VLA Scientific Memorandum No. 141

# Can CLEAN be improved ?

by

T.J.Cornwell NRAO/VLA

March 1982

## 1. Introduction:

The one major drawback to the use of the CLEAN deconvolution algorithm is its poor behaviour on regions of extended emission; particularly as manifested in the appearence of stripe-like artifacts in the CLEAN image. Advantages over more sophisticated deconvolution algorithms such as MEM are its speed, particularly in the Clark algorithm ( although a similar two stage approach to MEM is possible ) and simplicity ( especially of programming ). However, algorithms such as MEM are designed to treat correctly regions of extended emission. It is clear that an ideal deconvolution algorithm would merge the best attributes of both CLEAN and MEM. In this memo I will present details of an initial attempt at designing such an algorithm.

## 2. The CLEAN algorithm:

CLEAN utilises an iterative point source subtraction technique to minimise a chi-squared term which, in the u,v plane, can be written :

$$\chi^2 = \Sigma \quad w_k*[V_k-T_k]^2$$

-equation (2.1)

where $V_k$ = observed complex visibility at the kth sample point

$T_k$ = predicted complex visibility at the kth sample point

$w_k$ = weight attached to the kth sample point

and [ ] represents the absolute value. Here, and below, repeated indices are to be summed.

In the map plane $\chi^2$ can be written as :

$$\chi^2 = \Sigma\, p_{i,j}*f_i*f_j$$
$$- 2*\Sigma\, d_i*f_i + \Sigma\, w_k*[V_k]^2$$

-equation (2.2)

where $p_{i,j}$ = beam matrix

$d_i$ = dirty map vector

$f_i$ = predicted map vector

and the summations cover the entire map.

Choosing f to maximise $-\chi^2$ we find the usual convolution equation :

$$\Sigma\ p_{i,j}*f_j = d_j$$

-equation (2.3)

The CLEAN algorithm chooses one of the possible solutions of this equation. Preference is given to those images containing a number of point sources in a mainly empty field. The uniqueness of the predicted map and the asymptotic value of $\chi^2$ both depend on the number of independent sample points and number of beam areas of non-zero brightness allowed in the CLEAN map ( see Schwarz 1978).

## 3. A modification of CLEAN :

One would like to alter CLEAN such that regions of extended emission are treated properly and in particular so that stripes, not constrained by the data, are removed. One approach, which we will adopt here, is to change the dirty map or dirty beam in some way and then just use CLEAN as usual.

In general we may do this by maximising a combination of $-\chi^2$ and some other function which measures "good" maps. Let H(f) be this function; we then maximise

$$O = \alpha*H(f) - \chi^2$$

-equation (3.1)

where the variable $\alpha$ controls the balance between fitting the data and obtaining a "good" map.

The predicted map is found by solving :

$$\Sigma\ p_{i,j}*f_i = d_j + \alpha*dH/df_j$$

-equation (3.2)

In all interesting cases H will depend non-linearly on f. Perhaps the easiest method of solution is to use CLEAN to solve equation (3.5) and then calculate the correction, $\alpha*dH/df_j$, to the dirty map, iterating until convergence is achieved.

The optimum value of $\alpha$ can be estimated by multiplying equation (3.5) by $f_j$ and summing. We then find that :

$$0.5*(\chi^2+(\Sigma\ f_j*p_{i,j}*f_i - \Sigma\ w_k*[V_k]^2)) = \alpha*\ \Sigma\ f_i*dH/df_i$$

-equation (3.3)

The difference on the left hand side is related to the discrepancy in signal to noise of the observed and reconstructed visibilities. For a reasonably unbiased image we will assume that this vanishes. If the expected value of $\chi^2$ is $\sigma^2$ per pixel then :

$$\alpha = \sigma^2/(2*\langle f_i*dH/df_i\rangle)$$

-equation (3.4)

where < > denotes the average value.

The only missing ingredient is the goodness measure H; this we now consider.

What is desired of the goodness measure of a map ? Two attributes seem important :

1. Only positive brightness should be allowed although, of course, negative residuals will be permissible. This constraint should be dropped for Q and U maps.

2. Images having low dispersion in pixel values should be preferred; spurious stripes in the image should then be removed.

Infinitely many functions satisfy these criteria; the most interesting the various entropy measures. I use the term entropy merely to denote the lack of spread in pixel values, not any physical concept. Some of the entropy measures are :

$$H1 = -\sum f_i*\ln(f_i)$$

-equation (3.5)

$$H2 = \sum \ln(f_i)$$

-equation (3.6)

$$H3 = -\sum 1./f_i$$

-equation (3.7)

$$H4 = -\sum 1./f_i^2$$

-equation (3.8)

$$H5 = \sum \sqrt{(f_i)}$$

-equation (3.9)

and their cousins Hi', formed by normalising $f_i$ with respect to the total flux in the image. ( Wernecke and D'Addario used H2 whereas Gull and Daniell used H1'.)

All of these measures are maximised for images with low dispersions in pixel values and all require positivity. If we drop the positivity constraint then the smoothness measure S is available :

$$S = -\sum f_i^2$$

-equation (3.10)

We will now go on to consider the use of these goodness measures in a practical CLEAN-based algorithm.

## 4. The Maximum smoothness Method

Unfortunately, if smoothness is used in place of H then all the permissible solutions to equation (3.5) are very close to the principal solution of the ordinary convolution equation. However, if we use the CLEAN algorithm to find the solution then we introduce the extra constraint that the map be composed of a number of point sources. We should obtain a smoother map than the usual CLEAN solution and one which does not have the sidelobes found in the principal solution. One very convenient aspect is that to find the solution to this "maximum smoothness method" (MSM) we only need to modify the beam to be :

$$p_{i,j} + \sigma^2 * \delta_{i,j}/(2*<f^2{}_i>)$$

-equation (4.1)

If the sidelobes are reasonably small then the mean square signal can be estimated from the dirty map.

Since negative pixel values are allowed the zero spacing flux is not biased as it is if an entropy measure is used. The resolution is invariant over the field of view.

## 5. The Maximum Entropy Method:

Of the entropy measures H2 is the most convenient since $\alpha$ is independent of f ( see equation (5.2) ). We must then solve ( using CLEAN ) :

$$\Sigma\, p_{i,j} * f_i = d_j + \sigma^2/(2*f_j)$$

-equation (5.1)

In practice we go through the following sequence :

1. CLEAN dirty map to obtain initial CLEAN map. We then use this map to approximate the MEM map.
2. Correct dirty map using the MEM map, truncating below some arbitrary level e.g. $\sigma$ to avoid the forbidden negative values.
3. CLEAN the corrected dirty map
4. Goto 2. unless convergence is attained

The CLEAN beam may be chosen at will but in practice superresolution seems not to work well. and should be avoided just as it is in conventional CLEAN.

### 5.1. Pros and Cons:

Several desirable aspects of this general approach to MEM are apparent :

A1: Regions of good signal to noise ratio are less affected than the weaker regions. Ignoring the effect of sidelobes we find that the map is given by :

$$f_i = 0.5*(d_i+\sqrt{(d_i^2+2*\sigma^2)}$$

-equation (5.2)

which tends to $d_i$ for good signal to noise and to a fixed level $\sigma/\sqrt{(2)}$ as the signal vanishs.

A2: A stability analysis indicates that small sinusoids not required by the data are removed.

A3: Positivity is strongly encouraged.

A4: In the case of H2 only one pass through the image is required to find the correction to the dirty map.

A number of disadvantages are also involved :

D1: The predicted image is slightly biased. This bias is about $0.7\sigma$ for weak points and vanishs for strong points. For example, the zero spacing flux for an MEM map of a blank field is non-zero.

D2: The resolution varies with signal to noise, consequently simple interpretation of the image may be difficult.

D3: Several passes through the entire cycle are required, however each pass is only marginally more expensive than CLEAN.

D4: The r.m.s noise is a free parameter and can be chosen at will. Large values produce a ridiculously smooth map whereas small values have virtually no discernable effect. Such free parameters will appear in any non-linear deconvolution method such as MEM or regularisation. In fact the CLEAN windows play a similar role in CLEAN.

D5: The clipping below some arbitrary level is unsatisfactory in that it strongly affects the positivity of the final map. I can see no easy way in which this can be avoided in the present scheme.

Several of these disadvantages might affect the application of such pseudo-MEM maps to the estimation of spectral indices, percentage polarisation, optical depth etc. We will now examine these in further detail.

First we consider the bias. From equation (5.2) we see that for a signal of $x\sigma$ the bias is, ignoring sidelobes, :

$$0.5*(\sqrt{(x^2+2)}-x)*\sigma$$

-equation (5.3)

For a $5\sigma$ detection the bias is then about $0.1\sigma$ and for a $3\ \sigma$ detection the bias is about $0.4\sigma$; in most practical cases this effect will be negligible. Also by virtue of the positivity constraint MEM should provide a better estimate of the zero spacing flux than CLEAN and hence one may gain .

Secondly, we consider the variable resolution. The use of the CLEAN beam avoids the problems introduced by superresolution. The converse of superresolution, subresolution, which occurs on weak features may be more serious. Using equation (5.2) we find that, ignoring sidelobes, the increase in width of a $5\sigma$ (3) Gaussian is about 4.3 per cent (11.7 per cent). Hence one must be careful in quoting apparent sizes of weak sources. In most cases I would expect that this effect to be negligible compared to the uncertainties due to non-Gaussian profiles.

Thirdly, for maps of strong sources $\sigma$ will contain a contribution due to the limited dynamic range; this is probably best estimated either from a blank region of the I-map. If a stripe is present then convergence can be hastened by initially setting $\sigma$ to the amplitude of the stripe and subsequently decreasing it to the correct value.

The relative importance of the advantages and disadvantages will vary from case to case as happens with CLEAN and selfcalibration.

## 6. How do these methods work ?:

Suppose that after CLEANing a map we find that a small sinusoid corresponding to an unmeasured part of the uv plane is present in the map. The concave nature of the entropy measures and the smoothness measure ensures that the dirty map is altered by the addition of a small sinusoid phase shifted by 180 degrees.

We can now see that equation (3.2) simply uses ordinary feedback methods to stabilise the CLEAN algorithm and, as such, could be derived with no mention of entropy or smoothness.

## 7. An example:

Fig. 1 shows a dirty map of SAG A at 20cm courtesy of R.D.Ekers and J.van Gorkom. Figure 2 shows the CLEAN map (loop gain = 0.1,10000 iterations ). Stripes are present in the map running along pa 30 degrees with an amplitude of about 10 to 20 mJy per beam. A slice taken on a vertical line is shown in Fig. 3. I applied the pseudo-MEM algorithm to this data using values for $\sigma$ of 10,20,50 mJy per beam. Slices from the resulting maps are shown in Fig. 4. For $\sigma$=50 the sinewave has, as expected, been reversed in phase and amplified whereas for $\sigma$=10 and 20 it has decreased somewhat. After two more iterations with $\sigma$=10 the slice is as shown in Fig. 5. The zero level has changed by about 5-10 mJy per beam and the stripes have diminished considerably. It can be seen that , with the exception of the stripes, the final structure, shown in Fig 6, has changed very little. In Fig. 7 I show the usual slice through the MSM map made with $\sigma$=10. The smoothness seems comparable to that of the MEM map. The full MSM map is shown in Fig. 8.

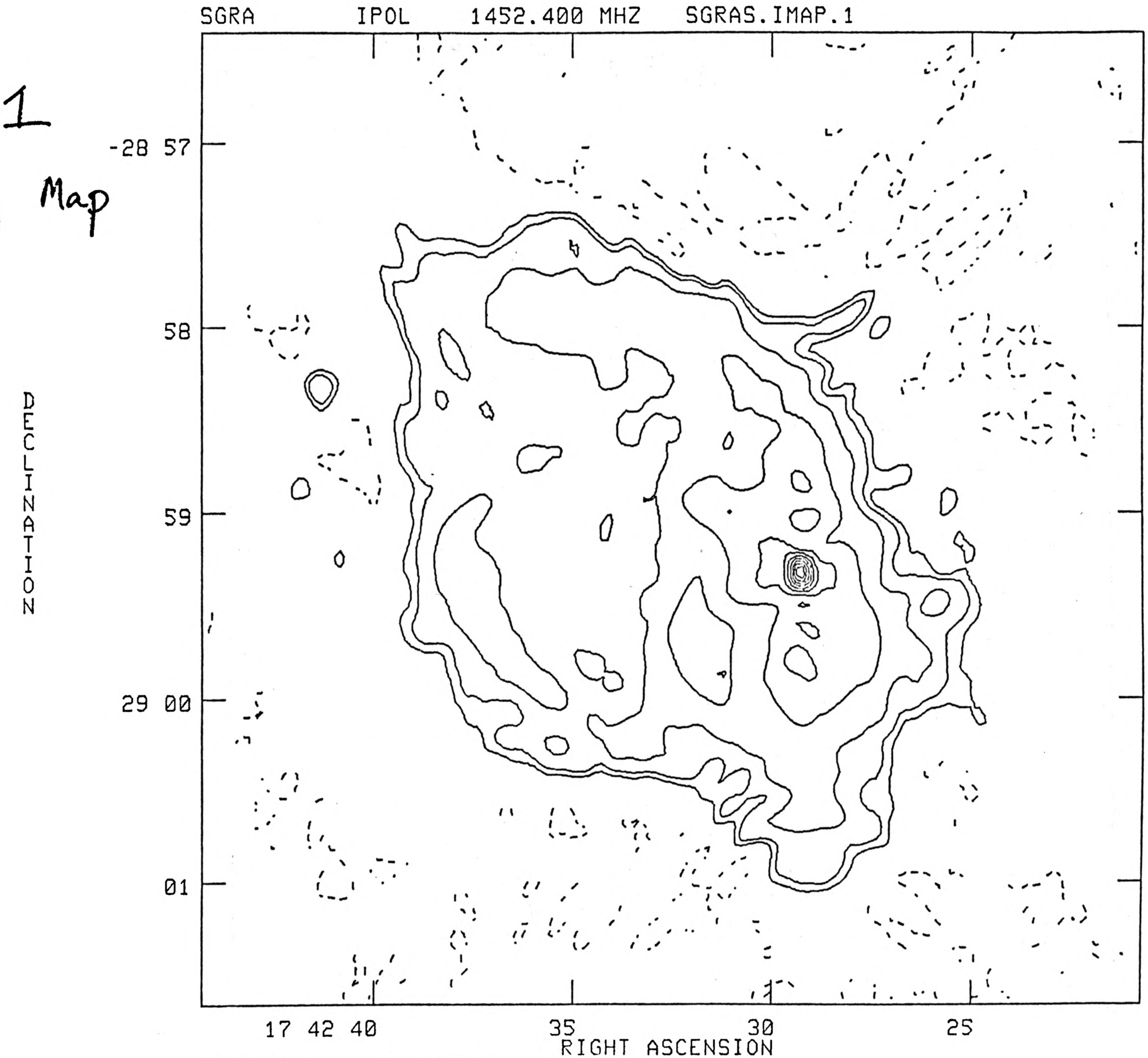
## 8. Does this really help ?:

The presence of stripes in a CLEAN map indicates that something has gone awry with the algorithm we all love and trust. Does this mean that we should rely on a completely unknown process to cobble together a reasonable looking map ? Well maybe, and maybe not. We could only CLEAN data which has no big holes in the uv coverage but this sort of conservatism is that which would prevent any use of CLEAN or selfcalibration.

It is possible that some other solution to the stripe problem exists relying on, say, a variable loop gain or adaptive boxes; however for those who would prefer this type of approach I would point out that these pseudo-MEM and pseudo-MSM algorithms should be regarded simply as means of stabilising the CLEAN algorithm. I have made no mention of the canonical ensemble of monkeys usually invoked in discussions of MEM; in fact I regard the various entropy measures and the smoothness measure as more or less arbitrary functions which are chosen primarily to stabilise CLEAN.

On purely practical grounds MSM appears to preferable over MEM, mainly because at most two passes through CLEAN are necessary and because no bias is introduced. It also treats Q and U maps correctly.

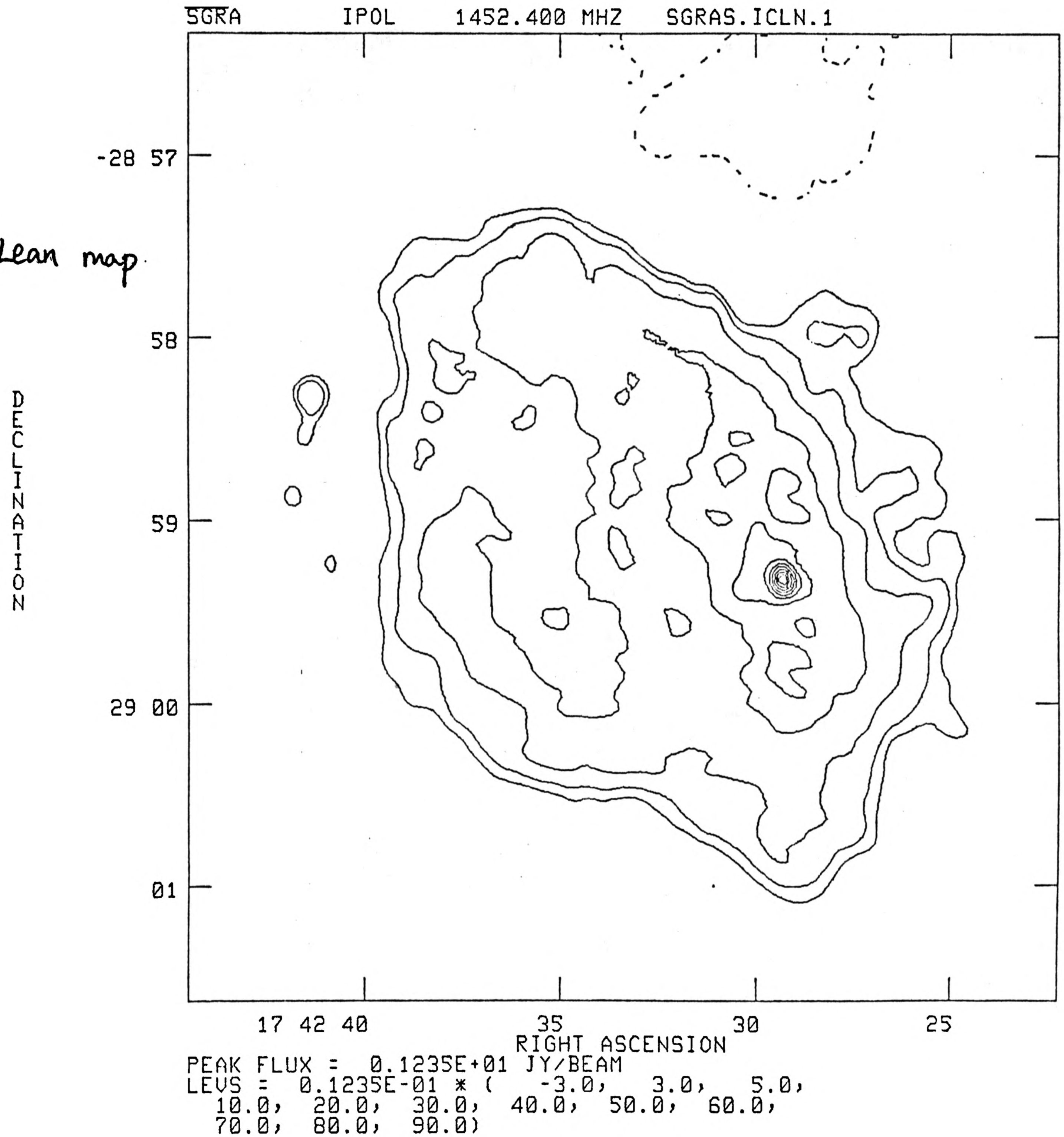
Fig 1
Dirty Map



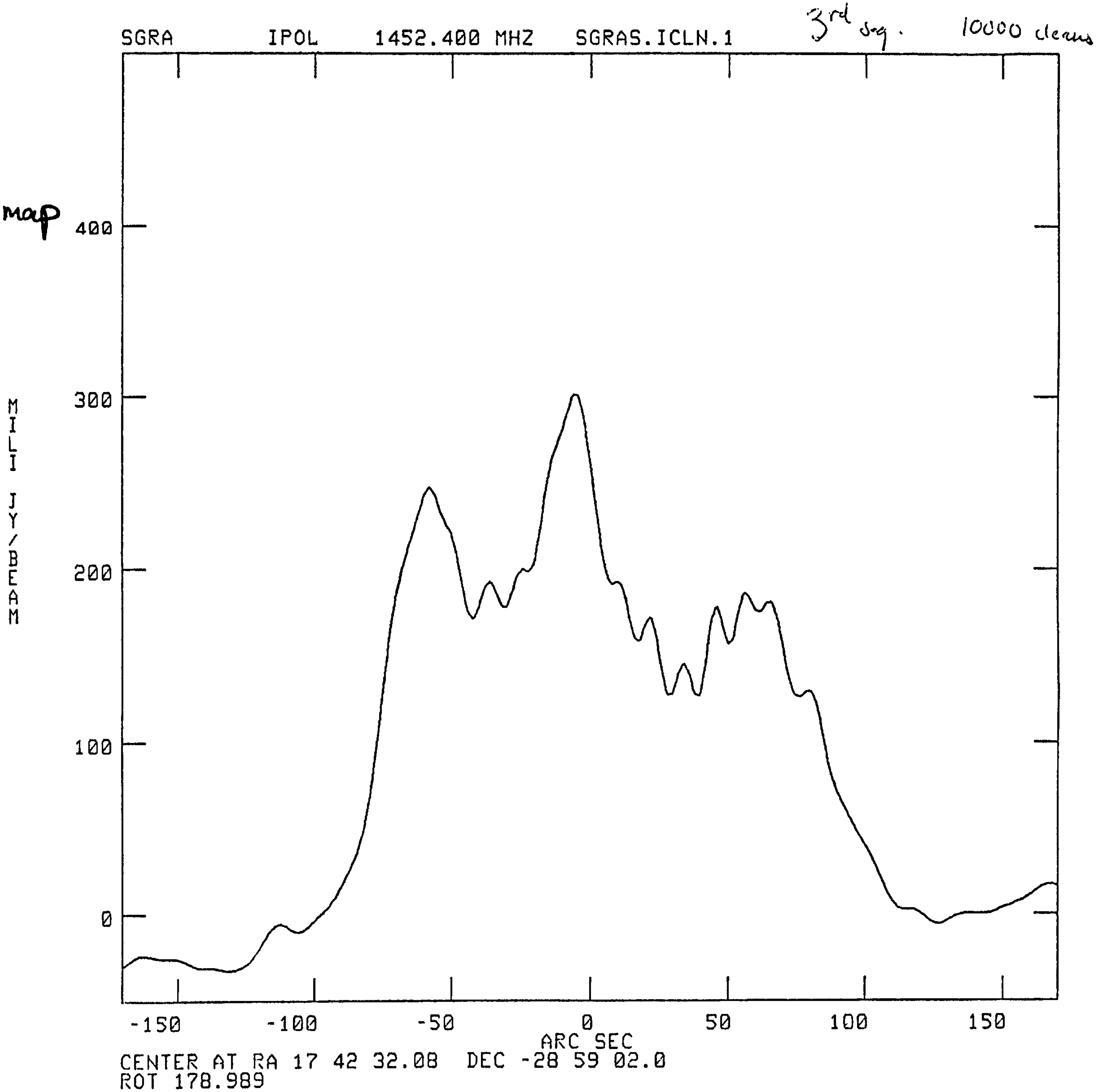

PEAK FLUX = 0.1158E+01 JY/BEAM
LEVS = 0.1158E-01 * ( -3.0, 3.0, 5.0,
10.0, 20.0, 30.0, 40.0, 50.0, 60.0,
[redacted]0, [redacted].0,[redacted]0.0)[redacted]

Fig 2

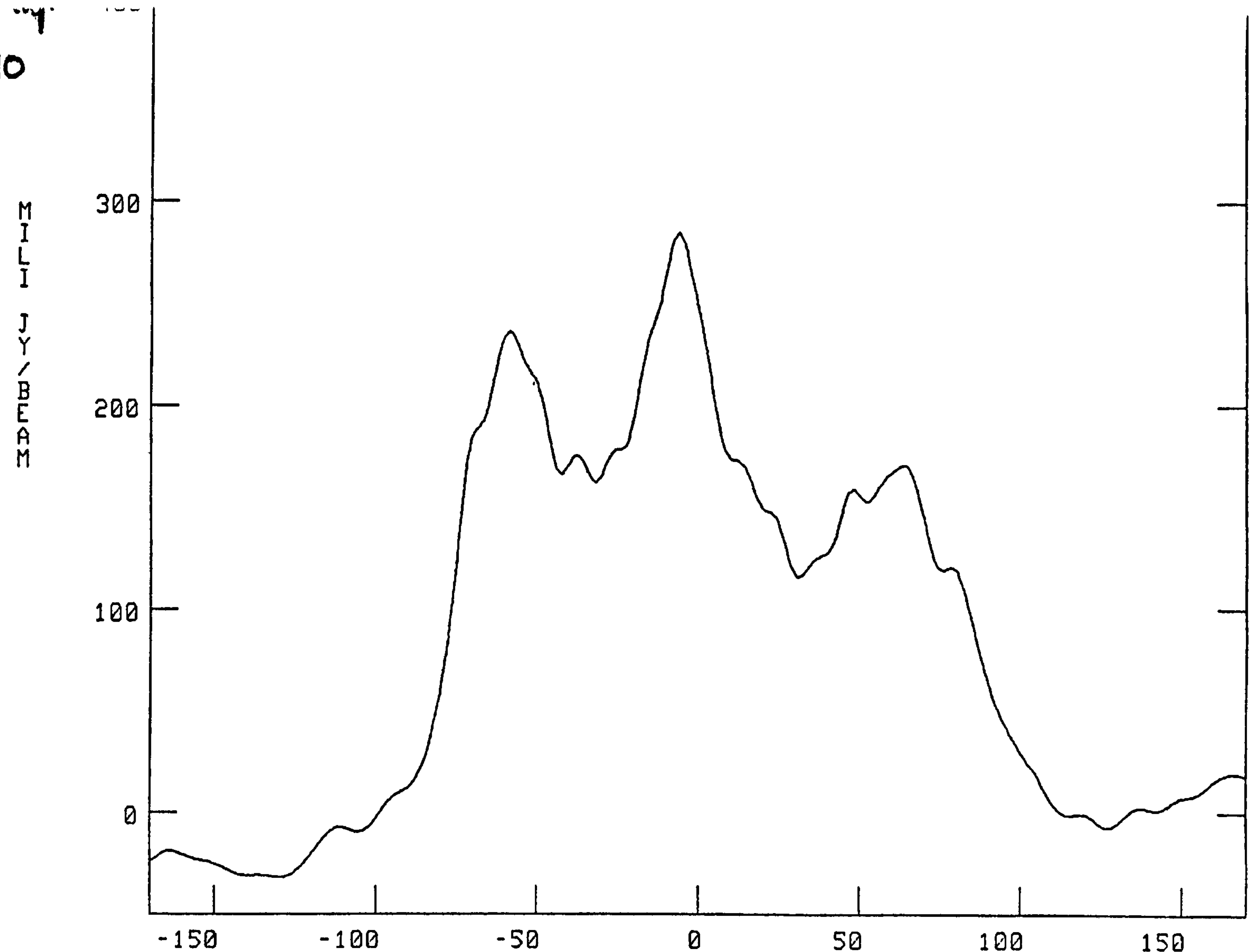
Initial CLean map



Fg 3.
Clean map



σ = 10



CENTER AT RA 17 42 32.08 DEC -28 59 02.0
ROT 178.989

Fig 4 (cont)
MEM map
$\sigma = 20$

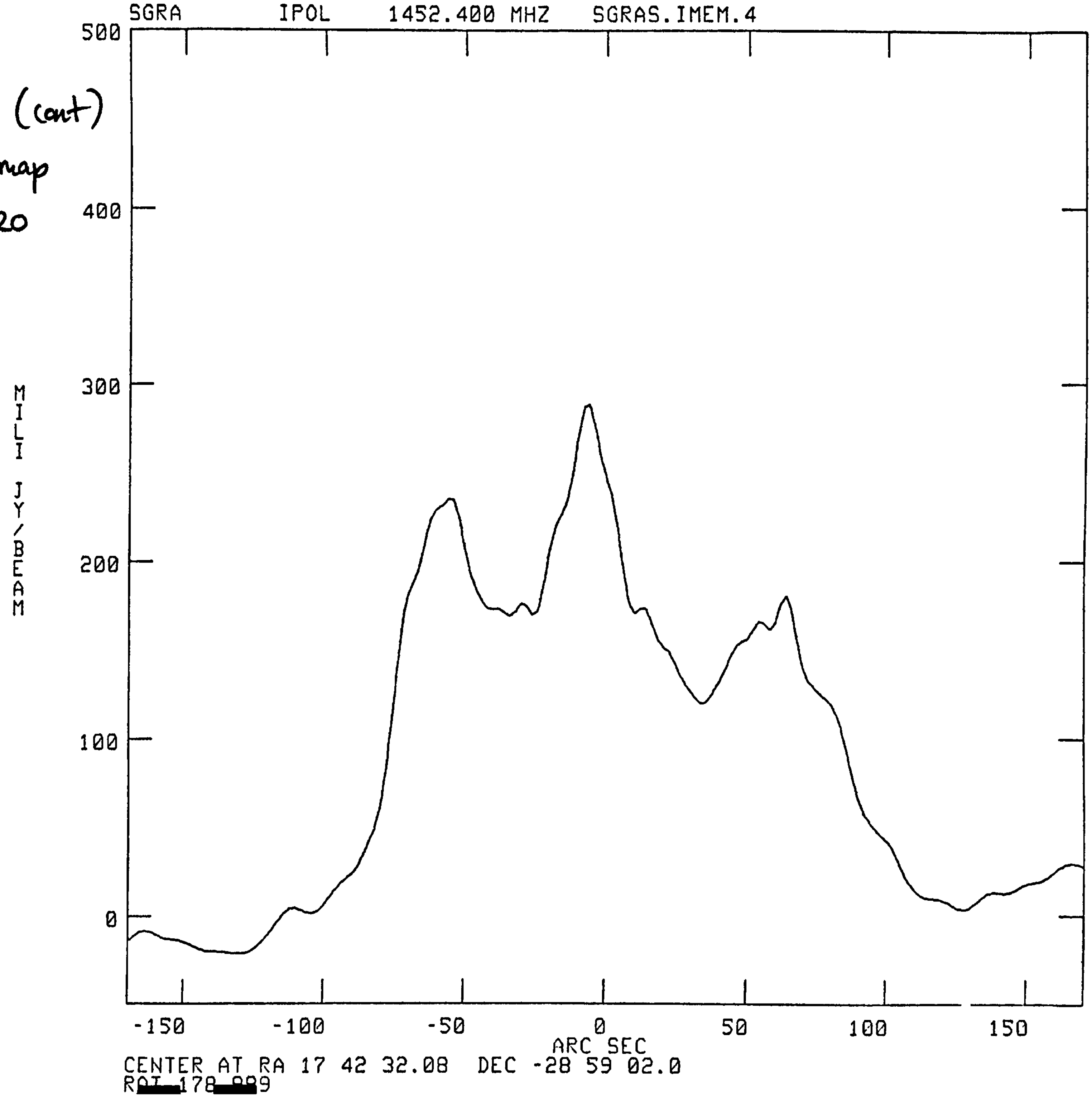
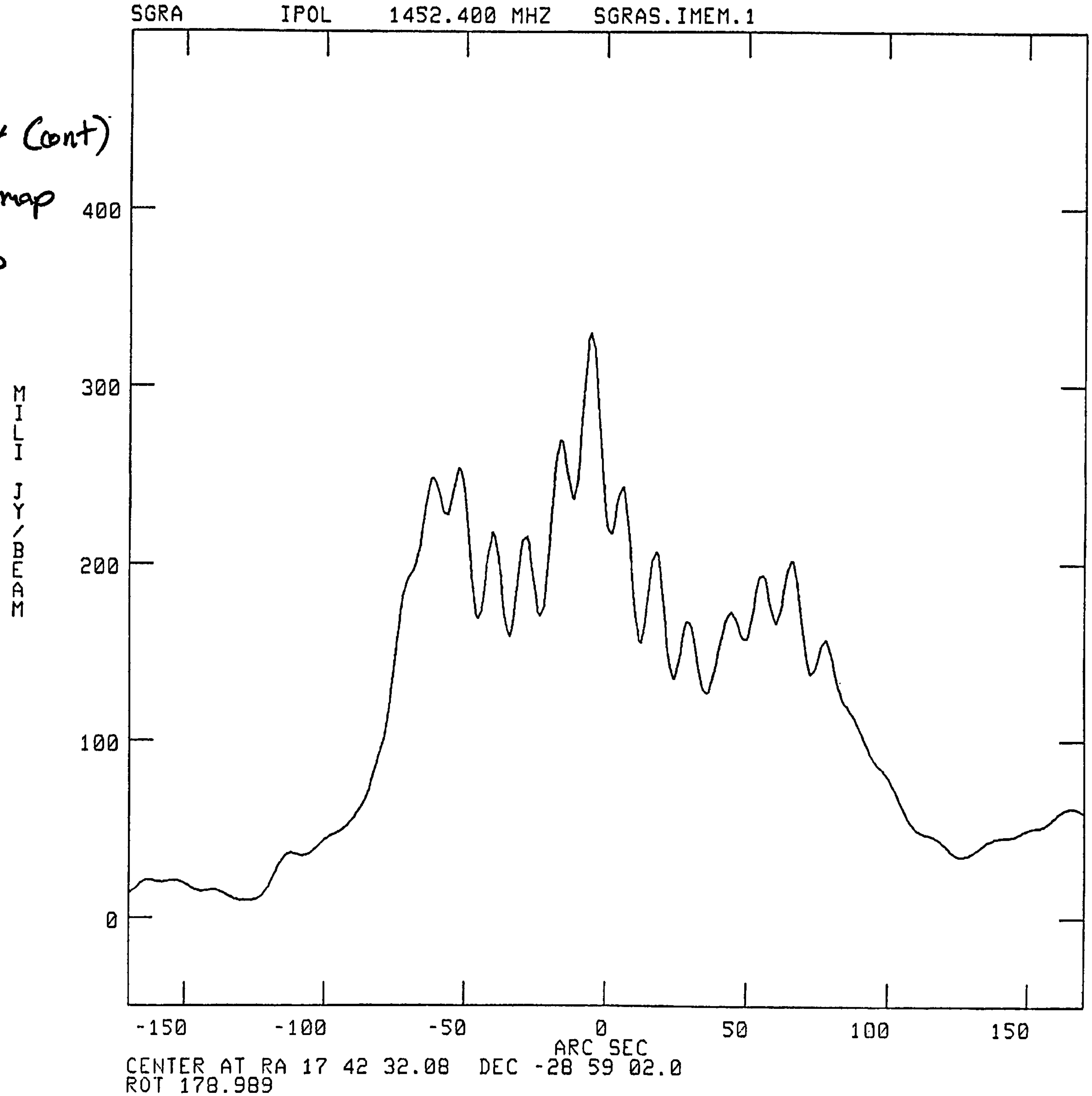


Fig 4 (cont)
MEM map
$\sigma = 50$




270

Fig 5

MEM map

$\sigma = 10$

Two more iterations

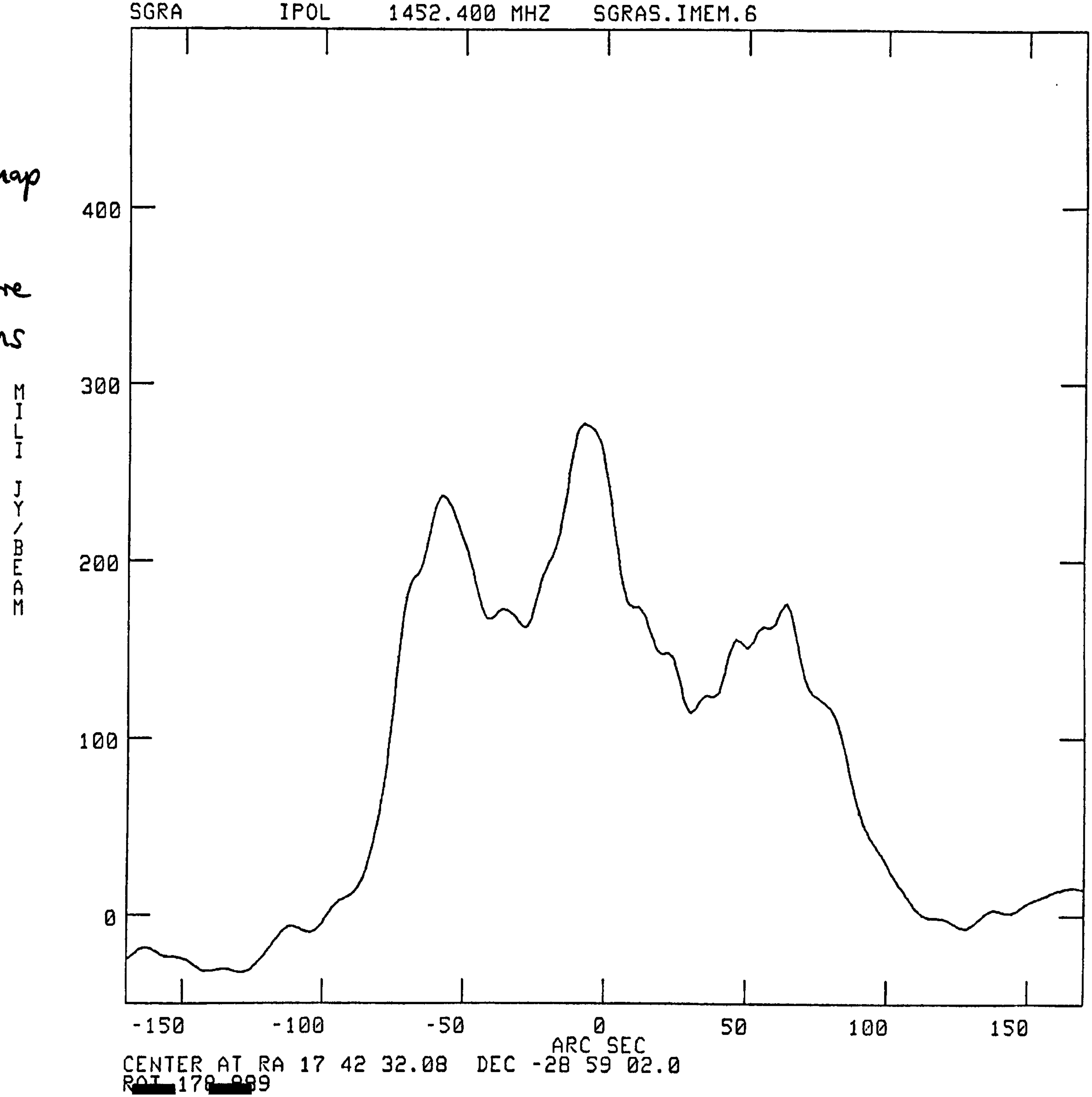
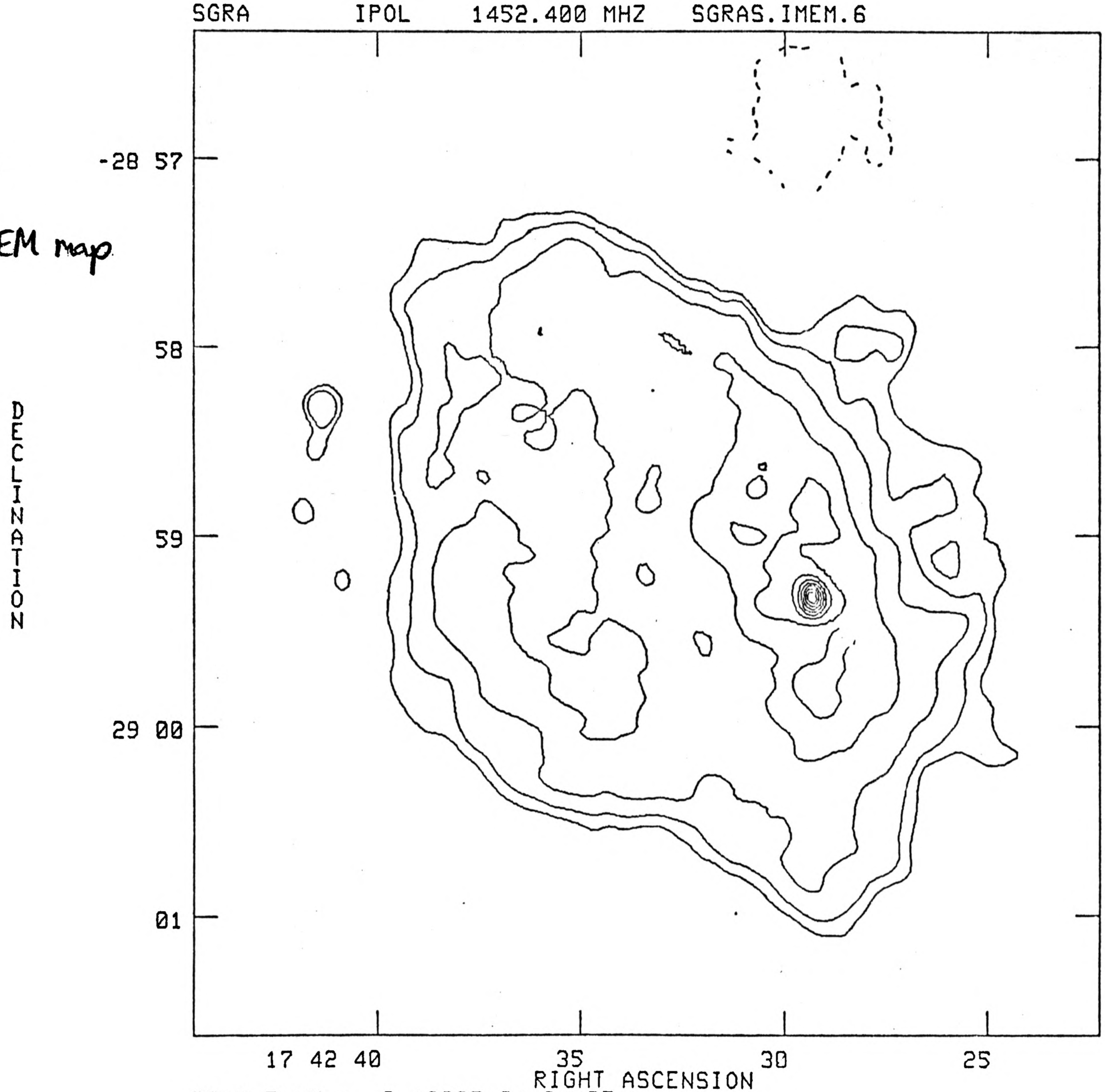


Fig 6

Final MEM map




PEAK FLUX = 0.1200E+01 JY/BEAM
LEVS = 0.1200E-01 * ( -3.0, 3.0, 5.0,
10.0, 20.0, 30.0, 40.0, 50.0, 60.0,
70.0, 80.0, 90.0)

Fig 7

MSM map

$\sigma = 10$

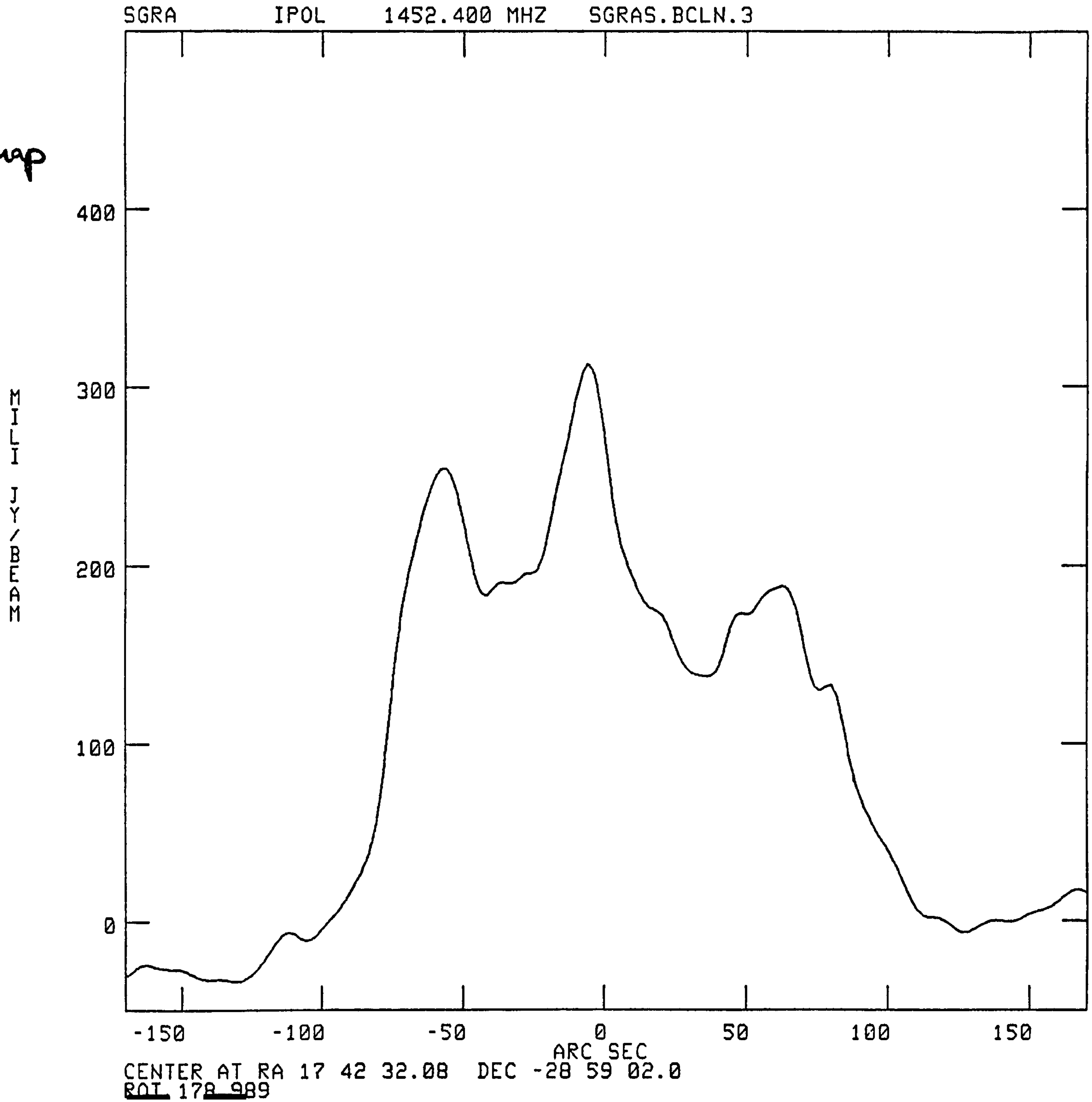


Fig 8.

MSM map

$\sigma = 10$