SF2A 2008
C. Charbonnel, F. Combes and R. Samadi (eds)

THE MEXICAN MILLION MODELS DATABASE: A VIRTUAL OBSERVATORY FOR GASEOUS NEBULAE

Morisset, C.[1]

Abstract. The 3MdB (Mexican Million Models database) is a large database of photoionization models for H II regions. The number of free parameters for the models is close to 15, incluing the description of the ionizing SED and the description of the ionized gas. The outputs of the models are more than 70 emission line intensities, the ionic fractions and temperatures. All the parameters and outputs are included in the MySQL database, giving the possibility to the user to search into the database for example for all the models that reproduce a given set of observations.

1 Introduction

The study of the ionized interstellar medium is mainly based on the analysis of the observed emission line intensities. From line ratios one may determine physical and chemical parameters of the nebulae such as the electron temperature, the electron density and the abundances of the most common elements. The characteristics of the ionizing spectrum can also be determined from the line intensities. The interaction between the ionizing source and the gas is computed a photoionization code (e.g. Cloudy, see Ferland et al. 1998) allowingto constructu numerical models of H II regions, including the intensities of the emission lines. Such models can then be compared to the observations and a fit can be defined. I present here a new database of photoionization models, which can be used to look for models that are reproducing a given observation or a given catalog of observations. This tool can be understand as a kind of H II regions virtual observatory where line intensities from millions of models can be mined.

2 P-space and O-space

One can describe a (photoionization) model as a link from the parameter-space (P-space) to the observable-space (O-space). The parameter-space is describing an object in terms of effective temperature, luminosity, size of the nebula, radial density variation, abundances, presence of dust, etc. This can be seen as the set of inputs required to compute the model. The object in the observable-space is described by the set of the emission line intensities. This is also the set of outputs of the photoionization model. The dimension of the P-space is the number of free parameters needed to describe a model, it can easily reach a value of 15 for 1D models (as when running Cloudy), many more for 3D models where the description of the density distribution is more complexe (using e.g. Cloudy_3D, see Morisset 2006). The dimension of the O-space is the number of emission line intensities that one can obtain from the photoionization code. It can be seven hundreds of lines! In the O-space we find the results of the modeling process (what we classically call the models, projections from the P-space into the O-space using a code) and also the observations of "real" objects. Actually, taking into account the error bars around each observed value of emission line intensity transform the observed objects to an hyper-boxes around the observed values (in the O-space). The relation between the shape in the P-space and the corresponding shape in the O-space is far from being linear. For example, a rectangule in the P-space does not transform into a rectangular plane in the O-space, but rather into a complex hyper-shape. This is illustrated by Fig. 2 in Stasińska et al. (2006) where a regular grid in the P-space (of 2 dimensions U and Z) transforms into a curved

[1] Instituto de Astronomía, Universidad Nacional Autónoma de México; Apdo. postal 70–264; Ciudad Universitaria; México D.F.; 04510 México. e-mail: `Chris.Morisset@gmail.com`

shape into the O-space. The reverse is also true: a rectangular shape into the O-space is not obtained by a rectangular shape in the P-space: this is why it is not possible to easily obtain the parameters of the models that adjust a given observation. The action of fitting an observation by some models is finding the models which are close to a given observation in the O-space. Considering the errors on the observations, this means finding the models that fall in the hyper-box around the point that represent the object in the O-space. Due to the high non-linearity of the transformation between the P- and the O-space, there is no simple way to go from an observation to the set of physical parameters that describe the object. There are various ways to find the set of values in the P-space that reproduce an observed object (a point in the O-space, or an hypercube if we take the error bars into account): By running models and figuring out what are the effect in O-space of changing something in the P-space or by automatic Khi2 method: for example Cloudy can optimize a set of parameter to fit a set of observations. Generally these methods lead to a definition of the "best" model fitting the observations of an object. One can also use regular grids of models: this method can be very useful to see the effects of changing one parameter on the observables. It gives the possibility of finding various models that fit the same observation (within the errors) The last method is to use irregular grids of models. This is the case of a grid that can be adapted to increase the density of models in P-space where this is useful. Such an approach needs observations to know which locus in the P-space is "good" (it falls in a "good" locus in the O-space : where there is observed objects). For this one can use a kind of genetic algorithm, see next section.

3 A genetic algorithm for the definition of new models

To define a genetic algorithm, we must considere two phases: a phase of selection of parents and a phase of reproduction with random evolution, leading to children The selection of the parent models is performed in the O-space, in the hyper-boxes around the observations, the sizes of the hypercube being the acceptable error on each observable (e.g. emission line intensity). Any model that falls within an hyper-box around an observation is a model selected for the reproduction (it is a parent model). A new generation of models is generated from the set of parent models. The values of the parameters for the children are determined randomly around the values of the parent models, within a given range. Each parent will generate a given number of children.

4 The 3MdB

The Mexican Million Models database (supported by CONACyT grant 49737) is a project of a huge photoionization model database, where the user can search easily and quickly for models that reproduce a given set of observations. There are more than 15 parameters that can be varied to describe a model. First the ionizing SED can be described as a Planck function, as a stellar atmosphere model, in this case the stellar metallicity and the surface gravity may also be provided. There is also a possibility to describe the SED in terms of stellar cluster. Secondly the ionized gas: the inner radius of the nebula, the hydrogen density, the abundances of the main elements, the presence of dust (composition, density), a filling factor for the gas. Once the model is computed the corresponding entry in the database is completed by adding to the parameters the intensities of more than 70 emission lines. A entry in the 3MdB is: a point in the P-space (defined by the values of all the parameters), the corresponding point in the O-space (the values of the observables, i.e. line intensities), plus a set of other characteristics of the models. The genetic algorithm described in Sec.3 is used to compute the values of the parameters for the new generation models. The observations that are used for the selection of the parent models are from various catalogs, such as part of the SDSS like the one used by Izotov et al. (2006). The database contains 1,350,000 models (October 2008). The increasing rate of the database is 350 models/hour. It actually run on a 2-double-core AMD 64 bits processors computer. The data are in MySQL tables, driven by IDL routines calling Cloudy, reading the outputs and filling the database.

References

Ferland, G. J. Korista, K.T. Verner, D.A. Ferguson, J.W. Kingdon, J.B. and Verner, E.M. 1998, PASP, 110, 761

Izotov, Y. I., Stasińska, G., Meynet, G., Guseva, N. G., & Thuan, T. X. 2006, A&A, 448, 955

Morisset, C. 2006, in IAU Symposium, Vol. 234, Planetary Nebulae in our Galaxy and Beyond, ed. M. J. Barlow & R. H. Mendez, 467–468

Stasińska, G., Cid Fernandes, R., Mateus, A., Sodré, L., & Asari, N. V. 2006, MNRAS, 371, 972