical enrichment history of the ISM.

The Spectra of Ionized Nebulae

Classical HII Regions

Classical HII Regions are diffuse ionized nebulae surrounding hot O and B stars. They range in size from parsec-scale regions ionized by a few O or B stars (e.g., the Orion Nebula), up to giant HII region complexes ionized by clusters containing hundreds of O stars (e.g., 30 Doradus in the LMC). The Orion Nebula is the nearest bright HII region (d≈450 pc), and is the prototype of an HII "blister" on the side of a giant molecular cloud. Most of the ionizing photons (~95%) are from the bright O6.5V star θ^1C Orionis, with the remaining ionization coming from the B0.5V star, θ^1D Orionis.

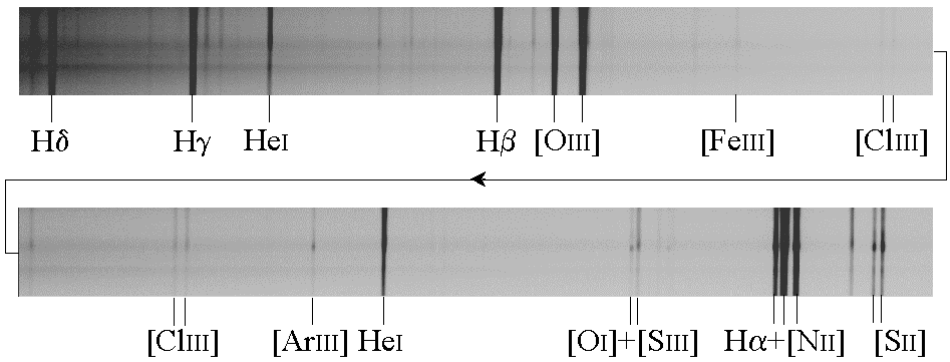
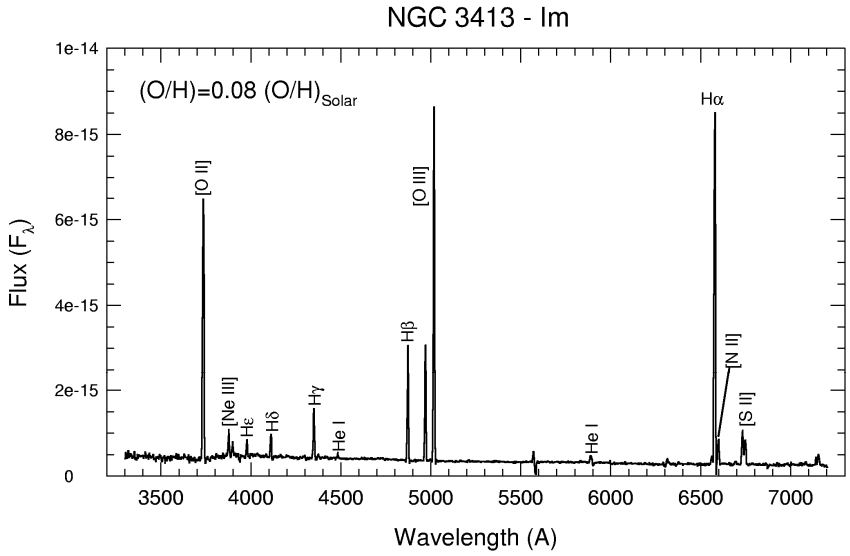


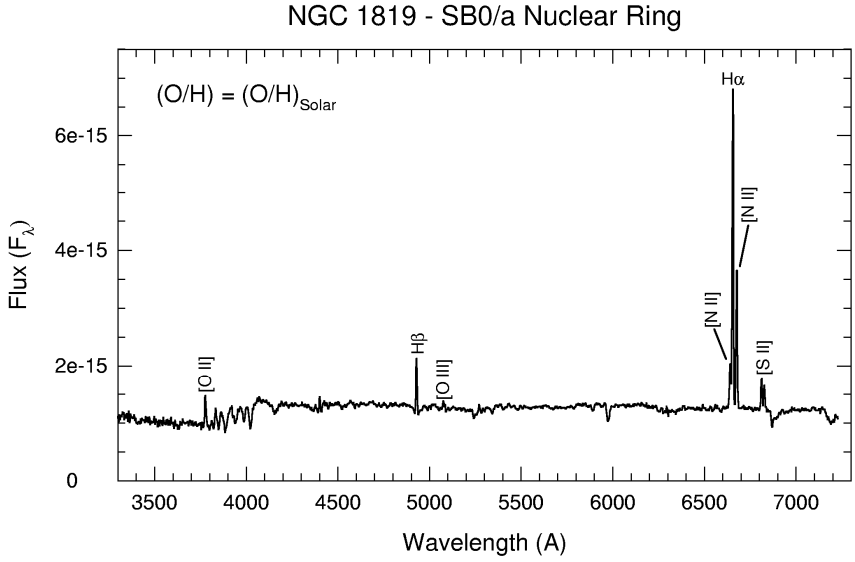
Figure III-11: Long-slit spectrum of the Orion Nebula taken in the region of the "Orion Bar" [1.8m Perkins Telescope, R. Pogge & S.R. Benfer]

This spectrum is dominated by H and He recombination lines, primarily Hα, Hβ, Hγ, and Hδ, and the HeI lines (λ4471Å and λ5876Å, plus an unlabeled line at λ6671Å), and bright lines of O^{++}, the [OIII]λλ4959,5007Å lines seen adjacent to Hβ, and N^{+}, the [NII]λλ6548,6583Å lines surrounding Hα. The [SII]λλ6716,6731Å doublet is also apparent, and is an important density diagnostic. Scans along the Bar proper are reproduced on the following pages for each of the two long-slit spectra in Figure III-11. In each, the bottom panel is scaled to show the bright emission lines visible in the long-slit spectra; this is how more distant HII regions often appear, with the fainter lines lost in the noise. The upper panels show a highly compressed intensity scale, showing the many faint lines, many with strengths <1% of Hβ, that appear throughout the spectrum at the high signal-to-noise ratios possible because Orion is so bright. Many faint lines from He^{+} O^{0}, S^{++}, Ar^{++}, Cl^{++}, Si^{++}, Fe^{+}, and Fe^{++} are visible, and only a few have been identified for clarity.

Below are spectra of the brightest HII complexes in the galaxies NGC 3413 and NGC 1819. The spectra appear very different because of the large difference in nebular abundances; NGC 3413 is an Irregular galaxy with O/H ~8% solar, while NGC 1819 is an S0 galaxy with roughly solar O/H.



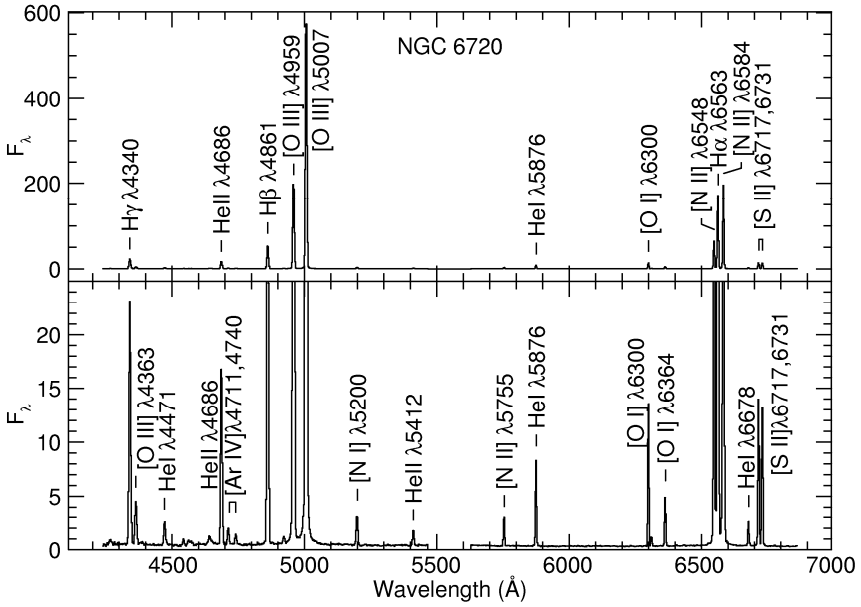
Because of the lower metallicity in NGC 3413, there is less cooling in the gas, resulting in a hotter nebula (T_e≈14000 K) and leading to greater excitation of metal-ion lines that are important for cooling like the [OIII] lines.

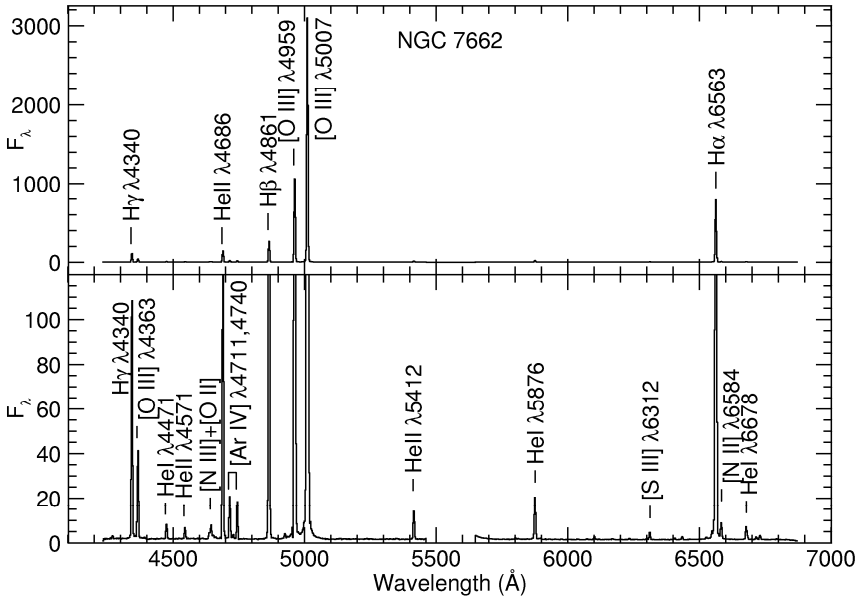


In NGC 1819, however, there are enough metals to effectively cool the nebula to T_e<7000 K, leading to weaker [OIII] lines (most of the cooling is being done by far-IR fine structure lines of O and N instead of visible O lines).

Planetary Nebulae

Planetary Nebulae (PNe) are the ejected outer envelopes of AGB stars ionized by UV photons from the remnant stellar core. These cores often have temperatures ranging from 50,000K up to 120,000K or more. Such stars are capable of doubly ionizing Helium (hν≥54.4 eV), as well as strongly ionizing many metals. The figures below (taken from N. J. Lame's OSU Ph.D. dissertation) show spectra of two PNe: the low-excitation PNe NGC 6720 (the Ring Nebula), and the high-excitation PNe NGC 7662 (the Saturn Nebula).





Notice how low-ionization potential ionic species like O^0, N^0, N^+ and S^+ are easily seen in NGC 6720, but nearly absent in NGC 7662.

Global Properties

A number of global parameters, averages over the ionized volume of the nebula, may be estimated from the emission lines.

The Emission Measure:

The "emission measure", EM, of a nebula is defined as the integral of the product of the electron and proton density along the line of sight:

$$EM \equiv \int_{los} n_e n_p ds \approx \left\langle n_e^2 \right\rangle L$$

Here L is the thickness of the nebula along the line of sight. The units of the emission measure are traditionally expressed in pc cm^{-6}. The emission measure defined in this way arises naturally from estimates of the total luminosities of recombination lines in nebulae. Consider, for example, the Hβ recombination line:

$$I(H\beta) = \int_{los} j_{h\beta} ds = \frac{1}{4\pi} \int_{los} n_e n_p h\nu_{H\beta} \alpha_{H\beta}^{eff}(H^0, T) ds$$

In general, the densities and temperature can all vary along the line of sight. In normal ionized nebulae, if the temperature is roughly constant in the fully ionized region (not a bad approximation), then setting T(s)≈⟨T⟩, we get:

$$I(H\beta) = \frac{h\nu_{H\beta} \alpha_{H\beta}^{eff}(H, \langle T \rangle)}{4\pi} \int_{los} n_e n_p ds$$

$$= \frac{h\nu_{H\beta} \alpha_{H\beta}^{eff}(H, \langle T \rangle)}{4\pi} \times EM$$

Similar expressions can be derived for all of the other recombination lines. Since in principle the emission measure is the same for all recombination lines, the HI recombination spectrum from a nebula simply scales as the emission measure, which is in principle an observable. It is more commonly derived at radio wavelengths for radio recombination lines.

Typical values of EM are:

1. Orion Nebula & some PNe: EM≈10^7 pc cm^{-6}
2. Faintest Radio HII Regions: EM≈200 pc cm^{-6}
3. Warm Interstellar Medium (WIM): EM down to 0.1 pc cm^{-6}

The latter represents about the current surface-brightness limits of the Wisconsin H-Alpha Mapping (WHAM) experiment in a 30-second integration in a 1° diameter beam with a 60-cm telescope. While we have largely concentrated on discrete HII regions in this chapter, keep in mind that 90% of the H^+ in the Galaxy is in the form of the low EM Hα-emitting clouds of the WIM.

The EM is often encountered in radio recombination-line studies of Galactic HII regions as it provides a reasonably robust estimate of the globally averaged electron density in the nebula. This is especially useful since most Galactic HII regions are heavily extinguished by dust, and so only visible at radio or far-infrared wavelengths, rendering the more familiar density diagnostics from derived from UV, optical or near-infrared forbidden lines useless.

Lyman Continuum:

In an ideal "ionization bounded nebula", all ionizing photons with hν≥13.6 eV (the "Lyman Continuum) will ionize H^0 and none escape the nebula. Since in photoionization equilibrium all photoionizations within the volume of the nebula are balanced by recombinations, we can use the strengths of the recombination lines to "count" the number of ionizing photons.

As we saw before for a simple spherical HII region (Strömgren Sphere):

$$Q(H^0) = \int_{\nu_0}^{\infty} \frac{L_\nu}{h\nu} d\nu = \int_0^{r_1} 4\pi n_e n_p \alpha_B(H^0, T) r^2 dr$$

$Q(H^0)$ is the number of ionizing photons/sec emitted by the central source, and r_1 is the radius of the Hydrogen Strömgren Sphere.

The total luminosity of a recombination line like Hβ is an integral over the ionized volume of the nebula:

$$L(H\beta) = \int_0^{r_1} 4\pi j_{H\beta} dV = \int_0^{r_1} n_e n_p h\nu_{H\beta} \alpha_{H\beta}^{eff}(H^0, T) dV$$

L(Hβ) is an observable, given an estimate of the distance to the nebula:

$$L(H\beta) = 4\pi d^2 F(H\beta)$$

Here F(Hβ) is the observed flux of Hβ from the nebula, and d is the distance. Assuming that T≈⟨T⟩ throughout the nebula is a good approximation, we find

$$Q(H^0) \approx \frac{\alpha_B(H^0, T)}{\alpha_{H\beta}^{eff}(H^0, T)} \frac{4\pi d^2 F(H\beta)}{h\nu_{H\beta}}$$

For typical nebular temperatures of $T \approx 10^4$ K,

$$Q(H^0) \approx 2.5 \times 10^{50} d_{pc}^2 F(H\beta) \text{ photons s}^{-1}$$

F(Hβ) is in flux units of erg s^{-1} cm^{-2} integrated over the line. In practice, you can use any recombination line, although Hα and Hβ are the most commonly used as they are easily observed. Brα (λ=4.05μm) and Brγ (λ=2.16μm) are sometimes used in highly obscured HII regions.

The thermal free-free continuum emission at radio wavelengths can provide a similar estimate of the Lyman Continuum flux:

$$Q(H^0) \approx 7.5 \times 10^{46} S_\nu d_{pc}^2 \nu_{GHz}^{0.1} T_e^{-0.45} \text{ photons s}^{-1}$$

The spectral flux density S_ν is measured in units of Janskys (10^{-23} erg s^{-1} cm^{-2} Hz^{-1}).

Once $Q(H^0)$ is calculated, it can be compared to the predicted $Q(H^0)$ for stars of different spectral types (or, more precisely, different effective temperatures and metallicities). An example of grids of simple LTE stellar atmospheres models for O and B stars that compute $Q(H^0)$ estimates are those of Panagia (1973, AJ, 78, 929). Common observational problems where such a comparison is made include

1. Estimations of the spectral types of stars embedded in dusty compact HII regions, where the star is effectively obscured and cannot have it spectral type determined by standard means, but it is possible to estimate $Q(H^0)$ from radio continuum measurements.
2. Estimating the number of O stars in a giant extragalactic HII regions by estimating $Q(H^0)$ from L(Hα). This is often a prelude to estimating the "instantaneous star formation rate" in a galaxy by predicting how many ionizing O stars should be present given an assumed initial mass function.

3. Estimating the temperatures of Planetary Nebula central stars using the Zanstra Method. This method compares estimates of $Q(H^0)$ from measurements of the integrated Hβ luminosity to continuum-band measurements of the brightness of the central star to give a distance independent estimate of the color temperature of the star. The traditional Zanstra method assumes a blackbody spectrum for the central star, but more sophisticated modern forms numerically integrate synthetic spectra from model atmospheres calculations.

The drawback of all of these techniques is that it really only gives a lower-limit on the number of Lyman continuum photons. If any ionizing photons escape from the nebula, or are reprocessed by dust grains thermal IR emission, the recombination-line luminosity or radio free-free continuum flux will underestimate the number of Lyman continuum photons. Nonetheless, reasonably well-defined lower limits are often better than no measurements at all

Nebular Temperatures

p^2 ions, like O^{++} and N^+, have a ground-state configuration like that sketched in Fig III-13. All of the levels above the lowest ground state are populated by electron-ion impact excitations. The populations of the ^1S and ^1D states depend critically on the temperature, and so the strengths of the emission lines arising from these levels will be temperature sensitive. In particular, the relative strengths of the $^1S \rightarrow {}^1D$ and $^1D \rightarrow {}^3P$ lines will provide a measurement of the nebular temperature, with a slight additional density dependence because the different levels have different critical densities.

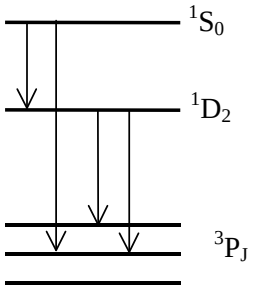


Figure III-12: Ground-state configuration for a p^2 ion like O^{++}, N^+, or S^{++}.

While a detailed treatment requires at least a 5-level atom calculation, we can learn a few things from a simple treatment as outlined below.

In the low-density limit, every collisional excitation into the ^1D level results in the emission of a photon by a downward transition into either the $^3P_2$ or $^3P_1$ states, in proportion to their transition probability (A). Similarly, every collisional excitation in the ^1S level results in the emission of a photon by downward transition into either the ^1D or $^3P_1$ states, again in proportion to their A values. Each $^1S \rightarrow {}^1D$ is also followed by a $^1D \rightarrow {}^3P$, but the contribution is small compared to direct excitation of the ^1D state.

The line ratio, in the limit $n_e \rightarrow 0$ is

$$\frac{j(^1D - {}^3P_1) + j(^1D - {}^3P_2)}{j(^1S - {}^1D)} = \frac{\gamma(^3P, {}^1D)}{\gamma(^3P, {}^1S)} \left[\frac{A_{^1S, {}^1D} + A_{^1S, {}^3P}}{A_{^1S, {}^1D}} \right] \frac{\bar{\nu}(^3P, {}^1D)}{\nu(^1D, {}^1S)} e^{-h\nu/kT_e}$$

The γ's are the effective collision strengths for the levels, and $\bar{\nu}$ is the mean frequency of the $^1D \rightarrow {}^3P$ transition weighted by the transition probabilities:

$$\bar{\nu}(^3P, ^1D) = \frac{A_{^1D,^3P_2}\nu_{^1D,^3P_2} + A_{^1D,^3P_1}\nu_{^1D,^3P_1}}{A_{^1D,^3P_2} + A_{^1D,^3P_1}}$$

and hν is the excitation energy of the $^1S \rightarrow ^1D$ transition. This simple form works for densities up to about the critical density for collisional de-excitation of the ^1D state, at which point you have to do the full statistical equilibrium calculation taking into account all collisional and radiative terms.

In the visible spectrum, there are two common density diagnostic ratios derived from (usually) readily observable line ratios of [OIII] and [NII]. The first-order solutions of the level populations result in numerical formulae for these ratios in terms of the temperature with a weak density dependence:

[OIII] Lines (Measures T_e in O^{++} Zone):

$$R_{[OIII]} = \frac{I(4959) + I(5007)}{I(4363)} \approx \frac{7.73e^{32900/T_e}}{1 + 4.5\times10^{-4}(n_e/T_e^{1/2})}$$

[NII] Lines (Measures T_e in N^+ Zone):

$$R_{[NII]} = \frac{I(6548) + I(6583)}{I(5755)} \approx \frac{6.91e^{25000/T_e}}{1 + 2.5\times10^{-3}(n_e/T_e^{1/2})}$$

These formulae are not exact, primarily because of resonances in the collision strengths, but in the temperature and density regimes of relevance for typical nebulae ($T_e \approx 5000–20000$ K, $n_e \approx 1–10^5$ cm^{-3}) they will differ from the results of detailed 5-level atom calculations by far less than the typical measurement uncertainties for this range of parameters. The observational uncertainties are primarily driven by the fact that the [OIII]λ4363Å and [NII]λ5755Å emission lines are very faint in typical nebulae, usually only a few percent the flux of Hβ or Hα. Measurement errors in the faintest lines thus drive the error in the final derived temperature.

A number of assumptions are at work in these measurements. Since the forbidden lines are optically thin, any measurement is an integral along the line of sight through the nebula. In a reasonably homogenous nebula, the temperature measures a mean temperature along the line of sight. In an inhomogeneous nebula, however, the integrated line fluxes are emissivity-weighted averages along the line of sight. The interpretation of the temperature so derived in such a case depends critically on the nature of the inhomogeneity. A number of methods have been suggested for trying to estimate the magnitude of the effect, especially as temperature estimates impact upon nebular abundance estimations.

Nebular Densities

Electron densities are measured by examining line ratios that are sensitive to collisional de-excitation. The usual method is to compare lines that arise in close-spaced doublet levels but which have very different critical densities. The best (simple) density diagnostics are p^3 ions like O^+ and S^+, which have ground state configurations like that shown in Figure III-13.

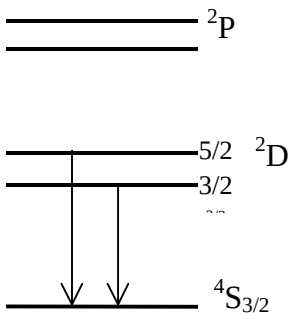


Figure III-13: Ground-state configuration of a p^3 ion like S^+ or O^+.

Particularly useful lines are those from the $^2D_{5/2,3/2} \rightarrow {}^4S_{3/2}$ transitions. Because the energy spacing of the doublet excited states is so small compared to the temperature, the exponential factor in the relative level populations is negligible, leaving behind only a weak ($T_e^{1/2}$) dependence from the collision strengths.

What matters for using these lines as a density diagnostic is that they have very different critical densities. Consider an idealized 3-level atom as shown in Figure III-14:

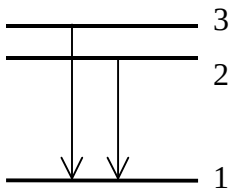


Figure III-14: A simple 3-level atom with a doublet excited state. The energy separation of the doublet levels (2 and 3) is exaggerated for clarity.

We measure the intensity ratio of the two lines arising from the excited doublet state, I_{31}/I_{32}.

Level 3 has a critical density

$$n_{crit,3} = \frac{A_{31} + A_{32}}{q_{31} + q_{32}}$$

Similarly, the critical density of level 2 is

$$n_{crit,2} = \frac{A_{21}}{q_{21}}$$

The intensity of a particular line, say I_{21}, is

$$\begin{aligned} I_{21} &= n_2 A_{21} h\nu_{21} \\ &\rightarrow n_e^2 \text{ as } n_e \rightarrow 0 \\ &\rightarrow n_e \text{ for } n_e \gg n_{crit,2} \end{aligned}$$

A similar condition holds for I_{31}. The behavior of these lines is best seen if we express the line intensities in terms of a "density normalized emissivity":

$$\frac{4\pi j_{21}}{n_e n_i} = \left(\frac{4\pi j_{21}}{n_e^2}\right)\frac{n_e}{n_i}$$

Here (n_i/n_e) is approximately the ionic abundance relative to H. A graph of the density normalized emissivities for the 3→1 and 2→1 transitions shows the behavior sketched in Figure III-15:

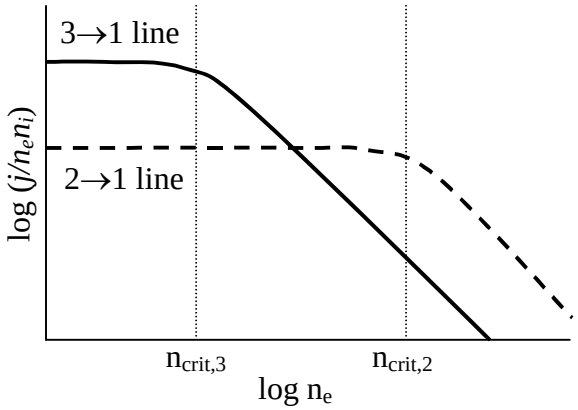


Figure III-15: Density dependence of the density-weighted emissivities for the doublet lines in our idealized 3-level atom shown in Figure III-14.

The density-normalized emissivity for a given line stays roughly constant with density until it reaches its critical density, above which it begins to fall off, eventually falling as $1/n_e$ when $n_e >> n_{crit}$. The line ratio, I_{31}/I_{21} therefore behaves as follows:

1. For $n_e \le n_{crit,3}$; I_{31}/I_{21} has a constant value
2. For $n_{crit,3} \le n_e \le n_{crit,2}$; I_{21} is roughly constant, but I_{31} drops, causing I_{31}/I_{21} to fall.
3. For $n_e \ge n_{crit,2}$; I_{21} begins to fall at the same rate as I_{31}, so that I_{31}/I_{21} becomes nearly constant again, but at a different value.

A more careful treatment shows that for the low-density limit, the line ratio tends towards the ratio of the statistical weights:

$$\frac{I_{31}}{I_{21}} \approx \frac{g_3}{g_2}$$

Far above the critical densities for both lines, the line ratio tends towards the limiting value give by Boltzmann statistics for "thermalized" level populations:

$$\frac{I_{31}}{I_{21}} \approx \frac{g_3}{g_2}\frac{A_{31}}{A_{21}}$$

The line ratio is sensitive to density (ratio $\propto 1/n_e$) within a relatively limited range of densities lying roughly between the critical densities of the two lines. At the low- and high-density limits the line ratios are effectively constant, and the line ratio only indicates which limit prevails.

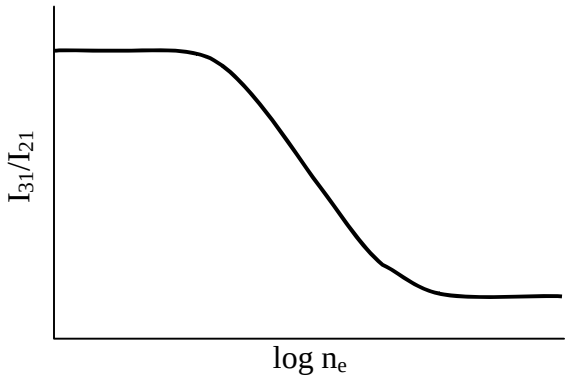


Figure III-16: Schematic density diagnostic curve for the doublet ratio in our 3-level atom.

Two density diagnostic line ratios of particular importance for the study of nebulae at visible wavelengths are:

[SII] λλ6716,6731Å Doublet (Measures n_e in S^+ Zone):

$$R_{[SII]} = \frac{I(6716)}{I(6731)}$$

These lines are well separated in wavelength (~15Å), so you don't need a high-dispersion spectrograph to measure them. It provides a reasonably good measurement of the nebular density in the range $n_e \approx 10^2$–10^4 cm^{-3} for T_e=5000–20000 K. The drawback is that S^+/H is small so the lines tend to be weak. A further drawback is that S^+ is the primary ionic species of S only within a narrow region just inside the boundary of the ionized zone of the nebula. Throughout most of the fully ionized zone the dominant ionic species of S is S^{++}.

[OII] λλ3726,3729Å Doublet (Measures n_e in O^+ Zone):

$$R_{[OII]} = \frac{I(3729)}{I(3726)}$$

This ratio works in the same regime as the [SII] doublet, but now the lines are more closely spaced, so you need a high-dispersion spectrometer to resolve the doublet. An advantage of using [OII] is that O^+ is an abundant ionic species and so the lines are bright. It does, however, share the drawback of the [SII] doublet in that O^+ is the minority ionic species of O in the nebula, and it does not provide an accurate estimate of the density in the fully ionized zone that makes up the bulk of the nebula.

Other ionic species can be used to measure densities. For example, in high-excitation PNe with very hot central stars ($T_* \geq 10^5$ K), the [ArIV]λλ4711,4740Å doublet in the visible and the [NeIV]λλ2425,2422Å doublet in the satellite UV have been used successfully and are sensitive to a similar range of nebular density as the low-excitation diagnostics above.

At much higher densities, lines of C^{++} in the satellite UV are useful for measuring densities. The [CIII]λ1907/CIII]λ1909 doublet is sensitive to densities between 10^3 and 10^6 cm^{-3}.

At near-IR wavelengths, [FeII] and [FeIII] lines are observed in the 1–2μm wavelength region that are sensitive to densities in the 10^3 to 10^6 cm^{-3} range. These have been used for measuring densities in nebulosity associated with the Galactic Center (e.g., by DePoy and collaborators).

Infrared Fine Structure Lines

The temperature and density diagnostics described so far are from different ionic states, thus while you can measure T_e in the O^{++} zone, the only accessible density diagnostic at visible wavelengths for O works in the O^+ zone. Ideally one would like diagnostics for the same ionic state, and hence the same physical region within the nebula.

The Far-IR fine structure lines of many abundant ionic species are sensitive to n_e but only very weakly sensitive to T_e. For example, for O^{++} there are two fine structure forbidden lines arising from transitions among the ^3P ground-state levels:

[OIII] $^3P_0 \rightarrow {}^3P_1$ at λ=88.4μm
[OIII] $^3P_1 \rightarrow {}^3P_2$ at λ=51.8μm

There are also lines of [SIII] that arise from the analogous transitions at 33.5μm and 18.7μm, respectively. These have been measured with spectrometers on the Kuiper Airborne Observatory, and with the ISO spacecraft.

The energy difference between the fine-structure levels is small, so that $e^{-\chi/kT} \approx 1$, and there is virtually no temperature dependence on the [OIII] I(52μm)/I(88μm), so it depends only on density (both levels have similar critical densities).

A transition from the $^1D_2 \rightarrow {}^3P_2$ level in O^{++} emits a forbidden line at λ5007Å, the decays via the $^3P_2 \rightarrow {}^3P_1$ transition to emit the λ52μm line. This ratio depends very strongly on temperature, but since $n_{crit}(^3P_2) < n_{crit}(^1D_2)$, it also has a strong density dependence.

By measuring two line ratios: I(5007Å)/I(52μm) and I(52μm)/I(88μm), we can "cross" the T_e and n_e dependence of each to make an estimate of both parameters. This was a technique pioneered by Harriet Dinerstein and her collaborators (Dinerstein, Lester, & Werner 1985, ApJ, 291, 561). It avoids many of the uncertainties due to nebular inhomogeneities when estimating n_e with one ionic species and T_e with another. Similar diagnostics have been applied to highly excited PNe using the [NeV] lines at 24.3μm, 14.3μm, and 3426Å (e.g., the ISO spectra of Pottasch et al. 1996, A&A, 315, L261 for NGC 6302). ISO and now Spitzer data have been particularly useful for extending the use of these diagnostics to more nebulae.

The particular disadvantage of the far-IR lines is their relative inaccessibility in a spectral region can only be observed from space or high-altitude airborne observatories. A further problem is that visible-light spectrometers permit angular resolutions of an arcsecond or better, whereas far-IR observations are on scales of arcminutes or a little less (the diffraction limit at 5000Å is 100 times smaller than that at 50μm for the same telescope), so in resolved objects it is difficult to compare far-IR and visible-light measurements in detail.

Abundances

Abundances refer to the relative proportions of atomic species in the gas phase in a nebula. These are usually expressed relative to Hydrogen. There are two types of nebular abundances that can be measured from the emission-line spectra of nebulae:

Ionic Abundances:

These are the relative abundances of particular ionic species, e.g., O^+/H^+. These are usually local measures as different ionization states of an element are found in different regions within a nebula due to its detailed ionization structure.

Total (Elemental) Abundances:

These are relative abundances of particular elements, e.g., O/H, expressed as the sum over all ionic forms present in the nebula. These provide a global estimate of the relative proportions of a given element in the gas-phase in the nebula.

Ionized nebulae are our primary probes of the gas-phase abundances in regions of recent star formation, and are the primary tracers of the chemical evolution of the ISM in our Galaxy and other galaxies.

Ionic Abundances

The observed intensity of a particular ionic emission line, I_{ul}, is given in general by

$$I_{ul} = \int_{los} j_{ul} ds = \int_{los} n_e n_i \varepsilon_{ul}(T_e) ds$$

where:

$n_i =$ density of the ion (e.g., O^+)

$\varepsilon_{ul}(T_e) =$ line emissivity

The line emissivity is function of the *local* electron temperature, T_e. For recombination lines, like H or He, the emissivity can be written in terms of the effective recombination coefficient, for example, the intensity of Hβ is

$$I(H\beta) = \frac{1}{4\pi} \int_{los} n_e n_p h\nu_{H\beta} \alpha^{eff}_{H\beta}(T_e) ds$$

Collisionally excited lines, like [OIII]λ5007Å, are expressed in terms of the collision strengths. In the low-density limit, this is given by:

$$I_{ul} = \frac{1}{4\pi} \int_{los} n_e n_i h\nu_{ul} q_{ul}(T_e) b_{ul} ds$$

Here b_{ul} is the fraction of excitations into the upper excited state that gives rise to the line (it is *not* a departure coefficient – beware!).

Line ratios are measured instead of total line luminosities, primarily because they are distance independent and do not depend on absolute calibration (which is always much harder than relative calibration in spectroscopy). For example, it is common practice at visible wavelengths to measure line strengths relative to Hβ because this line is readily observed and generally unblended with other lines (unlike Hα, which is often blended with [NII] lines). This ratio is usually written in terms of line emissivities, ε:

$$\frac{I_{ul}}{I_{H\beta}} = \frac{\int n_e n_i \varepsilon_{ul}(T_e) ds}{\int n_e n_p \varepsilon_{H\beta}(T_e) ds}$$

For a recombination line, like Hβ:

$$\varepsilon_{H\beta}(T_e) = \frac{4\pi j_{H\beta}}{n_e n_p} = h\nu_{H\beta} \alpha^{eff}_{H\beta}(T_e)$$

For a collisionally-excited line:

$$\varepsilon_{ul}(T_e) = \frac{8.63\times10^{-6}}{T_e^{1/2}} \frac{\gamma_{lu}}{g_u} e^{-h\nu_{ul}/kT} b_{ul} h\nu_{ul}$$

The simplest assumption to make is that the nebula is homogeneous along the line of sight. In this case, the density of an ion relative to the density of protons (Hydrogen ions) is:

$$\frac{n_i}{n_p} = \frac{I_{ul}}{I_{H\beta}} \times \frac{\varepsilon_{H\beta}(T_e, n_e)}{\varepsilon_{ul}(T_e, n_e)}$$

The emissivities are assumed to be known" functions of T_e, and I've included n_e to remind you of the weak density dependence of the recombination lines.

Recombination and collisionally excited lines have different temperature dependencies. In general:

$$\varepsilon_{H\beta} \propto T_e^{-0.9}$$

$$\varepsilon_{ul} \propto T_e^{-0.5} e^{-h\nu_{ul}/kT_e}$$

The exponential dependence on temperature of the collisionally excited lines is of particular concern, and one has to know T_e fairly accurately in order to derive good ionic abundances.

In real nebulae (i.e., those outside of textbooks and computer codes), T_e and n_e will vary with position, so our assumption of homogeneity must be modified if we are to believe our derived nebular abundances. A number of approaches have been proposed to account for the effects of inhomogeneities. The temperature fluctuations method proposed by Manuel Peimbert (1967, ApJ, 150, 825) has been used most frequently. This uses the observed line ratios to derive an estimate of the extent of fluctuations in T_e. The emissivity is expanded as a power series in T_e:

$$\varepsilon(T) = \varepsilon(T_0) + (T - T_0)\left(\frac{d\varepsilon}{dT}\right)_0 + \frac{1}{2}(T - T_0)^2\left(\frac{d^2\varepsilon}{dT^2}\right)_0 + \cdots$$

T_0 is the emission-measure-weighted average T along the line of sight:

$$T_0 = \frac{\int n_e n_i T ds}{\int n_e n_i ds}$$

Temperature fluctuations along the line of sight are quantified by the parameter t^2:

$$t^2 = \frac{\int n_e n_i (T - T_0)^2 ds}{\int n_e n_i ds}$$

This is the normalized second-moment of Temperature along the line of sight.

If all ions share a similar distribution in space in the nebula, then two line ratios that are diagnostic of temperature could in principle be used to measure both T_0 and t^2. For example the emission line ratios [OIII](4959+5007)/4363 and [SIII](9069+9532)/6312 (an analog of the [OIII] ratio) are sometimes used for this purpose because they are found in the same regions of a nebula.

Once these are estimated, we can then find the appropriate temperature to use for a given line. For Hβ, the appropriate temperature to use is

$$T_e(H\beta) \approx T_0[1 - 0.92t^2]$$

Thus, the recombination-line emissivities are weighted towards regions with cooler temperatures (recombination is more effective for lower temperatures).

For collisionally excited forbidden lines, the steeper temperature dependence results in emissivities that are weighted toward higher temperatures (higher T means more electrons sufficiently energetic to collisionally excite the lines):

$$T_e \approx T_0\left[1+\left(\frac{\chi_1+\chi_2}{kT_0}-3\right)\frac{t^2}{2}\right]$$

Where the χ's are the excitation potentials relative to the ground state for the lower (1) and upper (2) level responsible for the line of interest.

Observational work in the visible and Far-IR has produced values of the temperature fluctuation parameter of $t^2 \approx 0.035$ for Orion to $t^2 \approx 0.02$ for PNe, but in general t^2 is very hard to determine with any precision. A non-zero t^2 can change the inferred ionic abundances by factors of 2–3. A good illustrative example of this kind of work is Nollenberg et al. 2002 [ApJ, 581, 1002], who combined optical and ISO spectra to study sulfur abundances and temperature fluctuations in two extragalactic HII regions.

Total (Elemental) Abundances

If you could measure the abundances of all of the ions of a given element, you could estimate the total abundances of that element with respect to H. The elements that can be measured with various degrees of precision in nebulae are He, O, S, N, and Ne. With good UV spectroscopy (especially with FUSE and HST), C, Ar, and Cl abundances are also being measured with greater precision.

Oxygen is particularly important as the primary ionic species of O found in nebulae have strong optical and near-UV emission lines because O is relatively abundant and a primary coolant in most nebulae. It is more difficult to determine the abundances of other elements because only one or two of the primary ionic forms are observable in a single wavelength band, and because the lines are often relatively weak and blended with other features (e.g., stellar absorption lines and continuum mixed with the nebular spectrum).

Total abundances are estimated by adding up the contributions from all observed ionic species, and then correcting for the unobserved ions using an empirical Ionization Correction Factor (ICF). Most ICFs use of easily observed ionic species, chiefly O and He ions, and exploit similarities in the ionization potentials to make the appropriate corrections. The assumption is that ionic species with roughly similar ionization potentials should arise in the same physical regions. This is a reasonable assumption for most nebular conditions, but can become problematic if there are large density variations, as lines from the different species can often have very different critical densities.

Examples of commonly exploited ionization potential similarities are:

He^+ (IP=54.4eV) and O^{++} (IP=54.9eV), which differ by 0.5eV
O^+ (IP=35.1eV) and S^{++} (IP=34.8eV), which differ by 0.3eV

The near coincidence in IP between O^+ and S^{++}, for example, can be used to correct the observed S^{++} ionic abundances for the unseen S^{3+} lines in terms of the observed O^+ and O^{++} lines:

$$\frac{n(S^{++})}{n(S^{3+})} \approx \frac{n(O^{+})}{n(O^{++})}$$

Lines of [SII], [SIII], [OIII], and [OII] are readily observed at visible/near-IR wavelengths, but no strong [SIV] lines are seen (they appear at satellite UV wavelengths). Thus the S/H abundance, corrected for unseen ionic states of S should be:

$$\frac{S}{H} = \left(\frac{S^+ + S^{++}}{H^+}\right) \times ICF \approx \left(\frac{S^+ + S^{++}}{H^+}\right) \times \left(\frac{O^+ + O^{++}}{O^+}\right)$$

Another example is in highly excited PNe, where there is significant He^{++} emission. The fraction of O that is in the form of O^{3+} and O^{4+} should follow the relative fraction of He in He^{++}:

$$\frac{O}{H} = \left(\frac{O^{+} + O^{++}}{H^{+}} \right) \times ICF \approx \left(\frac{O^{+} + O^{++}}{H^{+}} \right) \times \left(\frac{He^{+} + He^{++}}{He^{+}} \right)$$

The O ionic fractions in both examples are estimated from the observed strengths of [OII] and [OIII] lines relative to Hβ. Some schemes for estimating ICFs work better than others, depending upon how much one wishes to stretch the definition of "coincidence" when comparing the ionization potentials of different ions. The best have usually been confirmed with detailed Monte Carlo analyses with model nebulae. In general, the more ions of different species you have, and the greater the range of IP's, the better you can constrain the derived abundances.