National Defence | Défense nationale

AN INVESTIGATION OF THE EXTENT OF THE FRESNEL REGION OF APERTURE ANTENNAS WITH ATTENTION TO SUPERDIRECTIVITY

by

Shawn Charland

DEFENCE RESEARCH ESTABLISHMENT OTTAWA
TECHNICAL NOTE 89-20

Canada

August 1989
Ottawa

National Defence / Défense nationale

AN INVESTIGATION OF THE EXTENT OF THE FRESNEL REGION OF APERTURE ANTENNAS WITH ATTENTION TO SUPERDIRECTIVITY

by

Shawn Charland
Radar Countermeasures Section
Electronic Warfare Division

DEFENCE RESEARCH ESTABLISHMENT OTTAWA
TECHNICAL NOTE 89-20

PCN
041LD

August 1989
Ottawa

ABSTRACT

This technical note describes the principles governing superdirective antennas and the factors limiting their realization. A criterion for determining the minimum distance to the far field of aperture antennas, including superdirective antennas, is presented. Under this condition it is shown that, at least for the marginally superdirective case, such antennas do not exhibit large Fresnel regions as compared with uniformly phased apertures. A modified criterion is also presented which defines the minimum distance to the far field in a given direction. Under this criterion it is shown that for the uniform distribution, the antenna pattern near nulls requires a large distance to converge to the far field form. This work has applications in determining antenna test range requirements.

RESUMé

Cette présente note technique décrit les principes régissant les antennes superdirectrices et les facteurs limitant leur réalisation. Un critère est présenté pour déterminer la distance minimum au champ lointain pour les antennes d'ouverture ainsi que pour les antennes superdirectrices. Dans cette condition on démontre que, du moins en ce qui concerne le cas d'une superdirectivité marginale, ces antennes ne déploient pas une grande région Fresnel en comparaison avec les ouvertures phasées uniformément. Un cirtère modifié est également présenté définissant la distance minimale au champ lointain pour une distribution uniforme, les zéros du diagramme polaire de gain de l'antenne représentent les directions où les distances minimales au champ lointain sont les plus grandes. Ce travail peut s'appliquer à la détermination des conditions requises pour un champ d'essai d'antennes.

EXECUTIVE SUMMARY

This investigation defines and describes superdirective aperture antennas. Several criteria are proposed for determining the maximum range of the Fresnel region of aperture antennas, and one is selected as best indexing this range. The selected criterion is based on a mean-squared convergence of the finite range pattern to the far field pattern. Using this criterion, the hypothesis that superdirective aperture antennas can exhibit a large Fresnel region is disproved, at least for the case of marginal superdirectivity. Application of this criterion over only a small sector of the antenna pattern shows that the nulls of a uniform aperture pattern converge to the far field form at relatively large range from the antenna. Also, apertures which are heavily weighted at the edges have a relatively large Fresnel region. Aperture distributions which are tapered at the edges have a relatively short Fresnel region.

<u>TABLE OF CONTENTS</u>

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES (cont.)

LIST OF SYMBOLS

γ	- supergain factor
a	- real part of angle of propagation
b	- imaginary part of angle of propagation
e	- base of the natural logarithm
β	- coefficient of the argument of the exponential in the Fresnel kernel
x,y	- variables of coordinate axes
j	- square root of -1
k_y	- y-component of the vector propagation constant
k_x	- x-component of the vector propagation constant
k	- vector propagation constant
δ	- delta function
θ,Θ	- angle variable (azimuth)
ϕ	- angle variable (elevation)
R	- radial distance from the centre of the aperture
ρ	- radial distance from a point on the aperture
ρ_o	- fixed radial distance from the aperture
D	- aperture width
λ	- wavelength
η, Z_o	- intrinsic admittance/impedance of the medium
ω	- angular frequency, rad./sec.
E	- electric field strength
H	- magnetic field strength
v	- velocity
f	- frequency, Hz
ψ	- see eqn. (106)
u	- direction of the linearly polarized electric field
n	- normal out of the aperture

1.0 INTRODUCTION

1.1 Background

The limitations of representing the physical world with mathematical models and relationships are well known, and the field of antenna theory is no exception. Although Maxwell's equations describe the variation of electromagnetic fields in time and space, mathematical approximations must generally be made to these expressions in order to formulate the relationships in a useful and tractable way. When considering the radiated (and non-radiated) fields of an aperture antenna, various approximations may be made in describing the radiation pattern, depending on the range from the antenna at which the radiation pattern is to be represented. There are generally three distinct regions over which the antenna pattern may be expressed in different forms, defined by the approximations which may be made to the solutions to Maxwell's equations. The region close to the antenna is known as the near field, in which reactive energy and boundary conditions around the antenna dominate. At slightly larger range, the Fresnel region is defined. The radiation pattern in the Fresnel region is dominated by the geometrical effects of the size of the antenna, and results in a radiation pattern which changes shape (as well as intensity) as a function of range. At very large range, the antenna may essentially be treated as a point source radiator. The far field is defined to include ranges over which the radiation pattern changes only in intensity (not in shape) as the range varies. The approximations which dominate in each of the regions blend into one another so it is difficult to delineate, in a rigorous fashion, where one set of approximation breaks down and another set begins to apply.

This investigation is concerned with the transition between the Fresnel region approximations and the far field approximations in representing the radiation pattern of an aperture antenna. The minimum distance to the far field is known to be a function of several variables including the physical aperture size and the aperture excitation. Superdirective antennas are of particular interest in this investigation because they have unusual aperture excitations which may influence the minimum distance to the far field. The excitation of these apertures produces a radiation pattern in visible space which has a directivity which excedes that of a uniformly excited aperture, which is the most directive completely real aperture excitation.

The purpose of this investigation is to derive a quantitative definition of the extent of the Fresnel region of an aperture antenna, and apply that definition to determine the extent of the Fresnel region of a superdirective aperture antenna. This definition of the extent of the Fresnel region is consistent with aperture dimensions, aperture excitation, and accurately reflects the degree of convergence of the antenna pattern, over the visible region, to the far field form. There is some evidence to suggest that the minimum distance to the far field for superdirective antennas may be large relative to that of apertures having completely real excitations. This will be discussed in a following section. For such antennas, geometrical considerations may not be the dominant factor in determining the extent of the Fresnel region. An investigation of these effects is important because it provides insight into the properties of superdirective antennas, and serves to expose the practicality (or lack thereof) and desirable attributes of realizing such antennas. This investigation also has applications concerning the minimum test range distance requirements for specific aperture antennas.

1.2 Technical Note Organization

Broadly speaking, this investigation is divided into two parts. The first deals with the definition, description, and realization of superdirective aperture antennas (sections 2 and 3), while the second part deals with the development of a quantitative definition of the minimum distance to the far field for a specific aperture distribution (sections 4, 5, and 6).

Section 2 presents the concept of describing the fields around an aperture antenna as a spectrum of plane waves. It also includes a discussion of admissible complex angles of propagation around an aperture, in the context of a plane wave spectrum.

Section 3 introduces the notion of superdirective aperture antennas, and describes the synthesis of such aperture excitations. The supergain factor is presented as a design aid, and some of the design constraints inherent to ensuring a realizable superdirective aperture are discussed.

Section 4 begins the second part of the investigation, and includes a derivation of the Fresnel kernel used in forming an approximate description of the fields in the Fresnel region of an aperture antenna. The conventional definition of the minimum distance to the far field, based on geometrical considerations alone, is derived. Three alternative criteria are presented for quantitatively defining the outer extent of the Fresnel region, along with an evaluation of the relative merit of each criterion. One of the three proposed criteria is selected as more appropriate to the investigation than the others, based on consistency, usefulness, and in reflecting the dependence of the extent of the fresnel region on the aperture excitation.

Section 5 presents an alternate form for the phase-sensitive mean-squared difference criterion. This leads to a discussion of a synthesis equation for specifying the minimum distance to the far field for an antenna. The relationship observed between half power beamwidth (HPBW) and the minimum distance to the far field (as determined by the phase sensitive criterion) is discussed.

Section 6 includes a discussion of radiation polarization and its effect on the validity of the mean squared difference criterion. Also, a modified mean squared difference criterion is introduced which is applied over only part of the radiation pattern. The fundamental properties of this new criterion are exposed by considering a special application to the case of a uniform aperture.

Section 7 is comprised of recommendations for further work in this area.

Section 8 summarizes the investigation.

2.0 DESCRIPTION OF APERTURE ANTENNA FIELDS

The electromagnetic fields associated with a radiating antenna may be represented as an angular spectrum of plane waves. This approach was first described in a paper by H. G. Booker and P. C. Clemmow [1]. Because of the relevance of this approach to further discussions, the plane wave spectrum approach is now developed, following the original paper.

Following Booker and Clemmow, several simplifications are adopted. Fig. 1 shows a diagram of a one-dimensional (line source) aperture antenna radiating into a half space $x > 0$. The electromagnetic fields are assumed to have a time harmonic variation, and the radiation pattern is represented in only a single plane. The magnetic field vector is taken parallel to the z-axis. By using Maxwell's equations, the charge and current conditions on a boundary of the source-free half space, and a specified time variation, the electromagnetic field throughout the half space may be determined. Simple ray tracing can be used (see fig. 2) to determine the phase of the contribution to the electromagnetic field from a Huygens source on the aperture to the phase of the total field at point P may be determined. Taking k as the free space vector propagation constant ($2\pi/\lambda$), the phase at point P is given by

$$\phi = e^{jk(x\cos\theta + y\sin\theta)} \tag{1}$$

where

$$R = (x\cos\theta + y\sin\theta) \tag{2}$$

or alternatively

$$\phi = e^{j(xk_x + yk_y)} \tag{3}$$

where k_x and k_y are the vector components of the propagation constant.

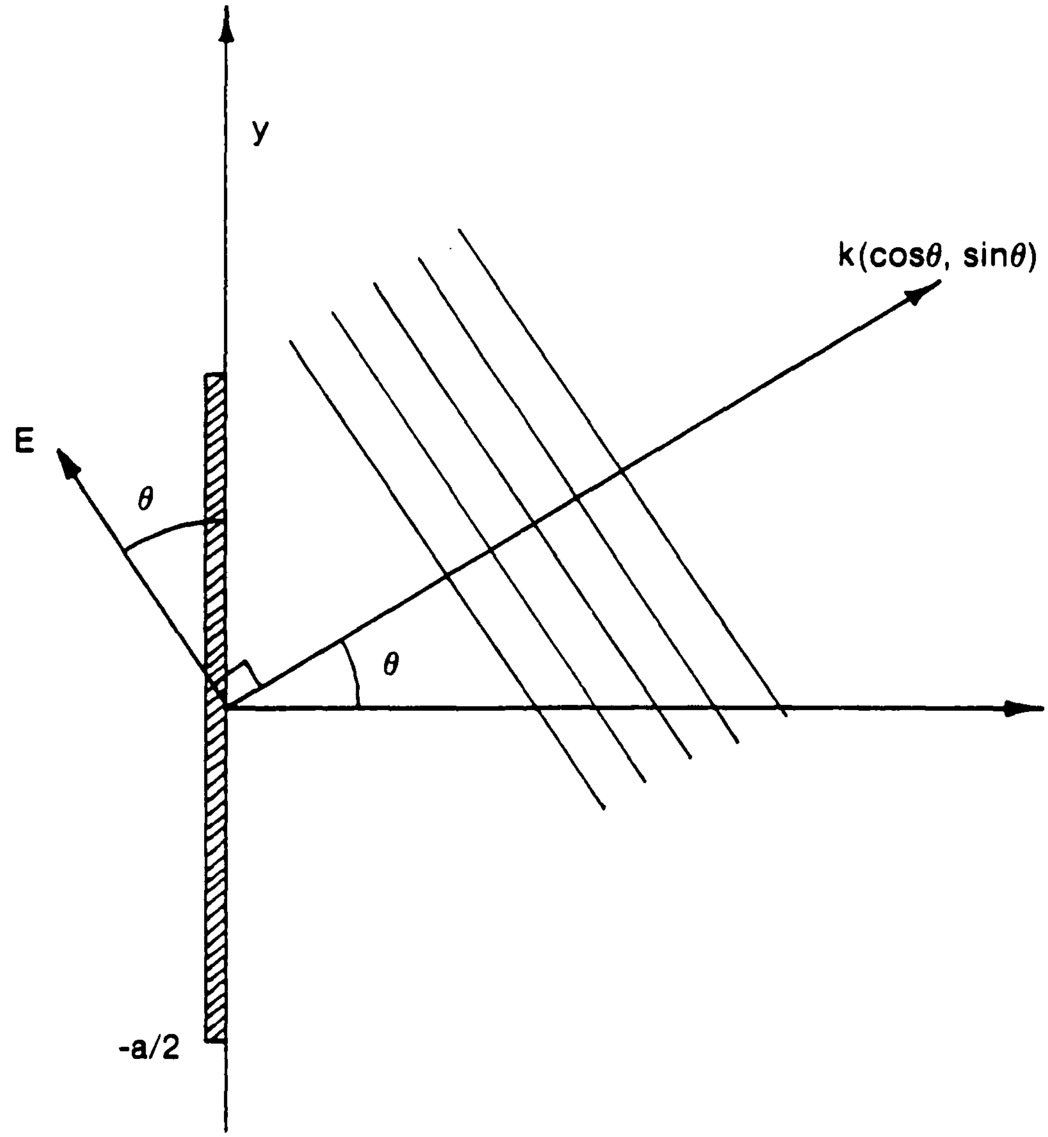


FIG. 1: DIAGRAM OF ONE DIMENSIONAL ANTENNA GEOMETRY

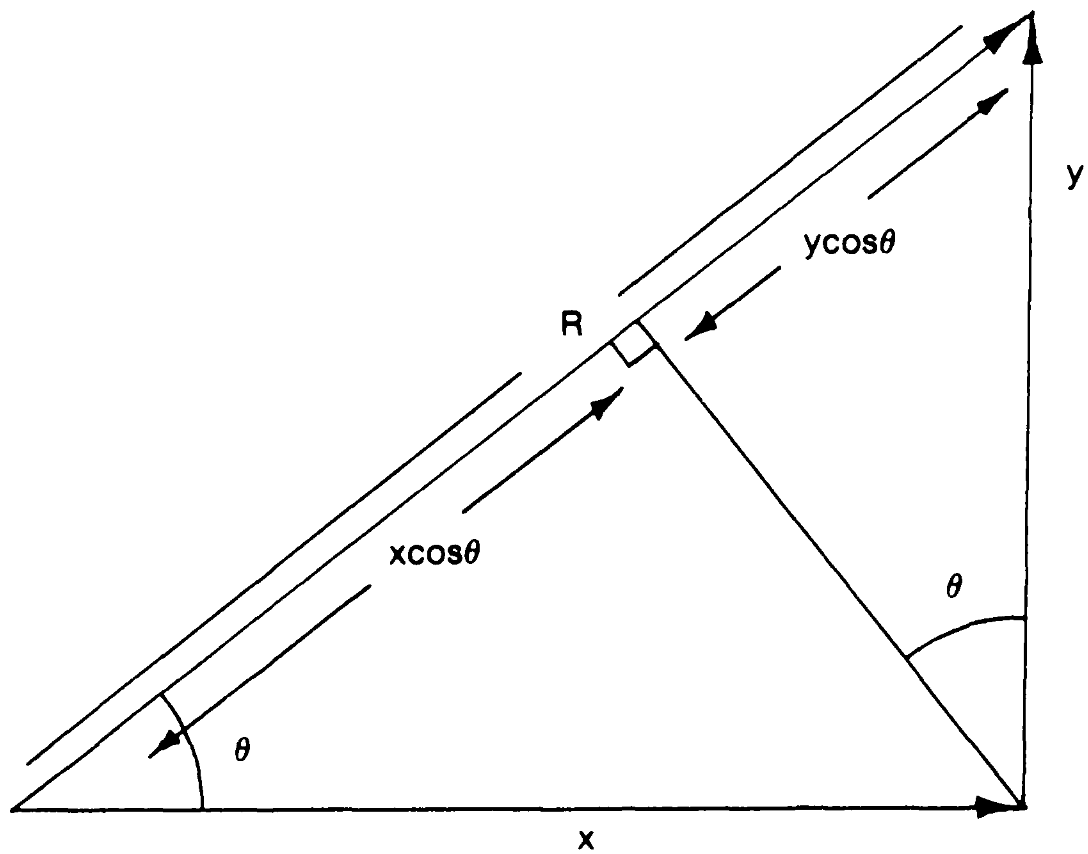


FIG. 2: GEOMETRY FOR THE PHASE EXPRESSION

As stated previously, the fields in the half space x > 0 are completely determined by the tangential field components on the aperture, considering the half space to be source free. The fields in the half space may be generally be written as:

$$E(x,y) = A(-\sin\theta, \cos\theta, O)e^{-jk(x\cos\theta + y\sin\theta)} \qquad (4)$$

$$H(x,y) = \eta A(0,0,1)e^{-jk(x\cos\theta + y\sin\theta)} \qquad (5)$$

where A(x,y,z) is an amplitude function of the three unit vectors. Note that, from fig. 1, the tangential component of the E field on the aperture is related to the cosine of the angle of propagation, and the component of the E field normal to the aperture is related to the (negative of the) sine of the angle. Setting x = 0 in eqns. (4) and (5) yields the expressions for the fields in the aperture required to support the specified fields in the half space. The expressions are:

$$E_y = A(\cos\theta)e^{-jyk\sin\theta} \qquad (6)$$

$$H(0,0,1) = Ae^{-jyk\sin\theta} \qquad (7)$$

These expressions represent a wave travelling over the aperture in the +y direction with vector propagation constant $k\sin\Theta$. Because this is a boundary condition, any such wave travelling over the aperture with vector propagation constant $k\sin\Theta$ (or k_y) leads to a plane wave travelling out of the aperture at an angle defined by the vector propagation constant k of the medium. The angle of propagation may be determined specifically be examining the transverse electromagnetic (TEM) plane wave constraints:

(i) $k_x + k_y = k$

(ii) $E . k = 0$

(iii) $H . k = 0$

Considering that the E-vector is in the plane defined to be the half space, condition (ii) may be expanded as:

$$E.k = (E_x + E_y).(k_x + k_y) = E_x k_x + E_y k_y = 0 \qquad (7a)$$

Hence once E_x, E_y and k_y are specified in the aperture plane, the value of k_x is also determined, and leads to a plane wave propagating out of the aperture at an angle whose cosine is k_x/x. It is now natural to ask what values of k_y are admissible in specifying the tangential fields of the aperture (or alternately, what values of $\sin\Theta$ are admissible).

The wave equation may be written, for the two-dimensional (single plane) case as:

$$\frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2} + k^2 F = 0 \qquad (8)$$

which implies that

$$k_x^2 + k_y^2 = k^2 \qquad (9)$$

Values of $\sin\Theta > 1$ are admissible and correspond to imaginary values of $\cos\Theta$. Values of $\sin\Theta > 1$ correspond to values of $k_y > k$. A discussion of complex angles of propagation is presented in Appendix A. The aperture component (y-component) k_y of the vector propagation constant is interpreted as the change of phase of the wave per free space wavelength across the aperture. For values of $k_y < k$, the change of phase is less than 2π radians per wavelength, corresponding to a plane wave propagating at some real angle Θ measured from the normal out of the aperture. However, values of $k_y > k$ specify more than 2π radians of phase change per free space wavelength across the aperture, implying a wave propagating in the aperture with a wavelength shorter than that of free space. Since the phase velocity of a wave in the aperture plane is given by:

$$v = \lambda f = \frac{\lambda\omega}{2\pi} = \frac{2\pi\omega}{2\pi k_y} = \frac{\omega}{k_y} \qquad (10)$$

waves corresponding to $k_y < k$ propagate at velocities above the free space value, while waves corresponding to $k_y > k$ are "slow" waves in the aperture, propagating at velocities below the free space value (because the same time harmonic variation applies in both cases).

However, referring to eqns. (3) and (4), slow waves in the aperture lead to a different kind of wave in the half space. Although $\sin\Theta$ is real, $\cos\Theta$ is imaginary for values of $\sin\Theta > 1$. This results in an exponential decay in the amplitude of these waves with distance from the aperture. This is reflected in the equations for the E and H fields in the half space, written below for the case of $k_y > k$:

$$E(x,y) = A(-\sin\theta,\cos\theta,0)e^{-k(\sin^2\theta-1)^{\frac{1}{2}}x}e^{-jyk_y} \tag{11}$$

$$H(x,y) = \eta A(0,0,1)e^{-k(\sin^2\theta-1)^{\frac{1}{2}}x}e^{-jyk} \tag{12}$$

These waves are known as "evanescent" or reactive waves, and do not propagate away from the aperture. Rather, they travel in the plane of the aperture. In summary, then, evanescent waves are slow moving, high frequency waves which are exponentially damped in the direction out of the aperture, and arise from considering all possible boundary conditions which could be imposed at the aperture.

These evanescent waves are associated with energy storage around the aperture. From the Poynting theorem, E X H gives the vector energy flow associated with an electromagnetic wave. To determine this vector in the context of this investigation, the E_x and E_y field components may be considered separately. E_x X H yields the component of energy flow along the y-axis. E_y X H yields the vector of the energy leaving the aperture in the x-direction. However, as can be seen from eqn. (3), E_y is related to the cosine of the angle of propagation of the wave out of the aperture. For evanescent waves, $\cos\Theta$ is imaginary, and so the component of the E field in the plane of the aperture is in time (or space) quadrature with the H field. Fig. 3 shows the instantaneous Poynting vector for the case of quadrature E and H fields. E_x X H changes direction every 1/4 cycle, and there is no net energy propagated away from the aperture. Rather, energy surges back and forth between electric and magnetic fields in front of the aperture. If evanescent waves are included in the distribution of aperture electromagnetic fields, it becomes clear that any time varying aperture distribution can, by Fourier analysis, be decomposed into a spectrum of sinusoidal waves propagating over the aperture. The evanescent waves are necessary in describing details of the aperture distribution which are finer than the free space wavelength. Since each harmonic in the aperture leads to a plane wave in the half space radiating at some angle dependent on the propagation constant of the harmonic, any arbitrary field distribution in the half space can be described as an angular spectrum of plane waves (Woodward and Lawson, [2]). The summation of these plane waves in the half space corresponds to integration over all harmonics in the aperture distribution-that is, over all possible values of k_y. Following Booker and Clemmow, arbitrary fields in the half space can be represented by the expressions below:

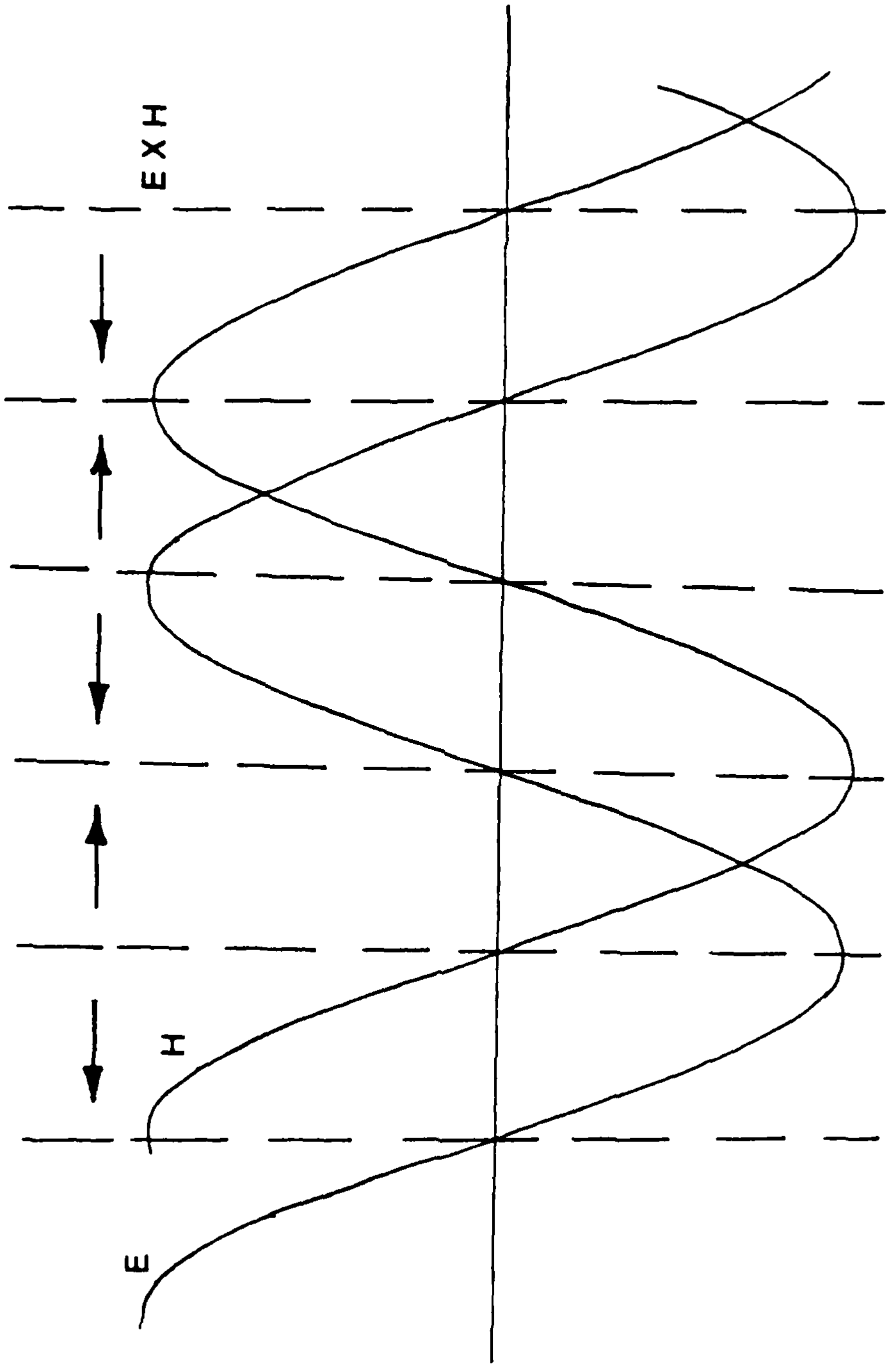


FIG. 3: POWER FLOW ASSOCIATED WITH EVANESCENT WAVES

$$E(x,y) = \frac{1}{\lambda} \int_{-\infty}^{\infty} P(\sin\theta)(-\sin\theta,\cos\theta,0)e^{-jk(x\cos\theta + y\sin\theta)}\frac{d\sin\theta}{\cos\theta} \tag{13}$$

$$H(x,y) = \frac{\eta}{\lambda} \int_{-\infty}^{\infty} P(\sin\theta)(0,0,1)e^{-jk(x\cos\theta + y\sin\theta)}\frac{d\sin\theta}{\cos\theta} \tag{14}$$

where

$$P(\sin\theta) = (\lambda)CA(\sin\theta) \tag{15}$$

since "A" is a function of the particular harmonic, ie. a function of $\sin\Theta$ (or k_y). The aperture distribution, completely specified by the electric field, is given by setting $x = 0$:

$$E(O,y) = \frac{1}{\lambda} \int_{-\infty}^{\infty} P(\sin\theta)e^{-jyk_y}d\sin\theta \tag{16}$$

This is a Fourier Transform expression.

As an aside, when the radiation pattern is described as the Fourier Transform of an aperture distribution, the pattern over $-1 < \sin\Theta < +1$ (or $-k < k_y < +k$) is said to lie in the visible region, since these values of $\sin\Theta$ correspond to real angles. The pattern over the remaining $\sin\Theta$ axis is said to lie in the invisible region. Understanding that the waves propagating at these (complex) angles are exponentially damped with range from the aperture gives a literal meaning to the pattern in the "invisible" region - these waves are invisible in that they do not propagate to the far field.

2.2 Discussion of Complex Angles of Propagation

As stated previously, the wave equation does not restrict values of $\sin\Theta$ to lie between +1 and -1. Thus values of $\sin\Theta > 1$ or $\sin\Theta < -1$ correspond to imaginary values of $\cos\Theta$ and complex angles of propagation. However, it does constrain the range of admissible complex angles. Suitable boundary conditions must be used to investigate the constraints on these complex angles of propagation.

Firstly, $\sin\theta$ must be real, but otherwise may take any value. If $\sin\theta$ were imaginary or had an imaginary part, the $e^{jk_y y}$ would have an exponential real part, and the field amplitude would go to infinity in either the +y or -y direction. Also, the $\cos\theta$ term must not have any positive imaginary part. A positive real part implies that the field amplitude goes exponentially to infinity in the +x-direction. If the complex angle of propagation is defined as $\theta = (a + jb)$, all admissible and inadmissible values of $\sin\theta$ may be represented by considering the four sign permutations of a and b:

$$\sin\theta = \frac{e^{j\theta} - e^{-j\theta}}{2j} \tag{17}$$

or, for (a + jb):

$$\sin\theta = \frac{e^{-b}(\cos(a) + j\sin(a))}{2j} + \frac{e^{b}(\cos(a) - j\sin(a))}{2j} \tag{18}$$

which may be written as:

$$\sin\theta = \frac{j(e^{b} - e^{-b})\cos(a)}{2} + \frac{(e^{-b} + e^{b})\sin(a)}{2} \tag{19}$$

The constraint that $\sin\theta$ has no imaginary part implies that $a = \pi/2$ or $-\pi/2$, unless $b = 0$. If $b = 0$ the angle of propagation is completely real, and $\sin\theta$ ranges between -1 and +1. The sign permutations of the complex angle gives values of $\sin\theta$ as:

$$(\tfrac{\pi}{2} + jb)\ldots\ldots\ldots\ldots\sin\theta = (e^{--b} + e^{b})/2 = \cosh(b) \tag{20}$$

$$(\tfrac{\pi}{2} - jb)\ldots\ldots\ldots\ldots\sin\theta = (e^{-b} + e^{b})/2 = \cosh(b) \tag{21}$$

$$(-\tfrac{\pi}{2} + jb)\ldots\ldots\ldots\ldots\sin\theta = -(e^{-b} + e^{b})/2 = -\cosh(b) \tag{22}$$

$$(-\tfrac{\pi}{2} - jb)\ldots\ldots\ldots\ldots\sin\theta = -(e^{-b} + e^{b})/2 = -\cosh(b) \tag{23}$$

To determine further constraints on the complex angle, values of $\cos\theta$ must be examined for the sign permutations of a and b. Recall that the existing constraint is that if b is nonzero, $a = \pm\pi/2$. $\cos\theta$ is defined as:

$$\cos\theta = \frac{e^{j\theta} + e^{-j\theta}}{2j} \tag{24}$$

or, substituting (a + jb) for the complex angle of propagation:

$$\cos\theta = \frac{(e^{-b} + e^{b})\cos(a)}{2} + \frac{(e^{-b} - e^{b})\sin(a)}{2} \tag{25}$$

and the permutations give:

$$(\tfrac{\pi}{2} - jb) \ldots\ldots\ldots \cos\theta = -j(e^{b} - e^{-b})/2 = -j\sinh(b) \tag{26}$$

$$(\tfrac{\pi}{2} - jb) \ldots\ldots\ldots \cos\theta = j(e^{b} - e^{-b})/2 = j\sinh(b) \tag{27}$$

$$(-\tfrac{\pi}{2} + jb) \ldots\ldots\ldots \cos\theta = j(e^{b} - e^{-b})/2 = j\sinh(b) \tag{28}$$

$$(-\tfrac{\pi}{2} - jb) \ldots\ldots\ldots \cos\theta = -j(e^{b} - e^{-b})/2 = -j\sinh(b) \tag{29}$$

Recalling that $\cos\Theta$ may not have a positive imaginary part, (a + jb) and (a - jb) are inadmissible complex angle of propagation. A similar discussion may be found in Woodward and Lawson [2]. Following Woodward and Lawson, a diagram of the contour of admissible angles in the complex plane is shown in fig. 4.

The above discussion lends insight into the nature of the invisible region of the radiation pattern and evanescent waves. From eqn. (4), the amplitude weighting of the components of the electric field are functions of the angle if propagation resolved on the unit vectors, as:

$$A = A(-\sin\theta, \cos\theta, 0) \tag{30}$$

The invisible region includes, by definition, values of $-1 > \sin\Theta > +1$, and in this region $\sin\Theta$ may be written as $\pm\cosh(b)$ and $\cos\Theta$ may be written as $-j\sinh(b)$. Thus the amplitude of the aperture fields may be written as, referring to eqn. (15):

$$A = A(\pm\cosh(b), -j\sinh(b), 0) \tag{31}$$

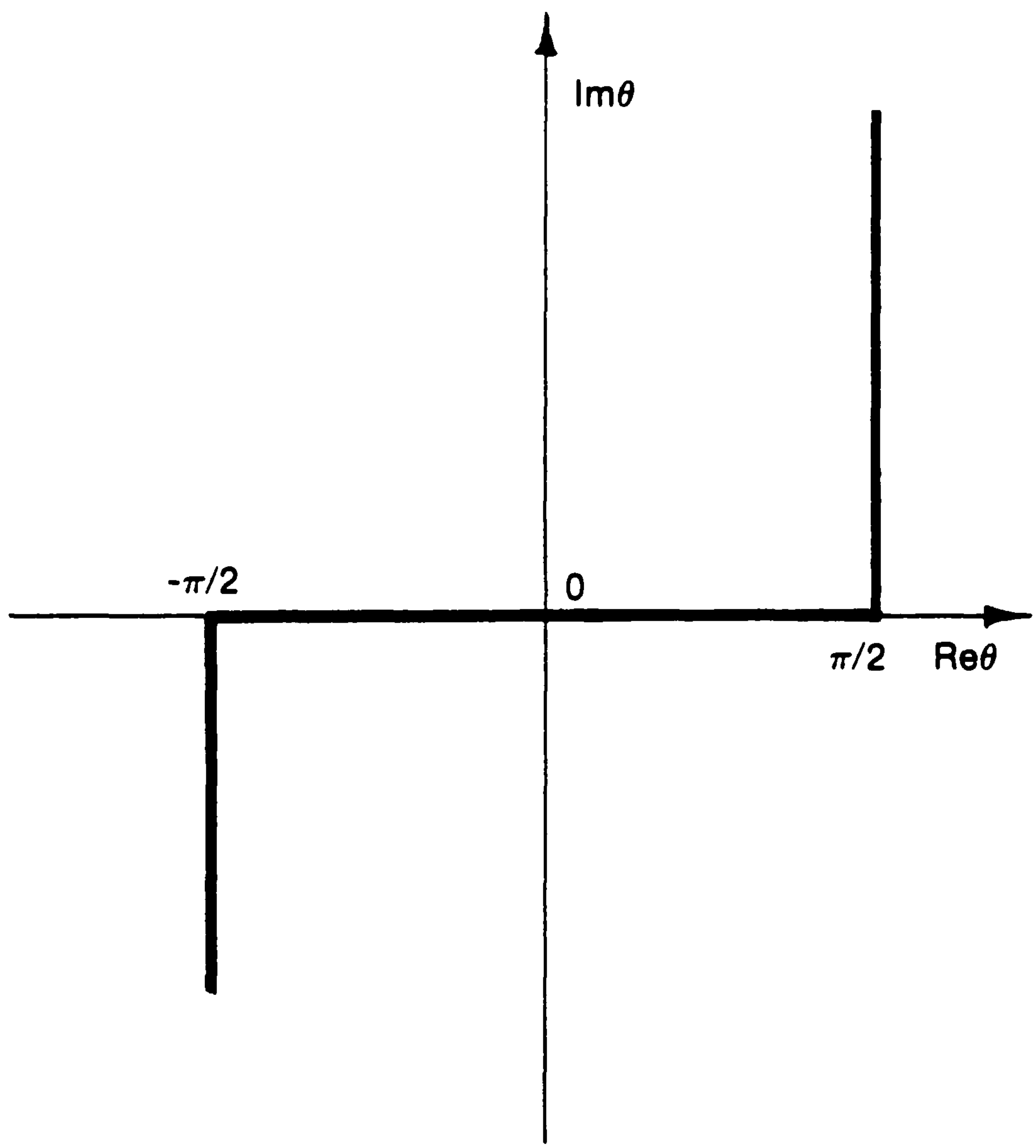


FIG. 4: CONTOUR OF ADMISSIBLE ANGLES OF PROPAGATION

By definition, cosh(b) and sinh(b) increase exponentially with increasing b. This means that the fields in the aperture plane required to support the evanescent waves increase exponentially as the waves become more evanescent, that is, as one moves farther into the invisible region.

3.0 THE SUPERDIRECTIVE ANTENNA

3.1 Superdirectivity

An antenna whose directivity is greater than that of a uniformly illuminated aperture is said to be superdirective. For the case of a uniform phase aperture distribution, the uniform amplitude distribution yields the highest directivity. This may be shown analytically using the Cauchy-Schwartz inequality and the expression for directivity. Directivity is defined as the maximum radiation intensity per unit solid angle, normalized to average radiation intensity.

$$G = \frac{4\pi \left| \int E(x,y)dxdy \right|^2}{\lambda^2 \int |E(x,y)|^2 dxdy} \tag{32}$$

The Cauchy-Schwartz inequality given as:

$$\left| \int fg dxdy \right|^2 \leq \int f^2 dxdy \int g^2 dxdy \tag{33}$$

Thus, using eqn. (32), an aperture width of A, and letting f(x,y) = 1:

$$\left| \int E(x,y)dxdy \right|^2 < A \int |E(x,y)|^2 dxdy \tag{34}$$

and the uniform case yields the highest directivity.

However, Woodward and Lawson [2] outline a synthesis procedure by which any arbitrary radiation pattern may be realized. The procedure was originally suggested in 1943 by Schelkunoff [3], who showed that the power gain of a linear array could be theoretically increased indefinitely, provided the number of elements in the array is correspondingly increased without increasing the length of the array. In 1948, Riblet [4] concluded that the directivity of a two dimensional current distribution could likewise be made arbitrarily large.

Fundamental to the concept of superdirectivity is that, while the radiation pattern is defined over the entire real $\sin\Theta$ axis, the visible region (or that part of the pattern which constitutes the far field) lies only over ± 1. Another important fact is that the aperture distribution corresponding to a sum of radiation patterns is given by the sum of the corresponding component aperture distributions (linearity of electromagnetic fields in a reciprocal medium). Further, it is possible to specify an infinite number of independent radiation patterns which can be realized by a given finite aperture. Such radiation patterns which can be realized exactly by finite length and finite energy apertures are known as "aperture limited" functions (Rhodes, [5]).

3.2 Superdirective Antenna Synthesis

A method for determining the aperture distribution required to synthesize a superdirective radiation pattern will now be presented. Following the method for synthesizing arbitrary radiation patterns laid out by Woodward and Lawson, the desired radiation pattern is first specified at n points over the visible region. To realize this, a sum of appropriately weighted (independent) aperture limited functions is used. The composite pattern has the form:

$$p = \sum^{s} A_s P_s \tag{35}$$

where A_s values are found from the simultaneous equations arising from specifying the desired radiation pattern at s discrete points. Woodward and Lawson offer no proof that as the number of specified points increases the composite pattern converges to a smooth curve, but rather rely on intuition and worked examples. They observe that as one attempts to force the composite pattern to an arbitrary shape, the pattern in the invisible region becomes greater. In the limit and in general, as the specified pattern is achieved, the pattern in the invisible region tends to infinity. An alternate method of synthesis was proposed by Rhodes [5] in which he came to the same conclusion. Rhodes' method centres on achieving an arbitrary radiation pattern in a best mean-square sense, subject to a constraint on the size of the pattern in the invisible region. Rather than specifying the pattern at n discrete points, he attempts to expand the desired pattern over the visible region in prolate spheroidal wave functions.

To further illustrate and clarify the idea that the pattern becomes large in the invisible region, an example is presented from Woodward and Lawson. The independent component patterns are chosen as the following function evaluated at integer values of s:

$$p_s(\sin\theta) = \frac{\sin(\pi W(\sin\theta - s/W))}{\pi W(\sin\theta - s/W)} \tag{36}$$

where W is the aperture width in wavelengths.

The desired radiation pattern is specified at values of

$$\sin\theta = r/W \tag{37}$$

where r takes positive and negative integer values. The simultaneous equations for the weighting coefficients take the form:

$$p(r/W) = \sum^{S} A_s p_s(r/W) \tag{38}$$

and using the orthogonality of sine functions, these reduce to:

$$A_r = p(r/W) \tag{39}$$

The composite (desired) radiation pattern thus has the form:

$$p(\sin\theta) = \sum^{S} p\left(\frac{s}{W}\right) \frac{\sin(\pi W(\sin\theta - s/W)}{\pi W(\sin\theta - s/W)} \tag{40}$$

Fig. 5 shows how the resulting radiation pattern is built up of the component patterns, and how the maximum of each component lies on the zeros of the other patterns. Thus if one wishes to specify the desired radiation pattern at values between points at r/W, sidelobes of the component patterns must be used, placing the main beam of these component patterns in the invisible region. Although this can be done theoretically, control of the visible region by using the sidelobes of patterns in the invisible region is very weak. Large patterns result in the invisible region for only marginal adjustments to the pattern over the visible region. This is illustrated by the radiation pattern shown in fig. 6. From discussions in previous sections, the pattern in the invisible region corresponds to evanescent waves and energy storage around the aperture. Also, the electric and/or magnetic fields in the aperture required to support these fields increase roughly exponentially as one moves into the invisible region. For this reason, large patterns in the invisible region generally mean that the specified aperture distribution is unrealizable. Because of this, and coupled with the fact that control of the visible region by patterns in the invisible region is so weak, it appears that it is impractical to realize significantly superdirective aperture distributions.

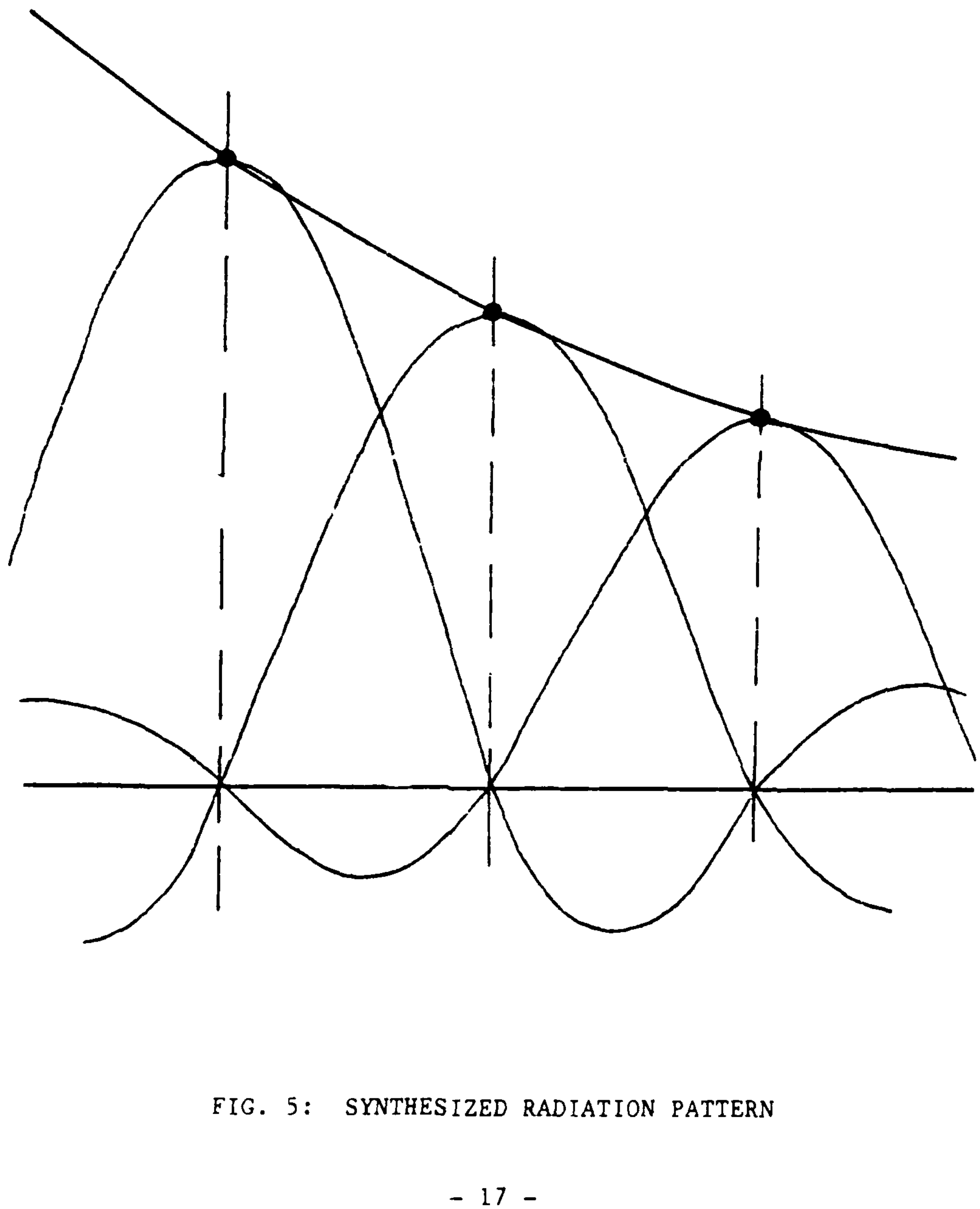


FIG. 5: SYNTHESIZED RADIATION PATTERN

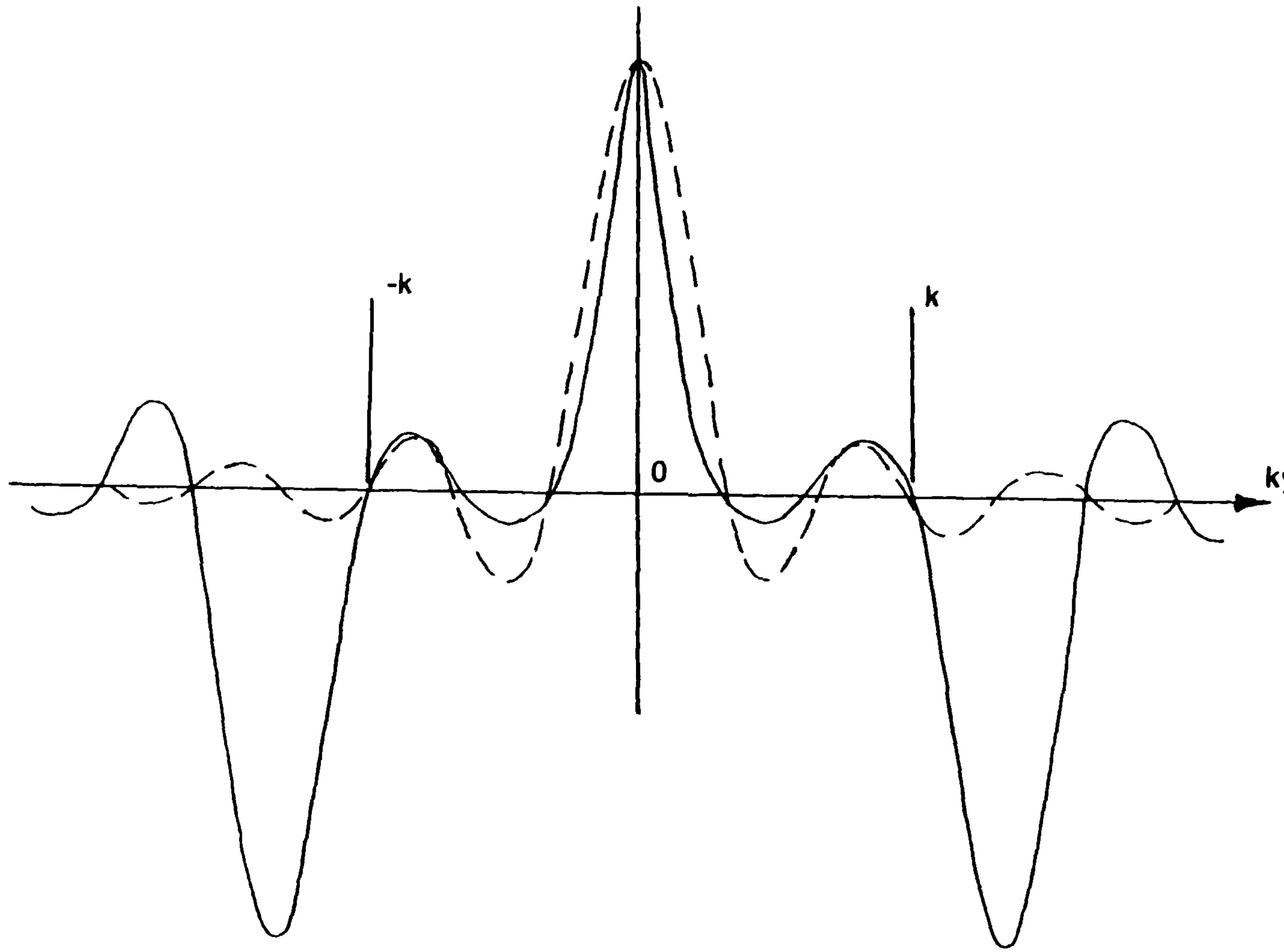


FIG. 6: SYNTHESIZED SUPERDIRECTIVE RADIATION PATTERN

3.3 Realizability and the Supergain Factor

Though superdirective apertures and apertures yielding patterns with arbitrarily narrow beams are theoretically possible, they are generally unrealizable for a number of practical reasons. The physical constraints of realizing a superdirective antenna become severe very quickly as one attempts to improve over the uniform case, so that even marginally superdirective antennas are difficult to realize. The physical limitations include high voltage and current requirements (with attendant high ohmic losses), large amounts of reactive energy, critical tolerances on elements spacing and phasing, narrow bandwidth, and low radiation resistance.

Specifically, in a note concerning the practical limitations of the directivity of antennas, Wilmotte [6] references the design of an antenna within 1 wavelength having a directivity comparable to an antenna 10 wavelengths long. However, its radiation resistance was calculated to be of the order of a millionth of an ohm. Such a low radiation resistance means that antenna efficiency is very low and most of the energy is being burned off as heat rather than radiation. He notes that the radiation resistance remains relatively steady for "conventional" designs (in a directivity sense) until the directivity in increased beyond a certain point, after which the radiation resistance falls off very quickly. Similarly Schelkunoff [7] notes that the assertion that there is no upper limit on the directivity obtainable from a linear array assumes there is no heat loss. He explains that a small radiation resistance in an antenna at resonance can allow for large energy storage (analogous to a resistor-capacitor-inductor circuit) and results in reactive fields extending a large distance from the antenna. The notion of a small radiation resistance is intrinsic to the idea of a superdirective antenna. Further, Schelkunoff [7] concludes that detuned superdirective antennas do not channel energy to themselves, and so reactive fields are likewise intrinsic to superdirective antennas.

Taylor [8] in 1955 introduced a parameter intended to indicate the practicality of a given design without the attendant complex calculations. Taylor proposed the notion of a "supergain ratio", which is defined as:

$$\gamma = \frac{\int_{\infty}^{\infty} f^2 dk_y}{\int_{-k}^{k} f^2 dk_y} \qquad (41)$$

A large supergain factor indicates an impractical design. The idea here is that large amounts of stored energy around an aperture indicate an unrealizable design, and the ratio can loosely be interpreted as

$$\gamma = 1 + \frac{P_j}{P_r} \qquad (42)$$

where P_j is the net reactive power and P_r is the total radiated power. From a discussion presented by Colin and Rothschild [9], a supergain factor only moderately greater than 1 does not necessarily imply a realizable design, but a supergain ratio of much greater than 1 ensures an unrealizable one. This is because P_j is the net reactive power, and may be zero at resonance, since it can be written as:

$$P_j = 2w(W_m - W_e) \tag{43}$$

where W_m and W_e are the maximum stored magnetic and electric energies respectively. Stored electric and magnetic energies are 180° out of phase with one another, so under conditions of resonnance the reactive stored energy may be very large while P_j is zero, and hence the design would be unrealizable. In answer to this problem, Rhodes [10] derived an expression for the net "observable" stored electric and magnetic energies. The term "observable energies" refers to the electric and magnetic energies which are pumped into the volume around the aperture. Except in the case of resonance, the volume around the aperture contains a surplus of either electric or magnetic energy per cycle, and P_j represents this surplus. Rhodes derives separate expressions for W_m and W_e, and defines P_j' as:

$$P_j' = 2w(W_m + W_e) \tag{44}$$

or, more explicitly,

$$2\omega W_m = \frac{4\pi^2}{2kZ_0} \int_{k_y^2+k_x^2>k^2} \frac{|k_xF_y - k_yF_x|}{k_y^2+k_x^2-k^2}\, dk_y dk_x \tag{45}$$

$$2\omega W_m = \frac{4\pi^2}{2kZ_0} \int_{k_y^2+k_x^2>k^2} \frac{k^2(|F_y|^2 + |F_x|^2)}{k_y^2 + k_x^2 - k^2}\, dk_y dk_x \tag{46}$$

where F_y and F_x are defined as

$$F_y(k_y,k_x,k) = \frac{1}{4\pi^2} \iint_s E_y(x,y,O)e^{j(yk_y+xk_x)}\, dydx \tag{47}$$

$$F_y(ky,kx,k) = \frac{1}{4\pi^2} \iint_s E_x(x,y,O)e^{j(yky+xkx)}\, dydx \tag{48}$$

for the general case of a two-dimensional aperture in the yz plane. This mean that Taylor's supergain factor may be redefined using P_j' rather than P_j, and so is of greater design interest. However, two points bear special mention: first, the supergain ratio remains a "figure of merit" for a prospective design, useful primarily as a consistent means of comparing the desirability of one design over another. Secondly, a large supergain ratio does not imply a superdirective antenna. As stated previously, by superimposing translated and scaled aperture limited patterns it is theoretically possible to synthesize any desired radiation pattern shape over the visible region (or any part thereof) at the price of generally increasing the stored energy around the aperture (ie. the size of the pattern in the invisible region). This statement holds true for arbitrarily obtuse radiation patterns as well as superdirective ones (for instance, large reactive fields are expected in attempting to synthesize a true isotropic source).

4.0 THE MINIMUM RANGE TO THE OUTER FRESNEL BOUNDARY

4.1 Definition of the Fresnel Region

Having described the notion of superdirective aperture antennas, an analysis is now carried out to determine the extent of the Fresnel region of such antennas. A general analysis of this problem for superdirective antennas can be usefully applied to specific cases of marginally superdirective antennas which can be realized, although strongly superdirective antennas may not be realizable. The critical phasing of superidrective aperture antennas suggest that such antennas may exhibit a large Fresnel region, compared to the case of uniforrr phase apertures of comparable physical size. The Fresnel region of an aperture antenna will now be defined and expressions for the fields in this region will be derived.

The volume surrounding an antenna in a source free space has been traditionally divided into three regions, in which various groups of simplifications may be used in solving for the electromagnetic field distribution. The Fresnel-Kirchoff scalar diffraction equation is an approximation of the solution of Maxwell's equations around a radiating aperture antenna, and may be used in calculating these fields. It is expressed as:

$$F(x,y,z) = \frac{1}{4\pi} \int_s G(\xi,\eta)\frac{e^{-jkr}}{r}\{jk - \frac{1}{r}\cos(n,r) - jk\cos(n,s)\}d\xi d\eta \qquad (49)$$

The region very close to the aperture is known as the near field or reactive field. No simplifications can be made to the Fresnel-Kirchoff equation here. In fact, as Skolnik [11] points out, the scalar diffraction equation is itself an approximation very close to the aperture because boundary conditions have been ignored. In this region the reactive field dominates. This reactive or evanescent field dies out with increasing range from the aperture. Beyond the range at which the reactive fields may be ignored, the scalar diffraction equation may be simplified somewhat, defining the Fresnel region of the antenna. The Fresnel region is characterized by the fact that the shape of the radiation pattern is a function of distance from the aperture. This effect is due to an effective phase curvature across the aperture, arising from path differences across the aperture. In the far field, the aperture can be treated as a point source, and the radiation pattern does not change shape with increasing distance.

The expressions for the fields in the Fresnel region and far field are of interest in this investigation. An expression for the fields in the Fresnel region may de derived by referring to fig. 7a. From the geometry of the diagram, the expression for the fields in front of the aperture may be written as:

$$E(y,r) = \int_{\text{line}} \frac{I(x)e^{-jk\rho}}{\rho} dx \tag{50}$$

The radial distance from a point on the aperture to an observation point can be written as:

$$\rho = \sqrt{r^2 + (y - x)^2} \tag{51}$$

or

$$\rho = \rho_0 \sqrt{1 - \frac{2xy}{\rho_0^2} + \frac{x^2}{\rho_0^2}} \tag{52}$$

By applying the approximation

$$\sqrt{1 + \Delta} = 1 + \frac{\Delta}{2} \tag{53}$$

the radial distance from a point on the aperture to the observation point may be written as:

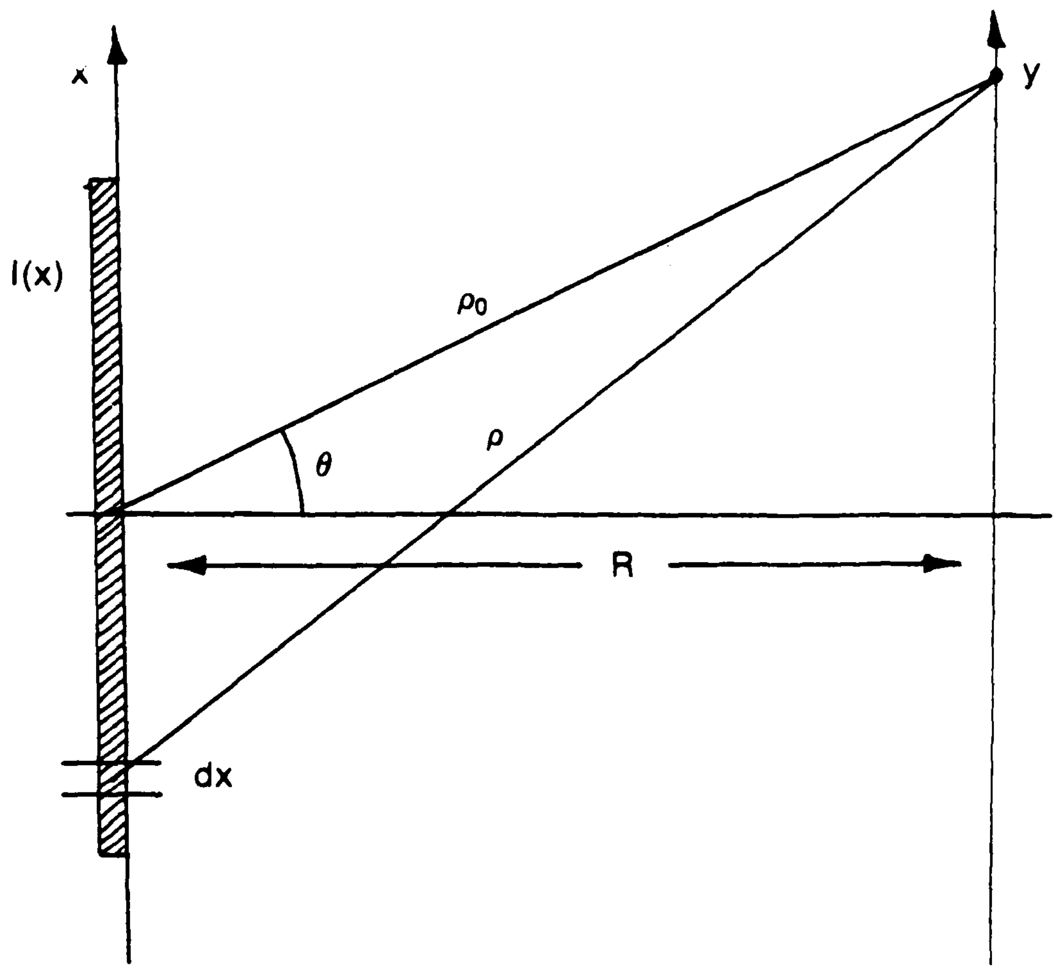


FIG. 7a: GEOMETRY FOR DERIVING THE FRESNEL KERNEL

$$\rho \simeq \rho_0\left(1 - \frac{xy}{\rho_0^2} + \frac{x^2}{2\rho_0^2}\right) \tag{54}$$

By referring to fig. 7a, the angle Θ of the point of observation relative to the normal out of the aperture can be stated as:

$$\frac{y}{\rho_0} = \sin\theta \tag{55}$$

Thus eqn. (51) can be written as:

$$E(\sin\theta) = \frac{ke^{-jk\rho_0}}{\rho_0} \int_{\text{line}} I(x) e^{jk\left(y\sin\theta - \frac{y^2}{2\rho^2}\right)} dx \tag{56}$$

Eqn. (56) is characterized by the exponential term, which may be considered in two parts. The first part, $\exp(+jyk\sin\Theta)$, is known as the Fourier kernel and is associated with the far field or Fraunhoffer region. The second term of the exponential, $\exp(-jy^2k/2\rho_0)$, is known as the Fresnel kernel, and is associated with the fresnel region. In the far field, the Fresnel kernel is negligible since ρ_0 tends to *infinity*. *It is* a quadratic phase front across the aperture, and is of the same form as the phase distribution one would expect from a circular wave (in the single plane case) impinging on the aperture from a point source at a finite distance in front of the aperture. Since ρ is the range from the aperture it is apparent that the radiation pattern in the Fresnel region is a function of range, whereas in the far field all dependence on range has disappeared (except the 1/r electric field variation).

There are no fixed, well-defined boundaries specifying the transition from the near field and Fresnel region and the Fresnel region and the far field. Skolnik [11] states that the inner Fresnel boundary is taken to be that distance from the antenna at which the reactive component of the field is not more that 1% of the radiated field. From this, he concludes that the Fresnel approximations should not be used to predict fields closer that 8 wavelengths from the aperture, since at this distance the reactive field intensity is about 40 dB below the radiated field intensity.

The outer Fresnel boundary has traditionally been taken as $2D^2/\lambda$ (used mainly for measruements of antenna patterns), where D is the maximum dimension of the aperture. This is an arbitrary but (generally) conservative minimum distance at which the far field approximations may be applied. This criterion arises by imposing the constraint that the maximum allowable edge phase error on the aperture is $\pi/8$ rad. The path difference between a ray from an observation point on boresight of the aperture and the centre of the aperture vs. the edge of the aperture is given by, referring to fig. 7b:

$$y - x = \lambda/16 \tag{57}$$

and also

$$y = \sqrt{\left(\frac{D}{2}\right)^2 + x^2} \tag{58}$$

Substituting eqn. (58) into eqn. (57), the path difference is expressed:

$$\sqrt{\left(\frac{D}{2}\right)^2 + x^2} - x = \frac{\lambda}{16} \tag{59}$$

Isolating x in eqn. (59) gives:

$$x = \frac{64D^2 - \lambda^2}{32\lambda} \tag{60}$$

and if $D \gg \lambda$, then

$$x \approx \frac{2D^2}{\lambda} \tag{61}$$

Note that this $2D^2/\lambda$ definition of the transition from the Fresnel region to the far field cannot be used for antennas smaller than 2λ because of the reactive field criterion (of 8λ from the aperture, Skolnik [11]) for the transition from the near field to the Fresnel region. For a 2λ aperture, the minimum distance to the far field is equal to the distance at which the reactive field is 1% of the radiated field.

Because the definition of the distance to the outer Fresnel boundary is generally taken as $2D^2/\lambda$, it is natural to ask if the aperture distribution in any way affects the minimum distance at which the far field approximations may be applied. Jull [12] states that the effect of finite range on a radiation pattern does depend on the aperture distribution, but in a qualitative way. He states that patterns highly tapered towards the edges are less affected by path differences than distributions weighted to the edges, and as a consequence $2D^2/\lambda$ may not be conservative enough in the latter case. The

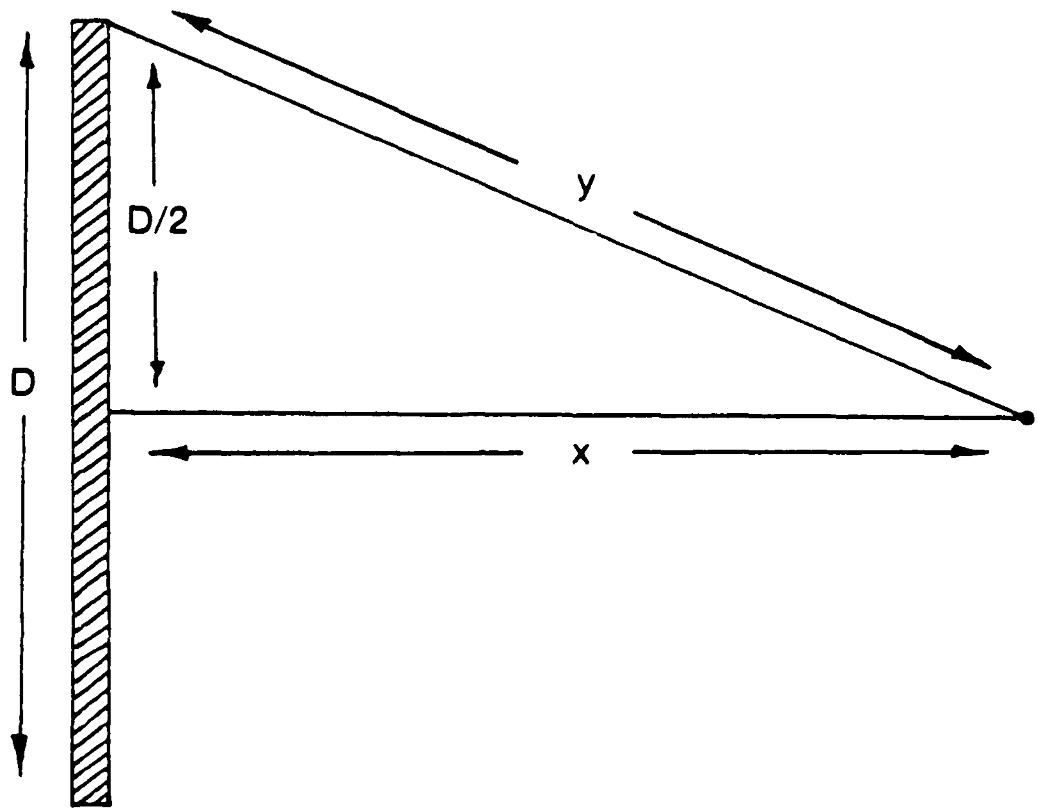


FIG. 7b: GEOMETRY FOR DERIVING THE MINIMUM DISTANCE TO THE FAR FIELD

finite range effects and properties of antennas are important when dealing with large antennas, such as reflectors. For example, consider a satellite station having a paraboloidal reflector 30 metres in diameter. The $2D^2/\lambda$ approximation with an uplink frequency of 6 GHz determines an outer Fresnel region limit at approximately 36 km. Thus all test measurements of the reflector are restricted to the Fresnel region. Even for small antennas, far field measurements may prove difficult or impossible. At 35 GHz a 1 metre reflector requires a test range of over 230 metres, which is difficult to arrange due to multipath problems. Such measurements might be more easily arranged if the $2D^2/\lambda$ criterion is unnecessarily conservative for certain aperture distributions. Thus it may be useful to specify the minimum distance to the far field as a function of aperture distribution in a rigorous way.

4.2 Minimum Far Field Criteria

4.2.1 The Fresnel Zones Criterion

In order to determine the effect of the aperture distribution on the minimum range at which the far field approximations may be applied, the implications of the far field approximations must be examined at a fundamental level. As stated previously, the form of the radiation pattern is a function of range in the Fresnel region. The radiation pattern converges in some sense to the far field pattern as the range increases through the Fresnel region. An aperture distribution based criterion for the minimum range at which the far field approximations may be applied (outer extent of the Fresnel region) should in some way reflect that the traditional conservative placement of the boundary is acceptable for most uniform phase aperture antennas. A standard or point of reference is required for such a criterion, and this reference should be related to the conservative placement of $2D^2/\lambda$. For the purposes of this investigation, however, the reference range is arbitrarily taken as $D^2/2\lambda$, to highlight the analytical nature of the investigation by divorcing the results from traditional definitions of the outer extent of the Fresnel region.

An example of a criterion which is not based on aperture distribution is presented by Silver. Silver ([13], pp. 196 ff) uses Fresnel zones to define the minimum distance at which the far field assumptions can be applied. To develop this criterion, the aperture is divided up into Fresnel zones by considering a point some distance away on boresight, surrounded by concentric spheres of radii increasing in steps of $\pi/2$. The intersections of these spheres and the aperture form rings on the aperture plane. Successive rings on the aperture, then, give alternately positive and negative contributions to the field received at the point on boresight - this is due to the path differences between the various rings and the field point. Following Silver, the field at the measurement point is given by the sum of the contributions of all the rings, or (assuming a whole number of Fresnel zones in the aperture)

$$F = s1 - s2 + s3 - \dots \tag{62}$$

where s1 is the magnitude of the contribution of the Ith Fresnel zone. This sum can be reduced to (approximately):

$$S = 0.5(s1 \pm sN) \tag{63}$$

depending on whether the Nth zone is odd or even. Thus sliding the field point along boresight causes the received boresight field to increase and decrease depending on the Fresnel zones which are impressed on the aperture. At distances beyond which the aperture is just a single Fresnel zone or part thereof, there is no more field fluctuation with distance. The distance at which the aperture decomposes to exactly one Fresnel zone is $D^2/4\lambda$. However, Silver points out that this distance cannot be taken as the minimum distance to the far field since the aperture edge is still 180° out of phase with the centre. At a distance of D^2/λ the aperture edge is $\pi/4$ out of phase with the centre, and this is determined to approximate the far field conditions sufficiently well (ie. phase differences across the aperture due to path differences to field points are negligible).

Silver ([13], p. 562) also quotes the radiation pattern beamwidth as an indication of the minimum distance at which the far field approximations apply. He notes that the 10 dB width of the radiation intensity pattern remains approximately constant for path distances greater than $1.5D^2/\lambda$, and notes that this agrees with the more conventional (and conservative) definition of $2D^2/\lambda$.

4.2.2 Half Power Beamwidth Criterion

There are a number of other criteria, based on a heuristic approach to this question, which could be used to quantitatively specify the minimum distance to the far field. One such heuristic criterion might be the measure of the half power beamwidth (HPBW) of the pattern at a given range from the aperture. The criterion might place a constraint on the HPBW of the pattern in the Fresnel region such that the far field approximations are reasonable if the far field and finite range HPBW values are close to each other (eg. the ratio of the Fresnel HPBW to far field HPBW evaluates to some predetermined, fixed convergence factor).

To pursue this possibility, note that there is an approximate relationship between HPBW and aperture dimension D:

$$\mathrm{HPBW} \cong \frac{\lambda}{D} = \Theta \tag{64}$$

Thus, using eqn. (64), eqn. (61) can be written as:

$$\frac{2D^2}{\lambda} = \frac{2D}{\Theta} = R \tag{65}$$

Stating Θ as $m\Theta'$, where m is a beam sharpening factor applied to a given beamwidth Θ, then eqn. (65) predicts that as the HPBW becomes smaller, the distance R to the far field increases very rapidly. For the standard or reference minimum distance to the far field (necessary in matching such a heuristic criterion to the $D^2/2\lambda$ value), it is reasonable to heuristically turn to the case of a uniform phase - uniform amplitude aperture. An estimate of a possible convergence factor may be formed by taking the minimum distance to the far field for this aperture to be $D^2/2\lambda$, calculating the HPBW at this range, and normalizing to the far field value of HPBW. In this way, the conventional definition of the minimum distance to the far field is incorporated in the heuristic HPBW criterion. The value of the convergence factor is approximately 0.982. With this criterion, the radiation patterns for several apertures may be evaluated at different distances HPBW ratio compared to the criterion of 0.982. This heuristic criterion, then, can be stated as follows: the minimum range at which the far field approximations may be applied corresponds to the range at which the ratio of the far field HPBW to the finite range HPBW evaluates to 0.982.

This criterion is only one of a number of possible criteria which could be proposed to define the outer extent of the Fresnel region. Other criteria might use some other parameter of the radiation pattern which may show a better correlation to far field distance (convergence of the finite range pattern to the far field form), while also reflecting the changes in HPBW. Such parameters could include the second moment of the main lobe (rather than HPBW), the second moment of the aperture distribution, or the second moment of the far field radiation pattern.

4.2.3 Phase-Insensitive Criterion

Another possible (heuristic) criterion for fixing the minimum distance to the far field draws on the notion that the radiation pattern is converging to the far field pattern as range increases through the Fresnel region. A mean squared difference criterion could be proposed over the visible region, between the magnitude of the finite range pattern and the magnitude of the far field pattern. This mean squared difference criterion, normalized using the far field radiation pattern, is written as:

$$\mathrm{MSD} = \frac{\int_{-k}^{k} \left|\left|\int_{-a/2}^{a/2} e^{-j\beta y^2} D(y) e^{jyk_y} dy\right|^2 - \left|\int_{-a/2}^{a/2} D(y) e^{jyk_y} dy\right|^2\right|^2 dk_y}{\int_{-k}^{k} \left|\int_{-a/2}^{a/2} D(y) e^{jyk_y} dy\right|^4 dk_y} \tag{66}$$

where $\beta = k/2\rho_o$, from eqn. (56). Eqn. (66) actually represents a normalized mean squared difference over the visible region between power patterns at finite and infinite ranges. Because power patterns are used, phase information in the radiation patterns does not affect the calculated minimum distance to the far field. For this reason, this criterion may be understood as a "phase-insensitive" criterion. This criterion applied over the visible region yields a measurably verifiable (observable) degree of convergence between the two patterns, since the evanescent fields have become attenuated to negligible value (by definition) in the Fresnel region. Including the invisible region in the criterion could alter the predicted minimum distance the far field in a way which might not bear a good correlation to the actual measurements of the finite range radiation pattern.

The reference for this phase-insensitive mean squared difference criterion may again be drawn from considering the uniform aperture case. Evaluating eqn. (66) for the uniform phase - uniform amplitude aperture at a range of $D^2/2\lambda$ yields a numerical result which defines a convergence factor which may be (heuristically) applied to other aperture distributions. Because this criterion is determined by integration over only the visible region, this numerical result is generally dependent on aperture size, since lengthening the aperture corresponds to compressing the radiation pattern in angle and so effectively increasing the amount of the radiation pattern which falls in the visible region. However, for apertures which have most of their patterns in the visible region already, the numerical result (convergence factor) is relatively insensitive to increases in aperture size.

4.2.4 Phase-Sensitive Criterion

A third (heuristic) criterion for specifying the outer extent of the Fresnel region bears a close resemblance to the phase-insensitive criterion. In this case, phase information is included in forming the mean squared difference, over the visible region, between the finite and infinite range patterns. The criterion thus uses the generally complex electric field antenna pattern rather than the power pattern, and is stated as:

$$\mathrm{MSD} = \frac{\int_{-\infty}^{\infty} \left| \int_{-a/2}^{a/2} e^{-j\beta y^2} - 1) D(y) e^{jyk_y} dy \right|^2 dk_y}{\int_{-\infty}^{\infty} \left| \int_{-a/2}^{a/2} D(y) e^{jyk_y} dy \right|^2 dk_y} \qquad (67)$$

Note that the normalization of the criterion is with respect to the radiated power of the aperture. Also, the expression includes integration over the visible and invisible regions. Although evanescent fields have disappeared at the outer extent of the Fresnel region, from a strictly theoretical point of view they have equal weight as the fields of the visible region. This argument could also be applied to the phase-insensitive criterion, though the interpretation of the power pattern in the invisible region becomes uncertain since there is no net power flow associated with the evanescent waves described by this region. By example, the phase-sensitive criterion seems to have some advantage over the phase-insensitive criterion. Considering the case of an aperture with delta functions at its edges, the phase-insensitive criterion (eqn. (66)) evaluates to:

$$\mathrm{MSD} = \frac{\int_{-k}^{k} \left| \left| e^{-j\beta\frac{a^2}{4}} \left(e^{jk_y\frac{a}{2}} + e^{-jk_y\frac{a}{2}}\right)\right|^2 - \left|e^{jk_y\frac{a}{2}} + e^{-jk_y\frac{a}{2}}\right|^2 \right|^2 dk_y}{\int_{-k}^{k} \left|\left(e^{jk_y\frac{a}{2}} + e^{-jk_y\frac{a}{2}}\right)\right|^4 dk_y} = 0 \qquad (68)$$

Clearly this is inconsistent with actual measurements that could be taken. This discrepancy is attributable to the fact that eqn. (66) ignores the phase information of the radiation pattern.

The above proposed criteria each have individual merits, and are quantitatively investigated for the cases of 16 different uniform phase aperture distributions. The results of this investigation constitute the following section.

4.3 Investigation of the Minimum Far Field Criteria

The three minimum far field criteria all have in common that the aperture distribution plays a determining role in specifying the minimum distance at which the far field approximations may be applied (the outer extent of the Fresnel region). The motivation behind this analysis is to determine the extent of the Fresnel region of superdirective aperture antennas, and so check the hypothesis that such antennas can exhibit large Fresnel regions. To facilitate this, 16 different uniform phase/nonuniform amplitude aperture distributions are chosen arbitrarily, and each of the minimum far field criteria of the previous section is applied to these apertures. The goal of this analysis is to draw generalizations concerning the relationship between aperture distribution and the minimum distance at which the far field approximations may be applied. Uniform phase apertures are, by definition, not superdirective, and so are well suited to determining the merit (in terms of conventional aperture distributions) of a given minimum far field criterion. The selected aperture distributions are specified in Appendix B, and are named and indexed below:

1 - UNIFORM
2 - DOUBLE COSINE
3 - TRIANGLE
4 - COSINE
5 - EDGE
6 - COMPLEMENT OF DOUBLE COSINE
7 - NOTCH
8 - FILLED EDGE
9 - EDGE + COMPLEMENT OF DOUBLE COSINE
10 - 1 + TRIANGLE
11 - DOUBLE TRIANGLE
12 - COMPLEMENT OF DOUBLE TRIANGLE
13 - KNIFE EDGE
14 - COSINE SQUARED
15 - TRIANGLE SQUARED
16 - TRIANGLE^4

The aperture model used in the analysis is that shown in fig. 1. That is, the aperture is one dimensional (line source) and lies along the y-axis. The aperture radiates into the half space $x > 0$, and the magnetic field lies along the z-axis. All simplifications drawn in Section 1 pertaining to this figure apply in the following analysis.

Fig. 8 shows a graph of normalized HPBW vs. distance from the aperture for the first four aperture distributions. From the graph, the rate at which the respective distributions converge (in a HPBW sense presented earlier) with increasing range. The far field HPBW values for the first four aperture distributions are given below:

UNIFORM	: 0.2783
DOUBLE COSINE	: 0.281
TRIANGLE	: 0.4009
COSINE	: 0.3747

(These HPBW values are measured in the k_y plane). The DOUBLE COSINE distribution has a wider main beam than the uniform distribution, and converges quicker than the uniform case. However, the cosine distribution has a narrower main beam than the triangle distribution, and converges faster. Because of this and by exception, a simple minimum far field criterion based on the value of the HPBW at a finite range normalized to the far field value is unworkable, at least as formulated here.

Intuitively, the reason for this is probably that HPBW is too coarse or arbitrary a parameter of a radiation pattern to be directly related to the minimum range to the far field. Analytically, there is no rigorous obvious reason to expect that the HPBW would be an indicator of the minimum range to the far field. The motivation for investigating this possibility is that, by approximating the HPBW by λ/D, the traditional definition of the minimum range to the far field ($2D^2/\lambda$) can be written in terms of the HPBW. The approximation of HPBW as a function of λ and D is quite rough - only a rule of thumb, as shown by the example distributions below:

Distribution	Half Power Beamwidth
UNIFORM	0.88 /D
TRIANGLE	1.27 /D
PARABOLIC	1.16 /D
SQUARED PARABOLIC	1.36 /D
SEMICIRCLE	1.02 /D

Thus criterion concerning the half power beamwidth ratio is excluded from further discussions. This does not necessarily mean that the minimum range to the far field might not be a function of HPBW by some other criterion, nor does the crudeness of the λ/D approximation devalue the hypothesis that narrow beam patterns have a large Fresnel region.

The second minimum distance criterion described in the previous section is given by eqn. (66), and is a (phase-insensitive) mean squared difference of finite and infinite range power patterns, normalized using the radiation pattern. Figs. 9a and 9b are graphs of this mean squared difference (MSD) criterion vs. distance from the aperture. Each curve represents a different aperture distribution, and shows how the respective power pattern converges to the far field pattern. Using this criterion, the far field is defined to begin at a range for which the phase-insensitive MSD expression evaluates to 0.04794 (the value of the expression for a uniform aperture at a range of $D^2/2\lambda$). All apertures in the numerical calculations have a dimension of 10λ.

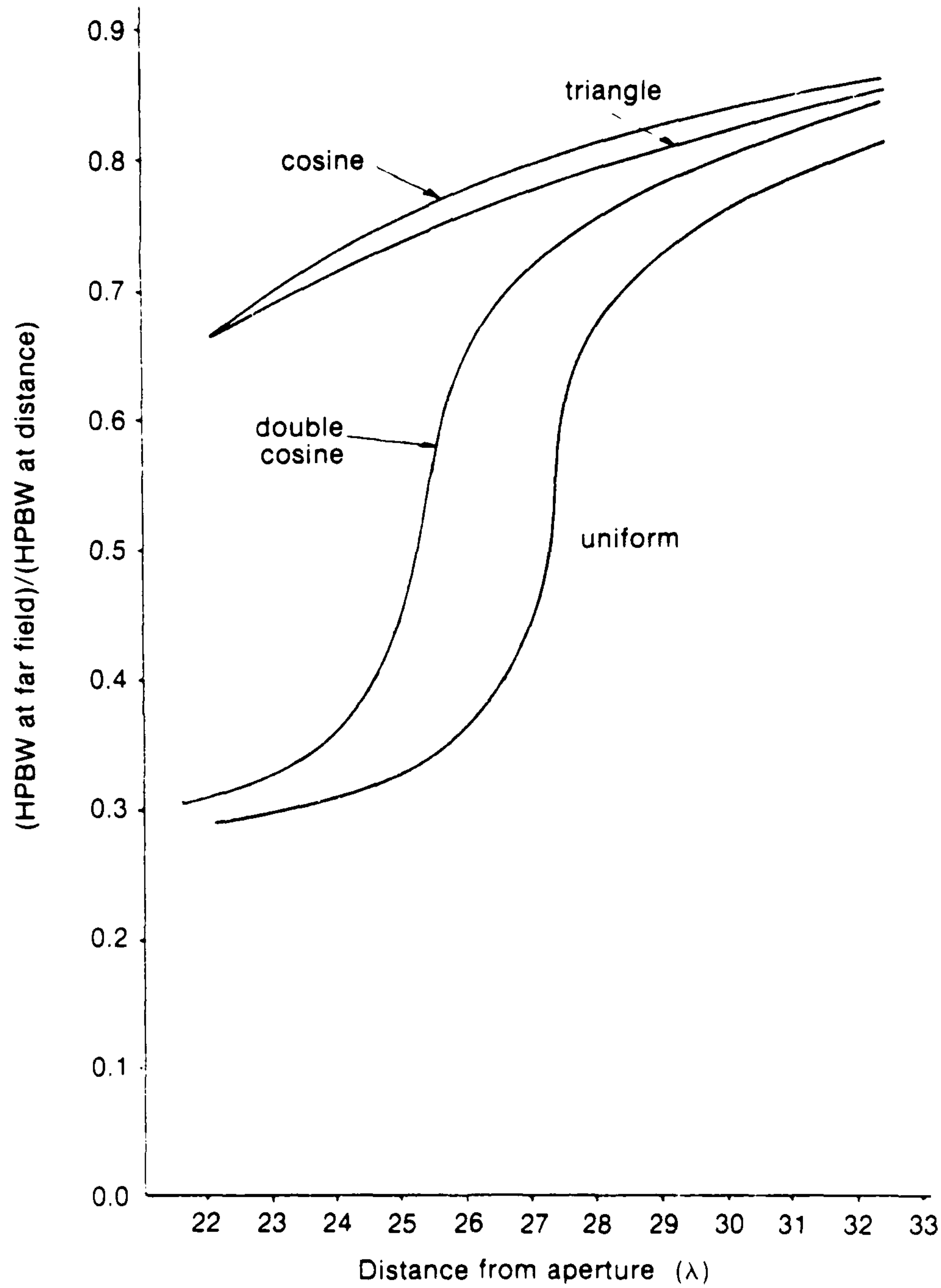


FIG. 8: GRAPH OF NORMALIZED HALF POWER BEAMWIDTH VS DISTANCE FROM THE APERTURE

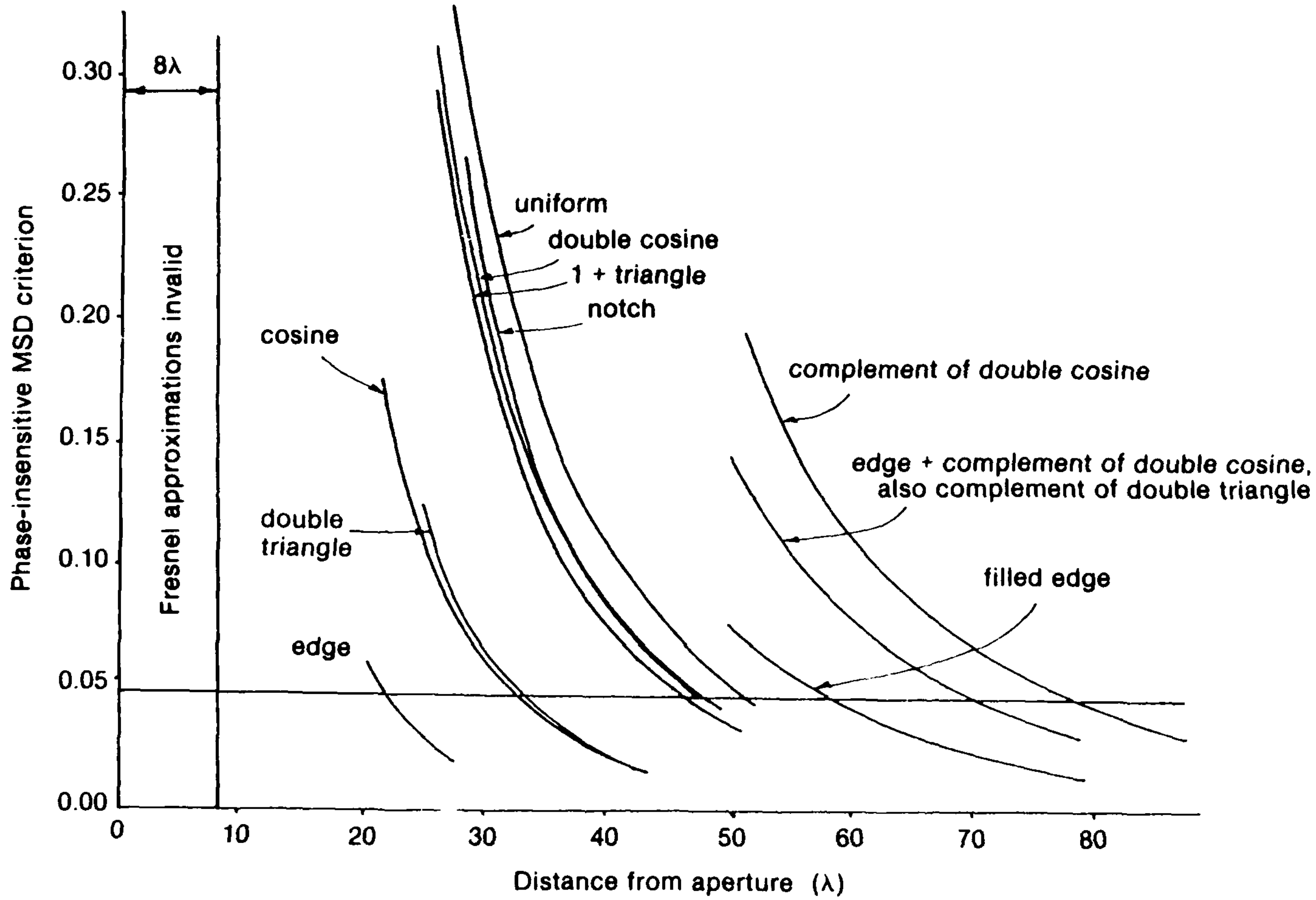


FIG. 9a: GRAPH OF MEAN SQUARED DIFFERENCE CRITERION VS. DISTANCE FROM THE APERTURE

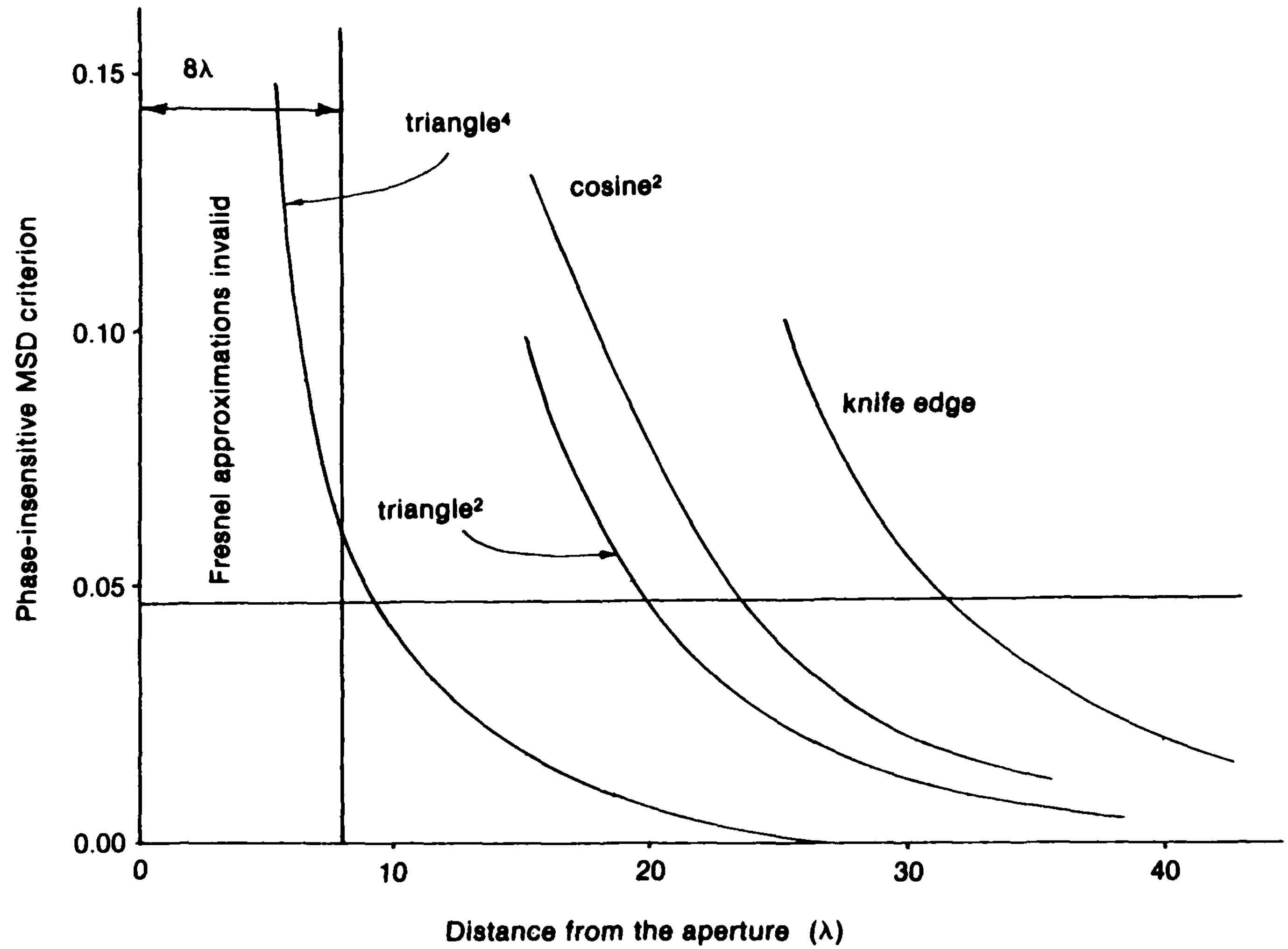


FIG. 9b: GRAPH OF MEAN SQUARED DIFFERENCE CRITERION VS. DISTANCE FROM THE APERTURE (CONTINUED)

Figs. 9a and 9b are used to determine the range at which the phase-insensitive MSD criterion is satisfied for each of the 16 aperture distributions. This range for each aperture is the range at which the respective MSD curve intercepts the line drawn at MSD = 0.04794. Fig. 10 shows a graph of HPBW vs. minimum range to the far field, according to this criterion. A general trend of increased minimum range to the far field with decreased HPBW is evident. Some data points directly refute this relationship for this criterion. In any case, the scatter of data points is too great to fit a best curve and draw meaningful quantitative conclusions about the minimum range to the far field from the HPBW of the a given radiation pattern.

In order to define a more rigorous index of aperture width, to reflect the aperture distribution, the physical width D is replaced by the second moment D_2 of the aperture distribution. A possible correlation between this second moment D_2 and the *minimum* range to the far field is then explored using the above data for the 16 aperture distributions. Fig. 11 shows a graph of HPBWxD_2 vs. distance from the aperture. Since HPBW is proportional to D for a given aperture distribution, it is also proportional to D_2, and scaling the vertical axis in this way implies that the graph of fig. 11 is independent of aperture size for apertures larger than 10λ (providing that the radiation pattern of a given aperture of 10λ lies principally in the visible region). The horizontal axis is scaled with respect to D^2/λ). The data points of fig. 11 appear randomly scattered, thus there does not seem to be a useful relationship between the second moment of the aperture distribution and the minimum range to the far field, for this phase-insensitive MSD criterion.

Still another possibility in finding a correlation between the beamwidth and the minimum range to the far field is to replace HPBW with the second moment of the radiation pattern, calculated over the visible region (since for the example distributions, the pattern in the invisible region is essentially zero). Fig. 12 shows a graph of the standard deviation of the power pattern vs. minimum range to the far field using the phase-insensitive MSD criterion. Again, the data points of the 16 aperture distributions are almost randomly scattered. A similar graph using the standard deviation of the main lobe also shows a random scattering of points. Thus, as for the previous minimum range criterion, using the phase-insensitive MSD criterion there appears to be no useful relationship between the beamwidth of the radiation pattern and the minimum range at which the far field approximations may be applied.

The last criterion for the outer extent of the Fresnel region to be considered is given by eqn. (67), and is the mean squared (magnitude) difference between the complex electric field pattern of the aperture, calculated at finite and infinite range from the aperture. This criterion, unlike the previous MSD criterion, includes phase information of the radiation pattern. Also, for reasons stated previously, this MSD criterion is applied over the entire radiation pattern. A brief discussion of the form that the phase-sensitive MSD criterion should have in the invisible region is found in Appendix A. As for the case of the phase-insensitive criterion, the convergence factor of the phase-sensitive criterion is taken as eqn. (67) evaluated at a range of $D^2/2\lambda$ for a uniform phase/uniform amplitude aperture of dimension 10λ.

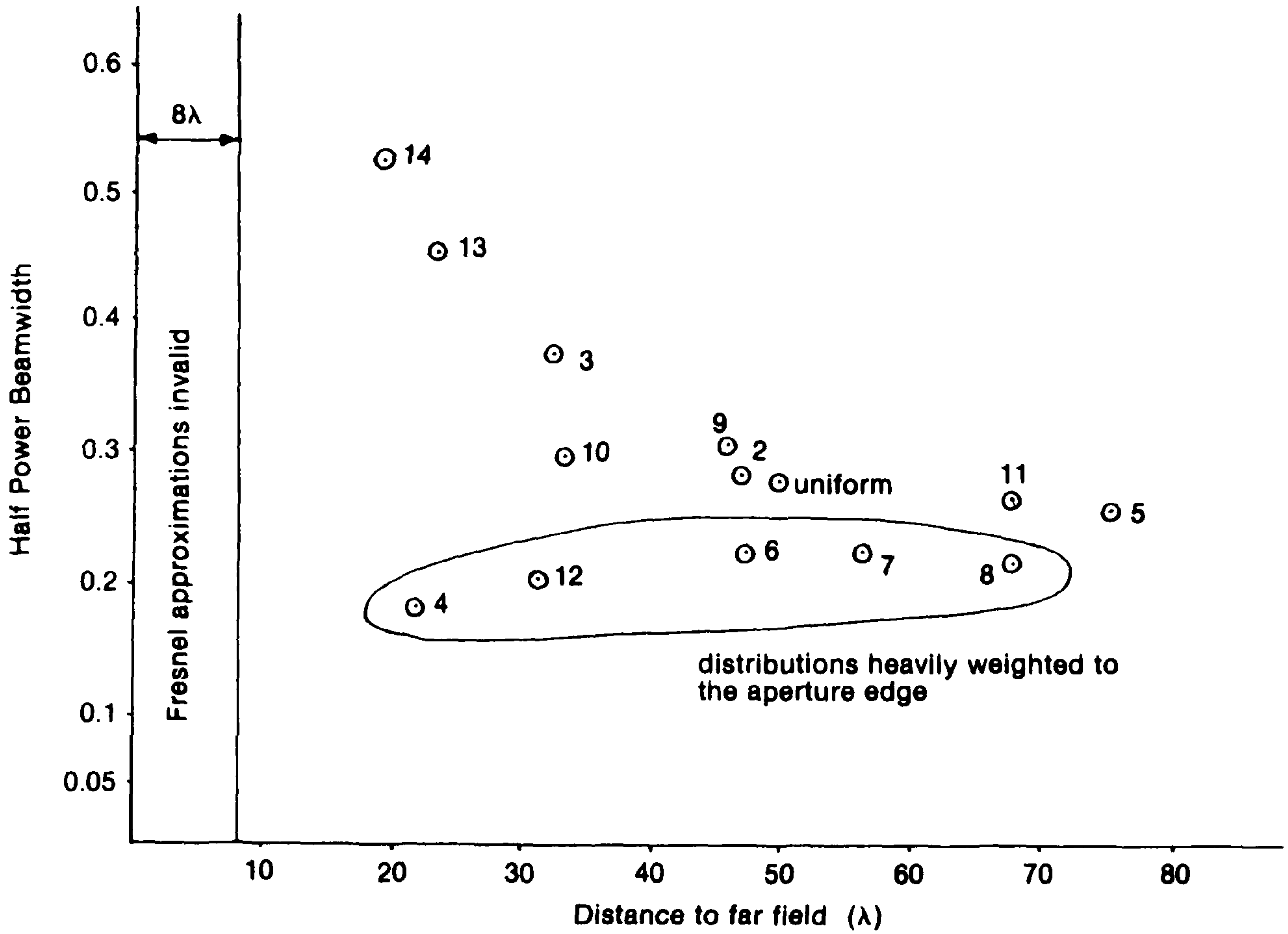


FIG. 10: GRAPH OF HPBW VS. DISTANCE TO THE FAR FIELD (PHASE-INSENSITIVE CRITERION)

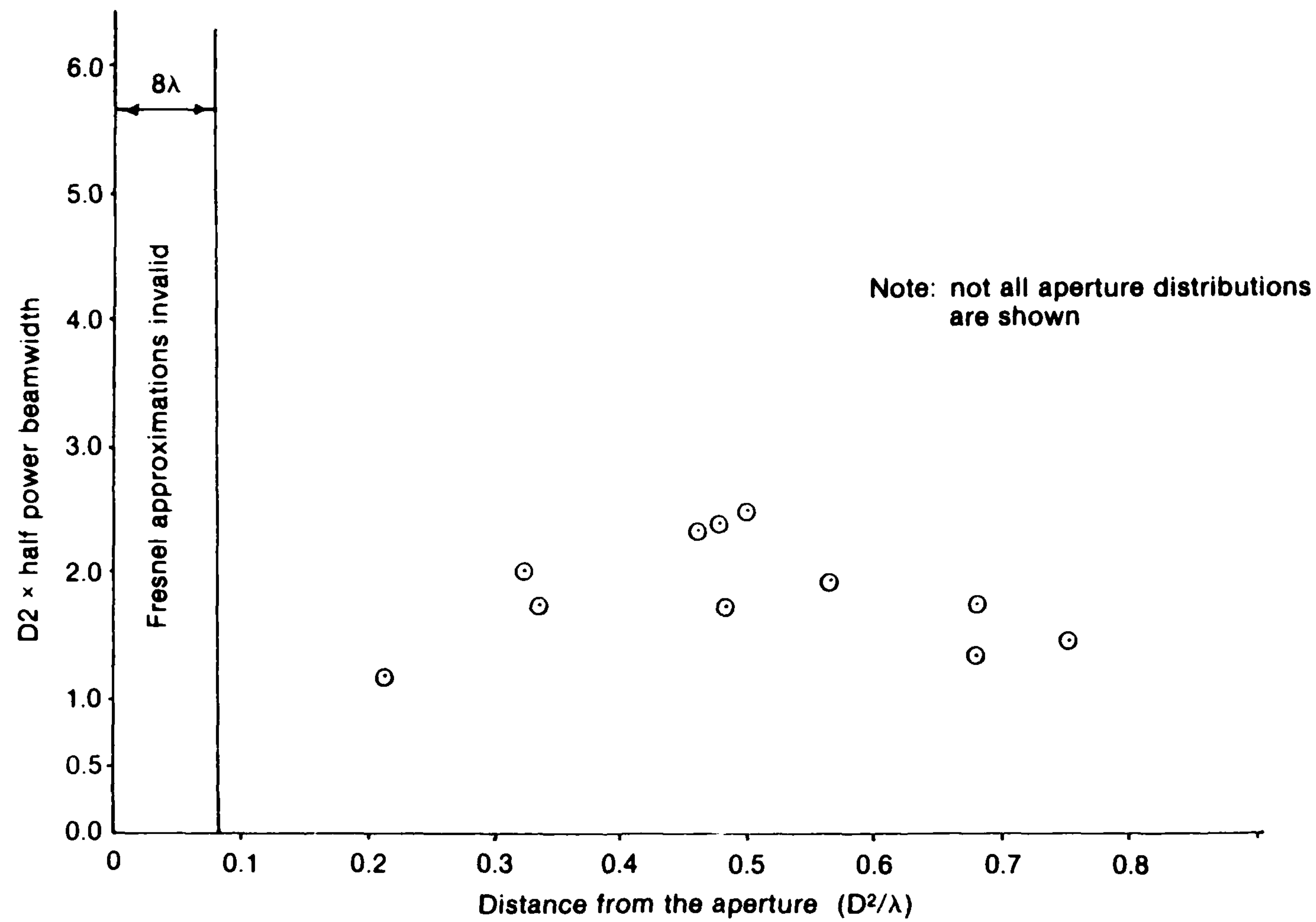


FIG. 11: GRAPH OF D2 X HPBW VS. DITANCE FROM THE APERTURE (PHASE-INSENSITIVE CRITERION)

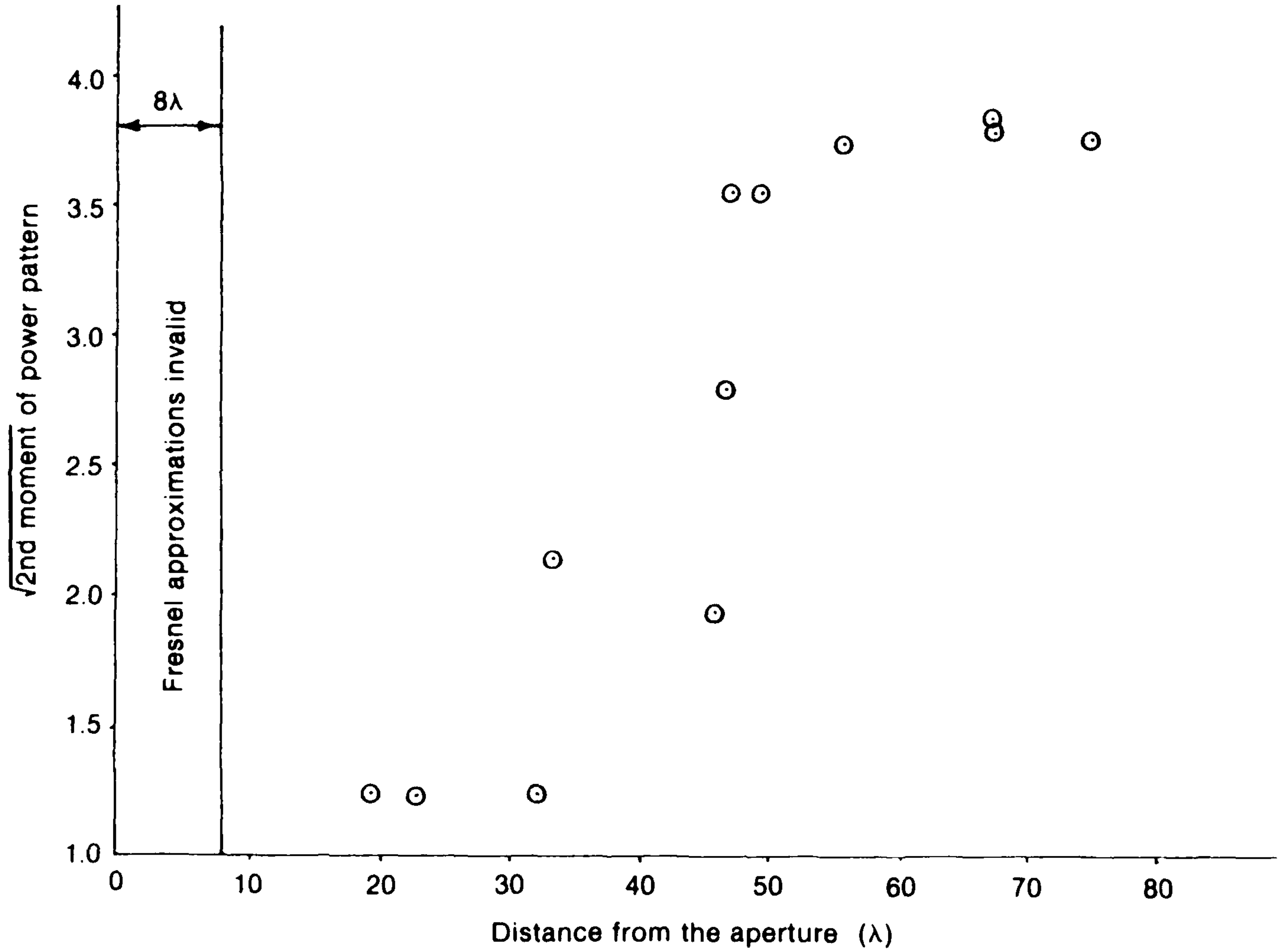


FIG. 12: GRAPH OF 2ND MOMENT OF THE POWER PATTERN VS. DISTANCE FROM THE APERTURE (PHASE-INSENSITIVE CRITERION)

The range at which the phase-sensitive MSD criterion satifies the convergence factor may be determined from fig. 13, for each of the 16 subject aperture distributions. Fig. 13 shows a graph of the phase-sensitive MSD criterion vs. range from the aperture for the 16 aperture distributions. The convergence factor for the phase-sensitive criterion is 0.44443. The curves have the same general shape as those of figs. 9a and 9b, but the relative positions of the curves are shifted. Fig. 14 shows a graph of HPBW vs. minimum range to the far field. Unlike the case of the phase-insensitive criterion, there here appears to be a good correlation between HPBW and the minimum range to the far field as specified by the phase-sensitive MSD criterion. Fig. 14 is independent of aperture size since the integration of eqn. (67) is over the entire radiation pattern. The apparent correlation between HPBW and minimum range to the far field is striking, and seems in some sense difficult to explain, since it appears to be better than the original approximation of HPBW by λ/D.

A graph of standard deviation of the main lobe vs. minimum range to the far field is presented in fig. 15. The correlation between beamwidth and minimum range to the far field is still good, though not as good as for HPBW. Fig. 16 shows a graph of the standard deviation of the radiation pattern vs. minimum range to the far field, and the scatter of the points is excessive. It appears that, in heuristically considering the possibility of a correlation between parameters of the radiation pattern and the minimum range to the far field, the correlation is strongest for the HPBW of the radiation pattern when using the phase-sensitive MSD criterion.

5.0 THE PHASE-SENSITIVE CRITERION

5.1 A Synthesis Form of the Phase-Sensitive Criterion

In the previous section, a correlation has been observed between HPBW and the minimum range to the far field (outer extent of the Fresnel region) if the phase-sensitive MSD criterion is used. However, the form of the relationship is not obvious from eqn. (67). In order to gain insight into this correlation and further investigate the properties of the phase-sensitive MSD criterion, Parseval's Theorem may be applied to eqn. (67). Parseval's Theorem relates power in the antenna pattern domain in this case, and may be expressed as:

$$\int_{-\infty}^{\infty} |x(t)|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |X(\omega)|^2 d\omega \qquad (69)$$

Since the integral expressions of the phase-sensitive MSD criterion include squared magnitudes, Parseval's Theorem is well suited to simplifying eqn. (67). The denominator of the phase-sensitive MSD criterion (eqn. (67)) may be written, using eqn. (69), as:

$$\int_{-\infty}^{\infty} \left| \int_{-a/2}^{a/2} D(y) e^{jyk_y} dy \right|^2 dk_y = 2\pi \int_{-a/2}^{a/2} |D(y)|^2 dy \qquad (70)$$

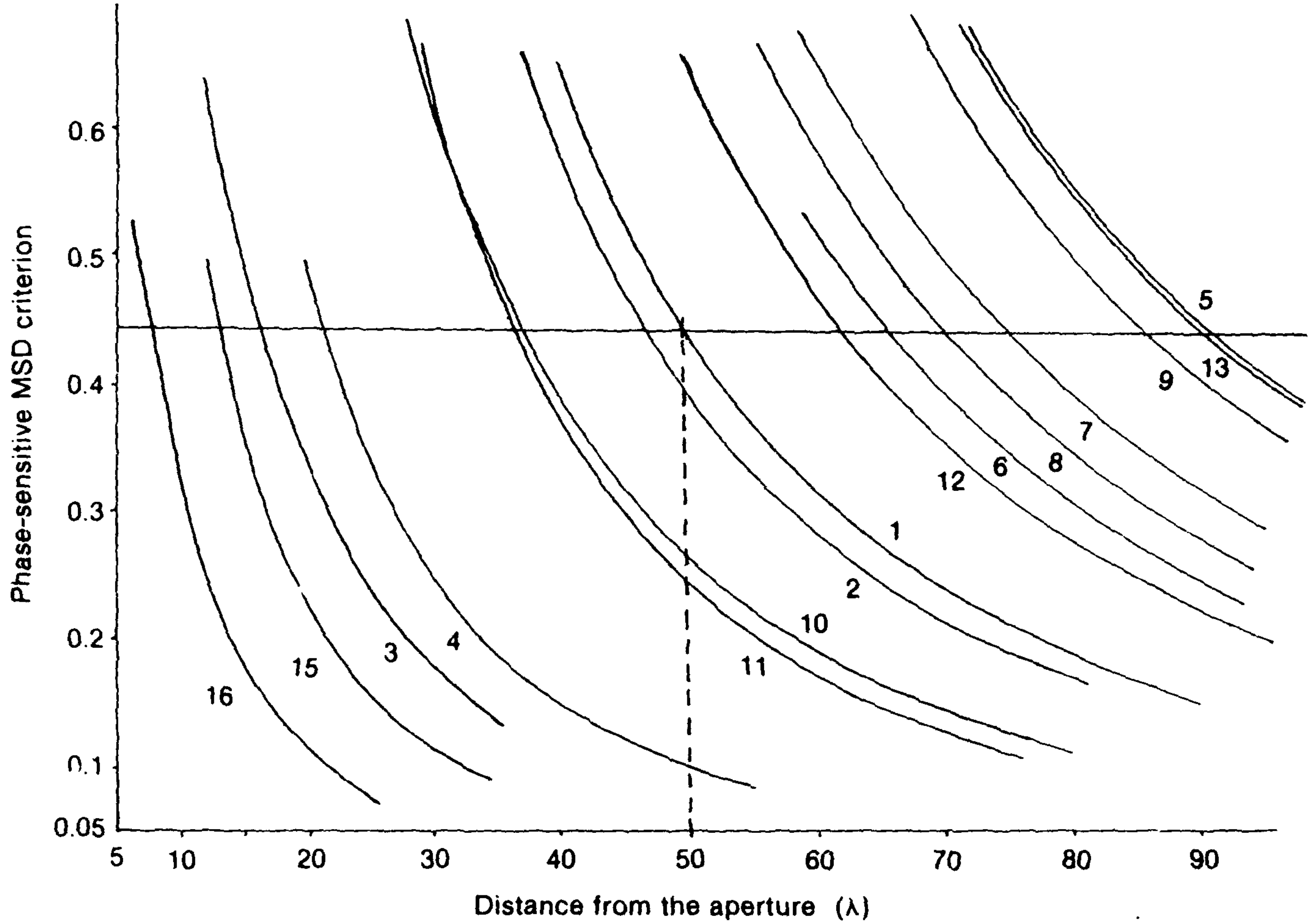


FIG. 13: *CONVERGENCE OF RADIATION PATTERNS*

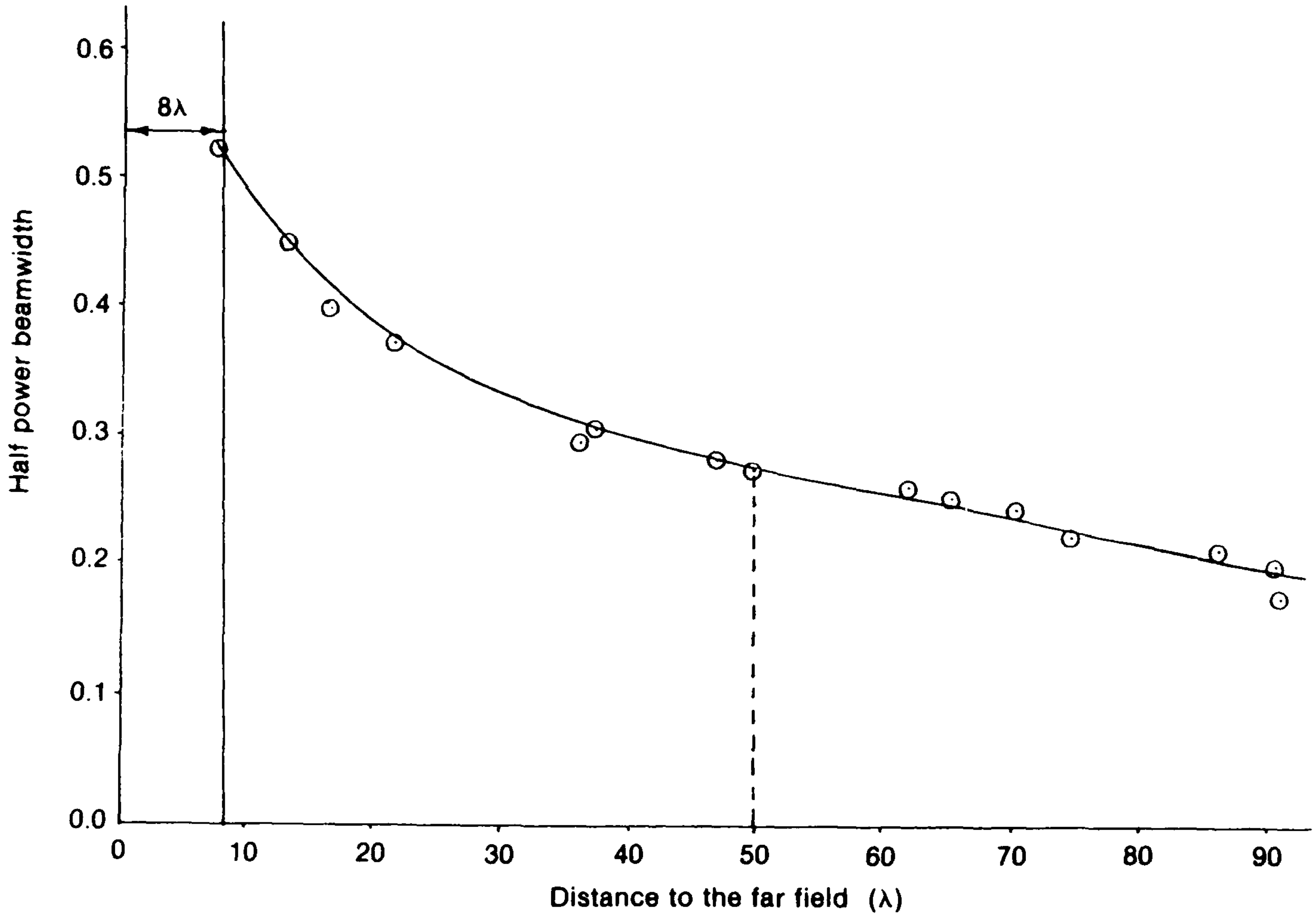


FIG. 14: GRAPH OF HPBW VS. DISTANCE TO THE FAR FIELD (PHASE-SENSITIVE CRITERION)

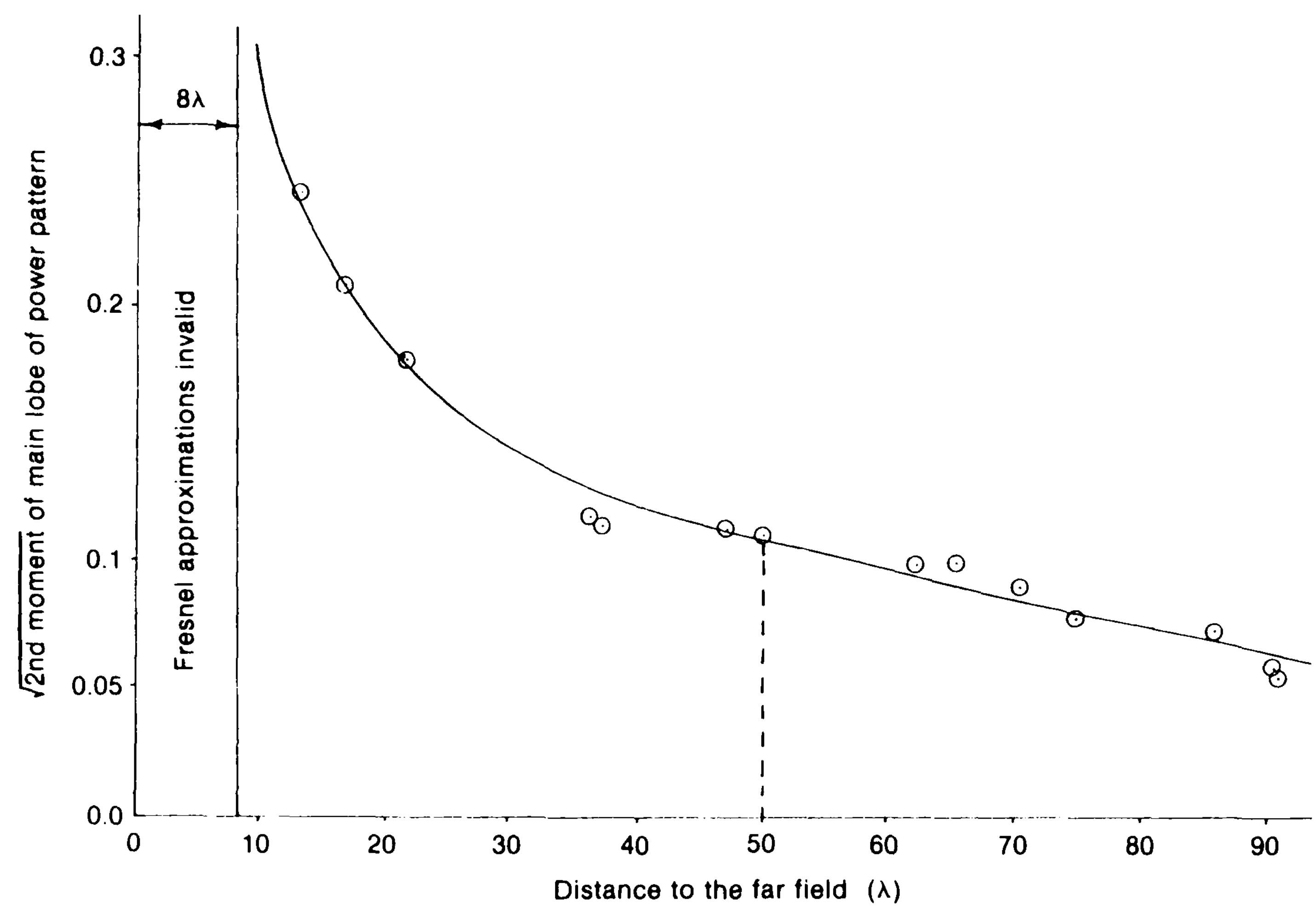


FIG. 15: GRAPH OF √2ND MOMENT OF MAIN LOBE OF POWER PATTERN VS. DISTANCE TO THE FAR FIELD (PHASE-SENSITIVE CRITERION)

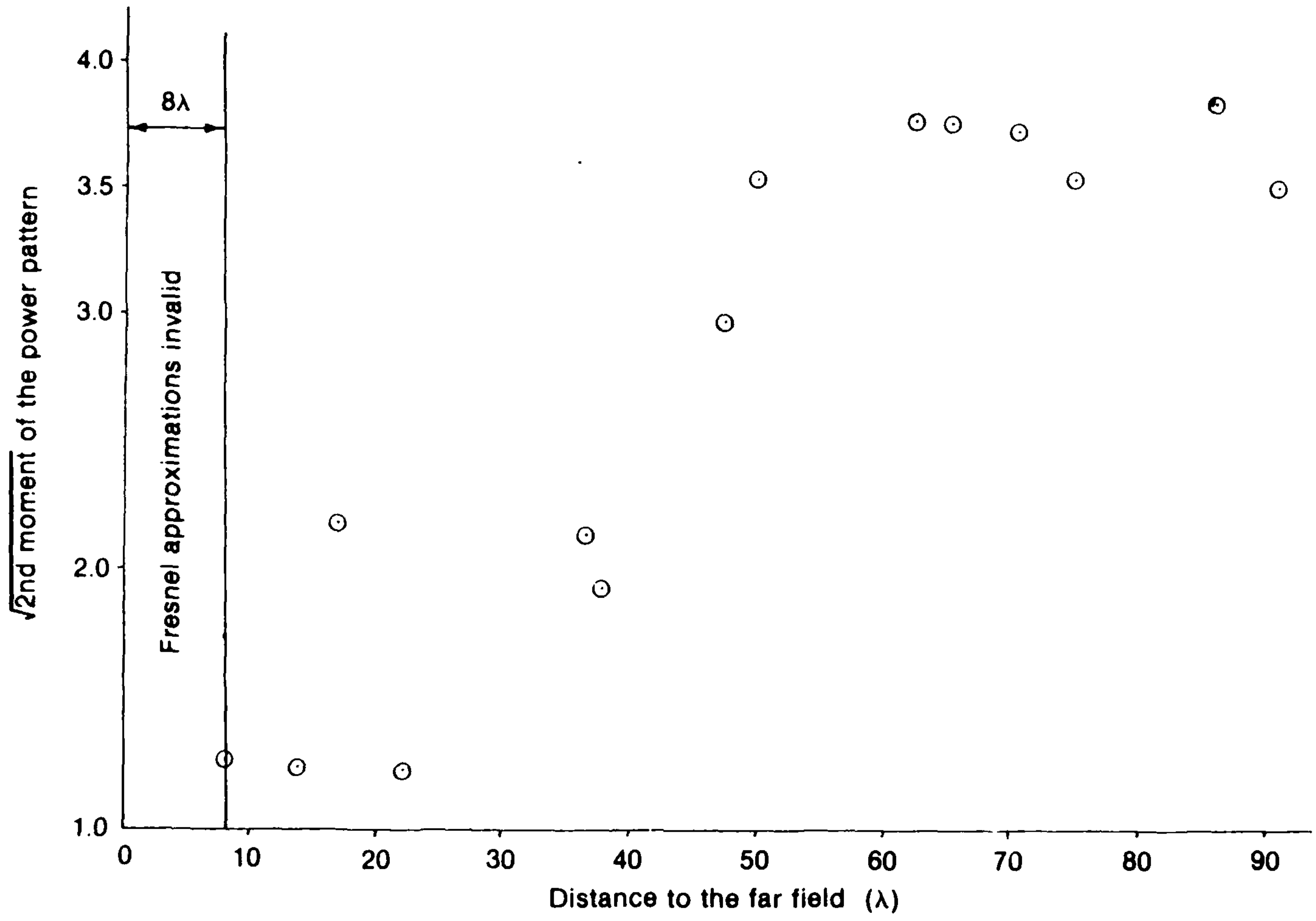


FIG. 16: GRAPH OF $\sqrt{\text{2ND MOMENT}}$ OF POWER PATTERN VS. DISTANCE TO THE FAR FIELD (PHASE-SENSITIVE CRITERION)

Similarly, the numerator of eqn. (67) may be written as:

$$\int_{-\infty}^{\infty} \left| \int_{-a/2}^{a/2} (e^{-j\beta y^2} - 1) D(y) e^{jyk_y} dy \right|^2 dk_y = 2\pi \int_{-a/2}^{a/2} \left| (e^{-j\beta y^2} - 1) D(y) \right|^2 sy \tag{71}$$

The limits of integration of integrals in both the numerator and the denominator are now finite. Recalling that the convergence factor for this criterion is 0.44443, the phase-sensitive criterion may be written as:

$$\frac{\int_{-a/2}^{a/2} \left| (e^{-jBy^2} - 1) D(y) \right|^2 dy}{\int_{-a/2}^{a/2} \left| D(y) \right|^2 dy} = 0.44443 \tag{72}$$

The problem of determining the minimum range at which the far field approximations may be applied has now been transferred from the antenna pattern domain, where intuition works well, to the aperture domain, where intuition is less effective. With this change of domain comes the possibility of determining aperture distributions which have a desired outer range of the Fresnel region. The aperture form of the criterion suggests a synthesis equation in this sense. Pursuing this, the phase-sensitive MSD equation can be written as:

$$\int_{-a/2}^{a/2} \left| (e^{-j\beta y^2} - 1) D(y) \right|^2 dy - 0.44443 \int_{-a/2}^{a/2} \left| D(y) \right|^2 dy = 0 \tag{73}$$

Merging the integral signs and factoring out D(y), eqn. (73) can be written as:

$$\int_{-a/2}^{a/2} \left\{ \left| (e^{-j\beta y^2} - 1) \right|^2 - 0.44443 \right\} \left| D(y) \right|^2 dy = 0 \tag{74}$$

Expanding the first squared magnitude term simplifies the expression to:

$$\int_{-a/2}^{a/2} \left\{ (\cos(\beta y^2) - 1)^2 + \sin^2(\beta y^2) - 0.44443 \right\} \left| D(y) \right|^2 dy = 0 \tag{75}$$

In some sense eqn. (75) may be interpreted as a synthesis equation for constructing aperture distributions which have a specified outer extent of the Fresnel region, as defined by the phase-sensitive MSD criterion. The outer extent of the Fresnel region is specified by choosing a value for β, since $\beta = k/2\rho_0$ (see eqn. (56)) is inversely proportional to the range at which the phase-sensitive MSD criterion is evaluated.

The integrand of eqn. (75) is the product of two functions, and the expression tells us that the area under the composite function must be zero at the range equal to the outer extent of the Fresnel region, for a given aperture. At ranges greater than the outer extent of the Fresnel region, eqn. (75) yields a negative value (ie. in the far field according to this criterion, eqn. (75) is negative). The second term of the integrand is a squared magnitude, and so is positive definite. Thus the balancing of the positive and negative contributions to the integral is controlled by the first term of the integrand, and by the limits on the integral. Adjusting the limits of the integral corresponds to altering the aperture size in order to achieve a certain range to the outer bound of the Fresnel region. Closer examination of the first term in the integrand shows that, as expected, large apertures imply a relatively large Fresnel region. This can be understood by noting that the first term of the integrand is a function of β, and so is determined by the choice of the range to the outer extent of the Fresnel region. For relatively large range from the aperture (relatively small values of β, see eqn. (56)), this function is a quadratic curve with a negative dc offset of 0.4443. The first term of the integrand is shown in fig. 17. As the value of β becomes small, the quadratic curve flattens out. The region of the integrand which contributes positively to the integral becomes pinched off at the edges of the aperture. With this in mind, the phase-sensitive MSD criterion can more properly be written as an inequality.

In the general case, a value for β is determined for which the criterion evaluates to zero. Since this range is by definition in the far field, values of range larger than this (smaller values of β) must also be in the far field of the aperture. Since larger values of range (smaller values of β) cause the integral of eqn. (75) to become negative, the "=" of eqn. (75) may be replaced with "<". This results in a subtle change in the interpretation of the phase-sensitive MSD equation (eqn. (75)): values of β for which the integral is negative lie in the far field, and the outer extent of the Fresnel region lies at a range corresponding to the largest value of β which satisfies the inequality.

With this more proper form for the criterion and returning to the observation that the positive contribution in the integral is pinched off at the edges of the aperture with decreasing β, additional conclusions may be drawn. Aperture weighting in the negative region of the integrand acts to move the outer extent of the Fresnel region closer to the aperture. Weighting the aperture distribution in the positive region of the integrand acts to move the outer extent of the Fresnel region farther from the aperture. Positive contributions are at and near the aperture edges, and negative contributions are centred near the centre of the aperture. The transition point between

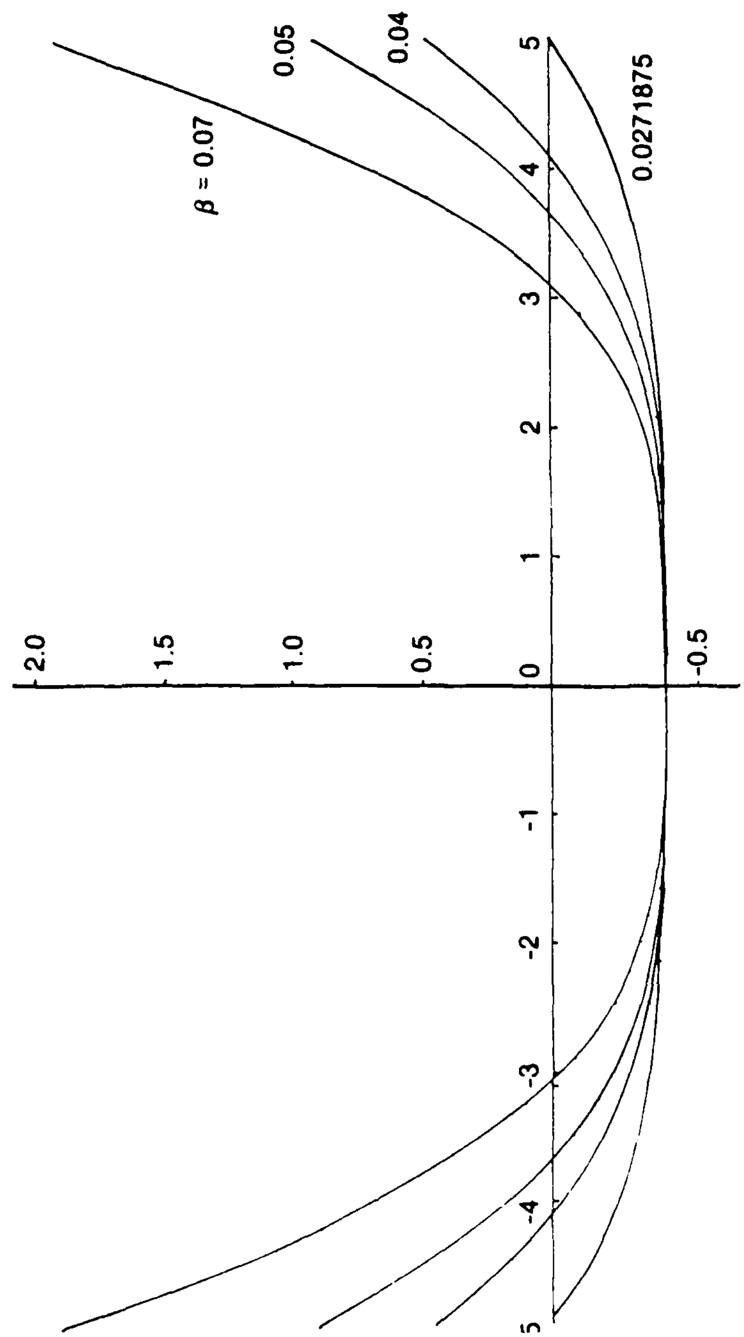


FIG. 17: GRAPH OF THE QUADRATIE PHASE CURVATURE ACROSS THE APERTURE FOR FINITE RANGES

positive and negative contributions is determined by the value of β. In order to synthesize an aperture with the largest possible range to the outer boundary of the Fresnel region, an aperture weighting should be selected which is heavily weighted at the aperture edges, with zero or reduced weighting near the centre. In the limit the aperture distribution with the largest range to the outer Fresnel region is one consisting of delta functions at the at the aperture edges, zero valued otherwise. The largest value of β which satisfies the phase-sensitive MSD criterion for such an aperture is that value for which the first term of the integrand of eqn. (75) is zero valued at the aperture edges. Due to the negative offset of 0.44443, the integrand is negative over the entire aperture, though for this special case this region is given zero weight. For a 10λ aperture the maximum value of β is 0.02719, which corresponds to an outer Fresnel region (minimum far field range) of $1.156D^2/\lambda$ (compared with the reference range of $D^2/2\lambda$ for the uniform aperture).

With this understanding, it is appropriate to reconsider the graph of fig. 14. The "delta" aperture has a radiation pattern corresponding to a sine function and has the narrowest main lobe of all 16 aperture distributions considered in this investigation. It contributes a data point in this graph at approximately 116λ, in agreement with the trend of the data. Further insight into the relationship between beamwidth and the outer extent of the Fresnel region is obtained by examining the form of the second term (positive definite term) of the integrand of eqn. (75).

The second term of the integrand is the squared magnitude of the aperture distribution. This means that the phasing of the aperture distribution is actually ignored by the phase-sensitive MSD criterion. (note that the term "phase-sensitive" criterion refers to the fact that the phase information in the radiation pattern domain is used in determining the degree of convergence of the finite range pattern to the far field form, and does not imply conditions of the criterion when translated to the aperture domain). The fact that the phase information of the aperture is ignored by this criterion implies that the graph of fig. 14 indicates a trend, but does not continue past a minimum range of approximately 116λ. According to the phase-sensitive MSD criterion and using the (artificial) convergence factor derived from $D^2/2\lambda$, the outer extent of the Fresnel region for an aperture antenna lies between the minimum range of validity of the criterion (8λ) and $1.1557D^2/\lambda$ (which occurs for the case of an aperture with delta functions at the aperture edges). This statement applies irrespective of aperture amplitude and/or phase distribution.

The fact that the phase of the aperture distribution is ignored by the phase-sensitive MSD criterion means that the original hypothesis, that a superdirective aperture has a large Fresnel region, is invalid in the context of this analysis. Such apertures rely on phasing effects to achieve narrow beamwidths, and it has become apparent that while such phasing does alter the shape of the radiation pattern, the mean squared difference between the finite range and far field pattern shapes is not affected by this phasing. This conclusion has implicit to it a qualifications which permeates all mathematical considerations in the analysis thus far. The qualification is that range to the outer extent of the near field should not excede the range to the outer extent of the Fresnel region. Theoretically there does exist the possibility that the range at which the reactive so-called "near field" could outstrip (in range) geometrical or path difference Fresnel effects. Such a condition might arise for the case of a strongly superdirective antenna. The effect of such a near field boundary and Fresnel boundary inversion on the phase-sensitive MSD criterion is beyond the scope of this investigation. Thus the conclusion that the hypothesis has been disproved may only be justified for the case of marginally superdirective antennas. Also, the conclusion that the hypothesis is incorrect is strictly in terms of the phase-sensitive MSD criterion, though this criterion is reasonable in the sense of indicating the validity of the far field approximations, and is conventional in the sense of using the normalized squared difference.

5.2 Half Power Beamwidth and the Minimum Far Field Distance

The previously presented (fig. 14) relationship between HPBW and the range of the outer extent of the Fresnel region is now investigated using the modified form (eqn. (75) of the phase-sensitive MSD criterion. The criterion, expressed appropriately as an inequality, is repeated below for convenience:

$$\int_{-a/2}^{a/2} \{(\cos(\beta y^2) - 1)^2 + \sin^2(\beta y^2) - 0.44443\} |D(y)|^2 dy \leq 0 \qquad (76)$$

As stated in the previous section, the first factor of the integrand is a quadratic curve with a negative dc offset, while the second term is positive definite. Thus the first factor controls those areas of the aperture distribution which contribute positively or negatively to the integral. Further, the shape of the quadratic curve is related to the range between the observer and the aperture. As the range increases, the quadratic curve becomes flatter over the integration interval - a factor which is important in explaining the observed relationship between HPBW and the range to the outer extent of the Fresnel region. This flattening of the quadratic curve over the integration interval is reflected in the plot of the trajectories of its zero crossings as β is varied, shown in fig. 19.

In order to carry out an investigation of this observed HPBW relationship, a family of continuously varying apertures (considered typical of the aperture distributions considered thus far) are considered. These distributions are shown in fig. 18.

In the first case, the aperture distribution under consideration has zero weight over a symmetrical interval centred on the aperture, and uniform amplitude and phase elsewhere. The width "b" of the nonzero intervals is a variable of the aperture distribution. In the second case, the nonzero interval of the aperture is symmetrical about the centre of the aperture and has width "c". Varying "c" corresponds to varying the width of a uniform aperture. At the range at which the inequality of eqn. (76) is just met, the zero crossings of the quadratic curve must occur somewhere in the nonzero

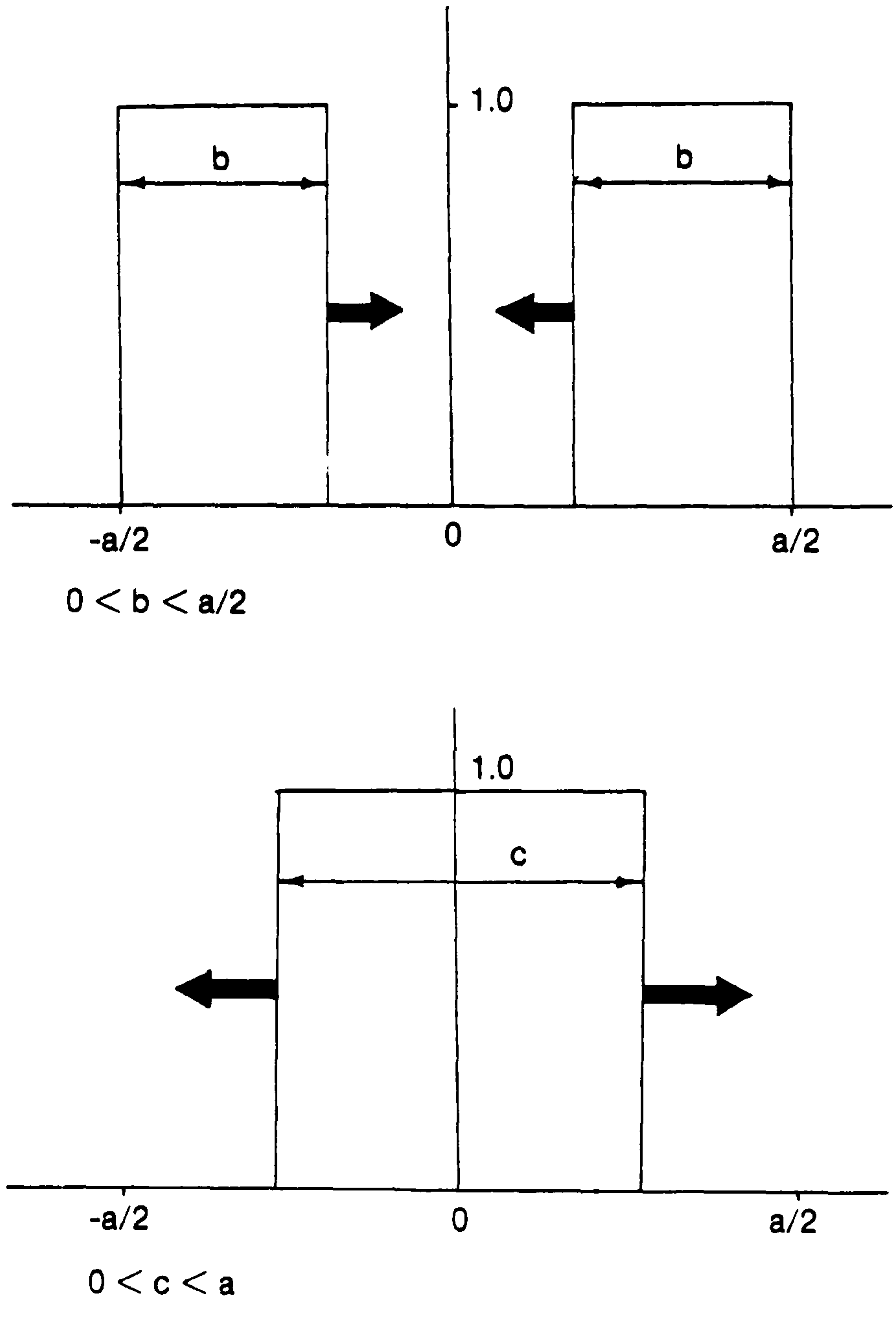


FIG. 18: A FAMILY OF CONTINUOUSLY VARYING APERTURE DISTRIBUTIONS

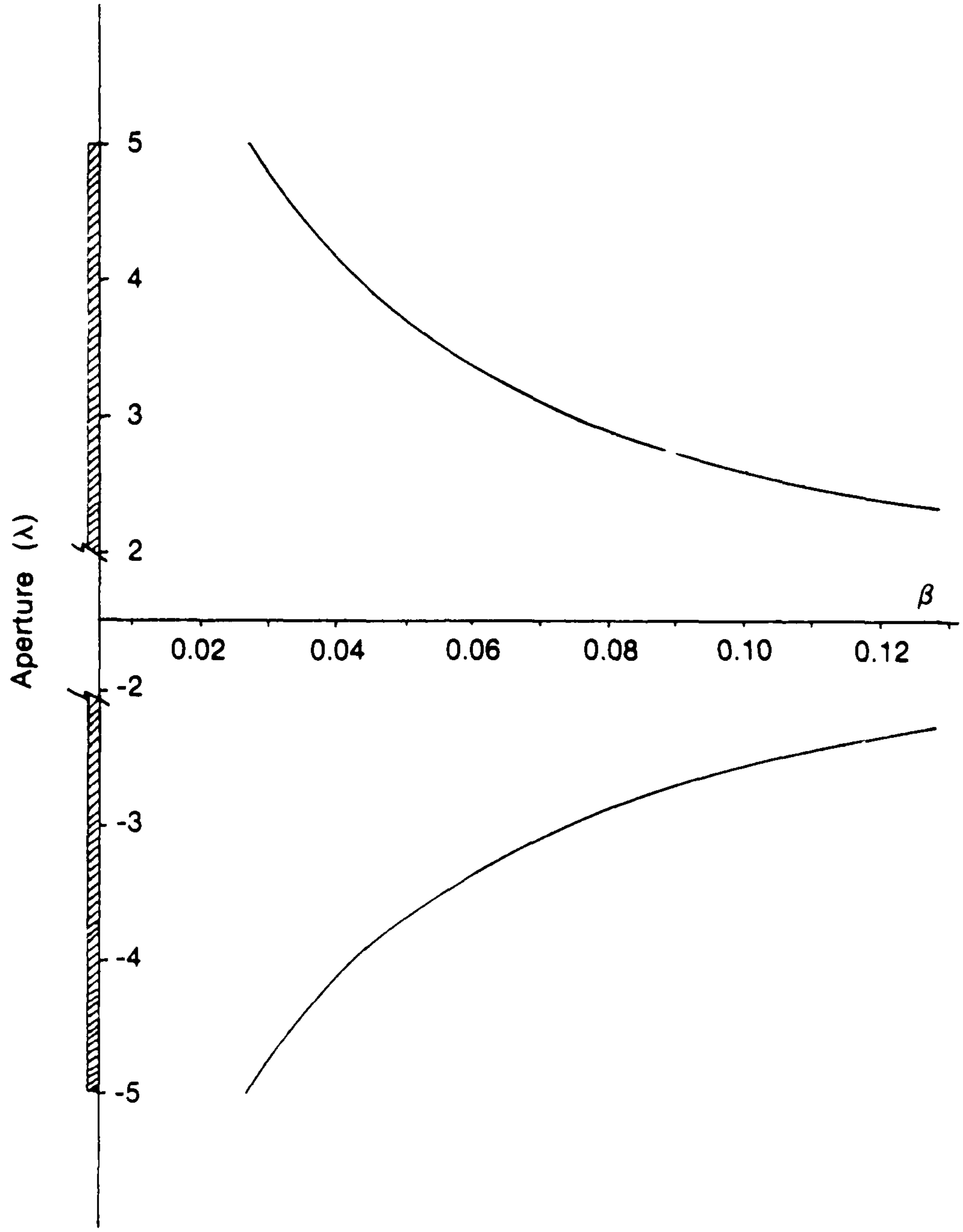


FIG. 19: TRAJECTORIES OF THE ZERO CROSSINGS OF THE QUADRATIC TERM IN THE MSD CRITERION

regions of the aperture distribution. This is because positive and negative areas of the integrand must cancel to yield a zero result of the integral. Considering the first family of continuously varying apertures, if the width "b" of the nonzero regions is small, then the quadratic curve is roughly linear over the nonzero region of the aperture, and the required position of the zero crossings is approximately at the midpoint of the nonzero regions. This approximate required location of the zero crossings may be used to determine a value of β, corresponding to a range from the aperture which yields such zero crossings, ie. by referring to fig. 19.

Graphs of HPBW vs. range to the outer extent of the Fresnel region may be compared for the 16 aperture distributions considered previously and this family of continuously varying apertures. Such a comparison is shown in fig. 20. Although the curves do not overlap, it is apparent that they both have the same shape. This supports the insight that the relationship between HPBW and the range to the outer extent of the Fresnel region hinges on the notion that the positive and negative areas of the MSD integrand must cancel at the minimum range. As the aperture weight is moved to the edges, the boundaries separating the positive and negative areas of the integrand must move toward the aperture edges, corresponding to flattening the quadratic curve and increasing the minimum range at which the MSD criterion is satisfied. Conversely, distributions which are tapered at the edges exhibit relatively small range to the outer Fresnel region.

It is well known that as edge taper of an aperture distribution is increased, its sidelobe levels fall and HPBW increases. Conversely, as the distribution is weighted to the edges, the sidelobe levels rise and HPBW decreases. Thus it is apparent that the unexpected good correlation between HPBW and range to the outer Fresnel region arises from the not so unexpected correlation between HPBW and edge taper. For this reason, the correlation between HPBW and range to the outer Fresnel region observed in fig. 14 is a misleadingly good one, and is qualitative in the sense that HPBW can be qualitatively related to edge taper. The investigation of the "family" of continuously varying aperture distributions indicates that including other arbitrary aperture distributions would only serve to increase the scatter about a best curve (ie. the full extent of the scatter by including other aperture distributions is not apparent in fig. 14).

6.0 POLARIZATION EFFECTS AND A DIRECTIONAL CRITERION

6.1 The Infinitessimal Current Patch

The vector nature of the electromagnetic field has been ignored in both qualitative and quantitative discussions of the range of the outer Fresnel boundary thus far. The phase-sensitive MSD criterion expresses mathematically the notion that the radiation pattern of an aperture changes with increasing range through the Fresnel region, converging to the far field form at large range. Although the shape of the radiation pattern as calculated from scalar equations may have converged closely to the far field pattern, it is possible that there are near field polarization effects that might cause this finite range pattern and the far field pattern to be quite different. The question

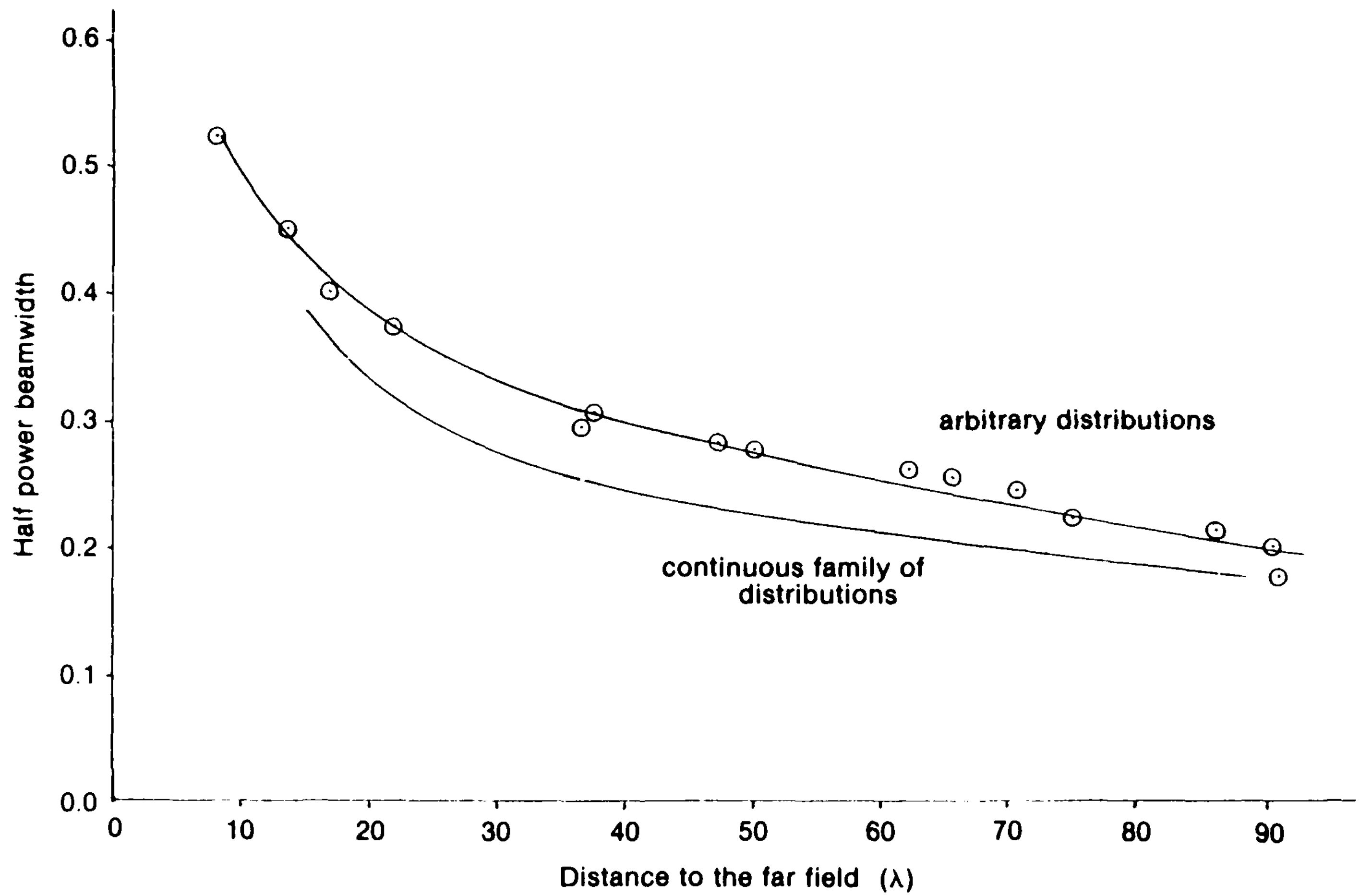


FIG. 20: COMPARISON OF CORRELATIONS BETWEEN HPBW AND MINIMUM DISTANCE TO THE FAR FIELD

becomes whether range dependent polarization effects are significant enough to warrant inclusion in a definition of the range to the outer boundary of the Fresnel region. A natural question to ask, then concerns the relationship between scalar and vector field expressions.

A useful way of approaching this question involves the analysis and description of the vector electromagnetic fields around an infinitessimal current element. It is well known that there is a radial component of the electric field which is significant at close range to the antenna, but vanishes in the far field. Investigating the mathematical foundation for this vanishing, space-dependent component may yield some insight into the issue of polarization and its relevance to discussions of the minimum range to the outer boundary of the Fresnel region.

An analysis of the vector fields around an infinitessimal current element is now presented, following Balanis ([14], pp 164 - 169). Calculation of the vector fields arising from a distribution of sources in a volume V is most easily done by first calculating the vector potential. The vector potential is given, in a (x,y,z) coordinate system, by:

$$\frac{e^{j\omega t}\mu_0}{4\pi}\int_{V'}\frac{J(x',y',z_,)}{r}e^{-jkr}dV' = A \tag{77}$$

where

$$r = \sqrt{(x-x')^2+(y-y')^2+(z-z')^2} \tag{78}$$

and primed (') coordinates are source points, unprimed coordinates are field (observation) points. For reasons which will be apparent in a further section, the infinitessimal current element is assigned an infinitessimal width dy, so that the subject of the analysis is actually an infinitessimal patch of linearly directed current. This adjustment of the model has only a trivial effect on the form of the field equations, but allows generalizations about planar apertures from the results of the analysis. The geometry of the infinitessimal current patch is presented in fig. 21. Because the element has only dz length, eqn. (77) reduces to:

$$\bar{A} = \frac{\mu_0 I dydz}{4\pi r}e^{-jkr}e^{j\omega t}\,\hat{a}_z \tag{79}$$

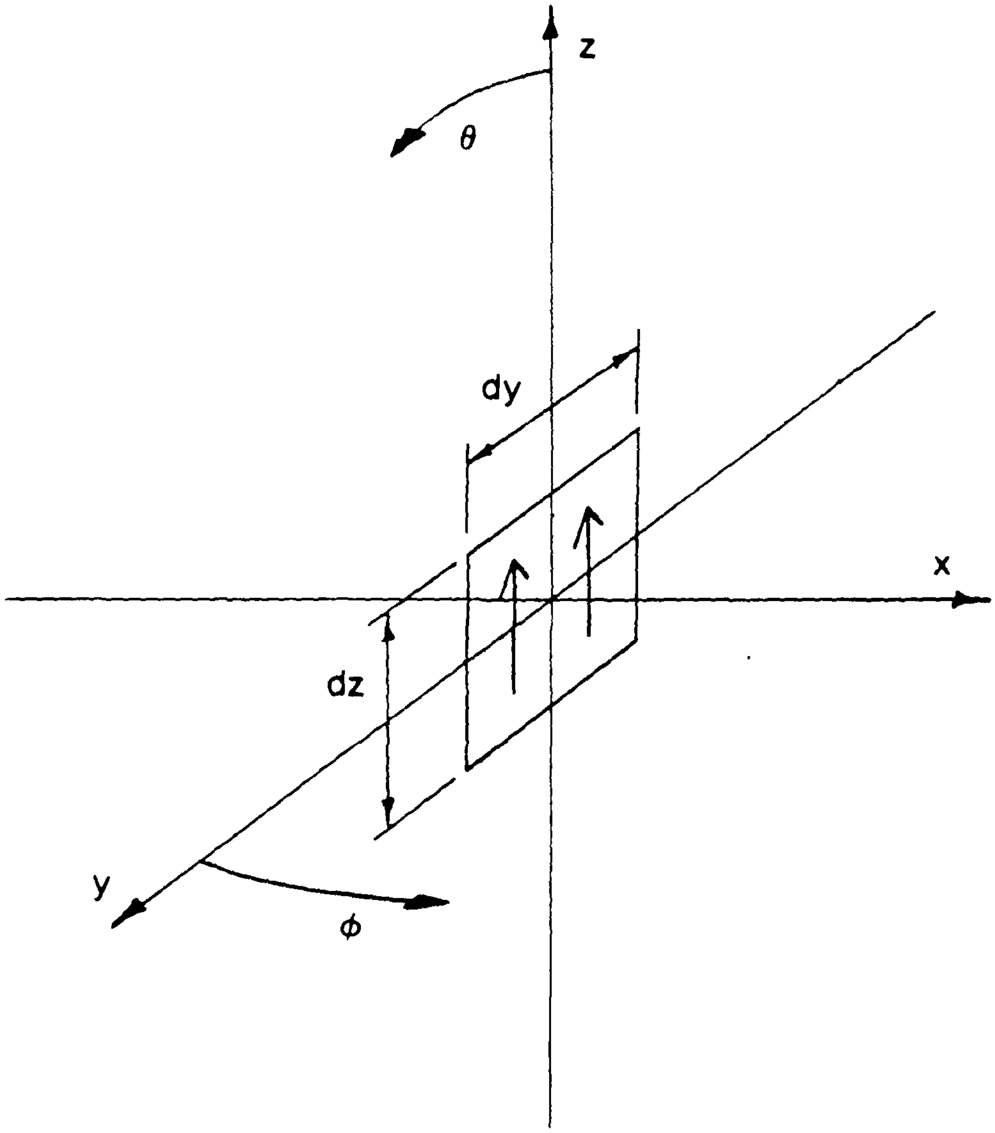


FIG. 21: GEOMETRY FOR THE INITESIMAL CURRENT PATCH

Moving to spherical coordinates, the unit vector in the z-direction can be written as:

$$\hat{a}_z = \hat{a}_r \cos\theta - \hat{a}_\theta \sin\theta \tag{80}$$

Having formed the vector potential of the fields, the magnetic field strength is found from the expression:

$$\bar{H} = \left(\frac{1}{\mu_0}\right) \bar{\nabla} \times \bar{A} \tag{81}$$

and has the form:

$$\bar{H}_\phi = \frac{Idzdy}{4\pi} \left\{\frac{jk}{r} + \frac{1}{r^2}\right\} \sin\theta e^{j(\omega t - kr)} \tag{82}$$

Thus the magnetic field has only a ϕ component, and has a $\sin\Theta$ dependence. From the magnetic field strength the electric field strength can be specified, using the relation:

$$\bar{E} = \frac{1}{j\omega\epsilon} \bar{\nabla} \times \bar{A} \tag{83}$$

The electric field is given by:

$$E_r = \frac{-j\mu_0 Idzdy}{2\pi k\sqrt{\mu\epsilon}} \left\{\frac{jk}{r^2} + \frac{1}{r^3}\right\} \cos\theta e^{j(\omega t - kr)} \tag{84}$$

$$E_\theta = \frac{-j\omega\mu_0 Idzdy}{4\pi k^2} \left\{\frac{-k^2}{r} + \frac{jk}{r^2} + \frac{1}{r^3}\right\} \sin\theta e^{j(\omega t - kr)} \tag{85}$$

The electric field has a radial component, which decreases as the square of the range from the current patch, and has a cosΘ dependence. This contrasts the remaining E and H fields, which have a sinΘ dependence. The angular variation of these fields is shown in fig. 22.

The polarization of the E field is now examined as a function of r and Θ. The radial dependence of E_Θ is examined first, repeated below for convenience:

$$E_\theta(r) \propto -\{\frac{-k^2}{r} + \frac{jk}{r^2} + \frac{1}{r^3}\} \tag{86}$$

If r > 1λ, then first term of the right hand side of eqn. (86) is the largest term. The field thus points in the positive direction, since k is positive. The radial component of the E field has a cosΘ dependence, thus the E field at the four principal positions around the source is parallel to the z-axis, as shown in fig. 23. Description of the E-field as a function of r and Θ between these four principal positions now follows.

Since the field is complex, the real and imaginary parts of the E field are considered separately. Considering r larger than 1 wavelength, a vector diagram of the addition of the r and Θ E field components is shown in fig. 24. The angle ϕ of the total real E field can thus be expressed as:

$$\phi = \theta - \tan^{-1}\{\frac{E_r}{E_\theta}\} = \tan^{-1}\{\frac{\frac{-1}{r^3}\cos\theta}{\frac{1}{2}\{\frac{k^2}{r} - \frac{1}{r^3}\}} \tag{87}$$

and at Θ = π/4, the expression simplifies to:

$$\phi = \tan^{-1}\{\frac{\frac{-1}{r^3}}{\frac{1}{2}\{\frac{k^2}{r} - \frac{1}{r^3}\}}\} \tag{88}$$

Performing a similar analysis for the imaginary component of the total E field, and again using the geometry of fig. 24, the angle of the imaginary component of the E field can be expressed as:

$$\phi = \theta - \tan^{-1}\{\frac{\frac{-k}{r^2}\cos\theta}{\frac{-k}{2r^2}\sin\theta}\} = \theta - \tan^{-1}\{\frac{\cos\theta}{2\sin\theta}\} \tag{89}$$

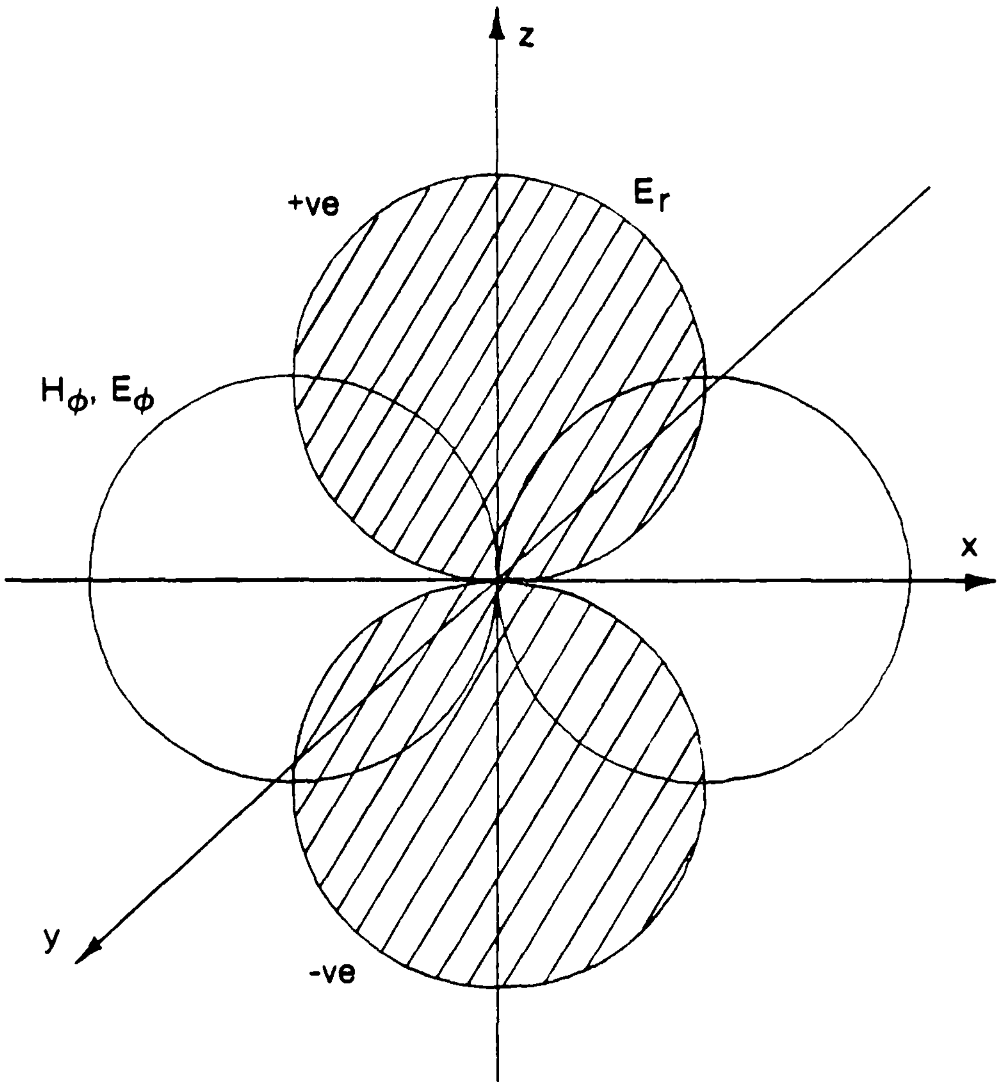


FIG. 22: DIAGRAM SHOWING THE ANGULAR VARIATION OF FIELDS AROUND THE INFINITESIMAL CURRENT PATCH

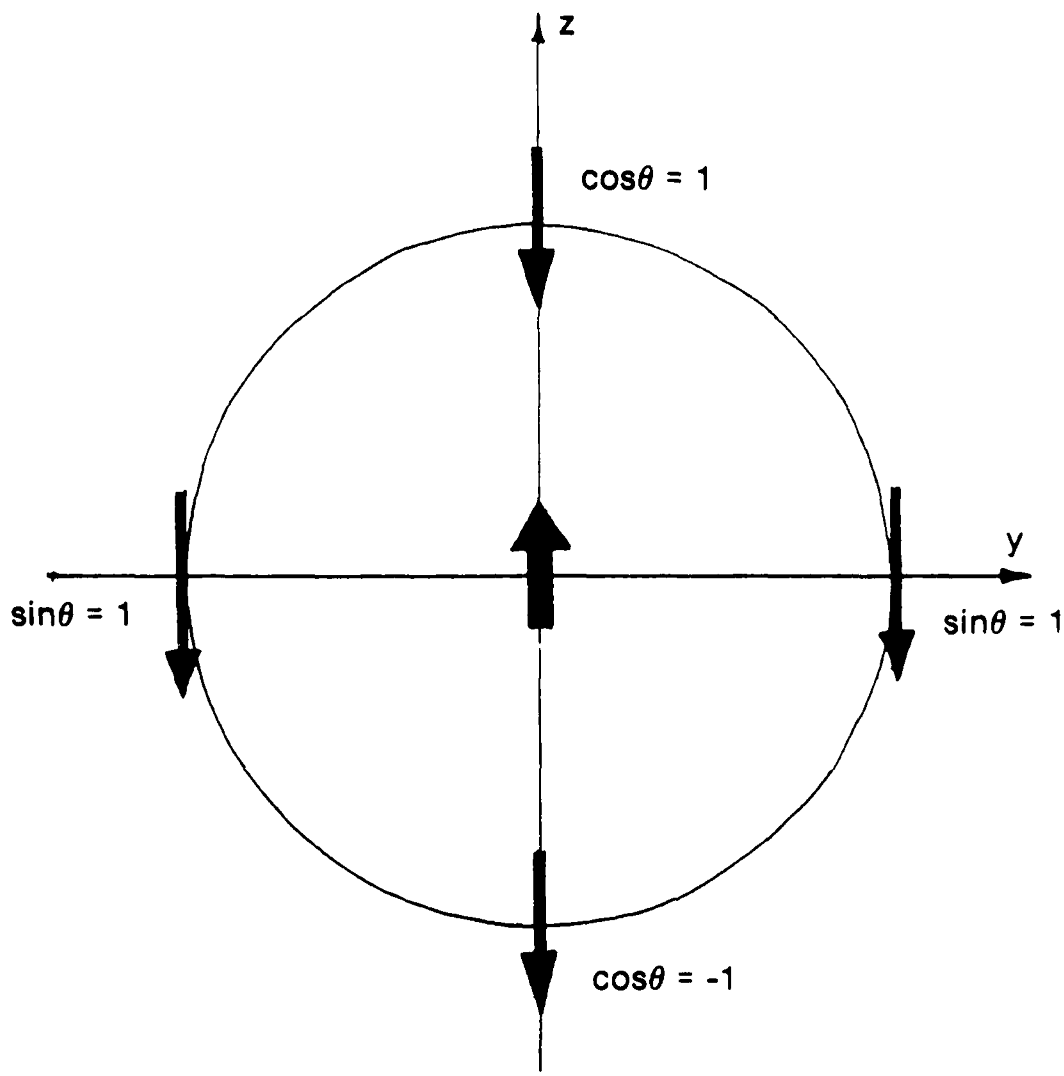


FIG. 23: POLARIZATION OF THE TOTAL E VECTOR AT THE FOUR PRINCIPAL ANGLES

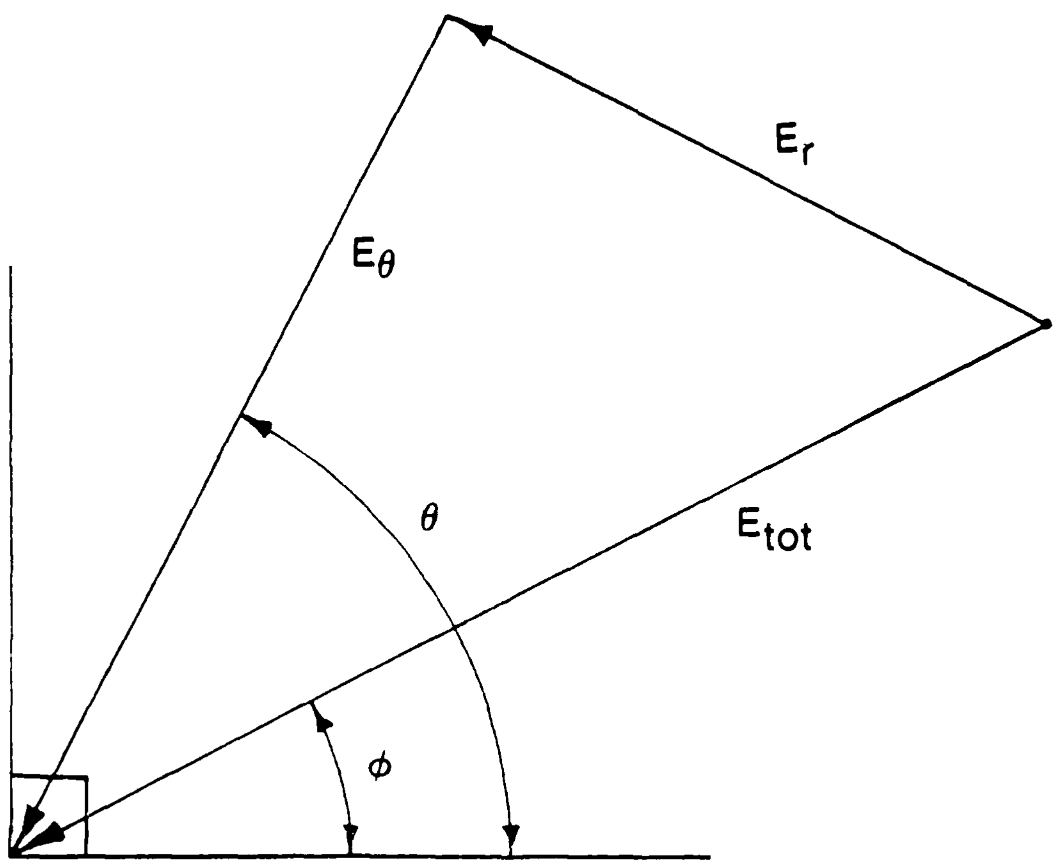


FIG. 24: GEOMETRY FOR DETERMINING THE ANGLE OF THE ELECTRIC FIELD VECTOR

Note that the angle of the total imaginary E field is not a function of range because both r and Θ E field components fall of as the square of range r. The variation of the real and imaginary total E vectors with radial distance r and angle Θ is presented in figs. 25 and 26.

These two diagrams, and especially that of the total real E field, show that the total E vector "tumbles" away from the source, twisting with range. The question remains whether this radial polarization variation plays a role in determining the range to the outer boundary of the Fresnel region, and more specifically, whether it invalidates the (scalar) phase-sensitive MSD criterion.

The radial variation of the total (real and imaginary) E vector arises from the $1/r^2$ and $1/r^3$ terms in the radial real E component. At large range from the source current patch, the components have died away (as has the imaginary E field components) and the remaining term has only a direction component. It is natural, then, to now specify a range at which this radial component may be ignored. Because we are dealing with a vector addition of r and Θ E field components and because radial E component can be thought of as a perturbation of the direction of the E_Θ component, the radial E component is taken to be negligible when it is 1/20 of the E_Θ component. That is to say, the perturbation of a vector by the addition of an orthogonal vector 1/20 of its magnitude is deemed to be negligible. Applying this criterion along the $\Theta = \pi/4$ contour where the coefficients of the radial and Θ components are equal ($\sin\Theta = \cos\Theta$), then the radial component may be ignored at ranges of more than 20 wavelengths. Thus, conservatively and in a polarization sense, the far field approximations may be applied at a range of greater than 20λ wavelengths from the current patch, along a lines of observation corresponding to $\Theta = \pm\pi/4$. This statement is more conservative for observation angles closer to $\Theta = \pi/2$ (normal to the direction of current flow), and less conservative for angles of observation near $\Theta = 0$ (along the direction of current flow), due to the $\cos\Theta$ variation of the radial component. To quantify the 20 wavelength criterion somewhat, the direction of the total real E vector is tabulated below at $\Theta = \pi/4$ for some selected range values:

r (wavelengths)	angle of real E at $\Theta = \pi/4$
1	47.98°
2	45.73°
3	45.32°
5	45.12°
7	45.06°
10	45.03°
infinity	45.00°

6.2 Vector vs. Scalar Aperture Field Expressions

The modelled source in the above analysis is an infinitessimal linearly polarized current patch, rather than a current element having zero width.

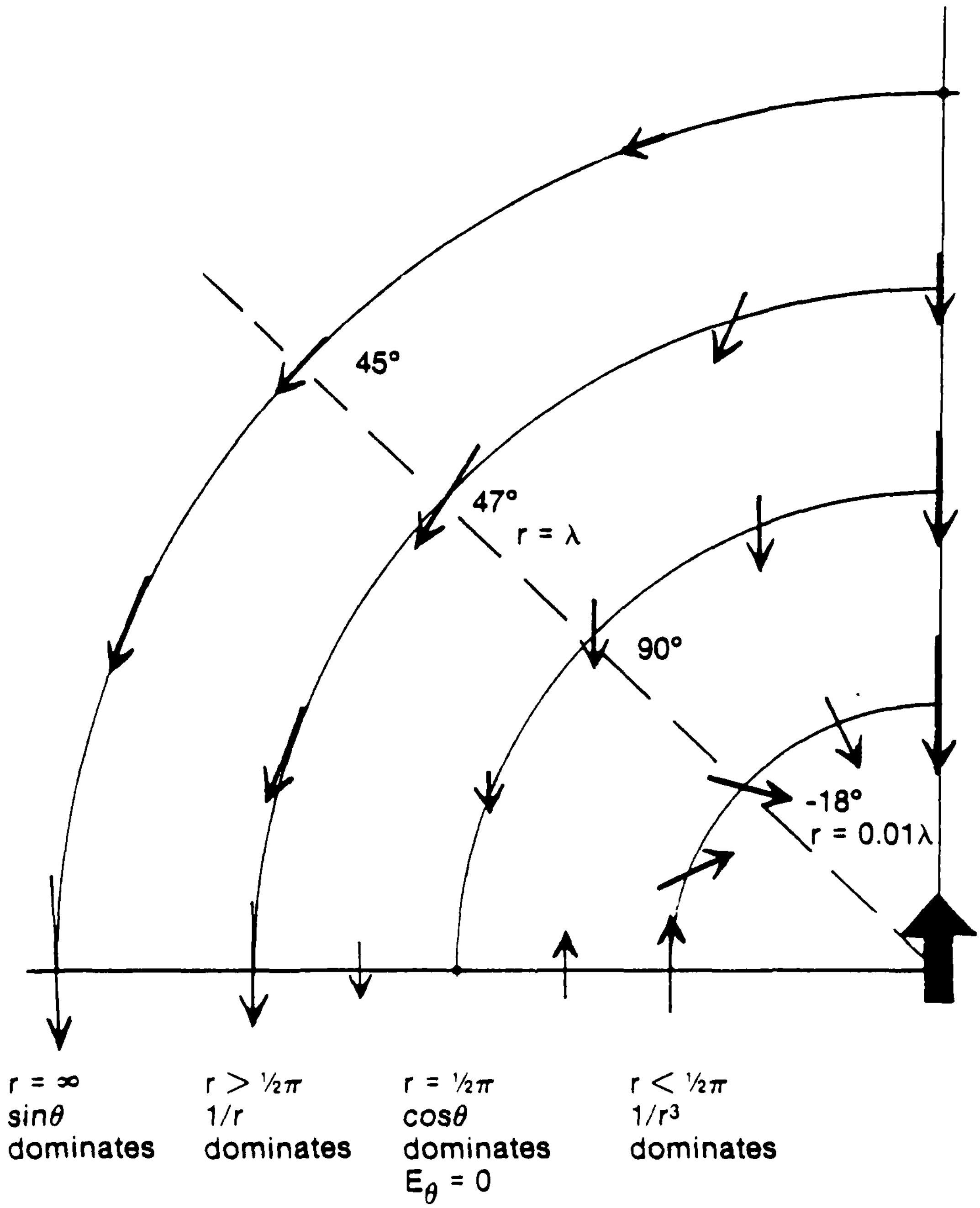


FIG. 25: DIAGRAM SHOWING THE POLARIZATION OF THE REAL E VECTOR THROUGH ONE QUADRANT

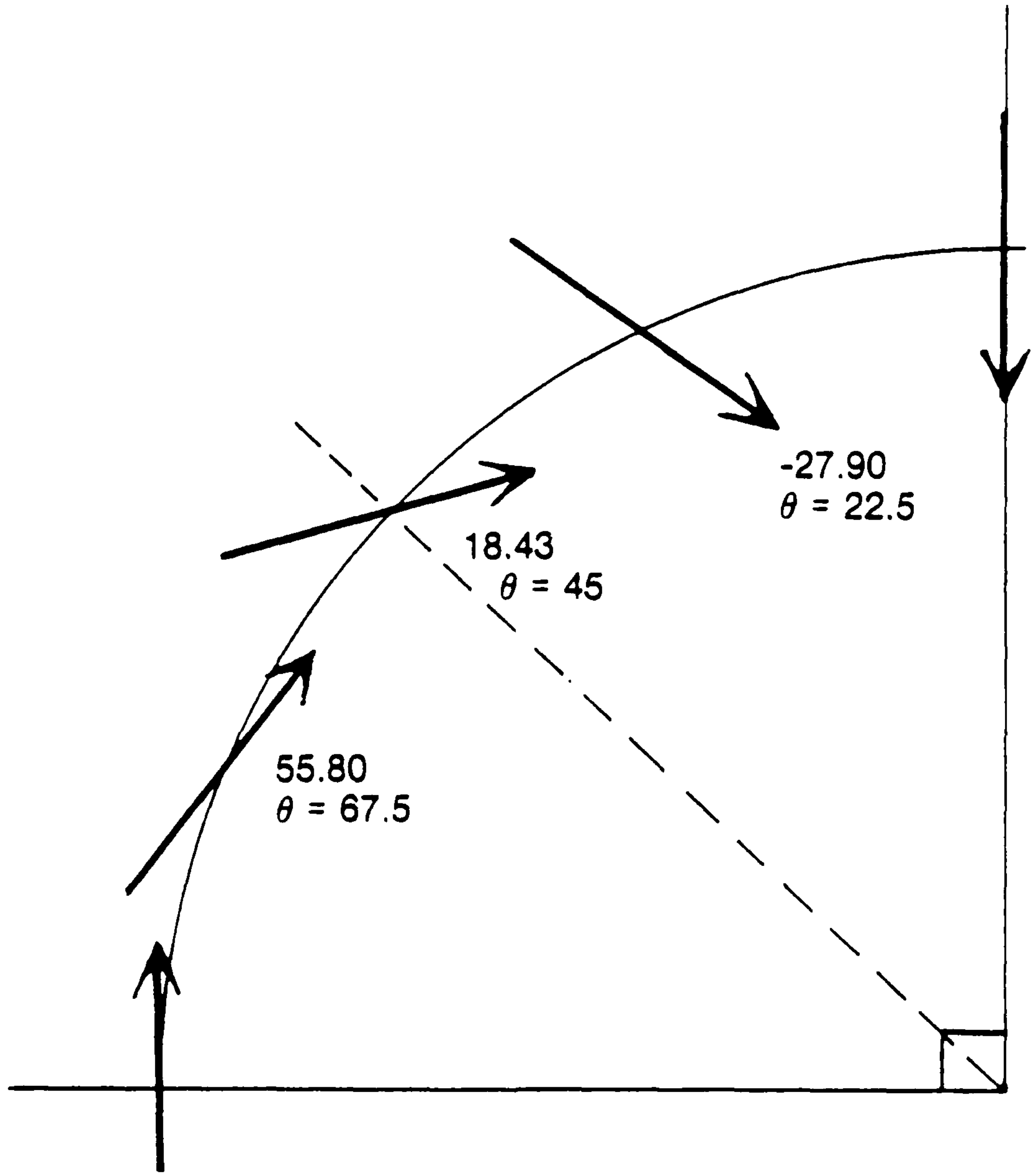


FIG. 26: DIAGRAM SHOWING THE POLARIZATION OF THE POLARIZATION OF THE IMAGINARY E VECTOR OVER ONE QUADRANT

This analysis can be extended to draw generalizations concerning planar apertures by applying the principle of superposition. Theoretically, and planar aperture can be synthesized by combining properly positioned, weighted and phased current patches. The conclusions of the previous section, concerning the significance of the radial component of the electric field, still apply, though the vector addition of the resulting fields and their spatial distribution implies that there is another possible source of a radial component of the total field. "Point source" polarization effects (as a function of range r) may be ignored beyond a range of 20λ, though polarization effects due to the actual antenna size and geometry deserve consideration. Fig. 26a shows the geometry describing this component.

If the infinitessimal current patches are spatially distributed, then a component of the E field from a single current patch may contribute a radial component to the total E field, as measured from the centre of the aperture. This additional radial component arises, then, from the spatial distribution of the current patches about the centre of the aperture. The radial component on boresight due to a single current patch is given by:

$$E_r = A\sin\phi = \frac{Ahdh}{\sqrt{h^2 + r^2}} \tag{89a}$$

where A is the (complex) weighting of the current patch of width dh, and h is the displacement of the current patch from the centre of the aperture. If r >> h, then the radial component due to a single current patch is given by:

$$E_r = \frac{Ahdh}{r} \tag{89b}$$

Integrating over the entire aperture, assuming a uniform phase/uniform amplitude aperture distribution, and normalizing to the weighting A of the current patch, the total radial E field on boresight at range r >> h is given by:

$$E_{r_{tot.}} = \int_{-D/2}^{+D/2} \frac{h}{r} dh = \frac{D^2}{4r} \tag{89c}$$

Similarly, the total E_θ field component from a single current patch can be expressed as:

$$E_\theta = A\cos\phi = \frac{Ardh}{\sqrt{h^2 + r^2}} \approx Adh \tag{89d}$$

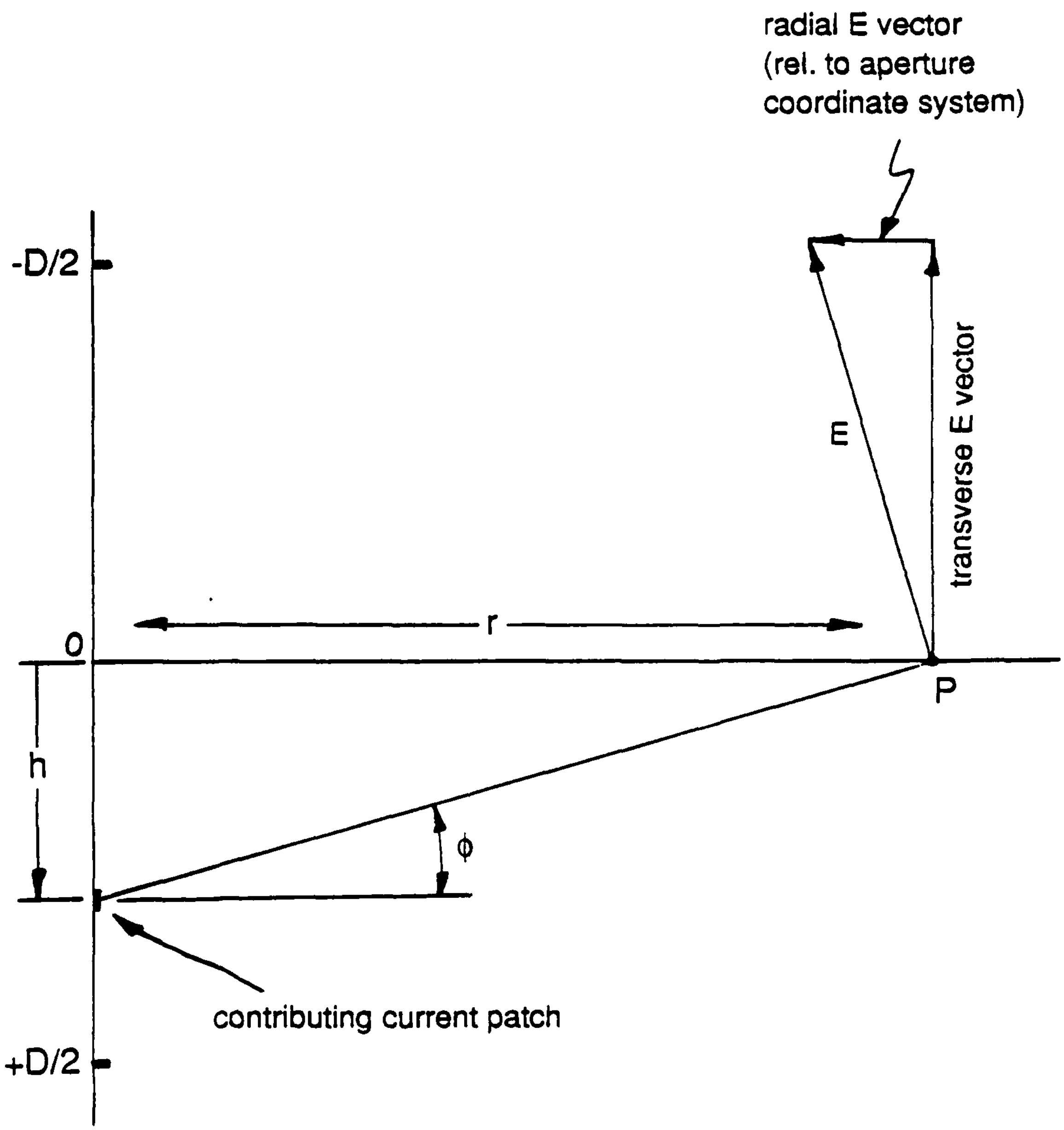


FIG. 26a: RADIAL E-FIELD COMPONENT DUE TO APERTURE GEOMETRY

Integrating over the entire aperture and normalizing to the current patch weighting A, the total E_θ field component on boresight for a uniform phase/uniform amplitude aperture at range r >> h is:

$$E_{\theta_{tot.}} = \int_{-D/2}^{+D/2} dh = D \qquad (89e)$$

Using eqns. (89c) and (89e), the constraint that the radial component of the E field must be less than 1/20 the θ component translates to a range of:

$$r > 5D \qquad (89f)$$

This criterion that the radial component of the polarization be negligible relative to the E_θ component implies a larger range than that determined for a single current patch if the aperture is larger than 4 wavelengths. For apertures of several wavelengths, this constraint on the range to the outer boundary of the Fresnel region is less than that determined by the phase-sensitive MSD criterion, and so we tentatively conclude that polarization effects do not invalidate the MSD criterion, at least for even aperture distribution functions.

To state this conclusion with absolute conviction would require a consideration of aperture distributions other than the uniform case considered here, and for directions other than boresight. A notable exception to this generalization might be an odd aperture distribution which, through destructive interference, cancels the θ field component on boresight but not the radial component. An example of such an aperture is one having a positive delta excitation on one side and a negative delta excitation on the other (and otherwise zero excitation), though the cancellation of the θ component of the E field only occurs at the nulls of the radiation pattern. Over the balance of the radiation pattern the MSD criterion is expected to be more conservative than the geometrical effects giving rise to the radial component of the total E field.

This geometrically derived radial component, as well as point source polarization effects, can be shown to be generally negligible at the range of the outer boundary of the Fresnel region (as predicted by the phase-sensitive MSD criterion) by using another approach. To demonstrate this, we return to the issue of reconciling the current patch analysis and the subsequent 20 wavelength criterion with the scalar method of calculating the electromagnetic fields of an aperture antenna. We wish to understand the nature of the simplifications which allow the vector nature of the fields to be ignored and instead calculated as a scalar field (see Silver [13], pp. 164, 165). If the fields are linearly polarized over the aperture, the vector expression for the diffraction field of the aperture (excluding the line integral which accounts for the distribution of charge around the boundary) can be reduced to:

$$\frac{1}{4\pi}\int_A \{(-j\omega\mu)\hat{n}\times\bar{H})\psi + (\hat{n}\times\bar{E})\times\bar{\nabla}\psi + (\hat{n}\bullet\bar{E})\bar{\nabla}\psi\}dS$$
$$= \frac{-1}{4\pi}\int_A \{\psi\frac{\partial\bar{E}}{\partial n} - \bar{E}\frac{\partial\psi}{\partial n}\}dS + \frac{1}{4\pi}\oint_{\Gamma_A} \psi\bar{E}\times\bar{\tau}dS \quad (90)$$

where:

$$\psi = \frac{e^{-jkr}}{r} \quad (91)$$

and the normal derivative d/dn is applied to each cartesian component of the electric field. τ is a vector (in the plane of the aperture) perpendicular to the polarization of the E field in the aperture. The complete expression for the electric field around the aperture can then be written as:

$$E(p) = \frac{-1}{4\pi}\int_A \{\psi\frac{\partial\bar{E}}{\partial n} - \bar{E}\frac{\partial\psi}{\partial n}\}dS + \frac{1}{4\pi}\oint_{\Gamma_A} \psi\bar{E}\times\bar{\tau}dS$$
$$- \frac{1}{4\pi j\omega\epsilon}\int_{\Gamma_A} \bar{\nabla}\psi(\bar{\tau}\bullet\bar{H})dS \quad (92)$$

Silver points out that the high frequency approximation used in analyzing appropriate electromagnetic diffraction problems is based on the assumption that the diffraction rays are confined to small angles away from the normal out of the aperture. Under this condition, the scalar surface integral may be taken alone in calculating the diffraction field, as shown below:

$$u(p) = \frac{1}{4\pi}\int_A \{\psi\frac{\partial u}{\partial n} - u\frac{\partial\psi}{\partial n}\}dS \quad (93)$$

where u is the proper component for the linearly polarized electric field. However, Silver points out that the scalar expression for the fields around the aperture can only yield qualitative results very close to the aperture because the line integrals (of eqn (92)) which were ignored make a significant contribution. The twisting of the field vectors in this region is attributed to the vector contributions of these line integrals. Silver states that the line integrals are significant throughout the near zone (ie. that region comprised of points in the immediate neighborhood of the aperture). Beyond this range the scalar representation of the fields is appropriate.

This, then, reconciles the vector nature of the electromagnetic fields of an aperture antenna with the scalar expressions for the fields and with the case of an infinitessimal current patch. In summary, only one component of the vector field dominates at large range from the source - that component which is transverse to the direction of propagation (this is actually the radiation condition). This means that at sufficiently large range, the angle of observation specifies the polarization of the radiated field, and so a scalar description of the fields can be adopted with the direction of polarization "understood". Also, at closer range, the polarization of the field exhibits a radial dependence and a scalar representation is inappropriate. Range dependent polarization effects do not appear to invalidate the phase-sensitive MSD criterion because of the conservative range specified by this criterion in estimating the outer extent of the Fresnel region.

6.3 A Modified Mean Squared Difference Criterion

The investigation thus far has included a mathematical and qualitative description of the nature of superdirective antennas, the hypothesis that such antennas have a large Fresnel region (large minimum range to acceptable application of the far field approximations), and has included the formulation of a rigorous definition of the range to the outer boundary of the Fresnel region. Having investigated this phase-sensitive mean squared difference (MSD) criterion, an apparent correlation was observed between half power beamwidth (HPBW) and the range to the outer boundary of the Fresnel region, which tended to support the hypothesis. However, a slight mathematical reworking of the criterion produced (in a loose sense), a synthesis form of the MSD criterion and showed that, using this criterion, the hypothesis is disproved. It is natural then to return to this criterion and again examine its formulation to gain insight into why the hypothesis is disproved. As pointed out in a previous section, the phase-sensitive MSD criterion is reasonable in that it indexes the quality of the far field approximations at finite range by expressing the degree of convergence of the finite range complex radiation pattern to the far field form. Further, the phase-sensitive MSD criterion is conventional in the sense that a normalized mean squared difference scheme is employed to compare the complex radiation patterns.

An important feature of the phase-sensitive MSD criterion is that, in the form presented previously, it is applied over the entire radiation pattern, including the invisible region. This formulation was chosen to ensure mathematical consistency, since any alternative would be to favour some part of the radiation pattern - a justification of which is absent in the original generally-formulated hypothesis. However, the disproval of the hypothesis (at least in the absence of an inversion of the near field/Fresnel boundary, ie. strongly superdirective antennas) gives some motivation to investigate the possibility of applying the phase-sensitive MSD criterion over only a specific sector. An application of the MSD criterion over a sector is interpreted as specifying the minimum range from an antenna at which the far field approximations may be applied over the given sector. This required minimum range can be shown to be a function of the position and size of the sector. Even sectors corresponding to complex angles of observation (the invisible region) are of interest since these sectors can be made visible by shifting the radiation pattern with a phase tilt across the aperture.

The original form of the phase-sensitive MSD criterion is repeated below for convenience:

$$\frac{\int_{-\infty}^{\infty} \left| \int_{-a/2}^{a/2} (e^{-j\beta y^2} - 1) D(y) e^{jyk_y} dy \right|^2 dk_y}{\int_{-\infty}^{\infty} \left| \int_{-a/2}^{a/2} D(y) e^{jyk_y} dy \right|^2 dk_y} \leq K \tag{94}$$

where K is the convergence factor against which the different aperture distributions are evaluated, and chosen to reflect traditional (distribution independent) definitions of the range to the outer boundary of the Fresnel region. Eqn. (94) can be written as:

$$\frac{\int_{-\infty}^{\infty} |H(k_y)|^2 dk_y}{\int_{-\infty}^{\infty} |G(k_y)|^2 dk_y} \leq K \tag{95}$$

This criterion may be applied over a sector of the radiation pattern by windowing the radiation pattern as indicated below:

$$\frac{\int_{-\infty}^{\infty} |W(k_y - k_o) H(k_y)|^2 dk_y}{\int_{-\infty}^{\infty} |W(k_y - k_o) G(k_y)|^2 dk_y} \leq K \tag{96}$$

where $W(k_y)$ is the window function in the antenna pattern domain. Parseval's theorem may be used to transform this expression from the radiation pattern domain to the aperture domain, converting multiplication to convolution:

$$\int_{-\infty}^{\infty} |e^{jk_y y}R(y)*(e^{-j\beta y^2}-1)D(y)|^2 dy - K\int_{-\infty}^{\infty} |e^{jk_y y}R(y)*D(Y)|^2 dy \leq 0 \tag{97}$$

where $R(y)$ is the inverse Fourier Transform of the window function, and the exponentials arise from the shifting property of the Fourier Transform. Eqn. (97) can be written as the products of complex conjugates:

$$\int_{-\infty}^{\infty}\{ \int_{-\infty}^{\infty} e^{jk_o}R(y-x)(e^{-j\beta x^2}-1)D(x)dx \int_{-\infty}^{\infty} e^{-jk_o}R^*(y-z)(e^{j\beta z^2}-1)D^*(z)dz$$

$$- K\int_{-\infty}^{\infty} e^{jk_o}R(y-x)D^*(x)dx \int_{-\infty}^{\infty} e^{-jk_o}R^*(y-z)D^*(z)dz\} \leq 0 \tag{98}$$

The only terms of eqn. (98) which are functions of y are the window functions, and so the order of integration can be changed. Collecting like terms, eqn. (98) can be rewritten as:

$$\int_{-a/2}^{a/2}\int_{-a/2}^{a/2}\{(e^{-j\beta x^2}-1)(e^{j\beta z^2}-1) - K\}e^{jk_o x}e^{-jk_o z}D(x)D^*(z)$$

$$\int_{-\infty}^{\infty} R(y-x)R^*(y-z)dxdydz \leq 0 \tag{99}$$

By examining the form of eqn. (99), a parallel becomes apparent with the synthesis form of the phase-sensitive MSD criterion presented in eqn. (91). Eqn. (91) identifies regions of the line aperture distribution which contribute positive or negative area to the MSD integral. Regions contributing positive area must be more heavily weighted in order to ensure a large range to the outer boundary of the Fresnel region. Eqn. (99), however, presents a plane, rather than a line, over which the weighting function of the aperture distribution is defined, which determines regions of positive and negative area contribution. The plane is that defined by the x and z variables. The x-z plane collapses to a line along the diagonal for the case of application of the MSD criterion over the entire (visible and invisible) radiation pattern. This is because the window function ranges over all real and complex angles and thus R(y) becomes a delta function. The integration contour thus lies along x = z for this special case.

Further simplifying assumptions could be introduced to the directional or windowed form of the phase-sensitive MSD criterion, such as specifying the window shape in the radiation pattern domain to be a sin(x)/x function. This would yield a rectangular R(y) function allowing integration with respect to y and a closed form expression of the result. However, the remaining integrand is still relatively complicated and so does not lead to a simple graphically clear weighting function which should be applied to an aperture function, to ensure a large range to the outer Fresnel boundary at a given observation angle. Whereas the omnidirectional form of the phase-sensitive MSD criterion was found to be insensitive to the phase profile of the aperture distribution, the directional form of the criterion is clearly a function of this phase profile.

Yet another simplifying assumption which could be introduced is to let the window function $W(k_y)$ become a delta function in the radiation pattern domain. This corresponds to determining the convergence of the radiated field, to the far field form, at a single angle of observation rather than over a sector. The phase-sensitive MSD criterion can be written in this case as:

$$\left| \int_{-a/2}^{a/2} (e^{-j\beta y^2}-1)D(y)e^{jk_o y}dy \right|^2 - k \left| \int_{-a/2}^{a/2} D(y)e^{jk_o y}dy \right|^2 \leq 0 \qquad (100)$$

This directional criterion is shown in fig. 27, evaluated over most of the visible region for the case of a uniform aperture distribution and for a convergence factor arbitrarily borrowed from the omnidirectional criterion. There is a large variation of the range to the outer boundary of the Fresnel region using the directional criterion - as much as an order of magnitude. Most striking is that the range to the outer boundary of the Fresnel region increases dramatically near the nulls in the radiation pattern. Also, it is clear that the radiation pattern near the first (and to a lesser degree the second) null requires a large range to converge to the far field form. In contrast, the mean beam converges quite quickly. This reflects the fact that the main beam and higher order sidelobes form quite quickly with increasing

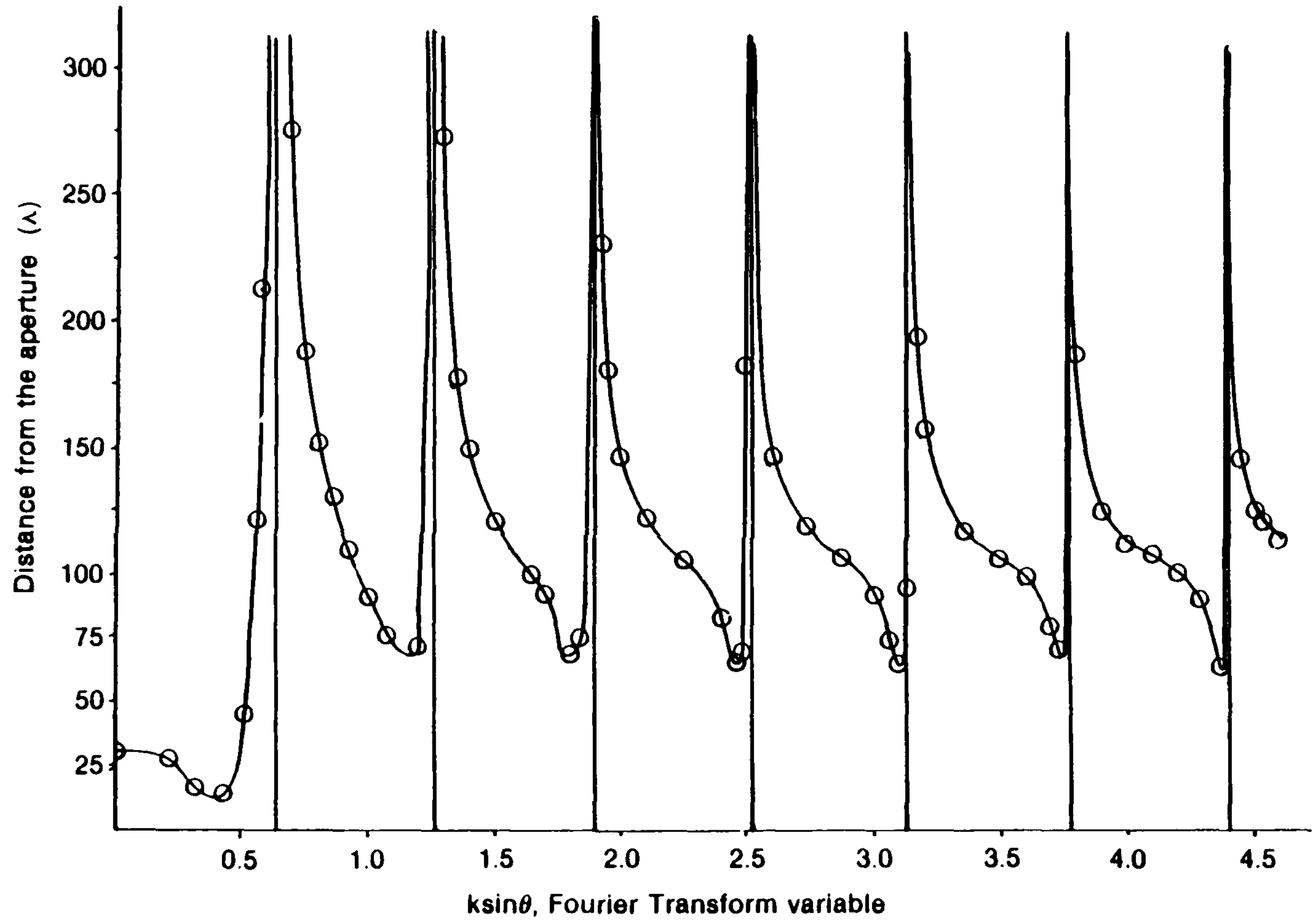


FIG. 27: CONVERGENCE OF THE UNIFORM APERTURE RADIATION PATTERN

range, whereas the first sidelobe begins as a "shoulder" of the main beam and gradually separates with range as the first null increases in depth. this convergence of the radiation pattern is hidden by the omnidirectional form of the phase-sensitive MSD criterion since is comments only on the total mean squared difference between the finite range and far field patterns. The omnidirectional form does not indicate where the difference between the two patterns is concentrated. Further, since the main beam constitutes the balance of the radiated power, it is reasonable that the omnidirectional version of the criterion predicts a range to the outer boundary of the Fresnel region which is biased toward lower values. Although the graph of fig. 27 does not extend into the invisible region, it is reasonable to assume that the trend of the graph continues, and that all higher order sidelobes form within a relatively short range from the antenna.

Finally, fig. 27 gives another insight into the convergence of the radiation pattern through the Fresnel region. The spikes near the nulls of the radiation pattern arise from the fact that the far field nulls do not exactly coincide with the nulls of the finite range pattern. That is, it appears that the nulls shift position slightly with increasing range through the Fresnel region. From the asymptotes of fig. 27, it appears that the nulls of the finite range pattern move away from boresight as one moves away from the antenna. The hypothesis that the angular placement of the nulls is a function of range in the Fresnel region has not been tested and is beyond the scope of this investigation.

The behavior of the radiation pattern near nulls suggests that the original hypothesis, that a superdirective antenna may exhibit a large range to the outer boundary of the Fresnel region, may be correct in some sense when considering the directional criterion. Fig. 27 indicates that the shape of the radiation pattern near a zero (or near a feature of the radiation pattern having a steep slope) may take large range to converge to the far field form. It is possible that antennas having a very narrow main beam may imply a large range to the outer boundary of the Fresnel region for the beam have a good approximation of the far field form. If this is so, then in this (directional) sense the hypothesis could be considered correct.

6.4 Applications Notes

Perhaps the most natural application of this analysis and the conclusions drawn concerns the range requirements in antenna pattern measurement. The conventional and generally conservative $2D^2/\lambda$ range to the outer boundary of the Fresnel region implies unachievably large test ranges for antennas of only moderate size (although it is possible to determine the far field pattern from near field measurements using elaborate mathematical and measurement techniques). For instance, as mentioned previously, at 35 GHz a 1 metre reflector requires a test range of over 230 metres. Multipath and environmental conditions (such as rain and terrain irregularities) make such a range difficult to realize and degrade the reliability of measurements when such a range is available. Thus it is sensible to use as small a range as possible when performing such antenna measurements.

The notion of determining the quality of the $2D^2/\lambda$ approximation has been examined by Hansen [15]. He notes that the $2D^2/\lambda$ rule is generally acceptable for antennas with moderate sidelobe levels (-25 dB), and that in such cases directivity errors are of the order of 0.1 dB and pattern errors are negligible. By using the Taylor n distribution (because it is a versatile, efficient, and robust optimum distribution) he shows that the first sidelobe of low sidelobe patterns requires a large range to converge to the far field form. This conclusion is also presented by Hacker and Schrank [16] by an analysis of several typical low and ultralow sidelobe radiation patterns.

Fig. 27 is in qualitative agreement with both the Hansen paper and the Hacker and Schrank paper. Fig. 27 shows that the first and to a lesser degree the second sidelobes require a large range to converge, whereas the main beam and higher order sidelobes form quickly. This is also confirmed qualitatively by examining the set of drawings of the radiation patterns of the uniform aperture at various ranges, found in [16]. The error in the level of the first sidelobe is graphed in [16] as a function of range, and it is shown that the lower the level of t e sidelobes, the greater the error in the first sidelobe (relative to the normalized far field level) at a given range. This criterion determines the required test range by examining only one point in the radiation pattern. The directional phase-sensitive MSD criterion extends the criterion used in [15] and [16] for application at any point over the entire radiation pattern.

The statement that the $2D^2/\lambda$ range is acceptable for patterns with high sidelobe levels seems at first to contradict the conclusions of the previously presented analysis using the phase-sensitive MSD criterion. Using the MSD criterion, it appears that high sidelobe patterns require relatively large range to converge to the far field form and so should require large test ranges, whereas cosine and edge tapered distributions converge quite quickly with range. This apparent discrepancy can be explained by closely examining the form of the omnidirectional MSD criterion, and exactly what it indexes. Although the sidelobes of low sidelobe patterns require a large range to form properly, they are only a small feature of the radiation pattern. Most of the radiated power is in the broad main beam. It has been shown that the main beam forms at relatively short range, and due to this concentration of energy the errors in the tiny sidelobes do not impact significantly on the omnidirectional criterion. Thus the predicted range to the outer boundary of the Fresnel region is biased toward the range at which the main beam is well formed. In the case of higher sidelobe patterns, less energy is concentrated in the main beam, so errors in the sidelobe patterns have a larger effect on the value of the MSD criterion. Because the main beam forms at shorter range than the sidelobes, even for higher sidelobe patterns, the net effect is that higher sidelobe patterns require (on a power average) a larger range to converge to the far field form. References [15] and [16] imply that if one wishes to measure the far field form of the first sidelobe of low and ultralow sidelobe patterns, then large test ranges must be used.

7.0 FURTHER WORK

Further work concerning the specification of the range to the outer boundary of the Fresnel region could concentrate on a more thorough investigation of the directional phase-sensitive mean squared difference criterion. The omnidirectional case resulted in a synthesis equation based on positive and negative integral contributions over a line. It is clear that in the directional case this line extends to a plane. However, the form of the equations does not allow for good intuitive insight into the behavior of a synthesis equation over this plane. This is because the weighting functions for the aperture distribution can be considered as the product of three independent functions: the Fresnel kernel component, the shifting exponentials, and the convolution of the transformed window function. Further mathematical manipulation of the criterion using integral operators bears investigation.

Though the original hypothesis is disproved in the context of the omnidirectional phase-sensitive criterion, it is possible that the steep slope features of a superdirective radiation pattern might require a large range to converge to the far field form, and in this sense perhaps the original hypothesis could be demonstrated to have some validity. This is suggested by the convergence of the uniform aperture pattern near the nulls since it is in this region that the slope of the radiation pattern is relatively large. Further investigation is necessary using a ıperdirective aperture larger than two wavelengths, since this is the minimum size for which the given Fresnel approximations are valid (assuming negligible reactive fields at the outer Fresnel boundary).

The hypothesis that the nulls trace non-radial trajectories through the Fresnel region should be tested. This could be checked most directly by determining the exact angular location of the nulls for the uniform aperture case. Although the nulls may translate only a slight amount in the Fourier Transform variable k_y, this may imply significant translation in real angle for nulls near the edge of the visible region, because the transformation between real angle (bearing) Θ and k_y is nonlinear.

8.0 CONCLUSION

The purpose of this investigation was to investigate the extent of the Fresnel region for the case of aperture antennas, and specifically to test the hypothesis that a superdirective antenna may exhibit a large range to the outer boundary of the Fresnel region. The first part of the investigation centred on describing the nature of superdirective antennas. The second part of the investigation involved a specification of a rigorous definition of the outer extent of the Fresnel region, and analyzing the implications of this criterion.

Fundamental to the concept of superdirective antennas is that the radiation pattern is defined over an infinite range of complex angles of propagation, with only the real angles corresponding to the "visible" region.

A superdirective aperture may be synthesized by mathematically constraining the radiation pattern over the visible region to have a superdirective shape (ie. a directivity greater than that of a uniform phase/uniform amplitude aperture having the same physical dimension). The desired shape is then decomposed into the sum of orthogonal, aperture limited functions. The required aperture excitation is then determined from this decomposition. The phase profile of the resulting aperture distribution is critical in realizing the superdirective radiation pattern. However, as the directivity of the specified pattern is increased even slightly above that of a uniform aperture, the pattern in the invisible region (corresponding to complex angles of propagation) becomes large. The literature explains that the large patterns in the invisible region are indicative of an unrealizable antenna due to large reactive fields, energy storage around the aperture, and high voltage fields in the aperture plane. Superdirective antennas are generally unrealizable because of large currents required in the elements (of an array realization), and critically close tolerances on element spacing, phasing, and amplitude excitation.

A rigorous definition of the range to the outer extent of the Fresnel region was presented. This criterion is based on the normalized mean squared difference between the complex radiation pattern at a finite range and the far field pattern. By using a mathematical reformulation of this phase-sensitive mean squared difference criterion the original hypothesis is disproved in the context of this criterion, and for at least marginally superdirective antennas. The notion that the hypothesis is disproved cannot be extended to the case of strongly superdirective antennas because there exists the possibility that the range at which the reactive near field criterion is met could be larger than the predicted range to the outer boundary of the Fresnel region. There is assumed to be no such inversion of the near field/Fresnel boundary in applying the phase-sensitive criterion. Also, although aperture phasing is important to superdirective antennas, this phasing is ignored by the phase-sensitive criterion (though phasing of the radiation pattern is not ignored). Polarization effects were concluded to be significant at ranges within the Fresnel region but to be insignificant at the range (as predicted by the phase-sensitive criterion) of the outer Fresnel boundary. A mathematical reformulation of the phase-sensitive criterion allows the range to the outer boundary of the Fresnel region to be specified, and an appropriate aperture distribution chosen.

Finally, the application of the mean squared difference criterion to only a sector of the radiation pattern was considered. Mathematical manipulation of this directional criterion demonstrated that, when considering the convergence of the finite range pattern to the far field form over an observation sector, the aperture phasing is important. A special case of the directional criterion was applied to a uniform aperture, and it was found that the radiation pattern near nulls requires a large range to converge to the far field form. Also, the first and second sidelobes require a relatively large range to converge, whereas the main beam converges to the far field form at relatively short range. Higher order sidelobes, like the main beam, converge relatively quickly. These findings, using the directional criterion, are in agreement with the literature.

9.0 REFERENCES

[1] Booker, H. B., Clemmow, P.C.; "The Concept of an Angular Spectrum of Plane Waves, and Its Relation to That of a Polar Diagram and Aperture Distribution"; JIEE, pt. III, vol. 97, January 1950; pp. 11 - 17

[25 Woodward, P. M., Lawson, J. D.; "The Theoretical Precision With Which An Arbitrary Radiation-Pattern May be Obtained From a Source of Finite Size"; JIEE, vol. 95, 1948; pp. 363 - 370

[3] Schelkunoff, S. A.; "A Mathematical Theory of Linear Arrays"; Bell Systems Technical Journal, vol. 22, January 1943; pp. 88 - 107

[4] Riblet, H. J.; "Note on the Maximum Directivity of an Antenna"; Proc. I.R.E., vol. 36, May 1948; pp. 620 - 624

[5] Rhodes, D. R.; "The Optimum Line Source for the Best Mean-Squared Approximation to a Given Radiation Pattern"; IEEE Trans. on Ant. and Prop., July 1963; pp. 440 - 446

[6] Wilmotte, R. M.; "Note on the Practical Limitations in the Directivity of Antennas";Proc. I.R.E., vol. 36, July 1948; p. 878

[7] Schelkunoff, S. A.; Antennas: Theory and Practice; John Wiley and Sons Inc., New York, 1952; pp. 195 - 198

[8] Taylor, T. T.; "Design of Line Source Antennas For Narrow Beamwidth and Low Sidelobes"; I.R.E Trans. on Ant. and Prop., vol. 3, January 1955; pp. 16 - 28

[9] Colin, R. E., Rothschild, S.; "Reactive Energy in Aperture Fields and Aperture Q"; Canadian Journal of Physics, vol 41, December 1963; pp. 1967 - 1979

[10] Rhodes, D. R.; "On the Stored Energy of Planar Apertures"; IEEE Trans. on Ant. and Prop., vol. AP-14, no. 6;November 1966

[11] Skolnik, M.; Radar Handbook; McGraw-Hill Book Co. Inc., New York, 1970

[12] Jull, E. V.; Aperture Antennas and Diffraction Theory; Peter Peregrinus Ltd., 1981; pp. 42 - 46

[13] Silver, Samuel; Microwave Antenna Theory and Design; McGraw-Hill Book Co. Inc., New York, 1949

[14] Balanis, C. A.; Antenna Theory: Analysis and Design; Harper and Row, Publishers, New York, 1982, Ch. 1 - 3, Ch. 11 pp. 446 - 486

[15] Hansen, R. C.; "Measurement Distance Effects on Low and Ultralow Sidelobe Patterns"; IEEE Trans. on Ant. and Prop., vol. AP-32, no. 6, June 1984; pp. 591 - 593

[16] Hacker, P. S., Schrank, H. E.; "Range Distance Requirements for Measuring Low and Ultralow Sidelobe Antenna Patterns"; IEEE Trans. on Ant. and Prop., vol. AP-30, no. 5, September 1982, pp. 956 - 965

APPENDIX A

Discussion of the Invisible Region

Over the visible and invisible regions there is a 1/r range variation in the electric field strength (see eqn. (49)). In comparing the radiation patterns over the visible region, this 1/r spatial variation is removed by multiplying the field expressions by r. However, as shown in Section 2.1, there is an additional spatial variation of the fields in the invisible region, of the form:

$$e^{-k\sqrt{(\sin^2\theta - 1)}\, r} \tag{A-1}$$

This spatial variation must also be removed when calculating the normalized mean squared difference between the finite range and far field antenna patterns. The question arises as to what if any special considerations should be made in removing this dependence. For instance, although the proper factor is:

$$e^{k\sqrt{(\sin^2\theta - 1)}\, r} \tag{A-2}$$

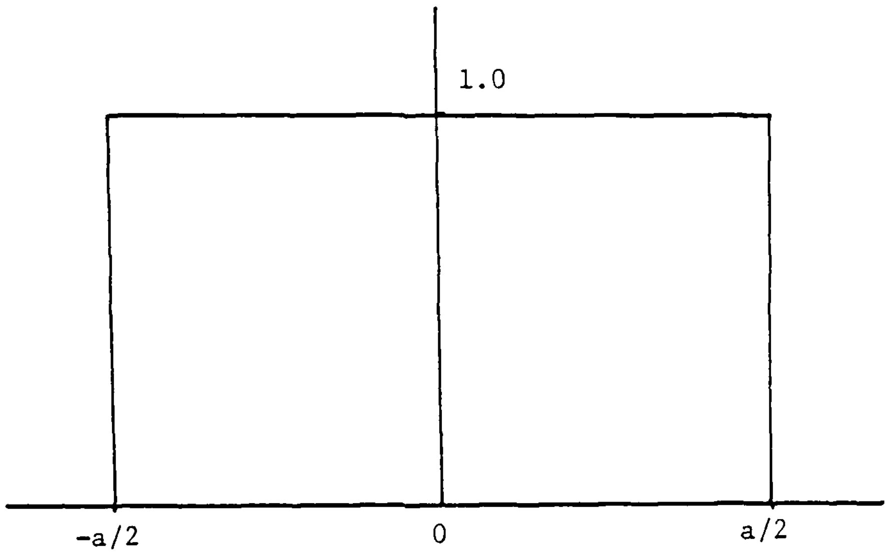
mathematical consistency demands consideration of the appropriate interval over which the multiplication should be applied.

Only the fields in the invisible region have this exponential variation with range, so to include the visible region in the range of this exponential normalizing factor (eqn. (A-2)) would be to introduce a range variation rather than remove one. This means that the Fourier Transform of the aperture distribution does indeed give the radiation pattern independent of range variation in both the real and invisible regions. Thus, in using the mean squared difference criterion over the entire radiation pattern, it is appropriate to use the Fourier Transform of the aperture distribution as the range-independent radiation pattern.

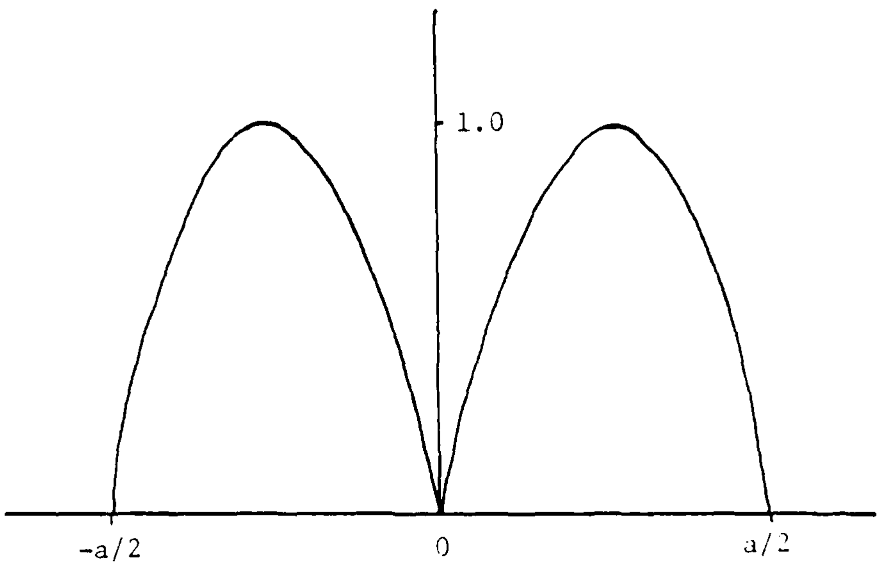
APPENDIX B

Definition of the Aperture Distributions

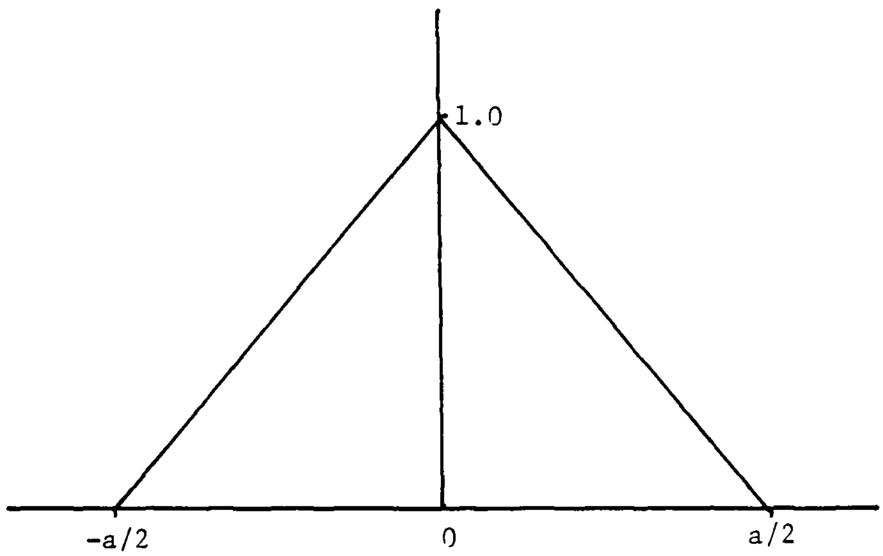
This appendix defines the 16 aperture distributions used to investigate the criteria for determining the range to the outer boundary of the Fresnel region.



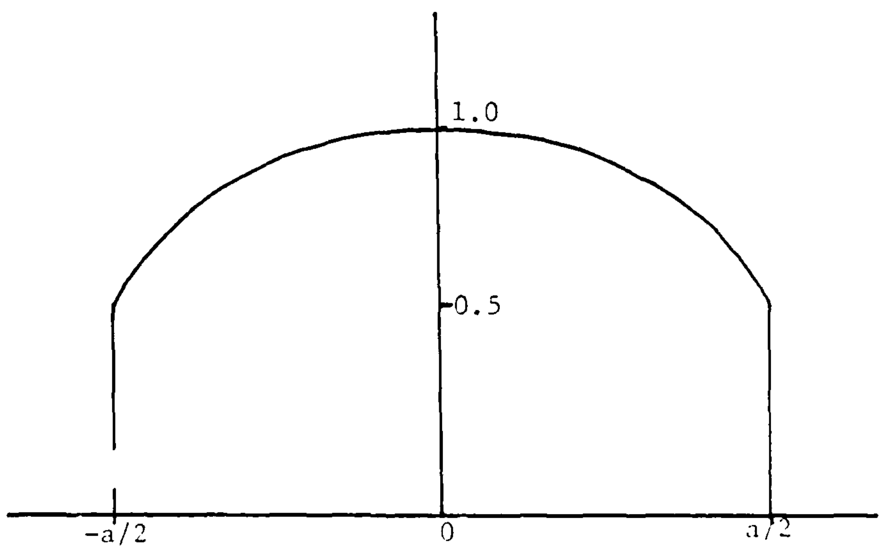
1. uniform



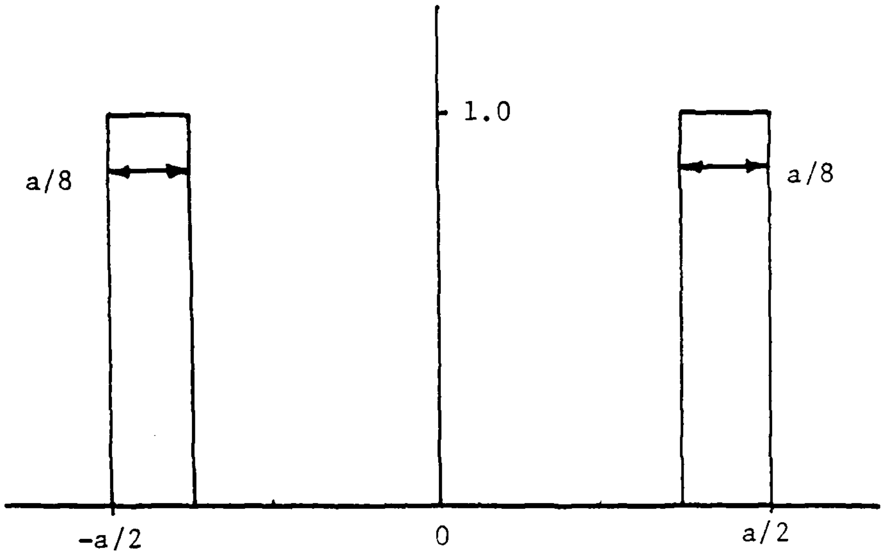
2. double cosine



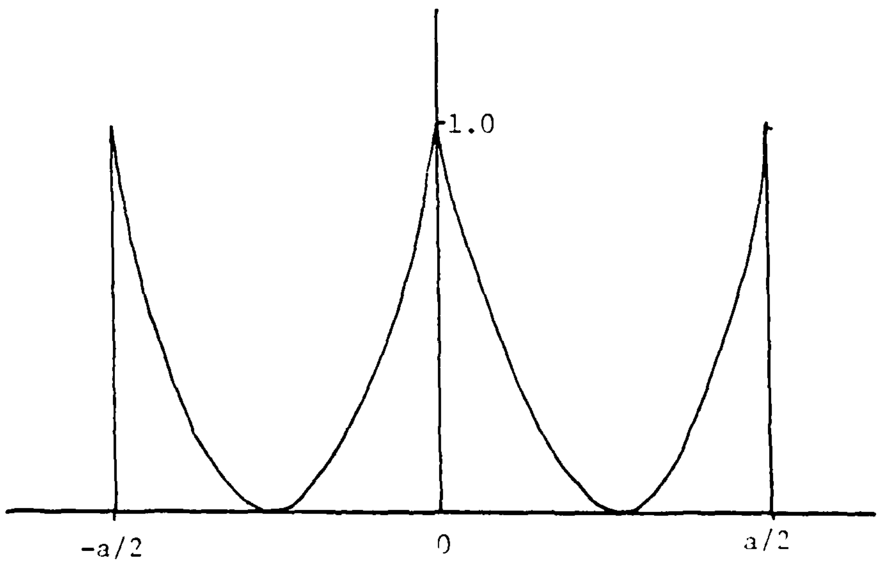
3. triangle



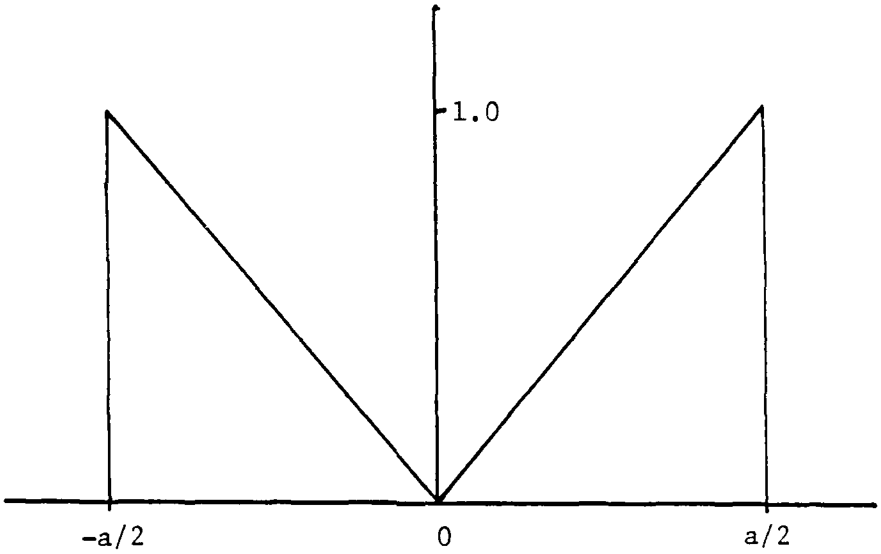
4. cosine



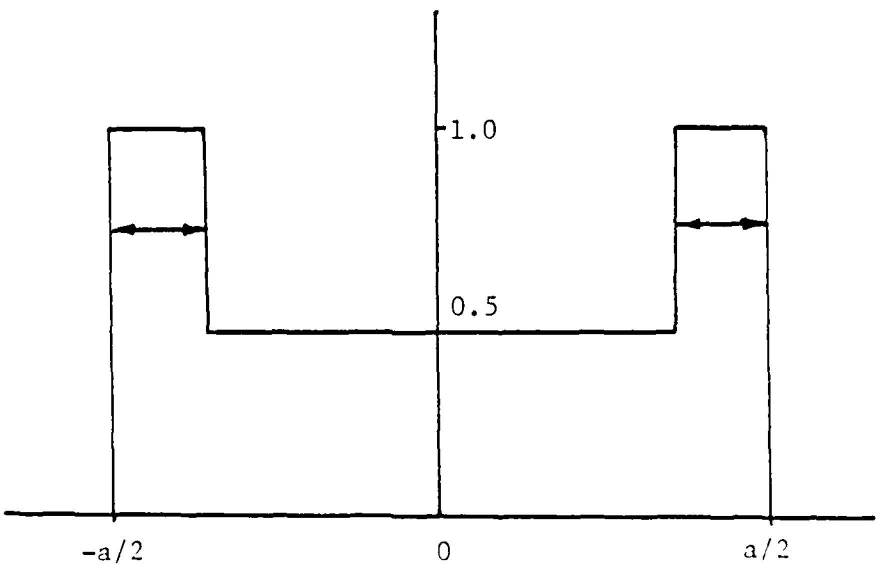
5. edge



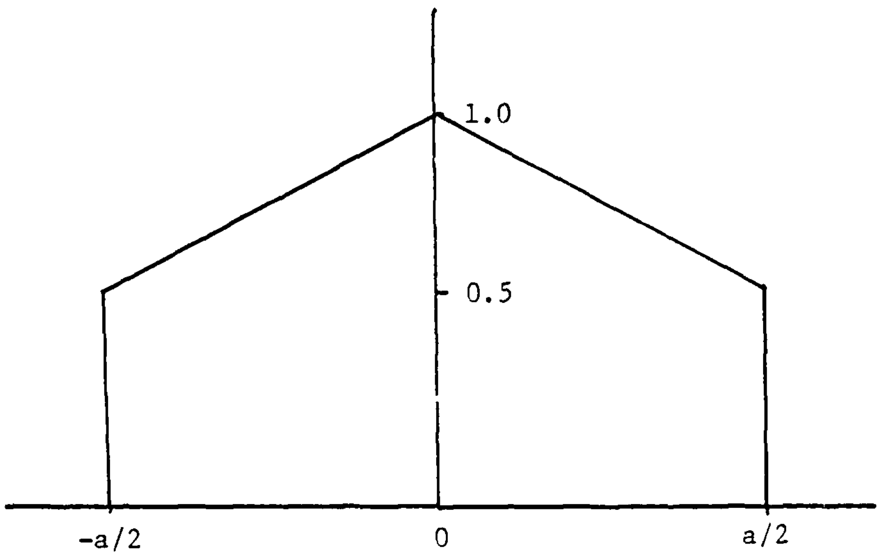
6. complement of double cosine



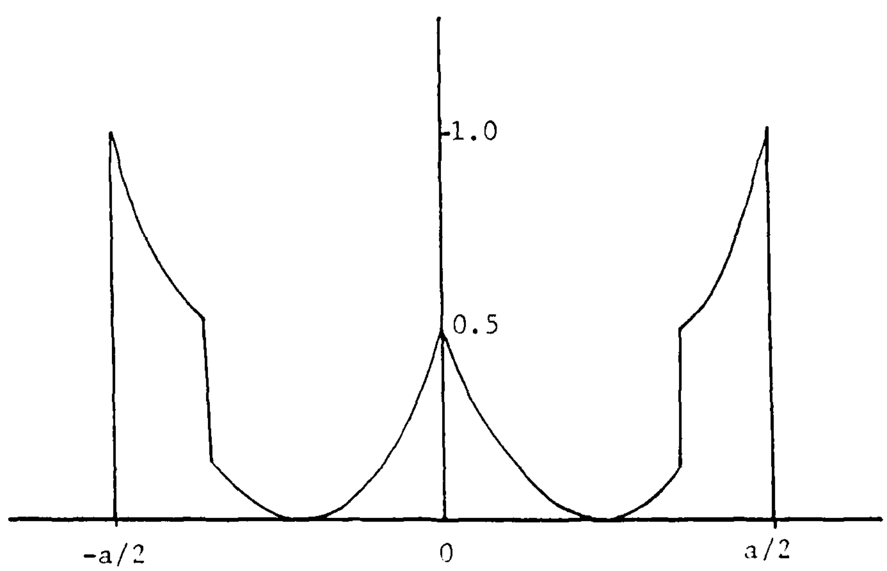
7. notch



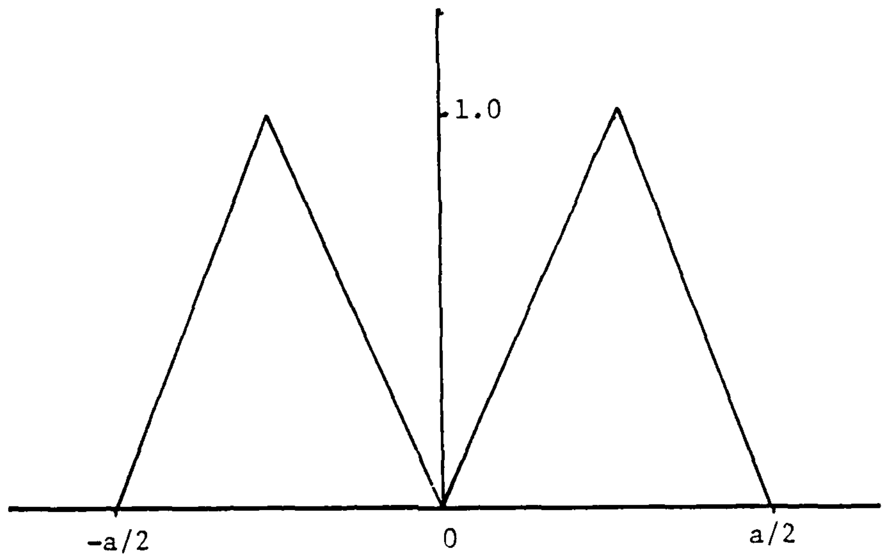
8. filled edge



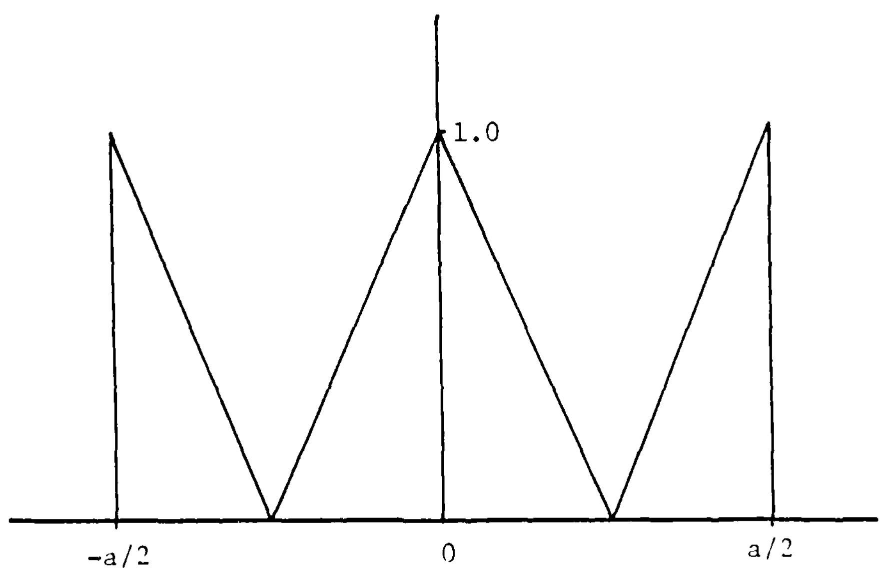
9. 1 + triangle



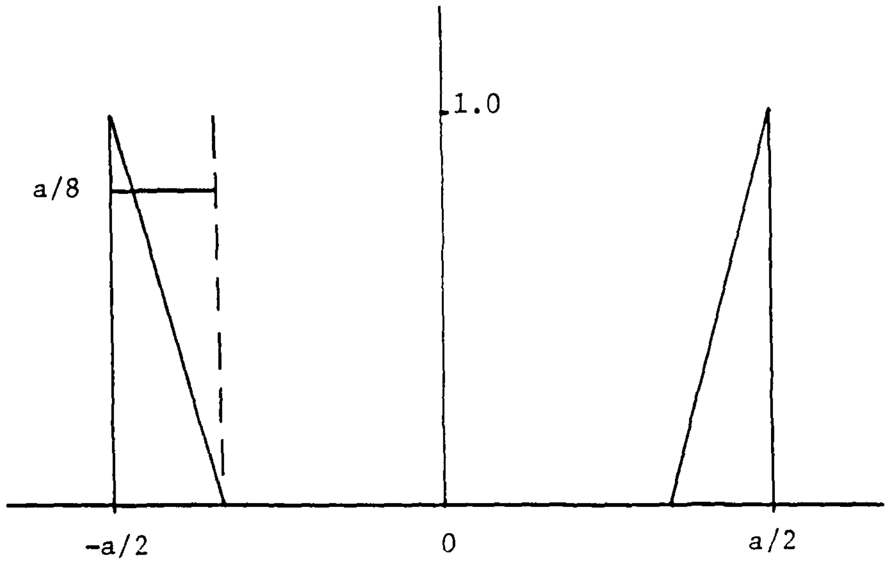
10. edge + complement of double cosine



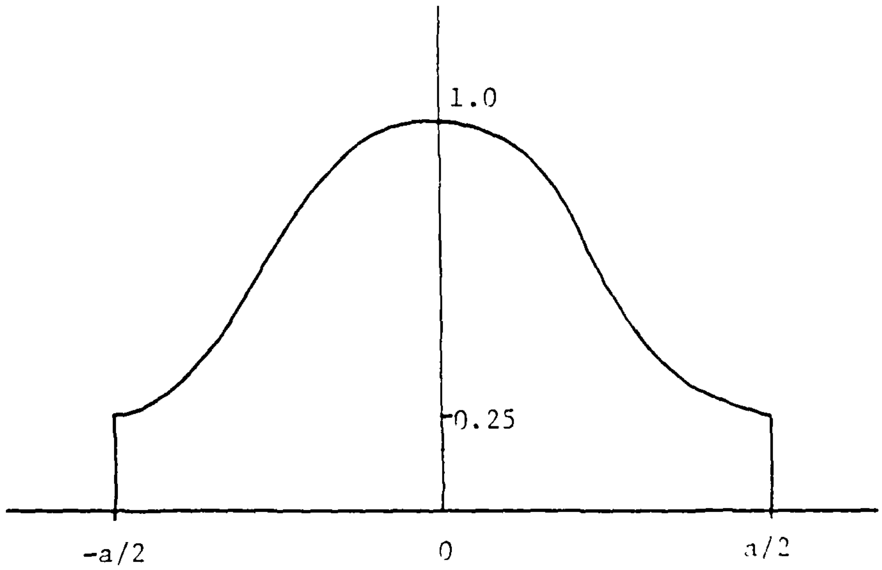
11. double triangle



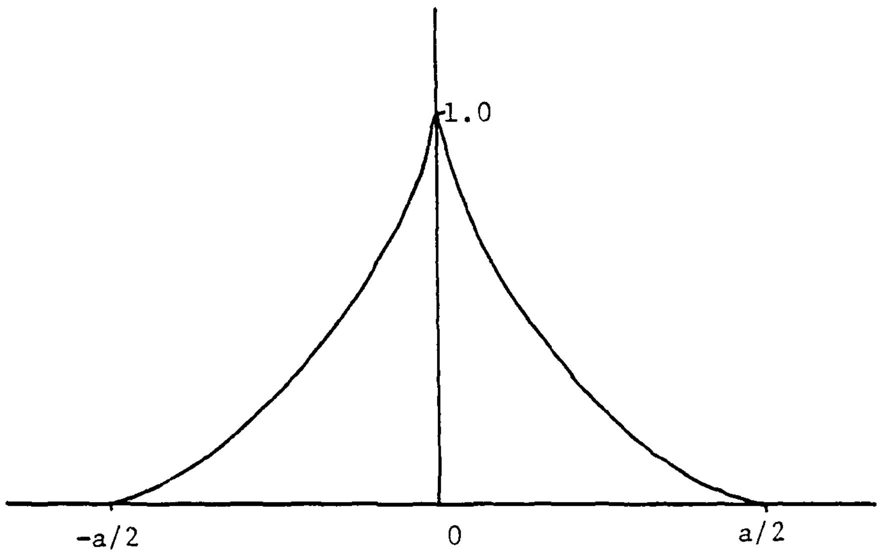
12. complement of double triangle



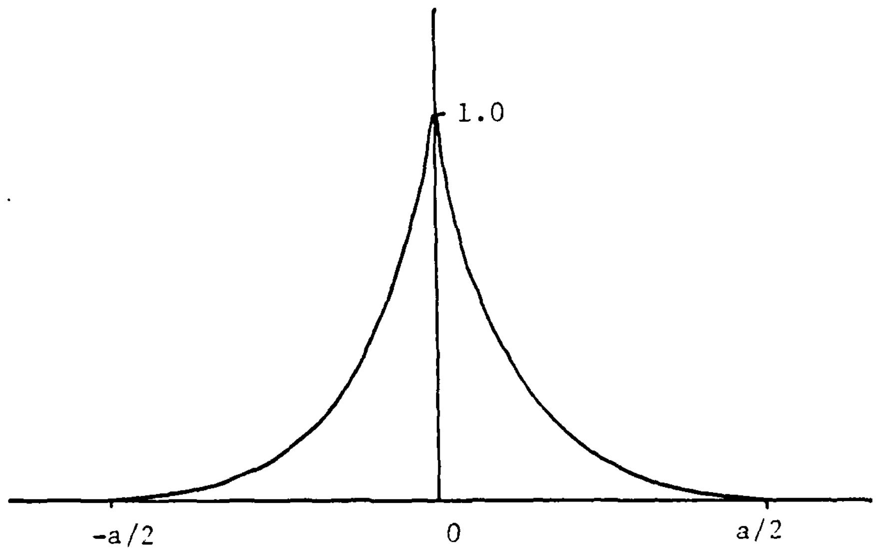
13. knife edge



14. cosine squared



15. triangle squared



16. (triangle)**4

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Establishment sponsoring a contractor's report, or tasking agency, are entered in section 8.) NATIONAL DEFENCE DEFENCE RESEARCH ESTABLISHMENT OTTAWA SHIRLEY BAY, OTTAWA, ONTARIO K1A 0Z4 CANADA	2. SECURITY CLASSIFICATION (overall security classification of the document, including special warning terms if applicable) UNCLASSIFIED

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S,C or U) in parentheses after the title.)

AN INVESTIGATION OF THE EXTENT OF THE FRESNEL REGION OF APERTURE ANTENNAS WITH ATTENTION TO SUPERDIRECTIVITY (U)

4. AUTHORS (Last name, first name, middle initial)

CHARLAND, SHAWN

5. DATE OF PUBLICATION (month and year of publication of document) AUGUST 1989	6a. NO. OF PAGES (total containing information. Include Annexes, Appendices, etc.) 88	6b. NO. OF REFS (total cited in document) 16

7. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

DREO TECHNICAL NOTE

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address.)

DEFENCE RESEARCH ESTABLISHMENT OTTAWA
Ottawa, Ontario
K1A 0Z4

9a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant)	9b. CONTRACT NO. (if appropriate, the applicable number under which the document was written)
10a. ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique to this document.) DREO TECHNICAL NOTE 89-20	10b. OTHER DOCUMENT NOS. (Any other numbers which may be assigned this document either by the originator or by the sponsor)

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification)

(XX) Unlimited distribution
( ) Distribution limited to defence departments and defence contractors; further distribution only as approved
( ) Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
( ) Distribution limited to government departments and agencies; further distribution only as approved
( ) Distribution limited to defence departments; further distribution only as approved
( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in 11) is possible, a wider announcement audience may be selected.)

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), or (U). It is not necessary to include here abstracts in both offical languages unless the text is bilingual).

(U) This technical note describes the principles governing superdirective antennas and the factors limiting their realization. A criterion for determining the minimum distance to the far field of aperture antennas, including superdirective antennas, is presented. Under this condition it is shown that, at least for the marginally superdirective case, such antennas do not exhibit large fresnel regions as compared with uniformly phased apertures. A modified criterion is also presented which defines the minimum distance to the far field in a given direction. Under this criterion it is shown that for the uniform distribution, the antenna pattern near nulls requires a large distance to converge to the far field form. This work has applications in determining antenna test range requirements.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Superdirectivity
Aperture Antenna
Fresnel Region