1983A&A...118..194N

Astron. Astrophys. 118, 194–196 (1983)

ASTRONOMY AND ASTROPHYSICS

*Research Note*

# Reconstruction of a polarized brightness distribution by the maximum entropy method

R. Nityananda and R. Narayan

Raman Research Institute, Bangalore-560080, India

Received June 23, accepted September 23, 1982

**Summary.** In the maximum entropy method of reconstructing a brightness distribution $B$ it has recently been recognized that a whole family of functions $f(B)$ can be used in place of the usual choices $f = \ln B$ and $f = -B \ln B$. In this note we suggest an analogous family of functions for polarized brightness distributions described by a correlation matrix $\boldsymbol{B}$. The entropy is shown to be the integral of the trace of $f(\boldsymbol{B})$ and reduces to the form given earlier by Ponsonby for the $\ln B$ case. We obtain an expression for the gradient of the entropy which is analogous to that for the unpolarized case and can be used in a similar way in an iterative scheme for finding the maximum entropy solution. Further, we show that the entropy is a convex function of the brightness matrix when $f'' < 0$, thus guaranteeing a unique maximum entropy solution consistent with the measurements.

**Key words:** maximum entropy – polarization – aperture synthesis

## 1. Introduction

The Maximum Entropy Method (MEM) of restoring a brightness distribution/spectrum from partial knowledge of its Fourier coefficients (Burg, 1967) was strongly advocated for astronomical applications by Ables (1972). In this technique one is required to optimize the unmeasured Fourier coefficients by maximizing the entropy expression $E$ given below, subject to the measurements as constraints.

$$E = \int_0^1 \int_0^1 f[B(x, y)]dxdy. \tag{1}$$

Here $x$ and $y$ are fractional sky coordinates running from 0 to 1 and $B(x, y)$ is the brightness distribution. Two forms of $f$ are common in the literature.

$$f_1 \equiv \ln(B); \quad f_2 \equiv -B \ln B \tag{2}$$

(Burg, 1967; Ables, 1974 for $f_1$; Frieden, 1972; Gull and Daniell, 1978 for $f_2$). In recent times there has been much debate on the foundations and relative merits of these two forms. However, Högbom (1978) and Subrahmanya (1978) both suggested that the precise form of the function $f$ is not crucial so long as it has the general feature of discouraging ripple and negative brightness values. Hefferman and Bates (1982), in the context of reconstruction from projections, have recently given examples to show that the detailed form of the function being maximized is not important. This viewpoint has received strong support in our recent studies (Nityananda and Narayan, 1982, henceforth NN) where we show that any form of $f$ with the general properties

$$d^2f/dB^2 \equiv f'' < 0; \quad d^3f/dB^3 \equiv f''' > 0 \tag{3}$$

will produce "good" reconstructions with flat baselines and sharp peaks. Burg (1975) proved the important theorem that the reconstructed $B(x, y)$ for any particular choice of $f$ is unique provided the condition $f'' < 0$ is satisfied. This has considerably helped in developing iterative numerical schemes to implement the MEM.

An extension of the MEM with the entropy form $f_1$ to polarized brightness distributions was given by Ponsonby (1973) who suggested that one should maximize

$$E = \int_0^1 \int_0^1 \ln[I^2(x, y) - Q^2(x, y) - U^2(x, y) - V^2(x, y)]dxdy \tag{4}$$

subject to the measurements as constraints. Here $I$, $Q$, $U$, and $V$ are the usual Stokes parameters (see e.g., Landau and Lifshitz, 1975). Ponsonby (1973) did not discuss schemes to obtain the MEM reconstruction. In this paper, we give the appropriate generalization of Ponsonby's (1973) entropy expression (4) for the case when one has a general form for $f$ and not just the logarithm. We then show that the numerical schemes that have been developed for the "scalar" brightness case can be directly generalized by introducing the "matrix" brightness. Finally we show that $f'' < 0$ is a sufficient condition for uniqueness of the polarized MEM reconstruction, thus completely establishing the close parallel between the unpolarized and polarized problems.

## 2. Matrix generalization of the MEM for the polarized case

A polarized brightness distribution can be described using a two by two matrix $\boldsymbol{B}(x, y)$ whose diagonal elements give the mean square electric field along two orthogonal (spatial) directions. The off diagonal elements give the complex correlation between the two orthogonal electric field components. The correlation matrix $\boldsymbol{B}$ is related to the Stokes parameters $I$, $Q$, $U$, and $V$ as follows (e.g., Landau and Lifshitz, 1975)

$$\boldsymbol{B} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} \dfrac{I+Q}{2} & \dfrac{U+iV}{2} \\ \dfrac{U-iV}{2} & \dfrac{I-Q}{2} \end{bmatrix}. \tag{5}$$

---

*Send offprint requests to:* R. Narayan

© European Southern Observatory • Provided by the NASA Astrophysics Data System

1983A&A...118..194N



All the quantities occurring in (5) are functions of the two sky coordinates $(x, y)$ and an interferometer with antennas separated by a baseline $(X, Y)$ measures the Fourier transform of $\boldsymbol{B}$ for this spacing. In this notation, the entropy expression (4) given by Ponsonby (1973) reads

$$\begin{aligned} E-\ln 4 &= \int_0^1\int_0^1 \ln[I^2-Q^2-U^2-V^2]dxdy-\ln 4 \\ &\equiv \int_0^1\int_0^1 \ln[\det \boldsymbol{B}]dxdy \equiv \int_0^1\int_0^1 \mathrm{Trace}\,[\ln \boldsymbol{B}]dxdy \\ &\equiv \int_0^1\int_0^1 [\ln\lambda_1+\ln\lambda_2]dxdy, \end{aligned} \tag{6}$$

where we have given four equivalent forms. Here $\lambda_1$ and $\lambda_2$ are the eigenvalues of the matrix $\boldsymbol{B}$, given by

$$\lambda_{1,2}=\tfrac{1}{2}[I\pm(Q^2+U^2+V^2)^{1/2}]; \quad \lambda_1>\lambda_2. \tag{7}$$

Physically, $\lambda_1$ and $\lambda_2$ represent the intensities of two orthogonally polarized, mutually incoherent components into which the incident radiation can be resolved. The final expression in (6) is just the sum of the usual $\ln B$ entropy for each of these components. This is precisely the argument used by Ponsonby (1973) to write down Eq. (4). The generalization to other forms of entropy is now straightforward and leads to

$$E=\int_0^1\int_0^1 [f(\lambda_1)+f(\lambda_2)]dxdy \quad \equiv \int_0^1\int_0^1 \mathrm{Trace}\,[f\{\boldsymbol{B}(x,y)\}]dxdy. \tag{8}$$

The above generalization is physically motivated, whereas an alternative generalization of (6) such as

$$E'=\int_0^1\int_0^1 f[\det\{\boldsymbol{B}(x,y)\}]dxdy \tag{9}$$

would have no such basis.

Because of the contraints imposed by the measurements, while maximizing $E$ we should consider variations in $\boldsymbol{B}$ corresponding only to unmeasured spatial frequencies i.e.,

$$\delta\boldsymbol{B}(x,y)=\sum_{X_u, Y_u}\delta\,\boldsymbol{\varrho}(X_u, Y_u)e^{2\pi i(xX_u+yY_u)}, \tag{10}$$

where the subscript $u$ stands for "unknown" and we denote the Fourier coefficients of $\boldsymbol{B}$ by $\boldsymbol{\varrho}(X, Y)$. For a variation of the type (10), the change in $E$ in Eq. (8) is

$$\delta E=\sum_{X_u}\sum_{Y_u}\int_0^1\int_0^1 \mathrm{Trace}\,[f'\{\boldsymbol{B}(x,y)\}\delta\boldsymbol{\varrho}(X_u, Y_u)e^{2\pi i(xX_u+yY_u)}]dxdy. \tag{11}$$

The above result is correct even when the matrices $\boldsymbol{B}$ and $\delta\boldsymbol{B}$ do not commute because the trace operation is invariant under cyclic permutation of a product of matrices. Writing Eq. (11) explicitly in terms of matrix elements and differentiating, we have

$$\frac{\partial E}{\partial\varrho_{kl}(X_u, Y_u)}=\sigma_{lk}(-X_u, -Y_u), \tag{12}$$

where we denote the Fourier coefficients of $f'(\boldsymbol{B})$ by $\boldsymbol{\sigma}(X, Y)$. Since $\boldsymbol{B}$ is a Hermitian matrix, so is $f'(\boldsymbol{B})$, and hence

$$\frac{\partial E}{\partial\varrho_{kl}(X_u, Y_u)}=\sigma^*_{kl}(X_u, Y_u). \tag{13}$$

Equation (13) shows that the component of the gradient of $E$ corresponding to any unknown Fourier coefficient of $\boldsymbol{B}$ is the complex conjugate of the corresponding Fourier coefficient of $f'(\boldsymbol{B})$. This is analogous to the result for the scalar brightness case. Equation (13) can be directly used to develop a gradient type algorithm to numerically maximize $E$ (see NN for the scalar case). Further (13) shows that when $E$ is a maximum, all the $\sigma_{kl}(X_u, Y_u)$ are identically equal to zero, implying that $f'(\boldsymbol{B})$ is band-limited. This result is again analogous to that in the scalar case and can be made the basis of fixed point algorithms to obtain the MEM solution (e.g., Gull and Daniell, 1978; Willingale, 1981, for the unpolarized problem).

## 3. Uniqueness of the polarized MEM reconstruction

Consider two polarized brightness distributions $\boldsymbol{B}_1(x, y)$ and $\boldsymbol{B}_2(x, y)$ both of which exactly fit the set of interferometer measurements. Let us assume that both maximize $E$ given in (8). Since the spatial correlations $\varrho_{kl}(X, Y)$ are linear in $\boldsymbol{B}(x, y)$, any linear combination

$$\boldsymbol{B}_p(x,y)=(1-p)\boldsymbol{B}_1(x,y)+p\boldsymbol{B}_2(x,y); \quad 0\leqq p\leqq 1 \tag{14}$$

automatically satisfies the measured constraints. Also, if $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ satisfy at each $(x, y)$ the physical realizability conditions $I\geqq 0$ and $I^2\geqq Q^2+U^2+V^2$, then so does $\boldsymbol{B}_p$.

Consider any point $(x, y)$. We have

$$\begin{aligned} \frac{d^2}{dp^2}[\mathrm{Trace}\,\{f(\boldsymbol{B}_p)\}] &= f''(\lambda_1)\left(\frac{d\lambda_1}{dp}\right)^2+f''(\lambda_2)\left(\frac{d\lambda_2}{dp}\right)^2 \\ &\quad +\left[f'(\lambda_1)\frac{d^2\lambda_1}{dp^2}+f'(\lambda_2)\frac{d^2\lambda_2}{dp^2}\right], \end{aligned} \tag{15}$$

where $\lambda_1$ and $\lambda_2$ are the eigenvalues of $\boldsymbol{B}_p$. With a little algebra it is easily shown that this quantity is negative at all $(x, y)$ if $f''<0$. Since this is clearly impossible if both $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ are maxima of $E$, we thus conclude that the MEM solution is unique.

The above uniqueness proof cannot be applied to the general $n(>2)$ multichannel case. This general problem (which, however, has no application in aperture synthesis) still awaits a solution.

## 4. Conclusions

The main results of this paper are:

a) We have obtained a general expression [Eq. (8)] for the entropy $E$ of a polarized brightness map $\boldsymbol{B}$ in terms of a general function $f$ [which should satisfy the constraints in (3) for useful results].

b) We have obtained an expression [Eq. (13)] for the gradient of $E$ with respect to unmeasured Fourier coefficients of $\boldsymbol{B}$ and shown that the map $\boldsymbol{B}_{\max}$ which maximizes $E$ has $f'(\boldsymbol{B}_{\max})$ band-limited. Iterative numerical schemes to obtain $\boldsymbol{B}_{\max}$ can be developed from these results as for the scalar brightness case (NN).

c) We have shown that if $f''<0$ then there is only one solution $\boldsymbol{B}_{\max}$ which maximizes $E$. This will lend robustness to numerical algorithms since convergence to the unique solution is assured.

d) In an earlier paper (NN) we have described schemes to control the resolution and sensitivity to noise of the MEM for scalar brightness maps. The same techniques can now be employed for polarized maps because of b) and c) above.

© European Southern Observatory • Provided by the NASA Astrophysics Data System

1983A&A...118..194N



**References**

Ables, J.G.: 1972, Symposium on the Collection and Analysis of Astrophysical Data, published as *Astron. Astrophys. Suppl.* **15**, 383 (1974)

Burg, J.P.: 1967, Paper presented at the Annual International Society of Exploration Geophysicists Meeting, Oklahoma City, reprinted in Modern Spectrum Analysis, ed. D. G. Childers, IEEE Press (1978)

Burg, J.P.: 1975, Thesis "Maximum Entropy Spectral Analysis", Dept. of Geophys., Stanford Univ.

Frieden, B.R.: 1972, *J. Opt. Soc. Amer.* **62**, 511

Gull, S.F., Daniell, G.J.: 1978, *Nature* **272**, 686

Hefferman, P.B., Bates, R.H.T.: 1982, *Optik* **60**, 129

Högbom, J.A.: 1979, in *Image Formation from Coherence Functions in Astronomy*, ed. C. van Schooneveld, Reidel, Dordrecht, Holland, p. 237

Landau, L.D., Lifshitz, E.M.: 1975, The Classical Theory of Fields, Pergamon Press, Oxford, U.K., p. 119

Nityananda, R., Narayan, R.: 1982 (to appear in *J. Astrophys. Astr.*)

Ponsonby, J.E.B.: 1973, *Monthly Notices Roy. Astron. Soc.* **163**, 369

Subrahmanya, C.R.: 1978, in *Image Formation from Coherence Functions in Astronomy*, ed. C. van Schooneveld, Reidel, Dordrecht, Holland, p. 287

Willingale, R.: 1981, *Monthly Notices Roy. Astron. Soc.* **194**, 359

© European Southern Observatory • Provided by the NASA Astrophysics Data System