Astrophysics, Vol. 51, No. 3, 2008

From Ambartsumian to our days

PLASMA DIAGNOSTICS OF PLANETARY NEBULAE

A. F. Kholtygin

Diagnostics for the rarefied plasmas in gaseous nebulae are reviewed, beginning with the pioneering papers of V. A. Ambartsumian. These papers, as well as the diagnostic techniques which have been developed on the basis of ideas contained in them, are discussed. Diagnostic techniques for homogeneous, as well as inhomogeneous, plasmas are described.

Keywords: *planetary nebulae: interstellar medium; plasma diagnostics*

1. Introduction

Research on gaseous nebulae has always played an enormous role in astrophysics, and continues to do so. Some nebulae can be seen in the sky as extended luminous objects surrounding or close to bright hot stars and are known as diffuse nebulae, of which the best known representative is the Orion nebula. It is now well known that diffuse nebulae are regions of ionized gas surrounding young hot stars of spectral class O and they are often referred to as ionized hydrogen regions (HII regions).

Nebulae of the other-- planetary-- type are compact luminous objects surrounding weak stars, which are the nuclei of planetary nebulae. At present, the number of observed planetary nebulae in our galaxy is approaching two thousand. Many planetary nebulae have also been found in other galaxies, including spiral (M31, M33, etc.) as well as irregular and elliptical galaxies.

V. V. Sobolev Institute of Astronomy, St. Petersburg State University, Russia; e-mail: afk@astro.spbu.ru

Original article submitted April 2, 2008; accepted for publication April 17, 2008. Translated from Astrofizika, Vol. 51, No. 3, pp. 363-383 (August 2008).

 0571-7256/08/5103-0294

The spectra of gaseous nebulae differ significantly from those of stars. They consist of a weak continuum spectrum and a large number of intense emission lines, of which the lines of hydrogen and ionized and neutral helium dominate in the visible range. However, the brightest lines in the spectra of nebulae, with wavelengths λ5007 Å and λ4959 Å were not identified for a long time, since they did not show up in the spectra of elements studied in earthbound laboratories. For this reason they were assumed to belong to an unknown element, nebulium. They were thus denoted by N_1 and N_2 (the ***nebulium 1*** and ***nebulium 2*** lines).

The nature of the nebular emission in these lines was understood only after the development of quantum mechanics. In 1928 the American astrophysicist Bowen showed [1] that the N_1 and N_2 lines are forbidden lines of the O^{2+} ion and are transitions from metastable levels of this ion. He also showed that the bright $\lambda 3726 + 3729$ Å doublet in the spectra of nebulae is associated with forbidden transitions of singly ionized oxygen O^+. This discovery was an indication of the low density of both matter and radiation in nebulae.

Rosseland established that the source of the radiation from nebulae is reprocessed ultraviolet radiation from hot stars located in the nebulae. At the end of the 1920's two basic mechanisms were proposed for the formation of the radiation from nebulae. The first was developed by the American astrophysicist Menzel and the Dutch astronomer Zanstra. This mechanism (recombination) begins with the ionization of atoms by ultraviolet radiation from the star which excites the nebular luminosity (sometimes several stars, in the case of diffuse nebulae). The subsequent recombination of the ions and capture of electrons into excited levels leads to cascade transitions to the ground state of the atoms through intermediate excited states. The low densities of the gas in nebulae means that the chain of cascade transitions is essentially uninterrupted. The cascade radiative transitions mainly generate radiation in the lines of the most abundant elements, hydrogen and helium. This fact served as the basis of Zanstra's method [2] of determining the temperature of the central stars of planetary nebulae based on the intensities of the emission from nebulae in the Balmer series of hydrogen.

The second mechanism, proposed by Bowen, explains the forbidden line emission from nebulae. According to this mechanism, metastable levels of the ions of oxygen, nitrogen, and other elements are excited during collisions with electrons. When the degree of ionization of hydrogen and helium in nebulae is high owing to the UV radiation from the stars which excite them, the nebulae contain many electrons with energies greater than 2-3 eV, which are capable of exciting the metastable levels.

In this paper we examine the situation which held in the beginning of the 1930's, when Ambartsumian's papers on the physics of gaseous nebulae began to appear, and briefly describe the development of that branch of the physics of gaseous nebulae associated with the use of Ambartsumian's methods for analyzing the forbidden line emission from nebular plasmas to determine the physical conditions in these nebulae. Owing to the limited space, we shall deal only with planetary nebulae, although the methods of spectral analysis developed by him also apply to and are used for a much more extensive range of objects.

The content of Ambartsumian's papers on nebulae are discussed briefly in section 2 of this review. The results of his 1933 paper [3] are discussed in section 3 in the light of the current state of research in this area of science. The general techniques for diagnostics of astrophysical plasmas, as well as the results of their use for determining the physical characteristics of planetary nebulae, including inhomogeneous ones, are described section 4.

2. V. A. Ambartsumian's papers on gaseous nebulae

Ambartsumian's papers on the spectra of planetary and diffuse nebulae (all are collected in Volume 1 of his ***Scientific Papers*** [4]) were published in the 1930's. This was a time when theoretical astrophysics was being founded and research on gaseous nebulae was at its leading edge.

His first paper in this area [5] is devoted to determining the temperature of the nuclei of planetary nebulae. He proposed an original method for determining the temperature based on comparing the energy $E_{\mathrm{H}\beta}$ emitted by the nebula in the Hβ line to the energy $E_{\lambda 4686}$ emitted in the HeII λ4686 line. Here it is assumed that all the UV photons emitted by the central star of the nebula, and capable of ionizing hydrogen atoms and helium ions, are absorbed in the nebula. This method, known as ***Ambartsumian's method***, yielded realistic estimates of the temperatures of the central stars that are close to their modern values.

This understanding of the closeness of the mechanisms for formation of the spectra of nebulae and of stars with emission lines in their spectra was typical of Ambartsumian. He used Zanstra's method for determining the temperatures of Wolf-Rayet stars [6], obtaining an estimate (T = 65000 K) for the temperature of the star HD 192163 that is close to the modern value.

Ambartsumian also studied radiative transport in the continuum and resonance lines of nebulae assuming a rectangular profile for the absorption coefficient in the line. He found that radiation pressure in the Lα line plays a decisive role in the expansion of nebulae [7,8]. These papers attracted widespread interest and were the origin of a detailed discussion of the role of radiation pressure in the gas dynamics of nebulae, as it turned out that radiation pressure in the Lα line causes an acceleration of the outer portions of a nebula by on the order of 1 km/s over 10 years, as well as a similar deceleration of the inner portions that has not been confirmed observationally.

The problem was elaborated by Ambartsumian's student, Sobolev, who showed that as photons diffuse out from the line center, they reach the line wings, where the absorption coefficient is small. This effect leads to a reduction in the radiative pressure force by roughly two orders of magnitude at the boundary of a nebula compared to that calculated by Ambartsumian. In a nebula that is expanding with a velocity gradient, the radiative pressure in the Lα line will be lower than for a stationary nebula owing to the Doppler shift of the photons in the spectrum line.

One direct result of the discussion initiated by Ambartsumian's papers was Sobolev's theory of radiative transport in media with large gradients of their large-scale motions [9], which is today's standard apparatus for interpreting the spectra of rarefied astrophysical objects.

In another paper Ambartsumian studied [10] the ionization of the gas in the shells surrounding hot stars. A method was proposed for determining the radial mass concentrations of the luminous gas in a nebula based on the energy emitted in the hydrogen Balmer lines [11].

Ambartsumian's paper [3] on the diagnostics of the rarefied gas in nebulae, which was noted above, is especially important. He studied the process for populating the metastable levels of atoms. Ultimately, he derived a formula relating the temperature of the gas in nebulae to the intensity ratio of forbidden lines observed in their spectra. The method of determining the temperatures of nebulae from the intensity ratios of forbidden lines in their spectra was proposed on the basis of these studies.

From the very beginning it was clear that this method is applicable for a wide range of other objects besides nebulae: the envelopes of nova and supernova stars, diffuse nebulae, the corona of the sun and other stars, and many other objects.

Ambartsumian generalized his work of the 1930's in his book, ***Theoretical Astrophysics*** [12], which became the first textbook in Russian on this subject. The classic course in theoretical astrophysics [13], which was the primary textbook in this discipline for many years, grew out of this text. It was translated into several languages. Many of Ambartsumian's ideas and methods were later included in Sobolev's ***Course in Theoretical Astrophysics*** [14].

In later years, Ambartsumian became involved in other areas of astrophysics; however, the ideas expounded in the papers mentioned above became the basis of many areas of research on planetary and diffuse nebulae [15].

The limited length of this article does not allow us to trace the evolution of all the ideas regarding the physics of gaseous nebulae advanced in Ambartsumian's articles from the 1930's. For this reason, we shall concentrate on analyzing his 1933 paper [3] about the physics of forbidden line emission from nebulae, the content of which was also summarized in his book [12].

3. Atoms and ions in metastable states. Ambartsumian's view.

3.1. The accumulation of atoms in metastable states. Calculations of the populations of excited states of atoms and ions are of decisive importance for determining the intensities of line emission in the spectra of nebulae. Ambartsumian examined [12] the problem of determining the populations of two- and three-level atoms. The solution of the first problem (Fig. 1, left) follows from the time-independent equation

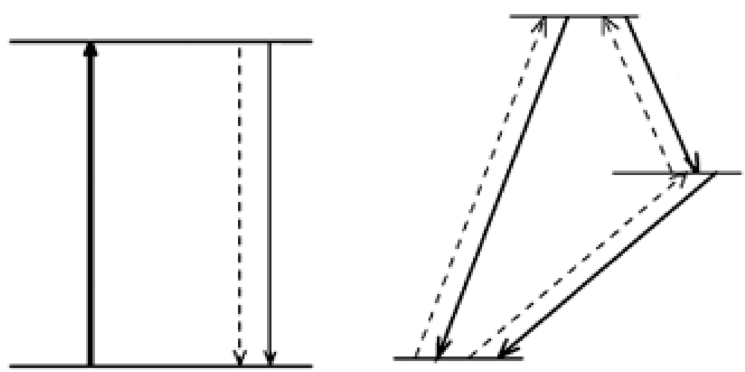

Fig. 1. Scheme for formation of lines in the spectra of planetary nebulae by electron impact excitation. Left: a two-level atom. The thin smooth line represents radiative $2 \rightarrow 1$ transitions; the dashed line-- deexcitation of level 2 by electron impact; the thick smooth line indicates electron-impact excitation of level 2. Right: a three-level atom. The smooth lines indicate transitions from upper to lower levels owing to spontaneous radiative transitions and deexcitation by electron impact; the dashed lines, excitation of the levels by electron impact.

$$b_{12}n_1 - (A_{21} + a_{21})n_2 = 0\,. \tag{1}$$

Here n_k is the population in level k, A_{21} is the Einstein coefficient for the probability of the $2 \to 1$ transition, $n_1 b_{12}$ is the number of atoms undergoing a transition into metastable state 2 owing to collisions per cm^3 per s, and $n_2 a_{21}$ is the number of downward collisional transitions. For the reader's convenience we shall use modern notation for the numbers of excitation and deexcitation events: $b_{12} = n_e q_{12}$ and $a_{21} = n_e q_{21}$, where q_{12} and q_{21} are the rates of collisional excitation and deexcitation for the $2 \to 1$ transition, and n_e is the electron density of the plasma.

Equation (1) determines the ratio of the populations in levels 1 and 2:

$$b_{12}n_1 - (A_{21} + a_{21})n_2 = 0\,. \tag{2}$$

Here T is the kinetic temperature of the plasma, which equals the electron temperature T_e for a nebula. From here on, we follow tradition and speak of the electron temperature of the nebula. An important conclusion, that was reached even before quantum mechanical calculations of the electron impact collisional cross sections and rates were carried out and was deduced in Ref. 3, is that for a low density of gas in a nebula, when $A_{21} \gg n_e q_{21}$, the energy emitted in the forbidden line $2 \to 1$ per second per unit volume of a rarefied plasma with electron density n_e at temperature T_e depends only on the number of electron-impact excitation events; i.e.,

$$E_{21} \approx n_e n_1 q_{12} h\nu_{12}\,. \tag{3}$$

Equation (3) is standard for low density plasma physics and is widely used for estimating the intensities of lines in the spectra of nebulae and other objects.

Ambartsumian also discussed the excitation of metastable levels in a model of a three level atom. In that case, the populations of the levels in the absence of radiative excitation events are determined by the following balance equations (Fig. 1, right):

$$\begin{cases} n_e q_{12} n_1 + (A_{32} + n_e q_{32})n_3 = (A_{21} + n_e q_{21} + n_e q_{23})n_2\,, \\ n_e q_{13} n_1 + n_e q_{23} n_2 = (A_{31} + A_{32} + n_e(q_{31} + q_{32}))n_3\,. \end{cases} \tag{4}$$

The solution of these equations is

$$\begin{cases} \dfrac{n_2}{n_1} = \dfrac{n_e q_{12}(A_{31} + A_{32} + n_e q_{31} + n_e q_{32}) + n_e q_{13}(A_{32} + n_e q_{32})}{(A_{21} + n_e q_{21})(A_{31} + A_{32} + n_e q_{31} + n_e q_{32}) + n_e q_{23}(A_{31} + n_e q_{31})}\,, \\ \dfrac{n_3}{n_1} = \dfrac{n_e q_{12} n_e q_{23} + n_e(A_{21} + n_e q_{21} + n_e q_{23})}{(A_{21} + n_e q_{21})(A_{31} + A_{32} + n_e q_{31} + n_e q_{32}) + n_e q_{23}(A_{31} + n_e q_{31})}\,. \end{cases} \tag{5}$$

Under the rarefied plasma conditions of planetary nebulae, the contribution of the radiative $3 \to 1$ transition and collisional deexcitation to the populations in levels 2 and 3 can be neglected. It then turns out that

$$\begin{cases} E_{21} \approx n_e (q_{12} + q_{13}) n_1 h\nu_{12} , \\ E_{32} \approx n_e q_{13} n_1 h\nu_{23} . \end{cases} \tag{6}$$

Rather than the energy E_{ki} emitted by the nebula in the $k \rightarrow i$ line, in the following we use the ***line intensity*** I_{ki}. I_{ki} is taken to mean the ratio, corrected for interstellar absorption, of the radiative flux in the given line to the radiative flux in the Hβ line.

3.2. Determining the electron temperature from the lines of nebulium

Equation (6) implies that the intensity ratio of the $2 \rightarrow 1$ and $3 \rightarrow 2$ lines is equal to

$$\frac{I_{21}}{I_{32}} = \frac{E_{21}}{E_{32}} = \frac{\nu_{12}}{\nu_{23}} \left(1 + \frac{q_{12}}{q_{13}} \right). \tag{7}$$

In the case of the customarily most intense forbidden lines of doubly ionized oxygen in the spectra of nebulae, the ratio $I_{21}/I_{32} = I(N_1 + N_2)/I([\mathrm{OIII}]\lambda 4363)$ is the ratio of the combined intensity of the nebular lines N_1 and N_2 to the intensity of the [OIII] λ4363 auroral line.

In order to find the intensity ratio (7) it is necessary to be able to calculate the ratio q_{12}/q_{13}. At the end of the 1930's, when Ambartsumian derived Eqs. (2)-(7), it was not possible even to estimate the electron-impact rates of excitation of atoms and ions. Calculating these rates was a very complicated problem. The first realistic calculations of the excitation cross sections for O^{2+} by electron impact were done only in 1940 [16].

To estimate q_{12}/q_{13}, Ambartsumian used the known relationship between the rates of electron impact collisional excitation and deexcitation for the levels and assumed that the ratio $(g_2 q_{21})/(g_3 q_{31})$ is equal to unity. Then

$$\frac{q_{12}}{q_{13}} = \frac{g_2\, q_{21}}{g_3 q_{31}} e^{h\nu_{23}/kT_e} \approx e^{h\nu_{23}/kT_e} , \quad \frac{I_{21}}{I_{32}} = \frac{\nu_{12}}{\nu_{23}} \left(1 + e^{h\nu_{23}/kT_e} \right) \approx 0.9\, e^{33000/T_e} \tag{8}$$

It should be emphasized that for at temperatures typical of nebulae, the dependence of the ratio I_{21}/I_{32} on T_e, which is determined by an exponential factor $\exp(h\nu_{23}/kT_e)$, is very strong. For $T_e = 7000\,\mathrm{K}$ this factor equals 111, while for $T_e = 10000\,\mathrm{K}$ it equals 27, and for $T_e = 15000\,\mathrm{K}$, 9.

Thus, having found the ratio I_{32}/I_{21} from observations, it is possible to make a reliable estimate of the corresponding electron temperature T_e. In this way Ambartsumian obtained an electron temperature of $T_e \approx 7000\,\mathrm{K}$ for nebulae and concluded that the temperature of the electron gas in nebulae is considerably lower than the temperatures of their central stars. This value of T_e is somewhat low from a modern perspective. However, given how little information was available about the collisions of atoms with electrons in the 1930's, the accuracy of this estimate seems amazing.

The reason that Ambartsumian's method still yields reasonable estimates of the temperature even in its original simplified form is the strong (exponential) temperature dependence of the number of electrons capable of exciting

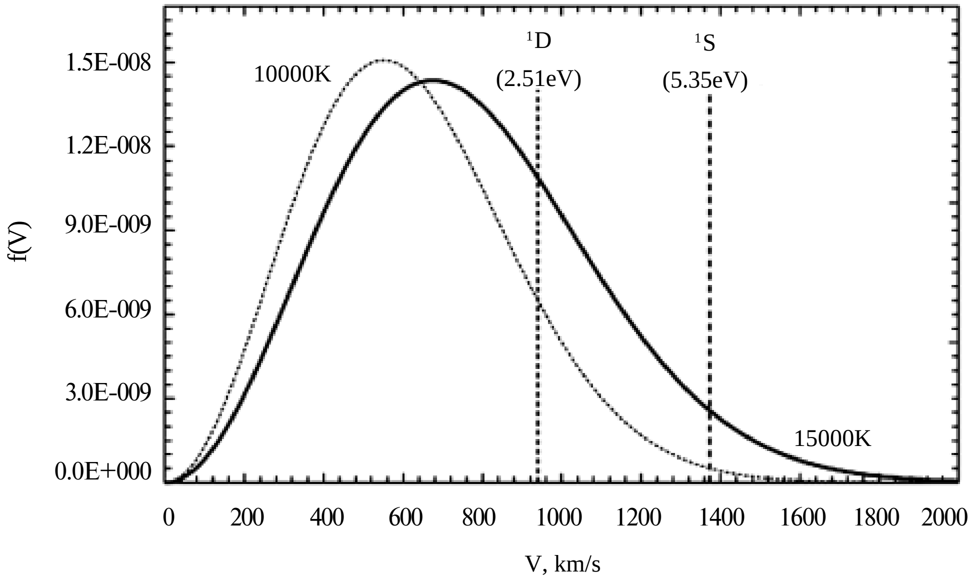

Fig. 2. The electron velocity distribution in an equilibrium plasma with temperature T_e =1 0000 K (dashed curve) and T_e = 15000 K (thick smooth curve). The vertical dashed lines indicate the thresholds for excitation of the ^{1}D and ^{1}S levels of the O^{2+} ion.

the upper levels of these atomic transitions.

As an illustration, Fig. 2 shows the electron velocity distribution function for electron temperatures of 10000 and 15000 K. The thresholds for excitation of the ^1D and ^1S levels of the O^{3+} ion are indicated in this figure. They serve as the upper levels for, respectively, the nebular lines N_1 and N_2, and for the auroral [OIII] $\lambda 4363$ line. It is obvious from the figure that the efficiency of exciting the levels of O^{3+} by electron impact, which is determined by the fractions of electrons with energies above the excitation threshold, must depend strongly on the temperature.

The ratio of the intensities of the auroral [OIII] $\lambda 4363$ line and the nebular [OIII] $\lambda 4959+\lambda 5007$ lines is essentially determined by the ratio of the numbers of electrons with velocities $\upsilon \geq 1370$ km/s and $\upsilon \geq 940$ km/s, respectively. Thus, the intensity ratio of the [OIII] lines makes it possible to determine the temperature dependent ratio of the concentrations of electrons with different energies in the plasma and, thereby, to estimate its electron temperature.

The simple idea of using the relative intensities of lines in the spectra of gaseous nebulae for determining the physical conditions in them was the seed from which grew the diagnostic techniques for rarefied astrophysical plasmas which developed rapidly beginning in the 1940's.

3.3. Collision strengths and improvements in the ratio I_{21}/I_{32}. Using modern data on the rates of atomic processes, we now examine the accuracy of Ambartsumian's original formula for the intensity ratio of the nebular and auroral lines of [OIII].

We return to Eq. (5) for the populations of the metastable levels. In the limit of low electron densities,

$n_e << A_{21}/q_{12}$, the line intensity ratio is

$$\frac{I_{21}}{I_{32}} = \frac{\nu_{12}}{\nu_{23}}\left[1 + \frac{q_{12}}{q_{13}}\left(1 + \frac{A_{31}}{A_{32}}\right)\right]. \tag{9}$$

In order to compare this exact (in the limit of low n_e) formula with the approximate Eq. (8) obtained assuming that $(g_2 q_{21})/(g_3 q_{31}) = 1$ and $A_{31} = 0$, we have to calculate the ratio q_{12}/q_{13} for the typical temperatures T_e = 10000-20000 K of planetary nebulae.

The electron-impact excitation and deexcitation rates q_{ij} and q_{ji} (for $j > i$) are given in terms of the corresponding reaction cross sections $\sigma_{ij}(\upsilon)$ and $\sigma_{ji}(\upsilon)$, which depend on the velocities υ of the electrons colliding with an atom, by

$$q_{ij} = \int_{\upsilon_0}^{\infty} \sigma_{ij}(\upsilon)\upsilon f(\upsilon)\,d\upsilon, \quad q_{ji} = \int_{0}^{\infty} \sigma_{ji}(\upsilon)\,\upsilon\, f(\upsilon)\,d\upsilon, \tag{10}$$

where υ_0 is the minimum (threshold) velocity of electrons capable of exciting the $i \to j$ transition and $f(v)$ is the maxwellian velocity distribution of the electrons. (See Fig. 2.) Modern quantum mechanical calculations show that the cross sections $\sigma_{ij}(\upsilon)$ and $\sigma_{ji}(\upsilon)$ depend very strongly on energy. However, in the case of metastable states, the cross sections averaged over energy have a simpler dependence on υ and are approximately inversely proportional to the electron energy. Hence, they are usually written in the form

$$\sigma_{ij}(\upsilon) = \frac{h^2}{4\pi m^2}\frac{\Omega_{ij}}{g_i \upsilon^2}, \tag{11}$$

where Ω_{ij} is an dimensionless effective cross section (usually on the order of unity) that depends on the electron velocity and is often referred to as the collision strength. The collision strengths for excitation and *deexcitation are equal*, i.e., $\Omega_{ij} = \Omega_{ji}$, while the deexcitation cross section $\sigma_{ij}(\upsilon)$ is given by a formula analogous to Eq. (11) with g_i replaced by g_j. Substituting Eq. (11) in Eq. (10), we obtain

$$q_{ij} = \frac{A_0}{g_i T_e^{1/2}}\langle\Omega_{ij}\rangle e^{-h\nu_{ij}/kT_e}, \quad q_{ji} = \frac{A_0}{g_j T_e^{1/2}}\langle\Omega_{ij}\rangle. \tag{12}$$

Here the dimensional constant $A_0 = 8.6287 \cdot 10^{-6}\,\mathrm{cm^3\,s^{-1}\,K^{1/2}}$ and $\langle\Omega_{ij}\rangle$ is the collision strength averaged over a maxwellian electron velocity distribution, usually referred to as the *effective collision strength.* It is often denoted by $\Upsilon_{ij} = \langle\Omega_{ij}\rangle$.

Using data from the tables of Golovatyi et al. [17], for $T_e = 10^4$ K we calculate the ratio

$$\left(\frac{g_2 q_{21}}{g_3 q_{31}}\right) = \frac{\Upsilon_{12}}{\Upsilon_{13}} \approx 7.9 \,. \tag{13}$$

This ratio has a very weak temperature dependence. For $T_e = 1.2 \cdot 10^4$ K it equals 7.7 and for $T_e = 1.5 \cdot 10^4$ K, 7.6. When the collision strengths obtained by Seaton [14] are used, the ratio in Eq. (13) is equal to 8.9.

Thus, we can see that the ratio (13) differs significantly from the value assumed by Ambartsumian. Nevertheless, because of the exponential dependence of the ratio q_{12}/q_{13} of the excitation rates on the parameter $h\nu_{23}/kT_e \approx 32970/T_e$, where T_e is in degrees Kelvin, the error in the temperature determination is not too large.

Using Eq. (12) for the rates of the collisional transitions and taking the ratio $A_{31}/A_{32} = 0.15$ [17], for $R = I_{21}/I_{32}$ for the O^{2+} ion we obtain

$$R = \frac{I(4959\,\text{Å}) + I(5007\,\text{Å})}{I(4363\,\text{Å})} = 0.9 + 7.9\, e^{32970/T_e} \,. \tag{14}$$

4. Diagnostics of the gas in nebulae

The ideas established in Ambartsumian's analysis of the processes for populating the metastable states of atoms have turned out to be unusually fruitful for solving problems in plasma diagnostics.

The state of plasma diagnostics for gaseous nebulae in the mid 1940's is described in the book *Physical Processes in Gaseous Nebulae*, by Menzel et al., which was translated into Russian [18]. The classics on the physics of nebulae and the diagnostics of nebular plasmas in the 1950-1960's were Aller's book [19] and review [20]. In the 1980's Pottasch's book [21] provided a quite detailed discussion of plasma diagnostic techniques for planetary nebulae and Kwock's book [22], which discusses these same questions in a way accessible even to undergraduates, was published in 2000. The most complete modern discussion of diagnostic techniques for astrophysical plasmas is by Osterbrock, whose book [23] appeared in a second edition in 2006 [24]. Interest in the theoretical analysis of the processes for populating the levels of atoms by electron collisions continues, as evidenced by a just-published article on this topic [25].

4.1. Elementary diagnostic techniques. The elementary techniques are applied to the analysis of the emission from homogeneous or almost homogeneous plasmas. A homogeneous plasma is one with constant values of the temperature and density throughout and, therefore, is characterized by an average electron temperature Te and an average electron density *ne*. In this case, the set of parameters describing the plasma state includes only the relative abundances of the elements present in the plasma, along with the parameters T_e and *ne*. The methods for finding the parameters of a homogeneous plasma, which describe the observed spectra of gaseous nebulae and other objects, are described elsewhere [21,23,24,26,27].

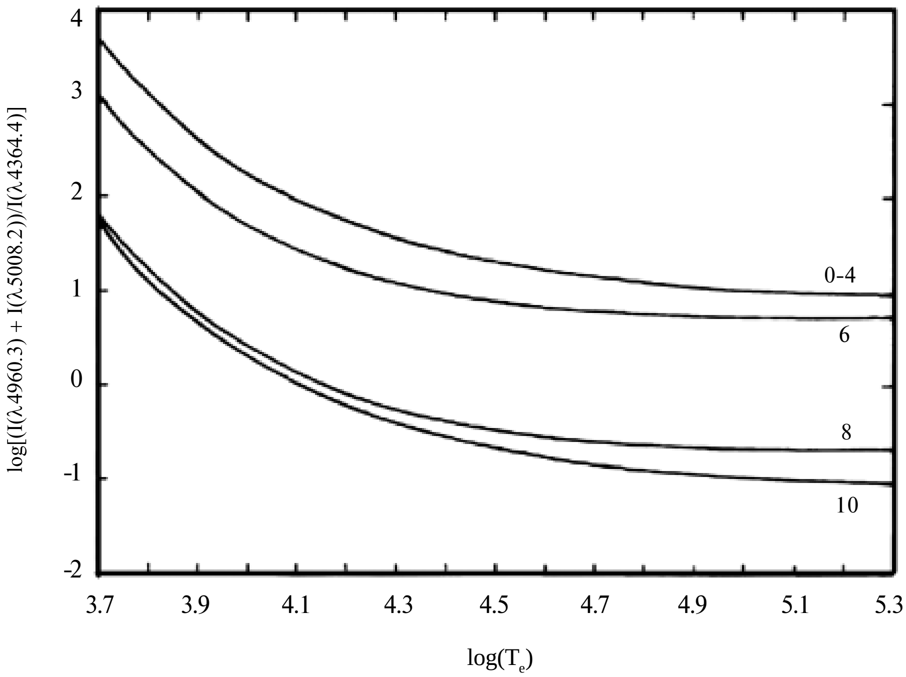

Fig. 3. The [OIII] line intensity ratio $(I(4959\text{Å})+I(5007\text{Å}))/I(4363\text{Å})$ as a function of T_e for different values of *ne* [28]. The logarithm of *ne* is indicated next to the corresponding curves.

4.1.1. Determination of the electron temperature. The lines of [OIII] are most often used for determining the electron temperature of nebulae. In the limit of low electron densities, $n_e < 10^5$, the ratio *R* of the intensities of the $N_1 + N_2$ and 4363 Å lines is given by Eq. (14). The constant in front of the exponential factor equals 7.6 [22]. The small difference from this value of K_0 given above are associated with differences in the atomic constants employed. The constant term 0.9 in Eq. (14) is often omitted.

For higher ne, the role of deexcitation (superelastic collisions) must be taken into account. Figure 3 shows the calculated [28] variation of the ratio *R* with *Te*. It is clear that for the $n_e \le 10^5$ cm^{-3} typical of planetary nebulae, *R* is insensitive to *ne*. The dependence on ne must be taken into account for denser gases.

Line intensity ratios of other elements can be used to determine the temperature. Lines of the C^0, N^+, F^{3+}, and Ne^{4+} ions, which are isoelectronic with O^{2+}, as well as of the Ne^{2+}, Ne^{4+}, Mg^{4+}, Ne^{6+}, etc., ions are used as electron temperature indicators. The electron temperatures of a large number of nebulae have been determined from the intensity ratios of forbidden lines. A large subset of these measurements can be found in Refs. 29-31.

4.1.2. Determination of the electron density. The ratios I_{21}/I_{31} of the intensities of doublets with close upper levels having an energy difference $h\nu_{23}/kT_e << 1$ K are often used for determining the electron densities in astrophysical objects. These include the [NI] $\lambda 5202/5199$, [OII] $\lambda 3729/3726$ and [SII] $\lambda 6731/6716$ lines, as well as lines of other ions corresponding to transitions from the doublet levels $^2D_{5/2}$ and $^2D_{3/2}$ of np^3 configurations with n = 2 – 4.

This method for determining the electron density was not examined by Ambartsumian, but the relevant expressions can easily be derived from equations given in his papers. Because $h\nu_{23}$ is small, the transition probability $A_{32} << A_{31}$, which means that, for small $n_e << A_{21}/q_{21}$, atoms from the excited state 3 mainly undergo transitions to the ground state 1. This implies that excitation of level 3 essentially has no effect on the population of level 2; that is, the populations of levels 2 and 3 can be treated independently, using Eq. (2) to determine their populations.

Then, assuming that the ratio of the transition energies $h\nu_{12}/h\nu_{13} = 1$, the line intensity ratio is given by

$$\frac{I_{21}}{I_{31}} = \frac{(n_2/n_1)h\nu_{12}\,A_{21}}{(n_3/n_1)h\nu_{13}\,A_{31}} = \left(\frac{g_2 q_{21}}{g_3 q_{31}}\right)\frac{1+n_e\,q_{31}/A_{31}}{1+n_e\,q_{21}/A_{21}}. \tag{15}$$

Using Eq. (12) we obtain

$$R = \frac{I_{21}}{I_{31}} \approx R_0\,\frac{1+\left(n_e/n_e^{(1)}\right)/\sqrt{T_e}}{1+\left(n_e/n_e^{(2)}\right)/\sqrt{T_e}}. \tag{16}$$

Here $R_0 = \Upsilon_{12}/\Upsilon_{13}$ is the limiting value of the ratio I_{21}/I_{31} for low *ne*. In the LS coupling approximation, the ratio $\Upsilon_{12}/\Upsilon_{13} = g_2/g_3$. When the J = 5/2 level lies below the J = 3/2 level (e.g., for the O^+ and N^{2+} ions), $R_0 = 6/4 = 1.5$. When the levels are otherwise positioned (for the S^+ ion), $R_0 = 2/3$.

The density $n_e^{(1)} = g_3 A_{31}/(A_0\,\Upsilon_{13})$, while $n_e^{(2)} = g_2 A_{21}/(A_0\,\Upsilon_{12})$. Given the weak dependence of the effective collision strengths on T_e, the densities $n_e^{(1)}$ and $n_e^{(2)}$ can be regarded as insensitive to the electron temperature.

Let $n_e^{min} = \min\left(n_e^{(1)}, n_e^{(2)}\right)$ and $n_e^{max} = \max\left(n_e^{(1)}, n_e^{(2)}\right)$. Then, for $n_e << n_e^{min}$, the ratio $R \approx R_0$ and for $n_e >> n_e^{max}$, $R \approx g_2 A_{21}/g_3 A_{31}$. The latter is obvious, since at high electron densities the populations of the levels follows a Boltzmann law, and the ratio $n_2/n_3 = (g_2/g_3)\exp(h\nu_{23}/kT_e) \approx g_2/g_3$. Within the interval $n_e^{min} < n_e < n_e^{max}$, the ratio R is sensitive to the electron density and can be used to determine it.

Substituting the transition probabilities and collision strengths for the O^+ ion [28] in Eq. (16) and carrying out some simple transformations, we obtain

$$\frac{I_{3729}}{I_{3727}} = 0.42\,\frac{1+207\sqrt{T_e}/n_e}{1+22\sqrt{T_e}/n_e}. \tag{17}$$

For high *ne*, electron impact transitions between levels 2 and 3 become important. These transitions can be included by using the exact solution of the population balance equations (5) for a three-level atom. Then

$$\frac{I_{3729}}{I_{3727}} = 0.42\,\frac{1+57\sqrt{T_e}/n_e}{1+13\sqrt{T_e}/n_e}. \tag{18}$$

In this way we see that including collisional transitions between levels 2 and 3 does not change the form of the doublet line intensity ratio as a function of T_e and *ne*, or its limiting values for very low and very high *ne*. However, the

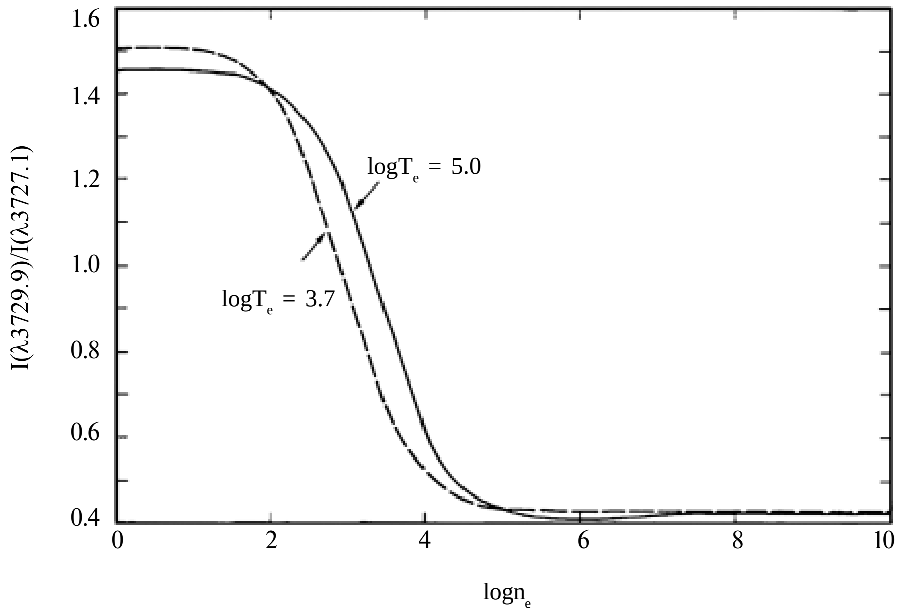

Fig. 4. The [OII] $I(3729\text{Å})/I(3726\text{Å})$ line intensity ratio as a function of ne for different values of Te [28].

coefficients in front of the ratio $\sqrt{T_e}/n_e$ in Eq. (18) do vary significantly. When the collision strengths given in Ref. 14 are used, these coefficients change to 31 and 10, respectively, and the coefficient in front of the fraction on the right hand side of this equation changes to 0.35. This is illustrated in Fig. 4 which shows the calculated [28] line intensity ratio for the forbidden [OII] doublet as a function of ne.

Lines of the O^+ and S^+ ions are mainly used for determining ne in planetary nebulae [29,30,32]. The most complete catalog of electron densities in planetary nebulae is given by Stangellini et al. [33].

4.2. Combined diagnostics of *ne* and T_e. The models of two and three level atoms with which the determination of the plasma parameters of gaseous nebulae began, provide only a crude description of the actual picture of excitation of atomic and ion lines by electron collisions. Even in the well studied case of the excitation of the nebular lines of doubly ionized oxygen O^{2+}, accounting for the fine splitting of the levels of the lower ^3P term (which was regarded as a single level by Ambartsumian and many others) and adding the populations of the metastable $^5S^0$ level of the $2s2p^3$ configuration makes it necessary to examine a six-level atomic model. A more complete description of the spectra of atoms and ions including transitions in the hard ultraviolet and x-ray regions of the spectrum often requires consideration of an atomic model that includes tens of levels. Thus, a model of the Fe^{5+} ion with 80 levels has been discussed [34].

The shift to multilevel atomic models requires the solution of a system with a large number of steady state balance equations. In addition, other processes besides electron collisions, in particular photo- and dielectronic recombination into high levels and cascade transitions involving them, can be involved in filling metastable levels.

When electron collisional and ion recombination transitions are taken into account, the steady state equations which determine the populations of the levels n_k of a given atom or ion X^{m+} can be written in the compact matrix form

$$T\ \boldsymbol{n} = n_e n^+ \alpha, \tag{19}$$

where $\boldsymbol{n}$ is the population vector for the levels of the atom and α is the recombination rate vector. The matrix T has the form

$$T_{ik} = \begin{cases} -(A_{ki} + n_e q_{ki}), & i < k, \\ \sum_{j=1}^{k-1} A_{kj} + n_e \sum_{j \neq k}^{N} q_{kj}, & i = k, \\ -n_e q_{ki}, & i > k. \end{cases} \tag{20}$$

Finding the level populations by solving the system of Eqs. (19) requires knowledge of the densities of the $\mathrm{X}^{(m+1)+}$ ion and, thereby, solving the ionization equilibrium equations for the X atom.

For resonance lines and transitions involving low-excited ion lines of the astrophysically significant elements, the contribution of recombination to the populations of levels excited by electron impact can be neglected and the right hand side of the system of Eqs. (19) can be set equal to zero. In that case, Eqs. (19) only determine the relative populations $n_k / n(\mathrm{X}^{m+})$ of the levels. In order to obtain the absolute magnitudes of the populations, the value of $n(\mathrm{X}^{m+})$, the total density of the X^{m+} ion, must be specified and the system of Eqs. (19) has to be supplemented by the equation

$$\sum_k n_k = n\left(\mathrm{X}^{m+}\right). \tag{21}$$

The increase in the number of equations included in the system of steady state equations leads to an increase in the number of transitions for which line intensity ratios can be calculated. The following ratios have been calculated [35]:

$$R_1 = \frac{I(4363\text{Å})}{I(4959\text{Å}) + I(5007\text{Å})}, \quad R_2 = \frac{I(1661\text{Å}) + I(1667\text{Å})}{I(4959\text{Å}) + I(5007\text{Å})}. \tag{22}$$

Figure 5 shows the range of possible values of the ratios R_1 and R_2 for the electron temperatures and densities typical of planetary nebulae.

When the line intensities of several ion species are found, it is also possible to simultaneously determine *ne* and T_e, as illustrated in Fig. 6. The axes of this plot correspond to *ne* and T_e, while the curves indicate the values of T_e and *ne* for which the intensity ratios of the lines indicated in the figure are equal to the observed line intensity ratios in the spectrum of the planetary nebula NGC7026. If *ne* and T_e are the same throughout the entire emitting

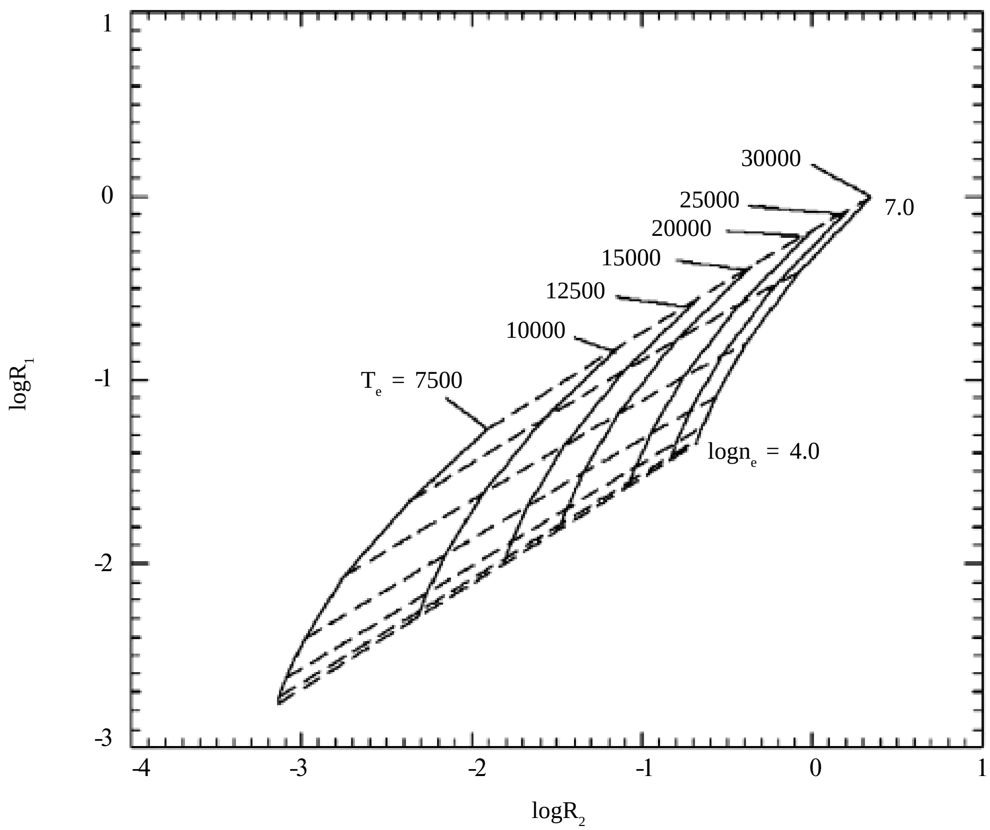

Fig. 5. The intensity ratios R_1 and R_2 of the [OIII] lines as functions of T_e and *ne*. The values of log*ne* increase from 4.0 to 7.0 with a step size of 0.5 on going from bottom to top.

plasma, then all the curves should intersect at a single point. In fact, because of the inhomogeneity of the nebula, there is a substantial spread in the values of T_e and *ne*, so this method can only provide estimates for these parameters.

The set of methods for determining the parameters of low density plasmas based on the intensities of forbidden lines in their spectra has also been used since the 1950's for analyzing the physical conditions in the coronae of the sun and other stars, whence it has come to be known as the *corona approximation* [36].

4.3. Diagnostics of inhomogeneous plasmas. The improvement in the quality of spectral observations associated with the launch of the Hubble telescope and the operation of large telescopes with mirror diameters of 8-10 m has led to the discovery of extreme inhomogeneity in planetary nebulae. It appears that images of nebulae contain a large amount of very small-scale details [37] associated with compact regions in the nebulae within which conditions differ from the average over these nebulae. It is currently possible to obtain spectra of compact regions in nebulae with sizes smaller than an arc second and, therefore, to perform detailed diagnostics of these small regions [38].

Researchers had assumed that nebulae are highly inhomogeneous even before detailed images with a high

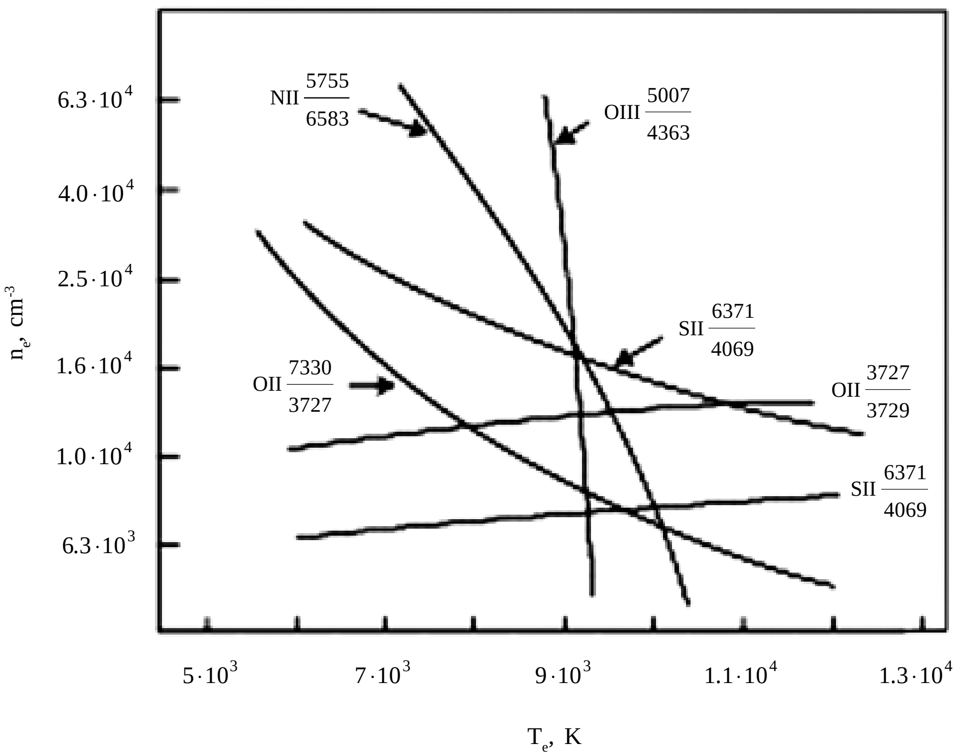

Fig. 6. A combined determination of the electron density *ne* and electron temperature T_e for the nebula NGC7026 [21].

angular resolution became possible. In the 1950's and 1960's it was suggested that small scale fluctuations in the temperature and density occur in nebulae [39-41] in order to explain some features of the spectra of nebulae that were not consistent with a model of a homogeneous plasma with constant temperature and density, and a method was proposed for calculating the spectra of nebulae taking these fluctuations into account. It was pointed out [42] that variations in the electron density can also have a significant effect on line intensities.

.A method of accounting simultaneously for the effect of small fluctuations in the temperature and density of the gas in nebulae on their spectra has been proposed [43-45]. It was shown that the energy radiated in a given line by an inhomogeneous nebula can be represented as the energy emitted in this line by a homogeneous nebula with average values $\overline{T_e}$ and $\overline{n_e}$ for the entire nebula, multiplied by a correction factor which depends on the amplitude of the fluctuations in the electron temperature and density.

The applicability of this approximation is limited by the assumption that the amplitude of the fluctuations in T_e and n_e is small. Calculations [43,44] show that this smallness can be interpreted very broadly. Even for deviations from the average temperature and density of ±25% in nebulae, the difference in the total energies emitted in visible and UV lines calculated assuming a small amplitude for the fluctuations in the electron temperature and density and with exact integration of the variable emission coefficients over the entire volume of the nebula was less than 3-5% in most cases.

The overall amplitude of the temperature variations in a nebula is made up of large scale variations associated with a relatively slow reduction in the average temperature with increasing distance from the central star (characteristic dimensions comparable to the size of the nebula itself) and small scale variations with dimensions considerably smaller than those of the nebula. Calculations [46] show that the amplitudes of the large scale variations are small. At the same time, in order to match the observed and theoretical intensities of lines in the spectra of nebulae it is necessary in many cases to use large-amplitude variations in T_e. Thus, the main contribution to the amplitude of the temperature variations in nebulae is from small-scale variations (fluctuations).

The intensities of lines excited by electron impact depend very strongly on the amplitude of fluctuations in T_e, while the intensities of recombination lines are insensitive to these fluctuations. For fluctuations in T_e with an amplitude of ±10% from the average temperature, the intensity of the [CIII] λ1907 line can increase by a factor of 1.5-2 relative to that for a plasma with a uniform temperature, while the intensities of the purely recombination Hβ and [CII] λ4267 lines are essentially invariant [44].

The technique of plasma diagnostics with fluctuations in both the temperature and density of atoms has been analyzed [44]. A maximum likelihood principle is used to find the optimum values of the parameters [47]. In doing this, the substantially different measurement accuracies for the intensities of different lines in the spectra of gaseous nebulae are taken into account [48,49]. The intensities I_λ of strong lines (comparable to that of the Hβ line) are measured considerably more precisely than those of the weak lines ($I_\lambda / I(\mathrm{H}\beta) \leq 1$). Whereas the error in measuring the intensities of strong lines is less than 5% [50], the corresponding error for weak lines may be as high as 30-50% or more. This circumstance is not usually taken into account in determinations of the parameters of nebulae.

This diagnostic technique for inhomogeneous plasmas proposed in Refs. 43 and 44 has made it possible to obtain a very good description of the spectra of nebulae. The differences between the observed and modelled line intensities is less than 5-10%. In addition, this method can be used to find the global characteristics of an entire nebula, and not just for a particular ion, as happens with the elementary diagnostic techniques. The use of this method for calculating ionization models of nebulae is described in a recent paper [51].

If, on the other hand, the deviations in the temperature and density from their averages cannot possibly be assumed small, then it is necessary to consider the actual distributions of a differential measure of the emission over the plasma volume. In this case, the set of parameters describing the state of the plasma must be supplemented by parameters that describe the distribution of the differential measure of the emission. A similar approach to the diagnostics of inhomogeneous plasmas has been described elsewhere [52].

4.4. Diagnostics of nebular plasmas and the carbon problem. Despite many years of work and the resulting understanding of the basic processes by which the spectra of nebula are formed, some unsolved problems remain. The most long-lived of these is the so-called ***carbon problem***. Essentially it is the following: when the abundance of carbon is determined from the intensities of recombination lines (***recombination abundance***), primarily the purely recombination [CII] λ4267 line, the resulting value is sometimes an order of magnitude or more higher than the abundances determined using the intensities of electron-impact excited lines [49] (***collisional abundances***). With further study it turned out that the carbon problem is just as much a nitrogen and oxygen problem, since the abundances of these elements obtained from recombination and collisional lines bear the same relation as for carbon [53].

Beginning with Ref. 39, the reason for the large difference between the recombination and collisional abundances has been seen to be the existence of temperature fluctuations. It was shown [49] that the intensities of weak recombination lines may be overestimated. This idea is supported by a model calculation [48] of the process for measuring the intensities of weak lines. We note that the problem of excess measured intensities of weak lines has been encountered repeatedly by researchers on planetary nebulae [20].

A solution of the carbon problem in terms of the combined action of two effects has been proposed [43]. The first arises in fluctuations in the electron temperature, which lead to an increase in the intensities of collisional lines and, accordingly, to a reduction in the carbon abundance determined using these lines relative to that obtained with a model of the nebula without electron temperature fluctuations. The second effect is an enhancement in the intensities of weak recombination lines, which leads to an overestimate of the recombination abundances.

It has been assumed that greater accuracy in measurements of the intensities of recombination lines would finally solve the carbon problem. Unfortunately, this hope has not been justified. The accuracy of the line intensity measurements has increased by 1-2 orders of magnitude, but the discrepancies between the recombination and collisional abundances remain [38]. The existence of small condensates (size ?1000 a.u.) with a hydrogen deficit and an elevated content of CNO and heavier elements in nebulae has been proposed as a solution to the problem [38]. Condensates of this sort are regarded as the main source of radiation in the recombination lines. Note, however, that estimates [53] show that the degree of ionization of the elements CNO in these condensates may be too low to produce significant recombination line radiation.

5. Conclusion

Our brief review shows that the ideas of V. A. Ambartsumian about the possibility of determining the parameters of a rarefied plasma using emission line intensities have fallen on fertile ground. The diagnostic techniques for rarefied plasmas which originated in an analysis of the spectra of gaseous nebulae have grown into an extensive branch of plasma physics, including astrophysical plasmas, and are used in almost half of published papers dealing with the analysis of the spectra of a wide class of astrophysical objects, from stars and the interstellar medium to the intergalactic gas.

REFERENCES

1. I. S. Bowen, *Astrophys. J.* **67**, 1 (1928).
2. H. Zanstra, *Publ. Dom. Astroph. Obs*. **4**, 209 (1931).
3. V. Ambarzumian, *Tsirk. Pulk. obs*., No. 6, 10 (1933).
4. V. A. Ambartsumian, *Scientific Papers*, Vol. 1 [in Russian], Izd. AN ArmSSR, Erevan (1960).
5. V. Ambarzumian, *Tsirk. Pulk. obs*., No. 4 (1932).
6. V. Ambarzumian, *Nature* **129**, 725 (1932).
7. V. Ambarzumian, *Mon. Notic. Roy. Astron. Soc*. **93**, 50 (1932).

8. V. A. Ambartsumian, Uch. Zap. LGU, No. 31, 5 (1939).
9. V. V. Sobolev, Moving Stellar Envelopes, Izd. LGU, Leningrad (1947).
10. V. Ambarzumian, *Mon. Notic. Roy. Astron. Soc*. **95**, 469 (1935).
11. V. Ambarzumian, *Zeit. für Astroph*. **6**, 107 (1933).
12. V. A. Ambartsumian, *Theoretical Astrophysics* [in Russian], GONTI, Leningrad-Moscow (1939).
13. V. A. Ambartsumian, É. R. Mustel', A. B. Severnyi, and V. V. Sobolev, *Theoretical Astrophysics* [in Russian], Gostekhizdat, Moscow (1952).
14. V. V. Sobolev, *A Course in Theoretical Astrophysics* [in Russian], Nauka, Moscow (1985).
15. V. V. Sobolev, ed., *A History of Astronomy in Russia and the USSR* [in Russian], Yanus-K., Moscow (1999).
16. M. Hebb and D. Menzel, *Astrophys. J.* **92**, 408 (1940).
17. V. V. Golovatyi, A. A. Sapar, T. Kh. Feklistova, and A. F. Kholtygin, Atomic Data for the Spectroscopy of Rarefied Astrophysical Plasmas. Gaseous Nebulae [in Russian], Valgus, Tallinn (1991).
18. D. Menzel, D. Baker, L. Aller, D. Shortley, M. Hebb, and L. Goldberg, Physical Processes in Gaseous Nebulae [Russian translation], Izd. Inostr. Lit., Moscow (1948).
19. L. H. Aller, Gaseous Nebulae, Chapman & Hall, London (1956).
20. L. H. Aller and W. Liller, *Planetary Nebulae* [Russian translation], Mir, Moscow (1972).
21. S. Pottasch, *Planetary Nebulae* [Russian translation], Mir, Moscow (1987).
22. S. Kwock, The Origin and Evolution of Planetary Nebulae, Cambridge Univ. Press, Cambridge (2000).
23. D. E. Osterbrock, Astrophysics of Gaseous Nebulae and Active Galactic Nuclei, Univ. Sci. Books, Univ. Minnesota (1989).
24. D. E. Osterbrock and G. J. Ferland, Astrophysics of Gaseous Nebulae and Active Galactic Nuclei, 2nd. ed. University Science Books, Sausalito, CA (2006).
25. M. Taylor and J. M. Vilchez, arXiv:0709.3473v3 (2008).
26. Z. B. Rudzikas, A. A. Nikitin, and A. F. Kholtygin, *Theoretical Atomic Spectroscopy* [in Russian], Izd. LGU, Leningrad (1990).
27. K. V. Bychkov and A. F. Kholtygin, *Elementary Processes in Astrophysical Plasmas* [in Russian], Moscow (2007).
28. M. Kafatos and J. P. Lynch, *Astrophys. J. Suppl. Ser*. **42**, 611 (1980).
29. M. J. Seaton, *Mon. Notic. Roy. Astron. Soc*. **114**, 154 (1954).
30. J. B. Kaler, *Astrophys. J.* **160**, 887 (1970).
31. J. B. Kaler, *Astrophys. J.* **308**, 322 (1970).
32. M. J. Seaton and D. E. Osterbrock, *Astrophys. J.* **125**, 665 (1956).
33. L. Stangellini and J. B. Kaler, *Astrophys. J.* **343**, 811 (1989).
34. G. X. Chen and A. K. Pradhan, *Astron. Astrophys. Suppl. Ser*. **147**, 111 (2000).
35. F. L. Crawford, F. P. Keenan, K. M. Aggarwal et al., *Astron. Astrophys.* **362**, 730 (2000).
36. P. L. Dufton, *Comp. Phys. Comm*. **13**, 25 (1977).
37. L. Bianchi, H. Ford, R. Bohlin, F. Paresce, and G. de Marchi, *Astron. Astrophys*. **301**, 537 (1997).
38. Y. G. Tsamis, J. R. Walsh, D. Pequignot et al., *Mon. Notic. Roy. Astron. Soc.* **355** (2008).
39. M. Peimbert, *Astrophys. J*. **150**, 825 (1967).
40. M. Peimbert and R. Costero, Bol. Obs. Tonantzintla y Tacubaya 5, 3 (1969).
41. R. H. Rubin, *Astrophys. J.* **155**, 841 (1969).

42. R. H. Rubin, *Astrophys. J. . Suppl. Ser.* **69**, 897 (1989).

43. A. F. Kholtygin, *Astron. Astrophys.* **329**, 691 (1998).

44. A. F. Kholtygin, *Astrofizika* **43**, 627 (2000).

45. A. F. Kholtygin, J. C. Brown, J. P. Cassinelli et al., *Astron. Astrophys. Trans.* **22**, 499 (2003).

46. R. Gruenwald and S. M. Viegas, *Astron. Astrophys.* **303**, 535 (1995).

47. Z. Brandt, Statistical Methods for the Analysis of Observations [Russian translation], Mir, Moscow (1975).

48. C. Rola and D. Pelat, *Astron. Astrophys*. **287**, 677 (1994).

49. C. Rola and G. Stasinska, *Astron. Astrophys.* **282**, 199 (1994).

50. W. A. Feibelman, S. Hyung, and L. H. Aller, *Mon. Notic. Roy. Astron. Soc*. **278**, 625 (1996).

51. M. Taylor and J. M. Vilchez, arXiv:0711.1474v1 (2008).

52. P. G. Judge, V. Hubeny, and J. C. Brown, *Astrophys. J. Suppl. Ser.* **475**, 275 (1997).

53. A. F. Kholtygin and T. Kh. Feklistova, *Astron. zh.* **69**, 960 (1992).