Astron. Astrophys. 65, 345—356 (1978)

ASTRONOMY AND ASTROPHYSICS

Mathematical-statistical Description of the Iterative Beam Removing Technique (Method CLEAN)

U. J. Schwarz

Kapteyn Laboratory, University of Groningen, Postbus 800, NL-8002 Groningen, The Netherlands

Received October 13, 1976; revised July 15, 1977

Summary. The CLEAN method (Högbom, 1974), which is a deconvolution method, is analysed mathematically for the 1-dimensional case. It is shown that the method is equivalent to solving a system of linear equations by an iterative method (Temple, 1938). A criterion of convergence is given. In typical applications of the method the solution of the system of equations is *not* unique and the consequences for the CLEAN solution are discussed.

By applying the analysis to maps which are obtained from Fourier transformed data the convergence criterion is shown to be equivalent to the condition that all weights used for the Fourier transform have to be non-negative. It is proven that the method is in fact a statistically correct least-squares fit of sine functions to the observed data (the visibility function). The choice of clean beam and the effects of adding residuals are analysed.

An error analysis is given which allows the errors in interpolated and extrapolated information to be determined. Some numerical examples of error calculations are given.

The extension of the present analysis to 2 dimensional distributions is briefly discussed in an Appendix.

Key words: CLEAN — data-processing — Fourier transform — statistics — radio astronomy

1. Introduction

In recent years there has been an increasing number of radio synthesis observations where the iterative beam removing technique, known under the name CLEAN has been used successfully (Högbom, 1974; Schwarz et al., 1973; Rogstad et al., 1971). This technique proves to be useful under a variety of different circumstances. CLEAN is in its simplicity intuitively appealing, but until now no criteria have been available to decide under which conditions the method may be applied and what the limitations are. Most of the users probably have made various tests to convince themselves that the method works for their application. The more widespread the use of the method gets, the more urgent it is to establish some criteria for its range of applications in order to avoid incorrect use which may lead to wrong results.

In this paper the method is analysed mathematically and statistically to give a basis from which criteria can be developed to decide under which conditions CLEAN works. Such an analysis provides the basis for further refinements of CLEAN. It can also be useful in comparing this method with other methods, such as the proposed "maximum entropy" method (Ables, 1974).

In Part I we discuss the method in its simple form, as seen by most users, namely as a deconvolution procedure. Although they mostly use Fourier transformed data it is not immediately apparent that this is important. But in Part II we shall see that CLEAN is especially suited for such Fourier transformed data. It will be shown that CLEAN is not just another deconvolution procedure, but is a statistically correct fitting procedure, when dealing with noisy Fourier transformed data.

Part I. CLEAN as a Deconvolution

2. The Mathematical Description of the Method

The foundation of the basic ideas of CLEAN is given in the paper by Högbom, 1974. In order to make the present paper more easily understood, we give a very brief summary of the method, introducing at the same time the required mathematical symbols.

2.1. Brief Description of the Method and Definitions of Terms

The method is used in the case that one has an observed map, the "*dirty*" map (d), which is the convolution of the *brightness distribution* (t_0) with an instrumental response, called the "*dirty*" beam (b) (this nomenclature comes from the application of the method in radio astronomy; in optics these terms correspond to "image", "object" and "point-spread function", respectively). The "dirty" beam may have some unwanted secondary responses, as

sidelobes, wings, etc. The aim of the method is to remove the effects of these secondary responses. This is done in two steps: first a deconvolution step in which the dirty map is decomposed in a set of scaled δ-functions, the *components*, t, which, when convolved with the "dirty" beam would reproduce the original "dirty" map. Second, the components are convolved with a hypothetical "*clean*" *beam*, h, which is free from the unwanted responses. This finally gives the "*clean*" *map*, c. We can also define a "*true clean*" map, c_0, which is the convolution of the ("true") brightness distribution with the "clean" beam h.

$$\begin{aligned} d &= b*t = \text{"dirty" map}, \\ c &= h*t = \text{"clean" map}, \\ c_0 &= h*t_0 = \text{"true clean" map}. \end{aligned} \tag{2.1}$$

The deconvolution need not be complete, i.e. leaving *residuals*, r, which often are added to the "clean" map (see Sect. 9). The "clean" map c can be regarded as an estimate of the "true clean" map c_0. We denote the "true" distributions and quantities by adding the subscript 0.

The deconvolution is not necessarily unique and this is the main problem. In order to overcome this, some extra information about the brightness distribution must be used. The CLEAN method is designed for the case that the brightness distribution contains only a few sources at well separated, small regions, i.e. *the brightness distribution is essentially empty*. CLEAN makes use of these characteristics in the following way: it searches for the maximum in the correlation between the "dirty" map and the "dirty" beam, which is in some applications identical (or in general close) to the absolute largest value in the "dirty" map. We can make the plausible (but not necessary) assumption that this response is mainly due to a real signal and only a minor part comes from the filter response from other sources placed further away. Some fraction g times the absolute largest value is accepted as a first approximation of the set of δ-functions. A dirty beam pattern scaled by this value and centered at the corresponding position is subtracted from the dirty map thereby removing a great deal of the unwanted secondary responses. The procedure is repeated on the remaining map and by successive iterations one builds up the set of δ-functions. Overcorrection in an early stage is later corrected automatically, because the method allows also corrections of negative intensities.

2.2. Simplifying Assumptions

The maps are in general continuous functions of one or more coordinates. However, in order to make the method directly applicable for pratical cases we assume that the maps are sampled at a finite number N points.

In this first part we will make another assumption, namely that the brightness distribution is sampled adequately. In Part II this requirement is not used anymore; there we assume only, that the "dirty" map is sampled adequately. For further simplification of the mathematical formulation we restrict ourselves to one-dimensional maps. The two-dimensional case can be treated as an extension of the one-dimensional case described here (see Appendix).

2.3. The Vector and Matrix Notation

The above restrictions enable us to formulate the method CLEAN in an algebraically simple way. The maps and the scaled delta functions, the components, will be written as vectors containing N elements ($\mathbf{d}$, $\mathbf{c}$, $\mathbf{c}_0$, $\mathbf{t}$, $\mathbf{t}_0$). An individual element is referred to as "element at a given position". The "dirty" and the "clean" beam are written in matrix notation, in the following way:

$$\mathbf{B} = \begin{bmatrix} 1 & b_1 & b_2 & b_3 & \dots & b_{N-1} \\ b_1 & 1 & b_1 & b_2 & \dots & b_{N-2} \\ b_2 & b_1 & 1 & b_1 & \dots & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ b_{N-1} & & \dots & & & 1 \end{bmatrix}. \tag{2.2}$$

Thereby a *regular* grid of coordinates is used and the "dirty" beam is assumed to be *symmetric*. This results in a "Toeplitz" matrix. The convolution can now be written as a multiplication of a square matrix with a vector, i.e.

$$\mathbf{d} = \mathbf{B}\mathbf{t}_0 . \tag{2.3}$$

If we consider a subset of the above equations, we get a matrix of the more general form ($W < N$)

$$\mathbf{B} = \begin{bmatrix} 1 & b_{0,1} & b_{0,2} & \dots & b_{0,W-1} \\ b_{1,0} & 1 & b_{1,2} & \dots & \dots \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ b_{W-1,0} & b_{W-1,1} & b_{W-1,2} & \dots & 1 \end{bmatrix}. \tag{2.4}$$

The symmetry of the beam leads to a symmetric matrix, that means that $b_{i,k} = b_{k,i}$. Such a subset of the equation will be discussed in Section 4. The above notation can also be used in the general case, when the sampling is at arbitrary coordinates. A simple example may illustrate this.

Example: $\mathbf{t}_0$ is given at 4 coordinates, x_1 to x_4. The beam at the relative distance $x_i - x_k$ is $b(x_i - x_k)$. We now have the following scheme:

coordinate:	x_1	x_2	x_3	x_4
$\mathbf{t}_0$	t_{0_1}	t_{0_2}	t_{0_3}	t_{0_4}
$\mathbf{d}$	d_1	d_2	d_3	d_4

The "dirty" map $\mathbf{d}$ is the convolution of $\mathbf{t}_0$ with the beam, written explicitely:

$$\begin{aligned} d_1 &= t_{0_1} + t_{0_2}b(x_2 - x_1) + t_{0_3}b(x_3 - x_1) + t_{0_4}b(x_4 - x_1) \\ d_2 &= t_{0_1}b(x_1 - x_2) + t_{0_2} + t_{0_3}b(x_3 - x_2) + t_{0_4}b(x_4 - x_2) \\ d_3 &= t_{0_1}b(x_1 - x_3) + t_{0_2}b(x_2 - x_3) + t_{0_3} + t_{0_4}b(x_4 - x_3) \\ d_4 &= t_{0_1}b(x_1 - x_4) + t_{0_2}b(x_2 - x_4) + t_{0_3}b(x_3 - x_4) + t_{0_4}. \end{aligned} \tag{2.5}$$

The coefficients of the t_{0_i} are the elements $b_{i,k}$ of the matrix **B**. A symmetric beam means that $b(x_i - x_k) = b(x_k - x_i)$, therefore the matrix **B** is symmetric and there are at most six independent elements, corresponding to the values of the beam at the relative distances of the four coordinates.

An additional symmetry of the matrix **B** can arise in Fourier transformed data (cf. Sect. 6.1).

In this paper we will use sometimes a special type of beam, the "*uniform weight*" beam. The name comes from the definition of a beam being the Fourier transform of a set of weights, see Part II. If the non-zero weights are all equal, this leads to a beam with the characteristic that its self-convolution is the original beam times a scale factor f; in matrix notation:

$$\mathbf{BB} = f\mathbf{B}\,. \tag{2.6}$$

Although it is quite a special type of beam, it makes many concepts of CLEAN more easily understandable in Part I. All those parts involving the "uniform weight" beam will be made more general in Part II.

2.4. The Algebraic Formulation of the Iteration Process

The iteration process we can now describe, introducing $\mathbf{t}'$ and $\mathbf{r}$, the components and residuals, respectively, after a number of iterations:

$$\mathbf{r} = \mathbf{d} - \mathbf{B}\mathbf{t}'\,. \tag{2.7}$$

A factor g times the absolute largest value of the residuals is considered as the next approximation of the set of components, and the scaled response of the dirty beam is subtracted from the dirty map. The remaining map is searched again for the absolute largest value, to be used for the correction in the next iteration, etc.

If we denote by r_m the largest absolute value of the residuals found at a position m, the correction to the vector of the components t' can be expressed as

$$\delta\mathbf{t}' = (0,\ldots,0,+gr_m,0,\ldots,0)\,. \tag{2.8}$$

The set of components is updated

$$\mathbf{t}'' = \mathbf{t}' + \delta\mathbf{t}'\,, \tag{2.9}$$

and the convolution with the dirty beam is subtracted from the residuals

$$\mathbf{r}' = \mathbf{r} - \mathbf{B}\delta\mathbf{t}'\,. \tag{2.10}$$

For the next iteration we get

$$\mathbf{r}'' = \mathbf{r}' - \mathbf{B}\delta\mathbf{t}'' = \mathbf{r} - \mathbf{B}(\delta\mathbf{t}' + \delta\mathbf{t}'')\,. \tag{2.11}$$

Working backwards substituting the residuals of the previous iterations, using the above recursion relations (2.9) and (2.10) we get

$$\mathbf{r} = \mathbf{d} - \mathbf{B}\sum\delta\mathbf{t}^{(i)}\,, \tag{2.12}$$

which is equivalent to Equation (2.7).

If the residuals converge to zero, then one solves

$$\mathbf{B}\mathbf{t} = \mathbf{d}\,, \tag{2.13}$$

a system of N linear equations, by an iterative method. This method is in fact a relaxation method discussed by Southwell (1935) and Temple (1938), see also Forsythe and Wasow (1960).

If the determinant of **B**, det (**B**) $\neq 0$, then the solution is unique and $\mathbf{t} = \mathbf{t}_0$. If, however, det (**B**) $= 0$, then there is only a solution if **d** is a linear combination of the vectors, corresponding to the lines of the matrix **B**, i.e. **d** is in the range of **B**; the solution is *not* unique.

The method is very similar to the well known Gauss-Seidel method, to solve linear equations by an interative method. There the corrections $\delta\mathbf{t}'$ are calculated systematically at the positions 1 to N, and $g = 1$.

3. Criterion of Convergence

Temple (1938) proved that the method converges starting from any initial approximation if the matrix **B** is symmetric and *positive definite*. The convergence is fastest if the gain factor $g = 1$.

We will outline the proof briefly: we define the *norm*

$$Q = (\mathbf{t}' - \mathbf{t}_0)^T\mathbf{B}(\mathbf{t}' - \mathbf{t}_0) = \Delta\mathbf{t}^T\mathbf{B}\Delta\mathbf{t}\,, \tag{3.1}$$

with

$$\Delta\mathbf{t} = \mathbf{t}' - \mathbf{t}_0\,. \tag{3.2}$$

The difference of the norm in two consecutive iterations is

$$\delta Q = \Delta\mathbf{t}'^T\mathbf{B}\Delta\mathbf{t}' - \Delta\mathbf{t}^T\mathbf{B}\Delta\mathbf{t} \quad (\text{with } \Delta\mathbf{t}' = \mathbf{t}'' - \mathbf{t}_0)\,. \tag{3.3}$$

For a correction $\delta\mathbf{t}' = \mathbf{t}'' - \mathbf{t}'$ this can be written, making use of the symmetry of **B**

$$\delta Q = 2\delta\mathbf{t}'^T\mathbf{B}\Delta\mathbf{t} + \delta\mathbf{t}'^T\mathbf{B}\delta\mathbf{t}'\,. \tag{3.4}$$

Using for the correction $\delta\mathbf{t}'$ as defined in Eq. (2.8) the expression simplifies to

$$\begin{aligned}\delta Q &= -2gr_m^2 + (gr_m)^2 \\ &= (g^2 - 2g)r_m^2\,.\end{aligned} \tag{3.5}$$

δQ is negative if

$$0 < g < 2\,. \tag{3.6}$$

In other words: the norm Q is a monotonically decreasing quantity with lower bound 0 (because **B** is assumed to be positive definite). It can become zero only if $\Delta\mathbf{t} = 0$. The above derivation holds more generally for any residual, instead of r_m.

From the above it follows that δQ converges to zero, therefore r_m, the maximum absolute residual, converges to zero too, i.e. all residuals converge to zero.

From Equations (2.7) and (3.2) we have

$$\mathbf{r} = -\mathbf{B}\Delta\mathbf{t}\,, \tag{3.7}$$

and solving the equation

$$\Delta \mathbf{t} = -\mathbf{B}^{-1}\mathbf{r}\,. \tag{3.8}$$

The inverse matrix $\mathbf{B}^{-1}$ exists, because $\mathbf{B}$ is positive definite. Substituting $\Delta\mathbf{t}$ in Equation (3.1) we get

$$Q = \mathbf{r}^T\mathbf{B}^{-1}\mathbf{r}\,. \tag{3.9}$$

The norm converges to zero, therefore $\Delta\mathbf{t}$ also has to converge towards zero and the Equation (2.3) is solved.

3.1. $\mathbf{B}$ *is Positive Semi-definite*

In the previous case where the matrix $\mathbf{B}$ is positive definite, the deconvolution step in CLEAN is just a way to solve the Equation (2.3), which gives a unique result. But as already mentioned, CLEAN is particularly suited for applications where the maps are the Fourier transform of the observed data. The matrix $\mathbf{B}$ can then be made to be positive semi-definite only, and the rank ($R < N$) of the matrix will be known.

The proof of convergence is first given in Forsythe and Wasow (1960).

The solution may contain any contribution z_0 which is a solution of the homogeneous equation

$$\mathbf{B}\cdot\mathbf{z}_0 = 0\,. \tag{3.10}$$

$\mathbf{z}_0$ we call a "zero-eigenvector", i.e. an eigenvector belonging to an eigenvalue 0. Its contribution to Q is zero.

Adding in Equation (3.7) $\mathbf{z}_0$ to $\Delta\mathbf{t}$ we get

$$Q = (\Delta\mathbf{t} + \mathbf{z}_0)^T\mathbf{B}(\Delta\mathbf{t} + \mathbf{z}_0) = \Delta\mathbf{t}^T\mathbf{B}\Delta\mathbf{t}\,. \tag{3.11}$$

Although $\mathbf{r}$ and Q converge to zero, $(\Delta\mathbf{t} + \mathbf{z}_0)$ does not necessarily converge towards zero.

We then have one of the infinitely many solutions, $\mathbf{t} = \mathbf{t}_0 + \mathbf{z}_0$, of the Equation (2.13). Q can now also be expressed by replacing $\mathbf{t}_0$ by $\mathbf{t}$:

$$Q = (\mathbf{t}' - \mathbf{t})^T\mathbf{B}(\mathbf{t}' - \mathbf{t})\,. \tag{3.12}$$

In the case of a uniform-weight beam Q can expressed in terms of $\mathbf{r}$ only, even though $\mathbf{B}^{-1}$ does not exist (cf. Eq. (3.9)). Multiplying Equation (3.7) by $\mathbf{B}$

$$\mathbf{Br} = \mathbf{BB}\Delta\mathbf{t} = f\mathbf{B}\Delta\mathbf{t}\,, \tag{3.13}$$

where f is the normalizing factor, we can substitute the right-hand side of above Equation into (3.1).

$$\begin{aligned} Q &= (1/f)\Delta\mathbf{t}\mathbf{Br} \\ &= (1/f)\mathbf{r}^2\,. \end{aligned} \tag{3.14}$$

In other words: in the case of a "uniform weight" beam the norm Q is just a constant factor times the sum of squares of the residuals.

Note: If the method CLEAN is tested with ad hoc numerical examples, as is done sometimes, the two criteria of convergence, i.e. that $\mathbf{B}$ has to be at least positive semi-definite and that $\mathbf{d}$ has to be in the range of $\mathbf{B}$, can easily be violated; the tests then can lead to divergence.

4. The Solution of a Subsystem of Equations ("Window" CLEAN)

In this section we will discuss the solution of a subsystem of the Equation (2.13). We assume that the N by N matrix $\mathbf{B}$ is positive semi-definite, $\det(\mathbf{B}) = 0$ and that the dirty map $\mathbf{d}$ is in the range of $\mathbf{B}$. Further, we assume that R, the *rank* of the matrix, is known (see Sect. 6.2).

Suppose the deconvolution process is started and components are found at various positions, say W different positions; we now allow no new positions to be selected and in this way we define a "*window*" of W positions. Or one can also define a priori a window of W positions, e.g. the positions within a certain area. The iteration process can be continued, thereby searching the maximum residual only within the "window" and improving the amplitudes of the W components. By this method, which we call the "*window*"-CLEAN method, one forces $N - W$ components of $\mathbf{t}$ to be exactly zero.

$$\mathbf{B}^{(W)}\mathbf{t} = {}^{(W)}\mathbf{d} - {}^{(N-W)}\mathbf{r}\,. \tag{4.1}$$

The upper index added to the vectors indicates the number of non-zero elements. In practice not all residuals have to be calculated at each iteration, once the set of W positions is defined. One can define a submatrix $\mathbf{B}_W$, consisting of the columns and rows corresponding to the non-zero positions of ${}^{(W)}\mathbf{t}$ and corresponding vectors $\mathbf{t}_W$ and $\mathbf{d}_W$ of length W. One then merely solves a subsystem of W equations:

$$\mathbf{B}_W\mathbf{t}_W = \mathbf{d}_W\,. \tag{4.2}$$

The solution $\mathbf{t}_W$ can be used to calculate the $N - W$ residuals according to Equation (4.1). The norm Q converges towards a value Q_W. If $W \geqq R$, then Q_W and all residuals ${}^{(N-W)}\mathbf{r}$ also become zero.

Even if $W < R$, then it is not always the case that, for a given window, $\mathbf{B}_W$ is positive definite. There is no simple criterion to decide in which cases it is, but special cases are discussed in Section 6.3.

5. The Ambiguity of the Deconvolution Procedure

Either the brightness distribution $\mathbf{t}_0$ or the solution $\mathbf{t}$ can contain a contribution of a zero-eigenvector. But using the basic assumption stated in the beginning that the map is essentially empty, we can put some limits on the occurrence of zero-eigenvectors.

First we notice that any zero-eigenvector has to be non-zero in at least $R + 1$ positions. This is because there are $N - R$ independent zero-eigenvectors (Cramer,

1946, p. 115), and therefore at most $N - R - 1$ positions can be made zero by a linear combination of these vectors.

We now discuss the uniqueness of the solution for the case that $Q = 0$ and that $\mathbf{t}_0$ contains less than R non-zero values. In general CLEAN is not bound to R or less non-zero values. If this number exceeds R, then the solution is not unique and can contain any zero-eigenvectors. If the solution is reached with exactly R positions, then the solution is unique *within that particular set of R positions* but it is not necessarily the correct one.

If the solution requires less than R positions, then it will in general be unique and correct. Mathematically speaking, however, there can be exceptions, when certain relationships hold between the brightness distribution $\mathbf{t}_0$ and the sampling of the visibility. In order to exclude these exceptions, one has to have *not more than* $R/2$ components. This can be proven in the following way: suppose there are two (exact) solutions, each with $\leqq R/2$ components. The difference is a vector, which has at most R non-zero positions. Since a zero-eigenvector must have at least $R + 1$ non-zero positions, the difference vector of the two solutions cannot be a zero-eigenvector, and hence the two solutions must be identical.

The uniqueness of the solution is however somewhat an academic concept, since in most practical applications of CLEAN the data are affected by noise. Q cannot be made exactly zero with less than R positions. One wants to find the most probable solution; this is a statistical rather than a mathematical concept, cf. Section 8.

Now we have established that, with some a priori knowledge of the true map, there is a unique solution. However, we must locate the non-zero values of the true map. Once the positions of these values are found, it is easy to solve the equation.

It is a particular characteristic of the method CLEAN that the evaluation of the solution and the location of the positions are intimately related.

a) Summarizing the evaluation of the solution:
with each iteration, where possibly a new position is defined, the norm

$$Q = \Delta\mathbf{t}_j^T \mathbf{B} \Delta\mathbf{t}_j \qquad (\Delta\mathbf{t}_j = \mathbf{t}_j - \mathbf{t}_0) \tag{5.1}$$

is diminished, improving the fit with respect to the true map $\mathbf{t}_0$. With the uniform weight beam defined in Section 2.3 we can interpret Q as

$$Q = \mathbf{r}^2/f, \tag{5.2}$$

thus the sum of residuals squared is reduced.

b) The location of positions:
the assumption is made that the absolute largest value of the dirty map (respectively the residuals) indicates the presence of a non-zero value. Högbom (1974) has shown that for the "uniform weight" beam, this value represents the maximum correlation between "dirty" beam and "dirty" map.

If indeed the brightness distribution is essentially empty, then the incorporation of a zero-value, a "false" position, in the solution does no harm since the final solution (assuming $Q = 0$) will make $\mathbf{t}$ at this position exactly zero. This situation can occur when the "true" map contains several sources, the secondary responses of which may add up. The CLEAN method using a loop gain $g = 1$ can then give rise to a damped oscillatory behaviour of the iteration process and to additional "false" positions. In this case a small loop gain is advisable. The loop gain can even be made infinitesimally small with finite computing effort; this will be discussed elsewhere.

Another interpretation of the procedure will be given in Section 7. In contrast to CLEAN the Gauss-Seidel method (cf. Sect. 2.4) selects all N positions and does not take into account the emptiness of the brightness distribution.

An often applied method is either to search only for the maximum (positive) residuals or to allow only positive components; this is legitimate, since Q decreases always, but does not necessarily lead to convergence for it can happen that all residuals become negative. Therefore this method, which makes use of the positiveness of the sky, can only be applied in the beginning of the phase b) to define a "window". The requirements of positiveness of the residuals and small loop gain can often be used in combination. Thereby the sum of the components (the total intensity) is steadily increased during the iteration process.

Part II. CLEAN as a Least Squares Fit

6. Method CLEAN Applied to Fourier Transformed Data

In radio astronomy the method CLEAN is often used in connection with interferometer and synthesis observations, where the maps are derived from a Fourier transform of the observed data. The discussion in this and the following sections is applicable to such data as well as to other data for which the Fourier transform is used to obtain the result, e.g. spectra from autocorrelators.

So far the method CLEAN has been described as a deconvolution procedure. Here we will show that the method is especially well suited for Fourier transformed data and that it can be looked at as a fitting procedure of sine-functions in a least square sense.

6.1. Definitions

We define the "true" map (the brightness-distribution multiplied by the primary beam pattern of the individual antennas) by its Fourier transform, the "true" visibility $V_0(u)$ (including noise). The "true" visibility is sampled at a finite number of points by interferometric observa-

tions, at values $u = u_k$, where u_k is the effective base-line. To each observation a weight $w(u_k)$ is attached.

Because the true map is real we do not have to observe at $u = -u_k$, but assign to this point the conjugate complex value of the observed visibility: $V(-u_k) = \overline{V}(u_k)$; we can also define $w(-u_k) = w(u_k)$. The number of observations R is the number of observed amplitudes and phases (counted separately) of the complex quantity. If $u = 0$, the so-called zero-spacing, then only an amplitude is measured, the phase being zero by definition. R is odd or even if the zero-spacing is or is not measured respectively.

The "dirty" map is the Fourier transform of the visibility times the weights:

$$\mathbf{d} \overset{\text{FT}}{\longleftrightarrow} (w(u_k)\cdot V_0(u_k)) = (\mathbf{w}\cdot V_0(u_k)), \tag{6.1}$$

$$d(x_n) = \sum_{k=1}^{R} w(u_k)\cdot V_0(u_k)\exp(-i2\pi x_n u_k), \quad n = 1,\ldots,N,$$

where we take by definition

$$\sum_{k=1}^{R} w(u_k) \equiv 1. \tag{6.2}$$

As before we write $d(x_n)$ as a vector $\mathbf{d}$, similarly we write $\mathbf{w}$.

The "dirty" beam has to be calculated for all differences

$$\Delta x_{nm} = x_n - x_m, \tag{6.3}$$

as the Fourier transform of the weights

$$b(\Delta x_{nm}) = \sum_{k=1}^{R} w(u_k)\exp(i2\pi\Delta x_{nm}u_k). \tag{6.4}$$

Since $w(-u_k) = w(u_k)$ the beam is symmetric and we have

$$b(-\Delta x_{nm}) = b(\Delta x_{nm}), \tag{6.5}$$

and also

$$b(0) = 1. \tag{6.6}$$

The $b(\Delta x_{nm})$ can be arranged in a matrix $\mathbf{B}$ as defined in Section 2, Equation (2.4).

Mostly the Fourier transform is performed into a regular grid of N points, therefore

$$\Delta x_{nm} = l\Delta x, \tag{6.7}$$

and the dirty beam can be written also as a vector, $\mathbf{b}$. For computational efficiency the visibilities are often also forced into a regular grid, especially when the fast Fourier transform algorithm is to be used. The observed points are either interpolated or assigned to the closest gridpoint. Assuming a correct interpolation the weights then represent a vector of length N, where R points have non-zero weight. With a grid-interval of Δu the "dirty" map becomes periodic with a period of $1/\Delta u$ and

$$\Delta x = 1/(N\Delta u). \tag{6.8}$$

This results in an extra symmetry of the beam

$$b(l\Delta x) = b((N - l + 1)\Delta x). \tag{6.9}$$

The matrix $\mathbf{B}$ becomes then circulant

$$\mathbf{B} = \begin{bmatrix} 1 & b_1 & b_2 & b_3 & \ldots & b_2 & b_1 \\ b_1 & 1 & b_1 & b_2 & \ldots & b_3 & b_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ b_1 & b_2 & & \ldots & & b_1 & 1 \end{bmatrix}, \tag{6.10}$$

with $b_n \equiv b(n\Delta x)$.

In this discrete case the Fourier transform, denoted by DFT, becomes

$$\mathbf{d} \overset{\text{DFT}}{\longleftrightarrow} (\mathbf{w}V_0),$$

$$d_n = \sum_{k=-N/2}^{N/2-1} w_k V_{0_k}\exp(-i2\pi kn/N), \tag{6.11}$$

and the inverse defined by

$$w_k V_{0_k} = (1/N)\sum_{n=1}^{N} d_n \exp(i2\pi kn/N). \tag{6.12}$$

In this discrete case the visibility can also be written as a vector, $\mathbf{V}_0$.

We define here a "function product", using the symbol $\hat{\times}$, for the simplification of the notation especially in connection with the convolution theorem. This product is defined by:

$$\mathbf{z} = \mathbf{x}\,\hat{\times}\,\mathbf{y} = (x_1y_1, x_2y_2, \ldots, x_Ny_N). \tag{6.13}$$

Summarizing the definitions, we have the following Fourier transform pairs

$$\mathbf{b} \overset{\text{DFT}}{\longleftrightarrow} \mathbf{w}, \tag{6.14}$$

$$\mathbf{d} \overset{\text{DFT}}{\longleftrightarrow} (\mathbf{w}\,\hat{\times}\,\mathbf{V}_0). \tag{6.15}$$

Taking the Fourier transform of both sides of the deconvolution Equation (2.3)

$$\mathbf{w}\,\hat{\times}\,\mathbf{V} = \mathbf{w}\,\hat{\times}\,\mathbf{V}_0, \tag{6.16}$$

where we define the Fourier transform of the components t

$$\mathbf{V} \overset{\text{DFT}}{\longleftrightarrow} \mathbf{t}. \tag{6.17}$$

One sees that $\mathbf{V}$ equals the true visibility $\mathbf{V}_0$ in those positions, where $\mathbf{w}$ is non-zero.

6.2. *Convergence*

The method CLEAN converges under three conditions (cf. Sect. 3): (i) the matrix $\mathbf{B}$ must be symmetric, (ii) $\mathbf{B}$ has to be positive definite or positive semi-definite, and (iii) the dirty map $\mathbf{d}$ has to be in the range of $\mathbf{B}$. In this section we will consider these conditions.

In Equation (6.5) the symmetry of the dirty beam has been proven, hence the matrix $\mathbf{B}$ is symmetric. The two

other conditions can be tested by applying the theory of quadratic forms, using the eigenvalues and eigenvectors of the matrix **B**. All eigenvalues are real, because **B** is symmetric. They are readily found by applying the convolution theorem of the Fourier transform. The eigenvalues λ_j are defined by the equation

$$\mathbf{B}\mathbf{z}_j = \lambda_j \mathbf{z}_j , \tag{6.18}$$

where $\mathbf{z}_j$ are the corresponding eigenvectors. Taking the Fourier transform one gets

$$N\mathbf{w} \hat{\times} \boldsymbol{\zeta}_j = \lambda_j \boldsymbol{\zeta}_j , \tag{6.19}$$

with

$$\mathbf{z}_j \overset{\text{DFT}}{\longleftrightarrow} \boldsymbol{\zeta}_j . \tag{6.20}$$

Equation (6.19) can be satisfied with

$$\boldsymbol{\lambda} = N\mathbf{w} , \tag{6.21}$$

and

$$\boldsymbol{\zeta}_j = (0, \ldots, 0, 1, 0, \ldots, 0) , \qquad j = 1, \ldots, N , \tag{6.22}$$

with the element 1 at the position j. The eigenvector $\mathbf{z}_j$ is the Fourier transform of $\boldsymbol{\zeta}_j$. One has now N independent eigenvectors, belonging to the eigenvalues $\lambda_j = Nw_j$. The eigenvalues of the matrix **B** are just a constant times the weights.

The theorems of quadratic forms allow us to distinguish the following three cases:

(i) **B** is positive definite if all weights $w_k > 0$.
(ii) **B** is positive semi-definite if $w_k \geqq 0$, with at least one $w_k = 0$.
(iii) **B** is indefinite if there is at least one $w_k < 0$.

The method CLEAN can therefore be used for the cases (i) and (ii). It is case (ii), however, for which the method CLEAN is primarily designed, since, when the observations are not complete (i.e. some weights are zero) the "dirty" beam has sidelobes and grating-responses. The rank R of the matrix is the number of non-zero weights, i.e. the number of observations as defined in Section 6.1. The determinant $|\mathbf{B}|$ is the product of the eigenvalues, it is zero for case (ii) ($R < N$) and the system of linear Equations (2.3) has no unique solution.

The eigenvector $\mathbf{z}_j$ is the Fourier transform of $\boldsymbol{\zeta}_j$

$$\mathbf{z}_j = (1, \exp(i2\pi j/N), \exp(i2\pi 2j/N, \ldots) . \tag{6.23}$$

Combining pairs of eigenvectors $\mathbf{z}_j$ and $\mathbf{z}_{N-j}$ by adding and subtracting respectively, one gets a new set of eigenvectors

$$\mathbf{z}'_j = (1, \cos(2\pi j/N), \cos(2\pi 2j/N), \ldots) ,$$
$$\mathbf{z}'_{N-j} = (0, \sin(2\pi j/N), \sin(2\pi 2j/N), \ldots) . \tag{6.24}$$

This means that to each weight $w_j = w_{N-j}$ there belongs a pair of eigenvectors which can be combined to give a sine-wave of given frequency of any phase.

The dirty map **d** is the Fourier transform of the visibility times the weights, therefore it is a sum of sine-waves belonging to the weights of non-zero value. This means that **d** is indeed in the range of the matrix **B**; this is in fact the convolution theorem.

In the more general case of the direct Fourier transform the matrix **B** is an infinite "Toeplitz"-matrix and there is a similar relation between the eigenvalues and the Fourier transform (Toeplitz, 1911).

6.3. Positive Definiteness of the Subsystem of Equations

It is important that in the "window" CLEAN the submatrix $\mathbf{B}_W$ is positive definite, as mentioned in Section 4.

If the sampling function **w** is non-zero at arbitrary grid points and the "window" CLEAN is applied at an arbitrary set of $W(\leqq R)$ points, there is no simple rule which allows us to predict if $\mathbf{B}_W$ is positive definite or not. However, in two complementary special cases it is possible to prove the positive definiteness of $\mathbf{B}_W$ (Buurema, 1974, private communication).

(i) **w** is arbitrary. For any set of $W(\leqq R)$ points within an interval of R adjacent grid points, $\mathbf{B}_W$ is positive definite. One can easily see that if the "dirty" beam has grating responses of 100%, then this interval is the distance between the main response of the beam and the first grating response.

(ii) **w** is non-zero at regularly spaced grid points with the spacing equal to ξ times the grid distance. If $\xi = 1$ then all sets of W "cleaned" points give a positive definite matrix $\mathbf{B}_W$. If $\xi = 2, 3, \ldots$, this gives rise to grating responses of amplitude 1 in the "dirty" beam. Any set of W points is legitimate, unless there are pairs of points which have a mutual distance equal to the distance of the grating response, i.e. N/ξ times the grid interval.

7. The Deconvolution Step as a Least Squares Fit

In this section we will look what the deconvolution step does in the domain of the Fourier transform. We will see that the norm Q introduced in Section 1 has a simple meaning; it is the weighted sum of squares of the Fourier transform of $\mathbf{t}' - \mathbf{t}_0$. These relations are derived in Section 7.1. This property of the norm Q can be used to show that the method CLEAN performs a least squares fit in the Fourier transform domain (Section 7.2), using a sybsystem of the system Equation (4.2). A necessary condition for the validity of the concept of a least squares fit is that the subsystem must have a unique solution, i.e. the submatrix $\mathbf{B}_W$ has to be positive definite (not only positive semi-definite!).

7.1. The Norm Q and the Visibility

If one writes the norm Q, Equation (3.12), in the form using relation (3.7)

$$Q = (\mathbf{t}' - \mathbf{t})^T \mathbf{r} , \tag{7.1}$$

and applies Parceval's theorem one can replace $(\mathbf{t}' - \mathbf{t})$ and $\mathbf{r}$ by their Fourier transforms

$$(\mathbf{t}' - \mathbf{t}) \overset{\text{DFT}}{\longleftrightarrow} \mathbf{V}' - \mathbf{V} = \Delta\mathbf{V}\,, \tag{7.2}$$

$$-\mathbf{r} = \mathbf{B}\mathbf{t}' - \mathbf{d} \overset{\text{DFT}}{\longleftrightarrow} (\mathbf{w} \hat{\times} \mathbf{V}' - \mathbf{w} \hat{\times} \mathbf{V}) = \mathbf{w} \hat{\times} \Delta\mathbf{V}\,, \tag{7.3}$$

and one gets

$$Q = \Delta\mathbf{V}(\mathbf{w}\Delta \hat{\times} \mathbf{V}) = \sum \mathbf{w}_k \Delta\mathbf{V}_k^2\,. \tag{7.4}$$

Equation (4.4) shows that Q is the weighted sum of squares of the differences $\Delta\mathbf{V}$ in the Fourier transform domain. These are the differences between the Fourier transform of the approximation $\mathbf{t}'$ of the CLEAN process and the Fourier transform of the solution $\mathbf{t}$, which equals the observed visibility $\mathbf{V}_0$ at those positions, where $\mathbf{w}$ is non-zero. Especially if $Q = 0$, then $\Delta\mathbf{V}$ is zero at the non-zero positions of $\mathbf{w}$.

7.2. Minimalization of the Sum of Square Deviations

In each step of the iteration process the norm Q is reduced. This means that the set of components $\mathbf{t}'$ is improved in each step, such that their Fourier transform approximates closer and closer the observed visibility function. If at a given number of positions, W, no new positions are allowed and the iteration process continued with W positions ("window" CLEAN), then the norm approaches $Q_W \geqq 0$, the residuals within the window becoming zero. If we assume the W by W matrix $\mathbf{B}_W$ is positive definite, then any change of the components would result in finite residuals within the window and the norm Q would increase; therefore Q_W is a minimum.

One can show that

$$\Delta Q_W = Q'_W - Q_W = (\mathbf{t}'_W - \mathbf{t}_W)^T \mathbf{B}_W (\mathbf{t}'_W - \mathbf{t}_W)\,. \tag{7.5}$$

Since $\mathbf{B}_W$ is positive definite, ΔQ_W is indeed always positive unless $\mathbf{t}'_W = \mathbf{t}_W$. Equation (7.5) is very similar to the general definition of Q, cf. Equation (3.1); the difference is that here (i) the summation takes place only over the W positions of the "window" and (ii) $\mathbf{t}_W$ is not a solution of the original system of equations, $\mathbf{t}$, but only a solution of the subsystem of equations, Equation (4.2).

The Fourier transform of ${}^{(W)}\mathbf{t}$ is the sum of sine-functions belonging to each of the components of ${}^{(W)}\mathbf{t}$. Because the norm Q_W is the weighted sum of squares of the difference of the superposition of sine-functions with the observed visibility $\mathbf{V}_0$, we can interpret CLEAN as a fit of sine-functions to the observed visibility in a least squares sense, where the components play the role of parameters, characterizing the sine-functions. If we take $W = R$ then we determine as many unknowns as there are observations. This solves not only the subsystem, but also the original system and therefore $\mathbf{r}$ becomes zero in all N points.

It may appear arbitrary that the CLEAN method fits sine-functions to the observed visibility. But if the real map consists only of pointlike sources, then the sine-function is the most appropriate choice. If a pointlike source lies between two grid points, then the observed visibility can be approximately fitted by two sine functions with slightly different periodicities corresponding to the two adjacent grid points. The smaller the grid interval, the better the fit. The deconvolution step will select (using a small loop gain) the two closest gridpoints; these two components will have the largest amplitudes.

We can now interpret the characteristic CLEAN procedure of searching the maximum in the dirty (or residual) map in a different way: if we want to fit a sine-wave to the observed visibility $V_0(u)$, we must know its period. In order to get the period one can perform a Fourier transform to obtain the spectrum and look for its maximum; this is just the maximum of the "dirty" map. Therefore, CLEAN can be interpreted as an iterative harmonic analysis.

8. Error Estimates

8.1. Error Analysis

If the method CLEAN is applied to data which are affected by noise, then the components $\mathbf{t}$ will also be affected by noise. If the experiment were to be repeated with different noise, then CLEAN would possibly find a different set of points. Under such circumstances it would be hard to make error estimates of the components. What we can do, however, is to make error estimates for a given set of points. In other words the sub-system of equations (Eq. (4.2)) would be solved for observations with different noise. The error estimates can be used in the standard tests of significance.

We assume Gaussian noise in the observed complex quantity $\mathbf{V}$, being independent in the real and imaginary part, and also being independent from point to point. But the noise enters the "dirty" map through the Fourier transform and therefore the noise is not independent from point to point. The correlation of the noise can be described by a matrix $\mathbf{E}$; this matrix can be constructed from a vector $\mathbf{e}$ in a similar way as $\mathbf{B}$ from the "dirty" beam $\mathbf{b}$. $\mathbf{e}$ is defined by

$$\mathbf{e} \overset{\text{DFT}}{\longleftrightarrow} \mathbf{w}_0 \hat{\times} \mathbf{p}^2\,, \tag{8.1}$$

where $\mathbf{w}_0$ are the natural weights, i.e. the noise in the observations is $\sigma \mathbf{w}_{0_i}^{-1/2}$ and $\mathbf{p}$ is a taper function. Especially, if $\mathbf{p} = 1$ then

$$\mathbf{w} = \mathbf{w}_0 \hat{\times} \mathbf{p} = \mathbf{w}_0\,, \tag{8.2}$$

and we have

$$\mathbf{e} = \mathbf{b}, \qquad \mathbf{E} = \mathbf{B} \quad \text{and} \quad \mathbf{E}_W = \mathbf{B}_W\,. \tag{8.3}$$

In order to calculate the errors of the components, $\mathbf{t}$ or linear combinations of those, etc., we can use the standard

theory on the propagation of errors (cf. W. E. Eming, Statistical Adjustment of Data, 1964). Below we will use $p = 1$ throughout.

Writing

$$\mathbf{t}_W = \mathbf{B}_W^{-1}\mathbf{d}_W\,, \tag{8.4}$$

the estimated error $\mu^2(\mathbf{t}_W)_j$ of the jth component becomes then

$$\mu^2(\mathbf{t}_W)_j = \sigma^2 \sum_{k,l} \mathbf{E}_{W_k,l}\mathbf{B}_{Wjk}^{-1}\mathbf{B}_{Wlj}^{-1} = \sigma^2 \sum_{k,l} \mathbf{B}_{Wkl}\mathbf{B}_{Wjk}^{-1}\mathbf{B}_{Wlj}^{-1}$$
$$= \sigma^2\mathbf{B}_{Wjj}^{-1}\,. \tag{8.5}$$

The error of the total intensity, $F_0 = \sum_j \mathbf{t}_j$, is

$$\mu^2(F_0) = \sigma^2 \sum_{l,k} \mathbf{B}_{kl}^{-1}\,. \tag{8.6}$$

The error of the average position (the first moment) and higher moments will be discussed elsewhere.

Another application of the theory is to estimate the error of the fitted visibility function at any point. For this purpose we define the quantity

$$U \equiv V_c + V_s\,, \tag{8.7}$$

where V_c and V_s are the real and imaginary parts, respectively, of the visibility. One can show that the error in U becomes, at a given point β:

$$\mu^2(U)_p = \sigma^2 \sum_{k,l} \cos|2\pi(m_k - m_l)\beta/N|\mathbf{B}_{k,l}^{-1}\,, \tag{8.8}$$

where m_k is the position of the component t_{W_k} in the map (i.e. the component number in the vector $\mathbf{t}$). This formula allows one to estimate the error in "holes" or of the extrapolation of the visibility. A simple example is to calculate the error for $\beta = 0$, which gives the error of the fitted visibility at the "zero"-spacing; this gives the same result as Equation 8.6, which is correct because the "zero"-spacing gives the total itensity.

The noise σ of the observations can be determined from the norm Q_W. If one has fitted $W < R$ components and the residuals outside the "window" are only random (i.e. the deviations from the measured visibilities V_{0_k} are random) then statistical theory predicts (cf. van der Waerden, 1959)

$$\sigma^2 = Q_W/(R - W)\,. \tag{8.9}$$

Practical applications of the error analysis to 2-dimensional real observations will be discussed elsewhere.

8.2. *Numerical Examples*

The examples given here are simple artificial noise-free cases, which illustrate the relation between the visibility and the set of components $\mathbf{t}$. Also the error estimates of the various quantities are shown.

Two sources have intensities 2 and 1 respectively and are shown as arrows in Figure 1. The true visibility ($N = 34$) is represented in Figure 3 as heavy lines. Three

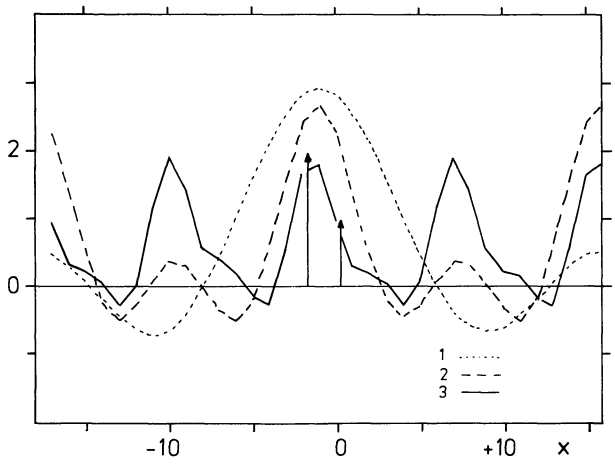


Fig. 1. Example of the brightness distribution of two close sources (arrows) and the convolution with 3 different "dirty" beams. The "dirty" beams are based on the weights shown in Figure 3c

different "dirty" beams are used, which are based on $R = 5$ observations, distributed regularly on baselines $u_m = 2$, 4 and 8 times Δu, respectively, as shown at the bottom of Figure 3. The resulting "dirty" maps are plotted in Figure 1. The results of the deconvolution step with $W = 4$ components are given in Figure 2. In order to make comparison easier, the solution has been forced into four adjacent points. The estimated visibilities (i.e. the discrete FT of the components) can be seen in Figure 3. For this particularly simple source structure

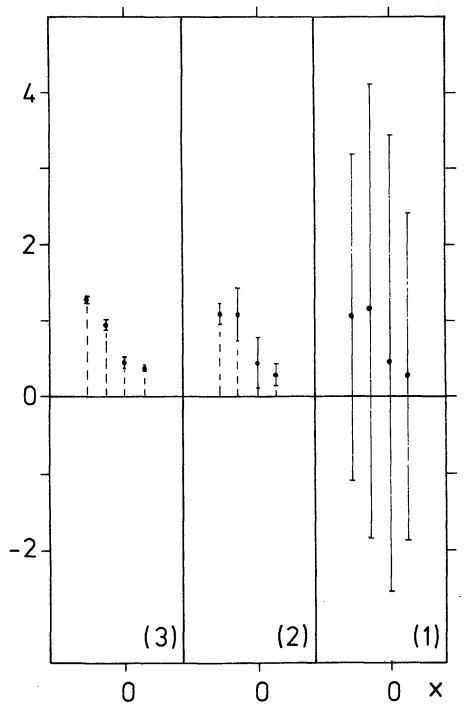
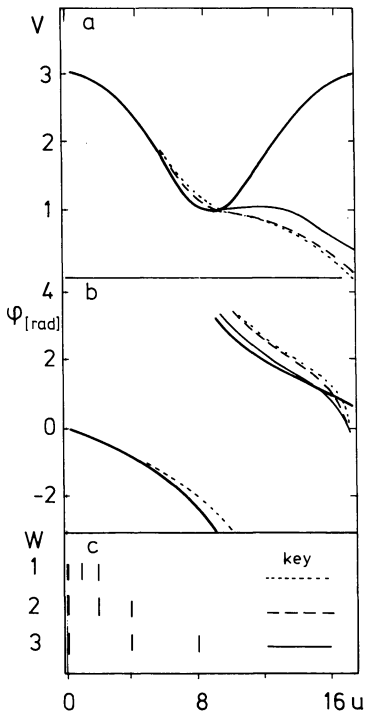


Fig. 2. Results of the deconvolution step based on the "dirty" maps of Figure 1. The errors bars indicate the errors based on noise of 0.01. Note the very large errors for case 1



Figs. 3a and b. The complex Fourier transform of the components shown in Figure 2. One sees how far unknown information could be extrapolated, assuming a simple source structure. **c:** The weights, i.e. the sampling function

even the case with the smallest baseline $u_m = 2$. Δu gives a rather good fit, which can be extrapolated to at least twice the used baseline. The error estimates of the components are given as error bars in Figure 2; the errors of the visibilities are plotted separately in Figure 4. The errors of the sum of the components are 1.6, 1.6, 1.7 times σ for $u_m = 2$, 4 and 8 respectively. All error estimates are based on a value $\sigma = 0.01$.

The errors show clearly the well known fact, that one needs an extremely high signal-to-noise ratio in order to extrapolate the visibility (i.e. to increase the resolution) even moderately.

In a second example, the two point sources are on the grating responses of each other. Their amplitudes are again 1 and 2, respectively. The grating response has an amplitude of 34%, see Figure 5b. As in the previous example the deconvolution components and the associated errors are displayed. The average amplitudes, positions and errors of the two groups of components are shown in Figure 5d. The errors of the fluxes of the individual sources and the total flux are 1.45, 1.55, and 2.20 times σ respectively. One sees that the error in flux (2.20σ) is only slightly higher than the combined error of the two sources $(1.45^2 + 1.55^2)^{1/2}\sigma = 2.12\sigma$, in spite of the strong

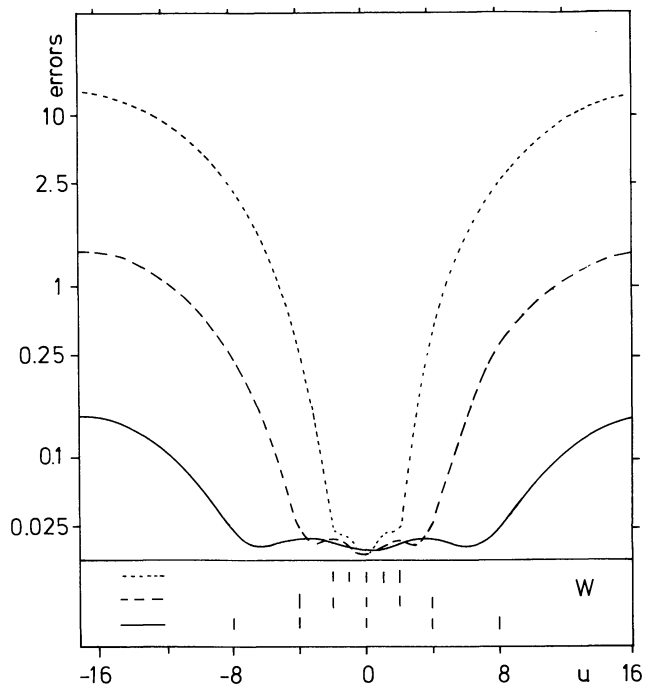


Fig. 4. The errors in the visibility (more correctly in $U = V_c + V_s$, see text) as function of the spacing u for the 3 cases. Note the very steep increase in errors for the extrapolated visibilities

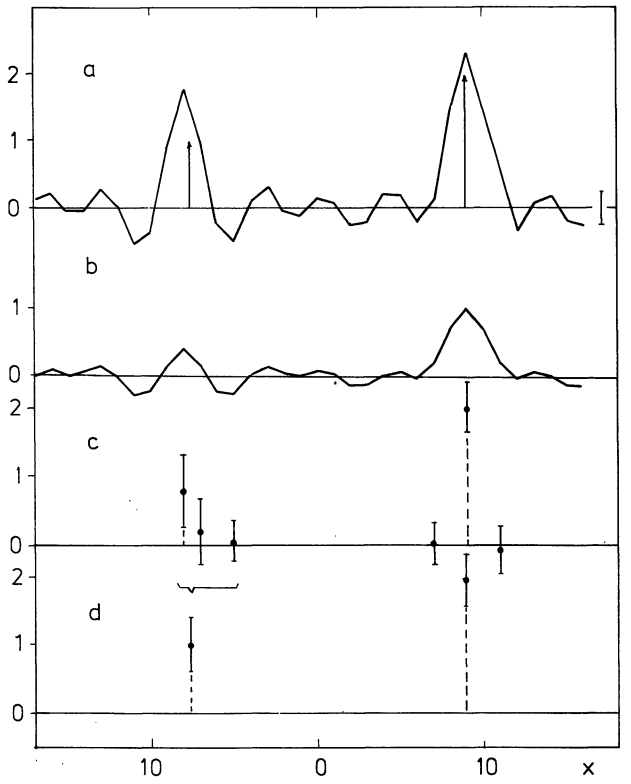


Fig. 5. a True brightness (arrows) and dirty map. **b** Dirty beam. **c** Components found in the deconvolution step with error bars, the errors being based on noise of amplitude as indicated by the bar in Fig. 5a. The noise is however not included in the dirty map **d** Combined errors for each source

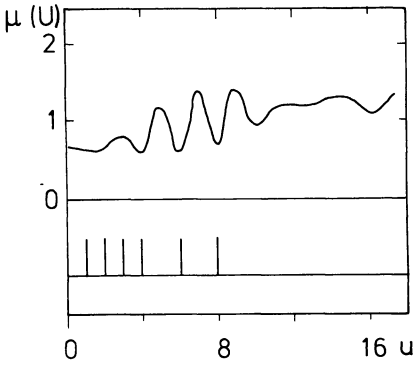


Fig. 6. a Error of the visibility (more correctly of $U = V_c + V_s$, see text). **b** Weights (sampling function)

grating response interfering directly. The same result is also obtained if the amplitudes of the sources are equal. The small additional uncertainty due to the fact that the two sources are on a grating response of each other manifests itself in the "ripple" of the error in the visibility (Fig. 6); a similar "ripple" is not seen in the previous example (Fig. 4).

This "clean" solution of $W = 6$ and $R = 13$ does not fall under the general classes, for which the positive definiteness can be proved; this particular example has also a positive definite matrix $\mathbf{B}_W$.

9. The "CLEAN" Map

9.1. Convolving the Components with a "clean" Beam

As mentioned in Section 2, a "clean" map is constructed by convolving the components found in the deconvolution step, $\mathbf{t}_W$, with a hypothetical beam, the "clean" beam, $\mathbf{h}$. The effect of this is to weight down the high frequency terms of the estimated visibility, which are the most uncertain ones (cf. Sect. 8).

The "clean" beam $\mathbf{h}$ is here defined as the Fourier transform of a normalized set of weights $\mathbf{w}_h$

$$\mathbf{h} \overset{\text{DFT}}{\longleftrightarrow} \mathbf{w}_h\,, \qquad \sum \mathbf{w}_{h_i} = 1\,. \tag{9.1}$$

The normalization ensures that the "clean" beam h has the same maximum amplitude of unity as the "dirty" beam b which it replaces. The "clean" map is produced by convolving the components with the "clean" beam ($\mathbf{H}$ in matrix notation). For simplicity we shall in this section drop the "window" sub- and superscripts W and write:

$$\mathbf{c} = \mathbf{H}\mathbf{t}\,, \tag{9.2}$$

$$\mathbf{c} \overset{\text{DFT}}{\longleftrightarrow} \mathbf{w}_h \hat{\times} \mathbf{V}\,. \tag{9.3}$$

With a window of $W \geqq R$ positions, the norm Q can be brought to zero. Then $\mathbf{c}$ will be in agreement with the (weighted) true transform $\mathbf{V}_0$ at the positions u where this has been measured. In between these positions the extrapolated values are used.

How does one choose a suitable "clean" beam? It should obviously have a satisfactory main lobe and as small sidelobes as possible. Its Fourier transform $\mathbf{w}_h$ must not stretch far outside the largest measured value $u_{\max}$ since this would lead to unacceptably large errors in the extrapolated visibilities (Sect. 8). The "clean" beam is often chosen as the Fourier transform of a Gaussian which is truncated just outside $u_{\max}$. Such a beam satisfies the above criteria. Truncated versions of the "dirty" beam however, are unsuitable, as their transforms will normally stretch far outside $u_{\max}$.

In certain circumstances it may be useful to define a "clean" beam in analogy with a Wiener filter by weighting the extrapolated visibilities according to their uncertainty (cf. Eq. (8.8))

$$\mathbf{h} \overset{\text{DFT}}{\longleftrightarrow} \begin{cases} w_k & \text{if } w_k \neq 0 \\ 1/\mu^2(U_k) & \text{if } w_k = 0 \end{cases} \tag{9.4}$$

The above discussion shows that one cannot in practice improve the resolution significantly by choosing a very narrow "clean" beam. Such a beam would have a transform stretching far outside the measured range of u. A significant increase in the resolution can be achieved in theory, but only under very special circumstances and with noise-free data.

9.2. Adding the Residuals

For practical reasons the deconvolution process is usually stopped before all residuals have disappeared. It is then common practice to add the residuals $\mathbf{r}'$ to the map obtained by convolving the components $\mathbf{t}'$ extracted so far with the "clean" beam $\mathbf{h}$. The purpose is to keep significant features that may still be left in $\mathbf{r}'$. The "clean" map $\mathbf{c}'$ is then given by:

$$\begin{aligned} \mathbf{c}' &= \mathbf{H}\mathbf{t}' + \mathbf{r}' \\ &= \mathbf{H}\mathbf{t}' + \mathbf{B}(\mathbf{t} - \mathbf{t}') \\ &= \mathbf{H}\mathbf{t} + (\mathbf{B} - \mathbf{H})(\mathbf{t} - \mathbf{t}')\,, \end{aligned} \tag{9.5}$$

which shows that, as expected, $\mathbf{c}'$ differs from the zero residual "clean" map $\mathbf{c}$ (Eq. (9.2)) by the sidelobes of those components which were missed because the iterations were interrupted. Taking the Fourier transform of the right-hand side we get:

$$\mathbf{c}' \overset{\text{DFT}}{\longleftrightarrow} \mathbf{w}_h \hat{\times} \mathbf{V} + (\mathbf{w} - \mathbf{w}_h) \hat{\times} (\mathbf{V} - \mathbf{V}')\,. \tag{9.6}$$

At the positions u_i of the measured Fourier components we have $w_i \neq 0$ and $V_i = V_{0_i}$. In general, since there are finite residuals $(\mathbf{t} - \mathbf{t}')$, we also have

$$(V_i - V_i') = (V_{0_i} - V_i') \neq 0\,, \tag{9.7}$$

and the second term in Equation (9.5) disappears for the measured components only if $w_{h_i} = w_i$. In practice one can arrange the weights so that $w_{h_i} = q w_i$, but the

constant q will be determined by the normalizing Equations (6.2) and (9.1). We see that $q \ll 1$ if, as is often the case, a limited number of measurements are distributed over a fine grid in u: there are then many more non-zero positions in $\mathbf{w}_h$ than there are in $\mathbf{w}$. Since q will normally be $\neq 1$, one should not expect the "clean" map $\mathbf{c}'$ to agree *exactly* with the measured (and weighted) Fourier components.

The map *can* be brought to exact agreement with the weighted measurements if the residuals are not added directly but in the form $q\mathbf{r}'$, i.e. scaled by the factor q. This, however, means that one changes the amplitude scale of the residuals, and the original purpose of the addition is lost.

Acknowledgements. I wish to thank many colleagues for many stimulating discussions and critical remarks regarding the various stages of the manuscript, a.o. Dr. R. D. Ekers (Groningen), Drs. J. P. Hamaker (Dwingeloo), Dr. J.-P. Marchand (Denver) and especially Dr. J. A. Högbom (Stockholm). I would like also to acknowledge the important contribution of Dr. H. J. Buurema (Groningen) for Section 6.3. The English has been much improved by Mrs. J. A. Ekers. The numerical experiments were done on the Cyber 74-16 computer of the Computing Centre of the University of Groningen.

Appendix

If the 2-dimensional field is a square regular grid of P^2 grid points, then the 2-dimensional "dirty" beam is a P^2 by P^2-matrix

$$\left\{\begin{array}{ccccccc} & & \cdots & & & & \\ \cdots & d_{-2} & d_{-1} & d_0 & d_1 & d_2 & \\ & c_{-2} & c_{-1} & c_0 & c_1 & c_2 & \\ & & & & & & \cdots \\ \cdots & b_{-2} & b_{-1} & 1 & b_1 & b_2 & \\ & a_{-2} & a_{-1} & a_0 & a_1 & a_2 & \\ & & & \cdots & & & \end{array}\right\}, \tag{A.1}$$

which can be put in a vector of length P^4

$$\ldots c_{-2}c_{-1}c_0c_1c_2\ldots b_{-2}b_{-1}1b_1b_2\ldots a_{-2}a_{-1}\ldots \tag{A.2}$$

In our application in synthesis observations the beam is derived from a 2-dimensional Fourier transform, resulting in the symmetries:

$$b_{-k} = b_k; \qquad a_{-k} = c_k, \quad \text{etc.} \tag{A.3}$$

Therefore the vector will be symmetric

$$\ldots c_{-2}c_{-1}c_0\, c_1\, c_2\ldots b_{-2}b_{-1}\, 1\, b_1\, b_2\ldots c_2c_1c_0c_{-1}c_{-2}\ldots \tag{A.4}$$

As in the 1-dimensional case, the convolution can be expressed in a system of P^4 equations. This system can be written in a symbolic form as follows:

$$\{\mathbf{B}_2\}\mathbf{t} = \left\{\begin{array}{ccccc} (B) & (C) & (D) & (D^-) & (C^-) \\ (C^-) & (B) & (C) & (D) & (D^-) \\ (D^-) & (C^-) & (B) & (C) & (D) \\ (D) & (D^-) & (C^-) & (B) & (C) \\ (C) & (D) & (D^-) & (C^-) & (B) \end{array}\right\} \cdot \mathbf{t} = \mathbf{d}\,, \tag{A.5}$$

where the submatrices (B), (C) have $(P)^2$ elements, defined as in the 1-dimensional case, Equation (2.2). The submatrices (C^-), (D^-) are matrices constructed analogously, but with the "mirror" images of the vectors instead, $\mathbf{c}'$, $\mathbf{d}'\ldots$, namely

$$\mathbf{c}' = (\ldots c_2c_1c_0c_{-1}c_{-2}\ldots)\,. \tag{A.6}$$

One sees easily that the matrix $\{\mathbf{B}_2\}$ is symmetric. In order that the method CLEAN works, $\{\mathbf{B}_2\}$ has to be a symmetric, positive semi-definite matrix. As we have seen the first condition is fulfilled, the second has to be proven. Note that the matrix $\{\mathbf{B}_2\}$ is not circulant any more, but has some higher order structure. To prove the non-negative definiteness of the matrix we use the norm Q (Eq. (3.1)), which also in the 2-dimensional case can be interpreted in terms of the visibility,

$$Q = \sum_{i,j}^{P,P} w_{ij}\Delta V_{ij}^2 \quad \text{with} \quad \sum w_{ij} = 1\,. \tag{A.7}$$

If all weights $w_{ij} \geqq 0$, then $Q \geqq 0$, hence $\{\mathbf{B}_2\}$ is positive semi-definite. Completely analogously to the 1-dimensional case, to each of the weights $w_{ij} = 0$ there belong eigenvalues $\lambda_{ij} = 0$ and eigenvectors which are 2-dimensional sine-waves. Also the error analysis is extendable; the sub-matrix $\{\mathbf{B}_{2w}\}$ must be inverted to derive the errors.

References

Ables, J. G.: 1974, *Astron. Astrophys. Suppl.* **15**, 383
Cramer, A.: 1946, Mathematical Methods of Statistics, Princeton
Denning, W. F.: 1964, Statistical Adjustment of Data, Dover, New York
Forsythe, G. E., Wasow, W. R.: 1960, Finite Difference Methods for Partial Differential Equations, Wiley, New York
Högbom, J. A.: 1974, *Astron. Astrophys. Suppl.* **15**, 417
Rogstad, D. H., Shostak, G. S.: 1971, *Astron. Astrophys.* **13**, 99
Schwarz, U. J., Cole, D. J., Morris, D.: 1973, *Austr. J. Phys.* **26**, 661
Southwell, R. V.: 1935, *Proc. Roy. Soc. A*, **151**, 56
Temple, G.: 1938, *Proc. Roy. Soc. A*, **169**, 476
Toeplitz, O.: 1911, *Math. Ann.* **70**, 351
Waerden, B. L. van der: 1959, Mathematical Statistics, Springer, Berlin-Heidelberg-New York