ON THE USE OF LUNAR OCCULTATIONS FOR INVESTIGATING THE ANGULAR STRUCTURE OF RADIO SOURCES

By P. A. G. Scheuer*

[*Manuscript received May 3, 1962*]

Summary

It is shown that it is possible, in principle, to recover the strip distribution of brightness across a radio source from the Fresnel diffraction pattern which is observed as a source is occulted. Practically useful methods of approximating to this ideal are indicated. Rather surprisingly, the resolution attainable in the presence of a given noise level is not related to the angular width of the first Fresnel zone, but is just the same as it would be if geometrical optics were valid.

I. Introduction

Hazard (1962) has demonstrated the practical possibility of obtaining precise positions of radio sources, and some information on their angular sizes, by observing lunar occultations. Some of his observations show considerable detail in the diffraction pattern, and the question arises: What information about the angular structure of the source can be deduced from these patterns? Hazard has calculated the diffraction patterns to be expected for a number of models of various diameters, which show that the pattern is sensitive to structure considerably smaller than the width of the first Fresnel zone (which, in Hazard's example, is about 10″ arc). He also writes: "From an analysis of the detailed structure of the diffraction pattern it should also be possible to obtain information on the brightness distribution across these intense sources." It turns out that the problem of recovering the strip brightness distribution across a source from a diffraction pattern has a surprisingly simple solution, and that the said brightness distribution can be found (in the presence of a given noise level) just as well as it could have been obtained from an occultation curve determined by geometrical optics alone.

The diffraction pattern on the Earth's surface due to a remote point source is the usual Fresnel diffraction pattern of a straight edge (over regions small compared with the Moon's shadow, and provided that the lunar mountains are much smaller than the first Fresnel zone, conditions which are satisfied in observations such as Hazard's), which gives (see Fig. 1):

$$\text{Amplitude} \propto \int_{-x}^{\infty} e^{\pi i y^2/\lambda D},$$

$$\text{Power} \propto p(x) = \int_{-x}^{\infty} e^{i\pi a y^2}\mathrm{d}y \int_{-x}^{\infty} e^{-i\pi a y^2}\mathrm{d}y, \tag{1}$$

where $a = 1/\lambda D$.

* Division of Radiophysics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

This pattern is swept past an aerial on the Earth's surface by the motion of the Earth and the Moon, so that the occultation curve of a point source is a constant multiple of (1) (Fig. 2 (*a*)), and the occultation curve of a more complicated brightness distribution is the true brightness distribution convolved with (1).

Now, in the approximation of geometrical optics, where (1) is replaced by a step function, the occultation curve gives, at each instant, the total flux density of the unocculted part of the source, and the brightness distribution in strips parallel to the Moon's limb is obtained simply by differentiating the occultation

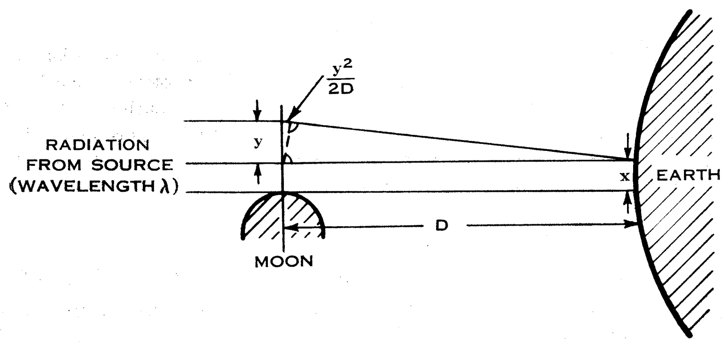

Fig. 1.—Illustrating the derivation of Fresnel's diffraction formula.

curve. Differentiation will therefore yield an approximation to the strip distribution ; the differentiated occultation curve of a point source is clearly the differential of (1), which is

$$q(x)=p'(x)=\mathrm{e}^{\mathrm{i}\pi ax^2}\int_{-x}^{\infty}\mathrm{e}^{-\mathrm{i}\pi ay^2}\mathrm{d}y+\mathrm{e}^{-\mathrm{i}\pi ax^2}\int_{-x}^{\infty}\mathrm{e}^{\mathrm{i}\pi ay^2}\mathrm{d}y, \tag{2}$$

(Fig. 2 (*b*)) and, generally, the differentiated occultation curve is the true strip distribution convolved with (2).

It will be shown in the following sections that :

(i) The differentiated occultation curve contains all the Fourier components of the true strip distribution (in fact it contains them in equal proportions), and hence

(ii) it is possible to recover the true strip distribution.

(iii) The " restoring function " $c(x)$ with which the differentiated occultation curve must be convolved to recover the true strip distribution is just

$$p'(-x)=q(-x).$$

(iv) Any attempt to find a perfect " restoring function " which applies directly to the occultation curve leads to a non-convergent process, but a practically useful " restoring function " can be found which leads to a strip distribution with any assigned finite degree of resolution.

(v) The presence of noise on the occultation curve limits the fineness of the resolution that can be achieved, but this limitation is precisely the same as that which would occur if there were no diffraction. The resolution attainable is not related to the width of the first Fresnel zone. This result implies that the resolution, and also the positional accuracy, attainable in a lunar occultation experiment is not determined by the frequency chosen for the observations, except in so far as the frequency affects the signal-to-noise ratio, receiver stability, interference, etc.

II. Notation

The *Fourier transform* (which will be abbreviated to F.T.) of $f(x)$ shall mean

$$F(\nu)=\int_{-\infty}^{\infty} f(x)\mathrm{e}^{-2\pi i\nu x}\mathrm{d}x.$$

With this notation, if $F(\nu)$ is the F.T. of $f(x)$, then $f(x)$ is the F.T. of $F(-\nu)$ (inversion theorem); also, the F.T. of the complex conjugate of $f(x)$ is the complex conjugate of $F(-\nu)$.

The *convolution* of $f(x)$ with $g(x)$, written $f(x)*g(x)$, or simply $f*g$, shall mean

$$f*g=\int_{-\infty}^{\infty} f(y)g(x-y)\mathrm{d}y.$$

With this notation, convolution is commutative and associative, that is,

$$f*g=g*f, \text{ and } f*(g*h)=(f*g)*h,$$

provided that $f*g$, $f*h$, and $g*h$ all exist. Also, the F.T. of $f(x)*g(x)$ is $F(\nu)G(\nu)$ (convolution theorem).

Functions such as delta-functions ($\delta(x)$) and Heaviside unit functions ($H(x)$) occur almost inevitably in the calculations below. These functions do not yield convergent integrals when one attempts to find their Fourier transforms according to the above formula. They may, however, be regarded as limits of sequences of better-behaved functions, which have Fourier transforms in the usual sense, and a consistent rigorous theory can be constructed on this basis (Lighthill 1958).

III. Calculation of the Restoring Function

The differentiated occultation curve, $g(x)$, is the true strip distribution, $t(x)$, convolved with $q(x)$:

$$g(x)=t(x)*q(x),$$

therefore,

$$G(\nu)=T(\nu)Q(\nu) \text{ (convolution theorem).}$$

Hence if $Q(\nu)\neq 0$ for all ν (which we shall find to be true), we can find

$$T(\nu)=\frac{1}{Q(\nu)}G(\nu).$$

Using the convolution theorem again,

$$\begin{aligned} t(x)&=\{\text{function whose F.T. is } 1/Q(\nu)\}*g(x)\\ &=\{\text{F.T. of } 1/Q(-\nu)\}*g(x) \text{ (inversion theorem).}\end{aligned}$$

Thus convolution with the restoring function

$$c(x) = \text{F.T. of } 1/Q(-\nu) \tag{3}$$

recovers $t(x)$ from $g(x)$, and we now have to calculate the expression (3).

Firstly, with a real and positive

$$\begin{aligned}\text{F.T. of } e^{\pm i\pi a x^2} &= \int_{-\infty}^{\infty} e^{\pm i\pi a x^2} e^{-2\pi i \nu x} dx \\ &= e^{\mp i\pi\nu^2/a} \int_{-\infty}^{\infty} e^{\pm i\pi a (x \mp \nu/a)^2} d(x \mp \nu/a) \\ &= |a^{-\frac{1}{2}}| e^{\pm\frac{1}{4}\pi i} e^{\mp i\pi\nu^2/a}.\end{aligned} \tag{4}$$

Secondly, $q(x)$ may be written in the form

$$\begin{aligned} q(x) &= e^{i\pi a x^2} \int_{-x}^{\infty} e^{-i\pi a y^2} dy + \text{complex conjugate} \\ &= e^{i\pi a x^2} \int_{-\infty}^{x} e^{-i\pi a y^2} dy + \text{conjugate,}\end{aligned}$$

since $e^{-i\pi a y^2}$ is even,

$$= e^{i\pi a x^2} \times (e^{-i\pi a x^2} * H(x)) + \text{conjugate.}$$

Therefore,

$$Q(\nu) = \{\text{F.T. of } e^{i\pi a x^2}\} * \{\text{F.T. of } (e^{-i\pi a x^2} * H(x))\} + \text{conjugate function of } -\nu,$$

using the convolution theorem and the theorem on the F.T. of the complex conjugate of a function,

$$= \{\text{F.T. of } e^{i\pi a x^2}\} * (\{\text{F.T. of } e^{-i\pi a x^2}\} \times \{\text{F.T. of } H(x)\}) + \text{conj } (-\nu)$$

(convolution theorem again)

$$= \int_{-\infty}^{\infty} |a^{-\frac{1}{2}}| e^{\frac{1}{4}\pi i} e^{-i\pi(\nu-\mu)^2/a} . |a^{-\frac{1}{2}}| e^{-\frac{1}{4}\pi i} e^{i\pi\mu^2/a} \{\text{F.T. of } H(x)\} d\mu + \text{conj } (-\nu),$$

using (4) in the definition of a convolution,

$$\begin{aligned} &= \frac{1}{a} e^{i\pi\nu^2/a} \int_{-\infty}^{\infty} \{\text{F.T. of } H(x)\} e^{+2\pi i (\nu/a)\mu} d\mu + \text{conj } (-\nu) \\ &= \frac{1}{a} e^{i\pi\nu^2/a} H(\nu/a) + \frac{1}{a} e^{-i\pi\nu^2/a} H(-\nu/a) \text{ (inversion theorem).}\end{aligned}$$

That is,

$$Q(\nu) = \frac{1}{a} e^{i(\pi\nu^2/a) \operatorname{sgn} \nu}. \tag{5}$$

Thus $q(x)$ contains all Fourier components with equal weight (for $|Q(\nu)| = 1/a$ is a constant) but different phases.

From (5),

$$1/Q(-\nu) = a e^{i(\pi\nu^2/a) \operatorname{sgn} \nu} = a^2 Q(\nu),$$

and hence

$$\begin{aligned} c(x) &= \text{F.T. of } 1/Q(-\nu) = a^2 \{\text{F.T. of } Q(\nu)\} \\ &= a^2 q(-x).\end{aligned} \tag{6}$$

That is, the restoring function $c(x)$ is the same as the function $q(x)$ that smeared out the true distribution.

The basic result is, then,

$$t(x)=a^2q(-x)*g(x). \tag{7}$$

IV. Modification of the Solution for Practical Computation

Equation (7) is, in principle, a complete solution of the problem. It says that the true strip distribution may be recovered from the occultation curve by differentiation, followed by convolution with $a^2q(-x)$, where $q(x)$ is the analytic expression of (2) (which can be computed easily from tables of Fresnel integrals).

In practical computation, both processes are undesirable: graphical differentiation of an observed curve, often with appreciable noise, is at best an uncertain process, and as for convolution with $q(-x)$, Figure 2 (*b*) shows how slowly such a process would converge. The amplitudes of the successive positive and negative lobes of $q(x)$ approach a constant value, and their areas decrease only because of the increasing frequency of the oscillation.

Let us try to avoid the differentiation, by finding the restoring function that must be applied directly to the occultation curve $f(x)=\int_{-\infty}^{x} g(t)\mathrm{d}t$. In the context of the methods used above, this is most neatly set out by using the fact that differentiating a function is equivalent to convolving it with $\delta'(x)$, and integration from $-\infty$ is convolution with $H(x)$. Thus, formally, we can obtain a restoring function for

$$f(x)=\int_{-\infty}^{x} g(t)\mathrm{d}t=\int_{-\infty}^{\infty} g(y)H(x-y)\mathrm{d}y=g*H$$

by writing

$$\begin{aligned} t&=c*g=c*\{\delta*g\}=c*\{(\delta'*H)*g\}=c*\{\delta'*(H*g)\}\\ &=c*(\delta'*f)=(c*\delta')*f=\frac{\mathrm{d}c}{\mathrm{d}x}*f\\ &=a^2p''(-x)*f, \end{aligned} \tag{8}$$

so that the required restoring function must be $p''(-x)$. Figure 2 (*c*) shows how the amplitude of $p''(x)$ increases linearly for large positive x, so that the areas of the positive and negative lobes do not approach zero; hence, for occultation curves f which, though bounded, do not approach zero for large x, the process represented by the right-hand side of (8) does not converge. The fallacy in the derivation of (8) occurs in the step

$$c*(\delta'*f)=(c*\delta')*f,$$

which is valid only if $c*f$ is meaningful. This statement can easily be checked by partial integration of

$$\int_{-\infty}^{\infty} c(x-y)\frac{\mathrm{d}f(y)}{\mathrm{d}y}\mathrm{d}y.$$

But, since neither $c(x)$ nor $f(x)$ approaches zero for large positive x, $c*f$ is meaningless.

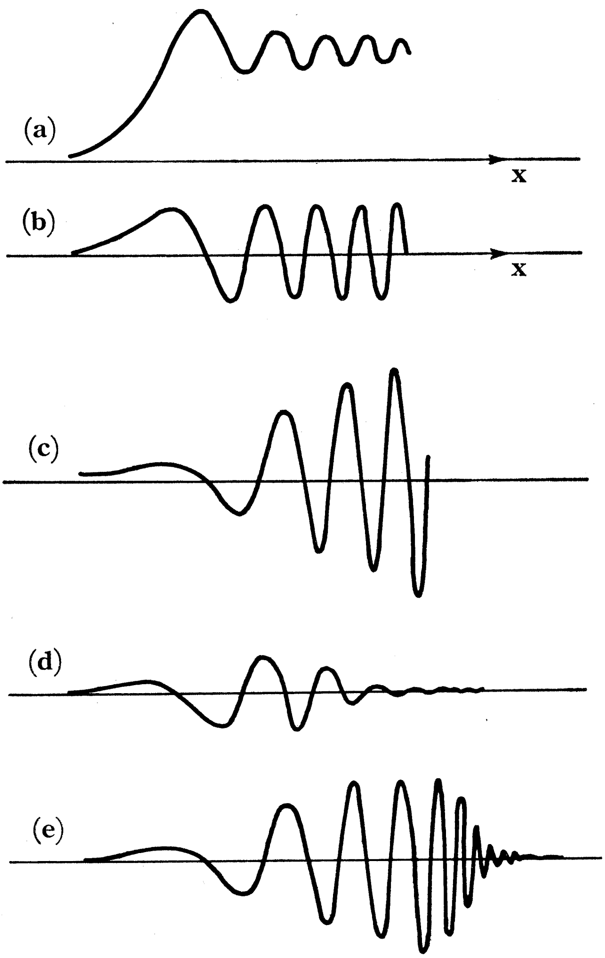

Fig. 2.—The variation of the illumination $p(x)$ near the edge of a shadow, and functions derived from it.
(*a*) $p(x)$; (*b*) $q(x)=p'(x)$; (*c*) $p''(x)$; (*d*) and (*e*) $\exp(-x^2/b^2)*p''(x)$.

The trouble is remedied if, instead of aiming at perfect restoration, one is content with a strip distribution as seen with an aerial beam of small but finite width—for example a Gaussian beam $\exp(-x^2/b^2)$, where b may be given any desired value. Then we obtain

$$\begin{aligned} e^{-x^2/b^2}*t &= e^{-x^2/b^2}*(c*g) = (e^{-x^2/b^2}*c)*g \\ &= (e^{-x^2/b^2}*c)*\delta'*f \\ &= \{(e^{-x^2/b^2}*c)*\delta'\}*f, \end{aligned}$$

a step which is now valid because

$$e^{-x^2/b^2}*c \to 0 \text{ as } x \to \pm\infty, \text{ though } c(x)=a^2q(-x) \text{ does not,}$$
$$\begin{aligned} &= \{e^{-x^2/b^2}*(\delta'*c)\}*f \\ &= a^2\{e^{-x^2/b^2}*p''(-x)\}*f. \end{aligned}$$

The " partial restoring function "

$$c_b(x) = e^{-x^2/b^2} * p''(-x) \tag{9}$$

is a perfectly reasonable function, vanishing like x^{-4} as $x \to +\infty$ and like $xe^{-a^2b^2x^2}$ as $x \to -\infty$.

Smaller and smaller values of b yield strip distributions of greater and greater resolution. The corresponding " partial restoring functions "

$$\exp(-x^2/b^2) * p''(-x)$$

will include more and more of the increasing oscillations of $p''(-x)$ before the cut-off appears (when the oscillations become narrower than $\exp(-x^2/b^2)$). Examples are sketched in Figures 2 (d) and 2 (e). Thus better and better resolution in the final strip distribution requires more and more precise observations, and a reasonable practical procedure would be to try several values of b in decreasing order, to find out what degree of restoration the data will stand—the appearance of strips of negative brightness is one clear danger signal.

I have not succeeded in finding expressions for $\exp(-x^2/b^2) * p''(-x)$ in terms of tabulated functions, and it seems they must be found numerically. There is, of course, nothing magical about the use of a Gaussian profile as the " aerial beam " aimed at; any sufficiently smoothly varying profile would do, and a suitable selection might yield functions like $c_b(x)$ more amenable to computation.

V. How Much Correction Should Be Attempted? The Effect of Noise

In most practical cases, correction procedures for instrumental broadening of various kinds are useful for narrowing the instrumental profile by a factor of two or so, and thereafter become unprofitable owing to the rapidly increasing precision required of the observed profile. In the present case, the natural scale of the instrumental profile is the width of the first Fresnel zone, and it is natural to assume that, in the restoration process, it may be profitable to aim at a resolution $1\frac{1}{2}$ or 2 times better, but no more. In the case of lunar occultations, this type of limitation does not arise.

In fact, the scale of the diffraction pattern is not a relevant quantity, and the actual occultation curve is just as useful as the occultation curve which would have been obtained if geometrical optics were valid ($\lambda \to 0$, or $a \to \infty$). This may be shown most easily by considering how the occultation curve given by geometrical optics

$$\int_{-\infty}^{x} t(y)\,dy = H(x) * t(x)$$

can be recovered from the actual curve $f(x)$.

By (7),

$$t(x) = a^2 q(-x) * f'(x)$$

so that

$$\begin{aligned} H(x) * t(x) &= H(x) * a^2 q(-x) * f'(x) \\ &= a^2 q(-x) * f(x). \end{aligned}$$

Thus convolution with $q(-x)$ turns the observed curve into the " geometrical " occultation curve. This process corresponds to passing the observed curve through a filter whose transfer function is $Q(\nu)$. But, as (5) shows, $|Q(\nu)|$ is constant, and only the phases of the various Fourier components of the noise are changed by passage through the filter. Since the phases of the noise components are in any case random, the r.m.s. noise level is left unchanged by convolution with $q(-x)$ (except for the scale factor $1/a$ in (5), which applies equally to the signal).

To make quite sure of a rather surprising result, the noise level produced by " partial restoration " with $c_b(x)$ will also be examined. So far, only the shape of the brightness distribution has been of interest, and constant factors have been disregarded. In assessing the signal-to-noise ratio, one must be more careful about normalization. Let us use a restoring function $Ac_b(x)$ such that an isolated source produces a bump on the " restored " curve whose area is numerically equal to the height of the step on the observed occultation curve. This condition requires that

$$\begin{aligned} f(\infty)-f(-\infty) &= \int_{-\infty}^{\infty} Ac_b(x)*f(x)\mathrm{d}x \\ &= \int_{-\infty}^{\infty} A\mathrm{e}^{-x^2/b^2}*p''(-x)*f(x)\mathrm{d}x \\ &= \int_{-\infty}^{\infty} A\mathrm{e}^{-x^2/b^2}*q(-x)*f'(x)\mathrm{d}x. \end{aligned}$$

The last expression is the value at $\nu=0$ of the F.T. of the integrand, and hence (by the convolution theorem) is the product of the F.T.'s of the three factors at $\nu=0$. Thus the condition is

$$\begin{aligned} f(\infty)-f(-\infty) &= A\sqrt{(\pi b)}.Q(0).\int_{-\infty}^{\infty} f'(x)\mathrm{d}x \\ &= A\sqrt{(\pi b)}.(1/a).(f(\infty)-f(-\infty)), \\ A &= a/\sqrt{\pi(b)}. \end{aligned}$$

When the observed curve with noise, $f(x)+n(x)$ say, is " restored " by convolving with $Ac_b(x)$, what happens to the noise? The noise becomes

$$Ac_b(x)*n(x)=A\mathrm{e}^{-x^2/b^2}*p''(-x)*n(x),$$

whose spectrum is

$$\begin{aligned} & A\{\text{F.T. of } \mathrm{e}^{-x^2/b^2}\}\{\text{F.T. of } q'(-x)\}N(\nu) \\ &= A\{\sqrt{(\pi b)}\mathrm{e}^{-\pi^2b^2\nu^2}\}\{-2\pi\mathrm{i}\nu Q(-\nu)\}N(\nu) \\ &= \frac{a}{\sqrt{(\pi b)}}\cdot\sqrt{(\pi b)}\mathrm{e}^{-\pi^2b^2\nu^2}.(-2\pi\mathrm{i}\nu)\frac{1}{a}\mathrm{e}^{\mathrm{i}\pi\nu^2/a\,\mathrm{sgn}\,\nu}.N(\nu) \\ &= -2\pi\mathrm{i}\nu\mathrm{e}^{-\pi^2b^2\nu^2}\mathrm{e}^{\mathrm{i}\pi\nu^2/a\,\mathrm{sgn}\,\nu}N(\nu). \end{aligned}$$

In this formula, a appears only in a phase factor, so that the power spectrum of the noise on the restored curve is independent of a, and in particular, it is the same as it would be in the limiting case $a\rightarrow\infty$ (geometrical optics).

Now, in the case of geometrical optics, the maximum resolution permissible in the presence of noise is very easy to estimate. Consider an occultation curve, such as Figure 3, determined by the geometrical shadow of the Moon. The highest useful resolution AB is reached when the total flux ΔS of that part of the source in the strip AB is detectable with the smallest useful signal-to-noise ratio in the presence of noise smoothed with the time constant AB. The resolution to be aimed at therefore depends on the source to be investigated; greater resolution can profitably be attempted on intense sources and on sources of small angular size than on faint or diffuse sources.

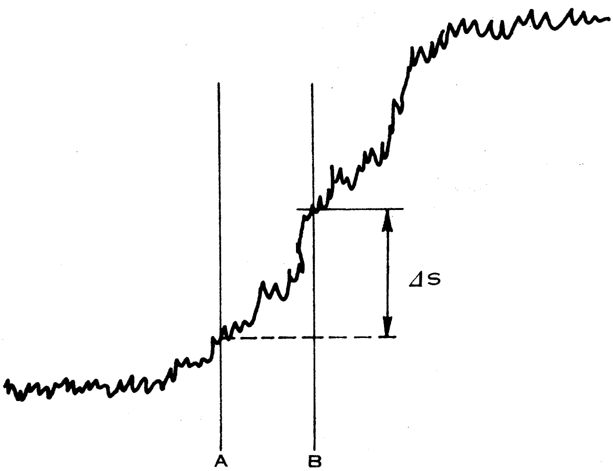

Fig. 3.—Illustrating the effect of noise.

VI. A Physical Model to Illustrate the Problem

Results as simple as those of Section III ought to be obvious from physical considerations, and indeed they are.

At any one time, we can replace the occulting system of Figure 1 by an equivalent system of aerials shown in Figure 4. An infinite linear array of dipoles (labelled α_1, α_2, α_3, . . .) is connected to a receiver through cables of varying lengths

$$\sqrt{(D^2+z^2)}\simeq D+\frac{z^2}{2D}=\frac{\lambda}{2\pi}\left(\frac{2\pi D}{\lambda}+\pi a z^2\right),$$

to simulate the varying path lengths in Figure 1, but all the array to the left of dipole α_1 is covered with an absorbing lid, which we shall still call the "Moon". This arrangement has the reception pattern $p(x)$ of Figure 2 (a). (The sensitivity is $p(x)$ at an angle x/D from the Moon's limb.) The change in the reception pattern when the "Moon" moves along a little way then gives the function $p'(x)=q(x)$, the reception pattern corresponding to the differentiated occultation curve.

We can now apply the usual theory of aerial arrays, which shows that the output from an aerial system is the sum of the outputs from the interferometers formed by all the possible pairs of aerial elements, and that such a pair with a base line z contributes to the Fourier component of the reception pattern with spatial frequency z/λ. Hence, when the " Moon " of Figure 4 slides along to expose one extra dipole (labelled α_0), the extra output is that due to all the possible pairs of dipoles that were not present before—the pairs $\alpha_0\alpha_0$, $\alpha_0\alpha_1$, $\alpha_0\alpha_2$, . . ., $\alpha_0\alpha_n$, . . . The fact that $q(x)$ has a flat Fourier spectrum now follows immediately from the fact that there is a uniform number of dipoles per unit length of the array. Also, the phases corresponding to the pairs $\alpha_0\alpha_n$ and $\alpha_n\alpha_0$ are respectively $\pi a z^2$ and $-\pi a z^2$ (ignoring the constant $2\pi D/\lambda$), so that all the features of equation (5) follow from the picture of Figure 4. (It will be noticed that the variable z used here is the same as ν/a.)

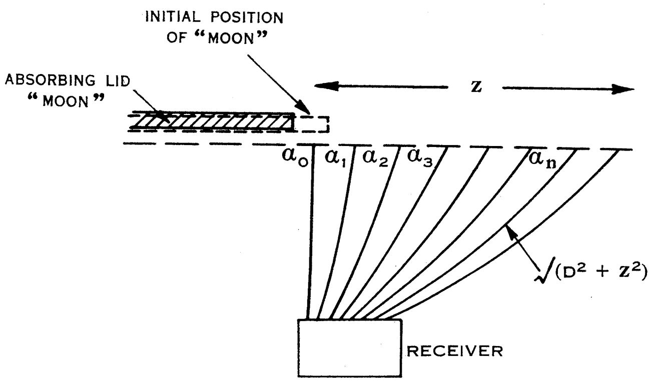

Fig. 4.—An aerial system whose response simulates a lunar occultation.

VII. Discussion

Some approximations made in the present paper should be pointed out.

(i) The irregularities on the Moon's limb have been ignored; these may amount to 1″ or 2″ of arc. The Moon's limb is also curved, but, since the first Fresnel zone is of the order of 10″ arc wide, the restoration process is not likely to be required for sources much wider than 1′ arc, and over this range the curvature introduces displacements smaller than the irregularities.

(ii) The treatment applies to a monochromatic receiver; the bandwidth of a real receiver will blur out the high order diffraction fringes. The result is that fine structure smaller than

$$\sqrt{(\text{fractional bandwidth})} \times (\text{width of first Fresnel zone})$$

is lost. Thus, with a 1% bandwidth, the limit of resolution due to bandwidth is of the order of 1″ arc.

(iii) In the derivation of the Fresnel diffraction pattern (1) it is assumed that the pattern subtends a very small angle. The small angle approximation is valid up to diffraction fringes much narrower than any that are likely to be used in observations of lunar occultations. However, it might appear from the analysis of this paper that one could construct an aerial capable of high resolution by placing a tin sheet (as diffracting edge) a few feet over a dipole, and it is the failure of the small angle approximation that prevents such an arrangement from yielding high resolution. (This point arose in discussion with Dr. R. F. Mullaly.)

Thus it appears that one may attain a limit of resolution of 1″ instead of 10″ by applying the restoration process described in this paper. In most cases, the limit is likely to be reached before 1″ resolution is attained, owing to inadequate signal-to-noise ratio.

I understand from Dr. Hazard that he intends to try out the restoration process on one or two of his observations.

VIII. Acknowledgments

This work was carried out during the tenure of a C.S.I.R.O. research fellowship. I am indebted to Professor R. N. Bracewell for his careful perusal of the manuscript; his comments have led to substantial improvements in the paper.

IX. References

Hazard, C. (1962).—*Mon. Not. R. Astr. Soc.* (in press).

Lighthill, M. J. (1958).—"Introduction to Fourier Analysis and Generalized Functions." (Cambridge Univ. Press.)