THE THEORETICAL PRECISION WITH WHICH AN ARBITRARY RADIATION-PATTERN MAY BE OBTAINED FROM A SOURCE OF FINITE SIZE

By P. M. WOODWARD, B.A., and J. D. LAWSON, B.A.

(*The paper was first received* 17*th July*, 1947, *and in revised form* 4*th March*, 1948.)

SUMMARY

It appears that it is possible to approximate as closely as desired to a specified radiation-pattern by a suitable distribution of field over an aperture of given size, though the necessary currents in the conducting elements of the source would in general be prohibitively large in comparison with the power radiated. The difficulty of obtaining a high degree of approximation, and in particular a power gain very much greater than that of a uniformly illuminated aperture, is thus a practical rather than a theoretical one. The same is true for the linear array of given length as for the continuous aperture if no limit is set to the number of elements. Even when this number is limited by the adoption of half-wavelength spacing, the broadside power gain is not a maximum when the amplitudes and phases of the elements are equal, unless the elements are ideal isotropic point-sources.

---

(1) INTRODUCTION

In a remarkable paper[1] on the theory of linear arrays, S. A. Schelkunoff has shown that the power gain obtainable from an aerial array of given length may be increased indefinitely, provided that the number of elements in the array is correspondingly increased. Until the paper appeared, and indeed since, it has been widely believed that the power gain of an aerial system is theoretically limited by its size. In particular, it has been held that the greatest power gain in the direction normal to a plane radiating "aperture" is obtained by means of an equi-phase field distribution of uniform amplitude across the aperture. This belief probably arose in the first instance from a misunderstanding of the role of those waves whose phase gradient across the aperture is even greater than that required to swing the principal direction of radiation away from the normal and into the aperture plane itself, i.e. greater than 2π radians per free-space wavelength. It is the purpose of the present paper to set out in mathematical form an exact two-dimensional theory of aperture-distributions and, by including these waves, to suggest the theoretical possibility not only of unlimited power gain in any given direction, but of producing, from an aerial system of given finite size, a radiation pattern approximating to any specified shape as closely as desired. The treatment differs essentially from Schelkunoff's, and it may be applied equally well to continuous source-distributions of field as to arrays of discrete source elements.

(2) GENERAL DEFINITION OF AN APERTURE IN AERIAL PROBLEMS

The term "aperture" is now widely used to denote the effective radiating area of a directional aerial, such as the mouth of a paraboloid reflector, or even the area covered by a broadside array of dipoles. Since the word is often used where no aperture in the physical sense exists, it is worth attempting a generalized definition.

Consider the ideal situation in which an infinite plane divides space into two regions, one being completely "free," while the other contains an electromagnetic source which is radiating power into the free half. Some power will, in general, flow across this plane at all points, diminishing to zero at very large distances from the source. Many practical sources, however, are such that the position of the plane may be so chosen that the power flux across it is zero (or substantially so) except over some finite region of the plane. Under these circumstances, that part of the plane through which the flow of power is confined may be termed an aperture even though no physical hole exists.

(3) AN ANGULAR SPECTRUM OF INFINITE PLANE WAVES

A quite general field in the free half of space bounded by the aperture-plane may be expressed as a combination of infinite plane-waves, provided that other than purely real directions of propagation are included. Whilst these other directions are included here, simplifications and idealizations of a different kind will nevertheless be made.

Let rectangular co-ordinates be chosen (Fig. 1) such that O*yz*

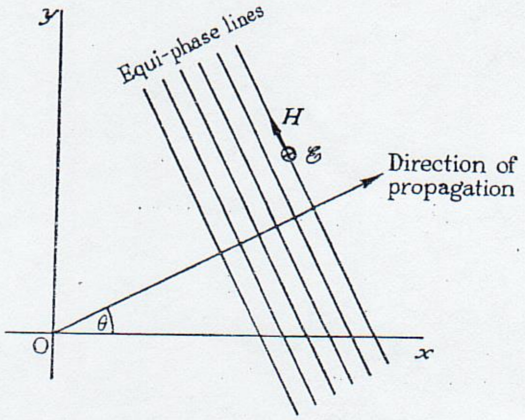

Fig. 1.—Co-ordinate system.
The *z*-axis is perpendicular to the plane of the paper.

represents the aperture-plane and the positive *x*-direction belongs to the free half of space. Only fields which do not vary with *z* will be considered, the entire analysis thus being rendered two-dimensional. For further simplicity, suppose that one of the field vectors is everywhere parallel to O*z*. The electric vector $\mathscr{E}$ will be so chosen, though symmetry between $\mathscr{E}$ and H enables Z_0H to be interchanged with $\mathscr{E}$ throughout, where Z_0 is the intrinsic impedance of free space.

An infinite plane-wave travelling in the direction making an angle θ with the axis of x may now be specified by the equations

$$\mathscr{E}_z\left(\frac{x}{\lambda}, \frac{y}{\lambda}\right) = -A \exp\left[-2\pi j\left(\frac{x}{\lambda}\cos\theta + \frac{y}{\lambda}\sin\theta\right)\right] \quad . \quad (1)$$

$$Z_0H_x\left(\frac{x}{\lambda}, \frac{y}{\lambda}\right) = -A \exp\left[-2\pi j\left(\frac{x}{\lambda}\cos\theta + \frac{y}{\lambda}\sin\theta\right)\right]\sin\theta \quad . \quad (2)$$

Radio Section paper. (U.D.C. 621.396.677)
Written communications on papers published without being read at meetings are invited for consideration with a view to publication.
Messrs. Woodward and Lawson are at the Telecommunications Research Establishment, Ministry of Supply.

$$Z_0H_y\left(\frac{x}{\lambda},\frac{y}{\lambda}\right) = A \exp\left[-2\pi j\left(\frac{x}{\lambda}\cos\theta + \frac{y}{\lambda}\sin\theta\right)\right]\cos\theta \quad . \quad (3)$$

omitting the time factor. A is an arbitrary complex amplitude and phase factor; λ is the wavelength, $\mathscr{E}_x$, $\mathscr{E}_y$, and H_z are everywhere zero.

Subject to the above limitations, a quite general field may be obtained by taking a combination of plane waves of different relative amplitudes, phases and directions of propagation. The dependence of amplitude and phase on direction will be denoted by $P(\theta)d\theta$, which now replaces the constant A. The non-zero components of $\mathscr{E}$ and H in this combination are

$$\mathscr{E}_z = -\int P(\theta) \exp\left[-2\pi j\left(\frac{x}{\lambda}\cos\theta + \frac{y}{\lambda}\sin\theta\right)\right]d\theta \quad . \quad (4)$$

$$Z_0H_x = -\int P(\theta) \exp\left[-2\pi j\left(\frac{x}{\lambda}\cos\theta + \frac{y}{\lambda}\sin\theta\right)\right]\sin\theta d\theta \quad . \quad (5)$$

$$Z_0H_y = \int P(\theta) \exp\left[-2\pi j\left(\frac{x}{\lambda}\cos\theta + \frac{y}{\lambda}\sin\theta\right)\right]\cos\theta d\theta \quad . \quad (6)$$

The limits of θ require careful consideration. Since we are concerned with positive values of x only, real values of θ will be confined to $-\pi/2 < \theta < \pi/2$. Values of θ lying in the other two quadrants would correspond to plane waves travelling in from infinity, and are excluded. For generality, however, certain complex values do need inclusion. These must be selected in such a way that the corresponding plane-wave does not tend exponentially to infinity either with positive x or with positive or negative y, i.e. the coefficient of x in the expression

$$-2\pi j\left(\frac{x}{\lambda}\cos\theta + \frac{y}{\lambda}\sin\theta\right) \quad . \quad . \quad . \quad . \quad (7)$$

must not be allowed to have a positive real part, nor must the coefficient of y have any real part, positive or negative. We therefore include only the complex values

$$\theta = \frac{\pi}{2} + \text{positive imaginary terms,}$$

$$\theta = -\frac{\pi}{2} + \text{negative imaginary terms,}$$

the contour of integration, C, being as shown in Fig. 2. These waves decrease exponentially in the positive x-direction. Further,

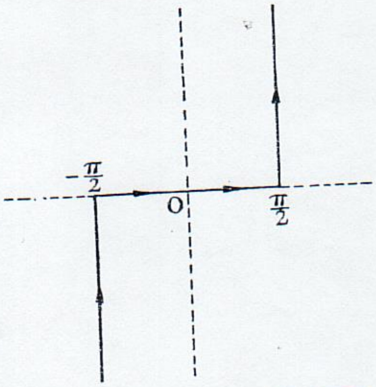

Fig. 2.—Contour of integration in the complex θ-plane.

the components of $\mathscr{E}$ and H in the Oyz-plane are in quadrature, so that there is no associated mean flow of power across this plane, though there is a continuous flow of power parallel to it.* Such types of wave, for which $\sin\theta$ is real but numerically greater than unity, are well known. They occur, for instance, outside a dielectric boundary at which total internal reflection is occurring ($\sin\theta$ being given as greater than unity by Snell's Law). Another well-known example is a cut-off waveguide carrying the H_{01} mode, in which two waves travel at angles $\pm$ arc sin $(\lambda/2d)$ with the walls of the guide, where d is the dimension parallel to H.

Having set up a field sufficiently general for the purposes in hand by superimposing infinite plane-waves of angular spectral density $P(\theta)$, it is of interest to discover the physical interpretation of this factor.

(4) THE ANGULAR SPECTRUM AS A RADIATION PATTERN

In order to identify $P(\theta)$, it is necessary to explore the field at large distances from the origin by deriving the asymptotic expansions for the integrals (4), (5) and (6). It will be sufficient to write down only one of them, $\mathscr{E}_z$ being the most convenient.

Let polar co-ordinates r, ϕ be chosen with the previous origin so that

$$x = r\cos\phi, \quad y = r\sin\phi.$$

Then
$$\mathscr{E}_z\left(\frac{r}{\lambda},\phi\right) = -\int_C P(\theta)\exp\left[-2\pi j\frac{r}{\lambda}\cos(\theta-\phi)\right]d\theta \quad . \quad (8)$$

If r is large, this is of a form suitable for approximate evaluation by the method of stationary phase.[2,3] Briefly, the important part of the integrand lies in the neighbourhood of $\theta = \phi$, since values of θ which differ significantly from ϕ make the rate of change of $r\cos(\theta - \phi)$ with θ very large along the real axis. The integrand is therefore highly oscillatory on the real axis and, except near $\theta = \phi$, cancels itself out under the integral sign† for sufficiently large values of r. Along those arms of the contour of integration which run parallel with the imaginary axis, the integrand is almost completely damped out when r is large. In the neighbourhood of $\theta = \phi$, $P(\theta)$ may be replaced by $P(\phi)$ and brought outside the integral.

Thus
$$\mathscr{E}_z\left(\frac{r}{\lambda},\phi\right) \simeq -P(\phi)\int_{-\pi/2}^{\pi/2}\exp\left[-2\pi j\frac{r}{\lambda}\cos(\theta-\phi)\right]d\theta \quad . \quad (9)$$

The remaining integral may now be evaluated by expanding the cosine,

$$\mathscr{E}_z\left(\frac{r}{\lambda},\phi\right) \simeq -P(\phi)\exp\left(-2\pi j\frac{r}{\lambda}\right)\int_{-\infty}^{\infty}\exp\left[\pi j\frac{r}{\lambda}(\theta-\phi)^2\right]d\theta \quad . \quad (10)$$

The fact that the integration is now performed beyond the region of validity of the expansion of the cosine is of no concern again owing to the oscillations of the integrand. Eqn. (10), which involves a standard definite integral, reduces to

$$\mathscr{E}_z\left(\frac{r}{\lambda},\phi\right) \simeq -\frac{P(\phi)\exp\left(j\frac{\pi}{4} - 2\pi j\frac{r}{\lambda}\right)}{\sqrt{(r/\lambda)}} \quad . \quad . \quad (11)$$

The function P gives the dependence of the field, at large distances from the origin, on bearing from the origin, and may therefore be identified as a conventional angular radiation-pattern or polar diagram.

* Perhaps surprisingly, this does not necessarily imply a source at infinity, because there is a continuous range of values of θ for which power is flowing parallel to Oy and a suitable choice of $P(\theta)$ could effect complete cancellation at infinity.

† Certain restrictions on the function P are implicit in this argument. If, however, the field is such as can be produced by a source of finite size these restrictions are automatically satisfied. (They would not be satisfied for a single infinite plane-wave, which makes P an impulse function.)

(5) THE RELATION BETWEEN RADIATION PATTERN AND APERTURE-DISTRIBUTION

In Fig. 1, Oyz represents the aperture-plane. The two tangential field components in this plane are, from eqns. (4) and (6),

$$\mathscr{E}_z\left(0, \frac{y}{\lambda}\right) = - \int_C P(\theta) \exp\left(-2\pi j \frac{y}{\lambda} \sin\theta\right) d\theta \quad . \quad (12)$$

$$Z_0 H_y\left(0, \frac{y}{\lambda}\right) = \int_C P(\theta) \exp\left(-2\pi j \frac{y}{\lambda} \sin\theta\right) \cos\theta d\theta \quad . \quad (13)$$

From this point onwards, it is convenient to omit the indication that x is zero in the aperture-plane, and to express the radiation pattern in the form $p(\sin\theta)$ instead of $P(\theta)$, since $\sin\theta$ rather than θ is the significant angular variable. The functions P and p are different, but the dependence on θ is the same. The contour of integration in the $\sin\theta$ plane is merely the real axis. Thus,

$$\mathscr{E}_z\left(\frac{y}{\lambda}\right) = - \int_{-\infty}^{\infty} \frac{p(\sin\theta)}{\cos\theta} \exp\left(-2\pi j \frac{y}{\lambda} \sin\theta\right) d(\sin\theta) \quad . \quad (14)$$

$$Z_0 H_y\left(\frac{y}{\lambda}\right) = \int_{-\infty}^{\infty} p(\sin\theta) \exp\left(-2\pi j \frac{y}{\lambda} \sin\theta\right) d(\sin\theta) \quad . \quad (15)$$

whence by Fourier's theorem are obtained the well-known equations

$$p(\sin\theta) = -\cos\theta \int_{-\infty}^{\infty} \mathscr{E}_z\left(\frac{y}{\lambda}\right) \exp\left(2\pi j \frac{y}{\lambda} \sin\theta\right) d\left(\frac{y}{\lambda}\right) \quad . \quad (16)$$

and

$$p(\sin\theta) = Z_0 \int_{-\infty}^{\infty} H_y\left(\frac{y}{\lambda}\right) \exp\left(2\pi j \frac{y}{\lambda} \sin\theta\right) d\left(\frac{y}{\lambda}\right) \quad . \quad (17)$$

Eqns. (14) and (15) are often used as a guide for the design of aerial systems to have prescribed radiation patterns in one plane. As they stand, however, the equations are too idealized, since substitution of an arbitrary function $p(\sin\theta)$ will give distributions of $\mathscr{E}$ and H which extend between $y/\lambda = \pm\infty$. If an additional postulate, such as

$$H_y(y/\lambda) = 0; \quad |y/\lambda| > \tfrac{1}{2}W \quad . \quad . \quad . \quad . \quad (18)$$

be made, the mathematics becomes more realistic, as this condition is a way of ensuring that the field is compatible with a source of finite size. But eqn. (18) cannot be satisfied for an arbitrary function $p(\sin\theta)$ in eqn. (15). As a result, it has often been supposed that it is impossible to find a field-distribution over a finite aperture which will exactly reproduce a specified radiation-pattern. Yet it must be remembered that the real radiation-pattern extends only from $\sin\theta = -1$ to $+1$. In what is conveniently termed the imaginary region of the pattern, for which $|\sin\theta| > 1$, any functional form may be assigned to $p(\sin\theta)$ without affecting the distant radiation field. There is therefore a considerable choice of distributions of H over the aperture plane which will produce the same radiated field. It is not now suggested that one such choice will exactly satisfy condition (18), but what the authors do suggest is that fields can be constructed which both satisfy eqn. (18) and approximate as closely as desired to a given $p(\sin\theta)$ in the range of real angles. The distinction is significant as the actual limit would not appear to be attainable, owing to lack of convergence of $p(\sin\theta)$ in the imaginary region and consequent infinite amplitudes in the aperture.

The foregoing remarks will become clearer in the light of the examples to be given in the following Section.

(6) A SYNTHESIS THEOREM AND ITS IMPLICATIONS

"By suitably distributing the field over an aperture of given finite width, specified values can be assigned to the radiation pattern in any finite number of directions."

The proof rests on two facts. First, radiation patterns and their corresponding aperture-distributions may be added linearly, and secondly, it is possible to define an unlimited number of linearly independent radiation-patterns and associated field-distributions over a given finite aperture. (One might, for example, divide the aperture into n sections, and illuminate each separately.) Let the radiation pattern be specified in n directions $\sin\theta_r$. Then if $p_s(\sin\theta)$ are n linearly independent patterns obtainable from the given aperture, a composite pattern having the required values is

$$p = \sum^{s} A_s p_s \quad . \quad . \quad . \quad . \quad . \quad . \quad (19)$$

where the constants A_s are determined from the simultaneous equations

$$p(\sin\theta_r) = \sum^{s} A_s p_s(\sin\theta_r), \; r = 1, 2, \ldots n \quad . \quad (20)$$

This result would be of but little value were it not possible to infer the behaviour of the composite radiation-pattern between the directions in which values are assigned, a behaviour which depends on the choice of the component patterns. In this connection, one should again observe that the pattern is of interest only over the finite range of $\sin\theta$ corresponding to real angles θ, and that it is therefore possible to choose the directions in which the pattern is specified at as close a spacing as desired. As the number of directions and specified values is increased, and the spacing decreased, it is not obvious that the composite pattern will converge to a smooth curve over the range of real angles; in fact it might fail to converge altogether between the assigned values. Whilst no proof of proper convergence has yet been devised, intuition, together with examples such as those following, would indicate that trouble lies not in the real range but outside it. The more one attempts to force the radiation pattern to converge to an arbitrarily assigned functional form by fixing more and more values within the real range, the greater does the pattern become in the imaginary region. This results in very large reactive fields over the aperture; fields which would probably become infinite, except in isolated cases, if the limiting radiation-pattern were attained.

Fig. 3(*a*) shows a beam 60° wide between zeros, obtained from an aperture one wavelength across. For comparison, the pattern from a uniformly illuminated aperture of the same width is shown dotted. The aperture-distribution (here defined in terms of H_y) which will produce the narrow beam is shown in Fig. 3(*b*), the scale being such that unit field over the aperture produces the dotted pattern in Fig. 3(*a*). If a distribution of -1 were superimposed on this, the radiation pattern would consist of the full curve minus the dotted curve, and the main beam would disappear! The narrow beam was obtained by adding together, suitably weighted, a number of patterns of the form

$$p_s(\sin\theta) = \frac{\sin[\pi W(\sin\theta - s/W)]}{\pi W(\sin\theta - s/W)} \quad . \quad . \quad (21)$$

This expression represents the pattern from an aperture of width $W\lambda$, illuminated uniformly in amplitude but with a linear phase slope amounting to a total slip of $2\pi s$ radians across the aperture. It was derived from eqn. (17). In the present example, W is

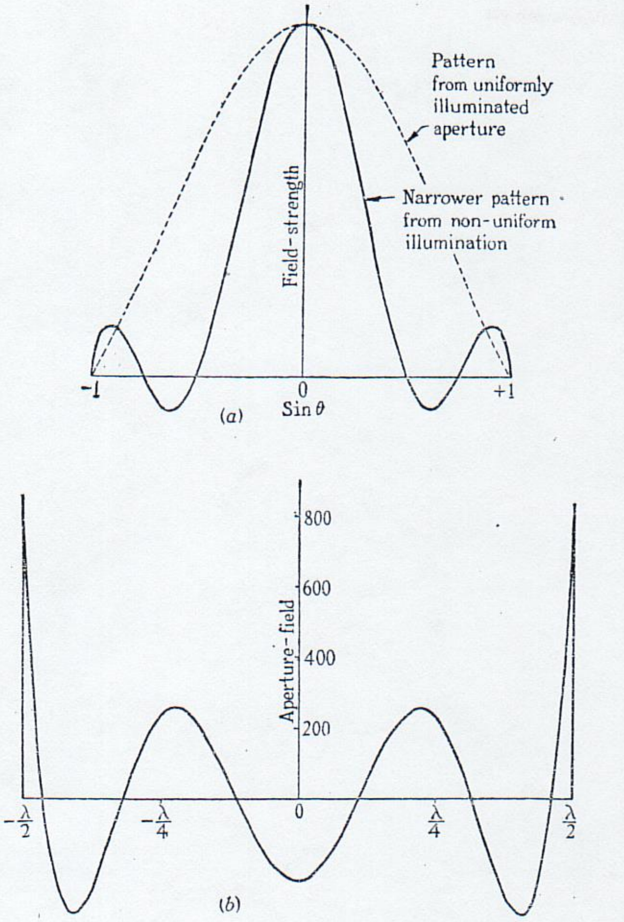

Fig. 3.—Radiation patterns from (*a*) an aperture of width λ, and (*b*) corresponding apperture distribution.

Apperture distribution shown in (*b*) produces the radiation pattern shown by full curve in (*a*). [The vertical scale of (*b*) is such that unit field produces the dotted pattern of (*a*).]

unity. The simultaneous equations (20) were used to force the composite pattern through unity at $\sin\theta = 0$ and through zero at $\sin\theta = \pm\frac{1}{2}, \pm\frac{3}{4}$, and ± 1. Values of s in eqn. (21), were chosen to make s/W correspond to the specified values of $\sin\theta$, a procedure which, judging from the authors' numerical work, appears to avoid any tendency for the composite pattern to oscillate violently between its fixed points. The solution yields the following expression for the resultant pattern:

$$\begin{aligned} p(\sin\theta) = {} & 50\,000{\cdot}00p_0 \\ & - 66\,815{\cdot}60(p_{-\frac{1}{2}} + p_{\frac{1}{2}}) \\ & + 58\,434{\cdot}91(p_{-\frac{3}{4}} + p_{\frac{3}{4}}) \\ & - 16\,736{\cdot}79(p_{-1} + p_1) \quad . \quad . \quad . \quad (22) \end{aligned}$$

The individual functions, p_s, never exceed unity, and in the real range of θ their resultant never exceeds unity. In the imaginary region, the expression has very large values, and it will be shown later how this results in a correspondingly large field over the aperture. By proceeding in a manner similar to this example, and using more equations of the form (20), the beamwidth could be made indefinitely small and the gain indefinitely large, or indeed one might approach any required shape of radiation pattern.

It has been mentioned that the component patterns may be selected quite arbitrarily, provided that they are linearly independent. There is one particular choice which gives a specially clear insight into the problem. It consists of eqns. (21) with integral values of s only. These represent component patterns from individual terms in a Fourier-series representation of the aperture distribution. If the values of the overall pattern $p(\sin\theta)$ be specified at

$$\sin\theta = r/W \quad . \quad . \quad . \quad . \quad . \quad . \quad (23)$$

for positive and negative integral values of r, the simultaneous equations are

$$p(r/W) = \sum^{s} A_s p_s(r/W) \quad . \quad . \quad . \quad . \quad (24)$$

and since

$$p_s(r/W) = \begin{cases} 1, & r = s \\ 0, & r \neq s \end{cases} \quad . \quad . \quad . \quad . \quad (25)$$

the eqns. (24) immediately reduce to

$$p(r/W) = A_r \quad . \quad . \quad . \quad . \quad . \quad (26)$$

The composite pattern is therefore

$$p(\sin\theta) = \sum^{s} p\left(\frac{s}{W}\right)\frac{\sin[\pi W(\sin\theta - s/W)]}{\pi W(\sin\theta - s/W)} \quad . \quad (27)$$

and the corresponding aperture distribution is

$$Z_0 H_y\left(\frac{y}{\lambda}\right) = \begin{cases} \dfrac{1}{W}\sum^{s} p\left(\dfrac{s}{W}\right)\exp\left(-2\pi j s\dfrac{y}{\lambda W}\right), & \left|\dfrac{y}{\lambda}\right| < \frac{1}{2}W \\ 0, & \left|\dfrac{y}{\lambda}\right| > \frac{1}{2}W \end{cases} \quad . \quad . \quad . \quad . \quad (28)$$

as may be verified most simply from eqn. (17). This particular method of synthesizing a specified pattern has been described in an earlier paper.[4] Fig. 4 shows how the resultant pattern is

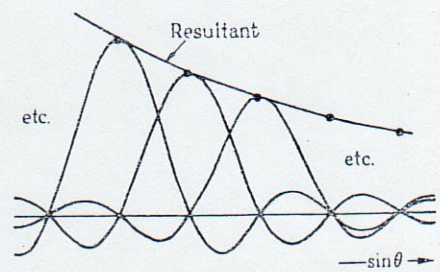

Fig. 4.—Component beams from Fourier-series analysis of aperture distribution.

built up from a series of beams, disposed at a uniform spacing of $1/W$ along the $\sin\theta$ axis. They do not interfere at any of the points where their zeros coincide, so that the single beam which does not have a zero at any one such point has the exact strength of the composite pattern in this direction. It will thus be seen that the behaviour of the radiation pattern from an aperture of width $W\lambda$ is uniquely determined for all values of $\sin\theta$ by its value at intervals $1/W$. The number of such values within the range of real angles is roughly equal to the number of half-wavelengths across the aperture. If it is required, within the range of real angles, to vary the behaviour of the pattern between these values without altering the values themselves, it is necessary to vary the strengths of the component beams lying in the imaginary region, whose side lobes make a small contribution to the pattern in the real range. Detailed structure of a radiation pattern, such as demands the specifying of values at closer spacing than $1/W$ in $\sin\theta$, may thus be obtained only by introducing comparatively large beams in the imaginary region.

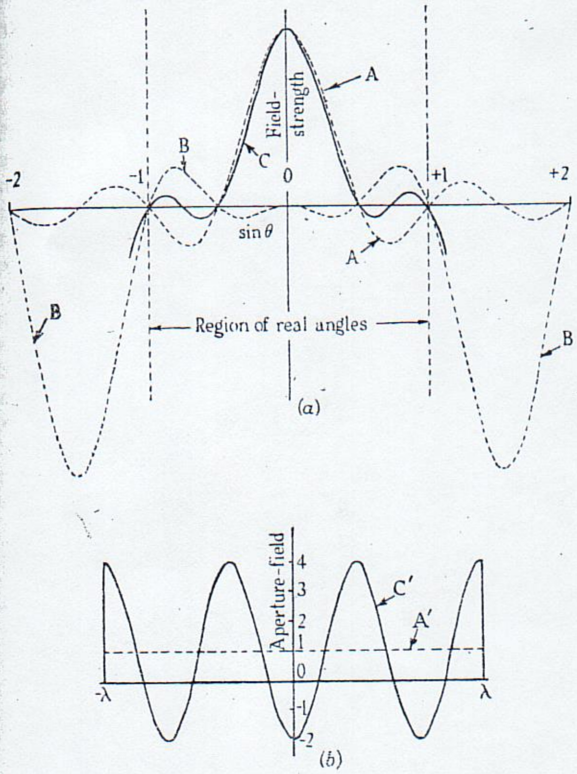

Fig. 5.—Effect of imaginary beams.

(a) Curve A: pattern from uniformly illuminated aperture of width 2λ.
Curve B: sum of two patterns whose main beams lie in the imaginary region.
Curve C: composite pattern, curves A and B.
(b) Aperture distributions corresponding to radiation patterns shown in (a).
Curve A′: aperture distribution to produce radiation pattern of curve A in (a).
Curve C′: aperture distribution to produce radiation pattern of curve C in (a).

Fig. 5(a) illustrates the above. Here, shown dotted, is the pattern from a uniformly illuminated aperture of width 2λ. It is required to eliminate the side lobes without broadening the main beam. To some extent this has been achieved as shown by the full curve, whose composition is

$$p(\sin\theta) = p_0 - 1\cdot 5(p_{-3} + p_3) \quad . \quad . \quad . \quad (29)$$

where p_s is given by eqn. (21) with $W = 2$. Fig. 5(b) shows the corresponding aperture-distributions. Note the considerable increase in aperture field necessary to give the small increase in gain. The reason for this increase may be seen by applying Parseval's formula[5] to eqn. (17), giving

$$\int_{-\frac{1}{2}W}^{\frac{1}{2}W} \left| Z_0 H_y\left(\frac{y}{\lambda}\right)\right|^2 d\left(\frac{y}{\lambda}\right) = \int_{-\infty}^{\infty} |p(\sin\theta)|^2 d(\sin\theta) \quad . \quad (30)$$

i.e. the integrated square of the aperture distribution is equal to the integrated square of the radiation pattern considered as a function of $\sin\theta$. In the above example, reducing the side lobes has slightly decreased $\int|p\sin\theta)|^2 d(\sin\theta)$ in the range $|\sin\theta| < 1$, but has increased it very considerably outside this range. It is interesting to note that if the aperture distribution is produced by currents in a conducting sheet, then the loss due to finite conductivity of the sheet is proportional to the left-hand side, and hence to the right-hand side of eqn. (30).

Eqn. (30) is not intended to be a power relation. In order to equate mean power leaving the aperture to the power flow at large distances from the aperture, Parseval's formula should be applied jointly to eqns. (16) and (17), giving

$$-\tfrac{1}{2}\int_{-\frac{1}{2}W\lambda}^{\frac{1}{2}W\lambda} \mathscr{E}_z H_y^* dy = \frac{\lambda}{2Z_0}\int_{-\infty}^{\infty} \frac{|p(\sin\theta)|^2}{\cos\theta} d(\sin\theta) \quad . \quad (31)$$

Taking real parts of both sides,

$$-\tfrac{1}{2}\mathscr{R}\int_{-\frac{1}{2}W\lambda}^{\frac{1}{2}W\lambda} \mathscr{E}_z H_y^* dy = \frac{\lambda}{2Z_0}\int_{-\pi/2}^{\pi/2} |p(\sin\theta)|^2 d\theta \quad . \quad . \quad (32)$$

since $\cos\theta$ is purely imaginary when $\sin\theta$ is numerically greater than unity. The left-hand side of eqn. (32) is the real part of the complex Poynting vector resolved normally to the aperture plane and integrated over it. It represents the mean power leaving the aperture. The right-hand side of the equation is the mean power flow integrated over the polar diagram. The imaginary part of eqn. (31) is also of physical significance. It reads

$$-\tfrac{1}{2}\mathscr{I}\int_{-\frac{1}{2}W\lambda}^{\frac{1}{2}W\lambda} \mathscr{E}_z H_y^* dy = -\frac{j\lambda}{2Z_0}\left[\int_{-\infty}^{-1} + \int_{1}^{\infty}\right] \frac{|p(\sin\theta)|^2}{\sqrt{(\sin^2\theta - 1)}} d(\sin\theta). \quad (33)$$

The left-hand side represents twice the angular frequency times the difference of the mean values of the magnetic and electric energies in the whole of the free half of space in front of the aperture-plane.[6] This difference is associated with energy stored in the vicinity of the aperture, as distinct from energy travelling outwards to infinity. That it will be large for radiation patterns which are large in their imaginary region may at once be seen from the right-hand side of the equation. In order that the aerial system giving rise to the aperture distribution should present a purely resistive load to the generator, a large balancing reactance would have to be incorporated behind the aperture plane, thus forming a highly resonant and therefore frequency-sensitive arrangement.

It may be concluded that a required radiation pattern in one plane may be specified in roughly as many directions as there are half-wavelengths across the aperture, uniformly spaced with respect to the sine of the angle with the normal to the plane of the aperture, without an excessively reactive aperture distribution. An unlimited number of further values may be specified, but the field over the aperture must then become very large. In particular, a radiation pattern giving a directivity very much greater than may be obtained from the uniformly illuminated aperture is always associated with a large aperture field. At a large distance from the aperture, even in the principal direction of radiation, the resultant effect is one of almost complete cancellation, the radiation fields from each portion of the aperture nearly cancelling each other out. If the radiation pattern is symmetrical about the direction normal to the aperture, the resulting aperture distribution, which is then purely real, contains portions of large equi-phase field almost balanced by large anti-phase portions. Yet it should be noted that even of purely equi-phase aperture distributions, uniform amplitude does not give the greatest directivity. For consider (Fig. 6) an aperture of width $\frac{1}{2}\lambda$: discrete elements at the extremities of the aperture are more directive than the uniform distribution.

(7) GENERALITY OF TWO-DIMENSIONAL TREATMENT

The purely two-dimensional treatment to which this paper has been confined is a mathematical idealization, for the entire field has been assumed uniform to infinity in the direction Oz (Fig. 1). One consequence is that at large distances from the source, the field falls off inversely as the square root of the distance

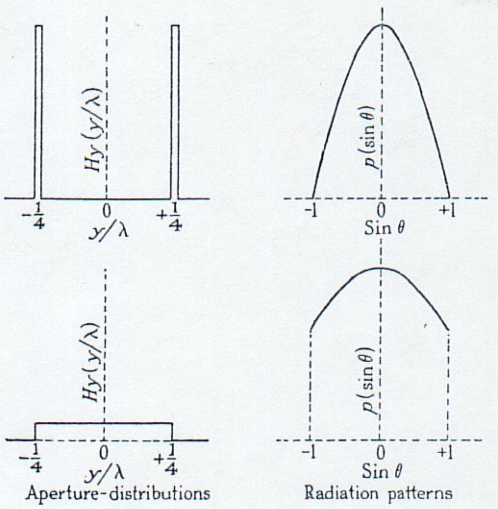

Fig. 6.—Comparison of directivities from equi-phase aperture-distributions.

(Section 4), whereas a finite source in three dimensions produces a field falling off as the inverse of distance. The shape of the radiation pattern in the xy-plane of an aerial radiating in three dimensions, however, is independent of the aperture distribution measured parallel to Oz, provided that the overall aperture distribution can be expressed as a function of y only multiplied by a function of z only. Consequently the two-dimensional analysis gives one cross-section of the three-dimensional polar diagram. Further, there is no reason why high gains should not be obtained in three dimensions, because the synthesis theorem (Section 6) may equally well be applied to three dimensions as to two.

(8) LINEAR ARRAYS OF DISCRETE ELEMENTS

Suppose that the aperture is defined, not as a continuous range but as a series of N discrete elements. It is then possible to produce only N linearly independent aperture distributions, and consequently only N points may be assigned to the radiation pattern. Schelkunoff[1] gives a very neat method of forcing a pattern from N elements through one finite and $N-1$ zero values, and shows how to keep the zero values in the region of real angles as the distance between the elements composing the array is decreased. By placing a large number of elements close together, very narrow beams can then be obtained from small aerials, and at the same time the side-lobe level can be kept extremely low. The pattern from a linear array may also be synthesized on lines very similar to those elaborated in Section 6 for the continuous aperture. It will suffice to draw attention to the single important distinction. The pattern from an array whose elements are equally spaced at intervals d is

$$p(\sin\theta) = \sum^{n} A_n \exp\left(2\pi j \frac{nd}{\lambda} \sin\theta\right) \quad . \quad . \quad . \quad (34)$$

where A_n is a complex number proportional to the amplitude and phase of the nth element. The above is a repeating function of $\sin\theta$ of period λ/d. Control of the pattern is thus effectively confined to a finite range, λ/d, of $\sin\theta$. For an array of N elements, this limitation is precisely equivalent to there being only N degrees of freedom, because there is room for only N component beams of the type illustrated in Fig. 4 within one period of the resultant pattern. It should also be observed that a single period covers the whole range of real angles only if the spacing between elements is $\frac{1}{2}\lambda$ or less. If the spacing is more than $\frac{1}{2}\lambda$, secondary beams may be produced because the pattern partly (if not wholly) repeats itself within the real range.

The special case of $\frac{1}{2}\lambda$ spacing is of some importance. It has commonly been held that[7] the greatest power gain of such an array, when used as a broadside, is obtained by feeding the elements with equal amplitudes and phases. Without qualification, this statement is fallacious, being based on the highly idealized assumption of fully isotropic sources in three dimensions. The authors have performed the simple calculation (see Appendix) for the power gain of a 3-element, $\frac{1}{2}\lambda$ spaced array, first of isotropic line-sources in two dimensions (for pedagogic reasons), and secondly of parallel Hertzian doublets in three dimensions pointing in a direction perpendicular to the length of the array, and have found in each case that the greatest broadside power-gain is obtained when the central element is fed with a slightly larger amplitude than the outer ones. That the broadside power gain of a $\frac{1}{2}\lambda$ spaced array of fully isotropic point-sources is a maximum when the amplitudes and phases of its elements are uniform may be deduced by noting first that eqn. (34) holds in three dimensions if θ is the angle of latitude in spherical

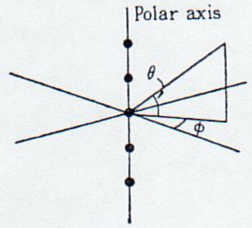

Fig. 7.—Special co-ordinate system suitable for isotropic point-sources.

co-ordinates with polar axis along the length of the array (Fig. 7). The total power radiated in one-half of space is proportional to

$$\int_{-\pi/2}^{\pi/2}\int_{-\pi/2}^{\pi/2} |p(\sin\theta)|^2 \cos\theta \, d\theta \, d\phi$$

where ϕ is longitude. This is proportional to

$$\int_{-1}^{1} |p(\sin\theta)|^2 d(\sin\theta) \quad . \quad . \quad . \quad . \quad . \quad (35)$$

Parseval's theorem applied to eqn. (34) states that

$$\int_{-\lambda/2d}^{\lambda/2d} |p(\sin\theta)|^2 d(\sin\theta) = \frac{\lambda}{d}\Sigma|A_n|^2 \quad . \quad . \quad . \quad (36)$$

the integral having been taken over one period of $p(\sin\theta)$. For $\frac{1}{2}\lambda$ spacing this is identical with expression (35). Power density in the direction $\theta = 0$ being proportional to $|\Sigma A_n|^2$, the gain is proportional to $|\Sigma A_n|^2/\Sigma|A_n|^2$. This is a maximum when all the values of A_n are equal, as may be verified by regarding the numbers A_n as vectors in the complex plane, maximizing the numerator with the denominator held fixed, and then minimizing the denominator while holding the numerator fixed.

It is perhaps worth remarking that the above proof holds for arrays spaced at any integral multiple of $\frac{1}{2}\lambda$.

(9) CONCLUSION

It has been seen that, in dealing with radiation problems, $\sin\theta$ rather than θ is in many ways the significant angular variable, and that, far from being a mere mathematical fiction, complex values of θ corresponding to values of $\sin\theta$ numerically greater than unity have a definite physical interpretation. In accordance with this idea, a radiation pattern $p(\sin\theta)$ may be treated as a function which exists over the entire range of real values of $\sin\theta$, though only that portion of the function whose argument lies in the interval $(-1, 1)$ may be interpreted in the conventional manner as energy travelling outwards to infinity from the source. If the source is such that radiation is virtually constrained to pass through a plane aperture of width $W\lambda$, and the distribution of field over the aperture can be varied at will, it is conveniently possible to assign arbitrary values to the radiation pattern in directions spaced at intervals $1/W$ in $\sin\theta$. There are roughly as many such real directions as there are half-wavelengths across the aperture. If directions for which $|\sin\theta| > 1$ are also included, the behaviour of the pattern may be controlled between the specified real directions. Such control is so weak that, to be effective, very large reactive aperture-fields are necessary. Indeed, convergence to an arbitrary radiation-pattern, the width of aperture being held fixed, is generally associated with divergence of the aperture field, making the absolute limit unattainable.

Similar considerations apply to linear arrays of finite length as to continuous apertures, except that an array of N elements has only N degrees of freedom and only N values may be assigned to the pattern. Convergence to an arbitrary radiation-pattern can therefore be achieved only by increasing the number of elements, and if the total length of the array is held constant in the process the currents in the elements become large as soon as their separation falls below $\frac{1}{2}\lambda$.

In practice, no startling improvements over the conventional type of aerial are likely, since even for small increases in performance (Fig. 5) the currents in the conducting elements of the source become very large. Besides the increased copper-loss thus incurred, the shape of the radiation pattern and the impedance of the aerial would be extremely sensitive to small changes of frequency, and the manufacturing tolerances for any practical aerial could become prohibitive. If, for instance, we decrease the amplitude distribution shown in Fig. 3(*b*) by one-fifth of the thickness of the line in the figure, the main beam would disappear entirely. These examples, however, may give too pessimistic a picture of the practical possibilities. End-fire arrays have their main beams near the imaginary region, and so may be controlled more readily than broadside arrays. F. K. Goward has shown[8] how to add two beams with $\sin\theta$ slightly greater than unity to produce a worthwhile improvement over the conventional end-fire array. This improved array would be difficult to make, though perhaps not impossible.

The authors do believe, however, that it is of some practical importance to understand clearly the theoretical nature of the limitations imposed on radiation patterns from sources of finite size, and that the method of apertures is as convenient a way as any of incorporating these limitations in a mathematical theory.

(10) ACKNOWLEDGMENT

The paper is published by permission of the Controller, H.M. Stationery Office. Crown copyright is reserved.

(11) REFERENCES

(1) SCHELKUNOFF, S. A.: "A Mathematical Theory of Linear Arrays," *Bell System Technical Journal*, 1943, **22**, p. 80.
(2) WATSON, G. N.: "A Treatise on Bessel Functions," (Cambridge University Press, 1944, Section 8.2).
(3) JEFFREYS, H., and JEFFREYS, B.: "Methods of Mathematical Physics" (Cambridge University Press, 1946, Section 17.05).
(4) WOODWARD, P. M.: "A Method of Calculating the Field over a Plane Aperture Required to produce a Given Polar Diagram," *Journal I.E.E.*, 1947, **93**, Part IIIA, p. 1554.
(5) TITCHMARSH, E. C.: "Introduction to the Theory of Fourier Integrals" (Clarendon Press, 1937, Section 2.1).
(6) STRATTON, J. A.: "Electromagnetic Theory" (McGraw-Hill Book Co., 1941), Section 2.20.
(7) DOLPH, C. L.: "A Current Distribution for Broadside Arrays which Optimizes the Relationship between Beam Width and Side-Lobe Level," *Proceedings of the Institute of Radio Engineers*, 1946, **34**, p. 335. (See top of p. 338.)
(8) GOWARD, F. K.: "An Improvement in End-Fire Arrays," *Journal I.E.E.*, 1947, **94**, Part III, p. 415.
(9) RIBLET, H. J.: "Note on the Maximum Directivity of an Antenna," *Proceedings of the Institute of Radio Engineers*, 1948, **36**, p. 620.

(12) APPENDIX

Power Gain of $\frac{1}{2}\lambda$-Spaced Elements

Two examples will be given to illustrate that the broadside power gain of a linear array whose elements are separated by $\frac{1}{2}\lambda$ is not necessarily a maximum when the amplitudes and phases are equal. The first example is two-dimensional and is the case of three isotropic line-sources. The second and more practical example is three-dimensional and consists of three Hertzian dipoles. As has been shown in Section 8, it is only for isotropic $\frac{1}{2}\lambda$-spaced point-sources that the gain is a maximum with uniform amplitudes and phases.

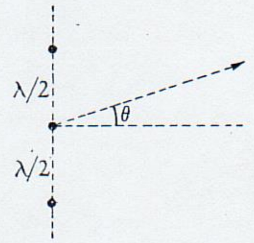

Fig. 8.—Linear array of three isotropic line-sources.

Fig. 8 represents a cross-section of three isotropic line-sources. Let the outer elements carry unit current and the central element a current I, all the currents being in phase. Then the radiation pattern is

$$p(\sin\theta) = 2\cos(\pi\sin\theta) + I.$$

The broadside power gain is

$$G = \frac{2\pi|p(0)|^2}{\int_0^{2\pi}|p(\sin\theta)|^2 d\theta}$$

Evaluation in terms of the Bessel function J_0 gives

$$G = (2 + I)^2/[2 + I^2 + 2J_0(2\pi) + 4IJ_0(\pi)]$$

This has a minimum when $I = -2$ and a maximum when

$$I = [2J_0(\pi) - J_0(2\pi) - 1]/[J_0(\pi) - 1] \simeq 1\cdot 40.$$

The gain for the uniform case, $I = 1$, is $4\cdot 05$, whereas the gain when $I = 1\cdot 40$ is $4\cdot 29$, an improvement of 6%.

Fig. 9 represents three Hertzian dipoles at $\frac{1}{2}\lambda$-spacing along Ox, carrying equi-phase currents $(1, I, 1)$ as in the first example. Let Oz be taken as the axis of spherical polar co-ordinates. The radiation pattern of each element is $\sin\theta$, whence the overall pattern of the array is

$$p(\sin\phi, \sin\theta) = [2\cos(\pi\sin\phi\sin\theta) + I]\sin\theta.$$

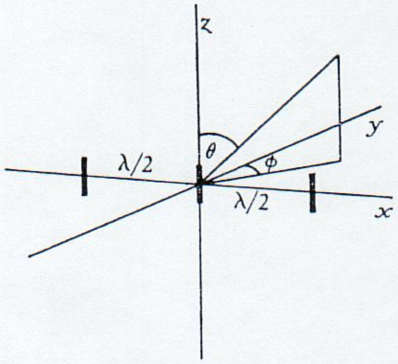

Fig. 9.—Linear array of three Hertzian dipoles.

The broadside gain is

$$G = \frac{4\pi|p(0, 1)|^2}{\int_0^{\pi}\int_0^{2\pi}|p(\sin\phi, \sin\theta)|^2 \sin\theta d\theta d\phi}$$

On carrying through the integrations, one obtains

$$G = [\pi(2 + I)^2]\Big/\left(\frac{2\pi}{3}I^2 - \frac{4}{\pi}I + \frac{4\pi}{3} + \frac{1}{2\pi}\right)$$

There is a minimum when $I = -2$ and a maximum when

$$I = (27 + 8\pi^2)/(12 + 8\pi^2) \simeq 1\cdot165$$

The gain for uniform feeding, $I = 1$, is $5\cdot47$, but the gain when $I = 1\cdot165$ is $5\cdot51$. The improvement is less than 1% and hence of merely theoretical interest.

DISCUSSION ON "IMPULSIVE INTERFERENCE IN AMPLITUDE-MODULATION RECEIVERS"*

Mr. C. C. Eaglesfield (*communicated*): The paper is marred by the continual use of the term "bandwidth." While the concept of bandwidth is useful in a rough engineering way, it is frequently misleading when applied to a general investigation.

The difficulty is to define the extent of the band in such a way that it is extricable from the general equations. This has led in the past to the postulation of "ideal filters," but, as is well known, this is indefensible mathematically. The transmission factor of a network is a complex number, and, with certain limitations, convenient assumptions about this complex number as a function of frequency may be made. But it does not seem possible to find any assumption which leads to convenient amplitude and phase characteristics.

The implicit assumption made by the author for the band-pass filter is that the amplitude characteristic is everywhere zero except between two frequencies which he terms the cut-off frequencies. (The difference between these frequencies he defines as the pass band.) This assumption is not tenable, so that his treatment has an appearance of rigour which will not deceive the mathematician.

In any case, the engineering interpretation is obscure. In practice the cut-off frequencies must be associated with large attenuation, but the question is: How much attenuation and with reference to what point on the band? The usual practice is to take as the pass band the frequency range over which the variations of attenuation are comparatively small.

The usefulness of formulae containing bandwidth, such as those given in the paper, is due to the similar characteristics of the filters used in practice, which may be said to differ mainly in frequency scale.

One of the main deductions that the author makes from his equations is that an increase of filter bandwidth does not change the shape of the impulse response, but expands the voltage scale and shrinks the time scale in the same proportion as the bandwidth expansion. If frequency scale is substituted for bandwidth, this property is very easily demonstrated for a low-pass filter as follows: Taking as the filter impedance an unspecified function of frequency, $A(jw)$, the voltage across it due to an impulse of current, $p1$, may be written operationally as $V = A(p)p1 = f(t)$, say. If now the frequency scale is expanded by a factor a, the new voltage is $= A(p/a)p1 = aA(p/a)p/a1$, and by an elementary proposition this is equal to $af(at)$, which is the result as stated.

Put crudely, the idea behind most noise suppressors is to keep the interfering pulses very sharp and tall and then chop off their heads by an amplitude limiter. A subsequent filter, just fast enough to pass the modulation, smoothes their rumps.

As has been shown, an expanded frequency scale in the circuits before chopping serves this purpose; it is possible that other conclusions of the paper could be reached by modifying the treatment on these lines.

Mr. D. Weighton (*in reply*): Mr. Eaglesfield raises the problem of the validity of the term bandwidth applied to radio receivers, where the mathematical abstraction of the ideal filter is not admissible. I am indebted to him for pointing out that use of the term "frequency scale" in place of bandwidth would make the results more precise. It would, however, imply a restriction of application which is not, in fact, necessary. Two of the main data of design in any transmission system are, generally, the limitation of frequency space available by interference from adjacent channels, and, when phase may be neglected, the highest modulating frequency. This leads the engineer to deal in terms of bandwidth and to adopt some convenient (though admittedly artificial) definition of the term. It was felt for this reason, therefore, that the formulae would be of little value unless written in terms of bandwidth, even at the cost of some loss of precision. It appears, in fact, that the formulae apply with fair accuracy without insistence on any greater similitude of transmission characteristics than is normally the case in radio receivers.

It should be noted that where the term "bandwidth" is first introduced (in dealing with the band-pass filter), the assumption involved is that it is possible to find two frequencies outside of which the response of the filter is very small, so that the integration may be limited to a finite range. This assumption is tenable in almost all cases, and the results obtained are generally true to a close approximation. The need for a more precise definition of bandwidth arises later and is discussed in Section 4, so that at no point is any greater rigour claimed than the argument deserves.

It is, of course, relatively simple to write the impulsive response of a network in terms of a complex transmission coefficient and to deduce certain properties of the output signal by operational means. There is, however, implicit in this calculation an assumption that the impulse is of infinitely short duration, and with a similar simplification the Fourier approach reduces to a problem of comparable complexity. It is, moreover, important to inquire into the limits within which the approximation is valid, and the use of the Fourier integral is more informative in this respect. In general, this method was used since it shows more readily the implication of the approximations which are necessary and leads to a clearer physical picture of the operation involved.

* Paper by D. Weighton (see 1948, 95, Part III p. 69).