Publ. Obs. Astron. Belgrade No. 49 (1995), 15-18

STAR COUNT STUDY OF THE INTERSTELLAR DUST CLOUDS

L. G. BALÁZS
Konkoly Observatory, Budapest

Abstract. A method is outlined which allows one to estimate the optical absorption and the distance of the interstellar dust clouds. The procedure is based on the maximum likelihood approach of parameter estimation. The accuracy of the parameter estimation is briefly outlined.

1. STAR COUNTS AS INDICATORS OF THE SPATIAL DISTRIBUTION OF STARS

The location of the stars on the celestial sphere does not offer a direct indication on their spatial distribution. Actually, the two dimensional distribution on the celestial sphere is a projection of a three dimensional sample of objects. There are, however, some measurable properties of the stars like apparent brightness and proper motion which depend on some intrinsic property of the object and on the distance as well. Let us denote the measurable quantity with e and its intrinsic counterpart with E. Then there is a functional relationship

$$e = e(E, r).$$

In a number of cases the dependence of e on the distance r is a power law : r^n where $n = 1$ in case of proper motions and $n = 2$ in case of apparent brightness if there is no interstellar absorption.

According to the Bayes theorem of conditional probabilities we obtain

$$\phi(e) = \int \phi(e|r)g(r)dr$$

where $\phi(e)$, $\phi(e|r)$ and $g(r)$ mean the probability density of e, the conditional probability of e, given r, and the probability density of r, respectively. In this way we get an integral equation connecting the spatial distribution of the stars with the observed distribution assuming that the conditional probability of e given r is known somehow. $\phi(e)$ is estimated from the observations counting the stars in a certain Ω region of the sky and in the interval of $(e; e + de)$. The number of stars obtained in this way is often called star counts.

Due to the integral expression given above the star counts yield some information on the spatial distribution of the stars. If the $\phi(e|r)$ conditional probability is given the integral expression is a Freedholm-type integral equation. The solution of this equation gives the probability density of the stars in the line of sight. The estimation of the spatial distribution from the star counts i.e. inverting the integral expression is called the inverse approach. The opposite way when one makes some prediction for the star counts, assuming that the spatial distribution is known, is called direct approach.

2. SOLUTION OF THE BASIC EQUATION OF STELLAR STATISTICS

In the most common case the measured quantity is the apparent magnitude m in a certain color range :

$$m = M + 5log_{10}r - 5 + a(r).$$

where M denotes the absolute brightness, r the distance in the line of sight and $a(r)$ is the optical absorption. If M and $\rho = 5log_{10}r - 5a(r)$ are stochastically independent i.e the distribution of the absolute magnitudes does not depend on the distance we obtain

$$A(m) = \int_{-\infty}^{\infty} \Delta(\rho)\Phi(M)d\rho = \int_{-\infty}^{\infty} \Delta(\rho)\Phi(m - \rho)d\rho.$$

In this expression A, Δ, Φ are the probability densities of m, ρ, r, respectively. The integral equation obtained is often called the basic equation of the stellar statistics. It is a general equation, called equation of the convolution, giving the probability density of the sum of two independent probability variables.

There are several approaches to the solution of the basic equation. The serious trouble of the solution appears in the structure of the equation. Since the probability density of the observed quantity is estimated from the observations the function is never given exactly but instead by a quantity having a certain level of stochasticity. This stochastic noise is amplified during the inversion and makes the solution very unstable against the high frequency noise. The common solutions of the basic equation used in the astronomical praxis are summarized in the text books (e.g. Kurth 1967, Mihalas and Binney 1981). A more comprehensive discussion of the problem is given by Balázs (1995).

3. MODELLING STAR COUNTS

The opposite way, called direct approach needs the knowledge of the spatial distribution of the stars and, in the case of computing the distribution of apparent magnitudes, the luminosity function and the spatial distribution of the interstellar absorption.

There are several models for predicting star counts. The most frequently used ones are the models of Bahcall and Soneira (1980) and more recently those of Wainscoat et al. (1992). The models differ from each other on the way how they approximate the spatial distribution of the stars and the interstellar matter and the luminosity function. The BS model approximates the spatial distribution by a series of exponential disks, the scale heights are dependent on the absolute magnitude, and the halo by a de Vacauleurs (1959) spheroid. The luminosity function used by the model has an analytical form obtained by fitting the empirical data.

The Wainscoat model also has a series of exponential disks but uses dependence of the scale heights not on the absolute magnitudes but on the spectral types. The luminosity function is, correspondingly, a sum of gaussian components belonging to the different spectral types. Besides the exponential disk and halo the Wainscoat model adds further structural elements : bulge, spiral arms and molecular ring.

The cloudy structure of the interstellar matter is not taken into account in these models

4. EFFECT OF DUST CLOUDS ON THE STAR COUNTS

The star count models assume a spatially smoothed distribution of the interstellar absorbing material. According to the usual assumption the spatial distribution of the absorbing material is approximated by an exponential disk with a scale height of 80-100 pc

The presence of the individual dust clouds in the line of sight causes lower star counts in the cloud area in comparison with the surrounding field on the sky. This difference in star counts depends on the absorption and the distance of the cloud. The thickness of the cloud can be neglected.

Approximating the cloud with spatially homogeneous absorption then the a_{cl} absorption and r_{cl} distance can be treated as unknown parameters to be estimated. The estimation can be proceed using the maximum likelihood approach.

Let us have an observed sample of magnitudes $m_1, ..., m_2$ and suppose that there is an absorbing cloud in the line of sight at a distance of r_{cl} and an absorption of a_{cl}. By definition the likelihood function is written in the form

$$L(a_{cl}, r_{cl}) = \sum_{i=1}^{n} log A(m_i | a_{cl}, r_{cl}).$$

Here $A(m|a_{cl}, r_{cl})$ is obtained from the basic equation of the stellar statistics when inserted the model stellar distribution using an absorption law where a cloud of absorption a_{cl} and at a distance of r_{cl} was added to the general smoothed absorption. a_{cl} and r_{cl} are two unknown parameters to be estimated by maximizing the likelihood function.

Maximizing proceeds by solving the following system of equations :

$$\frac{\partial L}{\partial a_{cl}} = \frac{\partial L}{\partial r_{cl}} = 0.$$

In a general case the star counts are not homogeneous over the whole cloud area on the sky. We may generalize the procedure applied in the homogeneous case by dividing the cloud area into subregions where the star counts can be taken uniform.

We may perform the ML analysis independently within the subregions. However, the distance is the same for all of these regions and therefore we may reduce the number of parameters to be estimated by maximizing the following likelihood function :

$$L = \sum_{l=1}^{k} L^{(l)}(a_{cl}^{(l)}, r_{cl})$$

where k is the number of subregions and their likelihood functions and parameters to be estimated labelled by l.

Estimating the parameters it is an important question how reliable they are. There is a general theorem in case of ML estimations giving some hint for the accuracy of parameters obtained in this way. Denoting the value of the likelihood function at the maximum with L_{max} and those at the true value of the parameters with L_0 then

$$2(L_{max} - L_0) \approx \chi_m^2$$

is true asymptotically when the sample size goes to infinity. In the above formula the degree of freedom of the chi square probability variable m is the number of the parameters to be estimated.

Acknowledgements

This work was partly supported by the OTKA grant no. T4341.

References

Bahcall, J. N. and Soneira, R. M. : 1980, *Astrophys. Journ. Suppl.*, **44**, 73.

Balázs, L. G. : 1995, Inverse Problems, accepted for publication.

de Vaucouleurs, G. : 1959, in *Handbuch der Physik*, Vol. **53**, ed. S. Flügge, Springer, Berlin, 311.

Kurth, R. : 1967, *Introduction to the Stellar Statistics*, Pergamon Press.

Mihalas, D. and Binney, J. : 1981, *Galactic Astronomy*, W. H. Freeman and Co.

Wainscoat, R. J., Cohen, M., Volk, K., Walker, H. J. and Schwartz, D. E. : 1992, *Astrophys. Journ. Suppl.*, **83**, 111.