The problem of the missing low spatial frequencies.

m Kesteven & j Caswell

15 March 1986

1 Introduction

A problem common to all synthesis telescopes is the matter of the missing short spacings: the central region of the visibility plane is unsampled because of the difficulty in operating two telescopes at spacings comparable to an antenna diameter. The shortest spacing in the compact array of the AT is 30.6m.

The consequences of this missing data can be serious if the source has significant amounts of extended structure. The map which results from the Fourier inversion of the visibility data will have a "negative bowl", of extent largely dictated by the dimensions of the missing spacings. This depression is undesirable on aesthetic grounds - the sky is, after all, positive. More seriously, deconvolution schemes such as CLEAN can become unstable when confronted by an unhappy mix of sidelobes and negative bowl. Finally, the map will be incorrect, since data is missing: attempts to compare this map with maps made at other frequencies in order to derive spectral index variations, for example, should be treated with caution until it is shown that the missing data represents a small fraction of the source, or until the missing data is restored.

A number of schemes have been proposed to counter this problem : Williams, Kenderdine & Baldwin (1966); Bajaja & Van Albada, (1979); Ekers & Rots, (1979); Braun & Walterbos, (1985). This note describes these schemes. and examines their relevance to the AT.

Notation

The discussion which follows involves representation of the data in both domains - map and visibility. These are a Fourier transform pair; we will use lower case for the visibility plane, and upper case for the map plane. As usual an asterisk (*) will be used to denote the convolution operation.

2 The Problem

Let $B_{true}(l)$ be the true sky brightness, $A_{samp}(l)$ the transform of the sampling function in the visibility plane. $A_{ant}(l)$ is the primary beam pattern.

The observed brightness is then given by :

$$B_{obs}(l) = [B_{true}(l) \cdot A_{ant}(l)] * A_{samp}(l)$$

For a regular, linear array such as the CA the sampling function $a_{samp}(\underline{u})$ will be more-or-less uniform from inner radius r_{min} out to a maximum baseline r_{max}. We can write for a_{samp}:

$$a_{samp}(\underline{u}) = a_{full}(\underline{u}) - a_{inner}(\underline{u})$$

where

$$a_{full}(\underline{u}) = 1 \qquad \text{for all } |u| < r_{max}$$

and

$$= 0 \qquad \text{beyond;}$$

$$a_{inner}(\underline{u}) = 1 \qquad |u| < r_{min}.$$

$$= 0 \qquad \text{beyond.}$$

Then

$$B_{obs}(\ell) = [B_{true}(\ell) \cdot A_{ant}(\ell)] * A_{full}(\ell)$$

$$- [B_{true}(\ell) \cdot A_{ant}(\ell)] * A_{inner}(\ell)$$

The effect of the missing samples is difficult to predict, as it depends on the morphology and distribution of the source(s) over the entire field of view. Some very rough guidelines are offered here.

The full beam is narrower than the inner beam; therefore every source must show a positive central peak embedded in a negative bowl. If the emission is from a point source then the ratio (positive peak/negative bowl) is substantial ~ $(r_{max}/r_{min})^2$ ~ $(6000/30)^2$. With extended structure the ratio drops: in the extreme case when the structure is greater in extent than the primary beam then no flux may be detected, as it all falls in the unsampled central region. The mean level will be zero if the zero spacing is missing.

A shell source could well display a positive perimeter surrounding a negative central region. All the sources in the field will contribute to the negative bowl, so an inoffensive source may be overwhelmed by strong nearby sources.

The strategies for coping with the missing spacings depend on the source structure: one might be able to interpolate into the central region if the source structure is all fine scale; ie. if one has *a priori* evidence that there is no large scale structure, or if one is not interested in the large scales. Interpolation is unlikely to be successful if large scale structure is present and important.

3 Interpolation schemes

3.1 Direct interpolation - Braun & Walterbos (1985)

If the map is largely complete - all spacings present from r_{min} (eg. 30m) to r_{max} (6 km), then the processing effort should be able to concentrate on the few missing samples (~ 4 points in the Westerbork case). Braun & Walterbos (1985) describe an interpolation procedure which they apply to WSRT data with good results. It is important to recognize the two preconditions: a source with structure well distributed in the visibility plane (the low spatial frequencies must not dominate if the interpolation is to succeed): and well sampled visibilities (with essentially all the spacings).

Their argument is that the region outside the source should not be negative. The corresponding minimum solution which they seek is therefore that set of amplitudes for the missing low frequency spacings which removes the negative bowl and produces a zero intensity baseline. The algorithm proceeds in two steps:

a. Filter out high spatial frequencies from the map. This is a critical step, and requires a certain act of faith - blank out the region of the source; then remove high frequency noise by a "$\kappa\sigma$" clip technique. (Iteratively remove points departing from the mean by more than some specified multiple of the variance). Let the resulting map be M.

b. Find the low frequency spacing amplitudes by minimizing :

$$C = \sum[(Sky*A_{inner}) - M]^2$$

This is a tractable operation since we are looking for a small number of spacings (4 or so).

This scheme requires the source to be small compared to the minimum size of the negative bowl. (Otherwise the blanking operation will fail). For the CA, for example, with a minimum spacing of 30m, the negative bowl size is approximately half the primary beam. Equivalently, this means that one could add arbitrary broad sources (half primary beam sized) to their solution.

3.2 CLEAN

Iterative deconvolution schemes such as CLEAN and MEM perform an (implicit) interpolation in the visibility plane. If the structure is essentially small scale the operation will automatically remove the bowl. (This is essentially an article of faith - the map is ok if it looks ok).

Manifest ,difficulties occur in complex fields (eg, galactic sources); a negative sidelobe plus the negative bowl may be greater (in absolute magnitude) than the next genuine peak, so that CLEAN will launch after fictitious negative sources. The CLEAN can be stabilised by adding a positive "lump" to the map before CLEANing; the operation is honest if the "lump" is removed at the end of the operation. The "lump" is chosen to cancel (approximately) the negative bowl.

The Braun & Walterbos procedure should perhaps be viewed as a "pre-CLEANing" operation, rather than the replacement of CLEAN that B&W envisaged. Some form of CLEAN or MEM will likely be needed on all maps in order to remove the near-in sidelobes, as well as to provide the model of the sky for self-calibration operations.

4 Direct observation of the missing spacings

Direct observation of the missing spacings is clearly the only unambiguously correct procedure. The questions examined below are: how to obtain the data and how to calibrate it; and how to combine it with the interferometer data.

4.1 Single dish observations. Bajaja & van Albada (1979)

Suppose we were to make a map with a large diameter telescope. For example, we may use Parkes 64m observations to complement the CA data. We need to ensure that both data sets are placed on the same flux scale, and we need to correct for the two primary beam patterns.

4.1.1 The data

CA observations

$$b_{meas}(\underline{u}) = [b_{true}(\underline{u}) * a_{CA}(\underline{u})] \cdot a_{samp}(\underline{u})$$

(The visibility measured at point $\underline{u}$ in the (u,v) plane is a convolution of the transform of the sky (at $\underline{u}$) with the transform of the CA primary beam pattern).

Parkes

$$B_{meas}(\ell) = B_{true}(\ell) * A_{PKS}(\ell)$$

(The map is the sky convolved with the Parkes primary beam pattern).

4.1.2 The algorithm

a. We centre the PKS map on the tracking centre used for the CA;

b. Fourier transform to the (u,v) plane:

$$b_{meas}(\underline{u}) = b_{true}(\underline{u}) \cdot a_{PKS}(\underline{u})$$

Remove the effects of the Parkes beam - we would like a unit point source to have unit visibility for all reasonable $\underline{u}$. $|u| < r_p$, where r_p is of order the half-power width, although one could presumably find a more sophisticated definition which would optimize the signal/noise.

thus, $b'_{meas} = b_{true}$ for $|u| < r_p$,

$= 0$ beyond.

Convolve b' with the transform of the CA primary beam, in order to ensure that these visibilities are consistent with the CA visibilities:

$$b''_{meas} = b'_{meas} * a_{CA}$$

We need to truncate b'' at a radius $r_c < r_p$, since the convolution will mimic interferometer visibilities only as long as a_{CA} falls in the region $|u| < r_p$.

$$b''_{meas} = 0 \qquad \text{for } |u| > r_c$$

c. We can now concatenate the two visibility data sets:

interferometer data: $b_{meas} = [b_{true} * a_{CA}] \cdot a_{samp}(\underline{u})$

and single dish : $b''_{meas} = [b_{true} * a_{CA}] \cdot \prod(u/r_C)$

The new beam will be the Fourier transform of $[a_{samp}(\underline{u}) + \prod(u/r_C)]$

The data concatenation could also be performed in the map plane - the maps would be added with scale factors given by the respective areas covered in the (u,v) plane.

Thus provided that the single dish that measured the missing spacings is substantially larger than the interferometer dishes, there are few risky approximations involved. We need a good representation of the primary beams, out to around the half-power points. The calibration amounts to requiring that a unit point source produce a known amplitude deflection in the single dish observations, and visibilities of known amplitude in the interferometer observations.

4.1.3 Practical details and problems: limitations

a. r_p and r_c

The critical quantity in this discussion is the magnitude of the scaling applied to the data: we will be dividing the observed visibility (the fourier transform of the map) by the antenna's response in the (u,v) plane. We will truncate the division operation when the response has fallen to some critical value, η. This critical value could be set by the signal to noise ratio; for illustration we will adopt $\eta = 0.5$ as a convenient compromise.

In figure 1 we show the response function for the Parkes 64m and for the Culgoora antennas. (The derivation is discussed in the appendix).

A reasonable value for r_p would be that value of r for which $\eta = 0.5$. That is, 27m (Parkes), and 8.5m (Culgoora).

r_c is set by the convolution operation. As a means of estimating r_c consider the following simplifying scheme: we convolve a pillbox function of radius r_P with the Culgoora beam. Out to some radius r_L the integral is constant, as the Culgoora beam falls entirely within the pillbox. Beyond r_L portions of the Culgoora beam lie outside the pillbox and the integral decreases. We set r_c to be the radius at which the integral has fallen to a fraction ε of the central value. $\varepsilon = 0.9$ is a possible choice. We then find $r_c \sim r_P - D_c/2$, or $r_c \sim 17$m. Figure 2 illustrates the r_P/r_c/beam relationship.

Thus Parkes, at 64m, is only just adequate for the Culgoora array, on this formulation. Matters are alleviated somewhat by extending r_P to greater scale factors. (Requiring correspondingly higher signal/noise ratios).

It means that a Culgoora dish could not be used.

This argument is perhaps overstating the case. We have a plausible estimate of the sky (b_T) from 0 to r_P; we have a measure of the sky convolved with the primary beam ($b_T{*}a_c$) beyond. We should therefore be able to estimate ($b_T{*}a_c$) from 0 to r_P. Given such an algorithm, the limitation is then simply the scaling of the data at r_P. If a 22m dish were used to estimate the missing 15m spacing the factor required is of order 5, which is probably higher than acceptable in most cases.

It still means that a Culgoora dish could not be used.

b. Beam switching ?

Single-dish observations (in the continuum) are frequently made with some form of beam switching, in order to overcome the baseline fluctuations due to atmospheric effects. Maps can be made with beam switching - that is, we can deconvolve the beam switching artefacts. Is such a map suitable for our purposes here? (We assume that the dish is large enough - ie. satisfies the criteria discussed above).

The beam switching operation amounts to making two maps:

$$M_1(\ell+\Delta) = \int B(\ell' - \ell - \Delta) A(\ell') d\ell'$$

and $$M_2(\ell-\Delta) = \int B(\ell' - \ell + \Delta) A(\ell') d\ell'$$

$$M(\ell) = M_1 - M_2$$

$$m(\underline{u}) = \sin(2\pi \underline{u}\Delta) b(\underline{u}) a(\underline{u})$$

We can then recover $(b(\underline{u})a(\underline{u}))$ by scaling $m(\underline{u})$ by the sine term.

The separation of the beams is generally of order 5 beamwidths; which means $\Delta = 2.5\theta \sim 3.3(\lambda/D)$. The wavelength of the sine term is comparable to the spacing we need (15m), so care will be needed, and some fitting operation may be necessary.

We conclude that beamswitching should introduce no serious obstacle to obtaining a good estimate of the 15m spacing. The zerospacing is somewhat more tricky; however it can be argued that it is not required as it cannot alter the shape of a map, but only the zero level.

4.2 Restoration of the missing spacings using an interferometer.

(The Williams et al/Ekers & Rots method)

In rough outline, the algorithm is:

a. Take the shortest baseline interferometer that is available. (30m for the CA).

b. Assume, for the moment, that the sky is stationary over the interferometer.

c. Keep the delay and local oscillators fixed, appropriate to the field centre.

d. Drive both antennas from the field centre, recording the interferometer output as a function of the pointing offset ($I(\ell)$).

f. Fourier transform $I(\ell)$ to obtain a "visibility" function from which short spacing information can be obtained.

The operation is one-dimensional - we attempt to estimate the visibility data over the range of radius 0 to r_m (the shortest spacing available), along the line normal to the interferometer fringes.

Figure 3 shows a simplified version of the argument. The upper panel shows one component of the sky brightness (corresponding to a spacing s). The second panel has the fringes set up for the interferometer spacing u. The third panel shows a stylized polar diagram.

The interferometer output is obtained from the product of panels 1 and 2, convolved with 3. The normal arrangement is to sample the result at one point only. (With the antenna pointing at the field centre).

The product 1 by 2 is :

$$P = B\cos(2\pi s \lambda).\cos(2\pi u \lambda)$$

$$= B[\cos(2\pi(u-s)\lambda) + \cos(2\pi(u+s)\lambda)]$$

The convolution operation is essentially a smoothing one - if the polar diagram A spans many cycles, then P*A will be small. However, there can be a range of values of s for which we can extract some significant information about B(s). Our interest is in the region u/2, if we choose u = 30m.

$$FT(P*A) = p.a = B\delta(u-s).a(s) = b(u-s).a(s)$$

For completeness we offer a more rigorous derivation. The interferometer output is:

$$X(\underline{u}, \ell') = \int B(\ell)A(\ell' - \ell)e^{j2\pi\underline{u}.\ell}d\ell$$

(defined for all ℓ'). Transform:

$$x(\underline{u},\underline{s}) = \int e^{j2\pi\underline{s}.\ell'}A(\ell' - \ell)d\ell' \int B.....$$

$$= a(\underline{s}) \int e^{j2\pi\underline{s}.\ell}B(\ell)e^{j2\pi\underline{u}.\ell}d\ell$$

$$= a(\underline{s})b(\underline{u}-\underline{s})$$

Thus provided that the antenna has reasonable response at u/2 (15m for the CA), we can measure b(u/2). But this is the crunch - at 15m the CA antenna is far down the polar diagram, to around 1/5. Thus a large scale factor is required, which in turn means that the signal/noise ratio must be high if the process is to succeed.

Some further manipulations are required before the visibility can be added to the general data set - we need it in the form : b*a, if we are subsequently to correct for the antenna primary beam.

$$b'(\underline{u}_1) = \int b(\underline{u}')a(\underline{u}_1-\underline{u}')d\underline{u}'$$

$$= \int x(\underline{u},\underline{u}-\underline{u}')a(\underline{u}_1-\underline{u}')/a(\underline{u}-\underline{u}')d\underline{u}'$$

$\underline{u}$ is the smallest spacing available (30m); $\underline{u}_1$ is the spacing we are after (15m). In the integral $\underline{u}'$ will range from :

u_1-u' = -15 to +15; ie, u' ranges from 0 to 30;

the denominator thus ranges from a(0) to a(30) - to very small values, leading to substantial amplification of noise. It seems therefore that this scheme is unlikely to be workable in the Culgoora context.

This conclusion is consistent with the observation that this scheme is unlikely to be an improvement over the use of a CA dish in the single dish mode. We find that a single dish has the same area providing a 15m spacing, as do two dishes placed 30m apart. (The area is given by:

$$\text{Area} = 2[\frac{\pi}{2} - \theta - \sin(2\theta)/2]$$

where,

$$\theta = \sin^{-1}[(D-S)/2R] \text{ (interferometer)}$$

$= \sin^{-1}(S/2R)$ (single dish)

where R is the dish radius, S is the spacing in question, and D is the interferometer baseline).

The essential problem as far as the AT (CA) is concerned is that the antennas must be positions at designated stations, which then imposes the 30m minimum spacing and the attendant problems. This type of limitation did not apply to the Williams et al case: they made the observations "as close together as possible without overlapping".

4.3 Restoration in the map plane

There has been some suggestion (T.Cornwell) that a single-dish map could be used as an *a priori* estimate for MEM, or even CLEAN. This means that the concatenation of the data sets takes place in the map plane, with some reduction in the computation load. The essence of the procedure is this:

a. convolve the single dish map (suitably tarted up) with the dirty beam of the interferometer map;

b. subtract the result from the dirty map;

c. CLEAN;

d. add the single dish map (unconvolved) to the CLEAN map.

There are a number of difficulties at present unresolved - just how do you tart up the map to account for the various polar diagrams? More serious, perhaps, you now have a map from which the positivity constraint has been removed. The CLEAN operation may well produce garbage, but the diagnostic tools have been largely destroyed.

references

Bajaja & van Albada (1979) Astr. & Ast. 75, p.251

Braun & Walterbos (1985) Astr. & Ast. 143, p.307

Ekers & Rots (1979) Image formation from coherence
functions. (van Schooneveld ed.) p.61

Williams, Kenderdine & Baldwin (1966) MNRAS 70, 53-110

Appendix

The visibility weighting function of a single dish

If we are to combine single dish and interferometer data in the visibility plane we need to have an accurate measure of the telescope's response in the (u,v) plane. We estimate here that transform.

1. Culgoora antennas

The design of the feed and subreflector is intended to produce a uniform illumination (voltage) out to the edge of the 22m aperture. (cf. G.James, AT/21.1.1/034 and 045). This should result in a beam of the form $(J_1(x)/x)^2$, and a transform:

$$f(u) = \cos^{-1}(u/2a) - (u/2a)\sqrt{(1-(u/2a)^2}$$

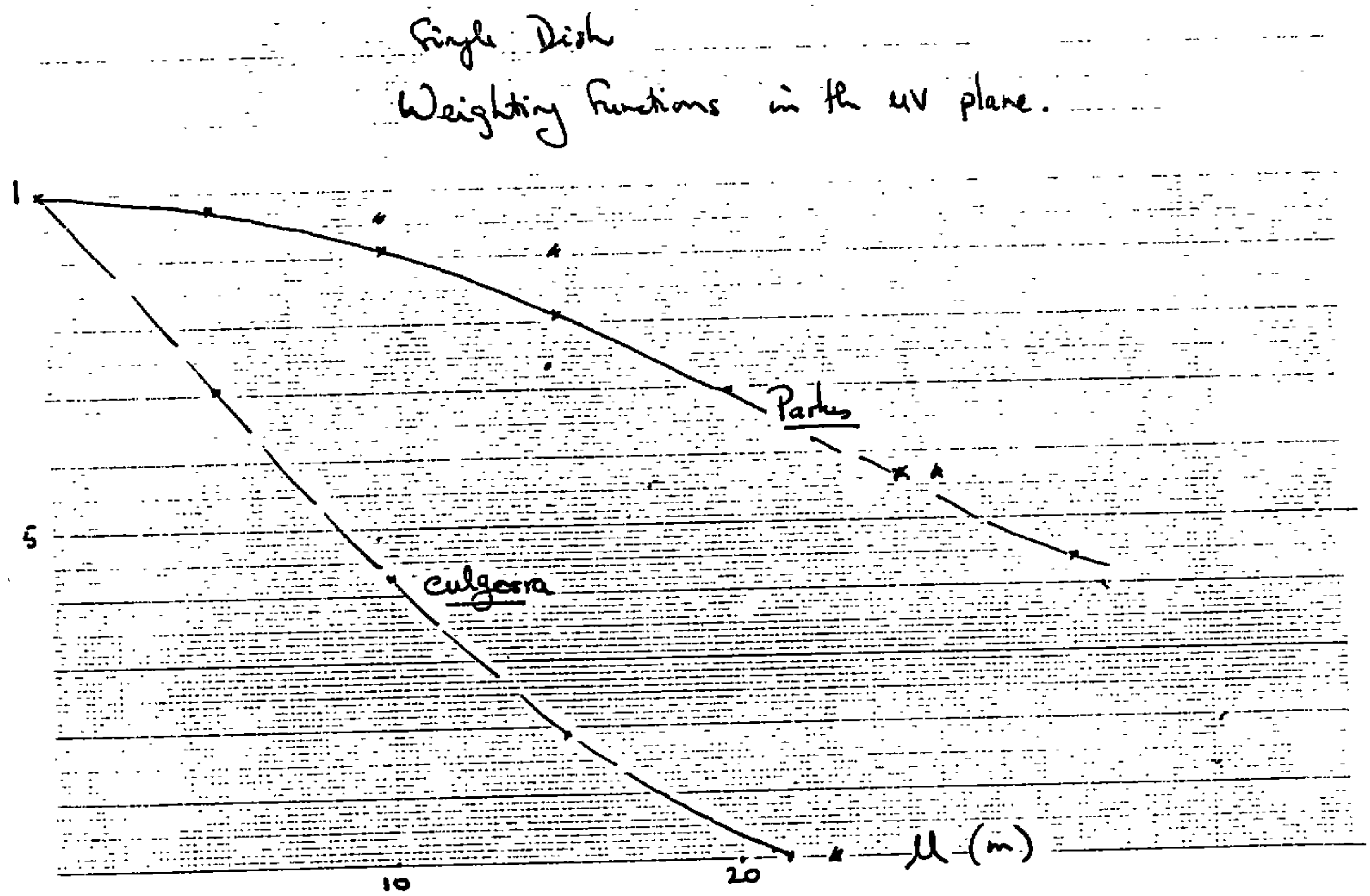
From the computed 5.9 GHz beamshape (G.James) we find a = 10.7m. The corresponding f(u) is shown in fig. 3

2. Parkes.

This function is less critical, as it is much broader than the Culgoora response. The illumination is said to be generally gaussian, down to 20% at the edge; (ie. down to 0.5 at r = 21m). Evaluating the transform of a gaussian truncated at the 20% level, squaring, then transforming, we find a gaussian which falls to 0.5 at r = 27m, shown in fig. 3

----

file: [mjk.scal]spacing.rno



Single Dish
Weighting Functions in the uv plane.
1
5
Parkes
Culgoora
10
20
u (m)

Fig 1

The Convention problem

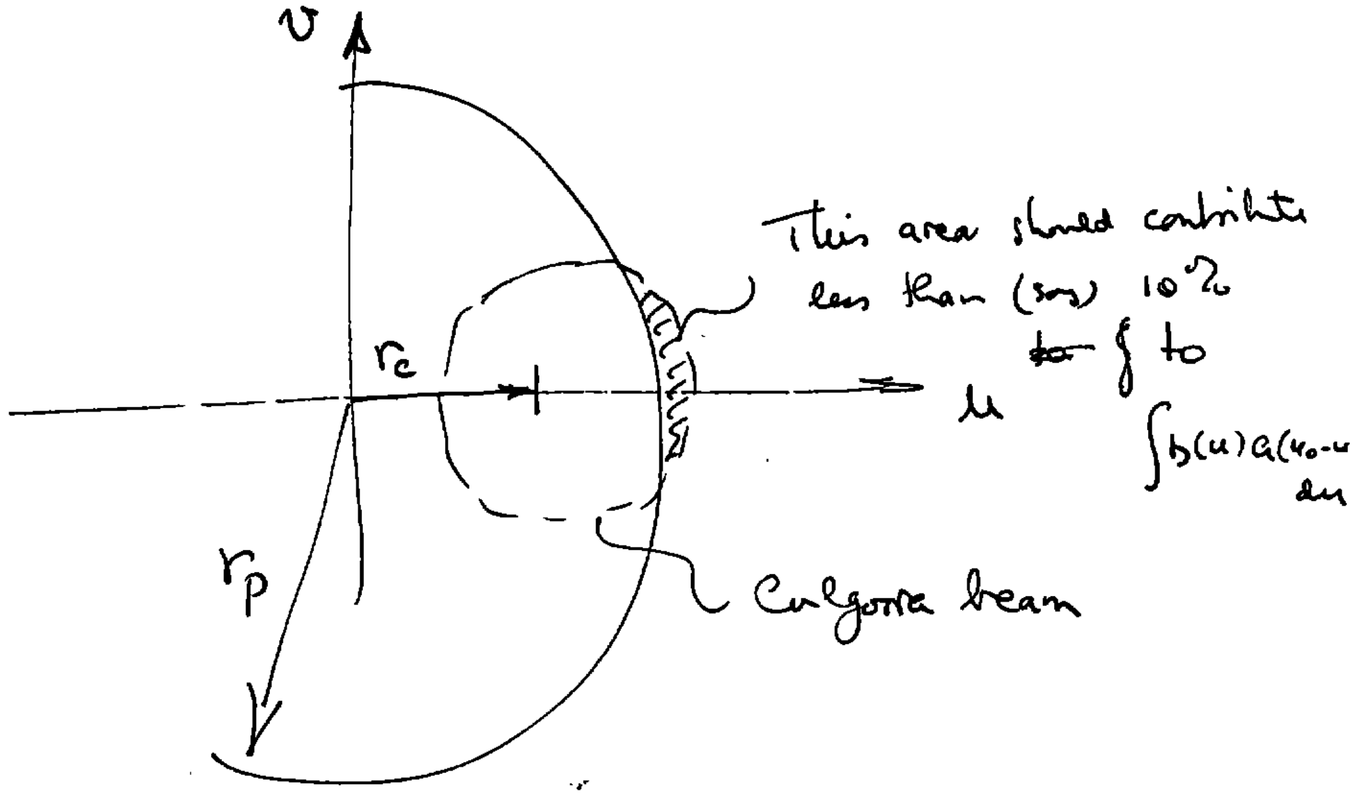


fig. 2

fig. 3

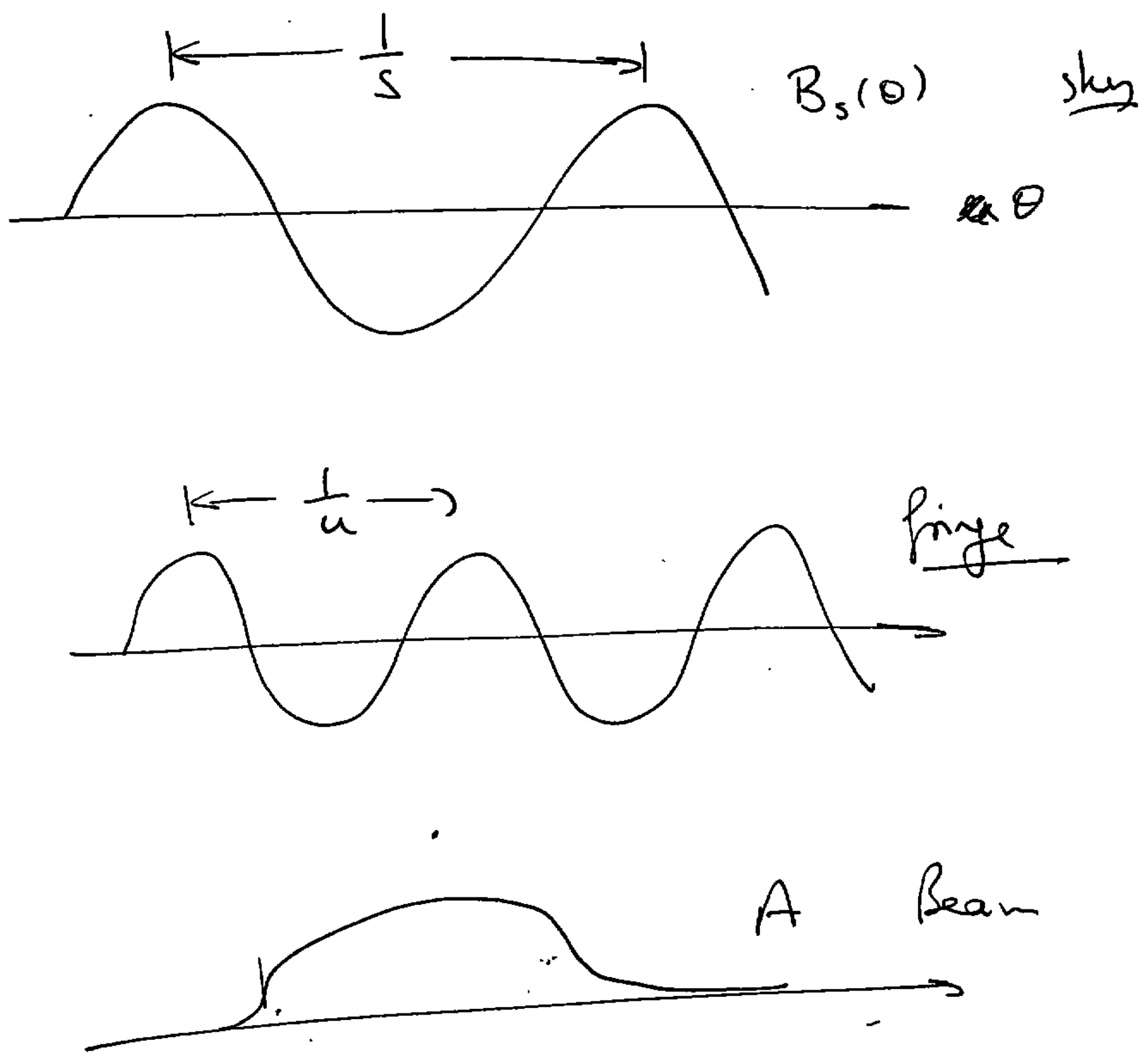


fig 3