Poster contributions of the "Advanced Lectures on the Starburst-AGN Connection", Eds. R. Mújica, I. Aretxaga & D. Kunth, 2001, INAOE Electronic Edition (http://www.inaoep.mx/∼agn00/posters.html), pp. 61-64

An inverse method for stellar population synthesis with error analysis

J. Moultaka and D. Pelat

DAEC, Observatoire de Meudon, 92195 Meudon, France

Abstract. In this paper we set out the problems of stellar population synthesis methods and describe an inverse method with its error analysis. This method allows to get the exact solutions of the synthesis if they exist and gives the unique solution that minimizes a "synthetic distance" when no exact solution exists. In any way, our method controls the set of all acceptable solutions.

1. Introduction

Stellar population synthesis in galaxies, as any physical problem, can be treated either by a direct method (i.e. going from the model parameters to the observables) or by an inverse one (i.e. going from the observables to the model parameters).
In the first and second sections, we discuss these two ways to solve the problem, in the third section, we describe briefly our method and finally, in section 4, we discuss our results.

2. Direct method:

In the direct method, one chooses the parameters of an "a priori" model (as the initial mass function, the star formation rate etc...) and deduces the predicted observables (as fluxes or equivalent widths) that are then compared with the real ones. The solution is accepted if they are sufficiently close to each other. This is the "model fitting" which is so commonly used in astrophysics.
The drawback of this method is that one cannot, in general, pretend finding *the* solution of the problem but *a* solution among a possibly wide set of acceptable solutions. Unfortunately, it is also the case in the inverse method and so we think that one should not present a single solution but has to give some kind of "error bars" around the adopted one.
This task is very difficult in the direct approach while it is naturally done in the inverse one.

3. Inverse method:

In the inverse problem, one starts from observables (in our case, the equivalent widths, hereafter EW) and deduces the parameters of the model (i.e. the stellar contributions to the total luminosity and all astrophysical quantities that follow).
In the inverse problem, we don't have to presuppose anything about the parameters we are looking for, and in the ideal situation:

1. *There exists a solution.*
2. *The solution is unique.*
3. *It is stable relatively to possible perturbations.*

However, this situation rarely happens in practice. Many "real" physical problems turn out to be what is called **"ill-posed problems".**
A problem is ill-posed if, at least, one of the three previous properties is not satisfied.

Often, it so happens that the stellar synthesis problem is an ill-posed one. Because of e.g. the limited signal-to-noise ratio (S/N), most of the time, the solution is inexistant or is not unique and/or a *small* variation of the observables induces a *large* variation on it.
A way to deal with ill-posed problems is to characterize their acceptable domain, to estimate the error zone and to regularize them by introducing some constraints.

4. Description of the inverse method

In our inverse method, the unique constraints on the stellar contributions are that the galactic stellar spectrum is composed only by the stars of our database and that these contributions are positive (otherwise the solution will not be physical).
The equation of synthesis is:

$$\boxed{W_{syn\,j} = \frac{\sum_{i=1}^{n_s} k_{\lambda_0 i} W_{ji} I_{ji}}{\sum_{l=1}^{n_s} k_{\lambda_0 l} I_{jl}}} \tag{1}$$

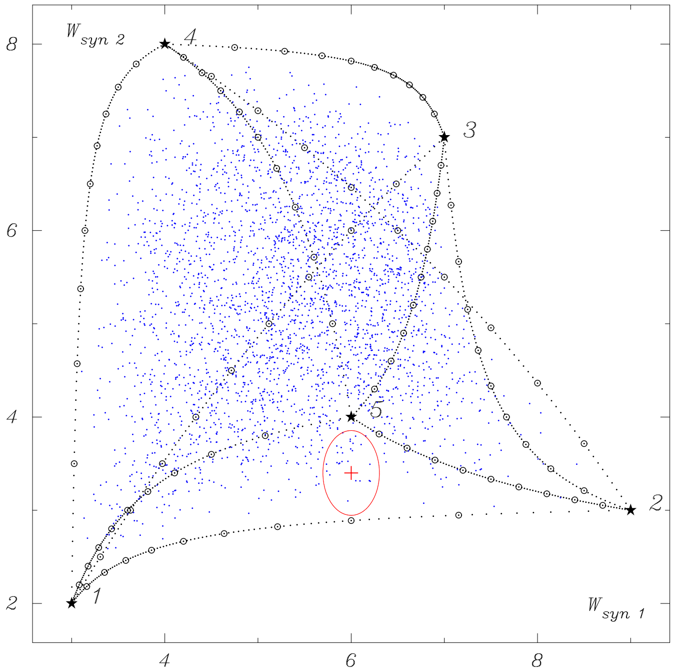

Fig. 1. The synthetic surface in a 2-dimensional equivalent widths' space (W_{syn}) created by five theoretical stars (under-determined case i.e. there will be several solutions).
Each star is characterized by one point in the 2-dimensional space symbolized by $\star$.
The cross symbol is the observed galaxy (in this case, **the galaxy can be synthesized**) and the ellipse represents the error region around the observation.
The points inside the synthetic surface are the image by the application of equation (1) of a uniform filling of the stellar contributions' space.
We use this uniform filling to calculate the **information** H brought by an observation and we define this information by $H = 1 - p$ where p is the probability that one point falls within the error ellipse.

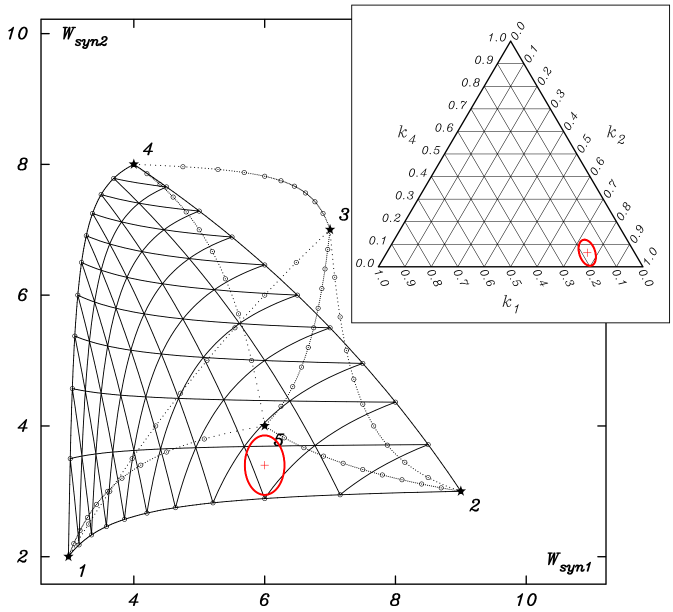

Fig. 2. First solution of the problem: the three stars which are contributing to the synthesis are stars number 1, 2 and 3. We show in the equivalent widths' space the synthetic domain (with its grid) formed by these three stars and in the stellar contributions' space the solution with its deduced error ellipse. In this solution, the standard deviations are moderate; **the solution is moderately determined.**

for $j = 1, ...n_l$,

where
n_l = number of lines,
n_s = number of stars constituting the database,
$W_{syn\,j}$ = synthetic EW at line number j,
W_{ji} = EW of line number j for star number i,
I_{ji} = continuum intensity of star i at line j,
$k_{\lambda_0 i}$ = contribution to luminosity of star number i at a reference wavelength λ_0.

This fundamental equation can be transformed in a linear one with positivity constraint on vector $\boldsymbol{k}$ of the stellar contributions as follows:

$$\left.\begin{aligned} \boldsymbol{k} &\geq 0 \\ \mathbf{A}\boldsymbol{k} &= \mathbf{0} \\ \boldsymbol{b}^T\boldsymbol{k} &= 1 \end{aligned}\right\} \tag{2}$$

where $\mathbf{A}$ is a matrix which is a function of galactic EW and stellar data and $\boldsymbol{b} = (\underbrace{1, 1, 1, ..., 1}_{Nbre\ of\ stars})$.

We now introduce **the synthetic surface** which is the domain of all equivalent widths that have an exact and physical solution. We are then confronted with two cases:

- The observed galaxy can be synthesized (i.e. its equivalent widths are on the synthetic surface):

In this case, the problem possesses a unique solution if it is **over-determined** ($n_l \geq n_s$) and several ones if it is **under-determined** ($n_l < n_s$)**.**
The error ellipsoid around the observed equivalent widths induces an error ellipsoid around the exact solution which gives a good estimation of the numerous acceptable solutions (see fig.1 to fig.4).

- The observed galaxy cannot be synthesized:

In this case, we search for the solution that minimizes a distance between the observed equivalent widths and the synthetic surface (see fig. 5) called **"the synthetic distance"** and defined by:

$$\boxed{D^2 = \sum_{j=1}^{n_l} (W_{obsj} - W_{synj})^2 P_j} \tag{3}$$

where
W_{obsj} = observed EW,

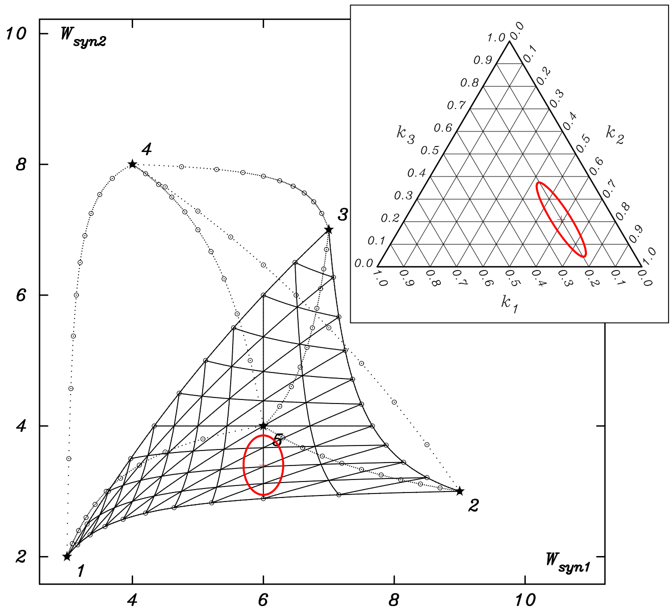

Fig. 3. Second solution of the problem: The three stars contributing to the synthesis are stars number 1, 2 and 5. Here the size of the deduced stellar error ellipse is too large. **The solution is badly determined.**

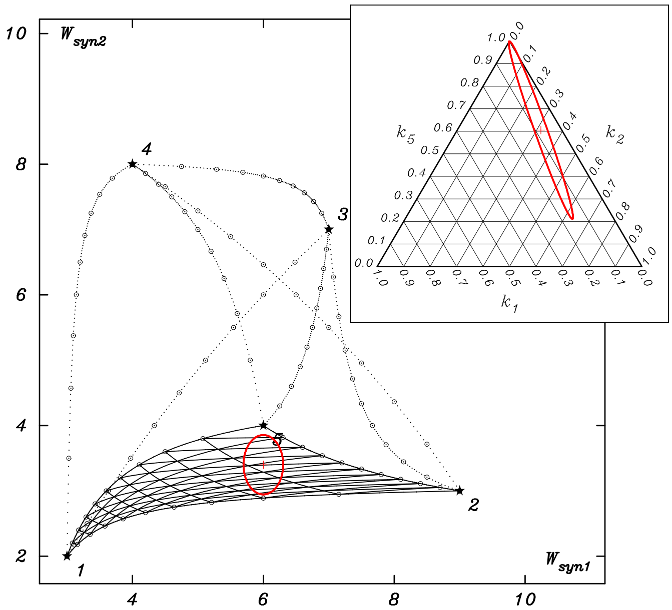

Fig. 4. Third solution of the problem: Here the three stars contributing to the synthesis are stars number 1, 2 and 4. We notice the small standard deviations in this solution; it is a **well determined solution.** The area (or the hyper-volume in a n-dimensional case) of the error region could be **a good criterion** for choosing **the best solution**

W_{synj} = synthetic EW,
P_j = weigths characterizing the quality of their measurements.
The error ellipsoid is computed by projecting (and then averaging) the deviation vectors around the mean observed EW vector upon an hyper-plane that approximates linearly the synthetic surface at the solution point.

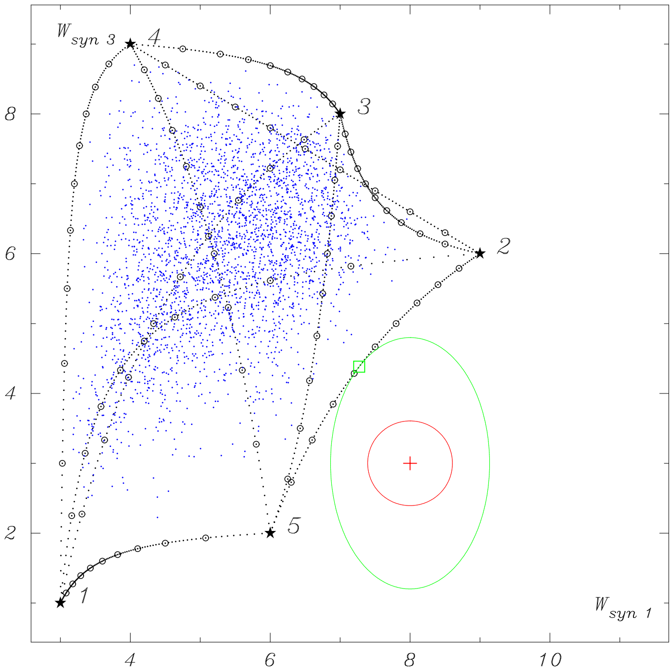

Fig. 5. We present the synthetic surface of the same stars of fig.1 but in the 2-dimensional space of their first and third equivalent widths. The symbols are the same as in fig.1. Here **the galaxy cannot be synthesized**, it falls outside the synthetic surface. The small circle is the error ellipse around it, the big ellipse represents the synthetic distance and the square is the synthetic vector that minimizes this distance.

The method is largely described in Pelat (1997) and (1998), the error analysis is exposed in Moultaka & Pelat (2000) and the application of this study upon the central regions of twelve galaxies is shown in Boisson et al. (2000).

5. Results and discussion

The error study shows:
- that the stability of the problem depends on the intensity continua I_{ji} of the database,
- that the stars of the database have to be relatively distant from each other in the equivalent width space in comparison with the error ellipsoid size.

Indeed, the signal-to-noise ratio (S/N) of the galaxy data is the key parameter: it must be high enough so that the error ellipsoid is smaller than the average distances between the stars on the synthetic surface.

It is obvious that the solution found by this method will not be necessarily astrophysical. In this case, more constraints inspired by the direct approach, have to be added. This work is in progress, it takes into account the theory of stellar evolution and uses two different

approaches for the IMF.
In our opinion, both direct and inverse approaches must be used in cooperation in order to solve this difficult problem.

The **most important conclusion** we bring out, is that great care has to be used when a solution is proposed through model fitting (direct approach). One must always present the set of acceptable solutions around it, for it may happen that this set is so large that any other solution could also be acceptable. In this case, the so called astrophysical conclusions are, in fact, contained in the hypothesis.
The inverse approach includes naturally the error analysis and therefore, stays away from this criticism.
Finally, we recall that one cannot get more reliable information than the data contain.

Acknowledgements. We would like to thank Catherine Boisson and Monique Joly for the very helpful comments they gave on this paper.

References

Pelat D., 1997, MNRAS, 284, 365
Pelat D., 1998, MNRAS, 299, 877
Moultaka J. & Pelat D. 2000, MNRAS, 314, 409
Boisson C., Joly M., Moultaka J., Pelat D. and Serote Roos M. 2000, A&A, 357, 850