OPTIMAL EXTRACTION OF MULTIPLE OVERLAPPING SPECTRA: A NEW ALGORITHM BASED ON THE MAXIMUM ENTROPY METHOD

S. V. Khmil[1] **and J. Surdej**[2,3]

[1]*Astronomical Observatory, Taras Shevchenko National University*
3 Observatorna St., 04053 Kyiv-53, Ukraine
e-mail: khmil@aoku.freenet.kiev.ua

[2]*Institut d'Astrophysique et de Géophysique, Université de Liège*
Avenue de Cointe 5, B-4000 Liège, Belgium
e-mail: surdej@astro.ulg.ac.be

[3]*Directeur de Recherches au FNRS (Belgium)*

A general algorithm of slit spectra extraction for a system of point-like sources (*e.g.*, multiple lensed images of the same quasar) has been developed, assuming that the point-spread function (PSF) of the measuring instrument and the positions of the spectra relative to the CCD frame are unknown. The main idea of the algorithm is to successively apply the maximum entropy method to each set of parameters, such as the spectra, the PSF, and the spectra positions, in order to iteratively improve their values. The algorithm uses all the a priori knowledge about the spectra (*e.g.*, flux positivity, relative magnitudes between the components, astrometry, *etc.*) to compute the initial parameter sets. After careful testing, the algorithm was applied to extract the spectra of the gravitationally lensed system Q1009-0252 A & B ($I_B/I_A = 0.1$, $\theta_{AB} = 1''.55$, $1\,\mathrm{pxl} = 0''.44$, typical seeing FWHM $= 0''.8$).

INTRODUCTION

Spectra of quasars multiply imaged by gravitational lensing contain unique information about the structure and the physical properties of the lensed sources, the intergalactic medium, and the lensing objects (galaxies and their clusters). Since the angular separation between such images is typically less than one or several arcseconds, their spectra very often overlap, are curved and noisy because of the atmospheric turbulence, instrument aberrations, detector noise, *etc.* Moreover, sometimes we are forced to process data which, for some reason, are incomplete. Thus, whenever we deal with observational data, some restoration technique is necessary to extract as much information as possible from them. In order to improve the reliability of such an extraction, it is very important to supplement the observational data with prior knowledge about the image. Powerful examples of such a priori information are the flux positivity, relative magnitudes of the image components, astrometric data, and so on.

We present here a new, sufficiently general, algorithm of slit spectra extraction for a system of point-like sources, which is based on the maximum entropy method (MEM). The algorithm uses all the a priori knowledge on the restored spectra and iteratively improves the spectra themselves, the numerical profile of the PSF, and the positions of the spectra within the CCD frame.

As an application of the algorithm, we present the extracted spectra for the gravitationally lensed quasar Q1009-0252 A & B.

DESCRIPTION OF THE ALGORITHM

Let us assume that we deal with standard, sky-subtracted, CCD observational data of slit spectra of an astronomical object which consists of $N_{\rm sp}$ point-like components. Typically, it may be a gravitationally lensed quasar or another system of multiple point sources. Hereafter, for convenience, spectra are labelled with the discrete index A ($A = 1, \ldots, N_{\rm sp}$) whereas the indices λ and i refer to pixel coordinates along the spectral and spatial directions of the CCD frame. In order to describe the spectra, let us introduce the following three sets of parameters: the set of spectral fluxes $C_{A\lambda}$; the profile of the PSF $\phi(x)$ which is supposed to be normalized to unity, generally nonsymmetric, and discretely given by its numerical values ϕ_n at $N_{\rm PSF}$ evenly separated points; the set of spectra positions relative to the CCD frame $x_{A\lambda}$.

Generally speaking, these three sets of parameters are unknown. *The main idea of the proposed method is to successively apply the MEM to each set of parameters in order to iteratively improve their values.* To implement this idea, we follow, with *some generalizations*, the simple maximum entropy deconvolution algorithm proposed by Cornwell & Evans [2]. The essence of the algorithm is directly bound to the concept of relative entropy.

Suppose that we are to reconstruct some object which is not obligatory an image and which is represented by *several* discrete sets of uniform *positive* parameters, like the three sets previously discussed. Following our main idea, consider separately one of the sets, say b_i ($i = 1, \ldots, M$), assuming that all the others are known. For convenience, we shall treat this set as a vector $\boldsymbol{b}$ in the M-dimensional space. Suppose also that a priori knowledge about the object can be expressed in the form of the vector $\boldsymbol{m}$ of the same dimension. The relative entropy

$$H(\boldsymbol{b}|\boldsymbol{m}) = -\sum_i b_i[\ln(b_i/m_i) - b_i + m_i] \tag{1}$$

is then a useful measure of the differences between these two objects. It is easy to see that the relative entropy has the maximum $H = 0$ at $\boldsymbol{b} = \boldsymbol{m}$, *i.e.*, in the case of no data, optimization of H just gives the a priori object.

If there are some observational data, they will pull the reconstructed object away from the a priori one. It can be shown that the MEM selects a single object from a variety of objects consistent with the data. The other advantages of the MEM are as follows: (i) the least assumption about unknown information, (ii) introducing no structure in the image without evidence in the data, (iii) the possibility to express a priori information in terms of a default image rather than the properties of an algorithm itself [2, 4].

The chi-square statistic is usually used to describe the deviations between the observed and the reconstructed data sets. To write the χ^2 function in our case, note that the contribution of every spectrum to the flux in the pixel (λ, i) is given by $C_{A\lambda}\phi(i - x_{A\lambda})$. Denoting the observed flux by $F_{\lambda i}$ and the corresponding standard deviation by $\sigma_{\lambda i}$, we can write the following expression for χ^2:

$$\chi^2 = \sum_{\lambda,i} \frac{1}{\sigma^2_{\lambda i}} \left(\sum_A C_{A\lambda}\phi(i - x_{A\lambda}) - F_{\lambda i} \right)^2 . \tag{2}$$

Thus, in order to reconstruct our object, we should optimize the relative entropy (1), subject to the constraint that the chi-square function (2) has a value which is statistically compatible with the data (see below). Sometimes, however, it is worth using another constraint $G =$ const with some function $G(\boldsymbol{b})$ whose explicit form depends on the nature of the parameters under consideration. According to the general method of Lagrange multipliers, we construct the objective function

$$Q = H - \alpha\chi^2 - \beta G, \tag{3}$$

where α and β are the Lagrange multipliers.

The optimization problem then consists in finding a vector $\boldsymbol{b}$ such that the gradient of Q, *i.e.*, the vector $\boldsymbol{\nabla} Q$ whose components are equal to $\partial Q/\partial b_i$, is zero. Following Cornwell and Evans [2], we use the Newton-Raphson iterative optimization of the objective function (3) with automatic adjustment of the Lagrange multipliers α and β at every step of iteration. The algorithm starts with $\boldsymbol{b} = \boldsymbol{m}$ and $\alpha = \beta = 0$. The step to the next trial vector $\boldsymbol{\Delta b}$ is then given by

$$\boldsymbol{\Delta b} = (-\nabla\nabla Q)^{-1} \cdot \boldsymbol{\nabla} Q. \tag{4}$$

According to Eq. (4), in order to obtain $\boldsymbol{\Delta b}$ it is necessary to calculate a matrix inverse to the Hessian $\nabla\nabla Q$. Fortunately, in our case it is possible to do this directly because $\nabla\nabla Q$ is block diagonal with block dimension of at most N_{sp} or N_{PSF}. For this purpose, we use the singular value decomposition algorithm. Note that the inverse matrix $g = (-\nabla\nabla Q)^{-1}$ provides a useful metric for the M-dimensional parameter space, which can be used to judge the closeness of the actual vector $\boldsymbol{b}$ to the true MEM one, for adjusting the Lagrange multipliers α and β, and for a diagnostic of successful iteration (for more details, see [2]).

As the criterion to stop the iterations, we request at least one of the following two conditions to be satisfied:

$$\chi^2 \approx N \quad \text{or} \quad |\Delta\chi^2| < \varepsilon\chi^2, \tag{5}$$

where N is the number of observations, and ε is usually chosen of order 0.01 or less. The former condition in (5) is the usual chi-square test whereas the latter one means stopping due to the slow convergence of iteration, which especially takes place, if the other sets of parameters are not yet fitted.

Let S, F, and X denote symbolically the implementations of the MEM to process the spectra, the PSF, and the spectra positions, respectively. Then the general scheme of our algorithm may be displayed as a sequence of MEM steps:

$$Start: \ S \to [F \to X \to S] \to \cdots \to [F \to X \to S] \to Stop. \tag{6}$$

Thus, we start with some initial values of the spectra, the PSF, and the spectra positions whose choice is dictated by the wide use of a priori information, *e.g.*, flux positivity, magnitudes of the components, astrometric data, an observational estimate of FWHM for atmospheric seeing, *etc.* Each step of the algorithm contains the MEM iteration, as previously described. We optimize the relative entropy subject to the following constraints (*cf.* the objective function (3)): χ^2 and flux conservation during the processing of spectra, χ^2 and the PSF normalizing condition during the PSF processing, only χ^2 while the positions are being processed. The transitions from one step to another, which are indicated with the arrows in the scheme (6), are carried out after successfully checking the stopping conditions (5). Every step ends with putting $\boldsymbol{m} = \boldsymbol{b}$, *i.e.*, with improvement of our prior knowledge about the corresponding set of parameters. The brackets symbolize undivided parts of the algorithm. In other words, computations have to be finished only after processing the spectra. We proceed in this way because we are mainly interested in obtaining the spectra.

The corresponding computer program was written in the C language. At first, all the different segments as well as the whole program have been carefully tested and continuously improved, using simulated observational data without noise and with noise. The latter one essentially consists of the photon (Poisson) noise due to the sources and to the sky, and of the readout (Gaussian) noise of the CCD detector. Testing and debugging of the program clearly showed that the algorithm is very reliable, robust and works very fine for a large range of the input parameters.

APPLICATION TO Q1009-0252 A & B

The gravitational lens system Q1009-0252 A & B (magnitudes $m_V = 17.9$ and 20.5, angular separation $1''\!.55$) was identified during a systematic search for multiply imaged quasars within a sample of Highly Luminous Quasars carried out at the European Southern Observatory (La Silla, Chile). Various spectra were obtained with EMMI at the NTT (seeing between $0''\!.7$ and $1''\!.2$, slit width of $1''$, pixel size of $0''\!.44$, exposure times between 15 and 30 min, spectral resolution of about 10Å). Photometry and spectroscopy of Q1009-0252 have first been reported by Surdej *et al.* [5] who have found that the redshifts of both the A and B components are identical ($z = 2.74$) and the line profiles of Lyα+NV, CIV, and CIII] look very similar but there is an excess in the continuous spectrum of A near $\lambda = 5000$Å. They have proposed that Q1009-0252 A & B are multiple lensed images of the same quasar, and that the blue excess is possibly due to microlensing of image A. An independent observational study by Hewett *et al.* [3] has confirmed the gravitational lens nature of this system.

The spectra presented by Surdej *et al.* [5] were extracted, using a complete artisanal technique. In order to check those results, we decided to re-extract the spectra using the algorithm described here. Fig. 1 displays the extracted spectra for images A and B obtained by co-adding the results after processing the observational data from three CCD frames (exposure time of 30 min, seeing between $0''\!.7$ and $0''\!.85$) which have been recognized as being the best. The obtained spectra are in very good agreement with those reported in [5].

Interpretation of the spectral differences observed between the two lensed images Q1009-025 A & B in terms of differential macro-lensing, micro-lensing, extinction arising in the lensing galaxy as well as intrinsic source variations combined with the time delay has been proposed by Claeskens *et al.* [1].

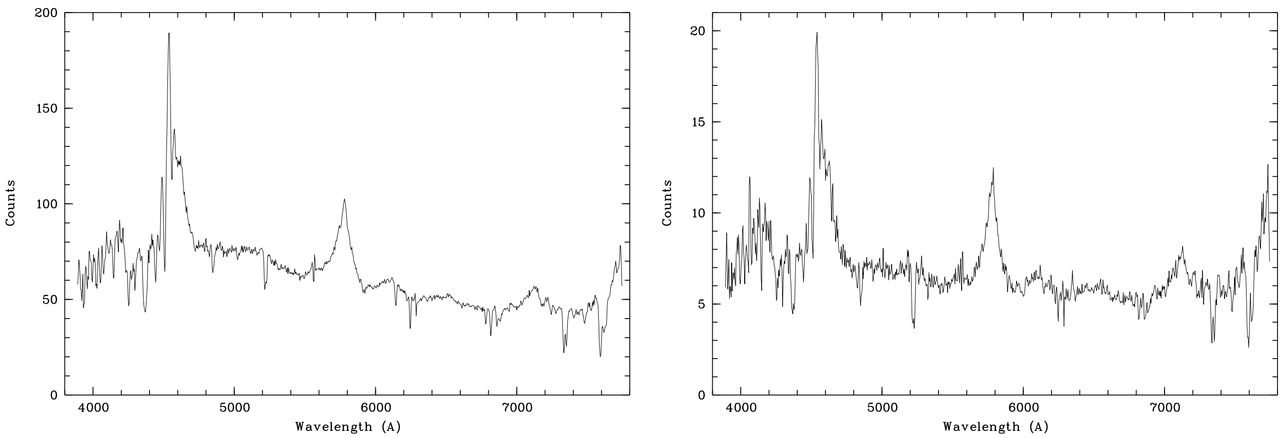

Figure 1: Spectra of Q1009-0252 A (left graph) and B (right graph) extracted using the proposed algorithm

CONCLUSIONS

The algorithm of optimal spectra extraction described in this report will be used in order to process some additional observations of gravitational lens systems observed in the framework of the ESO key program *Gravitational Lensing*. Finally, we intend to adapt this new algorithm to observations which will be obtained with the International 4-m Liquid Mirror Telescope (ILMT). This telescope ought to be constructed by the AMOS industrial company in Liège (Belgium) and be operated in Chile. A full description of this international project in which astronomers from Liège and from the Belgian Royal Observatory (Uccle) are involved is available at the URL: *http://vela.astro.ulg.ac.be/lmt*.

As to the gravitationally lensed quasar Q1009-0252 A & B, it should be emphasized that only a systematic spectro-photometric monitoring of this doubly imaged quasar will enable us to disentangle the various effects (macro-lensing, micro-lensing, differential extinction, time delayed variability, ...) which are probably at work for this gravitational lens system.

Acknowledgments. We are grateful to M. Remy for some fruitful discussions about the algorithm. One of us (S. Kh.) would like to thank the Belgian OSTC for the research fellowship granted to him within the programme of scientific and technical cooperation with Central & Eastern Europe. Our research was supported in part by PRODEX (Gravitational lens studies with HST), by contract P4/05 'Pôle d'attraction interuniversitaire' (OSTC, Belgium) and by the 'Fonds National de la Recherche Scientifique' (Belgium).

REFERENCES

[1] *Claeskens J.-F., Khmil S. V., Lee D.-W., Sluse D., Surdej J.* HST and ground-based observations of the gravitational lens Q1009-0252 A & B // Astron. and Astrophys.–2000.–accepted.

[2] *Cornwell T. J., Evans K. F.* A simple maximum entropy deconvolution algorithm // Astron. and Astrophys.–1985.–**143**, N 1.–P. 77–83.

[3] *Hewett P. C., Irwin M. J., Foltz C. G., et al.* The close-separation gravitational lens candidate Q1009-0252 // Astron. J.–1994.–**108**, N 5.–P. 1534–1541.

[4] *Skilling J., Bryan R. K.* Maximum entropy image restoration: general algorithm // Mon. Notic. Roy. Astron. Soc.–1984.–**211**, N 1.–P. 111–124.

[5] *Surdej J., Remy M., Smette A., et al.* Two new very close pairs of quasars with discordant redshift and a gravitational lens candidate // Gravitational Lenses in the Universe / Eds J. Surdej, D. Fraipont-Caro, E. Gosset, *et al.*–Université de Liège, 1994.–P. 153–159.