Astron. Astrophys. 89, 132–139 (1980)

ASTRONOMY AND ASTROPHYSICS

A New Method of Deconvolution and its Application to Lunar Occultations

C. R. Subrahmanya

Radio Astronomy Group, Tata Institute of Fundamental Research, Homi Bhabha Road, Bombay-400005, India

Received September 4, 1978; accepted January 18, 1980

Summary. A new method of deconvolution is described which uses our prior knowledge about the solution to derive some of the information obscured in the data because of the smoothing nature of convolution and the presence of noise. It uses a regularised least-squares criterion of agreement with the data, according to which the computed solution will lead to a minimum variance of noise and also be smooth in the sense of minimum variance of its second-differences. In addition, the present Optimum Deconvolution Method (ODM) also constrains this solution to satisfy our prior knowledge about it by using a combination of a new algorithm for incorporating bounds on the solution like positivity, and the Lagrange multiplier method for equality-constraints. The new algorithm is a rapidly converging sequence of iterations for minimising a weighted sum of squares of the deviations of the solution from the specified bounds.

For the sake of illustration, ODM is compared with the conventional method of Scheuer for deconvolving the lunar occultation data to derive the brightness distribution of a radio source. The required occultation data have been obtained both from computer simulations and from the observations of occultations with the Ooty radio telescope. A comparison of the restorations using the two methods indicates that a) ODM can be effectively applied even in very noisy situations; b) it leads to a superresolution, implying an improvement in resolution by about a factor of two over the conventional method; and c) ODM provides a "clean" output leaving all the effects of noise to the residuals. A practical procedure has also been discussed for obtaining the effective resolution and restoring errors from an analysis of the residuals, particularly their variance and power spectrum.

Key words: deconvolution – image reconstruction – lunar occultations

1. Introduction

Many measuring systems used to observe an object $q_{\rm obj}(x)$ respond to another function $r(x')$ which is a convolution of $q_{\rm obj}$ with a function $p(x)$ characteristic of the instrument. This is true of any linear space-invariant system like an antenna or any diffracting system. Deconvolution is the process of recovering a solution $q(x)$ such that

$$r(x') = p * q_{\rm obj} + \text{noise} = p * q + \varepsilon(x')$$

with the residuals $\varepsilon(x')$ statistically resembling noise. The asterisk (*) has been used above to denote convolution. In particular, we refer to situations where the observations can be supplemented by *a priori* information regarding the solution. For instance, if it represents the brightness distribution of a radio source, it must be positive everywhere.

The special importance attached to prior knowledge is because even a noise-free deconvolution is an ill-posed problem with no unique or stable solution. This can be seen from the fact that $p(x)$ often has vanishing or negligible Fourier components at certain frequencies leading to a complete loss of information on the corresponding Fourier components of the solution. For instance, the finiteness of an antenna results into a critical frequency beyond which the Fourier transform of $p(x)$ vanishes identically. It is meaningless to seek an exact solution to an ill-posed problem. In order to seek a meaningful approximation, one solves a "regularised" problem which includes a specification of a smooth transition of the Fourier components to the irrecoverable region. The conventional linear filter approach implies an *ad hoc* specification of the irrecoverable Fourier components, e.g., by assuming them to be zero as in the "principal solution" (Bracewell and Roberts, 1954). The various possible schemes have been discussed by Tikhonov and Arsenin (1977). However, the solution is often still unsatisfactory by violating our prior knowledge about it, e.g., by including negative values (sidelobes) when it is required to be positive throughout.

Such unrealistic solutions have been avoided in a new method – "Optimum Deconvolution Method" (ODM) – which constrains the solution suitably to ensure a least-squares agreement with observations as well as being consistent with all our prior knowledge about it (Subrahmanya, 1975, 1977). The aim of this paper is to describe this method and its application to the deconvolution of Fresnel-diffraction curves obtained in the lunar occultations of radio sources.

A fundamental specification of ODM is to seek a regularised least-squares solution (RLS) which simultaneously minimises the variance of residuals (least-squares) as well as the variance of the second-differences of the solution (regularization). This scheme of regularization is flexible enough to accept prior knowledge as externally imposed constraints on the solution. These could be either equalities (e.g., total intensity being known and specified) or inequalities [e.g., $q(x) \geqq 0$]. They are handled by using a combination of Lagrange multiplier method for equalities and a new fast-converging iterative algorithm for the inequalities.

Send offprint requests to: C.R. Subrahmanya, TIFR Centre, Indian Institute of Science Campus, Bangalore – 560012, India

A detailed description of the method is given in Sects. 2.2–2.4, and its application to lunar occultations is described in Sect. 3.1. In order to compare ODM with a classical method, a large number of noise-contaminated lunar occultation profiles have been restored by using both ODM and a classical method due to Scheuer (1962). It is inferred from these comparisons that ODM can lead to superresolution and also a "clean" output enabling more objective interpretation than the classical method.

2. Description of the Method

The desired solution $q(x)$ is obtained as a set of n quantities $q_i = q(x_i)$ which are its values at n successive points $x_1, x_2, \ldots, x_n$. It is assumed that (x_1, x_n) is a given range of x, e.g., the field of view of the instrument outside which the object does not contribute significantly to the observations. Thus, we can assume that $q(x \leqq x_1) = q(x \geqq x_n) = 0$. For simplicity, the expressions given below will correspond to a uniform spacing of x_i at an interval Δx. The observational data are denoted by $r_m = r(x'_m)$, $m = 1, 2, \ldots, n_d$. These measured values include an unknown contribution due to noise whose statistical properties like the mean and variance are generally known and must be satisfied by the residuals $\varepsilon_m = \varepsilon(x'_m)$ which can be computed from an acceptable solution. Thus the observational data are given by:

$$r_m = \int p(x'_m - x) q(x) dx + \varepsilon_m = \sum_{i=1}^{n} w_i p_{mi} q_i + \varepsilon_m \,, \tag{1}$$

where $p_{mi} = p(x'_m - x_i)$ and w_i depend on the formula used for numerical integration (e.g. Dahlquist and Bjorck, 1974, Sect. 7.4). For instance, for a uniform sampling as stated above, the frequently employed trapezoidal rule gives $w_i = \Delta x$ for all i.

2.1. *Regularised Least-squares Solution*

Using the second differences $\Delta^2 q_i = q_{i+1} - 2q_i + q_{i-1}$ for defining a regularisation criterion, we define an RLS as the one minimising

$$\varrho(q) = \sigma_r^2 + \gamma \sigma_s^2 = \sum_{m=1}^{n_d} \varepsilon_m^2 / n_d + \frac{\gamma}{n-2} \sum_{i=2}^{n} (\Delta^2 q_i)^2 \,,$$

where γ is an empirical parameter. $\varrho(q)$ is a weighted sum of the variances of residuals and second-differences of the solution. This will simultaneously satisfy the least-squares and smoothness (regularisation) requirements. This scheme of regularisation was suggested independently by Phillips (1962) and Tikhonov (1963) and their methods were subsequently developed and extended by Twomey (1965) and others (e.g. Tikhonov and Arsenin, 1977). The parameter γ controls the degree of smoothness imposed on the solution. The proper value has to minimise the ripples in the solution as much as possible without allowing σ_r^2 to depart from its minimum value. Usually, σ_r^2 has a broad minimum and the exact choice of γ is not critical to within about a factor of 5. It is a slowly decreasing function of signal-to-noise for any given problem.

The desired RLS is simply obtained by solving the set of "normal equations" defined by

$$\frac{\partial \varrho}{\partial q_i} = \frac{\partial}{\partial q_i} \left[\sum_m \left(r_m - \sum_i w_i p_{mi} q_i \right)^2 \Big/ n_d + \gamma \sum_i (q_{i+1} - 2q_i + q_{i-1})^2 / (n-2) \right] = 0$$

which can be written in the form

$$\sum_{j=1}^{n} (C_{ij} + \gamma H_{ij}) q_j = R_i \,, \quad i = 1, 2, \ldots, n \,, \tag{2}$$

where

$$C_{ij} = \sum_m w_i p_{mi} w_j p_{mj} / n_d \,; \quad R_i = \sum_m w_i p_{mi} r_m / n_d \tag{3}$$

and (H_{ij}), the term obtained by differentiating the variance of the second-differences, has the form identical to that given by Twomey (1963), i.e.,

$$(H_{ij}) = \frac{1}{n-2} \begin{pmatrix} 1 & -2 & 1 & 0 & 0 & 0 & 0 & . & . & . \\ -2 & 5 & -4 & 1 & 0 & 0 & 0 & . & . & . \\ 1 & -4 & 6 & -4 & 1 & 0 & 0 & . & . & . \\ 0 & 1 & -4 & 6 & -4 & 1 & 0 & . & . & . \\ \ldots & & \ldots & & \ldots & & \ldots & & \ldots & \\ \ldots & & \ldots & & \ldots & & \ldots & & \ldots & \end{pmatrix} . \tag{4}$$

Although a regularised solution is tailored to be sufficiently stable, it does not usually conform to all our prior knowledge about it. In particular, when the solution is a brightness distribution it must be positive everywhere whereas the traditional regularized solution leads to regions including spurious negative values ("sidelobes") and hence violates the positivity requirement. This is one of the major reasons for departing from the domain of linear filter theory and seeking new deconvolution methods to incorporate prior knowledge about the solution. The incorporation of positivity has been the most fruitful in giving superresolution compared to classical methods. The possibility of achieving super-resolution by using positivity has been demonstrated by Biraud (1969) and several others (see e.g. Frieden, 1975). In order to simplify the discussion, we will first describe ODM with positivity as the only prior knowledge and then consider a more general problem with other constraints.

2.2. *ODM in a Simple Case*

The suggested deconvolution method for positivity ($q_i \geqq 0$) as the prior information involves the following two stages:

(i) An initial solution is first obtained which is the RLS obtained by solving Eq. (2).

(ii) This initial solution is used as the first-approximation (zeroth order iteration) in the "positivity algorithm" described below which incorporates positivity iteratively retaining consistency with the criterion of RLS.

Positivity Algorithm. Positivity is incorporated iteratively in an RLS by minimising at each stage a weighted sum of those values of q_i which were negative (and hence violated the positivity constraint) in the previous iteration. More precisely, the normal equations for the k^{th} iteration are written as:

$$\sum_j q_j [C_{ij} + \gamma H_{ij} + \omega_j^{(k)} \delta_{ij} / \Omega^{(k)}] = R_i \,, \tag{5}$$

where

$$\omega_j^{(0)} = 0; \quad \omega_j^{(k+1)} - \omega_j^{(k)} = h_j^{(k)} \tag{6a}$$

and

$$\Omega^{(k)}=\begin{cases}\sum_{j=1}^{n} h_j^{(0)}, & k=1\\ \sum_j h_j^{(1)}, & k>1\end{cases} \tag{6b}$$

with

$$h_j^{(k)}=\begin{cases}0, & q_j^{(k)}\geqq 0\\ (q_j^{(k)})^2, & q_j^{(k)}<0\,.\end{cases} \tag{7}$$

These equations are much simpler than they appear at first sight. The constraint-violating (negative) values of the solution are controlled in the iterations by penalty terms which affect only the diagonal elements of the normal equation matrix, $(C_{ij}+\ldots)$. The definition of $h_j^{(k)}$ ensures that there is no penalty corresponding to the regions where the solution was already non-negative. In each iteration, the diagonal elements are incremented by an amount proportional to the constraint-violation as decided by the last term in the left-hand side of Eq. (5). The choice of weights is not unique and good results could be obtained even with a choice: $\Omega^{(k)}=1$ and $\omega_j^{(k+1)}-\omega_j^{(k)}=1/n$ or 0 depending respectively on whether $q_j^{(k)}$ was negative or not. However, we found that the iterations converged particularly rapidly for the choice suggested above.

The role of $\Omega^{(k)}$ is simply to provide a scaling factor for the weights, and their changes during the iterations are incorporated through $\omega_j^{(k)}$. However, for a rapid convergence, it is preferable to define $\Omega^{(k)}$ from the solution in the first iteration as done above, instead of the initial solution.

The criterion for convergence of iterations should be chosen according to the nature of the problem. For instance, if one has a rough estimate of the expected mean error in $q(x)$, then the iterations can be stopped when the *rms* of the negative values is significantly less than this expected error. However, in the case of lunar occultations to which we have applied ODM, one has an estimate of the *rms* error on the occultation curve which corresponds to that of the flux (occultation step) of the source. Since the flux is proportional to the area of the brightness profile, this means that one has an estimate of the error (say, σ) in the area under the solution. In this case the iterations can be stopped when the total contribution to the area by the negative values of the solution becomes $\lesssim 0.5\,\sigma$.

Since the successive iterations alter only the diagonal elements of the normal equation matrix, the matrix as a whole has to be computed only once. The equations are linear and can be easily solved by standard methods. The fact that the matrix is symmetric can be used to reduce the computation time for solving the system of equations. Considerable saving is also possible by using iterative methods for linear equations like Gauss-Seidel method (Dahlquist and Bjork, 1974, Sect. 5.6) which will converge very fast except perhaps for the initial solution.

2.3. ODM *in a General Situation*

The general formulation of a deconvolution problem should consider two aspects. First, all the systematic effects due to the instrument and observing conditions should be properly introduced. Secondly, the solution should be consistent with all our prior knowledge about it. These points will be elaborated below in order to explain the suggested ODM for a general problem.

2.3.1. Instrumental Effects

In general, there may be several instrumental effects like the beam shape, bandpass, time constant, etc., which independently convolve the incoming signal with their respective characteristic functions. The effective point-source response of the system is then a convolution of all such functions. Henceforth it will be assumed that $p(x)$ denotes this composite function. Most observing conditions also lead to continuous changes in the background like the baseline drifts or the spectral continuum. These can usually be expressed linearly in terms of a few, say K parameters $\{\alpha_s\}$, i.e., by a function of the form $\sum_s f_s(x')\alpha_s$, e.g., a polynomial with $f_s(x)=x^{s-1}$. Thus, a general representation of the observational values r_m is given by:

$$r_m=\sum_{i=1}^{N} w_i p_{mi} q_i+\sum_{s=1}^{K} f_s(x')\alpha_s+\varepsilon_m\,.$$

For simplicity, we rewrite this equation in the form

$$r_m=\sum_{l=1}^{N} f_{ml} q_l+\varepsilon_m\,,$$

where

$$N=n+K\,;\qquad q_{n+s}=\alpha_s$$

and

$$f_{ml}=\begin{cases}w_l p_{ml}, & l\leqq n\\ f_{l-n}(x'_m), & l>n\,.\end{cases} \tag{8}$$

The RLS will now have to minimise

$$\varrho(\underset{\sim}{q})=\sum_{m=1}^{n_d}\left(r_m-\sum_{l=1}^{N} f_{ml} q_l\right)^2\Big/ n_d+\gamma\sum_i (\Delta^2 q_i)^2/(n-2) \tag{9}$$

which has N unknowns q_l.

2.3.2. Prior Knowledge

Prior information should now be introduced as constraints on the minimisation of $\varrho(\underset{\sim}{q})$. These constraints can be either equalities like $F(\underset{\sim}{q})=0$, or inequalities like the bounds: $u_i\leqq g_i(q_i)\leqq v_i$, $i=1,2,\ldots,n$. Without any loss of generality we will consider one example for each of these forms and explain how they are incorporated in ODM.

Equality constraints are introduced by the classical method of Lagrange multipliers which can be summarised as follows. Minimisation of $\varrho(\underset{\sim}{q})$ under the constraint $F(\underset{\sim}{q})=0$ is achieved by minimising the augmented function

$$\varrho_a(\underset{\sim}{q})=\varrho(\underset{\sim}{q})+\lambda F(\underset{\sim}{q})\,,$$

where λ is a Lagrange multiplier, obtained along with the desired solution by solving the set of $N+1$ equations

$$\partial\varrho_a/\partial q_i=0\,,\qquad i=1,\ldots,N$$

and

$$F(\underset{\sim}{q})=0$$

which will all be linear if $F(\underset{\sim}{q})$ is linear in $\underset{\sim}{q}$.

Inequality constraints are typically encountered when one specifies upper and/or lower bounds on the solution or an arbitrary function of it, $g_i(q_i)$, $i=1,\dots,n$. These can be incorporated by a straightforward generalization of positivity algorithm which is obtained by simply redefining the terms $h_i^{(k)}$ occurring in Eq. (6) as follows:

$$h_i^{(k)}=\begin{cases} 0, & u_i\leqq g_i\leqq v_i \\ (u_i-g_i)^2, & g_i<u_i \\ (v_i-g_i)^2, & g_i>v_i\,, \end{cases} \tag{10}$$

where we have written g_i for $g_i(q_i^{(k)})$ for brevity. The specification of bounds on the solution itself would correspond to $g_i(q_i)=q_i$.

For the sake of completeness we will now summarise our statement of the general problem and the suggested ODM for it.

Problem. Obtain an RLS for

$$r_m=\sum_l f_{ml}q_l+\varepsilon_m\,,\qquad m=1,\dots,n_d$$

subject to

$$u_i\leqq q_i\leqq v_i$$

and

$$\sum_{i=1}^{n} t_iq_i=t$$

Solution. The solution is obtained through iterations as before and the equations for the k^{th} iteration can be written as:

$$\sum_{j=1}^{N} A_{ij}^{(k)}q_j^{(k)}+\lambda t_i=R_i\,,\qquad i=1,\dots,N \tag{11}$$

and

$$\sum_{i=1}^{n} t_iq_i=t\,, \tag{12}$$

where

$$R_i=\sum_{m=1}^{n_d} f_{mi}r_m\,,\qquad i=1,\dots,N\,, \tag{13a}$$

$$A_{ij}^{(k)}=\begin{cases}\sum_m f_{mi}f_{mj}+\gamma H_{ij}+\omega_j^{(k)}\delta_{ij}/\Omega^{(k)}\,, & i,j\leqq n\\ \sum_m f_{mi}f_{mj} & \text{otherwise}\,,\end{cases} \tag{13b}$$

and $\omega_j^{(k)}$, $\Omega^{(k)}$ are given by Eqs. (6a), (6b), and (10) with the substitution $g_i=q_i^{(k)}$ in Eq. (10).

2.4. Computational Simplifications

The major computations in ODM arise from the evaluation of the symmetric matrix (C_{ij}) and the iterations involving the solution of N linear equations. If both the input data and the solution are sampled at the same interval Δx, the computation of (C_{ij}) is simplified by the following relation:

$$\frac{1}{w_{i+1}w_{j+1}}C_{i+1,j+1}=\frac{1}{w_iw_j}C_{ij}+p_{n+1,i}p_{n+1,j}-p_{1i}p_{1j},\quad i,j<n\,.$$

Then the major burden arises from the solution of N linear equations. For this, the time taken is $\propto N^3$ for direct methods, but $\propto N^2$ for iterative methods like Gauss-Seidel method (see e.g. Dahlquist and Bjorck, 1974, Sect. 5.6). Further, in many situations, $p(x)$ decreases with x fast enough (e.g. in Fraunhoffer diffraction) so that it can be neglected in comparison with noise for large x. In these cases, (C_{ij}) will essentially have a band structure, with all the significant values concentrated in a narrow band parallel to the main diagonal. Many computer systems offer library packages with special routines to handle such band matrices leading to a considerable saving of computer time.

It may be remarked here that since $n=\text{range}/\Delta x$, computation in ODM is essentially $\propto 1/\Delta x^2$. Thus one should choose Δx as large as possible for minimum computation. This may sometimes conflict with Eq. (1) which assumes that a sampling of $p(x)$ at intervals Δx is close enough to justify the summation formula used for evaluating the convolution integral. Such an artificial lower limit on Δx set by $p(x)$ can be overcome in several ways. For instance, one can retain the n values of q_i separated at Δx as the unknowns in the equations, but in addition, define intermediate values of the solution in terms of q_i through a suitable interpolation formula. In other words, an interpolation formula is used to define an auxiliary set of values of $q(x)$ sampled at a closer interval $\Delta x_{\text{int}}<\Delta x$ in terms of the n basic unknowns q_i. When these are used in Eq. (1) to reduce the errors of integration formula, the basic form of the equation remains unchanged, but the coefficients of q_i will be different from w_ip_{mi}. All the subsequent equations are also valid after the new coefficients are used. The artificial dependence of Δx on $p(x)$ is also avoided in the specification of resolution as in the next section.

2.5. *Resolution*

The question of resolution poses a fundamental problem in methods like ODM which use known properties of the solution itself while deriving it. A classical deconvolution scheme is equivalent to the convolution of observed data with a predetermined "restoring function", and one can identify the "resolution" as a property of the restoring function. However, with the use of prior knowledge, the linear relation between the observations and the solution can no longer be maintained, and the conventional "resolution" is too gross a term to characterise such a method in which the derived solution is influenced by its own properties. Thus, in order to interpret the observed widths properly, one needs a new criterion which should be determined, *a posteriori*, from the derived solution. This is as yet an unsolved problem facing the use of prior knowledge in deconvolution. In the absence of an analytical procedure, it is worthwhile investigating the statistics of residuals for empirical clues regarding this aspect. Admittedly, this may not solve the problem fully since it is not clear if the residuals reflect properly the influence of quantities like Δx, γ and positivity which all contribute to the final solution obtained. However, their use is amplified below because of the possibility of an *a posteriori* definition of resolution which is rigorously valid in a linear method. The procedure is still being investigated through extensive numerical studies and further details will be published in due course.

2.5.1. Power-spectrum of Residuals and Resolution

It is usually possible to anticipate the general behaviour of the power-spectrum of observational noise by examining the noise in a typical region well away from the signal. If the solution obtained for a given deconvolution problem is indeed the best fit to the observations, one can expect that the power-spectrum of residuals should resemble that of the typical noise. But, in practice, one may find that even a broad resemblance in shape exists only upto a certain frequency ν_0 and that the power spectrum of residuals is quite erratic at higher frequencies. One can then hope that the Fourier components of the solution, too, have been recovered reasonably faithfully upto ν_0 and may not be reliable at higher frequencies. The effective resolution can then be inferred as being $\propto 1/\nu_0$, with the constant of proportionality decided by the general shape of the power-spectrum and convention followed in defining the term "resolution". For instance, if one defines resolution as the half-power-width of a restoring beam, and the noise has an approximate Gaussian power-spectrum, then the equivalent resolution is $\sim 1/\nu_0$. A more detailed investigation of this procedure for an *a posteriori* inference of the effective resolution is in progress and the results will be published in due course.

We wish to emphasize here that such a procedure is only intended to obtain a correction to the measured widths for resolution effects. In particular, no claim has been made for any inference of the shape of the restoring beam or its effects on the details of the restored profile. However, we feel that this can provide a uniform criterion for defining the angular size of a radio source from its restored profile which is important in any statistical investigation of the sizes as for cosmological studies.

2.5.2. Resolution Specification

Subject to a limiting resolution determined mainly by $p(x)$, noise, and possibly by the nature of the solution itself, one can restore a profile with any specified resolution. This simply implies a smoothing of the observed data to enhance the signal-to-noise ratio for the coarse properties of a source like its flux or centroid. Restoration with several specified resolutions has been found convenient in a commonly followed practical procedure for a routine reduction of lunar occultations of radio sources by a conventional method due to Scheuer (1962). In this way, one tries to obtain various parameters of a source by restoration with the optimum resolutions with respect to the relative errors with which the parameters can be determined. For instance, the intensity (area) is best determined by a coarse resolution whereas for estimating the size of a component, one would prefer a resolution of the order of the size itself, provided, of course, it is feasible under the given observing conditions (von Hoerner, 1964).

It is possible to allow a specification of resolution even in ODM provided it does not amount to surpassing the resolution limit mentioned above. A possible warning against such an attempt to specify an unattainable resolution may be obtained in a situation where the resolution inferred from the power-spectrum of residuals turns out to be much different from the specified value.

For convenience, we assume that the desired restoring beam is Gaussian such that the solution $q_s(x)$ with a resolution β_1 is $q*G_1$, where G_1 is a Gaussian of width β_1. At a first glance, one may like to recover it using pre-smoothed data $r*G_1$. However, a modification will be suggested below which will improve the computational efficiency considerably in many cases.

As pointed out in Sect. 3.4, it is computationally advantageous to overcome the upper limit on x set by $p(x)$. One should be able to choose it entirely on the basis of the required resolution say $\Delta x \sim 0.25\beta_1$. If this is too large to be allowed by the nature of $p(x)$, one can handle the situation without unduly sacrificing accuracy by pre-smoothing $p(x)$ with a Gaussian G_2 of width $\beta_2 < \beta_1$. The input data also need to be pre-smoothed to the extent given by

$$r*G_3=(p*G_2)*(q*G_1)+\varepsilon*G_3,$$

where $G_3=G_1*G_2$ is now a Gaussian of width $\beta_3=(\beta_1^2+\beta_2^2)^{1/2}$. This can be written as

$$r_s=p_s*q_s+\varepsilon_s$$

which is identical in form to Eq. (1) with r, p, ε and q replaced by the corresponding smoothed quantities $r_s=r*G_3$, etc. Hence the procedure for ODM described in the previous sections can be used for this revised problem. It is clear that the choice of β_2 or β_3 is not unique, although β_1 imposes an upper limit on β_2. An extreme but convenient choice would be: $\Delta x=0.25\beta_1$ and $\beta_2=3\Delta x$, implying $\beta_3=5\Delta x$.

3. Comparison with Other Methods

3.1. Application to Lunar Occultations

An evaluation of ODM will now be made by comparing it in detail with an established linear method for the specific case of lunar occultations of radio sources. This example has the particular advantage that the attainable resolution by a classical method is only limited by noise and not by the form of $p(x)$.

3.1.1. Lunar Occultations and the Classical Method

During the occultation of a radio source by the Moon (distant D from the observer), the lunar limb can be well-approximated by a straight edge diffracting the radio waves from the source. For the present discussion, we will assume monochromatic observations at a wavelength $\lambda_0=0.92$ m, corresponding to an observation with the Ooty Radio Telescope (Swarup et al., 1971a). The observations can be considered as the result of a one-dimensional scan across the source with a point-source-response $p(\theta)$ given by the Fresnel diffraction curve:

$$p(\theta)=\tfrac{1}{2}\{[\tfrac{1}{2}+C(\theta/\theta_0)]^2+[\tfrac{1}{2}+S(\theta/\theta_0)]^2\}, \tag{15}$$

where

$$C(x)=\int_0^x \cos(\pi u^2/2)du, \qquad S(x)=\int_0^x \sin(\pi u^2/2)du$$

and θ is an angular variable along the line of scan with a scale factor $\theta_0=(\lambda_0/2D)^{1/2}=7.2$ arc s. It was shown by Scheuer (1962) that the restoration with a resolution β_1 is simply obtained by convolving the observed occultation curve with a restoring function

$$\propto \frac{d^2}{d\theta^2}p(-\theta)*G_1(\theta)$$

and that the attainable resolution is limited only by the observational noise. Further details about the lunar occultation observations and the restoring techniques can be found in a review by Hazard (1976). We will assume that the restoring function is normalised such that the ideal occultation curve of a point source would be restored with a peak amplitude equal to its flux, F. As in any other linear method, the effects of noise are inseparably superposed on the restored output. If the rms-noise on the occultation curve is σ_0 for an integration over 1 arc s, it can be shown (von Hoerner 1964) that the rms-noise σ_1 on the restored output with a resolution β_1 is given by

$$\sigma_1=1.42\sigma_0/\sqrt{\beta_1}. \tag{16}$$

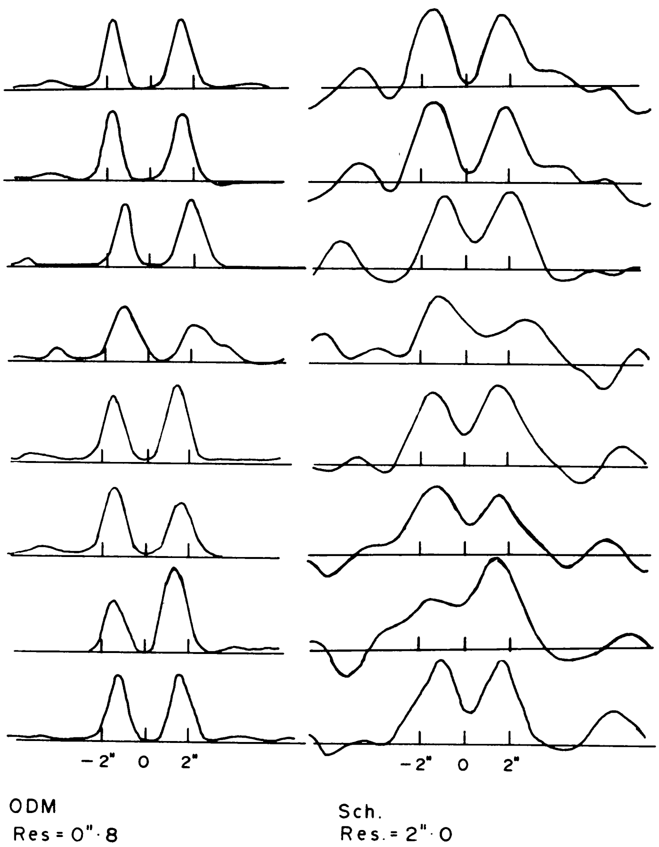

Fig. 1. Comparison of restorations of simulated occultations using ODM and Scheuer's method

By setting $F/\sigma_1 = 5$ in this relation von Hoerner (1964) defined the limiting resolution attainable for an unresolved source in Scheuer's method as $\beta_s = 52(\sigma_0/F)^2$.

3.1.2. Comparison of ODM with Scheuer's Method for Simulated Occultations

In a non-linear method like ODM, the resolution depends on the nature of the solution itself in addition to the observing conditions. Hence one cannot define a limiting resolution in ODM corresponding to β_s for Scheuer's method. This makes it difficult to assess theoretically the improvement obtained by ODM over Scheuer's method or any other deconvolution method. Hence a comparison was attempted earlier (Subrahmanya, 1977) by simulating about 150 noisy occultation curves of standard single and double Gaussian sources. Noise was represented by means of random numbers with standard normal distribution. Each of these simulations was analysed by ODM as well as Scheuer's method. This was done at an early stage in the development of ODM and the solution obtained was an RLS with positivity constraint as in Sect. 2.2. In particular, no attempt was made either to incorporate a specified resolution or to evaluate it from the residuals as in Sect. 2.5. However, this drawback was compensated by a detailed comparison of the distribution of the observed widths β_m (i.e. uncorrected for resolution effects) in both ODM and Scheuer's method. Instead of repeating the results on all the 150 simulations in this paper, we will confine ourselves to a representative set (called "*D*" in Subrahmanya, 1977) which corresponds to the simulations of 25 independent occultations of a source with two equal Gaussian components of width 1 arc s and amplitude 2.5 (i.e., flux F = area = 2.65) situated at $\theta = \pm 1.5$ arc s. The data were sampled at 0.25 arc s and corresponded to a value of $F/\sigma_0 = 5.3$ and $\beta_s = 1.9$ arc s. Restorations were performed with ODM without specifying resolution, and an output-sampling interval of 0.25 arc s was used. However, the measured total widths of all the 50 components averaged to $\bar{\beta}_m = 1.2$ arc s with an rms-deviation of $\Delta\beta_m = 0.5$ arc s. Since the true width of each component is 1 arc s, the ideal value of the measured width would be $(\beta_0^2 + 1)^{1/2}$, where β_0 is the resolution in ODM. In the absence of a better estimate of the resolution, we will assume $\beta_0^2 = \bar{\beta}_m^2 - 1 + \Delta\bar{\beta}_m^2$, which gives a value of $\beta_0 = 0.8$ arc s. Thus, the effective resolution in the ODM-restorations is about twice better than the limiting resolution $\beta_s = 52(\sigma_0/F)^2 = 1.9$ arc s for Scheuer's method. It is remarkable that such a super-resolution has been achieved in ODM without any sacrifice in the accuracy of any other parameter of the source. For the 25 cases, rms-fluctuations of the measured values of flux and position were respectively $\Delta F = 0.5$ and $\Delta\theta = 0.25$ arc s, compared to the expected values, 0.7 and 0.25 respectively for Scheuer's method at its limiting resolution.

The visual appearance of a typical ODM-output is striking and renders the difference between the two methods much more prominent than the effect of super-resolution. Examples are shown in Fig. 1, which gives a comparison of the two restorations for 8 of the 25 cases discussed above. The adjacent curves in these figures are the restorations of the same data by ODM (left) and Scheuer's method. The "cosmetic" improvement in ODM results mainly from positivity and does not, by itself, imply a quantitative improvement of a similar order in the restoration. It may be recalled here that the restoring function of a classical method simply convolves the noise-contaminated data and thus allows all the artefacts of restoration and the effects of noise to contaminate the solution visibly. This means that the accuracy of restoration can be judged by a visual inspection of the amplitudes of the ripples or sidelobes in the solution which can be easily recognised as spurious. On the other hand, the use of least-squares criterion and positivity make a recognition of the artefacts of an ODM-restoration much less straightforward on the solution than on the residuals. Thus, one can be confident about the detection of a weak feature in the solution only if it is significant above the level of wiggles in the *residuals*.

3.1.3. Analysis of Residuals

So far, we have not talked about the nature of residuals. In order to examine their behaviour, more simultations were performed and we will now summarize the major inferences drawn on the basis of this experience. As emphasized earlier, the concept of resolution is not straightforward in ODM and hence an attempt to derive an empirical relation like Eq. (16) between σ_1/σ_0 and β_1 is not expected to be successful. Further, since the non-uniqueness of deconvolution occurs mainly because the convolution is insensitive to the *high-frequency* Fourier components, a linear method should be expected to give faithful pictures of the object as one goes to resolutions coarser than the limiting resolution. Hence the difference between ODM and Scheuer's method is expected to become insignificant gradually for resolutions $> \beta_s$. As anticipated, our efforts at finding a simple relation analogous to Eq. (16) were not successful even for $\beta_1 \sim \beta_s$. The value of $\sigma_1 \sqrt{\beta_1}/\sigma_0$ which is 1.42 for Scheuer's method, usually ranged between 1.0 and 1.4 for ODM.

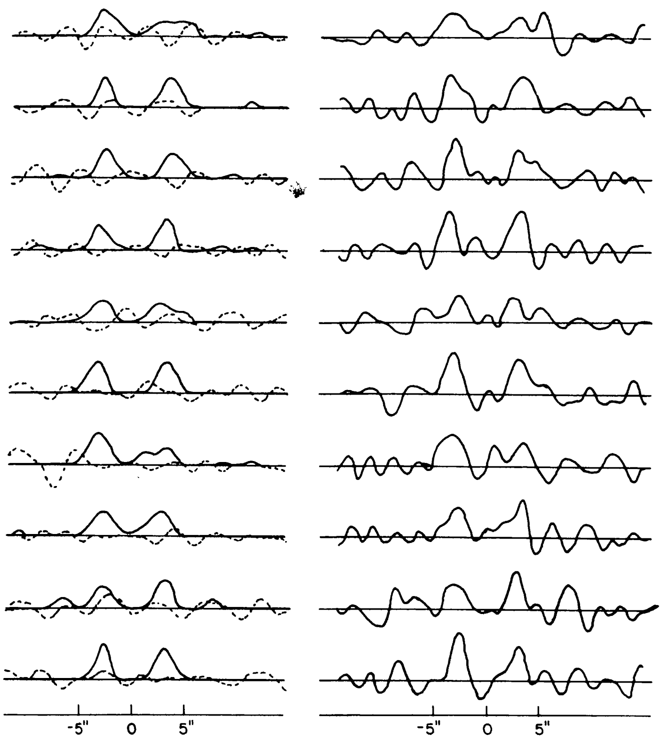

Fig. 2. Restoration of simulated occultations with ODM and Scheuer's method. Differentiated residuals are superposed on the profiles using ODM by dotted lines

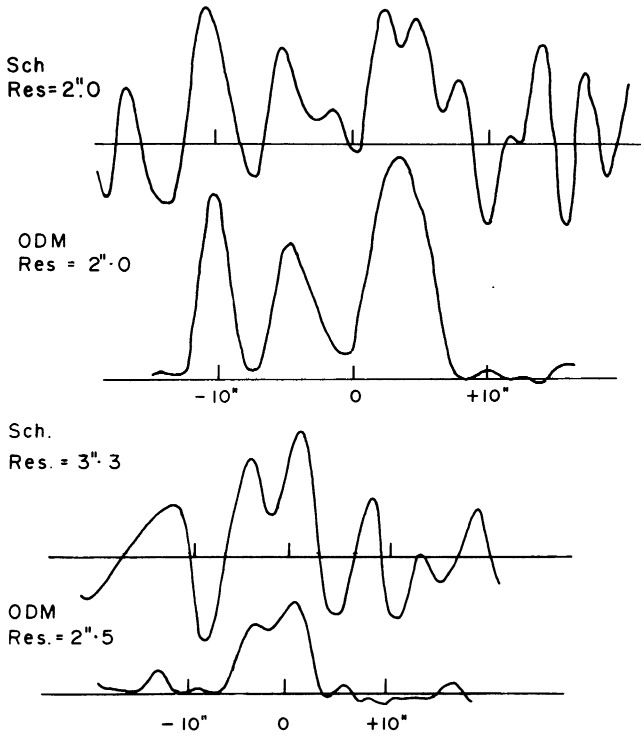

Fig. 3a and b. ODM and Scheuer's method for lunar occultations of **a** OTL 0500+270 along position angle 116°, and **b** OTL 1754−276 along 78°

The next step is to examine the behaviour of residuals themselves rather than confining to their variance. It may be remarked here that the residuals is ODM are similar to those obtained in any curve-fitting programme. Hence their conventional use to improve the solution is also applicable here. For instance, if a residual is particularly high (say, $>2\sigma_1$), then the corresponding value in the input can be eliminated from a second run of ODM to obtain a better solution. However, one cannot possibly hope to superpose error bars on the solution using the residuals since they are in a different domain altogether and the required transformation is again a deconvolution. But we believe that one can always find a compromise analogous to using the geometric optics approximation in place of diffraction in order to get a quick visual feeling for the reliability of a certain feature in the solution. This can be accomplished in lunar occultations by using the differentiated residuals, i.e., their first-differences to depict the errors in the solution domain. This corresponds to ignoring the fringes in $p(x)$ and regarding it as a step-function. It must be remembered that these provide only an *overestimate* of the actual errors involved since they include contributions both from an approximation made for $p(x)$ and also from the enhancement of errors in evaluating the numerical differences. Examples of differentiated residuals are given in Fig. 2, which shows some restorations with ODM on which the derivatives of residuals are superposed in dotted line. These are the result of 10 successive simulations of the occultations of equal double sources with components of width 1 arc s and separation 6 arc s, and corresponding to a signal-to-noise ratio $F/\sigma_0 = 5$. The limiting resolution in Scheuer's method is 2 arc s. The restorations in ODM were performed with a specified resolution of 1.5 arc s, but the effective resolution as inferred from the zero-crossings of residuals averaged to 1.8 arc s for the 10 simulations. Restored profiles using Scheuer's method for a resolution of 1.5 arc s are shown alongside each ODM profile in Fig. 2. The figure thus illustrates how ODM can extract out an "optimum" solution and separately provide the noise distribution.

4.1.4. Comparison for Actual Occultations Observed with the Ooty Radio Telescope

ODM has also been used successfully on more than 100 lunar occultations observed with the Ooty radio telescope (Subrahmanya, 1977; Subrahmanya and Gopal-Krishna, 1980; Venkatakrishna and Swarup, 1980). This application has led to a confidence in its general validity and practicality for this case. It is now being routinely used to supplement the information obtained from Scheuer's method for the lunar occultations being currently observed at Ooty. Examples of these restorations are provided in Figs. 3 and 4, in which the restorations with ODM are plotted below the corresponding restorations with Scheuer's method (Sch). The specified resolutions (Res.) are indicated alongside each profile. These profiles have only been given for a visual comparison between the two methods. The details on the position, structure and optical identification of the sources can be found in the following references:

OTL 1556–260: Swarup et al. (1971b);
2127–157: Kapahi et al. (1974);
0500–270 and 1754–276: Subrahmanya and Gopal-Krishna (1980).

The general conclusion drawn from our experience so far is that the use of ODM for live data does not pose any specific

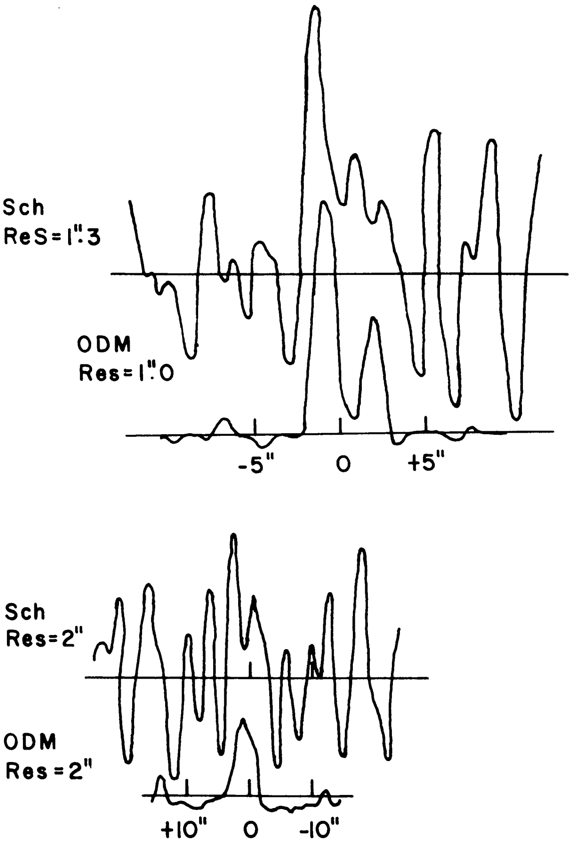

Fig. 4a and b. ODM and Scheuer's method for lunar occultations of **a** OTL 1556−260 along position angle 71°, and **b** OTL 2127−157 along 65°

problem. If a particular restored output were to seem unrealistic, it is usually due to the wrong choice of an empirical parameter and not because of any limitation of ODM as such. To this end, some general comments will now be made which will also serve as guidelines for judging whether a particular output is reliable or not, and also for improving it if necessary.

Because of the inherent nature of convolution, it may sometimes happen that a quick visual glance at the "clean" output of ODM may lead to an interpretation of a spurious step resulting from a local instability in the baseline as a genuine component of the source. But the suspicious nature of such a feature is usually evident from a closer examination of the behaviour of noise in the raw data or the differentiated residuals in the corresponding region. A good guideline to reject a spurious peak is also provided by comparing restorations performed with different specifications of ranges of the input data and the output profiles given by the solution. A genuine feature is stable and hence inferred equally significantly from such restorations. Further, it is recommended to obtain restorations with at least two specified of γ, differing from each other by about of 5 and select one by visual inspection. If the choice of γ is too low, too many marginally significant or insignificant peaks may appear on the restored profile. When the choice is too high, the change in the width of components with γ is usually prominent. Very often, a choice of too small a value of γ leads to components with measured widths narrower than the specified resolution and such outputs are easily rejected.

The use of second-differences for smoothness does not have any influence near the extremities, where only one-sided derivatives can be defined. Usually, these do not affect the solution away from the end-points. Moreover, since the assumed range of the solution implies that $q_1 = q_n = 0$, one can easily control the behaviour at the end-points by specifying bounds on such values so that their absolute value will be small.

If one also wants to find the best-fitting baseline along with the deconvolution, it is essential to introduce a constraint on the flux(area) of the source or any of its components. Otherwise, a vertical shift may occur in the baseline of the restored output. However, when the baseline is already known beforehand, the flux-constraint is generally redundant.

Acknowledgements. I acknowledge with grateful thanks the many suggestions, comments and advice I received from Prof. G. Swarup throughout this work. I have also received invaluable help from many of my colleagues through fruitful discussions at many stages of this work and I sincerely thank all of them. I also thank the referee whose comments have helped me in improving the text.

References

Biraud, Y.: 1969, *Astron. Astrophys.* **1**, 124

Bracewell, R.N., Roberts, J.A.: 1954, *Australian J. Phys.* **7**, 615

Dahlquist, G., Bjorck, A.: 1974, Numerical Methods, Transl. Anderson, N., Englewood Cliffs, Prentice Hall

Frieden, B.R.: 1975, in *Picture Processing and Digital Filtering*, ed. Huang, T.S., Springer, Berlin, Heidelberg, New York

Hazard, C.: 1976, in *Methods of Experimental Physics* **12** C, ed. Meeks, M.L., Academic Press, New York, Chap. 4.6

Hoerner, S. von: 1964, *Astrophys. J.* **140**, 65

Kapahi, V.K., Gopal-Krishna, Joshi, M.N.: 1974, *Monthly Notices Roy. Astron. Soc.* **167**, 299

Papoulis, A.: 1965, Probability, Random Variables, and Stochastic Processes, McGraw-Hill, New York, Sect. 14.4

Phillips, D.L.: 1962, *J. Assoc. Computing Machinery* **9**, 84

Scheuer, P.A.G.: 1962, *Australian, J. Phys.* **15**, 333

Subrahmanya, C.R.: 1975, *Bull. Astron. Soc. India* **3**, 13

Subrahmanya, C.R.: 1977, Ph.D. Thesis, Bombay University

Subrahmanya, C.R., Gopal-Krishna: 1980, *Mem. Astron. Soc. India* **1**, 2

Swarup, G., Sarma, N.V.G., Joshi, M.N., Kapahi, V.K., Bagri, D.S., Damle, S.H., Ananthakrishnan, S., Balasubramanian, V., Bhave, S.S., Sinha, R.P.: 1971a, *Nature Phys. Sc.* **230**, 185

Swarup, G., Kapahi, V.K., Sarma, N.V.G., Gopal-Krishna, Joshi, M.N., Rao, A.P.: 1971b, *Astrophys. Letters* **9**, 53

Tikhonov, A.N.: 1963, *Dokl. Akad. Nauk* **151**, 501

Tikhonov, A.N., Arsenin, V.Y.: 1977, Solutions of Ill-posed Problems, Wisconsin, Washington

Twomey, S.: 1963, *J. Assoc. Computing Machinery*, **10**, 97

Twomey, S.: 1965, *J. Franklin Inst.* **279**, 95

Venkatakrishna, K.L., Swarup, G.: 1980, *Mem. Astron. Soc. India* **1**, 25