Historical Remarks on the Spectroscopic Analysis of Planetary Nebulae

Lawrence H. Aller

Department of Astronomy and Astrophysics, UCLA, CA 90095, U.S.A.

Siek Hyung

Korea Astronomy Observatory, Yusong-gu, TaeJon 305-348, Korea

Abstract. To record the spectral lines of stars and nebulae, photographic photometry was routinely used in the early 1900s, to be later replaced by methods employing more advanced detectors. We review the progress of planetary nebular spectroscopic studies, i.e. from the early photographic photometry, through to the 1950s photoelectric scanner observations, and finally to recent charge-coupled-device spectroscopy.

1. Introduction

The study of planetary nebulae (PNe) was inspired largely by the challenge to physical theory presented by the excitation states of elements seen in their remarkable spectra; these are quite unlike laboratory spectra, or those seen from "normal" stellar atmospheres. A high-temperature energy source for the production of the H lines was identified by Zanstra in 1927. By far the strongest lines in most PNe are the so-called "forbidden" lines of the ions of abundant elements: notably O, N & Ne, which are produced by electron collisional excitations. Bowen, who was a pioneer in this field, explained the great strength of certain permitted UV O III lines in high-excitation PNe, by a fluorescence mechanism. Recently, many permitted lines of C II, N II and similar ions, presumably due to recombination, have been studied in PNe, particularly by Liu and his associates (2001). Ambarzumian analyzed the excitation of the He lines, in a preliminary survey of this problem.

Before recent times, analyses of the shapes and structures of PNe were greatly impeded by the large and unknown distances to almost all of them. Curtis (1918) completed the first photographic PNe survey with the 36-inch Crossley reflector at Lick Observatory, and found irregularities, blobs and ansae in the images. This fine-scale structure cannot be resolved with ground-based telescopes without adaptive optics (AO), for all but the some of the nearest PNe. With the advent of the Hubble Space Telescope (HST) and AO, this resolution problem was somewhat alleviated. By the late-twentieth century, before these refinements in observational techniques were made available, important progress was made by O. Wilson and R. Minkowski at Mt. Wilson Observatory, with the Coudé spectrograph, and by direct photography at the 100-inch. Large-scale, monochromatic images, limited only by seeing, became obtainable. The assess-

ment and analysis of the spectroscopic data, either by empirical extrapolations or by theoretical models, involves many uncertainties, which are often largely due to the unknown distances and intrinsic geometrical irregularities of PNe. Here we make some remarks on 20th-century planetary nebular spectroscopy, i.e. diagnostics, abundances, and structural analyses, made possible only by the technical improvements in: instrumentation, computational power and atomic theory.

2.   Observational Devices

2.1.   Photographic Spectrometry and its Problems

Line intensity ratios of high accuracy are required to obtain useful diagnostics (electron density N_ϵ and temperature T_ϵ) of the nebular plasma. Two problems present themselves in spectrophotometry of PNe. The first is the great range (well over 10,000) of the intensities of the lines in which we are interested. The second is the fact that a good spectral resolution is important for resolving the components of density-sensitive lines, e.g. 3726.0 & 3728.6 of [O II], and 5198 & 5200 of [N I].

Spectral line intensities in stars and nebulae were routinely observed by photographic photometry, around the middle of the last century. With the advent of the photoelectric scanner in the 1950s, the strongest lines could be measured with this newer instrument, but the spectral resolution was still low, except at the Coudé focus. With a Coudé spectrograph and a calibrated photographic plate, many of the weaker lines could be measured and placed on a more accurate intensity scale, provided the zero point had been established. For example, in Mt. Wilson spectrophotometry (circa the 1960's & 1970's), spectral line intensities were obtained using this type of hybrid scanner system.

The photographic plate is a non-linear device, so systematic errors could creep in when a large range of intensities is involved. To make matters worse, there seems to have been a persistent systematic error in the fundamental calibration of the Mt. Wilson Coudé, in the early 1950s, which was too small to affect stellar line intensity measurements, but sufficient to have caused much trouble with nebular photographic line data. This defect appears in the work of Aller, Bowen, & Minkowski (1955) on NGC 7027 .

A further troublesome disadvantage is that all photographic emulsions have low quantum efficiency, especially longward of 500nm. Panchromatic plates are used in this region; specially dyed and hyper-sensitized plates must be employed to investigate the region longward of about 700nm. The farther one proceeds into the near infra-red, the more difficult it is to do spectrophotometric work, as the emulsions become trickier to handle. An annoying feature of the non-linearity of the photographic plate is that the dependence of the logarithmic intensity on the emulsion density steepens with increasing wavelength, adding to the misery of the investigator. To cover increasingly fainter spectral lines, many more exposures were required.

2.2. Other Spectroscopic Devices

The development of advanced spectrographs must overcome the aforementioned two hindrances: small dynamic range in intensity measurement, and low spectral dispersion. New devices, that offer improved performance over the use of "unfortified" photographic plates, were required. These were indeed developed, and included (but were not limited to):

(a) the image tube camera,
(b) the Lallemand electronic camera,
(c) the image tube scanner (ITS),
(d) the Image Photon Counting System (IPCS), and
(e) the charge-coupled-devices (CCDs).

With the image tube camera, a phosphor screen is excited, e.g. by the image of a nebular spectrum, and the screen is placed in contact with a photographic plate, which was then able to record the image far more efficiently. A further advantage that the phosphor has over the blue-sensitive photographic emulsions, is that it can much more efficiently record green, yellow and red light.

The Lallemand electronic camera was adapted by M.F. Walker to operate at the Coudé focus of the 3m Lick telescope. These observations covered the region from 360nm to 800 nm at high resolution, for 33 PNe (Aller & Walker, 1965; 1970). Although this technique gives excellent high-resolution spectra and has high efficiency, it requires difficult laboratory skills, which tax the capabilities of most astronomers.

For many observatories established in the late 19th or 20th century, e.g. Mt. Wilson or Lick, night sky light pollution became a serious problem, especially for the study of faint objects such as nebulae and galaxies. A popular technique was to employ an image tube scanner, with 2 identical receivers, one set on the object, the other on the empty sky. The latter signal is subtracted from the former, leaving the nebular spectrum. To eliminate differences in the 2 receivers, the observations were repeated with the object placed in the other aperture. In this technique, a photographic plate is used; the radiation fluxes were then processed electronically. Because of the ease of sky elimination, the ITS method was popular. However, the noise becomes very large when the night sky emission became polluted by artificial emission, such as from the Hg 4359 line. This particular line falls close to the $\lambda 4363$ [O III] line, which is an important transition for determining T_e in gaseous nebulae.

The IPCS, when used with a high-dispersion spectrograph, can yield high-resolution spectra; but, it is most useful for low intensity lines, though it cannot accommodate a large brightness range. An example of its application to a PN is given in Aller et al. (1981, NGC 2867). Combined with an instrument capable of measuring strong fluxes, the IPCS can yield good results over the entire spectral range.

Starting in the 1980s, it became possible to secure high-dispersion nebular spectra of high accuracy, using a spectrograph equipped with a CCD. The CCD is a linear device, and it has a high dynamic range, and also enjoys high S/N over a wide range of wavelengths. By a series of exposures with carefully chosen integration times, one can now obtain a dynamic range of 100,000, i.e. 0.01 to a

few 10 times the Hβ intensity! Even higher dispersion spectra are achievable with the use of an echelle spectrograph. The development of new computer software reduction packages, e.g. IRAF and MIRA, now allow the efficient reduction of these CCD echelle data. Advancements were also made in modern computing techniques, and in the more accurate calculation of high quality atomic data (in particular, transition probabilities and collisional strengths).

3.   Remarks on Diagnostics, Structures, and Models

It has long been known that the 3726/3729 [O II] ratio is a good indicator of N_e in the O$^+$ zone, while the $\lambda\lambda$4363/5007 ratio yields the T$_e$ in the O^{++} zone. For many years, these two ratios gave us our principal clues as to the N_e and T$_e$ in gaseous nebulae. The original quantum mechanical theory was somewhat inadequate, but now, thanks to the efforts of many investigators, the situation has greatly improved.

Calculations of cross-sections for ions with 2 or 4 equivalent p electrons give T$_e$ discriminants over a large range of temperature. Some ionic cross-sections are relatively insensitive to N_e, over a wide range of densities, in PNe. To obtain both N_e and T$_e$ from, say [O III], we would require not only optical line observations, but also IR spectra uncontaminated by atmospheric absorption.

Ionic species with 3 equivalent p electrons show forbidden lines in [N I], [O II], [Ne IV], [S II], [Cl III] & [Ar IV], that fall mostly in the accessible optical region. Ratios of pairs of nebular line intensities, e.g. $\lambda\lambda$5537/5517 of [Cl III], are insensitive to T$_e$, and give a good estimate of N_e. From the relative intensities of the nebular type transitions, and of the auroral or transauroral lines, we can get N_e and T$_e$ at the same time. See, for example, Keenan et al. (1996, 1998, 1999). The quantum mechanical calculations required for ions with 3 equivalent p electrons are more difficult than those for 2 or 4 equivalent p electrons; only recently were properly calculated atomic data obtained for some of the ions with 3 equivalent p electrons.

Elemental abundances found from optical recombination lines (ORLs) tend to be higher than those from collisionally excited lines (CELs) by a factor of ~1.5 to ~20. What causes this discrepancy? One possibility is that there are huge variations in N_e and T$_e$ within the emitting layers, far in excess of what had been suspected before. Lines of p^3 configuration tend to indicate a higher T$_e$ than p^2 ions. Fluorescence may play an important role. Intensive further work on this problem is urgently needed.

Additional diagnostics for heavier atoms, e.g. iron, in various ionization stages are now becoming available: e.g. [Fe III] and [Fe VII] (Keenan et al. 2001). These should supply good diagnostics for nebular envelopes. Iron lines are particularly valuable diagnostic tools, as a wide range of ionization stages is often present.

The structural complexity inherent in PNe is well illustrated by the objects such as Hu 2–1 (Miranda et al. 2001a), IC 4846 (Miranda et al. 2001b) and NGC 6543 (Miranda & Solf 1992, Reed et al. 1999). It has been suggested that, in these systems, the central exciting source is a binary star in which mass transfer from a red giant produces an accretion disk around a secondary. From these PNe systems, collimated outflows appear (Mastrodemos & Morris 1998, 1999).

Sun Kwok's well-known PN model proposes that the outer envelope of the antecedent giant star expands slowly. The outer envelope is eventually struck by a swiftly moving wind from the hot central nucleus. This wind might have speeds of hundreds of km s^{-1}, while the slowly swelling envelope of what had been a giant star moves outward at a speed of about 4–5 km s^{-1}. This "bull-dozing" action of the fast wind compresses the slowly moving shell; shock waves are set up moving inward toward the star, and also outward through the compressed shell.

This simple, attractive picture encounters severe difficulties when confronted by the details seen in most observations, though it may provide a nice framework for the understanding of the basic structures of many PNe. In objects such as NGC 7662, where we have a bright, high-excitation inner ring surrounded by a fainter, low-excitation outer ring, there sometimes can also be a halo thousands of times fainter than the inner ring. This appendage arises from the giant star winds; and it may have a mass as high as $\sim$0.2 M$_\odot$ (Middlemass et al. 1991). This halo, which appears to be heated by the stellar winds, may actually have a higher T_e than the bright ring.

PNe structures frequently show many complexes, in addition to the simple non-uniformities in N_e and T_e. These include, e.g. ansae, which are structures well exhibited in NGC 7009, with blobs of material often ejected symmetrically on opposite sides of the PN; and FLIERs (Fast-moving Low Ionization Emission Regions), e.g. Balick (1987), Balick et al. (1993). The FLIERs do not show the [O III] or [Ne III] lines usually prominent in PNe. In most PNe, both ansae and FLIERs appear in pairs, going in opposite directions. In NGC 7662, they move with speeds of $\sim$50 km s^{-1} (much greater than normal nebular expansion speeds) and seem to have an age of about a thousand years.

The best known of the polypolar PNe are the Peimbert Type I (Peimbert 1967, 1978) object NGC 2440 (Lopez et al. 1998), which has been studied by many observers over the years. In such an object, an accretion disk in a binary system may be involved. For example, IC 4846, which is described as an elongated shell with 3 collimated pairs of jets at different orientations, is hard to explain in terms of any system with simple geometry. These principal outflows may be caused by a precessing bipolar jet. Miranda et al. (2001b) suggest that IC 4846 involves 3 different components: (a) 2 shells with different orientations; (b) an attached circular shell; and (c) 3 pairs of point-symmetric components at different orientations. A binary source would provide much of the complexity in this and similar systems. However, in view of the low frequency of such configurations, are such complicated PNe likely to be unusual? Modeling such complex, evolving PNe structures presents a real challenge to theorists, some of whom take into account the magnetohydrodynamics of expanding, radiating shells.

PNe were not supposed to be as complicated as this – at least so thought many investigators in the last century. But the improved observational insights have changed everything; we have here mentioned some examples, e.g. from the data secured with the Canary Island telescopes. The hope remains that we can obtain data covering a wide range of evolutionary stages, and structural complexity. Insofar as nebular line intensities are concerned, what we ideally need are isotopic maps for important lines, which involve a wide range in excitation

and intensity. We have the observational expertise to undertake such enterprises. It would be advisable to select a number of well-chosen candidate PNe, and to observe them intensively. Every student of PNe has a "little list", which may include such venerable, rich-line objects as: NGC 7027, NGC 7009 and NGC 7662. However, it is very important that we also include complex objects, and those at interesting stages of their evolution, especially in the pre-PN stage.

4. Conclusions

The question of theoretical models for an evolving, expanding, PN shell involves problems in hydrodynamics & probably magnetohydrodynamics. The complexities introduced by polypolar PNe make any treatment of the problem excruciatingly difficult. Diagnostics have been established for the determination of (N_e, T_e) within PNe. Originally, we were limited to lines in the optical region, but now more and more data are being secured in the UV and IR (e.g. Mathis et al. 1998). In all analyses where we combine data from different spectral regions, e.g. IUE, IR and optical data, we should ensure that the same spatial area of the PN is being covered in each of the datasets.

To interpret the data, we often use a theoretical nebular model. A good model gives a satisfactory representation of the observed line intensities; such a model may then give a fair estimate of the actual abundances in the PN, and also of the ionic distribution among the unobserved ionization stages. Numerous examples of the chemical compositions derived for PNe may be found in various papers (see, e.g. Harrington et al. 1982, NGC 7662; Hyung et al. 1999, IC 4634).

Another technique is to measure the monochromatic intensities at several places in the PN image. From the images at different wavelengths, the ionic ratios may be used to estimate $n_{element}/n_H$; see, e.g. Barker (1980, 1982, Ring Nebula; 1986, NGC 3242NGC 3242), Hawley & Miller (1977, NGC 6720). With these ratios, we can get abundances for prominent elements with well-defined emission lines, such as: He, C, N, O, Ne, S, Cl and Ar. This method is well-suited to large, rather extended objects, such as NGC 6853 (the Dumbbell Nebula), where theoretical models fail to accommodate the geometrical irregularities intrinsic to the PN. For small PNe, the total monochromatic flux for each line is usually measured with a photoelectric scanner; but for large PNe, we need to measure the distribution F(x,y) of brightness over the monochromatic image–in this case, we get the integrated emissivity along the line of sight. Thus, the specification of line (or continuum) intensities for PNe may be much more complicated than for the non-extended sources. Finally, for rare elements, which may be represented by only one or two ions, abundance estimates are only obtainable through detailed photoionization modeling.

In one respect, some of the older data, when properly calibrated and analyzed, may prove to be worthwhile. In 1937, Louis Berman pointed out that the [O III] 4363 line was variable in IC 4997, a bright, dense, presumably young PN. Photoelectronic photometry by Liller & Aller (1955) confirmed these intensity changes. Since then, variability has been found in other PNe. As the lifetime of the PN stage is measured in thousands of years, and that of the AGB stage in hundreds of thousands, or a few million years, we might expect these changes to be barely detectable within a century or two, even in a rapidly evolving PN.

With the marked increase in accuracy of currently available data, the chances of detecting changes in PNe "while you wait" will become ever greater. We should prepare for this possibility.

In summary, improvements in observational techniques have enabled us to find line and continuum intensities for selected points within PNe images, from which adequate structural and kinematical pictures may be built. Accurate atomic data for diagnostically important ions, such as C, N, O, Ne, S, Cl, Ar and Fe, are also required, in order to take full advantage of the new, improved observational data.

References

Aller, L.H. 1965, ApJ, 141, 1318

Aller, L.H., Bowen, I.S., & Minkowski, R. 1955, ApJ, 122, 62

Aller, L.H., & Walker, M.F. 1970, ApJ, 161, 917

Aller, L.H., Keyes, C.D., Ross, J.E., & Omara, B.J. 1981, MNRAS, 197, 647

Balick, B. 1987, AJ, 94, 671

Balick, B., Rugers, M., Terzian, Y. & Chengular, J.N. 1993, ApJ, 411, 773

Barker, T. 1980, ApJ, 240, 99

——. 1982, ApJ, 253, 167

——. 1986, ApJ, 308, 314

Curtis, H.D. 1918, Publ. Lick Obs. 13, 1

Harrington, J.P., Seaton, M.J., Adams, S., & Lutz, J.H. 1982, MNRAS, 199, 517

Hawley, S.A., & Miller, J.S. 1977, ApJ, 212, 94

Hyung, S., & Aller, L.H. 1997, ApJ, 491, 242

Hyung, S., Aller, L.H., & Feibelman, W.A. 1999, ApJ, 525, 294

Keenan, F.P., Aller, L.H., Bell, K.L., Hyung, S., McKenna, F.C., & Ramsbottom, C.A. 1996, MNRAS, 281, 1073

Keenan, F.P., Aller, L.H., Bell, K.L., Espey, B.R., Feibelman, W.A., Hyung, S., McKenna, F.C., & Ramsbottom, C.A. 1998, MNRAS, 295, 683

Keenan, F.P., Aller, L.H., Bell, K.L., Crawford, F.L., Feibelman, W.A., Hyung, S., McKenna, F.C., & McLaughlin, B.M. 1999, MNRAS, 304, 27

Keenan, F.P., Aller, L.H., Ryans, R.S.I., & Hyung, S. 2001, PNAS, 98 (17), 9476

Liller, W., & Aller, L.H. 1955, AJ, 60, 167

Liu, X.-W., Luo, S.-G., Barlow, M.J., Danziger, I.J., & Storey, P.J. 2001, MNRAS, 327, 141

Lopez, J.A., Meaburn, J., Bryce, M., & Holloway, A.J. 1998, ApJ, 493, 803

Mastrodemos, N., & Morris, M. 1998, ApJ, 497, 303

——. 1999, ApJ, 523, 357

Mathis, J.S., Torres Peimbert, S., & Peimbert, M. 1998, ApJ, 495, 328

Middlemass, D., Clegg, R.E.S., Walsh, J.R., & Harrington, J.P. 1991, MNRAS, 251, 284

Miranda, L.F., & Solf, J. 1992, A&A, 260, 397

Miranda, L.F., Torrelles, J.M., Guerrero, M.A., Vazquez, R., & Gomez, Y. 2001a, MNRAS, 321, 487

Miranda, L.F., Guerrero, M.A., & Torrelles, J.M. 2001b, MNRAS, 322, 195

Peimbert, M. 1967, ApJ, 150, 825

————. 1978, in IAU Symp. 76, Planetary Nebulae, ed. Y. Terzian (D. Reidel: Dordrecht, Holland), 215

Reed, D.S., Balick, B., Hajian, A.R., Klayton, T.L., Giovanardi, S., Casertano, S., Panagia, N., & Terzian, Y. 1999, AJ, 118, 2430

Zanstra, H. 1927, ApJ, 65, 50
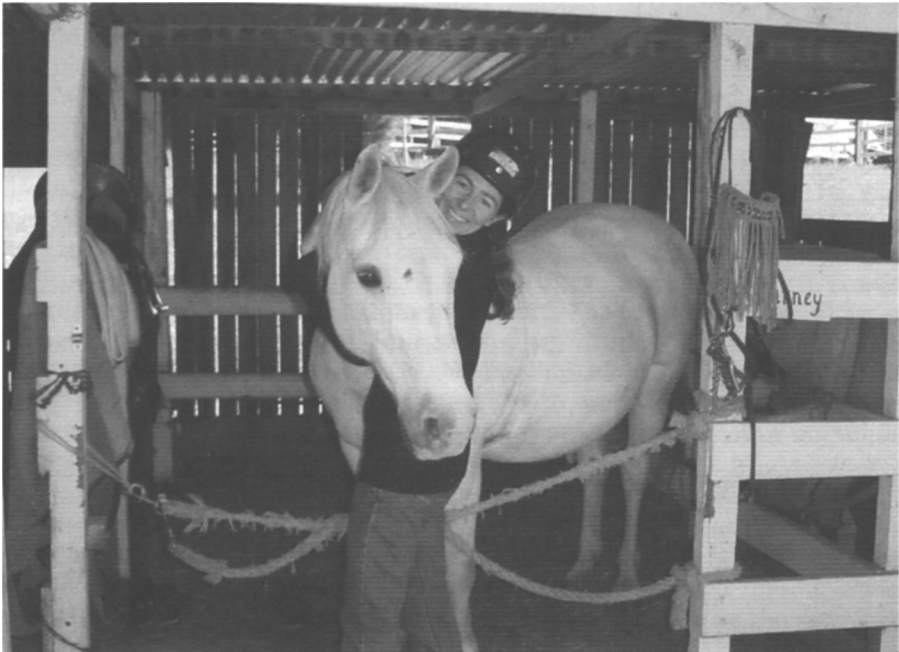
Orsola De Marco on the horse riding excursion.