
Universiteit
Leiden
The Netherlands

A solution to the short spacing problem in radio interferometry

Braun, R.; Walterbos, R.A.M.

Citation
Braun, R., & Walterbos, R. A. M. (1985). A solution to the short spacing problem in radio interferometry. *Astronomy And Astrophysics*, *143*, 307-312. Retrieved from https://hdl.handle.net/1887/6844

Version:	Not Applicable (or Unknown)
License:	Leiden University Non-exclusive license
Downloaded from:	https://hdl.handle.net/1887/6844

Note: To cite this publication please use the final published version (if applicable).

1985A&A...143..307B

Astron. Astrophys. 143, 307–312 (1985)

ASTRONOMY AND ASTROPHYSICS

A solution to the short spacing problem in radio interferometry

R. Braun and R.A.M. Walterbos

Sterrewacht Leiden, Postbus 9513, 2300 RA Leiden, The Netherlands

Received July 27, accepted October 4, 1984

Summary. An efficient and reliable method is developed for the derivation of short spacing data missing from a radio interferometer observation, without the use of single dish data. The same approach is shown to apply to other instances of localized missing or damaged spatial frequency data. The method is based upon a direct fit of the missing Fourier coefficients to the (non-linearly) isolated map plane response.

Key words: interferometry – radio telescopes – data analysis – image processing

1. Introduction

A problem that has long plagued users of synthesis radio telescopes is the difficulty of satisfactorily mapping very extended objects. Since it is often not possible to observe the lowest spatial frequencies (corresponding to baselines $\lesssim$ the dish diameter), the instrumental response typically contains a shallow negative "bowl", as illustrated in Fig. 1. When this response is convolved with an extended brightness distribution it results in large scale background variations which make the quantitative interpretation of the observed image exceedingly difficult. This is illustrated in Fig. 2 by a WSRT 608 MHz observation of the supernova remnant VRO 42.05.01. In some cases it has been possible to remedy this situation by extracting and inserting the short spacing information from maps made with a single dish instrument (e.g. Bajaja and Van Albada, 1979). The feasibility and success of such an operation are dependent on a number of factors. First of all, single dish data of comparable sensitivity with a well determined base level must be available at the same frequency. Secondly, to allow a good relative calibration of the single dish data with the interferometer data, there should be sufficient overlap in spatial frequency response. In practice this means observations with a dish of diameter $\gtrsim$60 m for comparison with the shortest WSRT baseline of 36 m.

However, since the negative bowl is evident in the instrumental response of the interferometer alone it may be possible to remove this effect directly with a non-linear deconvolution algorithm. This has been the standard approach to address this problem in the absence of appropriate single dish data. In practice, the effective deconvolution of an extended two dimensional image is exceed-

Send offprint requests to: R. Braun

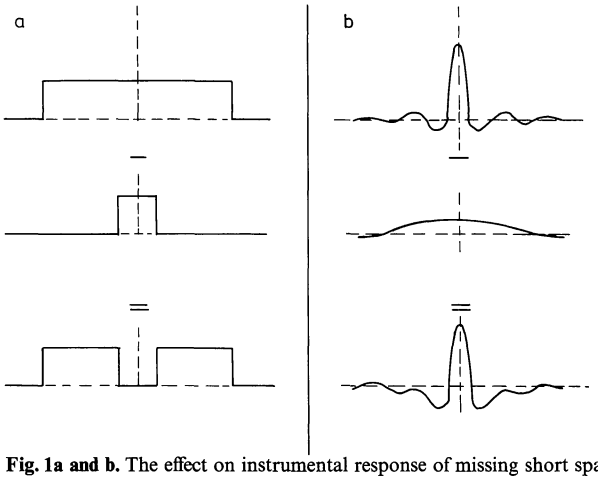


Fig. 1a and b. The effect on instrumental response of missing short spacings. **a** Observed spatial frequencies and **b** the corresponding instrumental response

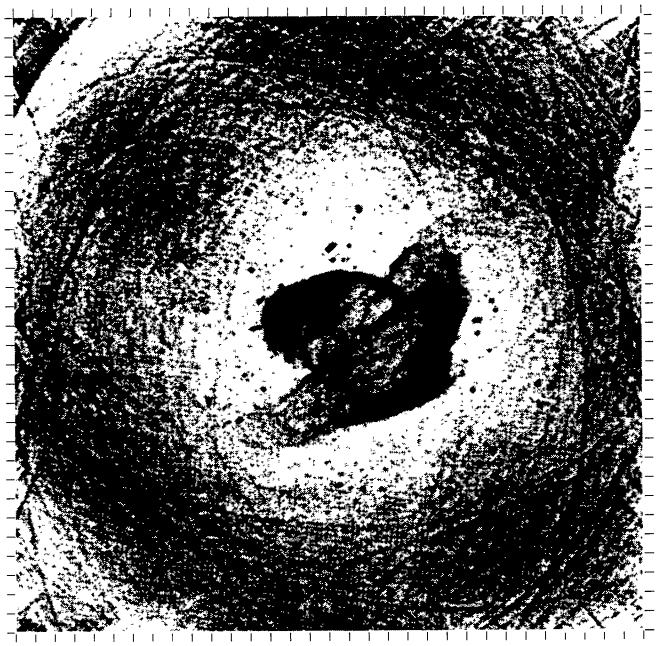


Fig. 2. The effect of missing short spacings on an observation of an extended object. Shown is a WSRT 608 MHz map (2° east-west) of the supernova remnant VRO 42.05.01 (Braun and Strom, in preparation) containing the spacings $B = n \cdot 18$ m for $n = 2, 3, 4, \ldots 153$. The grey-scale is based on a histogram of map intensities with mid-grey corresponding to zero, lighter shades negative and darker shades positive. The effect of missing short spacings is the negative bowl within which the source is immersed

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

1985A&A...143..307B

ingly difficult. The CLEAN algorithm (Högbom, 1974; Schwarz, 1978) is inefficient and in many cases ineffective in dealing with very extended sources. This algorithm tends to "converge" after a great deal of computing time to an unsatisfactory solution in which the bowl has not been effectively removed. An additional cause for concern are "CLEAN ripples", a processing artifact often generated in instances of extended sources with steep outer gradients. The original CLEAN algorithm has been extended by a number of authors to achieve higher efficiency (Clark, 1980) and diminish "CLEAN ripples" (Cornwell 1983a). Perhaps the most effective extension for application to very extended objects has been the "Multi-resolution CLEAN" algorithm developed by Brinks and Shane (1984) which achieves considerable gain in efficiency while recovering $\sim 80\%$ of the source flux. However, the fact remains, that a point by point deconvolution algorithm such as CLEAN is not the most appropriate method of modeling an extended brightness distribution.

Another approach to extended object deconvolution is provided by the various "Variational Methods (VM)". These methods attempt to determine the source model which gives statistical agreement with the data while satisfying an additional functional constraint. One of the earliest of these to receive astronomical application was the Maximum Entropy Method (MEM) (e.g. Gull and Daniell, 1978). In this case the constraint is the maximization of $-\Sigma f_i \log f_i$ for the source model pixels $\{f_i\}$. Extended source components grow most rapidly from the flat initial model, with subsequent sharpening as required by compact source components. This provides an intrinsic advantage over CLEAN in both reliability and (potential) efficiency when modeling extended sources. Refinements in numerical technique have led to versatile and efficient algorithms to solve the problem of constrained maximization in this context (Skilling, 1981; Cornwell, 1983b). Unfortunately, most researchers who have published data processed in this way have presented the source model as an end product. These methods are only useful for quantitative work when the source model has been appropriately smoothed to reflect the intrinsic spatial resolution of the instrument and the residuals of the model fit have been added back into the result (when these have significant information content) (Braun, in preparation). In practise, reasonable results can be obtained in this way with a factor of 2 to 5 improvement in computational efficiency over CLEAN. However, as with "Multi-resolution CLEAN" only $\sim 80\%$ of short spacing flux is typically recovered.

Given the difficulty that existing deconvolution algorithms have in recovering short spacing information, we will consider a more direct approach to this problem. The method we suggest is outlined in Sect. 2, while specific examples are presented in Sect. 3. In Sect. 4 we close with a discussion of the limitations of this approach and a summary.

2. The method

The problem, as already stated, is inadequate sampling of the inner portion of the spatial frequency plane. In the case of the WSRT, the spacings $B = n \cdot 18$ m are normally observed for $n = 2, 3, 4, \ldots, 153$, so that only the data at, $B = 0$ and $B = 18$ m (see Fig. 3) are missing from a complete (18 m increment) synthesis. (Similar gaps in the short spacings occur for other uniform arrays such as the 5 km telescope, Penticton, etc. while in the case of the VLA, the detailed shape and extent of the short spacing gap and grating response are determined by the array configuration, hour angle coverage and declination of the observed source. The general principles outlined below will also apply but must be adapted to the specific situation.)

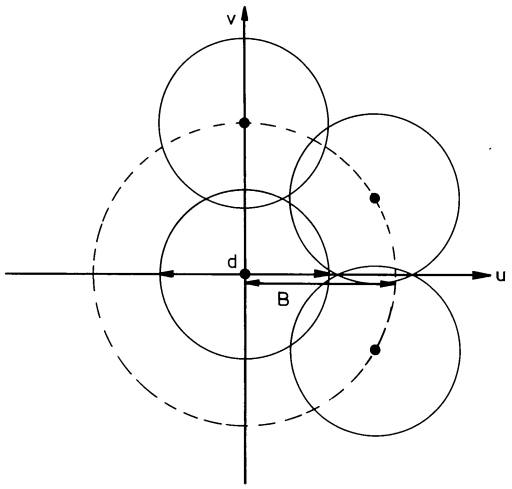


Fig. 3. The location of missing spatial frequency data and the number of independent samples needed to represent it. If the baselines $B = n \cdot \Delta B$ for $N = 2, 3, 4, \ldots$ have been observed, only the data at $B = 0$ and $B = \Delta B$ are missing from a uniform sampling of the spatial frequency plane. Since a correlated data point has FWHM d, the number of independent data points, N along a baseline track can be estimated by laying off this length along the track giving $N \approx \frac{\pi B}{d}$

We will now derive an approximate value for the number of independent data samples which were not observed. Due to the finite width of each dish of an interferometer pair, a range of spatial frequencies contributes to each correlation. The spatial frequency response of one dish is simply the autocorrelation of its aperture function, so that the spatial frequency response of an interferometer pair is the correlation of the two contributing aperture functions. In practical terms, if two dishes have approximately Gaussian primary beams with FWHM (Full Width at Half Maximum) corresponding to d_1 and d_2 meters, then their aperture functions will have FWHM $\sqrt{2}d_1$, and $\sqrt{2}d_2$ so that the spatial frequency response will have a FWHM of

$$d = \left[\frac{1}{2d_1^2} + \frac{1}{2d_2^2}\right]^{-1/2} \text{m}. \tag{1}$$

In the case of the WSRT and the VLA all dishes have an effective FWHM, $d_1 = d_2 \approx 20$ m so that a correlated data point also has FWHM, $d \approx 20$ m. A spatial frequency track for a baseline of B meters then has

$$N \approx \frac{\pi B}{d} \text{ independent data points} \tag{2}$$

as illustrated in Fig. 3. (Only half the plane need be measured since a real brightness distribution has a Hermitian Fourier transform.) Specifically, for $B = 18$ m and $d = 20$ m there are only $N = 3$ independent points which have not been observed. The number of points needed to completely determine such a track is then $\approx 2N$, i.e. 2 samples per "beam".

Let us now consider the map plane counterpart to a spatial frequency data point. A complex data point in the (u, v) plane can be expressed in radial coordinates as

$$f(B, \theta) = R \cdot [\delta(B - B')\delta(\theta - \theta') + \delta(B - B')\delta(\theta - \theta' + \pi)] + jI \cdot [\delta(B - B')\delta(\theta - \theta') - \delta(B - B')\delta(\theta - \theta' + \pi)] \tag{3}$$

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

1985A&A...143..307B

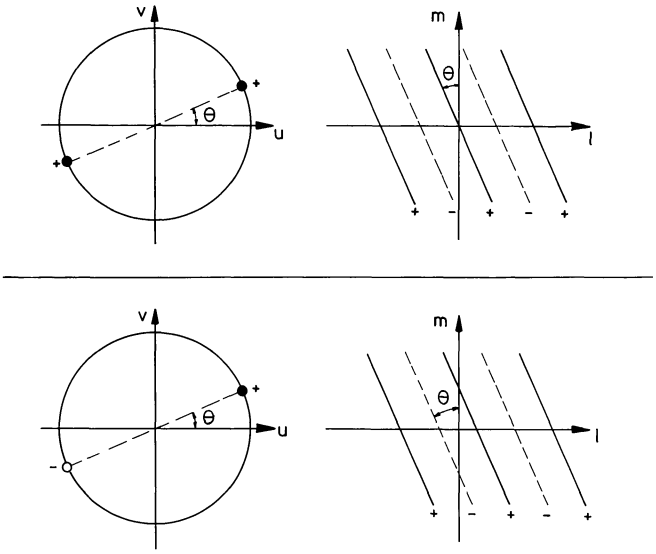


Fig. 4. The map plane counterpart to spatial frequency data. A real data point at (B, θ) gives rise to a cosine ripple with wavelength B while an imaginary data point results in a sine ripple

which has the Fourier transform,

$$F(l,m)=R\cdot\cos\left[\frac{2\pi}{B'}(l\cos\theta'+m\sin\theta')\right] +I\cdot\sin\left[\frac{2\pi}{B'}(l\cos\theta'+m\sin\theta')\right]. \quad (4)$$

This relationship is illustrated in Fig. 4.

We are now prepared to write a complete analytic description of the missing short spacing information in the map plane:

$$\text{i.e. } F(l,m)=A_0+\sum_{i=1}^{2N}\left\{R_i\cos\left[\frac{2\pi}{B_i}(l\cos\theta_i+m\sin\theta_i)\right] +I_i\sin\left[\frac{2\pi}{B_i}(l\cos\theta_i+m\sin\theta_i)\right]\right\}, \quad (5)$$

where A_0 represents the $B=0$ component and the sum is over the $2N$ spatial frequency points at (B_i, θ_i). The short spacing problem can therefore be directly addressed by determining the small number of coefficients, $\{A_0, R_i, I_i\}$ in Eq. (5).

As was clear from Fig. 2, the result of missing short spacing information is a large scale variation in the map background; especially the negative bowl within which the source is immersed. (The remnant of such a bowl is the typical indication that a deconvolution algorithm has not succeeded in fully recovering the source flux.) A reasonable demand on the short spacing coefficients is that they produce a flat background in the map area outside of any source, but within the first grating ring. In practice this condition provides an effective constraint which allows the coefficients of Eq. (5) to be derived from the minimization of

$$C=\sum_{l',m'}[F(l,m)-M(l,m)]^2, \quad (6)$$

where the summation is carried out for functional values $F(l,m)$ and input map values $M(l,m)$ over a subset of the map plane pixels $\{l',m'\}$.

The derivation of appropriate map plane values $M(l',m')$ for use in Eq. (6) is also a matter of some consequence. The effectiveness of the method depends upon the isolation in the map plane of only those spatial frequencies which are to be fit. Since most source and grating structure contains a large range of spatial frequencies, such map areas must not be used. The efficiency of the processing can be enhanced by effective data compression from the outset since only the long wavelength background variations of the map are desired in constraining the short spacing coefficients. We have adapted a $\kappa\sigma$ clipping algorithm (e.g. Herzog and Illingworth, 1977) developed for the processing of optical imagery for this purpose. This method is used to determine a local mean background level in subregions (e.g. 16×16 pixels) of the input map. The intensities most deviant ($>2\sigma$) from the mean are iteratively discarded, with the final mean intensity adopted for the entire subregion (consisting now of only one pixel). In one step this procedure provides compact source rejection, noise reduction (factor 10) and data compression (by a factor of ~100 in pixels). The appropriate map plane values can then be isolated from the compressed data set by masking out extended sources, grating response and any aliasing along the map edges, leaving $\sim10^3$ values for use in Eq. (6). Efficient algorithms for solving non-linear least squares problems of this size have become routinely available (e.g. Gill and Murray, 1978). We have made use of a NAG library (Version 10) subroutine (EO 4 FDF) for this purpose.

Once the short spacing coeffients have been determined, the functional form [Eq. (5)] can be added directly to the original image; or equivalently, the derived background variations (-1 times the short spacing data) can be subtracted from the original image.

3. Application

The practical use and effectiveness of this method will be illustrated by a number of specific examples. The effects of missing short spacing data can be clearly seen in the 512×512 pixel map shown in Fig. 2. This map was generated from a 608 MHz WSRT observation (Braun and Strom, in preparation) where only the $B=0$ and $B=18$ m spacings are missing from a complete (18 m increment) synthesis. Processing with the $\kappa\sigma$ algorithm produced the 64×64 pixel map shown in Fig. 5a. Blocks of 16×16 pixels of the input map (Fig. 2) were used to determine the local background level. The masks used to exclude the extended emission of the source as well as grating and aliased response along the map edge have been applied. The remaining points of this map were used to constrain the coefficients $\{A_0, R_i, I_i\}$ of Eq. (5) along the track $B=18$ m at angles $\theta=-60°, -30°, 0°, 30°, 60°$, and 90°. The functional form which was so derived has been added to the map in Fig. 5b. The "bowl" is quite effectively removed in this way, although a low frequency ripple with wavelength corresponding to 36 m remains at $\theta\approx45°$. An inspection of the spatial frequency data indicated that strong interference due to the sun was present on the 36 m baseline in this hour angle range. Since the problem of correcting bad data at a specific spatial frequency is very similar to the determination of unmeasured data we can address it in the same way. The same masks shown in Fig. 5a were applied to Fig. 5b to isolate appropriate data points. These points were used to constrain the set of coefficients $\{A_0, R_i, I_i\}$, but now along the track $B=36$ m at angles $\theta=-45°, 0°, 45°$, and 90°. (Since the bad data was concentrated at $\theta=45°$, the set of 12 angles needed to completely define the $B=36$ m track were not used). The derived functional form was added back into the map to give the map shown in Fig. 5c. Background variations are now virtually eliminated from the data. The derived background (the combined effect of missing or damaged data at $B=0$, 18, and 36 m) is

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

1985A&A...143..307B

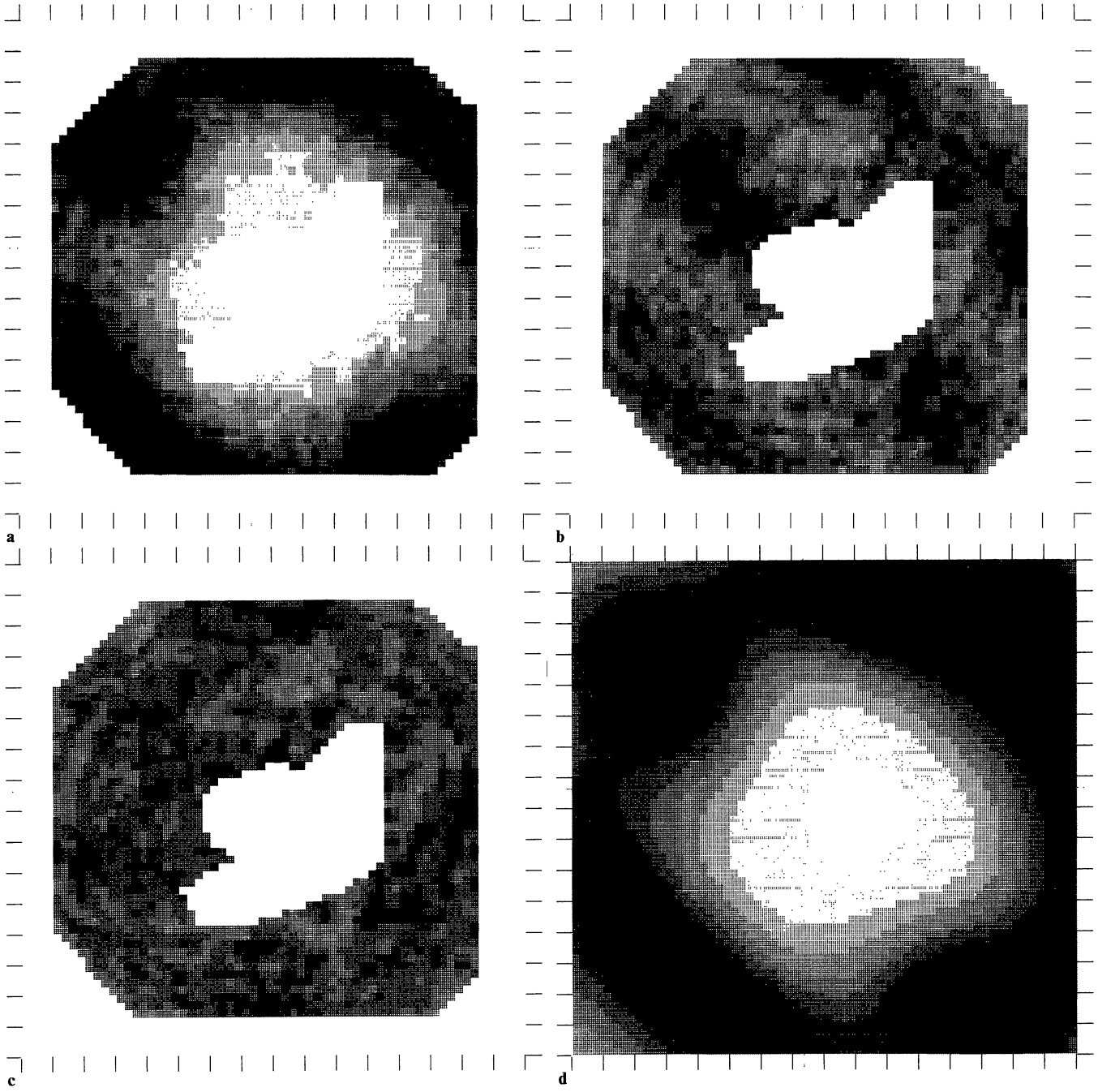


Fig. 5a-d. Illustration of the processing. **a** The map of Fig. 2 after kappa-sigma processing and masking of extended emission, grating and aliased response. The grey-scale is based on a histogram of the remaining map intensities as in Fig. 2 and is kept fixed for **a-d. b.** The map of **a** after removal of the derived background variations at $B=0$ m and $B=18$ m. The "bowl" is quite effectively removed but a ripple with wavelength corresponding to $B=36$ m remains. **c.** The map of **b** after removal of the derived background variations at $B=36$ m. This map is essentially flat and zero. **d.** The derived background variations due to $B=0$, 18, and 36 m

illustrated in Fig. 5d for comparison with the input in Fig. 5a. Finally, the derived background was oversampled (by padding its Fourier transform with zeroes) in order to obtain the 512×512 pixel array which was subtracted from the full resolution input. This result is illustrated in Fig. 6. A comparison with Fig. 2 shows the radical improvement achieved by this processing. What is perhaps equally impressive is that the entire procedure required less than 2 min computing time (on the AMDAHL V7B of the Leiden University Computing Institute). An estimate of the extent to which the source flux has been recovered is not possible for this source since its frequency spectrum is still too poorly determined.

Another observation which suffers from missing short spacing data is depicted in Fig. 7a. Shown are the central 512×512 pixels of one polarization channel of a 1412 MHz WSRT observation (Braun and Strom, in preparation) of the supernova remant DA 530. This observation also contains the spacings $B=n \cdot 18$ m for $n=2, 3, 4, \ldots 153$. Just as in the previous example, the complete field (in this case 1024×1024 pixels) was $\kappa\sigma$ processed, masked and

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

1985A&A...143..307B

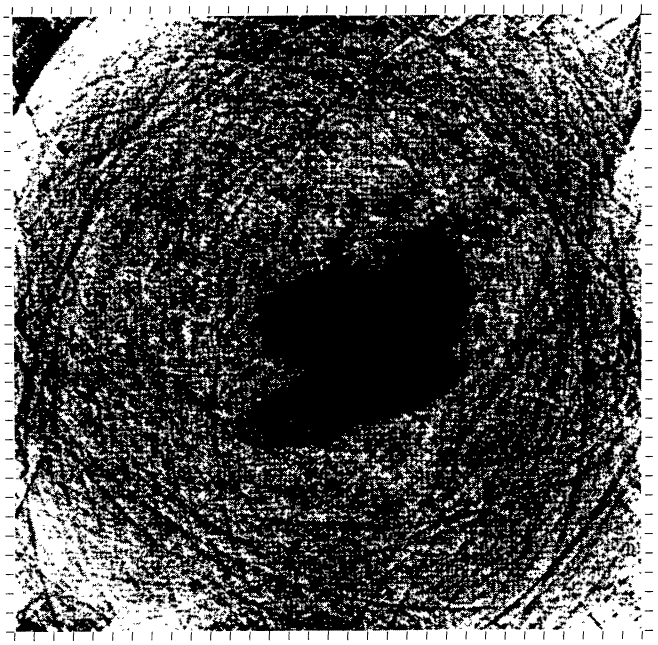
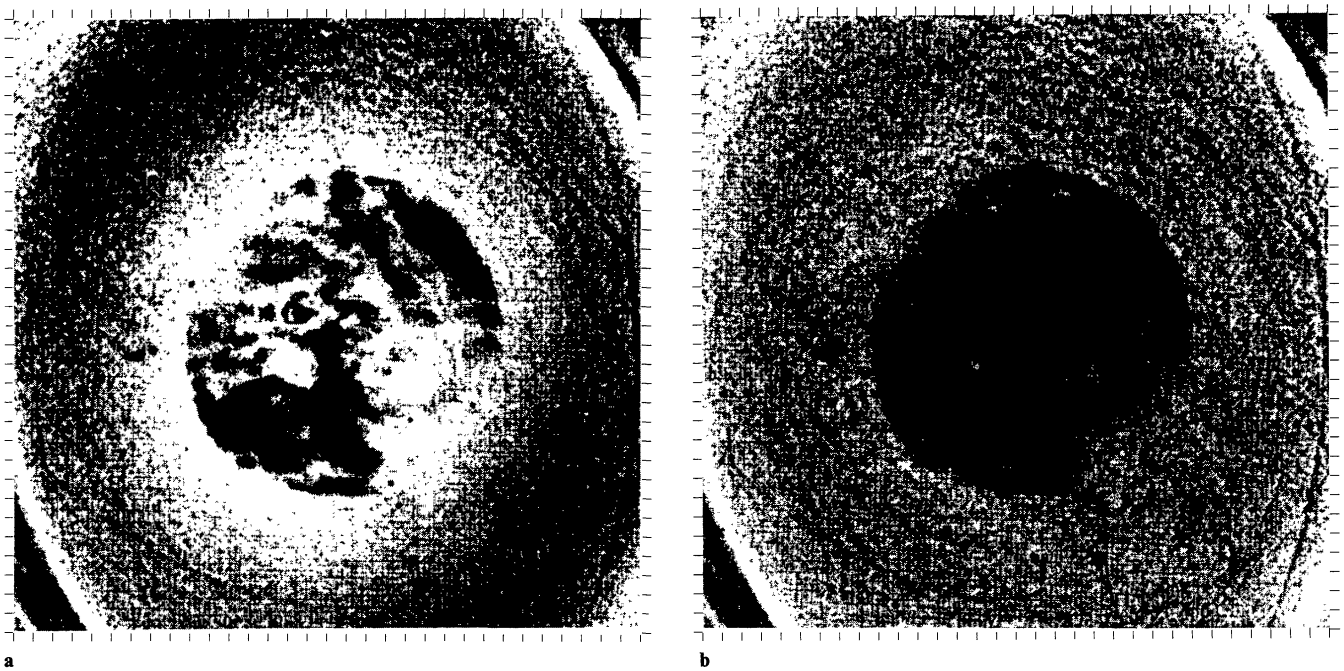


Fig. 6. The map of Fig. 2 after correction for missing data at $B=0$ and $B=18$ m and interference at $B=36$ m. The grey-scale used is that of Fig. 2

used to constrain the missing coefficients at $B=0$ m and along the $B=18$ m track. The result (after $2^1/_2$ min computing time) is shown in Fig. 7b. The other polarization channels at 1412 MHz as well as a 608 MHz WSRT observation of this source (Braun and Strom, in preparation) were similarly processed and used to derive source fluxes. These fluxes are plotted in Fig. 8 together with previously published flux determinations. The presently derived fluxes (from the interferometer data alone) are in excellent agreement with single dish determinations and are consistent with all of the source flux being accurately recovered.

4. Discussion

The examples of the previous section have illustrated the effectiveness of this approach in both the correction of damaged data which was observed and the derivation of short spacing data which was not observed. The success of this method is dependent on the availability of enough map plane area that contains exclusively those spatial frequencies which are to be derived so as to adequately constrain the choice of coefficients in the known functional form [Eq. (5)]. In one dimension the problem can be more easily visualized. We are attempting to determine the phase and *DC* offset of a cosine wave of fixed wavelength based on samples of the waveform over a limited spatial extent. In the absence of noise the sampling of even a small fraction of a wavelength would be sufficient to allow its accurate estimation, while in practise $\sim^1/_2$ wavelength should be ample. In two dimensions the problem is similar, although at each position angle the cosine ripples are attenuated with increasing distance from the source through interference with ripples of other orientations. In the absence of noise this would not change matters but in practise it becomes important to sample the waveform as close to the source position as possible, where its amplitude is largest. An annular region of thickness $\gtrsim^1/_4$ wavelength would provide sampling over $\gtrsim^1/_2$ wavelength at each positional angle, which should be sufficient for an accurate estimation. The regular sampling interval of the WSRT results in the concentration of grating response into rings at a radius corresponding to the wavelength of the sampling interval. Since this is typically the shortest spacing which is to be derived sources smaller than about 75% of the first grating radius can be expected to have a sufficiently well defined background. This limitation is closely



a **b**

Fig. 7a and b. Before and after prcessing. **a** A WSRT 1412 MHz map (0°85 east-west) of the supernova remnant DA 530 (Braun and Strom) containing the spacings $B=n\cdot 18$ m for $n=2, 3, 4, \ldots 153$. The grey-scale is based on a histogram of map intensities and is kept fixed for **a** and **b**. **b** The map of **a** after correction for missing data at $B=0$ and $B=18$ m

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

1985A&A...143..307B

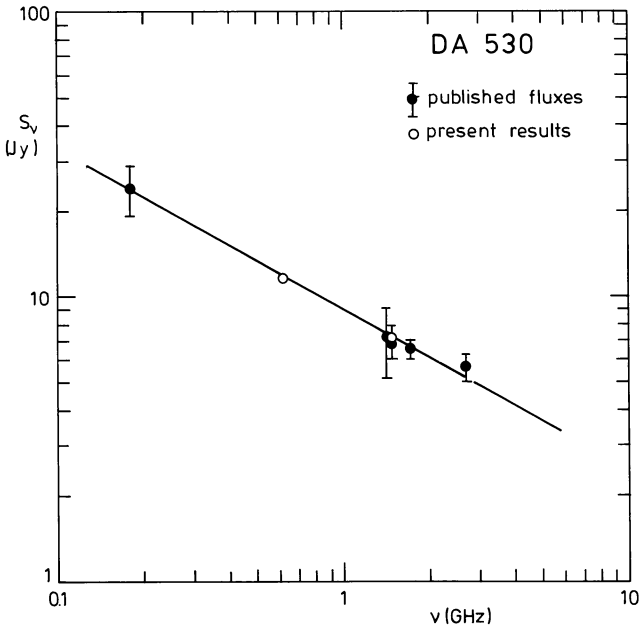


Fig. 8. Comparison of flux determinations. Published single dish fluxes for the supernova remnant DA 530 taken from Haslam et al. 1980, together with fluxes at 608 and 1412 MHz determined from WSRT data alone after short spacing processing. There is excellent agreement within the 5% absolute flux calibration of the WSRT

related to the demands of the sampling theorem which are also not satisfied for sources as large as the grating radius. On the other hand, the irregular sampling of the VLA makes it more appropriate to speak of far sidelobes rather than grating response. To the extent that sidelobe fluctuations average to zero on the scale size of the spatial frequencies of interest, these should not adversely affect the determination of the short spacings.

The great computational efficiency of this approach is a direct consequence of effective data compression. The specific problem with the spatial frequency data is addressed directly by way of its map plane counterpart. The $\kappa\sigma$ processing is used to isolate (together with masking) the long wavelength information while reducing the size of the problem by a factor $\sim 10^2$. Computing time is consequently reduced by a factor of 20 to 50 (with vastly superior results) over present implementations of either CLEAN or the *VM*.

In summary, we suggest approaching the problem of satisfactory interferometer imaging directly using the available constraints. Instead of determining an idealized source model as an intermediate step, we suggest the direct determination of only that spatial frequency data which was known to be poorly determined using the constraint that the map area free of sources become a flat background.

Acknowledgements. It is a pleasure to thank R. G. Strom and C. van Schooneveld for a critical reading of the manuscript. The authors were financially supported by the Netherlands Foundation for Astronomical Research (ASTRON). The Westerbork Synthesis Radio Telescope is operated by the Netherlands Foundation for Radio Astronomy (SRZM). Both Foundations are supported by the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

Referenes

Bajaja, E., van Albada, G.D.: 1979, *Astron. Astrophys.* **75**, 251
Brinks, E., Shane, W.W.: 1984, *Astron. Astrophys. Suppl.* **55**, 179
Clark, B.G.: 1980, *Astron. Astrophys.* **89**, 377
Cornwell, T.J.: 1983a, *Astron. Astrophys.* **121**, 281
Cornwell, T.J.: 1983b, VLA Scientific Memorandum 149
Gill, P.E., Murray, M.: 1978, SIAM *J. Numer. Anal.* **15**, 977
Gull, S.F., Daniell, G.J.: 1978, *Nature* **272**, 686
Haslam, C.G.T., Pauls, T., Salter, C.J.: 1980, *Astron. Astrophys.* **92**, 57
Herzog, A.D., Illingworth, G.: 1977, *Astrophys. J. Suppl.* **33**, 55
Högbom, J.: 1974, *Astron. Astrophys. Suppl.* **15**, 417
Schwarz, U.J.: 1978, *Astron. Astrophys.* **65**, 345
Skilling, J.: 1981 (preprint)

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