Nebular Spectroscopy: A Guide on H II Regions and Planetary Nebulae

Manuel Peimbert, Antonio Peimbert, and Gloria Delgado-Inglada
Instituto de Astronomía Universidad Nacional Autónoma de México Apartado Postal 70–264, C.P. 04510 México, CDMX, Mexico
Received 2017 March 31; accepted 2017 May 8; published 2017 June 20

Abstract

We present a tutorial on the determination of the physical conditions and chemical abundances in gaseous nebulae. We also include a brief review of recent results on the study of gaseous nebulae, their relevance for the study of stellar evolution, galactic chemical evolution, and the evolution of the universe. One of the most important problems in abundance determinations is the existence of a discrepancy between the abundances determined with collisionally excited lines and those determined by recombination lines: this is called abundance discrepancy factor (ADF) problem, and we review results related to it. Finally, we discuss the possible reasons for the large $r^2$ values observed in gaseous nebulae.

Key words: ISM: abundances – (ISM:) H II regions – (ISM:) planetary nebulae: general

Online material: color figures

1. Introduction

There are two main types of gaseous nebulae that we will discuss in this review: H II regions and planetary nebulae (PNe). Precise models of individual nebulae permit us to determine accurate abundances, which allow us to test models of stellar evolution, galactic chemical evolution, and the evolution of the universe as a whole.

H II regions are sites where star formation is occurring. Therefore, H II regions provide us with the initial abundances from which stars are made at present; abundances are paramount to test models of the chemical evolution of galaxies. For spiral galaxies they provide us with radial abundance gradients of heavy elements relative to hydrogen that have to be explained by models of galactic chemical evolution (Berg et al. 2013; Sánchez et al. 2014; Bresolin & Kennicutt 2015; Esteban et al. 2015; Magrini et al. 2016; Arellano-Córdova et al. 2016). Irregular galaxies that have a high fraction of their mass in the form of gas where almost no heavy elements are present, permit us to determine the primordial abundance of helium and hydrogen due to Big Bang nucleosynthesis (Ferland et al. 2010; Izotov et al. 2014; Aver et al. 2015; Peimbert et al. 2016).

PNe are produced by intermediate-mass stars, those in the 0.8 M$\odot$ to 8 M$\odot$ range, while they transit from the red giant stage to the white dwarf stage. The study of these nebulae show that intermediate-mass stars are responsible for most of the nitrogen, about half the carbon, and a non-negligible fraction of oxygen and helium present in the universe (e.g., Karakas & Lattanzio 2014). These stars also produce an important fraction of the slow neutron process elements, like rubidium, strontium, yttrium, zirconium, cesium, barium, lanthanum, and praseodymium. Big Bang nucleosynthesis created all of the hydrogen and deuterium; most of the helium; and a fraction of the lithium. Massive stars, those with more than 8 M$\odot$, produce most of the remaining elements, including the fast neutron process elements as well as a fraction of the helium, carbon, nitrogen, and oxygen abundances (e.g., Pagel 2009).

The classical textbook on the study of gaseous nebulae and active galactic nuclei is Osterbrock & Ferland (2006). This book discusses in depth the physical processes in PNe and H II regions. It also mentions previous books and many review papers that presented earlier ideas and results on these objects. Other texts that in the past were considered fundamental in the study of photoionized regions are Stromgren (1939, 1948), Seaton (1960), Osterbrock (1974, 1989), and Aller (1984).

Other relevant review articles and books related to the physical conditions of gaseous nebulae are Kwok (2000) on the origin and evolution of planetary nebulae, Ferland (2003) on quantitative spectroscopy of photoionized clouds; Dopita & Sutherland (2003) on the astrophysics of the ionized universe; Stasińska (2004) on cosmochemistry the melting pot of the elements; Pagel (2009) on nucleosynthesis and the chemical evolution of galaxies; and Stasińska et al. (2012) on oxygen in the universe.

In the last few years there have been four meetings dedicated only to planetary nebulae. Two of the symposia organized by the IAU (Manchado et al. 2012; Liu et al. 2017) and two conferences on asymmetric planetary nebulae (Zijlstra et al. 2011; Morisset et al. 2014). The IAU Planetary Nebula
Working Group produced a review paper on the present and future of PNe and their central stars research and related subjects (Kwitter et al. 2014).

Pérez-Montero (2017) recently published a tutorial that deals specifically with the determination of chemical abundances of H II regions through the so-called direct method.

In this paper, we present a combination of a tutorial and a review paper on recent results related to the study of H II regions and PNe. Section 2 describes some basic physical processes present in gaseous nebulae. Section 3 describes the main methods used to determine the physical conditions in gaseous nebulae. Section 4 discusses methods to determine the total abundances of the elements in gaseous nebulae, with special emphasis on the determination of the ionization correction factors to take into account the ions that are not observed. Section 5 presents recent results derived from the study of galactic and extragalactic H II regions. Section 6 presents recent results derived from planetary nebulae. Section 7 includes a general discussion on possible physical processes present in gaseous nebulae. Section 8 presents recent results derived from the study of planetary nebulae and HII regions. In this section we briefly explain some concepts that are necessary to understand the subsequent sections. It is beyond the scope of this paper to describe all of them in detail, and we suggest that the reader review the books and papers mentioned above (in particular, the book by Osterbrock & Ferland 2006) to understand the physics involved in these objects.

2. Brief Discussion of Physical Processes in Gaseous Nebulae

There are several processes occurring in the ionized gas of planetary nebulae and H II regions. In this section we briefly explain some concepts (such as photoionization, recombination, heating, cooling, and emission line mechanisms) that are necessary to understand the subsequent sections. It is beyond the scope of this paper to describe all of them in detail, and we suggest that the reader review the books and papers mentioned above (in particular, the book by Osterbrock & Ferland 2006) to understand the physics involved in these objects.

2.1. Photoionization and Recombination Processes: Local Ionization Equilibrium

The basis of the study of photoionized regions is to assume an equilibrium between ionization and recombination. It has long been known that in equilibrium a photoionized region will have a large volume of ionized gas surrounded by a relatively narrow transition zone where the gas goes from being nearly completely ionized to nearly completely neutral (e.g., Stromgren 1939; Osterbrock & Ferland 2006).

Because approximately 90% of the atoms of the ISM are hydrogen, for a first approximation the equilibrium is studied for a gas made up entirely of hydrogen atoms. Locally one must study the equilibrium between recombination and ionization:

\[
\nu_{\text{e}}n_{\text{p}}\alpha_{\text{A}}(H^0, T) = n(H^0) \int_{\nu_{0}}^{\infty} \frac{4\pi \nu}{h \nu} \sigma_{\nu} d\nu,
\]

where \(n_{\text{e}}, n_{\text{p}}, n(H^0),\) represent the electron, proton, and neutral hydrogen density, \(\alpha_{\text{A}}(H^0, T)\) represents the recombination coefficient for hydrogen at a given temperature, \(h \nu_{0}\) the energy required to ionize hydrogen (13.6 eV), \(J_{\nu}\) represents the local radiation, and \(a_{\nu}\) the ionization cross section of a given photon. In photoionized regions, \(\int_{\nu_{0}}^{\infty} \frac{4\pi \nu}{h \nu} \sigma_{\nu} d\nu\) dominates over \(n_{\text{e}}n_{\text{p}}\alpha_{\text{A}}(H^0, T)\) and the ionization fraction is nearly one for most of the volume.

Globally, one can estimate the maximum volume that can be ionized by a constant ionization source. The first thing to notice is that recombinations to the ground level will produce photons with energy greater that \(h \nu_{0}\), thus they can be considered an additional source of ionization; this source usually accounts for approximately 40% of the ionizing photons. One solution to avoid this problem is to ignore recombinations to the ground level (as the subsequent ionization will cancel out the recombination). The recombination rate to all levels but the ground level is usually referred to as \(\alpha_{\text{f}}(H^0, T)\). Because there will be a steady rate of recombinations, ionizing photons will be required to keep a volume of gas ionized; those photons will be exhausted when a volume of size \(4\pi r_{S}^3\) is ionized, where \(r_{S}\) (the Strömgren radius) can be estimated as

\[
Q(H^0) = \frac{4\pi}{3} r_{S}^3 n^2(H) \alpha_{\text{f}}(H^0, T),
\]

where \(Q(H^0)\) represents the rate of ionizing photons produced by the central star, and \(n(H)\) the total hydrogen density.

For nebulae with realistic chemical compositions, the ionization of helium needs to be modeled as well as those of heavier elements. In general, neither will affect much the results for hydrogen, as helium recombination will, in general, return hydrogen ionizing photons, and the abundance of heavy elements is very small. However, the degree of ionization of heavy elements cannot be neglected because it turns out to be very important for the temperature equilibrium.

2.2. Heating and Cooling: Local Thermal Equilibrium

The previous results cannot be properly determined without knowing the physical conditions of the gas of the nebula. While the density is generally considered to be given by the characteristics of the gas previous to the photoionization, the energetics of photoionized regions are dominated by photoionization, and the temperature strongly depends on the physics of photoionization. As such, in equilibrium the temperature of the gas comes from a balance of local cooling and heating processes.

Heating will come from photoionizing photons, and hydrogen photoionization will account for at least 90% of this
heating. Helium photoionization will account for 10%, or less of the total heating, while photoionization of heavy elements will only produce trace amounts of heating. Additional heating can come from dust photo-heating, or from free–free absorption, but neither will dominate the energetics.

The heating due to photoionization will be due to the excess energy of photoionizing photons beyond the ionizing threshold of the different atoms and ions (13.6 eV for hydrogen); consequently, it will be proportional to the number of photoionizations and to the temperature of the photoionizing star. On the other hand, in equilibrium, the number of photoionizations is equal to the number of recombinations, which in turn is proportional to the density squared for most of the object.

Other possible sources of heating are cosmic rays and shock waves; these are traditionally considered to be unimportant, thus they are ignored. However, while the former is not expected to be important, the latter can locally contribute with an important fraction of the heating (see Section 7).

To balance the heating, one must consider the possible sources of cooling. The most obvious source that must be considered is recombination where the kinetic energy of the captured photon is removed from the gas. This is, however, not the most important source of cooling.

If one were to consider a gas made up entirely of hydrogen (or hydrogen and helium), recombination would indeed be the best way to remove energy from the nebula. In this scenario, the heating is proportional to the temperature of the star and the cooling is proportional to the temperature of the nebula, so balance will occur when the gas has a temperature similar to that of the star (in fact, when done carefully, the gas ends up being hotter than the ionizing star). This is not seen in nature. Ionizing stars of H II regions have temperatures of $T_{\text{star}} \sim 30000$–$45000$ K, while H II regions often have temperatures of $T_e \sim 7000$–$15000$ K; central stars of PNe have temperatures of $T_{\text{star}} \sim 30000$–$200000$ K, while PNe often have temperatures of $T_e \sim 7000$–$20000$ K. The difference is due to the additional cooling produced by forbidden lines excited by collisions.

2.3. Emission-line Mechanisms

The spectra of ionized nebulae are characterized by a weak continuum and strong emission lines. Figure 1 shows some of the most conspicuous emission lines: [O II] $\lambda\lambda 3727, 3729$, H$\delta$, H$\gamma$, H$\beta$, [O III] $\lambda\lambda 4959, 5007$ in the blue part of the spectrum, and H$\alpha$, [N II] $\lambda\lambda 6548, 6583$, He I $\lambda 6678$, [S II] $\lambda\lambda 6717, 6731$, He I $\lambda 7065$, [Ar III] $\lambda 7135$ in the red part of the spectrum.

These lines are produced when an atom (or ion) is de-excited after being excited by photons or by collisions between atoms or atoms and electrons. The weak continuum is due to bound-free, free–free, two photon emission, dust scattered light, and starlight for extragalactic H II regions. The main mechanisms producing emission lines are recombination, collisional excitation, and photoexcitation. One particular line can be produced by several mechanisms, but usually one of them dominates the others. Most of the strongest lines in ionized nebulae are produced by collisional excitation.

Deeper spectra allow us to measure the considerable fainter recombination lines. Figures 2 and 3 show three published echelle spectra of the H II regions M8 and M17, obtained by García-Rojas et al. (2007) and the PN NGC 5315 obtained by Peimbert et al. (2004). These portions of the spectra show several recombination lines (such as C II, O I, N II, Ne II, and N III lines) together with some collisionally excited lines of [Fe III] and [Ar IV]. The broad emission feature in the spectra of NGC 5315 is due to the Wolf-Rayet type central star.
In general, the intensity of an emission line, \( I(\lambda) \), can be written as

\[
I(\lambda) = \int j_\lambda \, ds = \int n(X^+)n_e \epsilon_\lambda(T_e) \, ds,
\]

where \( j_\lambda \) is the emission coefficient, \( n(X^+) \) is the density of the ion that emits the line, \( n_e \) is the electron density, and \( \epsilon_\lambda \) is the emissivity. For example, the H\( \beta \) and [O III] \( \lambda 5007 \) intensity lines are given by

\[
I(H\beta) = \int n(H^+)n_e \epsilon(H\beta, T_e) \, ds,
\]

\[
I([O III]) = \int n(O^{++})n_e \epsilon(\lambda 5007, T_e) \, ds.
\]

The emissivities of recombination lines (RLs) and collisionally excited lines (CELs) are discussed below.

### 2.3.1. Recombination Lines

RLs are produced when free electrons are captured by ions and descend from excited to lower levels emitting photons in the process. These lines are also known as permitted lines because they typically satisfy all the selection rules for an electric dipole transition. Most of the bright RLs in emission spectra are from hydrogen and helium. Metals such as carbon, nitrogen, and oxygen also produce RLs, but they are much weaker due to their lower abundances with respect to hydrogen and helium.

The emission coefficient of a recombination line, \( j_{\text{int}} \), is given by

\[
j_{\text{int}} = \frac{\hbar \nu_{\text{int}}}{4\pi} n(X^+)n_e \alpha_{\text{int}}^\text{eff},
\]

where \( \hbar \nu_{\text{int}} \) is the energy difference between the two levels and \( \alpha_{\text{int}}^\text{eff} \) represents the effective recombination coefficient.

Some of the RLs that can be found in the spectra of ionized nebulae are H I lines (e.g., H\( \alpha \) at 6563 Å, H\( \beta \) at 4861 Å, H\( \gamma \) at 4340 Å), He I lines (e.g., 5875 and 4471 Å), He II lines (e.g., 4686 Å), O I lines (e.g., 8446 and 8447 Å), O II lines (e.g., 4639, 4642, 4649 Å), O III lines (e.g., 3265 Å), O IV lines (e.g., 4657 Å), C II lines (e.g., 4267 Å), C III lines (e.g., 4647 Å), C IV lines (e.g., 4657 Å), N II lines (e.g., 4237 and 4242 Å), N III lines (e.g., 4379 Å), N IV lines (e.g., 4606 Å), and Ne II lines (e.g., 3694 Å).

The H I photons emitted by recombination may or may not escape. In an optically thin nebula, all the emitted photons will escape; this is known as Case A (Baker et al. 1938). On the other hand, in an optically thick nebula, all of the hydrogen Lyman photons will be absorbed; this is called Case B. The intermediate situations between these two extreme cases are explained below.

### 2.3.2. Collisionally Excited Lines

In contrast with hydrogen and helium, the energies of the first excited levels of some ions of heavy elements are within a few eV of the ground level, which makes it relatively easy to reach them through collisions with electrons. CELs are produced when the atoms excited through collisions decay via radiative transitions. Although the transition probability of these lines is low, the relatively low density of ionized nebulae allows these transitions to occur. Some of these transitions, the
ones produced in the optical range, are forbidden by the parity selection rule ($\Delta L = \pm 1$), thus the emitted lines are often called forbidden lines. Others, called semi-forbidden lines, only violate the spin rule ($\Delta S = 0$).

The emission coefficient of a collisionally excited line produced by a radiative transition from level $k$ to level $l$ is given by

$$f_{kl} = \frac{h \nu_{kl}}{4\pi} f_k A_{kl} n(X^+)$$

where $f_k$ is the fraction of ions $X^{+i}$ in the upper level, $k$, and $A_{kl}$ is the spontaneous transition probability from level $k$ to $l$. To compute the emissivity of a CEL, one needs to know the population of the upper level. It is necessary to solve the statistical equilibrium equations: the rate of population of a level by radiative and collisional processes is balanced with the rate of de-population by these processes:

$$\sum_{l>k} f_l n_e q_{lk} + \sum_{l<k} f_k A_{kl} = \sum_{l>k} f_k n_e q_{kl} + \sum_{l<k} f_l A_{lk},$$

where $f_l$ and $f_k$ are the fraction of ions $X^{+i}$ in the levels $l$ and $k$, and $q_{kl}$ and $q_{lk}$ are the collisional de-excitation and excitation rates. The detailed expressions of $q_{kl}$ and $q_{lk}$ can be found in Osterbrock & Ferland (2006). To solve these equations, atomic data (collision strengths and radiative transitions) and an estimate of the physical conditions are necessary. Software such as PyNeb (Luridiana et al. 2015) and the IRAF package NEUBARL provide a solution for the equations for a number of levels depending on the ion.

The CELs are one of the main cooling factors within ionized nebulae, so the presence of intense CELs depends on having available levels a few eV above the ground level. Because the temperature in photoionized nebulae is usually equivalent to about 1 eV, these differences are only available within the same ground-state electron configuration and therefore require the presence of a sub-shell with at least two electrons and two empty spaces to have the possibility of fine structure interaction between equivalent electrons to produce multiple levels within such configuration.

### 2.3.3. Fluorescence

Some permitted lines of some elements, such as oxygen, are brighter than expected by pure recombination because they are excited by starlight or by other nebular lines. The Bowen lines are one particular case of these lines; they are produced because there is a coincidence between the wavelengths of two lines. In this example, the O III line at 303.80 Å and the Ne II line at 303.78 Å. Some of the photons emitted by the He$^{++}$ atoms are absorbed by the O$^{++}$ atoms and then are re-emitted. Because the interpretation of fluorescence lines is complicated, they are preferably not used to derive physical or chemical parameters in nebulae; indeed, it is important to identify this type of lines in order to avoid them in the calculations.

**Case C and Case D.** From the early studies of photoionized regions, it was recognized that the hydrogen spectrum strongly depends on the optical depth of the Lyman lines as well as on the effect of the possible presence of fluorescence of the same lines. Baker et al. (1938) and Aller et al. (1939) defined Case A as the simplest possible case, where the nebula is optically thin; Case B, when the size of the nebula is such that the optical depth of the Lyman lines is so large that the fraction of such photons that escape is negligible; and Case C when the spectrum of an optically thin nebula is affected by fluorescence of hydrogen due to non ionizing continuum (Ferland 1999).

Case B has been the most studied of these, as the physical conditions of most nebulae are closer to those of Case B than to those of cases A and C. Recently, it has been recognized that some objects are affected by both optical depth and fluorescence (Luridiana et al. 2009; Peimbert et al. 2016); this scenario is frequently called Case D.

In Case D, because the hydrogen transitions are optically thick, de-excitation occurs through higher series lines, in particular excitation to level $n$ ($n > 3$) produce transitions to $n'(n' > 2)$, Balmer emissivities are systematically enhanced above case B predictions. Moreover, the He I lines are also enhanced by fluorescence. Case D produces small effects in the H I and He I lines, but they might be important in the determination of the primordial helium abundance.

**Optical depth of levels other than the ground level.** The first excited level of H decays too fast to be of any significance: heavy elements are not abundant enough to have very large optical depths; however, He$^+$ is abundant enough and has a metastable level with a long enough half life to require special attention.

The effects on the helium abundance determination due to the optical depth of the 2$^2$S metastable level have been studied by Robbins (1968); Benjamin et al. (2002), and Aver et al. (2011).

### 3. Calculation of Physical Conditions and Ionic Abundances from Observations

Using the concepts described in Section 2 and the simple assumptions that the photoionized region is homogeneous in temperature, density, and chemical composition, it is possible to determine the physical conditions and chemical abundances of photoionized regions. To a first approximation, these three assumptions seem to be adequate (all the gas has a common origin, and photoionization models show that temperature should vary by only a few percent across most of each photoionized region). Careful study of the best observed objects, in particular the presence of abundance discrepancy factors (ADFs), show that at least one of those simplifications is not warranted.
In this section, we describe how, starting from observations of specific photoionized regions and atomic data sets, it is possible to determine the physical conditions (density and temperature) and the ionic chemical abundances. While presenting the determinations of ionic chemical abundances we will present two sets of chemical abundances: those derived from CELs and those derived from RLs; when both sets are available, RLs produce higher chemical abundances and their ratio is called the ADF.

### 3.1. Quality of the Available Spectra

Long-slit optical spectra have been widely used in the literature to derive the physical conditions and ionic abundances of ionized nebulae. There are spectra of several hundreds of Galactic and extragalactic PNe and H II regions with resolutions better than \( \sim 5 \) Å (see, for example, the compilations by Kwitter & Henry 2012; Maciel et al. 2017). These spectra allow the detection of nebular lines such as [O III] \( \lambda 4959, \lambda 5007 \), [N II] \( \lambda 6548, \lambda 6583 \), [S II] \( \lambda 6717, \lambda 6731 \), as well as RLs from hydrogen, helium, and carbon (C II \( \lambda 4267 \)). Deep long-slit spectra also allow the measurement of weaker lines such as the auroral lines [O III] \( \lambda 4363 \) and [N II] \( \lambda 5755 \). Some available long-slit spectrographs are LRIS at Keck I telescope, GMOS at Gemini telescope, FORS at Very Large Telescope (VLT), and OSIRIS at Gran Telescopio Canarias (GTC); many more spectrographs are available in smaller telescopes.

Echelle spectra provide a resolution better than 1 Å that allows us to separate and measure nearby lines, such as the multiplet 1 of O II at \( \sim 4650 \), used to compute O \( ^{++} \) abundances with RLs, whose lines can be contaminated with lines of N II, N III, [Fe III]. The number of ionized nebulae with deep and high-resolution data in the literature is around 50 (see, e.g., the sample used by Delgado-Ingla et al. 2014). Some examples of echelle spectrographs are HIRES at Keck I telescope, MIKE at Magellan Clay telescope, and UVES at VLT; dozens of other echelle spectrographs are available in smaller telescopes, but one requires echelles at large telescopes to obtain deep and high-resolution spectra needed to study faint lines.

The use of Integral Field Unit (IFU) and multi-object spectrographs (MOS) allows the spatial study of ionized nebulae, often at the expense of a poorer spectral resolution. Several long-slit spectrographs are also MOS and some instruments may operate in both modes (MOS and IFU). A few examples of abundance studies based on IFU or MOS data can be found in Magrini & Gonçalves (2009), Stasinska et al. (2013), Kehrig et al. (2016), and Zinchenko et al. (2016). Some examples of MOS and IFU spectrographs are GMOS at Gemini telescope; FORS and VIMOS at VLT; and PMAS at the CAHA 3.5 m telescope. With the arrival of instruments such as MEGARA (Gil de Paz et al. 2016) in GTC, a high-resolution IFU and MO spectrograph, it will be possible to study ionized nebulae with a high spectral and spatial resolution.

### 3.2. Calibrations

Spectral emission lines allow the determination of physical conditions and ionic abundances in ionized nebulae. Before using spectroscopic data, one should remove the effect of the instrument and the atmosphere and perform some calibrations. This is called data reduction and it depends on the type of spectra (long-slit, echelle, multi-fiber, and integral field unit) and on the instrument and telescope. It is beyond the scope of this tutorial to explain in detail the whole process of data reduction, and we refer the reader to the web pages of the observatories (such as GTC, ESO, and GEMINI) to find more information. One of the main packages used to reduce astronomical data is IRAF and there are many manuals available explaining how to use it.

In brief, the main steps in data reduction are bias and dark subtraction, cosmic rays removal, flat-field correction, wavelength, and flux calibration. The bias is the number of counts in the detector pixels for zero second exposures. The dark current is the number of counts in the pixels when no light is falling in the detector and it is caused by the movements of electrons in the electronics. Both signals need to be subtracted from all images. The flat-field image allows the correction of pixel-to-pixel variations due to differences in the sensitivity. Cosmic rays are high-energy particles that arrive into the detectors constantly and need to be removed, which is easy when there are various exposures of the same field. Wavelength calibration consists in the transformation of pixel scale into wavelength scale and it is done by using comparison spectra of lamps (such as Th-Ar, Ne lamps). And finally, flux calibration transforms the number of counts into intensities and for this, standard stars are required. Besides, one would like to trim the spectra, combine them into one final spectrum, then correct from: the Earth movement around the Sun, the interstellar reddening, and the earth atmospheric lines if possible.

The observation of the continuum and the emission lines of gaseous nebulae need to be corrected for interstellar extinction due to interstellar dust. We refer the reader to the discussion of this issue in Chapter 7 of Osterbrock & Ferland (2006).

### 3.3. Continuum Emission and Underlying Absorption in Nebular Spectra

The emission line spectra are superimposed on top of a weak continuum that has to be subtracted from the observations to recover the emission line intensities. The weak continuum is

---

2. http://www.gtc.iac.es/
3. http://www.eso.org
4. http://www.gemini.edu
5. http://iraf.noao.edu/docs/spectra.html
due to bound-free and free–free emission, two photon emission, dust scattered light, and starlight. For galactic objects it is possible to avoid the central star in PNe and the brightest stars in galactic H II regions, but not for extragalactic H II regions and PNe. Starlight will contain many permitted lines in absorption; of particular interest are the H i lines since they are used to calibrate most of the objects and can be significantly affected by underlying absorption. The errors produced by the underlying component can be minimized by studying objects with the highest equivalent width of H β in emission.

We consider that the best procedure to correct for underlying absorption in H II regions is to use the models by González-Delgado et al. (1999, 2005). According to these models, for young objects the EWβ(H β) is expected to be less than 2.5 Å, and at the same time the EWem(H β) is expected to be more than ~150; on the other hand, for older objects, EWβ(H β) is expected to be larger, while EWem(H β) is expected to be smaller. The correction for underlying absorption for objects with EWem(H β) > 150 Å is inversely proportional to EWem(H β), while for objects with EWem(H β) ≤ 150 Å the correction, and consequently the associated error, increases even faster due to the larger EWβ(H β) predicted by the models.

### 3.4. Atomic Data

The first step, when calculating the physical conditions and ionic abundances is to select the set of atomic data for your calculations. A different selection will translate into a different result. A compilation of some of the most used atomic data was provided by Mendoza (1983) and other recent compilations can be found on the CHIANTI6 and NIST7 web pages. Atomic physicists make a great effort providing us accurate atomic data, it is fair to give them the credit by citing the original papers where the atomic data are published.

Juan de Dios & Rodríguez (2017) recently discuss the impact of different sets of atomic data on the determination of the chemical abundances of O, N, S, Ne, Cl, and Ar.

The package PyNeb contains several sources of atomic data so that the user can choose their preferred ones. There is also a default set of atomic data recommended by the developers of PyNeb. The software IRAF does not allow a simple change of the atomic data, but the adopted sources can be checked. The C17 version of the photoionization code CLOUDY (Ferland et al. 2013) allows for an easier treatment of the atomic data because the files have been moved to external databases.

### 3.5. Determination of Physical Conditions

The spectra of ionized nebulae show bright emission lines that allow us to determine the physical conditions of the gas: electron temperatures (T_e) and densities (n_e).

The electron configurations with a complete sub-shell only have one specific configuration, thus one available level; ground configurations with a sub-shell with only one electron (as well as those lacking one electron to be complete), while theoretically able to have more than one configuration, have several configurations that are equivalent and thus have the same energy. For these three types of configurations, the first excited state requires at least one electron to move from one electronic shell to another. The amount of energy required for such transitions is (a) too high to be easily accessible via collisions and (b) lies in the ultraviolet. On the other hand, ions with ground-state electronic configurations of s^2p^2, s^2p^3, and s^2p^4 are easily observed (which have their first four excited levels with an energy ~kT, i.e., easily reachable via collisions and with some transitions in the optical range); the study of the intensities of these transitions has long been understood and many codes are available which can model the intensities of such lines as a function of temperature, density, and abundance. In principle, ions with electron configurations of the form s^2d^i (2 ≤ i ≤ 8) can also be studied in a similar manner, but ions with the required number of electrons are less abundant; also, because they have many more energy configurations available, software to model the behavior of such ions is less easy to obtain.

The p^2, p^3, and p^4 ions have line ratios sensitive to the electron temperature. The number of electrons required for those configurations are exactly six, seven, eight, 14, 15, or 16; in principle, ions with 32, 33, and 34 electrons have the same configurations, but the most abundant of these would be Se^{++}, which is five orders of magnitude less abundant than O^{+} and is too faint for its auroral lines to be seen. Some examples of observable ions are N^{+}, O^{+}, O^{++}, Ne^{++}, S^{+}, S^{++}, Cl^{++}, Ar^{++}, and Ar^{+++}.

The p^3 ions have line ratios sensitive to the n_e. The numbers of electrons required for this configurations are exactly seven and 15; in principle, ions with 33 electrons have the same configuration, but the most abundant of these would be Se^{+}, which is even less abundant than Se^{++} (although the relevant lines for traditional density determinations are nebular lines). Some examples of observable ions are: O^{+}, Ne^{+3}, S^{+}, Cl^{++}, and Ar^{+3}.

The Grotrian diagrams for O^{+}, O^{++}, and Ne^{++} obtained with PyNeb are presented in Figures 4–6.

In principle, we can also use d^i levels (2 ≤ i ≤ 8) to calculate physical conditions, but the number of excited energy levels that need to be considered as well as the atomic physical parameters makes it impractical, and only a few of these ions have been studied (e.g., Fe^{++}, Fe^{+3}, and Ni^{++}). Numerical packages used to study them are not widely distributed (PyNeb does contain these ions).
The intensity ratios of some forbidden lines are highly sensitive to $T_e$, thus they are very useful in its calculation. The reason for that is that electrons with very different energy are needed to populate the different ionic levels of an ion through collisions.

For example, the excitation energy of level $^3S_{01}$, from which the [O III] $\lambda$4363 line originates, is 5.35 eV, whereas the excitation energy of level $^1D_{2}$, from which the [O III] $\lambda$5007 line originates, is 2.51 eV (see Figure 4). Therefore, the ratio between these two lines tells us about the temperature of the plasma. We expect the intensity ratio of [O III] $\lambda$4363/$\lambda$5007 to be higher in hotter nebulae.

In Chapter 5 of Osterbrock & Ferland (2006), one can find the analytical expressions for the intensity ratio [O III] $(I(4959) + I(5007))/I(4363)$ and other intensity ratios that are commonly used in the literature to estimate $T_e$, such as [N II] $I(5755)/I(6548) + I(6583))$, [Ne III] $I(3343)/I(3869) + I(3968))$, and [S III] $I(6312)/I(9532) + I(9069))$.

The main disadvantage is that the auroral lines (such as [O III] $\lambda$4363 and [N II] $\lambda$5755) are weak, and if the observations are not deep enough, they are not detected. Alternative methods have been proposed to compute the chemical abundances when the $T_e$ cannot be derived, such as the so-called strong line methods that will be discussed in Section 4.3.

3.5.1. Electron Temperatures from CELs

The intensity ratios of some forbidden lines are highly sensitive to $T_e$, thus they are very useful in its calculation. The reason for that is that electrons with very different energy are needed to populate the different ionic levels of an ion through collisions.

For example, the excitation energy of level $^3S_{01}$, from which the [O III] $\lambda$4363 line originates, is 5.35 eV, whereas the excitation energy of level $^1D_{2}$, from which the [O III] $\lambda$5007 line originates, is 2.51 eV (see Figure 4). Therefore, the ratio

![Figure 4. Grotrian diagram for O$^{++}$ from PyNeb (Luridiana et al. 2015). (A color version of this figure is available in the online journal.)](image1)

![Figure 5. Grotrian diagram for O$^{+}$ from PyNeb (Luridiana et al. 2015). (A color version of this figure is available in the online journal.)](image2)

![Figure 6. Grotrian diagram for Ne$^{++}$ from PyNeb (Luridiana et al. 2015). (A color version of this figure is available in the online journal.)](image3)
errors are at least \( \sim 1000 \text{ K} \) for the best observed objects (e.g., Peimbert et al. 2000, 2007; Izotov et al. 2007) or to try to combine one RL with a CEL (Peimbert & Peimbert 2013); in this last scenario, although the sensitivity of the ratios is similar to the sensitivity of nebular to auroral CEL ratios, the intensity of RLs is frequently smaller than that of auroral lines, thus the observational errors tend to be large.

### 3.5.3. Electron Temperatures from the Balmer Continuum

In a gaseous nebula, it is possible to determine the temperature from the ratio of the Balmer jump to a Balmer line; this is because the intensity of any Balmer line as well as the total energy in the Balmer continuum is proportional to \( n(H) \times T^{-1} \). However, the Balmer continuum becomes wider (in wavelength range) with increasing temperature \( \propto T^{1/2} \); therefore, the height of the Balmer discontinuity is approximately proportional to \( T^{-1.5} \).

To determine the temperature from the Balmer continuum, it is necessary to estimate the underlying Balmer discontinuity in absorption due to the direct star contribution and to the dust scattered light. The effect of the underlying Balmer discontinuity due to the stellar continuum on the determination of the temperature of the nebulae becomes negligible for a radiation field dominated by stars hotter than about 45,000 K (i.e., it is important for all H II regions, but can be ignored for most PNe).

Balmer temperatures have been determined by several groups (e.g., Peimbert & Costero 1969; Liu & Danziger 1993).

### 3.5.4. Temperature Inhomogeneities (\( \tau^2 \))

As previously mentioned, photoionized nebula are frequently assumed to have uniform temperature; yet, photoionization codes already show some temperature variations across each nebula, and the existence of other physical processes suggests the existence of larger variations.

To a second order approximation, we can characterize the temperature structure of a gaseous nebula by two parameters: the average temperature, \( T_0 \), and the root mean square temperature fluctuation, \( \tau \), given by

\[
T_0(X^+) = \frac{\int T_e n_e n(X^+)dV}{\int n_e n(X^+)dV},
\]

and

\[
\tau^2(X^+) = \frac{\int (T_e - T_0(X^+))^2 n_e n(X^+)dV}{T_0(X^+)^2 \int n_e n(X^+)dV},
\]

respectively, where \( n_e \) and \( n(X^+) \) are the electron and the ion densities of the observed emission line and \( V \) is the observed volume Peimbert (1967).

To determine \( T_0 \) and \( \tau^2 \), we need two different measurements of \( T_e \): one that preferentially weights the high temperature regions and one that preferentially weights the low-temperature regions of the observed volume. For example, the temperature derived from the ratio of the \([O III] \lambda\lambda 4363, 5007\) lines, \( T_e(4363/5007) \), and the temperature derived from the ratio of the Balmer continuum to \( R(H/\beta) \), \( T_e(Bac/H/\beta) \); these temperatures are related to \( T_0 \) and \( \tau^2 \) by

\[
T_e(4363/5007) = T_0(O^{++}) \left[ 1 + \frac{1}{2} \left( \frac{91300}{T_0(O^{++})} - 2.68 \right)^2 \right]
\]

and

\[
T_e(Bac/H/\beta) = T_0(H^+)(1 - 1.70\tau^2),
\]

respectively (Peimbert et al. 2014). These two equations are very good approximations to \( T_0 \) and \( \tau^2 \), when terms of higher order in \( \tau \) can be neglected, which is when \( \tau^2 \ll 1.00 \).

It is also possible to determine a temperature from the intensity ratio of a collisionally excited line of an element \( p + 1 \) times ionized to a recombination line of the same element \( p \) times ionized; for example, the ratio of \([O III] \lambda 5007\) to the RLs of multiplet 1 of \( O II \), \( T_e(O II/[O III]) \). This ratio is independent of the element abundance and depends only on the electron temperature. By combining \( T_e(O II/[O III]) \) with a temperature determined from the ratio of two CELs, like \( T_e(4363/5007) \), it is also possible to derive \( T_0 \) and \( \tau^2 \) (e.g., Peña-Guerrero et al. 2012b; Peimbert & Peimbert 2013).

Another temperature, \( T_e \) (He I), can be obtained from the intensity of many pairs of \( He I \) RLs, because each line has a slightly different temperature dependence (e.g., Peimbert et al. 2000; Izotov et al. 2007; Porter et al. 2013).

Most of the H II regions observed in other galaxies are very bright and have most of their oxygen in the \( O^{++} \) stage and most of its helium in the \( He^+ \) stage. Therefore, the \( \tau^2 \) values derived from \( T_e(4363/5007) \) together with any of the following temperatures: \( T_e(Bac/H/\beta) \), \( T_e(O II/[O III]) \), or \( T_e(He I) \) are representative of the whole object.

According to its definition, \( \tau^2 \) can be determined even when the material is not chemically homogeneous (as is expected in some PNe). However, under such circumstances, the \( \tau^2 \) of each element can be completely different; moreover, when any element is not well mixed with hydrogen, it is not possible to define observationally the total abundances.

The net effect of \( \tau^2 \) on the temperatures is that temperatures derived from CELs are larger than \( T_0 \), which in turn is larger than temperatures derived from RLs (temperatures derived from the ratio of a CEL to a RL tend to be similar to \( T_0 \)). For CELs, this effect is largest at small \( T_e \), while for RLs this effect is larger for large \( T_e \).

### 3.5.5. Local Electron Densities from CELs

The \( n_e \) can be derived from the intensity ratio of CELs of the same ion originated from levels with nearly the same excitation energy. Thus, the intensity ratio does not depend on \( T_e \), but it...
does depend on the ratio of the collision strengths. If the involved lines have different transition probabilities or different collisional de-excitation rates, their intensity ratio will strongly depend on \( n_e \).

Some of the intensity ratios used to derive \( n_e \) are \([\text{O II}] \lambda 3726/\lambda 3729, [\text{S II}] \lambda 6717/\lambda 6731, [\text{Cl III}] \lambda 5518/\lambda 5538, \) and \([\text{Ar IV}] \lambda 4711/\lambda 4740\). Each intensity ratio is valid in a specific range of densities. In the low-density regime the intensity ratio is proportional to the collisionally excitation ratio, whereas in the high-density regimen \( n_e \) above the critical density, collisions dominate and the intensity ratio is proportional to the spontaneous transition probability ratio. Chapter 5 of Osterbrock & Ferland (2006) explains further details about the determination of \( n_e \).

In general, \( T_e \) and \( n_e \) from forbidden lines are computed together as they depend on each other. Software such as PyNeb (Luridiana et al. 2015), Neat (Wesson et al. 2012), and the routine TEMDEN from IRAF (Shaw & Dufour 1995) calculate the physical conditions from emission-line spectra.

It is important to mention that \( T_e \) and \( n_e \) computed from one particular intensity ratio are representative of the zone of the nebula where the involved intensity lines are emitted. When calculating ionic abundances, one has to make assumptions on the temperature and density structure of the nebula. For example, if \( T_e \) and \( n_e \) are homogeneous throughout the entire nebula or if there are different regions in the nebula with different representative physical conditions.

### 3.5.6. Local Electron Densities from RLs

There are at least two ways to determine densities from RLs: the metastable \( 2^3S \) level of \( \text{He I} \) or the density dependence of the lines of the multiplet 1 of \( \text{O II} \). Due to the specific characteristics of He, no other ion with the same characteristics is expected to be abundant enough to allow us to use the same technique; on the other hand, other \( \text{O II} \) multiplets, as well as multiplets of many other ions, should have similar density dependencies as the multiplet 1 of \( \text{O II} \); however, because they will be fainter, none of them has been studied so far.

When \( \text{He}^+ \) recombines, the new electron can be either parallel or anti-parallel to the old one, creating two families of excited \( \text{He}^0 \), while the anti-parallel family will quickly reach the ground state \( (1S) \); the helium that recombines with parallel electrons will arrive at the metastable \( 2^3S \) level and can easily remain there for a fraction of an hour. The energy structure is such that collisions with helium in the ground state cannot excite the atom, while collisions with helium in the metastable level can; this will affect each \( \text{He I} \) line differently and its efficiency depends on the density, which makes it possible to use this effect to determine the density. Detailed calculations of the atomic physics of \( \text{He}^+ \) recombination can be found in Porter et al. (2013).

The line ratios of the multiplet 1 of \( \text{O II} \) depend strongly on \( n_e \) (Peimbert et al. 2005a); this dependence arises in the \( \text{O}^{++} \) before its recombination. At high densities, \( \text{O}^{++} \) is expected to be distributed in a 3:2:1 ratio between its three lowest energy states, while at low densities all the ions are expected to be at the ground-energy level; even after recombination, these signatures are not completely erased (Bastin & Storey 2006).

The expected ratios for the lines of multiplet 1 of \( \text{O II} \) as a function of \( n_e \) can be found in Peimbert & Peimbert (2005b) and Storey et al. (2017).

#### 3.5.7. Root Mean Square Densities and Filling Factor

In their pioneering study of \( n_e \) in nebulae, Seaton & Osterbrock (1957) noted that density inhomogeneities produced a disagreement between densities derived from forbidden line ratios and those derived from \( \text{H}(\beta) \) surface brightness; an example quoted by them was NGC 7027.

Osterbrock & Flather (1959) studied in detail the density distribution in the Orion nebula and showed that the densities derived from radio fluxes were considerably smaller than those derived from the \([\text{O II}] \lambda 3726/\lambda 3729 \) line intensity ratio. They suggested a model in which only a fraction \( \epsilon = 0.03 \) of the nebula was filled with high-density material and the rest was empty. Since then, \( \epsilon \) has been called the filling factor and is defined by

\[
\epsilon = \frac{n_e^2(\text{rms})}{n_e^2(\text{CEL})},
\]

where \( n_e(\text{FL}) \) is the density determined from forbidden line ratios and \( n_e(\text{rms}) \) is the root mean square density determined from a Balmer line or from the radio continuum flux.

In the presence of a filling factor the mass of a nebula is given by

\[
M(\epsilon) = \epsilon^{1/2} M(\text{rms}).
\]

For various epsilon determinations of PNe see Torres-Peimbert & Peimbert (1977) and Mallik & Peimbert (1988) and references therein.

#### 3.5.8. Other Density Distributions

The filling factor is only an approximation. Real objects will not have \( \sim 97\% \) of their volume with perfect vacuum while the other \( \sim 3\% \) has a constant density; such a model can only be considered a first approximation to reality. With the widespread presence of ADFs, models within this approximation seem incapable of reproducing the observations. As long as the densities are below the critical density for all CELs and the material is well mixed, reality can be approximated by a filling factor plus constant density.

However, if there are regions with density above the critical density for nebular lines, the intensity of such lines will no longer be proportional to density squared; to ignore this will
cause us to measure an unexpectedly low-intensity nebular line, which will make us (a) underestimate the abundance of such an ion, (b) overestimate the temperature, and (c) underestimate the abundance (again) when using such a high temperature. Also, when enough nebular and infrared lines of different ions get suppressed, the cooling will be affected and a temperature structure will emerge.

If the gas is not well mixed and there is a density structure correlated with the chemical composition traditional line ratio analyses no longer make sense; instead, models of the expected line intensities have to be made. Many “two-phase” models are explored in the literature (e.g., Ercolano et al. 2003; Yuan et al. 2011). In general, these models include a semi-large hot “normal” phase, small droplets of a cold “hydrogen-poor” phase, and a large vacuum volume. Of course, these phases are required when neither phase is close enough to reproduce the observations, so the phases will be very different and the average (sum) of their emissions will be very different from the emission of a phase with average physical conditions; yet it is unlikely that in real objects all of the droplets will have the exact same density and chemical composition. Also, there is bound to be a transition layer between the droplets and the surrounding medium (as well as between the hot phase and the vacuum), so that when required, these two-phase models should be considered only as an approximation to reality.

3.6. Determination of Ionic Abundances

As in most other ISM environments, in photoionized regions, hydrogen is the most abundant element, representing \( \sim 90\% \) of the available atoms. These regions are considered ionized if most of the atoms are ionized, therefore in most photoionized regions, \( H^+ \) is the most abundant ion comprising \( \sim 90\% \) of the available ions. For this reason, most of the time, abundances are measured relative to \( n(H^+) \), as \( n(X^{++})/n(H^+) \) or as \( 12 + \log(n(X^{++})/n(H^+)) \). For simplicity, from now on we will use the following notation: \( X^{++}/H^+ \) and \( 12 + \log(X^{++}/H^+) \).

A side effect of this definition is that there are no pure CEL abundance determinations, as all abundances are made with respect to \( H/\beta \) (or some other hydrogen line).

There is an additional subtlety regarding ionic abundance determinations: when determining abundances, we make sure to use the best available density for \( X^{++} \) (the ion we are trying to determine), but we make use of the same density to represent \( H^+ \) when we know that most of the time the representative abundance for \( H^+ \) can be substantially different; we will even use a different density for \( H^+ \) when determining the abundance for a different ion. A simple example of this can be studied by considering an object with \( O/H = 0.001 \), where \( 50\% \) of the oxygen is singly ionized and the other \( 50\% \) is twice ionized; let us further assume that the representative density for \( O^+ \) is \( n_e(O^+) = 100 \, \text{cm}^{-3} \) and for \( O^{++} \) is \( n_e(O^{++}) = 400 \, \text{cm}^{-3} \).

Under these assumptions, \( O^+ \) occupies \( 80\% \) of the volume, while \( O^{++} \) occupies the other \( 20\% \); on the other hand, if we weigh by emission measure, we find that the \( O^+ \) region produces only \( 20\% \) of the intensity of \( H/\beta \), while the \( O^{++} \) region produces the other \( 80\% \). In this example, to obtain \( O^+/H^+ = O^{++}/H^+ = 0.0005 \), the density we should use for \( H^+ \) is \( n_e(H^+) = 200 \, \text{cm}^{-3} \). Instead, to determine \( O^+/H^+ \), we assume that \( n_e(O^+) = n_e(H^+) = 100 \, \text{cm}^{-3} \); when doing this we find that \( O^+/H^+ = 0.0002 \) (because at such density, a lot more \( H^+ \) is needed to produce the intensities observed in the high-density region). At the same time, to determine \( O^{++}/H^+ \), we assume \( n_e(O^{++}) = n_e(H^+) = 400 \, \text{cm}^{-3} \); when doing this, we find that \( O^{++}/H^+ = 0.0008 \) (now a lot less \( H^+ \) is needed to produce the intensities observed in the low-density region). Although the final result when determining total abundances is correct, \( O/H = 0.001; \) one should be careful of the true meaning of measuring \( O^+/O^+ = 4.0 \).

What the previous example shows is that we implicitly weigh our observations by emission measure; not only is this the most natural way of weighing observations, it has additional advantages because most physical processes in photoionized regions are proportional to density squared (e.g., emission, recombination, absorption, the rate at which ionizing photons are consumed, collisions, etc.), and at the end of the day, if the material is well mixed, it gives us the correct total abundance. However, there are still many processes that do not depend on the density squared (e.g., the degree of ionization partially depends in density, computer models cannot ignore density). In particular, in the presence of chemical inhomogeneities, we are not able to reproduce the total abundance with such ease; instead, detailed (complex) models are required.

3.6.1. From RLs

Although to some extent, the terms “permitted lines” and “recombination lines” are often used interchangeably, one must take care not to confuse fluorescent lines (which will tend to be permitted lines) with RLs.

To a first approximation, the intensities of most RLs are proportional to \( n_e \times n(X^{++}) \) and inversely proportional to \( T_e \); this remains so through more than an order of magnitude in \( T_e \), as well as for densities of \( 10^{10} \, \text{cm}^{-3} \) or even more (although there are a few important departures from such proportionality). This is because the number of interactions between ions and electrons is proportional to the density of both, because the electron capture efficiency increases for slower electrons and because the emission occurs very fast, so very high densities are required for the ions to be collisionally de-excited. Also, because they are proportional to the ionic abundance, the only intense RLs are the lines originating in the most abundant ions. Of them, the most important ones are the \( H_1 \) lines, without which it is not possible to determine abundances (the expected emissivities of \( H_1 \) lines can be found in Storey & Hummer 1995).
for most objects the atomic data to be utilized is the one for case B.

Other than hydrogen, helium lines are the RLs most utilized for abundance determinations, as helium is about 100 times more abundant than the third most abundant element; also, for temperatures below 50,000 K, it is very inefficient to collisionally excite either He$^0$ or He$^+$. He$^0$ has many more lines than H$^0$ because it does not have the degeneracy on the energy levels present (e.g., there are six He I lines that represent transitions of electrons from quantum number $n = 4$ to $n = 2$, while for H I, there is only one such line). However, one must be careful when choosing among these lines and select those that are relatively intense but not affected by collisions (e.g., He I $\lambda$6678 or 4921). The presence of He$^{++}$ produces a different set of problems, as many He II lines can be blended with the H I lines, and one must be sure the intensity that one measures at 4861Å truly corresponds to H$\beta$ (the intensity of He II $\lambda$5412 is a good indicator of contamination by He II lines to H$\alpha$ and H$\beta$); the abundance of He$^{++}$ is best determined from He II $\lambda$4686, which is the strongest He II optical line.

Recombination lines of heavier elements are much fainter, thus there are few spectra with enough signal-to-noise (S/N) to accurately determine heavy element abundance; some of the most famous are the C II doublet $\lambda\lambda4267$ and the O II octuplet $\lambda\lambda\lambda4650$. C II $\lambda\lambda$4267 has been important for many decades because not only is it the brightest of the optical RLs from heavy elements, but also because no carbon CELs are available in the visible part of the spectrum; C II $\lambda$4267 is the only way to determine carbon abundance. On the other hand, the O II $\lambda$4650 multiplet has been important because its lines are the brightest RLs from an ion with available CEL data, and as such, they represent the best way to measure an ADF. However, for a few objects, recombination lines of more than two dozen ions have been observed.

One big advantage of determining abundances from RLs, is that all have very similar dependencies with $T_e$ and $n_e$ as such, any error in the measured physical parameters, as well as the presence of a large $\tau$, does not affect the determinations. The main disadvantage is that unless one is observing a relatively bright gaseous nebula with very large telescopes, there are only a handful of ions for which RLs are available (of course, in the presence of chemical inhomogeneities, RL determinations become meaningless, but CEL determinations become meaningless too).

### 3.6.2. From CELs

Most of the emission lines emitted by ionized nebulae are optically thin, thus their intensities are proportional to the abundance of the ion that emits the line. The intensity of CELs highly depends on $T_e$ and therefore a reliable estimation of the physical conditions is needed to obtain reliable ionic abundances. The general expression of an emission line has been presented in Equation (3). The ratio between a CEL and H$\beta$ is

$$\frac{I(\lambda)}{I(H\beta)} = \frac{X^{+i} \nu_{\lambda} A_{\text{em}} f_i}{H^+ \nu_{H\beta} n_e \epsilon_{H\beta}^e},$$

where $f_i$ is the fraction of ions $X^{+i}$ in the upper level, $k$. Then, the ionic abundance can be expressed as

$$X^{+i} = \frac{I(\lambda)}{I(H\beta)} \epsilon_{H\beta}^e,$$

where the emissivities are derived by solving the statistical equilibrium equations to obtain the level populations.

When deriving ionic abundances, one has to adopt the adequate $T_e$ and $n_e$ for each ionic species. One simple assumption is that the whole nebula can be described with one $T_e$ and one $n_e$. One can also calculate each ionic abundance with different $T_e$ and $n_e$ values. The abundances derived with each approach may be very different (not to mention the differences associated to the adopted atomic data set).

### 3.7. Abundance Discrepancy Factors (ADFs)

In many objects, it is possible to measure abundances of the same ion, from both RLs and CELs; they never agree (unless the error bars are very large). It is systematically found that abundances derived from RLs are higher than those derived from CELs. The ADF was defined by Tsamis et al. (2003) as

$$ADF(X^{+i}) = \frac{X^{+i}\text{RLs}}{X^{+i}\text{CEls}},$$

and it is overwhelmingly determined to be greater than one.

A note of warning: because abundances are usually presented in a logarithmic scale, in some works the ADF is defined also in a logarithmic scale where

$$ADF^* = \log(ADF(X^{+i})) = \log\left(\frac{X^{+i}\text{RLs}}{H^+}\right) - \log\left(\frac{X^{+i}\text{CEls}}{H^+}\right);$$

when this convention is used, the values are usually in the $0.2 < ADF^* < 0.8$ range. We recommend the use of the definition of Equation (17).

Typical H II regions present $1.5 \lesssim ADF \lesssim 3$, while typical PNe present $2 \lesssim ADF \lesssim 5$. Chemically inhomogeneous PNe show ADF values in the 10–80 range, with some PNe showing knots, in their inner regions, with ADF values reaching as high as 800 (Corradi et al. 2015).

The presence of ADFs (different than 1.0) show that the simplest models, where the chemistry, temperature, and density are homogeneous (or even models where the chemistry is homogeneous, and there are two zones defined by their ionization degree) are not adequate to represent real photoionized regions.
Another important point about ADFs: in all objects where CEL and RL abundances can be simultaneously determined, ADFs have proven to be ubiquitous, and until their origin is not well understood; all CEL abundances are suspect (for most objects, RL abundances are probably a good approximation to the true abundances; yet for very few objects, CEL abundances are probably a good approximation (see Section 7)). Thus, for objects where only CEL abundances are available, these values should be considered a lower limit to the true abundances; therefore, a correction should be made (e.g. Peña-Guerrero et al. 2012a). Moreover, the corrections available at present are only crude corrections and much work should be done before having high-quality corrections at our disposal.

In Section 7, we will discuss the physical processes which may be responsible for the observed ADFs.

4. Calculation of Total Abundances

There are different approaches to compute chemical abundances of PNe and HII regions. One may compute the total abundances by adding up all the ionic abundances of each element. This can be done with RLs or CELs. Because there are significant differences between both ionic abundances, it is important not to mix them in the derivation of total abundances. When some of the ions are not observed, ionization correction factors (ICFs) need to be used. This is the more direct method to compute chemical abundances in ionized nebulae. As we mentioned above, the abundances from CELs and RLs are different, causing an ADF. One may use the $r^2$ formalism to solve the discrepancy. One may also compute total abundances without calculating first the ionic abundances. One option is fitting a photoionization model to the observations. Another is to use the so-called strong line methods. The following subsections explain in detail each of the aforementioned methods.

4.1. Adding Ionic Abundances

The total abundance of one particular element is given by the sum of the ionic abundances of all the ions present in a nebula. If we can measure emission lines of all these ions, then the calculation of the total abundance is straightforward:

\[
\frac{X}{H} = \sum \frac{X^{+i}}{H^+}.
\]

In general, the available observations cover only a particular wavelength range (optical, ultraviolet, or infrared) or some of the lines are too faint to be measured, thus the total abundances must be calculated only from a few ions. In this case, ionization correction factors (ICFs) must be used to take into account the contribution of those ions for which we cannot derive their ionic abundance:

\[
\frac{X}{H} = \sum_{\text{obs}} \frac{X^{+i}}{H^+} \times \text{ICF}.
\]

The first ICFs were derived according to similarities between ionization potentials (IP) of different ions (e.g., Peimbert & Costero 1969; Peimbert & Torres-Peimbert 1971). For example, the widely used relations N/O = N$^+$ / O$^+$ and Ne/O = Ne$^{++}$/O$^{++}$ are based on the similarities between the IP of N$^+$ and O$^+$ (29.6 eV and 35 eV, respectively) and Ne$^{++}$ and O$^{++}$ (63.4 eV and 54.9 eV, respectively). However, these simple relations have proved to be inadequate because do not take into account all of the physics involved in the ionized gas (for example, charge exchange reactions).

ICFs derived from photoionization models are more reliable because, in principle, they take into account all the physics involved in photoionization. The first important compilation of ICFs based on photoionization models (about a dozen models) is the one by Kingsburgh & Barlow (1994), but some of them have been improved considerably based on nets of more complex photoionization models. New ICFs derived in the last years from large grids of photoionization models are mentioned in Section 6.4.

Software such as PyNeb (Luridiana et al. 2015), NEAT (Wesson et al. 2012), and the routine TEMDEN from IRAF (Shaw & Dufour 1995) calculate the physical conditions and ionic abundances from emission line spectra. In addition, PyNeb contains several ICFs that can be used to determine total element abundances.

4.2. Fitting Photoionization Models

In some cases, there are no available ICFs to compute the total abundances (e.g., for fluorine, phosphorous, and germanium). In others, the available ICFs are not valid. For example, some ICFs are not adequate for objects with very low or very high degree of ionization. In these situations, the chemical composition of ionized nebulae can only be obtained by constructing a tailored photoionization model.

Photoionization models are the theoretical representations of real nebulae. They are computed with specific codes that include all of the physics involved in photoionized models, solve the ionization and energy equilibrium equations and calculate the radiation transfer. The most popular and used photoionization code is CLOUDY developed by a large group of people led by G. Ferland (Ferland et al. 2013). The library PyCloudy (Morisset 2014) can be used to compute pseudo-3D CLOUDY models. Other commonly used photoionization codes are MAPPINGS (Dopita et al. 2013) and MOCASSIN (Ercolano et al. 2003).

A detailed recent review on photoionization models of PNe has been recently published by Morisset (2016). A previous paper on this topic is Ferland (2003).
To compute a photoionization model, one should provide some information about the ionizing source and the ionized gas: the shape and intensity of the radiation; the chemical composition of the gas and dust in the nebula; and the geometry of the cloud. The output consists of the several physical quantities (such as the ionization structure and electron temperature at every position in the nebula) and the line intensities. The predicted intensities are then compared with the observations. In principle, a good agreement indicates that the input parameters are consistent with the observations. If the agreement is not reasonable, the input parameters need to be changed until the observations and predictions agree.

It is important to mention that high-quality spectra are needed to obtain a good photoionization model. Adjusting only a few observables (even if those are all one has) should not give one confidence in the model being created. The more constraints one has, the better the computed photoionization model will be. Along with the emission lines, it is crucial to have an idea of the morphology of a given nebula and its density structure. Some examples of tailored models constructed in the last years are IC 418 (Morisset & Georgiev 2009), TS01 (Stasińska et al. 2010), and NGC 6302 (Wright et al. 2011).

One sticking point from such models is that, when observations of RLs and CELs originating from the same ion are available, they are unable to reproduce the observed ADFs. Yet, for objects where the explanation for the observed ADF is a chemically homogeneous \( r^2 \), photoionization models still have advantages over the direct method: since the intensity of [O III] \( \lambda5007 \) is less affected by \( r^2 \) than the [O III] 5007/4363 ratio, models that adjust nebular lines will be less affected than calculations that use auroral to nebular ratios to determine abundances.

### 4.2.1. Databases

In addition to tailored models, one may compute grids of photoionization models covering a wide range of physical parameters to study general behaviors of ionized nebulae.

The Mexican Million Models database (3MdB, Morisset et al. 2015) contains several grids of models computed with CLOUDY: more than half a million of photoionization models for PNe (Delgado-Inglada et al. 2014), \(~40\)000 models of diffuse ionized gas (Flores-Fajardo et al. 2011), \(~85\),000 models based on CALIFA observations (Morisset 2016), and \(~30\),000 models of giant H II regions that have been used to derive oxygen and nitrogen abundances (Vale Asari et al. 2016). All of these models are available for the community and can be used to explore many open issues.

### 4.3. Strong Lines Methods

The last of the methods relies only in measuring the intensities of nebular lines, frequently normalized to H\( \beta \). These methods are not designed to determine abundances of individual elements, but rather, by assuming that there is a proportionality in the abundances of all heavy elements, they are designed to determine the overall abundance of heavy elements. These methods look for lines or line combinations whose overall intensity depends clearly in the total metallicity, but only marginally in other aspects such as degree of ionization or density.

The most famous of the strong line methods was proposed by Pagel et al. (1979) and is commonly known as Pagel’s method or \( R_{23} \), where

\[
R_{23} = \frac{I(5007) + I(4959) + I(3727)}{I(H\beta)}. \tag{21}
\]

One particular problem of \( R_{23} \) is that it is bivalued, with \( R_{23} \) reaching a maximum of \(~10\) when \( Z \sim 1/3 Z_\odot \), and decreasing for both larger and smaller metallicities; another problem, common to most methods, is that it has a (weak) dependence on ionization degree.

Besides \( R_{23} \), there are at least a dozen other strong line methods such as [O III]/H\( \beta \) (Aller 1942), [N II]/H\( \alpha \) (Storchi-Bergmann et al. 1994), [O II]/H\( \alpha \) (Thompson & Djorgovski 1991), [N II]/[O II] and [N II]/[S II] (Jensen et al. 1976), [O III]/[N II] (Alloin et al. 1979), [S II]/H\( \alpha \) (Denicolò et al. 2002), ([S II]+[S III])/H\( \alpha \) (Vilchez & Esteban 1996), ([S II]+[S III]+[S IV])/H\( \alpha \) (Oey & Shields 2000), ([O II]+[Ne III])/H\( \gamma \) (Pérez-Montero et al. 2007), [Ne III]/[O II] (Nagao et al. 2006), and [Ar III]/[O III] and [S III]/[O III] (Stasińska 2006). Many of them were selected to avoid the bivaluation of \( R_{23} \) or to avoid the necessity of the very blue [O II] \( \lambda3727 \) (or to require only a small part of the spectrum). However, they depend on fainter lines than \( R_{23} \) and tend to have larger scatter on their calibrations.

Another characteristic of the strong line methods is that they have to be calibrated by a different method. As such, any determination made by the strong line methods will have the uncertainties associated with the scattering within the particular strong line method, as well as with the biases and scattering associated with the method used for calibration. Specifically, it does not provide a solution to the ADF problem, but rather the ADF has to be properly understood so that any strong line method can be properly calibrated before accurate abundance determinations can be made.

As it stands today, different calibrations of Pagel’s method vary by nearly an order of magnitude (Kewley & Ellison 2008); however, most of this scatter disappears when one uses a calibration that also considers the degree of ionization, to make a more robust determination. Yet, even after that, determinations calibrated using CELs and the direct method (e.g., Pilyugin & Thuan 2005) differ from those using RLs or CELs + \( r^2 \) (e.g., Peña-Guerrero et al. 2012a) by a factor of about two.
4.4. Element Depletions

The abundances derived from emission lines following the steps described before provide information about the chemical composition of the ionized gas, but elements can be also deposited in dust grains. The underabundance found for many elements with respect to a reference value is generally interpreted as due to depletion onto dust grains. The depletion factor of an element $X$ is given by

$$\frac{X}{H} = \log(\frac{X}{H})_{\text{gas}} - \log(\frac{X}{H})_{\text{true}},$$

(22)

where $\log(\frac{X}{H})_{\text{gas}}$ is the observed gaseous abundance and $\log(\frac{X}{H})_{\text{true}}$ is the true abundance of the object. For many practical purposes, the true chemical abundances are not known; instead, a “cosmic” (or reference) abundance can be used. The composition of the Sun is often taken as the reference composition, but other abundances can be used, like those of young O- and B-type stars.

Noble gases such as helium, neon, and argon are not expected to be important contributors to the dust; volatile elements like carbon and oxygen are somewhat embedded in grains, and refractory elements such as iron and calcium are mostly deposited in the dust (Whittet 2003). The elements that contribute most to the mass of dust grains are carbon, oxygen, magnesium, silicon, and iron (e.g., Whittet 2010). A review on element depletion for 17 elements has been done by Jenkins (2009).

Generally, element depletions are only used to correct the gaseous abundances and compute total element abundances. For example, to compare the chemical compositions of PNe and H II regions with those obtained in other sets of objects. However, depletions can also be used to gain information about dust formation and evolution in different environments. Some examples of these two types of studies are presented in the next two sections.

5. Recent Results from H II Regions

5.1. Abundance Discrepancy Factor in H II Regions

There are three different methods to determine abundances in H II regions that are frequently used: (a) the strong lines method, which is based on the intensity ratio of the nebular lines (such as those of [O II], [O III], [N II], [S III], and [Ne III]), relative to those of hydrogen; this method is calibrated with a net of photoionized models such as those provided by CLOUDY (Ferland et al. 2013), (b) the direct method, which is based on the ratio of forbidden lines to hydrogen lines, and the temperature derived from the ratio of nebular lines to auroral lines like those of [O III], [N II], [S III], and (c) the RLs method, which is based on the intensities of the RLs of C, O, and helium relative to those of hydrogen, that is almost independent of the $T_e$. The intensity of the nebular lines of [O III] is typically two orders of magnitude higher than that of the auroral lines and three orders of magnitude higher than that of the O II RLs. Therefore, the telescope time to derive the intensity of the required lines increases considerably when going from method (a) to method (c).

There is a long-standing riddle when applying $T_e$ for the determination of abundances relative to hydrogen. One generally derives higher relative abundances from RLs than from the forbidden lines. This is called the “abundance discrepancy factor (ADF) problem”, and a possible origin lies in temperature inhomogeneities, parameterized as $\tilde{T}^2$ (Peimbert 1967; Peimbert et al. 1993). The presence of temperature inhomogeneities causes the abundances determined from the forbidden lines to be underestimated, accounting for the ADF.

The ADF in gaseous nebulae can be due to inhomogeneities in temperature, density, and chemical composition. Gaseous nebulae indeed show density and temperature inhomogeneities, and some of them also show chemical inhomogeneities. Therefore, these three possible causes for the ADF values should be studied to determine the real abundances in gaseous nebulae. Temperature inhomogeneities are also known as temperature fluctuations or temperature variations in the literature.

5.2. The $\kappa$ Electron Distribution

It was recently proposed that the free electrons in H II regions and PNe might have significant deviations from a Maxwellian velocity distribution due to the presence of super-thermal electrons and that their distribution can be represented by a generalized Lorentzian distribution given by the $\kappa$ formalism (Nicholls et al. 2012, 2013; Dopita et al. 2013). Ferland et al. (2016) have shown that the distance over which heating rates change are much longer than the distance suprathermal electrons can travel and that the timescale to thermalize these electrons is much shorter than the heating or cooling timescales. These estimates imply that suprathermal electrons will have disappeared into the Maxwellian velocity distribution long before they affect the collisionally excited forbidden and RLs, therefore the electron velocity distribution will be closely thermal and the $\kappa$ formalism can be ruled out for these objects.

5.3. Temperature Inhomogeneities in H II Regions

The difference between the O/H values obtained from oxygen forbidden lines and those derived from RLs can be due to the presence of temperature inhomogeneities in H II regions. The importance of the temperature inhomogeneities can be estimated by means of $\tilde{T}^2$, which can be estimated by comparing temperatures derived from two different methods: for example, by (a) comparing the $T$ derived form the ratio of the Balmer continuum to a Balmer line with the temperature derived from the ratio of two forbidden lines, (b) comparing the temperature derived from He I RLs with a temperature derived using forbidden lines, or also (c) comparing the abundances, for the
same ion, derived using RLs with those derived using CELs. From a group of 37 galactic and extragalactic H II regions observed by different authors Peimbert et al. (2012) found $r^2$ values in the 0.019 to 0.120 range with an average value of 0.044 based on the O II RLs and the [O III] CELs. The $r^2$ average value is considerably higher than the Orion nebula value that amounts to 0.028 ± 0.006.

5.4. Chemical Inhomogeneities in H II Regions

In a chemically inhomogeneous medium, CELs are expected to originate mainly in regions that are relatively metal-poor, temperature-high, and density-low, while the RLs are expected to originate mainly in regions that are relatively metal-rich, temperature-low, and density-high. Based on high-quality observations of multiplet V1 of O II and the NLTE atomic computations of O II, Peimbert & Peimbert (2013) study the density and temperature of a sample of H II regions. They find that the signature for oxygen-rich clumps of high density and low temperature is absent in all objects of their sample: one extragalactic and eight Galactic H II regions. The temperatures derived from (1) RLs of O II and (2) RLs of H I together with Balmer continua are lower than those derived from forbidden lines, while the densities derived from RLs of O II are similar or smaller than densities derived from forbidden lines. Electron pressures derived from CELs are about two times larger than those derived from RLs. These results imply that the proper abundances are those derived from RLs and that these nebulae are chemically homogeneous.

5.5. The Orion Nebula

5.5.1. The Gaseous O/H Value in the Orion Nebula

The gaseous O/H value in H II regions can be obtained from CELs or RLs of oxygen relative to RLs of H.

The Orion Nebula is the best observed H II region. Orion is in the plane of the Galaxy, about 400 pc farther away from the nucleus of the Galaxy than the Sun. There have been many O/H gaseous abundance determinations of the Orion nebula based on oxygen forbidden lines. Representative values derived during the last 40 years by different authors, in $12 + \log(O/H)$ units, are: 8.52 (Peimbert & Torres-Peimbert 1977); 8.49 (Osterbrock et al. 1992); 8.51 (Peimbert et al. 1993); 8.47 (Esteban et al. 1998); 8.51 (Deharveng et al. 2000); 8.51 (Esteban et al. 1998, 2004); results that have been nearly constant through four decades and several atomic physics calculations. From these results we estimate that: $12 + \log(O/H) = 8.50 ± 0.02$. Alternatively based on RLs Esteban et al. (1998, 2004) obtain that: $12 + \log(O/H) = 8.72 ± 0.07$ and $8.71 ± 0.03$ respectively.

The difference between both types of determinations is real and an explanation for this difference has to be sought. The abundances derived from the ratio of two RLs are almost independent of the temperature structure, while the abundances derived from the ratio of a forbidden line to a recombination line depend strongly on the temperature structure.

In the presence of temperature inhomogeneities over the observed volume, and adopting the temperatures derived from the ratio of two forbidden lines, the O/H abundances derived from oxygen forbidden lines become smaller than the real ones. While the O/H abundances derived from RLs are almost independent of the temperature structure and therefore are representative of the real O/H values.

The difference between the abundances derived from forbidden lines and those derived from RLs has been called the abundance discrepancy factor (ADF). See Section 6.1 for further discussion on the ADF.

5.5.2. Temperature Inhomogeneities in the Orion Nebula

The difference between the O/H values obtained from oxygen forbidden lines and those derived from RLs can be due to the presence of temperature inhomogeneities in H II regions. The importance of the temperature inhomogeneities can be estimated by the mean squared temperature fluctuation, $r^2$, frequently named the $r^2$ value or $r^2$ parameter, or $r^2$. $r^2$ can be estimated by comparing temperatures derived from two different methods; for example, by (a) comparing the $T_e$ derived from the ratio of the Balmer continuum to a Balmer line with the temperature derived from the ratio of two forbidden lines, (b) comparing the temperature derived from the ratios of a large number of He I RLs, that have slightly different temperature dependence, with the temperature derived from the ratio of two forbidden lines, or (c) comparing the intensities of a recombination, a nebular, and an auroral line originating from the same ion.

5.5.3. The Dust Contribution to the Total O/H Ratio in the Orion Nebula

In this branch of astrophysics, it is relatively easy to determine the ratio of oxygen to hydrogen atoms in the photoionized gas, an obvious followup of this study is to try to compare these determinations with those made in other sets of objects, especially as oxygen amounts to ~50% (by mass) of the heavy elements. Unfortunately, a fraction of the oxygen is trapped by dust.

Recently, Espíritu et al. (2017) found that the fraction of oxygen atoms embedded in dust in the Orion nebula amounts to 0.126 ± 0.024 dex in agreement with the previous value found by Mesa-Delgado et al. (2009).

Accurate abundance corrections due to depletion have only been done for oxygen. While depletion is not important for noble gases (He, Ne, and Ar), depletion of metallic elements (Fe, Ni, Mg, etc.) is so high that studies of the gas phase will not permit us to determine the true abundances of such atoms. Finally, depletions of elements such as C, N, S, or Cl have not
been studied in any detail and a depletion of 10%–40% is to be expected, so comparisons between stellar and nebular abundances of such elements should be taken with a grain of salt.

5.5.4. Comparison of the Total O/H Value in the Orion Nebula with the O/H Value of the B Stars in the Ori OB1 Association

Nieva & Simón-Díaz (2011) have determined the O/H ratio for 13 B stars of the Ori OB1 association and obtain a $12 + \log(O/H) = 8.77 \pm 0.03$ value. This value has to be compared with the log O/H value (gas+dust) in the Orion nebula that amounts to 8.84 ± 0.04. These results are in fair agreement and imply that the total O/H value for the Orion H II region and their associated B stars is about $12 + \log(O/H) = 8.80 \pm 0.05$. A value considerably higher than the O/H gas ratio in the Orion nebula derived from forbidden lines and the assumption of $\chi^2 = 0.00$ that amounts to $12 + \log(O/H) = 8.50$.

5.5.5. Comparison of the Orion Nebula O/H Values with Those of the Solar Vicinity

Asplund et al. (2009) find a protosolar $12 + \log(O/H) = 8.73 \pm 0.05$ value. To compare the solar value with the Orion nebula value, we have to consider the chemical evolution of the Galaxy during the last 4.6 Gyr and the galactocentric distance at which the Sun was formed. Carigi & Peimbert (2011) estimated from a chemical evolution model of the solar vicinity that the Sun originated at a similar galactocentric distance than the one it has now.

The present O/H ratio in the solar vicinity is 0.02 dex higher than in the Orion association because the O/H gradient of the Interstellar Medium (ISM) amounts to 0.044 dex per kpc (Esteban et al. 2005). In the solar neighborhood, according to models of galactic chemical evolution, the O/H ratio during the last 4.6 Gigayears has increased by 0.13 dex (Carigi & Peimbert 2011). By adding these two values to the protosolar value we expect for a recently formed star in the solar vicinity a $12 + \log(O/H) = 8.88$ value. From the Orion nebula, we expect a $12 + \log(O/H) = 8.84$ value for the solar vicinity, in fair agreement with the previous estimate.

There are two other estimates of the O/H value in the ISM that can be made from observations of F and G stars of the solar vicinity. According to Allende-Prieto et al. (2004), the Sun appears deficient by roughly 0.1 dex in O, Si, Ca, Sc, Ti, Y, Ce, Nd, and Eu, compared with its immediate neighbors with similar iron abundances; by assuming that the oxygen abundances of the solar immediate neighbors are more representative of the present-day local ISM than the solar one, and by adding this 0.10 dex difference to the oxygen photospheric solar value by Asplund et al. (2009) that amounts to 8.69, we obtain that the present value of the ISM has to be higher than $12 + \log(O/H) = 8.79$. A similar result is obtained from the data by Bensby & Feltzing (2006), who obtain for the six most O-rich thin-disk F and G dwarfs of the solar vicinity an average O/H value 0.16 dex higher than the solar photospheric value; by assuming their value as representative of the present-day ISM of the solar vicinity, we find $12 + \log(O/H) = 8.85$. Both results are in good agreement with the O/H Orion nebula value derived from the O II RLS method after adding the fraction of oxygen embedded in dust grains.

5.6. The Primordial Helium Abundance, $Y_P$

According to the Big Bang nucleosynthesis (BBN), the primordial abundances of helium and hydrogen by unit mass, $X_P$ and $Y_P$, produced during the first four minutes after the start of the expansion of the universe amount to about 0.25 and 0.75, respectively (e.g., Peimbert 2008; Pagel 2009; Cyburt et al. 2016, and references therein). The best objects to determine the primordial helium and hydrogen abundances are galaxies where star formation has been very scanty and therefore present almost no heavy elements.

A thorough discussion on the open problems related to the primordial light element abundances was presented by Charbonnel et al. (2010). In particular, the subject of the primordial helium abundance was reviewed on a round table discussion (Ferland et al. 2010).

The three most recent determinations of $Y_P$ based on observations of H II regions in metal-poor galaxies are those by Izotov et al. (2014), Aver et al. (2015), and Peimbert et al. (2016). Their results are presented in Tables 1 and 2.

The determination of $Y_P$ based on BBN depends on several input values, like the number of neutrino families $N_F$, and the neutron lifetime $\tau_n$. From the observed $Y_P$ value, assuming BBN and adopting $\tau_n = 880.3$ s (Olive et al. 2014) it is possible to determine $N_F$ (see Table 1). Alternatively, by adopting an $N_F$ of 3.046 (Mangano et al. 2005), it is possible to determine $\tau_n$ (see Table 2). The $Y_P$ values by Aver et al. (2015) and Peimbert et al. (2016) agree with each other, but are in

| $Y_P$ | $N_F$ | $\tau_n$ | $Y_P$ Source |
|-------|-------|---------|--------------|
| 0.2551 ± 0.0022 | 3.58 ± 0.16 | Izotov et al. (2014) |
| 0.2449 ± 0.0040 | 2.91 ± 0.30 | Aver et al. (2015) |
| 0.2446 ± 0.0029 | 2.89 ± 0.22 | Peimbert et al. (2016) |

| $Y_P$ | $\tau_n$ | $Y_P$ Source |
|-------|---------|--------------|
| 0.2551 ± 0.0022 | 921 ± 11 | Izotov et al. (2014) |
| 0.2449 ± 0.0040 | 872 ± 19 | Aver et al. (2015) |
| 0.2446 ± 0.0029 | 870 ± 14 | Peimbert et al. (2016) |
Table 3

| Element | Orion   | 30 Doradus |
|---------|---------|------------|
| O       | 24 ± 5  | 22 ± 5     |
| Mg      | 91 ± 3  | 72 ± 9     |
| Si      | 76 ± 7  | 81±9       |
| Fe      | 97±1    | 92±1       |

disagreement with the value by Izotov et al. (2014) by more than 3σ.

5.7. Dust in H II Regions

To compare the chemical abundances of H II regions with stellar abundances, it is necessary to add to the gaseous abundances; in H II regions, the fraction of the elements embedded in dust grains is expected to be negligible. On the other hand, the fractions of Mg, Si, and Fe, as well as for most true metals, are expected to be so large that the gas phase abundance is not useful to determine the true chemical abundance; between these two extremes, elements like C, N, O, S, and Cl are expected to be mostly in gaseous form, but with a substantial fraction of their atoms trapped in dust grains.

From the study of Galactic and extragalactic H II regions, Rodríguez & Rubin (2005) found that the iron depletion seems to increase at higher metallicities. In Table 3, we present the O, Mg, Si, and Fe depletions derived for the Orion nebula and 30 Doradus by Peimbert & Peimbert (2010). From the study of 78 H II regions, they find that the fraction of Fe embedded in dust grains increases with the total O/H ratio.

Peimbert & Peimbert (2010) based on the depletions of Mg, Si, and Fe atoms, estimated that in H II regions the fraction of oxygen atoms embedded in dust grains increases from about 0.08 dex, for the metal poorest H II regions known, to about 0.12 dex, for metal-rich H II regions. Recently, Espíritu et al. (2017) found that in the Orion nebula 25 ± 4% of the oxygen atoms are embedded in dust.

5.8. Calibration of Pagel’s Method to Determine the O/H Abundances

Pagel et al. (1979) found that the $I([\text{O III}] 4959 + 5007)/I(\text{H}\beta)$ versus $n(\text{O})/n(\text{H})$ diagram varied smoothly with the total O/H abundance ratio and proposed to use it to determine the O/H ratio for objects where the intensity of the [O III] λ4363 line, needed to derive the temperature, is not available. This diagram is now known as the $R_{23}$-O/H diagram. Peña-Guerrero et al. (2012a) calibrated the $R_{23}$-O/H diagram taking into account the fraction of oxygen trapped in dust grains and the presence of temperature inhomogeneities in H II regions. Based on 28 H II regions, they find that these two corrections increase the O/H values by 0.22–0.35 dex relative to the gaseous O/H values assuming constant temperature.

5.9. More on Abundances of Galactic and Extragalactic H II Regions

One of the problems of deriving abundances relative to hydrogen is the dependence on the temperature of the forbidden lines in the visual region. There are three ways to try to diminish the importance of the $T_e$ on the abundance determinations: (a) to derive the abundances from RLs relative to hydrogen, (b) to use CELs in the infrared that depend weakly on the temperature, and (c) to obtain the ratio of two forbidden lines of different elements from collisionally excited levels of similar energy.

5.9.1. Our Galaxy

Esteban et al. (2005) present C/H and O/H abundances for eight H II regions in our galaxy based on RLs and obtain well-defined galactocentric abundance gradients. Carigi et al. (2005) produce a chemical evolution model of the galaxy that fits the observed carbon, nitrogen, and oxygen galactic gradients and find that about half of the carbon in the ISM has been produced by massive stars and half by low- and intermediate-mass stars. Carigi & Peimbert (2011) present chemical evolution models that fit (a) the observed C/H and O/H galactocentric gradients obtained from H II regions; (b) the hydrogen, helium, carbon, and oxygen abundances of M17; (c) the protosolar abundances; and (d) the C/O, O/H, and C/Fe relations derived from solar vicinity stars.

Esteban et al. (2013), based on RLs have determined the abundances of carbon and oxygen for the Galactic H II region NGC 2579. They find that the chemical composition of this H II region is consistent with flattened carbon and oxygen gradients at its galactocentric distance of 12.4 ± 0.7 kpc. They also find that a levelling out of the star formation efficiency about and beyond the isophotal radius of the Galaxy can explain the flattening of the chemical gradients. They also obtain that $r^2 = 0.045 ± 0.007$ for NGC 2579, a value similar to those obtained for other galactic and extragalactic H II regions.

5.9.2. Other Spiral Galaxies

Esteban et al. (2009) derive C/H and O/H abundance gradients based on RLs for M33, M101, and NGC 2403; the C/H gradient is steeper than the O/H gradient. This result is similar to that found in the Milky Way and has important implications for chemical evolution models of spiral galaxies. Zurita & Bressolin (2012) based on the direct method (direct $T_e$-based method) find that the Ne/O, Ar/O, and S/O abundance ratios are consistent with a constant value across the M31 disk. Croxall et al. (2013) present observations of the
[O III] 88 micrometers fine structure line, which is insensitive to temperature. They combine these data with optical observations to obtain oxygen abundances, which are in agreement with estimates that assume that temperature inhomogeneities are present in H II regions.

Berg et al. (2015) based on the direct method obtain abundances for 45 H II regions in the spiral galaxy NGC 628, the S/O, Ne/O, and Ar/O are constant across the galaxy. Toribio San Cipriano et al. (2016) from RLs of O II and C II in H II regions derive C/H and O/H gradients for NGC 300 and M33, the C/H gradients are steeper than those of O/H leading to negative C/O galactocentric gradients. Bresolin et al. (2016) find that oxygen RLs yield nebular abundances that agree with stellar abundances for high-metallicity systems, but find evidence that in more metal-poor environments the oxygen RLs tend to be higher than those derived from stellar abundances. Croxall et al. (2016) present Large Binocular Telescope observations for 109 H II regions in M101; for 74 of them they determine abundances by the direct method. They study different ICFs in the literature and find a group of H II regions with low Ne/O ratios that defies explanation.

Toribio San Cipriano et al. (2017) have studied five H II regions in the Large Magellanic Cloud and four in the Small Magellanic Cloud. They derive the O/H, C/H, and C/O based on RLs. They find that the LMC seems to show a similar chemical evolution to that of the external zones of small spiral galaxies and that the SMC behaves as a typical star-forming dwarf galaxy. By comparing the nebulae and B-type stellar abundances, they find that they agree better with the nebular ones derived from CELs. Comparing the results with other galaxies, they find that the stellar abundances seem to agree better with the nebular abundances derived from CELs, in low metallicity environments, and from RLs in high metallicity environments. This behavior suggests that either (a) there are two processes creating the ADFs: one affecting RLs (mostly at low metallicities) and one affecting CELs (mostly at high metallicities) or (b) there are problems in the abundance determinations of B-type stars (either at low- or high-metallicities).

5.9.3. Irregular Galaxies

Esteban et al. (2014) have studied H II regions in star-forming dwarf irregular galaxies and found that the H II regions occupy a different locus in the C/O versus O/H diagram than those belonging to the inner disks of spiral galaxies, indicating their different chemical evolution histories, and that the bulk of carbon in the most metal-poor extragalactic H II regions should have the same origin than in halo stars. The comparison between the C/O ratios in these H II regions and stars of the Galactic thick and thin disks seems to give support to arguments about the merging scenario for the origin of the Galactic thick disk.

Berg et al. (2016), based on Hubble Space Telescope observations of 12 H II regions in low-metallicity dwarf galaxies, determined the C/O abundance ratio from the UV C III λλ1906+1909 and the O III] λ1666 lines ratio; this ratio is almost independent of dust absorption and the adopted T_e. They find that there is no increase of C/O with O/H at low metallicities; at higher O/H values, there seems to be a general increasing trend of C/O versus O/H. They also find that the C/N ratio appears to be constant over a large range in the O/H abundance.

6. Recent Results from Planetary Nebulae

6.1. Abundance Discrepancy Factor in PNe

The ADF in PNe can be due to inhomogeneities in temperature, density, and chemical composition. Gaseous nebulae indeed show density and temperature inhomogeneities (also known as fluctuations or variations) and some of them also show chemical inhomogeneities. Therefore, these three possible causes for the ADF values should be studied to determine the real abundances in a given PN.

6.2. Density Inhomogeneities in PNe

The effect of electron densities on the abundance determinations of gaseous nebulae can also mimic spurious ADF values. Rubin (1989), Viegas & Clegg (1994), and Tsamis et al. (2011) have studied the dependence of the line intensities on density when the upper energy levels producing forbidden lines are depopulated. The critical density for collisional de-excitation is different depending on the line of a given ion. If this effect is not taken into account, the temperatures derived from CELs are overestimated and the abundances are underestimated. This can be the case for high-density gaseous nebulae and for certain ions. This effect is particularly relevant when infrared lines are used to determine abundances and for objects of relatively high density. We do not expect this effect to be important for the determination of the O^+^+/H^+^ ratio, as the λ4363 and λ4959 [O III] lines have critical densities of 2.4 x 10^5 cm^-3 and 6.4 x 10^5 cm^-3, respectively, values that are considerably higher than the densities of the typical H II regions and PNe.

6.3. Temperature and Chemical Inhomogeneities in PNe

Georgiev et al. (2008) found that the central star of NGC 6543 is not H-poor and has a normal helium composition. Moreover, they also found that the T_e determined by five different methods indicates that the difference in nebular abundances between the RLs and CELs can be explained as due to temperature inhomogeneities in a chemically homogeneous medium. These results are in disagreement with those by Wesson & Liu (2004), who propose that NGC 6543
contains high-density H-poor inclusions that are rich in helium and heavy elements.

Liu et al. (2006) argue that the ADF in PNe is mostly due to chemical inhomogeneities, while Peimbert & Peimbert (2006) argue that the ADF in most planetary nebulae is due to temperature inhomogeneities in chemically homogeneous objects.

It has been suggested that the ADF problem can be due to the destruction of solid bodies inside PNe that produce cool and high-metallicity pockets (Henney & Stasińska 2010).

Liu et al. (2000) concluded that temperature fluctuations in NGC 6153 cannot account for its ADF. McNabb et al. (2016) studied three objects with very large ADF values and conclude that NGC 6153 is chemically inhomogeneous, while Peimbert et al. (2014) present results for 16 PNe and conclude that NGC 6153 is chemically homogeneous. The difference between such conclusions arises because McNabb et al. find very low recombination temperatures (∼3000 K) while Peimbert et al. find moderately low recombination temperatures (∼6000 K) and moderately low recombination densities (∼5500 cm⁻³), which is incompatible with the presence of high-density/low-temperature blobs.

Recently, Storey et al. (2017) computed the recombination coefficients for O⁺⁺ + e⁻ using an intermediate coupling treatment that fully accounts for the dependence of the distribution of population among the ground levels of O⁺⁺ on n_e and T_e. Based on these atomic data, it is possible to obtain from high-quality observations of the O II recombination lines the N(O⁺⁺)/N(H⁺), n_e and T_e values. By combining these results with the values derived from forbidden lines of [O II], it is possible to study the presence of temperature, density, and abundance inhomogeneities. Storey et al. present results for densities and temperatures based on the O II lines available in the literature and show that high ADF PNe present considerably lower temperatures than low ADF PNe and H II regions.

Further discussion on temperature and chemical inhomogeneities is presented in Section 7.

6.4. Ionization Correction Factors for PNe

In the last few years, new ICFs based on grids of CLOUDY photoionization models have been computed for He, C, N, O, Ne, S, Cl, and Ar (Delgado-Inglada et al. 2014), Zn (Smith et al. 2014), Se and Kr (Sterling et al. 2015), and Ni (Delgado-Inglada et al. 2016).

6.5. Element Production by PN Progenitors

The progenitor stars of PNe may go through several nucleosynthesis processes during their lives that alter their initial chemical composition. The mechanisms that take place and their efficiencies depend on the initial mass of the star and the metallicity; for example, they are generally more efficient at lower metallicities. In brief, all of the AGB stars produce some helium (³He) and nitrogen (¹⁴N), and destroy some carbon (¹²C) via the first dredge-up (FDU). The second dredge-up (SDU) takes place in stars with masses above 3.5–4 M☉ (depending on the metallicity and the model) and its major effect is a significant increase in the helium abundance. The third dredge-up (TDU) may occur in stars with masses above 1–1.5 M☉ and it is responsible for the carbon stars. Thermal pulses and dredge-ups bring to the surface carbon and s-process elements. Finally, the hot bottom burning (HBB) occurs in the most massive PN progenitors, those with masses above 3.5–4 M☉ (those stars where the SDU occurs). The net result of the HBB is a significant increase of nitrogen and the destruction of carbon (in the most massive AGB stars oxygen may also be destroyed). A thorough discussion on nucleosynthesis processes can be found in Karakas & Lattanzio (2014).

The comparison between PN abundances computed from observations and the theoretical predictions from stellar nucleosynthesis models is necessary to (a) understand the nucleosynthesis that occurs in low- and intermediate-mass stars and (b) their impact on the chemical enrichments of galaxies (see, e.g., García-Hernández & Górrn 2014; Delgado-Inglada et al. 2015; García-Hernández et al. 2016a, 2016b; García-Rojas et al. 2016b).

Traditionally, the PNe with high abundances of helium and nitrogen are associated with high-mass progenitors; these PNe are called Type I PNe according to the classification made by Peimbert (1978). However, recent results have shown that extra-mixing processes such as stellar rotation, thermohaline mixing, magnetic fields and other mechanisms could explain helium and nitrogen large abundances in low-mass stars that probably have not gone through the SDU and the HBB (Stasińska et al. 2013; Karakas & Lugano 2016).

Delgado-Inglada et al. (2015) found evidences of oxygen production in some PN progenitor stars. An oxygen enrichment of ∼0.3 dex is present in some Galactic PNe that arise from stars with masses ∼1.5 M☉ formed at sub-solar metallicity. This result can be explained as a consequence of oxygen being produced through the nuclear burning of helium and then dredged-up to the stellar surface via the TDU. The efficiency of this mechanism depends on the mass of the star and the metallicity, and it varies from model to model. Only the stellar models that include an extra mixing process, such as those computed by Ventura et al. (2013) and Pignatari et al. (2016), are able to reproduce the observations.

The abundances of several n-capture element abundances (such as P, F, Ge, Se, Br, Kr, Rb, Ca, and Xe) have been recently computed using high-resolution and deep spectra with new ICFs (Otsuka et al. 2011; Otsuka & Tajitsu 2013; García-Rojas et al. 2015; Sterling et al. 2015, 2016). Also a clear correlation has been found between the C/O values in PNe and the n-capture element enrichment; this indicates that n-capture elements and carbon are brought together to the surface in the same mechanism, the TDU (see, e.g., Sterling et al. 2015).
Moreover, it has been found that stars with different masses are enriched in a different amount in n-capture elements (see, e.g., Sterling et al. 2015).

Some interesting recent reviews on chemical abundances of PNe were made by Kwitter & Henry (2012); Magrini et al. (2012); Delgado-Inglada (2016); García-Rojas (2016).

6.6. Dust in PNe

In our Galaxy, AGB stars dominate dust production over SNe and, in general, these stars are one of the most efficient sources of dust in galaxies (Whitet 2003). As the leftovers of many AGB stars, PNe are adequate places to study the life cycle of dust grains in the circumstellar gas. The presence of dust in PNe is evidenced by the strong infrared excess produced by the emission from dust grains (Balick 1978; Kwok 1980). In addition, the gaseous abundances of refractory elements (such as Mg, Al, Si, Ca, Fe, and Ni) in PNe are much lower than the solar values and this is generally attributed as due to their depletion onto dust grains (see, e.g., Delgado-Inglada et al. 2009, and references therein).

The depletion factors of iron and nickel in Galactic PNe have been recently studied by Delgado-Inglada & Rodríguez (2014) and Delgado-Inglada et al. (2016) using optical spectra and ICFs. They found that in most of the PNe, up to \( \sim 90\% \) of the Ni and Fe atoms are condensed into dust grains. Delgado-Inglada et al. (2016) also found that (a) iron atoms tend to be more deposited in grains in PNe with C-rich dust than in PNe containing O-rich dust and (b) nickel atoms tend to be more attached to the grains than iron atoms in environments with a higher depletion.

As we mentioned in Section 5.7, the iron depletion in H II regions increases as the metallicity increases (Rodríguez & Rubin 2005; Peimbert & Peimbert 2010). To perform a similar analysis with PNe, deep spectra with high resolution of PNe at low metallicities are needed.

Another approach to study the dust present in PNe is through their infrared dust features. Traditionally, PNe are believed to have either carbon-rich dust or oxygen-rich dust. The grains are formed in the cool and dense atmospheres of AGB stars. One expects that stars with C/O > 1 in their envelopes would form carbon-rich grains, whereas those with C/O < 1 would form oxygen-rich grains. The value of C/O in the surface of the star depends on the nucleosynthesis that has occurred during the AGB phase. PNe arising from oxygen-rich AGB stars are expected to have oxygen-rich dust (such as oxides and silicates) and those coming from carbon-rich AGB stars would contain molecules associated with carbon-rich chemistry (such as SiC, MgS, and polycyclic aromatic hydrocarbons, PAHs).

Since the finding of the first carbon-rich star containing oxygen-rich dust (Waters et al. 1998), several PNe have been found that show at the same time PAHs (C-rich molecules) and amorphous and/or crystalline silicates (Perea-Calderón et al. 2009; Delgado-Inglada & Rodríguez 2014; García-Hernández & Gómez 2014; Guzmán-Ramírez et al. 2014; Cox et al. 2016; García-Rojas et al. 2017, in preparation). These objects are called PNe with double chemistry dust and the reason for this duality is still a mystery. Some authors suggest that AGB stars may suffer a late thermal pulse that changes the stellar outflow from O-rich into C-rich, making it possible for these stars to produce both oxygen- and carbon-rich dust (see, e.g., Perea-Calderón et al. 2009).

In this scenario, the observed C/O in the PNe is expected to be greater than 1. On the other hand, it has been proposed that the CO molecules may be dissociated by ultraviolet photons in an oxygen-rich environment and the free carbon atoms could be used to form carbon-rich dust (see, e.g., Guzmán-Ramírez et al. 2014). This could explain the observation of PAHs in O-rich PNe (C/O < 1). Further analysis is needed to solve this question.

6.7. Galactic and Extragalactic O/H Gradients from PNe and H II Regions

Stanghellini & Haywood (2010) have determined the O/H gradient for our galaxy using the direct method for PNe of types I, II, and III defined by Peimbert (1978). The ages and initial masses of the PNe progenitors are Type I ages smaller than about 1 Gyr and initial masses \( \gtrsim 2 M_\odot \), Type II ages in the \( \sim 1–5 \) Gyr range and masses in the \( \sim 1.2–2.0 \) \( M_\odot \) range, and Type III ages \( \gtrsim 5 \) Gyr and initial masses \( \lesssim 1.2 \) \( M_\odot \). In Table 4, we present the O/H gradients derived by Stanghellini & Haywood (2010) and we also show the O/H gradient derived for H II regions based on recombination lines by Esteban et al. (2005). H II regions have ages of at most a few million years. Table 4 shows that there is a clear trend in the sense that the older the set of objects the flatter the derived gradient.

In a study of photoionized regions in M81, Stanghellini et al. (2014), based on the direct method, have obtained for H II regions an O/H gradient of \( -0.088 \pm 0.013 \) dex kpc\(^{-1} \) and for PNe an O/H gradient of \( -0.044 \pm 0.007 \) dex kpc\(^{-1} \). These results can be interpreted as being due to a flattening of the interstellar O/H gradient with time. Magrini et al. (2016), based on the direct method, have derived the PNe and H II regions gradients for NGC 300, M 33, M31, and M81. For

| Objects  | \( \Delta \log(O/H)/\Delta R \) | Source |
|----------|-------------------------------|--------|
| PNe all  | \( -0.023 \pm 0.006 \)        | (1)    |
| PNe Type I | \( -0.035 \pm 0.024 \)         | (1)    |
| PNe Type II | \( -0.023 \pm 0.005 \)        | (1)    |
| PNe Type III | \( -0.011 \pm 0.013 \)       | (1)    |
| H II regions | \( -0.044 \pm 0.010 \)    | (2)    |

Note. (1) Stanghellini & Haywood (2010); (2) Esteban et al. (2005).
these spiral galaxies, the PN gradients are flatter or equal to those of H II regions. The O/H interstellar values increase with time. Moreover, M31 and M81 show larger increments with time in the (H II-PN) O/H ratios at a given galactocentric distance than M33 and NGC 300. Magrini et al. (2017) find that NGC 55 shows no trace of radial gradients; they discuss the differences between this galaxy and NGC 300 that does show an O/H gradient.

7. Why are the Observed $r^2$ Values in Gaseous Nebulae so Large?

To determine abundance ratios, $T_e(4363/5007)$ has been used very often under the assumption of constant temperature ($r^2 = 0.00$). In the presence of temperature inhomogeneities, however, the use of $T_e(4363/5007)$ yields oxygen abundances that are smaller than the real ones (e.g., Peimbert 1967; Peimbert & Costero 1969). In general, under the assumption of $r^2 = 0.00$, the abundance ratios derived from the ratio of a collisionally excited line to a hydrogen recombination line are underestimated, while those derived from the ratio of two collisionally excited lines with similar excitation energies or from the ratio of two RLs, are almost independent of $r^2$. Nevertheless, for some applications, like the determination of the primordial helium abundance which is based on the ratio of hydrogen and helium RLs, the errors introduced by adopting a $r^2 = 0.00$ value are small but non-negligible (e.g., Peimbert et al. 2016).

From 20 well-observed PNe, it has been found that $0.024 \leq r^2 \leq 0.128$ with an average value of 0.064 (Peimbert et al. 2014), while from 37 well-observed H II regions, it has been found that $0.019 \leq r^2 \leq 0.120$ with an average value of 0.044 (Peimbert et al. 2012). Moreover, Toribio San Cipriano et al. (2017) from five H II regions in the LMC have found that $0.028 \leq r^2 \leq 0.069$ with an average value of 0.038, and from four H II regions in the SMC found that $0.075 \leq r^2 \leq 0.107$ with an average value of 0.089.

The majority of the best $r^2$ determinations have been obtained from the ratio of the O III to [O III] line intensities in the visual region (see Section 3.5.4.). It is important to corroborate these high $r^2$ values by comparing them with other $r^2$ values based on other methods. The following $r^2$ determinations are in agreement with those derived from the O II/[O III] method: (a) Balmer continuum to H/β ratio temperatures together with forbidden line temperatures (Peimbert 1967; Peimbert & Costero 1969; Liu & Danziger 1993); (b) 4363/5007 temperatures together with 5007/52 micron temperatures (Dinerstein et al. 1985); (c) a high spatial resolution map of the columnar electron temperature in the Orion nebula taken with the Hubble Space Telescope (O’Dell et al. 2003); (d) RLs of O+, O++, C++, and Ne++ together with CELs of these ions (Esteban et al. 2004); and (e) the He I versus [O III] lines method by Peimbert et al. (2012) together with the atomic physics of (Porter et al. 2013) (Peimbert et al. 2016; Espíritu et al. 2017).

On the other hand, typical photoionization models of static H II regions chemically homogeneous and with constant density, like those derived with CLOUDY (Ferland et al. 2013), predict $0.000 \leq r^2 \leq 0.015$ values with a typical value of about 0.004.

The difference between the predicted $r^2$ values based on photoionization models and the observed $r^2$ values is an open problem and at least 10 different possible explanations have been presented in the literature. We will say a few words about them (see also Torres-Peimbert & Peimbert 2003).

7.1. Electron Velocity $\kappa$ Distributions

As mentioned in Section 4.2, the electron velocity distributions in H II regions and PNe are very close to Maxwellian, ruling out the possibility that $\kappa$ distributions can account for the large ADF values observed (Ferland et al. 2016). Therefore, other causes should be found to explain the $r^2$/ADF problem.

Nevertheless, $\kappa$ distributions have been used in many papers on gaseous nebulae, both estimating them from observations or computing atomic data for different $\kappa$ values. The $\kappa$ distributions can be represented by temperature inhomogeneities; the $r^2$ needed can be obtained by the following expression:

$$r^2 = \frac{0.96}{\kappa}.$$  \hspace{1cm} (23)

Therefore, values derived from observations or used in atomic physics computations, in the of $10 \leq \kappa \leq 50$ range, correspond to the observed values of $0.020 \leq r^2 \leq 0.100$ (Peimbert & Peimbert 2013; Ferland et al. 2016).

7.2. Chemical Inhomogeneities

Tsamis et al. (2003), Tsamis & Pequignot (2005), and Stasińska et al. (2007) suggested the presence of metal-rich droplets produced by supernova ejecta as predicted in the scenario of Tenorio-Tagle (1996). These droplets do not get fully mixed with the interstellar medium until they become photoionized in H II regions and they could be responsible for the ADF problem in H II regions. Moreover, these droplets, if present, had to be denser and cooler than the surrounding material in H II regions. Based on high-quality observations of multiplet V1 of O II of 8 galactic H II regions and one extragalactic H II region, Peimbert & Peimbert (2013) find that the signature of oxygen-rich droplets of high density and low temperature is absent, ruling out the possibility of chemical inhomogeneities in these H II regions.

Similarly, Peimbert et al. (2014), based on high-quality observations of multiplet V1 of O II in 16 galactic planetary nebulae, find ADF values in the 1.42 to 9.67 range (including NGC 6153 with an ADF of 9.67). They also find that there is
no sign of low-temperature, high-density droplets, implying that these objects are chemically homogeneous. For chemically inhomogeneous PNe, the relatively metal-rich regions are expected to be cooler than the relatively metal-poor ones.

Alternatively, Corradi et al. (2015), García-Rojas (2016), García-Rojas et al. (2016a, 2017), Jones et al. (2016), McNabb et al. (2016), and Wesson et al. (2017, and references therein) have found a group of PNe that show ADF values in the 10 to 80 range. This group includes Haro 2–33 (Hen 2–283), Fg 1, NGC 6778, NGC 6337, NGC 6337, Pe 1–9, MPA 1759, M1–42, Hf 2–2, Abell 63, Abell 46, and Abell 30. This group corresponds to binary stars that after a first ejection of gas produce a second one with a very small mass but with a very high overabundance of O/H relative to the first ejecta. The ADF in these objects is strongly centrally peaked. They propose that these large ADF values should be explained in the framework of close binary evolution and discuss the possibility that these systems could have gone through a common envelope phase.

Some giant extragalactic H II regions include WR stars and SNe remnants that will produce chemical inhomogeneities, like NGC 2363 (González-Delgado et al. 1994).

### 7.3. Deposition of Mechanical Energy (Shocks and Turbulence)

Peimbert et al. (1991) suggest that the high $r^2$ values observed in PNe and H II regions might be due to the presence of shock waves.

Peimbert et al. (1995) find that $T(C^{+}/C^{++})$, the temperature derived from the C III $\lambda(1906 + 1909)/\lambda(4267)$, ratio is in general considerably smaller than the $T(O^{+}/O^{++})$ temperature derived from the $\lambda(5007)/\lambda(4363)$ ratio. They also find that the objects with highest $T(O^{+}/O^{++}) - T(C^{+}/C^{++})$ values are those that show large velocities and complex velocity fields, and consequently suggest that the deposition of mechanical energy by the stellar winds of the PNe is the main responsible for the temperature differences.

González-Delgado et al. (1994) propose that stellar winds from WR stars play an important role explaining the high $r^2$ values derived for the giant extragalactic H II region NGC 2363.

Stasinska & Schaerer (1999) find that photoionization models of I Zw 18 yield too low [O III] $\lambda\lambda4363/5007$ ratio by about 30%, relative to the observed value. This difference might be due to the presence of shock waves and turbulence. Similar results were obtained by González-Delgado et al. (1994) and Luridiana et al. (1999) for NGC 2363, Luridiana & Peimbert (2001) for NGC 5461, and Relaño et al. (2002) for NGC 346.

O’Dell et al. (2015) discuss the presence of stellar outflows in the Orion Nebula. Outflows can produce shocks and turbulence. Arthur et al. (2016) discuss the presence of turbulence in the Orion Nebula. The Orion Nebula is an H II region ionized mainly by an O7 star, while many of the giant extragalactic H II regions, for example NGC 5471 (Skillman 1985), I Zw 18 (Skillman & Kennicutt 1993), and NGC 2363 (González-Delgado et al. 1994), present evidence of large velocity winds probably produced by Wolf-Rayet stars and supernovae. These powerful outflows might be partly responsible for the higher $r^2$ values derived for these objects than those observed in the Orion Nebula.

### 7.4. Time-dependent Ionization Might be Important for PNe, but Not for H II Regions

When a photoionization front passes through a nebula it heats the gas above the steady-state value, and some time is needed to reach thermal equilibrium. When the stellar ionizing flux decreases, the outer regions of the nebula become isolated from the stellar radiation field and will continue cooling before fully recombining, creating cold partially ionized outer regions (this might be the case in NGC 7009).

### 7.5. Density Inhomogeneities

Extreme density inhomogeneities are present in most PNe, as can be seen from optical images. Density inhomogeneities produce temperature inhomogeneities, because when the density is high enough, CELs can be de-excited collisionally reducing the cooling efficiency of the forbidden lines.

Viegas & Clegg (1994) have discussed this possibility and find that for typical, chemically homogeneous PNe, this effect is of the order of $t^2 = 0.005$ or less. For chemically inhomogeneous PNe, the effect of density inhomogeneities is crucial.

For steady-state photoionization models, moderate density inhomogeneities are not very important; however, for time-dependent processes, the regions of higher density will reach equilibrium sooner than those of lower density.

The relevance of the density inhomogeneities can be estimated from the filling factor (see Equation (14)), which is almost always considerably smaller than one (e.g., Mallik & Peimbert 1988).

### 7.6. Shadowed Regions

Mathis (1976) proposed the ionization by diffuse radiation of shadowed regions, which are those regions ionized by photons not directly coming from the ionizing star. The shadowed regions in PNe are expected to have a lower temperature than that of the directly ionized regions from the central star. For the Helix Nebula, NGC 7293, Cantó et al. (1998) and O’Dell et al. (2005) estimate the temperature of the shadowed regions to be from one-half to two-thirds of that of regions ionized directly by the central star. This process is present, but it is not the main one producing temperature inhomogeneities. For NGC 7293,
the high-density knots cover only about 0.05 of the total solid angle (covering factor) as seen from the central star; this could contribute to $r^2$ with about 0.005–0.010.

### 7.7. Cosmic Rays

Based on the very small upper limit that they obtain for the $I(4686, \text{He}^{++})/I(H\beta)$ ratio, Peimbert & Goldsmith (1972) conclude that cosmic rays are not important in the heating rate of the Orion nebula; they also suggest that the observed $I(4686)/I(H\beta)$ ratio in some extragalactic H II regions might be due to cosmic ray ionization. Giammanco & Beckman (2005) propose the ionization by cosmic rays as a key mechanism acting in giant H II regions.

### 7.8. Spatially Distributed Ionization Sources Inside H II Regions

Ercolano et al. (2007) study the effects of spatially distributed ionization sources on the temperature structure of H II regions. They find that metallicity indicators calibrated by grids of spherically symmetric photoionization models may suffer a systematic bias due to their dependence on the ionization parameter of the system. The O/H values derived by not considering this effect produce underestimates in the 0.1 to 0.3 dex range. These errors are likely to represent the worst case scenario, nevertheless their magnitude and systematic nature do not allow them to be ignored.

### 7.9. Overestimation of the Intensity of Weak Emission Lines

Rola & Pelat (1994) have shown that the intensity of a weak line with a S/N smaller than 5 is overestimated. This effect might be present in the observations of oxygen and carbon RLs that in a given object are typically three orders of magnitude fainter than the brighter RLs of H. This intensity overestimation could also be present in the [O III] 4363 line that is typically two or three orders of magnitude fainter than the [O III] 5007 line. The overestimation of the oxygen RLs produces higher O/H abundances than the real ones, while the overestimation of [O III] 4363 produces higher $T_e$ and lower O/H abundances than the real ones. These two results increase spuriously the ADF values. To avoid this possibility, it is important to measure line intensities with a higher than 5 S/N.

### 7.10. Magnetic Reconnection

García-Segura et al. (2001) propose that the multiple, regularly spaced concentric shells around some PNe could be due to the effects of a solar-like magnetic, cycle with a periodic polarity inversion in the slow wind of an asymptotic giant branch (AGB) star. Presumably these shells of alternating polarity could give rise to magnetic reconnection processes once that they are compressed in the formed PN (i.e., in the swept-up shell). Magnetic fields have been measured in the torus surrounding K3–35 (Miranda et al. 2001); also, the presence of a magnetic field has been inferred in the PN OH0.9 + 1.3 (Zijlstra et al. 1989) and in the pre-planetary nebula IRAS 17150–3224 (Hu et al. 1993).

### 7.11. Summary on $r^2$

From the previous discussion, we consider that the best abundance determinations are those derived from the oxygen and hydrogen RLs and that the $r^2$ values derived from observations should be used to estimate the abundances derived from forbidden lines of the other elements. H II regions have $r^2$ values in the 0.02 to 0.12 range, with a typical value of 0.045, and ADF values in the 1.5 to 3 range with a typical value of 2. We conclude that probably in H II regions, the main source of temperature inhomogeneities is due to the deposition of mechanical energy by shocks and turbulence produced by stellar winds, including those winds due to SNe and WR stars in giant extragalactic H II regions.

We divide PNe in two groups: (a) classical PNe with homogeneous chemical composition and (b) PNe with inhomogeneous chemical composition.

Classical PNe have the following properties: (a) they do not have gaseous inclusions of high density and low temperature; (b) they have $r^2$ values in the 0.024 to 0.128 range with a typical $r^2$ value of 0.065; (c) they have ADF values in the 1.4 to 10.0 range with a typical value of 2.3; and (d) they have typical ionized masses in the 0.05 to 0.40 $M_\odot$ range (e.g., Mallik & Peimbert 1988).

PNe with inhomogeneous chemical composition have the following properties: (a) they do have inclusions of high density and low temperature in the gaseous envelope, possibly the result of close binary evolution; (b) their O/H values decrease from the central star outward; (c) they present large temperature inhomogeneities that cannot be analyzed with the $r^2$ formalism in a one-phase model; (d) their ADF values are in the 10 to 100 range; and (e) they have typical ionized masses of about 0.01 $M_\odot$.

For chemically homogeneous PNe, the best abundance determinations are those derived from oxygen and hydrogen RLs. We also conclude that probably the main source of temperature inhomogeneities in PNe is due to the deposition of mechanical energy from strong stellar winds.

For chemical inhomogeneous PNe, the best approach to determine the chemical abundances is to construct two-phase models, but still a particular kind of $r^2$ should be considered when fitting such models. By their very essence, such models will not have a homogeneous temperature; moreover, with inhomogeneous chemistry, the $r^2$ value can be very different for each chemical element and for each ion. Most of this will be directly considered when studying the two phases of the model; however, the hot phase alone is expected to behave similarly to
a chemically homogeneous PN, and as such it is expected to have its own particular $r^2$ (and ADFs). Because most of the CEL intensities are expected to come from this phase, the $r^2$ corresponding to the hot phase alone is necessary and should be sufficient to reproduce the behavior of all the optical CELs observed in these objects.

8. Final Remarks

The determination of accurate chemical abundances of PNe and H II regions is a fundamental problem in astrophysics. The chemical abundances of PNe provide us with constrains for the formation and destruction of dust in gaseous nebulae. There is still no consensus on its origin and scope. The same can be said for the cosmology and particle physics.

Further efforts are needed to obtain data for H II regions and PNe in galaxies of different types. In spiral galaxies, observations of H II regions at different galactocentric distances to study which has been the evolution of the ISM as a function of time for these galaxies. The study of very metal-poor H II regions can help to study which has been the evolution of the ISM as a function of time for these galaxies. The study of very metal-poor H II regions will permit to increase the accuracy of the primordial helium determination; this value is important for cosmology and particle physics.

The ADF problem should be studied further to obtain a consensus on its origin and scope. The same can be said for the high $r^2$ values derived from observations of H II regions and PNe.

We thank the referee for many excellent suggestions. We are grateful to all those who have worked and are working in this field; they have inspired us over the years. M.P., A.P., and G.D.-I. received partial support from CONACyT grant 241732. M.P. and A.P. acknowledge support from PAPIIT (DGAPA-UNAM) grant no. IN-109716. G.D.-I. acknowledges support from PAPIIT (DGAPA-UNAM) grant No. IA-101517.

References

Allende-Prieto, C., Barklem, P. S., Lambert, D. L., & Kunha, K. 2004, A&A, 420, 183
Aller, L. H. 1942, ApJ, 95, 52
Aller, L. H. 1984, Physics of Thermal Gaseous Nebulae (Dordrecht: D. Reidel Publishing Co.)
Aller, L. H., Baker, J. G., & Menzel, D. H., 1939, ApJ, 89, 587
Alloin, D., Collin-Souffrin, S., Joly, M., & Viguéras, L. 1979, A&A, 78, 200
Arelano-Cordoba, K. Z., Rodríguez, M., Mayya, Y. D., & Rosa-González, D. 2016, MNRAS, 455, 2627
Arthur, S. J., Medina, S.-N. X., & Henney, W. J. 2016, MNRAS, 463, 2864
Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481
Aver, E., Olive, K. A., & Skillman, E. D. 2011, JCAP, 3, 43
Aver, E., Olive, K. A., & Skillman, E. D. 2015, JCAP, 7, 11
Balick, B. 1978, in Proc. IAU Symp. 76, Planetary Nebulae: Observations and Theory (Dordrecht: D. Reidel Publishing Co.), 275
Baker, J. G., Menzel, D. H., & Aller, L. H. 1938, ApJ, 88, 422
Bastin, R. J., & Storey, P. J. 2006, in Proc. IAU Symp. 234, Planetary Nebulae in our Galaxy and Beyond, ed. Michael J. Barlow & Roberto H. Méndez (Cambridge, UK: Cambridge Univ. Press), 369
Benjamin, R. A., Skillman, E. D., & Smits, D. P. 2002, ApJ, 569, 288
Bensby, T., & Feltzing, S. 2006, MNRAS, 367, 1181
Berg, D. A., Skillman, E. D., Garnett, D. R., et al. 2013, ApJ, 775, 128
Berg, D. A., Skillman, E. D., Croxall, K. V., et al. 2015, ApJ, 806, 16
Berg, D. A., Skillman, E. D., Henry, R. B. C., Erb, D. K., & Carigi, L. 2016, ApJ, 827, 126
Bresolin, F., & Kennicutt, R. C. 2015, MNRAS, 454, 3664
Bresolin, F., Kudritzky, R. P., Urbaneja, M. A., et al. 2016, ApJ, 830, 64
Cantó, J., Raga, A., Steffen, W., & Shapiro, P. R. 1998, ApJ, 502, 695
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Carigi, L., & Peimbert, M. 2011, RMxAA, 47, 139
Carigi, L., Peimbert, M., Esteban, C., & García-Rojas, J. 2005, ApJ, 623, 213
Charbonnel, C., Tosi, M., Primas, F., & Ciampi, C. (ed.) 2010, in Proc. IAU Symp. 268, Light Elements in the Universe (Cambridge, UK: Cambridge Univ. Press)
Corradi, R. L. M., García-Rojas, J., Jones, D., & Rodríguez-Gil, P. 2015, ApJ, 809, 99
Cox, N. L. J., Pilleri, P., Berné, O., Cernicharo, J., & Joblin, C. 2016, MNRAS, 456, L89
Croxall, K. V., Pogge, R. W., Beg, D. A., Skillman, E. D., & Moustakas, J. 2016, ApJ, 830, 4
Croxall, K. V., Smith, J. D., Brandl, B. R., et al. 2013, ApJ, 777, 96
Cyburt, R. H., Fields, B. D., Olive, K. A., & Yeh, T.-H. 2016, RevMP, 88, 015004
Deharveng, L., Peña, M., Caplan, J., & Costero, R. 2000, MNRAS, 311, 329
Delgado-Inglada, G. 2016, in IAU Symp. 323: Planetary Nebulae: Multi-wavelength Probes of Stellar and Galactic Evolution, in press (arXiv:1611.10246)
Delgado-Inglada, G., Mes-Delgado, A., García-Rojas, J., Rodríguez, M., & Esteban, C. 2016, MNRAS, 456, 3855
Delgado-Inglada, G., Morisset, C., & Stasińska, G. 2014, MNRAS, 440, 536
Delgado-Inglada, G., & Rodríguez, M. 2014, ApJ, 784, 173
Delgado Inglada, G., Rodríguez, M., Mampaso, A., & Viironen, K. 2009, ApJ, 694, 1335
Delgado-Inglada, G., Rodríguez, M., Peimbert, M., Stasińska, G., & Morisset, C. 2015, MNRAS, 449, 1797
Denicoló, G., Terlevich, R., & Terlevich, E. 2002, MNRAS, 330, 69
Dinerstein, H. L., Lester, D. F., & Werner, M. W. 1985, ApJ, 291, 561
Dopita, M. A., & Sutherland, R. S. 2003, Astrophysics of the Diffuse Universe (Berlin, New York: Springer)
Dopita, M. A., Sutherland, R. S., Nicholls, D. C., Kewley, L. J., & Vogt, F. P. A. 2013, ApJS, 208, 10
Ercolano, B., Barlow, M. J., Storey, P. J., & Liu, X.-W. 2003, MNRAS, 340, 1136
Ercolano, B., Bastian, N., & Stasińska, G. 2007, MNRAS, 379, 945
Esprítu, J. N., Peimbert, A., Delgado-Inglada, G., & Ruiz, M. T. 2017, RMxAA, 53, 95
Esteban, C., Bresolin, F., Peimbert, M., et al. 2009, ApJ, 700, 654
Esteban, C., Carigi, L., Copetti, M. V. F., et al. 2013, MNRAS, 433, 382
Esteban, C., García-Rojas, J., Carigi, L., et al. 2014, MNRAS, 443, 624
Esteban, C., García-Rojas, J., Peimbert, M., et al. 2005, ApJ, 618, 95
Esteban, C., García-Rojas, J., & Pérez-Mesa, V. 2015, MNRAS, 452, 1553
Esteban, C., Mesa-Delgado, A., Morisset, C., & García-Rojas, J. 2016, MNRAS, 460, 4038
Esteban, C., Peimbert, M., García-Rojas, J., et al. 2004, MNRAS, 355, 229
Esteban, C., Peimbert, M., Torres-Peimbert, S., & Escalante, V. 1998, MNRAS, 295, 401
Feinberg, G. J. 1999, PASP, 111, 1524
Feinberg, G. J. 2003, ARA&A, 41, 51
Feinberg, G. J., Henney, W. J., O'Dell, C. R., & Peimbert, M. 2016, RMxAA, 52, 261
Feinberg, G. J., Izotov, Y., Peimbert, A., et al. 2010, in Proc. IAU Symp. 268, Light Elements in the Universe, ed. C. Charbonnel et al. (Cambridge, UK: Cambridge Univ. Press), 163
