1

CANADIAN THESES ON MICROFICHE

THÈSES CANADIENNES SUR MICROFICHE

National Library of Canada
Collections Development Branch

Canadian Theses on
Microfiche Service

Ottawa, Canada
K1A 0N4

Bibliothèque nationale du Canada
Direction du développement des collections

Service des thèses canadiennes
sur microfiche

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

Canada

0-315-19907-5

National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses
on Microfiche Service

Services des thèses canadiennes
sur microfiches

Ottawa, Canada
K1A 0N4

67814

PERMISSION TO MICROFILM — AUTORISATION DE MICROFILMER

• Please print or type — Écrire en lettres moulées ou dactylographier

Full Name of Author — Nom complet de l'auteur

Shawn Douglas Chauland

Date of Birth — Date de naissance

02/03/61

Country of Birth — Lieu de naissance

Canada

Permanent Address — Résidence fixe

Box 32, R.R.#2
Richmond, Ont.
K0A 2Z0

Title of Thesis — Titre de la thèse

An Investigation of the Extent of the Fresnel region of
Aperture Antennas with Attention to Superdirectivity

University — Université

Carleton U.

Degree for which thesis was presented — Grade pour lequel cette thèse fut présentée

M. Eng

Year this degree conferred — Année d'obtention de ce grade

1985

Name of Supervisor — Nom du directeur de thèse

J. S. Wight

Permission is hereby granted to the NATIONAL LIBRARY OF
CANADA to microfilm this thesis and to lend or sell copies of
the film.

The author reserves other publication rights, and neither the
thesis nor extensive extracts from it may be printed or other-
wise reproduced without the author's written permission.

L'autorisation est, par la présente, accordée à la BIBLIOTHÈ-
QUE NATIONALE DU CANADA de microfilmer cette thèse et de
prêter ou de vendre des exemplaires du film.

L'auteur se réserve les autres droits de publication, ni la thèse
ni de longs extraits de celle-ci ne doivent être imprimés ou
autrement reproduits sans l'autorisation écrite de l'auteur.

Date

30 April/85

Signature

Shawn Chauland

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

by

SHAWN CHARLAND, B.ENG.

A thesis submitted to the
Faculty of Graduate Studies and Research
in partial fulfilment of the requirements
for the degree of
Master of Engineering

Faculty of Engineering
Department of Systems and Computer Engineering
Carleton University

© copyright

1985, Shawn Charland

2 April, 1985

The undersigned hereby recommends to the Faculty of Graduate
Studies and Research acceptance of the thesis

"An Investigation of the Extent
of the Fresnel Region of Aperture Antennas
With Attention to Superdirectivity"

Thesis Supervisor

Chairman,
Department of Systems and
Computer Engineering

Department of Systems and Computer Engineering

Faculty of Engineering

Carleton University

22 April, 1985

ABSTRACT

This thesis describes the principles governing super-directive antennas and the factors limiting their realization. A criterion for determining the minimum distance to the far field is presented. Under this criterion, it is shown that at least for the marginally superdirective case such antennas do not exhibit large Fresnel regions compared to uniformly phased apertures. A modified criterion is also presented which defines the minimum far field distance 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.

Acknowledgements

The author wishes to thank Dr. Jim Wight for his support and insight into the finer aspects of this thesis. Also the suggestions and gentle criticisms of Mr. Fred Symons are acknowledged with gratitude and pleasure.

TABLE OF CONTENTS

	Page
Acceptance Sheet.....	ii
Abstract.....	iii
Acknowledgements.....	iv
Table of Contents.....	v
List of Important Symbols.....	vii
List of Figures.....	ix
Chapter 1 INTRODUCTION.....	1-1
1.1 Background.....	1-1
1.2 Thesis Organization.....	1-3
Chapter 2 DESCRIPTION OF APERTURE ANTENNA FIELDS.....	2-1
2.1 An Angular Spectrum of Plane Waves.....	2-1
2.2 Discussion of Complex Angles of Propagation...	2-11
Chapter 3 THE SUPERDIRECTIVE ANTENNA.....	3-1
3.1 Superdirectivity.....	3-1
3.2 Superdirective Antenna Synthesis.....	3-2
3.3 Realizability and the Supergain Factor.....	3-8
Chapter 4 THE MINIMUM DISTANCE TO THE FAR FIELD.....	4-1
4.1 Definition of the Fresnel Region.....	4-1
4.2 Minimum Far Field Criteria.....	4-9

1

LIST OF IMPORTANT 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_0 - fixed radial distance from the aperture
- D - aperture width
- λ - wavelength
- η, Z_0 - intrinsic admittance/impedance of the medium

- ω - angular frequency, in rad./sec.
- E - electric field
- H - magnetic field
- v - velocity
- f - frequency, in cycles/sec.
- ψ - see equation (106)
- u - direction of the linearly polarized electric field
- $\hat{n}$ - normal out of the aperture

LIST OF FIGURES

Figure	Description	Page
1	Diagram of one-dimensional antenna geometry.....	2-2
2	Geometry for the phase expression.....	2-3
3	Power flow associated with evanescent waves....	2-9
4	Contour of admissible angles of propagation....	2-15
5	Synthesized radiation pattern.....	3-6
6	Synthesized superdirective radiation pattern...	3-7
7a	Geometry for deriving the Fresnel Kernel.....	4-3
7b	Geometry for deriving the minimum distance to the far field.....	4-7
8	Graph of normalized half power beamwidth vs. distance from the aperture.....	4-19
9a	Graph of the mean squared difference criterion vs. distance from the aperture.....	4-22
9b	Graph of the mean squared difference criterion vs. distance from the aperture, cont'd.....	4-23
10	Graph of HPBW vs. distance to the far field (phase-insensitive criterion).....	4-24
11	Graph of $D_2 \times$ HPBW vs. distance from the	

	aperture (phase-insensitive criterion).....	4-26
12	Graph of $\sqrt{2\text{nd moment}}$ of the power pattern vs. distance from the aperture (phase-insensitive criterion).....	4-27
13	Convergence of radiation patterns.....	4-29
14	Graph of HPBW vs. distance to the far field (phase-sensitive criterion).....	4-30
15	Graph of $\sqrt{2\text{nd moment}}$ of main lobe of the power pattern vs. distance to the far field (phase-sensitive criterion).....	4-32
16	Graph of $\sqrt{2\text{nd moment}}$ of the power pattern vs. distance to the far field (phase-sensitive MSD criterion).....	4-33
17	Graph of the quadratic phase curvature across the aperture for finite ranges.....	5-5
18	A family of continuously varying aperture distributions.....	5-11
19	Trajectories of the zero crossings of the quadratic term in the MSD criterion.....	5-12
20	Comparison of correlations between HPBW and minimum distance to the far field.....	5-13
21	Geometry for the infinitesimal current patch....	6-4
22	Diagram showing angular variation of fields around the infinitesimal current patch.....	6-6
23	Polarization of the total E vector at the four	

	principal angles.....	6-7
24	Geometry for determining the angle of the electric field vector.....	6-9
25	Diagram showing the polarization of the real E vector through one quadrant.....	6-11
26	Diagram showing the polarization of the imaginary E vector through one quadrant.....	6-12
27	Convergence of the uniform aperture radiation pattern.....	6-24

Chapter 1 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 tell us how the electromagnetic fields change with time and space, we find that in order to formulate these relationships in a useful way we must make some mathematical approximations. As one moves away from an antenna, the actual representations for the fields may be approximated by some more useful expressions. Thus one has separate expressions for three distinct regions around the antenna, as defined by the approximations which can be made to the actual field expressions. Close to the antenna there is the near field, in which reactive energy and boundary conditions around the antenna dominate the fields. At slightly larger distances we have the Fresnel region, where the geometrical size of the antenna and resulting path differences are significant and affect the expression for the fields. The radiation pattern is changing with distance through the Fresnel region. Finally, in the far field the

antenna may be treated as a point source and the radiation pattern is constant in shape with increasing distance, subject to a $1/r$ amplitude variation. Between these regions, one set of approximations blends into another so that it is difficult to say exactly where, for instance, the Fresnel region stops and the far field begins. Near to and in this "grey area" it is ambiguous which set of equations may be used. It is known that the "minimum distance to the far field" is a function of the type of aperture distribution, and we will concern ourselves, for reasons which will become apparent, with a class of aperture antennas which are said to be "superdirective". These aperture antennas are superdirective because their directivity exceeds that of a uniformly illuminated aperture, which is the most directive uniform phase aperture.

This thesis documents work done to determine the extent of the Fresnel region around superdirective aperture antennas. There is some evidence to suggest that the minimum distance to the far field for such antennas may be quite large, which contradicts the conventional notion that the outer extent of the Fresnel region (or the inner bound of the far field) can be fixed from purely geometrical arguments. Such an investigation is important because it provides insight into the properties of these antennas and sheds light on the desirability of realizing superdirective

and/or marginally superdirective antennas - something which will be shown to be nontrivial for physical reasons. This information also has applications concerning the minimum test range distance requirements for specific antennas, which is a practical consideration for almost any antenna realization.

1.2 Thesis Organization

Because the thesis deals with a relatively obscure type of antenna application and a specific quantitative question, it is divided into two major parts. The first part deals with the definition, description, and realization of superdirective antennas, and the second part deals with rigorously defining the minimum distance to the far field for a specific aperture distribution.

Chapter 2 presents the concept of describing the fields around an antenna as an angular spectrum of plane waves. It also comprises of a discussion of admissible complex angles of propagation of plane waves around the aperture.

Chapter 3 introduces the notion of superdirective antennas, and describes the idea behind synthesizing superdirective aperture distributions. The supergain factor is defined as a design aid, and some of the design constraints

inherent in realizing such antennas are discussed.

Chapter 4 begins the second part of the thesis, and derives the Fresnel kernel used to describe the fields in the Fresnel region of an antenna, as well as presents the traditional definition of the minimum distance to the far field. Also, three criteria are suggested for rigorously defining the outer extent of the Fresnel region. A quantitative evaluation of these three criteria (and some variations therein) is presented. One is selected as better than the others in the sense of consistency, usefulness, and being indicative of an aperture dependence of the minimum distance to the far field.

Chapter 5 presents an alternate form for the phase-sensitive mean squared difference criterion. This leads to a discussion of the possibility of a synthesis equation for specifying the minimum distance to the far field for an antenna. The curious relationship observed between half power beamwidth and the minimum distance to the far field, as determined by the phase sensitive mean squared difference criterion is also explained.

Chapter 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 of it to the case of a uniform aperture.

Finally Chapter 7 summarizes the accomplishments of the thesis investigation and produces recommendations for future work in this area.

Chapter 2 DESCRIPTION OF APERTURE ANTENNA FIELDS

2.1 An Angular Spectrum of Plane Waves

The concept of representing electromagnetic fields around an antenna as an angular spectrum of plane waves is important to further discussions. It was first described in a paper by H. G. Booker and P. C. Clemmow [1]. We will now develop this concept following closely the development in the original paper.

Fig. 1 shows a diagram of a 1- dimensional antenna radiating into the half space $x > 0$. Following Booker and Clemmow, several simplifications will be adopted. The electromagnetic fields will be assumed to have a time harmonic variation. We will consider only a two- dimensional field in front of the aperture, and the magnetic field will be taken parallel to the z - axis. From Maxwell's equations, we know that if we specify the electromagnetic fields on a boundary in a source free space, then we can calculate the fields over the entire space. From fig. 2 we can derive the expression for the phase of the fields in front of the aperture. If we take k to be the free space vector propagation

Fig. 1 : Diagram of one dimensional antenna geometry

Fig. 2 : Geometry for the phase expression

constant $(2\pi/\lambda)$ then the phase is given by

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

where

$$R = (x\cos\theta + y\sin\theta) \quad (2)$$

or alternatively

$$\phi = e^{-j(k_x x + k_y y)} \quad (3)$$

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

As stated previously, the fields in the half space $x > 0$ can be completely determined from the tangential field components on the aperture. The fields in the half space may be generally written as below:

$$E(x,y) = A(-\sin\theta, \cos\theta, 0)e^{-jk(x\cos\theta + y\sin\theta)} \quad (4)$$

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

with $A(x,y,z)$ 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) gives us the expressions for the fields in the aperture required to support the specified fields in the half space $x > 0$. These expressions are:

$$E_y = A(\cos\theta)e^{-jy k \sin\theta} \quad (6)$$

$$H(0, y, 1) = Ae^{-jy k \sin\theta} \quad (7)$$

These expressions represent a wave travelling over the aperture in the $+y$ direction with vector propagation constant $k \sin\theta$. Because this a boundary condition, we know that 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 of the medium. Specifically, the plane wave will propagate at an angle from the normal out of the aperture. We can see this by examining the TEM plane wave constraints:

- (i) $k_x + k_y = k$
- (ii) $E \cdot k = 0$
- (iii) $H \cdot k = 0$

Now

$$E \cdot k = (E_x + E_y) \cdot (k_x + k_y) = E_x k_x + E_y k_y = 0$$

Hence once we specify E_y , E_x and k_y (in the aperture plane) we have immediately specified k_x , which leads to a plane wave propagating out of the aperture at an angle whose cosine is (k_x/k) . It is now natural to ask what values of k_y are admissible in our discussion (or alternatively, what values of $\sin\theta$ are admissible).

The wave equation may be written as, for our two-dimensional case

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

which requires that

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

and we can see that 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 will be presented later. Now note that time harmonic dependence of the fields is implied, and that k_y is really the change of phase of the wave per free space wavelength. For values of $k_y < k$, this means the phase velocity of the wave in the aperture plane is greater than the free space velocity. But if $k_y > k$, then we have more than 2π radians of phase change per wavelength and we

have described a wave with a wavelength shorter than that of free space. Since phase velocity is given by

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

we can see that, just as waves with $k_y < k$ propagate at velocities above the free space value, so waves with $k_y > k$ propagate at velocities below it.

However, referring back to eqns. (3) and (4), we see that slow waves in the aperture lead to a different kind of wave in the half space. Although $\sin \theta$ is real, $\cos \theta$ has become imaginary, and this has resulted in an exponential damp in the fields with increasing distance away from the aperture, as reflected in the adjusted equations below:

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

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

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 they arise from considering all possible boundary

conditions we could impose at the aperture.

Finally, these waves are associated with energy storage around the aperture. From the Poynting theorem, we know that $\mathbf{E} \times \mathbf{H}$ gives us the vector of energy flow associated with an electromagnetic wave. To determine this vector, we can separately consider the components of the $\mathbf{E}$ field in the plane of the aperture. $E_x \times \mathbf{H}$ yields the component of the energy flow along the y -axis, and that energy leaving the "top" of the aperture is balanced by an identical vector pointing into the "bottom" of the aperture. $E_y \times \mathbf{H}$ gives 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 $\mathbf{E}$ field in the plane of the aperture is in time (or space) quadrature with the $\mathbf{H}$ field. From fig. 3, we see that $E_x \times \mathbf{H}$ 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 we include the possibility of evanescent waves in our aperture distribution, then it becomes clear that any arbitrary time varying aperture distribution can, by Fourier analysis, be decomposed into a spectrum of sinusoidal waves propagating over the aperture. The evanescent waves are

Fig. 3 : Power flow associated with evanescent waves

necessary in allowing for "details" in the aperture distribution finer than the free space fundamental. 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, we can describe any arbitrary field distribution in the half space as an angular spectrum of plane waves (Woodward and Lawson, [2]). The summation of these plane waves in the half space corresponds to integrating 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:

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

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

where

$$P(\sin\theta) = (\lambda) CA(\sin\theta) \quad (15)$$

since "A" is a function of the particular harmonic we are describing, i.e. a function of $\sin\theta$ (or k_y). The aperture

distribution, completely specified by the electric field, is given by setting $x = 0$, and is

$$E(0,y) = \frac{1}{\lambda} \int_{-\infty}^{\infty} P(\sin\theta) e^{-jyky} d\sin\theta \quad (16)$$

This can be seen to be 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 alternatively $-k < ky < +k$) 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 away from the aperture gives a literal meaning to saying the pattern here lies in the "invisible" region - these waves are "invisible" in 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 . It only requires that

$$\sin^2\theta + \cos^2\theta = 1$$

Thus values of $\sin\theta > 1$ or $\sin\theta < -1$ correspond to imaginary values of $\cos\theta$ and complex angles. However, it does constrain the range of admissible complex angles. From our antenna problem (as depicted in fig. no. 1) we draw on boundary conditions to further investigate admissible angles of propagation.

We can first see that $\sin\theta$ must be real, but otherwise of arbitrary value. If $\sin\theta$ were imaginary or had an imaginary part, then $e^{jky \cdot y}$ would have an exponential real part, and the fields would go to infinity in either the $+y$ or $-y$ direction. Also we can see that $\cos\theta$ must not have any positive imaginary part. If it did, the fields would go to infinity exponentially in the $+x$ direction. If we define our complex angle as $(a + jb)$ then we can examine $\sin\theta$ for the 4 possible sign permutations of the angle:

$$\sin\theta = (e^{j\theta} - e^{-j\theta})/2j \quad (17)$$

or we can write, for $(a + b)$

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

or alternatively,

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

Now in order for $\sin\theta$ to have no imaginary part, we must have that $a = \pi/2$ or $-\pi/2$, unless $b = 0$. If $b = 0$ we are dealing with only real angles, and $\sin\theta$ will be between +1 and -1. The sign permutations of the complex angle gives values of $\sin\theta$ as shown below:

$$(\frac{\pi}{2} + jb) \dots \sin\theta = (e^{-b} + e^b)/2 = \cosh(b) \quad (20)$$

$$(\frac{\pi}{2} - jb) \dots \sin\theta = (e^{-b} + e^b)/2 = \cosh(b) \quad (21)$$

$$(-\frac{\pi}{2} + jb) \dots \sin\theta = -(e^{-b} + e^b)/2 = -\cosh(b) \quad (22)$$

$$(-\frac{\pi}{2} - jb) \dots \sin\theta = -(e^{-b} + e^b)/2 = -\cosh(b) \quad (23)$$

To determine further constraints on the complex angle, we examine values of $\cos\theta$ for the four sign permutations. Recall that we have already constrained that if b is non zero, then $a = \pm\pi/2$.

$$\cos\theta = (e^{j\theta} + e^{-j\theta})/2 \quad (24)$$

or alternatively, for $(a+jb)$

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

and the permutations give:

$$(\frac{\pi}{2} + jb) \dots \cos\theta = -j(e^b - e^{-b})/2 = -j\sinh(b) \quad (26)$$

$$(\frac{\pi}{2} - jb) \dots \cos \theta = j(e^b - e^{-b})/2 = j \sinh(b) \quad \text{eqn (27)}$$

$$(-\frac{\pi}{2} + jb) \dots \cos \theta = j(e^b - e^{-b})/2 = j \sinh(b) \quad \text{eqn (28)}$$

$$(-\frac{\pi}{2} - jb) \dots \cos \theta = -j(e^b - e^{-b})/2 = -j \sinh(b) \quad \text{eqn (29)}$$

Recalling that $\cos \theta$ may not have a positive imaginary part, we can readily see that, to satisfy the boundary conditions, $(a-jb)$ and $(-a+jb)$ are inadmissible complex angles. A similar discussion may be found in the paper by Woodward and Lawson [2]. Following Woodward and Lawson, a diagram of the contour of admissible angles in the complex plane may be found in fig. 4.

The above discussion lends still more insight into the nature of the invisible region and evanescent waves. From eqn. (4) we saw that the amplitude weighting of the components of the electric field were functions of the angle of propagation resolved on the unit vectors, as

$$A = A(-\sin \theta, \cos \theta, 0) \quad (30)$$

When in the invisible region ($|\sin \theta| > 1$), we know by looking at the admissible angles that we can write $\sin \theta$ as $\pm \cosh(b)$ and $\cos \theta$ as $-j \sinh(b)$, i.e.

$$A = A(\pm \cosh(b), -j \sinh(b), 0) \quad (31)$$

By definition, $\cosh(b)$ and $\sinh(b)$ will essentially

Fig. 4 : Contour of admissible angles of propagation

increase exponentially with increasing b . This means that the fields in the aperture plane required to support the evanescent waves will increase exponentially as the waves become more evanescent, that is, as we move farther into the invisible region. This fact will be important in later discussions.

Chapter 3 THE SUPERDIRECTIVE ANTENNA

3.1 Superdirectivity

An antenna whose directivity is greater than that of a uniformly illuminated aperture is said to be superdirective. It is known that, for the case of uniform phase distribution, the uniform amplitude distribution yields the highest directivity. This becomes apparent by looking at the definition of directivity, and using the Cauchy- Schwartz inequality. Directivity is defined as maximum radiation intensity per unit solid angle normalized to average radiation intensity.

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

The Cauchy- Schwartz inequality is stated as

$$\left| \int f g dx dy \right|^2 \leq \int f^2 dx dy \int g^2 dx dy \quad (33)$$

Thus we can see that

$$\left| \int E(x,y) dx dy \right|^2 < A \int |E(x,y)|^2 dx dy \quad (34)$$

and the uniform case yields the highest directivity.

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

The basic idea here 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 radiation patterns and their corresponding aperture distributions may be added linearly. 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 by finite length and finite energy apertures are called "aperture limited" functions (Rhodes, [12]).

3.2 Superdirective Antenna Synthesis

Following the method for synthesizing arbitrary radia-

tion patterns laid out by Woodward and Lawson, we first specify the desired radiation pattern at n points over the visible region. To realize this, we use a sum of appropriately weighted (independent) aperture limited functions. The composite pattern will have the form

$$p = \sum_s A_s p_s \quad (35)$$

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 [12] in which he came to the same conclusion. His method centres on achieving an arbitrary radiation pattern in a best mean-square way 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, we will briefly follow an example worked by Woodward and Lawson. They begin by choosing as their independent component patterns the following function evaluated at integer values of s :

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

W is the aperture width in wavelengths.

If we specify the radiation pattern at values of

$$\sin\theta = r/W \quad (37)$$

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

$$p(r/W) = \sum_s A_s p_s(r/W) \quad (38)$$

and using the orthogonality of sine functions, these reduce to

$$A_r = p(r/W) \quad (39)$$

so the composite radiation pattern has the form

$$p(\sin\theta) = \sum_s p\left(\frac{s}{W}\right) \frac{\sin(\pi W(\sin\theta - s/W))}{\pi W(\sin\theta - s/W)} \quad (40)$$

Fig. 5 shows how the resulting radiation pattern is built up of the component patterns, and how the maximum of each component lies on zeros of the other patterns. Thus we can see that if we wish to specify the desired radiation pattern at values between points at r/W , we must use the sidelobes of the component patterns whose main beams lie in the imaginary region. Although this can be done theoretically, the control of the visible region by using the sidelobes of patterns in the invisible region is very weak. Large patterns will result in the invisible region for only marginal adjustments to the pattern over the visible region. This is nicely illustrated by the radiation pattern shown in fig. 6. From discussions in previous sections, we know that patterns in the invisible region correspond to evanescent waves and energy storage around the aperture. We also know that the electric and/or magnetic fields in the aperture required to support these fields increase roughly exponentially as we move into the invisible region. For this reason, large patterns in the invisible region generally mean that we have specified an unrealizable aperture distribution. Thus, coupled with the fact that the control of the

Fig. 5 : Synthesized radiation pattern

Fig. 6 : Synthesized superdirective radiation pattern

visible region by patterns in the invisible region is so weak, we conclude that realizing significantly superdirective aperture distributions is impractical.

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. These physical constraints arise so quickly when one attempts to improve over the uniform case that even marginally superdirective antennas are difficult to realize. These limitations include very high voltages and currents and attendant high ohmic losses, large amounts of reactive energy, critical tolerances on element spacing and phasing, narrow bandwidth, and low radiation resistance.

Specifically, in a note concerning the practical limitations in the directivity of antennas, R. M. Wilmette [5] tells of designing 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 ridiculously 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 is increased beyond a certain point, and then the radiation resistance falls off very quickly. Similarly Schelkunoff [28] 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 result in reactive fields extending a large distance from the antenna. These reactive fields "redirect" passing energy and channel it to the antenna. We see that the notion of a small radiation resistance is intrinsic to the idea of a superdirective antenna. Further, Schelkunoff [28] concludes that detuned superdirective antennas do not channel energy to themselves, and so reactive fields are likewise intrinsic to superdirective antennas.

T. T. 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

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

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 be loosely interpreted as

$$\gamma = 1 + P_j/P_r \quad (42)$$

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 of 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

$$P_j = 2\omega(W_m - W_e) \quad (43)$$

where W_m and W_e are the maximum stored magnetic and electric energies respectively. Stored electric and magnetic energies are 180 degrees out of phase (one is written with a negative sign), so under conditions of resonance the reactive stored energy may be very large (while P_j is zero) and hence our design would be unrealizable. In answer to this problem, Donald R. Rhodes [10] derived an expression for the net "observable" stored electric and magnetic energies. By observable energies we mean 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 will contain a surplus of either electric or magnetic energy per cycle, and P_j is in fact this surplus. He derives separate expressions for W_m and W_e , and defines P_j' as

$$P_j' = 2w(W_m + W_e) \quad (44)$$

or, more explicitly,

$$2\omega W_m = \frac{4\pi^2}{2kZ_0} \int_{ky^2+kx^2>k^2} \frac{|kxF_y - kyF_x|}{ky^2+kx^2-k^2} dk_y dk_x \quad (45)$$

$$2\omega W_e = \frac{4\pi^2}{2kZ_0} \int_{ky^2+kx^2>k^2} \frac{k^2(|F_y|^2 + |F_x|^2)}{ky^2+kx^2-k^2} dk_y dk_x \quad (46)$$

where F_y and F_x are defined as

$$F_y(ky, kx, k) = \frac{1}{4\pi^2} \iint_S E_y(x, y, 0) e^{j(yky + xkx)} dy dx \quad (47)$$

$$F_x(ky, kx, k) = \frac{1}{4\pi^2} \iint_S E_x(x, y, 0) e^{j(yky + xkx)} dy dx \quad (48)$$

for the general case of a two dimensional aperture in the yz plane. This means 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 pros-

pective 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 in a previous section, by superimposing translated and scaled aperture limited patterns we may tailor any radiation pattern we desire at the price of generally increasing the stored energy around the aperture (i.e. 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, consider the reactive fields resulting from synthesizing an isotropic source).

Chapter 4 THE MINIMUM DISTANCE TO THE FAR FIELD

4.1 Definition of the Fresnel Region

Having described the notion of superdirective apertures at some length and concluded that there are formidable problems in realizing them, we now turn to an aspect of these antennas which promises to be of some practical interest (if we could have a superdirective aperture). There is some evidence that suggests that the Fresnel region of superdirective antennas may be quite large compared to the case of uniform phase apertures. We will proceed to define the Fresnel region and derive expressions for the fields there.

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 fields. The Fresnel- Kirchhoff scalar diffrac-

tion field is given by

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

Very close to the aperture is the near field or reactive field. No simplifications can be made to the expression for the fields here. In fact, as Skolnik [11] points out, the scalar diffraction field is itself just an approximation very close to the aperture because boundary conditions have been ignored. In this region the reactive field dominates. As we move outward from the aperture, the reactive or evanescent field dies away and we enter the Fresnel region, where the radiation pattern is changing with distance from the aperture. This 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 to us. Referring to fig. 7a we can derive the expression for the fields in the Fresnel

Fig. 7a : Geometry for deriving the Fresnel Kernel

region. From the geometry of the diagram, we can write the expression for the fields in front of the aperture as

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

The radial distance from a point on the aperture can be written as

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

or

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

By appealing to the approximation

$$\sqrt{1 + \Delta} \approx 1 + \frac{\Delta}{2} \quad (53)$$

we can write

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

Now also from fig. 7,

$$\frac{y}{\rho} = \sin\theta \quad (55)$$

Thus eqn. (51) becomes

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

Eqn. (56) is characterized by the exponential. $e^{+jy\sin\theta}$ is called the Fourier kernel, and is associated with the far field or the Fraunhofer region. $e^{-jky^2/2\rho}$

is called the Fresnel kernel, and is associated with the Fresnel region. In the far field, the Fresnel kernel disappears since ρ 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 impinging on the aperture from a point source a finite distance in front of the aperture. Since ρ is the distance from the aperture it is apparent that the radiation pattern in the Fresnel zone is changing with distance, whereas in the far field all dependence on distance has been removed (except the $1/r$ variation in amplitude).

There are no fixed, well-defined boundaries stating where the near field and Fresnel region and the Fresnel region and far field begin and end. 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 than 1% of the radiated field. From this, he concludes that the Fresnel approximations should not be

used to predict fields closer than 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 historically been taken as $2D^2/\lambda$, where D is the maximum dimension of the aperture. This is an arbitrary but (generally) conservative minimum distance at which we can assume the far field approximations apply. It arises by imposing the constraint that the maximum allowable edge phase error on the aperture is $\pi/8$ rad. From fig. 7b,,

$$y - x = \lambda/16 \quad (57)$$

and also

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

Substituting, this brings us to

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

Isolating x in eqn. (59) we can write

$$x = \frac{64D^2 - \lambda^2}{128\lambda} \quad (60)$$

Fig. 7b : Geometry for deriving the minimum distance to the far field

and if $D \gg \lambda$, then

$$x \approx \frac{D^2}{2\lambda} \quad (61)$$

which is the Rayleigh distance of optics. As stated previously, the conventional minimum distance to the far field is $2D^2/\lambda$, (or 4 times the above derived distance) to be conservative. Note that with the $2D^2/\lambda$ definition and the reactive field criterion, we cannot use the Fresnel approximations for antennas smaller than 2 wavelengths. For this size of aperture the minimum distance to the far field and the distance at which the reactive field is 1% of the radiated field are the same.

We have seen that the definition of the distance to the outer Fresnel boundary is quite arbitrary. It is natural to ask if in fact the aperture distribution in any way affects the distance at which we can begin using the far field approximations. Jull [16] 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 will be 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 finite range effects and properties of antennas are important when dealing with

large antennas, such as reflectors. For example, a heavy route satellite station may have a paraboloidal reflector 30 metres in diameter. From our $2D^2/\lambda$ approximation, an up-frequency of 6GHz places the outer Fresnel region some 36km away, so all field measurement of the reflector will be restricted to the Fresnel region. Even for smaller antennas far field measurements may prove difficult or impossible. At 35GHz a 1 metre reflector requires a test range of over 230 metres, which is difficult to arrange due to multipath problems. 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 do this, we must examine the implications, at a fundamental level, of saying the far field begins at this fixed distance. As stated previously, the radiation pattern is changing with distance as we move through the Fresnel region. Thus we can think of the radiation pattern "evolving" or converging in some sense to the far field pattern as we move through the Fresnel region. Also, the criterion we use to rigorously define where the far field begins should in some way reflect that the traditional con-

servative placement of the boundary is acceptable for "most" uniform phase cases. A standard or point of reference is required, and this reference should be related to the conservative placement of $2D^2/\lambda$.

Silver [20] (pp. 196 ff) uses Fresnel zones to define the minimum distance at which the far field assumptions can be applied. The aperture may be divided up into Fresnel zones by considering a point some distance away on boresight, surrounded by concentric spheres of radii increasing in steps of $\lambda/2$. The intersections of these spheres and the aperture form rings on the aperture. Successive rings on the aperture, then, give alternately positive and negative contributions to the field at the point on boresight - this is due to the path differences between the various rings and the field point. Following Silver, we note that the field at our field point is given by the sum of the contributions of all the rings, or (assuming a whole number of fresnel zones in the aperture)

$$F = s_1 - s_2 + s_3 - \dots \quad (62)$$

where s_i is the magnitude of the contribution of the i th Fresnel zone. It can be shown that this sum may be reduced to (approximately)

$$S = 0.5(s_1 + s_N) \quad (63)$$

depending on whether the Nth zone is odd or even. Thus we can see that sliding our field point along boresight the on-axis field goes through maxima and minima 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 will be 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 degrees 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 [20, p 562] also quotes radiation pattern beamwidth as an indication of the minimum distance at which one can assume Fraunhofer conditions 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 we could use to quantitatively say that the radiation pattern at a finite distance is as good as a far field pattern. One criterion might be the measure of the half power beamwidth (HPBW) of the evolving pattern. We could say that once the HPBW of the radiation pattern is reduced to some factor of the far field HPBW (say, 1.1 times the far field value) then we are in the far field. This proposal arises from considerations which originally prompted this investigation. We have, as an approximate relationship,

$$\text{HPBW} = \lambda / D = \theta \quad (64)$$

Thus eqn. (61) can be written as

$$D^2 / 2\lambda = D / 2\theta = R \quad (65)$$

If we state 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 will increase very rapidly. For the standard or reference previously mentioned, it is reasonable to turn to the case of a uniform phase - uniform amplitude aperture and see what the ratio of far field to Fresnel boundary HPBW is, when we take the boundary to be at $D^2 / 2\lambda$. In this way we have incorporated the traditional definition of the

boundary in our standard, and find the ratio to be approximately 0.982. With this criterion, we can, for several aperture distributions, calculate the radiation patterns at different distances from the aperture and compare this ratio to our criterion of 0.982. We say that the far field begins at the distance at which the ratio of the far field HPBW to the finite distance HPBW evaluates to 0.982. Additionally, there are other alternate parameters of the radiation pattern which may show a better correlation to far field distance, while also reflecting the changes in HPBW. Such parameters would be more "rigorously" defined than HPBW, and could include such things as the second moment of the main lobe, 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 criterion for fixing the minimum distance to the far field draws upon the notion that the radiation pattern is "converging" to the far field pattern as we move through the Fresnel region. We could propose a mean squared difference criterion (over the visible region) between the magnitude of the finite distance pattern and the magnitude of the far field pattern. This mean squared

difference criterion, normalized to the far field radiation pattern, has the form:

$$\text{MSD} = \frac{\int_{-k}^k \left| \int_{-a/2}^{a/2} e^{-j\beta y^2} D(y) e^{jyky} dy \right|^2 - \left| \int_{-a/2}^{a/2} D(y) e^{jyky} dy \right|^2 dk}{\int_{-k}^k \left| \int_{-a/2}^{a/2} D(y) e^{jyky} dy \right|^2 dk} \quad (66)$$

(eqn. (66) in fact uses the mean squared difference between radiation power patterns, for programming reasons). The criterion applied over the visible region would yield "real world" measurably verifiable and interpretable results, since the fields predicted by the invisible region have by definition died away by the time we get to the Fresnel region. That is to say, including the invisible region in our criterion could alter the predicted minimum distance to the far field in a way which could not be explained or observed by actual measurements of the fields. The "cut-off" or reference for this criterion again will be drawn from considering the uniform case. Evaluating this mean squared difference criterion for the uniform case at $D^2/2\lambda$ yields a numerical result which will define the minimum distance to the far field. Because the criterion is determined by integration over only the visible region, this numerical result will generally be dependent on aperture size, since lengthening the aperture corresponds to narrowing the radiation pattern and so effectively increasing the amount of the

radiation pattern in the visible region. However, for apertures which have most of their patterns in the visible region already, the numerical result will be relatively insensitive to increases in aperture size.

4.2.4 Phase-Sensitive Criterion

A third criterion for specifying where the far field begins bears a close resemblance to the previously mentioned criterion. We can again propose a normalized mean square difference criterion, but include phase information. In the expression below, we work with the generally complex radiation intensity pattern rather than the power pattern :

$$\text{MSD} = \frac{\int_{-\infty}^{\infty} \left| \int_{-a/2}^{a/2} (e^{-j\beta y^2} - 1) D(y) e^{jyky} dy \right|^2 dk y}{\int_{-\infty}^{\infty} \left| \int_{-a/2}^{a/2} D(y) e^{jyky} dy \right|^2 dk y} \quad (67)$$

Note that we are normalizing with respect to the power pattern. Also, the expression includes integration over the invisible as well as the visible region, since it is known that superdirective apertures misbehave themselves most characteristically in this region. Strong arguments for consistency in the notion of defining where we can assume we have a far field pattern dictate using the invisible region too. That is to say, although the evanescent fields have disappeared at the minimum far field, from a strictly

theoretical point of view they are no less important than the visible region. Also, by example we can see that though there are good arguments for the form of the second presented criterion, something goes wrong for the case of an aperture with delta functions at its edges. In this case, eqn. (66) will evaluate to

$$\text{MSD} = \frac{\int_{-k}^k \left| e^{-j\beta \frac{a^2}{4}} (e^{jky \frac{a}{2}} + e^{-jky \frac{a}{2}}) \right|^2 - \left| e^{jky \frac{a}{2}} + e^{-jky \frac{a}{2}} \right|^2 dk}{\int_{-k}^k \left| (e^{jky \frac{a}{2}} + e^{-jky \frac{a}{2}}) \right|^2 dk} \quad (68)$$

$$= 0$$

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 were investigated for the cases of 16 different uniform phase distributions. The results of these investigations constitute the following section.

4.3 Investigation of the Minimum Far Field Criteria

The three minimum far field criteria were proposed in the hope of rigorously relating an aperture distribution or

its radiation pattern to the minimum distance at which the far field approximations could be used, and so fix the extent of the Fresnel region. The motivation behind this is that we would like to check the hypothesis that the minimum distance to the far field can be very large for the case of superdirective apertures. Since we are looking for a generalization concerning aperture distributions or radiation patterns, 16 different uniform phase aperture distributions were chosen at random and each criterion was applied to each aperture. The uniform phase cases were selected because it was felt that, if the minimum distance to the far field could be related directly to the aperture distribution, it would be easier to draw out the relationship if phase complications were left out for the time being. The selected aperture distributions are completely 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 and lying on the y-axis. The aperture radiates into the half space $x > 0$, and the magnetic field lies on the z-axis. All simplifications drawn in Section 1 pertaining to this figure will 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 we can see the rate at which the patterns from the relative distributions converge (in a HPBW sense presented earlier) as we move away from the aperture. The far field HPBW values for these distributions are shown below:

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

(These HPBW values are measured in the ky plane and are not converted to angles). We can see that 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. This is the opposite of what we expected, and leads us to conclude that by exception, a simple minimum far field criterion

based on how closely the HPBW of the evolving pattern approximates the far field HPBW is not workable. 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 distance to the far field. Analytically, there is no obvious reason to expect that HPBW would be an indicator of the minimum distance. The motivation for investigating this possibility was that, by approximating HPBW by λ/D , we could write the traditional definition of the distance to the far field in terms of HPBW. It turns out that the approximation of the HPBW as a function of λ and D is very rough - only a rule of thumb, as shown in the table below :

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

Thus we exclude the criterion of the ratio of HPBWs from further discussion. This does not necessarily mean that the minimum distance to the far field by some other criteria might not be a function of HPBW, nor does the crudeness of the λ/D approximation devalue the hypothesis that narrow beam patterns would have large Fresnel regions.

The second minimum distance criterion described in the

previous section is given by eqn. (66). It is a mean squared difference of power patterns calculated at a finite distance from the aperture and the far field. Figs. 9a and 9b are graphs of this 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. As stated previously, we define the far field to begin when the MSD criterion evaluates to 0.04794. All apertures are 10 wavelengths long, a fact which, for the aperture distributions selected, is relatively unimportant. Figs. 9a and 9b are used to determine exactly where the MSD criterion is satisfied for each of the 15 aperture distributions. This distance for each aperture is simply the distance at which the respective MSD curve intercepts the line drawn at $MSD = 0.04794$. Fig. 10 shows a graph of HPBW vs. distance to the far field, according to this criterion. We can see that, though a general form seems to be evident, it would be impossible to fit a best curve to the points and have an acceptable scatter. This scatter leads us to conclude that, as for the previous minimum distance criterion, there is no useful relationship to HPBW.

In order to be more rigorous, the definition of aperture width D was changed from the maximum distance between two nonzero points in the distribution to the second moment

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)

Fig. 10 : Graph of HPBW vs. distance to the far field
(phase-insensitive criterion)

of the aperture. Fig. 11 shows a graph of $(\text{HPBW}) \times (D^2)$ vs. distance from the aperture (where D^2 is our new definition of D). Since HPBW is proportional to D , it is also proportional to D^2 , and scaling our vertical axis in this way means that our graph is independent of aperture size for apertures larger than 10 wavelengths (the horizontal axis is scaled with respect to $D^2/2\lambda$). This is only true for the case of aperture distributions which lead to radiation patterns that are essentially zero in the invisible region for the case of a 10 wavelength aperture. At any rate, the points are effectively randomly scattered, and we must again conclude that there is no useful relationship here.

Still another possibility is to replace HPBW with the second moment of the radiation pattern, calculated over the visible region (recall that, for the case of a 10 wavelength aperture, all the distributions selected have power patterns that are essentially zero in the invisible region). Fig. 12 shows a graph of the standard deviation of the power pattern vs. distance to the far field. Again, we can see that the points are almost randomly scattered. A similar graph using the standard deviation of the main lobe also resulted in a fairly random scattering of points.

The last minimum distance criterion suggested is given by eqn. (67), and is the mean square (magnitude) difference

Fig. 11 : Graph of D2 x HPBW vs. distance from the aperture (phase-insensitive criterion)

between the complex radiation intensity patterns calculated at a finite distance and in the far field. This criterion, unlike the previous one, is sensitive to the phase of the radiation pattern. Also, for reasons stated previously, this MSD criterion is applied over the entire radiation pattern (for reasons of mathematical consistency a brief discussion of the mathematical form that the MSD criterion should take to include the invisible region is given in appendix A). Fig. 13 shows a graph of the MSD criteria vs. distance from the aperture for the various aperture distributions. As was described for figs. 9a and 9b, this graph is used to determine the distance at which each aperture distribution satisfies the MSD criterion (numerically, a value of 0.44443). The curves have the same general shape as are found in figs. 9a and 9b, but the relative positions of the curves have been shifted. Fig. 14 shows a graph of HPBW vs. distance to the far field. Here, without qualification, the graph is independent of aperture size since we have integrated over the entire radiation pattern. The apparent correlation between HPBW and distance to the far field by this criterion is striking. In fact, the degree of correlation is difficult to explain, since it would appear to be better than the original approximation of HPBW by λ/D . A graph of the standard deviation of the main lobe vs. distance to the far field was drawn, and is presented in

Fig. 13 : Convergence of radiation patterns

Fig. 14 : Graph of HPBW vs. distance to the far field (phase-sensitive criterion)

fig. 15. We can see that the correlation is still good, but perhaps not as good as for HPBW. This is puzzling because HPBW is such a comparatively "arbitrary" definition of beamwidth. Fig. 16 shows a graph of the standard deviation of the radiation pattern vs. distance to the far field, and we can see that the scatter of the points is excessive. We seem to have "fallen" upon a relationship (for the aperture distributions selected) which is as unexpected as it is remarkable.

Fig. 16 : Graph of $\sqrt{2\text{nd moment of power pattern}}$ vs. distance to the far field (phase-sensitive criterion)

Chapter 5 THE PHASE-SENSITIVE CRITERION

5.1 A Synthesis Form of the Phase-Sensitive Criterion

We have observed that, in some way, the phase-sensitive MSD criterion given by eqn. (67) is a function of HPBW for the aperture distributions chosen. However, in the form of eqn. (67) the relationship is not obvious. In order to gain some insight into this curious correlation and to further investigate the properties of the MSD criterion, we apply Parseval's Theorem to the MSD criterion. Parseval's Theorem is given by eqn. (69) shown below:

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

Since we are examining integrals of squared magnitudes when dealing with the MSD criterion, it is quite natural to use Parseval's Theorem to simplify the expression. We note that both the numerator and denominator are squared magnitudes of Fourier Transforms. Thus we can rewrite the denominator of the MSD criterion as

$$\int_{-\infty}^{\infty} \left| \int_{-a/2}^{a/2} D(y) e^{jyky} dy \right|^2 dk = 2\pi \int_{-a/2}^{a/2} |D(y)|^2 dy \quad (70)$$

Next we can write the numerator of the expression, again using Parseval's Theorem, as

$$\int_{-\infty}^{\infty} \left| \int_{-a/2}^{a/2} (e^{-j\beta y^2} - 1) D(y) e^{jyky} dy \right|^2 dk = 2\pi \int_{-a/2}^{a/2} |(e^{-j\beta y^2} - 1) D(y)|^2 dy \quad (71)$$

It is important to note that the integrals in both the numerator and denominator are now over a finite region, which will make their numerical calculation much more straightforward. Recalling the numerical threshold defined previously, the MSD criterion may be rewritten as

$$\frac{\int_{-a/2}^{a/2} |(e^{-j\beta y^2} - 1) D(y)|^2 dy}{\int_{-a/2}^{a/2} |D(y)|^2 dy} = 0.44443 \quad (72)$$

At this point we realize that we have taken the problem of determining the minimum distance to the far field from the radiation pattern domain (in which our intuition works well) to the aperture domain where our intuition is less effective. However, we are interested in determining the types of aperture distributions which will lead to large minimum distances to the far field, and the current form of the MSD criterion suggests the possibility of a synthesis equation in some sense. Pursuing this, we can write the MSD equation as

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

Merging the integral signs and factoring out $D(y)$ we can write

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

Expanding the first squared magnitude term we simplify the expression to

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

In some sense we can interpret this expression as a synthesis equation for constructing aperture distributions which have a specified minimum distance to the far field. We specify the minimum distance to the far field by choosing a value for β , since β is inversely proportional to the distance at which we calculate the Fresnel pattern in order to determine the mean squared fit of radiation patterns.

The integrand is the product of two functions, and the expression tells us that the area under the composite function must be zero. The second function is a squared magnitude and so is positive definite. Thus the balancing of positive and negative contributions to the integral is controlled by the first function, and by the limits on the integral. Adjusting the limits on the integral corresponds to adjusting the aperture size in order to achieve a certain minimum distance to the far field. This is not what we had in mind in terms of a synthesis procedure - such a result is

self evident and of little use. Examining the first function we see that it is a function of β , and so is controlled by a choice of minimum distance to the far field. For relatively large distances from the aperture (relatively small values of β) this function is a quadratic curve with a negative offset of 0.44443. This is shown in fig. 17. As the value of β becomes small (as the distance from the aperture becomes large) 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. Keeping this in mind, we note that the synthesis equation can more properly be written as an inequality. In the general case, we may determine the value of β for which the synthesis integral evaluates to exactly zero. Since by definition we assume that we are in the far field at this distance, smaller values of β must also satisfy the MSD criterion. Since such smaller values cause the synthesis integral to become negative, we can replace "=" in the expression with "<". This results in a subtle change in the interpretation of the synthesis equation: values of β for which the integral is negative lie in the far field, and the minimum distance to the far field is the largest value of β which will satisfy the inequality.

With this more proper form for the synthesis expression and returning to our observation that the positive

Fig. 17 : Graph of the quadratic phase curvature across the aperture for finite ranges

contribution to the integral is pinched off at the edges of the aperture with decreasing β , we draw conclusions. Aperture weighting in the "negative" region of the integrand acts to move the minimum distance to the far field toward the aperture. Weighting in the positive region of the integrand acts to move the minimum distance away 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 is determined by the value of β . In order to synthesize an aperture with the largest possible minimum distance to the far field, we should collect the aperture weighting to the edges of the aperture, and give zero weighting to the remaining areas. In the limit, the aperture distribution with the largest possible minimum distance to the far field is one consisting of delta functions at the aperture edges; zero valued otherwise. The largest value of β which will satisfy the MSD criterion is that value for which the previously mentioned quadratic function is zero valued at the aperture edges (ie. due to the 0.44443 negative offset of the function, it will be negative over the entire aperture, though this region is given zero weight). It was found that for a 10 wavelength aperture the maximum value of β was 0.0271875, which corresponds to a maximum minimum distance to the far field of $1.155685D^2/\lambda$.

It is now useful to refer back to fig. 14. We know that the "delta" aperture has a radiation pattern corresponding to a sine function and has the narrowest main lobe of all aperture distributions considered so far. It would contribute a point on this graph at approximately 115 wavelengths, in agreement with the trend of the graph. Returning to the synthesis expression, we note that the positive definite part of the integrand is the squared magnitude of the aperture distribution. This means that the phasing of the aperture distribution is actually ignored by our MSD criterion! This tells us that the graph in fig. 14 indicates a trend, but does not continue past a minimum distance of about 115 wavelengths. According to our definition of the minimum distance to the far field, the minimum distance to the far field lies between zero distance from the aperture (delta function at the aperture centre) and $1.155685D^2/\lambda$ (for an aperture with delta functions at its edges). Strictly speaking, however, the MSD criterion cannot be used to predict minimum distances which lie closer than the region of validity of the Fresnel approximations.

The fact that the phase of the aperture distribution is ignored by the MSD criterion means that the original hypothesis that superdirective apertures have large minimum distances to the far field is wrong. Such apertures rely on phasing effects to achieve narrow beamwidths, and it has

phasing effects to achieve narrow beamwidths, and it has become apparent that while such phasing does alter the shape of the radiation pattern, in some strange way the mean squared difference between the finite distance radiation pattern and the far field pattern remains unaffected. This conclusion, that the hypothesis is disproved, has implicit to it a qualification which permeates all mathematical consideration of the Fresnel fields presented thus far. This qualification is that the distance at which the near field criterion is met should not extend to our predicted outer Fresnel boundary. Theoretically there does exist the possibility that the distance at which the near field criterion is met could outstrip geometrical or path difference Fresnel effects. We anticipate such a situation existing for at least the case of strongly superdirective antennas. The effect of such a near field boundary and "geometrical path difference effects" boundary inversion on the MSD criterion has not been considered. Thus our conclusion that the hypothesis has been disproved may only be justified for the case of marginally superdirective antennas. Also it must be emphasized that the conclusion that the hypothesis is wrong is only as good as the Mean Squared Difference criterion from which it was drawn. However, as discussed in a previous section, the MSD criterion is reasonable in the sense of indicating the validity of an approximation, and conven-

tional in the sense of using the normalized difference of squares. In retrospect, it is reasonable that the MSD criterion would ignore the phase of the aperture distribution, since phase information is destroyed by using squared magnitudes in the criterion. When information is destroyed, a degree of freedom is obtained. As an aside, it should be pointed out that of the three criteria for determining the minimum distance to the far field the current one has been referred to as the "phase sensitive" criterion, and now we have concluded that it is in fact "phase insensitive". This is not a contradiction in nomenclature since the current MSD criterion is sensitive to the phase of the radiation pattern, but insensitive to the phase of the aperture distribution.

5.2 Half Power Beamwidth and the Minimum Far Field Distance

Armed with this new form for the MSD criterion, we will now re-examine the previously presented relationship between half power beamwidth (HPBW) and the minimum distance to the far field. The new formulation is given by eqn. (75) expressed appropriately as an inequality, and is repeated below for convenience:

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

As stated in the previous section, the first factor of the integrand is an offset quadratic curve, and the second factor is positive definite. Thus the first factor controls those areas of the aperture distribution which will contribute positively or negatively to the integral. Further, the shape of the quadratic curve is related to the distance of the observer from the aperture. The larger the distance one considers from the aperture, the flatter the quadratic curve becomes. This fact is focal in explaining the observed relationship between HPBW and minimum distance to the far field. The flattening of the quadratic curve is reflected in the plot of the trajectories of its zero crossings as is varied, shown in fig. 19.

In order to carry out the investigation, we will consider a family of continuously varying apertures which are considered typical of the aperture distributions already used to expose the relationship. These distributions are presented in fig. 18. We begin by considering the set of distributions which are weighted at the aperture edges. At the distance at which the inequality of eqn. (76) is just met, the zero crossings of the quadratic curve must occur somewhere in the nonzero regions of the distribution. This is because positive and negative areas of the integrand must cancel. If "b" is small, then the quadratic curve will be roughly linear over the nonzero region of the aperture, and

**Fig 18 : A family of continuously
varying aperture distributions**

Fig. 19 : Trajectories of the zero crossings of the quadratic term in the MSD criterion

Fig. 20 : Comparison of correlations between HPBW
and minimum distance to the far field

we may then approximate the required position of the zero crossings of the quadratic curve as the midpoints of the nonzero regions. We now use this assumed position of the zero crossings to determine a value of β which corresponds to a distance from the aperture. This is done by using the graph shown in fig. 19. Recalling that β is derived from the distance from the aperture, we can compare graphs of HPBW vs. minimum distance to the far field for our set of general aperture distributions and for our specific family of continuously varying distributions. This graph is presented in fig. 20. Although the curves do not overlap, it is apparent that they both have the same shape. This supports our insight that the relationship between HPBW and minimum distance to the far field hinges on the mathematical fact that positive and negative areas of the MSD integrand must cancel. We can see that the relationship between HPBW and distance from the aperture is similar in both cases, though one curve was derived from more basic, qualitative consideration of the behavior of the MSD criterion. What this is telling us is that as aperture weight is moved to the edges, the boundaries separating positive and negative areas of the integrand must move toward the aperture edges. This corresponds to moving the zero crossings of the quadratic curve towards the aperture edges, and so flattening the quadratic curve, and so increasing the minimum distance

from the aperture. We thus see that the more the distribution is weighted to the edges of the aperture the farther will be the minimum distance to the far field, and conversely distributions tapered at the edges exhibit relatively small minimum distances to the far field.

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 has become apparent from this fact and the above discussion that the unexpected good correlation between HPBW and minimum distance to the far field arises from the not so unexpected correlation between HPBW and edge taper. This insight allows us to conclude that the correlation between HPBW and minimum distance to the far field is a misleadingly good one, and is qualitative in the sense that HPBW can be qualitatively related to edge taper. Investigation of our "family of continuously varying apertures" has made it apparent that including other arbitrary aperture distributions in the graph would only serve to increase the scatter about a "best" curve, and in this sense the relationship is "missleadingly" good (ie. the full extent of the scatter is not yet apparent).

Chapter 6 POLARIZATION EFFECTS AND A DIRECTIONAL CRITERION

6.1 The Infinitesimal Current Patch

Up to this point, the vector nature of the electromagnetic field has been ignored in both qualitative and quantitative discussions of the minimum distance to the far field. Our rigorous definition of this minimum distance expresses mathematically the notion that the radiation pattern of an aperture changes with increasing distance through the Fresnel region, and converges to the far field pattern. For the sake of convention and mathematical consistency with the literature, we index the extent of this convergence (with distance) by using a mean squared difference expression. 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 would cause this finite distance pattern and the far field pattern to be quite different. A natural question to ask, then, concerns the relationship between scalar and vector field expressions. How are they related? Is the scalar field proportional to some component

of the vector field, or perhaps is it the magnitude of the vector field? Another approach to dealing with the issue of polarization concerns the classical case of calculating the vector fields around an infinitesimal current element. It is well known that this example yields a radial component of the electric field which is significant at small distances from 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 our discussions of the minimum distance to the far field.

We will begin by presenting the analysis of the vector fields around an infinitesimal current element. This analysis may be found in Balanis [19, 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 by

$$\frac{e^{j\omega t\mu_0}}{4\pi} \int_V \frac{J(x', y', z')}{r} e^{-jkr} dV' = \bar{A} \quad (77)$$

where

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

and primed (') coordinates are source points, unprimed coordinates are field points. For reasons which will become apparent in a further section, the "infinitesimal current

element" will be assigned an infinitesimal width dy , so that it will in fact represent an infinitesimal patch of linearly directed current. As will be seen, this adjustment of the model will have only a trivial effect on the form of the field equations, but allows us to draw generalizations about planar apertures. The geometry of the "infinitesimal current patch" is presented in fig. 21. Because the element has only dz length, eqn. (77) reduces to

$$\bar{A} = \frac{\mu I dy dz}{4\pi r} e^{-jkr} e^{j\omega t} \hat{a}_z \quad (79)$$

Moving to spherical coordinates, we note that

$$\hat{a}_z = \hat{a}_r \cos\theta - \hat{a}_\theta \sin\theta \quad (80)$$

Having formed the vector potential of the fields, the magnetic field intensity can be found from the relation

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

and is given by

$$H_\phi = \frac{I dz dy}{4\pi} \left\{ \frac{jk}{r} + \frac{1}{r^2} \right\} \sin\theta e^{j(\omega t - kr)} \quad (82)$$

We can see that it only has a ϕ component, and has a $\sin\theta$ dependance. From the magnetic field intensity we can find the electric field intensity, since

$$\bar{E} = \frac{1}{j\omega\epsilon} \nabla \times \bar{H} \quad (83)$$

2 2
OF / DE

Fig. 21 : Geometry for the infinitesimal
current patch

The electric field is given by

$$E_r = \frac{-j\omega I dz dy}{2\pi k \sqrt{\mu E}} \left\{ \frac{jk}{r^2} + \frac{1}{r^3} \right\} \cos\theta e^{j(\omega t - kr)} \quad (84)$$

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

The electric field has a radial component, which dies away with the square of distance, and has a $\cos\theta$ dependence. This contrasts the remaining E and H fields, which have $\sin\theta$ dependence. The angular variation of these fields is shown in fig. 22.

At this point we would like to examine the polarization as a function of r and θ . As a first pass, we examine the radial dependence of E_θ , repeated below for convenience:

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

We can see that if we are, say, more than 1 wavelength away from the source, the largest term is negative (since $k > 1$) so the field points in the positive direction. The radial component of the electric field has a $\cos\theta$ dependence, which allows us to present the diagram in fig. 23 showing the electric field at four principal positions around our source. This simple diagram shows us that the radial component of the electric field arises from the component of

Fig. 22 : Diagram showing the angular variation of fields around the infinitesimal current patch

Fig. 23 : Polarization of the total $\mathbf{E}$ vector at the four principal angles

the E field which gives rise to the current flow in our source. Examining the E field equations further suggests that the field vector may do strange things at angles between those four presented in fig. 23, so we turn to the problem of determining the angle of the electric field as a function of r and θ .

Since the actual field is complex, we will consider the real and imaginary parts separately for now. Considering r larger than 1 wavelength so that the real part of E has a fixed direction, we can draw a vector diagram of the addition of the two E field components as shown in fig. 24. From this diagram we can write

$$\phi = \theta - \tan^{-1} \left\{ \frac{E_r}{E_\theta} \right\} = \tan^{-1} \left\{ \frac{\frac{-1}{r^3} \cos \theta}{\frac{1}{2} \left(\frac{k^2}{r} - \frac{1}{r^3} \right) \sin \theta} \right\} \quad (87)$$

and at $\theta = \pi/4$ we can write

$$\phi = \tan^{-1} \left\{ \frac{\frac{-1}{r^3}}{\frac{1}{2} \left(\frac{k^2}{r} - \frac{1}{r^3} \right)} \right\} \quad (88)$$

Performing a similar analysis for the imaginary component of the total E field, and again using the geometry of vector addition shown in fig. 24, we can write

$$\phi = \theta - \tan^{-1} \left\{ \frac{\frac{-k}{r^2} \cos \theta}{\frac{-k}{2r^2} \sin \theta} \right\} = \theta - \tan^{-1} \left\{ \frac{\cos \theta}{2 \sin \theta} \right\} \quad (89)$$

Fig. 24 : Geometry for determining the angle of the electric field vector

Note that the angle of the total imaginary E vector is not a function of distance because both radial and angular E components fall off as the square of distance. The variation of the real and imaginary total E vectors with radial distance and angle around the source is presented in figs. 25 and 26.

These two diagrams, and especially that of the total real E field, show how the total E vector "tumbles" away from the source, twisting with distance. It was the possibility of just such an anomalous effect which prompted this vein of investigation. The question remains as to whether this polarization variation will invalidate our criterion for determining the minimum distance to the far field, which is based solely on scalar considerations.

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 distances from the source, these components have died away (as have the imaginary E components) and we are left with only a θ component. We are concerned that the field might still be "twisting" with radial distance at ranges predicted for our minimum far field. It is natural, then, to now specify a distance at which this radial component may be ignored. Because we are dealing with a vector addition (of radial and E_θ components) and

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

because the radial E component can be thought of as a perturbation of the direction of the E_θ component, we take the radial E component to be negligible when it is $1/10$ the E_θ component. That is to say, the perturbation of a vector by the addition of another vector $1/10$ its size is deemed negligible. However, to be conservative we halve this criterion and require that the radial component be not more than $1/20$ the E_θ component. Applying this criterion along the $\theta = \pi/4$ contour where the coefficients of the radial and θ components are equal ($\sin \theta = \cos \theta$), we specify that the radial component may be ignored at distances of more than 20 wavelengths. Thus we will conservatively say that, in a polarization sense, we are in the far field of the source at a distance of 20 wavelengths. This will be more conservative as one moves toward boresight and less conservative as one moves toward $\theta = 0$, due to the $\cos \theta$ variation of the radial component. To quantify our 20 wavelength criterion somewhat, the direction of the total real E vector is tabulated below at $\theta = \pi/4$ for some selected distances:

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

Our model source throughout the above investigation has been an infinitesimal current "patch" rather than a current element. Thanks to superposition, this means that our analysis can be extended to draw generalizations about planar apertures. We can theoretically build up any planar aperture and the resulting fields by a superposition of many such properly weighted current patches. However, because such superposition involves the vector addition of fields, our arguments concerning the significance of the radial component of the fields still applies. We conclude that the polarization of the fields from planar apertures can be considered constant (radially) beyond a distance of 20 wavelengths. At distances larger than this, we need only concern ourselves with antenna geometry so far as polarization effects (as a function of r) are concerned. That is to say, "point source" polarization effects may be ignored, though polarization effects due to actual antenna size deserve consideration. This arises from the fact that at a finite distance from the antenna, a θ field component viewed from one point on the aperture will have a radial component when viewed from another point on the aperture. These geometrical considerations will, in a following section, also be found negligible by following to the next logical step of the analysis: there still remains the issue of

reconciling the current patch analysis and the subsequent 20 wavelength criterion with the scalar method of calculating the fields.

We wish to understand the nature of the simplifications which were adopted that allowed us to ignore the vector nature of the fields and calculate them as scalar fields. To do this, we refer to Silver [20, pp. 164, 165]. Here he points out that 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 as indicated below :

$$\begin{aligned} & \frac{1}{4\pi} \int_A \{ -j\omega\mu(\hat{n} \times \vec{H})\Psi + (\hat{n} \times \vec{E}) \times \vec{\nabla}\Psi + (\hat{n} \cdot \vec{E})\vec{\nabla}\Psi \} dS \\ &= -\frac{1}{4\pi} \int_A \left\{ \Psi \frac{\partial \vec{E}}{\partial n} - \vec{E} \frac{\partial \Psi}{\partial n} \right\} dS + \frac{1}{4\pi} \oint_A \Psi \vec{E} \times \vec{\tau} dS \end{aligned} \quad (90)$$

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

$$E(p) = \frac{-1}{4\pi} \int_A \left\{ \Psi \frac{\partial \bar{E}}{\partial n} - \bar{E} \frac{\partial \Psi}{\partial n} \right\} dS + \frac{1}{4\pi} \int_A \Psi \bar{E} \times \tau dS$$

$$- \frac{1}{4\pi j \omega \epsilon} \int_A \nabla \Psi (\tau \cdot \bar{H}) dS \quad (92)$$

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:

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

where u is the proper component for the linearly polarized electric field. However, Silver [20, p. 171] points out that the scalar expression for the fields around the aperture can only yield qualitative results very close to the aperture because line integrals (of eqn. (92)) which were ignored make a significant contribution. We attribute the twisting of the field vectors in this close region (as observed for the case of an infinitesimal current patch) to the vector contributions of these line integrals. Silver states that the line integrals will be significant

throughout the near zone (ie. that region comprised of points in the immediate neighborhood of the aperture).

This, then, reconciles our knowledge of the vector nature of the electromagnetic fields around the antenna with our scalar expressions for the fields and our case of an infinitesimal current patch. Summarizing, we have found by analyzing the current patch that only one component of the field dominates at large distances 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 distances the angle of our observation determines the polarization of the radiation; and so we can adopt a scalar representation of the fields with the direction of polarization "understood". Also, at closer distances, the polarization of the radiation changes or "twists" with distance due to the presence of a radial component of the field. By referring back to the approximations made to arrive at the scalar field expression, we see that line integrals that were ignored for larger distances become significant in closer regions, the regions in which we have observed polarization "twisting". In this region, then, our scalar representation of the fields is invalid. Quantitatively, we have determined that this polarization twisting does not affect or invalidate our MSD criterion.

6.3 A Modified Mean Squared Difference Criterion

At this point, we have presented the mathematical nature of superdirective antennas, postulated that they might imply a large minimum distance to the far field, and formulated a rigorous definition of the minimum distance to the far field. Having investigated this mean squared difference criterion, we found an interesting relationship between half power beamwidth and the minimum distance to the far field for various aperture distributions, which tended to support the hypothesis. However, a slight mathematical reworking of the criterion produced (in a loose sense) a synthesis equation and showed that, using this minimum distance criterion, the hypothesis was wrong. The next natural step is to take a closer look at the minimum distance (mean squared difference) criterion. As pointed out in a previous section, the MSD criterion is reasonable in that it indexes the quality of a far field approximation at a finite distance. It does this by comparing the fit of the actual finite distance complex radiation pattern to the far field complex radiation pattern. Further, the MSD criterion is conventional in the sense that a mean or normalized squared difference scheme is employed to compare the complex radiation patterns.

Although the formulation of the MSD criterion is rea-

sonable, it is important to notice that it is applied over the entire radiation pattern, including the invisible region. This was done to ensure mathematical consistency, since any alternative would be to favour some part of the radiation pattern - a choice which could not be justified mathematically. However, the fact that one cannot simultaneously measure the field at all points around the aperture now gives us some motivation to investigate the possibility of applying the MSD criterion over only a specific sector. We interpret the results of such a restricted application of the MSD criterion as saying that we must move a specified distance away from the antenna in a certain direction in order to measure the far field relative field strengths over the sector. It is expected that this required minimum distance will 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 following analysis was first presented by Fred Symons of the Defence Research Establishment Ottawa, in Ottawa, Ont.)

In attempting to formulate a MSD criterion expression for a single sector of the radiation pattern, it is useful

to first recall the original form of the criterion as applied over the entire radiation pattern. It is repeated below, as an appropriate inequality, for convenience :

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

which we can write as

$$\frac{\int_{-\infty}^{\infty} |H(ky)|^2 dk}{\int_{-\infty}^{\infty} |G(ky)|^2 dk} \leq K \quad (95)$$

We may effectively apply this criterion over only a sector of the radiation pattern by windowing the radiation pattern as indicated below:

$$\frac{\int_{-\infty}^{\infty} |W(ky-k) H(ky)|^2 dk}{\int_{-\infty}^{\infty} |W(ky-k) G(ky)|^2 dk} \leq K \quad (96)$$

We can use Parseval's Theorem to transform this expression from the radiation pattern domain to the aperture domain, and multiplications become convolutions.

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

$R(y)$ is the inverse transform of the window function, and the exponentials arise from the shifting property of the

Fourier Transform. Eqn. (97) can be written as products of complex conjugates:

$$\int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} e^{jk_x x} R(y-x) (e^{-j\beta x^2} - 1) D(x) dx \int_{-\infty}^{\infty} e^{-jk_z z} R^*(y-z) (e^{j\beta z^2} - 1) D^*(z) dz \right. \\ \left. - K \int_{-\infty}^{\infty} e^{jk_x x} R(y-x) D(x) dx \int_{-\infty}^{\infty} e^{-jk_z z} R^*(y-z) D^*(z) dz \right\} < 0 \quad (98)$$

Now we can see that the only terms which are functions of y are the "window" functions, and so we can change the order of integration. Collecting like terms, we arrive at the form

$$\int_{-a/2}^{a/2} \int_{-a/2}^{a/2} \{ (e^{-j\beta x^2} - 1) (e^{j\beta z^2} - 1) - K \} e^{jk_x x} e^{-jk_z z} D(x) D^*(z) \\ \cdot \int_{-\infty}^{\infty} R(y-x) R^*(y-z) dx dy dz < 0 \quad (99)$$

By examining the form of eqn (99) we see a parallel with the synthesis form of the MSD criterion presented in eqn. (91). Eqn. (91) allowed us to identify regions of the (line of the) 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 minimum distance to the far field. Eqn. (99), however, presents us not with a line but a plane over which we can define the weighting function of the aperture distri-

bution, which determines regions of positive and negative area contribution. The plane is that defined by the x and z variables. One can see that the plane will collapse to a line along the diagonal for the case of applying the MSD criterion over the entire radiation pattern. This is because the window function becomes a constant function thus $R(y)$ becomes a delta function. The integration contour lies along $x = z$.

Further simplifying assumptions could be introduced, such as specifying the window in the radiation pattern domain as a $\sin(x)/x$ function. This would mean that our $R(y)$ function is a rectangle function, allowing us to evaluate our integral with respect to y and express the result in a closed form. However, the remaining integrand is still relatively complicated and so does not lead us to a simple, graphically clear weighting function which we should apply to the aperture distribution to ensure a large minimum distance to the far field. It will be recalled that the integrand for the "omnidirectional" MSD criterion (eqn. (91)) did lead us to a simple weighting function that told us the type of distribution which would exhibit the largest minimum distance to the far field. Whereas that MSD criterion was found to be insensitive to the phasing of the aperture, we can see that the directional MSD criterion is definitely a function of the phase of the distribution.

Thus we do not expect the directional criterion to yield as immediately clear and conclusive results as were found for the omnidirectional case.

Yet another simplifying assumption which could be introduced into the directional MSD criterion is to let the window function in the radiation pattern domain become a delta function. This corresponds to observation not over a sector, but from a single direction. The MSD criterion for this case may be simplified to

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

Fig. 27 shows a plot of this directional MSD criterion over most of the visible region for the case of a uniform aperture distribution. We can see that there is a large variation of the minimum distance to the far field, as much as an order of magnitude. Most striking is that the minimum distance 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 distance to converge. The main beam can be seen to converge quite quickly. This reflects something which becomes apparent by comparing plots of radiation patterns as one moves outward through the Fresnel region. The main beam and higher order sidelobes form quite quickly, whereas the first sidelobe starts as a "shoulder" of the main lobe and

Fig. 27 : Convergence of the uniform aperture radiation pattern

gradually separates with distance as the first null increases in depth. This fact is hidden by the omnidirectional MSD criterion since it comments only on the total mean squared difference between the finite distance and far field patterns - it does not tell 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 MSD criterion would predict a minimum distance biased toward a small distance. Although the graph of fig. 27 does not extend into the invisible region, it is reasonable to assume that the trend of the graph will continue, and that all higher order sidelobes form within a relatively small distance from the antenna.

Finally, fig. 27 gives us another insight into the convergence of the radiation pattern through the Fresnel region. We note that the minimum distance to the far field rises sharply near the nulls of the radiation pattern. We attribute this to the fact that the denominator of the MSD criterion has become zero. The spikes may arise because the nulls in the far field do not exactly sit on the nulls of the finite distance pattern. That is, it appears that the nulls are shifting slightly as we move outward through the Fresnel region. From the asymptotes shown in fig. 27, we conclude that the nulls of the finite distance pattern move

away from boresight (toward their far field positions) as we move away from the antenna. The hypothesis that the nulls are adjusting themselves through the Fresnel region has not been tested, but is a reasonable one in the light of fig. 27.

The behavior of the radiation pattern near nulls suggests that the original hypothesis, that a superdirective antenna requires a large minimum distance to the far field, may be correct in some sense when we consider the directional MSD criterion. Fig. 27 indicates that the shape of the radiation pattern near a zero (or near a feature of the radiation pattern having a large slope) may take a large distance to settle. It is possible that antennas having a very narrow main beam may require a large distance for their main beam to form correctly. If this is so, then in this sense the original hypothesis could be considered correct.

6.4 Applications Notes

The most natural application of the work and conclusions presented in this thesis concern range distance requirements in antenna pattern measurement. The conventional and generally conservative $2D^2/\lambda$ minimum distance to the far field implies unrealistically large test ranges for antennas of only moderate size. 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 dust, rain, terrain irregularities) make such a range difficult to realize and degrade the reliability of measurements when such a range is available. Clearly in such a case one should use the smallest range which would give correct results. The notion of determining the quality of the $2D^2/\lambda$ approximation is not a new one. R. C. Hansen [25] notes that the $2D^2/\lambda$ rule is acceptable for antennas with moderate sidelobe levels (-25 dB), and that in such cases directivity errors are of the order of 0.1dB 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 takes a large distance to converge to its far field form. This conclusion is also presented by Hacker and Schrank [27] by an analysis of several typical low and ultralow sidelobe radiation patterns. They graphically show that the higher order sidelobes form quite quickly. The first sidelobe, however, begins as a "shoulder" on the main lobe, and separates off with increasing distance as the first null forms.

Fig. 27 is qualitatively and quantitatively in agreement with the Hansen paper and the Hacker and Schrank paper. Fig. 27 shows that the first and to a lesser degree the

second sidelobes take a large distance 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 distances, found in the Hacker and Schrank paper. The same criterion is used in both papers to determine the proper range for measuring the first sidelobe accurately. The error in the height of the first sidelobe (in dB) is plotted as a function of distance, and it is shown that the lower the sidelobes the greater the error in the first sidelobe at a given distance. This criterion determines the required test range by examining only one point in the radiation pattern. The directional MSD criterion in effect extends the criterion (used by both papers) over the entire radiation pattern.

The statement that the $2D^2/\lambda$ range is acceptable for patterns with high sidelobe levels seems to contradict the conclusions of this thesis. It was found, using the MSD criterion, that exactly the opposite was true - that high sidelobe patterns required the largest test ranges, whereas cosine and edge tapered distributions converged quite quickly. This apparent discrepancy can be explained by closely examining the form of the (omnidirectional) MSD criterion, and exactly what it is indexing. Although the sidelobes of low sidelobe patterns take a long distance 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 quite quickly, and due to this concentration of energy the errors in the tiny sidelobes go unnoticed by the omnidirectional MSD criterion. Thus the predicted minimum distance to the far field is biased toward the distance 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 fit over the sidelobes have a larger effect on the MSD criterion. The net effect is that if one wishes to have a good fit over most of the radiation pattern, high sidelobe patterns require larger test ranges than low sidelobe patterns. The Hacker and Schrank paper (and the Hansen paper) are merely saying that if one wishes to accurately measure the far field form of the first sidelobe of low and ultralow sidelobe patterns, then large test ranges must be used. In fact, the Hacker and Schrank paper concludes that if measurements are made near $2D^2/\lambda$, only the first one or two sidelobes are "lost", with the remainder of the pattern having the far field form.

Chapter 7 CONCLUSIONS AND FUTURE WORK

The purpose of this thesis was to investigate the extent of the Fresnel region for the case of aperture antennas, and specifically to test the hypothesis that superdirective antennas imply a large minimum distance to the far field. The first part of the work centred on investigating the nature of superdirective antennas, and is comprised largely of a literature search. The second part of the work involved specifying a rigorous definition of the minimum distance to the far field and analyzing the implications of this criterion.

It was found that the idea behind superdirective antennas hinges on the fact 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. By summing orthogonal, aperture limited radiation patterns, it is possible to constrain the pattern over the visible region to have any desired shape, thus we can specify a superdirective antenna. Since these orthogonal patterns must be translated to specific positions in order to add appropriately, aperture phasing is vital to the

superdirective aperture. However, as one specifies radiation patterns with directivity even slightly greater than the uniform case, the pattern in the "invisible" region (complex angles of propagation) becomes large. The literature explained that the large patterns in the invisible region were indicative of an unrealizable antenna due to large amounts of reactive energy and large fields required in the aperture plane. It was shown that superdirective arrays are generally unrealizable because of large currents required in the elements, and because of critically close tolerances on element spacing, phasing, and amplitude excitation.

Having examined the nature of superdirective antennas, a rigorous definition of the minimum distance to the far field was presented. This criterion was based on the normalized mean squared fit between the (complex) radiation pattern at a finite distance and the far field pattern. A mathematical reformulation of this criterion showed that, in the context of the criterion, the original hypothesis was disproved (at least for the case of marginally superdirective antennas) if one is concerned with the convergence of the entire radiation pattern to its far field form. We cannot safely extend the notion that the hypothesis was disproved to the case of strongly superdirective antennas because there exists the possibility that the distance at

which the reactive near field criterion is met would be larger than our predicted outer Fresnel boundary. Nevertheless, aperture phasing is important to the superdirective antenna, but here we find that phasing is ignored by the mean squared difference criterion. An investigation of radiation polarization changes with distance showed that while field vectors twist with distance near the antenna, this twisting is a near field effect and does not extend through the Fresnel region, so polarization effects could also be neglected. The reformulated criterion resulted in a synthesis equation allowing one to specify a desired minimum distance to the far field and then select an appropriate aperture excitation.

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 failed to produce a "clear" synthesis equation analogous to the omnidirectional case, though it was found that aperture phasing was important. A special case of the directional criterion was applied to a uniform aperture, and it was found that the radiation pattern around nulls requires a large distance to converge to the far field form. Also, while the main beam forms quite quickly, the first and second sidelobes require the largest distance to converge. Higher order sidelobes, like the main beam, converge

relatively quickly. These (directional) findings are in agreement with the literature.

Applications of this work to the determination of test range distance requirements is recommended.

Further work in this area could centre on a more thorough investigation of the directional 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 is extended 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 looks promising.

Regarding superdirective antennas, it is possible that the more rapidly changing features across the radiation pattern would take large distances to converge to their far field form, and in this sense perhaps the original hypothesis could be considered correct. This is suggested

by the convergence of the uniform aperture near nulls, since it is in this region that the slope of the radiation pattern is relatively large. Further investigation is necessary using a superdirective aperture larger than two wavelengths, since this is the minimum size for which the given Fresnel approximations are valid.

Lastly, the hypothesis that the nulls trace nonradial trajectories through the Fresnel region should be tested. This could be checked most simply by turning to the uniform case and determining the exact positions of the nulls at distances suggested by fig. 27. Although the nulls may translate only slightly in the k_y (Fourier Transform) variable, this may result in significant translation in angle. This is because $\sin\theta$ is a nonlinear transformation and so angular measure is compressed for values of k_y approaching $+k$ or $-k$.

BIBLIOGRAPHY

1. H. G. Booker, P. C. Clemmow, "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, pp. 11 - 17; January, 1950
2. P. M. Woodward, J. D. Lawson, "The Theoretical Precision With Which an Arbitrary Radiation-Pattern May Be Obtained From a Source of Finite Size", JIEE, vol. 95, pp. 363 - 370; 1948
3. S. A. Schelkunoff, "A Mathematical Theory of Linear Arrays", Bell Systems Technical Journal, vol. 22, pp. 88 - 107; January, 1943
4. H. J. Riblet, "Note on the Maximum Directivity of an Antenna", Proc. I.R.E., vol. 36, pp. 620 - 624; May, 1948
5. R. M. Wilmotte, "Note on Practical Limitations in the Directivity of Antennas", Proc. I.R.E., vol. 36, p. 878; July, 1948

6. T. T. Taylor, "A Discussion of the Maximum Directivity of an Antenna", Proc. I.R.E., vol. 36, p. 1135; 1948
7. Nicholas Yaru, "A Note on Super-Gain Antenna Arrays", Proc. I.R.E., vol. 39, pp. 1081 - 1085; Sept., 1951
8. T. T. Taylor, "Design of Line-Source Antennas for Narrow Beamwidth and Low Sidelobes", I.R.E. Trans. on Ant. and Prop., vol. 3, pp. 16 - 28; January, 1955
9. R. E. Colin, S. Rothschild, "Reactive Energy in Aperture Fields and Aperture Q", Canadian Journal of Physics, vol. 41, pp. 1967 - 1979, December, 1963
10. D. R. Rhodes, "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. D. R. Rhodes, "The Optimum Line Source for the Best Mean-Squared Approximation to a Given Radiation Pattern", IEEE Trans. on Ant. and Prop., pp. 440 - 446; July, 1963
13. D. R. Rhodes, "On a Fundamental Principle in the Theory of Planar Antennas", Proc. IEEE, vol. 52, pp. 1013 - 1021; September, 1964

14. Clemmow, P. C., The Plane Wave Spectrum Representation of Electromagnetic Fields, Pergamon Press Ltd., New York, 1966
15. Stratton, J. A., Electromagnetic Theory, McGraw-Hill Book Co. Inc., New York, 1941, pp. 497 - 500
16. Jull, E. V., Aperture Antennas and Diffraction Theory, Peter Peregrinus Ltd., 1981, pp. 42 - 46
17. G. Borgiotti, "Fourier Transforms Method of Aperture Antennas Problems", (in english) Alta Frequenza, vol. 32, pp. 196 - 204; November, 1963
18. P. M. Woodward, "A Method of Calculating the Field Over a Plane Aperture Required to Produce a Given Polar Diagram", Journal IEE, part IIIA, pp. 1554 - 1558; 1947
19. Balanis, C. A., Antenna Theory - Analysis and Design, Harper and Row, Publishers, New York, 1982, Ch. 1 - 3, Ch. 11 pp. 446 - 486
20. Silver, Samuel, Microwave Antenna Theory and Design, McGraw-Hill Book Co. Inc., New York, 1949
21. Slater, John C. and Frank, Nathaniel H., Introduction to Theoretical Physics, McGraw-Hill Book Co. Inc., New York, 1933

22. Harrington, Roger F., Time-Harmonic Electromagnetic Fields, McGraw-Hill Book Co. Inc., New York, 1961
23. Heald, Mark A. and Marion, Jerry B., Classical Electromagnetic Radiation, Academic Press, New York, 1980
24. Papas, Charles H., Theory of Electromagnetic Wave Propagation, McGraw-Hill Book Co. Inc., New York, 1965
25. R. C. Hansen, "Measurement Distance Effects on Low and Ultralow Sidelobe Patterns", IEEE Trans. on Ant. and Prop., vol. AP-32, no. 6, pp. 591 - 593; June, 1984
26. D. R. Rhodes, "On Minimum Range for Radiation Patterns",
Proc. I.R.E., pp. 1408 - 1410; September, 1954
27. Philip S. Hacker and Helmut E. Schrank, "Range Distance Requirements for Measuring Low and Ultralow Sidelobe Antenna Patterns", IEEE Trans. on Ant. and Prop., vol. AP-30, no. 5, pp. 956 - 965; September, 1982
28. Schelkunoff, S. A., Antennas: Theory and Practice, John Wiley and Sons Inc., New York, 1952, pp. 195 - 198

APPENDIX A

Discussion of the Invisible Region

Over the visible and invisible regions we have a $1/r$ spatial variation in the radiation intensity pattern (see eqn. (49)). In comparing the radiation patterns over the visible region, we multiply by r to normalize the patterns. However, as shown in Chapter 2, Section 2.1, we have an additional spatial variation of the fields in the invisible region, of the form

$$e^{-k(\sin^2\theta - 1)^{\frac{1}{2}}r} \quad (A-1)$$

Reasonably, in the invisible region when we take the mean squared difference between two patterns we must normalize for this variation too. The question arises as to what if any special considerations should be made. For instance, although we must multiply by

$$e^{k(\sin^2\theta - 1)^{\frac{1}{2}}r} \quad (A-2)$$

it is not entirely clear the region over which the multiplication should be done, of what function should be multiplied.

We know that only the fields in the invisible region have this spatial variation, so to include the visible region in the range of our new normalizing factor would be to introduce a spatial variation rather than to remove one. Rather, multiplying only the invisible region by this factor will remove the unwanted variation in the invisible region. This means that the Fourier Transform of the aperture distribution does indeed give the radiation pattern independent of spatial variation in both the visible and invisible regions. Thus we conclude that, in using the mean squared difference criteria applied over the entire radiation pattern, it is correct to use the Fourier Transform of the aperture distribution as the space-independent radiation pattern.

APPENDIX B

Definition of Aperture Distributions

This appendix specifies the 16 aperture distributions used to investigate the minimum far field criteria.

1. uniform

2. double cosine

3. triangle

4. cosine

5. edge

6. complement of double cosine

7. notch

8. filled edge

9. $1 + \text{triangle}$ 10. $\text{edge} + \text{complement of double cosine}$

11. double triangle

12. complement of double triangle

13. knife edge

14. cosine squared

15. triangle squared

16. (triangle)**4

APPENDIX C

Numerical Example of a Superdirective Antenna

To obtain some feel for the physical limitations inherent in the design of superdirective antennas, we refer to a paper by Nicholas Yaru [7], in which he presents a numerical example of the design of a superdirective antenna. We will briefly parallel his development of the problem, and cite some his numerical results.

The radiation pattern of a linear array with element spacing d can be written in the following way, using the Fourier Transform relationship:

$$f(\sin\theta) = \int_{-\infty}^{\infty} \sum_{n=0}^{N-1} i_n \delta(y - nd) e^{jyky} dy \quad (C-1)$$

and this can then be written as

$$f(\sin\theta) = \sum_{n=0}^{N-1} i_n e^{jndky} \quad (C-2)$$

now if we make the substitution

$$z = e^{jkyd} \quad (C-3)$$

then we can write the radiation pattern as a polynomial in z , as shown below:

$$f(z) = i_0 z^0 + i_1 z^1 + i_2 z^2 + \dots + i_{N-1} z^{N-1} \quad (C-4)$$

It is important to note that the coefficients of the polynomial in eqn. (C-4) are also the relative amplitudes of the array elements. This polynomial will have $(N-1)$ real or complex conjugate zeros, and can be written in the form

$$f(z) = i_0 \prod_{i=1}^{N-1} (z - z_i) \quad (C-5)$$

So we see that the zeros of our polynomial determine the nulls in the radiation pattern of the array. From array design theory, we know that in order for the radiation pattern to go all the way to zero at a given root, the zero must lie on the unit circle in the z -plane. By considering an arbitrary pair of complex conjugate zeros and the polynomial from which they arose, we can write

$$P(z) = z^2 + bz + c \quad (C-6)$$

which has roots at

$$z = \frac{-b \pm j\sqrt{4c - b^2}}{2} \quad (C-7)$$

These roots have magnitude given by

$$|Z| = \sqrt{\frac{4c - b^2 - b^2}{4}} \quad (C-8)$$

and have magnitude 1 only if $c = 1$ (i.e. only if the polynomial is symmetrical). Since the product of symmetrical polynomials is also a symmetrical polynomial, we know that the array elements giving rise to a radiation pattern with zeros at its nulls are symmetrically excited in amplitude.

The visible region is defined by $-k < ky < +k$, that is, for values of Z given by

$$\begin{aligned} Z(ky = +k) &= e^{-jkd} \\ Z(ky = -k) &= e^{jkd} \end{aligned} \quad (C-9)$$

and the range of the visible region in argument around the unit circle is

$$\arg(Z(k)) - \arg(Z(-k)) = 2kd \quad (C-10)$$

This is the range on the unit circle over which we distribute our zeros to specify the far field radiation pattern. Yaru points out that, for a broadside array, if the nulls in this range on the unit circle are spaced according to a "Tchebyscheff distribution" a supergain pattern will result (element spacing d must be less than $\lambda/2$).

The order of the Tchebyscheff polynomial to be used is determined by the equation

$$2n + 1 = N \quad (C-11)$$

where n is the order of the polynomial and N is the number of elements in the array.

Yaru considers the case of a nine element linear broadside array with element spacing of $d = \lambda / 4$. He notes that the relative field strength of the radiation pattern is given by the real part of eqn. (C-4), or

$$|E| = i_0 - i_1 \cos(kyd) + i_2 \cos(2kyd) + \dots + i_{N-1} \cos((N-1)kyd) \quad (C-12)$$

With the substitution $x = \cos(kyd)$ eqn. (C-12) yields

$$|E| = i_0 - i_1 x - i_2 (2x^2 - 1) - \dots - i_{N-1} T_{N-1}(x) \quad (C-13)$$

where T_{N-1} is the $(N-1)$ th Tchebyscheff polynomial. The coefficients of this polynomial are determined so that the polynomial will have the shape of the Tchebyscheff polynomial T_4 from -1 to x_0 and back to -1 again. x_0 is some value such that $T_4(x_0)$ is the desired ratio of side lobe to main lobe. Next, $T_4(x)$ must be shifted into half the array's visible region, given by

$$x = \cos(\pm kd) = \cos(\pi / 16) \quad (C-14)$$

$$x = \cos(ky=0) = 1$$

using the linear transformation $x' = ax + b$. Since $T_4(x)$ is given by

$$T_4(x) = 8x^4 - 8x^2 + 1 \quad (C-15)$$

we can see that

$$8(x')^4 = 8a^4x^4 - 32a^3x^3b - 48a^2x^2b^2 - 32axb^3 - 8b^4 \quad (C-16)$$

$$8(x')^2 = 8a^2x^2 - 16axb - 8b^2 \quad (C-17)$$

At this point we can see that, with powers of 4 in these expressions, we must determine a and b to a very great accuracy in order to properly specify the coefficients of our array. A small error in either a or b will give grossly wrong values for our relative amplitude excitations. Yaru has determined that, for the relative amplitude excitation values to be accurate to only 1 decimal place, a and b for the example cited must be specified to an accuracy of 10^{-9} .

For the nine element array in the example, Yaru determines the relative current excitations to be

$$\begin{aligned} i_0 &= 8,893,659,368.7 \\ i_1 &= -14,253,059,703.2 \\ i_2 &= 7,161,483,126.6 \\ i_3 &= -2,062,922,999.4 \\ i_4 &= 260,840,226.8 \end{aligned}$$

These coefficients determine the relative excitations of all

the elements of the array because of the symmetry property mentioned above for zero valued nulls in a broadside array. The E field in the direction of maximum radiation (broadside) is given simply by the sum of the coefficients, and evaluates to 19.5. We have specified currents in the elements in the order of tens of millions of amperes to produce a maximum field intensity that could be achieved by 19.5 milliamperes in a single element. Efficiency is approximately 10^{-14} %, and element excitations must be determined to about 1 part in 10^{14} to obtain the T4 directivity and the 29 db sidelobe level.

END

21-02-86

FIN