Astron. Astrophys. 176, 159–170 (1987)

ASTRONOMY
AND
ASTROPHYSICS

Maximum entropy method for polarized images

R.K. Shevgaonkar

Indian Institute of Astrophysics, Bangalore-560034, India

Received September 16, 1985; accepted October 9, 1986

Summary. A simple maximum entropy method algorithm for deconvolving polarized images obtained from synthesis observations has been described here. A general purpose Fortran program is also developed. It is demonstrated that although the a priori positivity condition is imposed on the total intensity image only, the polarized component images also get improved remarkably.

Key words: image processing – maximum entropy – radio astronomy – Stokes parameters – polarized images

1. Introduction

Polarization measurements on the radio astronomical sources have provided a new dimension in the understanding of the astrophysical processes responsible for the cosmic radiation. The radio sources in the sky exhibit a wide variety of polarization characteristics. For example, the extended non-thermal emission from our galaxy, and localized emission from certain objects like supernova remnants show a low degree of linear polarization, whereas some radio sources like flare stars, Sun, Jupiter etc. exhibit a high degree of circular polarization and a very little or no linear polarization. To study the astronomical objects in detail high resolution polarization observations are required. The aperture synthesis technique has been extended for obtaining high resolution polarized images of the brightness distribution in the sky. Each interferometer pair used for synthesizing an aperture acquires four correlation coefficients corresponding to the four Stokes parameters of the brightness distribution. If all the four correlations are measured for all interferometer pairs one obtains identical spatial Fourier coverage or identical synthesized beam for all the Stokes parameters. However, due to practical reasons there could be cases when all four correlations could not be measured over all interferometer pairs. This gives different spatial coverages and therefore different synthesized beams for different polarization components. Due to different convolving beams for the four Stokes images it is not unlikely to encounter a situation when some features in the observed image exhibit greater than 100% polarization which by definition would be unacceptable. This situation is similar to the one in the simple intensity imaging where, due to improper and finite sampling of the Fourier plane, the observed image has negative excursions.

The deconvolution method CLEAN, proposed by Högbom (1974) has become a popular technique to take away the interference due to the sidelobes of the synthesized beam from the observed image. The initial version of CLEAN was suitable only for deconvolving the well-isolated point-like sources. However, with subsequent modifications CLEAN has become a routine technique for deconvolving all types of brightness distributions. Indeed, the performance of the method depends to large extent upon how empty a given field of view is.

Presently, due to lack of any other technique, CLEAN is applied to the polarized images also (Högbom and Carlsson, 1974). The four Stokes images are CLEANed independently with their respective synthesized beams. However, since the performance of the method deteriorates for extended complex distributions, even after CLEANing, the images do not guarantee the two essential conditions namely, the total intensity is positive definite at every point of the map and the degree of polarization is less than unity. Solar astronomers in particular would be well aware of the limitations of CLEAN, as invariably they encounter situations where the degree of polarization is greater than unity in CLEANed maps.

The Maximum Entropy Method (MEM) uses a radically new approach for reconstructing synthesis images (Burg, 1967). Apart from the fact that MEM is an unconventional nonlinear technique and has a better mathematical formulation, there is a provision to incorporate the a priori knowledge about the image in the MEM. In the past, MEM has been applied successfully to the total intensity images. Many efficient algorithms have been developed in different fields (Gull and Daniell, 1978; Willingale, 1981; Skilling and Bryan, 1984; Cornwell and Evans, 1985). However, MEM has not been applied to the polarized images so far. The measure of the entropy of a polarized distribution was derived by Ponsonby (1973) for $\ln B$ – form of entropy which subsequently has been generalized by Nityananda and Narayan (1983) for any arbitrary entropy function. Although, MEM was thought as an alternative to CLEAN for polarized images almost a decade ago, and few computer simulations of the method with polarized maps were presented recently (Narayan and Nityananda, 1984), until now no detail formulation and algorithm of the method has been presented. In this paper we present a detail formulation of MEM for the polarized images. We also give a simple numerical implementation of the method. By simulated examples we demonstrate the characteristics of the method. The main advantage of MEM over CLEAN is that it treats all the four Stokes images simultaneously and imposes the essential conditions on the reconstructed image automatically.

2. Entropy of a polarized image

A polarized brightness distribution in general can be represented by a two-by-two matrix as

$$\mathbf{B}(x,y)=\frac{1}{2}\begin{bmatrix} I(x,y)+Q(x,y) & U(x,y)+iV(x,y) \\ U(x,y)-iV(x,y) & I(x,y)-Q(x,y) \end{bmatrix} \tag{1}$$

where (x, y) are the spatial image coordinates and I, Q, U, V are the four Stokes parameters. As a general case the polarization characteristics of the brightness distribution are space variant and I, Q, U, V are arbitrary functions of (x, y). The space variant degree of polarization can be defined as

$$d(x,y)=\frac{[Q^2(x,y)+U^2(x,y)+V^2(x,y)]^{1/2}}{I(x,y)} \tag{2}$$

Following previous authors (Ponsonby, 1973; Nityananda and Narayan, 1983; Gull and Skilling, 1984) the entropy of an arbitrary polarized distribution is equal to the sum of the entropies of the two orthogonal polarization components of the polarized image. Since the eigen values of the polarized brightness matrix $\mathbf{B}(x, y)$ essentially indicate two orthogonal mutually incoherent polarized components, the entropy of the polarized brightness distribution $\mathbf{B}(x, y)$ is equal to the sum of the entropies of the two eigen values of matrix $\mathbf{B}(x, y)$, giving

$$E(x,y)=f[\lambda_1(x,y)]+f[\lambda_2(x,y)] \tag{3}$$

where $\lambda_1(x, y)$ and $\lambda_2(x, y)$ are two eigen values of the matrix $\mathbf{B}(x, y)$ and f is the entropy function. For Burg's definition of the entropy, the function $f(\psi)=\ln\psi$ whereas from information theoretic definition of the entropy $f(\psi)=-\psi\ln\psi$.

It is worth mentioning here that as far as the image processing is concerned, MEM has been justified on purely heuristic grounds. Recent work (Nityananda and Narayan, 1982; Cornwell and Evans, 1985; Shevgaonkar, 1986b) clearly indicates that the entropy of an image is some measure of goodness of the image and is just a mean of providing the a priori knowledge about the true brightness distribution. The present pragmatic view of MEM predicts a variety of entropy functions which are capable of giving more or less similar reconstructions. Therefore apart from the two entropy functions mentioned above, one could also have a power law form ($f(\psi)=\psi^m$) of entropy. The value of m is such that the second derivative of $f(\psi)$ with respect to ψ is less than zero and the third derivative of $f(\psi)$ is greater than zero (Nityananda and Narayan, 1982). However, it should be noted that the former condition on $f(\psi)$ is the uniqueness condition and should be strictly satisfied whereas the latter condition can be relaxed depending upon the type of brightness distribution (see Shevgaonkar, 1986b).

Integrating Eq. (3) over the image the entropy of the brightness distribution $\mathbf{B}(x, y)$ is

$$E=\iint\{f[\lambda_1(x,y)]+f[\lambda_2(x,y)]\}\,dx\,dy \tag{4}$$

or

$$E=\iint \mathrm{tr}\,[f\{\mathbf{B}(x,y)\}]\,dx\,dy \tag{5}$$

Equation (5) is similar to the expression obtained in the past by Nityananda and Narayan (1983).

3. MEM for polarized images

If $\rho^I(u,v)$, $\rho^Q(u,v)$, $\rho^U(u,v)$ and $\rho^V(u,v)$ are the spatial visibility functions or Fourier transforms of I, Q, U and V respectively, the spatial autocorrelation matrix or the visibility matrix of a polarized brightness distribution can be written as

$$\boldsymbol{\rho}(u,v)=\frac{1}{2}\begin{bmatrix} \rho^I(u,v)+\rho^Q(u,v) & \rho^U(u,v)+i\rho^V(u,v) \\ \rho^U(u,v)-i\rho^V(u,v) & \rho^I(u,v)-\rho^Q(u,v) \end{bmatrix} \tag{6}$$

where (u, v) are the spatial coordinates and the brightness matrix $\mathbf{B}(x, y)$ is related to the visibility matrix $\boldsymbol{\rho}(u, v)$ through a Fourier transform relationship i.e.,

$$\mathbf{B}(x,y)=\sum_u\sum_v \boldsymbol{\rho}(u,v)e^{-i2\pi(ux+vy)} \tag{7}$$

Due to finite and non-uniform sampling of the synthesized aperture, the visibility matrix $\boldsymbol{\rho}$ is unknown over certain position of the uv-plane. The maximum entropy method tries to estimate the unknown $\boldsymbol{\rho}$'s such that the entropy integral in Eq. (5) is maximized, or in other words MEM generates that image for which the derivative of E with respect to unknown $\boldsymbol{\rho}$'s is zero.

Substituting Eq. (7) in (5) and differentiating with respect to the unknown autocorrelations we get the gradient of the entropy as

$$g^k(u,v)\equiv\frac{\partial E}{\partial\rho^k(u,v)}=\iint \mathrm{tr}\left[\frac{\partial f(\mathbf{B})}{\partial\mathbf{B}}\cdot\frac{\partial\boldsymbol{\rho}}{\partial\rho^k}\right]\cdot e^{-i2\pi(ux+vy)}\,dx\,dy \tag{8}$$

superscript k could be any one of I, Q, U or V.

From Eq. (8) it is clear that the gradient of the entropy with respect to the unknown visibilities can be obtained just through a Fourier transform and therefore it appears straightforward to implement an iterative gradient technique to maximize the entropy. However, to compute the gradient, the first step would be to find the $\partial f(\mathbf{B})/\partial\mathbf{B}$ which is slightly cumbersome as $\mathbf{B}$ is a two-by-two matrix. A simple elegant way of computing the derivative of $f(\mathbf{B})$ with respect to $\mathbf{B}$ is as follows.

Let us split the brightness matrix $\mathbf{B}(x, y)$ into four two-by-two matrices having elements 0 or ±1 as

$$\mathbf{B}(x,y)=I(x,y)\begin{bmatrix}1&0\\0&1\end{bmatrix}+Q(x,y)\begin{bmatrix}1&0\\0&-1\end{bmatrix}+U(x,y)\begin{bmatrix}0&1\\1&0\end{bmatrix}+iV(x,y)\begin{bmatrix}0&1\\-1&0\end{bmatrix} \tag{9}$$

Without losing generality, the factor $\frac{1}{2}$ in Eq. (1) has been dropped out for simplicity.

If we now define a unit vector $\boldsymbol{n}$ as

$$\boldsymbol{n}\equiv(n_1,n_2,n_3)=\frac{(Q,U,V)}{(Q^2+U^2+V^2)^{1/2}}=\frac{1}{Id}(Q,U,V) \tag{10a}$$

and a vector $\boldsymbol{\alpha}\equiv(\alpha_1,\alpha_2,\alpha_3)$ where

$$\alpha_1\equiv\begin{bmatrix}1&0\\0&-1\end{bmatrix},\quad \alpha_2\equiv\begin{bmatrix}0&1\\1&0\end{bmatrix},\quad \alpha_3\equiv i\begin{bmatrix}0&1\\-1&0\end{bmatrix} \tag{10b}$$

the brightness matrix $\mathbf{B}$ can be expressed in terms of $\boldsymbol{n}$ and $\boldsymbol{\alpha}$ as

$$\mathbf{B}=I[\boldsymbol{I}+d\boldsymbol{\alpha}\cdot\boldsymbol{n}] \tag{11}$$

where, $\boldsymbol{I}$ is a two-by-two unit matrix. $\boldsymbol{I}$ and I should not be confused.

The advantage of Eq. (11) over the matrix notation is that due to the simplifying property of $(\boldsymbol{\alpha}\cdot\boldsymbol{n})$ i.e.,

$$(\boldsymbol{\alpha}\cdot\boldsymbol{n})^r = \boldsymbol{I} \quad \text{if } r \text{ is even} \tag{12a}$$

$$= \boldsymbol{\alpha}\cdot\boldsymbol{n} \quad \text{if } r \text{ is odd}, \tag{12b}$$

the derivation of $\partial f(\mathbf{B})/\partial\mathbf{B}$ is less cumbersome.

The basic trick is to first find an expression for $\partial f(\mathbf{B})/\partial\mathbf{B}$ assuming $\mathbf{B}$ is a scalar quantity and then using Eq. (12a,b) separate the coefficients of unit matrix $\boldsymbol{I}$ and $(\boldsymbol{\alpha}\cdot\boldsymbol{n})$. The coefficient of the unit matrix represents the gradient of the entropy with respect to the Stokes parameter I. Similarly the coefficients of the three terms of $(\boldsymbol{\alpha}\cdot\boldsymbol{n})$ give the derivatives of the entropy with respect to Q, U and V respectively.

To explain the technique more clearly let us take the well-known entropy function $\ln\mathbf{B}$. Then assuming $\mathbf{B}$ is just a scalar variable

$$\frac{\partial f(\mathbf{B})}{\partial\mathbf{B}} = \mathbf{B}^{-1} = I^{-1}[\boldsymbol{I} + d\boldsymbol{\alpha}\cdot\boldsymbol{n}]^{-1} \tag{13}$$

Since the degree of polarization $d < 1$ and $|\boldsymbol{\alpha}\cdot\boldsymbol{n}| = \boldsymbol{I}$, the bracket in Eq. (13) can be expanded binomially to give

$$\frac{\partial f(\mathbf{B})}{\partial\mathbf{B}} = I^{-1}[\boldsymbol{I} - d(\boldsymbol{\alpha}\cdot\boldsymbol{n}) + d^2(\boldsymbol{\alpha}\cdot\boldsymbol{n})^2 - \cdots] \tag{14}$$

Using identities (12a,b) relation (14) can be simplified to

$$\frac{\partial f(\mathbf{B})}{\partial\mathbf{B}} = I^{-1}[\boldsymbol{I} - d(\boldsymbol{\alpha}\cdot\boldsymbol{n}) + d^2\boldsymbol{I} - \cdots.] \tag{15}$$

Separating terms of $\boldsymbol{I}$ and $(\boldsymbol{\alpha}\cdot\boldsymbol{n})$ we get

$$\frac{\partial f(\mathbf{B})}{\partial\mathbf{B}} = \frac{\boldsymbol{I}}{I(1-d^2)} - \frac{d(\boldsymbol{\alpha}\cdot\boldsymbol{n})}{I(1-d^2)} \tag{16}$$

Substituting Eqs. (16) and (6) in (8) the component of the Trace along the Stokes parameter I is

$$\mathrm{tr}_I = \mathrm{tr}\left[\frac{\partial f(\mathbf{B})}{\partial\mathbf{B}}\cdot\frac{\partial\boldsymbol{\rho}}{\partial\rho^I}\right] = \frac{2}{D}I \tag{17}$$

where D is the determinant of the brightness matrix $= (I^2 - Q^2 - U^2 - V^2)$.

Similarly the trace components along Stokes parameters Q, U, V are

$$\mathrm{tr}_{Q,U,V} = \mathrm{tr}\left[\frac{\partial f(\mathbf{B})}{\partial\mathbf{B}}\cdot\frac{\partial\boldsymbol{\rho}}{\partial\rho^{Q,U,V}}\right] = -\frac{2}{D}(Q, U, V) \tag{18}$$

Expressions (17) and (18) are similar to the ones obtained by Ponsonby (1973) for computing the gradient of the entropy of a polarized distribution.

Following similar steps as mentioned above the components of the trace for other entropy functions like $-\mathbf{B}\ln\mathbf{B}$ or $\mathbf{B}^m$ can be derived. The formulation of the method for other entropy function is desired as depending upon the type of distribution other functions may be superior to the $\ln B$ entropy function.

For $f(\mathbf{B}) = -\mathbf{B}\ln\mathbf{B}$,

$$\mathrm{tr}_I = -2 - \ln D \tag{19a}$$

and

$$\mathrm{tr}_{Q,U,V} = -\frac{(Q, U, V)}{\Sigma}\ln\left\{\frac{I+\Sigma}{I-\Sigma}\right\} \tag{19b}$$

where $\Sigma = (Q^2 + U^2 + V^2)^{1/2}$.

The constant -2 in Eq. (19a) gives a uniform value over the image which, when Fourier transformed, contributes only to the zero lag spatial autocorrelation or the integrated power of the brightness distribution. Since the integrated power of the image has to be kept unchanged the constant -2 can be dropped from Eq. (19a) for the computation of the gradient.

Finally let us write the expression for the trace for entropy function $f(\mathbf{B}) = \mathbf{B}^m$. For power law form of entropy we get

$$\mathrm{tr}_I = \frac{m}{2}\{(I+\Sigma)^{m-1} + (I-\Sigma)^{m-1}\} \tag{20a}$$

and

$$\mathrm{tr}_{Q,U,V} = \frac{m}{2\Sigma}\{(I+\Sigma)^{m-1} - (I-\Sigma)^{m-1}\} \tag{20b}$$

Knowing the trace the gradient of the entropy can be computed through a Fourier transform of the trace.

4. Numerical implementation and results

We now describe the numerical implementation of the method formulated in the previous section. Since the computation of the entropy gradient requires only one Fourier transform operation we have adopted the gradient technique for the entropy maximization.

To start with we have measured visibility coefficients $\rho^k(u_i^k, v_i^k)$ $(k = I, Q, U, V)$ corresponding to the four Stokes parameters I, Q, U, V respectively over a limited spatial region, and we want to estimate the value of $\rho^k(u_i^k, v_i^k)$ over an unmeasured region, satisfying the condition that the integral (5) is maximized and the measured quantities are unchanged, i.e.,

$$\rho^k(u_i^k, v_i^k) = \iint_{\text{field}} \mathbf{B}(x,y)e^{i2\pi(u_i^k x + v_i^k y)}\,dx\,dy \tag{21}$$

$$u_i^k = -M^k, -M^k + 1, \ldots 0, \ldots, M^k$$

$$v_i^k = -N^k, -N^k + 1, \ldots 0, \ldots, N^k$$

are exactly fitted to the measurements. (u_i^k, v_i^k) is the location of the ith visibility coefficient of the kth Stokes parameter. M^k and N^k defines the maximum extent of the measured Fourier coverage for the kth Stokes parameter. This generalization relaxes the requirement that the Fourier coverages are identical for all the four Stokes parameters. The integral in Eq. (21) is over the total solid angle observed by the primary beam of the synthesizing elements.

We start with a trial polarized image which is usually the same as the one obtained by direct Fourier transform of the measured visibility coefficients. The unmeasured visibility coefficients are assumed to be zero. In practice the integrations are replaced by discrete summations over a grid of $N_0 \times M_0$ points where N_0 and M_0 are chosen to be binary numbers to implement conveniently the Fast Fourier Transform (FFT). Using Eq. (8) the gradient of the entropy is computed. Without disturbing the measured visibilities the unmeasured visibilities only are shifted in the direction of the entropy gradient giving

$$\rho_{j+1}^k(u_i^k, v_i^k) = \rho_j^k(u_i^k, v_i^k) + xg_j^k(u_i^k, v_i^k) \tag{22}$$

The suffixes j and $(j+1)$ represent the values of the corresponding parameters in jth and $(j+1)$th iteration. x is a suitable constant which can possibly be chosen by the following criterion.

Let us first displace ρ_j^k by a small amount Δ in the direction of the gradient g_j^k. Let the gradient of the entropy at the shifted point be $g_j^{k'}$. If the entropy E is a quadratic function of ρ^k a choice of

$$x = \frac{\Delta (g_j^k)^2}{(g_j^k)^2 - g_j^{k'} \cdot g_j^k} \tag{23}$$

will bring the entropy E to the maximum along the search direction g_j^k. However, since E is not a quadratic function in general, a value of x decided by Eq. (23) may provide inadequate or excessive shifts. The excessive shifts can unnecessarily drag the solution away from the true solution and therefore it is advisable to put an upper limit on the choice of x.

An improvement over the simple gradient method is the conjugate gradient method due to Fletcher and Reeves (1964). In this method the search direction depends not only upon the current gradient but also upon the previous search directions. In this method the shift in jth iteration is made along the conjugate gradient direction given by

$$cg_j^k = g_j^k + \left\{\frac{g_j^k}{g_{j-1}^k}\right\}^2 \cdot cg_{j-1}^k \tag{24}$$

The unmeasured visibility coefficients after jth iteration are modified to

$$\rho_{j+1}^k = \rho_j^k + ycg_j^k \tag{25}$$

Using the property that at the maximum of the entropy the conjugate gradient is orthogonal to the gradient and therefore the dot product of the two vectors g^k and cg^k is zero, an optimum value for the shift constant y can be obtained. Fourier transforming the modified unmeasured visibility coefficients we obtain a new image which will be the trial model for the next iteration. The iteration continues till a satisfactory convergence is reached. Here, the square of the gradient of the entropy has been used to decide the convergence.

Although the numerical implementation of the algorithm appears straightforward there are certain practical difficulties. It can be seen very easily that entropy of a polarized brightness distribution given by Eq. (5) is defined only for those distributions which satisfy two essential conditions namely $I > 0$ and $d < 1$. The entropy becomes imaginary if any of these conditions is not satisfied. Although, this means that the maximization of the entropy automatically imposes the essential conditions on the image, it becomes impossible to implement the method since in the initial stages of iteration the images invariably give imaginary entropy.

To get over this difficulty, one way is to clip those points in the distribution which violate the essential conditions (Cornwell and Evans, 1985). This scheme does not have any hold on the quality of the reconstruction and one always obtains the same good or bad reconstruction. The other way (Bhandari, 1978; Nityananda and Narayan, 1982) is to add a suitable floating constant c temporarily to the total intensity image such that the essential conditions are met. In this paper we have adopted the later scheme. The choice of c decides the sharpness of the features in the reconstructed image and therefore depending upon the image enhancement requirement the value of c can be chosen. However, we have chosen a value of c such that the ratio of the maximum to the minimum unpolarized intensities in the image is of the order of 100. It should be noted that the reconstruction does not depend very critically upon the choice of c provided it is not too small or too large. As the reconstruction progresses and the negative excursions are reduced the value of the floating c reduces rapidly. At the end of iterations the small left out value of c is subtracted from the zero spacing visibility coefficient.

The algorithm has been satisfactorily (flow diagram Fig. 1) tried on variety of brightness distributions. Some of these are presented here. It is observed that the method generally converges in 20–30 iterations i.e., the square of the entropy gradient becomes smaller than 10^{-6} of its initial value. The main improvement in

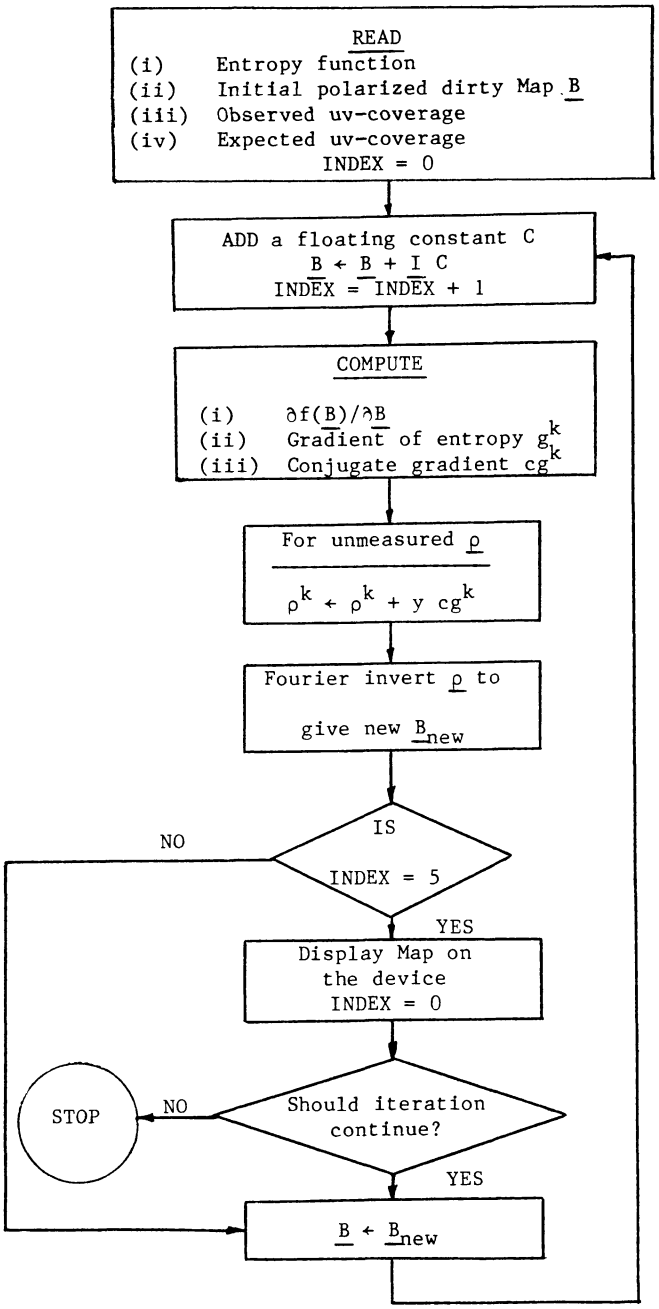


Fig. 1. Flow diagram of the maximum entropy method algorithm for polarized images

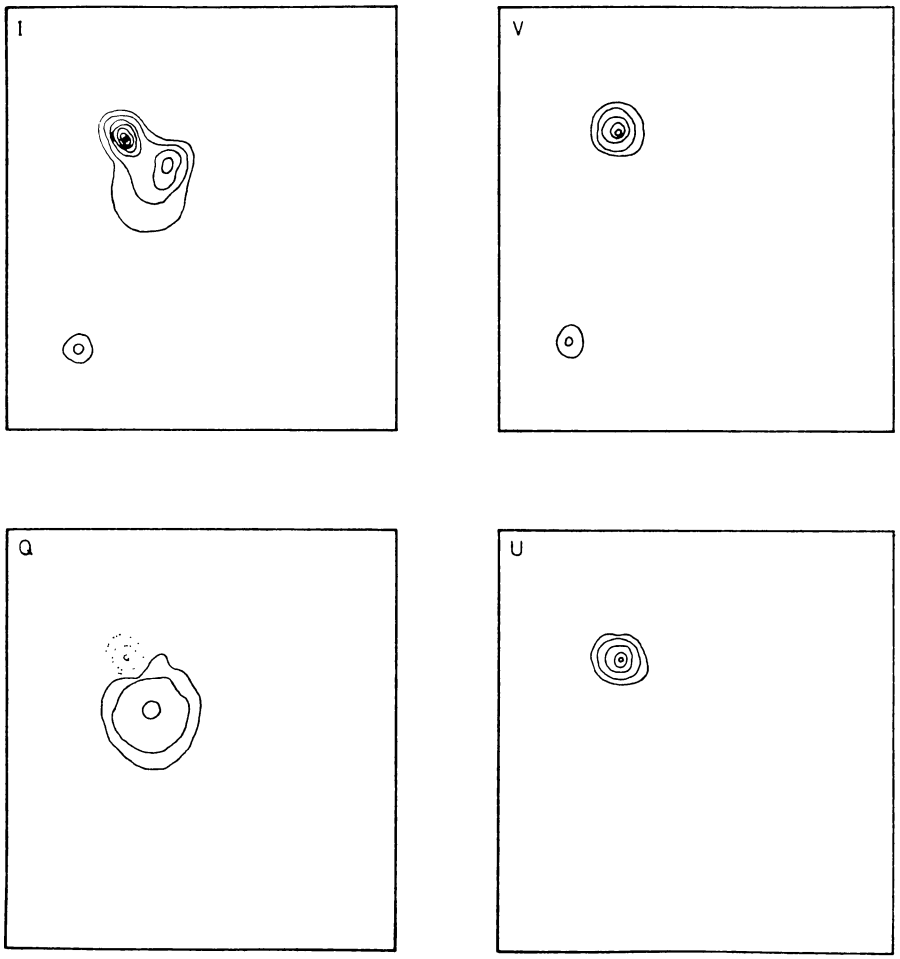


Fig. 2. True partially polarized brightness distribution. The four images indicated by I, Q, U, V correspond to the four Stokes parameters of the brightness distribution. Contours are at 5, 10, 20, 30, 40, 60, 80, 100% of the total intensity peak (100 units)

the image takes place within first few iterations only. The gradient of the entropy decreases very slowly after the first few iterations. Therefore the algorithm should have a provision to view the map periodically, say after every five iterations. If the improvement in five iteration is not significant the user may terminate the iteration forcibly. All the three forms of entropy are tried and found to give more or less similar results. However, it is seen that the solution is relatively stable for the $-B \ln B$ form of entropy than for the others.

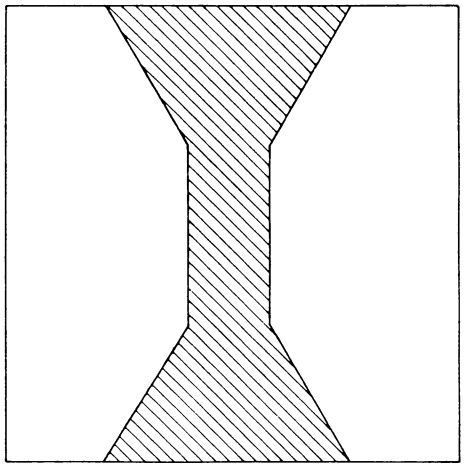


Fig. 3. Synthesized Fourier coverage. The visibility coefficients have been sampled only over hatched region. In the empty region the visibility coefficients are completely unknown

Figure 2 shows true maps for four Stokes parameters I, Q, U, V and the observed Fourier coverage is shown in Fig. 3. This type of Fourier coverage is commonly observed from a synthesis telescope like the Very Large Array (VLA) (Thompson et al., 1980) for a source at negative declinations. For this Fourier coverage the synthesized beam has a very poor resolution in the horizontal direction of the map. Convolution of the true maps (Fig. 2) with the synthesized beam gives the observed maps as shown in Fig. 4. It can be seen clearly that apart from the distortion of the central source the source in the left bottom corner is indistinguishable from the spurious features. Figure 5 shows the MEM reconstructed maps after 30 iterations using the $\ln B$-form of entropy. The reconstructed images are quite an improvement over the observed low resolution images. It is remarkable that although there is no explicit constraint like the positivity on the polarized components the images of Stokes parameters, Q, U, V are also improved appreciably.

Another example is of a bright active region on the Sun observed at 2 cm wavelength using the VLA in the C-configuration (Shevgaonkar and Kundu, 1985). Due to large Faraday rotation in the outer corona of the Sun linear polarization (if at all it is there) is generally averaged to zero over the observing bandwidth

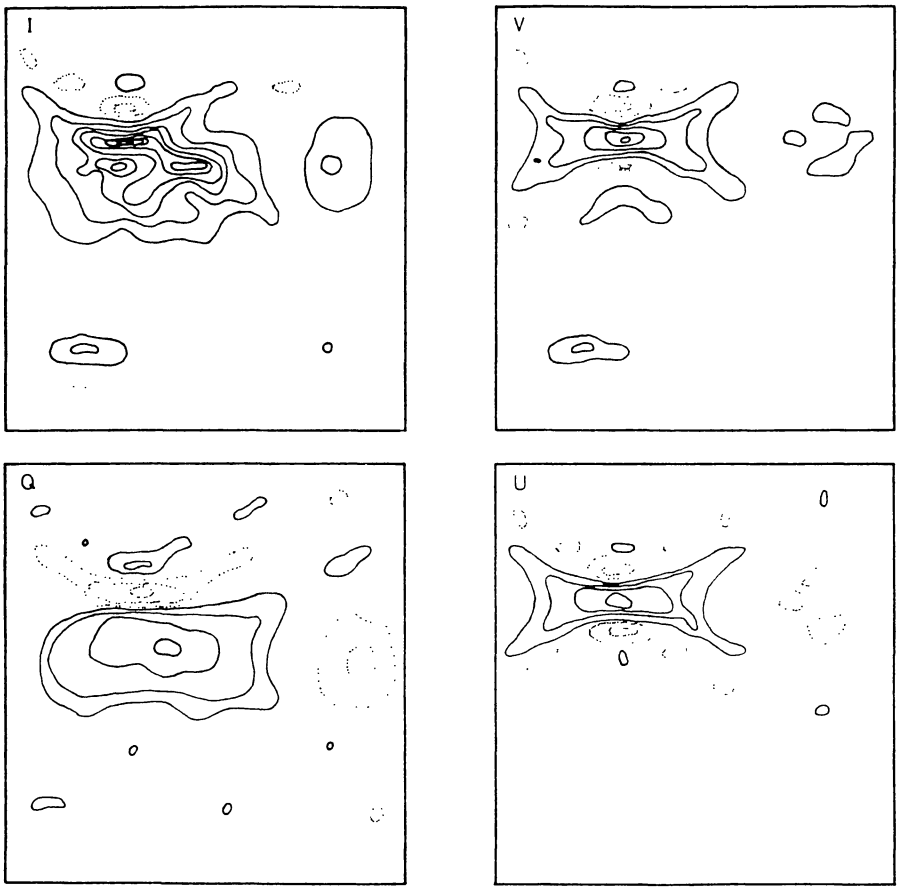


Fig. 4. Observed maps of the four Stokes images for the Fourier coverage in Fig. 3. Contours are at 5, 10, 20, 40, 60, 80, 100% of the total intensity peak (37 units)

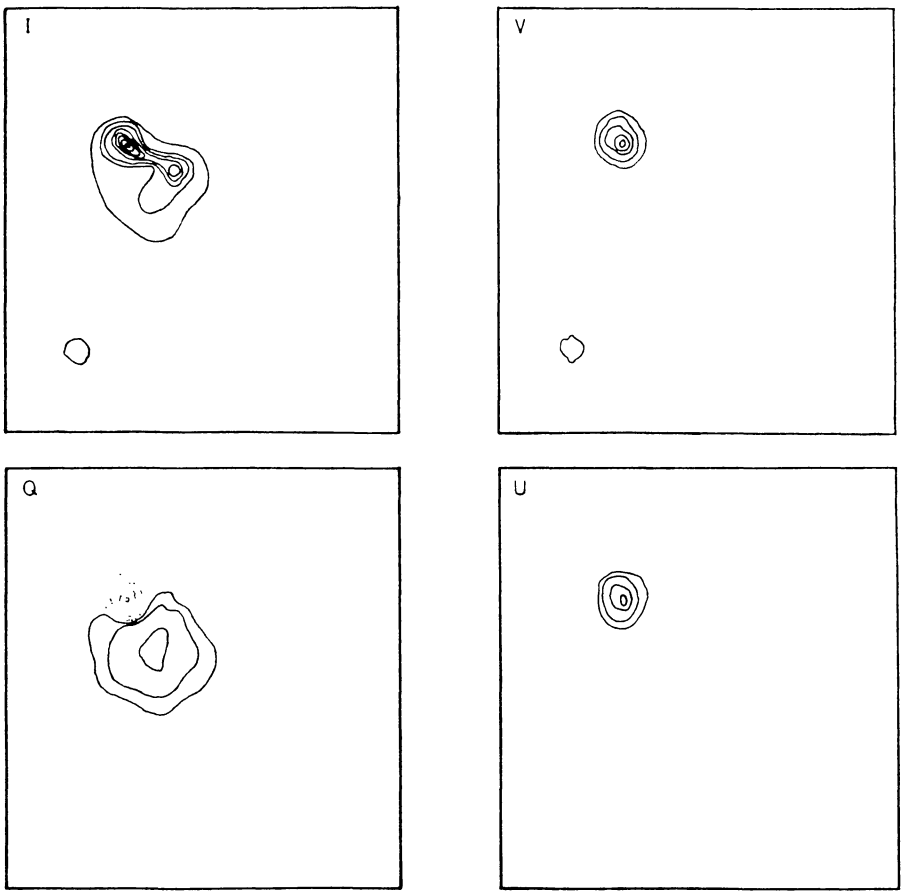
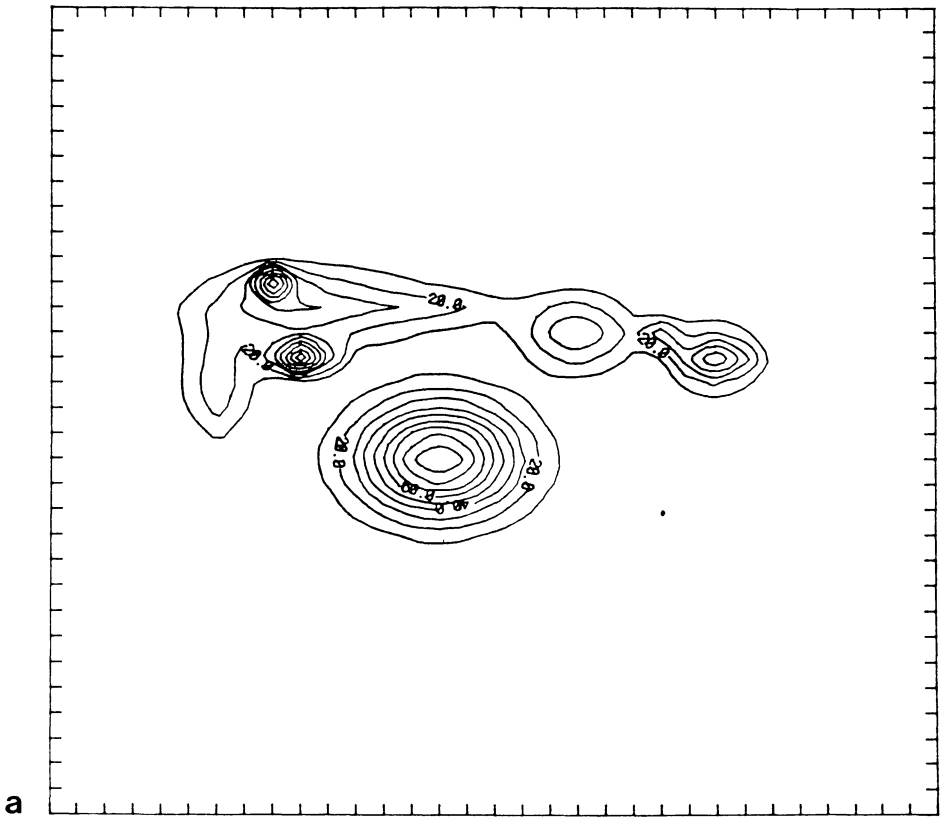


Fig. 5. MEM reconstructed polarized images after 30 iterations. Entropy function is $\ln B$. Contours are at 5, 10, 20, 40, 60, 80, 100% of the total intensity peak (86 units)



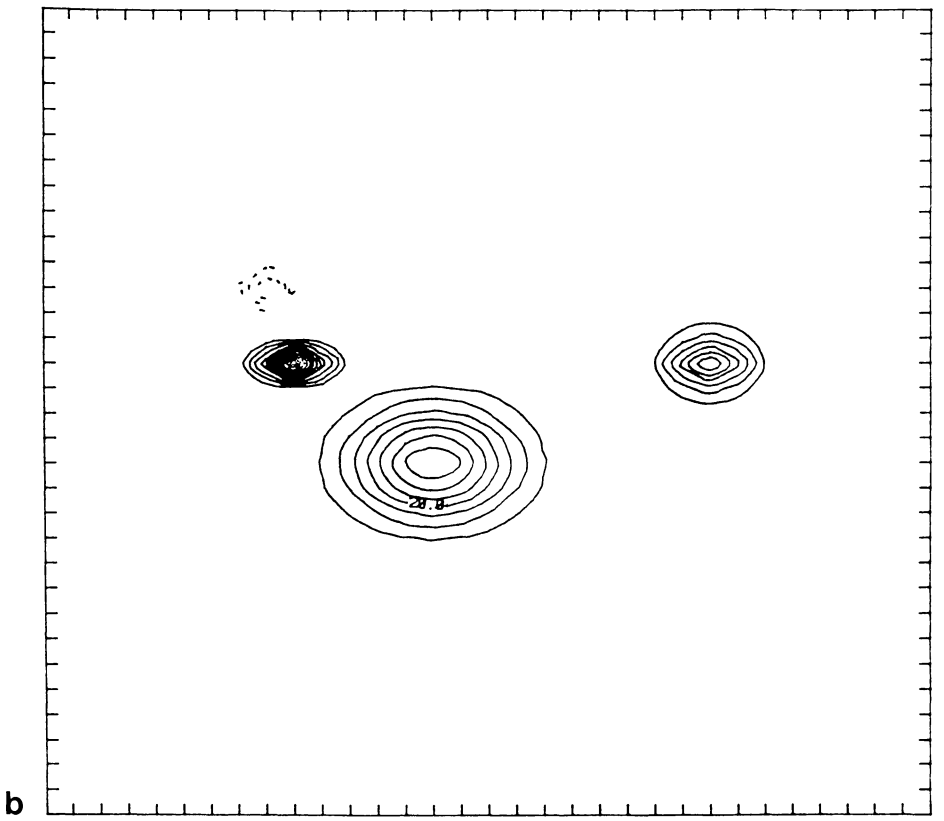


Fig. 6a and b. Total intensity (I) and circular polarization (V) maps of a solar active region at 2 cm wavelength for full day synthesis using the Very Large Array. In the circular polarization map dotted lines indicate left handed whereas solid lines indicate right handed polarization. Contours are drawn at equi-intensity interval

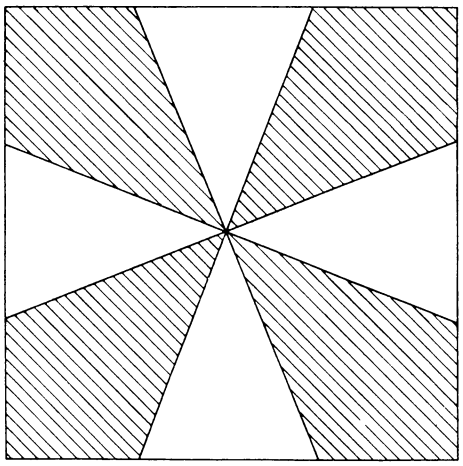


Fig. 7. Simulated Fourier coverage having four sectors. Data has been sampled over the hatched regions only.

(Kundu, 1965), and therefore one essentially observes the circular polarization only. Full-day synthesis data were used to obtain the maps of the solar active region. To remove the effect of the non-uniform *uv*-coverage of the VLA, the maps were first deconvolved using CLEAN. The deconvolved total intensity and circular polarization maps (shown in Fig. 6) are taken as the model maps for generating the visibility data for trying our algorithm. The maps are Fourier transformed to give the visibility function over a rectangular grid of 64 × 64. Assuming that the visibility function could not be observed in the four sectors of the rectangular grid of 64 × 64 (Fig. 7) the visibility coefficients over those four sectors are forced to zero. The Fourier transform of the truncated data provides the distorted images as shown in Fig. 8. Using the distorted images as the first trial model the MEM maps are reconstructed (Fig. 9). From the comparison of Fig. 6 and Fig. 9 it is convincing that the estimation of the unmeasured visibility coefficients using MEM can provide a reconstruction which is very close to the image obtained by actually measuring all the visibility coefficients.

For comparison's sake the CLEANed maps are also shown in Fig. 10. A loop gain of 0.3 and 300 iterations are used to get the CLEAN maps. It is clear that in the MEM maps the features are quite resolved whereas in the CLEAN maps although the sidelobes are suppressed the features are relatively blurred.

5. Treatment of noise

In a practical system the presence of noise is inevitable and therefore the measured quantities have an accuracy of the order of noise only. In the above formulation of MEM we see that during iteration the measured visibility coefficients are rigidly unaltered. However, in the presence of noise the measured visibilities could be modified within the error-bar of the noise. In other words the method should be modified to take into account the reliability of the measured visibility coefficients.

If we assume a zero mean gaussian noise with standard deviation σ^k (k represents the Stokes parameters I, Q, U, V; in general, σ^k can be function of baseline), the constraints given by Eq. (21) should be modified to

$$\frac{1}{N_0^k}\sum\left\{\frac{\rho^{k0}-\rho^k}{\sigma^k}\right\}^2 \leq 1; \quad k = I, Q, U, V \tag{26}$$

where N_0^k are total number of measured visibility coefficients for kth Stokes parameter. ρ^{k0} and ρ^k are actually measured and estimated visibility coefficients respectively.

Following previous authors (Ables, 1974; Wernecke and D'Addario, 1977; Willingale, 1981), for noisy data the constrained maximization of the entropy E can be changed to unconstrained maximization of an objective function given as

$$F \equiv E - \lambda \sum_k \frac{1}{N_0^k}\sum\left\{\frac{\rho^{k0}-\rho^k}{\sigma^k}\right\}^2 \tag{27}$$

where λ is a Lagrange multiplier which decides the weightage given to the maximization of the entropy relative to the observational constraints. The gradient of the objective function F with respect to the visibility coefficients is then given by

$$G^k \equiv \frac{\partial F}{\partial \rho^k} = g^k + \frac{2\lambda}{N_0^k}\left\{\frac{\rho^{k0}-\rho^k}{(\sigma^k)^2}\right\} \tag{28}$$

The algorithm given in Fig. 1 can be directly used for the noisy data with the gradient g^k replaced by G^k. A good discussion on how to choose the Lagrange multiplier has been presented in the past by Wernecke and D'Addario (1977). However, in practice the value of λ is chosen by trial and error to get the best results. If the data is good, it is appropriate to choose a value of λ which gives equal weightage to the two terms of Eq. (28).

The extension of MEM to noisy data, although looking straightforward, it has been pointed out in the past (Bryan and Skilling, 1980; Nityananda and Narayan, 1982) that the least – square data-fitting technique given above introduces a systematic bias in the reconstruction. The bias is independent of the actual noise in the observation. The later authors argue that to avoid the systematic bias in the reconstruction, the measurements, although corrupted by noise, must be fitted exactly. However, one may note that exact fitting of the noisy data may not give a reconstruction which satisfies essential conditions like the positivity and the degree of polarization less than unity. In fact for noisy data the least-squares technique along with the use of relative entropy (Cornwell and Evans, 1985; Shevgaonkar, 1986a,b) of the kind $\mathbf{B}\ln(\mathbf{B}/\mathbf{B}_0)$ with reasonably simple but correct default image $\mathbf{B}_0$ may provide much superior results.

6. Conclusion

A simple MEM algorithm for polarized images has been described here. The satisfactory application of the algorithm to polarized synthesis images has also been presented. Due to practical limitations when the visibility coefficients for the four Stokes parameters cannot be sampled over identical apertures, the independent CLEANing of the four Stokes images does not guarantee the essential conditions on the images (i.e., the intensity is positive definite and the degree of polarization is less than unity). The maximum entropy method treats all the Stokes images simultaneously and imposes the essential conditions automatically. For complex and extended structures MEM seems to be more promising compared to CLEAN. This holds also for wide field mapping, as pointed out by Cornwell and Evans (1985), since MEM is more computer efficient compared to CLEAN; the method certainly has a bright future.

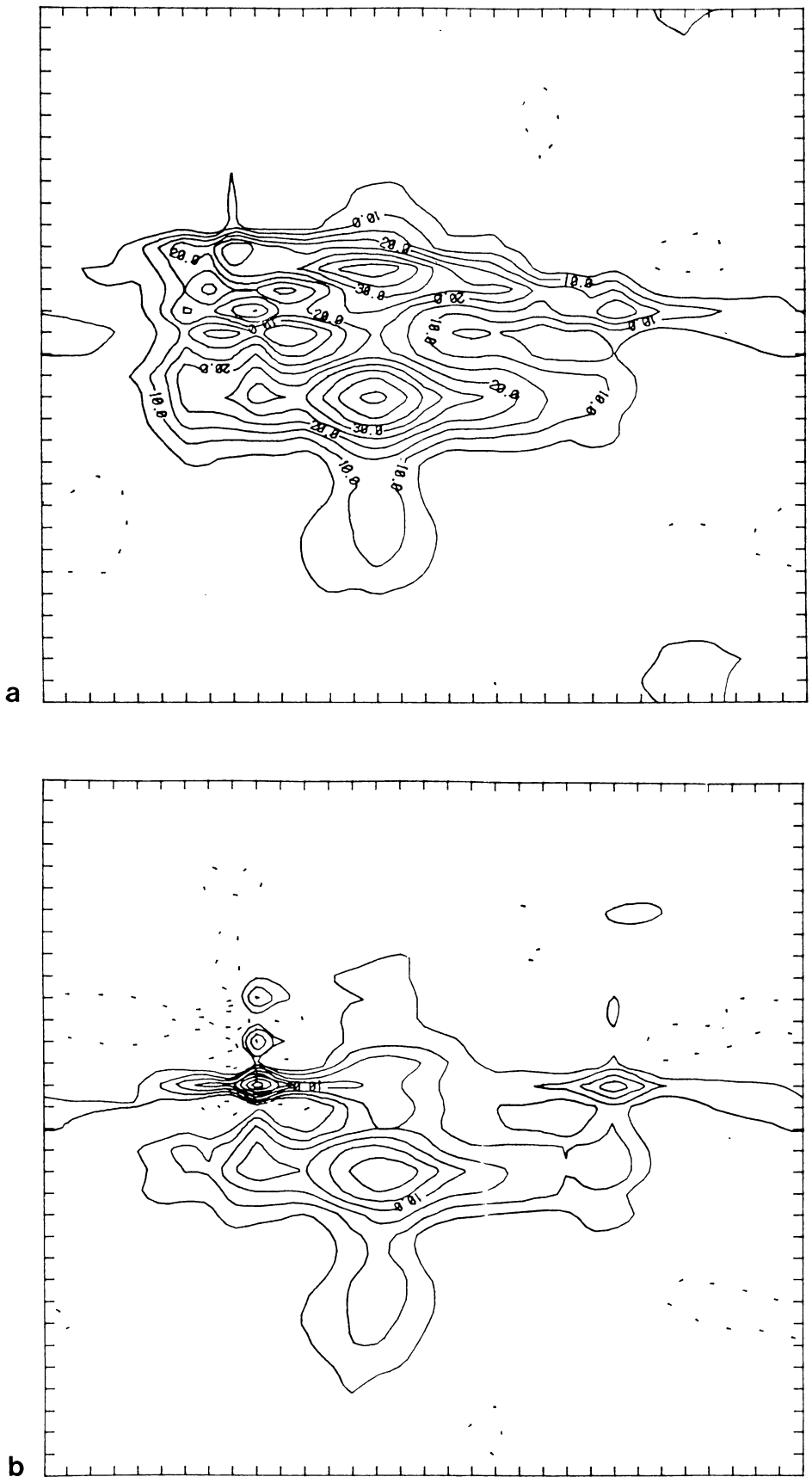


Fig. 8a and b. Distorted total intensity and circular polarization images due to truncated Fourier coverage. Contours are drawn at equi-intensity interval

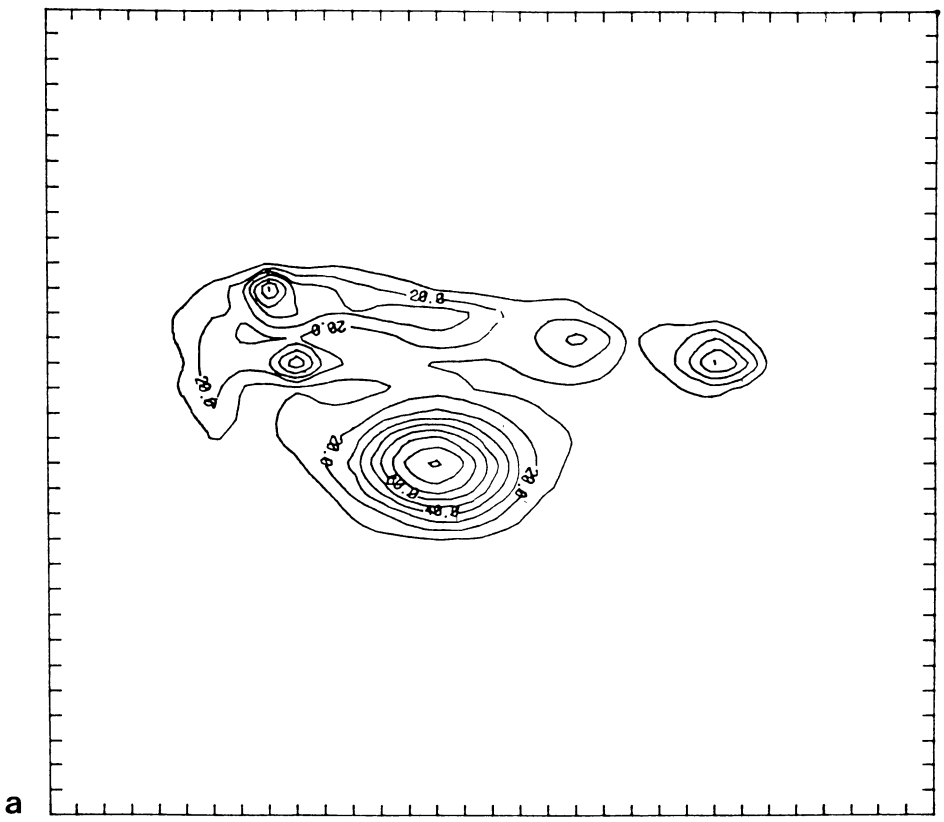
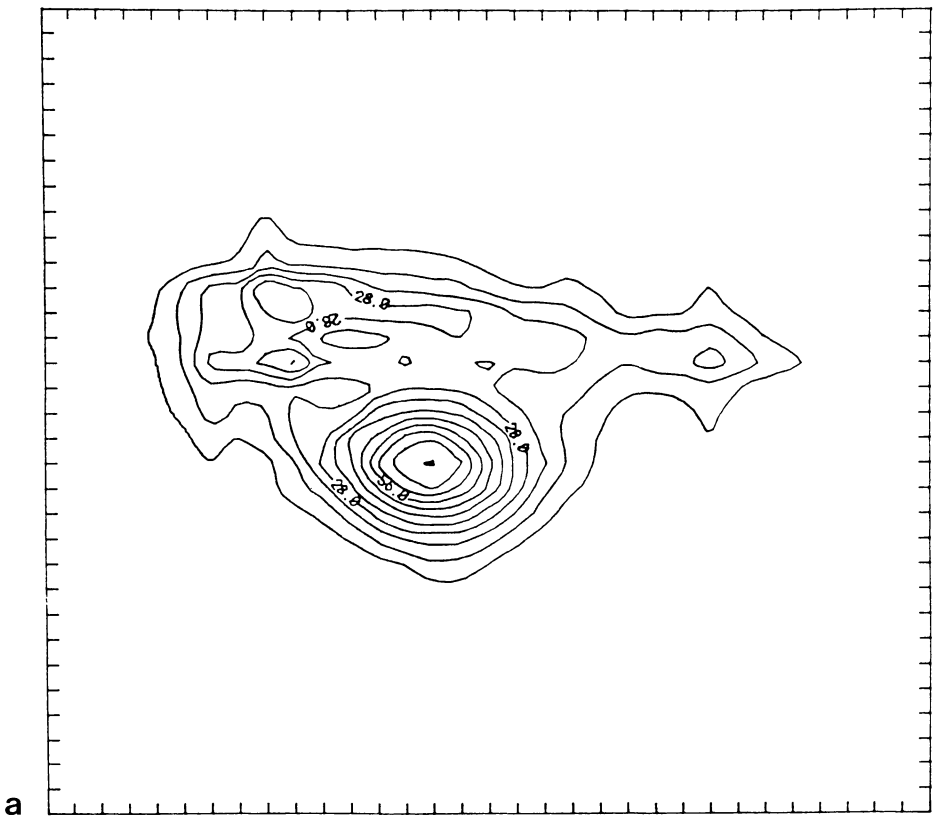
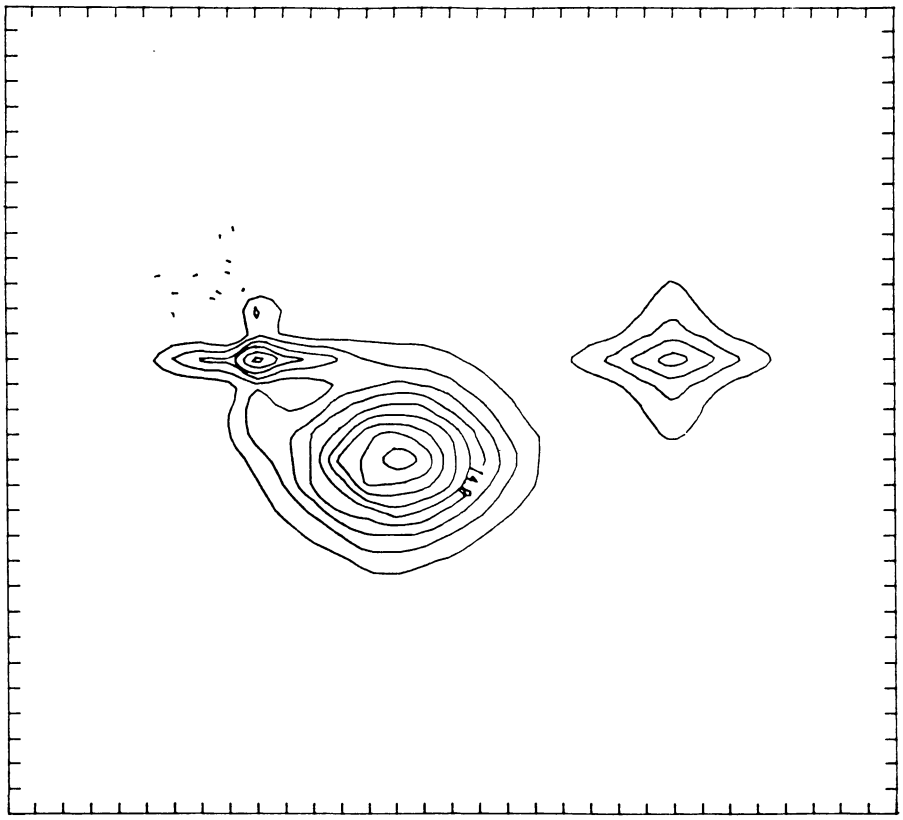


Fig. 9a and b. MEM reconstructed total intensity and circular polarization maps after 30 iterations. Entropy function is ln *B*. Contours are drawn at equi-intensity interval



a



b

Fig. 10a and b. Independently CLEANed total intensity (I) and circular polarization (V) images after 300 iterations each. Loop gain is 0.3. Contours are drawn at equi-intensity level

Acknowledgement. I thank Drs. R. Narayan, T.J. Cornwell and R. Nityananda for useful discussions.

References

Ables, J.G.: 1974, *Astron. Astrophys. Suppl.* **15**, 383
Bhandari, R.: 1978, *Astron. Astrophys.* **70**, 331
Bryan, R.K., Skilling, J.: 1980, *Monthly Notices Roy. Astron. Soc.* **191**, 69
Burg, J.P.: 1967, Proc. 37th Meeting of the International Soc. Exploration Geophysicist, reprinted in "Modern Spectrum Analysis", Ed. D.G. Childers, IEEE Press, New York (1978) p. 34
Cornwell, T.J., Evans, F.K.: 1985, *Astron. Astrophys.* **143**, 77
Fletcher, R., Reeves, C.M.: 1964, *Comput. J.* **7**, 149
Gull, S.F., Daniell, G.J.: 1978, *Nature* **272**, 686
Gull, S.F., Skilling, J.: 1984, in *Indirect Imaging*, IAU Proc. ed. J.A. Roberts, p. 267
Högbom, J.A.: 1974, *Astron. Astrophys. Suppl.* **15**, 417
Högbom, J.A., Carlsson, I.: 1974, *Astron. Astrophys.* **34**, 341
Kundu, M.R.: 1965, *Solar Radio Astronomy*, Interscience, New York
Narayan, R., Nityananda, R.: 1984, in *Indirect Imaging*, IAU Proc. ed. J.A. Roberts, p. 281
Nityananda, R., Narayan, R.: 1982, *J. Astrophys. Astron.* **3**, 419
Nityananda, R., Narayan, R.: 1983, *Astron. Astrophys.* **118**, 194
Ponsonby, J.E.B.: 1973, *Monthly Notices Roy. Astron. Soc.* **163**, 369
Shevgaonkar, R.K., Kundu, M.R.: 1985, *Astrophys. J.* **292**, 733
Shevgaonkar, R.K.: 1986a, *Astron. Astrophys.* **162**, 349
Shevgaonkar, R.K.: 1986b, *J. Astrophys. Astron.* **7**, 275
Skilling, J., Bryan, R.K.: 1984, *Monthly Notices Roy. Astron. Soc.* **211**, 111
Thompson, A.R., Clark, B.G., Wade, C.M., Napier, P.J.: 1980, *Astrophys. J. Suppl.* **44**, 151
Wernecke, S.J., D'Addario, L.R.: 1977, *IEEE Trans. Comput.* **C-26**, 351
Willingale, R.: 1981, *Monthly Notices Roy. Astron. Soc.* **163**, 369